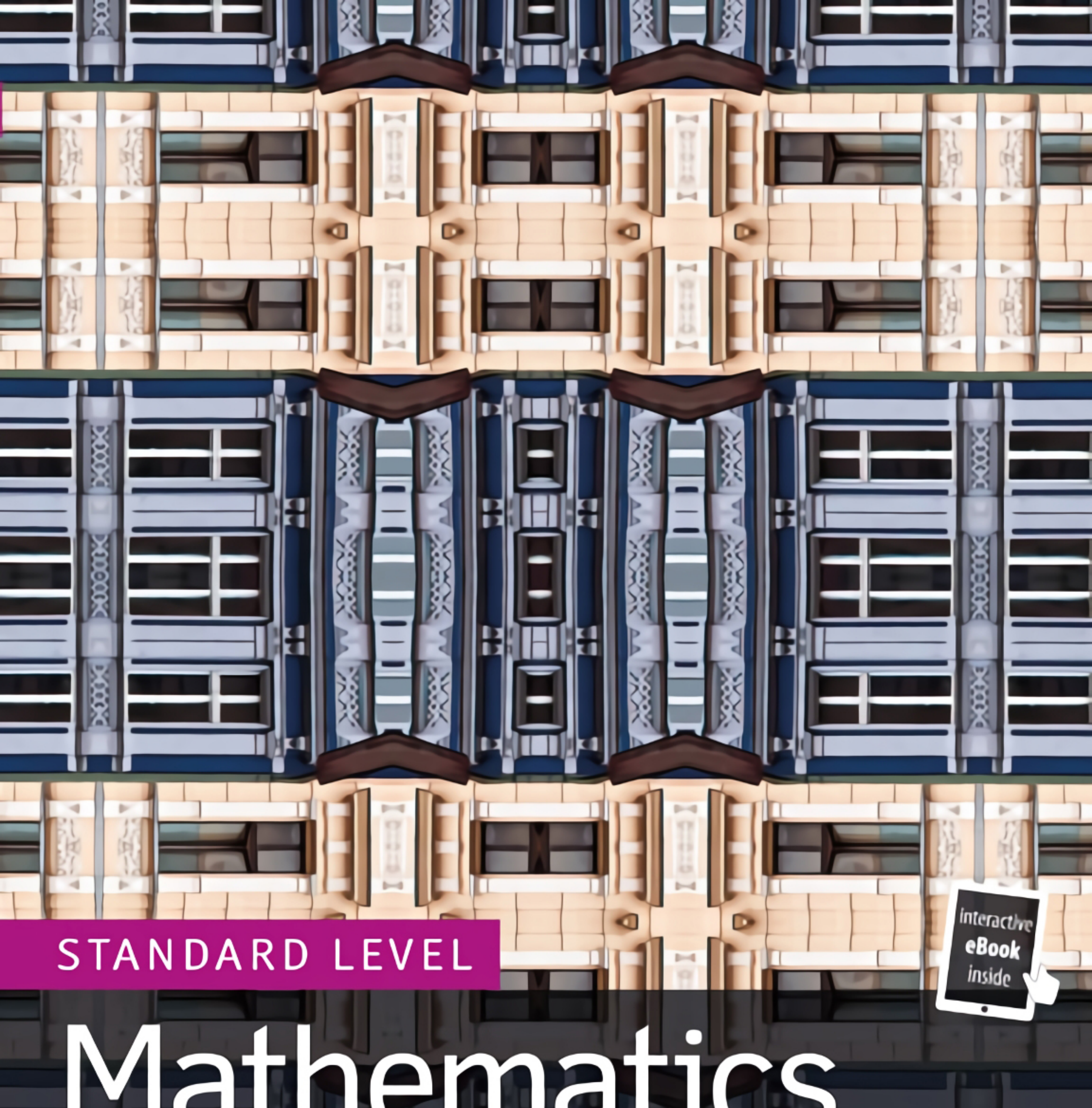


STANDARD LEVEL

interactive eBook inside

Mathematics
Applications and Interpretation

for the IB Diploma

IBRAHIM WAZIR
TIM GARRY
KEVIN FREDERICK
BRYAN LANDMANN

Pearson

STANDARD LEVEL

Mathematics

Applications and Interpretation

for the IB Diploma

IBRAHIM WAZIR
TIM GARRY
KEVIN FREDERICK
BRYAN LANDMANN

Published by Pearson Education Limited, 80 Strand, London, WC2R 0RL

www.pearsonglobalschools.com

Text © Pearson Education Limited 2019
Theory of knowledge chapter authored by Ric Sims
Development edited by Jim Newall
Copy edited by Linnet Bruce
Edited by Linnet Bruce and Sam Hartburn
Indexed by Georgie Bowden
Designed by © Pearson Education Limited 2019
Typeset by © Tech-Set Ltd, Gateshead, UK
Original illustrations © Pearson Education Limited 2019
Illustrated by © Tech-Set Ltd, Gateshead, UK
Cover design by © Pearson Education Limited 2019
Cover images: Front: © **Getty Images**: Busà Photography
Inside front cover: **Shutterstock.com**: Dmitry Lobanov

The rights of Ibrahim Wazir, Tim Garry, Kevin Frederick and Bryan Landmann to be identified as the authors of this work have been asserted by them in accordance with the Copyright, Designs and Patents Act 1988.

First published 2020

24 23 22 21 20
IMP 10 9 8 7 6 5 4 3 2 1

British Library Cataloguing in Publication Data
A catalogue record for this book is available from the British Library

ISBN 978 0 435 19345 4

Printed in Slovakia by Neografia

Acknowledgements
The authors and publisher would like to thank the following individuals and organisations for their kind permission to reproduce copyright material.

Photographs

(Key: b-bottom; c-centre; l-left; r-right; t-top)

123RF: Leonard Zhukovsky 148, **Alamy Stock Photo:** Llewellyn/Alamy Stock Photo 145, **Getty Images:** Kuninobu Sato/EyeEm/Getty Images 01, Wonsup Im/EyeEm 027, ©Daniela White Images/Moment 051, Alberto Manuel Urosa Toledano/Moment 091, Franco Tollardo/EyeEm 125, Domenico De Santo/500px 207, Sebastian-Alexander Stamatis/500px 253, Brasil2/E+ 297, Roc Canals Photography/Moment 325, Ditto/Image Source 395, Nobilis Bellator/500px 435, Vectorig/DigitalVision Vectors 461, **Shutterstock:** Salem Alforaih 363.

All other images © Pearson Education

We are grateful to the following for permission to reproduce copyright material:

Text

Text extracts relating to the IB syllabus and assessment have been reproduced from IBO documents. Our thanks go to the International Baccalaureate for permission to reproduce its copyright.

P 156 Transport for London: Tariffs, Transport of London, Privacy & cookies, Retrieved from https://tfl.gov.uk/modes/taxis-and-minicabs/taxi-fares/tariffs#on-this-page-0; **P 282 ACLI:** American Council of Life Insurers, Life Insurers Fact Book, 2017. © 2017, ACLI; **P 468 tylervigen.com:** Tyler Vigen, Spurious Correlations. Data 2000-2009. Retrieved from http://tylervigen.com/view_correlation?id=2323; **P 492 Scientific American Library:** McMahon, Thomas A; Bonner, John Tyler; On Size and Life, © 1983, Scientific American Library; **P 502 Hachette UK:** Stewart, Ian., Nature's Numbers: The Unreal Reality Of Mathematics, © 2008, Hachette UK. Used with Permission; **P 512, 513 Edge Foundation, Inc.:** What Kind of Thing Is A Number? A Talk with Reuben Hersh, October 2, © 1997, Edge Foundation, Inc. Used with Permission.

Dedications

First and foremost, I want to thank my wife, friend, and devotee, Lody, for all the support she has given through all of these years of my work and career. Most of that work occurred on weekends, nights, while on vacation, and other times inconvenient to my family. I could not have completed this effort without her assistance, tolerance, and enthusiasm.

Most importantly, I dedicate this book to my four grandchildren, Marco, Roberto, Lukas, and Sophia, who lived through my frequent absences from their events.

I also would like to extend my thanks to Catherine Barber, our Commissioning Editor at Pearson for all her support, flexibility and help.

Ibrahim Wazir

In loving memory of my parents.

I wish to express my deepest thanks and love to my wife, Val, for her unflappable good nature and support – and for smiling and laughing with me each day. I am infinitely thankful for our wonderful and kind-hearted children – Bethany, Neil and Rhona. My love for you all is immeasurable.

Tim Garry

I have inexpressible gratitude and appreciation for my wife Julie who has supported me and our family as I took on this second job. The only life raft here is love.

Lena and William, my heart ached every time I had to turn you away to get work done. Thank you for your acceptance and I hope you forgive me.

Ibrahim, thank you for your confidence in me and thinking I am worthy of this task.

Thank you to my grade 12 students who worked through some early chapter drafts and corrected plenty of typos. Slackerz unite!

Finally, to everyone else who picked up loose ends while I was otherwise occupied: thank you for easing the burden of this task.

Kevin Frederick

To my children Joshua and Jakob. May the book inspire them and many others to engage in mathematics.

Bryan Landmann

Contents

Introduction

IB Mathematics: Applications and Interpretation Standard Level **syllabus topics**

1. Number and Algebra
2. Functions
3. Geometry and Trigonometry
4. Statistics and Probability
5. Calculus

This textbook comprehensively covers all of the material in the syllabus for the two-year ***Mathematics: Applications and Interpretation Standard Level*** course of the International Baccalaureate (IB) Diploma Programme (DP). First teaching of this course starts in the autumn of 2019 with first exams occurring in May 2021. We, the authors, have strived to thoroughly explain and demonstrate the mathematical concepts and methods listed in the course syllabus.

Content

As you will see when you look at the table of contents, the five syllabus topics (see margin note) are fully covered, though some are split over different chapters in order to group the information as logically as possible. This textbook has been designed so that the chapters proceed in a manner that supports effective learning of the course content. Thus – although not essential – it is recommended that you read and study the chapters in numerical order. It is particularly important that you thoroughly review and understand all of the content in the first chapter, ***Algebra and function basics***, before studying any of the other chapters.

Other than the final two chapters (**Theory of knowledge** and **Internal assessment**), each chapter has a set of **exercises** at the end of every section. Also, at the end of each chapter there is a set of **practice questions,** which are designed to expose you to questions that are more 'exam-like'. Many of the end-of-chapter practice questions are taken from past IB exam papers. Near the end of the book, you will find answers to all of the exercises and practice questions. There are also numerous **worked examples** throughout the book, showing you how to apply the concepts and skills you are studying.

The Internal assessment chapter provides thorough information and advice on the required **mathematical exploration component**. Your teacher will advise you on the timeline for completing your exploration and will provide critical support during the process of choosing your topic and writing the draft and final versions of your exploration.

The final chapter in the book will support your involvement in the **Theory of knowledge** course. It is a thought-provoking chapter that will stimulate you to think more deeply and critically about the nature of knowledge in mathematics and the relationship between mathematics and other areas of knowledge.

eBook

Included with this textbook is an eBook that contains a digital copy of the textbook and additional high-quality enrichment materials to promote your understanding of a wide range of concepts and skills encountered throughout the course. These materials include:

- Interactive ***GeoGebra*** **applets** demonstrating key concepts
- **Worked solutions** for all exercises and practice questions
- Graphical display calculator (**GDC**) support

To access the eBook, please follow the instructions located on the inside cover.

Information boxes

As you read this textbook, you will encounter numerous boxes of different colours containing a wide range of helpful information.

Learning objectives

You will find learning objectives at the start of each chapter. They set out the content and aspects of learning covered in the chapter.

Learning objectives

By the end of this chapter, you should be familiar with...

- different forms of equations of lines and their gradients and intercepts
- parallel and perpendicular lines
- different methods to solve a system of linear equations (maximum of three equations in three unknowns)

Key facts

Key facts are drawn from the main text and highlighted for quick reference to help you identify clear learning points.

A function is **one-to-one** if each element y in the range is the image of exactly one element x in the domain.

Hints

Specific hints can be found alongside explanations, questions, exercises, and worked examples, providing insight into how to analyse/answer a question. They also identify common errors and pitfalls.

If you use a graph to answer a question on an IB mathematics exam, you must include a clear and well-labelled sketch in your working.

Notes

Notes include general information or advice.

Quadratic equations will be covered in detail in Chapter 2.

Examples

Worked examples show you how to tackle questions and apply the concepts and skills you are studying.

Example 1.5

Find x such that the distance between points $(1, 2)$ and $(x, -10)$ is 13 units.

Solution

$$d = 13 = \sqrt{(x-1)^2 + (-10-2)^2} \Rightarrow 13^2 = (x-1)^2 + (-12)^2$$
$$\Rightarrow 169 = x^2 - 2x + 1 + 144 \Rightarrow x^2 - 2x - 24 = 0$$
$$\Rightarrow (x-6)(x+4) = 0 \Rightarrow x - 6 = 0 \text{ or } x + 4 = 0$$
$$\Rightarrow x = 6 \text{ or } x = -4$$

How to use this book

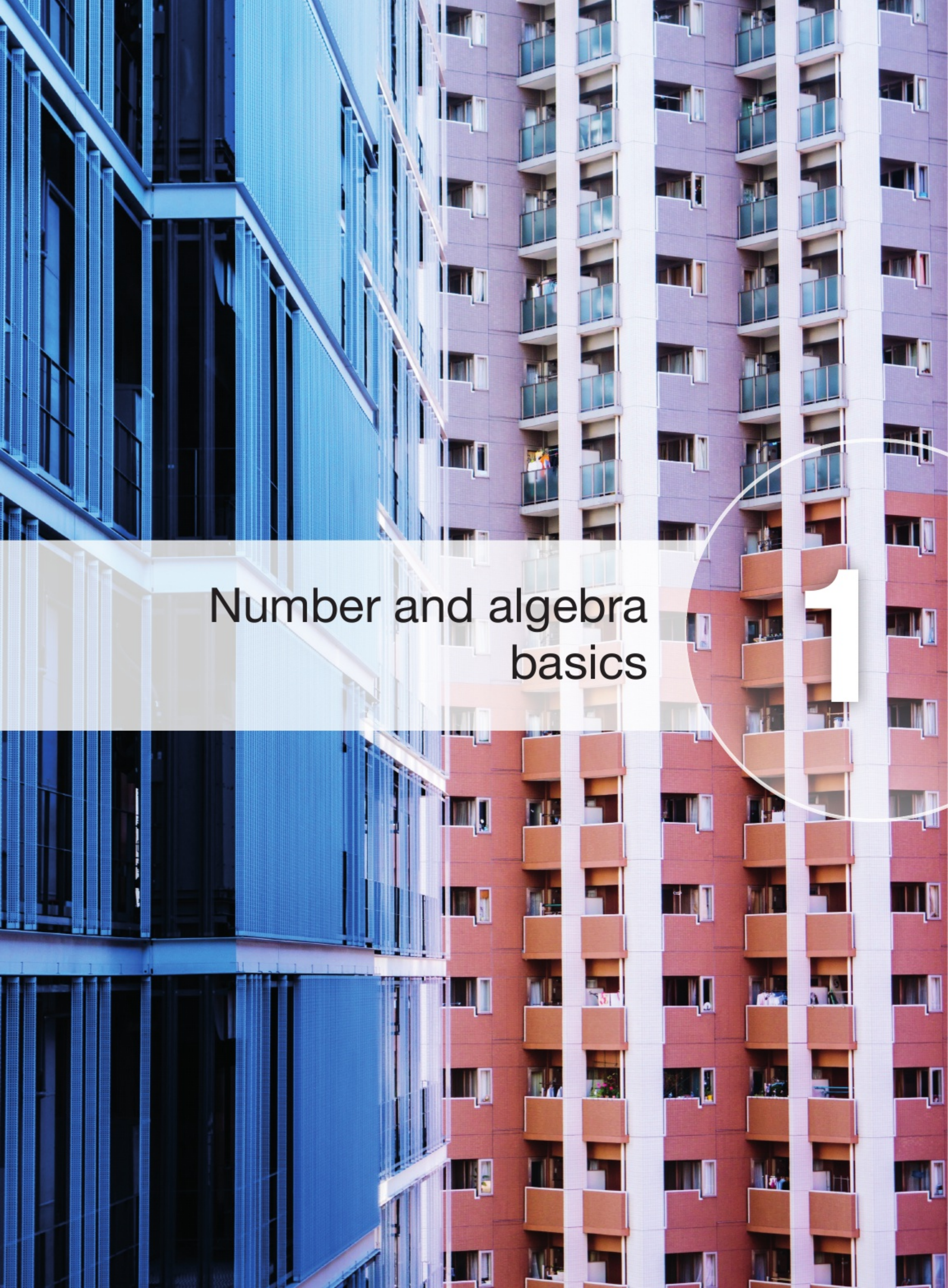
This book is designed to be read by you – the student. It is very important that you read this book carefully. We have strived to write a readable book – and we hope that your teacher will routinely give you reading assignments from this textbook, thus giving you valuable time for productive explanations and discussions in the classroom. Developing your ability to read and understand mathematical explanations will prove to be valuable to your long-term intellectual development, while also helping you to comprehend mathematical ideas and acquire vital skills to be successful in the *Applications and Interpretation Standard Level* course. Your goal should be understanding, not just remembering. You should always read a chapter section thoroughly before attempting any of the exercises at the end of the section.

Our aim is to support genuine inquiry into mathematical concepts while maintaining a coherent and engaging approach. We have included material to help you gain insight into appropriate and wise use of your GDC and an appreciation of the importance of proof as an essential skill in mathematics. We endeavoured to write clear and thorough explanations supported by suitable worked examples, with the overall goal of presenting sound mathematics with sufficient rigour and detail at a level appropriate for a student of HL mathematics.

Our thanks go to Kevin Frederick and Bryan Landmann who joined our team for this edition, helping us to add richness and variety to the series.

For over 10 years, we have been writing successful textbooks for IB mathematics courses. During that time, we have received many useful comments from both teachers and students. If you have suggestions for improving this textbook, please feel free to write to us at globalschools@pearson.com. We wish you all the best in your mathematical endeavours.

Ibrahim Wazir and Tim Garry



1 Number and algebra basics

1 Number and algebra basics

Learning objectives

By the end of this chapter, you should be familiar with…

- using approximations, bounds and percentage error to report numerical information to different levels of accuracy and evaluating the validity of calculations
- simplifying exponential expressions using the laws of exponents
- writing numbers in scientific notation and completing computations relevant in the real world
- logarithms as the inverse of exponents
- evaluating simple logarithmic expressions and solving exponential equations of the form $a^x = b$

In this chapter, the basic structures and concepts of numeracy and algebra used throughout the course will be introduced. To begin, methods for evaluating the accuracy and precision of measurements will be developed. These are particularly important in the sciences and architecture. We then look at the algebraic structure and applications of exponents, which are used to develop models in biology, physics and mathematical finance. As a special application of the exponent laws, scientific notation is examined as a practical means of representing quantities of very large or very small magnitudes, such as the mass of an electron (9.1×10^{-31} kg) or the number of particles in one mole of a substance, the Avogadro number (6.02×10^{23}). The chapter concludes with a treatment of logarithms, a mathematical structure used to invert exponentiation. Techniques for evaluating logarithms, as well as their use in solving exponential equations, will be discussed. Practical applications of logarithms include the quantitative description of acidity (pH scale) and earthquake intensity (Richter scale).

1.1 Approximations and error

Approximations

By rounding the answers to our calculations to a specific place value, number of decimal places (d.p.) or number of significant figures (s.f.), we obtain an approximation of the true value of the calculation. Estimation, where the initial values in a calculation are rounded, allows us to complete computations more easily. However, estimates of calculations can vary greatly from the true value, and the validity of such estimates should be evaluated.

To round a number to the 3rd place value for example, we look at the digit in the fourth place. If this digit is equal to or greater than 5, the 3rd position increases by 1 (we round up), otherwise the number remains unchanged (we round down). For instance, 4.375 rounded to the nearest whole number would be 4, as we round down due to the 3 in the tenth place. Rounded correctly to the nearest tenth (1 d.p.) we would obtain 4.4, as we round up due to the 7 in the hundredths place.

We can also use significant figures to approximate measured values.

1. All non-zero digits in a number are considered significant. For example, in 2300, the 2 and 3 are significant.
2. Zeros bounded by non-zero digits are significant, as we can assume that these zeros were measured. For example, 23001 has 5 significant figures, as the bounded zeros are significant.
3. Zeros to the right of a whole number, or to the left of a number in decimal expansions less than 1 are not significant, as they are required to report the magnitude of the number. For example, the zeros in 4500 or in 0.005 are not significant.
4. Zeros to the right of a number in a decimal point are significant, as we can assume that they represent additional levels of accuracy. For example, 2.30 has 3 significant figures, as the zero in the hundredthsposition is significant.

When completing calculations using quantities reported to varying numbers of significant figures, you should write your final answer using the number of significant figures of the least accurate measurement. It is often necessary to round the quantity to a specified number of significant figures.

Example 1.1

(a) Round the number 10.562 to

(i) the nearest whole number

(ii) 2 decimal places (2 d.p.)

(iii) 3 significant figures (3 s.f.).

(b) Complete the table.

Quantity	Number of significant figures (s.f.)	Round to …
25 400		2 s.f.
0.006		1 s.f.
0.3052		3 s.f.
40.00560		4 s.f.

Solution

(a) (i) $10.562 \approx 11$ to the nearest whole number

(ii) $10.562 \approx 10.56$ (2 d.p.)

(iii) $10.562 \approx 10.6$ (3 s.f.)

If 10.562 represented the number of cans of paint required to paint the walls of a classroom, then it would be reasonable to round this value to 11 cans. If 10.562 represented a monetary value in a currency such as dollars ($), one would write 10.56 to represent cents. On IB Examinations, it is often the case that numbers are rounded to three significant figures unless otherwise stated.

Calculations involving currencies such as dollars ($) or euros (€) are written in general to two decimal places.

(b)

Quantity	Number of significant figures (s.f.)	Round to …
25 400	3	2 s.f. 25 000
0.006	1	1 s.f. 0.006
0.3052	4	3 s.f. 0.305
40.00560	7	4 s.f. 40.01

Example 1.2

Use estimation to calculate each of the following, then do the full calculation and write your answers correct to 2 d.p. and 3 s.f.

(a) $2.67 + 3.94$ (b) $(5.23) \times (4.89)$ (c) $\dfrac{3.25 \times 7.52}{4.37}$

Solution

Note that by rounding the numbers to 3 and 4 respectively, one could have correctly obtained the answer of 7 correct to the nearest whole number. Also, when rounded to the nearest whole number, 6.61 is 7.

(a) Round the numbers to 3 and 4

$3 + 4 = 7$

$2.67 + 3.94 = 6.61$

6.61 is represented to 2 d.p. already and has 3 s.f.

Note that the estimated value need not coincide with the approximate value obtained upon rounding the final answer to the desired level of accuracy. In this case 25.5747 is 26 correct to the nearest whole number.

(b) Rounding 5.23 and 4.89 to 5 and 5 respectively:

$5 \times 5 = 25$

$5.23 \times 4.89 = 25.5747$

25.5747 is 25.57, correct to 2 d.p.

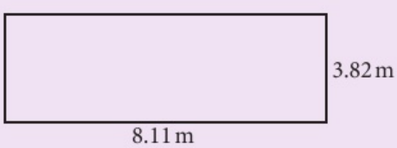
25.5747 is 25.6 correct to 3 s.f.

(c) Using estimation, we obtain $\dfrac{3 \times 8}{4} = 6$

Direct calculation yields an answer of 5.59 correct to 2 d.p. and 3 s.f.

Example 1.3

A rectangular wall of length 8.11 m and height 3.82 m is to be painted. A can of paint costs $5.69 and can cover an area of 5 m^2.



(a) Calculate the area of the wall, first by estimation, and then using a calculator. Write all the decimal places shown by your calculator. Then write the area correct to 2 d.p.

(b) Calculate the number of cans of paint required to paint the wall.

(c) Find the cost of painting the wall, reported to a suitable level of accuracy.

Solution

(a) By estimation, the length and width of the rectangular wall can be rounded to 8 m and 4 m respectively, giving an estimated area of 32 m^2.

Direct calculation gives 30.9802 m^2, which is 30.98 m^2 correct to 2 d.p.

(b) $\frac{30.9802 \text{ m}^2}{5 \text{ m}^2} \approx 6.2$ cans, correct to 1 d.p. Thus, 7 cans of paint are required.

(c) $5.69 \times 7 = 39.83$ Thus, the total cost for painting the wall is \$39.83.

Irrational numbers cannot be represented exactly as fractions of integers or as repeating or terminating decimals. We usually represent them by symbols. Calculations that involve irrational numbers, such as those using the circumference and area of circles or diagonals in rectangles, require approximation.

Example 1.4

(a) For the rectangle in Figure 1.1, find the exact length of the diagonal, d, using roots, the complete decimal expansion from your calculator and correct to 3 s.f.

(b) A circle has an area of 10 cm^2. Find the exact length of the radius, r, using roots and π, showing the complete decimal expansion from your calculator and correct to 3 s.f.

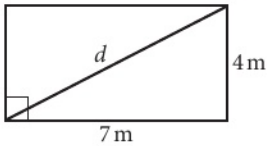


Figure 1.1 Rectangle for Example 1.4

Solution

(a) Using Pythagoras' theorem:

$d^2 = 7^2 + 4^2 = 65$

$d = \sqrt{65}$ m exactly.

A GDC output is $\sqrt{65} = 8.062257748$, which is 8.06, correct to 3 s.f.

(b) The radius, r, of a circle with area A is given by the formula $r = \sqrt{\frac{A}{\pi}}$.

The exact value is of r is $r = \sqrt{\frac{10}{\pi}}$

The value from a calculator is 1.784 124 116 cm, which is 1.78 cm correct to 3 s.f.

Bounds

In this section, we will consider the accuracy of measurements taken using devices such as rulers, scales and clocks. Since all measuring devices have a limited level of precision, any reading that we take using such a device will have a true value that lies in an interval of real numbers, called a **bound**.

If we take a measurement of a quantity and obtain the reading x, then the actual value of the quantity x will lie in the interval

$$x_{\min} \leqslant x < x_{\max}$$

where $x_{\min}$ and $x_{\max}$ are values that depend on the level of precision of our measuring instrument and the reading, x, that has been taken. x can take the minimum value of its bound. However, x can only take values up to, but not including, the maximum value of the bound, as this value would result in the reading of x being increased by one unit on the measuring device due to rounding

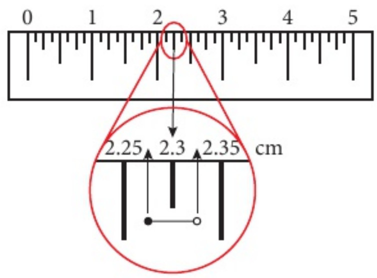


Figure 1.2 Ruler markings

If a measuring device gives a reading of x and has a level of precision p, then the bound is of the form $x_{\min} \leqslant x < x_{\max}$, where $x_{\min} = x - \frac{p}{2}$ and $x_{\max} = x + \frac{p}{2}$. The actual value of x can equal $x_{\min}$, and can take any value smaller than, but not equal to $x_{\max}$.

Suppose a ruler is used to measure a length, and has markings for every millimetre (correct to 0.1 cm). If the value of x is at least 2.25 cm, the nearest marking will be 2.3 cm. For any length x up to, but not including, 2.35, the nearest marking on the ruler is obtained by rounding down to 2.3 cm. Therefore, a reading of $x = 2.3$ cm could be anywhere in the interval $2.25\text{ cm} \leqslant x < 2.35\text{ cm}$. If the actual value of x was 2.35 cm, we would round its value up to 2.4 cm.

Example 1.5

Find the bound for each of the following measurements, given the level of precision, and draw the bound on a number line.

(a) Mass (M) of 42.3 kg, precision of 0.1 kg

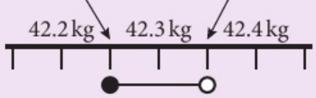
(b) Force (F) of 14 N, precision of 1 N

(c) Time (t) of 5.309 s, precision of 0.001 s

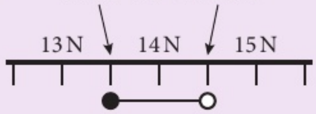
(d) Power (P) of 7.00 W, precision of 0.01 W

Solution

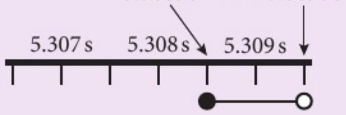
(a) $42.25\text{ kg} \leqslant M < 42.35\text{ kg}$



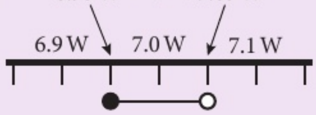
(b) $13.5\text{ N} \leqslant F < 14.5\text{ N}$



(c) $5.3085\text{ s} \leqslant t < 5.3095\text{ s}$



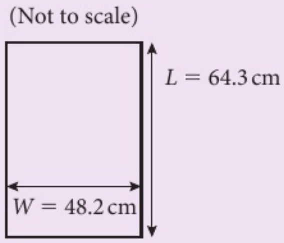
(d) $6.95\text{ W} \leqslant P < 7.05\text{ W}$



In some applications, it is necessary to use multiple measurements to calculate a desired quantity. For example, a calculation of the area of a rectangle requires measurements of length and width. In this case, the bound on the final calculation will depend upon the bounds of each individual measurement used.

Example 1.6

A rectangular picture is to be framed in glass, and has a length L and width W respectively of 64.3 cm and 48.2 cm, correct to 0.1 cm, as shown in the diagram.



(a) Find the bounds on the measurements of L and W.

(b) Find the bound on the area, A, of the picture.

(c) Compare the level of precision of A to that of L and W.

Solution

(a) The bounds for W and L are $48.15\text{ cm} \leqslant W < 48.25\text{ cm}$ and $64.25\text{ cm} \leqslant L < 64.35\text{ cm}$

(b) The maximum value of A is found by taking the product of the maximum values of W and L

$A_{max} = 48.25 \times 64.35 = 3104\text{ cm}^2$

As there are 4 s.f. in the values of W and L, the value of A should also have 4 s.f.

Similarly the minimum value of A is found by using the minimum values of W and L:

$A_{min} = 48.15 \times 62.35 = 3093\text{ cm}^2$

So the bounds on A are $3093\text{ cm}^2 \leqslant A < 3104\text{ cm}^2$.

(c) While L and W have a level of precision correct to 0.1 cm, the level of precision of A is $3104 - 3093 = 11\text{ cm}^2$.

In Figure 1.3 the dotted and dashed rectangles represent the smallest and largest possible rectangles to be constructed using the bounds of the measurements of L and W. Since area is calculated using the product of L and W, the level of precision of A is significantly lower than the precision of the component measurements of L and W.

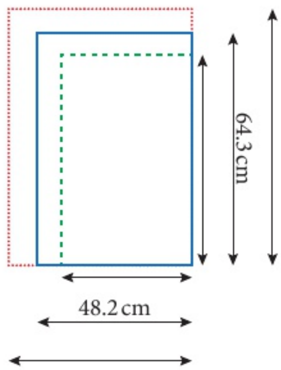


Figure 1.3 Solution to Example 1.6 (b)

Example 1.7

The body mass index (BMI) of an individual is used as an indicator of obesity, and is calculated using the formula:

$\text{BMI} = \dfrac{\text{Mass}}{(\text{height})^2}$, where mass is in kg and height is in m.

Calculate the bounds on the BMI of a person who has a mass of 65.5 kg, correct to 0.1 kg and a height of 1.76 m, correct to 0.01 m.

Solution

First state the bounds on mass, M in kg and height h in m. These bounds are: $65.45\text{ kg} \leqslant M < 65.55\text{ kg}$ and $1.755\text{ m} \leqslant h < 1.765\text{ m}$.

Notice that in this example, the level of precision of mass and height are different, due to the use of different measuring devices.

Since BMI is a fraction, the largest BMI will be found when the numerator, M, is largest and the denominator, h, is smallest:

$$\text{BMI}_{\text{max}} = \frac{65.55}{1.755^2} \approx 21.28\text{ kg m}^{-2}$$

The BMI is a minimum when the minimum mass M, and the maximum height h are used:

$$\text{BMI}_{\text{min}} = \frac{65.45}{1.765^2} \approx 21.01\text{ kg m}^{-2}$$

Combining these two results gives a bound of
$21.01\text{ kg m}^{-2} \leqslant \text{BMI} < 21.28\text{ kg m}^{-2}$

Percentage error of measurements

If we want to quantitatively judge how well a measurement approximates an exact value, we use the concept of percentage error of the measurement, written % error. The % error in a measurement states the difference between the measured and exact values as a percentage of the exact value. In symbols, $\%\text{ error} = \frac{v_e - v_a}{v_e}$, where v_e is the exact value of the quantity and v_a is the approximate value obtained through measurement using a device. Note that % error can be positive or negative, depending on the relative sizes of v_e and v_a.

Example 1.8

A metre stick correct to 1 cm is used to measure the height of a person as 176 cm, and is also used to measure the length of a piece of paper to be 12 cm. Find the magnitude of the maximum percentage error for each measurement and interpret its value.

Solution

For the height h of the person, the bound on the measurement is $175.5\text{ cm} \leqslant h < 176.5\text{ cm}$.

The maximum difference between v_e and v_a would occur if the exact value was the same as the minimum or maximum value of the bound. When

$v_e = 175.5\text{ cm}$, $v_a = 176\text{ cm}$

So $\%\text{ error}_\text{h} = \frac{175.5 - 176}{175.5} \approx -0.285\%$

When $v_e = 176.5$ cm, $v_a = 176$ cm, so % error$_h = \dfrac{176.5 - 176}{176.5} \approx 0.283\%$

The magnitude of the largest error in the measurement of h is 0.285%

For the length, L, of the piece of paper, the bound on the measurement is $11.5 \text{ cm} \leqslant L < 12.5 \text{ cm}$.

When $v_e = 11.5$ cm, $v_a = 12$ cm, % error$_L = \dfrac{11.5 - 12}{11.5} \approx -4.3\%$

When $v_e = 12.5$ cm, $v_a = 12$ cm, % error$_L = \dfrac{12.5 - 12}{12.5} \approx 4.0\%$

The magnitude of the largest error in the measurement of L is 4.3%.

Notice that since both h and L were measured with the same device, the difference between v_e and v_a for all of the examples was ±0.5 cm, corresponding to the precision of the metre stick. However, 0.5 cm is a small percentage of an exact value of 175.5 cm or 176.5 cm, while it is a larger percentage of an exact value of 11.5 cm or 12.5 cm.

If the level of precision of a device is high, the difference between v_a and v_e is low, within a given bound. We would therefore expect that the more precise the device, the lower the magnitude of the % error.

Example 1.9

A runner's time to complete a 100 m sprint is taken using two devices. The first timer measures the time, t_1, to be 11.5 s, correct to 0.1 s, while the second timer measures the time, t_2, to be 11.50 s, correct to 0.01 s. Find the magnitude of the maximum error in each of the measurements and compare their sizes.

Solution

For the first timer, the bound on the measurement is $11.45 \text{ s} \leqslant t_1 < 11.55 \text{ s}$

When $v_e = 11.45$ s, $v_a = 11.5$ s, % error$_{t_1} = \dfrac{11.45 - 11.5}{11.45} \approx -0.437\%$

When $v_e = 11.55$ s, $v_a = 11.5$ s, % error$_{t_1} = \dfrac{11.55 - 11.5}{11.55} \approx 0.433\%$

The maximum error in the measurement of t_1 has a magnitude of about 0.437%

For the second timer, the bound on the measurement is $11.495 \text{ s} \leqslant t_2 < 11.505 \text{ s}$

When $v_e = 11.495$ s, $v_a = 11.50$ s, % error$_{t_2} = \dfrac{11.495 - 11.50}{11.495} \approx -0.04350\%$

When $v_e = 11.505$ s, $v_a = 11.50$ s, % error$_{t_2} = \dfrac{11.505 - 11.50}{11.505} \approx 0.04346\%$

The maximum error in the measurement of t_2 has a magnitude of about 0.0435%; 10 times smaller than the error using the first timer.

Exercise 1.1

1. Write the number 207.6342765 correct to
   (a) the nearest integer
   (b) the nearest hundred
   (c) 2 decimal places
   (d) 3 significant figures
   (e) the nearest tenth
   (f) the nearest ten.

2. For each number
   (i) write the number of significant figures it has
   (ii) round the number correct to 3 significant figures.

   (a) 45 787
   (b) 25 300
   (c) 1001
   (d) 0.03502
   (e) 2.350
   (f) 4.0030
   (g) 20
   (h) 0.01
   (i) 0.0

3. For each measurement, write the bound in the form $x_{min} \leqslant x < x_{max}$
   (a) 23 kg correct to the nearest 1 kg
   (b) 2.7 s correct to the nearest 0.1 s
   (c) 5.23 N correct to the nearest 0.01 N
   (d) 0.020 W correct to 0.001 W

4. Find the largest % error for each quantity.
   (a) A measurement of 23 kg correct to the nearest 1 kg
   (b) A measured length of 5.8 cm, when the actual length is 5.9 cm
   (c) An estimated population of 1.6 million inhabitants in a city when the actual population is 1 567 100
   (d) 3.14 as an estimate of the number π, as shown on your calculator

5. The number $\varphi = \dfrac{1+\sqrt{5}}{2}$ is called the golden ratio and is related to the Fibonacci sequence.
   (a) Calculate φ correct to 2 decimal places and 3 significant figures.
   (b) If we use 1.618 as the exact value of φ, what is the percentage error in approximating φ using 1.6?
   (c) Research the golden ratio and explain how φ is found in nature, art and architecture.

6. A measurement of a quantity $x = 14.5$ is taken, correct to 0.1 units.
   (a) Find the value of x^2 exactly, and rounded to the appropriate number of significant figures.
   (b) Find the percentage error of the rounded answer in part **(a)** compared to the exact value.
   (c) Find the value of $\sqrt{x}$ correct to an appropriate number of significant figures. Why can you not calculate the exact value as in part **(a)**?
   (d) Write a bound for $\sqrt{x}$ to 3 significant figures

7. The diagonal d of a rectangle with length L and width W is given by Pythagoras' theorem $d = \sqrt{L^2 + W^2}$

   (a) A rectangle has $L = 18$ cm and $W = 10$ cm, both correct to the nearest 1 cm. Calculate d to 3 significant figures. How many significant figures would be appropriate for this calculation?

   (b) Find the difference between the maximum and minimum length of the diagonal d.

   (c) An estimate of d is given as 20 cm, and the exact value is the maximum found in **(a)** correct to 3 significant figures. Find the percentage error of the estimate.

8. A cuboid has a length of 0.89 m, correct to 0.01 m, a width of 2.1 m, correct to 0.1 m and a height of 2 m, correct to 1 m.

   (a) Write bounds for the length L, the width W and the height H.

   (b) Find a bound for the volume, V, of the cuboid.

   (c) Estimate the volume V of the cuboid and calculate the percentage error between your estimate and the dimensions given.

   (d) The cube is made from granite, which costs \$150.75 per m³. Find a bound for the total cost of the cuboid, reported to an appropriate level of accuracy.

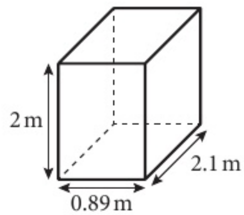


Figure 1.4 Cuboid for question 8

9. An individual has a mass of 75 kg, correct to the nearest kg and a height of 1.85 m, correct to the nearest 0.01 m.

   (a) Calculate the individual's BMI, correct to an appropriate number of significant figures. The formula for BMI is given in Example 1.7.

   (b) Find a bound for the individual's mass, M, and height, h, and hence calculate the individual's minimum and maximum BMI.

   (c) If the individual's actual BMI is a minimum, calculate the percentage error between the minimum BMI and the calculated BMI from part **(a)**.

   (d) **(i)** When two people are of equal height, what can you say about the mass of the individual with the lower BMI?

   (ii) When two people are of equal mass, what can you say about the height of the individual with the lower BMI?

   (iii) Compare the differences between BMI for people aged 16–20 in the country where you live and for the same group in a different country. What other factors could affect BMI?

10. The time it takes for a pendulum of length L to swing once is called the period (T) and is calculated using the formula $T = 2\pi\sqrt{\frac{L}{g}}$, where g is the acceleration due to gravity and has the value 9.8 m s^{-2}.

   (a) A pendulum has $L = 12.4$ cm, correct to 0.1 cm. Calculate T

   (i) correct to 3 significant figures

   (ii) correct to 2 decimal places.

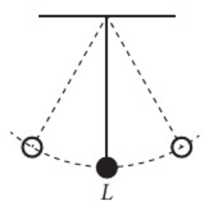


Figure 1.5 Pendulum for question 10

(b) Find a bound for T, correct to 3 decimal places.

(c) What is the minimum time taken for the pendulum to swing 60 times? What is the maximum time?

(d) A rough estimate of the period T is given as 0.68 s. Using your value from **(a)(i)** as the exact value, find the percentage error of the estimate and interpret this value.

11. The mass of a suitcase at an airport is measured using two sets of scales. The first set gives a reading of 23.5 kg, correct to the nearest 0.1 kg, while the second set gives a reading of 24 kg, correct to the nearest 1 kg.

(a) Calculate the maximum percentage error of the measurement of mass using the two sets of scales.

(b) Explain the relationship between level of precision of a measurement and the percentage error of the measurement using the values calculated in **(a)**.

(c) If 100 suitcases are transported in a plane, explain the consequences of using the second set of scales to measure the mass of all of the suitcases.

1.2 The exponent laws

An expression of the form a^b is called a **power** with **base** a and **exponent** b. In this section, rules for simplifying powers will be developed, and will be used to complete computations. a^3 can be thought of as the volume of a cube with side length a. Using this geometric interpretation of a^3, $a^1 = a$ is a length of a units, a^2 is the area of a square with side length a and a^3 is the volume of a cube of side length a.

Example 1.10

Evaluate

(a) 2^3 (b) $(-3)^2$ (c) $\left(\frac{2}{3}\right)^4$ (d) -0.2^5

Solution

(a) $2^3 = 2 \times 2 \times 2 = 8$

(b) $(-3)^2 = (-3)(-3) = 9$

(c) $\left(\frac{2}{3}\right)^4 = \left(\frac{2}{3}\right)\left(\frac{2}{3}\right)\left(\frac{2}{3}\right)\left(\frac{2}{3}\right) = \frac{16}{81}$

(d) $-0.2^5 = -(0.2 \times 0.2 \times 0.2 \times 0.2 \times 0.2) = -0.00032$

Example 1.11

(a) Write $3^4 \times 3^2$ as a single power.

(b) By using the repeated multiplication model of exponents, show that $a^m \times a^n = a^{m+n}$ where m and n are positive whole numbers.

Solution

(a) According to the rule, $3^4 \times 3^2 = 3^{4+2} = 3^6$

(b) Using the repeated multiplication model of exponents, we obtain:

$$a^m = \underbrace{a \times a \times a \times \ldots \times a}_{m \text{ times}} \text{ and } a^n = \underbrace{a \times a \times a \times \ldots \times a}_{n \text{ times}}$$

Therefore, $a^m \times a^n = \underbrace{(a \times a \times a \times \ldots \times a)}_{m \text{ times}}\underbrace{(a \times a \times a \times \ldots \times a)}_{n \text{ times}}$

Due to the associative law of multiplication, the brackets can be removed, and $a^m \times a^n$ can be written as a^{m+n}.

Example 1.12

(a) Simplify $(5^4)^2$

(b) Write the number 64 as a power in as many ways as possible.

(c) Show that any number that is a power with exponent 6 is a cube and a square. Interpret this result geometrically and give a numerical example.

Solution

(a) According to the second exponent law, $(5^4)^2 = 5^{4 \times 2} = 5^8$

(b) Since $64 = 2^6$, we obtain different power representations of 64 by using the different factorisations of the exponent 6. Since $6 = 1 \times 6$, 6×1, 2×3 and 3×2, we obtain the following power representations for 64:

$2^{1\times6} = (2^1)^6 = 2^6$

$2^{6\times1} = (2^6)^1 = 64^1$

$2^{2\times3} = (2^2)^3 = 4^3$

$2^{3\times2} = (2^3)^2 = 8^2$

(c) Using the second exponent law, $a^6 = (a^2)^3 = (a^3)^2$. Therefore, the number a^6 can be thought of as the area of a square with side length a^3 or the volume of a cube with side length a^2, as shown in Figure 1.6.

From part (b) it is clear that $64 = 2^6$ and is the volume of a cube of side length 4 units, or the area of a square of side length 8 units.

The multiplicative law of exponents states that when exponents have the same base, add the exponents: $a^m \times a^n = a^{m+n}$

The associative law for multiplication states that $(a \times b) \times c = a \times (b \times c)$. In words, the product is independent of the grouping of the terms to be multiplied, provided that their order does not change.

The second exponent law states that if a power is raised to an exponent, the result is a power in the original base, but with a new exponent that is the product of the original exponents: $(a^m)^n = a^{mn}$

The second exponent law allows us to simplify powers in which the base itself is a power with an exponent other than 1. This law is very useful in representing a given number as a power in multiple bases.

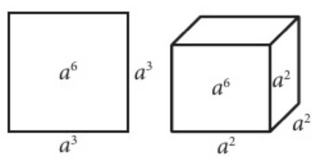


Figure 1.6 Solution for Example 1.12

We can establish the second exponent law using repeated multiplication:

$$(a^m)^n = \underbrace{(a^m)(a^m)\ \ldots\ldots\ldots\ (a^m)}_{n \text{ times}}$$

$$= \underbrace{\underbrace{(a \times a \times \ldots. \times a)}_{m \text{ times}}(a \times a \times \ldots. \times a)\ \ldots..\ (a \times a \times \ldots. \times a)}_{n \text{ times}}$$

Thus, we have n groups of m as being multiplied together. Removing the brackets using the associative law then gives a product of $m \times n$ as, having a value of a^{mn}

The third exponent law states that when the base of a power is a product of two constants a and b, the power can be represented as the product of two powers with base a and b, both using the same exponent, n: $(ab)^n = a^n b^n$

Example 1.13

(a) Write $(5y)^3$ as a product of powers.

(b) Verify that $(ab)^n = a^n b^n$ is true using $(15)^2$ and $3^2 5^2$

(c) Use $(ab)^n = a^n b^n$ to find the prime factor decomposition of $(12)^3$

Solution

(a) $(5y)^3 = 5^3 y^3 = 125y^3$

(b) $15^2 = 225$, $3^2 5^2 = 9 \times 25 = 225$

Thus both expressions are equal.

(c) Using $(a^m)^n = a^{mn}$ and $(ab)^n = a^n b^n$

$12 = (2^2 3)$, thus $12^3 = (2^2 3)^3 = (2^2)^3 (3)^3 = 2^6 3^3$

The commutative law for multiplication states that $ab = ba$

The product of two real numbers a and b is independent of the order in which they are multiplied.

The third exponent law can be extended to consider powers in which the base is a quotient of two constants a and b, $b \neq 0$.

Since $\left(\frac{a}{b}\right) = \left(a \times \frac{1}{b}\right)$, it can be written, for powers with fractional bases, as:

$$\left(\frac{a}{b}\right)^n = \left(a \times \frac{1}{b}\right)^n = a^n \times \frac{1}{b^n} = \left(\frac{a^n}{b^n}\right)$$

The third exponent law can be established using the repeated multiplication model of exponents:

$(ab)^n = (ab)(ab)(ab)\ldots..(ab)$

Using the commutativity law for multiplication, we can remove the brackets and rearrange the terms so that all as and bs are multiplied together to obtain

$$= \underbrace{aaaaa\ldots.aaaa}_{n \text{ times}}\underbrace{bbbbbb\ldots.bbb}_{n \text{ times}} = a^n b^n$$

Example 1.14

(a) Show by choosing values for a and b that $(a + b)^2 \neq a^2 + b^2$

(b) By expanding $(a + b)(a + b)$, show that $(a + b)^2 = a^2 + 2ab + b^2$

(c) Use a counter example to conclude that for $n > 2$ and $a, b \neq 0$, $(a + b)^n \neq a^n + b^n$

Solution

(a) Choosing $a = 1$ and $b = 1$ for example, gives $(1 + 1)^2 = 4$ and $1^2 + 1^2 = 2$

(b) As $ab = ba$:

$(a + b)(a + b) = a^2 + ab + ba + b^2 = a^2 + 2ab + b^2$

(c) If $n > 2$, and $a = 1$ and $b = 1$, then $(1 + 1)^n = 2^n$ but $1^n + 1^n = 1 + 1 = 2$. Thus, $(a + b)^n \neq a^n + b^n$

When using the third exponent law, it is important to realise that multiplication is being done within the bracket. Example 1.14 shows that if the base of a power is a sum, it is not correct to just add each sum and raise to the given exponent.

$(a + b)^n$ can be expanded using the binomial theorem, which is beyond the scope of this course. The binomial theorem has wide applications in statistics and probability.

When one power is divided by another and both have the same base, the quotient is a power in the same base whose exponent is the difference between the exponents in the numerator and denominator of the fraction: $\frac{a^m}{a^n} = a^{m-n}$

Exponent law 4 allows us to simplify quotients of powers in the same base, and in words states that if two powers in the same base are divided, the quotient is a power in the same base whose exponent is the difference between the exponents in the numerator and denominator of the fraction. As division by zero is undefined, it is assumed that $a \neq 0$

Example 1.15

Write each quotient as a single power.

(a) $\frac{3^6}{3^2}$ (b) $\frac{x^9}{x^7}$, $x \neq 0$ (c) $\frac{4^9}{2^6}$ (d) $\frac{8^5}{8^5}$ (e) $\frac{3^2}{3^3}$

Solution

(a) $\frac{3^6}{3^2} = 3^{6-2} = 3^4$

We can check this by calculating: $\frac{3^6}{3^2} = \frac{729}{81} = 81 = 3^4$

(b) $\frac{x^9}{x^7} = x^{9-7} = x^2$, $x \neq 0$

(c) Since 4 and 2 are different bases, we cannot simplify this directly. However, since $4 = 2^2$, we can rewrite the fraction using equal bases and then write it as a single power: $\frac{4^9}{2^6} = \frac{(2^2)^9}{2^6} = \frac{2^{18}}{2^6} = 2^{18-6} = 2^{12}$

(d) $\frac{8^5}{8^5} = 8^{5-5} = 8^0$ Notice that since numerator and denominator are equal, this also gives a value of 1.

(e) $\frac{3^2}{3^3} = 3^{2-3} = 3^{-1}$

Zero, negative and rational exponents

For a^b, the repeated multiplication model of exponents implies that b must be a positive whole number. However, Example 1.15 parts (d) and (e) show that we can get zero and negative exponents. We can also verify with a calculator that $4^{\frac{1}{2}} = 2 = \sqrt{4}$

When a power has a negative integer exponent, it is equal to the reciprocal of the base raised to the positive exponent: $a \neq 0, a^{-m} = \frac{1}{a^m}$

In this section, we develop appropriate models and methods to simplify expressions with zero, negative and fractional exponents, based on the exponent laws developed previously.

When $a \neq 0$, $a^0 = 1$. We can demonstrate this using the fourth exponent law: $a^0 = a^{m-m} = \frac{a^m}{a^m} = 1$ for any whole number m. From Example 1.15 part (e), $\frac{3^2}{3^3} = 3^{2-3} = 3^{-1}$. However, $\frac{3^2}{3^3} = \frac{3 \times 3}{3 \times 3 \times 3} = \frac{1}{3}$ so that $3^{-1} = \frac{1}{3^1}$

Example 1.16

Write each of these powers using positive exponents and evaluate where possible.

(a) 3^{-2} (b) 4^{-3} (c) $\left(\frac{1}{2}\right)^{-1}$ (d) $(3x^{-4})^2$ (e) $(x^{-1}y^3)^{-2}$

Solution

(a) $3^{-2} = \frac{1}{3^2} = \frac{1}{9} \approx 0.111$ (3 s.f.)

(b) $4^{-3} = \frac{1}{4^3} = \frac{1}{64} = 0.015\,625$

(c) $\left(\frac{1}{2}\right)^{-1} = \frac{1}{\left(\frac{1}{2}\right)^1} = 2$

(d) $(3x^{-4})^2 = 3^2x^{-8} = \frac{3^2}{x^8} = \frac{9}{x^8}$

(e) $(x^{-1}y^3)^{-2} = x^{(-1)(-2)}y^{-6} = \frac{x^2}{y^6}$

We can establish $a^{-m} = \frac{1}{a^m}$ using the repeated multiplication model and $\frac{a^m}{a^n} = a^{m-n}$:

$a^{-m} = a^{0-m} = \frac{a^0}{a^m}$

Provided $a \neq 0$, $a^0 = 1$, the result follows directly.

Example 1.17

(a) Write each of these powers using roots and evaluate.

(i) $16^{\frac{1}{2}}$ (ii) $27^{\frac{1}{3}}$ (iii) $64^{\frac{3}{2}}$ (iv) $(-8)^{\frac{5}{3}}$ (v) $\left(\frac{1}{49}\right)^{\frac{1}{2}}$

(b) For which values of n is $n^{\frac{1}{2}}$ defined? What about $(-1)^n$?

(c) Solve $p^{\frac{2}{3}} = 25$ and $49^p = 343$ for the variable p.

Solution

(a) (i) $16^{\frac{1}{2}} = \sqrt{16} = 4$

(ii) $27^{\frac{1}{3}} = \sqrt[3]{27} = 3$

(iii) $64^{\frac{3}{2}} = (\sqrt{64})^3 = 8^3 = 512$

(iv) $(-8)^{\frac{5}{3}} = \left(\sqrt[3]{-8}\right)^5 = (-2)^5 = -32$

(v) $\left(\frac{1}{49}\right)^{\frac{1}{2}} = \sqrt{\frac{1}{49}} = \frac{1}{7}$

(b) Since $n^{\frac{1}{2}} = \sqrt{n}$, $n \geqslant 0$. For $(-1)^n$, n can be any integer. However, if $n = \frac{a}{b}$ is a fraction in lowest terms, b cannot be an even integer.

(c)

$$\begin{aligned} p^{\frac{2}{3}} &= 25 \\ \left(p^{\frac{2}{3}}\right)^{\frac{1}{2}} &= 25^{\frac{1}{2}} \\ p^{\frac{1}{3}} &= 5 \\ p &= 5^3 = 125 \end{aligned}$$

$$\begin{aligned} 49^p &= 343 \\ (7^2)^p &= 7^3 \\ 7^{2p} &= 7^3 \\ 2p &= 3 \\ p &= \frac{3}{2} \end{aligned}$$

An exponent of the form $\frac{1}{m}$ has the same function as taking the mth root of the base a: $a^{\frac{1}{m}} = \sqrt[m]{a}$

Exponents of the form $\frac{n}{m}$ have the equivalent function as either taking the mth root of a and raising this result to the exponent n, or taking the mth root of a^n: $a^{\frac{n}{m}} = \left(\sqrt[m]{a}\right)^n = \sqrt[m]{a^n}$

We can establish $a^{\frac{n}{m}} = \left(\sqrt[m]{a}\right)^n = \sqrt[m]{a^n}$ using $(a^m)^n = a^{mn}$ and the fact that the root of a number is unique. If we consider $x^{\frac{1}{n}}$ and $\sqrt[n]{x}$, both expressions have the property that they give a result of x when raised to the exponent n and are therefore equal. From this result, it follows that $(x)^{\frac{m}{n}} = x^{m \times \frac{1}{n}} = (x^m)^{\frac{1}{n}} = \left(\sqrt[n]{x^m}\right)$. Alternatively, using the commutative law for multiplication, $x^{m \times \frac{1}{n}} = x^{\frac{1}{n} \times m} = \left(x^{\frac{1}{n}}\right)^m = \left(\sqrt[n]{x}\right)^m$ and we establish both expressions for $x^{\frac{m}{n}}$ in terms of roots of x.

We will now consider how to combine the exponent laws to simplify complex expressions including powers in the same base. The following table summarises the exponent laws.

Summary of exponent laws

Exponent law 1:	$a^m \times a^n = a^{m+n}$	
Exponent law 2:	$(a^m)^n = a^{mn}$	
Exponent law 3:	(i) $(ab)^n = a^n b^n$	(ii) $\left(\frac{a}{b}\right)^n = \frac{a^n}{b^n}, b \neq 0$
Exponent law 4:	$\frac{a^m}{a^n} = a^{m-n}$	
Exponent law 5:	$a \neq 0, a^{-m} = \frac{1}{a^m}$	
Exponent law 6:	(i) $a^{\frac{1}{m}} = \sqrt[m]{a}$	(ii) $a^{\frac{n}{m}} = \left(\sqrt[m]{a}\right)^n = \sqrt[m]{a^n}$

Example 1.18

Simplify

(a) $\dfrac{(x^2)^3(x^{-1})^0}{\left(x^2x^{\frac{1}{2}}\right)}, x \neq 0$ (b) $\dfrac{2^6 \times 4^{-1}}{16^{\frac{1}{4}} \times 32}$ (c) $\left(\dfrac{x^2}{y^{-3}}\right)^{-2}, x, y \neq 0$

Solution

(a) $\dfrac{x^6(1)}{x^{2+\frac{1}{2}}} = \dfrac{x^6}{x^{\frac{5}{2}}} = x^{6-\frac{5}{2}} = x^{\frac{7}{2}}$

(b) $\dfrac{2^6(2^2)^{-1}}{(2^4)^{\frac{1}{4}}(2^5)} = \dfrac{2^6 2^{-2}}{2^1 2^5} = \dfrac{2^4}{2^6} = \dfrac{1}{2^2} = \dfrac{1}{4}$

(c) $\left(\dfrac{x^{-4}}{y^6}\right) = \dfrac{\frac{1}{x^4}}{y^6} = \dfrac{1}{x^4y^6}$

Exercise 1.2

1. Simplify each expression completely, writing your answers as a power with positive exponents. Assume $x, y \neq 0$

(a) x^3x^4 **(b)** $\dfrac{x^3}{x^2}$ **(c)** $(x^3)^2$

(d) $(2x)^4$ **(e)** x^{-3} **(f)** $\sqrt[7]{x}$

(g) $\sqrt[3]{x^2}$ **(h)** x^0 **(i)** $x^2x^0(x^3)^4$

(j) $(2x^{-1})^3x^3$ **(k)** $\dfrac{x^2\left(x^{\frac{1}{2}}\right)}{(3x^4)}$ **(l)** $(x^4y^{-2})^3$

(m) $\left(\dfrac{x^5}{y}\right)^7$ **(n)** $\left(\dfrac{3x^{-2}}{2y^3}\right)^2$ **(o)** $\dfrac{1}{(\sqrt[4]{x})^7}$

2. Evaluate:

(a) 5^3 **(b)** $\left(\dfrac{2}{3}\right)^2$ **(c)** 4^{-3}

(d) $100^{\frac{1}{2}}$ **(e)** $27^{\frac{5}{3}}$ **(f)** $\left(\dfrac{1}{4}\right)^{-2}$

3. Solve:

(a) $p^2 = 36$ **(b)** $81^p = 1$ **(c)** $\left(\dfrac{1}{4}\right)^p = 4$

(d) $32^p = 2$ **(e)** $9^p = 27$ **(f)** $7^p = 7$

4. Write $\dfrac{2^2 4^3 8^{-1}}{64(32)}$ as a single power in base 2.

5. Find the value of each geometric quantity, correct to 3 significant figures.

(a) The area of a square with side length 7 cm

(b) The volume of a cube with side length $\sqrt{2}$ cm

(c) The length of one side of a cube with volume 20 cm^3

6. Give three examples of a number that is a square and a cube.

7. **(a)** Write the number 512 as a power in as many different ways as possible.
   (b) If a number x can be written as a power in only one way, a^p, what can you say about p?
   (c) If a number x can be written as a power in exactly two ways, what can you say about the exponent p?

8. For what values of x is the expression x^x defined?

9. Use your GDC to calculate $2^{\sqrt{2}}$ and 2^{π}, rounding your answers to 3 decimal places.
   Why do these calculations not obey any of the exponent laws from this section?

10. Use a GDC to calculate 0^0. Explain this answer by considering the expressions x^0 and 0^x for x values that approach zero.

11. A student claims that a^b can be thought of as repeated multiplication of a by itself b times. To what extent do you agree? Write a paragraph using examples from this section.

12. Using the repeated multiplication model for exponents, show that for $m > n, a \neq 0, \dfrac{a^m}{a^n} = a^{m-n}$

1.3 Scientific notation – numbers in the form $a \times 10^k$

A number written in the form $a \times 10^k$, where $1 \leqslant a < 10$, and k is a positive or negative integer, is said to be written in **scientific notation**. This form is often used in scientific calculations.

To motivate our study of scientific notation, consider these examples. The mass of subatomic particles, such as electrons, is a number less than one with thirty zeros after the decimal, while a proton's mass has approximately 27 non-zero decimal places. Similarly, the speed of light in a vacuum, c, written in m s^{-1} has nine digits. To simplify computations with very large or very small numbers, the size of the number is stored in the exponent of the power in base ten. We can then use the exponent laws of the previous section to simplify calculations.

If the number is whole, we place a decimal after the units digit.

For numbers larger than 10 we must move the digits to the right with respect to the decimal point until we achieve a number a, where $1 \leqslant a < 10$. The number of steps (jumps) required to the right gives the positive exponent, k (see Example 1.19).

Example 1.19

Write in scientific notation:

(a) 35 (b) 146 (c) 65 000 (d) 78.65 (e) 300 000 000

Solution

(a) $35 = 3.5 \times 10^1$ (b) $146 = 1.46 \times 10^2$

(c) $65\,000 = 6.5 \times 10^4$ (d) $78.65 = 7.865 \times 10^1$

(e) $300\,000\,000 = 3 \times 10^8$

Writing numbers less than 1 in scientific notation requires us to move the digits to the left to obtain a number a, where $1 \leqslant a < 10$. This results in the exponent being negative.

Example 1.20

Write in scientific notation:

(a) 0.1 (b) 0.045 (c) 0.00607

Solution

(a) $0.1 = 1 \times 10^{-1}$ (b) $0.045 = 4.5 \times 10^{-2}$ (c) $0.00607 = 6.07 \times 10^{-3}$

Multiplying and dividing numbers written in standard form is convenient, as the powers in base 10 can be simplified using the exponent laws. However, as the exponent laws do not apply to sums of powers, to add or subtract numbers written in standard form requires expanding the number and writing as a decimal.

Example 1.21

(a) Albert Einstein's famous equation $E = mc^2$ relates energy (J), mass (kg) and the speed of light $c = 3.0 \times 10^8\,\text{m s}^{-1}$

Calculate the energy of a proton. The mass of a proton is 1.67×10^{-27} kg

(b) The time, T, in seconds that it takes for light to travel a distance D in kilometres is given by the formula $T = \frac{D}{c}$, where $c = 3.0 \times 10^5\,\text{km s}^{-1}$ is the speed of light in a vacuum. The distance between the Sun and the Earth is 1.5×10^8 km

Calculate the time T taken for light from the Sun to reach the Earth

(i) in standard form (ii) to the nearest minute.

(c) The population of the German cities Munich and Frankfurt in 2017 were 1.43×10^6 and 7.76×10^5 respectively. How many more people live in Munich than in Frankfurt?

Solution

(a) $E = (1.67 \times 10^{-27})(3.0 \times 10^{8})^2$

Using $(a^m)^n = a^{mn}$

$= (1.67 \times 10^{-27})(9.0 \times 10^{16})$

Using the commutative law and $1\text{ J} = 1\text{ kg m}^2\text{ s}^{-2}$

$= 1.67 \times 9.0 \times 10^{-27} \times 10^{16}$

$= 15.03 \times 10^{-11}$

Change 15.03 to standard form

$= 1.503 \times 10^{1} \times 10^{-11}$

$= 1.503 \times 10^{-10}\text{ J}$

(b) (i) $T = \dfrac{1.5 \times 10^8}{3 \times 10^5} = \left(\dfrac{1.5}{3}\right) \times \left(\dfrac{10^8}{10^5}\right) = 0.5 \times 10^3 = 5 \times 10^{-1} \times 10^3$
$= 5 \times 10^2\text{ s}$

(ii) $T = 500\text{ s} = \dfrac{500}{60} \approx 8\text{ min}$

(c) Exponent laws do not allow us to simplify the subtraction of numbers. Write the numbers out in full and carry out the calculation:

$1430\,000 - 776\,000 = 654\,000$

Change the answer back to standard form: 6.54×10^5

```
1.67x10-27x(3x108)²
              1.503x10^-1 0
1.5x108÷3x105
               5.000x10^0 2
1.43x106-7.76x105
               6.540x10^0 5
```

Figure 1.7 GDC output for Example 1.21

Exercise 1.3

1. Write each number in standard form.

(a) 10 807 **(b)** 0.00983 **(c)** 345×10^2

2. Evaluate each of the following, writing your answer in standard form.

(a) $(4.5 \times 10^3) \times (2.1 \times 10^4)$

(b) $(3.2 \times 10^{-2})^2$

(c) $\dfrac{3.5 \times 10^5}{7.0 \times 10^6}$

(d) $5 \times 10^3 + 2.4 \times 10^4$

3. In chemistry, the mole is used to count the number of particles in a substance, and has the value $1\text{ mol} = 6.02 \times 10^{23}$ particles. Find the mass of 1 mole of electrons. Assume that the mass of an electron is 9.1×10^{-31} kg. Write your answers in grams.

4. A city has a population of 1.7×10^6. The population increases by 360 000 people.

(a) Write the new population in scientific notation.

(b) Calculate the percentage increase in population.

5. Newton's law of gravitation is used to calculate the force of attraction that any two bodies exert on each other: $F = G\frac{mM}{r^2}$
where F is the force in Newtons, m and M are the masses of the objects in kilograms, r is the separation distance between the objects in metres and $G = 6.67 \times 10^{11}\,\mathrm{N\,m^2\,kg^{-2}}$
Find the force that two electrons exert on each other when they are separated by a distance of 5×10^{-10} m, writing your answer in scientific notation. The mass of an electron is 9.1×10^{-31} kg.

6. A light year is defined as the distance travelled by light in one year, and is often used in astronomical calculations. Calculate the value of one light year, correct to the nearest km, in scientific notation. The speed of light is $3.00 \times 10^5\,\mathrm{km\,s^{-1}}$

1.4 Exponents and logarithms

$y = a^x \Rightarrow x = \log_a y$

Base	Expression
a	$\log_a m$
10	$\log m$
e	$\ln m$

Exponents and logarithms are inverses of each other. For instance, should a relation be expressed as $y = a^x$, then x is the exponent of the base a which yields the quantity y, written simply as $x = \log_a y$

When the base is 10, we typically do not write it. We can simply write **log *y***. A logarithm with base 10 is called a **common logarithm**.

When the base is **e**, we write **ln *y*.** A logarithm with base **e** is called a **natural logarithm**.

Example 1.22

Earthquake magnitudes (R) are measured on the Richter scale which is a base-10 logarithmic scale, and relative comparisons are useful. For example, an earthquake of magnitude $R = 4$ is ten times as strong as an earthquake of magnitude $R = 3$. What is the magnitude of an earthquake R_1 that is twice as strong as another of magnitude $R = 3$?

Solution

Comparing the relative magnitudes, the equation to be solved for R_1 is:

$$2 = \frac{10^{R_1}}{10^3}$$

$$\Rightarrow 2 \cdot 10^3 = 10^{R_1}$$

$$\Rightarrow R_1 = \log(2 \cdot 10^3) \text{ and using a GDC gives}$$

$$\Rightarrow R_1 \approx 3.301$$

Exercise 1.4

1. Write each equation in logarithmic form.

(a) $1000 = 10^3$ (b) $64 = 4^3$ (c) $100^{\frac{3}{2}} = 1000$

(d) $9^{\frac{1}{2}} = 3$ (e) $2\sqrt{2} = 8^{\frac{1}{2}}$ (f) $10^0 = 1$

(g) $e^0 = 1$ (h) $6^{-2} = \frac{1}{36}$ (i) $(\sqrt{2})^{-2} = \frac{1}{2}$

(j) $3^{-\frac{1}{2}} = \frac{1}{\sqrt{3}}$ (k) $\left(\frac{1}{2}\right)^{-3} = 8$ (l) $8^{-\frac{1}{2}} = \frac{\sqrt{2}}{4}$

(m) $(-2)^3 = -8$ (n) $(0.01)^{-1} = 100$ (o) $\left(\frac{\sqrt{2}}{2}\right)^3 = \frac{\sqrt{2}}{4}$

2. Express each equation in the form $x = \ldots$

(a) $y = 2^x$ (b) $y = 10^x$ (c) $y = e^x$

(d) $y = 2^{3x}$ (e) $y = 3 \cdot 2^x$ (f) $y = 5 - 2^x$

(g) $y = 3^{2x}$ (h) $y = 3^{\frac{x}{2}}$ (i) $y = e^{2x}$

(j) $y = 2^{x-3}$ (k) $y = e^{\frac{x}{2}}$ (l) $y = \frac{1}{2}e^{2x}$

3. Using the earthquake context of Example 1.7, find the magnitude of an earthquake that is:

(a) ten times as powerful as one of magnitude $R = 5.2$

(b) twice as powerful as one of magnitude $R = 5.2$

Chapter 1 practice questions

1. Given $p = x - \frac{\sqrt{y}}{z}$ $x = 1.775$ $y = 1.44$ $z = 48$

(a) calculate the value of p.

Barry first writes x, y and z correct to 1 significant figure and then uses these values to estimate the value of p.

(b) (i) Write down x, y and z, each correct to one significant figure.

(ii) Write down Barry's estimate of the value of p.

(c) Calculate the percentage error in Barry's estimate of the value of p.

2. A rectangle is 2680 cm long and 1970 cm wide.

(a) Find the perimeter of the rectangle, giving your answer in the form $a \times 10^k$, where $1 \leqslant a < 10$ and $k \in \mathbb{Z}$

(b) Find the area of the rectangle, giving your answer correct to the nearest thousand square centimetres.

3. A satellite travels around the Earth in a circular orbit 500 kilometres above the Earth's surface. The radius of the Earth is taken as 6400 kilometres.

   (a) Write down the radius of the satellite's orbit.

   (b) Calculate the distance travelled by the satellite in one orbit of the Earth. Give your answer correct to the nearest km.

   (c) Write down your answer to **(b)** in the form $a \times 10^k$, where $1 \leqslant a < 10, k \in \mathbb{Z}$

4. A shipping container is a cuboid with dimensions 16 m, $1\frac{3}{4}$ m and $2\frac{2}{3}$ m.

   (a) Calculate the exact volume of the container. Give your answer as a fraction.

   Jim estimates the dimensions of the container as 15 m, 2 m and 3 m and uses these to estimate the volume of the container.

   (b) Calculate the percentage error in Jim's estimated volume of the container.

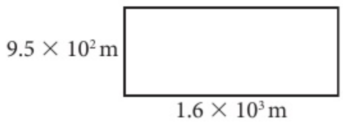


Figure 1.8 Rectangle for question 5

5. Figure 1.8 shows a rectangle with sides of length 9.5×10^2 m and 1.6×10^3 m.

   (a) Write down the area of the rectangle in the form $a \times 10^k$, where $1 \leqslant a < 10, k \in \mathbb{Z}$

   Helen's estimate of the area of the rectangle is 1 600 000 m^2.

   (b) Find the percentage error in Helen's estimate.

6. **(a)** What is the percentage error in estimating $\sqrt{2}$ with 1.41?

   (b) What is the percentage error in estimating π with 3.14?

   (c) Using your answers to parts **(a)** and **(b)**, comment on the accuracy of rounding irrational numbers to two decimal places.

7. A student's height is measured to be 1.89 m, correct to the nearest 0.01 m and the student's mass is 82 kg, measured to the nearest kilogram.

   (a) Find the bounds on the height and mass measurements of the student.

   (b) Calculate the maximum and minimum body mass index (BMI) using the formula $\text{BMI} = \dfrac{\text{mass(kg)}}{[\text{height(m)}]^2}$, rounded to 3 significant figures.

   (c) What is the maximum percentage error in the calculation of the student's BMI?

8. Two cities, A and B, have populations of 1.3×10^6 and 5.8×10^5 respectively.
   (a) How many times larger is city A than city B?
   (b) The two cities are combined to make a new city, C. Write the population of C in scientific notation.
   (c) By what percent has city A grown due to the addition of city B?

9. The distance from the Earth to another planet in the solar system is 5.8×10^8 km
   Light travels at the speed $c = 300\,000 \text{ km s}^{-1}$
   (a) Calculate the time taken for light leaving the other planet to reach the Earth, writing your answer in scientific notation.
   (b) Write your answer to **(a)** correct to the nearest minute.

10. Evaluate each power, writing your answer as a fraction.
   (a) 4^3 (b) $\left(\frac{4}{3}\right)^2$ (c) 2^{-3} (d) $81^{\frac{1}{2}}$

11. Write the expression $\dfrac{4^x 16^{-2x}}{32^x}$ as a single power in base 2.

12. Complete each calculation, writing your answer in scientific notation.
   (a) $(4.2 \times 10^3)(5.3 \times 10^{-1})$ (b) $\dfrac{4.6 \times 10^7}{9.2 \times 10^6}$
   (c) $(2.3 \times 10^3)^3$ (d) $8 \times 10^4 - 7 \times 10^2$

13. Simplify:
   (a) $x^7 x^4$ (b) $\dfrac{x^9}{x^2}$ (c) $(x^4)^2$ (d) $(-2x)^4$ (e) x^{-2}

14. Solve:
   (a) $p^2 = 49$ (b) $32^p = 1$ (c) $\left(\frac{1}{3}\right)^p = 3$
   (d) $81^p = 3$ (e) $8^p = 16$

15. Evaluate:
   (a) $\log_5 25$ (b) $\log 0.1$ (c) $\log_9 3$ (d) $\log_5 30$

16. Solve $4^x = 7$ using logarithms.

17. The formula $\pi(n) = \dfrac{n}{\ln(n)}$ is known as the prime counting function and gives the approximate number of primes less than a whole number n.
   The expression $\ln(n)$ is the logarithm of n in the base $e \approx 2.72$
   (a) Count the number of primes between 1 and 100.
   (b) Calculate $\pi(100)$ using a calculator.
   (c) Calculate the percentage error in using $\pi(100)$ to estimate the number of primes less than 100.

18. The formula $\text{pH} = -\log[H^+]$ is given to measure the acidity of chemicals.

(a) Find $[H^+]$ for sulfuric acid (pH value of 2) and hydrochloric acid (pH of 5).

(b) How much more acidic is sulfuric acid than hydrochloric acid?

19. Two earthquakes A and B have a Richter scale reading of 5.9 and 7.2. How many times more intense is earthquake B than A?

20. Make y the subject of the formula $x = a^y$ using logarithms.

2 Functions

2 Functions

Learning objectives

By the end of this chapter, you should be familiar with...

- different forms of equations of lines; gradients; intercepts; parallel and perpendicular lines
- the concept of a function, domain, range and graph
- function notation; the concept of a function as a mathematical model
- the informal concept of an inverse function as a reflection in $y = x$; the notation $f^{-1}(x)$
- the graph of a function and its key features including: x- and y- intercepts, vertical and horizontal **asymptotes**; using technology to graph a variety of functions
- using technology to find the intersection of two curves or lines.

This chapter focuses on subject material that is essential to understanding further content in this course. You may be familiar with some of the concepts discussed. Nevertheless, it is very important that you master the material in this chapter. Some content is related to the **prior learning topics** for the course. More prior learning topics can be found in the eText in the sections: **Sets, inequalities, absolute value and properties of real numbers**, and **algebraic expressions**.

2.1 Equations and formulae

Equations, identities and formulae

An equation is a statement equating two algebraic expressions that may be true or false depending upon what value(s) are substituted for the variable(s). The value(s) of the variable(s) that make the equation true are called the **solutions** or **roots** of the equation. All the solutions to an equation comprise the **solution set** of the equation. An equation that is true for all possible values of the variable is called an **identity**.

For example, $2x - 3 = 11$ is an **equation**. $x = 7$ is a solution or root of this equation, while $\{7\}$ is its solution set. $y^2 - 5y + 6 = 0$ is an equation which has two roots, $y = 2$ and $y = 3$. So, $\{2, 3\}$ is the solution set for this equation. $(x - 3)^2 = x^2 - 6x + 9$ is an **identity** because it is true for all values of x.

Many equations are referred to as a **formula** (plural: formulae) and typically contain more than one variable and, often, other symbols that represent specific constants or **parameters**. Formulae with which you should be familiar include: $A = \pi r^2$, $d = rt$, and $C = 2\pi r$

Most equations that we will encounter will have numerical solutions, but we can also solve a formula for one variable in terms of other variables – often referred to as changing the subject of a formula.

Example 2.1

(a) Solve for b: $a^2 + b^2 = c^2$

(b) Solve for l: $T = 2\pi\sqrt{\dfrac{l}{g}}$

(c) Solve for R: $I = \dfrac{nR}{R + r}$

Solution

(a) $a^2 + b^2 = c^2 \Rightarrow b^2 = c^2 - a^2 \Rightarrow b = \pm\sqrt{c^2 - a^2}$

If b is a length then $b = \sqrt{c^2 - a^2}$

(b) $T = 2\pi\sqrt{\dfrac{l}{g}} \Rightarrow \sqrt{\dfrac{l}{g}} = \dfrac{T}{2\pi}$

Square both sides and isolate l: $\dfrac{l}{g} = \dfrac{T^2}{4\pi^2} \Rightarrow l = \dfrac{T^2 g}{4\pi^2}$

(c) Multiply both sides of $I = \dfrac{nR}{R + r}$ by $(R + r)$ and simplify.

$I(R + r) = nR \Rightarrow IR + Ir = nR$

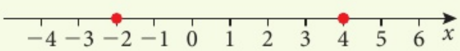
Collect the terms with R on one side and factorise.

$R(n - I) = Ir \Rightarrow R = \dfrac{Ir}{n - I}$

Equations and graphs

Two important characteristics of any equation are the number of variables (unknowns) and the type of algebraic expressions it contains (e.g. polynomials, rational expressions, trigonometric, exponential). Nearly all of the equations in this course will have either one or two variables.

Solutions for equations with a single variable consist of individual numbers that can be graphed as points on a number line. The **graph** of an equation is a visual representation of the equation's solution set.
For example, the solution set of the one-variable equation $x^2 = 2x + 8$ is $\{-2, 4\}$
The graph of the solution set for this equation is depicted on the real number line.



Graph of the solution set for the equation $x^2 = 2x + 8$

The solution set of a two-variable equation will be an **ordered pair** of numbers. An ordered pair corresponds to a location indicated by a point on a two-dimensional coordinate system, i.e. a **coordinate plane**. For example, the solution set of the **quadratic equation** $y = x^2$ will be an infinite set of ordered pairs (x, y) that satisfy the equation. Four ordered pairs in the solution set are shown in red in Figure 2.1. The graph of all the ordered pairs in the solution set forms the curve shown.

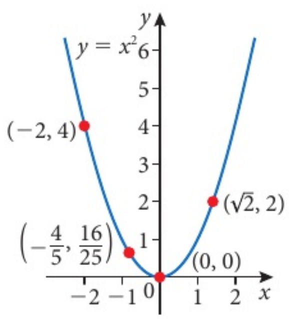


Figure 2.1 Graph of the solution set of the equation $y = x^2$

Linear equations – equations of lines

A **linear equation** in one variable, x, can be written in the form $ax + b = 0$, $a \neq 0$ and it will have exactly one solution $x = -\dfrac{b}{a}$

2 Functions

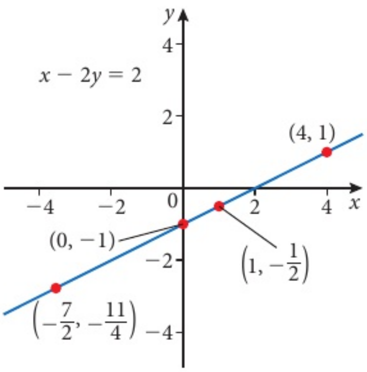


Figure 2.2 The graph of $x - 2y = 2$. A few ordered pairs are graphed in red

A linear equation in two variables x and y can be written in the form $ax + by = c$, where a, b, and c are real numbers. An example of a two-variable **linear equation in x and y** is $x - 2y = 2$. The graph of this equation's solution set (an infinite set of ordered pairs) is a **line** (Figure 2.2).

The **gradient** m **(slope)**, of a non-vertical line is defined by the formula

$$m = \frac{y_2 - y_1}{x_2 - x_1} = \frac{\text{vertical change}}{\text{horizontal change}}$$

For example, using the two points $\left(1, -\frac{1}{2}\right)$ and $(4, 1)$ on the line $x - 2y = 2$,

we find the gradient to be $m = \dfrac{1 - \left(-\frac{1}{2}\right)}{4 - 1} = \dfrac{\frac{3}{2}}{\frac{3}{1}} = \dfrac{1}{2}$

If we solve for y we can rewrite the equation in the form $y = \frac{1}{2}x - 1$

Note that the coefficient of x is the gradient of the line. If $x = 0$, then $y = -1$. Thus, the constant term in the equation is the y-coordinate of the point at which the line intersects the y-axis, that is, the ***y*-intercept**.

> Since all points on a vertical line have the same x-coordinate, the horizontal change between any two points is zero. Thus, vertical lines have undefined gradients because we cannot divide by zero.

There are several forms for writing linear equations.

General form	$ax + by + c = 0$ or $ax + by = c$	Every line has an equation in this form if both a and $b \neq 0$
Gradient–intercept form	$y = mx + c$	m is the gradient; $(0, c)$ is the y-intercept
Point–gradient form	$y - y_1 = m(x - x_1)$	m is the gradient; (x_1, y_1) is a known point on the line
Horizontal line	$y = c$	Gradient is zero; $(0, c)$ is the y-intercept
Vertical line	$x = c$	Gradient is undefined; unless the line is the y-axis it has no y-intercept

Table 2.1 Forms for equations of lines

Most problems involving linear equations and their graphs fall into two categories:

1. Given an equation, determine its graph.
2. Given a graph, or some information about it, find its equation.

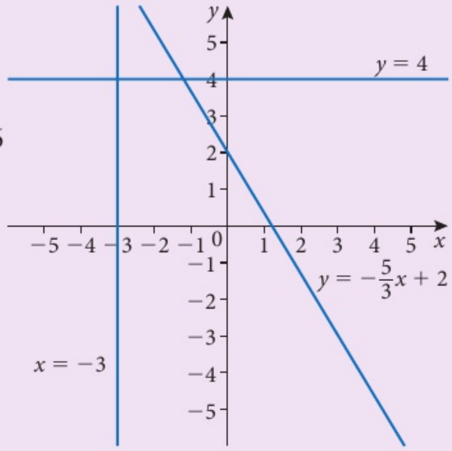
For lines, the first type of problem is often best solved by using the gradient–intercept form. Whereas for the second type of problem the point–gradient form is usually more useful.

Example 2.2

Without using a GDC, sketch the line that is the graph of each linear equation written in general form.

(a) $5x + 3y - 6 = 0$ (b) $y - 4 = 0$ (c) $x + 3 = 0$

Solution



(a) Solve for y to write the equation in gradient–intercept form.

$5x + 3y - 6 = 0 \Rightarrow 3y = -5x + 6$

$\Rightarrow y = -\frac{5}{3}x + 2$

The line has a gradient of $-\frac{5}{3}$ and a y-intercept of $(0, 2)$.

(b) The equation $y - 4 = 0$ is equivalent to $y = 4$ which, when graphed, is a horizontal line with a y-intercept of $(0, 4)$.

(c) The equation $x + 3 = 0$ is equivalent to $x = -3$ which, when graphed, is a vertical line. It has an x-intercept of $(-3, 0)$ but no y-intercept.

Example 2.3

(a) Find the equation of the line that passes through the point $(3, 31)$ and has a gradient of 12. Write the equation in gradient–intercept form.

(b) Find the linear equation in C and F knowing that when $C = 10$, $F = 50$ and when $C = 100$, $F = 212$. Solve for F in terms of C.

Solution

(a) Substitute $x_1 = 3$, $y_1 = 31$ and $m = 12$ into the point–gradient form.

$y - y_1 = m(x - x_1)$

$\Rightarrow y - 31 = 12(x - 3) \Rightarrow y = 12x - 36 + 31$

$\Rightarrow y = 12x - 5$

(b) The two points, (C, F), that are known to be on the line are $(10, 50)$ and $(100, 212)$. The variable C corresponds to the x variable and F corresponds to y in the definitions and forms stated above. The gradient of the line is $m = \frac{y_2 - y_1}{x_2 - x_1} = \frac{F_2 - F_1}{C_2 - C_1} = \frac{212 - 50}{100 - 10} = \frac{162}{90} = \frac{9}{5}$

Choose one of the points on the line, say $(10, 50)$, and substitute it and the gradient into the point–gradient form.

$F - F_1 = m(C - C_1) \Rightarrow F - 50 = \frac{9}{5}(C - 10) \Rightarrow F = \frac{9}{5}C - 18 + 50$

$\Rightarrow F = \frac{9}{5}C + 32$

The following examples demonstrate how the concept of gradient can be used to describe the orientation of straight lines and determine the coordinates of points on a straight line.

Moreover, lines with equal gradients are parallel, and lines for which the product of their gradients is -1 are perpendicular, meaning they intersect in a right angle.

The two lines shown in Figure 2.3 suggest that two distinct non-vertical lines are **parallel** if and only if their gradients are equal, $m_1 = m_2$

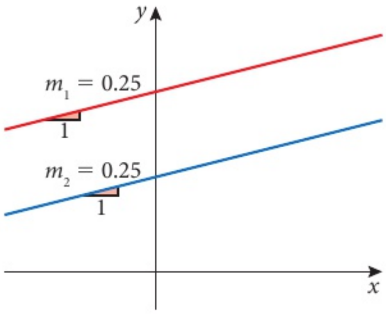


Figure 2.3 Two lines for which the gradients, m_1 and m_2 are equal. The lines are parallel

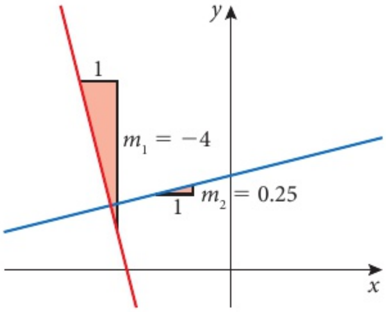


Figure 2.4 Two lines for which the product of the gradients, m_1 and m_2 is -1 are perpendicular

The two lines shown in Figure 2.4 suggest that two non-vertical lines are perpendicular if their gradients are negative reciprocals. That is, $m_1 = -\frac{1}{m_2}$ which is equivalent to $m_1 \cdot m_2 = -1$

Example 2.4

Consider the straight line with equation $y = \frac{3}{5}x + 2$

(a) Find the gradient and the y-intercept.

(b) Find the equation of a line that passes through the point $(1, 5)$ and parallel to the given line.

(c) Find the equation of a line that passes through the point $(1, 5)$ and perpendicular to the given line.

Solution

(a) The gradient, m is the coefficient of x, and hence $m = \frac{3}{5}$, and the constant 2 is the y-coordinate of the y-intercept, which is $(0, 2)$.

(b) The line parallel to the given line must have a gradient of $m_1 = \frac{3}{5}$

Using the form $y - y_1 = m(x - x_1)$, the equation of this line will be $y - 5 = \frac{3}{5}(x - 1) \Rightarrow y = \frac{3}{5}x + \frac{22}{5}$

(c) The line perpendicular to the given line must have a gradient of $m_2 = -\frac{1}{m} = -\frac{5}{3}$

Using the form $y - y_1 = m(x - x_1)$, the equation of this line will be $y - 5 = -\frac{5}{3}(x - 1) \Rightarrow y = -\frac{5}{3}x + \frac{20}{3}$

Example 2.5

Line L is given by the equation $3y + 2x = 9$ and point P has coordinates $(6, -5)$.

(a) Explain why point P is not on line L.

(b) Find the gradient of line L.

(c) (i) Write down the gradient of a line perpendicular to line L.

(ii) Find the equation of the line perpendicular to line L and passing through point P.

Solution

(a) If point $P(6, -5)$ lies on the graph of L then it must be a solution to the equation $3y + 2x = 9$

Substituting 6 for x and -5 for y gives:
$3(-5) + 2(6) = -15 + 12 = -3 \neq 9$

Since $(6, -5)$ is not a solution then P is not on the graph of line L.

(b) To find the gradient of L, rewrite $3y + 2x = 9$ in gradient–intercept form, $y = mx + c$

$3y + 2x = 9 \Rightarrow 3y = -2x + 9 \Rightarrow y = -\frac{2}{3}x + 3$

Thus, the gradient of L is $-\frac{2}{3}$

(c) (i) The line perpendicular to L has a gradient that is the negative reciprocal of $-\frac{2}{3}$. Thus, the gradient is $\frac{3}{2}$

(ii) Using the point–gradient form

$y - y_1 = m(x - x_1)$ the equation of the line perpendicular to L is $y - (-5) = \frac{3}{2}(x - 6) \Rightarrow y + 5 = \frac{3}{2}x - 9 \Rightarrow y = \frac{3}{2}x - 14$

Systems of linear equations

Many problems involve sets of equations with several variables, rather than just a single equation with one or two variables. Such a set of equations is often called a set, or system, of **simultaneous equations** because we find the values for the variables that solve all of the equations simultaneously. In this section, we only consider sets of simultaneous equations containing two linear equations with two variables. We will take a brief look at four solution methods:

1. Graphical method (with technology)
2. Substitution method
3. Elimination method
4. Technology (without graphing)

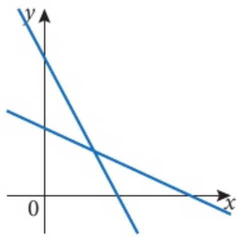
Graphical method

The graph of each equation in a **system of two linear equations in two unknowns** is a line. The graphical interpretation of such a system of equations corresponds to determining what point, or points, lie on both lines. Two lines in a coordinate plane can relate to one another in one of three ways:

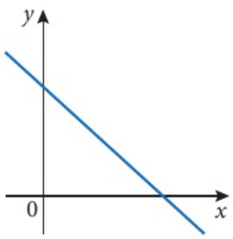
1. intersect at exactly one point
2. intersect at all points on each line (i.e. the lines are identical, or coincident)
3. not intersect (i.e. the lines are parallel).

These three possibilities are illustrated in Figure 2.5.

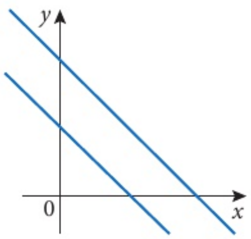
A graphical approach to solving a system of linear equations provides a helpful visual picture of the number and location of solutions. This can be performed on a GDC.



Intersect at exactly one point; exactly one solution



Identical – coincident lines; infinite solutions



Never intersect – parallel lines; no solution

Figure 2.5 Possible relationship between two lines in a coordinate plane

A system of equations that has no solution is **inconsistent**. A system of equations that has at least one solution is **consistent**.

Example 2.6

Use the graphical features of a GDC to solve each system of linear equations.

(a) $2x + 3y = 6$
$2x - y = -10$

(b) $7x - 5y = 20$
$3x + y = 2$

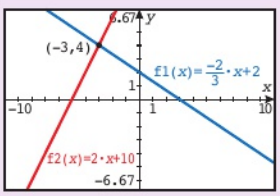
Solution

(a) First rewrite each equation in gradient–intercept form, i.e. $y = mx + c$. This is a necessity if we use our GDC and is also very useful for graphing by hand (manually).

$2x + 3y = 6 \Rightarrow 3y = -2x + 6 \Rightarrow y = -\frac{2}{3}x + 2$ and

$2x - y = -10 \Rightarrow y = 2x + 10$

Any of the popular GDC models can give the point of intersection shown in the diagram. The intersection point and solution to the system of equations is $x = -3$ and $y = 4$, or $(-3, 4)$



(b) $7x - 5y = 20 \Rightarrow 5y = 7x - 20 \Rightarrow y = \frac{7}{5}x - 4$ and

$3x + y = 2 \Rightarrow y = -3x + 2$

The solution to the system of equations is:

$x = 1.364$ and $y = -2.091$, or $\left(\frac{15}{11}, -\frac{23}{11}\right)$ (see GDC images in Figure 2.6)

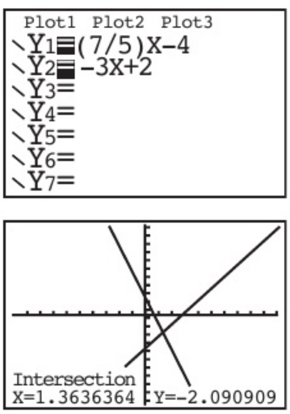


Figure 2.6 GDC screens for the solution to Example 2.6 (b)

In some situations, it may be useful or more efficient to solve a system of equations by using an analytic (or algebraic) approach. Two methods – elimination and substitution – are explained and illustrated here. However, because a GDC is allowed on all exams neither method is required.

Elimination method

To solve a system using the **elimination method**, we try to combine the two linear equations using sums or differences in order to eliminate one of the variables. Before combining the equations, we need to multiply one or both of the equations by a suitable constant so that one of the variables has coefficients that are equal (then subtract the equations), or that differ only in sign (then add the equations).

Example 2.7

Use the elimination method to solve the system of linear equations.

$5x + 3y = 9$

$2x - 4y = 14$

Solution

We can obtain coefficients for y that differ only in sign by multiplying the first equation by 4 and the second equation by 3. Then add the equations to eliminate the variable y.

$$\begin{array}{lcr} 5x + 3y = 9 & \rightarrow & 20x + 12y = 36 \\ 2x - 4y = 14 & \rightarrow & 6x - 12y = 42 \\ \hline & & 26x \qquad = 78 \\ & & x = \dfrac{78}{26} \\ & & x = 3 \end{array}$$

By substituting the value of $x = 3$ into either of the original equations we can solve for y.

$5x + 3y = 9 \Rightarrow 5(3) + 3y = 9 \Rightarrow 3y = -6 \Rightarrow y = -2$

The solution is $(3, -2)$.

Substitution method

The algebraic method that can be applied effectively to the widest variety of simultaneous equations, including non-linear equations, is the **substitution method**. Using this method, we choose one of the equations and solve for one of the variables in terms of the other variable. We then substitute this expression into the other equation to produce an equation with only one variable, for which we can solve directly.

Example 2.8

Use the substitution method to solve the system of linear equations.

$3x - y = -9$

$6x + 2y = 2$

Solution

Solve for y in the top equation.

$3x - y = -9 \Rightarrow y = 3x + 9$

Substitute $3x + 9$ in for y in the bottom equation.

$6x + 2(3x + 9) = 2 \Rightarrow 6x + 6x + 18 = 2 \Rightarrow 12x = -16 \Rightarrow x = -\frac{16}{12} = -\frac{4}{3}$

Now substitute $-\frac{4}{3}$ in for x in either equation to solve for y.

$3\left(-\frac{4}{3}\right) - y = -9 \Rightarrow y = -4 + 9 \Rightarrow y = 5$

The solution is $x = -\frac{4}{3}, y = 5$; or $\left(-\frac{4}{3}, 5\right)$

Technology

As shown in Example 2.6, we can use our GDC to graph the two lines which have equations that constitute a system of two linear equations and then apply an 'intersection' command to find the ordered pair that solves the system. Alternatively, a GDC should have a simultaneous equation solver (or systems of equations solver) that can be used to solve a system of linear equations without graphing.

If you use a graph to answer a question on an IB mathematics exam, you must include a clear and well-labelled sketch in your working. Thus, on an exam there is often more effort involved in solving a system of linear equations by graphing with your GDC compared to solving the system using a simultaneous equation solver (or systems of linear equations solver) on your GDC.

Example 2.9

When flying in still air (i.e. no head wind or tail wind) with its propellers rotating at a particular rate a small airplane has a speed of s kilometres per hour. With the propeller rotating at the same rate, the airplane flies 473 km in 2 hours as it flies against a head wind, and 887 km in 3 hours flying with the same wind as a tail wind. Find s and find the speed of the wind.

Solution

Let w represent the speed of the wind.

Applying distance = speed × time gives the following two equations.

$$\begin{cases} 473 = 2(s - w) \\ 887 = 3(s + w) \end{cases} \Rightarrow \begin{cases} 473 = 2s - 2w \\ 887 = 3s + 3w \end{cases}$$

This system of equations can be solved with a simultaneous equation solver on a GDC. One example of such a solver is shown in Figure 2.7.

Therefore, $s \approx 266\,\text{km h}^{-1}$ and the wind speed is $29.6\,\text{km h}^{-1}$, approximate to 3 significant figures.

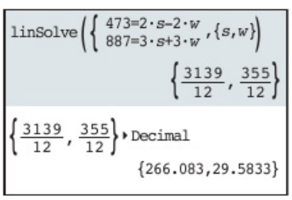


Figure 2.7 GDC simultaneous equation solver

Exercise 2.1

1. Solve each formula for the variable stipulated.

   (a) $m(h - x) = n$ solve for x

   (b) $v = \sqrt{ab - t}$ solve for a

   (c) $A = \frac{h}{2}(b_1 + b_2)$ solve for b_1

   (d) $A = \frac{1}{2}r^2\theta$ solve for r

   (e) $\frac{f}{g} = \frac{h}{k}$ solve for k

   (f) $at = x - bt$ solve for t

   (g) $V = \frac{1}{3}\pi r^3 h$ solve for r

   (h) $F = \frac{g}{m_1 k + m_2 k}$ solve for k

2. Find the equation of the line that passes through the two given points. If possible, write the line in gradient–intercept form.

   (a) $(-9, 1)$ and $(3, -7)$

   (b) $(3, -4)$ and $(10, -4)$

   (c) $(-12, -9)$ and $(4, 11)$

   (d) $\left(\frac{7}{3}, -\frac{1}{2}\right)$ and $\left(\frac{7}{3}, \frac{5}{2}\right)$

3. Find the equation of the line that passes through the point $(7, -17)$ and is parallel to the line with equation $4x + y - 3 = 0$. Write the line in gradient–intercept form.

4. Find the equation of the line that passes through the point $\left(-5, \frac{11}{2}\right)$ and is perpendicular to the line with equation $2x - 5y - 35 = 0$. Write the line in gradient–intercept form.

5. Use the elimination method to solve each system of linear equations.

   (a) $x + 3y = 8$
   $x - 2y = 3$

   (b) $x - 6y = 1$
   $3x + 2y = 13$

6. Use the substitution method to solve each system of linear equations.

   (a) $2x + y = 1$
   $3x + 2y = 3$

   (b) $3x - 2y = 7$
   $5x - y = -7$

7. Use a GDC to solve each system of two linear equations.

   (a) $3x + 2y = 9$
   $7x + 11y = 2$

   (b) $3.62x - 5.88y = -10.11$
   $0.08x - 0.02y = 0.92$

   (c) $2x - 3y = 4$
   $5x + 2y = 1$

8. The equation of the straight line L_1 is $y = 2x - 3$

   (a) Write down the y-intercept of L_1.

   (b) Write down the gradient of L_1.

   The line L_2 is parallel to L_1 and passes through the point $(0, 3)$.

   (c) Write down the gradient of L_2.

   The line L_3 is perpendicular to L_1 and passes through the point $(-2, 6)$.

   (d) Write down the gradient of L_3.

   (e) Find the equation of L_3. Give your answer in the form $ax + by + d = 0$ where a, b and d are integers.

2.2 Relations and functions

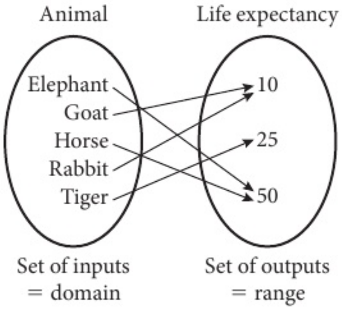


Figure 2.8 Mapping for a set of animals to a set of life expectancies

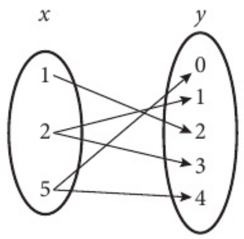


Figure 2.9 Mapping for a set of x values to a set of y values

When the elements in one set are linked to elements in a second set, we call this a **relation**. For example the life expectancy of some animals could be presented in a diagram called a **mapping**. To create the mapping we can look up the life expectancy of each animal. We would not start with a life expectancy and then try to determine the type of animal. This suggests that the direction in a relation from the elements in one set to the elements in another set is important.

In general, for a relation such that x is an element in the **input** (type of animal) and y is an element in the **output** (life expectancy), then we say that y **depends on** x, and we write $x \to y$ (or x determines y). We can also represent this relation as a set of ordered pairs (x, y), where x represents the input and y represents the output: {(elephant, 50), (goat, 10), (horse, 50), ...}

We can represent the mapping shown in Figure 2.9 as a set of ordered pairs: $\{(1, 2),(2, 1),(2, 3),(5, 0),(5, 4)\}$

A relation f defined by a set of ordered pairs (x, y) is a **function** if no two different ordered pairs have the same first coordinate and different second coordinates.

Notice what distinguishes the relation mapped in Figure 2.9 from the one shown in Figure 2.8. When we look at the arrows in the first relation (Figure 2.8), there is only one arrow leaving any element in the first set. This is not true for the second relation (Figure 2.9). When we look at the sets of ordered pairs, the first relation does not have any two pairs with the same first component; in other words, each animal has only one life expectancy. This is not true for the second relation. In the second relation, the number 2 in the input is linked to two different numbers in the output, 1 and 3. The first relation, where each element in the input is linked to only one element in the output, is called a **function**, while the second relation is not a function. The box on the left gives a more formal description of a function and the graphical interpretation of functions.

If the condition of not having different second coordinates is not met, then f is only a **relation** for the two given sets. For example, $\{(0, 1),(1, 2),(-1, 2),(2, 5),(-2, 5),(3, 10),(-3, 10)\}$ represents a function because no two ordered pairs have the same first coordinate and different second coordinates. On the other hand, $\{(0, 0),(1, 1),(2, -1),(1, 2),(3, 4),(5, 8)\}$ does **not** represent a function because $(1, 1)$ and $(1, 2)$ are two ordered pairs with the same first coordinate and different second coordinates.

For a set of ordered pairs to represent a function, all ordered pairs must have different first elements.

The largest possible set of values for the independent variable (the **input** set) is called the **domain** – and the set of resulting values for the dependent variable (the **output** set) is called the **range**. In the context of a mapping, each value in the domain is mapped to its **image** in the range. All the various ways of representing a mathematical function illustrate that its defining characteristic is that it is a rule by which each number in the domain determines a unique number in the range.

The graph of any relation consists of infinitely many ordered pairs each represented by a point on the graph. By analysing the graph of a function, we can often visually determine the domain and range of the function. See Figure 2.10.

For the graph of $y = x^2 + 1$ in Figure 2.11, we notice that there is no restriction on the values of x. Hence, the domain of the relation $y = x^2 + 1$ is the set of all real numbers. Also, the lowest point on the graph (the vertex of the parabola) is at (0, 1) which means that the range of the relation is the set of real numbers greater than or equal to 1. The graph in Figure 2.11 of $y = x^2 + 1$ represents a function because each value of x (domain) determines only one value of y (range). In other words, there are no two different points on the graph of $y = x^2 + 1$ which have the same x-coordinate.

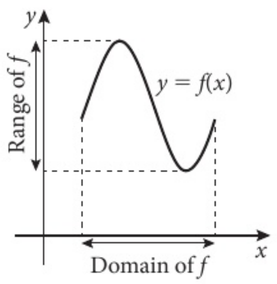


Figure 2.10 Visual depiction of domain and range of a function from its graph

Figure 2.12 shows the graph of the relation $x = y^2 + 1$, which is also a parabola but is 'sideways'. The domain of the relation is the set of real numbers greater than or equal to 1. If we extend the graph, we also notice that there is no limit on the values of y, so the range is the set of all real numbers.

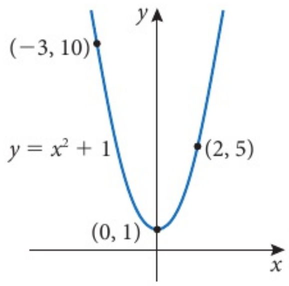


Figure 2.11 The graph of $y = x^2 + 1$

The graph in Figure 2.12 of $x = y^2 + 1$ does not represent a function. If $x = 5$, then y equals both 2 and -2. In other words, (5, 2) and (5, -2) both lie on the graph of $x = y^2 + 1$. Therefore, by definition, the set of ordered pairs represented by the graph of $x = y^2 + 1$ is not a function.

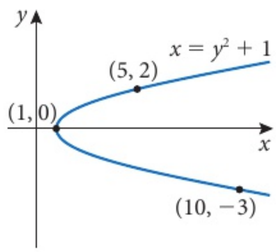


Figure 2.12 The graph of $x = y^2 + 1$

To determine if a relation (equation) is a function, use the **vertical line test**. If a vertical line exists that intersects the graph at more than one point, then the graph is **not** a function. This is because two ordered pairs would have the same first elements (x-coordinates). See Figure 2.13.

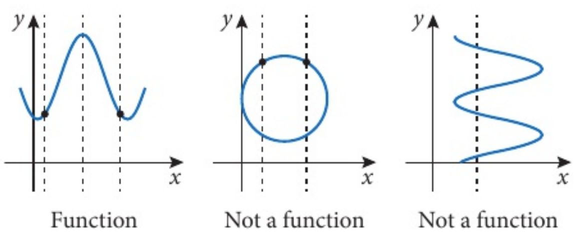


Figure 2.13 Vertical line test

A function is usually denoted by a letter such as f, g, or h. The set X, containing all values of x, is called the **domain** of f. The set of corresponding elements y in the set Y is called the **range** of f. For the animal life expectancy example, the set of animals is the domain and the set of numbers representing life expectancies constitute the range. The unique element y in the range that corresponds to a selected element x in the domain X is called the **value** of the function at x, or the **image** of x, and is written $f(x)$. The latter symbol is read 'f of x' or 'f at x', and we write $y = f(x)$. In many instances, x is also called the **input** of the function f and the value $f(x)$ is called the **output** of f. Since the value of y depends on the choice of x, y is called the **dependent variable**; x is called the **independent variable**. Unless otherwise stated, we shall assume hereafter that the sets X and Y consist of real numbers.

A simple pendulum consists of a heavy object hanging from a string of length L (in metres) and fixed at a pivot point. If we displace the suspended object

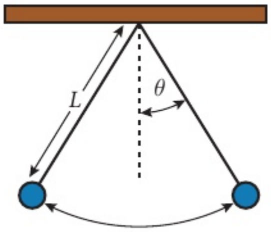


Figure 2.14 A simple pendulum

to one side by a certain angle θ from the vertical and release it, the object will swing back and forth under the force of gravity. The period T (in seconds) of the pendulum is the time for the object to return to the point of release. For a small angle θ the two variables T and L are related by the function $T = 2\pi\sqrt{\frac{L}{g}}$ where g is the gravitational field strength (acceleration due to gravity). Therefore, assuming the force of gravity is constant at a given elevation ($g \approx 9.81\,\mathrm{m\,s^{-2}}$ at sea level), the function can be used to calculate the value of T for any value of L. That is, L determines T.

A variety of functions – including quadratic functions – will be covered in Chapter 6 Modelling real-life phenomena.

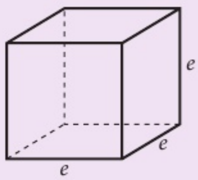
As with the period T and the length L for a pendulum, many mathematical relationships concern how the value of one **independent** variable determines the value of a second **dependent** variable. Here are some further examples:

- The area of a circle A is a function of its radius r: $A = \pi r^2$
- If walking at a constant speed of 5 km per hour, then distance d (km) is a function of time t (hours): $d = 5t$
- °F (degrees Fahrenheit) is a function of °C (degrees Celsius): $F = \frac{9}{5}C + 32$

Example 2.10

(a) Express the volume V of a cube as a function of the length e of each edge.

(b) Express the total surface area S of a cube as a function of the length e of each edge.



Solution

(a) V as a function of e is $V = e^3$

(b) The surface area of the cube consists of six squares each with an area of e^2. Hence, the surface area is $6e^2$; that is, $S = 6e^2$

Example 2.11

Find the domain of each function.

(a) $\{(-6, -3), (-1, 0), (2, 3), (3, 0), (5, 4)\}$ (b) Volume of a sphere: $V = \frac{4}{3}\pi r^3$

(c) $y = \frac{5}{2x - 6}$ (d) $y = \sqrt{3 - x}$

Solution

(a) The function consists of a set of ordered pairs. The domain of the function consists of all first coordinates of the ordered pairs. Therefore, the domain is the set $x \in \{-6, -1, 2, 3, 5\}$

(b) The physical context tells us that a sphere cannot have a radius that is negative or zero. Therefore, the domain is the set of all real numbers r such that $r > 0$

(c) Since division by zero is not defined for real numbers then $2x - 6 \neq 0$ Therefore, the domain is the set of all real numbers x such that $x \in \mathbb{R}$, $x \neq 3$

(d) Since the square root of a negative number is not real, then $3 - x \geqslant 0$ Therefore, the domain is all real numbers x such that $x \leqslant 3$

Example 2.12

Determine the domain and range for the function $y = 1 + \sqrt{x - 2}$

Solution

Using algebraic analysis:

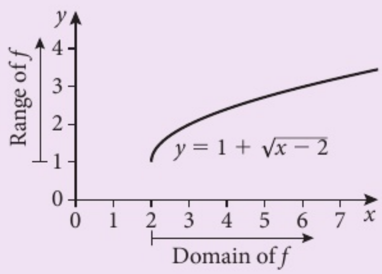
Squaring any real number produces another real number. Therefore, the domain of $y = 1 + \sqrt{x - 2}$ is the set of all real numbers such that the expression under the square root yields a real number. That is $x - 2 \geqslant 0$

Thus, the only values of x that can be part of the domain must be larger than or equal to 2. So, the domain is $x \geqslant 2$

Since $\sqrt{x - 2}$ is always positive, and it is added to 1, then the possible values of y are all numbers larger than or equal to 1. Thus, the range is $y \geqslant 1$

Using graphical analysis:

For the domain, the only possible values of x are from 2 onwards.
For the range the only possible values for y are the values from 1 onwards.



When determining the domain and range of a function, try to use both algebraic and graphical analysis. Do not rely too much on using just one approach. For graphical analysis of a function, producing a graph on a GDC that shows all the important features is essential.

Example 2.13

Find the domain and range of the function $h(x) = \dfrac{1}{x - 2}$

Solution

Using algebraic analysis:

The function produces a real number for all x, except for $x = 2$ when division by zero occurs. Hence, $x = 2$ is the only real number not in the domain. Since the numerator of $\dfrac{1}{x - 2}$ is a positive constant, the value of y cannot be zero. Hence, $y = 0$ is the only real number not in the range.

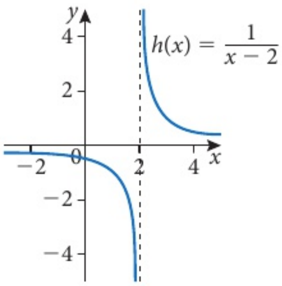


Figure 2.15 Graph for Example 2.13

Using graphical analysis:

The graph of the equation has a 'gap' at $x = 2$ that divides it into two branches that both continue indefinitely with no other gaps. Both branches are said to be **asymptotic** (approach but do not intersect) to the vertical line $x = 2$

This line is a vertical **asymptote** and is drawn as a dashed line (it is not part of the graph of the equation).

Similarly, there is a gap at $y = 0$ (x-axis) with both branches of the graph continuing indefinitely with no other gaps. Both branches are also **asymptotic** to the x-axis. Thus, the x-axis is a horizontal **asymptote**.

Both approaches confirm that the domain and range for $h(x) = \frac{1}{x-2}$ are:

domain: $\{x \mid x \in \mathbb{R}, x \neq 2\}$; range: $\{y \mid y \in \mathbb{R}, y \neq 0\}$

Example 2.14

Find the domain and range of the function $f(x) = \frac{1}{\sqrt{9 - x^2}}$

Solution

The graph of $y = \frac{1}{\sqrt{9-x^2}}$ on a GDC, shown in Figure 2.16, agrees with algebraic analysis indicating that the expression $\frac{1}{\sqrt{9-x^2}}$ will be positive for all x in the interval $-3 < x < 3$. Further analysis and tracing the graph reveals that $f(x)$ has a minimum at $\left(0, \frac{1}{3}\right)$

The graph on the GDC is misleading in that it appears to show that the function has a maximum value of approximately $y \approx 2.8037849$

Can this be correct? A lack of algebraic thinking and over-reliance on a GDC could easily lead to a mistake. The graph abruptly stops its curve upwards because of low screen resolution. Function values should get quite large for values of x a little less than three, because the value of $\sqrt{9 - x^2}$ will be small, making the fraction $\frac{1}{\sqrt{9-x^2}}$ large. Using a GDC to make a table showing values of $f(x)$ for values of x very close to -3 or 3 confirms that as x approaches -3 or 3, y increases without bound, i.e. y goes to $+\infty$. Hence, $f(x)$ has vertical **asymptotes** of $x = -3$ and $x = 3$

This combination of graphical and algebraic analysis leads to the conclusion that the domain of $f(x)$ is $\{x \mid -3 < x < 3\}$, and the range of $f(x)$ is $\left\{y \mid y \geqslant \frac{1}{3}\right\}$

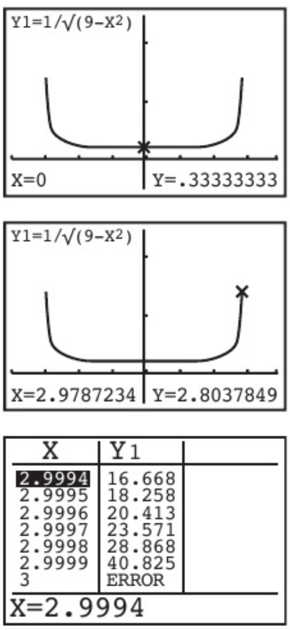


Figure 2.16 GDC screens for the solution to Example 2.14

Adding and subtracting functions

Two functions $f(x)$ and $g(x)$ can be added or subtracted within a common domain.

For example, let $f(x) = x^2 - 3$, and $g(x) = \sqrt{x-2}$

The domain of $f(x)$ is the set of all real numbers, while the domain of $g(x)$ is $x \geqslant 2$. If we add or subtract the functions, we can do that only over the set of numbers larger than or equal to 2 because outside this interval, $g(x)$ does not exist.

Figure 2.17 shows the graphs of f, g and $f + g$. Notice that $f + g$ only exists for $x \geqslant 2$

Let us look at a table of some values.

Notice how the values for $f + g$ only exist when both f and g are defined.

Subtraction of functions works similarly.

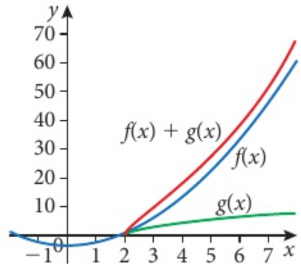


Figure 2.17 The graph of $y = 1 + \sqrt{x-2}$

x	$f(x)$	$g(x)$	$f+g$
-1	-2	None	None
0	-3	None	None
1	-2	None	None
2	1	0	1
3	6	1	7
6	33	2	35

Table 2.2 The values for $f + g$ only exist when both f and g are defined

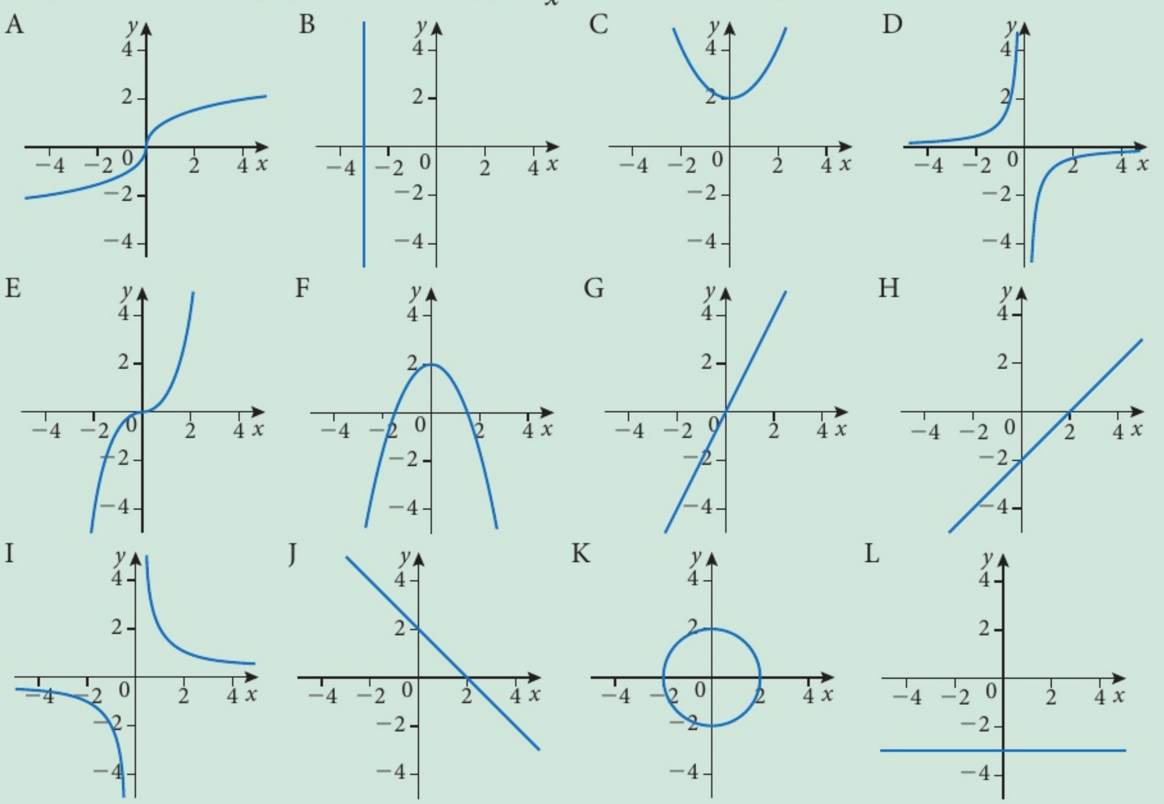
Exercise 2.2

1. **(i)** Match each equation to a graph.

(ii) State whether or not the equation represents a function. Justify your answer. Assume that x is the independent variable and y is the dependent variable.

(a) $y = 2x$ **(b)** $y = -3$ **(c)** $x - y = 2$ **(d)** $x^2 + y^2 = 4$ **(e)** $y = 2 - x$

(f) $y = x^2 + 2$ **(g)** $y^3 = x$ **(h)** $y = \dfrac{2}{x}$ **(i)** $x^2 + y = 2$



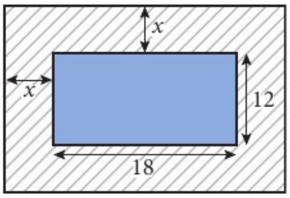


Figure 2.18 Diagram for question 4

2. Express the area, A, of a circle as a function of its circumference, C.

3. Express the area, A, of an equilateral triangle as a function of the length, l, of each of its sides.

4. A rectangular swimming pool with dimensions 12 metres by 18 metres is surrounded by a pavement of uniform width x metres. Find the area of the pavement, A, as a function of x.

5. In a right isosceles triangle, the two equal sides have length x units and the hypotenuse has length h units. Write h as a function of x.

6. The pressure P (measured in kilopascals, kPa) for a particular sample of gas is directly proportional to the temperature T (measured in kelvin, K) and inversely proportional to the volume V measured in litres, L). With k representing the constant of proportionality, this relationship can be written in the form of the equation $P = k\frac{T}{V}$

   (a) Find the constant of proportionality, k, if 150 L of gas exerts a pressure of 23.5 kPa at a temperature of 375 K.

   (b) Using the value of k from part **(a)** and assuming that the temperature is held constant at 375 K, write the volume V as a function of pressure P for this sample of gas.

7. In physics, Hooke's law states that the force F (measured in newtons, N) needed to extend a spring by x units beyond its natural length is directly proportional to the extension x. Assume that the constant of proportionality is k (known as the spring constant for a particular spring).

   (a) Write F as a function of x.

   (b) A spring that has a natural length of 12 cm is stretched by a force of 25 N to a length of 16 cm. Work out the value of the spring constant k.

   (c) What force is needed to stretch the spring to a length of 18 cm?

8. Find the domain of each function.

   (a) $\{(-6.2, -7), (-1.5, -2), (0.7, 0), (3.2, 3), (3.8, 3)\}$

   (b) Surface area of a sphere: $S = 4\pi r^2$

   (c) $f(x) = \frac{2}{5}x - 7$ **(d)** $h: x \mapsto x^2 - 4$

   (e) $g(t) = \sqrt{3 - t}$ **(f)** $h(t) = \sqrt[3]{t}$

   (g) $f: x \mapsto \frac{6}{x^2 - 9}$ **(h)** $f(x) = \sqrt{\frac{1}{x^2} - 1}$

Considering the function in 8 (g), the notation $f: x \mapsto \frac{6}{x^2 - 9}$ is equivalent to $f(x) = \frac{6}{x^2 - 9}$

9. Do all linear equations represent functions? Explain.

10. Consider the function $h(x) = \sqrt{x - 4}$

    (a) Find

    (i) $h(21)$ **(ii)** $h(53)$ **(iii)** $h(4)$

    (b) Find the values of x for which h is undefined.

    (c) State the domain and range of h.

11. For each function
   (i) find the domain and range of the function
   (ii) sketch a comprehensive graph of the function clearly indicating any intercepts or **asymptotes**.

(a) $f: x \mapsto \dfrac{1}{x-5}$ (b) $g(x) = \dfrac{1}{\sqrt{x^2 - 9}}$ (c) $h(x) = \dfrac{2x-1}{x+2}$

(d) $p: x \mapsto \sqrt{5 - 2x^2}$ (e) $f(x) = \dfrac{1}{x} - 4$

2.3 Inverse functions

Pairs of inverse functions

If we choose a number and cube it (raise it to the power of 3), and then take the cube root of the result, the answer is the original number. The same result would occur if we applied the two rules in the reverse order. That is, first take the cube root of a number and then cube the result, again the answer is the original number.

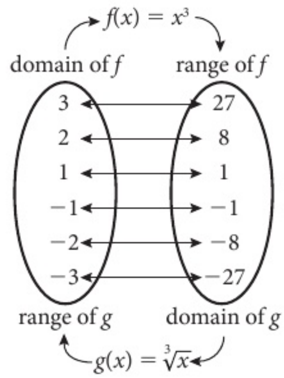


Figure 2.19 A mapping diagram for the cubing and cube root functions

Written as functions using function notation, the cubing function is $f(x) = x^3$, and the cube root function is $g(x) = \sqrt[3]{x}$

If we cube a number x, we get x^3. Then if we take the cube root of the result: $g(x^3) = \sqrt[3]{x^3} = x$

Thus, in essence g undoes what f did.

$f(-2) = -8$ and $g(-8) = -2$

Because function g has this reverse (inverse) effect on function f, we call function g the **inverse** of function f. Function f has the same inverse effect on function g [$g(27) = 3$ and then $f(3) = 27$], making f the inverse function of g. The functions f and g are inverses of each other. The cubing and cube root functions are an example of a pair of **inverse functions**. The mapping diagram for functions f and g (Figure 2.19) illustrates the relationship for a pair of inverse functions where the domain of one is the range for the other.

Do not mistake the -1 in the notation f^{-1} for an exponent. It is *not* an exponent. If a superscript of -1 is applied to the name of a function, as in f^{-1} or $\sin^{-1}$, then it denotes the function that is the inverse of the named function (e.g. f or sin). If a superscript of -1 is applied to an expression, as in 7^{-1} or $(2x+5)^{-1}$, then it is an exponent and denotes the reciprocal of the expression.

You should already be familiar with pairs of **inverse operations**. Addition and subtraction are inverse operations. For example, the rule of adding six ($x + 6$), and the rule of subtracting six ($x - 6$), undo each other. Accordingly, the functions $f(x) = x + 6$ and $g(x) = x - 6$ are a pair of inverse functions. The notation used to indicate the inverse of function f is f^{-1}. Since function g is the inverse of function f, we can write $f^{-1}(x) = x - 6$

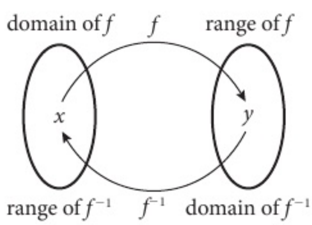


Figure 2.20 A mapping diagram for a pair of inverse functions

For a pair of inverse functions, the domain of f^{-1} is equal to the range of f; and the range of f^{-1} is equal to the domain of f.

It is clear that both $f(x) = x^3$ and $g(x) = \sqrt[3]{x}$ satisfy the definition of a function because for both f and g every number in its domain determines exactly one number in its range. Since they are a pair of inverse functions then the reverse is also true for both; that is, every number in its range is determined by exactly one number in its range. Such a function is called a **one-to-one function**. The phrase one-to-one is appropriate because each value in the domain corresponds to exactly **one** value in the range, and each value in the range corresponds to exactly **one** value in the domain.

A function is **one-to-one** if each element y in the range is the image of exactly one element x in the domain.

The existence of an inverse function

Determining whether a function is one-to-one is very useful because the inverse of a one-to-one function will also be a function. Analysing the graph of a function is the most effective way to determine whether a function is one-to-one. Let's look at the graph of the one-to-one function $f(x) = x^3$ shown in Figure 2.21. It is clear that as the values of x increase over the domain (i.e. from $-\infty$ to ∞) the function values are always increasing. A function that is always increasing, or always decreasing, throughout its domain is one-to-one and has an inverse function.

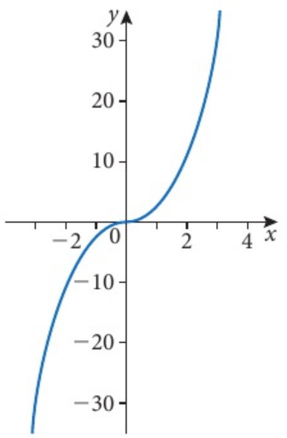


Figure 2.21 Graph of $f(x) = x^3$ which is increasing as x goes from $-\infty$ to ∞

Consider the 'reversing' effect that a pair of inverse functions have on each other. It must be true that if $f(a) = b$ then $f^{-1}(b) = a$. For example, consider again $f(x) = x + 6$ and $f^{-1}(x) = x - 6$

$f(-4) = 2$ and $f^{-1}(2) = -4$

The point $(-4, 2)$ will be on the graph of f, and the point $(2, -4)$ will be on the graph of f^{-1}. Hence, if the ordered pair (a, b) is a point on the graph of $y = f(x)$ then the reversed ordered pair (b, a) must be on the graph of $y = f^{-1}(x)$

Figure 2.22 shows that the point (b, a) can be found by reflecting the point (a, b) about the line $y = x$

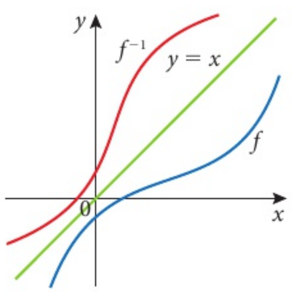


Figure 2.22 The graphs of f and f^{-1} are symmetrical about the line $y = x$

Example 2.15

Consider the function $g(x) = \dfrac{x}{x+4}$ with domain of $\{x \mid x \in \mathbb{R}, x \neq -4\}$

The function that is the inverse of g is $g^{-1}(x) = \dfrac{4x}{1-x}$

(a) Use a GDC to confirm that the domain and range of g are $\{x \mid x \in \mathbb{R}, x \neq -4\}$ and $\{y \mid y \in \mathbb{R}, y \neq 1\}$

(b) Write down the equations of all **asymptotes** for the graph of g.

(c) Write down the domain and range of g^{-1}

(d) The graphs of g and g^{-1} intersect at $(0, 0)$ and at point A. Use a GDC to find the coordinates of A.

(e) Write down the equation of the line that passes through points $(0, 0)$ and A.

Solution

(a) A graph of $g(x) = \dfrac{x}{x+4}$ shows that there is a 'gap' at $x = -4$ and at $y = 1$

(b) The graph of g has a vertical **asymptote** of $x = -4$ and a horizontal **asymptote** of $y = 1$

(c) The domain of g will be the range of g^{-1}, and the range of g will be the domain of g^{-1}

Therefore, the domain of g^{-1} is $\{x \mid x \in \mathbb{R}, x \neq 1\}$ and the range of g^{-1} is $\{y \mid y \in \mathbb{R}, y \neq -4\}$

(d) Using an intersection command in the graph window of a GDC shows that the graphs of g and g^{-1} intersect at $(0, 0)$ and $(-3, -3)$

Thus, the coordinates of A are $(-3, -3)$

(e) The equation of the line that passes through the points of intersection of g and g^{-1} is $y = x$

This is the line for which the graphs of g and g^{-1} are mirror images of each other.

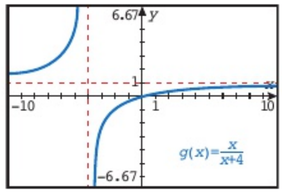


Figure 2.23 GDC screen for the solution to Example 2.15(a)

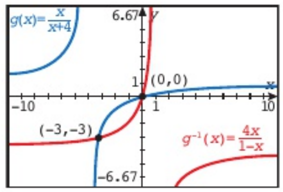


Figure 2.24 GDC screen for the solution to Example 2.15(d)

Exercise 2.3

In questions 1–4, assume that f is a one-to-one function.

1. **(a)** If $f(2) = -5$, then what is $f^{-1}(-5)$?
(b) If $f^{-1}(6) = 10$, then what is $f(10)$?

2. **(a)** If $f(-1) = 13$, then what is $f^{-1}(13)$?
(b) If $f^{-1}(b) = a$, then what is $f(a)$?

3. If $g(x) = 3x - 7$, then what is $g^{-1}(5)$?

4. If $h(x) = x^2 - 8x$, with $x \geqslant 4$, then what is $h^{-1}(-12)$?

5. For each pair of functions, show graphically that f and g are inverse functions by sketching the graphs of f and g on the same set of axes with equal scales on the x-axis and y-axis. Use a GDC to assist in making your sketches on paper.

(a) $f: x \mapsto x + 6$; $g: x \mapsto x - 6$
(b) $f: x \mapsto 4x$; $g: x \mapsto \dfrac{x}{4}$
(c) $f: x \mapsto 3x + 9$; $g: x \mapsto \dfrac{1}{3}x - 3$
(d) $f: x \mapsto \dfrac{1}{x}$, $g: x \mapsto \dfrac{1}{x}$
(e) $f: x \mapsto x^2 - 2, x \geqslant 0$; $g: x \mapsto \sqrt{x+2}, x \geqslant -2$
(f) $f: x \mapsto 5 - 7x$, $g: x \mapsto \dfrac{5-x}{7}$
(g) $f: x \mapsto \dfrac{1}{1+x}$, $g: x \mapsto \dfrac{1-x}{x}$

(h) $f{:}x \mapsto (6 - x)^{\frac{1}{2}}$; $g{:}x \mapsto 6 - x^2, x \geqslant 0$

(i) $f{:}x \mapsto x^2 - 2x + 3, x \geqslant 1$; $g{:}x \mapsto 1 + \sqrt{x - 2}, x \geqslant 2$

(j) $f{:}x \mapsto \sqrt[3]{\dfrac{x + 6}{2}}$; $g{:}x \mapsto 2x^3 - 6$

6. Consider the linear function $h(x) = 2x - 5$
   (a) The graph of h intersects the line $y = x$ at point P. Find the coordinates of P.
   (b) Point Q lies on the graph of h. The x-coordinate of Q is 1. Find the y-coordinate of Q.
   (c) Hence, write down the coordinates of two points that the inverse of h must pass through.
   (d) Find the equation for the inverse of h.

7. Consider the pair of inverse functions given by
   $f(x) = \sqrt{x + 3}, x \geqslant -3$ and $f^{-1}(x) = x^2 - 3, x \geqslant 0$
   Find the coordinates of the point where the graphs of f and f^{-1} meet. State the coordinates to an accuracy of 3 significant figures.

Chapter 2 practice questions

1. The equation of line L_1 is $y = 2.5x + k$. Point $A(3, -2)$ lies on L_1.
   (a) Find the value of k.
   The line L_2 is perpendicular to L_1 and intersects L_1 at point A.
   (b) Write down the gradient of L_2.
   (c) Find the equation of L_2. Give your answer in the form $y = mx + c$
   (d) Write your answer to part (c) in the form $ax + by + d = 0$ where a, b and $d \in \mathbb{Z}$

2. The golden ratio, r, was considered by the ancient Greeks to be the perfect ratio between the lengths of two adjacent sides of a rectangle.
   The exact value of r is $\dfrac{1 + \sqrt{5}}{2}$
   (a) Write down the value of r
      (i) correct to 5 significant figures;
      (ii) correct to 2 decimal places.
   Phidias is designing rectangular windows with adjacent sides of length x metres and y metres. The area of each window is $1\ \text{m}^2$
   (b) Write down an equation to describe this information.
   Phidias designs the windows so that the ratio between the longer side, y, and the shorter side, x, is the golden ratio, r.
   (c) Write down an equation in y, x and r to describe this information.
   (d) Find the value of x.

3. Consider the curve $y = 1 + \frac{1}{2x}$, $x \neq 0$
   (a) For this curve, write down
      (i) the value of the x-intercept;
      (ii) the equation of the vertical **asymptote**.
   (b) Sketch the curve over the interval $-2 \leqslant x \leqslant 4$

4. An iron bar is heated. Its length, L, in mm can be modelled by a linear function, $L = mT + c$ where T is the temperature measured in degrees Celsius. At 150 °C the length of the iron bar is 180 mm.
   (a) Write down an equation that shows this information.
   At 210 °C the length of the iron bar is 181.5 mm.
   (b) Write down an equation that shows this second piece of information.
   (c) Hence, find the length of the iron bar at 40 °C.

5. A building company has many rectangular construction sites, of varying widths, along a road. The area, A, of each site is given by the function $A(x) = x(200 - x)$. where x is the width of the site in metres and $20 \leqslant x \leqslant 180$
   (a) Site S has a width of 20 m. Write down the area of S.
   (b) Site T has the same area as site S, but a different width. Find the width of T.
   When the width of the construction site is b metres, the site has a maximum area.
   (c) (i) Write down the value of b.
      (ii) Write down the maximum area.
   The range of $A(x)$ is $m \leqslant A(x) \leqslant n$
   (d) Write down the value of m and of n.

6. A function $f(x) = p \times 2^x + q$ is defined by the mapping diagram shown.
   (a) Find the value of
      (i) p (ii) q
   (b) Write down the value of r.
   (c) Find the value of s.

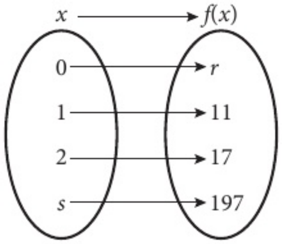


Figure 2.25 Mapping diagram for question 6

7. The straight line, L_1, has equation $2y - 3x = 11$
   The point A has coordinates $(6, 0)$
   (a) Give a reason why L_1 does not pass through A.
   (b) Find the gradient of L_1.
   L_2 is a line perpendicular to L_1. The equation of L_2 is $y = mx + c$
   (c) Write down the value of m.
   L_2 does pass through A.
   (d) Find the value of c.

8. The equation of a line L_1 is $2x + 5y = -4$

The point A has coordinates (6, 0)

(a) Write down the gradient of the line L_1.

A second line L_2 is perpendicular to L_1.

(b) Write down the gradient of L_2.

The point (5, 3) is on L_2.

(c) Determine the equation of L_2.

Lines L_1 and L_2 intersect at point P.

(d) Using a GDC, or otherwise, find the coordinates of P.

9. Water has a lower boiling point at higher altitudes. The relationship between the boiling point of water (T) and the height above sea level (h) can be described by the function

$T(h) = -0.0034h + 100$

where T is measured in degrees Celsius (°C) and h is measured in metres above sea level.

(a) Write down the boiling point of water at sea level.

(b) Use the function $T(h)$ to calculate the boiling point of water at a height of 1.37 km above sea level.

Water boils at the top of Mount Everest at 70°C.

(c) Use the function $T(h)$ to calculate the height above sea level of the top of Mount Everest.

3 Sequences and series

3 Sequences and series

Learning objectives

By the end of this chapter, you should be familiar with…

- arithmetic sequences and series; sums of finite arithmetic sequences; geometric sequences and series; and sums of finite geometric series
- sigma notation as applied to arithmetic and geometric series
- applications such as compound interest, depreciation, and population growth.

The heights of consecutive bounces of a ball in a lab experiment, the balance available after depositing an amount into an account earning interest, population growth, and annual depreciation of goods are only a few of the applications of sequences and series that you may experience. In this chapter you will review these concepts and consolidate your understanding of some specific types abundant in daily practices.

3.1 Sequences

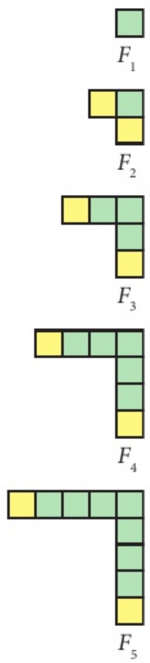


Figure 3.1 The first figure represents 1 unit square

If a_n is the nth term of a sequence, then a_{n-1} is the term before it and a_{n+1} is the term following it. For instance, if a_{10} is the 10th term of a sequence, then a_9 is the term before it and a_{11} is the term following it.

Look at the pattern in Figure 3.1.

The first figure represents 1 unit square, the second represents 3 unit squares, etc. This pattern can be viewed as a function, F, with domain $\{1, 2, 3, \ldots\}$ and range $\{1, 3, 5, 7, \ldots\}$ where:

$F(1) = 1$ 1 is the first number in the sequence.

$F(2) = 3$ 3 is the second number in the sequence.

$F(3) = 5$ 5 is the third number in the sequence.

… and so on…

In mathematics, rather than indicate the terms of the sequence as $F(1)$, $F(2)$, $F(3)$, we usually use subscript notation and write F_1, F_2, F_3, and so on. Thus,

$$F_1 = 1, F_2 = 3, F_3 = 5, \ldots$$

The number F_1 is called the first term, F_2 the second term, and so on. F_n is the nth term. The entire sequence is denoted as $\{F_n\}$.

In a sequence $\{F_n\}$, n is called the **index**. Unless otherwise indicated, we will start our indices at 1. Sometimes, it is more convenient to start at 0. Thus, we write $\{F_n\} = \left\{1, \frac{1}{4}, \frac{1}{9}, \frac{1}{16}, \frac{1}{25}\right\}$ to mean $F_1 = 1, F_2 = \frac{1}{4}, F_3 = \frac{1}{9}, F_4 = \frac{1}{16}, F_5 = \frac{1}{25}$.

The functional values $F_1, F_2, F_3, \ldots$ are called the **terms** of the sequence.

Here are some more examples of sequences:

$6, 12, 18, 24, 30$

$3, 9, 27, \ldots, 3^k, \ldots$

$\left\{\frac{1}{i^2}: i = 1, 2, 3, \ldots, 10\right\} = \left\{1, \frac{1}{4}, \frac{1}{9}, \frac{1}{16}, \cdots, \frac{1}{100}\right\}$

$\{b_1, b_2, \ldots, b_n, \ldots\}$, sometimes used with an abbreviation $\{b_n\}$.

The first and third sequences are **finite** and the second and fourth are **infinite**. Note that in the second and third sequences, we are able to define a rule that yields the nth number in the sequence (called the nth term) as a function of n, the term's number. In this sense you can think of a sequence as a function that assigns a unique number (a_n) to each positive integer n.

Example 3.1

Find the first five terms and the 50th term of the sequence $\{b_n\}$ such that $b_n = 2 - \frac{1}{n}$

Solution

Since we are given an explicit expression for the nth term as a function of its number n, find the value of the function for each term:

$$b_1 = 2 - \frac{1}{1} = 1; b_2 = 2 - \frac{1}{2} = 1\frac{1}{2}; b_3 = 2 - \frac{1}{3} = 1\frac{2}{3}; b_4 = 2 - \frac{1}{4} = 1\frac{3}{4};$$

$$b_5 = 2 - \frac{1}{5} = 1\frac{4}{5};$$

and $b_{50} = 2 - \frac{1}{50} = 1\frac{49}{50}$

So, informally, a sequence is an ordered set of real numbers. That is, there is a first number, a second, and so on. The notation used for these sets is shown in Example 3.1. The way the function was defined in Example 3.1 is called the **explicit definition** of a sequence.

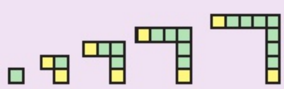
There are other ways to define sequences, one of which is the **recursive** definition. (It is also called **inductive** definition.) Example 3.2 shows you how this is used.

Example 3.2

Find the first 5 terms and the 20th term of the sequence $\{F_n\}$ given at the opening of this section.

Solution

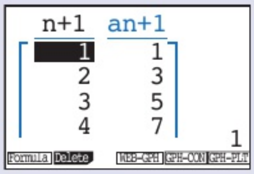
We can look at the sequence in terms of the relationship of neighbouring figures. Each figure is formed by adding two squares to the previous figure. We can organise the information like this:



$F_2 = F_1 + 2 = 3$ $\quad$ $F_3 = F_2 + 2 = 5$

$F_4 = F_3 + 2 = 7$ $\quad$ $F_5 = F_4 + 2 = 9$

You can use your GDC to perform the calculation. Here is an example.



So, the first five terms of this sequence are 1, 3, 5, 7, 9. However, to find the 20th term, we must first find all 19 preceding terms. This is one of the drawbacks of this type of definition, unless we can change the definition into explicit form.

The sequence represented above is a familiar one and it may be simple for you to guess the general form. Its terms are the first positive odd numbers and you can verify that their general term can be $F_n = 2n - 1$. So, in this case $F_{20} = 2(20) - 1 = 39$.

Recursive definition is not required by your IB syllabus. We mention it here for expository purposes only. On exams, you will not be asked questions using this definition.

A spreadsheet can help with calculations.

Column B represents the explicit form and column C represents the recursive form.

	A	B	C
1	*n*	*Fn*	*Fn*
2	1	1	1
3	2	3	3
4	3	5	5
5	4	7	7
⋮	⋮	⋮	⋮
20	19	37	37
21	20	39	39

Let this cell be 2*A2 − 1, then copy it down

Enter 1 into C1, then let this cell be C1 + 2, then copy it down

Example 3.3

A 180 litre bathtub contains 50 litres of water and is being filled at a rate of 10 litres per minute.

(a) How much water is in the bathtub after 5 minutes?

(b) When will the water flow over the top of the bathtub?

(c) The drain is opened when the bathtub contains 50 litres of water. It is still being filled at a rate of 10 litres per minute. The water drains out at a rate of 15 litres per minute. How long will it take until the tub is empty?

Solution

(a) Let w_t be the amount of water in t minutes. We have $w_0 = 50$ to start with. Then the sequence of water quantities is

$$w_1 = 60, w_2 = 70, \cdots, w_5 = 100$$

(b) We can just use simple arithmetic, find an explicit formula and apply it, or use a GDC or spreadsheet.

To overflow, we need to fill $180 - 50 = 130$ litres, which happens after 13 minutes.

If we set up a spreadsheet, as in Table 3.1, then we also get 13 minutes.

(c) In this case, the tub is being drained at a rate of $15 - 10 = 5$ litres per minute. Thus, the tub will be empty in 10 minutes.

w	t	w
50	0	50
60	1	45
70	2	40
80	3	35
90	4	30
100	5	25
110	6	20
120	7	15
130	8	10
140	9	5
150	10	0
160	11	
170	12	
180	13	

Table 3.1 Solution to Example 3.3(b)

Exercise 3.1

1. Find the first five terms of each infinite sequence.

   (a) $s(n) = 2n - 3$ (b) $g(k) = 2^k - 3$

   (c) $f(n) = 3 \times 2^{-n}$ (d) $\begin{cases} a_1 = 5 \\ a_n = a_{n-1} + 3; \quad \text{for } n > 1 \end{cases}$

   (e) $a_n = (-1)^n(2^n) + 3$ (f) $\begin{cases} b_1 = 3 \\ b_n = b_{n-1} + 2n; \quad \text{for } n \geqslant 2 \end{cases}$

2. Find the first five terms and the 50th term of each infinite sequence.

   (a) $a_n = 2n + 3$ (b) $b_n = 2 \times 3^{n-1}$

   (c) $u_n = (-1)^{n-1}\dfrac{2n}{n^2 + 2}$ (d) $a_n = n^{n-1}$

   (e) $a_n = 2a_{n-1} + 5$ and $a_1 = 3$ (f) $u_{n+1} = \dfrac{3}{2u_n + 1}$ and $u_1 = 0$

   (g) $b_n = 3 \cdot b_{n-1}$ and $b_1 = 2$ (h) $a_n = a_{n-1} + 2$ and $a_1 = -1$

3.2 Arithmetic sequences

Not all sequences have formulae. Some sequences are given by listing their terms. There are two types: arithmetic and geometric sequences.

Examine each sequence and the most likely formula for it.

7, 14, 21, 28, 35, 42, …	$a_1 = 7$ and $a_n = a_{n-1} + 7$, for $n > 1$
2, 11, 20, 29, 38, 47, …	$a_1 = 2$ and $a_n = a_{n-1} + 9$, for $n > 1$
48, 39, 30, 21, 12, 3, −6, …	$a_1 = 48$ and $a_n = a_{n-1} - 9$, for $n > 1$

Note that in each case, every term is formed by adding a constant number to the preceding term. Sequences formed in this manner are called **arithmetic sequences**.

For the sequences above, 7 is the common difference for the first, 9 is the common difference for the second and −9 is the common difference for the third.

Using this definition of the arithmetic sequence, it is possible to find the explicit definition of each sequence.

Applying the definition repeatedly will enable you to see the expression we are seeking:

$$a_2 = a_1 + d;$$
$$a_3 = a_2 + d = a_1 + d + d = a_1 + 2d;$$
$$a_4 = a_3 + d = a_1 + 2d + d = a_1 + 3d; \ldots$$

So, you can get to the nth term by adding d to a_1, $(n - 1)$ times.

This result is useful in finding any term of the sequence without knowing the previous terms.

A sequence $a_1, a_2, a_3, \ldots$ is an **arithmetic sequence** if there is a constant d for which $a_n = a_{n-1} + d$ for all integers $n > 1$, where d is called the **common difference** of the sequence, and $d = a_n - a_{n-1}$ for all integers $n > 1$

The **general (nth) term** of an arithmetic sequence, a_n with first term a_1 and common difference d may be expressed explicitly as $a_n = a_1 + (n - 1)d$

The arithmetic sequence can be looked at as a linear function: that is, for every increase of one unit in n, the value of the sequence will increase by d units. As the first term is a_1, the point $(1, a_1)$ belongs to this function. The constant increase d can be considered to be the gradient (slope) of this linear model. So the nth term, the dependent variable in this case, can be found by using the point–slope form of the equation of a line:

$$y - y_1 = m(x - x_1)$$

$$a_n - a_1 = d(n - 1) \Leftrightarrow a_n = a_1 + (n - 1)d$$

This agrees with our definition of an arithmetic sequence.

Example 3.4

Find the nth and the 50th terms of the sequence 2, 11, 20, 29, 38, 47, …

Solution

This is an arithmetic sequence whose first term is 2 and common difference is 9. Therefore,

$$a_n = a_1 + (n - 1)d = 2 + (n - 1) \times 9 = 9n - 7$$

$$\Rightarrow a_{50} = 9 \times 50 - 7 = 443$$

Example 3.5

(a) Find the explicit form of the definition of the sequence 13, 8, 3, -2, …

(b) Calculate the value of the 25th term.

Solution

(a) This is clearly an arithmetic sequence in the form $a_n = a_{n-1} - 5$, because $8 = 13 - 5$, $3 = 8 - 5$, $-2 = 3 - 5$, which indicates that the common difference is -5.

Explicit definition: $a_n = 13 - 5(n - 1) = 18 - 5n$

(b) $a_{25} = 18 - 5 \times 25 = -107$

Example 3.6

Find a definition for the arithmetic sequence whose first term is 5 and fifth term is 11.

Solution

Since the fifth term is given, using the explicit form, we have

$$a_5 = a_1 + (5 - 1)d \Rightarrow 11 = 5 + 4d \Rightarrow d = \frac{3}{2}$$

This leads to the general term

$$a_n = 5 + \frac{3}{2}(n - 1)$$

Example 3.7

Susie decides to deposit an initial amount of €1200 into a savings account on January 1.

The account pays 5% simple interest on the initial amount of money.

Assuming she does not deposit or withdraw any money, how much money will be in the account at the end of 10 years?

Solution

In simple interest, the interest earned every year is only applied to the original amount deposited.

The interest that is earned each year is $1200(0.05) = 60$

One way of finding the final amount is applying a formula for simple interest. To find out the amount, we can set up a table as follows:

Start of year 1	End of year 1	End of year 2	End of year 3	...	End of year 10
1200	$1200 + 60$ $= 1260$	$1260 + 60$ $= 1320$	$1320 + 60$	...	?

This is clearly an arithmetic sequence. However, the number of terms is 11 and not 10 as we may be misled to think. There are 10 year ends plus the beginning of year 1.

Thus, $u_n = u_1 + (n-1)d$, with $u_1 = 1200$, $n = 11$, $d = 10$, so the amount in question is $u_{11} = 1200 + (11-1)60 = €1800$

Example 3.8

Given the sequence, $u_{50} = 6$, $u_{51} = 11$, $u_{52} = 16$, find u_1

Solution

Draw a diagram. $\underset{u_1}{\_\_}, \quad \underset{u_2}{\_\_}, \quad \underset{u_3}{\_\_}, \quad \underset{\ldots}{\ldots}, \quad \underset{u_{50}}{6} \quad \underset{u_{51}}{11} \quad \underset{u_{52}}{16}$

Write an appropriate formula: $u_n = u_1 + (n - 1)d$

That is,

$$u_1 = ?$$

$$d = 11 - 6 = 16 - 11 = 5;\ n = 52;\ u_{52} = 16$$

Substitute these values into the general formula above:

$$16 = u_1 + (52-1) \times 5 = u_1 + 255$$

$$\therefore u_1 = 16 - 255 = -269$$

Exercise 3.2

1. State which of the following are arithmetic sequences.

(a) 10, 20, 30, 40, … **(b)** 3, 3.1, 3.14, 3.141, …

(c) 100, 50, 25, 12.5, … **(d)** 5, −2, −9, −16, …

2. State whether each given sequence is an arithmetic sequence. If it is, find the common difference and the 50th term. If it is not, say why not.

(a) $a_n = 2n - 3$ **(b)** $b_n = n + 2$

(c) $c_n = c_{n-1} + 2$, and $c_1 = -1$ **(d)** $u_n = 3u_{n-1} + 2$

(e) 2, 5, 7, 12, 19, … **(f)** 2, −5, −12, −19, …

3. For each arithmetic sequence find

(i) the 8th term

(ii) an explicit formula for the nth term.

(a) −2, 2, 6, 10, … **(b)** 29, 25, 21, 17, …

(c) −6, 3, 12, 21, … **(d)** 10.07, 9.95, 9.83, 9.71, …

(e) 100, 97, 94, 91, … **(f)** $2, \frac{3}{4}, -\frac{1}{2}, -\frac{7}{4}, \ldots$

4. Find all the missing terms in an arithmetic sequence with 13 as the first term and −23 as the 7th term.

5. Each of the following represents an arithmetic sequence. Find u_1.

(a) ___, ___, ___, ___, 17, ___, ___, 35

(b) ___, ___, ___, ___, ___, ___, −27, ___, ___, ___, ___, ___, ___, ___, −123

(c) $u_{37} = 145$, $u_{85} = 673$

(d) $u_{52} = 70$, $u_{125} = -149$

6. Which term of the arithmetic sequence 7, 3, −1, −5, … is −385?

7. In an arithmetic sequence, $a_5 = 6$ and $a_{14} = 42$

Find an explicit formula for the nth term of this sequence.

8. In an arithmetic sequence, $a_3 = -40$ and $a_9 = -18$

Find an explicit formula for the nth term of this sequence.

9. The first three terms and the last term are given for each sequence. Find the number of terms.

(a) 3, 9, 15, …, 525 **(b)** 9, 3, −3, …, −201

(c) $3\frac{1}{8}, 4\frac{1}{4}, 5\frac{3}{8}, \ldots, 14\frac{3}{8}$ **(d)** $\frac{1}{3}, \frac{1}{2}, \frac{2}{3}, \ldots, 2\frac{5}{6}$

(e) $1 - k, 1 + k, 1 + 3k, \ldots, 1 + 19k$

10. The 30th term of an arithmetic sequence is 147 and the common difference is 4. Find a formula for the nth term.

11. The first term of an arithmetic sequence is -7 and the common difference is 3. Is 9803 a term of this sequence? If so, which one?

12. The first term of an arithmetic sequence is 9689 and the 100th term is 8996.

    (a) Show that the 110th term is 8926.

    (b) Is 1 a term of this sequence? If so, which one?

13. The first term of an arithmetic sequence is 2 and the 30th term is 147. Is 995 a term of this sequence? If so, which one?

14. Find x if each of the following sequences is arithmetic.

    (a) $8 - x, x, x + 8$ (b) $x, 2x + 2, x - 5$

15. Ahmed deposits 2000 euros into an account on January 1 that pays 7% simple interest on the initial deposit. Assuming he does not deposit or withdraw any money, how much will be in the account at the end of

    (a) 5 years (b) 10 years (c) 30 years?

16. An artist decides to make a small sculpture of marbles in the shape of a triangle. She plans to have 45 marbles on the first (bottom) row of the triangle and one less marble in each successive row. How many marbles will be in the 24th row?

3.3 Geometric sequences

Examine the following sequences and the most likely formula for each of them.

$7, 14, 28, 56, 112, 224, \ldots$ $\quad a_1 = 7$ and $a_n = a_{n-1} \times 2$, for $n > 1$

$2, 18, 162, 1458, 13122, \ldots$ $\quad a_1 = 2$ and $a_n = a_{n-1} \times 9$, for $n > 1$

$48, -24, 12, -6, 3, -1.5, \ldots$ $\quad u_1 = 48$ and $u_n = y_{n-1} \times \frac{-1}{2}$, for $n > 1$

Note that in each case, every term is formed by multiplying a constant number with the preceding term. Sequences formed in this manner are called **geometric sequences**.

For the sequences above, 2 is the common ratio for the first, 9 is the common ratio for the second and $-\frac{1}{2}$ is the common ratio for the third.

According to this definition of the geometric sequence, it is possible to find the explicit form of the sequence.

A sequence a_1, a_2, a_3, is a **geometric sequence** if there is a constant r for which $a_n = a_{n-1} \times r$ for all integers $n > 1$ where r is the **common ratio** of the sequence, and $r = a_n \div a_{n-1}$ for all integers $n > 1$.

Applying the definition repeatedly will enable us to see the expression we are seeking:

$$u_2 = u_1 \times r = u_1 r$$
$$u_3 = u_2 r = (u_1 r)r = u_1 r^2$$
$$u_4 = u_3 r = (u_1 r^2)r = u_1 r^3; \ldots$$

The general (nth) term of a geometric sequence, u_n with common ratio r and first term u_1 may be expressed explicitly as $u_n = u_1 r^{(n-1)}$

You can see that you can get to the nth term by multiplying r by u_1, $(n-1)$ times.

This result is useful in finding any term of the sequence without knowing the previous terms.

Example 3.9

(a) Find the geometric sequence with $a_1 = 2$ and $r = 3$

(b) Describe the sequence $3, -12, 48, -192, 768, \ldots$

(c) Describe the sequence $1, \frac{1}{2}, \frac{1}{4}, \frac{1}{8}, \ldots$

(d) Graph the sequence $a_n = \frac{1}{4} \cdot 3^{n-1}$

Solution

(a) The geometric sequence is $2, 6, 18, 54, \ldots, 2 \times 3^{n-1}$. Note that the ratio of any two consecutive terms is 3: $\frac{6}{2} = 3 = \frac{18}{6} = \frac{54}{18} = \cdots$

(b) This is a geometric sequence with $u_1 = 3$ and $r = -4$

The nth term is $u_n = 3 \times (-4)^{n-1}$. Note that when the common ratio is negative, the sign of the terms of the sequence alternates.

(c) The nth term of this sequence is $a_n = 1 \times \left(\frac{1}{2}\right)^{n-1}$

Note that the ratio of any two consecutive terms is $\frac{1}{2}$ and that the terms decrease in value.

(d) Use a GDC to graph the sequence.

Note that the points in the graph lie on the graph of the exponential function $y = \frac{1}{4} \cdot 3^{x-1}$

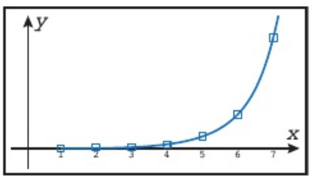


Figure 3.2 GDC screen for the solution to Example 3.9 (d)

Example 3.10

Find the 10th term of the geometric sequence $2, 4, 8, 16, \ldots$

Solution

This is a geometric sequence with $u_1 = 2$ and $r = 2$

Thus the 10th term is $u_{10} = u_1 r^{10-1} = 2 \times 2^9 = 1024$. A GDC can be used in such simple cases too. We just type the first stage 2×2 and then press EXE or ENTER 9 times.

```
2x2
                 4
Ansx2
                 8
Ansx2
                16
```

```
Ansx2
               256
Ansx2
               512
Ansx2
              1024
```

Figure 3.3 GDC screens for the solution to Example 3.10

Example 3.11

At 8:00 a.m., 1000 mg of medicine is administered to a patient. At the end of each hour, the amount of medicine is 60% of that present at the beginning of the hour.

(a) What portion of the medicine remains in the patient's body at 12 noon if no additional medication has been given?

(b) A second dose of 1000 mg is administered at 10:00 a.m. What is the total amount of the medication in the patient's body at 12 noon?

Solution

(a) Use the geometric model, as there is a constant multiple at the end of each hour. Hence, the amount at the end of any hour after administering the medicine is given by:

$a_n = a_1 \times r^{n-1}$, where n is the number of hours

So, at 12 noon $n = 5$ and $a_5 = 1000 \times 0.6^{(5-1)} = 129.6$

(b) For the second dosage, the amount of medicine at noon corresponds to $n = 3$

$a_3 = 1000 \times 0.6^{(3-1)} = 360$

So, the amount of medicine is $129.6 + 360 = 489.6$ mg

Compound interest

Compound interest is an example of a geometric sequence.

Interest compounded annually

When we borrow money, we pay interest, and when we invest money we receive interest. Suppose an amount of €1000 is put into a savings account that has a compound interest rate of 6%. How much money will we have in the bank at the end of 4 years?

It is important to note that, for compound interest, the 6% interest is given annually and is added to the savings account, so that in the following year it will also earn interest, and so on.

Time in years	Amount in the account
0	1000
1	$1000 + 1000 \times 0.06 = 1000(1 + 0.06)$
2	$1000(1 + 0.06) + (1000(1 + 0.06)) \times 0.06$ $= 1000(1 + 0.06)(1 + 0.06) = 1000(1 + 0.06)^2$
3	$1000(1 + 0.06)^2 + (1000(1 + 0.06)^2) \times 0.06$ $= 1000(1 + 0.06)^2(1 + 0.06) = 1000(1 + 0.06)^3$
4	$1000(1 + 0.06)^3 + (1000(1 + 0.06)^3) \times 0.06$ $= 1000(1 + 0.06)^3(1 + 0.06) = \mathbf{1000(1 + 0.06)^4}$

Table 3.2 Compound interest

The amount A_t accumulated after t years with compound interest rate r is:

$A_t = P(1 + r)^t$, where P is the amount invested at the start.

Note that since we are counting from 0 to t, we have $t + 1$ terms, and hence using the geometric sequence formula,

$u_n = u_0 r^{n-1}$
$\Rightarrow A_t = A_0(1 + r)^{(t+1)-1}$
$= A_0(1 + r)^t$

This appears to be a geometric sequence with five terms. The number of terms is five, as both the beginning and the end of the first year are counted. (Initial value, when time = 0, is the first term.)

In general, if a **principal** of €P is invested in an account that yields an interest rate r (expressed as a decimal) annually, and this interest is added at the end of the year, every year to the principal, then we can use the geometric sequence formula to calculate the **future value A,** which is accumulated after t years.

The compound interest formula is developed by repeating the steps in Table 3.3, with

$A_0 = P =$ initial amount

$r =$ annual interest rate

$t =$ number of years

Time in years	Amount in the account
0	$A_0 = P$
1	$A_1 = P + Pr = P(1 + r)$
2	$A_2 = A_1(1 + r) = P(1 + r)(1 + r) = P(1 + r)^2$
⋮	⋮
t	$A_t = P(1 + r)^t$

Table 3.3 Compound interest formula

You do not need to memorise these formulae. You can carry out all compound interest calculations using your GDC. You should familiarise yourself with the financial procedures of your GDC!

Interest compounded n times per year

Suppose that the principal P is invested as before but the interest is paid n times per year. Then $\frac{r}{n}$ is the interest paid every compounding period. Since every year we have n periods, therefore for t years, we have nt periods. The amount A in the account after t years is: $A = P\left(1 + \frac{r}{n}\right)^{nt}$

Example 3.12

€1000 is invested in an account paying compound interest at a rate of 6%. Calculate the amount of money in the account after 10 years when the compounding is

(a) annual (b) quarterly (c) monthly.

Solution

(a) The amount after 10 years is

$$A = 1000(1 + 0.06)^{10} = €1790.85$$

(b) The amount after ten years quarterly compounding is

$$A = 1000\left(1 + \frac{0.06}{4}\right)^{40} = €1814.02$$

(c) The amount after 10 years monthly compounding is

$$A = 1000\left(1 + \frac{0.06}{12}\right)^{120} = €1819.40$$

Our GDC can also give us the same results when we enter the correct values.

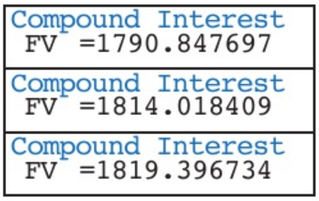


Figure 3.4 GDC screens for the solution to Example 3.12

Example 3.13

You invest €1000 at 6% compounded quarterly.

How long will it take this investment to increase to €2000?

Solution

Let $P = 1000$, $r = 0.06$, $n = 4$, and $A = 2000$ in the compound interest formula

$$A = P\left(1 + \frac{r}{n}\right)^{nt}$$

and then solve for t.

$$2000 = 1000\left(1 + \frac{0.06}{4}\right)^{4t} \Rightarrow 2 = 1.015^{4t}$$

Using a GDC, graph the functions $y = 2$ and $y = 1.015^{4t}$ and then find the intersection between their graphs.

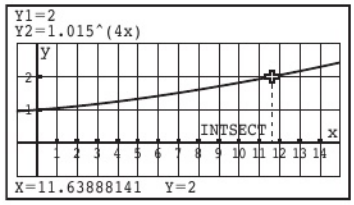


Figure 3.5 GDC screen for the solution to Example 3.13

It will take the €1000 investment 11.64 years to double to €2000. This translates into 47 quarters, as interest is added at the end of each quarter.

We can check our work to see that this is accurate by using the compound interest formula:

$$A = 1000\left(1 + \frac{0.06}{4}\right)^{47} = €2013.28$$

If you are familiar with logarithms, then solving the equation in the example is a straight forward application.

Example 3.14

(a) Samantha puts €15 000 in a bank account earning 6% annual interest compounded monthly. How much interest will she have earned after 20 years?

(b) You want to invest €1000. What interest rate is required to make this investment grow to €2000 in 10 years if interest is compounded quarterly?

Solution

(a) We use the compound interest equation to solve this. Note that we are asked for the interest only. We will need to subtract the principal from the total amount received. We also note that the monthly interest rate is $\frac{6}{12} = \frac{1}{2}$% and the number of compound periods is 240.

$$A = P\left(1 + \frac{r}{n}\right)^{nt} = 15\,000\,(1 + 0.005)^{240} = €49\,653.07$$

Thus, the interest earned $= 49\,653.07 - 15\,000 = €34\,653.07$

(b) Let $P = 1000$, $n = 4$, $t = 10$ and $A = 2000$ in the compound interest formula

$$A = P\left(1 + \frac{r}{n}\right)^{nt}$$

and solve for r:

$$2000 = 1000\left(1+\frac{r}{4}\right)^{40} \Rightarrow 2 = \left(1+\frac{r}{4}\right)^{40} \Rightarrow 1+\frac{r}{4} = \sqrt[40]{2}$$

$$\Rightarrow r = 4\left(\sqrt[40]{2}-1\right) = 0.0699$$

So, at a rate of 7% compounded quarterly, the €1000 investment will grow to at least €2000 in 10 years.

We can check to see whether our work is accurate by using the compound interest formula:

$$A = 1000\left(1+\frac{0.07}{4}\right)^{40} = €2001.60$$

Population growth

The same formulae can be applied when dealing with population growth.

$$u_n = u_1 r^{n-1} \Rightarrow P_t = P_0(1+r)^{t-1}$$

Example 3.15

The city of Baden in Lower Austria grows at an annual rate of 0.35%. The population of Baden in 1981 was 23 140. What is the estimate of the population of this city for 2025?

Solution

This situation can be modelled by a geometric sequence with first term 23 140 and common ratio 1.0035. Since we count the population of 1981 among the terms, the number of terms is 45.

2025 is equivalent to the 45th term in this sequence. The estimated population for Baden is therefore

Population(2025) $= P_{45} = 23\,140(1.0035)^{44} = 26\,985$

Depreciation

When you buy a new car, a new machine to be installed in a factory, or any other items whose value decreases with time, a record is kept to show how the value of the item changes. This is called a depreciation schedule.

Example 3.16

A new car is bought for €28 000 and depreciates in value each year by 17%.

(a) Represent the depreciation as a sequence.

(b) How much is the car worth after 5 years?

(c) Assuming the car is maintained properly, after how many years will it be worth €9000?

(d) Use the graph of this sequence to display and discuss the depreciation pattern.

Solution

(a) The value at the end of the year (at the beginning of a new year) is 83% of its value at the beginning of the year. Thus, if we let v_n represent the value of the car at the end of a year, then v_{n-1} represents the value of the car at the beginning of the year. That is $v_n = 0.83\, v_{n-1}$

So, this is a geometric sequence with first term 28 000 and common ratio 0.83. We will consider the initial value as v_0.

Thus, $v_n = 28\,000 \times 0.83^n$

(b) $v_5 = 28\,000 \times 0.83^5 = 11\,029$

(c) $9000 = 28\,000 \times 0.83^n \Rightarrow 0.83^n = \dfrac{9}{28}$

Using a GDC or software, $n \approx 6$ years.

(d) Depreciation per year decreases with time. Depreciation is faster at the beginning and slows down as time passes.

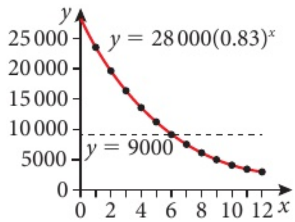


Figure 3.6 Solution to Example 3.16 (d)

Exercise 3.3

1. State which of the following are geometric sequences.

(a) 10, 20, 30, 40, …
(b) 1.5, 3, 6, 12, …
(c) 2, −4, 8, −16, 32, …
(d) 1, 4, 9, 16, 25, …

2. For each sequence

(i) determine whether the sequence is arithmetic, geometric, or neither

(ii) find the common difference for the arithmetic sequences and the common ratio for the geometric sequences

(iii) find the 10th term for each arithmetic or geometric sequence.

(a) $3, 3^{a+1}, 3^{2a+1}, 3^{3a+1}, \ldots$
(b) $a_n = 3n - 3$
(c) $b_n = 2^{n+2}$
(d) $c_n = 2c_{n-1} - 2$, and $c_1 = -1$
(e) $u_n = 3u_{n-1}$ and $u_1 = 4$
(f) 2, 5, 12.5, 31.25, 78.125 …
(g) 2, −5, 12.5, −31.25, 78.125 …
(h) 2, 2.75, 3.5, 4.25, 5, …
(i) $18, -12, 8, -\dfrac{16}{3}, \dfrac{32}{9}, \ldots$
(j) 52, 55, 58, 61, …
(k) −1, 3, −9, 27, −81, …
(l) 0.1, 0.2, 0.4, 0.8, 1.6, 3.2, …
(m) 3, 6, 12, 18, 21, 27, …
(n) 6, 14, 20, 28, 34, …
(o) 2.4, 3.7, 5, 6.3, 7.6, …

3. Find all the missing terms in a geometric sequence whose first term is 3 and fifth term is 96.

4. For each arithmetic or geometric sequence find

(i) the 8th term

(ii) an explicit formula for the nth term.

(a) $-3, 2, 7, 12, \ldots$

(b) $19, 15, 11, 7, \ldots$

(c) $-8, 3, 14, 25, \ldots$

(d) $10.05, 9.95, 9.85, 9.75, \ldots$

(e) $100, 99, 98, 97, \ldots$

(f) $2, \frac{1}{2}, -1, -\frac{5}{2}, \ldots$

(g) $3, 6, 12, 24, \ldots$

(h) $4, 12, 36, 108, \ldots$

(i) $5, -5, 5, -5, \ldots$

(j) $3, -6, 12, -24, \ldots$

(k) $972, -324, 108, -361, \ldots$

(l) $-2, 3, -\frac{9}{2}, \frac{27}{4}, \ldots$

(m) $35, 25, \frac{125}{7}, \frac{625}{49}, \ldots$

(n) $-6, -3, -\frac{3}{2}, -\frac{3}{4}, \ldots$

(o) $9.5, 19, 38, 76, \ldots$

(p) $100, 95, 90.25, \ldots$

(q) $2, \frac{3}{4}, \frac{9}{32}, \frac{27}{256}, \ldots$

5. Find all the missing terms in a geometric sequence whose first term is 7 and fifth term is 4375.

6. Find a such that 16, a, 81 forms a geometric sequence.

7. Find x, y, z, u such that 7, x, y, z, u, 1701 forms a geometric sequence.

8. Find a such that 9, a, 64 forms a geometric sequence.

9. The first term of a geometric sequence is 24 and the fourth term is 3. Find the fifth term and an expression for the nth term.

10. On January 1, 2009 Tara deposited £1500 into an account that offered 6% compounded monthly. If she did not make any further deposits or withdrawals, how much would be in the account at the end of 30 years?

11. Find the future value of ¥5000 that earns 6% annual interest for 15 years compounded

(a) annually **(b)** quarterly **(c)** monthly.

12. Which term of the geometric sequence 6, 18, 54, … is 118 098?

13. The 4th term and the 7th term of a geometric sequence are 18 and $\frac{729}{8}$.

Is $\frac{59\,049}{128}$ a term of this sequence? If so, which term is it?

14. The 3rd term and the 6th term of a geometric sequence are 18 and $\frac{243}{4}$.

Is $\frac{19\,683}{64}$ a term of this sequence? If so, which term is it?

15. Jim put €1500 into a saving account that pays 4% interest compounded semi-annually. How much will his account hold 10 years later assuming that he does not make any additional investments in this account?

16. In 7 years, Ricardo has earned £1300 interest on an initial investment of £3500. He is receiving interest compounded monthly.

(a) What is the annual interest rate of his investment?

(b) How much longer will it take Ricardo to at least double his investment?

17. How much money should you invest now if you wish to have an amount of €4000 in your account after 6 years when interest is compounded quarterly at an annual rate of 5%?

18. In 2017, the population of Switzerland was estimated to be 8 476 000. What will the Swiss population be in 2022 if it grows at a rate of 0.5% annually?

19. Each of the following represents a geometric sequence. Find u_1.

(a) ____, ____, ____, 135, ____, ____, 3645

(b) ____, ____, $\frac{4}{9}$, ____, ____, ____, ____, $\frac{128}{2187}$

(c) ____, ____, ____, ____, 1024, ____, 16 384

(d) ____, ____, ____, $\frac{1}{16}$, ____, $\frac{1}{64}$

20. On January 1, 2009, Mohammed deposited £1500 into an account paying 6% interest compounded annually. If he did not make any further deposits or withdrawals, how much would be in the account at the end of 2039?

21. Tim put €2500 into a saving account that pays 4% interest compounded semi-annually. How much will his account hold 10 years later assuming that he does not make any additional investments in this account?

22. Jane put £1000 into a savings account at her son William's birth. The interest she earned is 6% compounded quarterly. How much money will William have on his 18th birthday?

23. Praneel invests £18 000 in a bank account that pays 3.4% nominal annual interest, compounded quarterly.

Find the minimum number of years that Praneel must invest the money for his investment to be worth £27 000.

A nominal interest rate does not take inflation into account.

24. Sharlene buys a used car for $24 000. It is known that the value of such cars decreases at the rate of 10% per year. What will be the value of the car at the end of 5 years?

25. Lateasia buys a house in South Africa for 200 000 SAR. The value of houses in South Africa is increasing at the rate of 5% per year. What will be the value of the house in 10 years?

26. Ali, Bob and Connie each have 3000 USD (US dollar) to invest.

    Ali invests his 3000 USD in a firm that offers simple interest at 4.5% per annum. The interest is added at the end of each year.

    Bob invests his 3000 USD in a bank that offers interest compounded annually at a rate of 4% per annum. The interest is added at the end of each year.

    Connie invests her 3000 USD in another bank that offers interest compounded half-yearly at a rate of 3.8% per annum. The interest is added at the end of each half-year.

    (a) Calculate how much money Ali and Bob have at the beginning of year 7.

    (b) Show that Connie has 3760.20 USD at the beginning of year 7.

    (c) Calculate how many years it will take for Bob to have at least 6000 USD in the bank.

3.4 Series

The word series in common language implies much the same thing as sequence. But in mathematics when we talk of a series, we are referring in particular to sums of terms in a sequence.

Suppose a civil engineering graduate has a starting annual salary of £28 000 and receives a £1500 pay rise each year. Then,

$$28\,000, 29\,500, 31\,000, 32\,500, 34\,000$$

are terms of a sequence that describe this engineer's salaries over a 5-year period. The total earned is given by the **series**

$$28\,000 + 29\,500 + 31\,000 + 32\,500 + 34\,000 = £155\,000$$

A sequence can be used to define a series. For example, the infinite sequence

$$1, \frac{1}{2}, \frac{1}{4}, \frac{1}{8}, \frac{1}{16}, \ldots$$

defines the terms of the infinite series

$$1 + \frac{1}{2} + \frac{1}{4} + \frac{1}{8} + \frac{1}{16} + \cdots$$

The sum s_n of a sequence of values $a_1, a_2, a_3 \ldots$, is defined as the sum of the first n terms, or the **partial sum**. That is

$$s_n = a_1 + a_2 + \ldots + a_{n-1} + a_n$$

The sum of the terms of a sequence is called a series and is written using summation notation. The Greek letter sigma Σ is used to indicate a sum.

If the terms are in an arithmetic sequence, then the sum is an **arithmetic series**. If the terms are in a geometric sequence, then the sum is a **geometric series**.

Sigma notation

Any sum in a series s_k will be called a partial sum and is given by

$$s_k = a_1 + a_2 + \dots + a_{k-1} + a_k$$

This partial sum can be written using the sigma notation

$$s_k = \sum_{i=1}^{i=k} a_i = a_1 + a_2 + \dots + a_{k-1} + a_k$$

$\sum_{i=1}^{i=k} a_i$ can also be written as $\sum_{i=1}^{k} a_i$ or $\sum_{1}^{k} a_i$

Sigma notation is a concise and convenient way to represent long sums. The Greek capital letter sigma refers to the initial letter of the word sum. So, this expression means the sum of all the terms a_i where i, called the **index of summation**, takes the values from 1 to k. We can also write $\sum_{i=m}^{n} a_i$ to mean the sum of the terms a_i where i takes the values from m to n. In such a sum, m is called the lower limit and n the upper limit.

This indicates ending with $i = n$
This indicates addition $\rightarrow \sum_{i=m}^{n} a_i$
This indicates starting with $i = m$

For example, suppose we measure the heights of 6 children. Denote their heights by x_1, x_2, x_3, x_4, x_5 and x_6

The sum of their heights $x_1 + x_2 + x_3 + x_4 + x_5 + x_6$ is written more neatly as $\sum_{i=1}^{6} x_i$

Underneath Σ there is $i = 1$ and on top of it 6. This means that i is replaced by whole numbers starting at the bottom number, 1, until the top number, 6, is reached.

Thus $\sum_{i=3}^{6} x_i = x_3 + x_4 + x_5 + x_6$ and $\sum_{i=3}^{5} x_i = x_3 + x_4 + x_5$.

So, the notation $\sum_{i=1}^{n} x_i$ tells us:

- to add the scores x_i
- where to start: x_1
- where to stop: x_n

  (where n is some number greater than or equal to one; $\sum_{i=1}^{1} x_i$ is nothing but x_1)

Now take the heights of the children to be $x_1 = 112$ cm, $x_2 = 96$ cm, $x_3 = 120$ cm, $x_4 = 132$ cm, $x_5 = 106$ cm, and $x_6 = 120$ cm

Then the total height (in cm) is

$$\begin{aligned}\sum_{k=1}^{6} x_k &= x_1 + x_2 + x_3 + x_4 + x_5 + x_6 \\ &= 112 + 96 + 120 + 132 + 106 + 120 \\ &= 686\text{ cm}\end{aligned}$$

Note that we have used k instead of i in the formula above. i is a dummy variable: any letter can be used.

$$\sum_{k=1}^{n} x_k = \sum_{i=1}^{n} x_i$$

Example 3.17

Write each series in full.

(a) $\sum_{i=1}^{5} i^4$ (b) $\sum_{r=3}^{7} 3^r$ (c) $\sum_{j=1}^{n} x_j p(x_j)$

Solution

(a) $\sum_{i=1}^{5} i^4 = 1^4 + 2^4 + 3^4 + 4^4 + 5^4$

(b) $\sum_{r=3}^{7} 3^r = 3^3 + 3^4 + 3^5 + 3^6 + 3^7$

(c) $\sum_{j=1}^{n} x_j p(x_j) = x_1 p(x_1) + x_2 p(x_2) + \ldots + x_n p(x_n)$

Example 3.18

Evaluate $\sum_{n=0}^{5} 2^n$

Solution

$\sum_{n=0}^{5} 2^n = 2^0 + 2^1 + 2^2 + 2^3 + 2^4 + 2^5 = 63$

Example 3.19

Write the sum $\frac{1}{2} + \frac{2}{3} + \frac{3}{4} + \frac{4}{5} + \ldots + \frac{99}{100}$ in sigma notation.

Solution

The numerator and denominator in each term are consecutive integers, so they take on the value of $\frac{k}{k+1}$ or any equivalent form. We can therefore write the sum as $\sum_{k=1}^{99} \frac{k}{k+1}$

Properties of sigma notation

There are a number of useful results that we can obtain when we use sigma notation.

For example, suppose we have a sum of constant terms $\sum_{i=1}^{5} 2$

What does this mean? If we write this out in full, we get

$$\sum_{i=1}^{5} 2 = 2 + 2 + 2 + 2 + 2 = 5 \times 2 = 10$$

In general, if we sum a constant, k, n times then we can write

$$\sum_{i=1}^{n} k = k + k + \ldots + k = n \times k = nk$$

Suppose we have the sum of a constant times i. For example,

$$\begin{aligned}\sum_{i=1}^{5} 3i &= 3 \times 1 + 3 \times 2 + 3 \times 3 + 3 \times 4 + 3 \times 5 \\ &= 3 \times (1 + 2 + 3 + 4 + 5) = 45\end{aligned}$$

However, this can also be interpreted as

$$\sum_{i=1}^{5} 3i = 3 \times (1 + 2 + 3 + 4 + 5) = 3\sum_{i=1}^{5} i$$

which implies that $\sum_{i=1}^{5} 3i = 3\sum_{i=1}^{5} i$

In general, we can say $\sum_{i=1}^{n} ki = k\sum_{i=1}^{n} i$

The property $\sum_{i=1}^{n} ki = k\sum_{i=1}^{n} i$ tells us that the constant k can be moved 'in' and 'out' of the summation without changing the value of the sum.

Suppose that we need to consider the summation of two different functions, such as

$$\begin{aligned}\sum_{k=1}^{n}(k^2 + k^3) &= (1^2 + 1^3) + (2^2 + 2^3) + \ldots + (n^2 + n^3) \\ &= (1^2 + 2^2 + \ldots + n^2) + (1^3 + 2^3 + \ldots + n^3) \\ &= \sum_{k=1}^{n}(k^2) + \sum_{k=1}^{n}(k^3)\end{aligned}$$

In general, $\sum_{k=1}^{n}(f(k) + g(k)) = \sum_{k=1}^{n} f(k) + \sum_{k=1}^{n} g(k)$

At times it is convenient to change indices. For example, note how all three summations are the same:

$$\sum_{i=1}^{k} a_i = a_1 + a_2 + \ldots + a_{k-1} + a_k$$

$$\sum_{i=0}^{k-1} a_{i+1} = a_1 + a_2 + \ldots + a_{k-1} + a_k$$

$$\sum_{i=2}^{k+1} a_{i-1} = a_1 + a_2 + \ldots + a_{k-1} + a_k$$

Arithmetic series

In arithmetic series, we add the terms of arithmetic sequences. It is very helpful to be able to find an easy expression for the partial sums of these series.

For example, find the partial sum for the first 50 terms of the series

$$3 + 8 + 13 + 18 + \ldots$$

It is a tedious task to add the 50 terms individually.

Let $\{u_n\}$ be an arithmetic sequence with first term u_1 and a common difference d. Then,

$$u_n = u_1 + (n - 1)d$$

The sum of the first n terms of an arithmetic series is given by $S_n = \frac{n}{2}(u_1 + u_n)$

$1 + 4 + 7 + 10 + 13 + 16$

17

17

17

Figure 3.7 The terms can be arranged in three pairs

Take for example, the sum of the arithmetic sequence 1, 4, 7, 10, 13, 16.

There are 6 terms arranged into three pairs as shown in Figure 3.7. Thus, the sum is $S_6 = 3 \times 17$ which is simply what the formula above says: $S_6 = \frac{6}{2}(u_1 + u_6)$

$S_n = \frac{n}{2}(u_1 + u_n)$ can be changed to give an interesting perspective of the sum, i.e., $S_n = n\left(\frac{u_1 + u_n}{2}\right)$ which is n times the average of the first and last terms!

Moreover, if we substitute $u_1 + (n - 1)d$ for u_n then we get an alternative formula for the sum $S_n = \frac{n}{2}(u_1 + u_1 + (n - 1)d) = \frac{n}{2}(2u_1 + (n - 1)d)$

The partial sum S_n of an arithmetic series is given by one of the following:

$S_n = \frac{n}{2}(u_1 + u_n)$, or

$S_n = n\left(\frac{u_1 + u_n}{2}\right)$, or

$S_n = \frac{n}{2}(2u_1 + (n - 1)d)$

Example 3.20

Consider the following arithmetic series (50 terms)

$$3 + 8 + 13 + 18 + \ldots$$

(a) Write the series using sigma notation.

(b) Find the partial sum for the first 50 terms of the series.

Solution

(a) The first term is 3, and the common difference is 5. Thus, the nth term of the corresponding sequence is $u_n = u_1 + (n - 1)d = 3 + 5(n - 1) = 5n - 2$

This implies that $S_{50} = \sum_{1}^{50}(5n - 2)$

(b) Using the second formula for the sum we get:

$$S_{50} = \frac{50}{2}(2 \times 3 + (50 - 1)5) = 25 \times 251 = 6275$$

If we were to use the first formula, then we need to know the nth term. So, $u_{50} = 3 + 49 \times 5 = 248$ which can now be used.

$S_{50} = 25(3 + 248) = 6275$

Example 3.21

Consider the sequence in Figure 3.8.

(a) F_1 has 1 square. How many squares are in F_{20}?

(b) Is there a figure with 149 squares? If yes, which one and if not, why not?

(c) Find the number of squares in the first 880 figures.

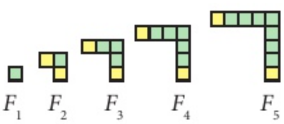


Figure 3.8 Diagram for Example 3.21

(d) F_1 has 4 line segments. How many line segments are in F_{20}?

(e) Is there a figure with 450 segments? If yes, which one and if not, why not?

(f) Find the number of segments in the first 880 figures.

Solution

(a) In every new figure, 2 squares are added. The sequence is an arithmetic sequence with first term 1, and a common difference of 2. Using the nth term form:

$$F_{20} = 1 + 2(20 - 1) = 39$$

(b) The term whose value is 149 satisfies the nth term form:

$$149 = 1 + 2(n - 1), \text{ thus } n = \frac{149 - 1}{2} + 1 = 75$$

Therefore, the 75th figure has 149 squares.

(c) Use one of the formulae for the arithmetic series:

$$S_{880} = \frac{880}{2}(1 + 1759) = 774\,400 \text{ or}$$

$$S_{880} = \frac{880}{2}(2 \times 1 + 2(880 - 1)) = 774\,400$$

(d) In every new figure, 6 line segments are added. The sequence is an arithmetic sequence with first term 4, and a common difference of 6. Using the nth term form

$$F_{20} = 4 + 6(20 - 1) = 118$$

(e) The term whose value is 450 must satisfy the nth term form

$$450 = 4 + 6(n - 1), \text{ thus } n = \frac{450 - 4}{6} + 1 = 75.33$$

This is not a whole number, and therefore, there is no figure with 450 segments.

(f) Use one of the formulae for the arithmetic series:

$$S_{880} = \frac{880}{2}(4 + 5278) = 2\,324\,080 \text{ or}$$

$$S_{880} = \frac{880}{2}(2 \times 4 + 6(880 - 1)) = 2\,324\,080$$

Example 3.22

An auditorium has 16 rows of seats with 28 seats in the first row, 36 seats in the second, 44 in the third and so on. Find the total number of seats.

Solution

The number of seats in this auditorium forms an arithmetic sequence with first term 28, and a common difference of 8. Thus, the total number of seats is $S_{16} = \frac{16}{2}(2 \times 28 + (16 - 1) \times 8) = 1408$

Thus, the auditorium has 1408 seats.

Geometric series

As is the case with the arithmetic series, we can find a general expression for the nth partial sum of a geometric series.

For example, find the partial sum for the first 20 terms of the series

$3 + 6 + 12 + 24 + \dots$

Let $\{a_n\}$ be a geometric sequence with first term a_1 and a common ratio $r \neq 1$. Then, the general term of a geometric sequence is:

$a_n = a_1 \times r^{n-1}$

The sum of the first n terms of a geometric series is given by $S_n = \dfrac{a_1 - r a_n}{1 - r}$; $r \neq 1$

If we substitute $a_1 \times r^{n-1}$ for a_n then we get an alternative formula for the sum

$$S_n = \frac{a_1 - ra_n}{1 - r} = \frac{a_1 - ra_1r^{n-1}}{1 - r} = \frac{a_1(1 - r^n)}{1 - r}; r \neq 1$$

When $r = 1$, then the partial sum S_n is given by

$S_n = a_1 + a_1 + \dots + a_1$
$= na_1$

The sum, S_n, of n terms of a geometric series with common ratio r ($r \neq 1$) and first term a_1 is:

$$S_n = \frac{a_1(1 - r^n)}{1 - r} \quad \left[\text{equivalent to } S_n = \frac{a_1(r^n - 1)}{r - 1}\right]$$

Example 3.23

Consider the partial sum for the first 20 terms of this series.

$3 + 6 + 12 + 24 + \dots.$

(a) Write the partial sum in sigma notation.

(b) Find the sum.

Solution

(a) The first term is 3 and the common ratio is 2.

The nth term of the corresponding sequence is $a_n = 3 \times 2^{n-1}$

Thus, the sum is $\sum_{1}^{20} 3 \times 2^{n-1}$

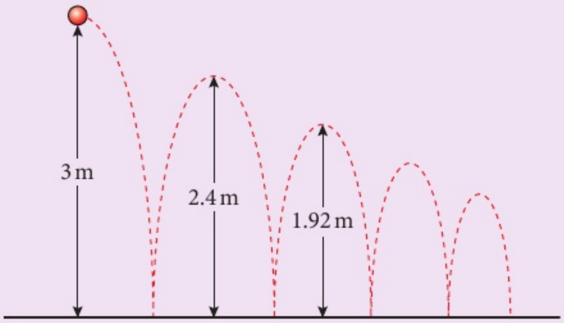
(b) We apply a formula for the sum as above

$$S_{20} = \frac{3(1 - 2^{20})}{1 - 2} = \frac{3(1 - 1\,048\,576)}{-1} = 3\,145\,725$$

Example 3.24

A ball has elasticity such that it bounces up 80% of its previous height. Find the total vertical distance travelled down and up by this ball when it is dropped from a height of 3 metres and after it bounces for the 10th time. Ignore friction and air resistance.

Solution



After the ball is dropped the initial 3 m, it bounces up and down a distance of 2×2.4 m. Each bounce after the first bounce, the ball travels 0.8 times the previous height twice – once upwards and once downwards. So, the total vertical distance, h is given by

$$h = 3 + 2(2.4 + (2.4 \times 0.8) + (2.4 \times 0.8^2) + \ldots) = 3 + 2 \times l$$

The amount l is a geometric series with $a_1 = 2.4$ and $r = 0.8$

The value of l is $l = \dfrac{2.4(1 - 0.8^{10})}{1 - 0.8} = 11.463$

Hence the total distance required is

$$h = 3 + 2(11.463) = 25.93 \text{ m}$$

Applications of series to compound interest calculations

Annuities

An **annuity** is a sequence of equal periodic payments. If you are saving money by depositing the same amount at the end of each compounding period, the annuity is called an **ordinary annuity**. Using geometric series, you can calculate the **future value (FV)** of this annuity, which is the amount of money you have after making the last payment.

You invest €1000 at the end of each year for 10 years at a fixed annual interest rate of 6%

Year	Amount invested	Future value
10	1000	1000
9	1000	$1000(1 + 0.06)$
8	1000	$1000(1 + 0.06)^2$
⋮	⋮	⋮
1	1000	$1000(1 + 0.06)^9$

Table 3.4 Calculating the future value

Notice how we filled the table in reverse order. We have the 10th year first. This is the last payment you make. It does not stay in the account to earn interest, so, its future value is €1000. Year 9's payment would have stayed 1 year, so, it earns interest for 1 year. Similarly, year 8's payment, earns interest for 2 years, and so on. Finally, the first payment earns interest for 9 years.

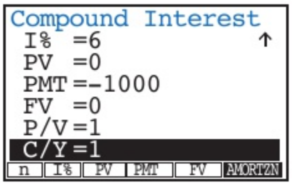


Figure 3.9 Using a GDC to calculate compound interest

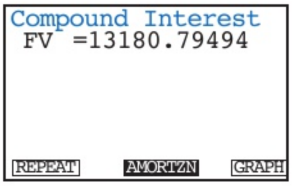


Figure 3.10 Using a GDC to calculate compound interest

The future value of this investment is the sum of all the entries in the third column, so it is

$$FV = 1000 + 1000(1 + 0.06) + 1000(1 + 0.06)^2 + \ldots + 1000(1 + 0.06)^9$$

This sum is a partial sum of a geometric series with $n = 10$ and $r = 1 + 0.06$

Hence

$$FV = \frac{1000(1 - (1 + 0.06)^{10})}{1 - (1 + 0.06)} = \frac{1000(1 - (1 + 0.06)^{10})}{-0.06} = 13\,180.79$$

The calculation above is meant to show you how the money accumulates. You will not be asked to calculate the future value manually. Instead, you can use a GDC to get this result.

We can generalise the previous formula in the same manner.

Let the periodic payment be R, and the periodic interest rate be i, i.e., $i = \frac{r}{n}$

Let the number of periodic payments be m.

The future value of this investment is the sum of all the entries in the last column, so it is

Period	Amount invested	Future value
m	R	R
$m - 1$	R	$R(1 + i)$
$m - 2$	R	$R(1 + i)^2$
⋮	⋮	⋮
1	R	$R(1 + i)^{m-1}$

Table 3.5 Formula for calculating the future value

$$FV = R + R(1 + i) + R(1 + i)^2 + \ldots + R(1 + i)^{m-1}$$

This sum is a partial sum of a geometric series with m terms and $r = 1 + i$

Hence

$$FV = \frac{R(1 - (1 + i)^m)}{1 - (1 + i)} = \frac{R(1 - (1 + i)^m)}{-i} = R\left(\frac{(1 + i)^m - 1}{i}\right)$$

You need to familiarise yourself with the financial menu of your GDC in order to work with annuities. The IB syllabus does not require that you use the formula. We include it here for information purposes only and for those of you who may be interested in working with the formula.

If the payment is made at the beginning of the period rather than at the end, then the annuity is called **annuity due** and the future value after m periods will be slightly different.

The future value of this investment is the sum of all the entries in the last column, so it is

$$FV = R(1 + i) + R(1 + i)^2 + \ldots + R(1 + i)^{m-1} + R(1 + i)^m$$

This sum is a partial sum of a geometric series with m terms and $r = 1 + i$. Hence

$$FV = R\left(\frac{(1 + i)^{m+1} - 1}{i} - 1\right)$$

If the previous investment is made at the beginning of the year rather than at the end, then in 10 years we have

$$FV = R\left(\frac{(1 + i)^{m+1} - 1}{i} - 1\right) = 1000\left(\frac{(1 + 0.06)^{10+1} - 1}{0.06} - 1\right) = 13\,971.64$$

Using your GDC, you enter $n = 11$ and then subtract 1000 from the result, or go to 'set up' and choose 'BEGIN' instead of 'END'.

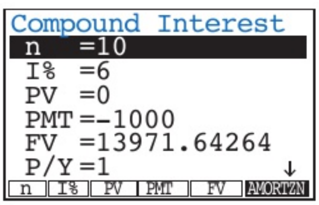


Figure 3.11 Annuity due calculation

Example 3.25

In order to save enough funds for her college education, Sophia's parents set up an annuity where they deposit €100 at the end of every month in an account which pays 6% annual interest. They started this the year she was born. How much money will be in the account on her 18th birthday?

Solution

The monthly interest rate is $\frac{0.06}{12} = 0.005$

Using the formula for the annuity:

$$FV = R\left(\frac{(1+i)^m - 1}{i}\right) = 100\left(\frac{(1+0.005)^{18\times 12} - 1}{0.005}\right) = 38\,735.32$$

We can also use a GDC or a spreadsheet, as shown in Figures 3.12 and 3.13.

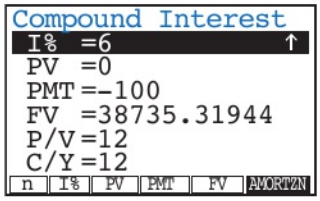


Figure 3.12 GDC screen for the solution to Example 3.25

A	B	C
n	Payment	Balance
1	100	100
2	100	200.5
3	100	301.5025
4	100	403.01
⋮	⋮	⋮
216	100	38735.32

Let C2 = B2 + C1*1.005, then copy it down

Figure 3.13 Spreadsheet for the solution to Example 3.25

Amortisation of a loan

When you borrow an amount of money, you usually promise to pay it back in monthly instalments. The process is called amortisation of loans. In this process, you need to know the periodic interest rate, i, the number of periods, n, and the amount of the loan to be paid, P. The periodic payment, R, required to pay the loan back on time is given by $R = \frac{i \times P}{1 - (1+i)^{-n}}$

However, your GDC can calculate this for you.

Example 3.26

You borrow €1000 at an interest rate of 8% and promise to pay it back within 6 months with monthly payments. Find the monthly payment required and set up a table showing the progress of the payment.

Solution

The monthly payment is

$$R = \frac{\frac{0.08}{12} \times 1000}{1 - \left(1 + \frac{0.08}{12}\right)^{-6}} = 170.58$$

Our GDC will give us the amount to be paid, as shown in Figure 3.14.

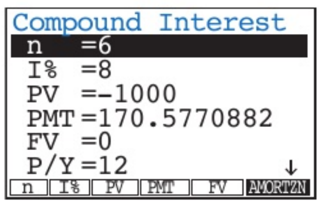


Figure 3.14 GDC screen for the solution to Example 3.26

For as long as you owe the money, interest will be charged. A schedule can show how you will end up paying the loan back; setting up a spreadsheet is an efficient way of doing this.

Period	Balance	Interest	Payment	Payment to principal
1	1000.00	6.67	170.58	163.91
2	836.09	5.57	170.58	165.00
3	671.09	4.47	170.58	166.10
4	504.98	3.37	170.58	167.21
5	337.77	2.25	170.58	168.33
6	169.45	1.13	170.58	169.45

Let B3 = B2 − E2

Let C2 = B2*(0.08/12)

Let E2 = D2 − C2

- Balance is the amount at the beginning of each period.
- Interest is the amount of interest on the amount owed for that period.
- Payment is the amount deposited.
- Payment to principal is the amount that goes into paying back the loan after deducting interest from it.

Note how the last cells in Balance and Payment to principal are the same, 169.45, indicating that the last payment covered the rest of the loan. If we were to continue the table beyond the sixth row, the cell in Balance would be zero because there is no money left to be paid.

In general, loan departments at financial institutions provide their employees with tables, similar to Table 3.6, where they list different possible monthly repayments according to the length of the loan term and the interest rate.

Loan term (years)	Monthly repayment in $, for every $10 000			
	Nominal annual interest rate			
	3%	5%	6%	7%
1	846.94	856.07	860.66	865.27
5	179.69	188.71	193.33	198.01
10	96.56	106.07	111.02	116.11
15	69.06	**79.08**	84.39	89.88
20	55.46	66.00	71.64	77.53
25	47.42	58.46	64.43	70.68

Table 3.6 Possible monthly repayments according to the length of the loan term and the interest rate

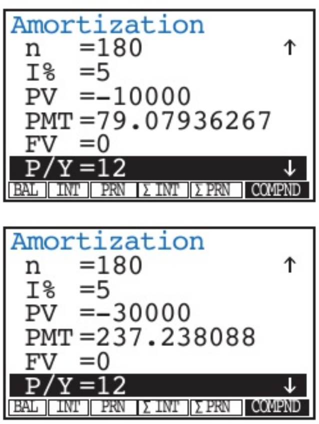


Figure 3.15 Periodic payment calculation

For example, if you borrow $30 000 for 15 years at 5% interest rate, then you choose the entry in the 15 years row and 5% column, $79.08, and multiply it by 3 (30 000 is 3 times 10 000). So, your monthly payment is $3 \times 79.08 = \$237.24$

Note that this amount is given by your GDC or by applying the formula given earlier.

Using the formula: $$R = \frac{\frac{0.05}{12} \times 30\,000}{1 - \left(1 + \frac{0.05}{12}\right)^{-12\times15}} = 237.238$$

Example 3.27

Alexander wants to borrow €180 000 to buy an apartment in his city. The rate offered by the bank is 6%, and the loan is taken for 20 years and will be paid in monthly instalments.

(a) Using Table 3.6, calculate his possible monthly payment.

(b) Calculate the amount of interest he would be paying.

(c) He can afford to pay a maximum of €1100 per month. What is the maximum possible amount he can borrow?

Solution

(a) Looking up the entry in 20 years row, 6% column: 71.64

Alexander's monthly payment is $\frac{180\,000}{10\,000} \times 71.64 = €1289.52$

(b) The amount he pays for interest is the total amount he pays back minus the amount he borrows: $1289.52 \times 12 \times 20 - 180\,000 = €129\,484.80$

(c) The amount he can afford: $\frac{11\,000}{71.64} \times 10\,000 = €153\,545.51$

Exercise 3.4

1. For each arithmetic series

(i) find the number of terms

(ii) find the sum of the series.

(a) $11 + 17 + \ldots + 365$

(b) $\sum_{k=0}^{13}(2 - 0.3k)$

(c) $9 + 13 + 17 + \ldots + 85$

(d) $8 + 14 + 20 + \ldots + 278$

(e) $155 + 158 + 161 + \ldots + 527$

2. For each geometric series, find

(i) the common ratio

(ii) the number of terms

(iii) the sum of the series.

(a) $2 - 3 + \frac{9}{2} - \frac{27}{4} + \ldots - \frac{177\,147}{1024}$

(b) $2 - \frac{4}{5} + \frac{8}{25} - \ldots + \frac{128}{15\,625}$

(c) $\frac{1}{3} + \frac{\sqrt{3}}{12} + \cdots + \frac{9}{4096} + \frac{9\sqrt{3}}{16\,384}$

(d) $2 - \frac{4}{3} + \frac{8}{9} - \frac{16}{27} + \ldots - \frac{4096}{177\,147}$

3. At the beginning of every month Maggie invests £150 in an account that pays 6% annual rate. How much money will there be in the account after 6 years?

4. How many terms should be added to exceed 678 in the series

$17 + 20 + 23 \ldots?$

5. How many terms should be added to exceed 2335 in the series

$-18 - 11 - 4 \ldots?$

6. Each month Susannah invests €100 in an account earning 4% annual interest compounded monthly.

(a) Find the value of her investment after 20 years, assuming she started with a zero balance.

(b) How much initial balance would she have had to start with if she makes no monthly payments, but the account has the same amount of money as your answer to part **(a)**?

7. Each month Maxine invests £150 in a fund for her retirement.

The fund earns 6% annual interest compounded monthly.

(a) Find the value of her retirement fund in 30 years.

(b) If she instead invests £200 per month, find the value of her retirement fund after 30 years.

8. Find the monthly payment required to save $4000 when earning 4% annual interest compounded monthly within 5 years.

9. The sum of the first n terms of an arithmetic sequence is given by

$S_n = 6n + n^2$.

(a) Write down the value of

(i) S_1 **(ii)** S_2

The nth term of the arithmetic sequence is given by u_n.

(b) Show that $u_2 = 9$

(c) Find the common difference of the sequence.

(d) Find u_{10}

(e) Find the lowest value of n for which u_n is greater than 1000.

(f) There is a value of n for which $u_1 + u_2 + \ldots + u_n = 1512$
Find the value of n.

10. Evaluate $\sum_{k=0}^{11}(3 + 0.2k)$

11. A ball is dropped from a height of 16 m. Every time it hits the ground it bounces 81% of its previous height.

(a) Find the maximum height it reaches after the 10th bounce.

(b) Find the total distance travelled by the ball when it hits the ground for the 10th time. (Assume no friction and no loss of elasticity.)

12. Find each sum.

(a) $7 + 12 + 17 + 22 + \ldots + 337 + 342$

(b) $9486 + 9479 + 9472 + 7465 + \ldots + 8919 + 8912$

(c) $2 + 6 + 18 + 54 + \ldots + 3\,188\,646 + 9\,565\,938$

(d) $120 + 24 + \frac{24}{5} + \frac{24}{25} + \ldots + \frac{24}{78\,125}$

13. A loan takes 10 years to repay at \$450 per month. The annual interest rate of the loan is 5% and the loan is compounded monthly.

(a) Find the initial value of the loan.

(b) Calculate how long it takes to repay this loan at \$900 per month.

14. On Saturdays Kevin goes to a running track to train. He runs the first lap of the track in 180 seconds. Each lap Kevin runs takes him 10 seconds longer than his previous lap.

(a) Find the time, in seconds, Kevin takes to run his fifth lap.

On a specific Saturday, Kevin ran his last lap in 280 seconds.

(b) Find how many laps he has run on this Saturday.

(c) Find the total time, in minutes, run by Kevin on this Saturday.

On one Saturday Kevin takes Liz to train. They both run the first lap of the track in 180 seconds. Each lap Liz runs takes 1.05 times as long as her previous lap.

(d) Find the time, in seconds, Liz takes to run her third lap.

(e) Find the total time, in seconds, Liz takes to run her first four laps.

Each lap Kevin runs again takes him 10 seconds longer than his previous lap. After a certain number of laps Kevin takes less time per lap than Liz.

(f) Find the number of the lap when this happens.

15. Give all answers in this question correct to two decimal places.

Arthur lives in London. On 1 August 2008 Arthur paid 37 500 euros (EUR) for a new car from Germany. The price of the same car in London was 34 075 British pounds (GBP).

The exchange rate on 1 August 2008 was 1EUR = 0.7234 GBP

(a) Calculate, in GBP, the price that Arthur paid for the car.

(b) Write down, in GBP, the amount of money Arthur saved by buying the car in Germany.

On 1 August 2008 Arthur invested the money he saved in a bank that paid 4.5% annual simple interest.

(c) Calculate, in GBP, the value of Arthur's investment on 1 August 2012.

Between 1 August 2008 and 1 August 2012 Arthur's car depreciated at an annual rate of 9% of its current value.

(d) Calculate the value, in GBP, of Arthur's car on 1 August 2009.

(e) Show that the value of Arthur's car on 1 August 2012 was 18 600 GBP, correct to the nearest 100 GBP.

On 1 August 2012 Arthur sold his car for 18 600 GBP and bought a new car from Germany for 30 500 EUR. He used the 18 600 GBP and the value of the investment he made on 1 August 2008 to buy the new car.

The exchange rate on 1 August 2012 was 1EUR = 0.8694 GBP.

(f) Calculate the amount of money remaining, in EUR, after the car had been bought.

16. Every month, since her 12th birthday, Hana has deposited \$200 in a bank account earning 6.2% annual interest compounded monthly.

(a) If Hana has just had her 32nd birthday, how much money is in the account?

(b) When Hana turns 40, she doubles the amount of money invested. How much money, to the nearest thousand, will she have when she retires at age 65?

Chapter 3 practice questions

1. In an arithmetic sequence, the first term is 4, the 4th term is 19 and the nth term is 99. Find the common difference and the number of terms n.

2. How much money should you invest now if you wish to have an amount of €3000 in your account after 6 years if interest is compounded quarterly at an annual rate of 6%?

3. Two students, Nick and Charlotte, decide to start preparing for their IB exams 15 weeks ahead of the exams. Nick starts by studying for 12 hours in the first week and plans to increase the amount by 2 hours per week. Charlotte starts with 12 hours in the first week and decides to increase her time by 10% every week.

(a) How many hours did each student study in week 5?

(b) How many hours in total does each student study for the 15 weeks?

(c) In which week will Charlotte exceed 40 hours per week?

(d) In which week does Charlotte catch up with Nick in the number of hours spent on studying per week?

4. Two diet schemes are available for people to lose weight. Plan A promises the patient an initial weight loss of 1000 grams the first month with a steady loss of an additional 80 grams every month after the first. In the second month the patient will lose 1080 grams and so on for a maximum duration of 12 months.

   Plan B starts with a weight loss of 1000 grams the first month and an increase in weight loss by 6% more every following month.

   (a) Write down the number of grams lost under plan B in the second and third months.

   (b) Find the weight lost in the 12th month for each plan.

   (c) Find the total weight loss during a 12-month period under
   (i) plan A **(ii)** plan B.

5. You start a savings plan to buy a car where you invest €500 at the beginning of the year for 10 years. Your investment scheme offers a fixed rate of 6% per year compounded annually.

   Calculate, giving your answers to the nearest euro(€),

   (a) how much the first €500 is worth at the end of ten years

   (b) the total value your investment will give you at the end of the ten years.

6. The first three terms of an arithmetic sequence are 6, 9.5, 13.

   (a) What is the 40th term of the sequence?

   (b) What is the sum of the first 103 terms of the sequence?

7. Marco starts training for running his city marathon.
   On the first day he runs 3 km. Every day he runs 10% more than the day before.

   (a) Write down the distance he runs on the second day of training.

   (b) Calculate the total distance Marco runs in the first seven days of training.

   Marco stops training on the day his total distance exceeds 150 km.

   (c) Calculate the number of days Marco has trained for the running race.

8. A marathon runner plans her training program for a 20 km race.
   On the first day she plans to run 2 km, and then she wants to increase her distance by 500 m on each subsequent training day.

   (a) On which day of her training does she first run a distance of 20 km?

   (b) By the time she manages to run the 20 km distance, what is the total distance she would have run for the whole training program?

9. In a certain country, cellular phones were first introduced in the year 2010. During the first year, the number of people who bought a cellular phone was 1600. In 2011, the number of new purchasers was 2400, and in 2012 the new purchasers numbered 3600.

(a) Note that the trend is a geometric sequence. Find the common ratio.

(b) Assuming that the trend continues:

(i) how many purchasers will join in 2022?

(ii) in what year would the number of new purchasers first exceed 50 000?

Between 2010 and 2012, the total number of purchasers reaches 7600.

(c) What is the total number of purchasers between 2010 and 2022?

During this period, the total adult population of the country remains approximately 800 000.

(d) Use this information to suggest a reason why this trend in growth would not continue.

10. In an arithmetic sequence, the first term is -5, the fourth term is 13, and the nth term is 11 995. Find the common difference d and the number of terms n.

11. An auditorium for a city is built in the form of a horseshoe and has 25 rows.

The number of seats in each row increases by a fixed number, d, compared to the number of seats in the previous row. The number of seats in the sixth row is 100, and the number of seats in the tenth row is 124. a_1 represents the number of seats in the first row.

(a) Write down the value of

(i) d **(ii)** a_1

(b) Find the total number of seats in the auditorium.

A few years later, a second level was added to increase the auditorium's capacity by another 1600 seats. Each row has four more seats than the previous row. The first row on this level has 70 seats.

(c) Find the number of rows on the second level of the auditorium.

12. Frieda is a dedicated swimmer. She goes swimming once every week. She starts with the first week of the year by swimming 200 metres. Each week after that she swims 20 metres more than the previous week. She does this for the whole year (52 weeks).

(a) How far does she swim in the final week?

(b) How far does she swim altogether?

13. You and your classmates are conducting an experiment. You receive a text message at 12:00. Five minutes later you forward the text message to three people. Five minutes later, those three people forward the text message to three new people. Assume this pattern continues and each time the text message is sent to people who have not received it before.

(a) Find the number of people who will receive the text message at 12:30.

(b) Find the total number of people who will have received the text message by 12:30.

(c) At what time will a total of 29 523 people have received the text message?

14. Two IT companies offer apparently similar salary schemes for their new appointees. Kell offers a starting salary of €18 000 per year and then an annual increase of €400 every year after the first. YBO offers a starting salary of €17 000 per year and an annual increase of 7% for the rest of the years after the first.

(a) (i) Write down the salary paid in the 2nd and 3rd years for each company.

(ii) Calculate the total amount that an employee working for 10 years will accumulate over 10 years in each company.

(iii) Calculate the salary paid the tenth year in each company.

(b) Tim works at Kell and Merijayne works at YBO.

(i) When would Merijayne start earning more than Tim?

(ii) What is the minimum number of years that Merijayne requires so that her total earnings exceed Tim's total earnings?

15. A theatre has 24 rows of seats. There are 16 seats in the first row, and each successive row increases by two seats, one on each side.

(a) Calculate the number of seats in the 24th row.

(b) Calculate the number of seats in the whole theatre.

16. The amount of €7000 is invested at 5.25% annual compound interest.

(a) Write down an expression for the value of this investment after t full years.

(b) Calculate the minimum number of years required for this amount to become €10 000.

(c) For the same number of years as in part **(b)**, would an investment of the same amount be better if it were at a 5% rate compounded quarterly?

17. With S_n denoting the sum of the first n terms of an arithmetic sequence, we are given that $S_1 = 9$ and $S_2 = 20$

(a) Find the second term.

(b) Calculate the common difference of the sequence.

(c) Find the 4th term.

18. Consider an arithmetic sequence whose second term is 7. The sum of the first four terms of this sequence is 12. Find

(a) the first term

(b) the common difference of the sequence.

19. In an arithmetic sequence of positive terms, let a_n represent the nth term. Given that $\frac{a_5}{a_{12}} = \frac{6}{13}$ and $a_1 \times a_3 = 32$, find $\sum_{i=1}^{100} a_i$

20. In an arithmetic sequence $a_1 = 5$ and $a_2 = 13$

(a) Write down, in terms of n, an expression for the nth term, a_n.

(b) Find n such that $a_n < 400$

21. Minta deposits 1000 euros in a bank account. The bank pays a nominal annual interest rate of 5%, compounded quarterly.

(a) Find the amount of money that Minta will have in the bank after 3 years. Give your answer correct to two decimal places.

Minta will withdraw the money from her bank account when the interest earned is 300 euros.

(b) Find the time, in years, until Minta withdraws the money from her bank account.

22. Consider the arithmetic sequence 85, 78, 71, …. Find the sum of its positive terms.

23. A geometric sequence is defined by

$u_n = 3(4)^{n+1}$, $n \in \mathbb{Z}^+$, where u_n is the nth term.

(a) Find the common ratio r.

(b) Hence, find S_n, the sum of the first n terms of this sequence.

24. An arithmetic sequence has 32 as its 10th term and the common difference is -6

(a) Find the second term of the sequence.

(b) Find the 30th term of the sequence.

(c) Find the sum of the first 50 terms of the sequence.

25. Given an arithmetic series $S_n = 2 + 5 + 8 + \ldots$,

(a) find an expression for the partial sum S_n, in terms of n

(b) for what value of n is $S_n = 1365$?

26. Veronica wants to make an investment and accumulate €25 000 over a period of 18 years. She finds two investment options.

Option 1 offers simple interest of 5% per annum.

(a) Find out the exact amount she will have in her account after 18 years, if she invests €12 500 with this option.

Option 2 offers a nominal annual interest rate of 4%, compounded monthly.

(b) Find the amount that Veronica has to invest with option 2 to have €25 000 in her account after 18 years. Give your answer correct to two decimal places.

27. Prachi is on vacation in the United States. She is visiting the Grand Canyon.

She drops a coin from the top of a cliff. The coin falls a distance of 5 metres during the first second, 15 metres during the next second, 25 metres during the third second and continues in this way.
The distances that the coin falls during each second forms an arithmetic sequence.

(a) **(i)** Write down the common difference, d, of this arithmetic sequence.

(ii) Write down the distance the coin falls during the fourth second.

(b) Calculate the distance the coin falls during the 15th second.

(c) Calculate the total distance the coin falls in the first 15 seconds. Give your answer in kilometres.

Prachi drops the coin from a height of 1800 metres above the ground.

(d) Calculate the time, to the nearest second, the coin will take to reach the ground.

Prachi visits a tourist centre nearby. It opened at the start of 2018 and in the first year there were 17 000 visitors. The number of people who visit the tourist centre is expected to increase by 10% each year.

(e) Calculate the number of people expected to visit the tourist centre in 2019.

(f) Calculate the total number of people expected to visit the tourist centre during the first 10 years since it opened.

28. Antonio and Rania started work at the same company at the same time. The starting annual salary for each was €36 000. Depending on which division of the company they work in, the company gives them a salary increase following the completion of each year of employment according to different plans. Antonio, an engineer, is paid using plan A and Rania, an accountant, is paid using plan B.

Plan A: The annual salary increases by €1200 each year.

Plan B: The annual salary increases by 3% each year.

(a) Calculate

(i) Antonio's annual salary during his second year of employment

(ii) Rania's annual salary during her second year of employment.

(b) Write down an expression for

(i) Antonio's annual salary during his nth year of employment

(ii) Rania's annual salary during her nth year of employment.

(c) Determine in which year Antonio's annual salary will become smaller than or equal to Rania's annual salary.

Both Antonio and Rania plan to work at the company for a total of 15 years.

(d) **(i)** Calculate the total amount that Rania will be paid during these 15 years.

(ii) Determine the amount by which Rania's salary exceeds Antonio's during these 15 years.

29. 300 rabbits were introduced to an isolated island in the Pacific. One week later the number of rabbits was 315. The number of rabbits, N, can be modelled by a geometric sequence. $N(t)$ is the number of rabbits t weeks since their introduction to this island.

$N(0) = 300$

(a) Find an expression for $N(t)$.

(b) Find the number of rabbits on the island after 10 weeks.

A scientist estimates that the island has enough food to support a maximum population of 2000 rabbits.

(c) Estimate the number of weeks it takes for the rabbit population to reach this maximum

30. The first four terms of an arithmetic sequence are 2, $a - b$, $2a + b + 7$, and $a - 3b$, where a and b are constants. Find a and b

31. Three consecutive terms of an arithmetic sequence are: a, 1, b. The terms 1, a, b are consecutive terms of a geometric sequence. If $a \neq b$, find the values of a and of b.

32. Lukas deposited \$8000, in a bank account which pays a nominal annual interest rate of 5%, compounded quarterly.

(a) Find how much interest Lukas has earned after 2 years.

Sophia also deposited money in a bank account. Her account pays a nominal annual interest rate of 7%, compounded semi-annually. After three years, the total amount in Sophia's account is \$10 000.

(b) Find the amount that Sophia deposited in the bank account.

33. Roberto invests 5000 euros in a fixed deposit account that pays an annual interest rate of 4.5%, compounded monthly, for seven years.

(a) Calculate the value of Roberto's investment at the end of this time.

Helena has 7000 euros to invest in a fixed deposit account which is compounded annually.

(b) Calculate the minimum annual interest rate needed for Helena to double her money in 10 years.

34. One of the locations in the 2016 Olympic Games is an amphitheatre. The number of seats in the first row of the amphitheatre, u_1, is 240. The number of seats in each subsequent row forms an arithmetic sequence. The number of seats in the sixth row, u_6, is 270.

(a) Calculate the value of the common difference, d.

There are 20 rows in the amphitheatre.

(b) Find the total number of seats in the amphitheatre.

Anisha visits the amphitheatre. She estimates that the amphitheatre has 6500 seats.

(c) Calculate the percentage error in Anisha's estimate.

35. You invest \$5000 at an annual compound interest rate of 6.3%.

(a) Write an expression for the value of this investment after t full years.

(b) Find the value of this investment at the end of five years.

(c) After how many full years will the value of the investment exceed \$10 000?

36. The sum of the first n terms of an arithmetic sequence $\{u_n\}$ is given by the formula $S_n = 4n^2 - 2n$

Three terms of this sequence, u_2, u_m and u_{32}, are consecutive terms in a geometric sequence. Find m.

37. The sum of the first 16 terms of an arithmetic sequence $\{u_n\}$ is 12. Find the first term and the common difference if the ninth term is zero.

38. **(a)** Write down how many integers between 10 and 300 are divisible by 7.

(b) Express the sum of these integers in sigma notation.

(c) Find the sum above.

(d) Given an arithmetic sequence with first term 1000 and a common difference of -7, find the smallest n so that the sum of the first n terms of this sequence is negative.

39. The sum of the first n terms of an arithmetic sequence is given by $S_n = 4n + 2n^2$

(a) Find the value of

(i) S_1

(ii) S_2

Let a_n be the nth term of this arithmetic sequence.

(b) Show that $a_2 = 10$

(c) Find the common difference of the sequence.

(d) Find a_{10}.

(e) Find the smallest possible value of n for which $a_n > 2000$

(f) For what value of n is the series $a_1 + a_2 + \dots + a_n = 3040$?

40. The table shows the monthly repayments for every £10 000 borrowed at the nominal annual interest rates and for the loan terms shown.

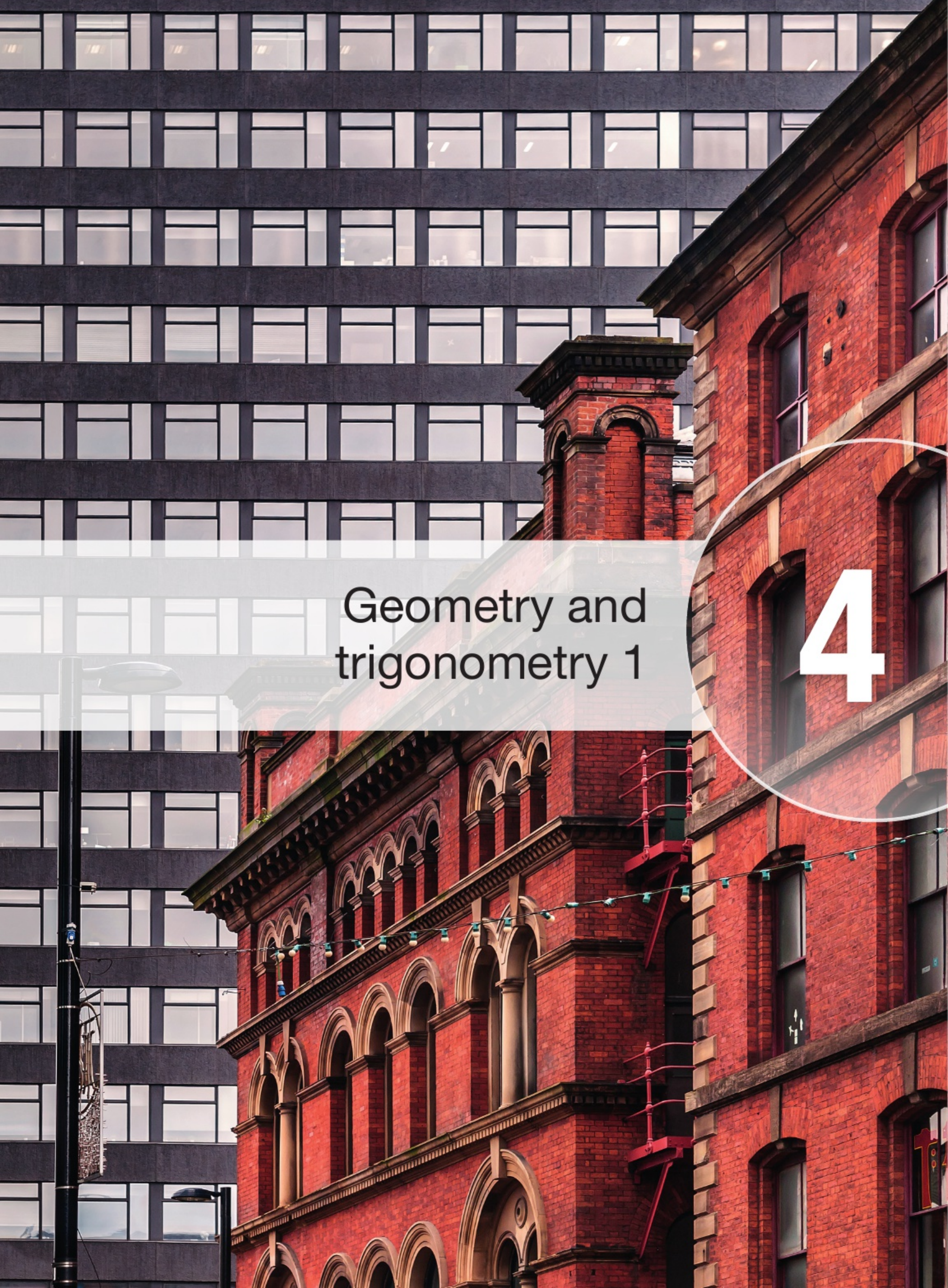
Loan term (years)	**Monthly repayment in £, for every £10 000**		
	Nominal annual interest rate		
	5%	6%	7%
5	188.72	193.33	198.01
10	106.07	111.02	116.11
15	79.08	84.39	89.88
20	66.00	71.64	77.53

Rami borrows £80 000 at a nominal annual interest rate of 6%, to be repaid over a 15 year term.

(a) Using the table, calculate the total interest that Rami pays.

Rana borrows a sum of money at a nominal annual interest rate of 5% over a 10-year term. Her monthly repayment must not be more than £600.

(b) Using the table, calculate the maximum amount that Rana can borrow. Give your answer correct to the nearest £1000.



4 Geometry and trigonometry 1

4 Geometry and trigonometry 1

Learning objectives

By the end of this chapter, you should be familiar with…

- representing locations on the coordinate plane and calculating distances between points, midpoints of line segments and intersections of lines
- finding missing side lengths and angles in right-angled triangles using the Pythagorean theorem and the primary trigonometric ratios sine, cosine and tangent
- using the sine and cosine rules to calculate missing side lengths and angles in non-right-angled triangles and application to real-life problems
- determining the surface area and volume of right pyramids, right cones, spheres, hemispheres and solids composed of these shapes.

In this chapter, the fundamental concepts and structures required to solve geometrical problems will be introduced and applied. We begin with the celebrated Pythagorean theorem and use it to develop a distance formula for line segments in the Cartesian coordinate plane. We also develop equations of straight lines and consider how they intersect. The introduction of the primary trigonometric ratios demonstrates the relationships between side lengths and interior angles of a right-angled triangle and provides a powerful tool for solving problems involving navigation and architecture. Although only applicable to right-angled triangles, the primary trigonometric ratios can be extended to characterise non-right-angled triangles. These are the sine and cosine rules. Our introduction to geometry culminates in the calculation of areas and volumes of 2- and 3-dimensional figures.

4.1 Basic coordinate geometry

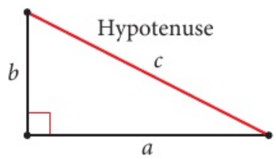


Figure 4.1 $a^2 + b^2 = c^2$

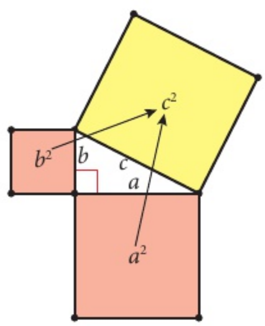


Figure 4.2 The squares formed using the base and the height of a right-angled triangle as side lengths fit perfectly in the square formed by using the hypotenuse as a side length

The foundational result of this section is the Pythagorean theorem, which states:

In a right-angled triangle, the square of the length of the hypotenuse (the longest side, located opposite the right angle) is equal to the sum of the squares of the lengths of the other two sides.

Figures 4.1 and 4.2 show the geometric interpretation where $a^2 + b^2 = c^2$

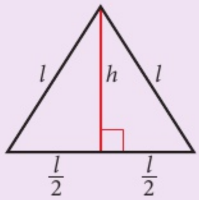
According to the Pythagorean theorem, the squares formed by using the base and the height of a right-angled triangle as side lengths fit perfectly in the square formed by using the hypotenuse as a side length.

Example 4.1

An equilateral triangle has a side length of l. Find the triangle's height h and area A in terms of l.

Solution

Using the Pythagorean theorem on one of the component right-angled triangles in the diagram gives:



$$\left(\frac{l}{2}\right)^2 + h^2 = l^2$$

$$\frac{l^2}{4} + h^2 = l^2$$

Collecting like terms and simplifying,

$$h^2 = \frac{4l^2}{4} - \frac{l^2}{4} = \frac{3l^2}{4}$$

Thus the height, $h = \frac{\sqrt{3}\,l}{2}$

Using the formula area $= \frac{1}{2}$base $\times$ height for the area of a triangle yields:

$$A = \frac{1}{2}l\left(\frac{\sqrt{3}\,l}{2}\right) = \frac{\sqrt{3}\,l^2}{4}$$

Some problems involving distance in 3 dimensions can be solved by breaking up the task into multiple 2-dimensional problems, as Example 4.2 demonstrates.

Example 4.2

The box in Figure 4.3 has a length of 80 cm, a width of 60 cm and a height of 90 cm. Calculate the length of the segment AB and hence, the diagonal AC.

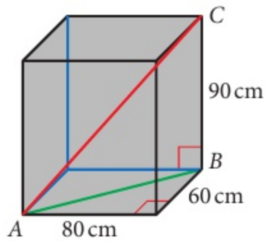


Figure 4.3 Diagram for Example 4.2

Solution

The length AB in the base of the box can be calculated as follows:

$AB^2 = 80^2 + 60^2$

$AB^2 = 10\,000$

$AB = 100$ cm

The length AC can now be found by using the right-angled triangle consisting of the base AB, the height of the box and the main diagonal as follows:

$AC^2 = 100^2 + 90^2$

$AC^2 = 10\,000 + 8100$

$AC = \sqrt{18\,100} \approx 135$ cm

Notice that this 3-dimensional problem is solved by considering two different 2-dimensional triangles within the diagram, one lying within the base and one perpendicular to the base.

In the IB, $[AB]$ denotes a line segment connecting the points A and B, AB is the length of this segment and (AB) is the line passing through points A and B.

The distance formula

Using the Pythagorean theorem: the distance between the two points $A(x_1, y_1)$ and $B(x_2, y_2)$ is given by the formula: $AB = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2}$

The Pythagorean theorem can be applied within coordinate geometry to enable us to calculate the distance between two points $A(x_1, y_1)$ and $B(x_2, y_2)$ in the Cartesian coordinate plane. This distance is denoted AB (see Figure 4.4). The line segment $[AB]$ forms the hypotenuse of a right-angled triangle with a base of length $x_2 - x_1$ and height $y_2 - y_1$

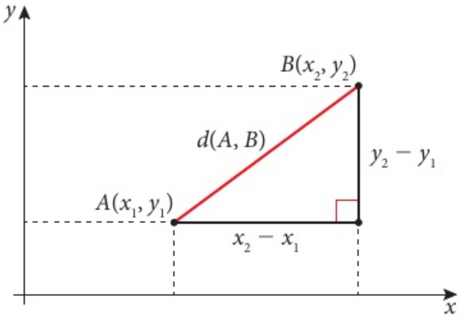


Figure 4.4 The line segment $[AB]$ forms the hypotenuse of a right-angled triangle

Example 4.3

(a) Find the length of the line segment connecting the points $A(-1, 4)$ and $B(3, 2)$

(b) How does the distance formula change when the points A and B lie on

(i) a horizontal line

(ii) a vertical line?

Solution

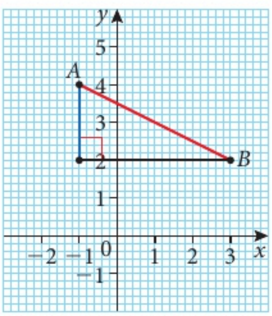
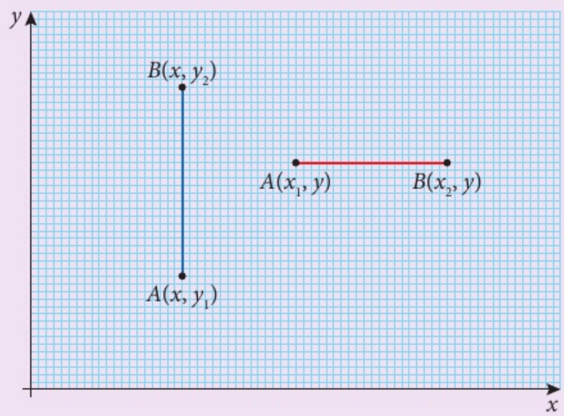


Figure 4.5 Solution to Example 4.3 (a)

(a) $AB = \sqrt{(3 - (-1))\,2 + (4 - 2)^2)} = \sqrt{4^2 + 2^2} = \sqrt{20}$

(b) (i) and (ii)



Two points A and B that lie on a vertical line have the same x coordinate. In this case

$$AB = \sqrt{(x - x)^2 + (y_2 - y_1)^2} = \sqrt{0 + (y_2 - y_1)^2} = |y_2 - y_1|$$

For A and B on the same horizontal line, the y coordinate is the same for both points and

$$AB = \sqrt{(x_2 - x_1)^2 + (y - y)^2} = |x_2 - x_1|$$

The midpoint of a line segment

In addition to calculating the length of a line segment, it is often important to locate the midpoint of a line segment. The **midpoint**, M, of the line segment $[AB]$ connecting two points $A(x_1, y_1)$ and $B(x_2, y_2)$, has the coordinates $\left(\frac{x_1 + x_2}{2}, \frac{y_1 + y_2}{2}\right)$(see Figure 4.6).

The coordinates of the midpoint of the line segment $[AB]$ are found by taking the average of the x and y coordinates of the points A and B.

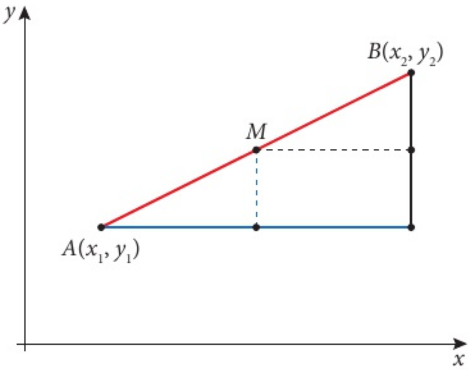


Figure 4.6 The midpoint of a line segment

Example 4.4

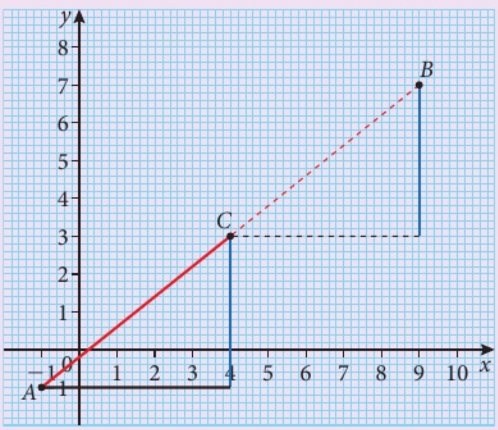
(a) Find the midpoint of the segment $[AB]$ connected by the points $A(2, 4)$ and $B(6, 12)$

(b) The point $C(4, 3)$ is the midpoint of the segment connecting $(-1, -1)$ and $B(m, n)$. Find the coordinates of the point B.

Solution

(a) Using the average of the x and y coordinates, $\left(\frac{2 + 6}{2}, \frac{4 + 12}{2}\right) = (4, 8)$

(b) Using the diagram, we obtain $B(9, 7)$



Example 4.4 (b) can be worked out using the midpoint formula. That is, $4 = \frac{-1 + m}{2} \Rightarrow m = 9$ and $3 = \frac{-1 + n}{2} \Rightarrow n = 7$

Equations of lines and intersection of lines

In this section, we revisit the equation for a straight line, introduced in Chapter 3, and use it to describe the intersection behaviour of pairs of straight lines. To do so, we must first define the **gradient**, m, of a line through the points $A(x_1, y_1)$ and $B(x_2, y_2)$ as $m = \frac{y_2 - y_1}{x_2 - x_1}$. The gradient is a measure of how much the height (y) of the straight line changes as we move one unit to the right along the x-axis. Lines with equal gradients or slopes are **parallel**, and lines for which the product of their gradients is -1 are **perpendicular**, meaning they intersect in a right angle. Lines that lie exactly on top of each other are called **coincident**.

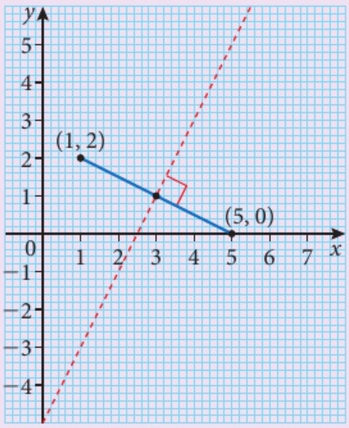
The words 'gradient' and 'slope' mean the same.

Perpendicular bisectors of line segments

The **perpendicular bisector** of a line segment is the line perpendicular to the segment passing through its midpoint. Every point on the perpendicular bisector is equidistant to the endpoints of the line segment.

Example 4.5

Find the equation of the perpendicular bisector (shown as dashed line in the graph) of the line segment connecting the points (1, 2) and (5, 0).



Solution

The midpoint of the segment is given by $\left(\frac{1+5}{2}, \frac{2+0}{2}\right) = (3, 1)$

The gradient of the line segment is $m = \frac{0-2}{5-1} = -\frac{1}{2}$, and the perpendicular bisector therefore has a gradient of 2. The equation of the perpendicular bisector is $y = 2x + b$, where b is the value of the unknown y-intercept. Substituting (3, 1) for (x, y) yields $1 = 2(3) + b$ and $b = -5$. Therefore, the perpendicular bisector has the equation $y = 2x - 5$

We conclude this section by showing how the equations of a pair of lines can be used to determine their point of intersection, when it exists.

Example 4.6

(a) Find the point of intersection of the lines with equations $y = -2x + 1$ and $y = 3x - 4$ algebraically.

(b) Draw both lines on the same graph and label the point of intersection.

Solution

(a) The x- and y-coordinates of the point of intersection are the same for both lines. Thus,

$-2x + 1 = 3x - 4$

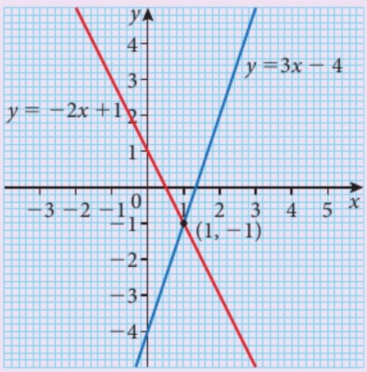
This simplifies to

$$5x = 5$$
$$x = 1$$

Substituting $x = 1$ into either of the equations yields the y-coordinate of the intersection point, which is $y = -1$.

The point of intersection is therefore $(1, -1)$.

(b) The lines and their point of intersection are shown on the graph.



The last example of this section considers the conditions under which a pair of lines possesses a unique point of intersection.

Example 4.7

(a) Find the x-coordinate of the point of intersection of the lines with equations $y = mx + b$ and $y = Mx + B$

(b) Interpret your answer to (a), considering different values of m, M, b, and B

Solution

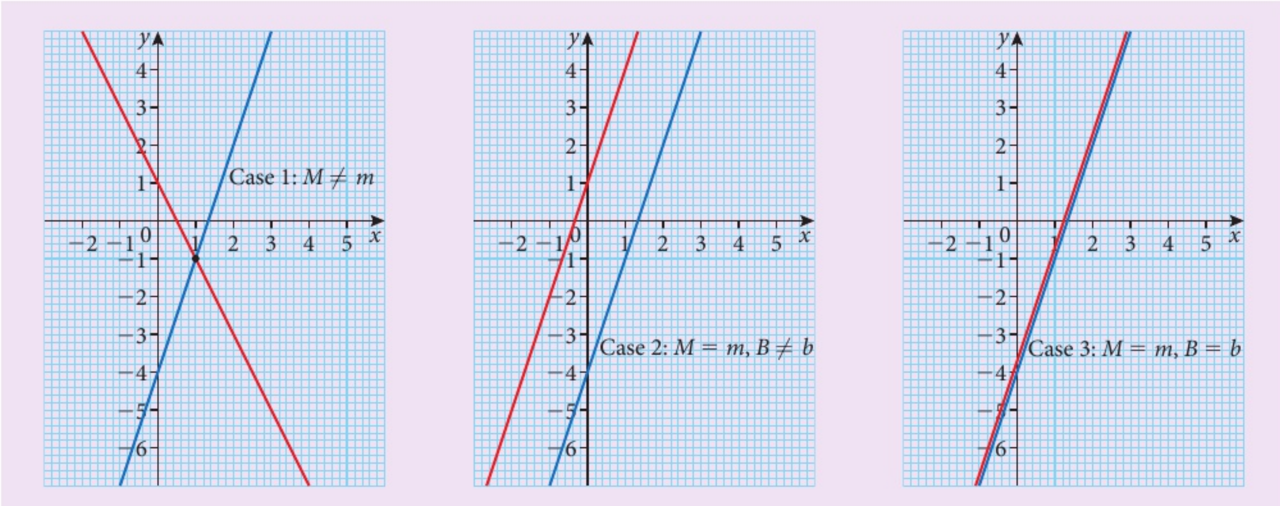
(a) At intersection, $mx + b = Mx + B$

When we simplify this, we obtain

$(m - M)x = B - b$

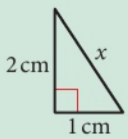
$$x = \frac{B - b}{m - M}$$

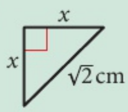
(b) If $m \neq M$ the two lines have different gradients and intersect at a unique point whose x coordinate is given by the expression in (a). If $m = M$, $b \neq B$, the two lines have equal gradients but different y intercepts and have no point of intersection. If $m = M$, $b = B$, the two lines are coincident, and intersect at all points. The three possibilities are shown in the graphs.

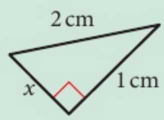


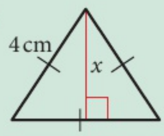
Exercise 4.1

1. Find the value of x in each triangle as a root and correct to 3 significant figures.

   (a)
   

   (b)
   

   (c)
   

   (d)
   

2. A Pythagorean triple is a set of three integers x, y and z that satisfy the equation $x^2 + y^2 = z^2$

   (a) Show that 3, 4 and 5 form a Pythagorean triple.

   (b) Show that $3n$, $4n$ and $5n$ form a Pythagorean triple for any integer n.

   (c) List all the square numbers between 1 and 200, and hence, find two more Pythagorean triples.

   (d) How can Pythagorean triples be interpreted using triangles and why are they special?

   (e) Show, using algebraic manipulation, that $m^2 - n^2$, $2mn$ and $m^2 + n^2$ form a Pythagorean triple for any whole numbers m and n, $m > n$

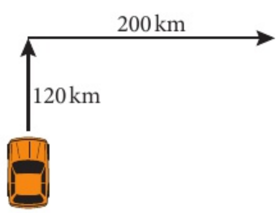
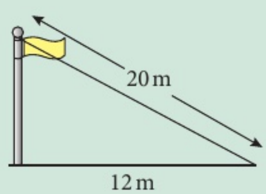


Figure 4.7 Diagram for question 3

3. A car drives 120 km north and then turns east, driving an additional 200 km. How far is the starting position from the destination?

4. A flag pole is supported by a sturdy rope that is fastened to a hook on the ground and the top of the flag pole. The rope is 20 m in length and the hook is located 12 m from the base of the flagpole. How high is the flagpole, correct to the nearest metre?



5. A cuboid is a 3-dimensional shape whose opposite faces are congruent rectangles. All adjacent edges are perpendicular to each other.
   (a) Find the lengths of [*CD*] and [*AB*] in the cuboid in Figure 4.8.
   (b) A cuboid-shaped closet has a length of 10 m, a width of 6 m and a height of 2.5 m. What is the maximum length of a broom that is to be stored in the closet?

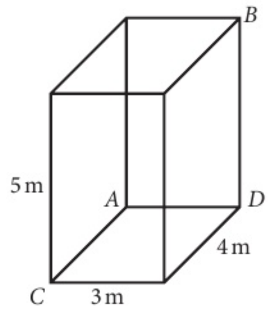


Figure 4.8 Diagram for question 5

6. For each of the following line segments, calculate the
   (i) length **(ii)** midpoint **(iii)** gradient.
   (a) The segment connecting (2, 5) and (9, 12)
   (b) The segment connecting (4, 9) and (7, 0)
   (c) The segment connecting (−2, 4) and (5, 3)
   (d) The segment connecting (4, 7) and (8, 7)
   (e) The segment connecting (−3, 0) and (−3, 5)

7. Draw line segments with the following gradients on a set of coordinate axes:
   (a) 1 and −1 **(b)** 2 and −0.5
   (c) −4 and 0.25 **(d)** $\frac{2}{3}$ and −1.5
   Explain the numerical relationship between the gradients in each pair.
   How do the line segments in each pair relate to each other?

8. Find the equation of each of the following lines in the form $y = mx + c$
   (a) The line passing through (0, 7) with gradient 3.
   (b) The line passing through (3, 6) with gradient −2.
   (c) The line passing through (2, 5) and parallel to $y = 3x - 8$
   (d) The line passing through the origin (0, 0) perpendicular to $y = \frac{1}{5}x - 7$
   (e) The perpendicular bisector of the segment connecting (0, 0) and (9, 3)

9. **(a)** Make y the subject of the formula $ax + by + c = 0$
   (b) Hence, find the gradient and y-intercept of the straight line $ax + by + c = 0$ in terms of a,b and c.

10. The lines $y = 2x + 1$, $y = -x + 4$ and $y = \frac{1}{2}x + 1$ bound a triangle *ABC* as shown in Figure 4.9.
   (a) Find the coordinates of *A*,*B* and *C*.
   (b) Find the lengths *AB*, *AC* and *BC* and hence, find the perimeter of the triangle *ABC*.
   (c) Find the gradient and midpoint of [*BC*].
   (d) Hence, find the equation of the perpendicular bisector of [*BC*].
   (e) Using graphing software or geometry instruments, draw the perpendicular bisectors of [*AC*] and [*BC*]. What do you notice about the perpendicular bisectors of all three sides of the triangle *ABC*?

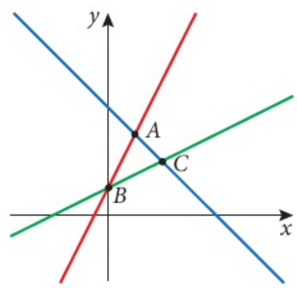


Figure 4.9 Diagram for question 10

4.2 Right-angled trigonometry

The primary trigonometric ratios

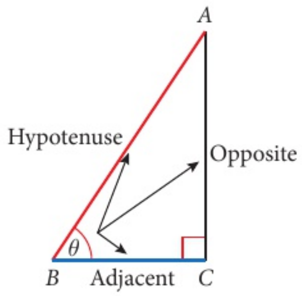


Figure 4.10 The sides of a triangle can be labelled as adjacent, opposite and hypotenuse

Greek letters such as θ (theta) are often used in texts to denote angles.

In this section, we develop the primary trigonometric ratios sine, cosine and tangent, which establish relationships between the side lengths of a right-angled triangle and its angles. These ratios are fundamental in solving problems in geometry, navigation and architecture and have far-reaching applications in the sciences.

Consider the right-angled triangle in Figure 4.10. Relative to the interior angle θ, the sides of the triangle can be labelled as adjacent (blue), opposite (black) and the hypotenuse (red), which is always the longest side, located opposite the right angle.

For an acute angle, θ, in a right-angled triangle, we define the following ratios:

$\sin\theta = \dfrac{\text{opposite}}{\text{hypoteneuse}}$ This is called the **sine ratio**.

$\cos\theta = \dfrac{\text{adjacent}}{\text{hypoteneuse}}$ This is called the **cosine ratio**.

$\tan\theta = \dfrac{\text{opposite}}{\text{adjacent}}$ This is called the **tangent ratio**.

The first example will show how to use a calculator to calculate the trigonometric ratios of known angles and will demonstrate some basic properties of these ratios.

Example 4.8

Make sure that your calculator is in 'degree' mode when you calculate trigonometric ratios.

(a) Use your calculator to find the following trigonometric ratios, correct to 3 decimal places:
   (i) $\sin 30, \sin 45, \sin 60, \sin 90$
   (ii) $\cos 30, \cos 45, \cos 60, \cos 90$
   (iii) $\tan 30, \tan 45, \tan 60, \tan 90$

(b) Explain why for any angle θ, $\sin\theta \leqslant 1$ and $\cos\theta \leqslant 1$
How is this different from $\tan\theta$?

Solution

Figure 4.11 Very small θ

(a) (i) $\sin 30 = 0.5$, $\sin 45 \approx 0.707$, $\sin 60 \approx 0.866$, $\sin 90 = 1$
   (ii) $\cos 30 \approx 0.866$, $\cos 45 \approx 0.707$, $\cos 60 = 0.5$, $\cos 90 = 0$
   (iii) $\tan 30 \approx 0.577$, $\tan 45 = 1$, $\tan 60 \approx 1.732$, $\tan 90$ is undefined

(b) Since $\sin\theta$ and $\cos\theta$ are fractions with the hypotenuse in the denominator, and the hypotenuse is the longest side of a right-angled triangle, the fraction must have a value less than or equal to 1. For very small θ, the opposite side is small and the adjacent side is quite large compared to the hypotenuse (Figure 4.11). Therefore, $\sin\theta$ is small while $\cos\theta$ is close to 1. The opposite is true for large values of θ (Figure 4.12).

Figure 4.12 Large θ

For the tangent ratio, as we see in Figure 4.13, if $\theta = 45°$, then the adjacent and opposite sides have equal value, and $\tan 45° = 1$.
For $\theta \geqslant 45°$, the opposite side is equal to or exceeds the adjacent side and $\tan\theta \geqslant 1$. As θ approaches $90°$, the adjacent side approaches zero and $\tan\theta$ approaches infinity. Hence, $\tan 90$ is undefined.

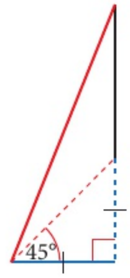


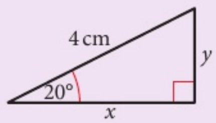
Figure 4.13 When $\theta = 45°$ the adjacent and opposite sides have equal value and $\tan 45 = 1$. As θ gets larger, $\tan\theta$ approaches infinity

In many applications, we may know the values of some side lengths or angles of a right-angled triangle. We then use trigonometric ratios to calculate the value of the missing quantities.

Example 4.9

Find the value of each missing side length, correct to 2 decimal places.

(a)



(b)

Solution

(a) $\sin 20 = \dfrac{y}{4} \Rightarrow y = 4\sin 20 \approx 1.37\text{ cm}$

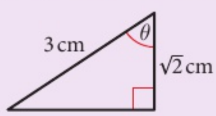
$\cos 20 = \dfrac{x}{4} \Rightarrow x = 4\cos 20 \approx 3.76\text{ cm}$

(b) $\sin 15 = \dfrac{2}{y} \Rightarrow y\sin 15 = 2 \Rightarrow y = \dfrac{2}{\sin 15} \approx 7.23\text{ cm}$

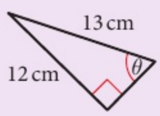
Example 4.10

For each of the following right-angled triangles, find the value of the interior angle θ, correct to the nearest degree.

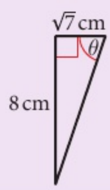
(a)



(b)



(c)



Solution

To find the value of a missing angle, the **inverse trigonometric ratios** must be used. These are called arccos, arcsin and arctan. Remember that the inverse of a function 'undoes' the action of the function. The inverse trigonometric functions allow us to input a given trigonometric ratio and return the value of the corresponding angle.

(a) $\theta = \arccos\left(\dfrac{\sqrt{2}}{3}\right) \approx 62°$ (b) $\theta = \arcsin\left(\dfrac{12}{13}\right) \approx 67°$ (c) $\theta = \arctan\left(\dfrac{8}{\sqrt{7}}\right) \approx 72°$

On most GDCs, arccos, arcsin, and arctan are also referred to as $\cos^{-1}$, $\sin^{-1}$, and $\tan^{-1}$

The area of a triangle

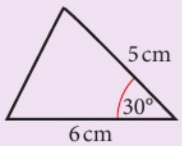
Recall that the area of a triangle is given by the formula area $= \frac{1}{2}\text{base} \times \text{height}$

In some problems, we may not know the height of the triangle, but we may know an angle and the enclosing side lengths. First we will consider the case where the angle is acute.

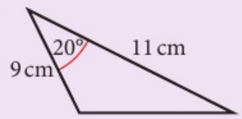
Example 4.11

Find the area of each of these triangles.

(a)



(b)



Solution

(a) Since we have two side lengths that enclose an angle, we can draw a perpendicular (dashed line in Figure 4.14) from one of the vertices of the triangle to the base, and thus calculate the height h. We obtain

$\sin 30 = \frac{h}{5} \Rightarrow h = 5 \sin 30 \text{ cm}$

The area is then $\frac{1}{2}(6)(5\sin 30) = 7.5\text{ cm}^2$

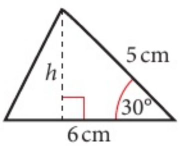


Figure 4.14 Solution to Example 4.11 (a)

(b) By drawing the height, h, of the triangle, we obtain the diagram in Figure 4.15.

$\sin 20 = \frac{h}{9} \Rightarrow h = 9 \sin 20$

The area of the triangle is $\frac{1}{2}(11)(9\sin 20) \approx 17\text{ cm}^2$

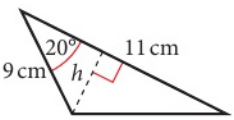


Figure 4.15 Solution to Example 4.11 (b)

In general, the area of a triangle with an acute angle θ enclosed by two sides of length a and b is $\frac{1}{2}ab\sin\theta$

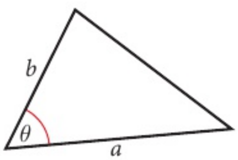


Figure 4.16 Area of triangle with two sides and the included angle

Applications of trigonometry

The following examples demonstrate how the primary trigonometric ratios can be used in calculations involving navigation and architecture. Recall that a 3-digit bearing is an angle measurement that uses 0° for north and states angles measured in a clockwise direction. Bearings are frequently used in navigation problems.

Example 4.12

A aeroplane flies 400 km on a bearing of 078° from city A to city B.

(a) How far north or south is B from A?

(b) How far east or west is A from B?

In order to avoid rounding errors, it is very important that you enter the expression as it is written into your calculator and only round the final answer. Do not calculate the trigonometric ratios and round them and then enter the rounded value into the expression.

Solution

(a) By letting y be the distance travelled north,

$\cos 78 = \frac{y}{400} \Rightarrow y = 400\cos 78 \approx 83\,\text{km}$

(b) By letting x be the distance travelled east,

$\sin 78 = \frac{x}{400} \Rightarrow x = 400\sin 78 \approx 391\,\text{km}$

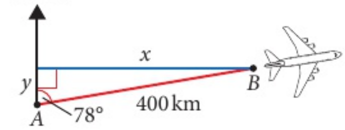


Figure 4.17 Solution to Example 4.12

Angles of elevation and depression

In land surveying and architecture, angles are often used to calculate lengths that are too large to measure. A clinometer is a device that helps measure such angles. Angles of **elevation** or **inclination** are angles above the horizontal, like looking up from ground level towards the top of a flagpole. Angles of **depression** or **declination** are angles below the horizontal, like looking down from your window to the base of the building opposite to you.

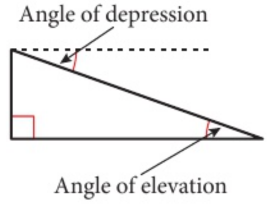


Figure 4.18 Angle of depression and angle of elevation

Example 4.13

The Burj al-Arab luxury hotel in Dubai, United Arab Emirates, is located on an island 280 m from Dubai beach. A land surveyor who is 1.78 m tall standing at the shore of Dubai beach measures the angle of elevation to the roof of Burj al-Arab to be 36.6°. How tall is the Burj al-Arab?

Solution

Using the diagram in Figure 4.19, $\tan 36.6 = \frac{h}{280} \Rightarrow h = 280\tan 36.6$

$\Rightarrow h \approx 208\,\text{m}$

Adding the height of the land surveyor now gives a total height of about 210 m.

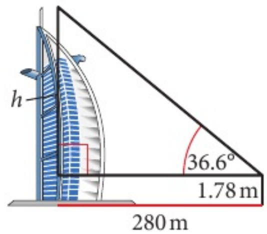


Figure 4.19 Solution to Example 4.13

Exact values of trigonometric ratios

We conclude this section by considering how to obtain exact values for trigonometric ratios of given angles. It is often difficult to obtain exact values for the primary trigonometric ratios in terms of fractions and roots, but for 30°, 45° and 60°, we can construct appropriate right-angled triangles to enable us to find the exact values for sine, cosine and tangent.

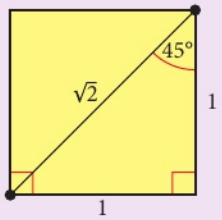
Example 4.14

By constructing appropriate right-angled triangles, find the exact values of $\sin\theta$, $\cos\theta$, $\tan\theta$, $\theta = 30°, 45°, 60°$

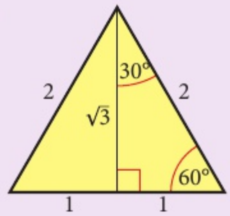
Solution

A right-angled triangle with an interior angle of 45° can be constructed by dividing a square of side length 1 unit along the diagonal (a). A right-angled triangle with interior angles of 30° and 60° can be constructed by dividing an equilateral triangle with side length 2 units into two congruent right-angled triangles (b).

(a)



(b)



Using the triangles shown in diagrams (a) and (b), we obtain

$\sin 45 = \cos 45 = \frac{1}{\sqrt{2}}$

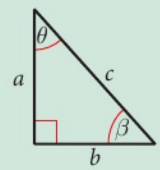
$\tan 45 = 1$

$\sin 30 = \frac{1}{2}, \cos 30 = \frac{\sqrt{3}}{2}, \tan 30 = \frac{1}{\sqrt{3}}$

$\sin 60 = \frac{\sqrt{3}}{2}, \cos 60 = \frac{1}{2}, \tan 60 = \sqrt{3}$

Exercise 4.2

1. Use the right-angled triangle to state the required trigonometric ratios:



(a) $\sin(\theta)$ **(b)** $\cos(\theta)$ **(c)** $\tan(\theta)$

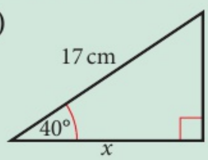
(d) $\sin(\beta)$ **(e)** $\cos(\beta)$ **(f)** $\tan(\beta)$

(g) How are the angles θ and β related to each other?

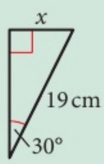
(h) Clearly describe the link between the trigonometric ratios of the angle θ and those of the angle β.

2. Find the missing side length in each triangle, writing your answer correct to 2 decimal places.

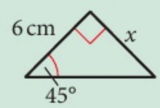
(a)



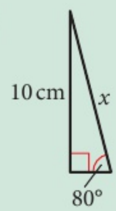
(b)



(c)

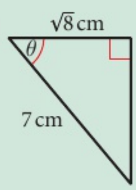


(d)

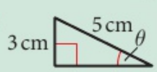


3. Find the value of the interior angle θ, correct to the nearest degree.

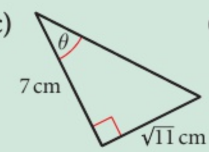
(a)



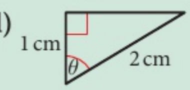
(b)



(c)



(d)

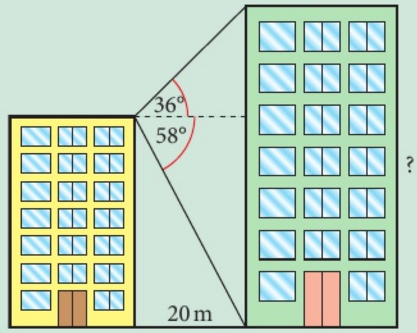


4. (a) A car drives 300 km from city A to city B on a bearing of 065°. How far north/south has the car driven? How far east/west has the car driven?

(b) Another car leaves city A and drives 300 km on a bearing of 110°. How far north/south has the car driven? How far east/west has the car driven?

(c) How do your calculations from **(a)** and **(b)** differ from each other?

5. The green and yellow buildings shown are 20 m apart along a street. The angle of elevation from the roof of the shorter building to the roof of the taller building is 36°. The angle of depression from the roof of the shorter building to the base of the taller building is 58°.



Find the height of the two buildings to the nearest metre.

6. A ladder that is 5 m long is to be placed against a wall that is 4.5 m high. The base of the ladder can be placed anywhere along a concrete strip that is 2 m in width. What is the smallest possible angle that the ladder can make with the ground?

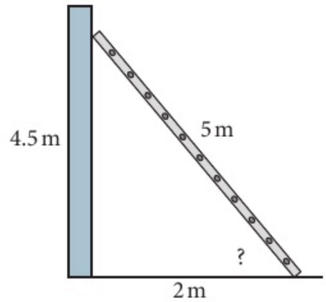
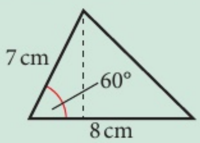


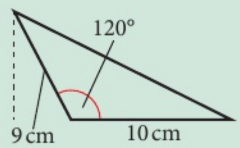
Figure 4.20 Diagram for question 6

7. Find the area of the following triangles.

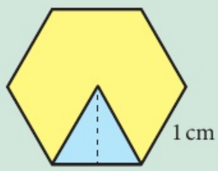
(a)



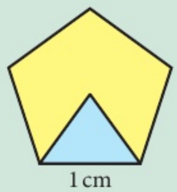
(b)



8. **(a)** Using trigonometric ratios, find the exact area of a regular hexagon with side length 1 cm.



(b) Using a similar procedure to **(a)**, find a formula, using trigonometric ratios, for the area of a regular pentagon with side length 1 cm.



(c) Explain clearly how you can find the area of a regular polygon with n sides.

9. By using $\theta = 10, 5, 2, 1, \frac{1}{2}$, explain how the values of $\sin\theta$ and $\cos\theta$ change as θ approaches zero.

4.3 Non-right-angled trigonometry

In this section, we extend the primary trigonometric ratios to triangles that contain no right angles. Although the primary trigonometric ratios cannot be directly applied to such triangles, we can develop relationships between the side lengths and interior angles of these triangles. These relationships form the sine and cosine rules. First, we develop the trigonometric ratios of angles larger than 90°.

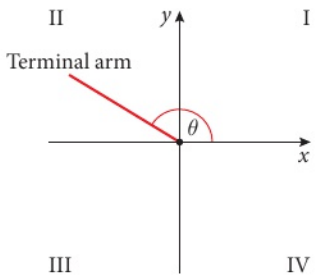


Figure 4.21 Angles larger than 90°

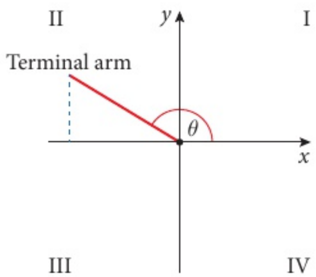


Figure 4.22 Trigonometric ratios

The angle θ will be defined as the angle that a terminal arm (red) makes with the positive x-axis, moving anti-clockwise. To calculate trigonometric ratios of angles larger than 90 degrees, draw the terminal arm of the angle and create a right-angled triangle in that quadrant. Label the base and height using the x and y values (which could now be negative!). The interior angle made by the terminal arm and the x-axis is acute and serves as the angle used to calculate the trigonometric ratio.

In this way, we can uniquely draw any angle θ, $0° \leqslant \theta \leqslant 360°$, in one of the four quadrants, labelled I, II, III, IV. The trigonometric ratios can then be calculated by projecting a line perpendicular to the x-axis (dashed line in Figure 4.22) and creating a right-angled triangle in the respective quadrant.

Example 4.15

By drawing an appropriate triangle, calculate

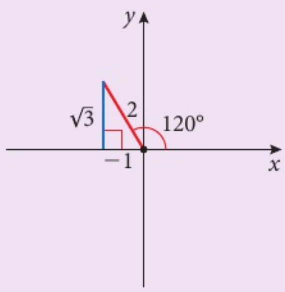
(a) $\sin 120°$ (b) $\cos 135°$ (c) $\tan 210°$ (d) $\sin 315°$

Solution

(a) By drawing a 120° angle, we create a right-angled triangle in quadrant II with an interior angle of 60°. Labelling the side lengths appropriately, we can calculate

$\sin 120 = \frac{\sqrt{3}}{2}$

Notice that $\cos 120 = -\frac{1}{2}$ since the adjacent side is on the negative x-axis.



(b) By drawing a 135° angle, as in Figure 4.23, we create a right-angled triangle in quadrant II with an interior angle of 45°. Labelling the side lengths appropriately, we can calculate $\cos 135 = -\frac{1}{\sqrt{2}}$

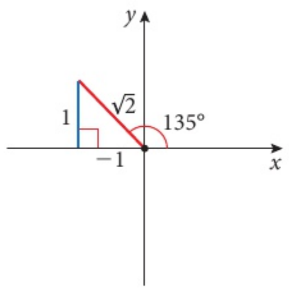
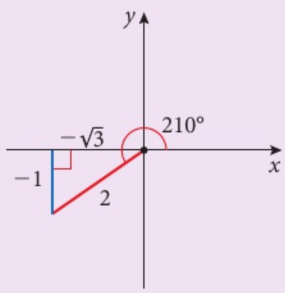


Figure 4.23 Diagram for Example 4.15 (b)

(c) By drawing a 210° angle, we create a right-angled triangle in quadrant III with an interior angle of 30°. Labelling the side lengths appropriately, we can calculate

$\tan 210 = \frac{1}{\sqrt{3}}$



(d) By drawing a 315° angle, as in Figure 4.24, we create a right-angled triangle in quadrant IV with an interior angle of 45°. Labelling the side lengths appropriately, we can calculate $\sin 315 = -\frac{1}{\sqrt{2}}$

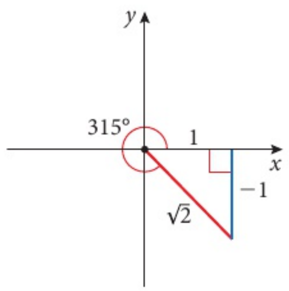


Figure 4.24 Diagram for Example 4.15 (d)

Example 4.16

(a) Show that for any angle θ, $0° \leqslant \theta \leqslant 90°$, $\sin\theta = \sin(180 - \theta)$

(b) Find two different angles θ with $\sin\theta = 0.65$

(c) Show that in a triangle containing an obtuse interior angle θ and two enclosing sides with lengths a and b, the area is given by the formula $\frac{1}{2}ab\sin\theta$

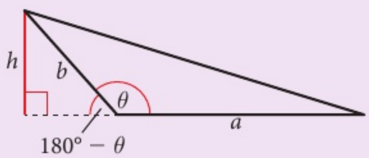
Solution

(a) In Figure 4.25, we see that $\sin\theta = \sin(180 - \theta) = \frac{a}{c}$

(b) Using a GDC, we obtain $\theta = \sin^{-1}0.65 \approx 40.5$

Using part (a), $\theta = 180 - 40.5 = 139.5°$ is another angle satisfying $\sin\theta = 0.65°$

(c) In the triangle shown, $h = b\sin(180 - \theta) = b\sin\theta$ and the area is then $\frac{1}{2}ab\sin\theta$



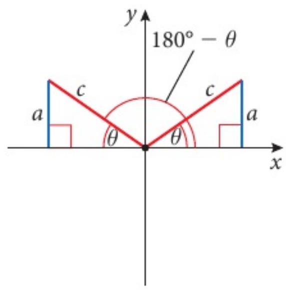


Figure 4.25 Solution to Example 4.16 (a)

Now that the trigonometric ratios for obtuse angles have been defined, we can consider how to find missing side lengths and interior angles of triangles that contain no right angles, or those containing obtuse angles.

The sine rule

Consider the following triangle:

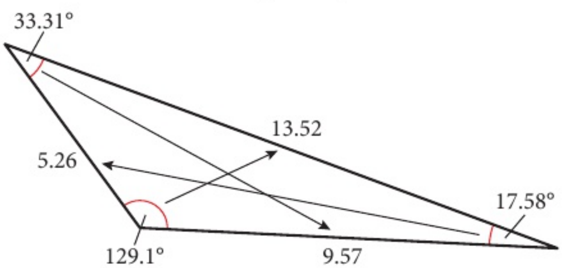


Figure 4.26 Notice the remarkable relationship

If we calculate the ratios: $\dfrac{13.52}{\sin 129.1} \approx 17.42$, $\dfrac{9.57}{\sin 33.31} \approx 17.42$, $\dfrac{5.26}{\sin 17.58} \approx 17.42$

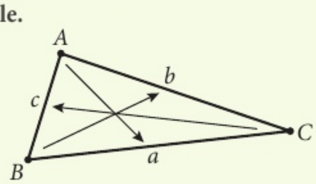
we notice a remarkable relationship that turns out to be true for all triangles.

The ratio of a side length of a triangle to the sine of the interior opposite angle of the triangle is constant. This relationship is called **the sine rule.**

The sine rule

In any triangle with side lengths, a,b,c and opposite interior angles A,B,C

$$\frac{a}{\sin A} = \frac{b}{\sin B} = \frac{c}{\sin C} \quad \text{or} \quad \frac{\sin A}{a} = \frac{\sin B}{b} = \frac{\sin C}{c}$$



The sine rule allows us to find missing side lengths or interior angles of a triangle, provided we know a side length and its opposite interior angle.

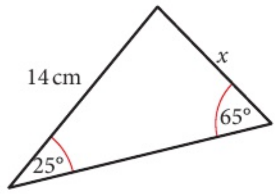


Figure 4.27 Diagram for Example 4.17

Example 4.17

Find the value of the missing side length x in the triangle shown in Figure 4.27, correct to 2 decimal places.

Solution

Since we have a side length and the opposite interior angle, we can apply the sine rule to obtain:

$$\frac{14}{\sin 65} = \frac{x}{\sin 25}$$

$$x = \frac{14 \sin 25}{\sin 65}$$

$$x \approx 6.53\,\text{cm}$$

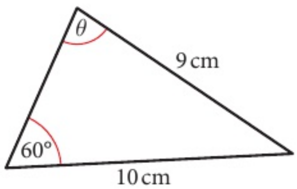


Figure 4.28 Diagram for Example 4.18

Example 4.18

Find the missing value of θ in the triangle shown in Figure 4.28.

Solution

Using the sine rule, we obtain $\dfrac{\sin 60}{9} = \dfrac{\sin\theta}{10}$

$$\sin\theta = \frac{10 \sin 60}{9}$$

$$\theta \approx 74.2°$$

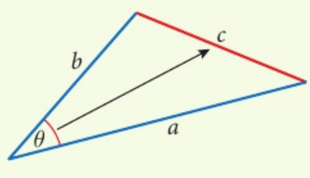
The cosine rule

If in a non-right-angled triangle, only two side lengths a and b and the enclosed angle θ between them are known, we cannot use the sine rule. In this case, we can calculate the side length c (red) located opposite to the known interior angle. If we know all three sides of the triangle, we can also calculate the value of any of the interior angles of the triangle. The formula that enables us to do so is called the **cosine rule.**

The cosine rule

In a triangle in which two side lengths a and b, as well as the enclosed interior angle θ, are known, the side c opposite to θ is given by

$$c^2 = a^2 + b^2 - 2ab\cos\theta$$



If all three sides are known, then an interior angle θ can be found using

$$\cos\theta = \frac{a^2 + b^2 - c^2}{2ab}, \text{ or } \theta = \cos^{-1}\left(\frac{a^2 + b^2 - c^2}{2ab}\right).$$

where c is the side of the triangle opposite to θ and a and b are the sides enclosing θ.

One way to understand the cosine rule is to notice that the $c^2 = a^2 + b^2$ portion is similar to the Pythagorean theorem for right-angled triangles. If $\theta < 90°$, then $a^2 + b^2 > c^2$ and $2ab\cos\theta$ is the amount that must be subtracted to create equality (Figure 4.29). If $\theta > 90°$, then $a^2 + b^2 < c^2$ and $2ab\cos\theta$ is the amount that must be added to create equality (Figure 4.30).

If $\theta = 90°$, then $\cos\theta = 0$ and the cosine rule collapses to the Pythagorean theorem for right-angled triangles.

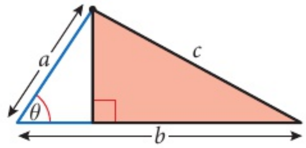


Figure 4.29 Interpretation of the cosine rule when $\theta < 90°$

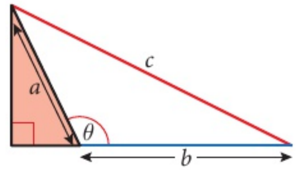
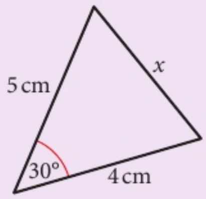


Figure 4.30 Interpretation of the cosine rule when $\theta > 90°$

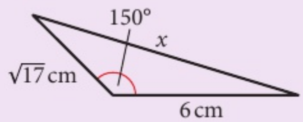
Example 4.19

Find the value of the missing side length x, correct to 2 decimal places.

(a)



(b)



Solution

(a) Using the cosine rule,

$$x^2 = 4^2 + 5^2 - 2(4)(5)\cos 30$$

$$x = \sqrt{16 + 25 - 40\cos 30}$$

$$x \approx 2.52\text{ cm}$$

(b) Using the cosine rule,

$$x^2 = 6^2 + \sqrt{17}^2 - 2(6)(\sqrt{17})\cos 150$$

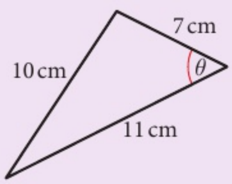
$$x = \sqrt{36 + 17 - 12\sqrt{17}\cos 150}$$

$$x \approx 9.79\text{ cm}$$

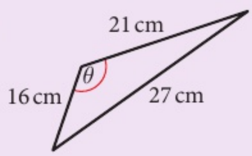
Example 4.20

In each of these triangles, find the angle θ, to the nearest degree.

(a)



(b)

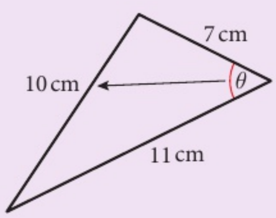


Solution

(a) Using the cosine rule, we obtain

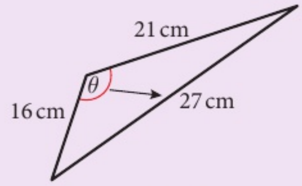
$$\theta = \cos^{-1}\left(\frac{7^2 + 11^2 - 10^2}{2 \times 7 \times 11}\right)$$

$$\theta \approx 64°$$



Notice the importance of recognising the side opposite to θ.

(b) Using the cosine rule, we obtain



$$\theta = \cos^{-1}\left(\frac{21^2 + 16^2 - 27^2}{2 \times 21 \times 16}\right)$$

$$\theta \approx 93°$$

Notice that if $\theta > 90°$, $\cos\theta < 0$, so that the expression in brackets is negative. One could calculate the expression in brackets before using $\cos^{-1}$, but this could lead to inaccuracy due to rounding.

When using the cosine rule to calculate missing angles, the expression in the brackets must take a value between -1 and 1.

We conclude this section by considering examples in which both the sine and cosine rules are required in order to obtain missing information in triangles. Example 4.22 is an application of non-right-angled triangle trigonometry to navigation, and requires us to construct the appropriate triangle first before we can calculate missing information.

Example 4.21

For the triangle shown in Figure 4.31, find

(a) the side length x (b) the angle θ (c) the area of the triangle.

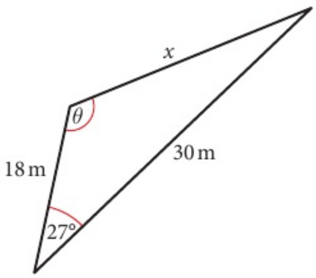


Figure 4.31 Diagram for Example 4.21

Solution

(a) Using the cosine rule,

$$x^2 = 30^2 + 18^2 - 2(30)(18)\cos 27$$

$$x = \sqrt{900 + 324 - 1080\cos 27}$$

$$x \approx 16.2\text{ m}$$

(b) We can use the answer from (a) in the sine rule to obtain

$$\frac{\sin 27}{16.2} = \frac{\sin\theta}{30}$$

$$\sin\theta = \frac{30\sin 27}{16.2}$$

$$\theta \approx 57.2°$$

However, from the diagram, we see that the angle is obtuse, thus $\theta = 180 - 57.2 = 123.8°$

(c) The area of the triangle is $\frac{1}{2}(18)(30)\sin 27 \approx 123\text{ m}^2$

Example 4.22

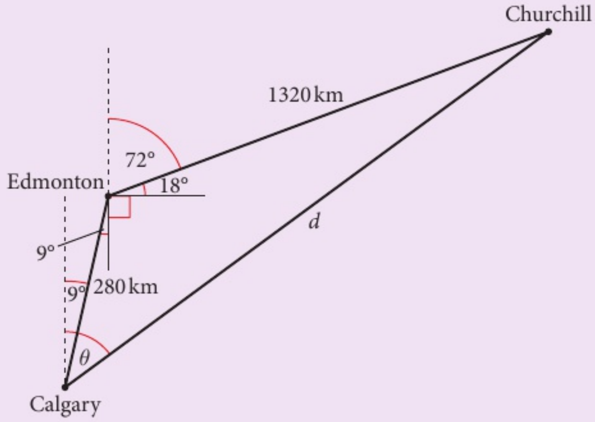
An aeroplane flies 280 km on a bearing of 009° from Calgary, Alberta in Canada to the provincial capital Edmonton. From Edmonton, the plane flies 1320 km on a bearing of 072° to Churchill, Manitoba, located on the shores of Hudson Bay.

(a) How far is Churchill from Calgary?

(b) If an aeroplane flies at an average speed of $720\,\text{km}\,\text{h}^{-1}$, how long is a direct flight from Calgary to Churchill to the nearest hour?

Solution

We can create the following diagram, labelling the cities and the given information:



(a) The distance, d, between Calgary and Churchill can be found using the cosine rule with the triangle in the diagram, using the interior angle of $18 + 90 + 9 = 117°$

$d^2 = 280^2 + 1320^2 - 2(280)(1320)\cos(117)$

$d \approx 1468\,\text{km}$

(b) The time taken to complete the flight is $\frac{1468}{720} \approx 2.04$

Exercise 4.3

1. Fill in the table with a + or − as necessary. The first row is filled in for you.

	$\sin\theta$	$\cos\theta$	$\tan\theta$
$0° < \theta < 90°$	+	+	+
$90° < \theta < 180°$			
$180° < \theta < 270°$			
$270° < \theta < 360°$			

2. Find two different angles θ for which $\sin\theta = 0.707$, correct to the nearest degree.

3. If $\theta > 0$, we define the angle $-\theta$ to be the angle a terminal arm makes with the positive x-axis, moving clockwise. By drawing an appropriate right-angled triangle, show that

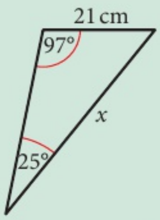
(a) $\cos\theta = \cos(-\theta)$

(b) $\sin(-\theta) = -\sin\theta$

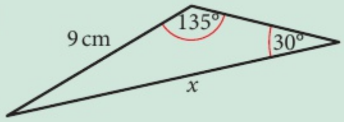
4. By drawing an appropriate right-angled triangle, find exact values of cos240, sin240, tan240 using fractions with roots.

5. Use the sine rule to find the missing side length, x, correct to 3 significant figures.

(a)

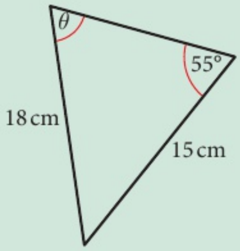


(b)

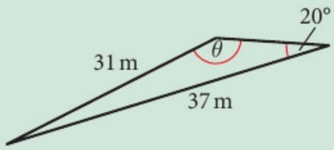


6. Use the sine rule to find the value of the angle θ, correct to the nearest degree.

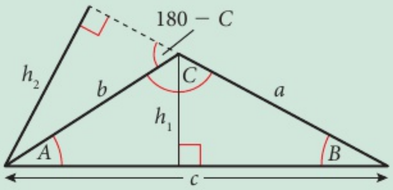
(a)



(b)



7. Using the diagram, complete the following steps to derive the sine rule.



(a) Show that $h_1 = a\sin B = b\sin A$

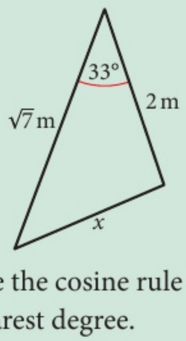
(b) Conclude that $\dfrac{\sin B}{b} = \dfrac{\sin A}{a}$

(c) Show that $h_2 = c\sin B = b\sin(180 - C)$

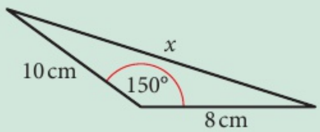
(d) Conclude that $\dfrac{\sin B}{b} = \dfrac{\sin C}{c}$, which completes a derivation of the sine rule.

8. Use the cosine rule to find the missing side length, x, in each triangle, correct to 3 significant figures.

(a)

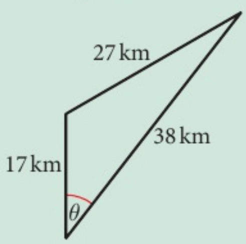


(b)

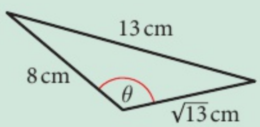


9. Use the cosine rule to find the value of the missing angle θ, correct to the nearest degree.

(a)

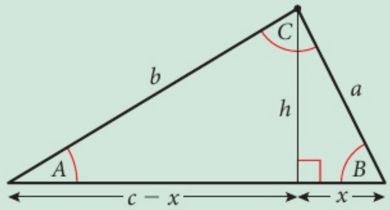


(b)



10. Clearly explain the conditions under which you would use the sine rule and the cosine rule to find out information about a triangle.

11. Using the diagram, complete the following steps to derive the cosine rule.



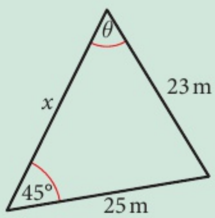
   (a) Show that $x = a\cos B$

   (b) Using the Pythagorean theorem in two right-angled triangles, find two different, equivalent expressions for h^2 in terms of x.

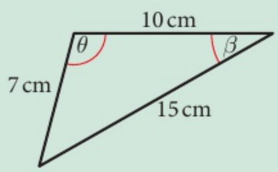
   (c) Using your answers from (a) and (b), show that $b^2 = a^2 + c^2 - 2ac\cos B$

12. For each triangle, find the missing angles and side lengths, giving your answers to 1 decimal place.

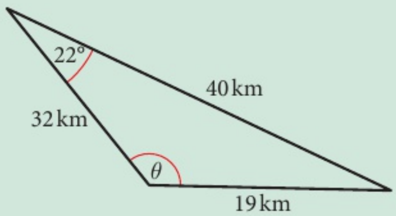
   (a)



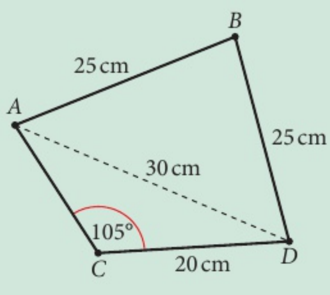
   (b)



   (c)



13. Find the area of quadrilateral $ABCD$.



14. A car drives 100 km from city A to city B on a bearing of 050°. The car then drives an additional 250 km from city B to C on a bearing of 030°.

   (a) Find the distance between cities A and C if there is a direct road connecting the two cities.

   (b) On what bearing must one drive from A to C?

   (c) If a car drives an average speed of $110\,\text{km}\,\text{h}^{-1}$, how long does it take to drive directly from A to C, correct to the nearest minute?

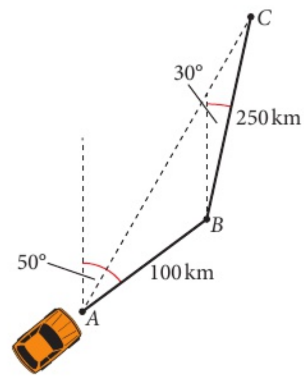
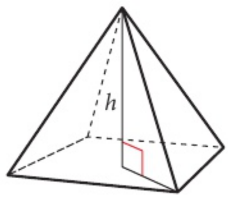
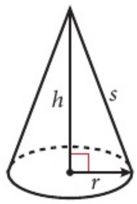
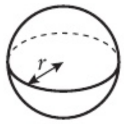
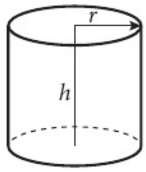


Figure 4.32 Diagram for question 14

4.4 Surface area and volume

In this section, we consider the geometric properties of 3-dimensional solids such as right pyramids, right cones, cylinders, spheres, hemispheres, cuboids and prisms, as well as compound shapes composed from them. Remember that **volume**, V, is the amount of space that is contained in a solid, while **surface area**, S, is the amount of material required to create the surface of the solid. The following table shows the solids we will be using, and important formulae needed to compute with them.

				
Name of solid	Right (square) pyramid	Right circular cone	Sphere	Cylinder
Volume	$V = \frac{1}{3}(\text{Area}_{\text{base}})h$	$V = \frac{1}{3}\pi r^2 h$	$V = \frac{4}{3}\pi r^3$	$V = \pi r^2 h$
Surface area	$S = \text{Area}_{\text{base}} + \text{Area}_{\text{triangles}}$	$S = \pi r^2 + \pi rs$	$S = 4\pi r^2$	$S = 2\pi r^2 + 2\pi rh$

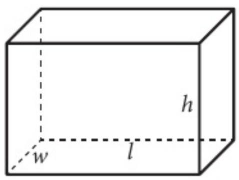
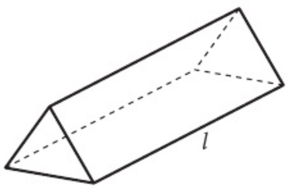
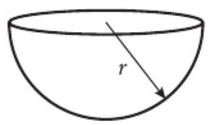
			
Name of solid	Cuboid	(Triangular) Prism	Hemisphere
Volume	$V = lwh$	$V = \text{Area}_{\text{base}} \times l$	$V = \frac{2}{3}\pi r^3$
Surface area	$S = 2lw + 2wh + 2lh$	$S = 2\,\text{Area}_{\text{base}} + \text{Area}_{\text{rectangles}}$	$S = 2\pi r^2$

Table 4.1 Examples of 3D solids

Example 4.23

An ice cream cone has the shape of a right circular cone with radius 3 cm and a height of 8 cm.

(a) Calculate the volume of the ice cream cone.

(b) Calculate the surface area of the ice cream cone.

Solution

(a) Using $r = 3\,\text{cm}$, $h = 8\,\text{cm}$, the volume is given by
$V = \frac{1}{3}\pi(3)^2(8) \approx 75.4\,\text{cm}^3$

(b) To calculate the surface area, we require the slanted height, s.
This can be found using the Pythagorean theorem as follows:

$$3^2 + 8^2 = s^2$$
$$s = \sqrt{9 + 64}$$
$$s = \sqrt{73}\,\text{cm}$$

We now use the formula $S = \pi rs$, as the ice cream cone is open at the top and does not have a circular base. Thus, $S = \pi(3)(\sqrt{73}) \approx 80.5\,\text{cm}^2$

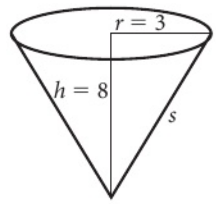


Figure 4.33 Solution to Example 4.23

Example 4.24

A spherical rubber ball has a radius of 10 cm. Calculate the volume and surface area of the ball and interpret your answer.

Solution

The volume of the ball is $V = \frac{4}{3}\pi 10^3 \approx 4189\,\text{cm}^3$. This is approximately the amount of space available inside the ball for air. The surface area of the ball is $A = 4\pi 10^2 \approx 1257\,\text{cm}^2$. If the ball was made from a thin sheet of rubber, this would be the area of the rubber sheet out of which the ball was made.

The following example involves a compound shape, composed of the body of a cylinder and hemisphere.

Example 4.25

A grain container has the shape of a cylinder with radius 3.5 m and a height of 7 m.

The roof of the container is a hemisphere, as shown in Figure 4.34

(a) Find the volume of the entire grain container.

(b) Find the surface area of the grain container.

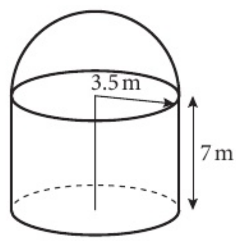


Figure 4.34 Diagram for Example 4.25

Solution

(a) The volume of the grain container is the sum of the volumes of the cylinder and hemisphere, $V = V_{\text{cylinder}} + V_{\text{hemisphere}}$. Notice that the cylinder and hemisphere have the same radius of 3.5 m. The volume is therefore $V = \pi(3.5)^2(7) + \frac{2}{3}\pi(3.5)^3 \approx 359\,\text{m}^3$

(b) In calculating the surface area, we notice that the base of the cylinder and hemisphere are not included. The surface area of the grain container is the sum of the surface areas of the bodies of the cylinder and hemisphere.

$S = S_{\text{cylinder}} + S_{\text{hemisphere}} = 2\pi(3.5)(7) + 2\pi(3.5)^2 \approx 231\,\text{m}^2$

In the following examples, trigonometry must be used on triangles within our solids to enable us to complete the necessary calculations.

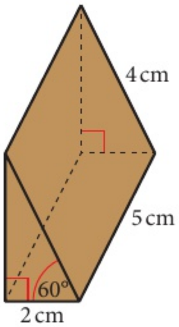


Figure 4.35 Diagram for Example 4.26

Example 4.26

A chocolate bar is in the shape of a right triangular prism as shown in Figure 4.35.

(a) Calculate the volume of the chocolate bar.

(b) Calculate the surface area of the chocolate bar.

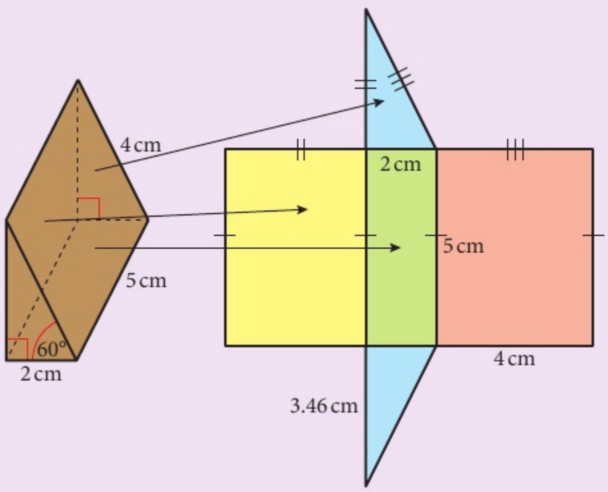
Solution

(a) In order to calculate the volume and surface area of the chocolate, we require the height, which can be calculated using

$\sin 60 = \dfrac{\text{height}}{4} \Rightarrow \text{height} \approx 3.46\,\text{cm}$

The area of the triangular base is $\frac{1}{2}(2)(3.46) = 3.46\,\text{cm}^2$ and the volume is $V = (3.46)(5) = 17.3\,\text{cm}^3$

(b) To calculate the surface area, we use the following diagram, called a **net**, that shows all the faces of the solid.



From the net, the surface area is:

$S = 2\left(\frac{1}{2}(2)(3.46)\right) + (4)(5) + (2)(5) + (5)(3.46)$

$S = 6.92 + 20 + 10 + 17.3$

$S = 54.22\,\text{cm}^2$

The volume represents how much chocolate is in the chocolate bar, and the surface area gives the minimum area of the paper used to package the chocolate bar.

Example 4.27

The Cheops Pyramid in Giza, Egypt, shown in Figure 4.36, is a right square pyramid with a base length and width of 230 m and a height of 147 m. Three of the base vertices are labelled B, C and D, while the apex is labelled A. The projection of the apex on the base is labelled O.

(a) Find the volume of the Cheops Pyramid.

(b) Find the length of the diagonal BD.

(c) Find the length of AB.

(d) Find the size of angle BAC and angle OBA.

(e) Hence, find the area of the triangle ABC and the lateral area of the Cheops Pyramid (lateral area is the total area of the triangular faces of the pyramid).

Diagram NOT to scale

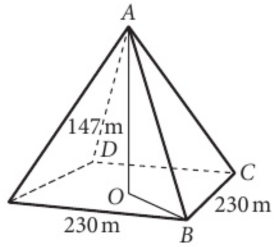


Figure 4.36 Diagram for Example 4.27

Solution

(a) The volume of the pyramid is

$V = \frac{1}{3}(230)^2(147) = 2\,592\,100\,\text{m}^3$

(b) Using the Pythagorean theorem on the right-angled triangle BCD yields

$BD^2 = 230^2 + 230^2$

$BD = \sqrt{2(230)^2} \approx 325\,\text{m}$

(c) AB can be obtained using the Pythagorean theorem on the triangle AOB, with $OB = \frac{1}{2}(BD) \approx 162.5\,\text{m}$, and $AB^2 = 162.5^2 + 147^2$, and

$AB \approx 219\,\text{m}$

(d) Angle $B\hat{A}C$ can be found by using the cosine rule on the isosceles triangle BAC, where $BA = CA = 219\,\text{m}$ and $BC = 230\,\text{m}$.

As a result, $B\hat{A}C = \cos^{-1}\left(\frac{2(219)^2 - 230^2}{2(219)^2}\right) \approx 63°$

$O\hat{B}A$ can be found using the right-angled triangle OBA with

$O\hat{B}A = \sin^{-1}\left(\frac{147}{219}\right) \approx 42°$

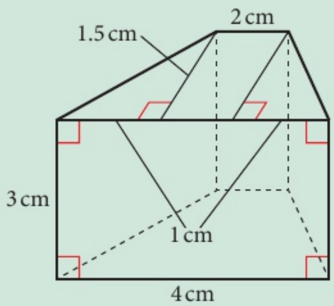
(e) The area of triangle ABC is $\frac{1}{2}(219)^2 \sin 63 \approx 21\,367\,\text{m}^2$ and the lateral area of the pyramid is given by

$4 \times \text{Area}_{\text{triangle}ABC} = 4(21\,367) = 85\,468\,\text{m}^2$

Exercise 4.4

1. For each of the solids described below, calculate
   (i) the volume
   (ii) the surface area.
   (a) A cuboid with length 3 cm, width 4 cm and height 7 cm.
   (b) A cylinder with radius 1.2 cm and a height of 2.3 cm.
   (c) A sphere with radius 6 cm.
   (d) A triangular prism 5 cm long with an equilateral triangular base and side length 2 cm.
   (e) A right circular cone with a radius of 5.4 m and a slanted height of 8 m.
   (f) A hemisphere with diameter 8 cm.

2. Find the volume and surface area of each of these compound shapes.
   (a)



   (b)

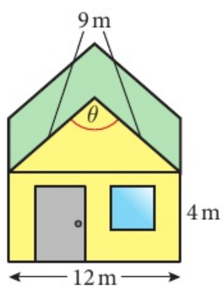
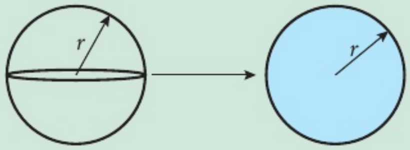


Figure 4.37 Diagram for question 3

3. Consider the house drawn in Figure 4.37, which is 17 m long. Find
   (a) the angle θ
   (b) the volume of the house
   (c) the surface area of the house.
   (d) For what practical purposes might you need to know the surface area and volume of a house?

4. The Earth has a radius of 6300 km and Jupiter has a radius of 69 900 km. Both planets are approximately spherical in shape.
   (a) Calculate the volume of the Earth and Jupiter, writing your answer in standard form.
   (b) How many times larger is the volume of Jupiter than the volume of Earth?
   (c) Calculate the surface area of the Earth and Jupiter, writing your answers in standard form.
   (d) How many times larger is the surface area of Jupiter than the surface area of Earth?

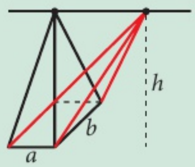
5. **(a)** The equatorial circle is the cross section of a sphere resulting from cutting it in half along the equator (see circle on right in the diagram). How does the surface area of a sphere compare to the area of one of its equatorial circles?



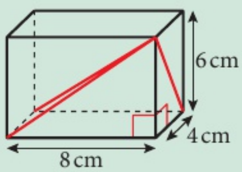
(b) How does the volume of a sphere change if you double its radius?

(c) How does the surface area of a sphere change if you double its radius?

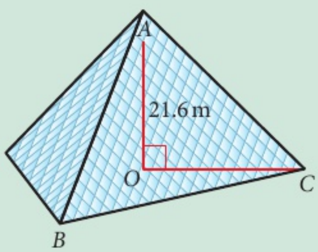
6. **(a)** What can you say about two pyramids that have the same base area and the same height (as shown in the diagram)?



(b) Find the volume and surface area of the pyramid in the diagram.



7. The glass pyramid shown in the diagram is a representation of part of the Louvre Museum in Paris, France. Its square base has a length and width of 34 m and a height of 21.6 m. The apex is labelled A, while B and C are corners on the base and O is the projection of A on the base.



(a) Find the length of OC.

(b) Find angle OCA, the slant of the pyramid.

(c) Find $B\hat{A}C$.

(d) Find the area of triangle BAC and hence, the surface area of the pyramid.

(e) Find the volume of the pyramid.

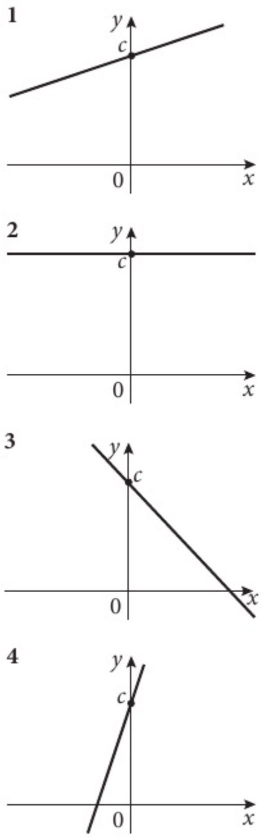


Figure 4.38 Diagrams for question 1

Chapter 4 practice questions

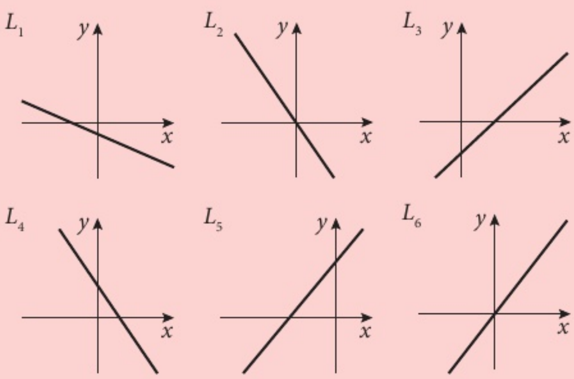
1. The four diagrams in Figure 4.38 show the graphs of four different straight lines, all drawn to the same scale. In each of them, c is a positive constant.

Copy the table and write the number of the diagram whose straight line corresponds to the equation in the table.

Equation	Diagram number
$y = c$	
$y = -x + c$	
$y = 3x + c$	
$y = \frac{1}{3}x + c$	

2. The following diagrams show six lines with equations of the form

$y = mx + c$



In the table there are four possible conditions for the pair of values m and c.

Match each of the given conditions with one of the lines drawn above.

Condition	Line
$m > 0$ and $c > 0$	
$m < 0$ and $c > 0$	
$m < 0$ and $c < 0$	
$m < 0$ and $c = 0$	

3. The straight line, L, passes through the points $A(-1, 4)$ and $B(5, 8)$.

(a) Calculate the gradient of L.

(b) Find the equation of L.

The line L also passes through the point $P(8, y)$.

(c) Find the value of y.

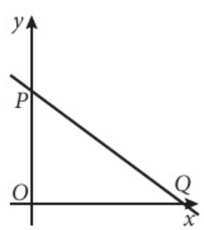


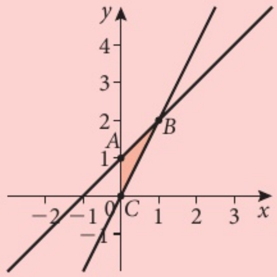
Figure 4.39 Diagram for question 4

4. The diagram shows the line PQ, whose equation is $x + 2y = 12$. The line intercepts the axes at P and Q respectively.

(a) Find the coordinates of P and of Q.

(b) A second line with equation $x - y = 3$ intersects the line PQ at the point A. Find the coordinates of A.

5. A triangle ABC is created using the intersections of the lines $x = 0$, $y = x + 1$ and $y = 2x$, as shown in the diagram:



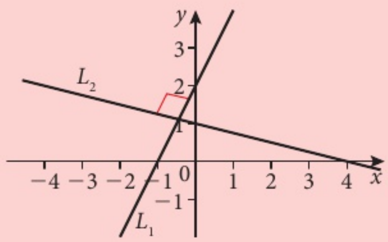
(a) Find the points A, B and C algebraically.

(b) Find the lengths AB, AC and BC

(c) Calculate the angle $A\hat{C}B$

(d) Hence, find the area of the triangle ABC

6. A student has drawn the two straight line graphs L_1 and L_2 and marked in the angle between them as a right angle, as shown. The student has drawn one of the lines incorrectly.



Consider L_1 with equation $y = 2x + 2$ and L_2 with equation $y = -\frac{1}{4}x + 1$

(a) Write down the gradients of L_1 and L_2 using the given equations.

(b) Which of the two lines has the student drawn incorrectly?

(c) How can you tell from the answer to part **(a)** that the angle between L_1 and L_2 should not be 90°?

(d) Draw the correct version of the incorrectly drawn line on the diagram.

7. A projector is mounted on a vertical wall AB. The beam of light from the projector makes an angle of 20° at A. The light makes a circle with diameter 9 m on the horizontal ground, as shown in Figure 4.40. The distance from the projector to the closest point on the circle, C, is 12 m. The distance between B and C is 10 m.

(a) Find $C\hat{D}A$. Give your answer correct to 2 decimal places.

(b) A 3.6 m high wall has to be erected between B and C. Find the furthest point from B at which this wall can be placed without blocking the light from reaching the circular area. Give your answer to the nearest metre.

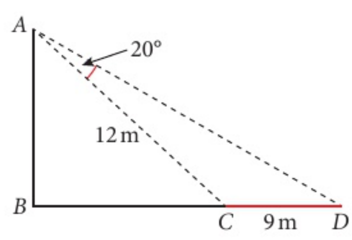
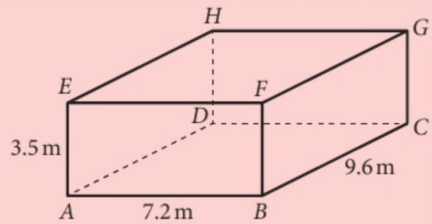


Figure 4.40 Diagram for question 7

8. A room is in the shape of a cuboid. Its floor measures 7.2 m by 9.6 m and its height is 3.5 m.



(a) Calculate the length of AC.

(b) Calculate the length of AG.

(c) Calculate the angle that AG makes with the floor.

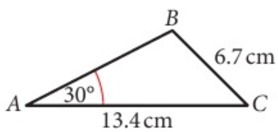


Figure 4.41 Diagram for question 9

9. Figure 4.41 shows triangle ABC in which $B\hat{A}C = 30°$, $BC = 6.7$ cm and $AC = 13.4$ cm

(a) Calculate the size of angle $A\hat{C}B$

Nadia makes an accurate drawing of triangle ABC. She measures angle $B\hat{A}C$ and finds it to be 29°.

(b) Calculate the percentage error in Nadia's measurement of $B\hat{A}C$

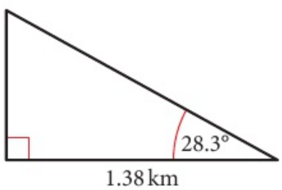


Figure 4.42 Diagram for question 10

10. José stands 1.38 km from a vertical cliff.

(a) Express this distance in metres.

José estimates the angle between the horizontal and the top of the cliff as 28.3° and uses it to find the height of the cliff.

(b) Find the height of the cliff according to José's calculation. Express your answer in metres, to the nearest whole metre.

(c) The actual height of the cliff is 718 metres. Calculate the percentage error made by José when calculating the height of the cliff.

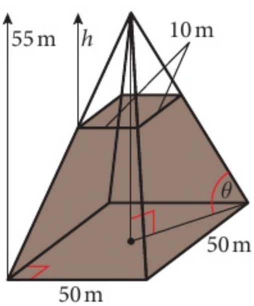


Figure 4.43 Diagram for question 11

11. A monument is shaped as a frustum (shaded in Figure 4.43) with a square base of length and width 50 m. The top of the frustum is a square with a length and width of 10 m. The black point is the projection of the apex on the base.

(a) Show that $h = 11$ m.

(b) Find the volume of the frustum.

(c) Calculate the angle θ, the slant of the frustum.

The notation $\sin^2\theta$ is used instead of $(\sin\theta)^2$.

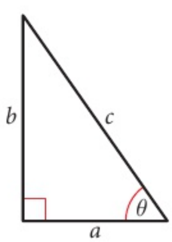


Figure 4.44 Diagram for question 12

12. A **trigonometric identity** is a formula relating trigonometric ratios that is true for all angles θ.

(a) Using $\theta = 30°, 45°$, verify the following trigonometric identities:

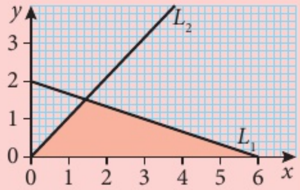
(i) $\tan\theta \equiv \dfrac{\sin\theta}{\cos\theta}$

(ii) $\sin^2\theta + \cos^2\theta \equiv 1$

(b) Using the triangle given, show that $\tan\theta \equiv \dfrac{\sin\theta}{\cos\theta}$ and $\sin^2\theta + \cos^2\theta \equiv 1$ for any angle θ, $0° < \theta < 90°$.

(c) The identity $\sin^2\theta + \cos^2\theta \equiv 1$ is often called the Pythagorean Identity. Using your working from **(b)**, explain why this is an appropriate name for the identity.

13. The diagram shows the straight lines L_1 and L_2. The equation of L_2 is $y = x$



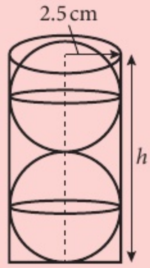
(a) Find
   (i) the gradient of L_1
   (ii) the equation of L_1
(b) Find the area of the shaded triangle.

14. 75 metal spherical cannon balls, each of diameter 10 cm, were excavated from a Napoleonic War battlefield.
(a) Calculate the total volume of all 75 metal cannon balls excavated.

The cannon balls are to be melted down to form a sculpture in the shape of a cone. The base radius of the cone is 20 cm.

(b) Calculate the height of the cone, assuming no metal is wasted.

15. Tennis balls with a radius of 2.5 cm are sold in pairs in a cylindrical can as shown.



(a) Find the height, h, of the can.
(b) Find the volume of the cylindrical can.
(c) Find the total volume of the tennis balls.
(d) What percentage of the can is filled with air?

16. A spherical statue is created by welding two hemispheres together with a metal band of length 12.56 m as shown in Figure 4.45.
(a) Calculate the radius of the sphere.
(b) Hence, calculate the volume and surface area of the sphere.

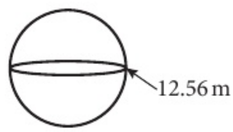
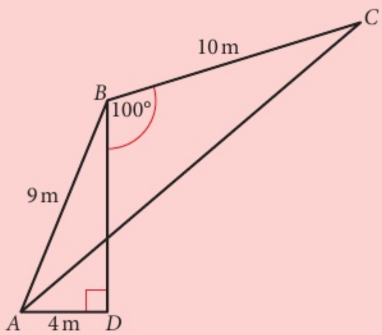


Figure 4.45 Diagram for question 16

17. In the diagram, $AD = 4\,\text{m}$, $AB = 9\,\text{m}$, $BC = 10\,\text{m}$, $B\hat{D}A = 90°$ and $D\hat{B}C = 100°$
(a) Calculate the size of $A\hat{B}C$.
(b) Calculate the length of AC.



18. Figure 4.46 shows a triangle ABC in which $AC = 17\,\text{cm}$. M is the midpoint of AC. Triangle ABM is equilateral.
(a) Write down
   (i) the length of BM in cm
   (ii) the size of angle BMC
   (iii) the size of angle MCB.
(b) Calculate the length of BC in cm.

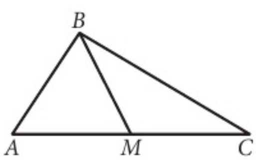


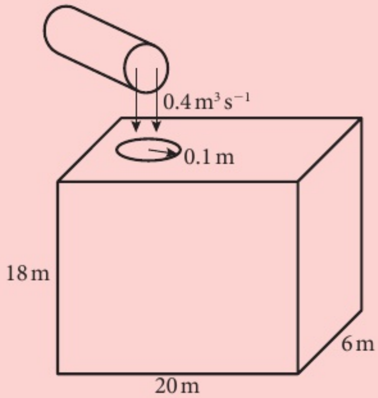
Figure 4.46 Diagram for question 18

19. A car drives 80 km from city A to city B on a bearing of 070°. From city B, the car drives an additional 200 km directly east to city C.

(a) Calculate the distance between city A and C.

(b) What is the bearing of C from A?

20. In a factory, contaminated water is pumped into a cuboid shaped container with dimensions $20\,\text{m} \times 18\,\text{m} \times 6\,\text{m}$ at a rate of $0.4\,\text{m}^3\,\text{s}^{-1}$, where it is purified.

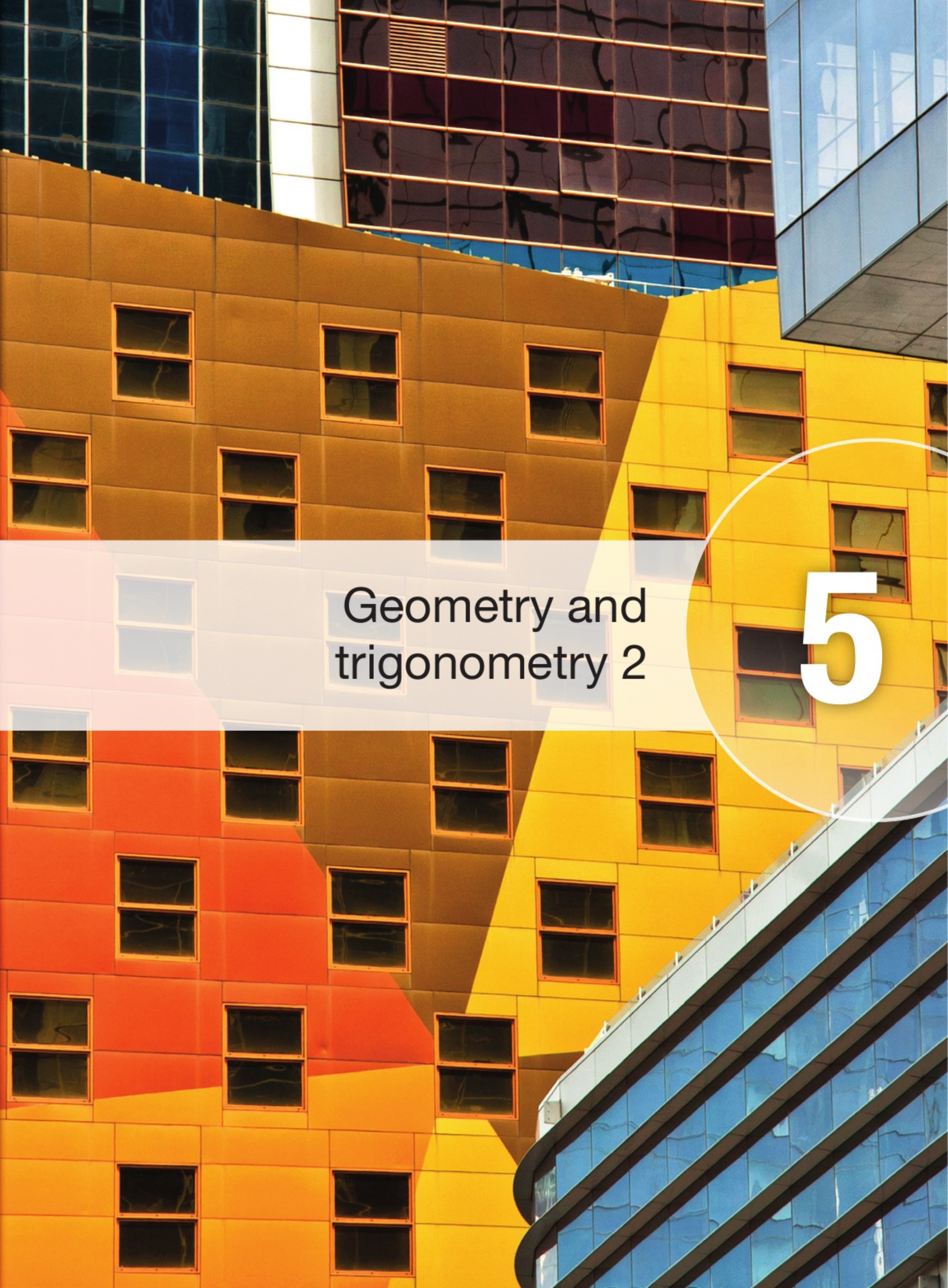


(a) Find the volume of the cuboid container.

(b) Calculate the time required to fill the container, correct to the nearest minute.

(c) The inner surface of the container is to be covered with an anti-corrosion spray. The opening to the container is a circle with radius 0.1 m. Calculate the surface area to be covered with the spray.

(d) After being purified, the water is pumped into a pond that is approximately cylindrical, with a diameter of 80 m. How much does the height of the pond increase by adding the water from the container, in cm?



5 Geometry and trigonometry 2

5 Geometry and trigonometry 2

Learning objectives

By the end of this chapter, you should be familiar with…

- calculating arc length and area of a sector of circle
- Voronoi diagrams, including terminology, interpolation, how to create them, and their applications.

You have already learned how to find many areas and lengths in triangles and other polygons. In this chapter we examine two important applications of geometry: measurement of arcs and sectors in a circle and Voronoi diagrams.

Voronoi diagrams find uses as varied as describing soap bubbles and animal dominance, analysing geographic markets, and partitioning regions for mobile phone coverage.

5.1 Arc length and area of a sector

A sector can be thought of as a pizza slice. This is an intuitive way to begin to understand how to calculate the area of a sector.

Example 5.1

Suppose you have a pizza with a total area of 120 cm². The pizza is cut into 8 equal slices as shown in the diagram.

(a) What is the area of one slice?

(b) What is the total area of 3 slices?

Solution

Use a ratio to solve the problem.

(a) Since we have cut the pizza into 8 equal slices, one slice is $\frac{1}{8}$ of the total area. Therefore, the area of one slice is $\frac{1}{8}(120) = 15\text{ cm}^2$

(b) Now we have 3 slices, hence $\frac{3}{8}$ of the total area: $\frac{3}{8}(120) = 45\text{ cm}^2$

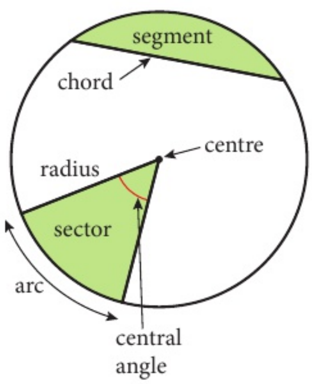


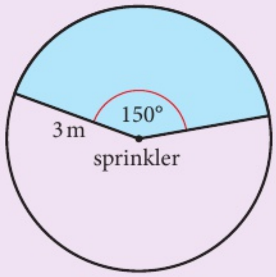
Figure 5.1 The parts of a circle

Area of a sector

Mathematically, a pizza slice is a **sector** of a circle. We are often interested in calculating the area of sectors, and we can use the method of Example 5.1 to calculate the area of any sector.

Example 5.2

On golf courses, the grass is carefully watered to make sure that the grass receives neither too much nor too little water. On a certain area of a golf course, a lawn sprinkler is set to spray an arc of 150° with a spray distance of 3 metres.



Calculate the area of grass that is watered by this sprinkler.

Solution

Since the sprinkler sprays an arc of 150°, the watered area is $\frac{150}{360}$ of the total circle area. To find the total circle area, we use the formula for the area of a circle: $A = \pi r^2$.

Therefore, the watered area is equal to the sector area

$$= \frac{150}{360}\pi(3)^2 = 3.75\pi \approx 11.8 \text{ m}^2$$

Length of an arc

Sometimes we are interested in the length of an **arc** on a circle.

We use the same logic we used for area of a sector: we need the size of the central angle for the arc, and then we calculate the portion of the circumference using a ratio.

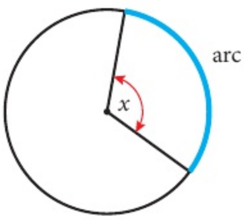
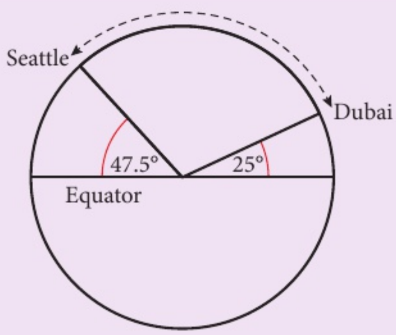


Figure 5.2 An arc and its central angle

Example 5.3

Some airplane routes come very close to flying over the North Pole. The route from Dubai, UAE, to Seattle, USA, comes very close to the North Pole. In this case, the distance of the flight can be estimated by the arc length.

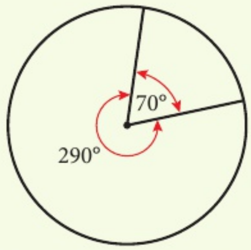


Given that Dubai's airport is at 25°N and Seattle's airport is at 47.5°N, as shown in the diagram, estimate the length of the flight between these two cities. The radius of the Earth is approximately 6370 km.

(a) Calculate the distance along the Earth's surface that the plane must travel.

(b) Assuming the plane travels at a constant elevation of 10 km above ground level, estimate the actual flight distance.

The area of a sector can be found using a fraction of the area of a circle. The numerator of the fraction is the central angle; the denominator is the whole circle (360°). The figure below shows two central angles, with measurements of 70° and 290°



Remember that a central angle is an angle with its vertex on the centre of a circle.

Area of sector $= \frac{\theta}{360}A$

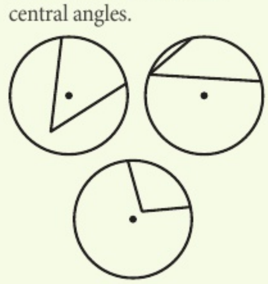
where A is the area of the circle and θ is the measure of the sector's central angle.

This can be written as

$A_{\text{sector}} = \frac{\theta}{360}\pi r^2$

where r is the radius of the circle and θ is the measure of the sector's central angle.

Beware! This method only works for central angle measurements. The angles below are not central angles.



Solution

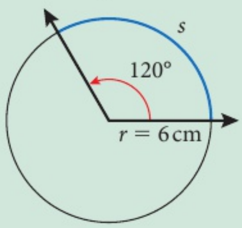
(a) First we must calculate the arc measure. Since the semicircle formed by the equator measures 180°, the arc measure must be $180 - 47.5 - 25 = 107.5°$
We are given that the radius of the Earth is 6370 km, so the distance along the Earth's surface is $\frac{107.5}{360}(2\pi)(6370) \approx 11\,952$ km
$= 12\,000$ km (3 s.f.)

(b) Our work in the previous question assumes that the plane travels along the Earth's surface – but planes fly above the Earth's surface! So, we must increase the radius of the circle to get a better estimate of the actual flight distance. The new radius will be $6370 + 10 = 6380$ km. Therefore we can estimate the actual flight distance will be
$\frac{107.5}{360}(2\pi)(6380) \approx 11\,970$ km $= 12\,000$ km (3 s.f.)
Note that our estimate of flight distance, to 3 significant figures, has not changed!

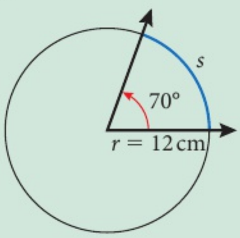
Exercise 5.1

1. Find the length of the arc s in each diagram.

(a)



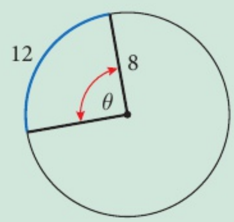
(b)



2. Find the area of the sector and the arc length determined by the given radius and central angle θ. Give your answer correct to 3 significant figures.

(a) $\theta = 30°, r = 10$ cm
(b) $\theta = 45°, r = 8$ m
(c) $\theta = 52°, r = 180$ mm
(d) $\theta = n, r = 15$ cm

3. Find the measure of angle θ in the diagram.



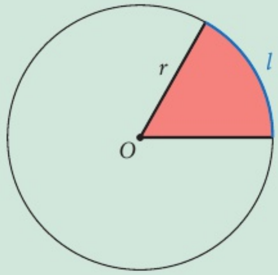
4. A bicycle with tyres 70 cm in diameter is travelling such that its tyres complete one and a half revolutions every second.

(a) The angular velocity of a rotating object represents the speed at which that object is turning. What is the angular velocity of this bicycle wheel in degrees per second?

(b) At what speed is the bicycle travelling along the ground, in km h^{-1}?

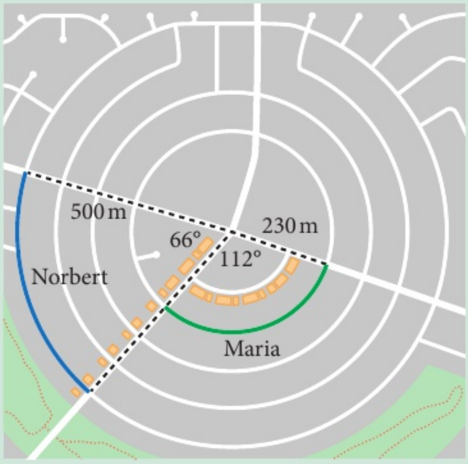
5. A sector of a circle with radius 4 cm has area $\frac{16}{5}\pi\,\text{cm}^2$. Find the measure of the central angle of the sector.

6. In circle O, the value of the area of the shaded sector is equal to the value of the length of arc l. Find the radius of the circle.



7. Many of the streets in Sun City, Arizona, USA, are formed by concentric circles. Maria and Norbert each go for a walk. Maria walks the path shown by the green arc, while Norbert walks along the path shown by the blue arc.

Who walks further, and by how much?



8. A circular irrigation system consists of a 400-metre pipe that is rotated around a central pivot point. The irrigation pipe makes one full revolution around the pivot point in a day. How much area, in square metres, is irrigated each hour?

9. Use the information given in Example 5.2 to complete the following exercise. In an arid climate, a golf course requires 2.5 mm of water per square metre of grass per day.

   (a) Calculate the volume of water required by the sector of grass in Example 5.2 in cm^3.

   The sprinkler delivers $800\,\text{cm}^3\,\text{min}^{-1}$ (flow rate).

   (b) Calculate the amount of time the sprinkler should run for in order to water the grass correctly.

   The sprinkler is now adjusted to spray in a 100° arc.

   (c) Assuming the same flow rate and spray radius, calculate the amount of time the sprinkler should run for in order to water the grass correctly.

5.2 Voronoi diagrams

A line goes through two points and has no endpoints.

A ray goes through two points and has exactly one endpoint.

A line segment goes through two points and has exactly two endpoints.

Emergency location services sometimes try to locate individuals based on their mobile-phone signal. The first step in this process can be to locate the individual based on the **nearest** mobile-phone antenna. A map to identify service regions for each mobile-phone antenna would look something like Figure 5.3.

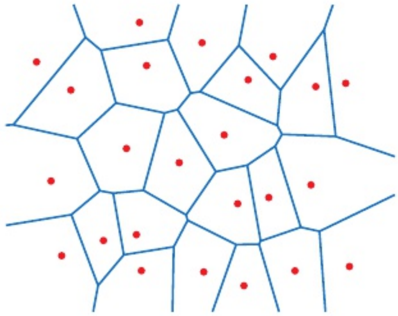


Figure 5.3 Each red dot indicates a mobile phone tower, and the blue lines are the boundaries of each region served by that tower

The map in Figure 5.3 is an example of a **Voronoi diagram**. The polygons on the map, each of which shows the service area for a particular mobile-phone antenna, are called **cells**. The dots represent mobile-phone antennas and are called **sites**. The line segment boundaries between each region are called **edges** and places where edges intersect are **vertices**. Notice that the some of the edges of cells around the outermost sites are actually rays, not line segments.

Geometrically, we know that the set of points that are equidistant from two points lies on the perpendicular bisector of the line segment between those two points. Therefore, every edge in a Voronoi diagram is a perpendicular bisector of a line segment connecting adjacent sites. We can use this fact to generate the equations of edges algebraically and to reason about points on the edges of a Voronoi diagram.

A **Voronoi diagram** divides a plane into a number of regions called **cells**. Cells are divided by boundaries called **edges**. Each cell contains exactly one **site**, such that every point in a given cell is closer to that cell's site than any other site. Points on edges are equidistant from two or more sites, **vertices** formed by edge intersections are equidistant from at least three sites.

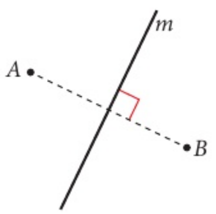
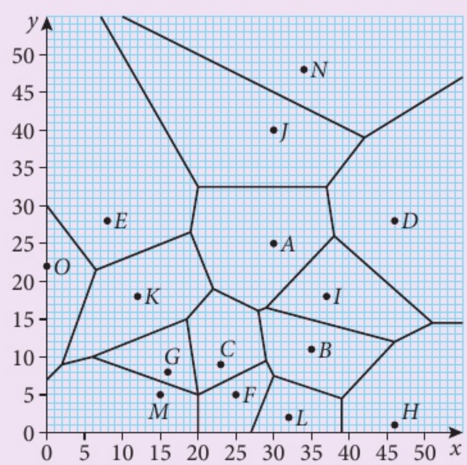


Figure 5.4 Recall that the set of points equidistant to A and B is the perpendicular bisector of the line segment with endpoints A and B. In other words, any point on the perpendicular bisector of segment AB is equidistant to points A and B

Example 5.4

Using the Voronoi diagram:



(a) identify the site closest to (30, 20)

(b) give the coordinates of a vertex equidistant to four sites

(c) estimate the coordinates of a point equidistant to sites A, E, and K

(d) explain why the diagram cannot give us a point equidistant to sites C, F, I, and L

(e) the point (a, a) is equidistant to sites J and N. Find the value of a

(f) find the equation of the edge between sites $E(8, 28)$ and $K(12, 18)$ in the form $y = mx + c$

Edges in Voronoi diagrams mark points that are equidistant from at least two adjacent sites.

Solution

(a) The point (30, 20) is within the cell containing A, so A must be the closest site.

(b) The vertex located at (20, 5) is formed by the intersection of four edges. Therefore, it must be equidistant to four sites, namely, C, G, M, and F.

(c) The vertex located at approximately (19, 26.5) is equidistant to sites A, E, and K since it is on the edges of the cells for all three sites.

(d) In this diagram, the edges of cells for sites C, F, I, and L do not share a common vertex. Therefore, this diagram cannot give us a point equidistant to those four sites.

(e) Points equidistant from J and N must be on the edge separating the cells of J and N. On that edge, there is only one point with equal x- and y-coordinates: the point (40, 40). Therefore, $a = 40$.

(f) To find the equation of the edge, we must find the midpoint of segment $\overline{EK}$ and the gradient of a line perpendicular to $\overline{EK}$. First, the midpoint:

$$\text{Midpoint}_{EK} = \left(\frac{8+12}{2}, \frac{28+18}{2}\right) = (10, 23)$$

Next, we find the gradient of $\overline{EK}$:

$$m_{EK} = \frac{28-18}{8-12} = \frac{10}{-4} = -\frac{5}{2}$$

The gradient of the edge is therefore $\frac{2}{5}$

Finally, use point–slope form to generate the equation of the line:

$$y - 23 = \frac{2}{5}(x - 10)$$

$$\Rightarrow y = \frac{2}{5}x + 19$$

For part (e), it may be helpful to draw the line $y = x$ on the diagram, since it contains all points with equal x and y coordinates.

Constructing Voronoi diagrams

There are a variety of **algorithms** (methods) to construct a Voronoi diagram from a list of sites. We will first look at how to add a site to an existing Voronoi diagram, and then we will use this algorithm to build a Voronoi diagram 'from scratch.'

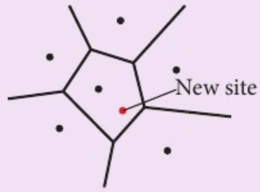
Incremental insertion algorithm

To add a new site p to an existing Voronoi diagram, we use the following process:

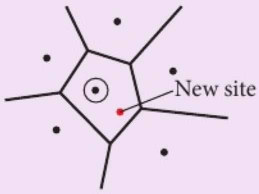
1. Find the nearest existing site. Call this site s. (The nearest site is quickly found because the new site p must be in its cell. In the case that the new site is on an edge, start with either adjacent site.)
2. Construct the perpendicular bisector of the segment joining p and s.
3. On this perpendicular bisector, create a segment with endpoints on the edges of the cell containing s. This segment is the first edge of the new cell containing p.
4. Build the remaining edges of the cell containing p by repeating the following:
   (a) At each endpoint of a new edge, divide the adjacent cell using a new perpendicular bisector between the new site and the site for the adjacent cell.
   (b) On the new perpendicular bisector, create a segment with endpoints on existing edges. (If only one existing edge is intersected, create a ray with its endpoint on that edge. If no existing edges are intersected, the entire perpendicular bisector becomes the new boundary.)
   (c) Continue until there are no unused new-edge endpoints.
5. Discard the edges inside the newly-created cell containing p.

Example 5.5

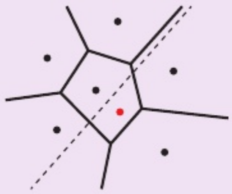
Show the process for adding the site shown in red to the Voronoi diagram shown.



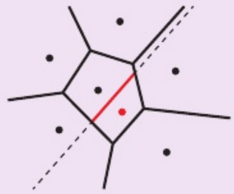
Solution



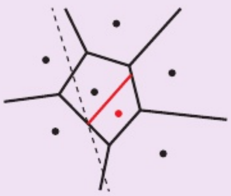
Step 1: Identify the nearest site.



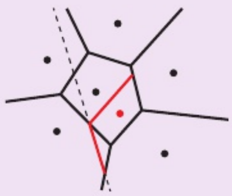
Step 2: Construct a perpendicular bisector.



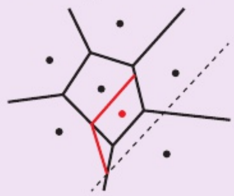
Step 3: Create a line segment with endpoints on existing cell boundaries. This is the first edge of the new cell.



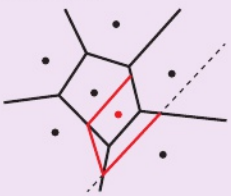
Step 4a: Continue on to the next cell adjacent to the newly-added line segment and create a new perpendicular bisector.



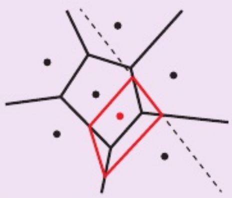
Step 4b: Create the next new cell edge.



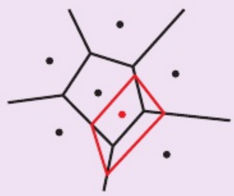
Step 4a: Continue this process for the next adjacent cell.



Step 4b: Continue creating new line segments to form the edges of the new site's cell.

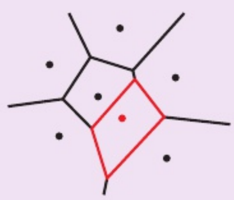


Step 4a/b: Continue until there are no more adjacent cells.



Step 4c: Stop creating new edges when there are no more unused new-edge endpoints.

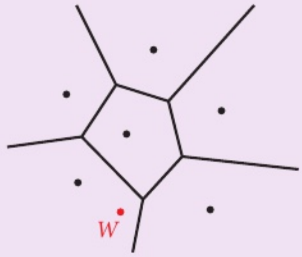
Step 5: Remove any edges inside the new cell.



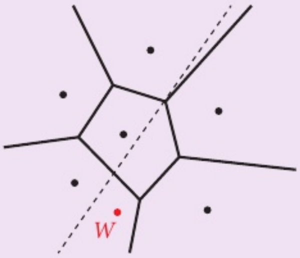
Note that in Step 4b of Example 5.5, it may happen that a new edge is a ray instead of a line segment. This is the case when the new site is on the outer edge of the Voronoi diagram, and then the new cell may be unbounded – two of the edges will be rays instead of line segments.

Example 5.6

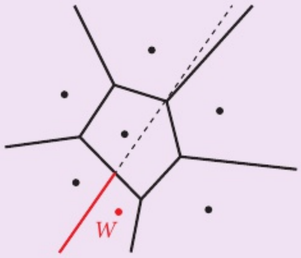
Show the steps to add site W to this Voronoi diagram.



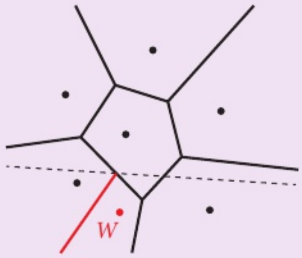
Solution



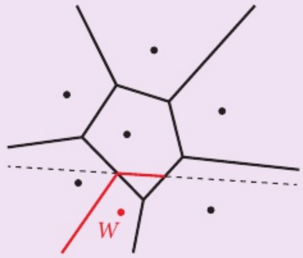
Find the perpendicular bisector between W and the containing cell's site.



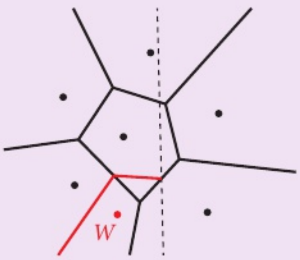
Add ray along perpendicular bisector.



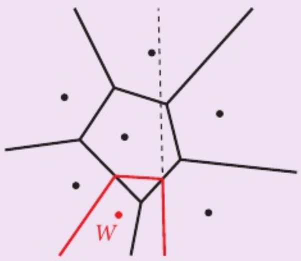
Find perpendicular bisector between W and next adjacent site.



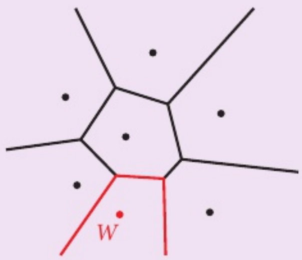
Add line segment along perpendicular bisector.



Find perpendicular bisector between W and next adjacent site.



Add ray along perpendicular bisector.



Remove edges inside W's new cell.

To create a new Voronoi diagram 'from scratch,' we simply apply the incremental insertion algorithm repeatedly.

Creating a new Voronoi diagram

To create a new Voronoi diagram from a list of sites $S_1, S_2, \ldots, S_n$:

1. Construct the perpendicular bisector between sites S_1 and S_2. (This creates the cells for S_1 and S_2.)
2. For each additional site $S_3, S_4, \ldots S_n$ repeat the incremental insertion algorithm until there are no more vertices to insert.

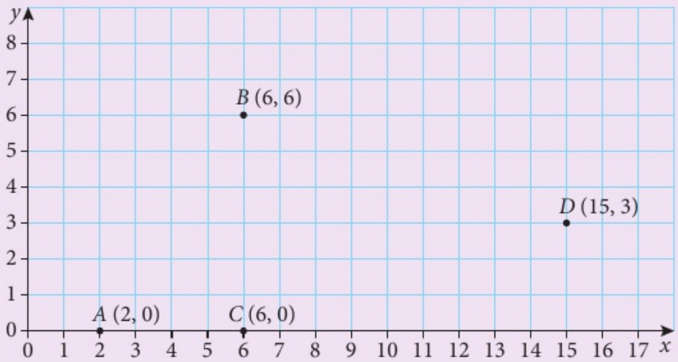
It doesn't matter which two sites you start the Voronoi diagram with, nor does it matter in what order the remaining sites are added. When showing this process, however, it's usually clearer to start with sites near to each other and then work on the next-closest site.

Example 5.7

Create a Voronoi diagram for the points $A(2, 0)$, $B(6, 6)$, $C(6, 0)$, and $D(15, 3)$.

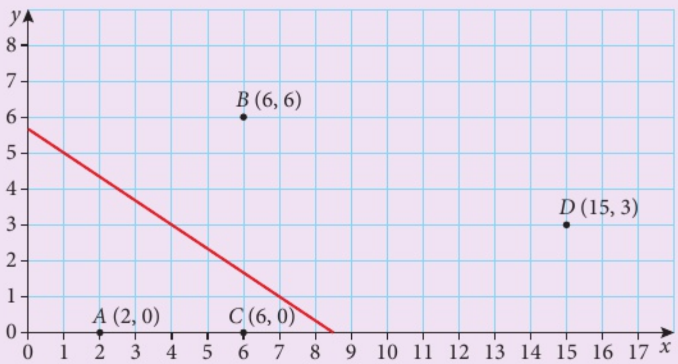
Solution

We begin by plotting the points and choosing two points to begin.



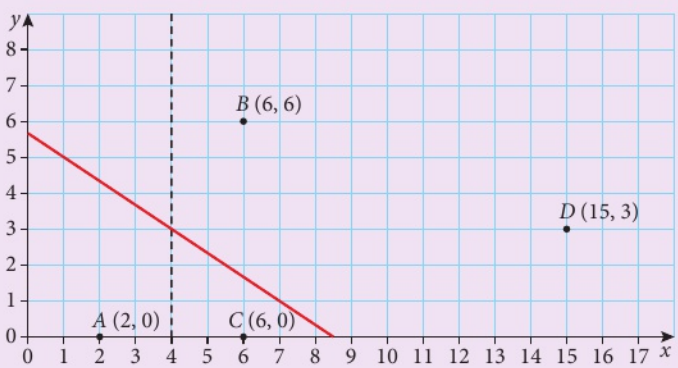
The initial set of four points

Since it doesn't matter which pair of points we begin with, we will start with A and B for simplicity. We first construct the perpendicular bisector of segment AB:

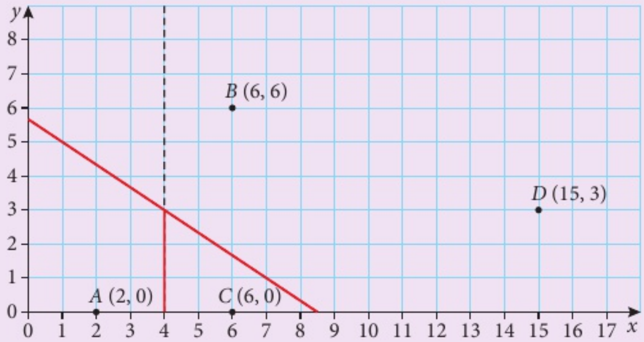


The first perpendicular bisector, shown in red

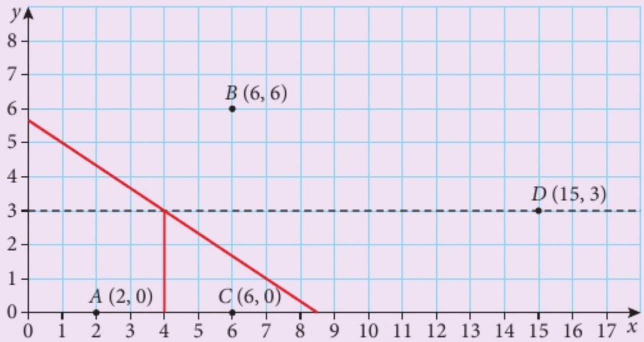
Then, choose the next point to insert. We will use point C. Using the incremental insertion algorithm, we create the perpendicular bisector of segment AC and then work until there are no more adjacent cells to divide:



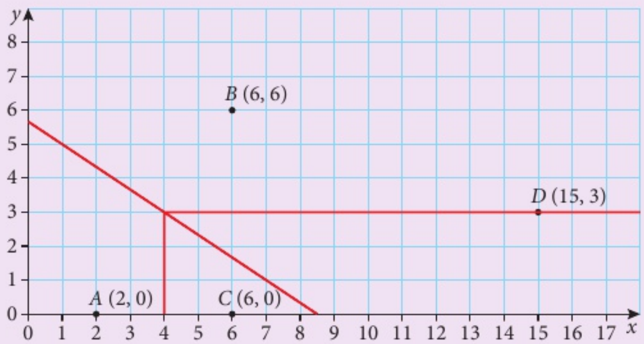
We divide the cell containing A and C



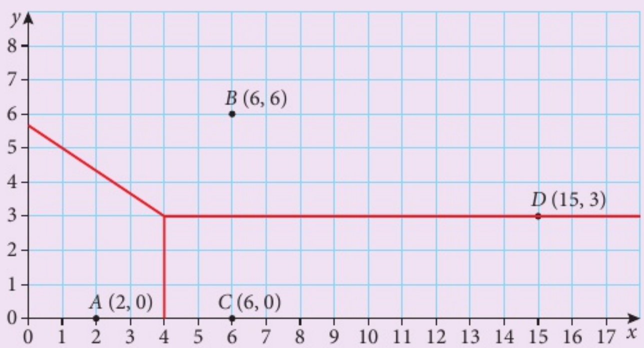
Add a line segment or ray to divide the cell containing A and C



Add another perpendicular bisector to divide the next adjacent cell (the cell of site B in this case).

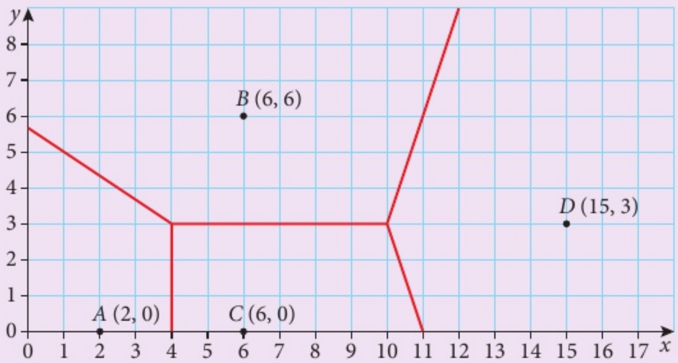


We are almost finished creating the cell for site C, we just need to remove edges inside the new cell.



We have finished inserting site C, now we can insert site D

For brevity, we will not show all steps for inserting site D. The process for inserting site D is the same as for site C. Note that since D is on an existing edge, we can choose to start by dividing the cell for either B or C. Again, we follow the steps for the incremental insertion algorithm to obtain the following finished Voronoi diagram.



After using the insertion algorithm for point D, the Voronoi diagram is complete.

Nearest-neighbour interpolation

Using a Voronoi diagram to assign values to points is a **function** in the general sense, just like the functions you learned about in Chapter 3. In this case, the domain is a set of points and the range is the set of values or labels. Every point is assigned to exactly one value or label based on which site is closest to that point.

A common use of Voronoi diagrams is to assign values to points based on the nearest site. This is called **nearest-neighbour interpolation**.

For example, consider the map in Figure 5.5. In the map, we can see measurements from rain-collection gauges on 27 July 2017 in Pueblo, Colorado, USA. The measurements are given in inches. If we want to estimate the rainfall at a location near one of the rainfall gauges, we can simply use the nearest gauge.

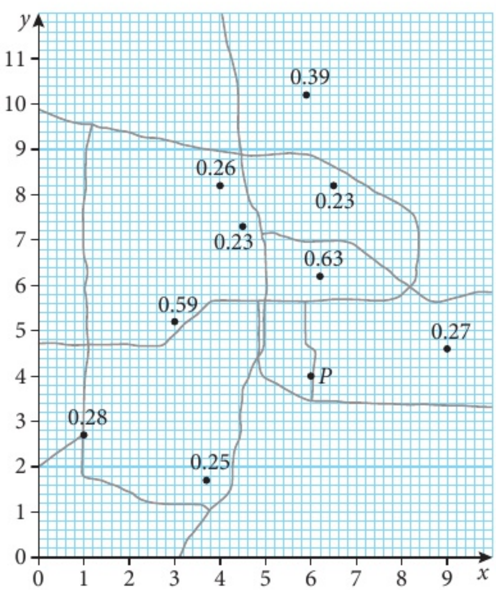


Figure 5.5 A map showing measurements of rain-collection gauges in inches from Pueblo, Colorado, USA. The lines shown are major roads.

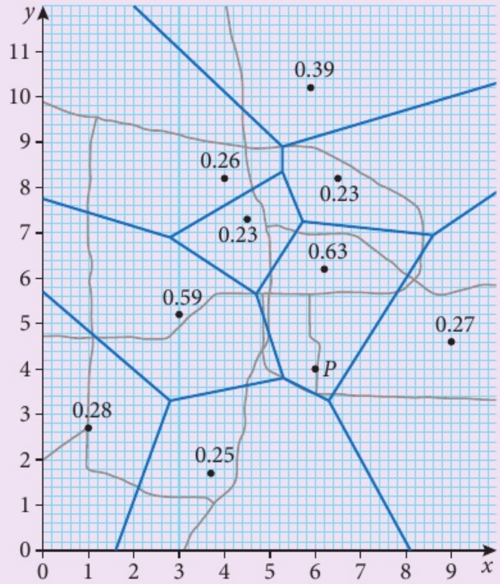
Example 5.8

Use the map of rainfall measurements in Figure 5.5 to estimate the rainfall at the coordinates $P(6, 4)$

Solution

To estimate the rainfall at $P(6, 4)$, we will use the **nearest neighbour** method. To determine the nearest neighbour, we will first build a Voronoi diagram using each point as a site.

Now, to determine the nearest neighbour to point $P(6, 4)$, we simply look at which cell the point is contained within. We can see from the diagram that the nearest site has a reading of 0.63 inches. Therefore, we estimate that the rainfall at $P(6, 4)$ is 0.63 inches.



The map of rainfall gauges with a Voronoi diagram added

You might argue that 0.63 inches is not the best estimate for the rainfall at point P: wouldn't it be better to try to make some sort of average of the rainfall gauges surrounding P? Indeed, there are different methods for interpolating unknown values. A common method that uses an area-weighted average of the nearest sites is called **natural neighbour** interpolation, but it is beyond the scope of this text.

Largest empty circle

One interesting application of Voronoi diagrams concerns finding the point(s) that are farthest from any site. For example, consider a Voronoi diagram where the sites are towns and you are looking for a place to put a toxic waste dump. Of course, no one wants the toxic waste dump in their backyard, so it makes sense to find the location farthest from any of the towns. This is why the **largest empty circle (LEC) problem** is sometimes called the **'toxic waste dump' problem**.

On a more positive note, suppose that a restaurant is looking for a new location within a city. They might be interested to see which place is furthest from competing restaurants, but still within the city. So, they draw a Voronoi diagram where the competing restaurants in the city are sites, as shown in Figure 5.6.

Visually, it appears that the largest area without a restaurant, that is, the point furthest from any existing restaurant, is the area between sites B, C, F, and K, shown in green in Figure 5.7.

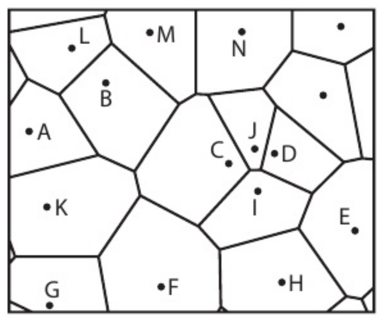


Figure 5.6 A map of existing competitor restaurants in a city

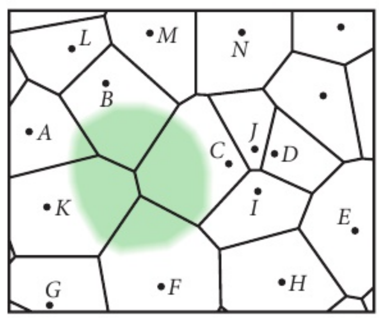


Figure 5.7 General area for a new restaurant location shown in green

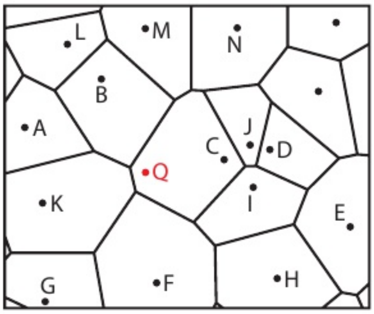


Figure 5.8 Point Q is one possible location. Is point Q the farthest point from any existing restaurant?

But where exactly should the restaurant go? Suppose we decide to put it at point Q, shown in Figure 5.8. But, since Q is located inside the cell of site C, then it is closer to C than any other site. So, we should move it away from site C.

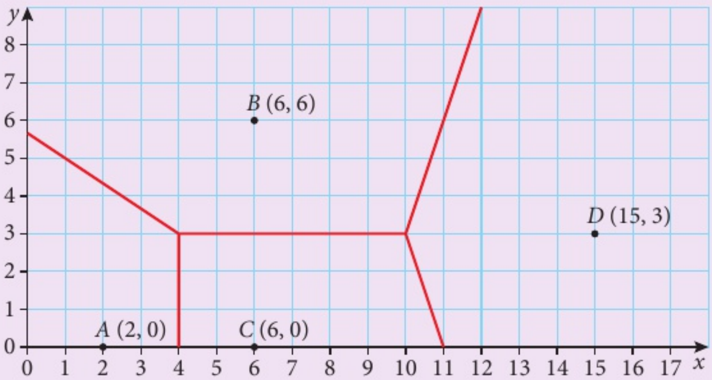
When we move point Q away from site C, we need to be careful not to get closer to a different site. By this reasoning, it should be apparent that Q must be on an edge. Otherwise, Q would be closer to one site than another! By the same reasoning, to maximise distance from all existing sites, Q must be on an intersection of 3 or more boundaries. Thus, Q must be on a vertex.

Actually, it is possible for the centre of the LEC to be just on a boundary, not a vertex, if it is near the outside of the Voronoi diagram. To handle this case, we would need to learn about something called a **convex hull**. This topic is not included in the IB syllabus, so we will only look at cases where we want the LEC to be 'inside' the Voronoi diagram.

The largest empty circle (LEC) that is contained within a Voronoi diagram must be centred on a vertex.

Example 5.9

Find the coordinates of the centre and radius of the largest empty circle (LEC) in the Voronoi diagram.



Solution

Since the centre of the LEC must be on a vertex of the Voronoi diagram, we only have to consider two points: (4, 3) and (10, 3). Although it is clear visually that (10, 3) must be the centre of the LEC, we can check algebraically by using the distance formula.

The vertex (4, 3) is equidistant to sites A, B, and C, so we can choose any of those coordinates to find the radius. Using site $A(2, 0)$, we have
$d = \sqrt{(4-2)^2 + (3-0)^2} = \sqrt{4+9} = \sqrt{13}$

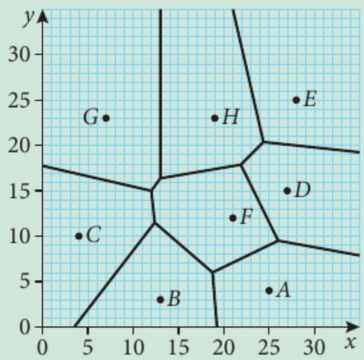
The vertex (10, 3) is equidistant to sites B, C, and D, so we can choose any of those coordinates to find the radius. Using site $C(6, 0)$, we have
$d = \sqrt{(10-6)^2 + (3-0)^2} = \sqrt{16+9} = \sqrt{25}$

Since $\sqrt{25} > \sqrt{13}$, the point (10, 3) must be the centre of the LEC with radius 5.

To find the LEC, we examine all vertices and calculate the distance from each to a site adjacent to that vertex. The LEC is centred at the vertex with the largest distance to an adjacent site.

Exercise 5.2

1. A new site named W will be placed at (15, 15) in the Voronoi diagram shown.



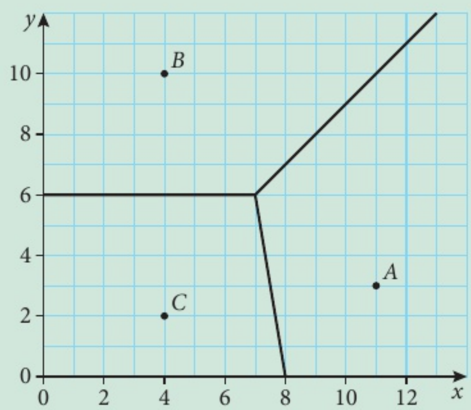
   (a) Write down the letter of the site whose cell will contain W.

   (b) Find the equation of the edge separating W from site G.

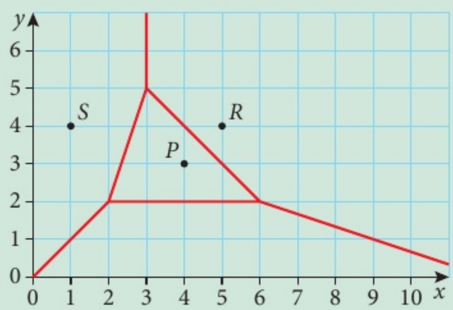
   (c) Sketch the perpendicular bisectors required to adjust the diagram after site W is included. Then write down the letters of the sites whose cells will be changed after the Voronoi diagram is adjusted for site W.

   (d) An additional site X is to be placed at (5, 5). Will every edge of X's cell be a line segment?

2. Copy the Voronoi diagram given on to graph paper. Show the steps to add a new site at (5, 5).



3. The sites S, P, Q (not shown), and R in the Voronoi diagram represent the locations of mobile phone antennas. The grid units are in kilometres. Use the Voronoi diagram to answer the following questions.

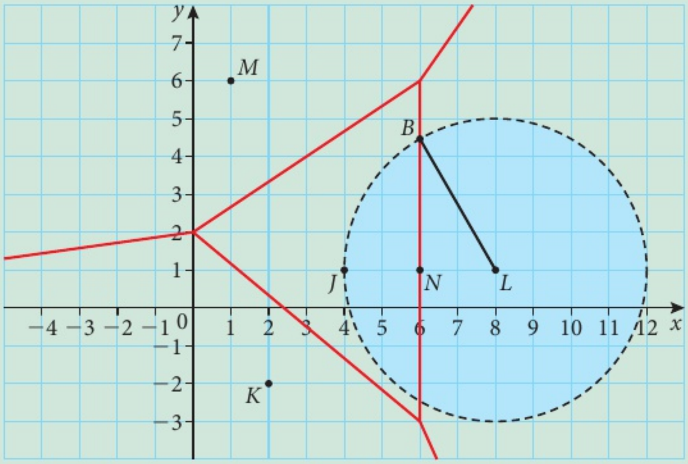


   (a) Site Q is missing from the diagram. Write down the coordinates of site Q.

   (b) Calculate the area covered by the mobile phone antenna located at site P.

   (c) Find the distance from point $X(9, 2)$ to the nearest mobile phone antenna.

4. In less populated areas, internet service providers may use WiMAX antennas to provide service. The Voronoi diagram represents four WiMAX antennas located at sites *J*, *K*, *L*, and *M*. The axis units are in kilometres.



   (a) Calculate the area for which the antenna at site *J* is the nearest antenna.

   (b) Calculate the angle *BLN*.

   (c) The antenna at site *L* has a maximum range of 4 km (shown by the shaded area). Calculate the area for which antenna *L* is the nearest antenna that is also within the range of 4 km.

   (d) Of the three vertices in the Voronoi diagram shown, which is the centre of the largest empty circle? Interpret your findings.

5. A soil scientist takes samples of the soil in a certain area and records the data as shown in Table 5.1. Assume the units are metres.

Site	Location	Soil type
A	(1, 2)	clay
B	(6, 3)	sand
C	(2, 4)	silt
D	(4, 5)	chalk

Table 5.1 Data for question 5

   (a) Beginning with sites *A* and *B*, and continuing alphabetically, draw a Voronoi diagram for sites *A*, *B*, *C*, and *D*, using the incremental insertion algorithm. Show your method by drawing:

   (i) the Voronoi diagram after sites *A* and *B* are added

   (ii) the Voronoi diagram after site *C* is added

   (iii) the finished Voronoi diagram.

   (b) Use nearest neighbour interpolation to determine the likely soil type at point *M*(3.5, 2.5).

   (c) Give a reason why it is not possible to use nearest neighbour interpolation to determine the likely soil type at point *N*(5, 4).

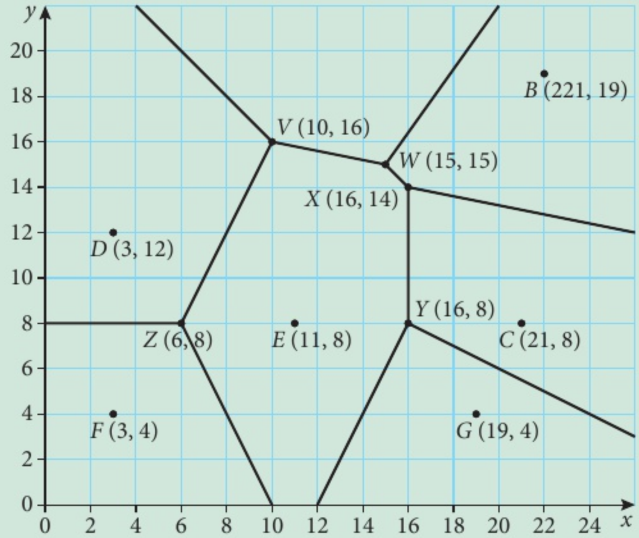
   The soil scientist takes an additional sample at location *E*(6, 0) and determines that the soil type at *E* is loam.

   (d) If site *E* is added to the Voronoi diagram, which cells would be divided? Give a reason for each cell that would be divided.

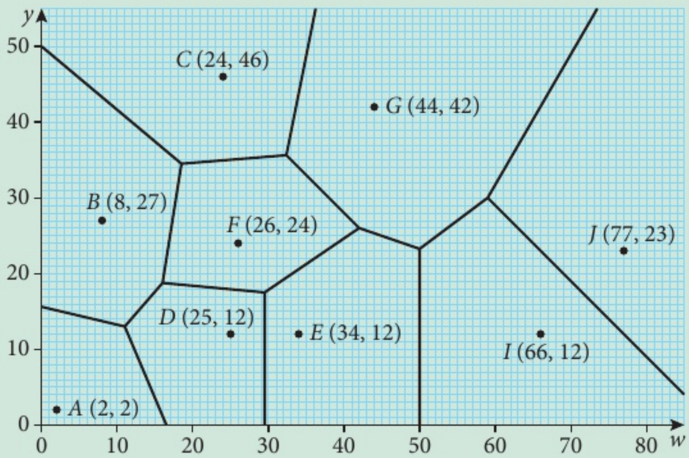
   (e) Given the new sample at site *E*, does your answer to question **(b)** change? Give a reason why or why not.

   (f) The study area is bounded by the area formed by the lines $x = 0$, $y = 0$, $x = 7$, and $y = 6$. Determine the area of the study that is likely to be loam.

6. Using the following Voronoi diagram, find the centre and radius of the largest empty circle.



7. In the following Voronoi diagram, each labelled point is the location of a bank branch in a city.



The bank decides to construct a new branch at $H(66, 28)$.

(a) Add this branch as a site and modify the Voronoi diagram accordingly.

(b) The bank wants to add one more branch in the area bounded by branches E, G, H, and I. Find the optimal location for a new branch based on the largest empty circle.

Chapter 5 practice questions

1. The planet Earth is approximately 150 million km from the Sun, on average. Assume that the orbit of the Earth is approximately circular and that one year is 365.25 days.

   (a) Find the distance travelled by the Earth in one day.

   Kepler's second law of planetary motion states that a line segment joining a planet and the Sun will sweep out equal areas (equal sectors) during equal intervals of time.

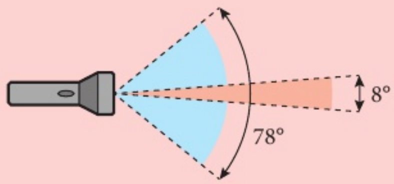
   (b) Calculate the area of the sector swept by Earth in one day.

   The average distance from Venus to the Sun is 108 million km and Venus completes one orbit in about 225 Earth days.

   (c) Assuming a circular orbit, how far does Venus travel in one Earth day?

   (d) What area does Venus sweep in one Earth day?

2. The light emitted by a source can be measured in lumens (lm), and the illumination on an area can be measured in lux (lx) which is equivalent to lumens per square meter (lm m^{-2}). A particular flashlight has an adjustable beam that can emit 1000 lm. The angle of the beam is adjustable from 8° to 78° as shown.



   Illumination is usable if it is greater than 30 lx. Assume that the entire sector is evenly illuminated.

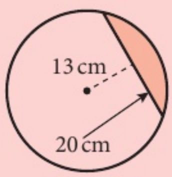
   (a) Find the area of the sector when the beam angle is 8° and the radius is 5 m. Hence, find the illumination in lx that this flashlight can produce for this area.

   (b) When the beam angle is set to 78°, what is the maximum radius, in metres, of the illuminated sector in order for the illumination to be at least 30 lx?

   (c) When the beam angle is set to 8°, what is the maximum radius in m of the illuminated sector in order for the illumination to be at least 30 lx?

   (d) The engineers designing the flashlight would like at least 50 lx with a beam angle of 8° and a distance of 22 m. How many lumens are required for this goal?

3. A 20 cm cut is made across a circular pizza. The shortest distance from the cut to the centre of the pizza is 13 cm as shown.



   (a) Find the diameter of the pizza.

   (b) Find the area of the shaded cut off piece.

4. A cone is made from a sector of a circle as shown in Figure 5.9. The radius of the circle is 15 cm. The vertical height h of the cone is equal to the radius r of its base.
   (a) Find the value of r, the radius of the base of the cone.
   (b) Hence find the central angle θ of the sector.

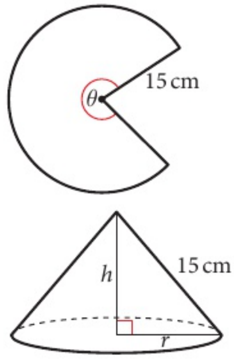
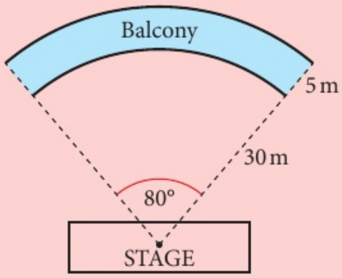


Figure 5.9 Diagram for question 4

5. The Earth is approximately spherical with a radius of 6371 km. Find the shortest distance between:
   (a) the North Pole and Novosibirsk, Russia, located at 55°N
   (b) the North Pole and Lusaka, Zambia, located at 15.5°S of the equator
   (c) the equatorial cities Macapá, Brazil, located at 51.1°W, and Entebbe, Uganda, located at 32.5°E
   (d) the equatorial cities Quito, Equator, located at 78.6°W, and Pontianak, Indonesia, located at about 109°E.

6. A concert hall has a balcony that follows the arc of a circle centred on the stage. The front edge of the balcony is 30 m from the stage centre and the balcony is 5 m deep. The measure of the angle from the stage centre is 80°.



   (a) A new guard rail will be installed at the front edge of the balcony. Calculate the length of the guard rail.
   (b) The floor of the balcony must be replaced. Calculate the floor area of the balcony in m^2.
   (c) The two side walls and the back wall of the balcony will be covered with sound-absorbing panels to reduce echoes. Each panel is 0.5 m wide and 2 m tall. Find the maximum number of panels that could fit along these three walls. (Assume there are no doorways or other obstructions.)

7. For this exercise, use a grid with $0 \leqslant x \leqslant 18$ and $0 \leqslant y \leqslant 12$
   (a) Create a Voronoi diagram with sites at $A(7,0)$, $B(3,4)$, $C(11,8)$, and $D(15,4)$.
   (b) Add the site $X(7,4)$ to the diagram.
   (c) Find the equation of the edge between sites X and A.
   (d) Find the equation of the edge between sites X and C.
   (e) Add the site $Y(7,8)$ to the diagram.
   (f) Find the centre of the largest empty circle with coordinates (a, b) such that $3 \leqslant a \leqslant 15$ and $0 \leqslant b \leqslant 8$
   (g) Add site Z at (a, b).
   (h) Find the equation of the edge between sites Z and D. Write your answer in the form $ax + by = c$, where $a, b, c \in \mathbb{Z}$

8. For this exercise, use a grid with $-10 \leqslant x \leqslant 10$ and $-10 \leqslant y \leqslant 4$

(a) Create a Voronoi diagram with sites at $A(0, 1)$, $B(-7, -2)$, $C(3, -2)$, and $D(6, -5)$.

(b) Add the site $X(0, -2)$ to the diagram.

(c) Find the equation of the edge between sites X and C.

(d) Add the site $Y(3, -6)$ to the diagram.

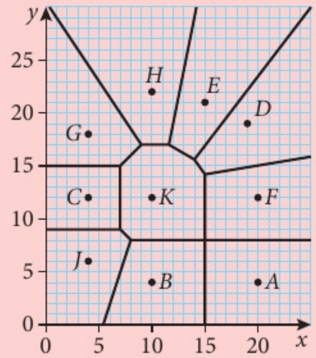
(e) Find the equation of the edge between sites X and Y. Write your answer in the form $ax + by = c$, where $a, b, c \in \mathbb{Z}$

9. In the Voronoi diagram shown each site is the location of a petrol station in a city.

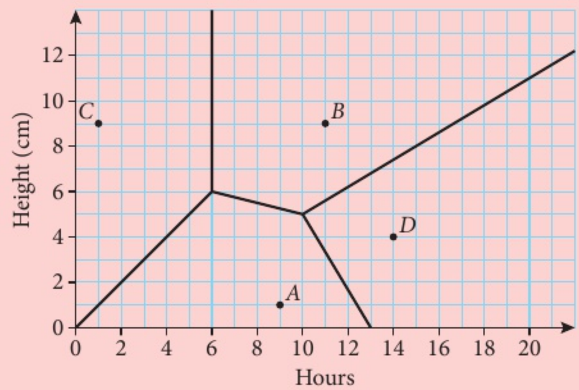
Given that a new petrol station is to be located as far as possible from existing petrol stations:

(a) give a reason why the new location should be on a vertex of the Voronoi diagram

(b) give a reason why (15, 8) is the best location for the new station.



10. Use the Voronoi diagram shown to answer the questions.



(a) Find the centre and radius of the largest empty circle. Give two justifications for your answer.

(b) Insert a new site X at the centre of the largest empty circle.

11. Sites A, B, C, and D are collinear. Explain why the Voronoi diagram for A, B, C, and D has no vertices.

12. Voronoi diagrams of regular polygons:

(a) Sites A, B, and C are located at the vertices of an equilateral triangle. Sketch the Voronoi diagram for these sites.

(b) Sites A, B, C, and D are located at the vertices of an square. Sketch the Voronoi diagram for these sites.

(c) Give a description for the Voronoi diagram for a regular n-gon including the position of the edges and the number of vertices.

6 Modelling real-life phenomena

6 Modelling real-life phenomena

Learning objectives

By the end of this chapter, you should be familiar with…

- modelling linear, linear piecewise, quadratic, cubic, exponential, direct/inverse variation, and trigonometric phenomena
- developing and fitting models (recognising the context, choosing an appropriate model, determining a reasonable domain and range, using technology to find parameters)
- testing and reflecting upon models (commenting on appropriateness and reasonableness of a model, justifying the choice of a model)
- using models (reading, interpreting, and making predictions, avoiding extrapolation).

Mathematical models help us to describe the world around us. In this chapter, we will look at several different kinds of mathematical models. We will examine how to choose, develop, test, apply, and extend a model. Here are some examples of the different kinds of models we see in the world around us:

- On a long flight, the airspeed of a plane is constant, so the distance remaining to the destination can be described by a **linear** model.
- In a situation where the price to manufacture x units of some product decreases **linearly**, the revenue from selling x units can be described by a **quadratic** model.
- The volume of a balloon relative to its diameter can be described by a **cubic** model.
- The spread of algae in a polluted lake can be described by an **exponential** model.
- A DJ charges a fixed amount to provide music for a party. The cost is spread equally among everyone who attends the party. The cost per person can be described by an **inverse variation** model.
- The price of electricity is often billed per kilowatt-hour (kWh), so the cost of powering the lights of a stadium relative to the time the lights are on can be described by a **direct variation** model.
- The height of a person above the ground on a Ferris wheel can be described by a **trigonometric** model.

The process of mathematical modelling is a design process and is illustrated in Figure 6.1.

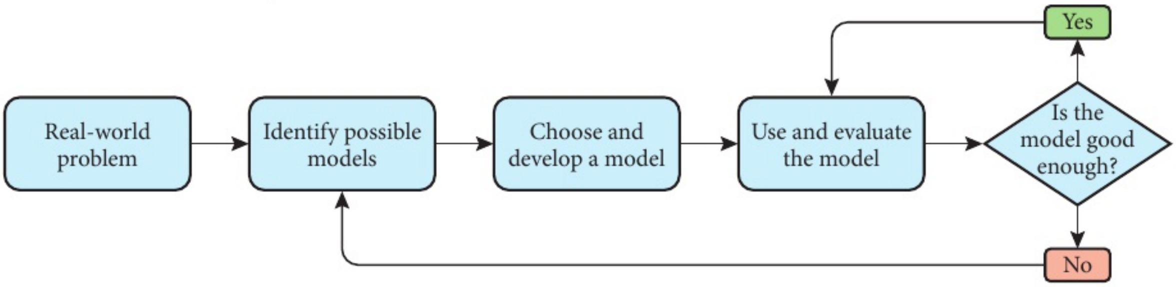


Figure 6.1 Mathematical modelling process

In this chapter we will learn how to develop models both by hand and by using technology. We will discuss how to test a model and then how to reflect upon or analyse the validity of the model. Finally, we will talk about how to use and, if needed, extend or revise a model.

It is important to remember that all models are simplifications of reality. We use models to tell us something about the way a system behaves and to make predictions. The goal of a model is to simplify and approximate a real system so that we can learn, predict, and analyse the behaviour. Part of our job is to use models wisely and be sure we understand the limitations and assumptions of any model we use.

6.1 Linear models

Linear models are used to describe situations where one quantity (the **dependent variable**) increases at a fixed rate relative to another quantity (the **independent variable**).

Developing and testing a linear model

Suppose that you need a plumbing repair in your home. You call a plumber to ask about how much it will cost. Of course, the plumber cannot give you an exact cost but does give you the estimates shown in Table 6.1.

Time required (hours)	Cost of repair (€)
1	110
2	185
3	260

Table 6.1 Repair cost estimates

Since we know that the cost must depend on the time required, **cost** is the dependent variable and **time required** is the independent variable.

To decide if this situation can be described by a linear model, we need to see if the rate of change is constant. To do this, check the gradient (in this context, the cost per hour) between two or more pairs of points:

$$\frac{185-110}{2-1}=75 \quad \text{and} \quad \frac{260-185}{3-2}=75$$

Since the cost per hour is constant, a linear model is appropriate for this situation. In addition, we have discovered that the cost per hour is €75. However, there seems to be another part to the cost. We can use slope–intercept form to find a linear model. We will use y for the dependent variable, cost, and x for the independent variable, time:

$$y-y_1=m(x-x_1) \;\Rightarrow\; y-110=75(x-1) \;\Rightarrow\; y=75x+35$$

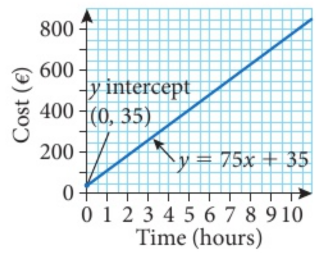


Figure 6.2 The graph of a linear model of the cost of hiring a particular plumber to come to your home

We can see that the plumber's hourly rate is €75, and he adds a fixed amount of €35. This is probably to compensate him for travelling to your home! Finally, it's a good idea to test the model to make sure it describes the situation:

For 1 hour: $y=75(1)+35=110$

For 2 hours: $y=75(2)+35=185$

This matches the estimates given so we can be confident that our model is appropriate. Now we can see what we might have to pay if it takes the plumber a whole 8-hour day to fix our problem:

For 8 hours: $y=75(8)+35=€635$

Notice also that graphs of linear functions are always lines.

Remember that in a graph the independent variable is usually placed on the horizontal axis, and the dependent variable is placed on the vertical axis.

Linear models describe situations where the rate of change (gradient) is constant. The graph of a linear function is a line.

Extending and revising models

Figure 6.3 New York City tax rates decal from the late 1970s. Linear models are very well suited to this sort of situation

Sometimes we can make an initial simple model but need to revise it to be more useful. Here is an example.

Consider the taxi costs shown in Figure 6.3. Because the cost increases at a fixed rate relative to the distance driven (10 cents for each $\frac{1}{7}$ miles driven), this is a good candidate for a linear model. Since the cost depends on the number of miles driven, we will make the cost the dependent variable and the distance driven the independent variable. It often helps to make a table showing the independent and dependent variables:

Miles driven	Calculation	Cost ($)
$\frac{1}{7}$	$0.75 + 0$	0.75
$\frac{2}{7}$	$0.75 + 0.10(1)$	0.85
$\frac{3}{7}$	$0.75 + 0.10(2)$	0.95
$\frac{4}{7}$	$0.75 + 0.10(3)$	1.05

Table 6.2 Independent and dependent variables

By explicitly showing our calculations, we get a good idea of how to develop our model. From the table, it appears that an appropriate model is Cost $= 0.75 + 0.10x$. But be careful! In this case, what is x? Notice that the number we are multiplying by 0.10 is not the number of miles driven – it is the number of $\frac{1}{7}$ miles after the first $\frac{1}{7}$ mile!

Since most people don't think in terms of $\frac{1}{7}$ miles, it would be useful if our independent variable was simply distance in miles. Let's try to revise our table:

Miles driven	The number of $\frac{1}{7}$ miles after the first $\frac{1}{7}$ mile	Calculation	Cost ($)
$\frac{1}{7}$	$\frac{1}{7} - \frac{1}{7} = 0$	$0.75 + 0.10(0)$	0.75
$\frac{2}{7}$	$7\left(\frac{2}{7} - \frac{1}{7}\right) = 1$	$0.75 + 0.10(1)$	0.85
$\frac{3}{7}$	$7\left(\frac{3}{7} - \frac{1}{7}\right) = 2$	$0.75 + 0.10(2)$	0.95
$\frac{4}{7}$	$7\left(\frac{4}{7} - \frac{1}{7}\right) = 3$	$0.75 + 0.10(3)$	1.05
m	$7\left(m - \frac{1}{7}\right)$	$0.75 + 0.10x$	C

Table 6.3 Revised table

To change 'miles driven' into 'the number of $\frac{1}{7}$ miles after the first $\frac{1}{7}$ mile,' we need to subtract $\frac{1}{7}$ (representing the first $\frac{1}{7}$ mile) and then multiply by 7 (so that each $\frac{1}{7}$ mile is counted as one unit).

So now we know that $C = 0.75 + 0.10x$ **and** $x = 7\left(m - \frac{1}{7}\right)$ where m is the number of miles.

By substituting the equation for x into the equation for C, we get

$$C = 0.75 + 0.10\left(7\left(m - \frac{1}{7}\right)\right) \Rightarrow C = 0.7m + 0.65$$

A graph of this function is shown in Figure 6.4.

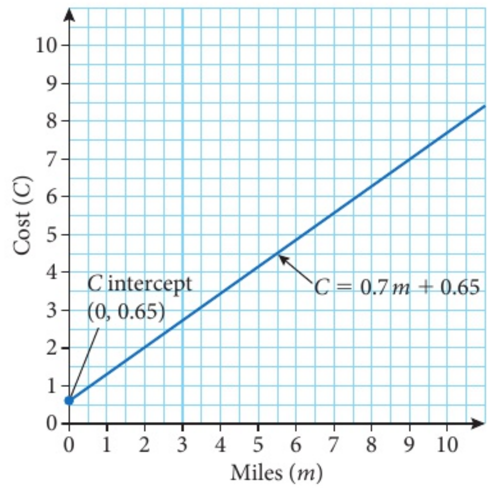


Figure 6.4 A graph of the linear model for the cost of hiring a New York taxi in the 1970s

Models don't always capture reality perfectly

If we look at the graph of our model from the taxi cost example, we see that the model suggests that for a journey of 0 miles, we will pay \$0.65. However, we know that we will always pay at least \$0.75. What went wrong?

The problem here is that the model assumes that the incremental cost (represented by the gradient of the function) is continuous – that is, that the taxi will charge us for any increment of a mile. However, we know that the taxi will charge for each $\frac{1}{7}$ of a mile. To make this point clear, consider what happens if we drive 0.5 miles. The model suggests that the cost would be

$$C = 0.7(0.5) + 0.65 = \$1.00$$

But we know that 0.5 miles is more than $\frac{3}{7}$ of a mile and less than $\frac{4}{7}$ so we would actually get charged

$$C = 0.7\left(\frac{4}{7}\right) + 0.65 = \$1.05$$

That is, our model works as long as we round the miles up to the nearest $\frac{1}{7}$ of a mile. As stated in the introduction, all models are a simplification of reality. This model can help us see how the cost of the ride relates to the length of the ride, but we need to be careful in using it to predict exact costs.

Example 6.1

The I&T Fitness Centre charges a one-time joining fee of \$100 and then charges \$2 per visit.

(a) Develop a model for the total cost after v visits.

(b) Use your model to find the cost of 20 visits.

(c) Draw a graph of your model.

Solution

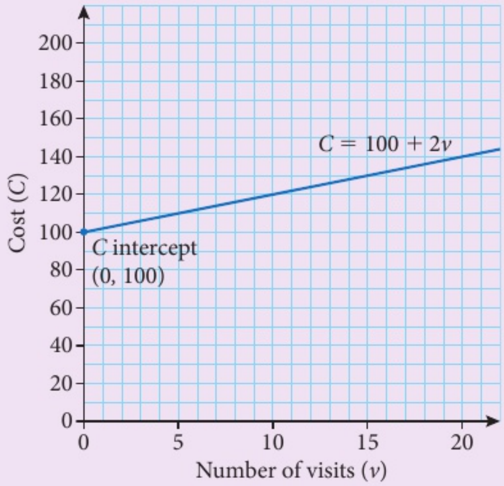
(a) Since the rate is already given ($2 per visit), plus a one-time cost of $100, we can write the model directly:

$$C = 100 + 2v$$

where C is the total cost and v is the number of visits.

(b) Using our model, 20 visits would cost $C = 100 + 2(20) = \$140$

(c)



We have drawn the graph with a continuous line, even though it is only possible to visit a whole number of times – e.g., we can't visit 1.5 times. For convenience, we often draw the graph as a continuous line even though that may not strictly represent reality.

Example 6.2

A visitor to the I&T Fitness Centre decides not to buy a membership and instead just pays the daily rate. The first visit costs her $12. At the end of the month, she notices she has visited 5 more times and paid a total of $60 for those 5 visits.

(a) Develop a linear model for the total cost after v visits.

(b) Use your model to predict the cost of 20 visits.

(c) After how many visits is it better to buy the membership described in Example 6.1?

Solution

(a) One visit costs $12, and 5 visits cost $60. To be sure there are no extra fees, check that the cost per visit for 5 visits is the same as the cost for 1 visit. For 5 visits, the cost per visit is $\frac{60}{5} = \$12$

Therefore, the rate of change is constant, and we can develop the model:

$C = 12v$ where C is the total cost and v is the number of visits.

(b) Using our model, 20 visits would cost $C = 12(20) = \$240$

This is much more expensive than the membership plan in Example 6.1!

(c) To find out at which point the membership plan in Example 6.1 becomes the better option, we first find at which point the two plans are equal. We can do this algebraically or graphically.

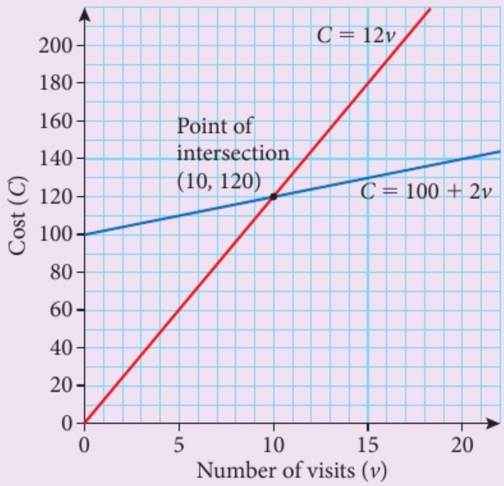
Algebraically, we are looking for the number of visits (v) that produces the same cost. Therefore, we can write $C = 100 + 2v = 12v$ and solve for v:

$$100 + 2v = 12v \Rightarrow v = 10$$

Therefore, the two plans are the same for 10 visits. Since we know that the membership plan costs only \$2 per visit, we know that it will be cheaper for any number of visits more than 10.

We can graph both models and look for the intersection.

The intersection point of the two graphs tells us that the two plans will cost the same (\$120) at 10 visits. We can see clearly that the membership plan increases much more slowly: \$2 per visit instead of \$12 per visit, so it is cheaper after 10 visits.



Interpreting and evaluating linear models

It is important to be able to recognise the structure of a linear model and interpret its meaning. Also, we must be careful to recognise limitations of linear models.

Example 6.3

The number of items that a clothing store sells can be modelled by the function $N = 1000 - 5p$, where N is the number of items sold and p is the price of the jeans in euros.

(a) Use the model to predict the number of jeans sold when the jeans are priced at €100.

(b) Interpret the value of the gradient and N-intercept in context.

(c) Interpret the value of the p-intercept in context.

(d) Use the model to predict the number of jeans sold when the jeans are priced at €500. Give a reason why this prediction is not reasonable.

Solution

(a) The model predicts that the number of jeans sold when the jeans are priced at 100 euros is: $N = 1000 - 5(100) = 500$ jeans.

(b) The gradient of -5 represents that for each euro that the price increases, the number of jeans sold decreases by 5 jeans. The N-intercept of (0, 1000) theoretically represents that the number of jeans sold when the jeans are free will be 1000 jeans – probably not realistic!

(c) The p-intercept occurs when $N = 0$. Therefore, we must solve $0 = 1000 - 5p$ to obtain $1000 = 5p \Rightarrow p = 200$

This tells us that when the price of jeans is 200 euros, the number of jeans sold will be 0.

(d) The model predicts that the number of jeans sold when the jeans are priced at €500 will be: $N = 1000 - 5(500) = -1500$ jeans

This is not reasonable because this suggests that customers will be giving back their jeans!

Example 6.3 shows that linear models often become nonsensical for certain extreme values of the independent variable. For this reason, it's best to give a limitation on the domain of the model to avoid these silly results. In the model above, it is sensible to limit the domain of the model to $1 \leqslant p \leqslant 200$

The upper bound of $p = 200$ is when the model predicts that N will be zero.

Piecewise linear models

Sometimes a real-life situation is not modelled by a single linear function, but is linear in parts. Consider the following example:

Example 6.4

A phone company charges a rate of \$0.24 for the first minute of a call, then \$0.12 per minute for the next 9 minutes, then \$0.06 per minute thereafter. Calls are charged on a per-second basis.

(a) Develop a piecewise linear model for the cost C of a call lasting t seconds.

(b) Use your model to calculate the cost of calls lasting:
(i) 45 seconds (ii) 4 minutes (iii) 15 minutes.

Solution

(a) For $0 < t \leqslant 60$, the cost is \$0.24 per minute, or $\frac{0.24}{60} = \$0.004$ per second. Then, for $60 < t \leqslant 600$, the cost is \$0.12 per minute or $\frac{0.12}{60} = \$0.002$ per second. However, we must also add in the cost for the first minute, and not charge twice for the first minute, so in total the cost will be $0.24 + 0.002(t - 60)$ for a call of t seconds. Likewise, for calls more than 10 minutes, we have the cost of the first minute, plus the cost of the next 9 minutes, plus the remaining cost of $\frac{0.06}{60} = 0.001$ per minute, which gives us $0.24 + 0.12(9) + 0.001(t - 600)$, which simplifies to $1.32 + 0.001(t - 600)$ for a call of t seconds.

We express this mathematically with the following notation:

$$C = \begin{cases} 0.004t, & 0 < t \leqslant 60 \\ 0.24 + 0.002(t - 60), & 60 < t \leqslant 600 \\ 1.32 + 0.001(t - 600), & t > 600 \end{cases}$$

In this notation, we define the value of C through a **piecewise function**. To evaluate the function, we select the piece of the function that applies to the value of t we want, as in part (b).

(b) For each value of t, we simply select the appropriate piece of the function.

(i) Since $t = 45$ is between 0 and 60 seconds, we use $C = 0.004t$ to give us $C = 0.004(45) = \$0.18$

(ii) Since 4 minutes is 240 seconds, and $t = 240$ is between 60 and 600 seconds, we use $C = 0.24 + 0.002(t - 60)$ to give us $C = 0.24 + 0.002(240 - 60) = \$0.60$

(iii) Since 15 minutes is $15 \times 60 = 900$ seconds, and $t = 900$ is more than 600 seconds, we use $C = 1.32 + 0.001(t - 600)$ to give us $C = 1.32 + 0.001(900 - 600) = \$1.62$

Many calculators can graph piecewise functions, as shown in Figure 6.5.

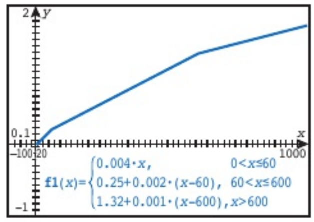


Figure 6.5 GDC piecewise function

Exercise 6.1

1. A plane is currently 3000 km from its destination, travelling at a constant speed of 900 km h^{-1}.

(a) Develop a linear model for the distance d remaining after t hours of travel.

(b) Interpret the d-intercept of your model in context.

(c) Interpret the t-intercept of your model in context.

(d) State a reasonable domain and range for your model.

2. A plane is currently 5000 km from its destination. 1.5 hours later, it is 3800 km from its destination.

(a) Develop a linear model for the distance d remaining after t hours of travel.

(b) Interpret the gradient of your model in context.

(c) Interpret the d-intercept of your model in context.

(d) Interpret the t-intercept of your model in context.

(e) State a reasonable domain and range for your model.

3. The table shows a comparison between EU and USA shoe sizes.

USA (Men's)	7	8	9
EU	40	41	42

(a) Develop a linear model to find the EU shoe size given the USA shoe size.

(b) Use your model to predict the EU shoe size for a USA Men's shoe size of 12.

(c) Use your model to predict the USA Men's shoe size for an EU shoe size of 44.

(d) Interpret the gradient of your model in context.

(e) Given that USA Men's shoe sizes typically run from 6 to 16, calculate a reasonable domain and range for your model.

4. Julie has collected data on how long it takes her to read books, based on the number of pages. The data she collected are shown in the table.

Number of pages	340	290	500
Time to read (minutes)	490	420	714

(a) Develop a linear model for the time required to read n pages.

(b) Use your model to predict the time required to read 1000 pages. Give your answer to the nearest 10 minutes.

(c) Interpret the gradient and y-intercept of your model in context.

(d) State a reasonable domain and range for your model.

5. Given that 68°F = 20°C and 212°F = 100°C:

(a) Develop a linear model for degrees Fahrenheit (F) in terms of degrees Celsius (C).

(b) Explain, in context, what the gradient of your model represents.

(c) Interpret the F-intercept of your model in context.

(d) Interpret the C-intercept of your model in context.

(e) Use your model to convert 10°C into °F.

(f) Use your model to find the numerical value in °C that is the same in °F.

(g) Given that absolute zero (the lowest possible temperature) is −273°C, calculate a reasonable domain and range for your model.

6. The DJ, IB Cool, charges a flat fee of \$150 per party plus \$75 per hour. The DJ, MC Numbers, charges \$120 per party plus \$80 per hour.

(a) Find linear models for each DJ as a function of the length of the party in hours.

(b) For parties longer than n hours, IB Cool is less expensive. Find the value of n.

7. A cyclist pedals at the rate of $300\,\text{m min}^{-1}$ for 20 minutes, then slows down to $150\,\text{m min}^{-1}$ for 16 minutes, then races at $400\,\text{m min}^{-1}$ for 4 minutes.

(a) Find the distance travelled after:

(i) 20 minutes (ii) 36 minutes (iii) 40 minutes.

(b) Write a piecewise linear function for the distance $D(t)$ in terms of the time t in minutes.

(c) Find the distance travelled after:

(i) 30 minutes (ii) 38 minutes.

(d) When has the cyclist travelled:

(i) 8 km (ii) 9 km?

8. The Athabasca Glacier in Alberta, Canada, has been slowly shrinking for many years. Photos from 1844 show the glacier about 2 km longer than it was in 2018.
   (a) Calculate the average rate at which the glacier is shrinking.
   (b) In 1844, the glacier was 8 km long. Assuming the present rate continues, in what year will the glacier be gone?

   In the year 1900, the end of the glacier was about 300 km from the US–Canada border; as it shrinks it moves farther from the border.
   (c) Write a function for the distance d in km between the end of the glacier and the US–Canada border, in terms of time t, in years, since 1900.
   (d) Assuming that the rate of change has been constant for many years, at what time did the glacier reach the US–Canada border?

9. Two plastic cup factories, Cups R Us and Cupomatic, can produce cups printed with the image of your choice. At Cups R Us, the mandatory setup and design cost is ZAR350 and the cost per cup is ZAR8.50.
   (a) Develop a linear model for the cost, C, of an order at Cups R Us based on the number of cups, n.
   (b) Write down a reasonable domain and range for your model.
   (c) Use your model to calculate the cost of an order of:
   (i) 100 cups (ii) 200 cups (iii) 400 cups.
   (d) Calculate the average cost per cup for:
   (i) 100 cups (ii) 200 cups (iii) 400 cups.
   (e) Hence, give a reason why, in general, it is more cost-effective to order more cups.

   Cupomatic charges ZAR2150 for 200 cups and ZAR3750 for 400 cups.
   (f) Develop a linear model for the cost, D, of an order at Cupomatic based on the number of cups, n.
   (g) Write down a reasonable domain and range for your model.
   (h) Interpret the gradient of your model in context.
   (i) Use your linear model to predict the cost of 600 cups.
   (j) For orders of more than x cups, it is more cost-effective to order from Cupomatic. Find the value of x.

10. Continuing from the previous question, Cups R Us will waive the setup and design cost if the order is at least 500 cups.
   (a) Develop a piecewise model for the cost, C, of an order at Cups R Us based on the number of cups, n.
   (b) The main competitor of Cups R Us is Cupomatic. You are given that the model for the cost, D, of ordering n cups from Cupomatic is $D = 8n + 550$. It is less expensive to order from Cupomatic if x, the number of cups ordered, is in the intervals $a \leqslant x < b$ or $x > k$. Find the values of a, b, and k.

11. As of 2018, taxi cab tariffs for working hours in London, England, are as follows.

- For the first 234.8 metres or 50.4 seconds (whichever is reached first) there is a minimum charge of 2.60 GBP.
- For each additional 117.4 metres or 25.2 seconds (whichever is reached first), or part thereof, if the distance travelled is less than 9656.1 metres there is a charge of 0.20 GBP.
- Once the distance has reached 9656.1 metres then there is a charge of 0.20 GBP for each additional 86.9 metres or 18.7 seconds (whichever is reached first), or part thereof.

(a) Develop a piecewise linear model for the cost, C, of a taxi ride based on the distance travelled, m, in metres.

(b) Find the cost of 0.2 km, 5 km, and 15 km rides.

(c) Develop a piecewise linear model for the cost, D, of a taxi ride based on the time taken, t, in seconds, ignoring distance.

(d) Find the cost of 0.5 minute, 5 minute, and 15 minute rides.

(e) Given that the actual taxi fare is always the greater of the two models, find:

(i) the cost of a ride that takes 10 minutes to go 4 km

(ii) the cost of a ride that takes 5 minutes to go 4 km.

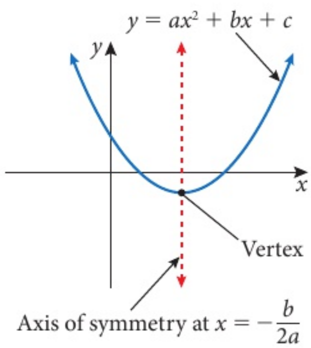


Figure 6.6 If $a > 0$ then the parabola is concave up

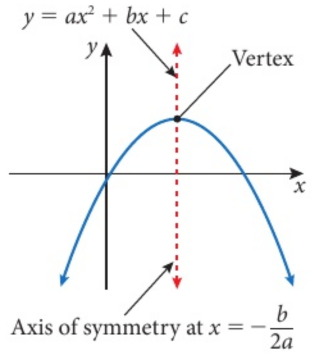


Figure 6.7 If $a < 0$ then the parabola is concave down

6.2 Quadratic models

Quadratic models appear frequently in real-world situations involving area, economics, projectile motion, and falling objects, among others. In this section we will look primarily at quadratic models of the form $y = ax^2 + bx + c$

Quadratic models are used in situations where the rate of change in the dependent variable changes linearly with respect to the independent variable. For example:

- Falling objects due to gravity, projectile motion: the acceleration is constant, the velocity follows a linear model, and the displacement (position) follows a quadratic model.
- Revenue models: the number of items sold of some item based on the price follows a linear model; the revenue from selling the number of items follows a quadratic model.

Quadratic models also have some geometric properties that make them well-suited to designing satellite dishes and to modelling bridge spans.

Before we begin, we will restate key properties of quadratic functions that you have seen before.

For a quadratic function of the form $y = ax^2 + bx + c$, where $a \neq 0$

- The graph of a quadratic function is roughly U-shaped and is called a **parabola**.
- Concavity: The graph of the function is **concave up** if and only if $a > 0$
  It is **concave down** if and only if $a < 0$
- Symmetry: The graph of the function is symmetrical about the vertical line with equation $x = -\frac{b}{2a}$
  This line is called the **axis of symmetry**.
- Maximum/minimum: The function has a **maximum** (when concave down) or **minimum** (when concave up) where the graph intersects the axis of symmetry. This point is called the **vertex** of the parabola. The x coordinate of the vertex is therefore given by $x = -\frac{b}{2a}$
- The ***y*-intercept** of the graph is found at $(0, c)$
- The ***x*-intercepts** of the graph, also called the **zeros** of the function, can be found with the quadratic formula: $x = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}$
  Remember that a graph may have zero, one, or two x-intercepts.

Example 6.5

A ball is thrown upwards from the top of a building. The height of the ball from ground level can be modelled by the function $h(t) = -4.9t^2 + 11t + 50$ where $h(t)$ is the height of the ball in metres and t is the time in seconds after the ball was thrown.

(a) Sketch a graph of the function.

(b) Write down the height of the building.

(c) Find the time when the ball reaches its maximum height.

(d) Find the maximum height reached by the ball.

(e) Find the time when the ball hits the ground.

(f) Describe a reasonable domain and range for this model.

Solution

We can solve this problem either by using an algebraic approach ('by hand') or by using our GDC to analyse the graph.

Algebraic approach

(a) We know the ball travels upwards before coming back down and we know the parabola is concave down since $a < 0$

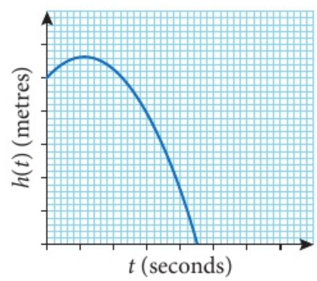


Figure 6.8 Solution to Example 6.5 (a)

(b) The height of the building is the initial height of the object, which is the height of the ball when $t = 0$, in other words, the y intercept. Therefore, the height of the building is 50 m.

(c) The time when the ball reaches its maximum height is the t-coordinate of the vertex of the parabola. We start by finding the t coordinate, which is also the location of the axis of symmetry:

$$t = -\frac{b}{2a} = -\frac{11}{2(-4.9)} = 1.12 \text{ s (3 s.f.)}$$

(d) The maximum height reached by the ball is found by substituting the t value from (c) back into the function:

$$h(1.12) = -4.9(1.12)^2 + 11(1.12) + 50 = 56.2\text{ m (3 s.f.)}$$

(e) The ball hits the ground when height is zero. Therefore, we must solve the equation $0 = -4.9t^2 + 11t + 50$

We can use the quadratic formula to solve this:

$$t = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a} = \frac{-11 \pm \sqrt{11^2 - 4(-4.9)(50)}}{2(-4.9)} = 4.51\text{ s (3 s.f.)}$$

(f) We should limit the domain to non-negative values. Also, the model doesn't make sense after the ball hits the ground since the graph suggests that the ball goes underground. Therefore a reasonable domain for this model is $0 \leqslant t \leqslant 4.51$

The range indicates the possible heights of the ball. Here, $0 \leqslant h(t) \leqslant 56.2$ makes sense since we know the maximum height of the ball and we presume the ball does not go underground.

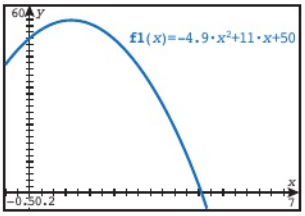


Figure 6.9 GDC approach (a)

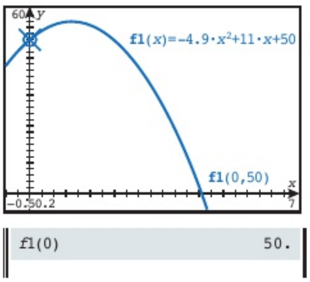


Figure 6.10 GDC approach (b)

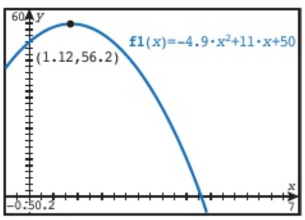


Figure 6.11 GDC approach (c)

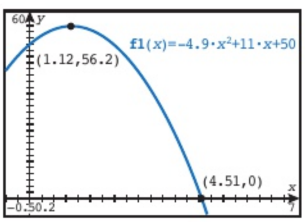


Figure 6.12 GDC approach (e)

GDC approach

(a) Here we use the GDC to obtain a graph, as shown in Figure 6.9. We need to be careful to adjust the viewing window appropriately.

On our calculator, we need to use x in place of t. It is important to note that this graph is height versus time – it is not the trajectory (path) of the object. The horizontal axis represents time, not distance.

(b) The Graph Trace feature can be used to evaluate the function at $t = 0$ to find the initial height, as shown in Figure 6.10. We could also evaluate the function directly.

(c) We can use our GDC to find the maximum height, as in Figure 6.11.

We see that the x coordinate is 1.12, so the ball reaches a maximum height at 1.12 s.

(d) Using the same point from part (c), we see that the y coordinate is 56.2, so the ball's maximum height is 56.2 m.

(e) We can use our GDC to find the positive zero of the function, as in Figure 6.12. The ball hits the ground after 4.51 s.

(f) We use the same logic as in the algebraic approach to conclude that a reasonable domain is $0 \leqslant t \leqslant 4.51$ and a reasonable range is $0 \leqslant h(t) \leqslant 56.2$

In the next few examples we will look at building a quadratic function.

Example 6.6

The number of jeans sold by a clothing store can be modelled by the function $N = 1000 - 5p$, where N is the number of items sold and p is price of the jeans in euros.

(a) Develop a model for the revenue earned from selling these jeans based on the selling price p.

(b) Use your model to find the price the jeans should be sold at in order to maximise revenue.

(c) Find the maximum revenue predicted by your model.

(d) Give a reasonable domain and range for your model.

Solution

(a) In general, revenue = selling price $\times$ number of units sold. Therefore, we can develop a model for revenue by multiplying the selling price, p, by the expression for the number of jeans sold:

$$R = (p)(1000 - 5p) \Rightarrow R = -5p^2 + 1000p$$

(b) Since this is a quadratic model with a concave-down graph ($a < 0$), we know there will be a maximum at the vertex. In this case, the p coordinate of the vertex is the selling price and the R coordinate is the revenue. To find the price, we use the formula for the p coordinate of the vertex:

$$p = -\frac{b}{2a} = -\frac{1000}{2(-5)} = 100 \text{ euros}$$

(c) To find the maximum revenue, we need to evaluate the model for 100 jeans (from part (b)):

$$R = -5(100)^2 + 1000(100) = 50\,000 \text{ euros}$$

(d) Clearly, our model doesn't make much sense if the price p is less than zero. But what about a maximum price? Since we know that this is a concave-down quadratic function with a vertex above the p axis, we know there will be two p intercepts. Solve for them algebraically:

$$0 = -5p^2 + 1000p \Rightarrow 0 = (-5p)(p - 200) \Rightarrow p = 0,\ p = 200$$

Therefore the p intercepts are 0 and 200. So $0 \leqslant p \leqslant 200$ is a reasonable domain. The range is given by the minimum and maximum values of the function on this domain: $0 \leqslant R \leqslant 50\,000$

It's often helpful to generate a graph of a model to get a picture of it. In this case, we can clearly see the maximum revenue and can use our GDC to verify our results in (c) and (d) , as shown in Figure 6.13.

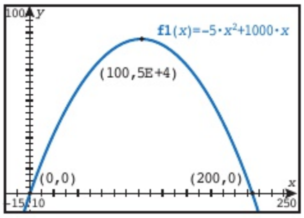
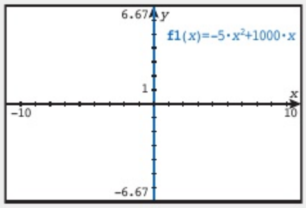


Figure 6.13 The graph of the model for revenue vs price of jeans

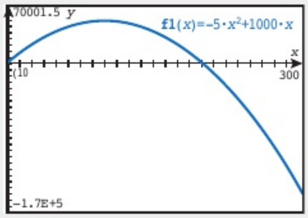
Note that 5E+4 for the y coordinate of the vertex is the GDC's version of scientific notation. We should read it like this: $5\text{E}{+}4 = 5 \times 10^4 = 50\,000$

When using a GDC to analyse a model graphically, it's important to take care when setting the viewing window. The default view for many GDCs is $-10 \leqslant x \leqslant 10$ and $-10 \leqslant y \leqslant 10$ or smaller. This view can be very misleading. For example, the model in Example 6.5 looks like this with the default viewing window.



The graph is there but with the current settings it is almost indistinguishable from the y axis. We can use our knowledge of the function to choose more appropriate settings.

After setting the x-axis to a suitable domain, in this case $0 \leqslant x \leqslant 300$, and using the Zoom Fit feature to scale the y-axis to fit the function, we get the image shown.



The negative part of the graph isn't very useful, but you can then use Zoom Box to examine the function more accurately.

Example 6.7

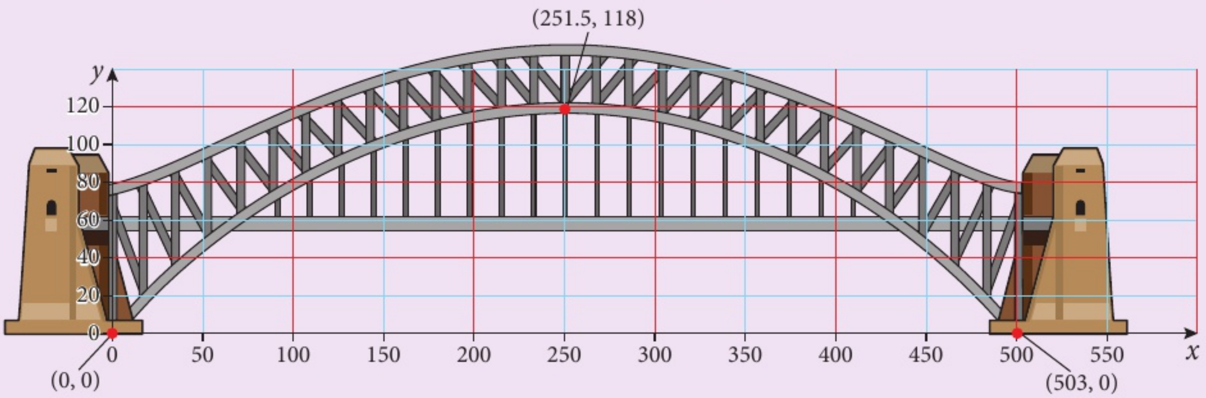
The Sydney Harbour Bridge is supported by two spans that can be modelled by quadratic functions.

The lower span is approximately 503 m wide and 118 m tall at its highest point.

Develop a quadratic model for the lower span such that one end of the span is positioned at (0, 0).

Solution

We can start by drawing our axes and locating the vertex at the maximum point of the lower span. Since the vertex must be halfway along the length of the bridge, its x coordinate must be 251.5.



Then we can use the general quadratic model $y = ax^2 + bx + c$ and some algebra to find the model for the lower span. First, since the graph must pass through (0, 0), it must be true that

$$0 = a(0)^2 + b(0) + c \Rightarrow c = 0$$

Upon reflection, we could have deduced the value of c from recognising that the y intercept of the graph is at (0,0). So far, our model is therefore

$$y = ax^2 + bx$$

Next, since the graph must also pass through (251.5, 118) and (503, 0), we can generate a system of equations by substituting each coordinate pair into the model:

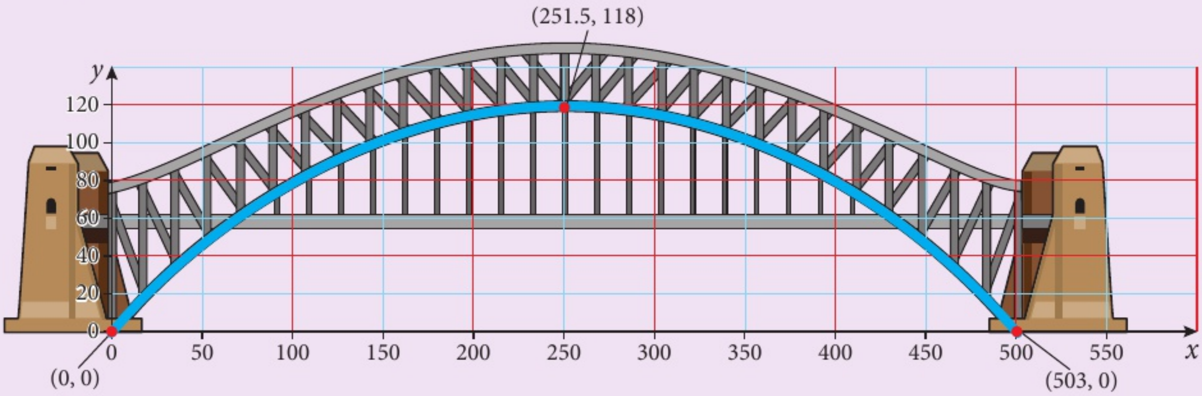
$$\begin{cases} 118 = a(251.5)^2 + b(251.5) \\ 0 = a(503)^2 + b(503) \end{cases} \Rightarrow \begin{cases} 63\,252.25a + 251.5b = 118 \\ 253\,009a + 503b = 0 \end{cases}$$

We can solve this system (algebraically, or by using a GDC) to obtain:

$$a = -0.001866,\ b = 0.9384$$

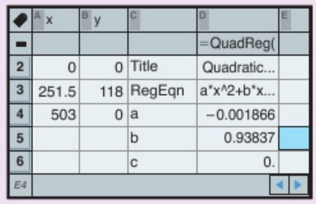
Therefore, our model is $y = -0.001866x^2 + 0.9384x$

If we graph this function with appropriate axes, we can see that it fits the lower span very well:



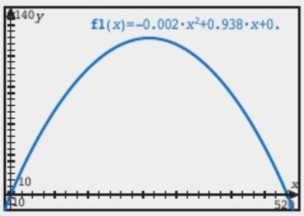
Note that most GDCs have a **quadratic regression** feature that can also find the model.

Notice that the values for a and b calculated by the GDC agree with our values.

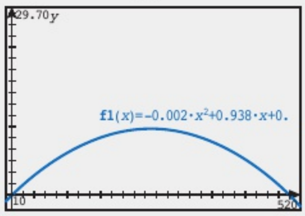


	A x	B y	C	D	E
▬				=QuadReg(	
2	0	0	Title	Quadratic...	
3	251.5	118	RegEqn	a*x^2+b*x...	
4	503	0	a	−0.001866	
5			b	0.93837	
6			c	0.	
E4					

Remember that the graph may appear differently depending on the viewing window chosen on your graph. For example, in the first screenshot, the scale of the units on the x and y axes is not 1:1, so the function appears taller than it should.



A graph can appear distorted if the x and y axes are not scaled in a 1:1 ratio



Using the Zoom Square feature changes the x and y axes to a 1:1 ratio

Using the GDC's Zoom Square function fixes this, as shown in the second screenshot.

Exercise 6.2

1. On Earth, the position of a falling object can be modelled by the function $h(t) = -4.9t^2 + v_0t + h_0$ where $h(t)$ is the height in metres after t seconds, v_0 is the initial velocity and h_0 is the initial height.

(a) Write a model for the height of a ball thrown upwards with an initial velocity of $5\ \text{m s}^{-1}$ from the roof of a 60-metre tall building.

(b) Use your model to find:

(i) the maximum height of the ball

(ii) the time until the ball hits the ground

(iii) the interval of time for which the ball is more than 50 metres above the ground.

2. A small manufacturing company makes and sells x machines each month. The monthly cost, C, in dollars, of making x machines is given by

$$C(x) = 0.35x^2 + 3200$$

The monthly revenue, R, in dollars, obtained by selling x machines is given by $R(x) = 180x - 0.55x^2$

Remember that profit = revenue − cost.

(a) Show that the company's monthly profit can be calculated using the quadratic function $P(x) = -0.9x^2 + 180x - 3200$

(b) The maximum profit occurs at the vertex of the function $P(x)$. How many machines should be made and sold each month for a maximum profit?

(c) If the company does maximise profit, what is the selling price of each machine?

(d) Find the smallest number of machines the company must make and sell each month in order to make a positive profit.

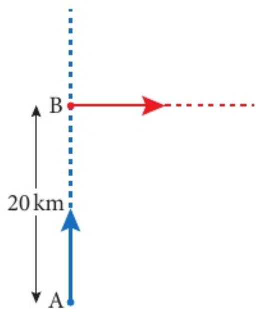


Figure 6.14 Diagram for question 3

3. The diagram shows two ships, A and B. At noon, ship A was 20 km due south of ship B. Ship A was moving north at 10 km h^{-1} and ship B was moving east at 4 km h^{-1}

Find the distance between the ships at:

(a) 13:00

(b) 14:00

Let $s(t)$ be the distance between the ships t hours after noon, for $0 \leqslant t \leqslant 4$

(c) Show that $s(t) = \sqrt{116t^2 - 400t + 400}$

(d) Sketch the graph of $s(t)$

(e) Due to poor weather, the captain of ship A can see another ship only if they are less than 9 km apart.

(i) Find the values of t during which ship A can see ship B.

(ii) Write down the times between which ship A can see ship B.

4. Worldwide grain production for the years 1965 to 2000 can be modelled by the function $G = -0.144t^2 + 6.88t + 266$, where G is the number of kilograms per person and t is years since 1965.

(a) Based on this model, what was the grain production in 1975?

(b) What was the maximum level of grain production and when did it occur?

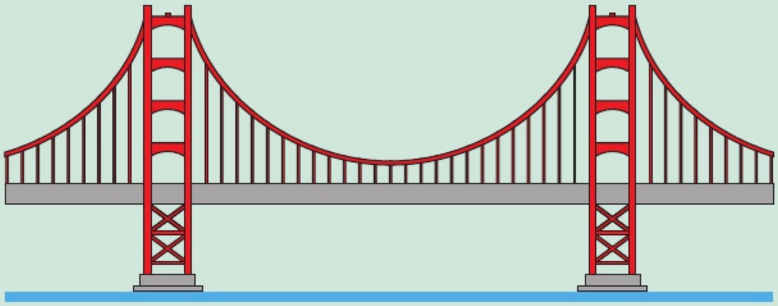
(c) The actual worldwide grain production in 2005 was 10 kg per person. What does this model predict for 2005?

(d) If worldwide grain production drops below 100 kg per person, major economic and health consequences are possible. Based on this model, when might this occur?

(e) Give a reason why this model will not (we hope) continue to be true.

5. The density of water based on temperature follows a quadratic model. The maximum density of water is 1 g ml^{-1} at 4°C. At 80°C, the density is 0.97183 g ml^{-1}.
   (a) Find a quadratic model for density, D, in terms of temperature, T, in the form $D = a(T - h)^2 + k$ where a, h, and k are constants to be determined.
   (b) Use your model to find the density of water at 0°C to 5 significant figures.
   (c) At what temperature $t > 0$ does the density of water drop below 0.960 g ml^{-1}?

6. The main span of the Verrazzano-Narrows bridge in New York City has a central span that is about 1300 m wide.



   The main cable supporting this span is about 150 m above the roadway at the top of each suspension tower, and about 6 m above the roadway at its lowest point in the centre.
   (a) Find a quadratic function to model the height h of the cable above the roadway in terms of the distance d from the left suspension tower.
   (b) Write down the domain and range of your function.
   (c) Calculate the height of the cable at a point 100 m from the left suspension tower.
   (d) At what distances from the left suspension tower is the cable less than 50 m above the roadway?

7. A farmer wants to fence two identical adjacent fields as shown in Figure 6.15.
   He has 1200 m of fencing to enclose the two identical regions.
   (a) Write down an expression for the total area, A, in terms of x.
   (b) Find the maximum total area for the two fields and the dimensions x and y.

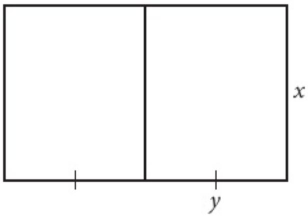


Figure 6.15 Diagram for question 7

8. A shop sells t-shirts for €16 each and ls 40 t-shirts per day. Analysing past sales shows that for every euro increase in price, the shop sells 2 fewer t-shirts per day. Let x be the price increase in euros.
   (a) Write an expression for the price of a single t-shirt in terms of x.
   (b) Write an expression for the number of t-shirts the shop can expect to sell per day in terms of x.

(c) For what value of x should the shop expect to sell no t-shirts?

(d) Hence write an expression for R, the expected revenue per day, in terms of x.

(e) Find the maximum value of R, and the price the shop should set in order to attain that maximum.

9. A decorative archway follows a parabolic curve. The inside height is 5 m and the inside width is 6 m. A large truck must pass through this archway. The truck is 4.3 m tall and 2.6 m wide.

(a) Show that the truck will not fit through the archway.

(b) Find the maximum height of a 2.6 m wide truck that can fit through the archway.

(c) Find the maximum width of a 4.3 m tall truck that can fit through the archway.

10. A tennis player hits a ball straight up. The height of the ball above the ground is described by the model $h(t) = -4.9t^2 + kt + 1$, where $h(t)$ is the height in metres at time t seconds after the ball is hit.

(a) Find the h intercept and interpret in context.

(b) Given that the tennis ball hits the ground after 3.2 seconds, find the value of k.

(c) The tennis ball is more than 10 m above the ground during the time $a < t < b$. Find the values of a and b.

(d) Find the average speed of the tennis ball for the first 0.5 seconds.

(e) Find the maximum height of the tennis ball and the time at which this occurs.

6.3 Cubic models

Cubic models are slightly more complex than quadratic models. We often encounter cubic models when we are dealing with quantities based on volume (such as optimising the volume or surface area of a package, or calculating forces from wind or water). Cubic models are also widely used in computer graphics, in everything from modelling the curves of the letters in this textbook to smoothing computer-generated effects and animation. In this section, we will examine cubic models of the form

$$f(x) = ax^3 + bx^2 + cx + d$$

Note that a simple cubic model, as in the next example, may have b, c, and d equal to zero. In that case, it can also be considered a direct variation model, which we will study later in this chapter.

Example 6.8

The maximum theoretical power that can be generated by a wind turbine can be modelled by the function

$$P = 0.297AdV^3$$

where P is the power in watts

A is the area swept by the turbine blades (swept area), in m^2

d is the air density, in kg m^{-3}

V is the wind speed in m s^{-1}

A certain wind turbine has a swept area of $80\ \text{m}^2$ and is located at sea level, where the air density is $1.225\ \text{kg m}^{-3}$

(a) Find the cubic model for this wind turbine.

(b) Use your model to calculate the maximum theoretical power generated when the wind speed is $10\ \text{m s}^{-1}$

(c) Given that wind speeds above $20\ \text{m s}^{-1}$ are strong enough to cause damage, give a reason why this turbine will not produce more than 300 000 watts.

(d) Determine a reasonable domain for this model.

Solution

(a) We substitute the known values to find the model for this particular turbine. Therefore, the cubic model for this turbine is

$$P = 0.297AdV^3 = 0.297(80)(1.225)V^3 \Rightarrow P = 29.1V^3$$

(b) The maximum theoretical power generated when the wind speed is $10\ \text{m s}^{-1}$ is

$$P = 29.1V^3 = 29.1(10)^3 = 29\,100 \text{ watts (3 s.f.)}$$

(c) Using the model, we can find the wind speed required to produce 300 000 watts.

$$P = 29.1V^3 \Rightarrow 300\,000 = 29.1V^3 \Rightarrow V = \sqrt[3]{10\,309} = 21.8\ \text{m s}^{-1}$$

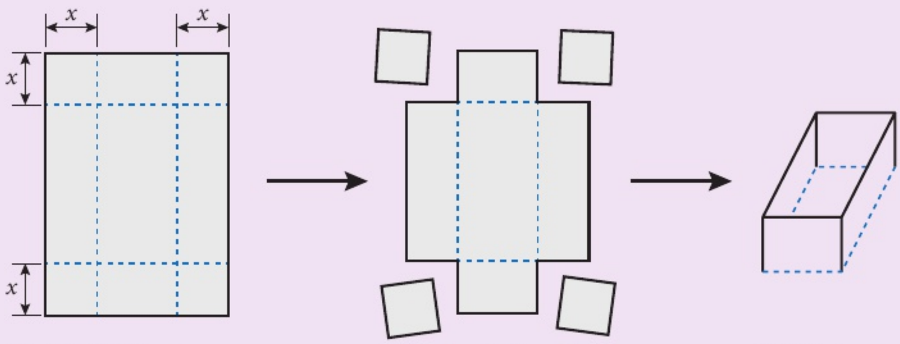
Since this is greater than the wind speed we are told will cause damage, it is not reasonable to expect this turbine to generate more than 300 000 watts.

(d) We are told that wind speeds above $20\ \text{m s}^{-1}$ are strong enough to cause damage, so that can be the upper limit for our domain. It doesn't make sense to predict power for negative wind speeds, so a reasonable domain is $0 \leqslant V \leqslant 20$

We often use models to find **optimal** solutions: that is, we are interested in maximising or minimising a quantity. Example 6.9 examines a classic problem.

Example 6.9

The dimensions of a piece of A4 paper, to the nearest centimetre, are 21×30 cm. It is possible to create an open box by cutting out square corners and folding the remaining flaps up, as shown in the diagram.



Mark 4 fold lines, equidistant from the edges of the paper

Cut out and discard corners

Fold into an open box

(a) Develop a cubic model for the volume of the open box that can be created using this method.

(b) Determine a reasonable domain for your model.

(c) Find the dimensions of the open box with the largest volume that can be created using this method.

(d) Calculate the maximum volume of the open box.

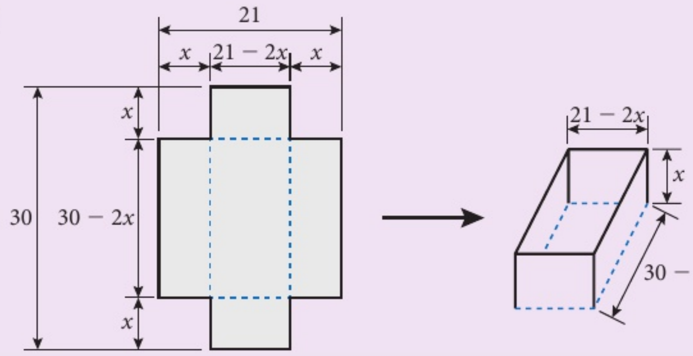
Solution

The first step in solving this problem is to find an appropriate model. Since we are interested in the volume of the box, it seems appropriate to start with a model for volume:

$$V = lwh$$

Then we need to think about what we know. The paper starts as 21×30 cm but those are not the dimensions of the open box. If we look at the second step in the first diagram, we can see that part of the width and length of the paper becomes the height for the box

(a)



The width and length of the paper get reduced by twice the length of the corners we cut out. Therefore, we know that

$$\text{length} = 30 - 2x, \text{ width} = 21 - 2x, \text{ height} = x$$

Now we can use those to develop a model for the volume of the box by substituting into the general model for volume:

$$V = lwh = (30 - 2x)(21 - 2x)(x)$$

(b) It doesn't make sense to remove a square with negative or zero length, so $x > 0$.

Is there an upper bound? What is the largest corner we can cut out? We are limited by the width of the paper. Since the paper is 21 cm wide, we need to cut less than $\frac{1}{2}(21) = 10.5$ cm from the edge in order to make an open box. Therefore, a reasonable domain is $0 < x < 10.5$

(c) Our goal is to find the maximum volume. To do this, we can use our GDC to graph the model and look for a maximum value.

Remember, when using a GDC to analyse a model, it's important to choose the viewing window carefully. Since the x axis represents the distance of each fold from the edge of the paper (which is also equal to the size of each square we cut out), we are only interested in positive x values. Also, because our paper is only 21 cm wide, x must be less than $\frac{21}{2} = 10.5$

Therefore, we set the window to $0 \leqslant x \leqslant 10.5$ and use the GDC's Zoom Fit to scale the y axis accordingly.

From the graph shown in Figure 6.16, we conclude that the value of x that produces the maximum volume is 4.06 cm. To find the dimensions, we need to go back to our expressions for the length, height, and width of the box. We can use these to obtain the missing dimensions:

$$\text{length} = 30 - 2x = 30 - 2(4.06) = 21.9 \text{ cm}$$

$$\text{width} = 21 - 2x = 21 - 2(4.06) = 12.9 \text{ cm}$$

$$\text{height} = x = 4.06 \text{ cm}$$

(d) The volume of the open box is given in the GDC output since the volume is the y coordinate in our GDC. Therefore, $V = 1140 \text{ cm}^3$.

If we expand the model for volume, we get

$V = (30 - 2x)(21 - 2x)(x)$
$\Rightarrow V = 4x^3 - 102x^2 + 630x$

This shows us that this is indeed a cubic model. However, since we are going to use our GDC to analyse this model, there is no need to expand the model.

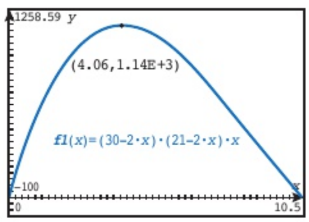


Figure 6.16 GDC screen for Exercise 6.9 (c)

Remember that $1.14\text{E}{+}3 = 1.14 \times 10^3 = 1140 \text{ cm}^3$

Exercise 6.3

1. Re-do Example 6.9 using a piece of paper with dimensions 8.5×11 inches (such as standard letter paper).

2. Re-do Example 6.9 using a piece of A3 paper, with approximate dimensions 30×42 cm.

t (seconds)	h (metres)
1	105
2	98
3	84
4	60
5	26

Table 6.4 Data for question 3

3. A rock falls off the top of a cliff. Let h be its height above ground in metres, after t seconds. Table 6.4 gives values of h and t.

   Jane thinks that the function $f(t) = -0.25t^3 - 2.32t^2 + 1.93t + 106$ is a suitable model for the data. Use Jane's model to:

   (a) write down the height of the cliff

   (b) find the height of the rock after 4.5 seconds

   (c) find after how many seconds the height of the rock is 30 m.

   Kevin thinks that the function $g(t) = -5.2t^2 + 9.5t + 100$ is a better model for the data.

   (d) Use Kevin's model to find the point at which the rock hits the ground.

   (e) Create graphs of f, g, and the data given. By comparing the graphs of f and g with the plotted data, explain which function is a better model for the height of the falling rock.

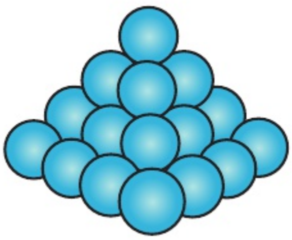


Figure 6.17 Diagram for question 4

4. An efficient way to stack cannonballs is as a pyramid with a square base, as shown in Figure 6.17.

   The balls are stacked such that there is 1 ball on the first layer, 4 balls on the second layer, 9 balls on the third layer, and so on.

   (a) Copy and complete the table.

Number of layers in stack, n	1	2	3	4
Total number of balls, B	1	5		

   (b) The total number of balls, B, can be modelled by a cubic function $B = an^3 + bn^2 + cn$. Using your GDC, or otherwise, find the exact values of a, b, and c.

   (c) Find the total number of balls in a stack with 10 layers.

   (d) Find the number of layers required to stack 819 balls.

5. The cumulative number of HIV AIDS cases reported in the United States from 1983 to 1998 follows the cubic model

   $$C = -222t^3 + 7260t^2 - 12\,700t + 13\,500$$

   where C is the cumulative number of HIV AIDS cases, and t is the number of years since 1983.

   (a) Find the cumulative number of reported cases in 1990.

   (b) Find the year in which the number of cases exceeds 500 000.

   (c) If this pattern continues, in what year will the maximum number of cases be reported?

   (d) Give a reason why this model will probably not be correct past 1998.

6. A colony of bacteria is exposed to a drug that stimulates reproduction. The number of bacteria is given by the model $P = 1200 + 17t^2 - t^3$ where P is the number of bacteria present t minutes after the drug is introduced, $0 \leqslant t \leqslant 20$

   (a) Find the number of bacteria present when the drug is first introduced.

   (b) Find the number of bacteria present after 5 minutes.

   (c) At what time are there 1000 bacteria?

   (d) Find the maximum number of bacteria and the time at which this occurs.

   (e) At what time are there no bacteria remaining?

7. Researchers are monitoring how a particular drug causes patients' body temperatures to change. They measure the patients' body temperatures just before and 1 hour after the drug is given. After x milligrams per kg of body mass of a drug are given, body temperature increases according to the model $\Delta T(x) = 0.25x^2(1 - 0.1x)$, $0 \leqslant x \leqslant 10$, where ΔT is the change in °C.

   (a) How much will a patient's body temperature change when 4 mg kg^{-1} are given?

   (b) Dosages between a and b mg kg^{-1} will increase patients' body temperatures by at least 3°C. Find the values of a and b.

   (c) What is the maximum body temperature increase from this drug, and at what dosage does it occur?

6.4 Exponential models

Exponential models arise in situations where the rate of change is a constant **factor**, that is, when the next value is found by multiplying by a constant factor. This sort of change can produce surprising results.

Just how fast is exponential growth?

A classic fable about the game of chess goes something like this:

A great long time ago a wise man invented the game of chess. This king of this land was so pleased with this new game that he offered the wise man the riches of his kingdom as a gift. The wise man replied, 'I am a simple man, and my needs are modest. Instead of your riches, please place one grain of rice on the first square of the chessboard, two grains of rice on the second square, four grains of rice on the third square, and so on, doubling the number of grains of rice for each of the remaining squares.'

The king replied, 'What a silly man! I offer him the riches of my kingdom and all he asks for is a few grains of rice!'

How much rice did the wise man ask for? To investigate this, let's build a table. Since we are multiplying by 2 each time, we can write the multiplication using exponents as shown in Table 6.5.

Square	Grains of rice	Grains of rice
1	1	$1 = 2^0$
2	2	$1 \times 2 = 2^1$
3	4	$1 \times 2 \times 2 = 2^2$
4	8	$1 \times 2 \times 2 \times 2 = 2^3$

Table 6.5 Building the table

Now, we could keep multiplying by 2 each time until we get to the 64th square, but that seems like a lot of work. If we look carefully, we can see a pattern: the exponent of 2 is equal to one less than the number of the square. Now we can add a few more rows as shown in Table 6.6.

Square	Grains of rice	Grains of rice
1	1	1
2	2	2^1
3	4	2^2
4	8	2^3
…	…	…
n		2^{n-1}
64		2^{63}

Table 6.6 Adding rows to the table

So, we can conclude that the king must place 2^{63} grains of rice on the last square. How much rice is that? If we estimate that one kilogram of rice contains approximately 50 000 grains of rice, then we have

$\frac{2^{63}}{50\,000} = 184\,467\,440\,737\,095 = 1.84 \times 10^{14}$ kg of rice. Is that a lot?

According to the Food and Agriculture Organisation of the United Nations, the estimated worldwide production of rice in 2017 was 759.6 million tonnes, or 7.596×10^{11} kg of rice. Therefore, the amount of rice on the last square alone is $\frac{1.84 \times 10^{14}}{7.596 \times 10^{11}} \approx 240$ times more than the entire worldwide harvest of rice in 2017!

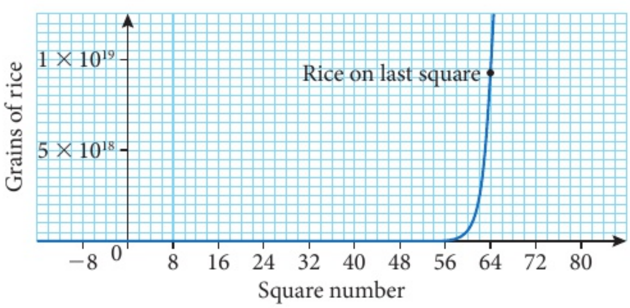


Figure 6.18 $y = 2^{x-1}$

So, in situations where exponential models apply, we can expect to see very fast increases or decreases. To see this more clearly, let's graph the model we developed, $y = 2^{x-1}$, shown in Figure 6.18.

Most calculators can't even graph numbers this large! At the scale of this graph, it looks like there are almost no grains of rice until somewhere around the 56th square. Only if we zoom in to the very first few squares can we see some of the initial growth, as shown in Figure 6.19.

So, both visually and numerically, we can see that exponential functions can describe very rapid change. We will need to keep this in mind when we consider the reasonableness of our models.

Of course, doubling the number of grains of rice on each square of a chessboard might seem like an extreme example. However, many real-life phenomena double in a fixed interval. For example, a single bacterium in ideal conditions can divide itself into two bacteria every 15 minutes! Furthermore, any amount that increases by a fixed fraction will also double in a fixed interval. For example, the mean university tuition increases in the USA is 4.2% per annum. This doesn't sound like much, but it means that the tuition rates have doubled about every 17 years. (Think: for a student at a US university, the cost of their children's education will be at least twice as much as their own, if they have children right away.)

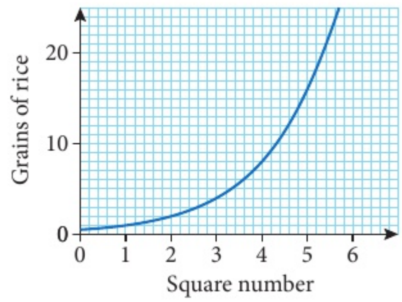


Figure 6.19 Zooming in on the graph

Developing exponential models

Exponential models can sometimes be developed from examining a table, as in the case of the grains-of-rice-on-a-chessboard example. Other times we know about or can hypothesise a percentage growth, doubling time, or growth factor. Any of these can be used to develop an exponential model. Example 6.10 looks at a situation where the growth rate is known.

Example 6.10

Suppose that a population of wombats is increasing at 7% per annum. The population at the beginning of 2018 is recorded to be 240 individuals. Assume that this rate of growth remains constant.

(a) Develop a model for the population of wombats over time.

(b) Use your model to predict the number of wombats at the beginning of 2025.

(c) How long will it take the population of wombats to double?

(d) An ecologist estimates that this region can sustain at most 1000 wombats. In what year will the population reach 1000 wombats?

Solution

(a) To develop a model, we start by making a table and see if we can see a pattern. We start by simply adding 7% for each year. However, to simplify our model, we start with year 0 as 2018.

Year	Years since 2018	Calculation of number of wombats	Number of wombats
2018	0	240	240
2019	1	$240 + 240 \times 0.07$	256.8
2020	2	$256.8 + 256.8 \times 0.07$	274.8
2021	3	$274.8 + 274.8 \times 0.07$	294.0

It is common in models where we are looking over time to think of the starting date as time 0. This could be the starting hour, day, month, or year, but doing so will simplify the calculations in our models. However, we have to be careful to convert back to actual times or dates when interpreting our results.

Although it doesn't make sense to have 256.**8** wombats, we keep the decimal value in order to make our subsequent calculations more accurate. On your GDC, you should store the previous results to maintain full precision in your calculations, as shown below.

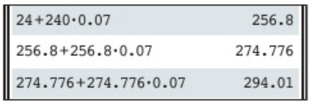


Figure 6.20 Store results to maintain full precision

There doesn't seem to be an obvious pattern so far. However, if we make a change to how we calculate the number of wombats we might have better luck. The key insight relies on a clever factorisation of our calculation:

$240 + 240(0.07)$	Original expression
$= 240(1) + 240(0.07)$	Multiplying by 1 does not change the value of the expression
$= 240(1 + 0.07)$	Distributive property
$= 240(1.07)$	Simplify

Now we rewrite our table:

Year	Years since 2018	Calculation of number of wombats	Number of wombats
2018	0	240	240
2019	1	240(1.07)	256.8
2020	2	256.8(1.07)	274.8
2021	3	274.8(1.07)	294.0

Finally, notice that we are multiplying 1.07 by the previous result each time. But, each previous result is also from multiplying by 1.07. So, we can then write the table like this:

Year	Years since 2018	Calculation of number of wombats	Number of wombats
2018	0	240	240
2019	1	240(1.07)	256.8
2020	2	240(1.07)(1.07)	274.8
2021	3	240(1.07)(1.07)(1.07)	294.0

Now we see that we are repeating the same multiplication, adding another factor of 1.07 in each row. This means that the exponent of the 1.07 factor is equal to the year since 2018. This allows us to write an exponential model in the final row of our table:

Year	Years since 2018	Calculation of number of wombats	Number of wombats
2018	0	240	240
2019	1	$240(1.07)^1$	256.8
2020	2	$240(1.07)^2$	274.8
2021	3	$240(1.07)^3$	294.0
$2018 + n$	n	$240(1.07)^n$	$240(1.07)^n$

Expressed as a function, our model is $P(n) = 240(1.07)^n$ where $P(n)$ is the population of wombats n years since 2018.

(b) Since 2025 is $2025 - 2018 = 7$ years since 2018, our model predicts that the population of wombats will be $P(7) = 240(1.07)^7 = 385$ wombats.

(c) To double, we need to solve the equation $P(n) = 480$, i.e., $480 = 240(1.07)^n$. We can use a GDC to solve this equation. There are two common methods.

GDC graphical method

Graph the left-hand side and the right-hand side of the equation $480 = 240(1.07)^n$ as two separate functions. We need to think carefully about a suitable viewing window. For the x axis, a domain of $0 \leqslant x \leqslant 20$ is a good start. Since the y axis is the number of wombats, and we know we are starting with 240 and looking for the point when the population is 480, we could choose $200 \leqslant y \leqslant 600$ as a place to start and use the Intersect feature to find the solution, as shown in Figure 6.21.

The population of wombats will reach 480 after 10.2 years. Therefore, the population will double every 10.2 years.

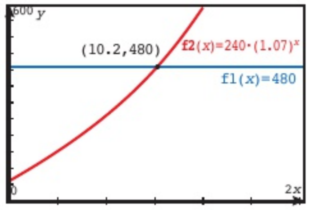


Figure 6.21 GDC graphical method

GDC numerical solver method

We could also use the numerical solver on a GDC (see Figure 6.22).

Again, we see that the population of wombats will double after 10.2 years.

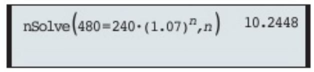


Figure 6.22 GDC numerical solver method

(d) To find the year during which the population reaches 1200 wombats, solve $P(n) = 1200 \Rightarrow 1200 = 240(1.07)^n$ to obtain $n = 23.8$ (3 s.f.)

Therefore, the population of wombats will reach 1200 after 23.8 years. But what year is that? It is $2018 + 23.8 = 2041.8$, that is, during the year 2041.

Using a numerical solver feature on a GDC is a quick way to find a solution to an equation if you know that there is only one solution. (You can tell numerical solvers to look near a starting value that you guess, but it can be hard to guess where to start!) So, be sure that you have thought about the equation you are trying to solve and are confident that there is, in fact, only one solution. Otherwise, a numerical solver can give you a solution that doesn't make sense in the context of the equation. In this case, it is often worth taking the time to use the GDC to make a good graph of the model by setting the viewing window appropriately.

We've now looked at two examples of exponential models relating to growth. The technique we used for developing the model is general. (In fact, it even works when there is a constant rate of decrease.) If we look back at the previous two examples, we can see that the general form of an exponential model is as follows.

General form of an exponential model

A quantity that is changing at a fixed fractional rate p per fixed period can be modelled by the exponential function

$$f(x) = k(1 + p)^x$$

where k is the initial value of the quantity.

In the general model, the fixed period we refer to may be days, hours, years, generations, chessboard squares, etc. Remember that p is the fraction of the previous value that is added each time. For example, if a quantity is doubling every period, we are adding 100% so $p = 1$. A quantity that increases by 10% every period has $p = 0.10$

A quantity that decreases by 15% every period has $p = -0.15$ (we call this **exponential decay**).

Sometimes, we add a constant c to the exponential model to obtain $f(x) = k(1 + p)^x + c$

The value of c represents the asymptotic value of the quantity. We will see this in the next example.

Interpreting exponential models

As with any model, we need to be able to understand what an exponential model can tell us and how it might be useful. In this example, we will look at how to interpret an exponential model.

Example 6.11

The value of a certain new car decreases according to the model $V(t) = 39\,000(0.72)^t + 1000$ where $V(t)$ represents the value in euros of the car t years after is it purchased.

(a) Find the original purchase price of the car.

(b) Interpret the meaning of the value 1000 in the model.

(c) Calculate when the value of the car will be half its original value.

Solution

(a) The original purchase price is $V(0) = 39\,000(0.72)^0 + 1000 = 39\,000 + 1000 = 40\,000$ euros.

(b) The expression $39\,000(0.72)^t$ tends toward zero as t increases. Therefore, the value 1000 in the model represents the eventual or residual value of the car. (This is probably the value of the car as scrap!)

(c) Since the original price of the car was 40 000, we are looking for $V(t) = 20\,000$ hence $20\,000 = 39\,000(0.72)^t + 1000$

Our GDC numerical solver is useful here (see Figure 6.23).

Therefore, the car will be worth half its original value after about 2.19 years.

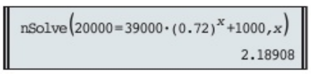


Figure 6.23 GDC numerical solver

Approaching a constant value
If a quantity is approaching a constant value c from an initial value $k + c$, and the difference $k - c$ is changing at a fixed fractional rate p per fixed period, the growth or decay in the quantity can be modelled by the exponential function $f(x) = k(1 + p)^x + c$

Notice that when we use an exponential model of the form $f(x) = k(1 + p)^x + c$, the value k is no longer the initial value.

The generic statement of the model makes sense: if we consider Example 6.11, the original purchase price of €40 000 is not decreasing by a fixed percentage. Rather, it is the amount above the residual value of €1000 that is decreasing at a fixed percentage per period. Table 6.7 illustrates this idea.

Years since purchase	Value (€)	Percent change in total value of car	Value (€) above residual value of €1000	Percent change in value above residual value of car
0	40 000		39 000	
1	29 080	−27.3%	28 080	−28.0%
2	21 218	−27.0%	20 218	−28.0%
3	15 557	−26.7%	14 557	−28.0%
4	11 481	−26.2%	10 481	−28.0%

Table 6.7 The amount above the residual value decreases at a fixed percentage per period

It's common in science applications to use a different form of the exponential model. Instead of changing the base of the exponent by adding the fractional rate p to 1, we add a parameter r as a factor in the exponent and use the special base e: $f(x) = ke^{rx}$. The value of r is often determined through experimental evidence. It turns out that the two models are equivalent, as Example 6.12 will show.

Example 6.12

A student measures the temperature of a cup of coffee at regular intervals in a room where the temperature is 20°C. Table 6.8 shows her data.

(a) Develop an exponential model of the form $f(x) = ke^{rx} + c$, where k, r, and c are constants to be determined.

(b) Use your model to find the temperature of the coffee after 8 minutes.

(c) Use your model to predict when the temperature of the coffee will reach room temperature, to 3 significant figures.

(d) Find the rate of decrease in the difference between the coffee temperature and room temperature as a percentage per minute.

Time (minutes)	Temperature (°C)
0	95
1	79.00
2	66.41
3	56.51
4	48.72

Table 6.8 Data for Example 6.12

Solution

(a) We know the coffee is cooling, so our model must be a decreasing exponential function. We also know that a decreasing exponential function will tend towards zero unless a constant is added, so the expression ke^{rx} must represent the (decreasing) difference between the room temperature and the coffee, while c represents room temperature (20°C). Therefore, $c = 20$. Since $f(0) = 95$, we have

$$95 = ke^{r(0)} + 20 \quad \Rightarrow \quad k = 75$$

Thus, so far we have $f(x) = 75e^{rx} + 20$

Using the data value from the first minute, we know $f(1) = 79.0$ so

$79.0 = 75e^{r(1)} + 20$	Substitute known values
$0.787 = e^{r}$	Subtract 20 and divide by 75
$\ln 0.787 = r \ln e$	Apply natural logarithm and use $\ln a^n = n \ln a$
$-0.240 = r$	Evaluate with GDC

Therefore, the model is

$$f(x) = 75e^{-0.240x} + 20$$

(b) After 8 minutes, the coffee temperature is $f(8) = 75e^{-0.240(8)} + 20 = 31.0°C$

(c) Theoretically, the coffee cup will never reach room temperature according to the model. However, it will be within three significant figures of room temperature when the temperature is less than 20.05°C. So, we are looking for the time when $f(x) = 20.05$

Algebraic approach

$20.05 = 75e^{-0.240x} + 20$	Substitute known values
$6.67 \times 10^{-4} = e^{-0.240x}$	Subtract 20 and divide by 75
$\ln(6.67 \times 10^{-4}) = -0.240x \ln e$	Apply natural logarithm and use $\ln a^n = n \ln a$
$30.5 = x$	Evaluate with GDC, divide by -0.240

Therefore the coffee temperature will be within 3 significant figures of room temperature after 30.5 minutes.

GDC approach

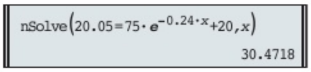


Figure 6.24 GDC numerical solver

We can enter the equation directly into the numeric solver to obtain the same result as above.

(d) At first, it might seem that we could simply calculate the percentage decrease from the data table. If we try this, we get this table.

Time (minutes)	Temperature (°C)	Percentage change from previous minute
0	95	
1	79.00	−16.8%
2	66.41	−15.9%
3	56.51	−14.9%
4	48.72	−13.8%

This doesn't make sense, since the percentage change must be constant. We need to read the question more carefully: 'Find the rate of decrease in the difference between the coffee temperature and room temperature…' Let's try another table.

Time (minutes)	Temperature (°C)	Difference between coffee temperature and room temperature	Percentage change from previous minute
0	95	75	
1	79.00	59.00	−21.3%
2	66.41	46.41	−21.3%
3	56.51	36.51	−21.3%
4	48.72	28.72	−21.3%

Therefore, the percentage change per minute in the difference between the coffee temperature and room temperature is −21.3%

Notice that in part (d), what we are doing is removing the effect of the room temperature and considering the cooling rate of the coffee. Algebraically, that means we are considering the expression $75e^{-0.240x}$. If we are looking for a percentage rate, we could convert this expression into a percentage-decrease model by using a law of exponents:

$75e^{-0.240x}$	Original expression
$= 75(e^{-0.240})^x$	Using the law $(x^a)^b = x^{ab}$
$= 75(0.787)^x$	Evaluate $e^{-0.240}$ using GDC

Notice that $1 - 0.213 = 0.787$, confirming our result that the coffee temperature difference is decreasing by a fixed rate of 21.3%

Graphical interpretation

It's worth looking at the graphs of the last few models. Notice that in each case, the location of the horizontal asymptotes is given by the value of c in the model $f(x) = k(1 + p)^x + c$

For the wombats in Example 6.10, $c = 0$ so the asymptote is at $y = 0$. In that context, the horizontal asymptote is not meaningful because it occurs in negative 'years since 2018,' so there doesn't seem to be a useful interpretation of it. For the car in Example 6.11 (not shown), $c = 1000$ so the asymptote is at $y = 1000$. In that context, it represents the eventual (residual) value of the car.

For the coffee in Example 6.12 (Figure 6.26), $c = 20$ so the asymptote is at $y = 20$. In that context, it represents the room temperature, which is the eventual temperature of the coffee.

A common use of exponential functions is half-life calculations.

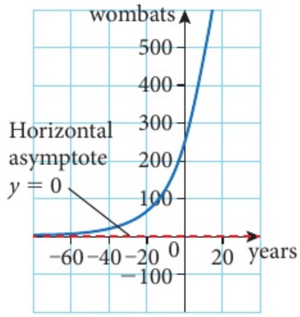


Figure 6.25 Horizontal asymptote at $y = 0$

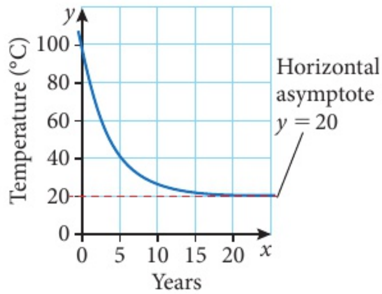


Figure 6.26 Horizontal asymptote at $y = 20$

> Half-life refers to a process of radioactive decay. A substance's half-life is the average amount of time it takes for the activity of a substance to decrease to half its previous value. Activity refers to the number of nuclei decaying at any time, measured in decays per second or Becquerel (Bq).
>
> For example, the half-life of fermium-253 is 3 days. If the activity of a sample of fermium-253 is 100 Bq, then after 3 days the activity will have decreased to 50 Bq. After another 3 days, the activity will be 25 Bq.

Horizontal asymptotes

For an exponential model of the form

$f(x) = k(1 + p)^x + c$ or
$f(x) = ke^{rx} + c$

the graph of f has a horizontal asymptote at $y = c$

Example 6.13

All living things on Earth continuously take in carbon. Since a tiny fraction of the carbon on Earth is the naturally occurring radioactive isotope carbon-14, the fraction of carbon-14 in all living things is constant. However, once an organism dies, it stops taking in carbon and the carbon-14 in its body begins to decay. Because of this, scientists can use carbon-14 dating to determine how many years have passed since an organism died.

A model for carbon-14 dating is $A(t) = ke^{-0.000121t}$, where $A(t)$ is the activity of carbon-14 remaining from an initial activity k, after t years.

(a) Show that the half-life of carbon-14 is 5730 years, to 3 significant figures.

(b) A tissue sample from a body discovered on the border between Austria and Italy has an average activity of 6.92 Bq. A sample of live tissue of the same size has an average activity of 12 Bq. Determine the age of the sample.

Solution

(a) Using the model, $A(t) = ke^{-0.000121t}$, a half-life is the time taken for the activity of carbon-14 to halve.

Therefore, we can set $k = 2$ and $A(t) = 1$ to find
$1 = 2e^{-0.000121t} \Rightarrow 5728.5 = t$

Therefore, to 3 significant figures, the half-life of carbon-14 is 5730 years.

(b) Using the model, we obtain $6.92 = 12e^{-0.000121t} \Rightarrow 4550 = t$

Therefore, the age of the tissue sample is 4550 years.

Exercise 6.4

1. When a skydiver is falling towards the Earth, they will accelerate until the force of gravity is equal to air resistance. The velocity of the skydiver at this point is called terminal velocity. A skydiver records the difference between her velocity and terminal velocity every 5 seconds and obtains the data shown in the table.

Free fall time (s)	0	5	10	15	20
Difference between velocity and terminal velocity (m s^{-1})	56	23.4	9.79	4.10	1.71

   (a) Develop an exponential model for this data.

   (b) Find the difference between velocity and terminal velocity at 7 seconds.

   (c) When is the skydiver first within 5 m s^{-1} of terminal velocity?

   (d) Predict the time when the skydiver will be within 1 m s^{-1} of terminal velocity.

2. The number of bacteria in two colonies, A and B, starts increasing at the same time. The number of bacteria in colony A after t hours is modelled by the function $A(t) = 12e^{0.4t}$

   (a) Find the number of bacteria in colony A after four hours.

   (b) How long does it take for the number of bacteria in colony A to reach 400?

   The number of bacteria in colony B after t hours is modelled by the function $B(t) = 24e^{kt}$ After four hours, there are 60 bacteria in colony B.

   (c) Find the value of k.

   (d) The number of bacteria in colony A first exceeds the number of bacteria in colony B after n hours, where $n \in \mathbb{Z}$. Find the value of n.

3. The concentration of medication in a patient's bloodstream is given by $C(t) = 9(0.5)^{0.021t}$, where C is in milligrams per litre t minutes after taking the medicine.

   (a) Write down $C(0)$.

   (b) Find the concentration of medication left in the patients' bloodstream after 40 minutes.

   (c) A patient takes the medicine at 14:00. The patient should take the medicine again when the concentration of medication reaches 0.350 milligrams per litre. What time should the patient take the medicine again?

4. A large lizard known as a Gila Monster is about 16 cm long at birth. For the first 8 years of its life, the Gila Monster's length increases by about 8% each year.

   (a) Write a function to model the length L of a Gila Monster t years after birth.

(b) Estimate the length of a 3-year-old Gila Monster.

(c) Estimate the age of a 25-cm long Gila Monster.

5. DDT is a toxic insecticide that was widely used in the past. The function $A = A_0 e^{kt}$ can be used to describe the amount A of DDT left in an area t years after an initial application of A_0 units.

(a) Given that the half-life of DDT is about 15 years, find the value of k.

(b) Find the amount of DDT left after 2 years when 50 units were initially applied to an area.

(c) A sample of soil is tested and 35 units of DDT are found. It is known that the last application of DDT was 20 years ago. Find the initial amount of DDT applied to this soil.

(d) In one area, 120 units of DDT are applied. Given that the safe level of DDT is 40 units, for how long will the area be unsafe?

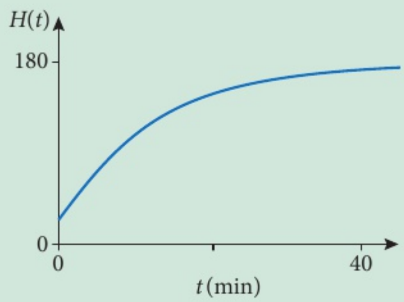
6. In certain soil conditions, the half-life of the pesticide glyphosate is about 45 days. A scientist studying how this chemical decays wrote in his notes that the data in an experiment could be modelled with the function $A(t) = 500(0.5)^t$

(a) Find $A(0)$ and interpret in context.

(b) Explain what the variable t represents in this context.

(c) Find $A(1)$ and interpret in context.

7. Sameera is baking a birthday cake. She places the cake mix in a preheated oven. The temperature in the centre of the cake mix in °C is modelled by the function $H(t) = 180 - a\,(1.08)^{-t}$ where time t is in minutes after the mix is placed in the oven. The graph of $H(t)$ is given.



(a) Write down what the value of 180 represents in the context of the question.

(b) The temperature in the centre of the cake mix was 22°C when placed in the oven. Find the value of a.

(c) The cake is removed from the oven 29 minutes after the temperature in the centre of the cake has reached 150°C.

Find the total time that the baking tin is in the oven.

8. The number of fish, N, in a pond is decreasing according to the model $N(t) = ab^{-t} + 40$, $t \geqslant 0$ where a and b are positive constants, and t is the time in months since the number of fish in the pond was first counted. The fish are first counted in January. Some data collected is shown in the table.

Month	January	May	September
Fish	850	100	?

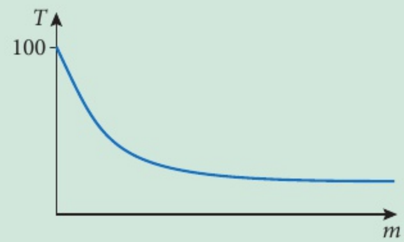
(a) Find the value of a and of b.

(b) Use your model to estimate the number of fish in September.

(c) Use your model to find the first month when there were less than 50 fish.

(d) The number of fish in the pond will not decrease below p. Write down the value of p.

9. A cup of boiling water is placed in a room to cool. The temperature of the room is 20°C. This situation can be modelled by the exponential function $T = a + b(k^{-m})$ where T is the temperature of the water, in °C , and m is the number of minutes for which the cup has been placed in the room. A sketch of the situation is given.



(a) Explain why $a = 20$

Table 6.9 shows some data about the temperature of the water.

Time (minutes)	Temperature (°C)
0	100
2	85
5	?
?	35

Table 6.9 Data for question 9

(b) Find the value of b and of k.

(c) Find the temperature of the water 5 minutes after it has been placed in the room.

(d) Find the total time needed for the water to reach a temperature of 35°C. Give your answer in minutes and seconds, correct to the nearest second.

10. A potato is placed in an oven heated to a temperature of 200°C. The temperature of the potato, in °C, is modelled by the function $p(t) = 200 - 190(0.97)^t$, where t is the time, in minutes, that the potato has been in the oven.

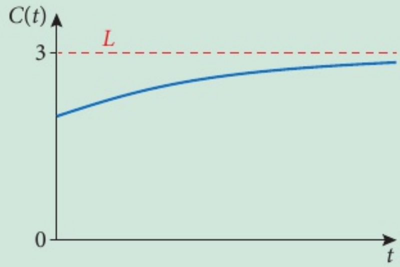
(a) Write down the temperature of the potato at the moment it is placed in the oven.

(b) Find the temperature of the potato half an hour after it has been placed in the oven.

(c) After the potato has been in the oven for k minutes, its temperature is 40°C.

Find the value of k.

11. The amount of electrical charge in a rechargeable battery is modelled by the function $C(t) = 3 - 2.4^{-t}$ where $C(t)$ is the amount of charge after t hours of charging.



(a) Write down the amount of electrical charge in the battery at $t = 0$

(b) The line L is the horizontal asymptote to the graph. Write down the equation of L.

(c) The charging of the battery should stop when it is within 1% of the maximum charge possible. After how many hours should the charging stop?

12. The number of cells, C, in a culture is given by the equation $C = p \times 2^{0.5t} + q$, where t is the time in hours measured from 12:00 on Monday and p and q are constants. The number of cells in the culture at 12:00 on Monday is 47. The number of cells in the culture at 16:00 on Monday is 53.

(a) Write down two equations in p and q .

(b) Calculate the value of p and of q.

(c) Find the number of cells in the culture at 22:00 on Monday.

6.5 Direct and inverse variation

In general, **direct variation** occurs when one quantity increases along with another quantity. **Inverse variation** occurs when one quantity decreases along with another quantity.

Direct variation and inverse variation are two types of models that occur so often in real life that they deserve a special look.

In the following two sections, we will examine direct and inverse variation models more carefully.

Direct variation

Suppose a ball is dropped from the roof of a building. The distance the ball has fallen varies directly with the square of the time since it was dropped. Symbolically, we write

$$d = at^2$$

where d is the distance travelled, t is the time elapsed, and a is a **constant of variation** that we will determine.

At the instant the ball is dropped, it has travelled 0 metres. One second later, it has travelled 4 metres.

If you have studied the falling-object model or physics, you might protest that the constant of variation must be equal to $\frac{1}{2}g = 4.91\ \text{m s}^{-1}$! However, remember that the falling-object model only perfectly describes objects falling in a vacuum or with negligible air resistance. In Earth's atmosphere, the air resistance will decrease the acceleration of the object, so that its acceleration will be less than $4.91\ \text{m s}^{-1}$. The ball in this example could be a lightweight perforated plastic ball.

By two seconds, it has travelled 16 metres. By three seconds, it has travelled 36 metres. We can express this data as ordered pairs:

$$(0, 0),\ (1, 4),\ (2, 16),\ (3, 36)$$

How can we find the constant of variation? We simply need to substitute one of the ordered pairs into the model:

$$4 = a(1)^2 \ \Rightarrow\ a = 4$$

Therefore the direct variation model is

$$d = 4t^2$$

To be sure, we can check the other points for agreement:

$$(2, 16) \ \Rightarrow\ d = 4\,(2)^2 = 16\ \checkmark$$

$$(3, 36) \ \Rightarrow\ d = 4\,(3)^2 = 36\ \checkmark$$

Notice that we cannot use the point (0, 0) to find the constant of variation. Why not? The point (0, 0) produces the equation $0 = 0$

Direct variation models are simplified polynomial models of the form

$$y = ax^n \text{ where } n \in \mathbb{Z},\ n > 0$$

The constant a is sometimes called **the constant of variation**.

In common usage, direct variation is often used only to refer to linear direct variation models, where $y = ax$

Here we use a broader definition, as shown in the key fact box.

Example 6.14

The cost C of a phone call varies directly with the length of the call m in minutes. A recent 5-minute phone call cost \$0.60.

(a) Write a direct variation model for this situation.

(b) Use your model to predict the cost of a 7-minute phone call.

(c) Find the length of a call that cost \$1.56.

(d) Write down the cost per minute for phone calls.

Solution

(a) Since a 5-minute phone call cost \$0.60, we have

$$C = am \ \Rightarrow\ 0.60 = a(5) \ \Rightarrow\ a = 0.12 \text{ therefore the model is } C = 0.12m$$

(b) A 7-minute phone call would cost $C = 0.12(7) = \$0.84$

(c) A call that cost \$1.56 must have $1.56 = 0.12m \ \Rightarrow\ m = 13$ minutes.

(d) From the model, we see that $a = 0.12$. Therefore, calls cost \$0.12 per minute.

Example 6.15

The volume of a sphere varies directly with the cube of the radius of the sphere. You are given that a sphere with a radius of 14 cm has a volume of 11 500 cm^3 (3 s.f.).

(a) Find a direct variation model.

(b) Use your model to find the volume of a sphere with a radius of 7 cm, to 3 significant figures.

(c) Using 6 significant figures, find the percentage error between your model and the volume given by the formula $V = \frac{4}{3}\pi r^3$ for a sphere with radius 7 cm, to three significant figures.

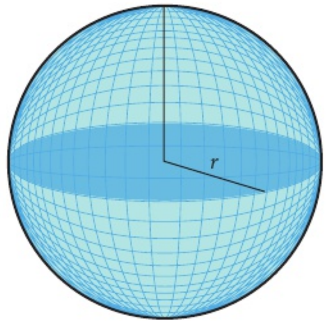


Figure 6.27 Diagram for Example 6.15

Solution

(a) Since the volume varies directly with the cube of the radius, we have

$$V = ar^3$$

Given that a sphere with radius 14 cm has a volume of 11 500 cm^3, we can solve for the constant of variation:

$$11\,500 = a(14)^3 \;\Rightarrow\; a = 4.19 \text{ (3 s.f.)}$$

Therefore, our model is $V = 4.19r^3$

(b) For a sphere of radius 7 cm, we have

$$V = 4.19(7)^3 = 1437.17 \approx 1440 \text{ cm}^3 \text{ (3 s.f.)}$$

(c) The theoretical volume of a 7 cm sphere is $V = \frac{4}{3}\pi(7)^3 = 1436.76 \text{ cm}^3$

There percentage error is therefore $\frac{|1436.76 - 1437.17|}{1436.76} \times 100 = 0.0285\%$

Inverse variation

Suppose you want to hire a DJ for a party. The cost of the DJ is \$500 for the night. To cover the cost, you plan to sell tickets. What price should you set for the tickets, based on how many you think you will sell? This is a case of inverse variation, because the more tickets you sell, the lower the price per ticket can be in order to cover the cost of hiring the DJ. Consider a few data points: if you only sell one ticket, the ticket price will need to be \$500 to cover costs (but it would be a lonely party!). If you can sell two tickets, then each one can be \$250 (party for two!). If you can sell 10 tickets, then each one is $\frac{500}{10} = \$50$

For 20 tickets, each one is $\frac{500}{20} = \$25$. Ah-ha! We have already developed the model without even thinking (too hard) about it: to cover costs, the ticket price is the total cost divided by the number of tickets you can sell. Symbolically:

$$P = \frac{500}{n}$$

where P is the ticket price and n is the number of tickets you must sell.

In Figure 6.28, we have drawn the graph as a series of dots but have not connected them. Why not? It's not possible to sell 1.5 tickets – we can only sell whole tickets. So, we plot the graph as a series of points rather than a smooth function. However, for convenience, we often draw the graph as a smooth function even though that may not strictly represent reality.

This produces an interesting graph, shown in Figure 6.28. Price per ticket varies inversely with the number of tickets sold.

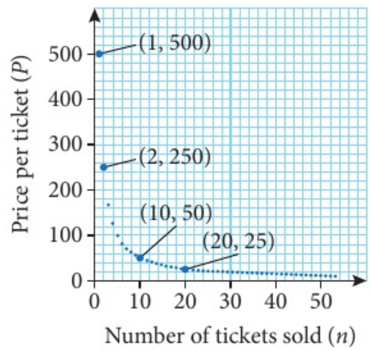


Figure 6.28 Price per ticket varies inversely with the number of tickets sold

Notice that in Figure 6.28 the graph approaches – but does not intersect – the horizontal axis. This makes sense, because as we sell more and more tickets, we could reduce the price per ticket. But, if we give away the tickets for free, we can't possibly cover the cost of the DJ.

In this example, we can rewrite our model using a negative exponent:

$$P = \frac{500}{n} \Rightarrow P = 500n^{-1}$$

Inverse variation models are of the form $y = ax^n$ where $n \in \mathbb{Z}, n \leqslant 0$

The constant a is sometimes called the **constant of variation**.

This leads us to the definition of inverse variation models shown in the key fact box.

Notice that the only difference between direct and inverse variation models is the sign of the exponent.

While it is quite common to state a weight in kg, it is technically incorrect. Kilogram (kg) is a unit of mass, not of weight. Weight is a force, which can be measured in Newtons (N) or kilogram-force (kgf); 1 kgf is equal to 9.81 N. However, kgf is a non-standard unit.

Example 6.16

The weight of an object varies inversely with the square of the distance from the centre of the Earth. At sea level (6370 km from the centre of the Earth), an astronaut weighs 600 N.

(a) Find a model for the weight of the astronaut based on the distance from the centre of the Earth.

(b) The International Space Station orbits at an altitude of 408 km. Find the weight of the astronaut while on board the International Space Station.

(a) Find the required distance from the centre of the Earth for the astronaut to weigh half what she weighs at sea level.

Solution

The answer to (b) might seem counter-intuitive: if the astronaut still weighs 530 N on the International Space Station, why do we always see videos of astronauts floating? The answer is that the space station and the astronauts inside it are actually falling towards the Earth at the same rate, so the astronauts seem to be 'floating' inside the space station.

(a) Since the weight varies inversely with the square of the distance, we can write $w = \frac{a}{d^2}$ where w is the weight in kg, d is the distance from the centre of the Earth, and a is the constant of variation that we will determine. We find a by substituting known values:

$$w = \frac{a}{d^2} \Rightarrow 600 = \frac{a}{6370^2} \Rightarrow a = 2.43 \times 10^{10}$$

Therefore our model is $w = \dfrac{2.43 \times 10^{10}}{d^2}$

(b) At an altitude of 408 km, the distance from the centre of the Earth is $6370 + 408 = 6778$ km

Therefore the weight of the astronaut would be $w = \dfrac{2.43 \times 10^{10}}{6778^2} = 530$ N

(c) We need to solve

$$300 = \frac{2.43 \times 10^{10}}{d^2} \Rightarrow 300d^2 = 2.43 \times 10^{10}$$

$$\Rightarrow d = \sqrt{\frac{2.43 \times 10^{10}}{300}} = 9010\,\text{km (3 s.f.)}$$

Exercise 6.5

1. Choose always/sometimes/never:

(a) Direct variation models always/sometimes/never pass through the origin (0, 0).

(b) Direct variation models are always/sometimes/never a type of polynomial model.

(c) Direct variation models are always/sometimes/never a type of linear model.

(d) Direct variation models are always/sometimes/never a type of exponential model.

2. Choose always/sometimes/never:

(a) Inverse variation models always/sometimes/never pass through the origin (0, 0).

(b) Inverse variation models are always/sometimes/never a type of polynomial model.

(c) Inverse variation models are always/sometimes/never a type of linear model.

(d) Inverse variation models are always/sometimes/never a type of exponential model.

3. Given that y varies directly with x, and $y = 462$ when $x = 11$, find:

(a) y when $x = 5$ **(b)** x when $y = 672$

4. Given that y varies directly with the square of x, and $y = 10$ when $x = 5$, find:

(a) y when $x = 20$ **(b)** x when $y = 40$

5. Given that y varies directly with the cube of x, and $y = 250$ when $x = 5$, find:

(a) y when $x = 8$ **(b)** x when $y = 128$

6. Given that y varies inversely as x, and $y = 10$ when $x = 5$, find:

(a) y when $x = 20$ **(b)** x when $y = 0.5$

7. Given that y varies inversely as the square of x, and $y = 10$ when $x = 5$, find:

(a) y when $x = 20$ **(b)** x when $y = 2.5$

8. Given that y varies inversely as the cube of x, and $y = 54$ when $x = 5$, find:

(a) y when $x = 15$ **(b)** x when $y = 250$

9. The height of an image produced by a projector varies directly with the distance from the screen. The image is 1.5 metres tall when the projector is 2 metres from the screen.

(a) Find a direct variation model for the size of the image S given the distance d from the screen.

(b) Hence predict the size of the image for a projector 7 metres from the screen.

(c) Use your model to find the distance required to project a 10-metre image.

10. The velocity of a falling object varies directly with the amount of time it has been falling.

(a) Given that an object that has been falling for two seconds has a velocity of 19.6 m s^{-1}, calculate the constant of variation.

(b) Calculate the velocity of an object that has been falling for 4 seconds.

(c) Terminal velocity for a skydiver is approximately 200 km h^{-1}. Calculate the approximate number of seconds it takes a skydiver to reach terminal velocity according to this model.

11. The position of a falling object varies directly with the square of the number of seconds it has been falling. We assume the initial velocity is zero.

(a) Given that an object that has been falling for 10 seconds towards the surface of the Moon travels 162 m, find the constant of variation.

(b) Calculate the distance an object falls in 5 seconds.

(c) Calculate the time required for an object to fall 200 m.

12. The radius of a satellite's orbit around the Earth varies inversely with the square of the velocity of the satellite.

(a) Given that a satellite that travels at a velocity of 7700 m s^{-1} has an orbital radius of 6.75×10^6 m, calculate the constant of variation.

(b) Find the orbital velocity of a satellite with an orbital radius of 7.0×10^6 m.

(c) Find the orbital radius of a satellite with a velocity of 8000 m s^{-1}.

A regular dodecahedron has 12 pentagonal faces.

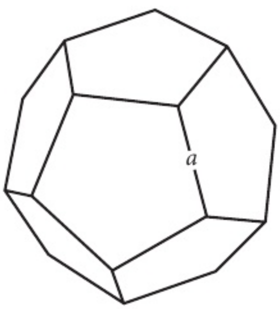


Figure 6.29 Dodecahedron

13. The volume of a regular dodecahedron varies directly with the cube of the length of an edge a.

(a) Give that the volume of a dodecahedron is 958 cm^3 with the edge length $a = 5$, find the constant of variation.

(b) Find the volume of a dodecahedron when the edge length is 8 cm.

(c) Find the edge length of a dodecahedron with volume 100 m^3.

14. The power a wind turbine can generate varies directly with the cube of the wind speed. A certain turbine can generate 314 W when the wind speed is 8 m s^{-1}.
    (a) Find the constant of variation.
    (b) Find the power generated when the wind speed is 12 m s^{-1}.
    (c) Find the wind speed necessary to generate 2000 W.

6.6 Trigonometric models

Trigonometric models are well-suited to describing phenomena that repeat themselves: tides, seasonal temperatures, motion of wheels, etc. In this section, we will examine the models $y = a\sin(bx) + d$ and $y = a\cos(bx) + d$

Exploration

To be able to use the sine and cosine functions to model effectively, you need to understand how the parameters a, b, and d affect the graphs of the functions and the difference between the graphs of sine and cosine. This is best done by using a graphing application such as Geogebra and experimenting with different values of these parameters. Your goal is to observe carefully the effect each parameter has on the shape of the graph. Follow the suggestions below to help you to understand what a, b, and d do. Make sure you work in **degree** mode.

1. What does *a* do?

Use your graphing program to generate a graph of $y = a\sin x$. Experiment with different values of a and notice the effect on the graph. Make sure to set a to at least the following four values: $a = 1,\ \frac{1}{2},\ 2,\ -2$

You should notice that $|a|$ is equal to the **amplitude** of the graph, which is half of the **wave height**, as shown in Figure 6.30. Also, if a is negative, then the graph is reflected across the x axis.

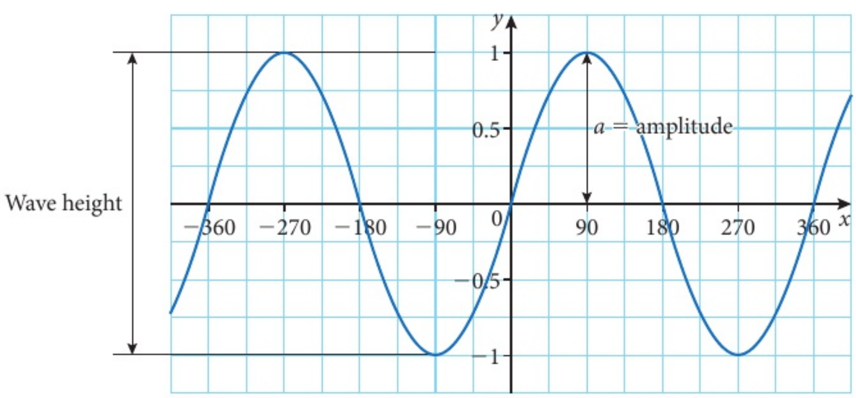


Figure 6.30 Amplitude is half of wave height

Value of b	Length of period (p)
1	360°
2	180°
$\frac{1}{2}$	720°
−2	180°

Table 6.10 Values of b and length of period

2. What does b do?

Use your graphing program to generate a graph of $y = \sin(bx)$. Experiment with different values of b and notice the effect on the graph. Make sure to set b to at least the following four values: $b = 1, \frac{1}{2}, 2, -2$

You should notice that b controls the **wavelength** of the graph. In mathematics and science, we refer to this as the **period** of the function. The period is the distance it takes for the graph to repeat itself – the distance between two adjacent local maxima or local minima, for the sine and cosine functions.

But, how exactly does b affect the period? If we write down the values of b and the length of the period, we get Table 6.10. What can we observe? First we notice that the sign of b does not seem to make a difference to the period since the period is the same for both 2 and −2. Next, we see that as the absolute value of b increases, the period gets shorter. This suggests that the length of the period and the value of b vary inversely. We can also see from the table that the product of $|b|$ and the period is a constant 360°. Therefore, we conclude that $|b|p = 360° \Rightarrow p = \frac{360°}{|b|}$. This is shown in Figure 6.31.

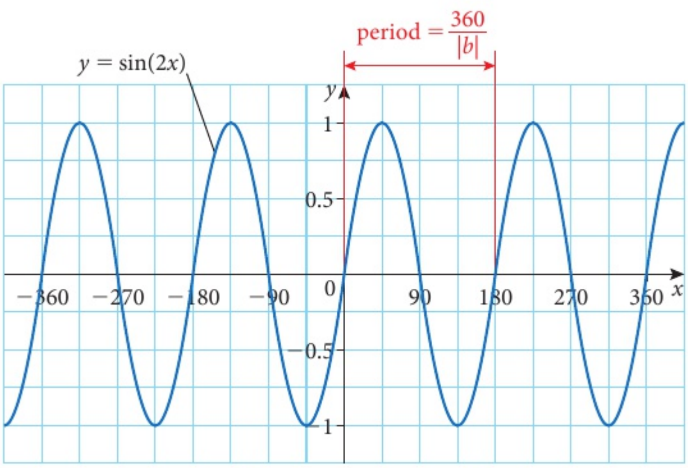


Figure 6.31 Relationship between period and b

Also, notice that if b is negative, then the graph is reflected across the y axis.

3. What does d do?

Use your graphing program to generate a graph of $y = \sin(x) + d$. Experiment with different values of d and notice the effect on the graph. Make sure to set d to at least the following four values: $d = 0, \frac{1}{2}, 2, -2$

You should notice that d controls the vertical position of the graph. Specifically, the value of d gives us the position of the **principal axis** of the function. The principal axis is a horizontal line with equation $y = d$ (Figure 6.32).

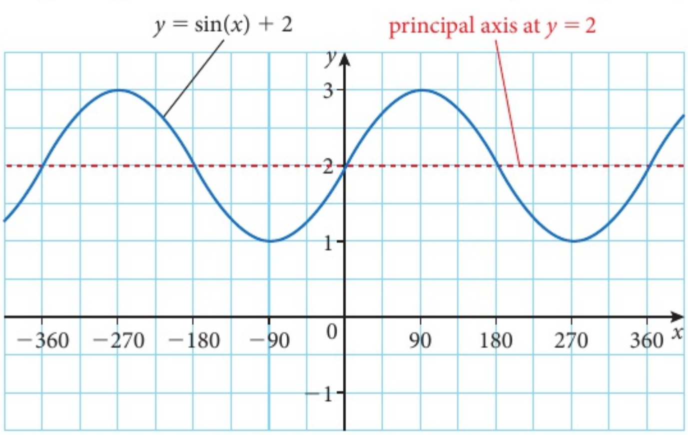


Figure 6.32 The principal axis of a periodic function

What is the difference between the graphs of sine and cosine?

Notice that in the questions above, we have only used sine. This is because the graphs of sine and cosine are very similar – the properties of a, b, and d are the same for both. However, there is an important difference between the graphs of sine and cosine. Graph the functions $y = \cos x$ and $y = \sin x$ and look carefully at the graphs.

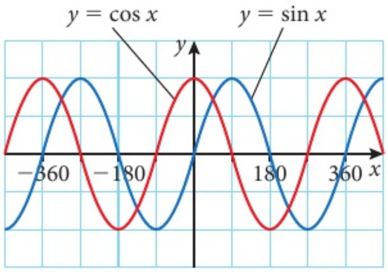


Figure 6.33 Comparison of sine and cosine graphs

Look at the behaviour of both functions when $x = 0$, as shown in Figure 6.33. Cosine has a maximum at $x = 0$ and then decreases. For sine, when $x = 0$ the graph increases from where the principal axis meets $x = 0$. Is that always true? It's not – remember how the parameters a, b, and d affect the graphs. For the cosine function, because the value of a can cause the graph to be reflected across the principal axis, when $a < 0$ cosine will have a minimum at $x = 0$ and then increase. For the sine function, when $a < 0$ the curve will decrease from where the principal axis meets $x = 0$. These graphs are shown in Figure 6.34.

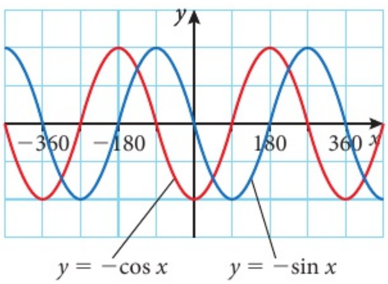


Figure 6.34 When $a < 0$, the graphs of sine and cosine are reflected across the x axis. This causes a change in behaviour around $x = 0$ that is important for modelling

The key findings of the exploration are summarised here.

For the trigonometric models $y = a\sin(bx) + d$ and $y = a\cos(bx) + d$

- Amplitude $= |a|$
  The amplitude is equal to half the wave height.
- For $a < 0$, the graph is reflected across the x axis.
- Period $= \dfrac{360°}{|b|}$
- For $b < 0$, the graph is reflected across the y axis.
- The principal axis is at $y = d$
- Use a sine model when you want the initial value of the model to be equal to the location of the principal axis (half-way between the maximum and minimum values). If the model should initially increase, let $a > 0$. If the model should initially decrease, let $a < 0$.
- Use a cosine model when you want the initial value of the model to be a maximum (let $a > 0$) or minimum (let $a < 0$).

Developing trigonometric models

Now that we have an understanding of the parameters of trigonometric models, we can model some real-life periodic phenomena.

Example 6.17

A reflector is attached to a bicycle wheel at a point about 17 cm from the centre of the wheel. The radius of the wheel is approximately 30 cm.
The wheel rotates clockwise and takes 4 seconds to complete one revolution.

(a) Assuming that the reflector is at the top-most position at time $t = 0$ develop a model for the height h of the reflector above the ground at any time t in seconds.

(b) Revise your model so that the reflector is at the bottom-most position at time $t = 0$

(c) Revise your model so that the reflector is at the right-most position at time $t = 0$

(d) Revise your model for a point on the outer edge of the wheel. Assume that the point is at the left-most position at time $t = 0$

Solution

(a) Since the radius of the wheel is 30 cm, and the reflector is 17 cm from the centre of the wheel, we know that the minimum height of the reflector is $30 - 17 = 13$ cm. At the top-most position, the reflector will be $30 + 17 = 47$ cm from the ground. Since the amplitude is half of the wave height, it must be $\frac{47 - 13}{2} = 17$ cm, so $a = 17$ or $a = -17$.

That makes sense since the reflector is located 17 cm from the centre of the wheel! Also, the principal axis must be halfway between the minimum and maximum values, so it must be at $\frac{47 + 13}{2} = 30$ cm, hence $d = 30$. That makes sense because the centre of the wheel is 30 cm above the ground. Since it takes 4 seconds for the wheel to complete one revolution, the period of our model must equal 4. We can then find $|b|$:

$$\text{Period} = \frac{360}{|b|} \quad \Rightarrow \quad 4 = \frac{360}{|b|} \quad \Rightarrow |b| = \frac{360}{4} = 90$$

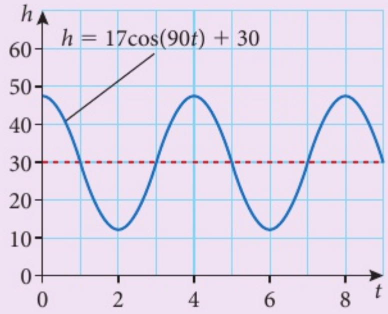
So $b = 90$ or $b = -90$

Finally, we need to decide if a sine or cosine model is better suited for this situation. Since we are told that the reflector begins at the top-most point, our model must begin at a maximum. Therefore, we choose the cosine function, and we keep both a and b positive since no reflection across the x or y axis is needed. Thus, we have $a = 17$, $b = 90$, and $d = 30$, and our model is

$$h = 17\cos(90t) + 30$$

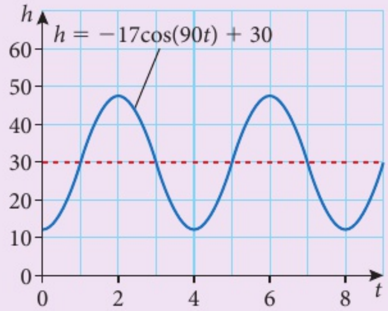
We can check by graphing two complete periods ($0 \leqslant t \leqslant 8$)

The graph appears to fit the behaviour we expect.



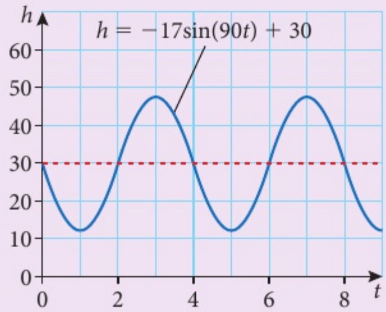
(b) For the reflector to start at the bottom-most position, the model must have a minimum as an initial value. In this case, we simply let a be negative, and revise our model accordingly:

$h = -17\cos(90t) + 30$
as shown in the graph.



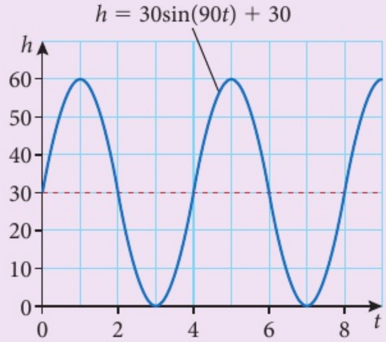
(c) For the reflector to start at the right-most position, the model must have an initial value equal to the location of the principal axis. Therefore, we must choose a sine model. Also, we are told that the wheel rotates clockwise, which will cause the height of the reflector to initially decrease. Therefore, we must let $a < 0$ and so our model is:

$h = -17\sin(90t) + 30$
as shown in the graph.



(d) A point on the outer edge of the wheel will have a similar model, but the amplitude of the model is now equal to the radius of the wheel. Also, since the point is in the left-most position at time $t = 0$, the model must increase initially. Therefore, we use a sine model with $a = +30$:

$h = 30\sin(90t) + 30$
as shown in the graph.



Example 6.18

The height of a person on Vienna's famous Riesenrad ferris wheel can be modelled by the function

$$h(t) = a\cos(36(t - 5)) + d$$

where $h(t)$ is the height in feet at time t minutes.

(a) Given that the diameter of the Riesenrad is 200 feet, deduce the value of a.

(b) The boarding platform at the lowest point of the Riesenrad is 12 feet above the ground. Find the value of d.

(c) Calculate the number of minutes until a person riding the Riesenrad reaches the top of the wheel.

(d) Calculate the time required for one complete rotation.

(e) Determine a reasonable domain and range for this model if each ride is exactly one complete revolution.

(f) In one 10-minute ride, during what interval of time will a person on the Riesenrad be at least 100 feet above the ground?

(g) Comment on the reasonableness of this model.

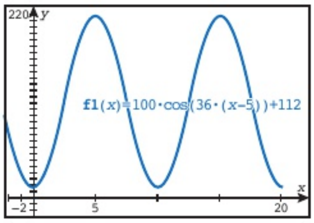


Figure 6.35 GDC screen for solution to Example 6.18 (c)

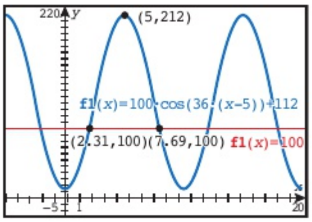


Figure 6.36 GDC screen for solution to Example 6.18 (f)

We do not consider the next period, since the question asked only for the first ten minutes.

Solution

(a) Since the diameter of the Riesenrad is 200 feet, the wave height of the model must be 200, so the amplitude $a = 100$

(b) Since the minimum value of the function is 12 feet, the principal axis must be located at $12 + 100 = 112$. Therefore $d = 112$

(c) By graphing the model, we obtain the GDC screen shown in Figure 6.35.

We can see that a local maximum is at $t = 5$ minutes.

(d) Since the first maximum is reached 5 minutes after the first minimum (when the person boards the Riesenrad), we can conclude that one complete rotation takes 10 minutes. This agrees with the graph.

(e) Since one complete revolution takes 10 minutes, the domain for this model should be $0 \leqslant t \leqslant 10$. In that time, the person will travel from 12 feet to 212 feet, so the range is $12 \leqslant h(t) \leqslant 212$

(f) Again, we can use our GDC to solve this, by adding a second function with the constant value of 100, as shown in Figure 6.36.

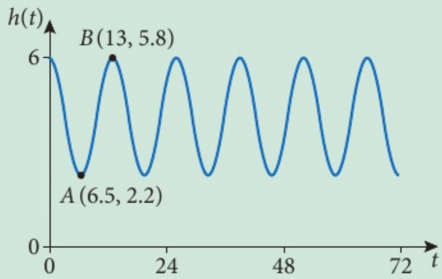
Then, by using the Intersection command, we can find the times when a person goes above and below 100 feet.

We see that a person would be at least 100 feet above the ground for $2.31 \leqslant t \leqslant 7.69$ minutes.

(g) Although this model is reasonable in principal, it assumes that the rotation speed of the wheel is constant. In reality, the wheel would need to stop at regular intervals to allow passengers to get on and off the ride.

Exercise 6.6

1. At Cabot Cove the height of the water in metres is read from a depth gauge in the water. The depth can be modelled by the function $h(t) = p\cos(q \times t) + r$, where t is the number of hours after high tide at 15:00 on 10 March 2018. The diagram shows the graph of h, for $0 \leqslant h \leqslant 72$



The point $A(6.5, 2.2)$ represents the first low tide and $B(13, 5.8)$ represents the next high tide.

   (a) How much time is there between the first low tide and the next high tide?

   (b) Find the difference in height between low tide and high tide.

   (c) Find the value of p, q, and r.

   (d) There are two high tides on 11 March 2018. At what time does the second high tide occur?

   (e) Calculate the number of minutes where the height of the water is at least 5.5 metres on 11 March 2018.

2. The depth of water at a dock can be modelled by the function $d(t) = a\cos(bt) + 6.2$, for $0 \leqslant t \leqslant 18$, where $d(t)$ is the depth in metres t hours after high tide. At high tide, the depth is 9.0 m. Low tide is 6.25 hours later, at which time the depth is 3.4 m.

   (a) Find the value of a and the value of b.

   (b) Use the model to find the depth of the water 9 hours after high tide.

   (c) A certain sailboat needs water at least 4.5 m deep. How many hours after high tide can it stay at the dock before it must leave?

3. The height of a seat on a Ferris wheel can be modelled by the function $h(t) = -22\cos(40t) + 23$, for $t \geq 0$, where $h(t)$ is the height in metres after t minutes.

   (a) Find the height of the seat when $t = 0$

   (b) The seat first reaches a height of 35 m after k minutes. Find k

   (c) Calculate the time needed for the seat to complete a full rotation.

4. The London Eye, a large Ferris wheel, has a diameter of 120 metres. One revolution of the wheel takes 30 minutes. Let A be a point at the bottom of the wheel, at ground level.

   (a) Write down the height of A above ground level after:
   (i) 15 minutes **(ii)** 20 minutes.

   (b) The function $h(t) = a\cos(bt) + c$ gives the height in metres of point A after t minutes. Find a , b, and c.

   (c) Find the number of minutes that A is more than 100 metres above the ground during one rotation.

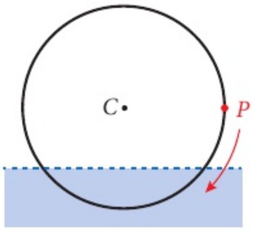


Figure 6.37 Diagram for question 5

5. A water-wheel rotates clockwise. Point P is at the right-most position on the wheel.

The diameter of the wheel is 16 metres. The centre of the wheel, C, is 5 metres above the water level. The height of point P above the water level is modelled by $h(t) = a\sin(bt) + c$

(a) Find the value of a

(b) It takes 90 seconds for the wheel to rotate once. Find the value of b

(c) Calculate the number of seconds point P is underwater during one rotation.

6. The number of hours of daylight in San Diego, California, can be modelled by the function $d(t) = a\sin(bt) + c$, where $d(t)$ is the number of hours of daylight t days after 21 March. The longest day has 14.4 hours of daylight and the shortest day has 9.6 hours of daylight. On 21 March there are 12 hours of daylight. Assume 365 days in one year.

(a) Find the value of a and of c

(b) Assuming that the cycle of day lengths repeats exactly every year, find the value of b in the form $\frac{p}{q}$, where $p, q \in \mathbb{Z}$

(c) Find how many days in a year have more than 14 hours of daylight.

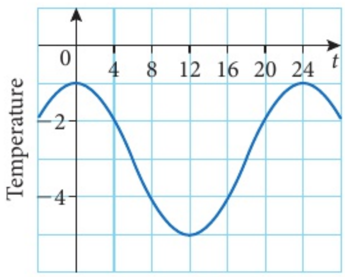


Figure 6.38 Diagram for question 7

7. The temperature in degrees Celsius during a 24-hour period is shown on the graph in Figure 6.38 and is given by the function $f(t) = a\cos(bt) + c$, where a, b, and c are constants, t is the time in hours, and (bt) is measured in degrees.

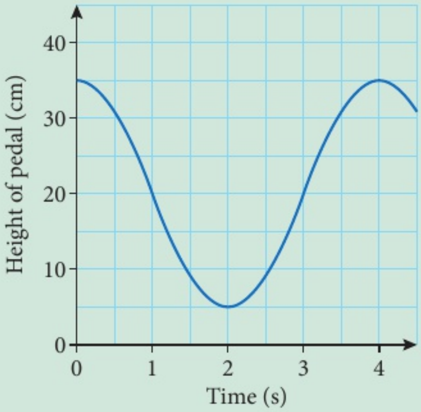
(a) Write down the value of a.

(b) Find the value of b.

(c) Write down the value of c.

(d) Write down the interval of time during which the temperature is increasing from -4°C to -2°C.

8. The height, $h(t)$, in centimetres, of a bicycle pedal above the ground at time t seconds is a cosine function of the form $h(t) = A\cos(bt) + C$, where (bt) is measured in degrees. The graph of this function for $0 \leqslant t \leqslant 4.3$ is shown here.



(a) Write down the maximum height of the pedal above the ground.

(b) Write down the minimum height of the pedal above the ground.

(c) Find the amplitude of the function.

(d) Hence, or otherwise, find the value of A and of C.

(e) Write down the period of the function $h(t)$, including units.

(f) Hence find the value of b.

(g) Calculate the first value of t for which the height of the pedal above the ground is 30 cm.

(h) Calculate the number of times the pedal rotates in one minute.

9. The depth, in metres, of water in a harbour is given by the function $d = 4\sin(0.5t) + 7$, where t is in minutes, $0 \leqslant t \leqslant 1440$

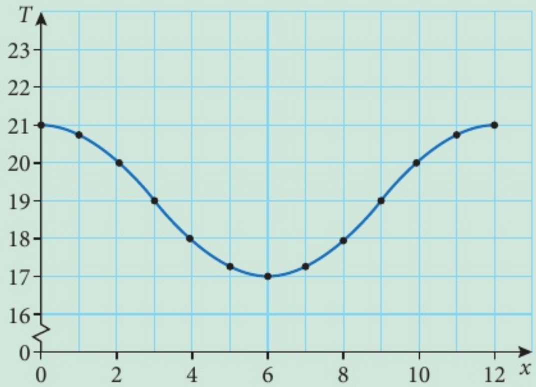
(a) Write down the amplitude of d.

(b) Find the maximum value of d.

(c) Find the period of d. Give your answer in hours.

(d) On Tuesday, the minimum value of d occurs at 14:00. Find when the next maximum value of d occurs.

10. The graph represents the temperature ($T°$ Celsius) in Washington measured at midday during a period of thirteen consecutive days starting at Day 0. These points also lie on the graph of the function $T(x) = a + b\cos(cx)$, $0 \leqslant x \leqslant 12$, where a, b, $c \in \mathbb{R}$

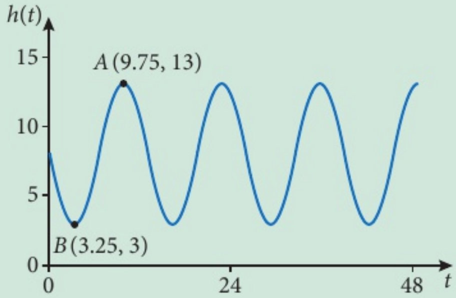


(a) Find the value of a and of b.

(b) Find the value of c.

(c) Using the graph, or otherwise, write down the part of the domain for which the midday temperature is less than 18.5°C.

11. At Kabulonga Beach the height of the water in metres is modelled by the function $h(t) = p\sin(q \times t) + r$, where t is the number of hours after 8:00 hours on 19 September 2018. The diagram shows the graph of h, for $0 \leqslant h \leqslant 48$



The point $A(9.75, 13)$ represents the first high tide and $B(3.25, 3)$ represents the first low tide.

(a) How much time is there between the first low tide and the first high tide?

(b) Find the difference in height between low tide and high tide.

(c) Give a reason why p must be negative.

(d) Find the values of p, q, and r.

(e) There are two high tides on 20 September 2018. At what time does the second high tide occur?

6.7 Modelling skills

In the real world, it's critical to be able to identify, develop, evaluate, and use models responsibly. In the preceding sections of this chapter we have examined different models and touched on many of the uses and limitations of models. However, it's also important to be able to choose an appropriate model.

One of the authors recorded the mass of his young baby at a several points after she was born. The data is given in Table 6.11.

Days since birth	Mass (kg)
0	3.36
4	3.4
17	3.69
45	5.29
73	6.15

Table 6.11 Baby mass recordings

A graph of this data is shown in Figure 6.39.

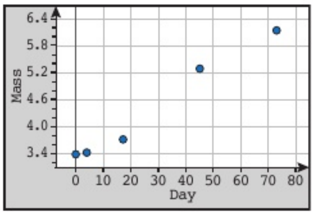


Figure 6.39 Mass versus days since birth for a young child

The data appears approximately linear, so the author used technology to generate a linear model for this data:

$m = 0.0408d + 3.24$, where m is mass in kg, d days after birth.

The model suggests that she is gaining 0.0408 kg per day, with an initial mass of 3.24 kg. Graphically, the model appears to fit the data reasonably well, as shown in Figure 6.40.

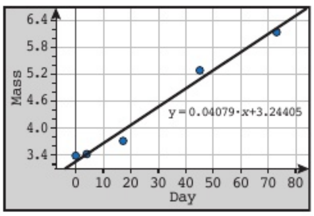


Figure 6.40 Linear model for mass in terms of days after birth

Of course, the author was very interested in the eventual mass of his young baby girl. How much would she weigh she was one year old? Five years old? Ten years old? He used the model to make some predictions, shown in Table 6.12 (he even added leap days to be sure).

Days since birth	365 (1 year)	1826 (5 years)	3652 (10 years)
Mass (kg)	18.1	77.7	152

Table 6.12 Mass projections

Are these results reasonable? A mass of 77.7 kg for a five-year-old girl is alarming, and a mass of 152 kg for a ten-year-old girl is definitely a cause for concern!

What went wrong? The mathematical mechanics are perfectly correct.

There are two flaws in our reasoning. One, our **model choice** (linear model) assumes that the rate of change is constant (in this case, 40.8 g per day), which is not correct; we know that the growth rate will decrease as the child gets older. Two, we **extrapolated** beyond known data. We will discuss both of these issues in this section.

Choosing a model

One of the most challenging parts of modelling in the real world is choosing an appropriate model. Often, we try to use our understanding of a situation when we are choosing and developing a model. Other times we may look at the shape of the data (as seen on a graph) to try to help us choose a model.

Here are some guidelines to help, with examples from the introduction to this chapter:

- Consider the rate of change.
  - Is it constant? Try a linear model.

    Example: On a long flight, the airspeed of a plane is constant, so the distance remaining to the destination can be described by a **linear** model.
  - Is the quantity increasing/decreasing by a fixed percentage or ratio? Try an exponential model.

    Examples: Algae in a polluted lake doubles every 3 days; a bank account earns 5% quarterly interest, compounded annually; the value of a car is decreasing by 25% per year; what will the activity of a radioisotope be after a given time.
  - Is the quantity increasing/decreasing at a linearly increasing/decreasing rate? Try a quadratic model.

    Examples: The price to manufacture x units of some product decreases linearly, the revenue from selling x units can be described by a quadratic model; the velocity of a falling object changes linearly, the position follows a quadratic model.

- Consider the nature of the phenomenon.
  - Does it relate volume to a linear quantity? Try a cubic model.

    Examples: The volume of a balloon relative to its diameter can be described by a cubic model; electricity generated by a wind turbine based on wind speed.
  - Is it cyclical/periodic/repeating? Try a trigonometric model.

    Examples: tide height; person on Ferris wheel; seasonal average temperatures.
  - Does one variable appear to be in constant ratio to a positive or negative power of the other variable? Consider a direct or inverse variation model.

    Examples: A DJ charges a fixed amount to provide music for a party, the cost per person can be described by an inverse variation model; the distance a dropped object has travelled varies with the square of elapsed time, use a direct variation model.

- Consider the shape of the data.

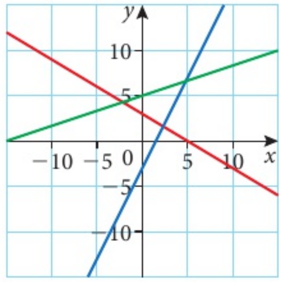


Figure 6.41 Linear models

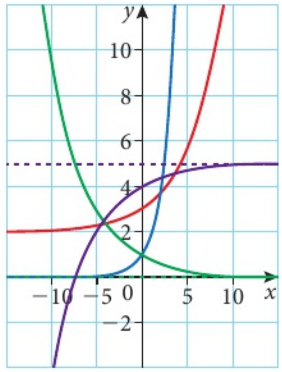


Figure 6.42 Exponential models

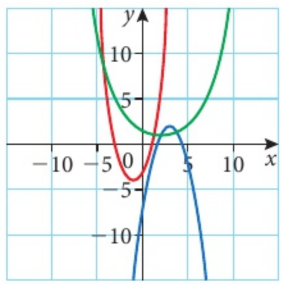


Figure 6.43 Quadratic models

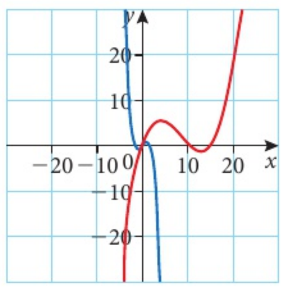


Figure 6.44 Cubic models

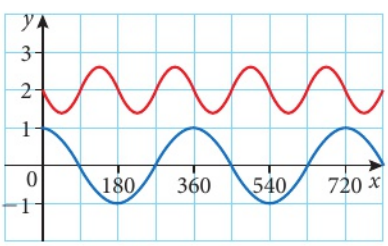


Figure 6.45 Trigonometric models

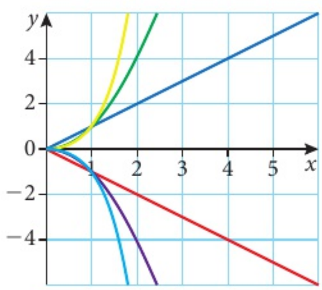


Figure 6.46 Direct variation models

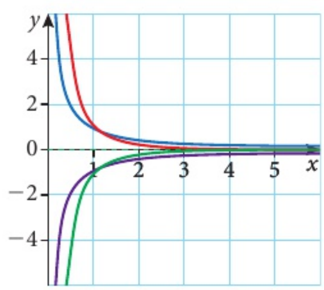


Figure 6.47 Inverse variation models

Example 6.19

For each of the following examples, choose an appropriate model and state a reason why you chose it.

(a) Modelling a person's mass as a function of their height.

(b) The number of hours of daylight in Tokyo each day of the year.

(c) The population of bacteria in a Petri dish over time.

(d) The resale value of a computer over time.

(e) The distance from Earth for a Voyager space probe travelling away at a constant rate.

(f) The height of the golf ball Alan Shepard hit on the moon in 1971.

(g) The number of payments required to repay a loan as a function of the amount of the monthly payment.

Solution

(a) Since weight is based on volume, and volume varies with the cube of height, we should use a cubic direct variation model of the form $W = kh^3$

(b) Number of hours of daylight per day is a periodic phenomenon, so a trigonometric model is appropriate.

(c) The population of bacteria in a Petri dish will grow exponentially (until the maximum population for the dish is reached), so an exponential model is appropriate.

(d) A computer's value will depreciate exponentially, so an exponential model is appropriate.

(e) Since the Voyager is travelling away from Earth at a constant rate, a linear model is appropriate.

(f) The golf ball will follow a quadratic falling-object model as it is still subject to gravity (just less than on Earth).

(g) If we increase the amount of the monthly payment, the number of payments we have to make decreases, and vice-versa. Since the total repayment amount (the value of the loan) is the product of the number of payments and the payment amount, this will be an inverse variation model of the form $V = an \Rightarrow n = \frac{V}{a}$ where V is the value of the loan (a constant), a is the amount of each payment, and n is the number of payments.

Testing a model

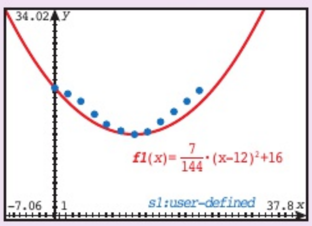
Once a model is developed (using the techniques in this chapter), it is important to test it. In fact, in real life, it is often the case that after testing a model you decide to develop an alternative model or revise your existing model. How do we test a model? We can test a model by looking at the fit of the model for the known data. If the model does not seem to fit well, it can be a sign that we do not fully understand the phenomenon we are trying to model. In that case, we may decide to try a different model, or perhaps more research is necessary!

Developing useful models is the work of many people: economists, doctors, aid workers, actuaries, and many other researchers. This chapter is a basic introduction to the art and science of modelling.

It is important to understand that, in the real world, for many topics, there is often no 'right' model. Instead, for real-life applications we often judge a model by its usefulness. What does it tell us about the data we have? Does it tell us something about the nature of the phenomena we are modelling? Does it allow us to make meaningful predictions? Are the predictions surprising but plausible? What assumptions are we making? Asking these sorts of questions is a crucial skill for the practicing mathematical modeller. In the context of this course, we can only simulate the sort of scenarios you might encounter in the real world in the hope that you might be critical and thoughtful in your future encounters with mathematical models.

Example 6.20

Anna is attempting to model the temperature of a pond at her school. She has collected 12 points of data, measuring the temperature of the pond every 2 hours. Based on her data, she believes a quadratic model may be useful so she has used the vertex and y-intercept to fit a model. The graph shows her data points and model.



(a) Give two reasons why Anna's choice of model may not be appropriate.

(b) Suggest a better model and give a reason why it would be more appropriate.

Solution

(a) Anna's model doesn't seem to fit the data very well. Also, the temperature of the pond probably varies during the day and night in a periodic way, which would not suggest a quadratic model.

(b) Since the temperature of the pond probably varies periodically, a trigonometric model would be more appropriate.

Interpolation versus extrapolation

As in the example at the start of the section, we can quickly get into trouble when we extrapolate. But what exactly is extrapolation? How can we tell when we are using a model appropriately and when we are extrapolating? One way to check is to look at the range of known data, as in Example 6.21.

Example 6.21

A used-car dealer has recently sold five of the same kind of car, called the Canyonero. All five were in similar condition, but the ages of the cars varied. The data are shown in the table.

Age of car (years)	Sale price ($)
2	20 000
3	14 500
4	12 000
6	6100
10	2000

The dealer decides to use the model $P = 20\,000(0.75)^{t-2}$ to describe the price of the Canyonero (P represents the price and t represents the age in years).

(a) Use the model to predict the price of a new Canyonero. Comment on the usefulness of your prediction.

(b) Use the model to predict the price of an 8-year-old Canyonero. Comment on the usefulness of your prediction.

(c) Use the model to predict the price of a 15-year-old Canyonero. Comment on the usefulness of your prediction.

Solution

(a) According to the model, the price of a new Canyonero would be $P = 20\,000\,(0.75)^{0-2} = \$35\,600$ (3 s.f.). Although this may seem like a reasonable prediction, our data only includes cars from 2 to 10 years old, so we cannot be confident in this prediction. Therefore, it is not a useful prediction.

(b) According to the model, the price of an 8-year-old Canyonero would be $P = 20\,000\,(0.75)^{8-2} = \$3\,560$ (3 s.f.). Since 8 years is within the range of the age of cars in our data, we can be reasonably confident in this prediction. Therefore, this is a useful prediction.

(c) According to the model, the price of a 15-year-old Canyonero would be $P = 20\,000\,(0.75)^{15-2} = \$475$ (3 s.f.). Not only does this value seem like a small amount, but it is not within the range of the age of cars in our data. Therefore, it is not a useful prediction.

As you can see in Example 6.21, it is important to note that we decide whether a prediction is interpolation or extrapolation based on the range of our independent variable. This is easier to see when we look at a graph.

As shown in Figure 6.48, any predictions based on values within the range of the independent variable are called **interpolations**. Therefore, any predictions for cars between 2 and 10 years old would be interpolations and therefore relatively safe and useful.

On the other hand, any predictions based on values outside the range of the independent variable are called **extrapolations**. Therefore, any predictions for cars newer than 2 years old or older than 10 years old are considered extrapolations. Extrapolated predictions are less reliable and therefore less useful, since we simply can't be sure what the value of the car might be outside the range that we have observed.

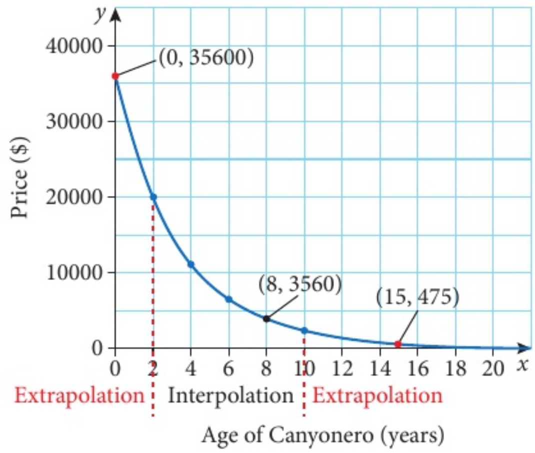


Figure 6.48 Interpolation is based on the range of the independent variable

Although it is tempting to look at the value of a prediction to decide whether a prediction is useful or not, it is critical to consider first whether the value of the independent variable is within the range of known values of the independent variable. That is, we can decide whether a value is interpolation or extrapolation before we make a prediction. Prediction based on extrapolation should always be used with caution!

Interpolations are predictions based on values within the range of known values of the independent variable. **Extrapolations** are predictions based on values outside the range of known values of the independent variable.

Given a model $f(x)$ based on data with minimum x value a and maximum x value b, $a \leqslant x \leqslant b$

- If $a \leqslant c \leqslant b$, then $f(c)$ is an **interpolated value**.
- If $c < a$ or $c > b$, then $f(c)$ is an **extrapolated value**.

Exercise 6.7

1. Identify appropriate models for the following situations.

(a) The acceleration produced by exerting a constant force on various masses.

(b) The total cost of fuel purchased at a fixed price per litre.

(c) The area of a rectangle where the length is twice as large as the width.

(d) Average monthly temperature over a year in Tokyo.

(e) The volume of a cube as a function of its side length.

(f) The value of a savings account growing at a fixed percentage rate.

(g) The height of tide at a beach.

(h) The value of a new car over time.

(i) The cost for hiring a bus, per person.

(j) The cost of a wedding as a function of the number of guests.

2. For the following data sets, identify the valid domain for interpolation. Assume we are treating the first row in the table as the independent variable.

(a)

Time in dive (s)	2	4	6
Velocity (m s^{-1})	10	53	90

(b)

x	41	33	100	76	11
y	33.4	31.8	52.3	47.3	16.6

(c)

n	8.2	3.9	5.5	1.3	8.5
$P(n)$	15.5	11.5	12.3	6.0	15.2

3. In general, what are the key feature(s) that we might see to suggest that we should use each of the following models:

(a) Linear **(b)** Quadratic **(c)** Trigonometric

(d) Exponential **(e)** Inverse

Chapter 6 practice questions

1. A city is concerned about pollution, and decides to look at the number of people using taxis. At the end of the year 2000, there were 280 taxis in the city. After n years the number of taxis, T, in the city is given by $T = 280 \times 1.12^n$

(a) Find the number of taxis in the city at the end of 2005.

(b) Find the year in which the number of taxis is double the number of taxis there were at the end of 2000.

At the end of 2000 there were 25 600 people in the city who used taxis. After n years the number of people, P, in the city who used taxis is given by $P = \dfrac{2\,560\,000}{10 + 90e^{-0.1n}}$

(c) Find the value of P at the end of 2005, giving your answer to the nearest whole number.

(d) After seven complete years, will the value of P be double its value at the end of 2000? Justify your answer.

Let R be the ratio of the number of people using taxis in the city to the number of taxis. The city will reduce the number of taxis if $R < 70$

(e) Find the value of R at the end of 2000.

(f) After how many complete years will the city first reduce the number of taxis?

2. The profit (P) in Swiss Francs made by three students selling homemade lemonade is modelled by the function $P = -\frac{1}{20}x^2 + 5x - 30$ where x is the number of glasses of lemonade sold.

   (a) Copy and complete the table.

x	0	10	20	30	40	50	60	70	80	90
P		15			90			75	50	

   (b) On graph paper draw axes for x and P, placing x on the horizontal axis and P on the vertical axis. Use suitable scales. Draw the graph of P against x by plotting the points. Label your graph.

   (c) Use your graph to find:

   (i) the maximum possible profit

   (ii) the number of glasses that need to be sold to make the maximum profit

   (iii) the number of glasses that need to be sold to make a profit of 80 Swiss Francs

   (iv) the amount of money initially invested by the three students.

   (d) The three students Baljeet, Jane and Fiona share the profits in the ratio of 1:2:3 respectively. If they sold 40 glasses of lemonade, calculate Fiona's share of the profits.

3. The function $Q(t) = 0.003t^2 - 0.625t + 25$ represents the amount of energy in a battery after t minutes of use.

   (a) Write down the amount of energy held by the battery immediately before it was used.

   (b) Calculate the amount of energy available after 20 minutes.

   (c) Given that $Q(10) = 19.05$, find the average amount of energy used per minute for the interval $10 \leqslant t \leqslant 20$

   (d) Calculate the number of minutes it takes for the energy to reach zero.

4. Jashanti is saving money to buy a car. The price of the car, in US Dollars (USD), can be modelled by the equation $P = 8500(0.95)^t$

   Jashanti's savings, in USD, can be modelled by the equation $S = 400t + 2000$

   In both equations t is the time in months since Jashanti started saving for the car.

   (a) Write down the amount of money Jashanti saves per month.

   (b) Use your graphic display calculator to find how long it will take for Jashanti to have saved enough money to buy the car.

(c) Jashanti does not want to wait too long and wants to buy the car two months after she started saving. She decides to ask her parents for the extra money that she needs. Calculate how much extra money Jashanti needs.

5. A building company has many rectangular construction sites, of varying widths, along a road. The area, A, of each site is given by the function $A(x) = x(200 - x)$ where x is the width of the site in metres and $20 \leqslant x \leqslant 180$

(a) Site S has a width of 20 m. Write down the area of S.

(b) Site T has the same area as site S, but a different width. Find the width of T.

(c) When the width of the construction site is b metres, the site has a maximum area.

(i) Write down the value of b

(ii) Write down the maximum area.

(d) The range of $A(x)$ is $m \leqslant A(x) \leqslant n$
Write down the value of m and of n.

6. Water has a lower boiling point at higher altitudes. The relationship between the boiling point of water (T) and the height above sea level (h) can be described by the model $T = -0.0034h + 100$ where T is measured in degrees Celsius (°C) and h is measured in metres above sea level.

(a) Write down the boiling point of water at sea level.

(b) Use the model to calculate the boiling point of water at a height of 1.37 km above sea level.

(c) Water boils at the top of Mt. Everest at 70°C. Use the model to calculate the height above sea level of Mt. Everest.

7. The temperature in °C of a pot of water removed from a cooker is given by $T(m) = 20 + 70 \times 2.72^{-0.4m}$, where m is the number of minutes after the pot is removed from the cooker.

(a) Show that the temperature of the water when it is removed from the cooker is 90°C.

(b) Calculate the temperature of the water after 10 minutes.

(c) Calculate how long it takes for the temperature of the water to reach 56°C.

(d) Write down the temperature approached by the water after a long time. Justify your answer.

Consider the function $S(m) = 20m - 40$ for $2 \leqslant m \leqslant 6$

The function $S(m)$ represents the temperature of soup in a pot placed on the cooker two minutes after the water has been removed. The soup is then heated.

(e) Comment on the meaning of the constant 20 in the formula for $S(m)$ in relation to the temperature of the soup.

(f) Solve the equation $S(m) = T(m)$. Interpret the solution in context.

(g) Hence describe by using inequalities the set of values of m for which $S(m) > T(m)$

8. Shiyun bought a car in 1999. The value of the car, V, in USD, is depreciating according to the exponential model $V = 25\,000 \times 1.5^{-0.2t}$, $t \geqslant 0$, where t is the time, in years, that Shiyun has owned the car.

(a) Write down the value of the car when Shiyun bought it.

(b) Calculate the value of the car three years after Shiyun bought it. Give your answer correct to two decimal places.

(c) Calculate the time for the car to depreciate to half of its value when Shiyun bought it.

9. In an experiment it is found that a culture of bacteria triples in number every four hours. There are 200 bacteria at the start of the experiment.

(a) Write down the number of bacteria after 8 hours.

(b) Calculate how many bacteria there will be after one day.

(c) Find how long it will take for there to be two million bacteria.

10. The cost per person, in euros, when x people hire an airplane can be determined by the function $C(x) = x + \frac{200}{x}$

(a) Calculate the cost per person when 40 people hire the airplane.

(b) When the number of people who hire the airplane is a or b, the cost per person is 33 euros. Find the value of a and of b.

(c) When n people hire the airplane, the cost per person is the minimum possible value. Find n.

(d) Find the minimum cost per person, to the nearest 0.01 euro.

11. In Fairbanks, Alaska, there are 3.7 hours of daylight on 21 December. This is the shortest day of the year. The number of hours of daylight can be modelled by the function $L(d) = p\cos(q \times d) + r$, where $L(d)$ is the number of hours of daylight d days after 21 December.

(a) Write down the number of hours of daylight for the longest day of the year.

(b) Give a reason why p must be negative.

(c) Hence find the values of p and r.

(d) Assuming that the cycle of day length repeats exactly every 365 days, find the value of q in the form $\frac{a}{b}$, where $a, b \in \mathbb{Z}$

(e) Find the number of days during a year that have less than 6 hours of daylight.

7 Descriptive statistics

7 Descriptive statistics

Learning objectives

By the end of this chapter, you should be familiar with...

- concepts of population, sampling, random sampling, outliers (introduction), discrete and continuous data
- reliability of data sources
- sampling techniques and their effectiveness and bias in sampling
- presenting data: frequency tables for discrete and continuous data, histograms, and cumulative frequency tables and graphs
- how to find median, quartiles, percentiles, range and interquartile range (IQR) and identify outliers
- interpreting box-and-whisker plots, and comparing distributions using box-and-whisker plots
- central tendency and spread of data; calculation of mean and standard deviation; effects of constant changes on the original data.

Statistics are used daily in newspapers, on television, on the internet, in advertisements, and in ordinary conversations. Consider the following examples.

- In a certain large city, people complain that police take a long time to respond to calls for help. A news reporter collected some data over a month and discovered the following: for 1200 victims of a major assault, it took an average of 5 minutes and 20 seconds for the victim to call the police. Once the police were called, she found out that on average a police car was on the scene within 2 minutes and 40 seconds.
- A person sent 10 job applications with his original name and 10 applications with a popular name. The applications with the original name received no response. Those with the popular name received 8 responses.
- *Business Insider* (19 July 2016) reported that just four countries are home to over 60% of the world's high net worth individuals, according to the World Wealth Report by the consulting firm Capgemini.
- The UK Office for National Statistics reported that households without children spent a higher proportion of their total spending on transport than households with children. (Statistical Bulletin, 18 January 2018.)
- Antibacterial soaps are no better than regular soap. This is the result of a study carried out in a large city with 448 households where 224 were given antibacterial soaps and the other 224 were given normal liquid soap. The families used the soap for a year. All participants' hands were cultured for bacteria at the beginning and the end of the study. By the end of the year, tests revealed that they had the same number of bacteria regardless of which soap they used.

Each of these news clips involves information in the form of numbers – length of time, number of responses, percent of wealth, etc. In each of the news items, conclusions can be drawn from the data presented. Statistics is a way to get information from data.

Statistics is gathering, organising, and drawing conclusions from data.

7.1 Data and variables

Data are numerical evidences. To learn from data, we need more than just the numbers. The numbers in a medical study, for example, mean little without some knowledge of the goals of the study and of what blood pressure, heart rate, and other measurements contribute to those goals. That is, data are numbers with a context, and we need to understand the context if we are to make sense of the numbers. On the other hand, measurements from the study's several hundred subjects are of little value to even the most knowledgeable medical expert until you organise, display, and summarise them. We start with examining data. This branch of statistics is called **descriptive statistics**.

Example 7.1

A new advertisement for an ice-cream in May last year resulted in a 35% increase in sales. Therefore, the advertisement was effective. Do you agree?

Solution

A major flaw in this argument is looking at the numbers only. A 35% increase is good. However, knowing the purpose and background is also important. In the northern hemisphere, general sales of ice-cream increase in the months of June, July, and August regardless of advertisements.

Variables

Any set of data contains information about some group of individuals. The information is organised in variables.

Take for example part of the school records by the end of school year, which may look like Figure 7.1.

Variable

A unit, or individual, or object

ID	Gender	Class	Final	Fee
1101	F	6	85	25640
1102	M	8	72	31200
1103	M	2	93	18700

Observation, value, measurement

Figure 7.1 A sample of school records

Individuals are the objects described in a set of data. Individuals are sometimes people. When the objects that we want to study are not people, we often call them **cases**. Units, objects and other names are frequently used too.

A **variable** is any characteristic of an individual. A variable can take different values for different individuals.

The data set is constructed to keep track of some information about the students. The individuals are the students in the class. There are five variables in this data set. These include an identifier for each student – ID, gender, class, A **qualitative (categorical** or **attribute) variable** places an individual into one of two or more groups or categories.

A **quantitative variable** takes numerical values representing counts or measurements for which arithmetic operations such as adding and averaging make sense. Quantitative variables can be split into two types.

Discrete variables result from a finite (countable) number of possible values. Number of children in a family, number of customers waiting to be served, the number of eggs a hen can lay per week, etc. are discrete variables.

Continuous variables can have infinitely many possible values that can be associated with points on a continuous scale in such a way that there are no gaps or interruptions. The time it takes to finish your homework, the thickness of a hard disk, the amount of milk a cow can produce per day, etc. are continuous variables.

The **distribution** of a variable tells us what values it takes and how often it takes these values.

final score, and fee. There are no units of measurement for ID, gender or class. The other variables both have a unit. For example, Final could be a number of marks and Fee is probably a currency. Some variables, like gender and class, simply place individuals into categories. Others, like Final and Fee, take numerical values for which we can do arithmetic and find averages.
It makes sense to give an average grade or fee, but it does not make sense to give an average gender. We can, however, count the numbers of female and male students and do arithmetic with these counts.

Example 7.2

Identify each variable as qualitative or quantitative, and if quantitative, whether it is discrete or continuous.

(a) A cigarette contains 16.13 mg of tar.

(b) There were 15 females and 21 males in the statistics course last term.

(c) Major earthquakes happened in 1952, 1960, 1964, 2004, and 2011.

(d) The average rainfall in Rio de Janeiro for the first 6 months of the year is: 114 mm, 105 mm, 103 mm, …

(e) Upon completion of training for his upcoming bicycle race, Kevin's mass was 3.26 kg less than when he started.

(f) Zip codes (post codes) of a country.

Solution

(a) Quantitative, continuous

(b) Qualitative

(c) Quantitative, discrete

(d) Quantitative, continuous

(e) Quantitative, continuous

(f) Qualitative (There is no numerical meaning, they are just numbers and sometimes letters.)

Populations and samples

Population is the complete collection of all elements (scores, people, measurements, etc.) to be studied.

A **sample** is a sub-collection of elements drawn from a population.

Descriptive statistics is a range of methods for planning the collection of, collecting, organising, summarising, and presenting data. There are also tools to analyse, interpret, and draw conclusions based on the collected data. This is called **inferential** statistics. At the core of these two branches we have the concepts of **population** and **sample**.

If we are interested in predicting which party will win the next elections, then our population will be all registered voters. An opinion poll organisation will choose a certain number of voters and ask them about their preferences. This is a sample.

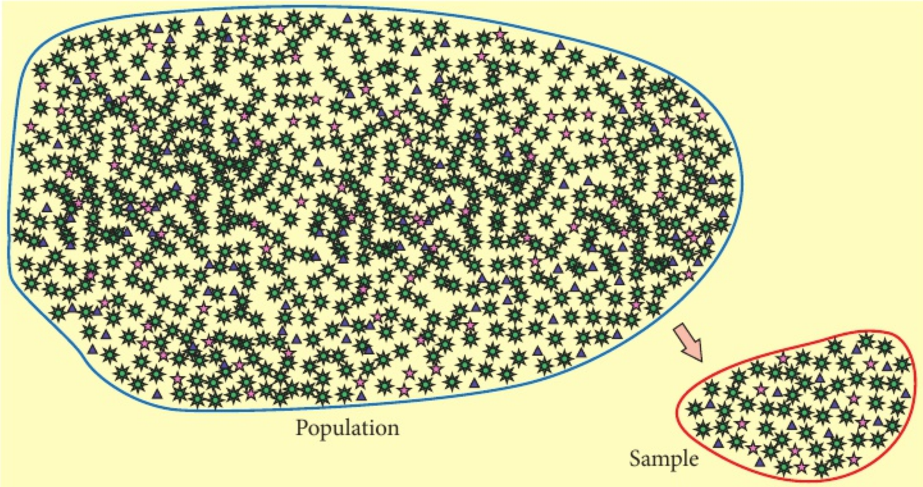


Figure 7.2 Population and sample

A computer company buys a shipment of hard disks from a supplier. The contract with the supplier states that less than 1% of the disks may be defective. The company cannot test all the shipment. It is time-consuming, costly, and also partly damages the units. They perform what is called acceptance sampling by taking a sample, say of 100 disks, and if more than 1 defective disk is found, the shipment is returned. The whole shipment is the population and the 100 disks selected is a sample.

A **parameter** is a numerical characteristic of the population.
A **statistic** is a numerical characteristic of the sample.

Information contained in the sample is usually used to make an inference concerning a **parameter**, which is a numerical characteristic of the population. For instance, in predicting the winning party in elections, we are interested in the proportion of voters choosing that party on election day. That is a parameter. If we are interested in the average time commuters spend on using public transport, then that is a parameter. A parameter is estimated by computing a similar characteristic of the sample. This is called a **statistic**.

Example 7.3

150 000 European Union students attend university in the UK. To estimate the average yearly expenditure of these students a service provider plans to conduct a study by interviewing 500 students.

What is the population, the parameter, the sample and the statistic?

Solution

The population is the 150 000 EU students in the UK.

The parameter is the average yearly expenditure.

The 500 students who will be interviewed constitute a sample.

The average expenditure of this sample is a statistic that could be used to estimate the parameter of interest.

Sampling

As you know by now, any study concerning populations needs data to be collected. Usually we do not collect data from the entire population. For statistical studies, data from samples is used. Schemes similar to the one below are followed:

- Specify the population of interest.
- Choose an appropriate sampling method.
- Collect the sample data.
- Analyse the pertinent information in the sample.
- Use the results of the sample analysis to make an inference about the population.
- Provide a measure of the inference's reliability.

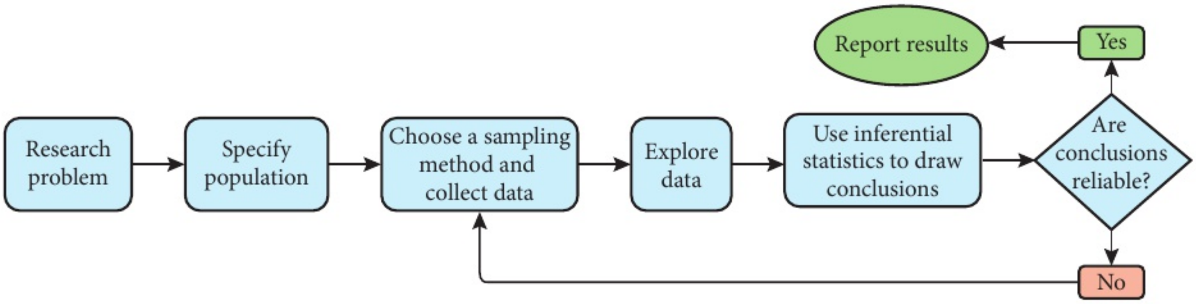


Figure 7.3 Sampling process

Reasons for sampling

Taking a sample instead of conducting a census offers several advantages.

- The sample can save money and time. If an eight-minute interview is being undertaken, conducting the interviews with a sample of 100 people rather than with a population of 100 000 is obviously less expensive. In addition to the cost savings, the significantly smaller number of interviews usually requires less total time.
- For given resources, the sample can broaden the scope of the study. With fixed resources, more detailed information can be gathered by taking a sample than gathering information from the whole population. Concentrating on fewer individuals or items, the study can be broadened in scope to allow for more specialised questions.
- Some research processes are destructive to the product or item being studied. For example, if light bulbs are being tested to determine how long they burn or if candy bars are being taste tested to determine whether the taste is acceptable, the product is destroyed.
- If accessing the entire population is impossible, a sample is the only option. If sampling is deemed to be appropriate, it must be decided how to select a sample. Since the sample will be employed to draw conclusions about the entire population, it is crucial that the sample is **representative** – it should reflect as closely as possible the relevant parameter of the population under consideration.

A sample that represents the characteristics of the population as closely as possible is called a **representative sample**.

Random and non-random sampling

The two main types of sampling are random and non-random. In **random sampling** every unit of the population has the same probability of being selected into the sample. Random sampling implies that chance enters into the process of selection.

In **non-random sampling** not every unit of the population has the same probability of being selected into the sample.

Sometimes random sampling is called **probability sampling** and non-random sampling is called **nonprobability sampling**. Because every unit of the population is not equally likely to be selected in non-random sampling, assigning a probability of occurrence is impossible. The statistical methods presented and discussed in your syllabus assume that the data comes from random samples.

Non-random sampling methods are not appropriate techniques for gathering data to be analysed by most of the statistical methods presented in this book. However, several non-random sampling techniques are described in this section, primarily to alert you to their characteristics and limitations.

Random sampling

Three basic random sampling techniques – simple random sampling (SRS), stratified random sampling, and systematic random sampling – are discussed here. Each technique offers advantages and disadvantages. Some techniques are simpler to use, some are less costly, and others show greater potential for reducing sampling error.

Sampling or chance error

Generally, all samples selected from the same population will give different results because they contain different elements of the population. Additionally, the results obtained from any one sample will not be exactly the same as the ones obtained from a census. The difference between a sample result and the result we would have obtained by conducting a census is called the **sampling error**, assuming that the sample is random and no non-sampling error has been made.

Non-sampling errors can occur both in a sample survey and in a census. Such errors occur because of human mistakes and not chance.

Sampling error occurs when, by chance, the sample does not represent the population.

Simple random sampling

The most elementary random sampling technique is **simple random sampling**. Simple random sampling can be viewed as the basis for the other random sampling techniques. With simple random sampling, each unit of the frame is numbered from 1 to N (where N is the size of the population). Next, a random number generator (or a table of random numbers, which is an outdated technique) is used to select n items into the sample.

Suppose it has been decided to interview 20 out of 659 high school students in a school to form an understanding of their views of a new block-scheduling the school wants to adopt.

We number the students from 001 (or simply 1) to 659 and have a random generator choose 20 of them. The numbers chosen will be your sample. The GDC output gives us a list of chosen numbers. The screenshot (Figure 7.4) shows five of them.

```
RanInt#(1,659,20)
{217,100,191,518,252▸
```

Figure 7.4 GDC random number generator

In a **stratified random sample**, we first divide the population into subpopulations, which are called strata. Then, one sample is selected from each of these strata. The collection of all samples from all strata gives the stratified random sample.

Stratified random sampling

In **stratified random sampling**, the population is divided into non-overlapping subpopulations called strata. The researcher then extracts a sample from each of the subpopulations. The main reason for using stratified random sampling is that it has the potential for reducing sampling error.

With stratified random sampling, the potential to match the sample closely to the population is greater than it is with simple random sampling because portions of the total sample are taken from different population subgroups. However, stratified random sampling is generally more costly than simple random sampling because each unit of the population must be assigned to a stratum before the random selection process begins.

Strata selection is usually based on available information. Such information may have been obtained from previous censuses or surveys. Stratification benefits increase as the strata differ more. Internally, a stratum should be relatively homogeneous; externally, strata should contrast with each other.

In statistics, the convention is that when we say between 20 and 30, 20 is included but 30 is not.

One advantage of stratified sampling is that in addition to collecting information about the entire population, we can also compare different strata. For example, in Figure 7.5, the information we get will also help us compare the different age groups among each other.

For example, in FM radio markets, age of listener is an important determinant of the type of programming used by a station. Figure 7.5 contains a stratification by age with three strata, based on the assumption that age makes a difference in preference of programming. This stratification implies that listeners of 20 to 30 years of age tend to prefer the same type of programming, which is different from that preferred by listeners of 30 to 40 and of 40 to 50 years of age. Within each age subgroup (stratum), there is homogeneity; between each pair of subgroups there is a difference, or heterogeneity. A sample from each stratum is taken and together they constitute a representative sample of the whole population.

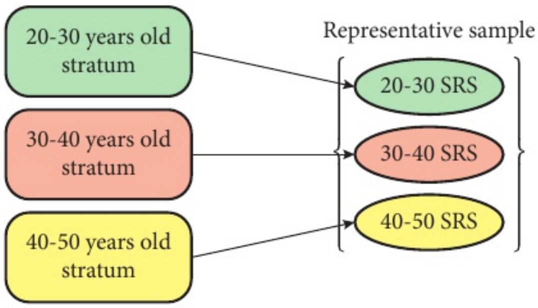


Figure 7.5 Stratified random sampling

Systematic sampling

Unlike stratified random sampling, **systematic sampling** is not done in an attempt to reduce sampling error. Rather, systematic sampling is used because of its convenience and relative ease of administration. With systematic sampling, every kth item is selected to produce a sample of size n from a population of size N. The value of k, sometimes called the sampling cycle, can be determined using

$$k = \frac{N}{n}$$

If k is not an integer value, then a whole-number value should be used.

For example, suppose we need a sample of 20 students from the 659 high school students using systematic sampling. $n = 20$ and $N = 659$, so:

$$k = \frac{659}{20} \approx 32$$

From the list of 659 students we randomly choose a starting number between 1 and 32, say 11 for example, and then after that we choose every 32nd number, i.e., $11 + 32 = 43, 75, 107$

Besides convenience, systematic sampling has other advantages. Because systematic sampling is evenly distributed across the population, a knowledgeable person can easily determine whether a sampling plan has been followed in a study.

Non-random sampling

Sampling techniques used to select elements from the population by any mechanism that does not involve a random selection process are called **non-random sampling techniques**. Because chance is not used to select items for the samples, these techniques are non-probability techniques and are not desirable for use in gathering data to be analysed by standard methods of inferential statistics. Sampling error cannot be determined objectively for these sampling techniques. Two non-random sampling techniques are presented here: convenience sampling, and quota sampling.

Convenience sampling

In **convenience sampling**, elements for the sample are selected for the convenience of the researcher. The researcher typically chooses elements that are readily available, nearby, or willing to participate. The sample tends to be less variable than the population because in many environments the extreme elements of the population are not readily available. The researcher will select more elements from the middle of the population. For example, a convenience sample of homes for door-to-door interviews might include houses where people are at home, houses with no dogs, houses near the street, first-floor apartments, and houses with friendly people. In contrast, a random sample would require the researcher to gather data only from houses and apartments that have been selected randomly, no matter how inconvenient or unfriendly the location. If research is carried out in a mall, a convenience sample might be selected by interviewing only shoppers who pass the location and look friendly.

Quota sampling

Quota sampling appears to be similar to stratified random sampling. However, instead of selecting a sample from each stratum, we use a non-random sampling method to gather data from one stratum until the desired quota of samples is filled. Quotas are described by setting the sizes of the samples to be obtained from the subgroups. Generally, a quota is based on the proportions of the subclasses in the population.

For example, a company is test marketing a new soft drink and is interested in how different age groups react to it. The researchers go to a shopping mall and interview shoppers aged 16–20, for example, until enough responses are obtained to fill the quota. In quota sampling, an interviewer would begin by asking a few filter questions; if the respondent represents a subclass whose quota has been filled, the interviewer would stop the interview.

Quota sampling can be useful if no previous information is available for the population. For example, suppose we want to stratify the population into cars using different types of winter tyres but we do not have lists of users of a particular brand of tyres. Through quota sampling, we would proceed by interviewing all car owners and rejecting all users who do not use the particular brand until the quota of users of the particular brand is filled.

Quota sampling is less expensive than most random sampling techniques because it essentially is a technique of convenience. Another advantage of quota sampling is the speed of data gathering. We do not have to call back or send out a second questionnaire if we do not receive a response; we just move on to the next element.

The problem with quota sampling is that it is a non-random sampling technique. Some researchers believe that a solution to this issue can be achieved if the quota is filled by randomly selecting elements and discarding those not from a stratum. This way quota sampling is essentially a version of stratified random sampling. The object is to gain the benefits of stratification without the high costs of stratification. However, it remains a nonprobability sampling method.

Exercise 7.1

1. Identify the experimental units, sensible population and sample on which each variable is measured. Then indicate whether the variable is quantitative or qualitative.

(a) Gender of a student.

(b) Number of errors on a final exam for 10th grade students.

(c) Height of a newly born child.

(d) Eye colour for children aged less than 14.

(e) Amount of time it takes to travel to work.

(f) Rating of a country's leader: excellent, good, fair, poor.

(g) Country of origin of students at international schools.

2. For each situation:

(i) identify the population

(ii) give two examples of how an appropriate sample of the population can be obtained.

(a) A large high school with students in grades 10–12 has 42 classes with a total of 1176 students. Information about sizes of students are needed for the school to stock enough sweatshirts for those interested in purchasing one. The sample size required is 30.

(b) In a machine parts factory, bolts are produced on an assembly line in three shifts and are collected in large containers. For quality control purposes, bolts are to be tested for their size and strength using a digital sensor.

3. Identify each variable as numerical (quantitative) or categorical (qualitative) data. If the data is numerical, classify it as discrete or continuous.
   (a) The blood type of a group of blood donors.
   (b) The number of cars each family in a community has.
   (c) The length of a fish caught from a pond.
   (d) The amount of time spent per evening studying mathematics.
   (e) The volume of liquid in a canned drink.
   (f) The number of languages spoken in a given community.
   (g) 100-metre race times.
   (h) Colours used by maths teachers to write on their whiteboards.
   (i) The rating of trumpet solos as superior, excellent or good.

4. Group the set of situations as categorical or numerical. For the numerical data, identify each as discrete or continuous.
   (a) Volume of paint in a can.
   (b) Number of children in a family.
   (c) Time taken to finish a marathon.
   (d) The length of an average children's film.
   (e) The height of the students in an IB Mathematics class.
   (f) The average carbon dioxide emissions in a large city.
   (g) The religion(s) practiced in each household of a city.

5. Classify each as an example of descriptive or inferential statistics.
   (a) Gathering data to determine if the average mass of 16-year-old football players in the UK is 60 kg.
   (b) Gathering data to determine a basketball player's free throw percentage for a season of 10 games.

6. State whether inferential or descriptive statistics would be needed for the following.
   (a) Finding the average mass of all 12-year-old boys in Spain.
   (b) Finding the average amount of money you spend on entertainment for the next 3 months.
   (c) Finding the number of heads that show after flipping a coin 100 times.
   (d) Determining if the proportion of women who wear seat belts is greater than the proportion of men who wear seat belts.
   (e) Determining if the average IB Mathematics score for boys is greater than the average IB Mathematics score for girls.

7. A professor wanted to select 20 students from his class of 300 students to collect detailed information on their profiles. He used his knowledge and expertise to select these 20 students.
   (a) Is this sample a random or a non-random sample? Explain.
   (b) Is this an SRS, a systematic sample, a stratified sample, a convenience sample, a judgment sample, or a quota sample? Explain your answer.
   (c) What kind of error, if any, will be made with this kind of sample? Explain.

   Suppose the professor enters the names of all the students enrolled in his class on a computer. He then selects a sample of 20 students at random using a statistical software package.
   (d) Explain whether this sample is random or non-random.
   (e) Explain what kind of sample it is.
   (f) Do you think any error will be made in this case? Explain your answer.
8. A company has 1000 employees, of whom 58% are men and 42% are women. The research department at the company wanted to conduct a quick survey by selecting a sample of 100 employees and asking them about their opinions on an issue. They divided the population of employees into two groups, men and women, and then selected 58 men and 42 women from these respective groups. Explain what kind of sample it is.
9. A large university would like to determine the views of its students on an increase in fees. They would also like to compare the faculties of business and arts and sciences as well as the graduate school. Describe a sampling plan to achieve this goal.
10. An opinion poll is to be given to a sample of 90 members of a local gym. The members are first divided into men and women, and then a simple random sample of 45 men and a separate simple random sample of 45 women are taken. What sampling method has been used? Explain.

7.2 Displaying distributions using graphs

Statistical tools and ideas help us examine data in order to describe their main features. This examination is called **exploratory data analysis.** Like an explorer crossing unknown lands, we want first to simply describe what we see.

One efficient approach is to begin with a graph or graphs. Then add numerical summaries of specific aspects of the data. This section presents methods for describing a single variable.

There are several ways of summarising and describing data. Among them are tables and graphs and numerical measures.

I just purchased a bag of milk chocolate sweets that have six different colours. There are 55 sweets in the bag: 16 brown, 13 red, 9 yellow, 7 green, 4 blue, and 6 orange. Table 7.1 shows these counts.

Colour	Frequency
Brown	16
Red	13
Yellow	9
Green	7
Blue	4
Orange	6

Table 7.1 Colours of sweets

This table is called a frequency table and it describes the distribution of sweet colour frequencies. Not surprisingly, this kind of distribution is called a **frequency distribution**. Often a frequency distribution is shown graphically as in Figure 7.6.

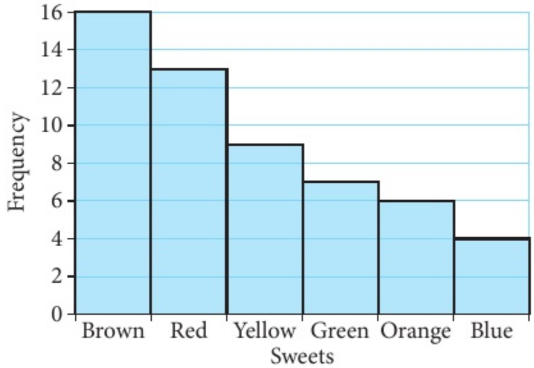


Figure 7.6 A bar graph of colours of sweets

The distribution shown in Figure 7.6 concerns just one bag of sweets. What about the distribution of colours for all sweets?

The manufacturer of the sweets provides some information about this matter, but they do not say exactly how many sweets of each colour they have ever produced. Instead, they report proportions rather than frequencies. Figure 7.7 shows these proportions.

Note that the distributions in Figures 7.6 and 7.7 are not identical. Figure 7.6 portrays the distribution in a **sample** of 55 of the sweets. Figure 7.7 shows the proportions for all of the sweets. Chance factors involving the machines used by the manufacturer introduce random variation into the different bags produced. Some bags will have a distribution of colours that is close to Figure 7.7; others will be further away.

Since every sweet is one of the six familiar colours, the six proportions shown in the figure add to 100. We call Figure 7.7 a relative frequency distribution (or probability distribution) because if you choose a sweet at random, the probability of getting, say, a brown sweet is equal to the proportion of sweets that are brown (0.30).

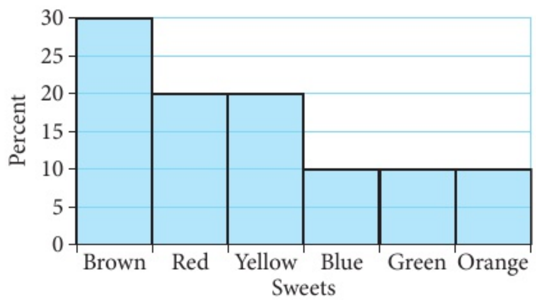


Figure 7.7 Relative frequency distribution

Bar graphs and similar charts help an audience grasp a distribution quickly. They are, however, of limited use for data analysis because it is usually easy to understand data on a single categorical variable, such as highest frequency, without a graph. We will move on to quantitative variables, where graphs are essential tools.

Graphs similar to the ones in Figures 7.6 and 7.7 are mainly used for qualitative variables and are called **bar graphs**.

Raw data, or data that have not been summarised in any way, are sometimes referred to as ungrouped data. Table 7.2 contains 110 data points of raw data of the unemployment rates for 110 countries for 2017 as reported by the International Monetary Fund (IMF). Data that have been organised into a frequency distribution are called **grouped data.**

0.7	3.1	4.0	4.4	5.1	5.8	6.7	7.7	9.0	11.3	14.6
1.0	3.2	4.0	4.4	5.2	5.8	6.7	8.0	9.1	11.7	15.3
1.1	3.2	4.0	4.5	5.4	5.8	6.8	8.0	9.3	11.8	16.5
2.0	3.4	4.0	4.6	5.4	6.0	6.9	8.1	9.4	12.2	17.2
2.2	3.4	4.0	4.7	5.5	6.0	6.9	8.3	9.4	12.2	18.9
2.2	3.7	4.2	4.9	5.6	6.1	6.9	8.4	9.8	12.2	19.6
2.8	3.7	4.2	5.0	5.6	6.2	7.1	8.7	10.1	12.2	20.5
2.9	3.8	4.2	5.0	5.7	6.3	7.1	8.7	10.2	12.5	21.5
2.9	3.8	4.2	5.0	5.7	6.7	7.2	8.9	11.0	12.8	22.5
3.0	3.9	4.4	5.0	5.7	6.7	7.4	9.0	11.3	13.9	27.5

Table 7.2 Unemployment rates of 110 countries for 2017 as reported by IMF

Table 7.3 presents a frequency distribution for the data displayed in Table 7.2, where the data have been grouped. The distinction between ungrouped and grouped data is important because the calculation of statistics differs between the two types of data.

When grouping data, there are some steps to be followed.

Step 1 Determine the range of the raw data. The range is often defined as the difference between the largest and smallest numbers. The range for the data in Table 7.2 is 26.8 $(27.5 - 0.7)$

Rate	Frequency
$0 \leqslant x < 3$	9
$3 \leqslant x < 6$	44
$6 \leqslant x < 9$	26
$9 \leqslant x < 12$	14
$12 \leqslant x < 15$	8
$15 \leqslant x < 18$	3
$18 \leqslant x < 21$	3
$21 \leqslant x < 24$	2
$24 \leqslant x < 27$	0
$27 \leqslant x < 30$	1

Table 7.3 Distribution for grouped unemployment rates

Step 2 Determine how many classes it will contain. A good general rule is to select between 5 and 15 classes. If the frequency distribution contains too few classes, the data summary may be too general to be useful. Too many classes may result in a frequency distribution that does not aggregate the data enough to be helpful. The final number of classes is arbitrary.

By examining the range we can determine a number of classes that will span the range adequately and also be meaningful to the user. The data in Table 7.2 were grouped into 9 classes for Table 7.3.

Step 3 Determine the width of the class interval. An approximation of the class width can be calculated by dividing the range by the number of classes.

For this data, this approximation would be $\frac{26.8}{9} = 2.98$

Normally, the number is rounded up to the next whole number, which in this case is 3. The frequency distribution must start at a value equal to or lower than the lowest number of the ungrouped data and end at a value equal to or higher than the highest number. The lowest unemployment rate is 0.7 and the highest is 27.5, so we start the frequency distribution at 0 and end it at 30. Table 7.3 contains the completed frequency distribution for the data in Table 7.2. Class endpoints are selected so that no value of the data can fit into more than one class.

The midpoint of each class interval is called the **class midpoint** and is sometimes referred to as the **class mark**. It is the value halfway across the class interval and can be calculated as the average of the two class endpoints. For example, in the distribution of Table 7.3, the midpoint of the class interval $3 \leqslant x < 6$ is 4.5, or $\frac{3+6}{2}$

The class midpoint is important, because it becomes the representative value for each class in most grouped statistics calculations.

Relative frequency is the proportion of the total frequency that is in any given class interval in a frequency distribution. Relative frequency is the individual class frequency divided by the total frequency. For example, from Table 7.3, the relative frequency for the class interval $3 \leqslant x < 6$ is $\frac{44}{110} = 0.4$ (Table 7.4).

Cumulative frequency is a running total of frequencies through the classes of a frequency distribution. The cumulative frequency for each class interval is the frequency for that class interval added to the preceding cumulative total. In Table 7.4 the cumulative frequency for the first class is the same as the class frequency: 9. The cumulative frequency for the second class interval is the frequency of that interval (44) plus the frequency of the first interval (9), which gives a cumulative frequency of 53. This process continues through the last interval, at which point the cumulative total should equal the sum of the frequencies (110). Table 7.4 gives cumulative frequencies for the data in Table 7.3.

Rate	Frequency	Class midpoint	Relative frequency	Cumulative frequency
$0 \leqslant x < 3$	9	1.5	0.0818	9
$3 \leqslant x < 6$	44	4.5	0.4000	53
$6 \leqslant x < 9$	26	7.5	0.2364	79
$9 \leqslant x < 12$	14	10.5	0.1273	93
$12 \leqslant x < 15$	8	13.5	0.0727	101
$15 \leqslant x < 18$	3	16.5	0.0273	104
$18 \leqslant x < 21$	3	19.5	0.0273	107
$21 \leqslant x < 24$	2	22.5	0.0182	109
$24 \leqslant x < 27$	0	25.5	0.0000	109
$27 \leqslant x < 30$	1	28.5	0.0091	110

$\boxed{0.0818 = \frac{9}{110}}$

$\boxed{0.2364 = \frac{26}{110}}$

Table 7.4 Cumulative frequency distribution for unemployment data

Example 7.4

The table shows the scores of 80 students in an IB class achieved in an assessment, out of a total of 100 marks.

42	50	61	70	81	90	57	65	76	47
41	50	61	71	82	92	57	65	78	47
42	50	61	71	83	93	57	66	79	48
42	52	61	72	83	94	57	67	86	49
45	52	62	72	83	98	57	67	88	49
45	53	62	72	84	98	57	67	89	59
45	53	63	74	84	45	58	67	89	69
45	56	64	75	84	46	58	68	59	69

(a) Develop a frequency and relative frequency table for the data.

(b) Develop a cumulative and relative cumulative frequency table of the data.

Solution

(a) Since the lowest potential grade is 40 and the highest is 100, we will choose 6 classes. We may choose another number too. However, separating grades into classes of multiples of 10 is a sensible choice: $\frac{98 - 41}{6} \approx 10$

Scores achieved (x)	Number of students (Frequency)	Relative frequency	Cumulative frequency	Relative cumulative frequency
$40 \leqslant x < 50$	15	18.7%	15	18.7%
$50 \leqslant x < 60$	18	22.5%	33	41.2%
$60 \leqslant x < 70$	18	22.5%	$33 + 18 = 51$	63.7%
$70 \leqslant x < 80$	11	13.8%	62	77.5%
$80 \leqslant x < 90$	12	15.0%	74	92.5%
$90 \leqslant x \leqslant 100$	6	7.5%	80	100%
Total	**80**	**100%**		

Each cell in the relative frequency column is the ratio of the cell's frequency to the total number. For example, the first cell is $\frac{15}{80} = 0.187 = 18.7\%$ and the last cell is $\frac{6}{80} = 0.075 = 7.5\%$

(b) Each cell in the cumulative frequency column is the sum of the cumulative frequency of the previous cell and the frequency of the interval itself . The last column shows relative cumulative frequency.

Histograms

Although frequency distribution tables are useful for analysing large sets of data, their table format may not be as visually informative as a graph. If a frequency distribution has been developed from a quantitative variable, a **histogram** can be constructed directly from the frequency distribution. A histogram is a graph that consists of vertical bars constructed on a horizontal line that is marked off with intervals for the variable being displayed. The intervals correspond to those in a frequency distribution table. The area of each bar is proportional to the number of observations in that interval. In many cases, the histogram enables the data to be interpreted more easily.

You cannot use histograms for qualitative variables. However, you can use a bar chart instead.

A histogram shows three general types of information:

1. It provides a visual indication of where the approximate centre of the data is.
2. We can gain an understanding of the degree of spread (or variation) in the data. The more the data cluster around the centre, the smaller the variation in the data. If the data are spread out from the centre, the data exhibits greater variation.

3. We can observe the shape of the distribution. Is it reasonably flat, is it skewed to one side or the other, is it balanced around the centre, or is it bell shaped?

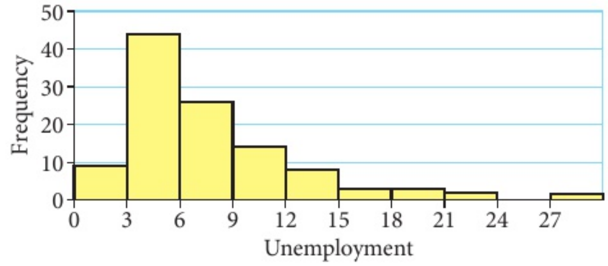


Figure 7.8 Histogram of the unemployment data

We can construct a histogram using the data in Table 7.3 using technology (Figure 7.8).

Each class is represented by a rectangle with a height that corresponds to the frequency. For example, the first class has a frequency of 9, and so, the height of the rectangle is 9.

Figure 7.8 shows that the data are skewed to the right: there are fewer countries with high unemployment. We can also see that very few countries have an unemployment rate less than 3% and that a majority of the countries have rates between 3% and 9%.

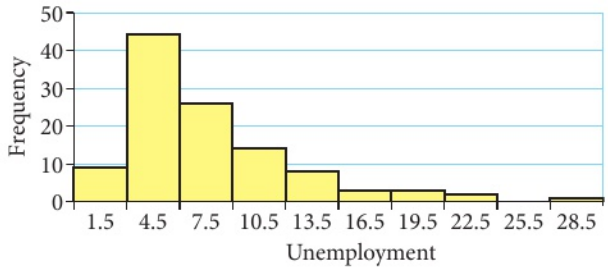


Figure 7.9 Histogram for unemployment data using midpoints

Histograms produced using technology can be presented in different forms. In Figure 7.8, we used the standard convention for histograms where the class boundaries separate the classes with the lowest boundary included in the class while the upper one is included in the next class.

Some software presents the histogram where the rectangles are placed directly above the class midpoints (Figure 7.9).

GDCs can also produce histograms of relatively good quality. The screenshot in Figure 7.10 shows an example where class boundaries have been entered in one list and the frequencies in another list. The GDC then draws the histogram.

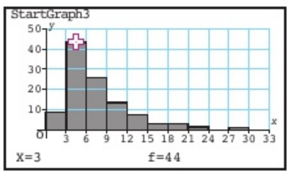


Figure 7.10 GDC histogram

Cumulative and relative frequency histograms and graphs

For many practical situations, it is more efficient to look at histograms representing the relative frequencies instead of frequencies. This is especially true when we compare two samples of different sizes. The shape of the histogram will not change.

For example, with the same unemployment data, Figure 7.11 shows a relative frequency histogram. Note that the shape is still the same but the numbers on the vertical axis shows percent (relative frequency) instead.

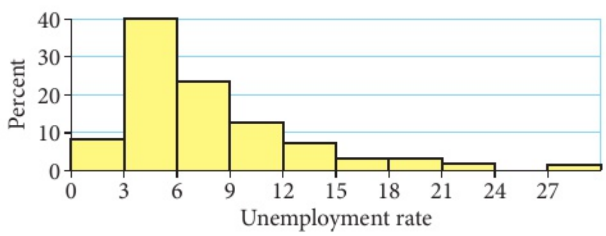


Figure 7.11 Relative frequency histogram for unemployment data

A **frequency graph** (also called frequency polygon) shows the same information as the histogram with the rectangles replaced by points at the midpoint of each class (Figure 7.12).

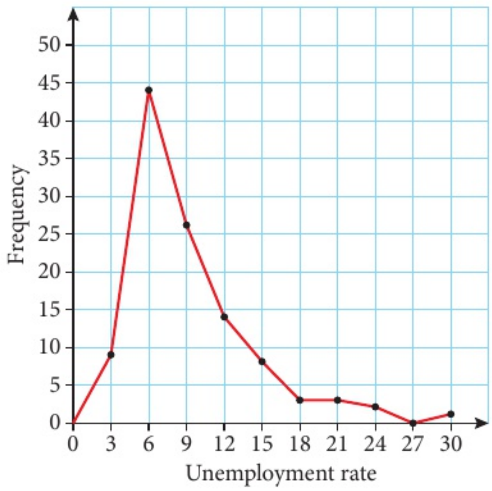


Figure 7.12 Frequency graph for unemployment data

Figure 7.13 Cumulative frequency graph the same data

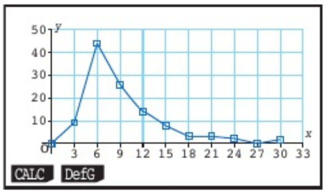


Figure 7.14 GDC frequency graph

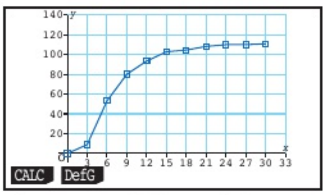


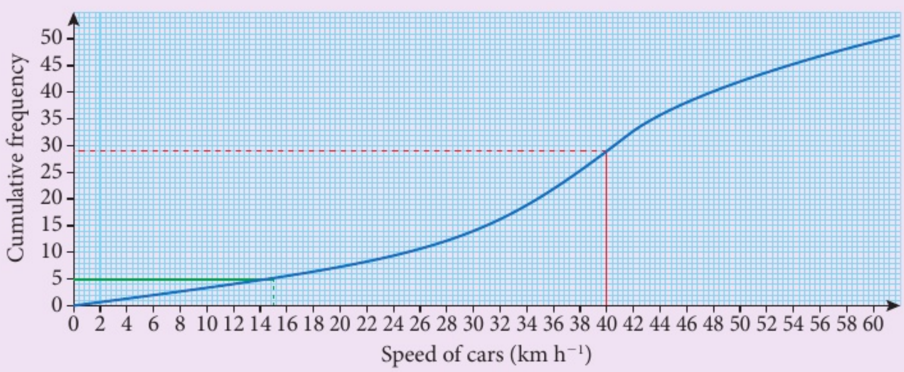
Figure 7.15 GDC cumulative frequency graph

A **cumulative frequency graph** (polygon) for each class. It has to be drawn at the upper limit for each class because its height represents the cumulative frequency up to that point. For example, in Figure 7.13 the height of the graph corresponding to 6 is 44, which is the cumulative frequency up to 6.

Figures 7.14 and 7.15 show two graphs produced by a GDC.

Example 7.5

At a busy intersection, the speed of passing cars is recorded and the information is represented in a cumulative frequency graph.



(a) How many cars were included in the sample?

(b) The speed limit at the intersection is 40 km h^{-1}. How many cars are driving at or below the speed limit? What percentage of the cars is that?

(c) If a car is travelling more than 10 km h^{-1} above the speed limit, they receive a speeding ticket. How many cars would receive a speeding ticket?

(d) The slowest 10% of all drivers are seen as a slow safety hazard, and are issued a fine. What is the maximum speed considered to be a slow safety hazard?

Solution

(a) From the graph, the speed of 50 cars were measured in the sample.

(b) Drawing a vertical line to the graph at $40\,\text{km h}^{-1}$ (red solid line) we can project horizontally onto the cumulative frequency axis (red dashed line) and we find that the slowest 29 cars out of 50 drive at or below the speed limit. This is 58% of the cars.

(c) $10\,\text{km h}^{-1}$ above the limit is $50\,\text{km h}^{-1}$. Drawing a vertical line to the graph at $50\,\text{km h}^{-1}$, we can project horizontally onto the cumulative frequency axis and we find that 41 cars out of 50 drive at or below $50\,\text{km h}^{-1}$, thus, 9 cars are potential recipients of a ticket.

(d) Calculating 10% of 50 gives 5, meaning that the 5 slowest cars are driving at a hazardously slow speed. Drawing a horizontal line to the graph for a cumulative frequency of 5 (green solid line) and projecting to the speed axis (green dashed line) gives the maximum speed of $15\,\text{km h}^{-1}$ for hazardously slow driving.

58% of the data are at 40 or below, so we call $40\,\text{km h}^{-1}$ the 58th **percentile**.

82% of the data are at 50 or below, so we call $50\,\text{km h}^{-1}$ the 82nd **percentile**.

When p% of the data lie on P or below, we call P the pth **percentile**.

Example 7.6

Data is collected for the number of text messages that 50 randomly selected high school students sent during one day. The results are listed below.

8, 52, 38, 48, 42, 9, 15, 36, 36, 53, 10, 8, 46, 46, 9, 11, 12, 24, 49, 34, 10, 11, 9, 11, 45, 25, 25, 37, 14, 16, 20, 22, 12, 43, 36, 23, 23, 26, 27, 16, 21, 29, 29, 38, 30, 47, 34, 39, 48, 46

(a) Set up a frequency and relative cumulative frequency table for the data with 7 classes.

(b) Draw a frequency histogram of the data.

(c) Draw a relative cumulative frequency graph.

(d) Estimate the 10th as well as the 90th percentiles of the data from your graph in part (c).

Solution

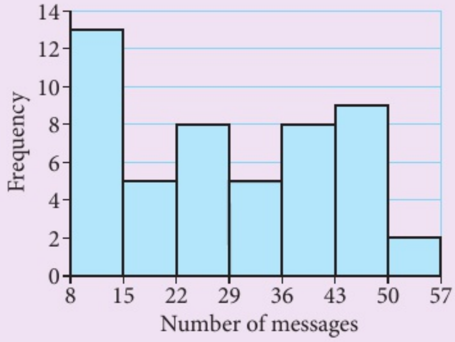
(a) Since the minimum value is 8 and the maximum value is 53, then

the class size would be $\dfrac{53-8}{7} = 6.4$

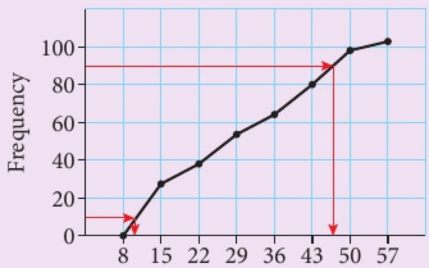
Round this to 7 and start the table with the minimum 8.

Class	Class midpoint	Frequency	Cumulative frequency	Relative cumulative frequency
$8 \leqslant t < 15$	11.5	13	13	13/50 =0.26
$15 \leqslant t < 22$	18.5	5	13 + 5 = 18	0.36
$22 \leqslant t < 29$	25.5	8	26	0.52
$29 \leqslant t < 36$	32.5	5	31	0.62
$36 \leqslant t < 43$	39.5	8	31 + 8 = 39	0.78
$43 \leqslant t < 50$	46.5	9	48	0.96
$50 \leqslant t < 57$	53.5	2	50	1.00

(b)



(c)



(d) As shown in the diagram, the 10th percentile is at 9 or 10 messages, while the 90th percentile is at 46 or 47 messages.

Exercise 7.2

1. A group of college students were selected at random and their ages were recorded. The table shows a summary.

Age	17	18	19	20	21	22	23–25	Over 25
Number of students	3	72	62	31	12	9	5	6
Cumulative frequency	3	75	137	x	180	189	194	y

(a) What are the values of x and y?

(b) How many students are younger than 21?

(c) Find the value of the 25th percentile.

2. The heights (in metres) of 30 students randomly chosen in a large school are recorded in the table.

0.92	1.32	1.76	1.63	1.79	1.28	1.77	1.62	1.611	1.85
1.26	1.67	1.77	1.78	1.93	1.73	1.55	1.52	1.89	1.59
1.78	1.73	1.15	1.76	1.69	1.72	1.04	1.53	1.58	2.00

(a) Choose an appropriate interval width and create a frequency table to represent the data. Is the data discrete or continuous?

(b) Draw a frequency histogram representing the data.

(c) Draw a cumulative frequency graph for the data.

(d) Using your graphs, summarise the data in a short paragraph.

3. The table gives the frequency and cumulative frequency of the amount of time spent by 60 students, in hours, studying per evening.

Time spent studying (h)	Frequency	Cumulative frequency
$0 \leqslant h < 1$	7	7
$1 \leqslant h < 2$	p	18
$2 \leqslant h < 3$	28	q
$3 \leqslant h \leqslant 4$	r	60

(a) State the values of p, q and r.

(b) Draw a frequency histogram and cumulative frequency graph for the data.

(c) Summarise the data in a short paragraph, using your graphs and tables.

4. The frequency bar graph in Figure 7.16 shows IB assessment scores for a group of students.

(a) How many student scores were included in the sample?

(b) What percentage of students achieved a score of 6?

(c) How many scores below 3 were there and what percentage of the total does this represent?

(d) The bar chart is a good representation of a cohort of 800 students. How many students in the cohort would you expect to receive a grade of 7?

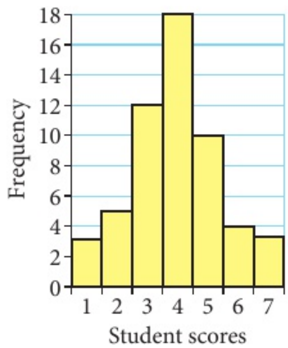
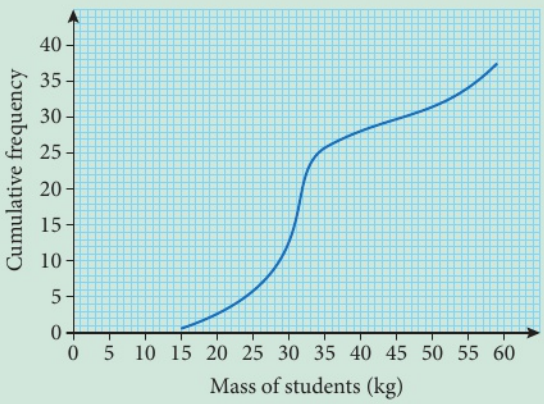


Figure 7.16 Graph for question 4

5. The cumulative frequency graph shows the information about the mass (in kg) of a group of students selected at random from a large public school.



(a) Copy and complete the frequency table.

Mass of students, M (kg)	Frequency
$15 \leqslant M < 30$	
$30 \leqslant M < 45$	
$45 \leqslant M \leqslant 60$	

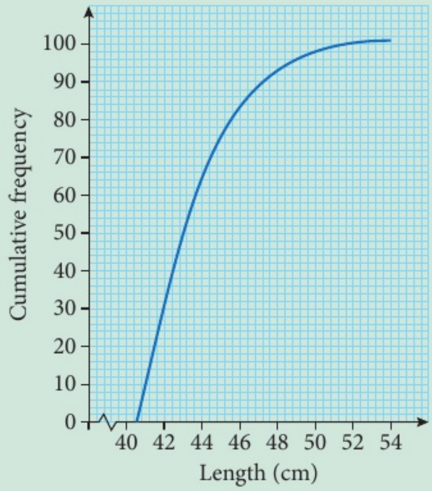
(b) How many students were included in the data sample?

(c) For which interval of masses does the frequency increase the most?

(d) How many students have a mass under 40 kg?

(e) Students in the bottom 25% by mass are given the option to participate in a specialised nutritional program. What is the maximum mass a student could have to be eligible for the program?

6. During an ecological survey, 100 pike are caught from a lake, and their lengths (in cm) are measured. The data is shown in the cumulative frequency graph.



(a) What is the range of lengths that were observed in the survey?

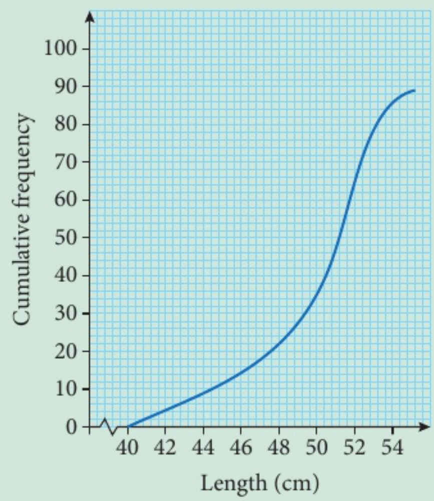
(b) How many pike had a length smaller than 42 cm?

(c) How many pike had a length greater than 47 cm?

(d) Pike that are longer than 50 cm are considered big. What percentage of the pike are big?

(e) A second similar survey is completed at a different location in the same lake, with the cumulative frequency graph shown.

Compare the two pike samples using the cumulative frequency graphs.



7. A large company has a report for days employees are absent per year. Table 7.5 shows a summary of absences.
   (a) Set up a cumulative frequency table.
   (b) Draw a cumulative frequency graph.

Absences (days)	Frequency
$0 \leqslant x \leqslant 5$	30
$6 \leqslant x \leqslant 11$	62
$12 \leqslant x \leqslant 17$	61
$18 \leqslant x \leqslant 23$	30
$24 \leqslant x \leqslant 29$	17

Table 7.5 Data for question 7

8. Grade point averages (GPA) in several colleges are on a scale of 0 to 4. Here are the GPAs of 45 students at a certain college.

1.8	1.9	1.9	2.0	2.1	2.1	2.1	2.2	2.2	2.3	2.3	2.4	2.4	2.4	2.5
2.5	2.5	2.5	2.5	2.5	2.6	2.6	2.6	2.6	2.6	2.7	2.7	2.7	2.7	2.7
2.8	2.8	2.8	2.9	2.9	2.9	3.0	3.0	3.0	3.1	3.1	3.1	3.2	3.2	3.4

   Prepare a frequency histogram, a relative frequency histogram, and a cumulative frequency graph. Describe the data in two to three sentences.

9. The table shows the grades of an IB course with 40 students on a 100-point test. Use graphical methods you learned so far to describe the grades.

61	62	93	94	91	92	86	87	55	56
63	64	86	87	82	83	76	77	57	58
94	95	89	90	67	68	62	63	72	73
87	88	68	69	65	66	75	76	84	85

10. The length of time in months between repeated speeding violations of 50 young drivers are given in the table.

2.1	1.3	9.9	0.3	32.3	8.3	2.7	0.2	4.4	7.4
9	18	1.6	2.4	3.9	2.4	6.6	1	2	14.1
14.7	5.8	8.2	8.2	7.4	1.4	16.7	24	9.6	8.7
19.2	26.7	1.2	18	3.3	11.4	4.3	3.5	6.9	1.6
4.1	0.4	13.5	5.6	6.1	23.1	0.2	12.6	18.4	3.7

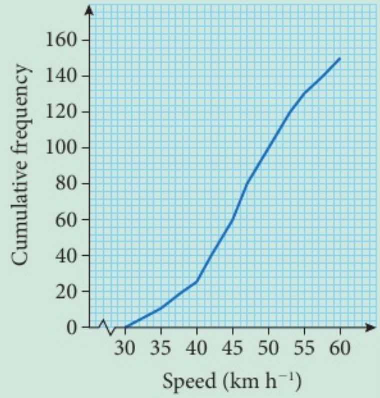
   (a) Construct a histogram for the data.
   (b) Would you describe the shape as symmetric?
   (c) The law in this country requires that the driving license be taken away if the driver repeats the violation within a period of 10 months. Use a cumulative frequency graph to estimate the fraction of drivers who may lose their license.

11. To decide on the number of counters needed to be open during rush hours in a supermarket, the management collected data from 60 customers for the time they spent waiting to be served. The times in minutes are given in the table.

3.6	0.7	5.2	0.6	1.3	0.3	1.8	2.2	1.1	0.4
1	1.2	0.7	1.3	0.7	1.6	2.5	0.3	1.7	0.8
0.3	1.2	0.2	0.9	1.9	1.2	0.8	2.1	2.3	1.1
0.8	1.7	1.8	0.4	0.6	0.2	0.9	1.8	2.8	1.8
0.4	0.5	1.1	1.1	0.8	4.5	1.6	0.5	1.3	1.9
0.6	0.6	3.1	3.1	1.1	1.1	1.1	1.4	1	1.4

   (a) Construct a relative frequency histogram for the times.
   (b) Construct a cumulative frequency graph and estimate the number of customers who have to wait 2 minutes or more.

12. The cumulative frequency graph shows the speeds of cars (km h^{-1}) passing through an intersection.



(a) State the minimum speed of a car travelling through the intersection.

(b) What percentage of cars drive at a speed higher than $55\,\text{km h}^{-1}$ through the intersection?

(c) Given that 40% of the cars travelling through the intersection have a speed higher than $k\,\text{km h}^{-1}$, what is the value of k?

(d) Find the 60th percentile.

13. The waiting time, t, in seconds, that it takes 300 customers at a supermarket cash register are recorded in Table 7.6.

Time, t (s)	Frequency
$t < 60$	12
$60 \leqslant t < 120$	15
$120 \leqslant t < 180$	42
$180 \leqslant t < 240$	105
$240 \leqslant t < 300$	66
$300 \leqslant t < 360$	45
$t \geqslant 360$	15

Table 7.6 Data for question 13

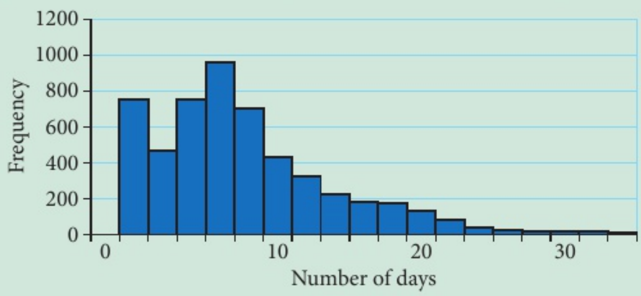
(a) Draw a histogram of the data.

(b) Construct a cumulative frequency graph of the data.

(c) Use the cumulative frequency graph to estimate the waiting time that is exceeded by 25% of the customers.

(d) Find the 75th percentile.

14. The bar graph shows the number of days in hospital spent by heart patients in Austrian hospitals in the 2017–2018 period.



(a) Describe the data in a few sentences.

(b) Draw a cumulative frequency graph for the data.

(c) What percent of the patients stayed less than 6 days?

15. Radar devices are installed at several locations on a main highway. Speeds, in km h^{-1} of 400 cars travelling on that highway are measured and summarised in Table 7.7.

Speed, s (km h^{-1})	Frequency
$60 \leqslant s < 75$	70
$75 \leqslant s < 90$	110
$90 \leqslant s < 105$	150
$105 \leqslant s < 120$	70
$120 \leqslant s < 135$	40
$s \geqslant 135$	10

Table 7.7 Data for question 15

(a) Construct a frequency table for the data.

(b) Draw a histogram to illustrate the data.

(c) Draw a cumulative frequency graph for the data.

(d) The speed limit in this country is $130\,\text{km h}^{-1}$. Use your graph from part **(c)** to estimate the percentage of the drivers driving faster than this limit.

7.3 Measures of central tendency and spread

Help Wanted

Falsead, a fast-growing company, is looking for a lab assistant in its research division.

Minimum requirement:

Master's degree and 2 years of experience.

All benefits. Average salary is €36,000

Call 01-234 567

Figure 7.17 Is it true that the average salary is €36,000?

Catherine was very interested in this advertisement and decided to find out more. After some research she found last year's financial report and the salary structure given in Table 7.8. It is true that the average salary is €36,000. The total salary earned by the 9 employees of the company is €108,000.

So, the average salary is $\frac{324\,000}{9} = 36\,000$. But she also discovered that the salaries are bunched up at the low end of the scale. More than half of the employees earn under €16,000. The salary of €36,000 quoted in the advertisement is not at all typical of the pay at Falsead.

Function	Salary	Number
CEO	€180,000	1
Manager	€24,000	1
Lab coordinator	€21,000	2
Assistants	€15,600	5

Table 7.8 Salary structure

In this section, we will discuss how to describing characteristics of data like this one in more meaningful ways.

Measures of location and centre

One of the first things we want to know about a data set such as salaries, prices of items, number of visitors, and so on is 'about how much?' or 'about how many?'

About how much do people earn at Falsead? About how much is the price of a mountain bike, about how many viewers of a TV program, etc.

When we ask, 'about how much?' we probably want to capture with a single number what is typical of the data.

What single number is most representative of an entire list of numbers? We will study three common measures of location: the **mean**, the **median**, and the **mode**. The mean, median and mode are all 'most representative,' but for different, related notions of representativeness.

The two most common measures of centre are the mean and the median. The mean is the 'average value' and the median is the 'middle value.' These are two different ideas for 'centre,' and the two measures behave differently.

Mean

The average or mean of a data set is usually simple to find. Just add all the values and divide by the number of values.

For example, the average salary at Falsead is

$$\frac{180\,000 + 24\,000 + 2 \times 21\,000 + 5 \times 15\,600}{9} = \frac{324\,000}{9} = 36\,000$$

The **arithmetic mean** or **average** of a set of n measurements (data set) is equal to the sum of the measurements divided by n.

$$\text{Mean} = \bar{x} = \frac{x_1 + x_2 + \cdots + x_n}{n} = \frac{\sum_{i=1}^{n} x_i}{n} = \frac{\sigma x}{n}$$

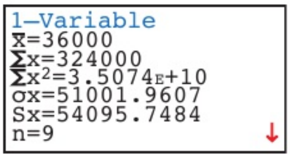


Figure 7.18 GDC mean value

In practice, you can key the data into your GDC which will give you the value of the mean.

In the IB syllabus, all data is considered as a population. Thus, the mean of the data set is a parameter. Usually, the mean of the population is called μ. The calculation of μ is the same as before, $\mu = \frac{x_1 + x_2 + \cdots + x_N}{N} = \frac{\sum_{i=1}^{N} x_i}{N}$ where N is the population size.

The salaries in Table 7.8 are misleading since they contain the salary of the CEO as an employee. This salary is an outlier, and does not belong to the same category of employee salaries. If we exclude the CEO's salary, the average salary in Falsead is

$$\bar{x} = \frac{24\,000 + 2 \times 21\,000 + 5 \times 15\,600}{8} = \frac{144\,000}{8} = 18\,000$$

This illustrates an important weakness of the mean as a measure of centre: the mean is sensitive to the influence of a few extreme observations. These may be outliers, but a skewed distribution that has no outliers will also pull the mean towards its long tail. Because the mean cannot resist the influence of extreme observations, we say that it is not a **resistant measure** of centre.

A measure that is resistant does more than limit the influence of outliers. Its value does not respond strongly to changes in a few observations, no matter how large those changes may be. The mean fails this requirement because we can make the mean as large as we wish by making a large enough increase in just one observation.

Median

A second measure of central tendency is the median, which is the value in the middle position when the measurements are ordered from smallest to largest. The median of this data can only be calculated if we first sort them in ascending order.

The **median, *M*** is the midpoint of a distribution. Half the observations are smaller than the median and the other half are larger than the median. Here is a rule for finding the median:

1. Arrange all observations in order of size, from smallest to largest.
2. If the number of observations n is odd, the median M is the centre observation in the ordered list. Find the location of the median by counting $\frac{(n+1)}{2}$ observations up from the bottom of the list. For example, if $n = 5$, then $\frac{(5+1)}{2} = 3$. That is, the median is the third point. In general, if n is odd, then it can be written as $n = 2k + 1$ for some integer k, and thus $\frac{n+1}{2} = \frac{2k+1+1}{2} = k + 1$. This implies that when we arrange data in ascending order, we get something similar to the set-up below:

$$\underbrace{x_1, x_2, x_3, \cdots, x_k,}_{k} \underset{\text{Median}}{\overset{M}{\overbrace{x_{k+1}}}} \underbrace{, x_{k+2}, x_{k+3}, \cdots, x_n}_{k}$$

3. If the number of observations, n, is even, the median M is the mean of the two centre observations in the ordered list. The location of the median is again $\frac{(n+1)}{2}$ from the bottom of the list. For example, if $n = 6$, then $\frac{(6+1)}{2} = 3.5$. That is, the median is the point with position between 3 and 4. In general, if n is even, then it can be written as $n = 2k$ for some integer k, and thus $\frac{n+1}{2} = \frac{2k+1}{2} = k + \frac{1}{2}$. In this case, the median is the average of the two middle points. This implies that when we arrange data in ascending order, we get something similar to the set-up below:

$$\underbrace{x_1, x_2, x_3, \cdots, x_k,}_{k} \underset{\text{Median}}{\overset{\frac{x_k + x_{k+1}}{2}}{\overbrace{M}}} \underbrace{, x_{k+1}, x_{k+2}, \cdots, x_n}_{k}$$

Example 7.7

The following are the five closing prices of NASDAQ's stock for the first business week in October 2018: 19.96, 20.08, 20.74, 21.12, 21.04. Find the median and the mean stock price for that week.

Solution

There are five values, so the position is given by $\frac{(n+1)}{2} = \frac{(5+1)}{2} = 3$

The median is the 3rd value.

19.96 20.08 20.74 21.04 21.12

Median

$$\text{Mean} = \bar{x} = \frac{19.96 + 20.08 + 20.74 + 21.12 + 21.04}{5} = 20.59$$

Medians require little arithmetic, so they are easy to find by hand for small sets of data. Arranging even a moderate number of observations in order takes a long time, so that finding the median by hand for larger sets of data is tedious. You can use your GDC to find both the mean and the median.

```
Mean(List 1)
                 20.588
Median(List 1)
                  20.74
```

In the previous example we calculated the sample median by finding the third measurement to be in the middle position. Let us demonstrate the case of even numbers with the following.

Let us assume that you took six tests last term and your marks were, in ascending order,

$$52 \quad 63 \quad \underbrace{74 \quad 78}_{\Uparrow} \quad 80 \quad 89$$

There are two 'middle' observations here. To find the median, choose a value halfway between the two middle observations:

$$m = \frac{74 + 78}{2} = 76$$

Although both the mean and median are good measures for the centre of a distribution, the median is less sensitive to extreme values or outliers. For example, the value 52 in the six tests example is lower than all your test scores and is the only failing score you have. The median, 76, is not affected by this outlier even if it were much lower than 52. Assume, for example that your lowest score is 12 rather than 52, the median's calculation

$$12 \quad 63 \quad \underbrace{74 \quad 78}_{\Uparrow} \quad 80 \quad 89$$

still gives the same median of 76. If we were to calculate the mean of the original set, we would get

$$\bar{x} = \frac{\sum x}{6} = \frac{436}{6} \approx 73$$

While the new mean with 12 as lowest score is

$$\bar{x} = \frac{\sum x}{6} = \frac{396}{6} = 66$$

Clearly, the low outlier 'pulled' the mean towards it while leaving the median untouched. However, because the mean depends on every observation and uses all the information in the data, it is generally, wherever possible, the preferred measure of central tendency.

Mode

A third way to locate the centre of a distribution is to look for the value of x that occurs most. This measure of the centre is called the **mode**.

The **mode** is the value that occurs with the highest frequency in a data set.

Mode is a French word that means *fashion* – an item that is most popular or common.

Example 7.8

The following data give the speeds (in km h^{-1}) of eight cars that were stopped for speeding violations on a street with speed limit of 50 km h^{-1}:

67, 72, 64, 71, 69, 64, 64, 78

Find the mode.

Solution

In this data set, 64 occurs three times, and each of the remaining values occurs only once. Because 64 occurs with the highest frequency, it is the mode.

In the case of a tie for the most frequently occurring value, two modes are listed. Then the data are said to be **bimodal**. Data sets with more than two modes are referred to as **multimodal**.

In applications, the concept of mode is often used in determining sizes. As an example, manufacturers who produce cheap rubber bands might only produce them in one size in order to save on machine setup costs. In determining the one size to produce, the manufacturer would most likely produce bands in the modal size. The mode is an appropriate measure of central tendency for nominal-level data.

We cannot say for sure which of the three measures of central tendency is a better measure overall. Each of them may be better under different situations. Probably the mean is the most-used measure of central tendency, followed by the median. The mean has the advantage that its calculation includes each value of the data set. The median is a better measure when a data set includes outliers. The mode is simple to locate, but it is not of much use in many practical applications.

Relationships between the mean, median and mode

Some histograms and frequency distributions can be symmetric and others can be skewed.

Knowing the values of the mean, median, and mode can give us some idea about the shape of a frequency distribution curve.

For a symmetric histogram and frequency distribution curve with one peak, the values of the mean, median, and mode are the same, and they lie at the centre of the distribution.

For a histogram and a frequency distribution curve skewed to the right (positively skewed), the value of the mean is the largest, that of the mode is the smallest, and the value of the median lies between these two. Note that the mode always occurs at the peak point. The value of the mean is the largest in this case because it is sensitive to extreme values that occur in the right tail. These extremes pull the mean to the right.

When a histogram and a frequency distribution curve are skewed to the left (negatively skewed), the value of the mean is the smallest and that of the mode is the largest, with the value of the median lying between these two. In this case, the extreme values in the left tail pull the mean to the left.

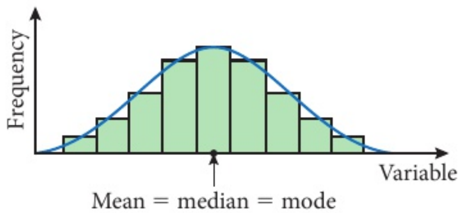


Figure 7.19 Symmetric distribution

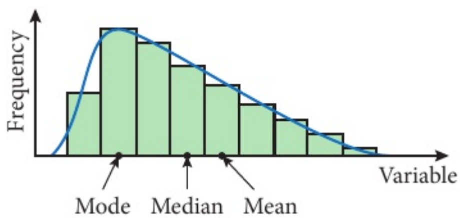


Figure 7.20 Positively skewed distribution

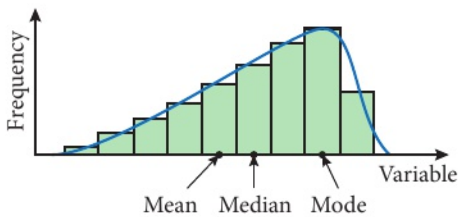


Figure 7.21 Negatively skewed distribution

Measures of variability and spread

Measures of location summarise what is typical of elements of a list, but not every element is typical. Are all the elements close to each other? Are most of the elements close to each other? What is the biggest difference between elements? On average, how far are the elements from each other? The answer lies in the measures of spread or variability.

It is possible that two data sets have the same mean, but the individual observations in one set could vary more from the mean than the observations in the second set do. It takes more than the mean alone to describe data. Measures of variability (also called measures of dispersion or spread) which include the range, the variance, the standard deviation, and interquartile range help to summarise the data.

For example, Bryan's and Jim's test scores in their mathematics class are:

Bryan: 60, 65, 70, 75, 80

Jim: 45, 55, 70, 85, 95

To compare the performance of these students, we calculate their means and find out that both have a mean grade of 70. The question is, which 70 is more typical of each student's performance?

As we can see from Figure 7.22, Bryan's 70 is more typical of his performance than Jim's because his test scores are closer to the 70 average than Jim's.

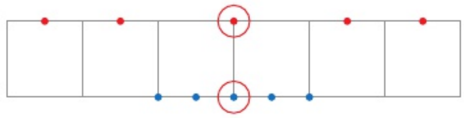


Figure 7.22 Comparison of scores

The simplest useful numerical description of a distribution consists of both a measure of centre and a measure of spread.

Range and quartiles

Finding the range for ungrouped data: range = largest value − smallest value.

The **range** is the simplest measure of dispersion to calculate. It is obtained by taking the difference between the largest and the smallest values in a data set.

The range for Jim's test scores is $95 - 45 = 50$ while Bryan's is $80 - 60 = 20$

With a range of 20 in comparison to 50, Bryan's grade of 70 is more typical for his performance than Jim's.

The range shows the full spread of the data, but it depends on only the smallest and largest observations, which may be outliers. We can improve our description of dispersion by also looking at the spread of the middle half of the data. The **quartiles** mark out the middle half of the data.

Quartiles are the summary measures that divide ranked data sets into four equal parts. Three measures will divide any data set into four equal parts. These three measures are the **first quartile** (Q_1), the **second quartile** (Q_2), and the **third quartile** (Q_3). The data should be ranked in increasing order before the quartiles are determined. The quartiles are defined as follows.

Quartiles are three summary measures that divide a ranked data set into four equal parts. The second quartile is the same as the median of a data set. The first quartile is the value of the middle term among the observations that are less than the median, and the third quartile is the value of the middle term among the observations that are greater than the median.

Approximately 25% of the values in a ranked data set are less than Q_1 and about 75% are greater than Q_1. The second quartile, Q_2, divides a ranked data set into two equal parts; hence, the second quartile and the median are the same. Approximately 75% of the data values are less than Q_3 and about 25% are greater than Q_3.

Bryan's grades:	60		65	70	75		80	$\Rightarrow$ IQR $= 77.5 - 62.5 = 15$
		$\uparrow$ Q_1 62.5		Median Q_2		$\uparrow$ Q_3 77.5		
Jim's grades:	45		55	70	85		95	$\Rightarrow$ IQR $= 90 - 50 = 40$
		$\uparrow$ Q_1 50		Median Q_2		$\uparrow$ Q_3 90		

25%	25%	25%	25%
	Q_1	Q_2	Q_3

Note that the parts do not have to have the same length, but each should contain 25% of the data points.

The difference between the third and the first quartiles gives the **interquartile range**; that is, IQR = interquartile range $= Q_3 - Q_1$
IQR describes the middle half of the data.

Example 7.9

The table gives the 2017 profits of the top 10 companies in the world in US$ billions.

Company	Profit
Walmart, US	486
State Grid, China	315
Sinopec, China	268
China National Petroleum	263
Toyota, Japan	255
Shell, UK-NL	240
VW, Germany	240
Berkshire Hathaway, US	224
Apple, US	216
GE, US	205

(a) (i) Find the values of the three quartiles.
(ii) Where does the 2017 profit of Apple fall in relation to these quartiles?
(b) Find the interquartile range.

The value of $Q_1 = \$224$ billion indicates that 25% of the companies in this sample had profits less than \$224 billion and 75% of the companies had profits higher than \$224 billion. Similarly, we can state that half of these companies had profits less than \$247.5 billion and the other half had profits greater than \$247.5 billion since the second quartile is \$247.5 billion. The value of $Q_3 = \$268$ billion indicates that 75% of the companies had profits less than \$268 billion and 25% had profits greater than this value.

Solution

(a) (i) First, we rank the given data in increasing order. Then we calculate the three quartiles.

$$\overbrace{205 \quad 216 \quad \underset{Q_1}{224} \quad 240 \quad 240}^{\text{Values less than the median}} \qquad \underset{\underbrace{\text{Median} = Q_2}_{\frac{240+255}{2}}}{247.5} \qquad \overbrace{255 \quad 263 \quad \underset{Q_3}{268} \quad 315 \quad 486}^{\text{Values greater than the median}}$$

The value of Q_2, which is also the median, is given by the value of the middle term in the ranked data set. For the data of this example, this value is the average of the 5th and 6th terms. Consequently, Q_2 is \$247.5 billion.

The value of Q_1 is given by the value of the middle term of the five values that fall below the median (or Q_2). So, Q_1 is \$224 billion.

The value of Q_3 is given by the value of the middle term of the five values that fall above the median. For the data of this example, Q_3 is \$268 billion.

(ii) The profit of Apple is \$216 billion, which is less than Q_1. So the profit of Apple is in the bottom 25% of the profits for 2017.

(b) The interquartile range is given by the difference between the values of the third and the first quartiles. Thus,

$$\text{IQR} = \text{interquartile range} = Q_3 - Q_1 = 268 - 224 = \$44 \text{ billion}$$

The first quartile Q_1 is the median of the observations whose position in the ordered list is to the left of the location of the overall median.

The third quartile Q_3 is the median of the observations whose position in the ordered list is to the right of the location of the overall median.

You can calculate the values in Example 7.9 by entering the data into a GDC as a list and then finding the descriptive statistics for them.

```
1-Variable
 x̄=271.2
 Σx=2712
 Σx²=795496
 σx=77.4606997
 Sx=81.6507467
↓n=10
```

```
1-Variable
↑MinX=205
 Q1=224
 Med=247.5
 Q3=268
 MaxX=486
↓Mod=240
```

Figure 7.23 Calculating the values in Example 7.9 on a GDC

Five-number summary – box plot

To get a quick summary of both centre and spread for a data set, combine all five numbers.

The **five-number summary** of a set of observations consists of the smallest observation, the first quartile, the median, the third quartile, and the largest observation, written in order from smallest to largest. In symbols, the five-number summary is:

$$\text{minimum} < Q_1 < M < Q_3 < \text{maximum}$$

As you have seen in Example 7.9, these five numbers offer a reasonably complete description of centre and spread.

The five-number summary leads to another visual representation of a distribution, the **box plot**. Figure 7.24 shows box plots for both Bryan's and Jim's grades.

The diagram shows clearly that both samples have the same centre, and both are symmetric around the centre, however, Jim's grades show more spread around the median.

We can put the information from the five-number summary together in one graphical display called a **box plot**, or a **box-and-whisker** plot.
It has upper and lower fences which are usually at 1.5 times the interquartile range from the upper and lower quartiles, respectively.
An **outlier** is a value which is lower than the lower fence or higher than the upper fence.

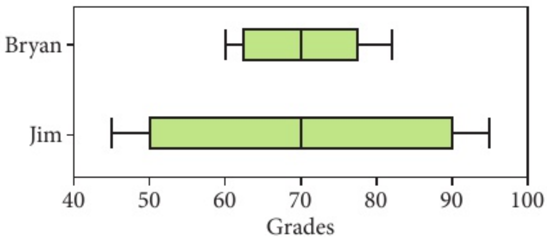


Figure 7.24 Box plot of grades

An outlier is an unusual observation. It lies at an abnormal distance from the rest of the data. There is no unique way of describing what an outlier is. A common practice is to consider any observation that is further than 1.5 IQR from the first quartile or the third quartile an outlier. Outliers are important in statistical analysis. Outliers may contain important information not shared with the rest of the data. Statisticians look very carefully at outliers because of their influence on the shapes of distributions and their effect on the values of the other statistics.

Example 7.10

A consumer agency is interested in studying weekly food and living expenses of college students. A survey of 80 students yielded the following expenses to the nearest euro.

38	50	55	60	46	51	58	64	50	49	48	65	58	61	65	53
39	51	56	61	48	53	59	65	54	54	54	59	65	66	47	49
40	51	56	62	47	55	60	63	60	59	59	50	46	45	54	47
41	52	57	64	50	53	58	67	67	66	65	58	54	52	55	52
44	52	57	64	51	55	61	68	67	54	55	48	57	57	66	66

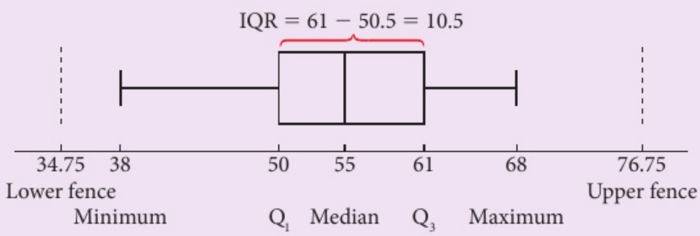
(a) Find the five-number summary.

(b) Draw a box plot.

Solution

(a) Using a GDC, we find the five numbers:
Minimum $= 38$, $Q_1 = 50.5$, Median $= 55$, $Q_3 = 61$, and Maximum $= 68$

(b) Draw an axis spanning the range of the data; mark the numbers corresponding to the median, minimum, maximum, and the lower and upper quartiles.



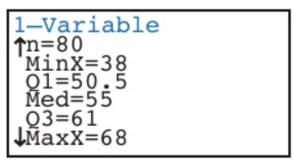


Figure 7.25 GDC screen for the solution to Example 7.10

Draw a rectangle with lower end at Q_1 and upper end at Q_3, as shown.

To help us consider outliers, mark the points corresponding to lower and upper fences. Mark them with a dashed line since they are not part of the box. The fences are constructed at the following positions:

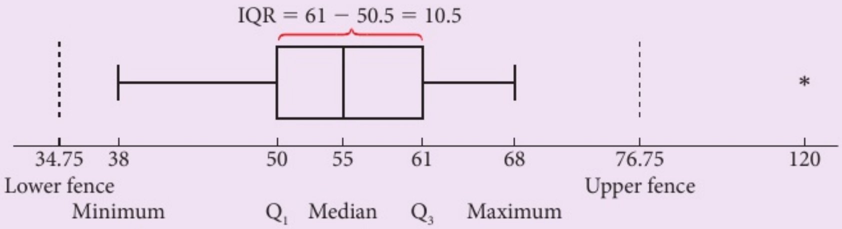
Lower fence: $Q_1 - 1.5 \times \text{IQR}$ (in this case: $50.5 - 1.5\,(10.5) = 34.75$)
Upper fence: $Q_3 + 1.5 \times \text{IQR}$ (in this case: $61 + 1.5\,(10.5) = 76.75$)

Any point beyond the lower or upper fence is considered an **outlier**.

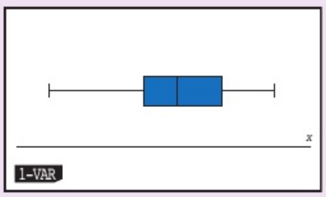
Extend horizontal lines called 'whiskers' from the ends of the box to the smallest and largest observations that are not outliers. In this case these are 38 and 68.

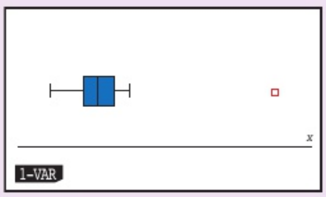
If the data has an outlier, mark it with an asterisk (∗) on the graph.

To demonstrate this point, consider what would happen if our maximum was 120 and not 68. Since the whisker is 76.75 and $120 > 76.75$, then 120 is an outlier as is clear by the box plots.



Here is the output of a GDC for both cases.





As you see, the box contains the middle 50% of the data. The width of the box is nothing but the IQR! Now we know that the middle 50% of the students' expenditure is €10.50

The box plot is not the only graph we use to explore data. You can also use the cumulative frequency polygon or ogive. Here is the ogive for the expenses data.

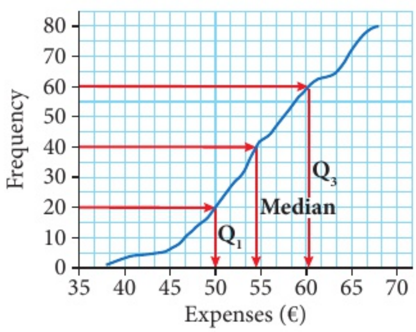


Figure 7.26 Ogive for expense data

Note how we locate the first quartile. Since there are 80 observations, the first quartile is approximately at the $\frac{n+1}{4} = \frac{81}{4} \approx 20$th position, which appears to be around 50.

The median is at the $\frac{n+1}{2} = \frac{81}{2} \approx 40$th position, i.e., approximately at 55.

Similarly, the third quartile is at $\frac{3(n+1)}{4} = \frac{243}{4} \approx 61$st, which happens here at approximately 61.

Example 7.11

The data below shows the heart rates of randomly chosen Females and Males taken from a large group of college students:

Females: 78, 70, 52, 55, 68, 60, 66, 90, 71, 87

Males: 70, 64, 68, 80, 77, 71, 100, 89, 40, 55

(a) Find the five-number summaries.

(b) Find the range and IQR for each.

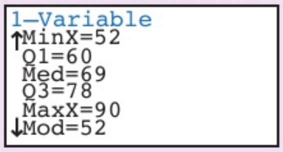
(c) Draw box plots of both data sets.

(d) Compare the two sets.

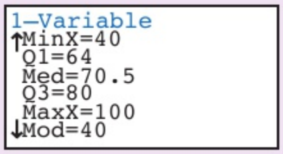
Solution

(a) Using a GDC here are the summaries:

Females:
Minimum $= 52$, $Q_1 = 60$
Median $= 69$, $Q_3 = 78$, and Maximum $= 90$



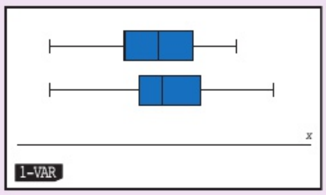
Males:
Minimum $= 40$, $Q_1 = 64$
Median $= 70.5$, $Q_3 = 80$
and Maximum $= 100$



(b) Females' range $= 90 - 52 = 38$
IQR $= 78 - 60 = 18$

Males' range $= 100 - 40 = 60$
IQR $= 80 - 64 = 16$

(c) Box plots are shown in the screenshot.



(d) The Females' data have a smaller range, a lower median rate, but a larger IQR than the non-smokers. Neither data set appears to have outliers.

Variance and standard deviation

The range, like the mean, has the disadvantage of being influenced by outliers. Look at the example of students' expenses. In the original data, the range was $68 - 38 = 30$, while when we replaced 68 by 120, the range became $120 - 38 = 82$. Just one extreme value almost tripled the range. Moreover, another disadvantage of using the range as a measure of dispersion is that its calculation is based on two values only: the largest and the smallest. All other values in a data set are ignored when calculating the range. Thus, it is not a very satisfactory measure of dispersion.

The **standard deviation** is the most frequently used measure of dispersion. The value of the standard deviation tells us how closely the values of a data set are clustered around the mean. In general, a lower value of the standard deviation for a data set indicates that the values of that data set are spread over a relatively smaller range around the mean, Conversely, a higher value of standard deviation indicates that the values are spread out more around the mean.

The standard deviation measures spread by looking at how far the observations are from their mean.

The standard deviation is obtained by taking the positive square root of the **variance**. The variance calculated for population data is denoted by σ^2 (read as sigma squared), and the variance calculated for sample data is denoted by s^2. Consequently, the standard deviation calculated for population data is denoted by σ and the standard deviation calculated for sample data is denoted by $\boldsymbol{s}$.

The basic formulae that are used to calculate the variance are:

$$\sigma^2 = \frac{\sum(x-\mu)^2}{N} \text{ and } s^2 = \frac{\sum(x-\bar{x})^2}{n}$$

The quantity $x - \mu$ or $x - \bar{x}$ is called the deviation of the x value from the mean.
The sum of the deviations of the x values from the mean is always zero; that is
$\sum(x-\mu) = 0$ and $\sum(x-\bar{x}) = 0$
We will use the sample notation for the mean and standard deviation.

For example, in Bryan's and Jim's test scores in their mathematics class:

Bryan: 60, 65, 70, 75, 80

Jim: 45, 55, 70, 85, 95

In both cases, $\bar{x} = 70$

Bryan: $\sum(x-\bar{x}) = (60-70) + (65-70) + (70-70) + (75-70) + (80-70)$
$= -10 - 5 + 0 + 5 + 10 = 0$

Jim: $\sum(x-\bar{x}) = (45-70) + (55-70) + (70-70) + (85-70) + (95-70)$
$= -25 - 15 + 0 + 15 + 25 = 0$

For this reason, we square the deviations to calculate the variance and standard deviation.

Thus, for Bryan,

$$s^2 = \frac{\sum(x-\bar{x})^2}{5}$$

$$= \frac{(60-70)^2 + (65-70)^2 + (70-70)^2 + (75-70)^2 + (80-70)^2}{5}$$

$$= \frac{250}{5} = 50$$

and $s = \sqrt{50} \approx 7.07$

For Jim,

$$s^2 = \frac{\sum(x-\bar{x})^2}{5}$$

$$= \frac{(45-70)^2 + (55-70)^2 + (70-70)^2 + (85-70)^2 + (95-70)^2}{5}$$

$$= \frac{1700}{5} = 340$$

and $s = \sqrt{340} \approx 18.44$

We can observe from this that the spread in Jim's scores is more than double that of Bryan.

You can use your GDC to calculate these values, as shown in Figure 7.27.

If you are interested in the standard deviation only, most GDCs have a command for it.

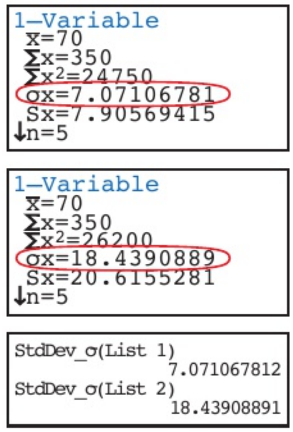


Figure 7.27 Using a GDC to calculate standard deviation

Very important
In all GDCs, the standard deviation we need so far in this course is σx and not Sx.

More important than the details of hand calculation are the properties that determine the usefulness of the standard deviation:

- s measures spread about the mean and should be used only when the mean is chosen as the measure of centre.
- $s = 0$ only when there is no spread. This happens only when all observations have the same value. Otherwise, s is greater than zero. As the observations become more spread out about their mean, s gets larger.
- s has the same units of measurement as the original observations. For example, if you measure wages in dollars per hour, s is also in dollars per hour.
- Like the mean $\bar{x}$, s is not resistant. Strong skewness or a few outliers can greatly increase s.

The five-number summary is usually better than the mean and standard deviation for describing a skewed distribution or a distribution with extreme outliers. Use $\bar{x}$ and s only for reasonably symmetric distributions that are free of outliers.

Measures of centre and spread for grouped data

The calculation of the mean and variance for grouped data is essentially the same as for raw data. The difference lies in the use of frequencies to save typing (writing) all numbers. Table 7.9 shows a comparison.

Statistic	Raw data	Grouped data	Grouped data with intervals
$\bar{x}$	$\bar{x} = \dfrac{\sum_{\text{all } x} x}{n}$	$\bar{x} = \dfrac{\sum_{\text{all } x} x_i \times f(x_i)}{\sum f(x_i)}$	$\bar{x} = \dfrac{\sum_{\text{all } m} m_i \times f(m_i)}{\sum f(m_i)}$ First find the midpoint of each class and then multiply the midpoints by the frequencies of the corresponding classes. The sum of these products gives an approximation for the sum of all values. To find the value of the mean, divide this sum by the total number of observations.
s^2	$s^2 = \dfrac{\sum_{\text{all } x} (x_i - \bar{x})^2}{n}$	$s^2 = \dfrac{\sum_{\text{all } x} (x_i - \bar{x})^2 \times f(x_i)}{\sum f(x_i)}$	$s^2 = \dfrac{\sum_{\text{all } m} (m_i - \bar{x})^2 \times f(m_i)}{\sum f(m_i)}$

Table 7.9 Grouped data calculations

When we estimate the mean of a data set from its frequency table, answers may differ from the real mean. In this example, the real mean is 55.475. The reason is that we are replacing the real values with estimates. For example, in the interval $35 \leqslant x < 45$, we replace the real values with 40. This way, we are assuming that all numbers in this interval are equal to 40.

For example, when we group the students' expense data into a frequency table, we get the data shown in Table 7.10.

Expenses	Midpoint m	Number of students
$35 \leqslant x < 45$	40	5
$45 \leqslant x < 55$	50	32
$55 \leqslant x < 65$	60	30
$65 \leqslant x < 75$	70	13

Table 7.10 Frequency table for grouped expense data

$$\bar{x} \approx \frac{\sum_{\text{all } m} m_i \times f(m_i)}{n}$$

$$= \frac{40 \times 5 + 50 \times 32 + 60 \times 30 + 70 \times 13}{80} = 56.375$$

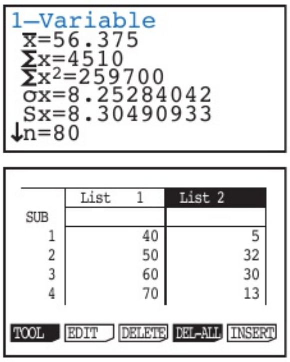


Figure 7.28 You should organise your data into two lists on your GDC

Your GDC will give you the result in the same way as in the non-grouped data. The difference here is that you need to organise your data in two lists, one for the class midpoints and one for the frequency.

Effect of constant changes to the original data on centre and spread measures

Consider Bryan's test scores again. The teacher decides to add 5 marks to each score at the end of the term. What are the new measures for Bryan's scores?

Bryan's new scores are: 65, 70, 75, 80, 85. Table 7.11 gives both the old and new measurements.

Measure	Old	New
Median	70	The middle observation is 75. Median is 75.
Mean	70	$\bar{x} = \frac{65 + 70 + 75 + 80 + 85}{5} = 75$
Quartiles	$Q_1 = 62.5, Q_3 = 77.5$	$Q_1 = 67.5, Q_3 = 82.5$
IQR	15	$82.5 - 67.5 = 15$
St. deviation	7.07	7.07

Table 7.11 Effect of adding 5 to each score

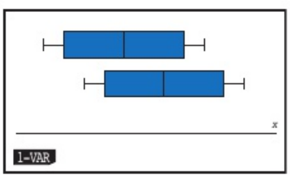


Figure 7.29 New box plot shifted to the right by 5 units

Note that 5 is added to each of the median, mean, and quartiles, but the IQR and standard deviation are not affected. This is so because, as you observe from the box plots, we moved the data by 5 units to the right, keeping the spread as it was.

Example 7.12

Five students took a trial test on paper 1 of an IB exam.

Their scores, out of 120 are 111, 96, 87, 72, 60

(a) Find the median, IQR, mean and standard deviation of the scores.

(b) The scores must be scaled to be out of a total of 40 instead.
Find the new median, IQR, mean and standard deviation of the scores.

Solution

Measure	Old	New
Median	87	29
Mean	85.2	28.4
Quartiles	$Q_1 = 66, Q_3 = 103.5$	$Q_1 = 22, Q_3 = 34.5$
IQR	37.5	12.5
St. deviation	17.86	5.95

Note that the new measures are equal to one third of the old ones.

When data values are multiplied by a constant k, then all centre and spread statistics are multiplied by k.

Example 7.13

Instruments that measure blood sugar level (glucose) may use one of two systems of measurement: mmol/L or mg/dL.
Every mmol/L is equivalent to 18 mg/dL.

Tim keeps track of his blood sugar level. Last month's readings are as follows.

The average reading was 5.2 mmol/L, the standard deviation was 2.3 mmol/L, and the range was 3.2 mmol/L.

Find Tim's readings in mg/dL.

Solution

Since every 1 mmol/L = 18 mg/dL, then Tim's readings in mg/dL are:

Average $= 5.2 \times 18 = 93.6$ mg/dL

Standard deviation $= 2.3 \times 18 = 41.4$ mg/dL

Range $= 3.2 \times 18 = 57.6$ mg/dL

Exercise 7.3

1. The number of visitors per week (thousands) for a book exhibition were

   9, 7, 8, 11, 9, 6, 10, 8, 12, 6, 8, 13, 7, 9, 10, 9, 10, 11, 12, 8, 7, 13, 10, 7, 7

   The data has been organised in the frequency table.

Number of visitors	6	7	8	9	10	11	12	13
Frequency	2	5	n	4	4	2	2	2

   (a) Write down the value of n.

   (b) Calculate the mean number and standard deviation of visitors per week, from the raw data and from the table.

   (c) What percentage of the weeks had more than 10 000 visitors?

   (d) What is the modal number of visitors?

2. The following data give the number of shoplifters apprehended during 9 weeks preceding Christmas 2017 at a large department store.

   1, 1, 8, 3, 1, 1, 7, 26, 51

   (a) Find the mean, median, and mode for these data.

   (b) Calculate the deviations of the data values from the mean. What is the sum of these deviations?

   (c) Calculate the range, variance, and standard deviation.

   (d) Draw a box plot. Can you explain why no whisker appears on the lower side?

   (e) Decide if there are any outliers. Explain.

3. The following data give the numbers of car thefts that occurred in a large city in the past 12 months.

   60, 30, 70, 10, 140, 30, 80, 70, 20, 60, 90, 10

   (a) Find the mean, median, and mode for these data.

   (b) Calculate the range and standard deviation.

   (c) Draw a box plot.

   (d) Decide if there are any outliers. Explain.

4. The pulse rates of 15 patients chosen at random from visitors of a local clinic are 72, 80, 67, 68, 80, 68, 80, 56, 76, 68, 71, 76, 60, 79, 71

   (a) Calculate the mean and standard deviation of the pulse rate of the patients at the clinic.

   (b) Draw a box plot of the data and indicate the values of the different parts of the box.

   (c) Check if there are any outliers.

5. The number of passengers on 50 flights from Washington to London on a commercial airline is given in the table.

165	173	158	171	177	156	178	210	160	164
141	127	119	146	147	155	187	162	185	125
163	179	187	174	166	174	139	138	153	142
153	163	185	149	154	154	180	117	168	182
130	182	209	126	159	150	143	198	189	218

(a) Calculate the mean and standard deviation of the number of passengers on this airline between the two cities.

(b) Develop a cumulative frequency graph. Estimate the median, first and third quartiles. Draw a box plot.

(c) Find the IQR and use it to check whether there are any outliers.

6. The table shows the frequency distribution of the daily commuting times (in minutes) from home to work for all 50 employees of a company.

Daily commuting time (minutes)	Number of employees
$0 \leqslant t < 10$	8
$10 \leqslant t < 20$	18
$20 \leqslant t < 30$	12
$30 \leqslant t < 40$	8
$40 \leqslant t < 50$	4

(a) Set up a cumulative frequency table and graph.

(b) Estimate the median, first and third quartiles as well as IQR.

(c) Find the mean and standard deviation.

7. The table gives the frequency distribution of the amounts of telephone bills for October 2018 for a sample of 50 families.

Amount of telephone bill (€)	Number of families
$40 \leqslant m < 70$	9
$70 \leqslant m < 100$	11
$100 \leqslant m < 130$	16
$130 \leqslant m < 160$	10
$160 \leqslant m < 190$	4

(a) Find the average value of a telephone bill for families in this community.

(b) Find the standard deviation.

(c) The amount of every bill contains a €24 flat monthly rate plus the cost of telephone calls. Find the average cost of calls and the standard deviation.

8. Every year, the faculty of a major university's business school chooses 10 students from the current graduating class that they feel will be most likely to succeed. Five years later, they evaluate their choices by looking at the annual incomes of the group of 10. The following is the list of the annual incomes of the class of 2007 (thousands of US$):

59, 68, 84, 78, 107, 382, 56, 74, 97, 60

(a) Find the mean, median, standard deviation, and IQR.

(b) Does this data set contain any outlier? If yes, ignore the outlier and recalculate the statistics. Which of these measures changes by a greater amount? Explain.

(c) Which measures provide a better summary for these data? Explain.

9. Students in one section of an SL maths class scored the following marks on a trial exam paper:

5, 5, 4, 6, 3, 7, 7, 3, x

(a) The section's average score is 5. What is the value of x?

(b) Find the median and standard deviation of this group.

(c) One student was absent on the day of the trial exam. What is the minimum score required of that student for the average score of the whole group to be 6?

(d) The student from part **(c)** received a mark of 5. Another section of the class has 12 students, and their average score was 4.5. What is the average of both sections?

10. For restaurants, the time customers linger over coffee and dessert negatively affects profit. To learn more about this variable, a sample of 14 days was observed, and the average time spent by different groups was recorded (to the nearest minute):

26, 21, 28, 28, 56, 45, 40, 32, 32, 29, 30, 27, 20, 25

(a) Find a detailed summary of the data.

(b) What do the results tell you about the amount of time spent in this restaurant?

11. The following data represent the ages in years of a sample of 25 jockeys from a local race track:

31, 43, 56, 23, 49, 42, 33, 61, 44, 28, 48, 38, 44,
35, 40, 64, 52, 42, 47, 39, 53, 27, 36, 35, 20

(a) Find the 5-number summary, and IQR.

(b) Construct a box plot for the ages and identify any extreme values. What does the box plot tell you about the distribution of the data?

(c) Construct a relative frequency distribution for the data, using five class intervals and the value 20 as the lower limit of the first class.

(d) Construct a relative frequency histogram for the data. What does the histogram tell you about the distribution of the data?

Chapter 7 practice questions

1. The temperatures in °C, at midday in a German city during summer of 2017, were measured for eight days and the results are recorded below.

$21, 12, 15, 12, 24, T, 30, 24$

The mean temperature was found to be 21°C.

(a) Find the value of T.

(b) Write down the mode.

(c) Find the median.

2. The age (in months) when a child starts to walk is observed for a random sample of children from a town in France. The results are:

$14.3, 11.6, 12.2, 14.0, 20.4, 13.4, 12.9, 11.7, 13.1$

(a) **(i)** Find the mean of the ages of these children.

(ii) Find the standard deviation of the ages of these children.

(b) Find the median age.

3. The following data are listed in ascending order: $2, b, 3, a, 6, 9, 10, 12$

The mean is 6 and the median is 5. Find:

(a) the value of a

(b) the value of b.

4. The table shows the frequency distribution of the number of dental fillings for a group of 50 children.

Number of fillings	0	1	2	3	4	5
Frequency	4	3	8	x	4	1

(a) Find the value of x.

(b) Find:

(i) the mean number of fillings

(ii) the median number of fillings

(iii) the standard deviation of the number of fillings.

(c) The first row in the table was entered by mistake. It should have been 2, 3, 4, 5, 6, and 7 instead.

Without going through the calculations required for part **(b)**, find the mean, median and standard deviation.

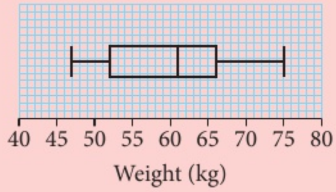
5. During the first 240 games played in 2017 in a National football league the number of goals scored are given in the table.

Number of goals	0	1	2	3	4	5
Number of games	48	66	57	51	3	15

(a) Find the mean number of goals scored per game.

(b) Find the median number of goals scored per game.

6. The masses, in kg, of 80 adult males were collected and are summarised in the box-and-whisker plot.



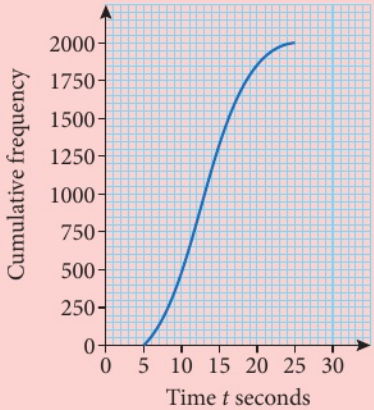
(a) Write down the median mass of the males.

(b) Calculate the interquartile range.

(c) Estimate the number of males who have a mass between 61 kg and 66 kg.

(d) Estimate the mean mass of the lightest 40 males.

7. The diagram shows the cumulative frequency graph for the time t taken to perform a certain task by 2000 men.



(a) Use the diagram to estimate:

(i) the median time

(ii) the upper quartile and the lower quartile

(iii) the interquartile range.

(b) Find the number of men who take more than 11 seconds to perform the task.

(c) 55% of the men took less than p seconds to perform the task. Find p.

Time	Frequency
$5 \leqslant t < 10$	500
$10 \leqslant t < 15$	850
$15 \leqslant t < 20$	a
$20 \leqslant t < 25$	b

(d) Write down the value of:

(i) a **(ii)** b

(e) Find an estimate of:

(i) the mean time

(ii) the standard deviation of the time.

Everyone who performs the task in less than one standard deviation below the mean will receive a bonus. Pedro takes 9.5 seconds to perform the task.

(f) Does Pedro receive the bonus? Justify your answer.

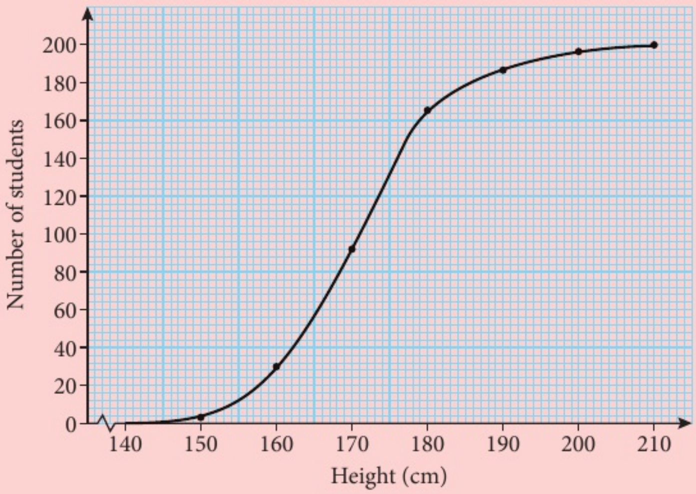
8. The heights of 200 students are recorded in the table.

Height (h) in cm	**Frequency**
$140 \leqslant h < 150$	2
$150 \leqslant h < 160$	28
$160 \leqslant h < 170$	63
$170 \leqslant h < 180$	74
$180 \leqslant h < 190$	20
$190 \leqslant h < 200$	11
$200 \leqslant h < 210$	2

(a) Write down the modal group.

(b) Calculate an estimate of the mean and standard deviation of the heights.

The cumulative frequency curve for this data is shown.

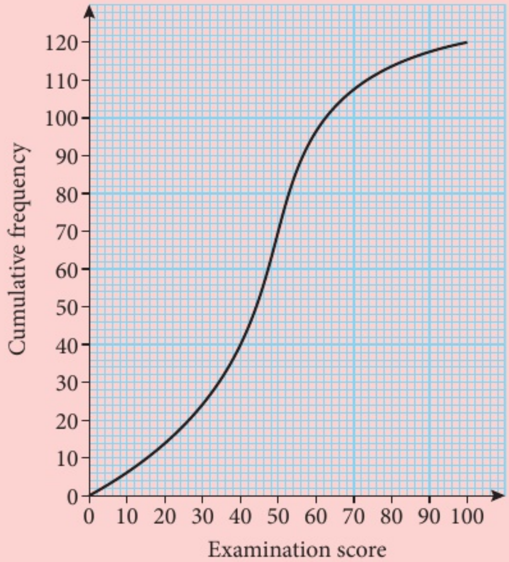


(c) Write down the median height.

(d) The upper quartile is 177.3 cm. Calculate the interquartile range.

(e) Find the percentage of students with heights less than 165 cm.

9. 120 Mathematics students in a school sat an examination. Their scores (given as a percentage) were summarised on a cumulative frequency diagram, as shown.



(a) Copy and complete the grouped frequency table for the students.

Examination score x (%)	$0 \leqslant x < 20$	$20 \leqslant x < 40$	$40 \leqslant x < 60$	$60 \leqslant x < 80$	$80 \leqslant x < 100$
Frequency	14	26			

(b) Write down the mid-interval value of the $40 \leqslant x < 60$ interval.

(c) Calculate an estimate of the mean examination score of the students.

8 Probability

8 Probability

Learning objectives

By the end of this chapter, you should be familiar with…

- concepts of trial, outcome, equally likely outcomes, sample space (U) and event
- the probability of an event A as $\mathrm{P}(A) = \dfrac{n(A)}{n(U)}$
- complementary events A and A' (not A), and performing calculations using the identity $\mathrm{P}(A) + \mathrm{P}(A') = 1$
- working with combined events and using the formula: $\mathrm{P}(A \cup B) = \mathrm{P}(A) + \mathrm{P}(B) - \mathrm{P}(A \cap B)$
- $\mathrm{P}(A \cap B) = 0$ for mutually exclusive events
- conditional probability: $\mathrm{P}(A|B) = \dfrac{\mathrm{P}(A \cap B)}{\mathrm{P}(B)}$
- calculating probabilities of independent events: $\mathrm{P}(A|B) = \mathrm{P}(A) = \mathrm{P}(A|B')$
- using Venn diagrams, tree diagrams and tables of outcomes to solve problems.

Uncertainty plays an important role in our daily lives and activities as well as in business, science, and almost all fields. Investors cannot be sure which stocks will deliver the best growth over the next year. Engineers try to reduce the likelihood that a machine will break down. Publishers cannot be sure of how sales of a new book are going to be over its life cycle.

Here are some other examples of the role of uncertainty in our lives:

- In the last 5 minutes of a world cup match, the manager of the German team replaces a defence player with a forward player.
- Tim is standing at a pedestrian traffic light, facing the 'Don't Walk' light. After looking left and right, he crosses the street.
- John is a 75 year old retired professor with high blood pressure and an incident of a minor heart attack. When he applies for life insurance, he is turned down as an unacceptable risk.

In each of the cases described, chance is involved. The team manager selected a forward who was more likely to get a much-needed goal. Tim is trying to reduce the chance that something undesirable would happen to him. John's insurance company felt that his age and health conditions greatly reduce the likelihood of his surviving for many more years.

How can we use sample data to draw conclusions about the populations from which we drew our samples? The techniques we use in drawing conclusions are part of inferential statistics, which uses **probability** as one of its tools.

The variables we discussed in Chapter 7 can now be redefined as **random variables**, whose values depend on the chance selection of the elements in the sample. Using probability as a tool, you will be able to create **probability distributions** that serve as models for random variables. You can then describe these using a mean and a standard deviation as you did in Chapter 7.

8.1 Randomness and probability

The reasoning in statistics rests on asking, 'How often would this method give a correct answer if I used it very many times?' When we produce data by random sampling or by experiments, the laws of probability enable us to answer the question 'What would happen if we did this many times?'

Probability is the study of randomness and uncertainty.

Flipping a coin, rolling a dice and observing the number on the top surface, counting cars at a traffic light when it turns green, measuring daily rainfall in a certain area, etc., are all examples of **experiments**. Every flip of the coin is a **trial**.

What does 'random' mean? In ordinary speech, we use 'random' to denote things that are unpredictable. Events that are random are not perfectly predictable, but they have long-term regularities that we can describe and quantify using probability. Take, for example, the flipping of an unbiased coin and observing the number of heads that appear. This is random behaviour.

When you flip the coin, there are only two outcomes, heads or tails. Figure 8.1 shows the results of the first 50 trials of an experiment where the coin was flipped 5000 times. Two experiments are shown. The red graph, with points marked by crosses, shows the result of an experiment where the first flip was a head, followed by a tail, making the proportion of heads to be 0.5. The third and fourth flips were also tails. So, the proportion of heads is 0.33, then 0.25. On the other hand, the other experiment, shown in green with points marked by dots, starts with a series of tails (0.00 proportion of heads), then a head, which raises the proportion to 0.2, etc.

An **experiment** is the process by which an observation (or measurement) is obtained. A single attempt or realisation of the experiment is a **trial**.

The proportion of heads is quite variable at first. However, in the long run, and as the number of flips increases, the proportion of heads stabilises at around 0.5. We say that 0.5 is the **probability** of a head.

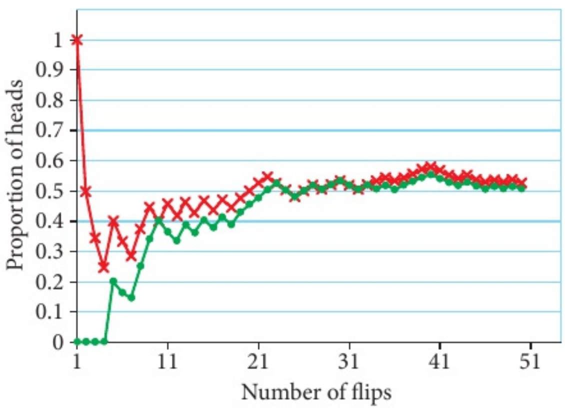


Figure 8.1 Two experiments for flipping a coin

It is important that you know that the proportion of heads in a small number of flips can be far from the probability. Probability describes only what happens in the long run. How a fair coin lands when it is flipped is an example of a random event. One cannot predict perfectly whether the coin will land heads or tails. However, in repeated flips, the fraction of times the coin lands heads will tend to settle down to a limit of 50%. The outcome of an individual flip is not perfectly predictable, but the long-term average behaviour is predictable. Thus, it is reasonable to consider the outcome of flipping a fair coin to be random.

We call an experiment **random** if individual outcomes are uncertain but there is nevertheless a regular distribution of outcomes in a large number of repetitions.

The **probability** of any outcome of a random experiment is the proportion of times the outcome would occur in a very long series of repetitions.

8 Probability

Basic definitions

Data are obtained by observing either uncontrolled events in nature or controlled situations in a laboratory. We use the term **experiment** to describe either method of data collection.

A description of a random phenomenon in the language of mathematics is called a **probability model**. For example, when we flip a coin, we cannot know the outcome in advance. What *do* we know? We are willing to say that the outcome will be either heads or tails. Because the coin appears to be balanced, we believe that each of these outcomes has probability 0.50. The description of coin flipping has two parts:

- A list of possible outcomes
- A probability for each outcome.

This two-part description is the starting point for a probability model.

The notation for sample space could also be *S* or any other letter. For example, for one flip of a coin, the sample space is $U = \{\text{heads, tails}\}$, or simply $\{\text{h, t}\}$

The **sample space** U of a random experiment (or phenomenon) is the set of all possible outcomes.

Example 8.1

Flip a coin twice (or two coins once each) and record the results. Represent the sample space

(a) in list form (b) in table form (c) as a tree diagram.

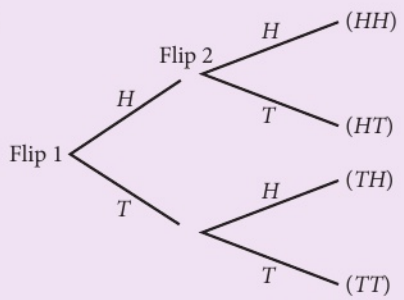
Solution

(a) $U = \{\text{HH, HT, TH, TT}\}$

(b)

Flip 1	Flip 2
H	H
H	T
T	H
T	T

(c)



If you flip a coin twice (or two coins once) and count the number of heads showing, the sample space is: $S = \{0, 1, 2\}$

Set theory provides a foundation for all of mathematics. The language of probability is much the same as the language of set theory. Logical statements can be interpreted as statements about sets.

A **simple event** is the outcome we observe in a single repetition (**trial**) of the experiment.

An **event** is an **outcome** or a **set of outcomes** of a random experiment.

We can also look at an event as a **subset** of the sample space or as a collection of simple events.

Example 8.2

When rolling a standard six-sided dice, what are the sets of event A: 'observe an odd number', and event B: 'observe a number less than 5'.

Solution

Event A is the set $\{1, 3, 5\}$; event B is the set $\{1, 2, 3, 4\}$.

Sometimes it helps to visualise an experiment using a **Venn diagram**. Figure 8.2 shows the outcomes of the dice rolling experiment from Example 8.2.

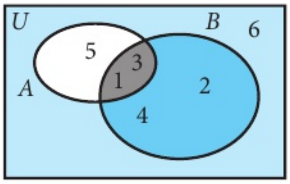


Figure 8.2 Sample space and events of rolling a dice

In general, in this book, we will use a rectangle to represent the sample space and closed curves to represent events.

To understand these definitions more clearly, let's look at the following example.

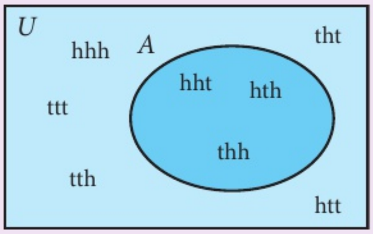
Example 8.3

Flip a coin three times and record the results. Show the event 'observing exactly two heads' as a Venn diagram and a table.

Solution

The sample space is made up of 8 possible outcomes such as: hhh, hht, tht, etc.

Observing exactly two heads is an event with three elements: {hht, hth, thh}



Flip 1	Flip 2	Flip 3
H	H	H
H	H	T
H	T	H
H	T	T
T	H	H
T	H	T
T	T	H
T	T	T

Some useful set theory results

Set operations have a number of properties, which are basic consequences of the definitions. Some examples are:

$A \cup B = B \cup A$
$(A')' = A$
$A \cap U = A$
$A \cup U = U$
$A \cap A' = \varnothing$
$A \cup A' = U$

U is the sample space and $\varnothing$ is the empty set

Two valuable properties are known as De Morgan's laws, which state that

$(A \cup B)' = A' \cap B'$

$(A \cap B)' = A' \cup B'$

Tree diagrams, tables and grids

In an experiment to check the blood types of patients, the experiment can be thought of as a two-stage experiment: first we identify the type of the blood and then we classify the Rh factor + or −.

We can represent the events on a **tree diagram**.

The sample space in this experiment is the set {A+, A−, B+, B−, AB+, AB−, O+, O−}

The same simple events can also be arranged in a probability table (Table 8.1).

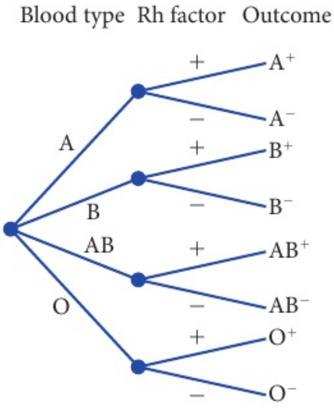


Figure 8.3 Tree diagram

		Blood type			
		A	B	AB	O
Rh factor	**Positive**	A+	B+	AB+	O+
	Negative	A−	B−	AB−	O−

Table 8.1 Probability table

Or using a 2-dimensional grid (Figure 8.4).

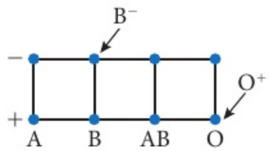


Figure 8.4 2D grid for blood types

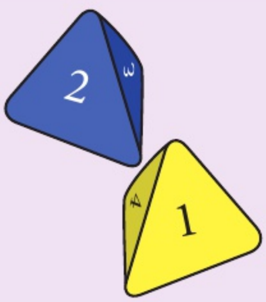
Example 8.4

Two tetrahedral dice, with sides numbered 1 to 4, one blue and one yellow, are rolled. List the elements of the following events:

T = {at least one dice lands on 3}

B = {the blue dice lands on 3}

S = {sum of the dice is a six}



Solution

We can use a sample space diagram, such as that in Figure 8.5, to help us.

$T = \{(1, 3), (2, 3), (3, 3), (4, 3), (3, 2), (3, 1), (3, 4)\}$

$B = \{(1, 3), (2, 3), (3, 3), (4, 3)\}$

$S = \{(2, 4), (3, 3), (4 ,2)\}$

Or we can use a table:

		Blue dice			
		1	2	3	4
Yellow dice	1	(1, 1)	(1, 2)	(1, 3)	(1, 4)
	2	(2, 1)	(2, 2)	(2, 3)	**(2, 4)**
	3	(3, 1)	(3, 2)	**(3, 3)**	(3, 4)
	4	(4, 1)	**(4, 2)**	(4, 3)	(4, 4)

For T, row 3 and column 3.

For B, column 3 only.

For S, red (bold) cells.

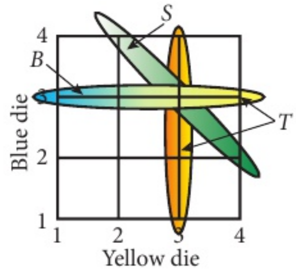


Figure 8.5 Solution to Example 8.4

Basic terms

A summary of what we learned in this section is:

Experiment: An activity or measurement that results in an outcome.

Sample space: All possible outcomes of an experiment.

Event: One or more of the possible outcomes of the experiment; a subset of the sample space.

Probability: a number between 0 and 1 which expresses the chance that an event will occur.

Exercise 8.1

1. In a large school, a student is selected at random. Give a reasonable sample space for answers to each of the following questions.
   (a) Are you left-handed or right-handed?
   (b) What is your height in centimetres?
   (c) How many minutes did you study last night?

2. We flip a coin and roll a standard six-sided dice and we record the number and the face that appear in that order. For example, (5, h) represents a 5 on the dice and a head on the coin. Find the sample space.

3. We draw cards from a deck of 52 playing cards.
   (a) List the sample space if we draw one card at a time.
   (b) List the sample space if we draw two cards at a time.
   (c) How many outcomes do you have in each of the experiments above?

4. Tim carried out an experiment where he flipped 20 coins together and observed the number of heads showing. He repeated this experiment 10 times and got the following results:

   11, 9, 10, 8, 13, 9, 6, 7, 10, 11

   (a) Use Tim's data to get the probability of obtaining a head.

   (b) He flipped the 20 coins for the 11th time. How many heads should he expect to get?

   (c) He flipped the coins 10 000 times. How many heads should he expect to see?

5. In a game a four-sided dice with sides marked 1, 2, 3, and 4 is used. The intelligence of the player is determined by rolling the dice twice and adding 1 to the sum of the numbers observed on the face of the side it lands on.

   (a) What is the sample space for rolling the dice twice?

   (b) What is the sample space for the intelligence of the player?

6. A box contains three balls, blue, green, and yellow. You run an experiment where you draw a ball, look at its colour and then replace it and draw a second ball.

   (a) What is the sample space of this experiment?

   (b) What is the event of drawing yellow first?

   (c) What is the event of drawing the same colour twice?

7. Repeat the same exercise as in Question 6, without replacing the first ball.

8. Nick flips a coin three times and each time he notes whether it is heads or tails.

   (a) What is the sample space of this experiment?

   (b) What is the event that heads occur more often than tails?

9. Franz lives in Vienna. He and his family decided that their next vacation will be either Italy or Hungary. If they go to Italy, they can fly, drive, or take the train. If they go to Budapest, they will drive or take a boat. Letting the outcome of the experiment be the location of their vacation and their mode of travel, list all the points in the sample space. Also list the sample space of the event 'fly to destination.'

10. A hospital codes patients according to whether they have health insurance, or no insurance, and according to their condition. The condition of the patient is rated as good (g), fair (f), serious (s), or critical (c). The clerk at the front desk marks 0, for non-insured patients, and 1 for insured, and one of the letters for the condition. So, (1, c) means an insured patient with critical condition.

(a) List the sample space of this experiment.

(b) What is the event: not insured, in serious or critical condition?

(c) What is the event: patient in good or fair condition?

(d) What is the event: patient has insurance?

11. A social study investigates people for different characteristics. One part of the study classifies people according to gender (G_1 = female, G_2 = male), drinking habits (K_1 = abstain, K_2 = drinks occasionally, K_3 = drinks frequently), and marital status (M_1 = married, M_2 = single, M_3 = divorced, M_4 = widowed).

(a) List the elements of an appropriate sample space for observing a person in this study.

(b) Three events are defined as:

A = the person is a male, B = the person drinks, and C = the person is single

List the elements of each of A, B, and C.

(c) Interpret each event in the context of this situation:

(i) $A \cup B$ (ii) $A \cap C$ (iii) C'

(iv) $A \cap B \cap C$ (v) $A' \cap B$

12. Cars leaving the highway can take a right turn (R), left turn (L), or go straight (S). You are collecting data on traffic patterns at this intersection and you group your observations by taking four cars at a time every 5 minutes.

(a) List a few outcomes in your sample space U. How many are there?

(b) List the outcomes in the event that all cars go in the same direction.

(c) List the outcomes that only two cars turn right.

(d) List the outcomes that only two cars go in the same direction.

13. You are collecting data on traffic at an intersection for cars leaving a highway. Your task is to collect information about the size of the car: truck (T), bus (B), car (C). You also have to record whether the driver has a safety belt on (SY) or no safety belt (SN), as well as whether the headlights are on (O) or off (F).

(a) List the outcomes of your sample space, U.

(b) List the outcomes of the event SY, that the driver has the safety belt on.

(c) List the outcomes of the event C, that the vehicle you are recording is a car.

(d) List the outcomes of the events in $C \cap SY$, C', and $C \cup SY$

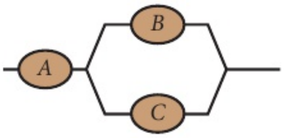


Figure 8.6 Diagram for question 14

14. Many electric systems use a built-in back-up so that the equipment using the system will work even if some parts fail. Such a system in given in Figure 8.6.

Two parts of this system are installed in parallel, so that the system will work if at least one part works. If we code a working system by 1 and a failing system by 0, then one of the outcomes would be (1, 0, 1), which means parts A and C work while B failed.

(a) List the outcomes of your sample space, U.

(b) List the outcomes of the event X, that exactly 2 of the parts work.

(c) List the outcomes of the event Y, that at least 2 of the parts work.

(d) List the outcomes of the event Z, that the system functions.

(e) List the outcomes of these events.

(i) Z' **(ii)** $X \cup Z$ **(iii)** $X \cap Z$

(iv) $Y \cup Z$ **(v)** $Y \cap Z$

15. Your school library has five copies of George Polya's *How To Solve It* book. Copies 1 and 2 are first edition, and copies 3, 4, and 5 are second edition. You are searching for a first edition book, and you will stop when you find a copy. For example, if you find copy 2 immediately, then the outcome is 2. Outcome 542 represents the outcome that a first edition was found on the third attempt.

(a) List the outcomes of your sample space, U.

(b) List the outcomes of the event A, that two books must be searched.

(c) List the outcomes of the event B, that at least two books must be searched.

(d) List the outcomes of the event C, that book 1 is found.

8.2 Probability assignments

There are a few theories of probability that assign meaning to statements like 'the probability that A occurs is p%.' In this book we will primarily examine only the **relative frequency concept**. In essence we will follow the idea that probability is 'the long-run proportion of repetitions on which an event occurs.' This allows us to 'merge' two concepts into one.

Equally likely outcomes

If a given experiment or trial has n possible outcomes among which there is no preference, they are equally likely. The probability of each outcome is then $\frac{100\%}{n}$ or $\frac{1}{n}$. For example, if a coin is balanced well, there is no reason for it to land heads in preference to tails when it is flipped, so accordingly, the probability that the coin lands heads is equal to the probability that it lands tails, and both are $\frac{100\%}{2} = 50\%$. Similarly, if a dice is fair, the chance that when it is rolled it lands with the side with 1 on top is the same as the chance that it shows 2, 3, 4, 5, or 6: $\frac{100\%}{6}$ or $\frac{1}{6}$. If an event consists of more than one possible outcome, the chance of the event is the number of ways it can occur, divided by the total number of things that could occur. For example, the chance that a dice lands showing an even number on top is the number of ways it could land showing an even number (2, 4, or 6), divided by the total number of things that could occur (6, namely showing 1, 2, 3, 4, 5, or 6).

Relative frequency

Probability is the limit of the relative frequency with which an event occurs in repeated trials. So, 'the probability that A occurs is $p\%$' means that if you repeat the experiment over and over again, independently and under essentially identical conditions, the percentage of the time that A occurs will converge to p. For example, to say that the chance that an individual is left handed is roughly 12%, means that if you ask people over and over again, independently, the ratio of the number of times you find a left handed individual to the total number of people you ask tends to be around 12% as the number of people you ask grows. Because the ratio of left-handers to total population is always between 0% and 100%, when the probability exists it must be between 0% and 100%.

Probability is on a scale of 0% to 100% or 0 to 1.

The probability of any event is the number of elements in an event A divided by the total number of elements in the sample space U.

$$P(A) = \frac{n(A)}{n(U)}$$

where $n(A)$ represents the number of outcomes in A and $n(U)$ represents the total number of outcomes.

Probability rules

The basic rules of probability are given below.

Rule 1

- Any probability is a number between 0 and 1, i.e. the probability $P(A)$ of any event A satisfies $0 \leqslant P(A) \leqslant 1$. If the probability of any event is 0, the event never occurs. Likewise, if the probability is 1, it always occurs. In rolling a standard dice, it is impossible to get the number 9, so $P(9) = 0$. Also, the probability of observing any integer between 1 and 6 inclusive is 1.

Remember that there is no such thing as negative probability or a probability greater than 1.

Rule 2

- All possible outcomes together must have probability 1, i.e., the probability of the sample space U is 1: $P(U) = 1$. Informally, this is sometimes called the 'Something has to happen rule.'

We have to be careful with these rules. By the 'something has to happen' rule, the total of the probabilities of all possible outcomes must be 1. This is because they are disjoint, and their sum covers all the elements of the sample space. Suppose someone reports the following probabilities for students in your high school (4 years). If the probability that a grade 1, 2, 3 or 4 student is chosen at random from the high school is 0.24, 0.24, 0.25, and 0.19 respectively with no other possibilities, you should know immediately that there is something wrong. These probabilities add up to 0.92. Similarly, if someone claims that these probabilities are 0.24, 0.28, 0.25, 0.26 respectively, there is also something wrong. These probabilities add up to 1.03, which is more than 1.

Rule 3

- If two events have no outcomes in common, the probability that one or the other occurs is the sum of their individual probabilities. Two events that have no outcomes in common, and hence can never occur together, are called **disjoint** events or **mutually exclusive** events.

  $$P(A \text{ or } B) = P(A) + P(B)$$

  This is the **addition rule for mutually exclusive events**.

  For example, in flipping three coins, the events of getting exactly two heads or exactly two tails are disjoint, and hence the probability of getting exactly two heads or two tails is

  $$\frac{3}{8} + \frac{3}{8} = \frac{6}{8} = \frac{3}{4}$$

  Additionally, we can always add the probabilities of **outcomes** because they are always disjoint. A trial cannot come out in two different ways at the same time. This will give you a way to check whether the probabilities you assigned are legitimate.

Rule 4

- The event that contains the outcomes **not in *A*** is called the **complement** of A and is denoted by $\boldsymbol{A'}$.

  $$P(A') = 1 - P(A), \text{ or } P(A) = 1 - P(A')$$

Example 8.5

When people create codes for their cell phones, the first digits follow distributions very similar to those shown in the table.

First digit	0	1	2	3	4	5	6	7	8	9
Probability	0.009	0.300	0.174	0.122	0.096	0.078	0.067	0.058	0.051	0.045

(a) Find the probabilities of the events

$A = \{\text{first digit is } 1\}$

$B = \{\text{first digit is more than } 5\}$

$C = \{\text{first digit is an odd number}\}$

(b) Find the probability that the first digit is

(i) 1 or greater than 5

(ii) not 1

(iii) an odd number or a number larger than 5.

Solution

(a) From the table:

$$\begin{aligned} P(A) &= 0.300 \\ P(B) &= P(6) + P(7) + P(8) + P(9) \\ &= 0.067 + 0.058 + 0.051 + 0.045 = 0.221 \\ P(C) &= P(1) + P(3) + P(5) + P(7) + P(9) \\ &= 0.300 + 0.122 + 0.078 + 0.058 + 0.045 \\ &= 0.603 \end{aligned}$$

(b) (i) Since A and B are mutually exclusive, by the addition rule, the probability that the first digit is 1 or greater than 5 is

$$P(A \text{ or } B) = 0.300 + 0.221 = 0.521$$

(ii) Using the complement rule, the probability that the first digit is not 1 is: $P(A') = 1 - P(A) = 1 - 0.300 = 0.700$

(iii) The probability that the first digit is an odd number or a number larger than 5:

$$\begin{aligned} P(B \text{ or } C) &= P(1) + P(3) + P(5) + P(6) + P(7) + P(8) + P(9) \\ &= 0.300 + 0.122 + 0.078 + 0.067 + 0.058 + 0.051 \\ &\quad + 0.045 \\ &= 0.721 \end{aligned}$$

Note that $P(B \text{ or } C)$ is not the sum of $P(B)$ and $P(C)$ because B and C are not mutually exclusive.

Equally likely outcomes

In some cases, we are able to assume that individual outcomes are equally likely because of some balance in the experiment. Flipping a balanced coin renders heads or tails equally likely, with each having a probability of 50%, and rolling a standard balanced dice gives the numbers from 1 to 6 as equally likely, each having a probability of $\frac{1}{6}$

Suppose in Example 8.5, we consider all the digits to be equally likely to happen, then our table would be as in Table 8.2.

First digit	0	1	2	3	4	5	6	7	8	9
Probability	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1

Table 8.2 Considering all digits to be equally likely to happen

$$\begin{aligned} P(A) &= 0.1 \\ P(B) &= P(6) + P(7) + P(8) + P(9) = 4 \times 0.1 = 0.4 \\ P(C) &= P(1) + P(3) + P(5) + P(7) + P(9) = 5 \times 0.1 = 0.5 \end{aligned}$$

Also, by the complement rule, the probability that the first digit is not 1 is

$$P(A') = 1 - P(A) = 1 - 0.1 = 0.9$$

2-dimensional grids are also very helpful tools that are used to visualise 2-stage or sequential probability models. For example, consider rolling a normal unbiased cubical dice twice. Here are some events and how to use the grid in calculating their probabilities (Figure 8.7).

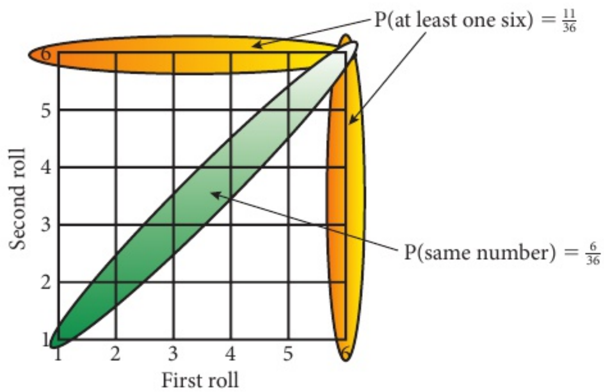


Figure 8.7 How to use the grid in calculating event probabilities

If we are interested in the probability that at least one roll shows a 6, we count the points on the column corresponding to 6 on the first roll and the points on the row corresponding to 6 on the second roll, observing naturally that the point in the corner should not be counted twice.

If we are interested in the number showing on both rolls to be the same, then we count the points on the diagonal as shown.

If we are interested in the probability that the first roll shows a number larger than the second roll, then we pick the points below the diagonal.

Hence P(first number > second number) $= \dfrac{15}{36}$

Exercise 8.2

1. For each of the following, list the sample space and say whether you think the events are equally likely.

(a) A family has 3 children; record the number of girls.

(b) Roll 2 dice; record the larger number.

(c) A family has 3 children; record each child's gender in order of birth.

(d) Flip a coin 20 times; record the longest run of tails.

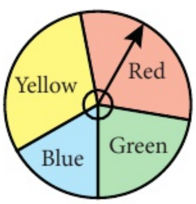


Figure 8.8 Spinner in question 2

2. The arrow on a spinner for a child's game stops rotating to point at a colour. (It does not stop at an edge!)

Which of the probabilities shown in the table are valid?

	Probability			
	Red	**Blue**	**Yellow**	**Green**
a	0.25	0.25	0.25	0.25
b	0.3	0.2	0.4	0.1
c	0.4	0.4	0.2	0.1
d	0	0	1	0
e	0.4	0.4	−0.3	0.5

3. A study to relate students' high school performance to their college performance was conducted at a large university. Here is a table for third-year college students and their high school academic ranking.

Rank	Top 20%	Second 20%	Third 20%	Fourth 20%	Lowest 20%
Number	615	345	435	90	15

A third-year student is selected at random. What is the probability that the student

(a) is not in the top 20% of the class

(b) is in the top 40% of the class

(c) is not in the lowest 20% of the class?

4. In a simple experiment, 20 chips with integers 1–20 inclusive are placed in a box and one chip picked at random.

(a) What is the probability that the number drawn is a multiple of 3?

(b) What is the probability that the number drawn is not a multiple of 4?

5. The probability an event *A* happens is 0.37.

(a) What is the probability that it does not happen?

(b) What is the probability that it may or may not happen?

6. You are playing with an ordinary deck of 52 cards by drawing cards at random and looking at them.

(a) Find the probability that the card you draw is:

(i) the ace of hearts **(ii)** the ace of hearts or any spade

(iii) an ace or any heart **(iv)** not a face card.

(b) Now you draw the ten of diamonds and put it on the table and draw a second card. What is the probability that the second card:

(i) is the ace of hearts **(ii)** is not a face card?

(c) Now you draw the ten of diamonds and return it to the deck and draw a second card. What is the probability that the second card:

(i) is the ace of hearts **(ii)** is not a face card?

7. In the morning on Monday, my class wanted to know how many hours students spent studying on Sunday night. They stopped schoolmates at random as they arrived and asked each: 'How many hours did you study last night?' Here are the answers of the sample they chose on Monday, 14 January 2018.

Number of hours	0	1	2	3	4	5
Number of students	4	12	8	3	2	1

Estimate the probability that a student:

(a) spent less than three hours studying Sunday night

(b) studied two or three hours

(c) studied less than six hours.

8. We flip a coin and roll a standard six-sided dice and record the number and the face that appear. Find:

(a) the probability of having a number larger than 3

(b) the probability that we get a head and a 6.

9. A dice is constructed in such a way that a 1 has the chance to occur twice as often as any other number.

(a) Find the probability that a 5 appears.

(b) Find the probability an odd number will occur.

10. You are given two fair dice to roll in an experiment.

(a) Your first task is to report the numbers you observe.

(i) What is the sample space of your experiment?

(ii) What is the probability that the two numbers are the same?

(iii) What is the probability that the two numbers differ by 2?

(iv) What is the probability that the two numbers are not the same?

(b) In a second stage, your task is to report the sum of the numbers that appear.

(i) What is the probability that the sum is 1?

(ii) What is the probability that the sum is 9?

(iii) What is the probability that the sum is 8?

(iv) What is the probability that the sum is 13?

11. An estimate of the world population in 2018 is shown in the table. Numbers are in millions.

Age	Male	Female	Total
14 years or less	964	899	1863
15–24	611	572	1183
25–54	1523	1488	3011
55–64	307	322	629
65 and over	284	352	636
Total	3689	3633	7322

A person is randomly chosen. What is the probability that this person is:

(a) a female

(b) a male between the age of 15 and 24

(c) at least 65 years old.

12. In each of the following situations, state whether or not the given assignment of probabilities to individual outcomes is legitimate. Give reasons for your answer.

(a) A dice is loaded such that the probability of each face is according to the following assignment, where x is the number of spots on the upper face and $P(x)$ is its probability.

x	1	2	3	4	5	6
$P(x)$	0	$\frac{1}{6}$	$\frac{1}{3}$	$\frac{1}{3}$	$\frac{1}{6}$	0

(b) A student at your school categorised in terms of gender and whether they are diploma candidates or not.

P(female, diploma candidate) = 0.57

P(female, not a diploma candidate) = 0.23

P(male, diploma candidate) = 0.43

P(male, not a diploma candidate) = 0.18

(c) Draw a card from a deck of 52 cards where x is the suit of the card and $P(x)$ is its probability.

x	Hearts	Spades	Diamonds	Clubs
$P(x)$	$\frac{12}{52}$	$\frac{15}{52}$	$\frac{12}{52}$	$\frac{13}{52}$

13. In Switzerland, there are three official mother tongues, German, French, and Italian. You choose a Swiss at random and ask what their mother tongue is. The table shows the distribution of responses.

Language	German	French	Italian	Other
Probability	0.58	0.24	0.12	?

(a) What is the probability that a Swiss person's mother tongue is not one of the official ones?

(b) What is the probability that a Swiss person's mother tongue is not German?

(c) What is the probability that a Swiss person's mother tongue is French or Italian?

14. The majority of email messages are now spam. The table shows the distribution of topics in spam emails.

Topic	Adult	Financial	Health	Leisure	Products	Scams
Probability	0.165	0.142	0.075	0.081	0.209	0.145

(a) What is the probability of choosing a spam message that does not concern these topics?

Parents are usually concerned with spam messages with 'adult' content.

(b) What is the probability that a randomly chosen spam email falls into one of the other categories?

15. An experiment involves rolling a pair of dice, 1 white and 1 red, and recording the numbers that come up. Find the probability that:

(a) the sum is greater than 8

(b) a number greater than 4 appears on the white dice

(c) at most a total of 5 appears.

16. **(a)** A box contains 8 chips numbered 1 to 8. Two are chosen at random and their numbers are added together. What is the probability that their sum is 7?

(b) A box contains 20 chips numbered 1 to 20. Two are chosen at random. What is the probability that the numbers on the two chips differ by 3?

(c) A box contains 20 chips numbered 1 to 20. Two are chosen at random. What is the probability that the numbers on the two chips differ by more than 3?

17. A wooden cube has its faces painted green. The cube is cut into 1000 small cubes of equal size. We mix the small cubes thoroughly. One cube is drawn at random. What is the probability that the cube:

(a) has two faces coloured green

(b) has three coloured faces

(c) does not have a coloured face at all?

18. Government statistics list leading causes of death. Numbers of deaths by gender in the US for 2015 are summarised in the table.

Cause of death	**Male**	**Female**	**Total**
Heart diseases	335 002	298 840	
Cancer	313 818	282 112	
Accidents	92 919	53 652	
Diabetes	43 123	36 412	
Other causes			
Total	1 373 404	1 339 226	

(a) Copy the table and fill in the blank cells.

(b) Find the probability that a randomly chosen death is:

(i) a male

(ii) a male with heart disease

(iii) due to cancer.

(c) What proportion of female deaths is caused by accidents?

(d) What proportion of accident deaths were female?

19. The social networking site Facebook is popular around the world. The following table shows the number of Facebook users and populations in four countries in 2017 (in millions).

Country	Facebook	Population
India	294	1339
Indonesia	131	264
USA	204	324
Mexico	84	129

(a) If we select a Facebook user from these at random, what is the probability that the person lives in Asia?

(b) If we pick a person at random in any of these countries, in which country is the probability of picking a Facebook user lowest? Highest?

(c) If we pick randomly a person in each country, what is the probability that all four are not users of Facebook?

8.3 Operations with events

In Example 8.5, we talked about the following events

$B = \{\text{first digit is more than } 5\}$

$C = \{\text{first digit is an odd number}\}$

We also claimed that these two events are not mutually exclusive. This brings us to another concept for looking at combined events.

The **intersection** of two events B and C, denoted by the symbol $\boldsymbol{B \cap C}$ or simply BC, is the event containing all outcomes common to B and C.

In Figure 8.9, $B \cap C = \{7, 9\}$ because these outcomes are in both B and C. Since the intersection has outcomes common to the two events B and C, they are not mutually exclusive.

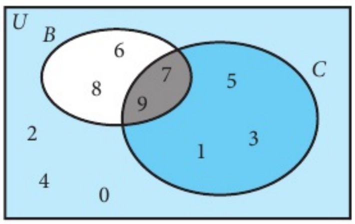


Figure 8.9 Intersection of events

The probability of $B \cap C$ is $0.058 + 0.045 = 0.103$

The probability of B or C is not simply the sum of the two probabilities. How can we find the probability of B or C when they are not mutually exclusive? We need to define the union (see Key Fact box).

The **union** of two events B and C, $B \cup C$, is the event containing all the outcomes that belong to B or to C or to both.

Here $B \cup C = \{1, 3, 5, 6, 7, 8, 9\}$. In calculating the probability of $B \cup C$, we observe that the outcomes 7 and 9 are counted twice. To remedy the situation, if we decide to add the probabilities of B and C, we subtract one of the incidents of double counting. So, $P(B \cup C) = 0.221 + 0.603 - 0.103 = 0.721$, which is the result we received with direct calculation. In general, we can state the following probability rule:

For any two events A and B, $P(A \cup B) = P(A) + P(B) - P(A \cap B)$.

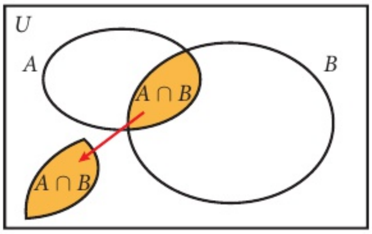


Figure 8.10 Union of events

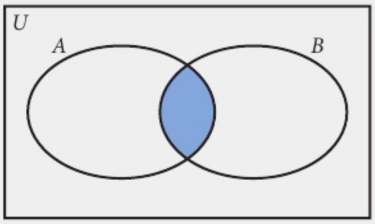
As you see from Figure 8.10, $P(A \cap B)$ has been added twice, so the 'extra' one is subtracted to give the probability of $(A \cup B)$

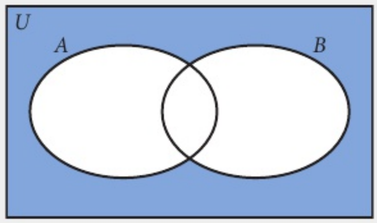
This general probability addition rule applies to the case of mutually exclusive events too. Consider any two events A and B. The probability of A or B is given by

$$P(A \cup B) = P(A) + P(B) - P(A \cap B) = P(A) + P(B)$$
since $P(A \cap B) = 0$

Some useful results

1. $P(A \cap B) = P(B \cap A)$
2. $P(A' \cap B') = P(A \cup B)' = 1 - P(A \cup B)$





The simple multiplication rule

Consider the following situation: In a large school, 55% of the students are male. It is also known that the percentage of smokers among males and females in this school are the same, 22%. What is the probability of selecting a student at random from this population that is a male smoker?

Two events A and B are **independent** if knowing that one of them occurs does not change the probability that the other occurs.

The multiplication rule for independent events: If two events A and B are **independent**, then $P(A \cap B) = P(A) \times P(B)$

Applying common sense only, we can think of the problem in the following manner. Since the proportion of smokers is the same in both groups, smoking and gender are independent of each other in the sense that knowing that the student is a male does not influence the probability that he smokes!

The chance we pick a male student is 55%. From those 55% of the population, we know that 22% are smokers, so by simple arithmetic the chance that we select a male smoker is

$$0.22 \times 0.55 = 12.1\%.$$

This is an example of the multiplication rule for independent events.

Example 8.6

The blood types of people can be one of the four types: O, A, B, or AB. The distribution of people with these types differs from one group of people to another. Here are the distributions of blood types for randomly chosen people in the USA, China, and Russia.

Country \ Blood type	O	A	B	AB
USA	0.43	0.41	0.12	?
China	0.36	0.27	0.26	0.13
Russia	0.39	0.34	?	0.09

(a) What is the probability of type AB in the USA?

(b) Dirk lives in the USA and has type B blood. Type B can only receive from O and B. What is the probability that a randomly chosen USA citizen can donate blood to Dirk?

(c) What is the probability of randomly choosing a person from the USA and a person from China with type O blood?

(d) What is the probability of randomly choosing three people, one from the USA, one from China and one from Russia, with type O blood?

(e) What is the probability of randomly choosing three people, one from the USA, one from China and one from Russia, with the same blood type?

Solution

(a) Since the complement of AB, AB′ is O, A, or B, then

$$P(AB) = 1 - (0.36 + 0.27 + 0.26) = 0.04$$

(b) This is the union of two mutually exclusive events, O or B:

$$P(O \cup B) = 0.43 + 0.12 = 0.55$$

(c) This is an intersection of two independent events (O in USA and China)

$$P(USA \cap China) = P(USA) \times P(China) = 0.43 \times 0.36 = 0.1548$$

(d) This is an extension of the two independent events

$$\begin{aligned} P(USA \cap China \cap Russia) &= P(USA) \times P(China) \times P(Russia) \\ &= 0.43 \times 0.36 \times 0.39 = 0.060372 \end{aligned}$$

(e) All three individuals could have O, A, B or AB types.

We found P(O) in part (d): $P(O) = 0.060372$

Using the complement for the missing entry for Russia

$$P(A) = 0.41 \times 0.27 \times 0.34 = 0.037638,$$

$$P(B) = 0.12 \times 0.26 \times (1 - (0.39 + 0.34 + 0.09)) = 0.005832$$

$$P(AB) = 0.04 \times 0.13 \times 0.09 = 0.000468$$

So, adding all the probabilities,

P(person from USA, China and Russia with the same blood type)
$= 0.104031$

This rule can also be extended to more than two independent events. Parts (d) and (e), on the assumption of independence, demonstrate the general rule:
$P(A \cap B \cap C \cap \ldots) = P(A) \times P(B) \times P(C) \times \ldots$

Do not confuse independent with mutually exclusive. Mutually exclusive (or disjoint) means that if one of the events occurs, then the other does not occur; 'independent' means that knowing one of the events occurs does not influence whether the other occurs or not.

Example 8.7

It is estimated that 17% of computers bought from one computer manufacturer require one repair job during the first month after purchase, 7% will need repairs twice during the first month, and 4% require three or more repairs.

(a) What is the probability that a computer chosen at random from this manufacturer will need:

(i) no repairs

(ii) no more than one repair

(iii) some repair?

(b) If you buy two computers from this manufacturer, what is the probability that:

(i) neither will require repair

(ii) both will need repair?

Solution

(a) Since all of the events listed are mutually exclusive, we can use the addition rule.

(i) $$\begin{aligned} \text{P(no repairs)} &= 1 - \text{P(some repairs)} \\ &= 1 - (0.17 + 0.07 + 0.04) = 1 - (0.28) = 0.72 \end{aligned}$$

(ii) $$\begin{aligned} \text{P(no more than one repair)} &= \text{P(no repairs)} + \text{P(one repair)} \\ &= 0.72 + 0.17 = 0.89 \end{aligned}$$

(iii) $$\begin{aligned} \text{P(some repairs)} &= \text{P(one repair)} + \text{P(two repairs)} \\ &\quad + \text{P(three or more repairs)} \\ &= 0.17 + 0.07 + 0.04 = 0.28 \end{aligned}$$

(b) Since repairs on the two computers are independent from one another, the multiplication rule can be used. Use the probabilities of events from part (a) in the calculations.

(i) $\text{P(neither will need repair)} = (0.72)(0.72) = 0.5184$

(ii) $\text{P(both will need repair)} = (0.28)(0.28) = 0.0784$

Conditional probability

In probability, conditioning means incorporating new restrictions on the outcome of an experiment: updating probabilities to take into account new information. We will look at how conditional probability can be used to solve complicated problems.

Example 8.8

A public health department wants to study the smoking behaviour of high school students. They interview 768 students from grades 10–12 and asked them about their smoking habits. They categorised the students into the following two categories: smokers and non-smokers. The results are summarised in the table.

	Smoker	Non-smoker	Total
Male	169	245	414
Female	145	209	354
Total	314	454	768

A student is selected at random from this study. What is the probability that we select:

(a) a female (b) a male smoker (c) a non-smoker?

Solution

(a) Since we have 354 females in the study,

$$P(\text{female}) = \frac{354}{768} = 0.461$$

So, 46.1% of our sample are females.

(b) Since we have 169 males categorised as smokers, the chance of a male smoker will be

$$P(\text{male smoker}) = \frac{169}{768} = 0.220$$

(c) We have 454 non-smokers

$$P(\text{non-smoker}) = \frac{454}{768} = 0.591$$

In Example 8.8, what if we know that the selected student is a female? Does that influence the probability that the selected student is a non-smoker? Yes it does!

Knowing that the selected student is a female changes our choices. The revised sample space is not made up of all students anymore. It is made up of only the female students. The chance of finding a non-smoker among the females is $\frac{209}{354} = 0.590$, i.e., 59.0% of the females are non-smokers as compared to the 59.1% of non-smokers in the whole population.

This is an example of conditional probability. We write this as

$$P(\text{non-smoker}|\text{female}) = \frac{209}{354} = 0.590$$

We read this as 'probability of selecting a non-smoker **given that** we have selected a female'. Interpreted differently, 59% of the females are non-smokers.

Looking further at this, if we asked the question differently: what is the probability of selecting a female knowing that a non-smoker has been selected? Our sample space will be the non-smoker group. Thus,

$P(\text{female}|\text{non-smoker}) = \frac{209}{454} = 0.460$. 46% of the non-smokers are females.

The conditional probability of A given B, $P(A|B)$, is the probability of the event A, updated on the basis of the knowledge that the event B occurred.

Remember that the probability we assign to an event can change if we know that some other event has occurred. This idea is the key to understanding conditional probability.

Imagine the following scenario.

You are playing cards and your opponent is about to give you a card. What is the probability that the card you receive is a queen?

As you know, there are 52 cards in the deck, 4 of these cards are queens. So, assuming that the deck was thoroughly shuffled, the probability of receiving a queen is

$$P(\text{queen}) = \frac{4}{52} = \frac{1}{13}$$

This calculation assumes that you know nothing about any cards already dealt from the deck.

Suppose now that you are looking at the five cards you have in your hand, and one of them is a queen. You know nothing about the other 47 cards (52 − the 5 cards you have!) except that exactly three queens are among them. The probability of being given a queen as the next card, given what you know, is

$$P(\text{queen}|\text{5 cards including a queen in hand}) = \frac{3}{47} \neq \frac{1}{13}$$

So, knowing that there is one queen among your five cards changes the probability of the next card being a queen.

Consider Example 8.8 again. We want to express the table frequencies as relative frequencies or probabilities. This is called a joint probability table.

Joint probabilities

	Smoker	Non-smoker	Total
Male	0.220	0.319	0.539
Female	0.189	0.272	0.461
Total	0.409	0.591	1.000

Marginal probabilities

Table 8.3 Joint probability table

To find the probability of selecting a student at random and finding that student is a non-smoker female, we look at the intersection of the female row with the non-smoking column and find that this probability is 0.272.

The probabilities in the right column and the bottom row are called marginal probabilities. For example, looking at the right column, you can read from it that 53.9% of the individuals asked are male and 46.1% are female. Looking at the bottom row, 40.9% of the individuals are smokers, while 59.1% are non-smokers.

Looking at this calculation from a different perspective, we can think about it in the following manner.

We know that the percentage of females in our sample is 46.1, and among those females, in Example 8.8, we found that 59.0% of those are non-smokers. So, the percentage of female non-smokers in the population is the 59.0% of those 46.1% females, i.e., $0.590 \times 0.461 = 0.272$!

In terms of events, this can be read as:

$$\text{P(non-smoker|female)} \times \text{P(female)} = \text{P(female} \cap \text{non-smoker)}$$

The previous discussion is an example of the **multiplication rule** of any two events A and B.

The multiplication rule
Given any events A and B, the probability that both events happen is given by
$P(A \cap B) = P(A|B)\,P(B)$

Example 8.9

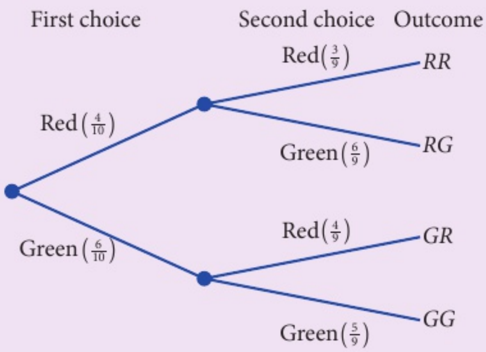
In a psychology lab, researchers are studying the colour preferences of young children. Six green toys and four red toys (identical apart from colour) are placed in a container. The child is asked to select two toys at random. What is the probability that the child chooses two red toys?

Solution

To solve this problem, we draw a tree diagram.

Each entry on the branches has a conditional probability. So, Red on the second choice is either Red|Red or Red|Green.

We are interested in RR, so the probability is:



$$P(RR) = P(R) \times P(R|R) = \frac{4}{10} \times \frac{3}{9} = 13.3\%$$

When $P(A \cap B) = P(A|B) \times P(B)$, and $P(B) \neq 0$, we can rearrange the multiplication rule to produce a definition of the conditional probability $P(A|B)$ in terms of the unconditional probabilities $P(A \cap B)$ and $P(B)$.

When $P(B) \neq 0$, the conditional probability of A given B is
$P(A|B) = \frac{P(A \cap B)}{P(B)}$

Why does this formula make sense?

First of all, note that it does agree with the intuitive answers we found above.

In Example 8.8, recall that $\text{P(non-smoker|female)} = \frac{209}{354} = 0.590$

Now applying the new definition we have

$$\text{P(non-smoker|female)} = \frac{\text{P(non-smoker} \cap \text{female)}}{\text{P(female)}} = \frac{0.272}{0.461} = 0.590 \text{ or}$$

$$\text{P(female|non-smoker)} = \frac{\text{P(non-smoker} \cap \text{female)}}{\text{P(non-smoker)}} = \frac{0.272}{0.591} = 0.460$$

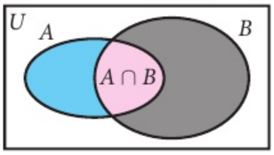


Figure 8.11 $P(A|B)$

Now, if we learn that B occurred, we can restrict attention to just those outcomes that are in B, and disregard the rest of U, so we have a new sample space that is just B (see Figure 8.11). For A to have occurred in addition to B, requires that $A \cap B$ occurred, so the conditional probability of A given B is $\frac{P(A \cap B)}{P(B)}$, just as we defined it above.

Example 8.10

In an experiment to study the phenomenon of colour-blindness, researchers collected information concerning 1000 people in a small town and categorised them according to colour-blindness and gender. The table shows a summary of the findings:

	Male	Female	Total
Colour-blind	40	2	42
Not colour-blind	470	488	958
Total	510	490	1000

What is the probability that a person chosen at random is:

(a) colour-blind, given that the person is a female

(b) a colour-blind female

(c) a colour-blind individual?

Solution

Let C stand for colour-blind, F for female, M for male and NC for not colour blind.

(a) Note that we do not have to search the whole population for this event. We limit our search to the females. There are 490 females. As we only need to consider females, then when we search for colour-blindness, we only look for the females who are colour-blind, i.e., the intersection. Here we only have two females. Therefore, the chance we get a colour-blind person, given the person is a female is

$$P(C|F) = \frac{P(C \cap F)}{P(F)} = \frac{n(C \cap F)}{n(F)} = \frac{2}{490} = 0.004$$

Note that we used the frequency rather than the probability. However, these are equivalent since dividing by $n(U)$ will transform the frequency into a probability.

$$\frac{n(C \cap F)}{n(F)} = \frac{\frac{n(C \cap F)}{n(U)}}{\frac{n(F)}{n(U)}} = \frac{P(C \cap F)}{P(F)} = P(C|F)$$

(b) To have a colour-blind female means that we are looking for the probability of the intersection, $P(F \cap C)$. Using the multiplication rule

$$P(F \cap C) = P(C|F) \times P(F) = \frac{2}{490} \times \frac{490}{1000} = 0.002$$

(c) From the table, $P(C) = \frac{n(C)}{n(U)} = \frac{42}{1000} = 0.042$

However, in cases where only probabilities are available, not the original data, the tree diagram provides us with a way to find this probability (Figure 8.12).

To have a colour-blind individual, the person is either colour-blind female or a colour-blind male, i.e., $(F \cap C) \cup (M \cap C)$. Thus, and since the events in question are mutually exclusive, we have

$$P(C) = P(F \cap C) + P(M \cap C) = 0.002 + \frac{510}{1000} \times \frac{40}{510} = 0.042$$

An alternative way to look at the problem is through a tree diagram. The method makes it easier for us to visualise the situation and carry out the solutions to the rest of this question.

The first two branches distinguish our people into males and females. The second set of branches distinguishes people according to their colour blindness. Changing from numbers into probabilities provides us with an efficient tool to perform the calculations we need.

Note that the probabilities in the first two branches use the whole sample space of 1000 people. In the second set, and because we already know the gender, the sample space changes to either females with 490 or males with 510. Thus the probabilities on the second set of branches are conditional probabilities.

For example, on the dashed branches, $\frac{2}{490} = P(C|F)$, and $\frac{488}{490} = P(NC|F)$

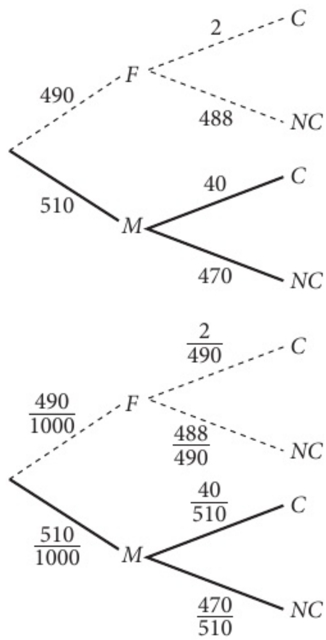


Figure 8.12 Tree diagrams for colour blindness

Example 8.11

One national airline is known for its punctuality. The probability that a regularly scheduled flight departs on time is $P(D) = 0.83$, the probability that it arrives on time is $P(A) = 0.92$, and the probability that it arrives and departs on time, $P(A \cap D) = 0.78$. Find the probability that a flight:

(a) arrives on time given that it departed on time

(b) departs on time given that it arrived on time.

Solution

(a) The probability that a flight arrives on time given that it departed on time is

$$P(A|D) = \frac{P(A \cap D)}{P(D)} = \frac{0.78}{0.83} = 0.94$$

(b) The probability that a flight departs on time given that it arrived on time

$$P(D|A) = \frac{P(D \cap A)}{P(A)} = \frac{0.78}{0.92} = 0.85$$

Independence

Two events are **independent** if learning that one occurred does not affect the chance that the other occurred. That is, if $P(A|B) = P(A)$, and vice versa.

Two events are independent if either $P(A \cap B) = P(A) \times P(B)$, or $P(A|B) = P(A)$

Otherwise, the events are not independent.

This means that if we apply our definition to the general multiplication rule, then

$$P(A \cap B) = P(A|B) \times P(B) = P(A) \times P(B)$$

which is the multiplication rule for independent events we studied earlier.

These results give us some helpful tools in checking the independence of events.

Example 8.12

Take another look at Example 8.11. Are the events of arriving on time (A) and departing on time (D) independent?

Solution

We can answer this question in two different ways:

It is given that $P(A) = 0.92$, and we found that $P(A|D) = 0.94$. Since the two values are not the same, then we can say that the two events are not independent.

Alternately, $P(A \cap D) = 0.78$ and
$P(A) \times P(D) = 0.92 \times 0.83 = 0.76 \neq P(A \cap D)$

Example 8.13

In many countries, the police stop drivers on suspicion of drunk driving. The stopped drivers are given a breath test, or a blood test, or both. In a country where this problem is vigorously dealt with, the police records show the following:

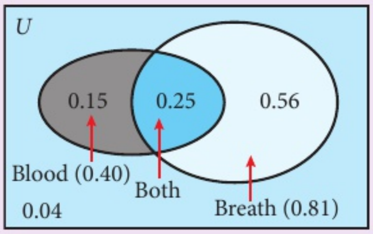
81% of the drivers stopped are given a breath test, 40% a blood test, and 25% both tests.

(a) What is the probability that a suspected driver is given:
  (i) a test
  (ii) exactly one test
  (iii) no test?

(b) Are the events giving the driver a breath test and giving the driver a blood test independent?

Solution

A Venn diagram can help explain the solution.



(a) (i) The probability that a driver receives a test means that they receive either a blood test, or a breath test, or both tests. The probability as such can be calculated directly from the diagram or by applying the addition rule. The diagram shows that if 81% receive the breath test and 25% are also given the blood test, then the remaining 56% do not receive a blood test. Similarly 15% of the blood test receivers do not get a breath test. So, the probability to receive a test is: $0.56 + 0.25 + 0.15 = 0.96$

Also, if we apply the addition rule

$$\begin{aligned} \text{P(breath or blood)} &= \text{P(breath)} + \text{P(blood)} - \text{P(both)} \\ &= 0.81 + 0.40 - 0.25 = 0.96 \end{aligned}$$

(ii) To receive exactly one test is to receive a blood test or a breath test, but not both! So, from the Venn diagram it is clear that this probability is $0.15 + 0.56 = 0.71$. To approach it differently, since we know that the union of the two events still contains the intersection, we can subtract the probability of the intersection from that of the union. i.e., $0.96 - 0.25 = 0.71$

(iii) To receive no test is equivalent to the complement of the union of the events. Hence, P(no test) $= 1 -$ P(1 test) $= 1 - 0.96 = 0.04$

(b) To check for independence, we can use any of the two methods we tried before. Since all the necessary probabilities are given, we can use the product rule:

If they were independent, then

$$\text{P(both tests)} = \text{P(breath)} \times \text{P(blood)} = 0.81 \times 0.40 = 0.324$$

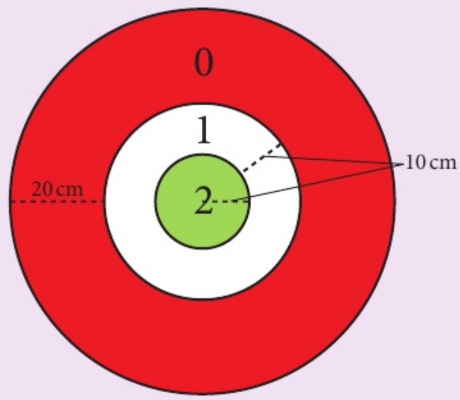
but P(both tests) $= 0.25$. Therefore, the events of receiving a breath and a blood test are not independent.

Example 8.14

The diagram shows a target for a dart game. The radius of the board is 40 cm and is divided into three regions as shown. You score 2 points if you hit the centre, 1 point for the middle region and 0 points for the outer region.

(a) What is the probability of scoring a 1 in one attempt?

(b) What is the probability of scoring a 2 in one attempt?



Solution

(a) $P(1) = \dfrac{\pi(20^2 - 10^2)}{\pi(40^2)} = \dfrac{3}{16}$

(b) $P(2) = \dfrac{\pi(10^2)}{\pi(40^2)} = \dfrac{1}{16}$

Some applications

Here are a few examples of applications of probability.

Life insurance

The life insurance industry relies heavily on relative frequencies in determining whether a person should be accepted for cover and the premiums that must be charged. Life insurance premiums are based largely on relative frequencies that reflect overall death rates for people of various ages. The table shows a portion of a mortality table of the type used by insurance companies.

	Male		Female			Male		Female	
Age	Deaths per 1000	Life expectancy (years)	Deaths per 1000	Life expectancy (years)	Age	Deaths per 1000	Life expectancy (years)	Deaths per 1000	Life expectancy (years)
0	1.6	91.2	1.6	92.8	50	2.0	38.4	1.1	40.4
1	0.4	90.3	0.4	91.9	60	4.9	28.5	3.3	30.3
10	0.1	80.7	0.1	82.5	70	10.9	19.3	8.7	20.9
20	0.4	70.0	0.2	71.9	80	31.8	11.3	23.9	12.6
30	0.7	59.4	0.3	61.3	90	107.7	5.6	86.8	6.5
40	0.8	48.9	0.5	50.8	100	267.0	2.8	229.3	3.1

Table 8.4 A portion of a mortality table of the type used by insurance companies

The table shows some interesting facts. The death rates start relatively high for new-borns, decrease slightly for the younger years, and then increase gradually afterwards.

Each death rate can be changed into a probability by dividing by 1000. The resulting probability is not essentially the probability of death for a specific individual but rather the probability for a typical individual in that age group.

For example, a typical male policy holder of age 60 would have a $\frac{4.9}{1000}$ or 0.0049 probability of passing away within the coming year.

A review of male versus female death rates shows the rates for males to be higher for each age group. Especially significant is the increase between 10 and 30 years, which is two times higher for males.

Airlines overbooking flights

Airlines overbook flights, that is they book passengers to more seats on a particular aeroplane than are available. They know that a certain percentage of people will not show up for the flight, particularly as it is likely that these passengers will have refundable tickets. So they want to fly with as few empty seats as possible.

Airlines use probability to determine exactly how many tickets to sell, too few and they're wasting seats, too many and they face penalties such as the cost of other flights, hotel stays and annoyed customers.

However, by using probability on past experiences airlines can predict the probability of passengers showing up for particular flight routes. For example, on a flight with capacity of 180 seats, an airline may sell up to 195 tickets. Using probability, it is estimated that there is almost a zero chance that all 195 will show up, there is about 1.11 percent chance that 184 show up.

Police checking for drunk drivers

In some countries, police conduct roadside stop points to check for drunk drivers. Normally drivers are required to roll down their window and speak to the police where police will look for signs of impairment. This often includes the smell of alcohol and dilation of the pupils (in reaction to them shining a flashlight in your eyes). Police will tend to give suspected drivers a breath test, or a blood test. Like any test, there is always a chance of an error. Probability is helpful in calculating risks of being at fault in accusing the wrong drivers.

Exercise 8.3

1. U is the set of positive integers less than or equal to 20.

   A, B and C are subsets of U.

   A = {even integers}; B = {multiples of 3}; $C = \{6, 8, 9, 12, 13, 15\}$

   (a) Find

   (i) $\mathrm{P}(A)$

   (ii) $\mathrm{P}(B)$

   (b) Copy and complete the Venn diagram in Figure 8.13 with all the elements of U, and find $\mathrm{P}(A \cup B)$

   (c) Find $\mathrm{P}(C')$

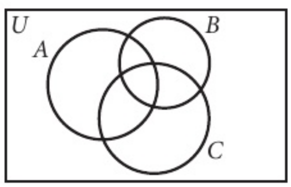
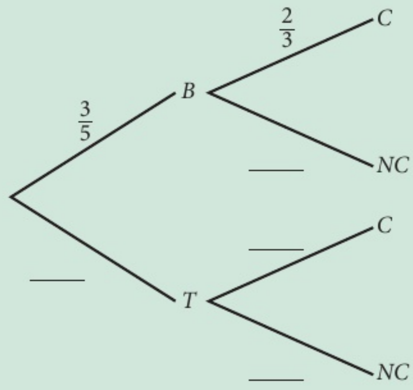


Figure 8.13 Venn diagram for question 1 (b)

2. Rami travels to work each day, either by bus or by train. The probability that he travels by bus is $\frac{3}{5}$. If he travels by bus, the probability that he buys coffee is $\frac{2}{3}$. If he travels by train, the probability that he buys coffee is $\frac{3}{4}$.

(a) Copy and complete the tree diagram.
C = Coffee,
NC = No coffee,
B = bus,
T = train

(b) Find the probability that Rami buys coffee when he travels to work.



3. Alex has an unbiased cubical (six faced) dice on which are written the numbers 1, 2, 3, 4, 5 and 6.

Ben has an unbiased tetrahedral (four faced) dice on which are written the numbers 2, 3, 5 and 7.

(a) Set up a Venn diagram with the numbers written on Alex's dice (A) and Ben's dice (B).

(b) Find $n(B \cap A')$

(c) Alex and Ben are each going to roll their dice once only. Alex looks at the number showing on top and Ben looks at the number his dice lands on. Find the probability that both show the same number.

4. Franz has an unbiased cubical (six faced) dice on which are written the numbers 1, 3, 3, 6, 6 and 6.

Gaby has an unbiased disc with one side yellow and one green.

The dice is rolled and the disc flipped at the same time.

(a) List the elements of the sample space.

(b) Find the probability that:

(i) the number shown on the dice is 3 and the disc is green

(ii) the number shown on the dice is 3 or the disc is green

(iii) the number shown on the dice is even given that the colour shown on the disc is green.

5. Events A and B are given such that $P(A) = \frac{3}{4}$, $P(A \cup B) = \frac{4}{5}$ and $P(A \cap B) = \frac{3}{10}$. Find $P(B)$.

6. Events A and B are given such that $P(A) = \frac{7}{10}$, $P(A \cup B) = \frac{9}{10}$ and $P(A \cap B) = \frac{3}{10}$

(a) Draw a Venn diagram representing the situation.

(b) Hence or otherwise, find:

(i) $P(B)$ **(ii)** $P(B' \cap A)$ **(iii)** $P(B \cap A')$

(iv) $P(B' \cap A')$ **(v)** $P(B|A')$

7. Events A and B are given such that $P(A) = \frac{1}{3}$, $P(A \cup B) = \frac{4}{9}$ and $P(B) = \frac{2}{9}$

Show that A and B are neither independent nor mutually exclusive.

8. Events A and B are given such that $P(A) = \frac{3}{7}$, and $P(A \cap B) = \frac{3}{10}$

If A and B are independent, find $P(A \cup B)$

9. Driving tests in a certain city are not easy to pass the first time you take them. After going through training, 60% of new drivers pass the test first time. If a driver fails the first test, there is chance of passing it on a second try. Two weeks later, 75% of the second-chance drivers pass the test. Otherwise, the driver has to retrain and take the test after 6 months. Draw a tree diagram and find the probability that a randomly chosen new driver will pass the test.

10. People with O-negative blood type are universal donors, i.e., they can donate blood to individuals with any blood type. Only 8% of people have O-negative blood.

(a) One person randomly appears to give blood. What is the probability that the person does not have O-negative blood?

(b) Two people appear independently to give blood. What is the probability that:

(i) both have O-negative blood

(ii) at least one of them has O-negative blood

(iii) only one of them has O-negative blood?

11. To understand which products are downloaded from the internet most, a study is conducted where students from several international schools in a big city are surveyed. The results of the questions asked are given below where books (B), music (M) or films (F), are the main products concerned.

100 students downloaded music
95 students downloaded films
68 students downloaded films and music
52 students downloaded books and music
50 students downloaded films and books
40 students downloaded all three products
8 students downloaded books only
25 students downloaded none of the three products

(a) Use this information to draw a Venn diagram.

(b) Find the number of students who were surveyed.

(c) (i) On your Venn diagram, shade the set $(F \cup M) \cap B'$

(ii) Find $n((F \cup M) \cap B')$

(d) A student who was surveyed is chosen at random.

Find the probability that:

(i) the student downloaded music

(ii) the student downloaded books, given that they had not downloaded films

(iii) the student downloaded at least two of the products.

The city has 2550 students attending international schools.

(e) Find the expected number of students in this city that download music.

12. 180 people were interviewed and asked what types of transport they had used in the last year from a choice of aeroplane (A), train (T) or bus (B). The following information was obtained.

47 had travelled by aeroplane
68 had travelled by train
122 had travelled by bus
25 had travelled by aeroplane and train
32 had travelled by aeroplane and bus
35 had travelled by train and bus
20 had travelled by all three types of transport

(a) Draw a Venn diagram to show this information.

(b) Find the number of people who, in the last year, had travelled by

(i) bus only

(ii) both aeroplane and bus but not by train

(iii) at least two types of transport

(iv) none of the three types of transport.

A person is selected at random from those who were interviewed.

(c) Find the probability that the person had used only one type of transport in the last year.

(d) Given that the person had used only one type of transport in the last year, find the probability that the person had travelled by aeroplane.

13. Two dice are rolled and the numbers on the top face are observed.

(a) List the elements of the sample space.

(b) Let x represent the sum of the numbers observed. Copy and complete the table.

x	2	3	4	5	6	7	8	9	10	11	12
P(x)		$\frac{1}{18}$									

(c) What is the probability that at least one dice shows a 6?

(d) What is the probability that the sum is at most 10?

(e) What is the probability that a dice shows 4 or the sum is 10?

(f) Given that the sum is 10, what is the probability that one of the dice shows a 4?

14. A large school has the numbers categorised by grade and gender, as shown in the table.

Grade / Gender	Grade 9	Grade 10	Grade 11	Grade 12	Total
Male	180	170	230	220	800
Female	200	130	190	180	700

(a) What is the probability that a student chosen at random will be a female?

(b) What is the probability that a student chosen at random is a male grade 12 student?

(c) What is the probability that a female student chosen at random is a grade 12 student?

(d) What is the probability that a student chosen at random is a grade 12 or female student?

(e) What is the probability that a grade 12 student chosen at random is a male?

(f) Are gender and grade independent of each other? Explain.

15. Some young people do not like to wear glasses. A survey asked a large number of teenaged students whether they needed glasses to correct their vision and whether they used the glasses when they needed to. The table shows the results.

		Used glasses when needed	
		Yes	**No**
Need glasses for correct vision	**Yes**	0.41	0.15
	No	0.04	0.40

(a) Find the probability that a randomly chosen young person from this group:

(i) is judged to need glasses

(ii) needs to use glasses but does not use them.

(b) From those who are judged to need glasses, what is the probability that they do not use them?

(c) Are the events of using and needing glasses independent?

16. Copy this table. Fill in the missing entries.

P(A)	P(B)	Conditions for events A and B	P($A \cap B$)	P($A \cup B$)	P($A \mid B$)
0.3	0.4	Mutually exclusive			
0.3	0.4	Independent			
0.1	0.5			0.6	
0.2	0.5		0.1		

17. In a large graduating class there are 100 students taking the IB examination. 40 students are doing Math/SL, 30 students are doing Physics/SL, and 12 are doing both.

(a) A student is chosen at random. Find the probability that this student is doing Physics/SL, given that they are doing Math/SL.

(b) Are doing Physics/SL and Math/SL independent?

18. A market chain in Germany accepts only Mastercard and Visa. It estimates that 21% of its customers use Mastercard, 57% use Visa, and 13% use both cards.

(a) What is the probability that a customer will have an acceptable credit card?

(b) What proportion of their customers has neither card?

(c) What proportion of their customers has exactly one acceptable card?

19. 132 of 300 patients at a hospital are signed up for a special exercise program which consists of a swimming class and an aerobics class. Each of these 132 patients takes at least one of the two classes. There are 78 patients in the swimming class and 84 in the aerobics class. Find the probability that a randomly chosen patient at this hospital is:

(a) not in the exercise program

(b) enrolled in both classes.

20. An ordinary unbiased 6-sided dice is rolled three times. Find the probability of rolling:

(a) three twos **(b)** at least one two **(c)** exactly one two.

21. An athlete is shooting arrows at a target. She has a record of hitting the centre 30% of the time. Find the probability that she hits the centre:

(a) with her second shot

(b) exactly once with her first three shots

(c) at least once with her first three shots.

22. Two unbiased dodecahedral (12 faces) dice, with faces numbered 1 to 12, are rolled. The scores are the numbers on the top side. Find the probability that:
   (a) at least one 12 shows
   (b) a sum of 12 shows on both dice
   (c) there is a total score of at least 20
   (d) a total score of at least 20 is achieved, given that a 12 shows on one dice.

Figure 8.14 Unbiased dodecahedral dice

23. In April, weather in Germany is unpredictable. On 1 April, Bryan, who lives in Dusseldorf, has a meeting in Munich and plans to travel by aeroplane. Due to weather conditions, there is a 60% chance that his flight will be cancelled. If his flight is cancelled, he plans to drive.

    If he travels by aeroplane, the chances of him being late to the meeting are 10%. If he decides on the car travel, the chances of being late are 25%.
   (a) Set up a complete tree diagram for the possible options.
   (b) Find the probability that Bryan will not be late for the meeting.

    On 1 April Bryan made it to the meeting on time.
   (c) Find the probability that he went by aeroplane.

24. Circuit boards used in electronic equipment are inspected more than once. The process of finding faults in the solder joints on these boards is highly subjective and prone to disagreements among inspectors. In a batch of 20 000 joints, Nick found 1448 faulty joints while David found 1502 faulty ones. All in all, among both inspectors 2390 joints were judged to be faulty. Find the probability that a randomly chosen joint is:
   (a) judged to be faulty by neither of the two inspectors
   (b) judged to be defective by David but not Nick.

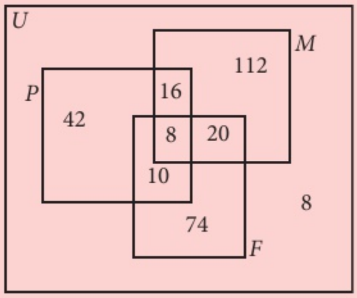
25. An estimate of the world population in 2018 is given in the table. Numbers are in millions.

Age	Male	Female	Total
14 years or less	964	899	1863
15–24	611	572	1183
25–54	1523	1488	3011
55–64	307	322	629
65 and over	284	352	636
Total	3689	3633	7322

   (a) What proportion of the population are females between the age of 15 and 64?
   (b) What proportion of the males are younger than 25?
   (c) What proportion of the 65 years or older generation are females?

Chapter 8 practice questions

1. In a study of least favourite subjects for students in a high school, students mentioned physics (P), mathematics (M) and French (F). The results of the study are shown in the Venn diagram.



   (a) Write down the total number of students in this school.

   (b) A student is chosen at random. What is the probability that the least favourite subject is:

   (i) mathematics only

   (ii) physics or French

   (iii) all three subjects

   (iv) physics or mathematics, but not French.

   (c) What is the probability that a student chosen at random does not include French as one of the least favourites?

2. Beartown has three local newspapers: *The Art Journal*, *The Beartown News*, and *The Currier*.

   A survey shows that

   32% of the town's population read *The Art Journal*,
   46% read *The Beartown News,*
   54% read *The Currier*,
   3% read *The Art Journal* and *The Beartown News* only,
   8% read *The Art Journal* and *The Currier* only,
   12% read *The Beartown News* and *The Currier* only, and
   5% of the population reads all three newspapers.

   (a) Draw a Venn diagram to represent this information. Label A the set that represents *The Art Journal* readers, B the set that represents *The Beartown News* readers, and C the set that represents *The Currier* readers.

   (b) Find the percentage of the population that does not read any of the three newspapers.

   (c) Find the percentage of the population that reads exactly one newspaper.

   (d) Find the percentage of the population that reads *The Art Journal* or *The Beartown News* but not *The Currier*.

A local radio station states that 83% of the population reads either *The Beartown News* or *The Currier*.

(e) Use your Venn diagram to decide whether the statement is true. Justify your answer.

The population of Beartown is 120 000. The local radio station claimed that 34 000 of the town's citizens read at least two of the local newspapers.

(f) Find the percentage error in this claim.

3. Two independent events A and B are given such that

$P(A) = k$, $P(B) = k + 0.3$ and $P(A \cap B) = 0.18$

(a) Find k

(b) Find $P(A \cup B)$

4. Many airport authorities test prospective employees for drug use. This procedure has plenty of opponents who claim that it creates difficulties for some people and that it prevents some others from getting these jobs even if they were not drug users. The claim depends on the fact that these tests are not 100% accurate. To test this claim, assume that a test is 98% accurate in that it identifies a person as a user or non-user 98% of the time. Each job applicant takes this test twice. The tests are done at separate times and are designed to be independent of each other. What is the probability that:

(a) a non-user fails both tests

(b) a drug user is detected (i.e., they fail at least one test)

(c) a drug user passes both tests?

5. In a group of 200 students taking the IB examination, 120 take Spanish, 60 take French and 10 take both.

(a) If a student is selected at random, what is the probability that the student:

(i) takes either French or Spanish

(ii) takes either French or Spanish but not both

(iii) does not take any French or Spanish?

(b) Given that a student takes the Spanish exam, what is the chance that they take French?

6. In a factory producing disk drives for computers, there are three machines that work independently to produce one of the components. In any production process, machines are not 100% fault free. The production after one batch from each machine is listed in the table.

	Defective	Non-defective
Machine I	6	120
Machine II	4	80
Machine III	10	150

(a) A component is chosen at random from the batches. Find the probability that the chosen component is:

(i) from machine I

(ii) a defective component from machine II

(iii) non-defective or from machine I

(iv) from machine I given that it is defective.

(b) Is the quality of the component dependent on the machine used?

7. At a school, the students are organising a lottery to raise money for their community. The tickets consist of coloured envelopes with a small note inside. The note says: 'You won!' or 'No prize.' The envelopes have several colours. They have 70 red envelopes that contain two prizes, and the rest (130 tickets) contain four other prizes.

(a) You want to help this class and you buy a ticket hoping that it does not have a prize. You pick your ticket at random by closing your eyes. What is the probability that your ticket does not have a prize?

(b) You are surprised – you picked a red envelope. What is the probability that you did not win a prize?

8. Two events A and B have the conditions:

$$\mathrm{P}(A|B) = 0.30, \quad \mathrm{P}(B|A) = 0.60, \quad \mathrm{P}(A \cap B) = 0.18$$

(a) Find $\mathrm{P}(B)$.

(b) Are A and B independent? Why?

9. In several ski resorts in Switzerland, the local sports authorities use senior high school students as 'ski instructors' to help deal with the surge in demand during vacations. To become an instructor, you have to pass a test and have to be a senior at your school. Here are the results of a survey of 120 students who are training to become instructors. In this group, there are 70 boys and 50 girls. 74 students took the test, 32 boys and 16 girls passed the test, and the rest, including 12 girls, failed. 10 of the students, including 6 girls, were too young to take the ski test.

(a) Copy and complete the Table 8.5.

	Boys	Girls
Passed	32	16
Failed		12
Training		
Too young		

Table 8.5 Table for question 9

(b) Find the probability that:

(i) a student chosen at random has taken the test

(ii) a girl chosen at random has taken the test

(iii) a randomly chosen boy and randomly chosen girl have both passed the ski test.

10. Events X and Y have the conditions $P(X) = 0.6$, $P(Y) = 0.8$ and $P(X \cup Y) = 1$

Find:

(a) $P(X \cap Y)$

(b) $P(X' \cup Y')$

11. In a survey, 100 managers were asked 'Do you prefer to watch the news or play sport?' Of the 46 men in the survey, 33 said they would choose sport, while 29 women also made this choice.

	Men	Women	Total
News			
Sport	33	29	
Total	46		100

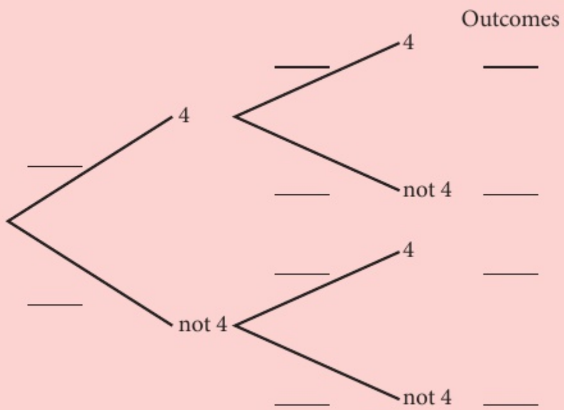
Find the probability that:

(a) a manager selected at random prefers to watch the news

(b) a manager prefers to watch the news, given that the manager is a man.

12. Two unbiased, six-sided dice are rolled, and the total score is noted.

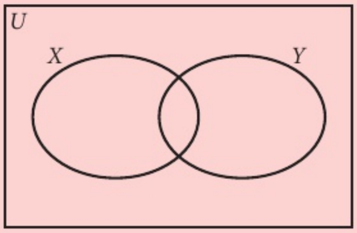
(a) Copy and complete the tree diagram by entering probabilities and listing outcomes.



(b) Find the probability of getting one or more fours.

13. The Venn diagram shows a sample space U and events X and Y.

$n(U) = 36$, $n(X) = 11$, $n(Y) = 6$ and $n(X \cup Y)' = 21$



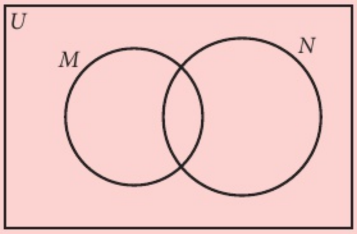
(a) Copy the diagram and shade the region $(X \cup Y)'$

(b) Find:

(i) $n(X \cap Y)$ **(ii)** $P(X \cap Y)$

(c) Are events X and Y mutually exclusive? Explain why or why not.

14. The Venn diagram shows the universal set U and the subsets M and N.



(a) Copy the diagram and shade the area in the diagram which represents the set $M \cap N'$

$n(U) = 100$, $n(M) = 30$, $n(N) = 50$, $n(M \cup N) = 65$

(b) Find $n(N \cap M')$

(c) An element is selected at random from U. What is the probability that this element is in $N \cap M'$?

15. Two fair dice are thrown and the number showing on each is noted. Find the probability that:

(a) the sum of the numbers is less than or equal to 7

(b) at least one dice shows a 3

(c) at least one dice shows a 3, given that the sum is less than 8.

16. For events A and B, the probabilities are $P(A) = \frac{3}{11}$, $P(B) = \frac{4}{11}$

Calculate the value of $P(A \cap B)$ if:

(a) $P(A \cup B) = \frac{6}{11}$

(b) events A and B are independent.

17. Consider events A, B such that $P(A) \neq 0$, $P(A) \neq 1$, $P(B) \neq 0$, and $P(B) \neq 1$.

For each of the below, state whether A and B are mutually exclusive, independent, or neither.

(a) $P(A|B) = P(A)$

(b) $P(A \cap B) = 0$

(c) $P(A \cap B) = P(A)$

18. Sophia is a student at an IB school.

The probability that she will be woken by her alarm clock is $\frac{7}{8}$

If she is woken by her alarm clock the probability that she will be late for school is $\frac{1}{4}$

If she is not woken by her alarm clock the probability that she will be late for school is $\frac{3}{5}$

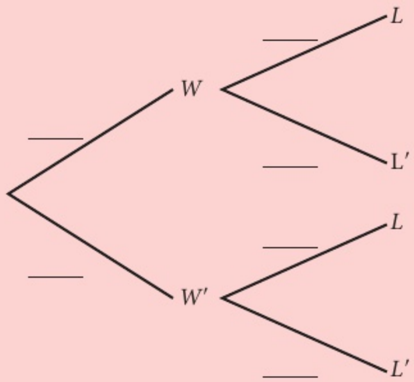
Let W be the event 'Sophia is woken by her alarm clock'.

Let L be the event 'Sophia is late for school'.

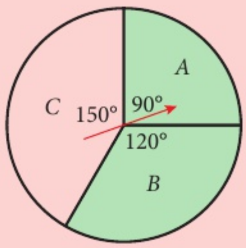
(a) Copy and complete the tree diagram.

(b) Calculate the probability that Sophia will be late for school.

(c) Given that Sophia is late for school what is the probability that she was woken by her alarm clock?



19. The diagram shows a circle divided into three sectors A, B and C. The angles at the centre of the circle are 90°, 120° and 150°. Sectors A and B are shaded as shown.



The arrow is spun. It cannot land on the lines between the sectors. Let A, B, C and S be the events defined by

A: Arrow lands in sector A

B: Arrow lands in sector B

C: Arrow lands in sector C

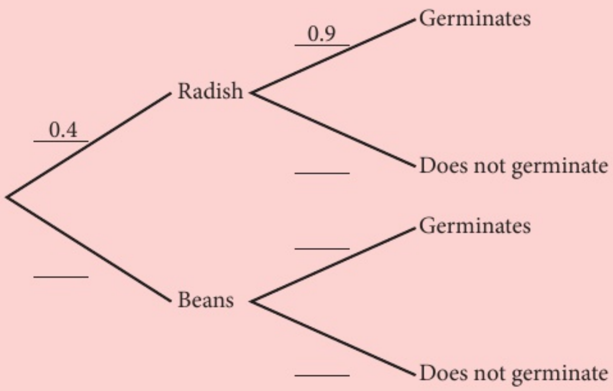
S: Arrow lands in a shaded region.

Find:

(a) $P(B)$ (b) $P(S)$ (c) $P(A|S)$

20. A packet of seeds contains 40% radish seeds and 60% bean seeds. The probability that a radish seed germinates is 0.9, and that a bean seed germinates is 0.8. A seed is chosen at random from the packet.

(a) Copy and complete the probability tree diagram.



(b) **(i)** Calculate the probability that the chosen seed is radish and germinates.

(ii) Calculate the probability that the chosen seed germinates.

(iii) Given that the seed germinates, calculate the probability that it is radish.

21. Two unbiased six-sided dice of different colours are rolled.

Find:

(a) P(the same number appears on both dice)

(b) P(the sum of the numbers is 10)

(c) P(the sum of the numbers is 10 or the same number appears on both dice).

22. The table shows the subjects studied by 210 students at a college.

	Year 1	Year 2	Totals
History	50	35	85
Science	15	30	45
Art	45	35	80
Totals	110	100	210

(a) A student from the college is selected at random.

Let A be the event the student studies art.

Let B be the event the student is in Year 2.

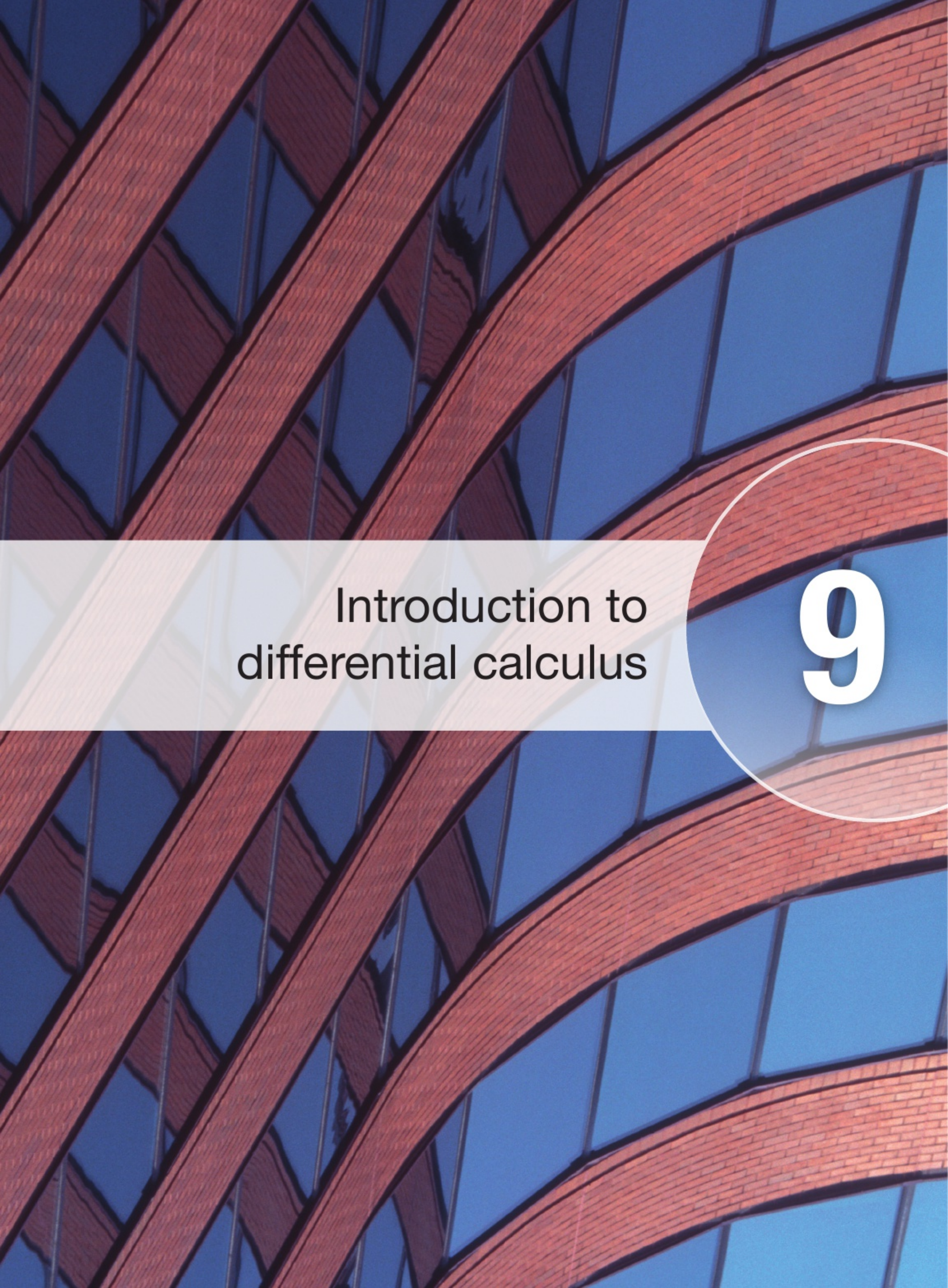
(i) Find P(A)

(ii) Find the probability that the student is a Year 2 art student.

(iii) Are the events A and B independent? Justify your answer.

(b) Given that a history student is selected at random, calculate the probability that the student is in Year 1.

(c) Two students are selected at random from the college. Calculate the probability that one student is in Year 1, and the other in Year 2.



9 Introduction to differential calculus

9 Introduction to differential calculus

Learning objectives

By the end of this chapter, you should be familiar with…

- the concept of a limit, and how to find limits graphically or numerically with technology
- derivatives as functions that give us the instantaneous rate of change (gradient) at any point on a function
- identifying increasing and decreasing intervals and how the sign of the derivative tells us about the graph of a function
- using the power rule to obtain the derivatives of polynomial functions and functions with integer exponents.

Calculus is the mathematics of continuous change.

We deal with change all the time in life, of course, but in this context when we say **change** we really mean **rate of change**.

Calculus is divided into two main parts: **differential calculus** and **integral calculus**. In this chapter we will introduce differential calculus.

By the end of this section, you should be able to answer the following questions:

- What is the difference between **average** rate of change and **instantaneous** rate of change?
- How is the rate of change of a linear function different from the rate of change of other functions?
- How do we use the idea of a limit to reason about the **instantaneous** rate of change?
- How can we estimate the **instantaneous** rate of change?

We will start by introducing the idea of a limit to lay the foundation for answering the remaining questions.

9.1 Limits and instantaneous rate of change

The idea of a limit

A paradox is a contradictory situation which appears to have some logical basis.

Consider the following situation:

Suppose your mathematics teacher offers to play a game with you. If you win, you may leave class. If you lose, you must stay and learn about limits. You decide to try and play the game, which is very simple: you stand one metre from the door to exit the classroom. You may exit the classroom, but each step you take towards the door may cover only half of the remaining distance to the door. Thus, if you start one metre from the door, the first step is $\frac{1}{2}$ metre long, the second step is $\frac{1}{4}$ metre long, the third step is $\frac{1}{8}$ metre long, and so on. Will you ever succeed in leaving the classroom?

The 'game' above is a version of one of several paradoxes attributed to the Greek philosopher Zeno (5th century BCE). Zeno suggested that since there is always a non-zero distance remaining, the student will never be able to leave the classroom. In Figure 9.1 we see the 1 metre distance to the door and the first 4 steps shown:

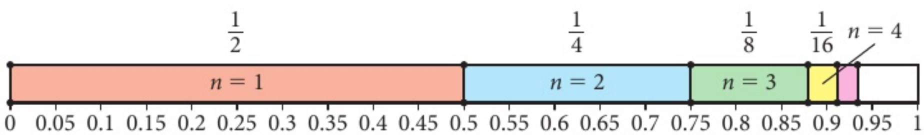


Figure 9.1 We see the 1 metre distance to the door and the first 4 steps

This paradox is useful for introducing the concept of a limit.

Consider Table 9.1, which shows each step, the distance for that step, and the total distance D_n:

Step number (n)	**Distance covered in nth step (in metres)**	**Total distance (D) covered after n steps (in metres)**
1	$\frac{1}{2}$	$\frac{1}{2}$
2	$\frac{1}{4}$	$\frac{1}{2}+\frac{1}{4}=\frac{3}{4}$
3	$\frac{1}{8}$	$\frac{1}{2}+\frac{1}{4}+\frac{1}{8}=\frac{7}{8}$
4	$\frac{1}{16}$	$\frac{1}{2}+\frac{1}{4}+\frac{1}{8}+\frac{1}{16}=\frac{15}{16}$
$\vdots$	$\vdots$	$\vdots$
n	$\frac{1}{2^n}$	$\frac{1}{2}+\frac{1}{4}+\frac{1}{8}+\cdots+\frac{1}{2^n}=?$

Table 9.1 The distance covered by each step and the total distance covered

In the nth step, the student covers a distance of $\frac{1}{2^n}$, and the total distance covered is the sum $D_n = \frac{1}{2}+\frac{1}{4}+\frac{1}{8}+ .... +\frac{1}{2^n}$. We can recognise this as a geometric series with common ratio of $\frac{1}{2}$ and a first term of $\frac{1}{2}$. Therefore, we can calculate the sum as

$$D_n = \frac{a(1-r^n)}{1-r} = \frac{\frac{1}{2}\left(1-\left(\frac{1}{2}\right)^n\right)}{1-\frac{1}{2}} = \frac{\frac{1}{2}\left(1-\left(\frac{1}{2}\right)^n\right)}{\frac{1}{2}} = 1-\left(\frac{1}{2}\right)^n$$

> Recall that for a geometric series with first term a and common ratio r,
> $a + ar + ar^2 + \cdots + ar^{n-1} = \frac{a(1-r^n)}{1-r}$

The functions $D_n = 1 - \frac{1}{2^n}$ and $D = 1$ have been plotted in Figure 9.2, with integer values of n represented by points.

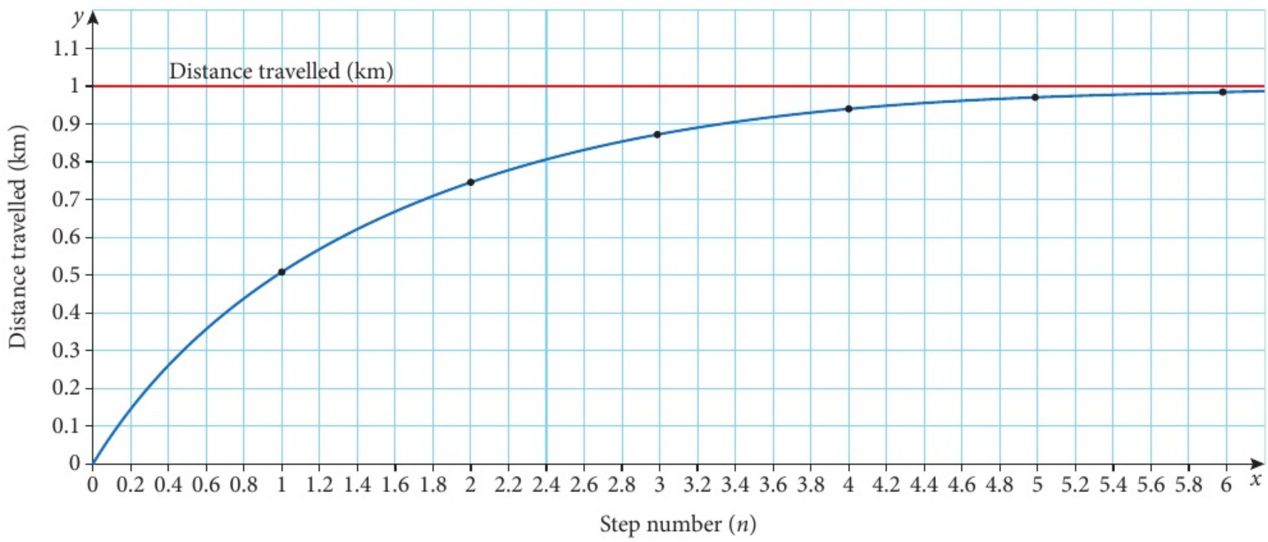


Figure 9.2 $D_n = 1 - \frac{1}{2^n}$ and $D = 1$

So, is Zeno right? Is the poor student trapped forever in the mathematics classroom? What we see is that, in one sense, Zeno is correct: there will always be some distance remaining between the student and the door. However, a clever student can make the following argument: How close do I need to get to the door? Whatever distance we choose, the student will always be able to take some number of steps to get there.

If the value of a function approaches a fixed value, then we say that the **limit** of that function or sequence is the fixed value. For a function $f(x)$ that approaches a fixed value L for some value $x = a$, we write:
$\lim\limits_{x \to a} f(x) = L$

Therefore, since we can get *arbitrarily close* to the door, we say that the value of D_n **approaches 1** (or, equivalently, the distance remaining **approaches zero**).

For our geometric series above, we could write:

$$\lim_{n \to \infty} D_n = 1 \text{ or } \lim_{n \to \infty} \frac{1}{2} + \frac{1}{4} + \frac{1}{8} + \cdots + \frac{1}{2^n} = 1$$

Example 9.1

Find $\lim\limits_{x \to 1} (x + 1)$

Solution

The function $f(x) = x + 1$ is plotted in Figure 9.3, and by inspecting the table of values, as x approaches 1 from both right and left, the value of y approaches 2. Therefore, we can write $\lim\limits_{x \to 1} (x + 1) = 2$

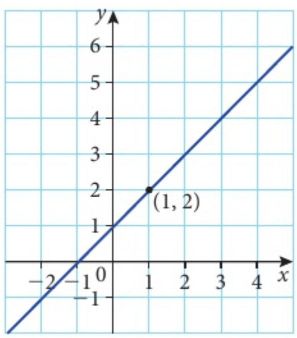


Figure 9.3 Graph for Example 9.1

x	y
0.90	1.90
0.99	1.99
1.00	2.00
1.01	2.01
1.10	2.10

Table 9.2 Table for Example 9.1

Of course, in Example 9.1, the value of the function at $x = 1$ is equal to the limit as x approaches 1. This is not always the case, as Example 9.2 shows.

Example 9.2

Find $\lim\limits_{x \to 3} \dfrac{x^2 - 9}{x - 3}$

Solution

Notice that the function $f(x) = \frac{x^2 - 9}{x - 3}$ is not defined at $x = 3$ (we cannot divide by zero). However, the graph and table of values indicate that as x approaches 3 from both the left and right, f approaches 6. Therefore, we can write $\lim_{x \to 3} f(x) = 6$

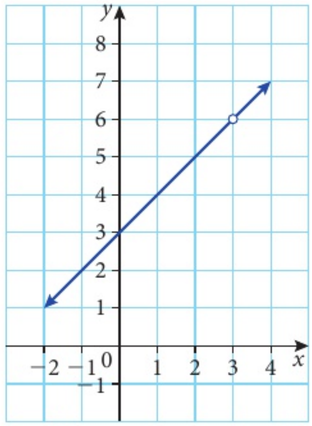


Figure 9.4 Graph for Example 9.2

x	y
2.9	5.9
2.99	5.99
3	undefined
3.01	6.01
3.1	6.1

Table 9.3 Table for Example 9.2

Instantaneous rate of change

What is the difference between **average** rate of change and **instantaneous** rate of change? To answer this question, we start with an example.

Example 9.3

The distance between Lusaka and Livingstone in Zambia is approximately 500 km. On a vacation, it took one of the textbook authors 5 hours to drive from Lusaka to Livingstone. What was the author's average rate of change?

Solution

In this context, average rate of change refers to the author's average speed, given by $\frac{\Delta \text{ distance}}{\Delta \text{ time}}$

So the author's average speed is $\frac{500\,\text{km}}{5\,\text{h}} = 100\,\text{km}\,\text{h}^{-1}$

Does that mean that the author was always driving at $100\,\text{km}\,\text{h}^{-1}$? Of course not. The number $100\,\text{km}\,\text{h}^{-1}$ represents an **average** rate of change. As we will see, we are also interested in the **instantaneous** rate of change. In a car you can see your instantaneous rate of change at any moment by simply looking at the speedometer (see Figure 9.5). But how can we find the instantaneous rate of change when a speedometer is not available?

Figure 9.5 According to this speedometer, this vehicle's instantaneous rate of change is $90\,\text{km}\,\text{h}^{-1}$

The fundamental question of differential calculus is how can we find the instantaneous rate of change?

Recall that we are accustomed to finding the slope of a line, as in the next example.

Example 9.4

Write down the (instantaneous) rate of change for the function $f(x) = 3x - 2$

Solution

For a linear function, the average rate of change and instantaneous rate of change are the same constant value, which is simply the gradient of the line. For a linear function, the gradient is the coefficient of the variable (x in this case), so the instantaneous rate of change is 3.

Remember that gradient and slope mean the same thing.

The word 'curve' is often equivalent to 'function', even if the function is linear.

We are used to finding the slope of a line, but not the slope of a curve, which leads us to this question: What do you think will be different about the gradient of a curve compared to the gradient of a straight line?

To think about this, consider an experiment where a toy car rolls down an inclined ramp (Figure 9.6).

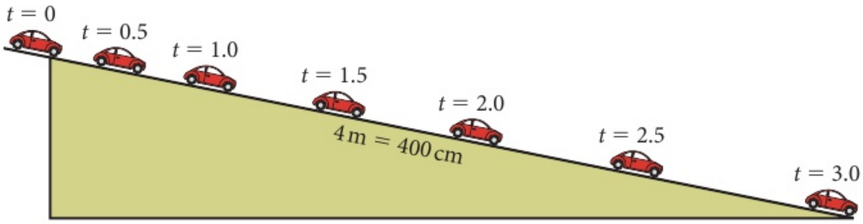


Figure 9.6 Toy car rolling down an inclined ramp

Time (sec)	Distance (cm)
0	0
0.5	28
1.0	69
1.5	127
2.0	203
2.5	294
3.0	400

Table 9.4 Time and distance data for toy car rolling down an incline

It takes the toy car 3 seconds to travel the entire length of the 4-metre ramp. At intervals of half a second, the distance (cm) of the car from the top of the ramp is recorded in Table 9.4.

Velocity is the **rate of change** of position with respect to time: velocity $= \dfrac{\Delta \text{ position}}{\Delta \text{ time}}$. We already know how to calculate the average velocity of the car. For example, the average velocity of the car for the interval from $t = 0$ to $t = 3$ is $\dfrac{\Delta \text{ position}}{\Delta \text{ time}} = \dfrac{400 \text{ cm}}{3 \text{ s}} \approx 133 \text{ cm s}^{-1}$. But, how do we find the instantaneous velocity at $t = 1$?

You should be familiar with expressing speed as $\dfrac{\Delta \text{ distance}}{\Delta \text{ time}}$. Here we use velocity and position instead of speed and distance because we want to allow negative values.

The Greek letter Δ (delta) is commonly used to represent change. For example, the expression Δx means 'the change in x'; Δ distance means 'the change in distance.'

A secant (or chord) is a line or line segment that intersects the graph of a curve at two points.

Figure 9.7 shows a **distance–time graph** for the motion of the toy car from the results in Table 9.4. The average velocity we computed for the interval from $t = 0$ to $t = 3$ is the gradient (slope) of the secant line that passes through the points (0, 0) and (3, 400).

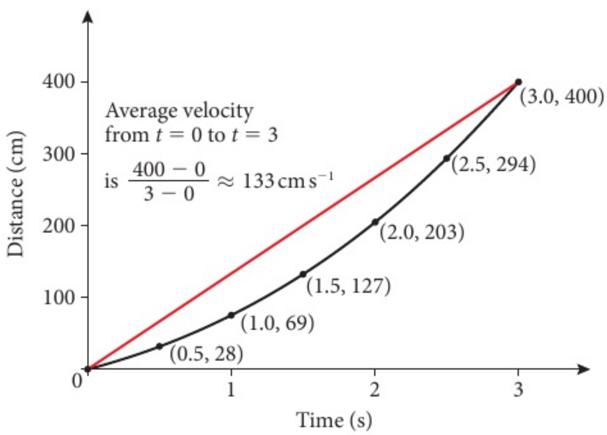


Figure 9.7 Distance–time graph; average velocity from $t = 0$ to $t = 3$

You may have carried out this experiment in a science course. How does the average velocity for the toy car vary over different time intervals?

Suppose that we want to find the instantaneous velocity at time $t = 1$

Figures 9.8 and 9.9 show secant lines whose gradients are the average velocity for the intervals $1 \leqslant t \leqslant 3$ and $1 \leqslant t \leqslant 2.5$ respectively. As we decrease the upper endpoint of the interval from $t = 3$ to $t = 2.5$, we move it closer to $t = 1$. We can continue this process by moving the upper endpoint ever closer to $t = 1$ (Figures 9.10 and 9.11). As we do this, the gradient of the secant line (average) approaches the gradient of the line that is **tangent** to the curve at $t = 1$ (Figure 9.12).

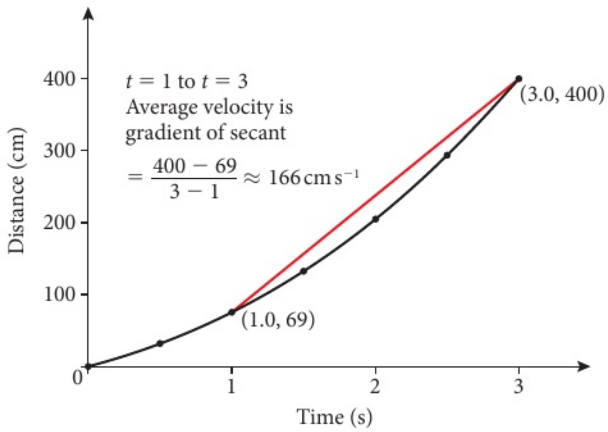


Figure 9.8 Average velocity from $t = 1$ to $t = 3$

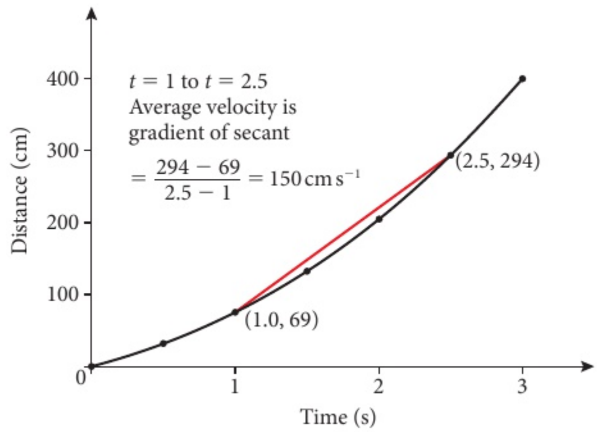


Figure 9.9 Average velocity from $t = 1$ to $t = 2.5$

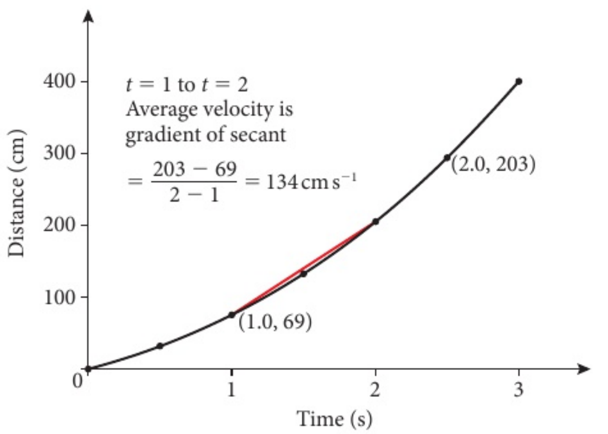


Figure 9.10 Average velocity from $t = 1$ to $t = 2$

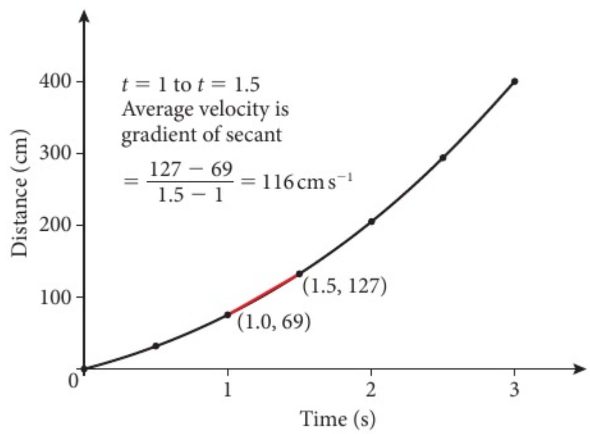


Figure 9.11 Average velocity from $t = 1$ to $t = 1.5$

Using the language of limits, we can say that the gradient of the line **tangent** to the distance–time graph at $t = 1$ (shown in Figure 9.12) is the limit of the gradient of the secant line as the time interval gets narrower and narrower, focusing on $t = 1$

A line that is **tangent** to a curve 'just touches' the curve at a point (the **point of tangency**) but may intersect with the curve at another point, as shown in Figures 9.13 and 9.14.

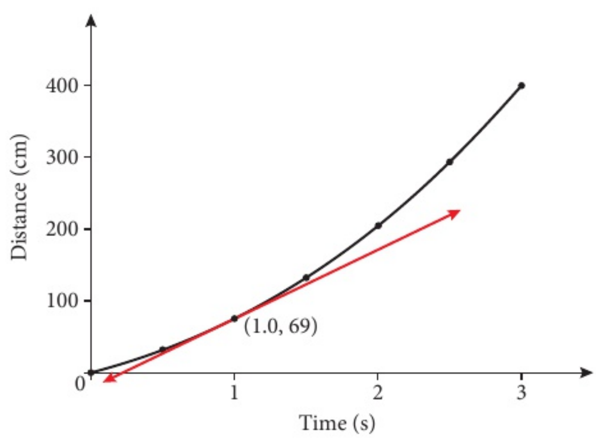


Figure 9.12 Velocity at $t = 1$ is limit of average velocity computed over narrowing time interval focused on $t = 1$

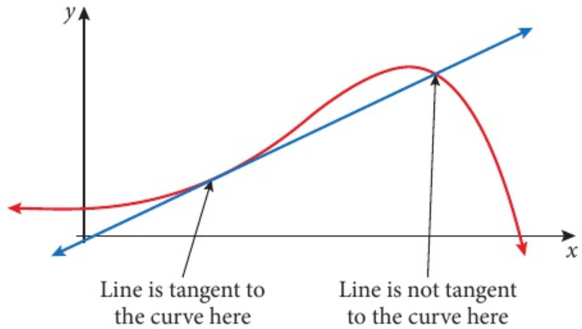


Figure 9.13 Graphical illustration of a tangent line

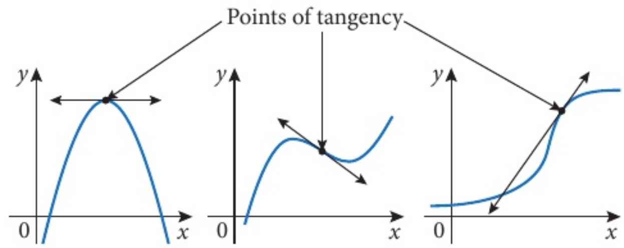


Figure 9.14 Different configurations of tangent lines

So, we are interested in finding the gradient of the line tangent to the curve at $t = 1$. We can approximate the gradient by finding the gradient of secants of progressively smaller intervals. However, the sequence of intervals will approximate the gradient only at $t = 1$; we need a different sequence for each instant we want to find the instantaneous rate of change.

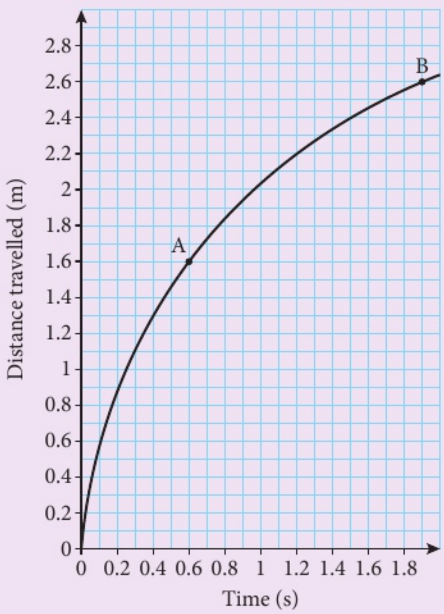
Recall the question we asked earlier: What do you think will be different about the gradient of a curve compared to the gradient of a line? Now, we can answer: a line (or linear function) has a constant gradient – we can represent the gradient for all points on the line with a single number. For a non-linear function, the gradient can have many different values. Therefore, we cannot represent the gradient for all the points on a curve with a single number. Instead, we will need another function – called the **gradient function** or **derivative** – that will give us the gradient at any point we specify. We will examine the gradient function in the next section. We can also estimate the gradient by drawing a tangent by eye and estimating the gradient of the tangent line.

Example 9.5

The distance, d, in metres, covered by a child travelling along a flat waterslide along the ground after time t, in seconds, is shown in the graph.

(a) Describe the speed of the child during the first two seconds.

(b) By copying the graph and drawing suitable tangents to the graph, find the speed of the child at points A and B to two decimal places. Include correct units.



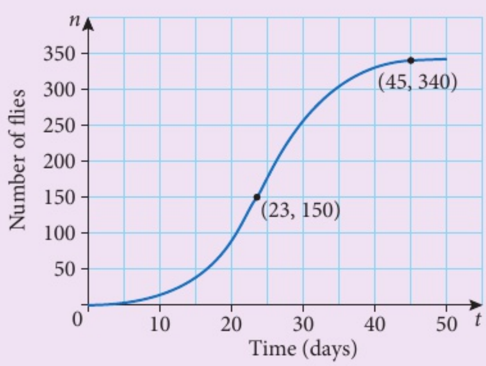
Solution

(a) The distance–time graph is steepest at $t = 0$, and its gradient decreases as t increases. This can be seen using the tangents to the graph at $t = 0,\ 0.1,\ 0.25$ and 0.425, shown on the graph. The child moves fastest upon touching the waterslide, and is subsequently slowed down by friction throughout the duration of the ride.

(b) Since distance is measured in metres, and time in seconds, the units are metres per second (m s^{-1}). By drawing suitable tangents at points A and B and calculating the gradient, the speed of the child at A is approximately $1.26\ \text{m s}^{-1}$ and at B is approximately $0.48\ \text{m s}^{-1}$

Example 9.6

In a laboratory experiment, a group of biologists studies a population of fruit flies over a period of 50 days. The number of fruit flies is recorded at regular intervals and points are plotted on a graph with time t in days on the horizontal axis and number n of flies on the vertical axis. A smooth curve is drawn through the plotted points as shown in the graph.



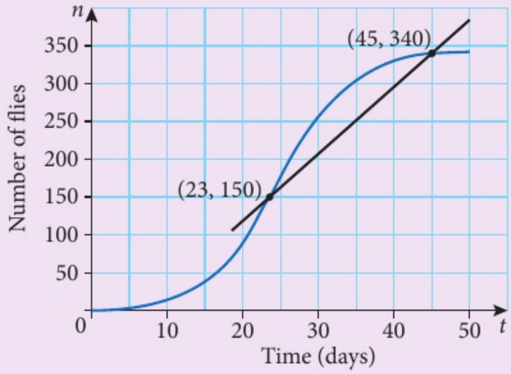
On the 23rd day of the experiment there were 150 flies and on the 45th day there were 340 flies.

(a) Compute the average rate of change of the number n of flies from day 23 to day 45. Give your answer to 3 significant figures and include units.

(b) Copy the graph and draw a line tangent to the graph at the point (23, 150). Use this tangent line to estimate the instantaneous rate of change of the number of flies on day 23.

Solution

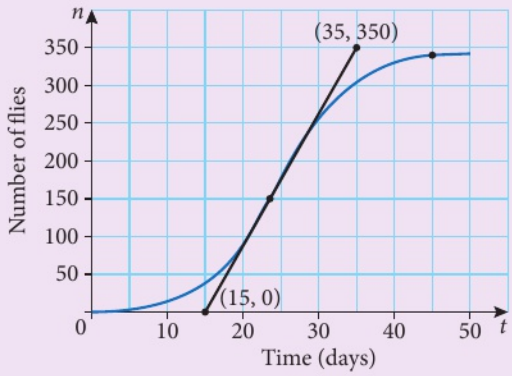
(a)



From day 23 to day 45, $\Delta n = 340 - 150 = 190$ and $\Delta t = 45 - 23 = 22$

Thus, the average rate of change in the number n of fruit flies from day 23 to day 45 is $\dfrac{\Delta n}{\Delta t} = \dfrac{190}{22} \approx 8.64 \dfrac{\text{flies}}{\text{day}}$

(b)



A reasonably accurate tangent line is shown in the diagram, passing through the points (15, 0) and (35, 350). Using these points gives the gradient for the tangent line at the point (23, 150) as $\dfrac{350 - 0}{35 - 15} = 17.5 \dfrac{\text{flies}}{\text{day}}$

Thus, the instantaneous rate of change of the fly population on day 23 is approximately $17.5 \dfrac{\text{flies}}{\text{day}}$. Answers will vary depending on where the tangent line is drawn.

Exercise 9.1

1. Another of Zeno's paradoxes is the following:

   Achilles is in a footrace with a tortoise. Since Achilles knows he runs 10 times as fast as the tortoise, he allows the tortoise a 100 m head start. Then Achilles begins. Achilles quickly runs the first 100 m, during which time the tortoise crawls 10 m.

   Achilles quickly runs the next 10 m, during which time the tortoise crawls 1 m. Achilles quickly runs the next 1 m, during which time the tortoise crawls 0.1 metres. Achilles continues to chase the tortoise, but for each interval he runs the tortoise crawls one-tenth of that distance, so the tortoise is always ahead of Achilles.

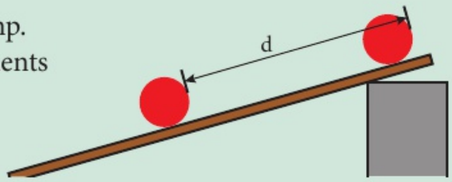
   Can Achilles catch the tortoise? Use the logic of limits to argue that he can.

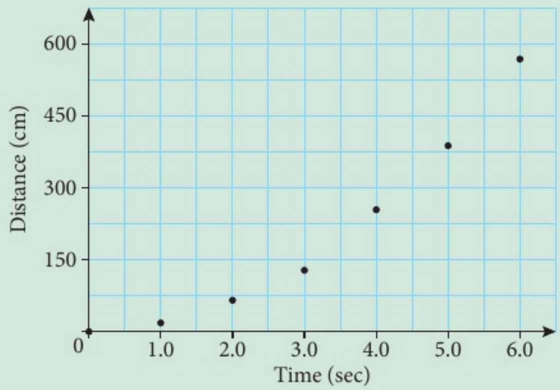
2. Using either a table or a graph on your GDC, evaluate each limit. Give approximate answers to 3 significant figures.

   (a) $\lim_{x \to 2} \dfrac{x^2 - 4}{x - 2}$ **(b)** $\lim_{x \to -1} \dfrac{x^2 - 3x - 4}{x + 1}$ **(c)** $\lim_{x \to 9} \dfrac{9 - x}{3 - \sqrt{x}}$

   (d) $\lim_{x \to 1} \dfrac{x^3 - 1}{x - 1}$ **(e)** $\lim_{x \to 0} \dfrac{\sqrt{x + 2} - \sqrt{2}}{x}$

3. A ball rolls down an inclined ramp. The time and distance measurements are recorded for a 6-second interval. The table and graph show the measurements.





Time t (s)	Distance (cm)
0	0
1	13
2	62
3	136
4	247
5	386
6	553

   (a) Compute the average velocity of the ball during the time intervals:

   (i) $t = 0$ to $t = 2$

   (ii) $t = 2$ to $t = 4$

   (iii) $t = 4$ to $t = 6$

   Give your answers to 3 significant figures, with units.

   (b) Estimate (to 3 significant figures) the instantaneous velocity of the ball at $t = 3$

   Briefly explain your procedure for obtaining the estimate.

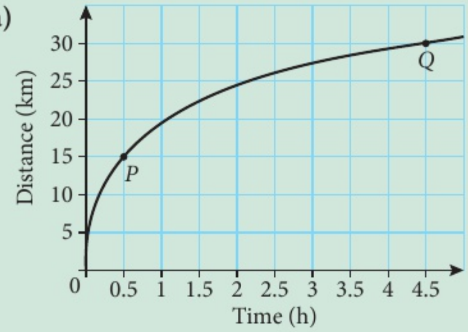
4. In each graph, the distance, d, in km, from a starting point for a bicyclist is plotted against time, t, in hours.

(i) Calculate the average speed of each bicyclist between points P and Q by drawing an appropriate secant.

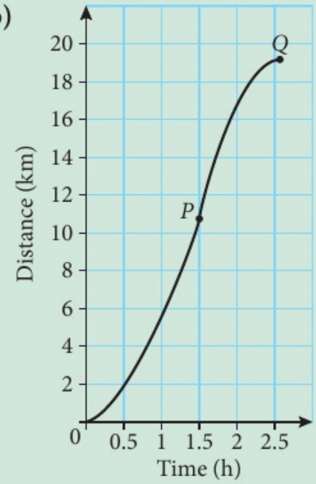
(ii) Estimate the instantaneous speed at point P for each bicyclist by copying the graph and drawing an appropriate tangent.

(iii) Are any of the bicyclists not moving at any time during their ride? If so, which bicyclist and during what time interval? Justify your answer.

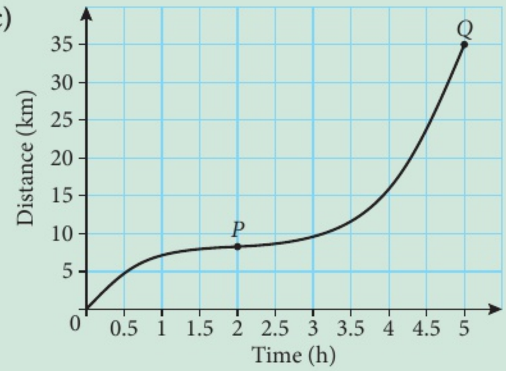
(a)



(b)



(c)

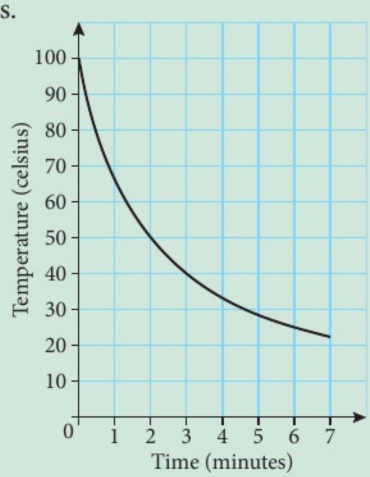


5. Boiling water is poured into a pan. The temperature (in degrees Celsius) of the water is measured regularly during the first seven minutes after the water was poured into the pan. The data points (time, temperature) are plotted on a graph and a smooth curve is drawn through these points. Give all answers to three significant figures.

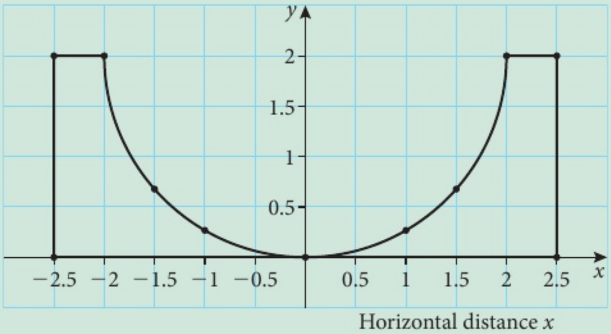
(a) Approximate the average rate of change of the temperature per minute for the interval from $t = 0$ to $t = 7$

(b) Approximate the instantaneous rate of change of the temperature per minute when $t = 1$

(c) Approximate the instantaneous rate of change of the temperature per minute when $t = 5$



6. A half-pipe is a skateboard ramp composed of a semi-circular surface as shown in the diagram.



(a) Describe the gradient of the half-pipe from left to right.

(b) Calculate the slope of the half-pipe at $x = -1.5, -1, 0, 1, 1.5$. How does the slope at x compare to the slope at $-x$?

(c) How would you describe the slope at $x = 2$ and $x = -2$?

7. The points $A\left(-1, \frac{1}{e}\right)$, $B(0, 1)$ and $C(1, e)$ lie on the graph $y = e^x$, as shown. Estimate (to 3 significant figures) the slope of line tangent to the graph of $y = e^x$ at each of the points A, B and C.
Using your GDC find approximate values (to 3 significant figures) for $\frac{1}{e}$ and e. Comment on your results.

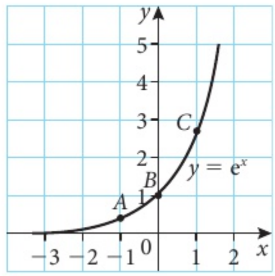


Figure 9.15 Graph for question 7

9.2 Derivative of a function

The **derivative** of a function, also known as the **derivative function** or **gradient function**, is a function that gives us the gradient of the tangent line at any point. We will now look at using and interpreting the derivative function.

Given a function $y = f(x)$, we will refer to its **derivative function** with the notation y', $f'(x)$, or $\frac{dy}{dx}$. Remember, the derivative function is a rule for computing the gradient of the line tangent to the graph of the original function $y = f(x)$ at a particular point, i.e., the instantaneous rate of change of $f(x)$ at a particular value of x.

There are many ways to denote the derivative of a function $y = f(x)$

The most common notations are:

Notation	Read as	Comment
y'	y prime	This notation is brief but does not name the independent variable.
$\frac{dy}{dx}$	'dy by dx' or 'the derivative of y with respect to x'	Names both dependent and independent variables and uses d for derivative.
$f'(x)$	'f prime of x' or 'the derivative of f of x'	Emphasises the name of the function and names the independent variable.

Table 9.5 Common notations for denoting the derivative of a function $y = f(x)$

The notation $\frac{dy}{dx}$ does not indicate d multiplied by y or x. Instead, you can think of dy and dx as names for 'a small change in y' and 'a small change in x.' When we say, 'a small change' we mean a quantity smaller than any real number, but not equal to zero. This notation helps us remember that the derivative function is giving us the instantaneous rate of change in y with respect to x. In this way, it is analogous to how the gradient of a line, m, is defined to be change in y compared to change in x: $m = \frac{\Delta y}{\Delta x}$

Remember that we sometimes use other variables instead of x and y, and not all functions are named f. For example, you may see $h'(t)$, $P'(n)$, $\frac{dh}{dt}$, $\frac{dP}{dn}$, etc.

To see why the derivative function is useful, consider the following example.

Example 9.7

A ball is thrown upwards from the roof of a building. The vertical height of the ball from the ground in metres after t seconds is given by the function $h(t) = -4.9t^2 + 18t + 50$

The derivative function is given as $h'(t) = -9.8t + 18$

(a) Find the initial velocity of the ball.

(b) Find the time at which the velocity of the ball is zero. Interpret the meaning of this.

(c) Find the velocity of the ball at the moment it hits the ground.

Solution

Notice that you need to be careful in choosing when to use the derivative function and when to use the original function. Remember that the derivative function will give you a rate.

(a) Since $h(t)$ is the position function at a given time, then $h'(t)$ represents the change in position with respect to time, i.e., the velocity. Therefore, the initial velocity of the ball is given at time $t = 0$, thus $h'(0) = -9.8(0) + 18 = 18\text{ m s}^{-1}$

(b) The ball has zero velocity when $h'(t) = 0 \Rightarrow -9.8t + 18 = 0 \Rightarrow t = 1.84$ s. This is the time when the ball reaches its maximum height, since it is stationary for an instant between moving up and falling down.

(c) The ball hits the ground when $h(t) = 0$

Therefore we must solve the quadratic equation $-4.9t^2 + 18t + 50 = 0$ to obtain $t = 5.52$ or $t = -1.85$

The negative solution is not meaningful in this context. Therefore, the ball hits the ground at time $t = 5.52$ s, and has a velocity of $h'(5.52) = -9.8(5.52) + 18 = -36.1\text{ m s}^{-1}$

Exercise 9.2

1. Identify the units for the value of the derivative indicated in each of the situations below.

(a) $h'(t)$, where $h(t)$ where h is altitude in km and t is time in hours.

(b) $\frac{dV}{dr}$ for the volume of a cylinder as $V = \pi r^2 h$; r and h are in cm.

(c) $\frac{dF}{ds}$, where F is a force in newtons at a distance s in metres.

(d) $P'(t)$, where P is the population in thousands at a time t in years.

(e) $\frac{dC}{dn}$ where C is the cost in euros to produce n shirts.

(f) $R'(n)$ where $R(n)$ is the revenue in dollars from selling n items.

2. A toy car rolling down a 400 cm ramp travels a distance in cm given by $d(t) = 35t^2 + 40t$ where t is the time in seconds. The derivative function is given as $d'(t) = 70t + 40$

   (a) Find the time it takes the toy car to reach the end of the 400 cm ramp.

   (b) Calculate the average velocity of the toy car.

   (c) Complete the table.

t	Distance travelled	Velocity
0		
1		
2		
3		

   (d) Give a reason why the distance at time $t = 3$ seconds does not make sense in the context of this exercise.

   (e) Write down the velocity of the toy car at the moment it reaches the end of the ramp.

   (f) Write down the velocity of the toy car the moment it begins rolling down the ramp.

   (g) Calculate the time when the toy car has travelled 200 cm.

   (h) Calculate the velocity of the toy car when it is half-way down the ramp.

   (i) Calculate the time when the velocity is $145\,\text{cm s}^{-1}$.

3. An oil leak underneath a car creates a circular puddle. As the oil leaks, the area of the puddle is given by $A = \pi r^2$ where r is the radius of the puddle in cm. The derivative is given by $\frac{dA}{dr} = 2\pi r$

   (a) Find the area of the puddle when the radius is 10 cm.

   (b) What are the units of $\frac{dA}{dr}$?

   (c) Find the average rate of change in the area of the puddle from the time the leak begins until the area is $400\pi\,\text{cm}^2$.

   (d) Find the instantaneous rate of growth of the puddle when the area is $400\pi\,\text{cm}^2$.

4. A runner goes for a long run in the mountains covering 25 kilometres in five hours. Her distance is described by the function

   $d(t) = \frac{2}{3}t^3 - \frac{35}{6}t^2 + \frac{35}{2}t$

   where $d(t)$ is in km and t is in hours, $0 \leqslant t \leqslant 5$

   The derivative function is given as

   $d'(t) = 2t^2 - \frac{35}{3}t + \frac{35}{2}$

(a) Find the distance covered after 2.5 hours.

(b) Verify that the runner completes 25 km at 5 hours.

(c) Find the runner's average speed during the first five hours.

(d) Give the units of $d'(t)$.

(e) Find the runner's speed at the start and end of the run.

(f) Find the time when the runner is going the fastest and the runner's maximum speed.

(g) Find the runner's minimum speed and the time when this occurs.

5. The distance, s (in metres), travelled by a child sliding along a flat waterslide after t seconds can be modelled by the function $s = 2\sqrt{3t}$, where $0 \leqslant t \leqslant 3$

The derivative is given as $\frac{\mathrm{d}s}{\mathrm{d}t} = \sqrt{\frac{3}{t}}$

(a) Give the units of $\frac{\mathrm{d}s}{\mathrm{d}t}$

(b) Find the average speed of the child.

(c) Find the speed of the child at 1, 2, and 3 seconds.

(d) Find the distance travelled by the child after 1, 2, and 3 seconds.

9.3 Interpretations of the derivative

It is important to be able to reason abstractly about the sign of the derivative. Specifically, what does the **sign of the derivative** tell us about the function? For this, it is sufficient to revisit linear functions and consider positive, zero, and negative gradients, as shown in Figure 9.16.

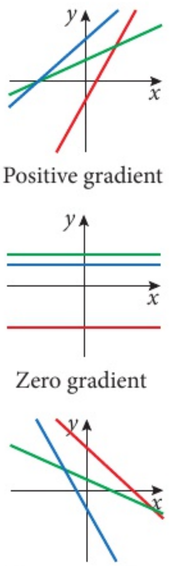


Figure 9.16 Gradients of linear functions

When a gradient is positive, the function is increasing. When a gradient is zero, the function is constant (neither increasing nor decreasing). When a gradient is negative, the function is decreasing. This straightforward reasoning is no different for non-linear functions: we simply use the sign of the derivative to determine if a function is increasing, constant, or decreasing.

For a given function $f(x)$, when

- $f'(x) > 0$, the gradient is positive, the function is increasing at x
- $f'(x) = 0$, the gradient is zero, the function is constant (neither increasing nor decreasing) at x
- $f'(x) < 0$, the gradient is negative, the function is decreasing at x

This simple understanding is critical for our use of the derivative, since knowing whether a function is increasing, constant, or decreasing is very important in many applications.

Example 9.8

Given the derivative $f'(x) = x^2 - x - 12$, identify intervals where the function f is increasing or decreasing.

Solution

There are two approaches. We can analyse the derivative algebraically or graphically with our GDC.

Algebraic approach

Recall that this is a quadratic function and is concave up. So, if we find the x intercepts, then we can reason about the sign of the derivative. The x intercepts are found by setting the derivative function equal to zero and solving:

$$f'(x) = x^2 - x - 12$$
$$0 = (x - 4)(x + 3)$$
$$x = 4 \text{ or } x = -3$$

Hence, the derivative crosses the x axis in two places. Since the derivative function is concave up we can sketch the graph, as shown in Figure 9.17.

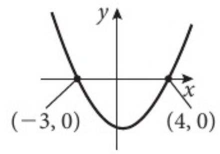


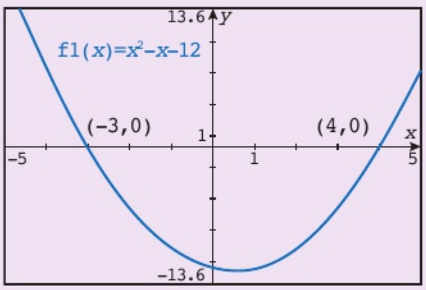
Figure 9.17 The derivative function is concave up

The derivative is positive for $x < -3$ or $x > 4$ and negative for $-3 < x < 4$

Therefore, the function f is increasing for $x < -3$ or $x > 4$ and decreasing for $-3 < x < 4$

Graphical approach

If we graph the derivative using a GDC, we can quickly come to the same conclusion as before.



Looking at the graph, we see that $f'(x)$ is positive for $x < -3$ or $x > 4$ and negative for $-3 < x < 4$. Therefore, the function f is increasing for $x < -3$ or $x > 4$ and decreasing for $-3 < x < 4$

Remember, when you are looking at the graph of a derivative, you need to look **only** at whether the graph is positive (above the x-axis), zero (intersecting the x-axis), or negative (below the x-axis).

Example 9.9

The rate of change in a population is given by the function $\frac{dP}{dt} = \frac{4}{t + 1}$ where P is the population in thousands and t is time in years, $t \geqslant 0$

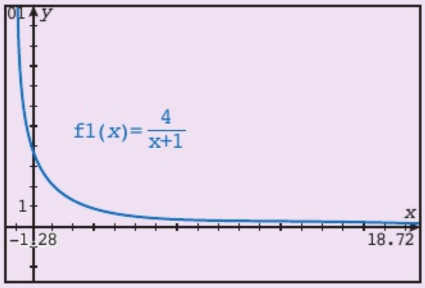
(a) Find the initial rate of change in the population.

(b) Describe how the population changes over time.

Solution

(a) The initial rate of change in the population is $\frac{dP}{dt} = \frac{4}{0 + 1} = 4 \Rightarrow$ increasing by 4000 individuals per year.

(b) To see how the population changes, examine a graph of the derivative on a GDC.



We see that the derivative is positive everywhere, so the population is growing. However, the derivative approaches zero, which means that population growth rate is decreasing. Therefore, the population is levelling off as time progresses.

Example 9.10

Given the graph of the **derivative** $\frac{dy}{dx}$ shown, write down intervals where y is increasing and intervals where y is decreasing.

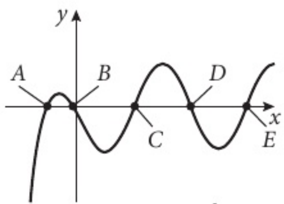


Figure 9.18 Graph of $\frac{dy}{dx}$ for Example 9.10

Solution

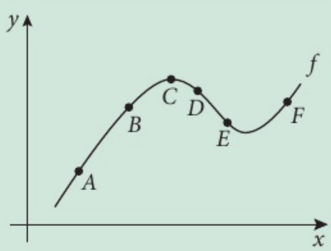
We see that $\frac{dy}{dx}$ is positive for $A < x < B$ and $C < x < D$ and $x > E$

Therefore, y is increasing on those intervals. Likewise, $\frac{dy}{dx}$ is negative for $x < A$ and $B < x < C$ and $D < x < E$, so y is decreasing on those intervals.

One of the important uses of the derivative is to help sketch graphs of functions. While we often use a GDC to generate graphs of functions, understanding the connection between the derivative of a function and the graph of the function is still important.

Exercise 9.3

1. Use the graph of f to answer each question.



(a) Between which two consecutive points is the average rate of change of the function greatest?

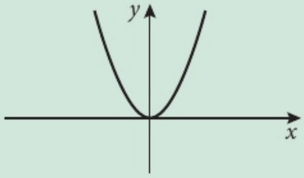
(b) At what points is the instantaneous rate of change of f

(i) positive **(ii)** negative **(iii)** zero

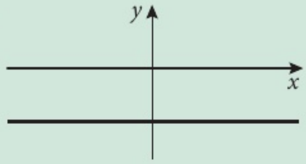
(c) For which two pairs of points is the average rate of change approximately equal?

2. The diagrams in the left-hand column show the graphs of the functions f_1, f_2, f_3, f_4. The diagrams in the right-hand column include the graphs of the derivatives of the functions shown in the left-hand column.

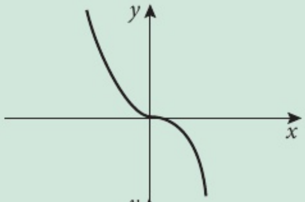
f_1



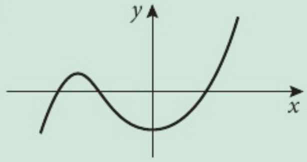
(a)



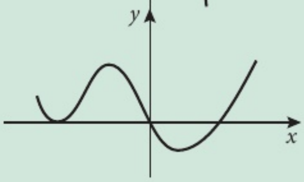
f_2



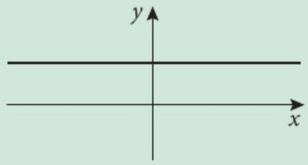
(b)



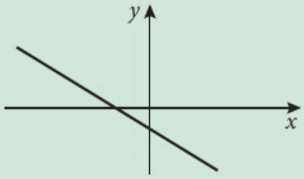
f_3



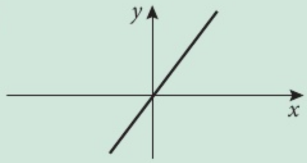
(c)



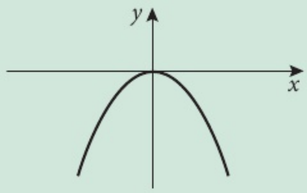
f_4



(d)



(e)

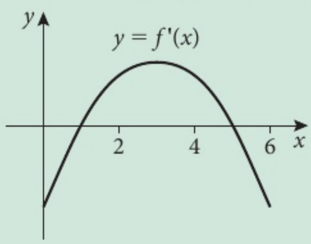


Copy and complete the table by matching each function with its derivative.

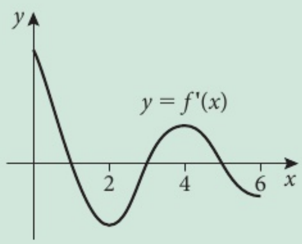
Function	Derivative diagram
f_1	
f_2	
f_3	
f_4	

3. The graphs show the **derivative** of a function f. Identify the intervals where f is increasing or decreasing.

(a)



(b)



4. The function $g(x)$ is defined for $-3 \leqslant x \leqslant 3$

The behaviour of $g'(x)$ is given in the table.

x	$-3 < x < -2$	-2	$-2 < x < 1$	1	$1 < x < 3$
$g'(x)$	negative	0	positive	0	negative

(a) For what intervals is g decreasing?

(b) Given that $g(0) = 0$, make a rough sketch of the graph of g.

5. The torque in newton metres produced by a Chevrolet 383 engine varies according to the rotational speed (rotations per minute, RPM) of the engine according to the model $T(r) = -0.00002720r^2 + 0.08459r + 440.0$ where $T(r)$ is the torque in Nm and r is revolutions min^{-1} (RPM).

Give your answers to 4 significant figures.

(a) Explain the meaning of a positive value for $T'(r)$.

(b) Use your GDC to obtain a graph of $T'(r)$ for the interval $0 \leqslant r \leqslant 10\,000$. On what intervals is the torque increasing? Decreasing? Justify your answer using $T'(r)$.

(c) Use your GDC to determine the maximum torque and the RPM at which this occurs. What is the value of $T'(r)$ at this point?

6. The rate of change in the concentration of a certain drug in a patient's bloodstream is given by the function $\frac{\mathrm{d}C}{\mathrm{d}t} = (-0.646t^{1.15} + 36.8t^{0.15})(0.98^t)$ where C is in nanograms per ml and t is in minutes since the patient was given the drug, for $0 \leqslant t \leqslant 480$

(a) Use your GDC to generate a graph of $\frac{\mathrm{d}C}{\mathrm{d}t}$.

(b) Find intervals where the concentration of the drug is increasing and decreasing.

(c) Find the time at which the concentration is decreasing the fastest and the rate at this time.

(d) Predict when the maximum concentration of the drug occurs.

7. The rate of change in the population of Windelberg can be modelled by the function $P'(t) = \dfrac{12000}{8500(0.82^t) + 1.22^t + 184}$ where P is the population in thousands and t is years since 1980.

(a) Give the units of $P'(t)$.

(b) Use your GDC to generate a graph of this function.

(c) Give a reason why the population of Windelberg does not decrease.

(d) Describe the general pattern of population change in Windelberg.

(e) Find the instantaneous rate of change in the population in 1990.

(f) Find the interval when the population is increasing by at least 20 000 people per year.

(g) Find the time when the population is increasing the fastest and the rate of increase at this time.

9.4 Derivatives of functions of the form $f(x) = ax^n$

So far in this chapter, we have examined the uses and interpretations of the derivative. However, how can we actually find the derivative? It depends on the function. In this section, we will develop a rule for finding the derivative of a function in the form $f(x) = ax^n$ where $a \in \mathbb{R}$ and $n \in \mathbb{Z}$

This rule – called the **power rule** – will allow us to find the derivative of a variety of functions containing terms in the form ax^n, including any polynomial function.

The process of finding a derivative is called **differentiation**.

Exploration of the power rule

Your GDC is able to graph the derivative of any function. Look up instructions for your particular GDC to learn how to do this.

Your GDC cannot tell you what the derivative of a function actually is (in a symbolic form), but it can calculate its numerical value at any point in the function's domain and produce a graph. This is called numerical differentiation.

Example 9.11

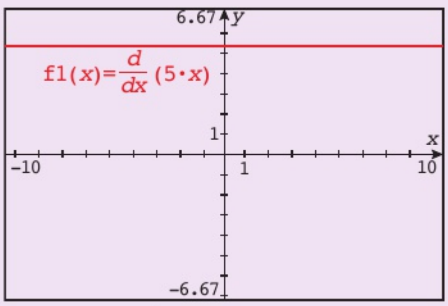
Graph the derivative of each function. For each, attempt to find the derivative function by examining the graph.

(a) $y = 5x$ (b) $y = x^2$ (c) $y = 3x^2$ (d) $y = x^3$ (e) $y = \frac{1}{2}x^3$

Then, summarise your findings and make a conjecture.

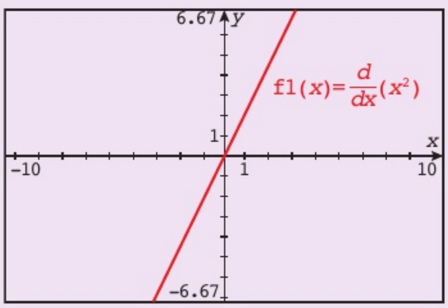
Solution

(a)



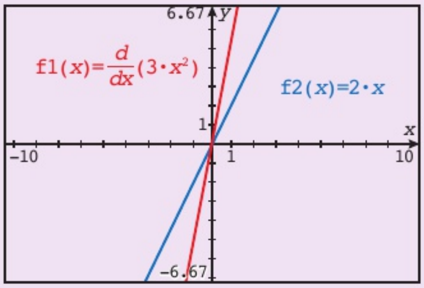
We see that $\frac{\mathrm{d}y}{\mathrm{d}x} = 5$, which agrees with our understanding that the gradient of linear functions of the form $y = mx + c$ is the value of m

(b)



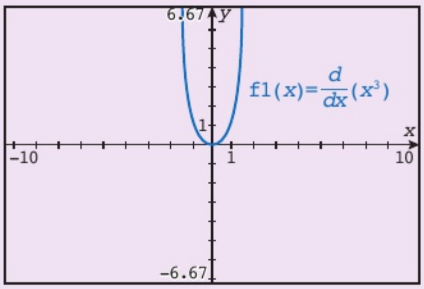
It seems that the derivative of the quadratic function (degree = 2) is a linear function (degree = 1). Look at the gradient of the line. It has a gradient of 2, so $\frac{dy}{dx} = 2x$

(c) The graph of the derivative of $y = 3x^2$ is shown here, but we will also compare it with the derivative from the previous question, $\frac{dy}{dx} = 2x$

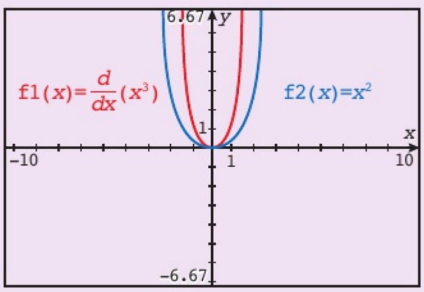


Interestingly, the derivative of $y = 3x^2$ is quite a bit steeper than that of $y = 2x$. If you examine some points or a table of values, you should be able to see that the derivative of $y = 3x^2$ is exactly 3 times as large at the derivative of $y = x^2$. Hence, we have $\frac{dy}{dx} = 3(2x) = 6x$

(d)



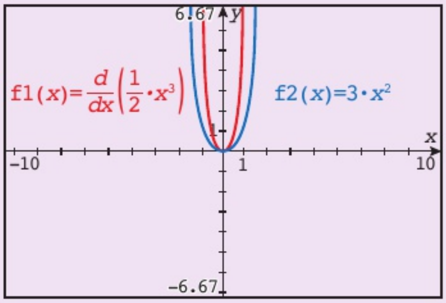
It appears that the derivative of a cubic function (degree = 3) is a quadratic function (degree = 2). Does $\frac{dy}{dx} = x^2$? We can add $y = x^2$ to the graph and compare.



So the derivative of $y = x^3$ is not simply $\frac{dy}{dx} = x^2$. Again, by examining a table of values or checking some of the points on the derivative function, you should be able to convince yourself that the derivative of $y = x^3$ is actually $\frac{dy}{dx} = 3x^2$

(e) The graph of the derivative of $y = \frac{1}{2}x^3$, along with the graph of $y = 3x^2$ for comparison.

As in part (d), the coefficient of $\frac{1}{2}$ has changed the derivative as well.



If you examine it closely, you will see that the derivative is a scaled-down version of $y = 3x^2$. In fact, it appears to be scaled by exactly half.

Hence, we have $\frac{dy}{dx} = \frac{1}{2}(3x^2) = \frac{3}{2}x^2$

Function	Derivative
(a) $y = 5x$	$\frac{dy}{dx} = 5$
(b) $y = x^2$	$\frac{dy}{dx} = 2x$
(c) $y = 3x^2$	$\frac{dy}{dx} = 6x$
(d) $y = x^3$	$\frac{dy}{dx} = 3x^2$
(e) $y = \frac{1}{2}x^3$	$\frac{dy}{dx} = \frac{3}{2}x^2$

Table 9.6 Summary of findings of Example 9.11

Table 9.6 summarises the findings of Example 9.11. The results of Table 9.6 lead us to the power rule.

The power rule
For a function of the form $f(x) = ax^n$, the derivative $f'(x) = anx^{n-1}$
In words: To differentiate a power expression multiply by the exponent, then reduce the exponent by one. In this course, n will be an integer (i.e. $n \in \mathbb{Z}$).

The **power rule** is given that name because it works on power functions; that is, functions of the form $f(x) = ax^n$ where $a \in \mathbb{R}$, $n \in \mathbb{Z}$

What about functions that are more than one term, such as $f(x) = 4x^2 - 10x$?

There is one more rule that helps us: **the derivative of a sum is the sum of the derivatives**. So, since the derivative of $4x^2$ is $8x$ and the derivative of $-10x$ is -10, it must be true that $f'(x) = 8x - 10$

The **derivative of a sum** (or difference) is the sum (or difference) of the derivatives. That is, if $h(x) = f(x) \pm g(x)$, then $h'(x) = f'(x) \pm g'(x)$.

Example 9.12

Use the power rule to find derivatives of each function.

(a) $y = x^4$ (b) $y = 10x^2$ (c) $y = 15$ (d) $y = x^{-2}$

(e) $f(x) = -\frac{1}{3}x^3$ (f) $s = -5t^2$ (g) $g(x) = x^4 - \frac{1}{3}x^3 + 10x^2 + 15$

(h) $y = 5x^2 - 4x + \frac{8}{x}$ (i) $h(x) = \frac{x^2 + 5}{x}$

Solution

(a) $\frac{dy}{dx} = 4 \times x^{4-1} = 4x^3$ (b) $\frac{dy}{dx} = 10 \times 2x^{2-1} = 20x$

(c) For $y = 15$, remember that, by definition, a constant doesn't change. Therefore, it must be true that the rate of change is zero.

Hence, $\frac{dy}{dx} = 0$. Or, you can think of this as $y = 15x^0$ so that the derivative is $\frac{dy}{dx} = 15 \times 0x^{0-1} = 0$

By convention, we usually write expressions with positive exponents, as shown in Example 9.12, part (d), $-\frac{2}{x^3}$ rather than $-2x^{-3}$

(d) A negative power is treated exactly the same:

$$\frac{dy}{dx} = -2 \times x^{-2-1} = -2x^{-3} = -\frac{2}{x^3}$$

(e) $f'(x) = -\frac{1}{3} \times 3x^2 = -x^2$

(f) $\frac{ds}{dt} = -5 \times 2t^{2-1} = -10t$

(g) We already know the derivative of each piece of this polynomial function, we just need to put them together using the sum rule:

$$g'(x) = 4 \times x^{4-1} - 3 \times \frac{1}{3}x^{3-1} + 2 \times 10x^{2-1}$$
$$= 4x^3 - x^2 + 20x$$

(Notice that the $+15$ disappears since its derivative is zero.)

(h) While the first two terms ($5x^2$ and $-4x$) are straightforward, what about $\frac{8}{x}$? Recall that we can rewrite this term because $\frac{8}{x} = 8x^{-1}$

Then, the power rule applies:

$$y = 5x^2 - 4x + \frac{8}{x}$$
$$= 5x^2 - 4x + 8x^{-1}$$
$$\Rightarrow \frac{dy}{dx} = 5 \times 2x^{2-1} - 1 \times 4x^{1-1} + (-1) \times 8x^{-1-1}$$
$$= 10x - 4 - \frac{8}{x^2}$$

Note that the last term becomes negative because of the negative exponent in the original function.

(i) First, rewrite to eliminate the fraction:

$$h(x) = \frac{x^2 + 5}{x} = \frac{x^2}{x} + \frac{5}{x} = x + 5x^{-1}$$

Then find the derivative:

$$h'(x) = 1 - 5x^{-2} = 1 - \frac{5}{x^2}$$

Although the power rule is relatively straightforward, it can be applied to many functions.

Exercise 9.4

1. For each function:

(i) find the derivative

(ii) compute the slope of the graph of the function at the indicated point.

Use a GDC to confirm your results.

(a) $y = 3x^2 - 4x$ point $(0, 0)$

(b) $y = 1 - 6x - x^2$ point $(-3, 10)$

(c) $y = \frac{2}{x^3}$ point $(-1, 2)$

(d) $y = x^5 - x^3 - x$ point $(1, -1)$

(e) $y = (x + 2)(x - 6)$ point $(2, -16)$

(f) $y = 2x + \frac{1}{x} - \frac{3}{x^3}$ point $(1, 0)$

(g) $y = \frac{x^3 + 1}{x^2}$ point $(-1, 0)$

2. Find the coordinates of any points on the graph of the function where the slope is equal to the given value.

(a) $y = x^2 + 3x$ slope $= 3$

(b) $y = x^3$ slope $= 12$

(c) $y = x^2 - 5x + 1$ slope $= 0$

(d) $y = x^2 - 3x$ slope $= -1$

3. The slope of the curve $y = x^2 + ax + b$, where $a, b \in \mathbb{R}$, at the point $(2, -4)$ is -1. Find the value of a and the value of b.

4. Use the power rule to show that the gradient of $y = mx + b$ where m and b are constants, is m.

5. The cumulative number of HIV AIDS cases reported in the United States from 1983 to 1998 follows the cubic model

$C = -222t^3 + 7260t^2 - 12700t + 13500$

where C is the cumulative number of HIV AIDS cases, and t is the number of years since 1983.

(a) Write down the units of $\frac{\mathrm{d}C}{\mathrm{d}t}$

(b) Find the derivative $\frac{\mathrm{d}C}{\mathrm{d}t}$

(c) Find the value of $\frac{\mathrm{d}C}{\mathrm{d}t}$ for 1990 and interpret in context.

(d) Find the number of new cases reported in 1990. Compare to your answer from **(c)**: should these two values be the same? Explain.

(e) Within the years 1983 to 1998, during what intervals was the cumulative number of HIV AIDS cases increasing and decreasing? Justify your answer using $\frac{\mathrm{d}C}{\mathrm{d}t}$

6. A colony of bacteria is exposed to a drug that stimulates reproduction. The number of bacteria is given by the model $P = 1200 + 17t^2 - t^3$ where P is the number of bacteria present t minutes after the drug is introduced, $0 \leqslant t \leqslant 20$

(a) Find $\frac{\mathrm{d}P}{\mathrm{d}t}$

(b) Find intervals where the population of bacteria is increasing and decreasing. Justify your answer using $\frac{\mathrm{d}P}{\mathrm{d}t}$

(c) What is the rate of change in the population of bacteria 10 minutes after the drug is introduced?

(d) Find the time when the bacteria are dying at a rate of 165 per minute.

(e) Use your GDC and the derivative to find when the population of bacteria is increasing the fastest. What is the rate of increase at this time?

7. A tennis player hits a ball straight up. The height of the ball above the ground is described by the model $h(t) = -4.9t^2 + 16t + 1$, where $h(t)$ is the height in metres at time t in seconds after the ball is hit.

(a) Find $h'(t)$, including units, and interpret its meaning.

(b) Find the velocity of the tennis ball 1 second after it is hit.

(c) At what time is the tennis ball descending at a rate of $10\,\mathrm{m\,s^{-1}}$?

(d) At what time is the velocity of the tennis ball equal to zero? What is special about this point?

(e) Find the maximum speed of the tennis ball and the time at which this occurs.

Chapter 9 practice questions

1. Find the derivative $f'(x)$ for each function.

(a) $f(x) = 25x^2 + 25$ (b) $f(x) = 30(x^2 - 4x + 1)$

(c) $f(x) = \dfrac{x^2 - 4x}{x^3}$ (d) $f(x) = \dfrac{1}{x} + \dfrac{x}{2}$ (e) $f(x) = \dfrac{7}{3x^{13}}$

2. The mass on the end of a spring moves up and down. Its height h (in cm) above the ground is given by the function $h(t) = t^3 - 7t^2 + 7t + 21$ for $0 \leqslant t \leqslant 5.5$ seconds.

(a) Find the two values of t when the height of the mass is 6 cm above the ground.

(b) Determine the function $v(t)$ that expresses the velocity of the mass.

(c) Find the two values of t when the velocity of the mass is zero (i.e. momentarily at rest).

(d) Determine the function $a(t)$ that expresses the acceleration of the mass.

(e) Find the acceleration (in $\mathrm{cm\,s^{-2}}$) of the mass at

(i) $t = 0$ (ii) $t = 5.5$

3. The curve $y = ax^3 - 2x^2 - x + 7$ has a gradient of 3 at the point where $x = 2$. Determine the value of a.

4. Consider the function $y = \dfrac{3x - 2}{x}$

(a) Find $\dfrac{dy}{dx}$

(b) Indicate the intervals for which the curve is increasing or decreasing.

5. The displacement s metres of a car, t seconds after leaving a fixed point A, is given by the function $s(t) = 10t - \frac{1}{2}t^2$
   (a) Calculate the velocity of the car when $t = 0$
   (b) Calculate the value of t when the velocity is zero.
   (c) Calculate the displacement of the car from A when the velocity is zero.

6. The cost in thousands of Thai Baht (THB) to produce baseball caps is given by the function $c(x) = 2x^3 - 12x^2 + 30x$ where x is in hundreds of caps. The derivative $c'(x)$ gives the marginal cost.
   (a) Find the derivative $c'(x)$
   (b) Find the marginal cost when 100 ball caps are produced and interpret in context.
   (c) Use $c'(x)$ to show that the cost function increases for all $x > 0$

   The revenue in thousands of THB generated from selling x baseball caps is given by the function $r(x) = 15x$ where x is in hundreds of caps.
   (d) Given that profit = revenue − cost, find a function $p(x)$ for the profit in THB from selling x hundred baseball caps.
   (e) Find the derivative $p'(x)$
   (f) Find intervals where $p(x)$ is increasing and decreasing.
   (g) The derivative $p'(x)$ gives the marginal profit. An optimum production level may be found when marginal profit is zero and $p(x)$ is positive. Find the optimum production level and the expected profit at this level.

7. The number of litres in a large tank of water t minutes after a drain is opened is given by $Q(t) = 300(20 - t)^2, 0 \leqslant t \leqslant 20$
   (a) Find the average rate at which the water is draining in the first 10 minutes.
   (b) Find the rate at which the water is draining at 10 minutes.
   (c) Is the rate at which the water is draining increasing or decreasing? Speeding up or slowing down?

8. The spread of measles in a particular school is modelled by the function $P(t) = \dfrac{200}{1 + 199(1.2)^{-t}}$ where $P(t)$ is the number of students who have measles and t is the number of days since the measles first appeared.
   (a) Use your GDC to predict the maximum number of students infected.
   (b) Use your GDC to obtain a graph of the derivative.
   (c) Find the rate of infection 20 days after the first appearance.
   (d) Find the average rate of infection for the first 20 days.
   (e) Find the maximum rate of infection and the day at which this occurs.
   (f) On which day after the maximum infection rate has passed does the infection rate drop to less than 1 person per day?

9. The value M in euros of a motorcycle t years after it is purchased can be modelled by the function $M(t) = \dfrac{4500}{t} + 750,\ t \geqslant 1$

(a) Find $M'(t)$

(b) Show that $M(t)$ is a decreasing function using $M'(t)$

(c) Find the time at which the value of the motorcycle is decreasing by €1000 per year.

(d) Find the time, according to this function, at which the value of the motorcycle is decreasing the fastest.

(e) Find the time at which the value of the motorcycle is decreasing by less than €50 per year, and the value of the motorcycle at this point.

10. The profit $P(n)$ in thousands of dollars for a company producing t-shirts can be modelled by the function $P(n) = -2n^3 + 6n^2 + n$ where n is the number of t-shirts produced (in thousands).

(a) Find the intervals where the profit is increasing. Justify your response using $P'(n)$

(b) The optimum production level can be found by looking for when the marginal profit, given by $P'(n)$, changes from positive to negative. Find the optimum production level and the expected profit at this level.

11. The fuel efficiency in miles per gallon for vehicles in the USA can be modelled by the function $E(t) = -0.0007t^3 + 0.0278t^2 - 0.0843t + 12$ where t is years since 1970, $0 \leqslant t \leqslant 35$

(a) Find $E'(t)$

(b) Fuel efficiency was improving (increasing) between the years a and b. Find the values of a and b.

12. The rate of change in the number of CD sales from 1991 to 2015 can be modelled by the function $S'(t) = 2.11t^2 - 74.4t + 498$ where $S(t)$ is the sales in millions per year at t years since 1990, $1 \leqslant t \leqslant 25$

(a) Describe the trend of CD sales during this time period.

(b) Find the year in which the CD sales increased by 250 million.

(c) In what year did CD sales first begin to decline?

(d) In which year did the biggest decrease in CD sales happen, and how much did CD sales decrease?

(e) Which graph in Figure 9.19 shows the number of CD sales from 1991 to 2015?

13. An experiment is carried out in which the number n of bacteria in a liquid is given by the formula $n = 650e^{kt}$, where t is the time in minutes after the start of the experiment and k is a constant. The number of bacteria doubles every 20 minutes. Find:

(a) the value of k

(b) the rate of change of the number of bacteria when $t = 90$

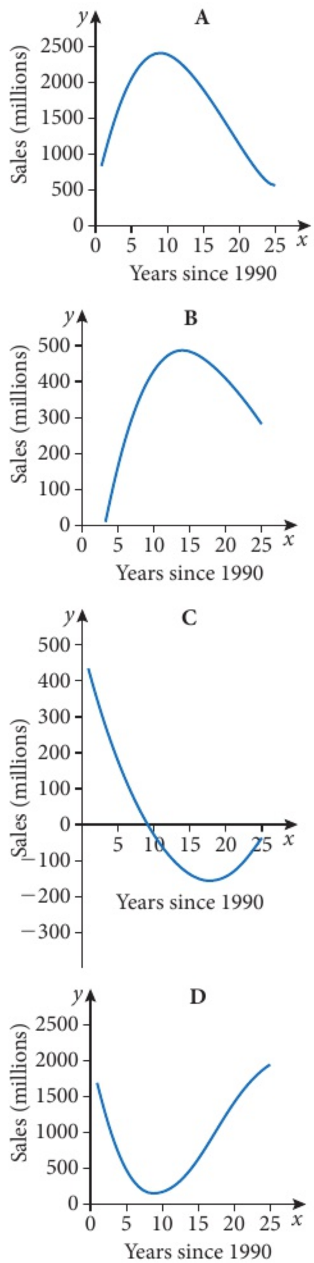


Figure 9.19 Graphs for question 12 (e)

10 Further differential calculus

10 Further differential calculus

Learning objectives

By the end of this chapter, you should be familiar with…

- the relationship between the graph of a function and its derivatives
- finding and testing for maximum and minimum points
- finding the equation of a tangent or a normal at a given point
- finding the optimum solution to a problem, i.e. a maximum or minimum (optimisation).

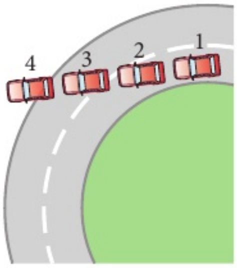


Figure 10.1 A car that loses traction completely in a curve will travel along a tangent to the curve of the road

In Chapter 9, you were introduced to the derivative of a function and learned rules for calculating the derivative. In this chapter, we will apply that knowledge to analyse functions, including finding maxima and minima. We will also learn how to find the equations of tangents and normals. Finally, we will look at an important and powerful use of the derivative: optimisation.

10.1 Maxima and minima – first derivative test

In real life, we often want to find the minimum or maximum value of a function. We will explore applications of this idea in Section 10.3. In this section, we will develop the calculus necessary to successfully analyse these situations.

Consider a function $f(x)$. Recall from Chapter 9 that the first derivative tells us the instantaneous rate of change at a given point on the function. Therefore:

- If $f'(x) > 0$, then the gradient is **positive**; the function is **increasing** at x.
- If $f'(x) = 0$, then the gradient is **zero**; the function is **neither increasing nor decreasing** at x.
- If $f'(x) < 0$, then the gradient is **negative**; the function is **decreasing** at x.

A **local** maximum (or minimum) is the largest (or smallest) value of a function in a particular interval. Informally, we can think of these as the peaks or valleys of the function.

Maxima and minima are referred to together as **extrema**.

This is shown in Figure 10.2. In this section, we will use the first derivative to find **local maxima or minima**.

For a local maximum to occur, the function must increase, then stop (at the maximum), then decrease. Likewise, for a local minimum to occur, the function must decrease, then stop (at the minimum), then increase (Figure 10.2).

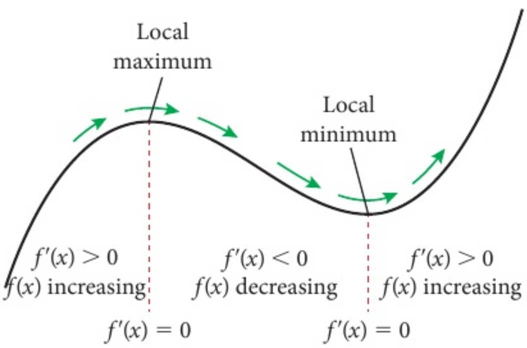


Figure 10.2 Local extrema

From Figure 10.2 we can see that a maximum or minimum occur when the first derivative equals zero and there is a change in the sign of the first derivative.

As shown in Figure 10.2, if $f'(x) = 0$, then:

- a local **maximum** occurs at c if $f'(x)$ changes sign from positive to negative at $x = c$
- a local **minimum** occurs at c if $f'(x)$ changes sign from negative to positive at $x = c$

This is often called the **first derivative test**.

Example 10.1

Find the coordinates of the local maximum or minimum of the function

$f(x) = 2x^2 - 12x + 15$

and identify it as a maximum or minimum.

Solution

While we could use our knowledge of quadratic functions to solve this problem, we will apply our knowledge of calculus to reach the same result. We begin by finding the first derivative:

$f'(x) = 2(2x) - 12 = 4x - 12$

We need to find places where the first derivative is equal to zero:

$f'(x) = 0 \Rightarrow 4x - 12 = 0 \Rightarrow x = 3$

Therefore, there is a **stationary point** at $x = 3$. We find whether it is a maximum or minimum by applying the first derivative test. To do this, we must find the sign of the first derivative for intervals to the left and right of $x = 3$. For this it is helpful to make a **sign diagram**. A sign diagram simply shows the sign of the derivative on various intervals.

A **stationary point** is a point on a function where it is neither increasing nor decreasing, that is, $f'(x) = 0$

We construct the sign diagram by first noting that the value of $f'(3) = 0$. Since 3 is the only value of x where $f'(x) = 0$, it divides our number line into two intervals, greater than 3 and less than 3.

Next, we check the sign of $f'(x)$ by testing values of x to the left and right of $x = 3$.

When choosing values of x to test in the first derivative, pick values within the interval that are easy to calculate.

To the left, we can test the sign of $f'(x)$ at $x = 0$:

$f'(0) = 4(0) - 12 = -12$, so $f'(x)$ is **negative**.

To the right, we can test the sign of $f'(x)$ at $x = 10$:

$f'(10) = 4(10) - 12 = 28$, so $f'(x)$ is **positive**.

Then we label our sign diagram accordingly.

Now we simply interpret the sign diagram to answer the question: Since the sign of $f'(x)$ changes from negative to positive at $x = 3$, there must be a **local minimum** at $x = 3$

To find the coordinates of the point, find the value of $f(x)$ at $x = 3$

$f(3) = 2(3)^2 - 12(3) + 15 = 18 - 36 + 15 = -3$

Therefore, the local minimum occurs at $(3, -3)$

Why is it only necessary to test one point in each interval? Remember that as long as the derivative is continuous (no breaks in its curve) then the only way it can change from positive to negative (or vice-versa) is for it to pass through zero. Thus, we find the places where the derivative is zero and then we only need to test one point in each interval created by those values.

You may have noticed in the previous example that we avoided using the familiar formula for the x-coordinate of the vertex of a parabola, $x = -\frac{b}{2a}$

Now that you understand calculus, you can see where this formula comes from.

Example 10.2

Prove that the x-coordinate of the vertex of $f(x) = ax^2 + bx + c$ is given by $x = -\frac{b}{2a}$

Solution

Since the x-coordinate of the vertex must occur at a local minimum or maximum, it must be at a stationary point. Hence, the x-coordinate of the vertex must be at a point where $f'(x) = 0$. So, we find the first derivative of $f(x)$:

$f'(x) = (2)ax + b = 2ax + b$

Since the first derivative must equal zero, we have

$f'(x) = 0 \Rightarrow 2ax + b = 0 \Rightarrow 2ax = -b \Rightarrow x = -\frac{b}{2a}$

Therefore, the x-coordinate of the vertex of $f(x) = ax^2 + bx + c$ is given by $x = -\frac{b}{2a}$

As Example 10.3 shows, when our original function is more complicated, we may need to find more than one stationary point.

Example 10.3

Find the x-coordinates of all local maxima and minima for the function

$f(x) = x^3 - 4x^2 - 3x + 12$

and identify them as maxima or minima.

Solution

We will proceed as in Example 10.1, by finding the first derivative:

$f'(x) = 3x^2 - 8x - 3$

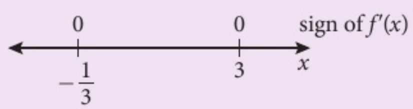
And then finding values of x which make it zero:

$f'(x) = 0 \Rightarrow 3x^2 - 8x - 3 = 0$

This quadratic equation can be factorised into

$(3x + 1)(x - 3) = 0 \Rightarrow x \in \left\{-\frac{1}{3}, 3\right\}$

Then we create the sign diagram:



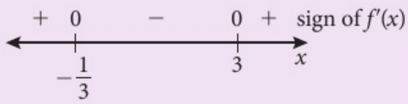
We can test at $x = -1, 0, 10$:

$f'(-1) = 3(-1)^2 - 8(-1) - 3 = 3 + 8 - 3 = 8 \Rightarrow +$

$f'(0) = 3(0)^2 - 8(0) - 3 = -3 \Rightarrow -$

$f'(10) = 3(10)^2 - 8(10) - 3 = 300 - 80 - 3 = 217 \Rightarrow +$

And now we can label our sign diagram accordingly.



By interpreting the sign diagram, we see:

- There is a local maximum at $x = -\frac{1}{3}$
- There is a local minimum at $x = 3$

Although the first derivative test can identify **local** maxima and minima, sometimes we are looking for the **absolute** maximum or minimum, that is, the largest or smallest value of a function within some domain. To do this, we need to find all the local maxima and minima, and then compare to see which one is the greatest/least value for the entire domain of f. When the domain is limited, we must also include values at the endpoints of the domain, because the values at the endpoints of the domain are always **local** maxima or minima even though the first derivative may not equal zero. Suppose the domain endpoint is at $x = c$, then we can use the first derivative test in a 'one-sided' manner (see key point).

For functions where the domain has endpoints, the endpoints are local extrema as shown in the graphs in Figures 10.3 and 10.4.

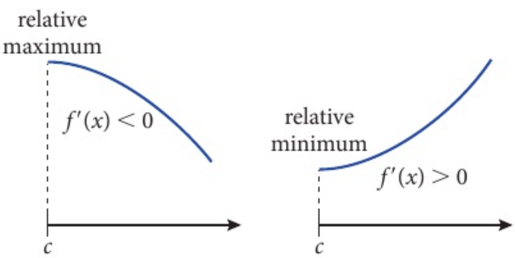


Figure 10.3 Left-hand endpoints of domain (lower bounds)

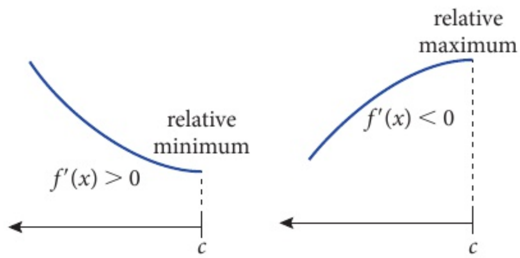


Figure 10.4 Right-hand endpoints of domain (upper bounds)

Example 10.4

Given $h(t) = 4t^3 + 3t^2 - 6t$ with domain $-\frac{3}{2} \leqslant t \leqslant 2$

(a) Find the coordinates of all local extrema and classify them as maxima or minima.

(b) Write down the coordinates of the absolute maximum and absolute minimum.

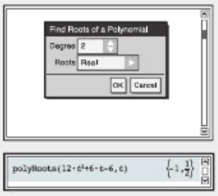
Solution

(a) Find the first derivative:

$h'(t) = 12t^2 + 6t - 6$

Then, find values of t for which the first derivative is zero (finding stationary points):

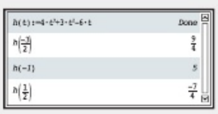
Your GDC has built-in tools for solving polynomials (including quadratics). Learn to use them! Here is one example, for the quadratic in Example 10.4:



The solutions to this quadratic are $t = -1$ or $t = \frac{1}{2}$

Remember that we don't need to find the exact value of $h'(t)$. We just need to find out whether it is positive or negative.

Most GDCs can define functions, saving you repetitive calculations and reducing error. Here is a screenshot of an example:



$$12t^2 + 6t - 6 = 0$$
$$2t^2 + t - 1 = 0$$

Solve the quadratic with factoring, quadratic formula or use your GDC (see margin box).

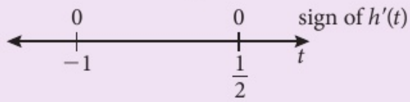
$(2t - 1)(t + 1) = 0$

$2t - 1 = 0$ or $t + 1 = 0$

$t = \frac{1}{2}$ or $t = -1$

Therefore, we have stationary points at $t = -1$ or $t = \frac{1}{2}$

We can draw a sign diagram to visualise the intervals we need to test for the first derivative test.



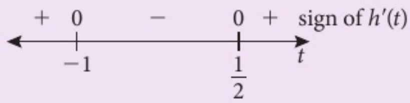
For example, test $t = -2, 0, 1$ to obtain the sign of $h'(t)$:

$h'(-2) = 12(-2)^2 + 6(-2) - 6 = 48 - 12 - 6 > 0$

$h'(0) = 12(0)^2 + 6(0) - 6 = 0 + 0 - 6 < 0$

$h'(1) = 12(1)^2 + 6(1) - 6 = 12 + 6 - 6 > 0$

Now, we update our sign diagram.



We see there is a local maximum at $t = -1$ and a local minimum at $t = \frac{1}{2}$

However, we have not yet considered the endpoints. Since the first derivative is positive at both the lower and upper bound of the domain, the lower bound will be a local minimum, while the upper bound will be a local maximum. So, we have four local extrema to consider:

Local minimum (lower bound of domain)	Local maximum	Local minimum	Local maximum (upper bound of domain)
$t = -\frac{3}{2}$	$t = -1$	$t = \frac{1}{2}$	$t = 2$

To find the coordinates, we must find the value of the function at each t value:

$h\left(-\frac{3}{2}\right) = 4\left(-\frac{3}{2}\right)^3 + 3\left(-\frac{3}{2}\right)^2 - 6\left(-\frac{3}{2}\right) = \frac{9}{4}$

$h(-1) = 4(-1)^3 + 3(-1)^2 - 6(-1) = 5$

$h\left(\frac{1}{2}\right) = 4\left(\frac{1}{2}\right)^3 + 3\left(\frac{1}{2}\right)^2 - 6\left(\frac{1}{2}\right) = -\frac{7}{4}$

$h(2) = 4(2)^3 + 3(2)^2 - 6(2) = 32$

So, the coordinates and types of local extrema are:

Local minimum (lower bound of domain)	Local maximum	Local minimum	Local maximum (upper bound of domain)
$\left(-\frac{3}{2}, \frac{9}{4}\right)$	$(-1, 5)$	$\left(\frac{1}{2}, -\frac{7}{4}\right)$	$(2, 32)$

(b) Now that we have analysed all local extrema, it is easy to see that $(2, 32)$ is an absolute maximum and $\left(\frac{1}{2}, -\frac{7}{4}\right)$ is an absolute minimum.

Example 10.5

A toy rocket is launched upwards into the air. Its vertical position, s metres, above the ground at t seconds is given by $s(t) = -5t^2 + 18t + 1$.

(a) Find the average velocity over the time interval from $t = 1$ second to $t = 2$ seconds.

(b) Find the instantaneous velocity at $t = 1$ second.

(c) Find the maximum height reached by the toy rocket and the time at which this occurs.

Solution

(a) $v_{avg} = \dfrac{s(2) - s(1)}{2 - 1} = \dfrac{[-5(2)^2 + 18(2) + 1] - [-5 + 18 + 1]}{1} = 3$ metres per second (or m s^{-1})

(b) $s'(t) = -10t + 18 \Rightarrow s'(1) = -10 + 18 = 8\,\text{m s}^{-1}$

(c) $s'(t) = -10t + 18 = 0 \Rightarrow t = 1.8$ Thus, s has a stationary point at $t = 1.8$

We know t must be positive and ranges from time of launch ($t = 0$) to when the rocket hits the ground, i.e. $h = 0$

Since it doesn't make sense to include negative heights, we should check the position function to establish the domain of the function:

Remember, your GDC can be used to solve quadratic equations.

$s(t) = -5t^2 + 18t + 1 = 0$

$$\Rightarrow t = \frac{-18 \pm \sqrt{18^2 - 4(-5)(1)}}{2(-5)} \Rightarrow t \approx -0.0547 \text{ or } t \approx 3.66$$

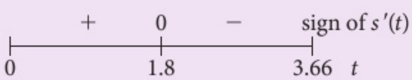
So, the rocket hits the ground about 3.66 seconds after the time of launch. Hence, the domain for the position (s) and velocity (v) functions is $0 \leqslant t \leqslant 3.66$

Therefore, the function s has three points we should check for a maximum: $t = 0$, $t = 1.8$ and $t \approx 3.66$

Applying the first derivative test, we check the sign of the derivative, $s'(t)$, for values on either side of $t = 1.8$, for example $t = 0$ and $t = 2$

$s'(0) = 18 > 0$ and $s'(2) = -2 < 0$

Our sign diagram looks like this:



Both of the domain endpoints, $t = 0$ and $t \approx 3.66$, are local minima so we can ignore them. Since the function changes from increasing to decreasing at $t = 1.8$ and $s(1.8) = -5(1.8)^2 + 18(1.8) + 1 = 17.2$, the toy rocket reaches a maximum height of 17.2 metres at 1.8 seconds after it is launched.

Sometimes we may encounter functions for which it is difficult to find a derivative algebraically. Your GDC can be a powerful tool for analysing these functions, because it has the ability to calculate **numerical** derivatives. That is, your GDC can tell you the value of the derivative at any point, and can therefore make a graph of the derivative, but it cannot give you an algebraic expression for the derivative.

For example, consider the previous example. We can type the position function into a GDC and generate a graph:

Remember that most GDCs expect that 'x' is the independent variable for a graph, so we have replaced t with x here.

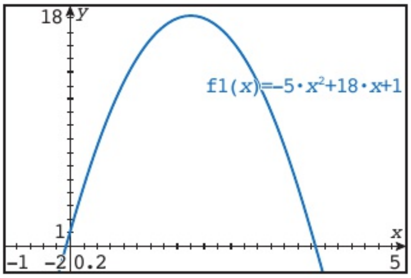


Figure 10.5 Type in the position function to generate a graph

Then, use a GDC to plot the derivative of that function:

You will need to look up how to use your GDC to find the numerical derivative. It may have mathematical notation as shown here $\left(\frac{\mathrm{d}}{\mathrm{d}x}\right)$, or it may have a command like nDeriv(f1(x),x). In either case, you will need to tell the calculator to differentiate with respect to a variable you specify. In this case, we are differentiating with respect to x.

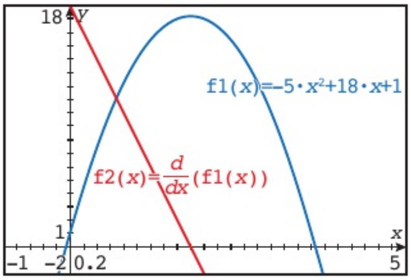


Figure 10.6 Plot the derivative of the function

Once we see the derivative function, we can use the GDC to find the zero. Now, we just need to interpret the results. Since $s'(1.8) = 0$, and the derivative is positive to the left and negative to the right of $t = 1.8$, we can conclude from the first derivative test that there must be a local maximum at $t = 1.8$. Furthermore, by looking at the graph of the position function (in blue), we see that the position does indeed reach a maximum at $t = 1.8$. Since we have already entered the position function, we can also use our GDC to find the maximum height.

Be careful reading graphs with both a function and its derivative. Strictly speaking, we should **not** graph a function and its derivative on the same axes, because the units on the y axes are different: remember that the 'y' value of the derivative is the rate of change in the original function.

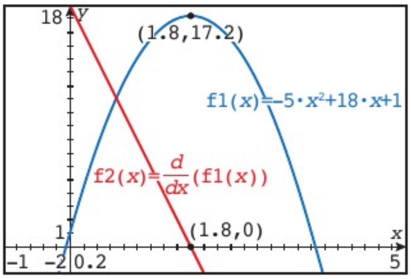


Figure 10.7 Finding the maximum height

This confirms our work above, showing that the toy rocket reaches a maximum height of 17.2 metres, 1.8 seconds after launch.

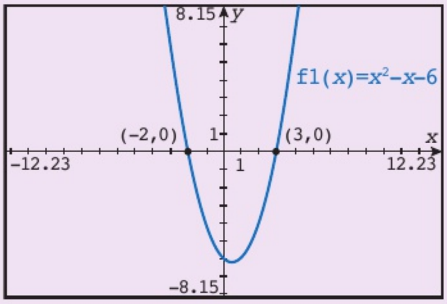
In fact, sometimes it is the case that we know the derivative function, but not the original function. In this case, a GDC is a powerful tool for finding extrema.

Example 10.6

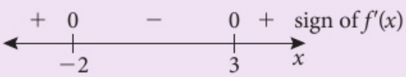
Given $f'(x) = x^2 - x - 6$, find the x-coordinates of all local extrema and identify them as maxima or minima.

Solution

We use the GDC to obtain a graph of the derivative, and find the x values for which $f'(x) = 0$



This graph can be interpreted in the same way as a sign diagram; it is equivalent to the following sign diagram:



Therefore, we conclude that there is a local maximum at $x = -2$ and a local minimum at $x = 3$

Be careful! Don't look at maxima or minima of the derivative graph. Instead, look for zeros, and where the graph is positive (above the x-axis) and negative (below the x-axis).

More complicated functions can be treated in the same way.

Example 10.7

A fruit distributor has noticed that the weekly change in the price of apples can be modelled by the function

$$PC = -\frac{13}{52}\sin\left(\frac{90}{13}(x - 5)\right)$$

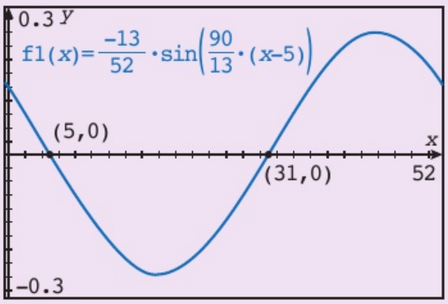
where PC is the change in the price of a kilogram of apples and x is the number of weeks since that start of the year.

(a) The fruit distributor would like to plan a sale when the price of apples will be at a minimum. What week should the sale be?

(b) Sales decrease when the price of apples approaches its maximum. In what week will the price of apples be highest?

Solution

(a) Since the given function predicts the **change** in the cost of apples, we can apply the first derivative test to find the minimum price of apples. Use a GDC to graph and locate the points where $PC = 0$



Since the price change (derivative) function is negative to the left of $x = 31$ and positive to the right, we conclude that the minimum price of apples is during week 31. The sale should be planned for week 31.

(b) With the same GDC graph, we see that the price change (derivative) function is positive to the left of $x = 5$ and negative to the right, so we conclude that the maximum price of apples is at week 5.

Exercise 10.1

1. For the following quadratic functions, find the vertex of the parabola using differentiation.

(a) $y = x^2 - 2x - 6$ **(b)** $y = 4x^2 + 12x + 17$

(c) $y = -x^2 + 6x - 7$

2. For each of the functions below:

(i) find the coordinates of any stationary points for the graph of the equation

(ii) state, with reasoning, whether each stationary point is a minimum, maximum or neither

(iii) sketch a graph of the equation and indicate the coordinates of each stationary point on the graph.

(a) $y = 2x^3 + 3x^2 - 72x + 5$ **(b)** $y = \frac{1}{6}x^3 - 5$

(c) $y = -x(x - 3)^2$ **(d)** $y = x^4 - 2x^3 - 5x^2 + 6$

(e) $y = -x^3 + 6x^2 - 15x - 2$ **(f)** $y = x - \sqrt{x}$

3. For each of the functions below, find any local extrema and give the coordinates of each extremum.

(a) $f(x) = x^3 - 12x$ **(b)** $f(x) = \frac{1}{4}x^4 - 2x^2$

(c) $f(x) = x + \frac{4}{x}$ **(d)** $f(x) = -3x^5 + 5x^3$

(e) $f(x) = 3x^4 - 4x^3 - 12x^2 + 5$

4. For each of the derivative functions below:

(i) use your GDC to graph the derivative given

(ii) hence, find the locations of all local extrema and justify each using the first derivative test.

(a) $f'(x) = x^2 + 3x - 4$ **(b)** $h'(t) = t^2 - 7t + 10$

(c) $\frac{dy}{dx} = \frac{-x}{\sqrt{25 - x^2}}$

5. For each of the functions below:

(i) use your GDC to sketch a graph of the derivative of the function

(ii) hence, find the locations of all local extrema and justify each as a maximum or minimum using the first derivative test.

Verify your results by examining the graph of the function.

(a) $y = \sqrt{36 - x^2}$ (b) $f(x) = \sqrt{x^2 + 5} + \sqrt{25 - x^2}$

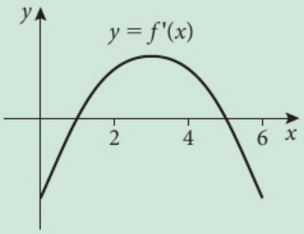
(c) $y = x^4 - 2x^3 - 5x^2 - 6$

6. The graphs of the derivative of a function f are shown below.

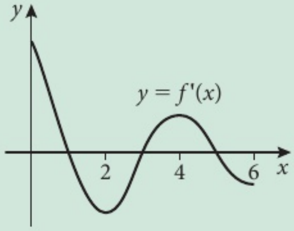
(i) For what intervals is f increasing or decreasing?

(ii) At what value(s) of x does f have a local maximum or minimum?

(a)



(b)



7. An object moves along a line such that its displacement s metres from the origin O is given by $s(t) = t^3 - 4t^2 + t$

(a) Find a function for the object's velocity in terms of t.

(b) For the interval $-1 \leqslant t \leqslant 3$, find the maximum displacement and the time at which this occurs.

(c) In words, accurately describe the motion of the object during the interval $-1 \leqslant t \leqslant 3$

8. For each function, use your GDC to graph the **derivative** of the function and find any relative extrema. State the coordinates of any such points and classify as maxima or minima. Graph the original function on your GDC to verify graphically.

(a) $f(x) = -\frac{1}{2}x^4 + 4x^2 + 10$ **(b)** $f(x) = x^2 + \frac{16}{x}$

(c) $f(x) = x^2 + \frac{81}{x^2}$ **(d)** $f(x) = 2x^6 - 6x^2$

(e) $f(x) = \cos(x^2), 0 \leqslant x \leqslant 3$ **(f)** $f(x) = xe^x$

(g) $f(x) = x\sin(x), -1 \leqslant x \leqslant 6$

9. An object moves along a line such that its displacement s metres from a fixed point P is given by $s(t) = t(t - 3)(8t - 9)$

(a) Find the initial velocity of the object.

(b) Find the velocity of the object at $t = 3$ seconds.

(c) Find the values of t for which the object changes direction. What significance do these times have in connection to the displacement of the object?

10. The delivery cost per tonne of bananas, D (in thousands of dollars), when x tonnes of bananas are shipped is given by $D = 3x + \frac{100}{x}, x > 0$
Find the value of x for which the delivery cost per tonne of bananas is a minimum, and find the value of the minimum delivery cost. Use the first derivative test to justify that this value is in fact a minimum. Verify your results by checking the graph of the function.

11. The displacement s metres of a car, t seconds after leaving a fixed point A, is given by $s(t) = 10t - \frac{1}{2}t^2$

(a) Calculate the velocity when $t = 0$

(b) Calculate the value of t when the velocity is zero.

(c) Calculate the displacement of the car from A when the velocity is zero.

12 A ball is thrown vertically upwards from ground level such that its height h at t seconds is given by $h = 14t - 4.9t^2$, $t \geqslant 0$

(a) Write a function for the ball's velocity.

(b) Find the maximum height the ball reaches and the time it takes to reach the maximum.

(c) At the moment the ball reaches its maximum height, what is the velocity of the ball?

13. The cost in Chinese renminbi to produce x fidget spinners can be modelled by the function $C(x) = 0.004x^2 + 5x + 2000$
Revenue from x fidget spinners can be modelled by the function $R(x) = -0.00016x^3 + 0.1x^2 + 20x$

Given that profit $P(x)$ = revenue − cost, find the maximum profit and the number of fidget spinners that should be produced to maximise profit.

14. The total cost, in euros, of producing n bicycle wheels can be modelled by the function $C(n) = 20\,000 + 20n - n^2 + 0.005n^3$

(a) Find the number of bicycle wheels that should be produced to minimise total cost, and the minimum total cost.

(b) The average production cost per bicycle wheel can be expressed as $\frac{C(n)}{n}$. Find the minimum average production cost and the number of wheels that should be produced to achieve this minimum.

15. A chemical reaction is modelled by the function $R(t) = 100 - \frac{100}{1 + 60e^{-0.25t}}$ where $R(t)$ is the percentage of the reactant remaining at time t in minutes, after a catalyst is introduced. Find the time at which the reaction rate is fastest.

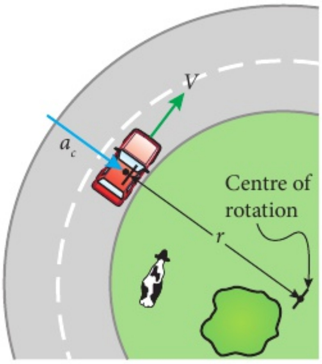


Figure 10.8 When a car is travelling through a turn, the momentum of the car is trying to carry the car in a direction tangent to the curve of the road. The car's tyres must generate a force normal to the curve of the road in order to steer the car around the curve

10.2 Tangents and normals

We are often interested in finding a **tangent** or **normal** to a curve.

Since we know how to find the derivative and how to use it to find the gradient at any point on a function, we can use the derivative to find the gradient of the tangent (or normal) at a point on a function. Then, we simply need to remember how to write the equation of a line, and with these skills we can write equations of tangents and normals.

Example 10.8

Given the function $g(x) = -x^2 + 8x - 14$

(a) Find the gradient of $g(x)$ at $x = 3$

(b) Hence write the equation of the tangent to $g(x)$ at $x = 3$

(c) Write the equation of the normal to $g(x)$ at $x = 3$

Solution

(a) To find the value of the gradient, we must first find the derivative of $g(x)$

$g'(x) = -2x + 8$

Then, we evaluate at $x = 3$: $\quad g'(3) = -2(3) + 8 = 2$

(b) Since we know the gradient of the tangent, we only need to find a point on the tangent. Since the tangent must intersect $g(x)$ at the point of tangency, we can find the y-coordinate by using $g(x)$:

$g(3) = -(3)^2 + 8(3) - 14 = -9 + 24 - 14 = 1$

Therefore $g(x)$ and the tangent both pass through (3, 1). Now we can write the equation of the tangent, using point–slope form:

$$y - 1 = 2(x - 3)$$
$$\Rightarrow y = 2x - 5$$

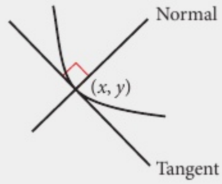
(b) Since a normal line is perpendicular to a tangent, recall that the product of the gradients must be -1

$$m_{\text{normal}} \cdot m_{\text{tangent}} = -1$$
$$\Rightarrow \quad m_{\text{normal}} \cdot 2 = -1$$
$$\Rightarrow \quad m_{\text{normal}} = -\frac{1}{2}$$

Since the normal must also pass through the point of tangency (3, 1), we can now write the equation of the normal, using point–slope form:

$$y - 1 = -\frac{1}{2}(x - 3)$$
$$\Rightarrow y = -\frac{1}{2}x + \frac{5}{2}$$

A **tangent** is a line that intersects a function and has the same gradient as the function at the point of intersection.
A **normal** is a line that intersects a function and is perpendicular to the tangent at the point of intersection.



Sometimes we are interested in developing functions that have specific properties of the tangent or normal. For example, since the tangent has the same gradient as the curve, it can provide a 'smooth' transition from one curve to another.

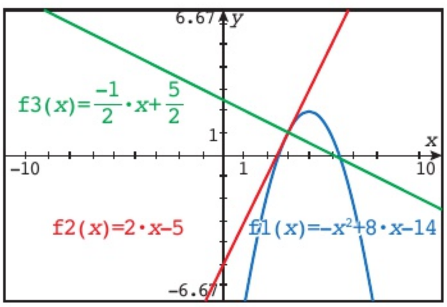


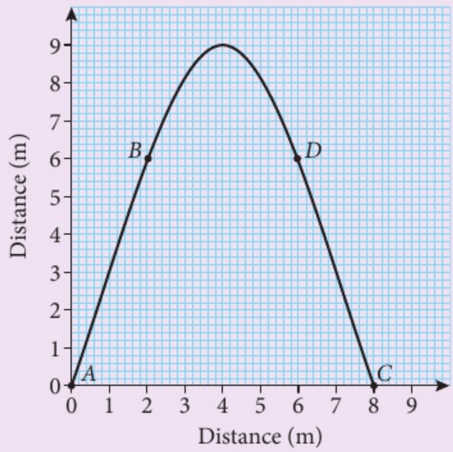
Figure 10.9 The tangent has the same gradient as the curve at the point of tangency

We can graph the function, the tangent, and the normal to verify our work (make sure your x and y scales are the same, or the lines may not appear perpendicular).

Example 10.9

Lena is designing an archway of a building. The archway should be approximately triangular, but with a rounded top part defined by a parabola.

She has sketched the design, as shown.



The line segments AB and CD are straight lines, while the curve above B (2, 6) and D (6, 6) will be a parabola. The overall height of the archway is unknown. The archway is symmetric about the line $x = 4$

(a) Write down the gradient of line segment AB.

(b) Write down the gradient of the curve at the maximum point of the archway.

(c) Given that the parabolic curve is modelled by the function $f(x) = ax^2 + bx + c$, find the values of a, b, and c, assuming that the gradient of $f(x)$ must be equal to the gradients of AB and CD where they intersect.

Solution

(a) Line segment AB has gradient $m_{AB} = \dfrac{6-0}{2-0} = 3$

(b) The gradient at the maximum point must be zero.

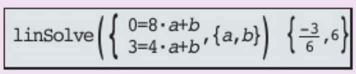
(c) Since the gradient of the curve must be zero at the maximum point, we know that $f'(4) = 0$. Likewise, to match the gradients of the line segment AB, we know that $f'(2) = 3$. We can write the derivative of $f(x)$ as $f'(x) = 2ax + b$ and hence write:

$$f'(4) = 0 \Rightarrow 0 = 2a(4) + b \Rightarrow 0 = 8a + b \qquad (1)$$

$$f'(2) = 3 \Rightarrow 3 = 2a(2) + b \Rightarrow 3 = 4a + b \qquad (2)$$

Now we have a simultaneous linear system to solve:

Using a GDC



Algebraically

$$\begin{cases} 0 = 8a + b \\ 3 = 4a + b \end{cases} \Rightarrow \begin{cases} 0 = 8a + b \\ -6 = -8a - 2b \end{cases}$$

$\Rightarrow -6 = -b \Rightarrow 6 = b$

$\Rightarrow 0 = 8a + 6 \Rightarrow -\frac{3}{4} = a$

So far we have $f(x) = -\frac{3}{4}x^2 + 6x + c$

Now find the value of c. Since we know that the function must pass through (2, 6), we can solve for c using that point:

$f(2) = 6 \Rightarrow 6 = -\frac{3}{4}(2)^2 + 6(2) + c \Rightarrow c = -3$

Therefore, the equation of the curve must be $f(x) = -\frac{3}{4}x^2 + 6x - 3$

Sometimes we need to find tangents and normals to a more complicated function. In this case, you should know how to use your GDC to find the numerical derivative.

Example 10.10

Given the function $h(x) = \frac{x^3 + 4x}{2x^2 - 3x}$

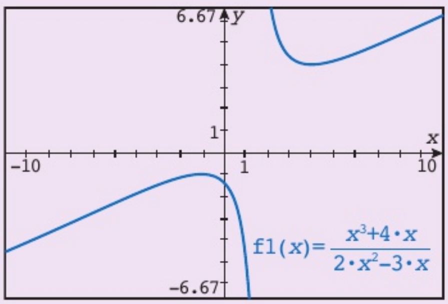
(a) Find the equations of the two vertical lines that are normal to h.

(b) Given that the line L_1: $y = -2x + k$ is tangent to h, find all possible value(s) of k.

Solution

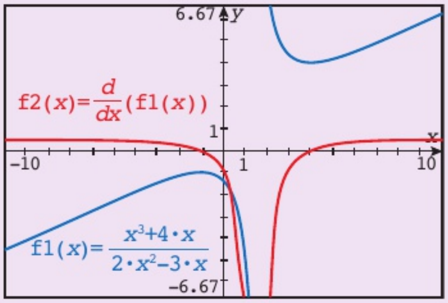
(a) For a vertical line to be normal to h at a point, it must be true that the tangent is horizontal at that point. Hence, the value of h' must be zero at that point.

To start, draw a graph of h

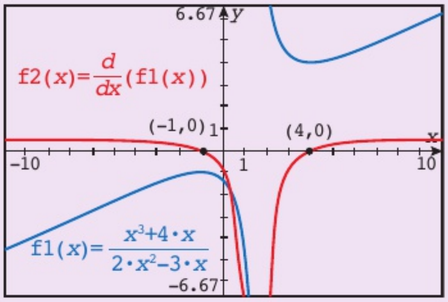


There are two possible horizontal tangents (and hence vertical normal lines): at approximately $x = -1$ and $x = 4$. Now, let us graph the derivative function h' so that we can find where it is equal to zero.

Remember that your GDC is a powerful tool for finding derivatives numerically. Unless you are required to find exact answers, you should use it where possible.



It appears that h' has zeros at $x = -1$ and $x = 4$; we confirm by using the GDC's **Zero** tool (twice) to find those zeros.



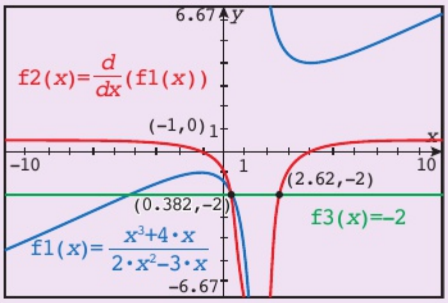
We are now confident that the vertical normal lines to h must pass through $x = -1$ and $x = 4$

Since these are vertical lines, the equations of the normal lines are also $x = -1$ and $x = 4$

(b) We need to find the points on h where the gradient is -2. If we can find those points on h, then we can solve for the value(s) of k.

Since we have already used the GDC to generate the graph of the derivative function h', we can also use it to find the places where the gradient is equal to -2. We can do this by adding the line $y = -2$ to the graph and finding where it intersects h'.

Use a GDC's **Intersection tool** (twice).



From this display, we see that $h'(x) = -2$ when $x \approx 0.382$ or $x \approx 2.62$

Thus, the gradient of h is -2 at those x values; the tangents $y = -2x + k$ must intersect h at those points. To find k, we must find the y-coordinates of those points and then solve for k in each tangent.

We have already entered the function into the GDC, so we can use the GDC to calculate the y values quickly. Our GDC gives that the y-coordinates for each of these points are approximately -1.85 and 4.85, respectively.

Now we know that we have two tangents, one of which must pass through $(0.382, -1.85)$ while the other passes through $(2.62, 4.85)$

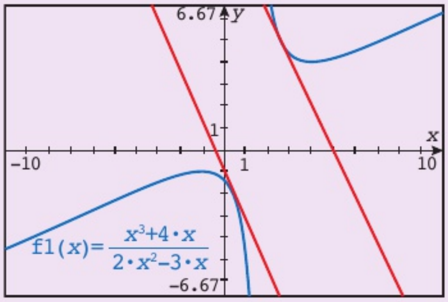
Finally, we solve for k, once for each point.

$$\begin{aligned} y &= -2x + k \\ -1.85 &= -2(0.382) + k \\ k &\approx -1.09 \end{aligned} \qquad \begin{aligned} y &= -2x + k \\ 4.85 &= -2(2.62) + k \\ k &\approx 10.1 \end{aligned}$$

Thus, the two tangents must be:

$y = -2x - 1.09$ and $y = -2x + 10.1$

We can verify visually on our GDC:



The final graph shows the two tangents we found, and we can see visually that indeed they are tangent to h

Learn how to use your GDC to store results from the graph, so that you avoid losing precision and do not have to retype long decimals. If you are retyping an approximate result into your GDC, you are not using it efficiently or effectively.

Exercise 10.2

1. For the functions below:

(i) Find the derivative of the function.

(ii) Calculate the gradient of the tangent of the function at the indicated point.

(iii) Find the equation of the tangent at that point.

(iv) Find the equation of the normal at that point.

Use a GDC to confirm your results.

(a) $y = 3x^2 - 4x$ point $(0, 0)$

(b) $y = 1 - 6x - x^2$ point $(-3, 10)$

(c) $y = \dfrac{2}{x^3}$ point $(-1, 2)$

(d) $y = x^5 - x^3 - x$ point $(1, -1)$

(e) $y = (x + 2)(x - 6)$ point $(2, -16)$

(f) $y = 2x + \dfrac{1}{x} - \dfrac{3}{x^3}$ point $(1, 0)$

(g) $y = \dfrac{x^3 + 1}{x^2}$ point $(-1, 0)$

(h) $y = x^2 + 2x + 1$ $x = -3$

(i) $y = x^3 + x^2$ $x = -\dfrac{2}{3}$

(j) $y = 3x^2 - x + 1$ $x = 0$

(k) $y = 2x + \dfrac{1}{x}$ $x = \dfrac{1}{2}$

2. For the functions below, find the coordinates of any points on the graph of the function where the tangent has the given gradient.

(a) $y = x^2 + 3x$ gradient $= 3$

(b) $y = x^3$ gradient $= 12$

(c) $y = x^2 - 5x + 1$ gradient $= 0$

(d) $y = x^2 - 3x$ gradient $= -1$

3. Find the equations of the lines tangent to the curve $y = x^3 - 3x^2 + 2x$ at any point where the curve intersects the x-axis.

4. Find the equation of the tangent to the curve $y = x^2 - 2x$ which is perpendicular to the line $x - 2y = 1$

5. The line $y = 4x - 13$ is tangent to $y = x^2 + ax + b$ at the point $(3, -1)$, where $a, b \in \mathbb{R}$. Find the value of a and the value of b

6. The line $y = -\frac{1}{3}x + 3$ is normal to $y = x^2 + ax + b$ at the point $(3, 2)$, where $a, b \in \mathbb{R}$. Find the value of a and the value of b

7. Find the coordinates of the point on the graph of $y = x^2 - x$ at which the tangent is parallel to the line $y = 5x$

8. Find the equation of the tangent(s) to the curve $y = x^3 - 5x$ which are perpendicular to the line $x + 7y = 38$.

9. Find the equation of the normal to the curve $y = x^2 + 4x - 2$ at the point where $x = -3$

Find the coordinates of the other point where this normal intersects the curve again.

10. Consider the function $g(x) = \dfrac{1 - x^3}{x^4}$

(a) Use your GDC to find the value of the derivative when $x = 1$

(b) Find the equation of both the tangent and the normal to the graph of g at the point $(1, 0)$.

11. Given the function $f(x) = x^3 + \frac{1}{2}x^2 + 1$

(a) Find the equation of the tangent to f at the point $\left(-1, \frac{1}{2}\right)$

(b) Find the coordinates of another point on the graph of f where the tangent is parallel to the tangent found in **(a)**.

12. Find the equation of both the tangent and the normal to the curve $y = \sqrt{x}(1 - \sqrt{x})$ at the point where $x = 4$

13. Using your GDC for assistance, make accurate sketches of the curves $y = x^2 - 6x + 20$ and $y = x^3 - 3x^2 - x$ on the same set of axes. The two curves have the same slope at an integer value for x somewhere in the interval $0 \leqslant x \leqslant 7$

(a) Find this value of x

(b) Find the equation for the tangent to each curve at this value of x.

14. The line L is tangent to $f(x) = -x^2 - 6x - 4$ and passes through the origin.

(a) Find the two possible points of tangency.

(b) Find the corresponding two equations for L

15. Find equations of both lines through the point $(2, -3)$ that are tangent to the parabola $y = x^2 + x$

10.3 Optimisation

Optimisation is the process of finding values that will maximise or minimise another quantity. For example, the following questions are typical optimisation problems:

- What production level or price should we choose in order to maximise profit?
- What production level should we choose to minimise costs?
- At what time is the maximum height or maximum velocity of a projectile reached?
- How can we fold a piece of paper to create an open box with maximum volume?
- How can we design a box to minimise packaging cost for a given volume?

Part of the process in solving optimisation problems is no different than finding maximum and minimum values of functions, which you have already seen. However, in this section we will focus on the problem-solving skills necessary to get to that point. We start by revisiting an example from Chapter 6 in order to highlight some key steps.

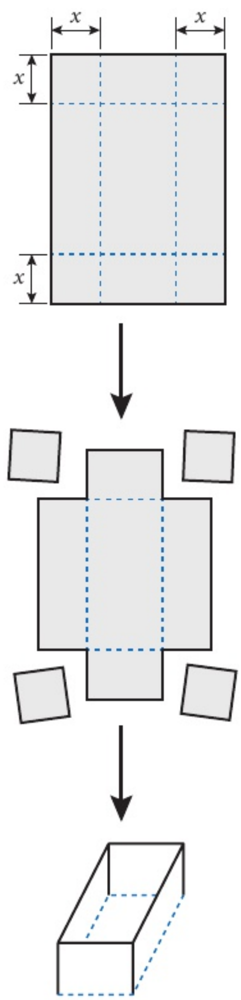


Figure 10.10 Creating an open box from a piece of paper (Example 10.11)

Example 10.11

The dimensions of a piece of A4 paper, to the nearest cm, are 21×30 cm. It is possible to create an open box by cutting out square corners and folding the remaining flaps up, as shown in the diagrams in Figure 10.10.

(a) Calculate the maximum volume of the open box.

(b) Find the dimensions of the open box with the largest volume that can be created using this method.

Solution

(a) **Choose an appropriate model for the quantity to be optimised.**
As discussed in Chapter 6, models come in many forms. In this section, our models often start with a simple, known formula. In this case, since we are interested in maximising the volume of the box, it seems appropriate to start with a model for volume: $V = lwh$

Draw an accurate diagram of the situation.
A diagram is helpful because it often uncovers relationships that will help us adapt our general model to the specific situation. In this case, we consider that the paper starts as 21×30 cm. However, those are not the dimensions of the open box. It seems like we should think about the dimensions of the open box because they are critical to finding the volume!

If we look at the second diagram in Figure 10.10, we can see that part of the width and length of the paper becomes the height for the box. Also, an important part of this step is to decide what the independent variable really is. In this case, it's the size of the square we cut out from each corner. So, let's label that x and then label the diagrams carefully as shown in Figure 10.11.

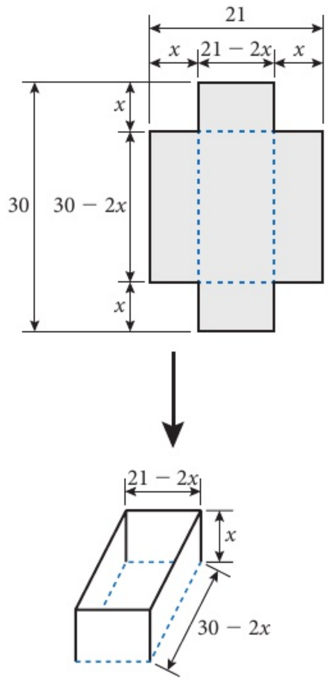


Figure 10.11 Careful labelling of a diagram is important (Example 10.11)

Use the diagram to write algebraic relationships.
After we label our diagram, we see that the width and length of the paper get reduced by twice the length of the corners we cut out. Therefore, we know that

length $= 30 - 2x$, width $= 21 - 2x$, height $= x$

Use algebraic relationships to express the general model in a single variable.
We can use the algebraic relationships we found above to develop a model for the volume of the box in terms of x by substituting into the general model for volume:

$$V = lwh = (30 - 2x)(21 - 2x)(x)$$

$$V = 4x^3 - 102x^2 + 630x$$

Use calculus to find extrema of the model.
Finally, we can now look for maxima by taking the derivative of the volume model:

$$V' = 12x^2 - 204x + 630$$

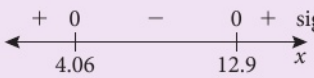
To find local maxima, we look for the place where the first derivative is zero and changes from positive to negative:

$$12x^2 - 204x + 630 = 0$$

You can then use a GDC polynomial solver.

```
polyRoots(12·x²-204·x+630,x)
                    {4.0559,12.9441}
```

Hence $x = 4.06$ or $x = 12.9$ (3 s.f.). Then, construct a sign diagram and test intervals.



Therefore, it appears that a local maximum for volume occurs when $x = 4.06$ (3 s.f.)

Note that values of $x > 10.5$ are nonsensical, since we can't cut out a square that is more than half the width of the paper.

Interpret your findings and answer the question in context.

The question asked us to find the maximum volume. So far, we have found the size of the square we should cut out is $x = 4.06$ cm.

To find the volume, we substitute this value into our model:

$$V = 4x^3 - 102x^2 + 630x = 4(4.06)^3 - 102(4.06)^2 + 630(4.06)$$
$$= 1144\text{ cm}^3$$

Therefore the maximum volume possible by creating a box with this method is 1144 cm^3.

(b) Use the algebraic relationships we found above:

length $= 30 - 2x = 30 - 2(4.06) = 21.9$ cm

width $= 21 - 2x = 21 - 2(4.06) = 12.9$ cm

height $= x = 4.06$ cm

Here are the key steps for solving optimisation problems:

- **Draw an accurate diagram of the situation.**
  A diagram can help you see geometric and algebraic relationships you may have missed otherwise. These relationships are crucial for the next steps.
- **Choose an appropriate model for the quantity to be optimised.**
  This often starts with a general, well-known formula or model, or even a simple equation. More than one model or general formula may be needed.
- **Decide what the independent variable is.**
  There is usually one independent variable that all the other variables are related to. It is our job to decide what that independent variable is, and then relate other variables to it in the next step.
- **Write algebraic relationships.**
  We write algebraic relationships between unknown quantities so that you can express your general model in terms of the single independent variable you have chosen. Often, each relationship will express the unknown quantity in terms of the single independent variable, but sometimes additional substitution may be necessary.

- **Express the general model in a single variable.**
  This step is a good one to set as your goal, because it can help you decide on an independent variable. Also, unless you can express the general model in a single variable, you will not be able to use the calculus you have learned to optimise the quantity. Don't forget to consider the domain of the model, since some extrema may be nonsensical in the context of the problem.
- **Use calculus or graphical analysis on your GDC to find extrema of the model.**
  This step is usually the most straightforward piece of solving the problem (if you are confident in your calculus or GDC skills). Once you have a model for the quantity to be optimised in terms of a single independent variable, then use the first derivative test to find local maxima and minima. Be careful to consider endpoints of the domain as well, as you saw in Section 10.2.
- **Interpret your findings and answer the question in context.**
  Finally, don't forget to answer the question, in context! Optimisation problems are sometimes long and you can forget the question you were trying to answer – check and make sure you answered the original question when you think you are finished. (Don't forget to check your units as well!)

Note that these steps are not strictly in order – sometimes a model may be obvious from the beginning, sometimes it may be helpful to decide on the independent variable first, etc. These steps are the same regardless of context.

Example 10.12

Two vertical posts, with heights of 7 m and 13 m, are secured by a rope going from the top of one post, to a point on the ground between the posts, and then to the top of the other post. The distance between the two posts is 25 m.

(a) How far from the base of the shorter post should the rope be anchored to the ground in order to minimise the length of the rope?

(b) What is the minimum length of rope necessary?

Solution

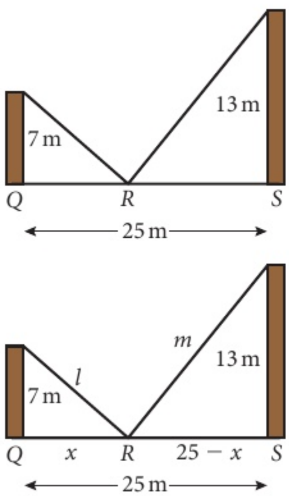


Figure 10.12 Diagrams for Example 10.12

(a) It seems like a good idea to start with an accurate diagram (Figure 10.12).

We have labelled the key points. Now, we need to think of some of the quantities in this diagram. Since the question asks for the length from the base of the shorter post (Q) to the anchor point (R), we will treat this as the independent variable and label it x. Also, we are trying to minimise the total length of the rope, so we will assign variables to each part of it as well, labelling them l and m. Finally, since the posts are 25 m apart, we can label the distance from R to the taller post (S) as $25 - x$

Now, let's use our diagram to try to find a general model. Since the quantity being optimised (the length of the rope) is the diagonal of a right triangle, the Pythagorean theorem $c^2 = a^2 + b^2$ will probably be useful. Using the Pythagorean theorem and the algebraic relationship $RS = 25 - x$, we can write

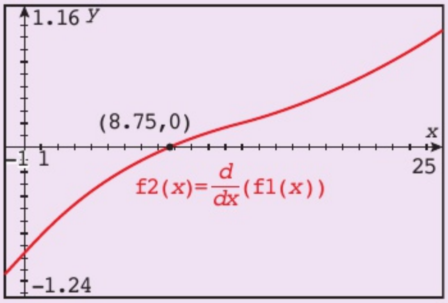
$$l^2 = x^2 + 7^2 \quad \text{and} \quad m^2 = (25 - x)^2 + 13^2$$

$$l = \sqrt{x^2 + 49} \qquad m = \sqrt{(25 - x)^2 + 169}$$

Let's call the total length T. Since we know that $T = l + m$,

$T = \sqrt{x^2 + 49} + \sqrt{(25 - x)^2 + 169}$

We now have a model for the quantity being optimised (T) expressed in a single variable (x). We are now ready to use calculus to find extrema, in this case the minimum value of T. Since this is a complicated expression, use a GDC to graph the derivative. To make the graph useful, consider the domain of our model. The length x could be anywhere from 0 to 25, so we can scale the graph accordingly (set the x-axis from 0 to 25, then choose a reasonable y-axis to fit the derivative). Finally, use the GDC's Zero tool to find the x value where $T' = 0$



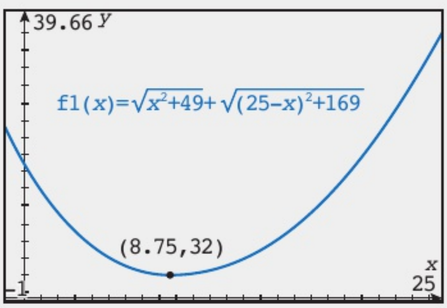
The GDC tells us that $T' = 0$ when $x = 8.75$. Furthermore, we can see from the graph of the derivative that T' is negative to the left of $x = 8.75$ and positive to the right of 8.75, so we conclude that $x = 8.75$ is a local minimum for T. Therefore, to answer the question in context, the rope should be anchored 8.75 m from the base of the shorter post in order to minimise the length of the rope.

(b) Use our model to find the length of the rope.

$$T = \sqrt{x^2 + 49} + \sqrt{(25 - x)^2 + 169}$$
$$= \sqrt{(8.75)^2 + 49} + \sqrt{(25 - 8.75)^2 + 169}$$

We can use our GDC to calculate this without losing precision (remember to store the value from the graph page). Therefore, when $x = 8.75$ m, the length of the rope is $T = 32.0$ m.

Example 10.12 can be solved without using calculus. Instead of using a GDC to graph the derivative function, we could analyse the length function, directly, to find the minimum. This method requires us to be more careful about considering the domain $0 \leqslant x \leqslant 25$ and setting the viewing window appropriately to obtain an appropriate display. Then, we can use the GDC's Minimum tool to find the minimum length and distance x at the same time.



Therefore the rope should be anchored 8.75 m from the base of the shorter post, and the minimum length is 32 m.

Example 10.13

To manufacture video game consoles requires the following costs per day, in US$, in a factory capable of producing up to 1000 units per day: $1580 fixed costs, $120 per console for production costs. Equipment maintenance costs vary directly with the square of the number of consoles produced.

(a) Given that equipment maintenance costs are $11 to produce 10 consoles, find an expression for equipment maintenance costs in terms of the number of consoles produced, x

(b) Find an expression for $C(x)$, the total costs for manufacturing x game consoles per day.

(c) Show that total manufacturing costs per day increases as production increases for all $x > 0$

(d) The average cost per game console is given by the function $\overline{C}(x)$. Find an expression for $\overline{C}(x)$

(e) Find the minimum average manufacturing cost and the number of video game consoles which should be manufactured to minimise the average manufacturing cost.

Solution

(a) Since equipment maintenance costs vary directly with the square of x, we have $y = kx^2$, so $11 = k(10)^2 \Rightarrow k = 0.11$

(b) Total cost per day is a sum of total costs, so our model is simply $C(x) = 0.11x^2 + 120x + 1580$

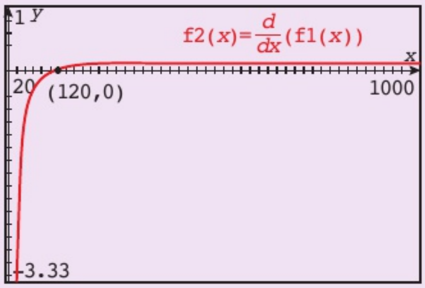
(c) Here we use calculus to find extrema – or, rather, to show that there are none within our domain. The first derivative is $C'(x) = 0.22x + 120$

This function is positive for all $x > 0$ so production costs are increasing for all $x > 0$

(d) Here we need a new model. Average cost per video game console is the total cost divided by the number of units, hence

$$\overline{C}(x) = \frac{0.11x\,2 + 120x + 1580}{x}$$

(e) Use a GDC to graph $\overline{C}'(x)$



We see that costs are decreasing for $0 < x < 120$ and increasing for $120 < x < 1000$ so there must be a local minimum at $x = 120$

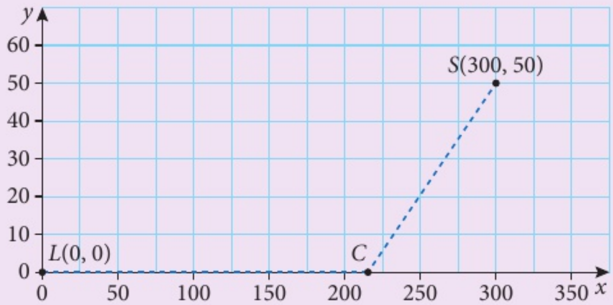
Therefore, average manufacturing cost is minimised when 120 units per day are produced.

The average cost per console at this level is $\overline{C}(120) = \$146$

Example 10.14

When a lifeguard on a beach sees a swimmer in distress, she must make a decision about the fastest way to get to the swimmer: should she run along the beach until she is as close to the swimmer as possible, then swim, or jump right in the water and swim? Of course, it depends on how fast she can swim compared to how fast she can run. (Ignore any currents in the water.)

Suppose that the lifeguard is positioned at point $L(0, 0)$ and a swimmer in distress is at point $S(300, 50)$ (distances in metres). The edge of the water is along the x-axis. The lifeguard will leave point L, run along the water's edge (the x-axis) to point C, and then swim to the swimmer at point S, as shown in the diagram.



(a) Given that the lifeguard can run at a rate of 3.5 minutes per km and swim at a rate of 1 minute per 100 m, where should point C be located in order to reach the swimmer as quickly as possible?

(b) What is the minimum time it will take the lifeguard to reach the swimmer?

Solution

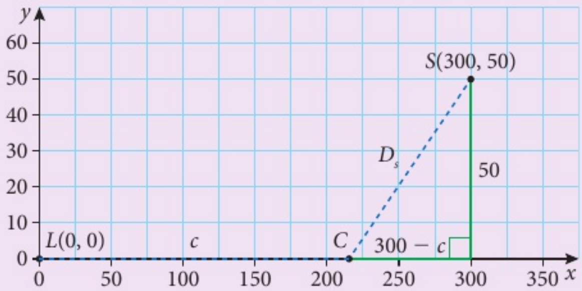
Since we are interested in minimising the total time, we should start with a very general model of $T = T_R + T_S$, where T is the total time, T_R is the running time, and T_S is the swimming time. We will use seconds as the units for all times. Also, since we are interested in finding point C, we should assign a variable to the x-coordinate (the y-coordinate will be zero). Let C have coordinates $(c, 0)$; this is our independent variable.

Next, we have to think about how to find algebraic relationships for T_R and T_S in terms of c. Since both of these have to do with rate and time, we can probably use the model rate $= \frac{\text{distance}}{\text{time}}$, rearranged to time $= \frac{\text{distance}}{\text{rate}}$; we just need to be careful that the rate has units of distance (metres) per time (seconds). The running part seems straightforward: wherever point C is, the running distance will be exactly c metres. So, we can write:

$$T_R = \frac{c}{1000\text{ m}/210\text{ s}} = \frac{c}{100/21} = \frac{21c}{100}$$

Notice that we converted the rate of 3.5 minutes per 1 km into 1000 m per 210 seconds.

Next, we need to find an expression for T_S. For this part, we see that the lifeguard will swim the diagonal of a right triangle. We can add some more information to our diagram, and we can use the Pythagorean theorem to model the distance the lifeguard will swim, D_S, as shown in the graph.



Therefore an expression for D_S is

$$D_S^2 = (300 - c)^2 + 50^2 \Rightarrow D_S = \sqrt{(300 - c)^2 + 50^2}$$

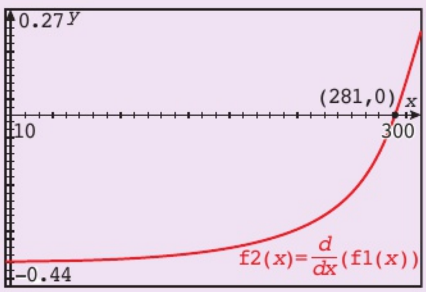
Using the general distance–time model, we can now write:

$$T_S = \frac{\sqrt{(300 - c)^2 + 50^2}}{\dfrac{100\text{ m}}{60\text{ s}}} = \frac{3}{5}\sqrt{(300 - c)^2 + 50^2}$$

Substituting the expressions for T_R and T_S into our total-time model, we obtain an equation in a single variable:

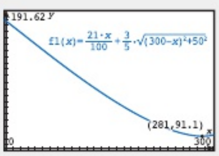
$$T = T_R + T_S$$

$$T = \frac{21c}{100} + \frac{3}{5}\sqrt{(300 - c)^2 + 50^2}$$



Now we are ready to use calculus to find the value of c which will minimise the total time it takes the lifeguard to reach the swimmer. For this expression, we will rely on our GDC to find the numerical derivative. Since c can be anything in the interval [0, 300], we will scale our x-axis accordingly and then use the GDC Zero tool to find the value of c for which $T' = 0$

Again, we could use a GDC to analyse the function directly.



The GDC tells us that the lifeguard should run for 281 m and the total time taken to reach the swimmer will be 91.1 seconds.

Now we need to interpret the results. For part (a), the lifeguard can minimise the time it takes to reach the swimmer by running for 281 metres and swimming the remaining distance. To find the minimum time this will take, use a GDC to evaluate the function for total time at $x = 281$ and obtain 91.1 seconds. Therefore, to answer question (b), it will take about 91 seconds for the lifeguard to reach the swimmer.

In other problems, we need to be careful to consider the endpoints of the domain as possible optimum solutions.

Example 10.15

A supply of four metres of wire is to be used to form a square and a circle. How much of the wire should be used to make the square and how much should be used to make the circle in order to enclose the greatest amount of area?

Solution

Start with an accurate diagram (see Figure 10.13). Let r equal the radius of the circle and x equal the side length of the square; we can decide which one should be our independent variable later.

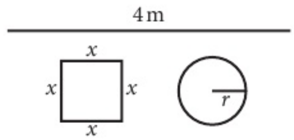


Figure 10.13 Diagram for Example 10.15

An appropriate general model for the total area is the sum of the area of a square and a circle:

$A = x^2 + \pi r^2$

Since we have two variables in this model, we need to find an algebraic relationship to express the model in a single variable. Recall that the square and the circle are made from a single piece of wire 4 metres long. So, we can write the algebraic relationship

$$4 = (\text{perimeter of square}) + (\text{circumference of circle})$$

$$\Rightarrow 4 = 4x + 2\pi r$$

This relationship can be solved for r, so that x is the independent variable:

$$4 = 4x + 2\pi r \Rightarrow 2\pi r = 4 - 4x \Rightarrow r = \frac{2(1 - x)}{\pi}$$

Substitute into our general model $A = x^2 + \pi r^2$, so that we now have a single variable:

$$A = x^2 + \pi r^2$$

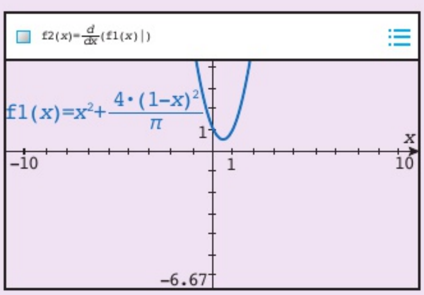
$$\Rightarrow A = x^2 + \pi\left[\frac{2(1 - x)}{\pi}\right]^2 = x^2 + \frac{4(1 - x)^2}{\pi}$$

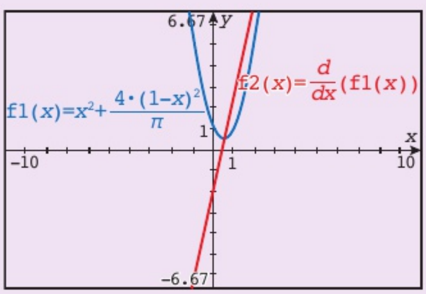
What about the domain? Clearly, $x \geqslant 0$ since it is a distance. Also, since the square's perimeter is $4x$, and the total length of wire is 4 m, we have

$4x \leqslant 4 \Rightarrow x \leqslant 1$

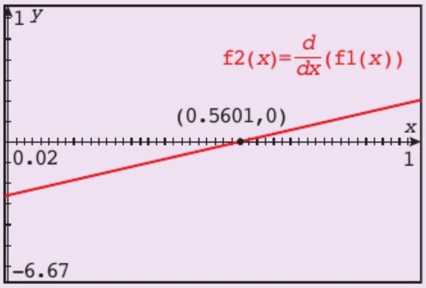
So, the domain for our area model is $0 \leqslant x \leqslant 1$

Although this function could be written as a quadratic and solved by hand, we will turn to our GDC again. Enter the function and then use a GDC to graph the numerical derivative:





Then, we hide the area function, set our viewing window to the domain $0 \leqslant x \leqslant 1$ and find the zero of the derivative.



Now we need to interpret our results and answer the question. We see that there is a local minimum for the area at $x = 0.560$ since the derivative is negative to the left and positive to the right. But we want to maximise area!

Recall our rules for endpoints from Section 10.1:

- The left endpoint of the domain, at $x = 0$, has a negative derivative to the right, so it is a local maximum.
- The right endpoint of the dom ain, at $x = 1$, has a positive derivative to the left, so it is also a local maximum.

So, to find out which of the local maxima is the absolute maximum, we need to check the value of the area at each:

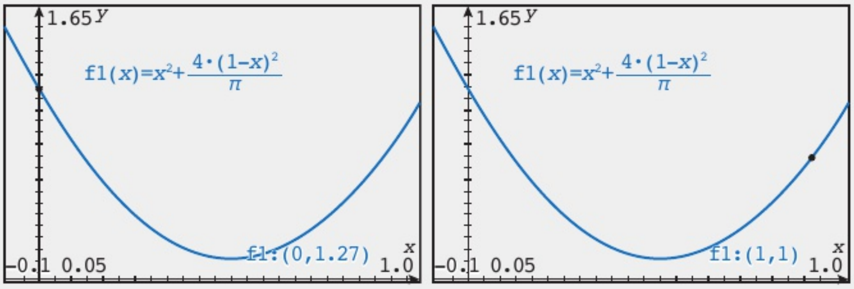
- For $x = 0$, we have $A = (0)^2 + \dfrac{4(1-0)^2}{\pi} = \dfrac{4}{\pi} = 1.27\text{ cm}^2$
- For $x = 1$, we have $A = (1)^2 + \dfrac{4(1-1)^2}{\pi} = 1\text{ cm}^2$

Therefore, we conclude that the maximum area is enclosed when the side length of the square is 0, and the entire wire is used to make the circle.

What if we wanted to minimise the area enclosed? We've already done the work. The side length of the square would be $x = 0.560$ cm in order to minimise the area enclosed.

We could analyse the area function directly if we wanted to avoid using calculus. After obtaining a graph of the area function, we see the minimum and two local maxima at the endpoints of the domain.

We can use the Trace tool and type in '0' and '1' to find approximations for the area at each endpoint:



Again, we see that the area is maximised ($A = 1.27\ \text{cm}^2$) when the side length of the square is zero.

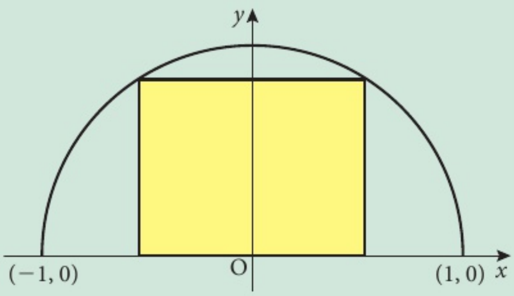
Exercise 10.3

1. Repeat Example 10.11 with a piece of US letter paper, which is 8.5 inches wide by 11 inches tall.

2. Repeat Example 10.12 with posts that are 8 m and 14 m tall with a distance between them of 30 m.

3. Repeat Example 10.14 where the lifeguard can run at 3 minutes and 40 seconds per 1 km and swim at 70 seconds per 100 m.

4. Find the dimensions that give the maximum area of the rectangle that can be inscribed in a semicircle with radius 1 cm. Two vertices of the rectangle are on the semicircle and the other two vertices are on the x-axis, as shown in the diagram.



5. A rectangular piece of aluminium is to be rolled to make a cylinder with open ends (a tube). Regardless of the dimensions of the rectangle, the perimeter of the rectangle must be 40 cm. Find the dimensions (length and width) that give a maximum volume for the cylinder.

6. Find the minimum distance between the graph of the function $y = \sqrt{x}$ and the point $\left(\frac{3}{2}, 0\right)$

7. Figure 10.14 consists of a rectangle $ABCD$ and a semicircle on each end. The rectangle has an area of $100\,\text{cm}^2$. If x represents the length of the rectangle AB, then find the value of x that makes the perimeter of the entire figure a minimum.

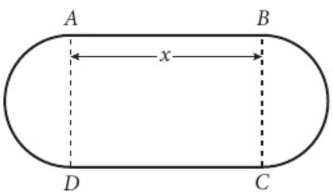
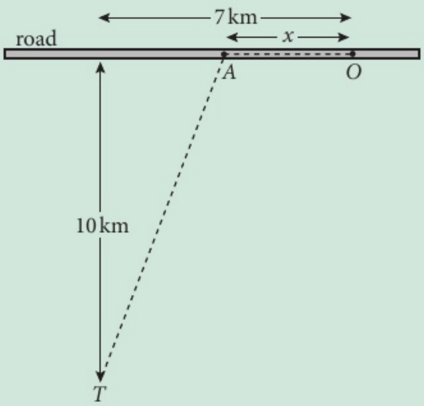
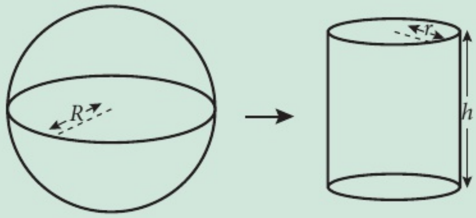


Figure 10.13 Diagram for question 7

8. Charlie is walking from the wildlife observation tower (point T) to the Desert Park office (point O). The tower is 7 km due west and 10 km due south from the office. The road to the office is 10 km due north from the tower. Charlie can walk at a rate of 2 kilometres per hour (kph) through the sandy terrain of the park, but she can walk a faster rate of 5 kph on the road. The point A is the point Charlie should walk to, in order to minimize the total walking time from the tower to the office. Find the value of x such that point A is x km from the office.



9. A cylinder is created by cutting away a sphere. Find the height, h, and the base radius, r, that will maximise the volume of a right circular cylinder, given that the sphere has a radius of $R = 10$ cm.

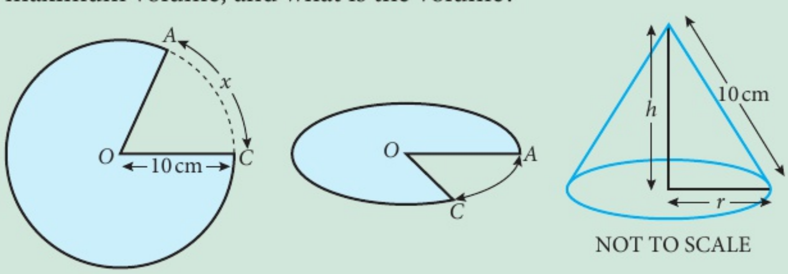


10. A manufacturer produces closed cylindrical cans of radius r cm and height h cm. Each can has a total surface area of 54π cm^2.
   (a) Solve for h in terms of r, and hence find an expression for the volume, V cm^3, of each can in terms of r.
   (b) Find the value of r for which the cans have their maximum possible volume.

11. The curve $y = ax^2 + bx + c$ has a maximum point at $(2, 18)$ and passes through the point $(0, 10)$. Find a, b and c.

12. To manufacture GPS devices requires the following costs per day, in US dollars, in a factory capable of producing up to 4000 GPS devices per day: \$2150 fixed costs, \$85 per GPS device for production costs. Equipment maintenance costs vary directly with the square of the number of GPS devices produced.
   (a) Given that equipment maintenance costs are \$4 to produce 5 GPS devices, find an expression for equipment maintenance costs in terms of the number of GPS devices produced, x.
   (b) Find an expression for $C(x)$, the total costs for manufacturing x GPS devices per day.
   (c) Show that total manufacturing cost per day increases as production increases for all $x > 0$.
   (d) The average cost per GPS device is given by the function $\overline{C}(x)$. Find an expression for $\overline{C}(x)$.
   (e) Find the minimum average manufacturing cost and the number of GPS devices which should be manufactured to minimise the average manufacturing cost.

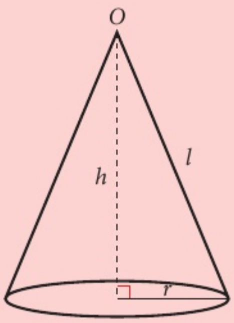
13. A cone of height h and radius r is constructed from a circle with radius 10 cm by removing a sector AOC of arc length x cm and then connecting the edges OA and OC. What arc length x will produce the cone of maximum volume, and what is the volume?



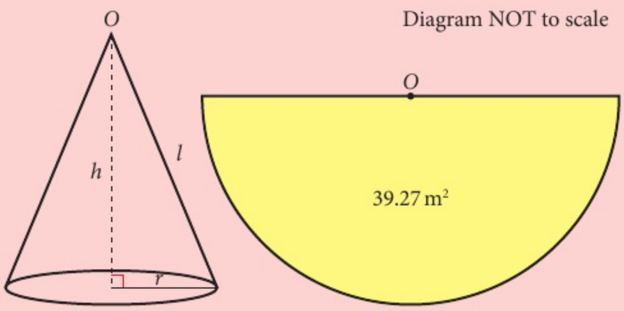
14. A ship sailing due south at 16 knots is 10 nautical miles north of a second ship going due west at 12 knots. Find the minimum distance between the two ships.

Chapter 10 practice questions

1. Tepees were traditionally used by nomadic tribes who lived on the Great Plains of North America. They are cone-shaped dwellings and can be modelled as cones, with vertex O. The cone has radius, r, height, h, and slant height, l.



A model tepee is displayed at a Great Plains exhibition. The curved surface area of this tepee is covered by a piece of canvas that is 39.27 m^2, and has the shape of a semicircle, as shown in the diagram.



(a) By first calculating the following values, show that the slant height, l, is 5 m, correct to the nearest metre.

(i) Find the circumference of the base of the cone.

(ii) Find the radius, r, of the base.

(iii) Find the height, h.

2. Let $f(x) = x^5$

(a) Write down $f'(x)$

Point $P(2, 32)$ lies on the graph of f.

(b) Find the gradient of the tangent to the graph of $y = f(x)$ at P

(c) Find the equation of the normal to the graph at P. Give your answer in the form $ax + by + d = 0$ where a, b and d are integers.

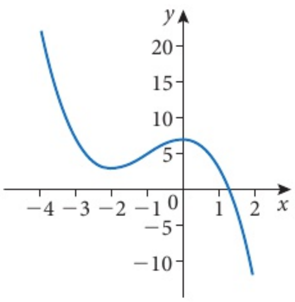


Figure 10.15 Graph of f for question 4

3. Consider the curve $y = x^3 + kx^2, k \in \mathbb{R}$

(a) Write down $\dfrac{\mathrm{d}y}{\mathrm{d}x}$

The curve has a local minimum at the point where $x = 3$

(b) Find the value of k

(c) Find the value of y at this local minimum.

4. Consider the graph of the function $f(x) = -x^3 - 3x^2 + 7$

(a) Estimate, to 1 significant figure, the coordinates where the maximum value of the function occurs.

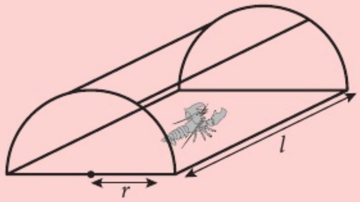
5. A function is given as $f(x) = 3x^3 - 7x + \dfrac{4}{x} + 10, -5 \leqslant x \leqslant 5, x \neq 0$

(a) Write down the derivative of the function.

(b) Use your GDC to find the coordinates of the local minimum point of f in the given domain.

(c) Use your GDC to find the coordinates of the local maximum point of f in the given domain.

6. A lobster trap is made in the shape of half a cylinder. It is constructed from a steel frame with netting pulled tightly around it. The steel frame consists of a rectangular base, two semi-circular ends and two further support rods, as shown in the diagram.



The semi-circular ends each have radius r and the support rods each have length l

Let T be the total length of steel used in the frame of the lobster trap.

(a) Write down an expression for T in terms of r, l and π

The volume of the lobster trap is $0.75\ \text{m}^3$.

(b) Write down an equation for the volume of the lobster trap in terms of r, l and π

(c) Show that $T = (2\pi + 4)r + \dfrac{6}{\pi r^2}$

(d) Find $\dfrac{\mathrm{d}T}{\mathrm{d}r}$

The lobster trap is designed so that the length of steel used in its frame is a minimum.

(e) Show that the value of r for which T is a minimum is 0.719 m, correct to 3 significant figures.

(f) Calculate the value of l for which T is a minimum.

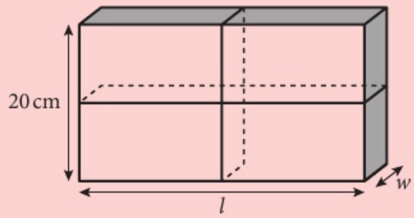
(g) Calculate the minimum value of T

7. A parcel is in the shape of a rectangular prism. It has a length l cm, width w cm and height of 20 cm. The total volume of the parcel is 3000 cm^3.

(a) Express the volume of the parcel in terms of l and w.

(b) Show that $l = \frac{150}{w}$

The parcel is tied up using a length of string that fits exactly around the parcel, as shown in the diagram.



(c) Show that the length of string, S cm, required to tie up the parcel can be written as

$S = 40 + 4w + \frac{300}{w}, 0 < w \leqslant 20$

(d) Draw the graph of S for $0 < w \leqslant 20$ and $0 < S \leqslant 500$, clearly showing the local minimum point. Use a scale of 2 cm to represent 5 units on the horizontal axis w (cm), and a scale of 2 cm to represent 100 units on the vertical axis S (cm).

(e) Find $\frac{\mathrm{d}S}{\mathrm{d}w}$

(f) Find the value of w for which S is a minimum.

(g) Write down the value of the length, l, of the parcel for which the length of string is a minimum.

(h) Find the minimum length of string required to tie up the parcel.

8. (a) Expand the expression $x(16x^3 - 27)$

(b) Differentiate $f(x) = x(16x^3 - 27)$

(c) Find the x-coordinate of the local minimum of the curve $y = f(x)$

9. Consider the function $f(x) = \frac{3}{4}x^4 - x^3 - 9x^2 + 20$

(a) Find $f(-2)$

(b) Find $f'(x)$

The graph of the function f has a local minimum at the point where $x = -2$

(c) Using your answer to part **(b)**, show that there is a second local minimum at $x = 3$

(d) Sketch the graph of the function $f(x)$ for

$-5 \leqslant x \leqslant 5$ and $-40 \leqslant y \leqslant 50$

Indicate on your sketch the coordinates of the y-intercept.

(e) Write down the coordinates of the local maximum.

Let T be the tangent to the graph of the function $f(x)$ at the point $(2, -12)$.

(f) Find the gradient of T.

The line L passes through the point $(2, -12)$ and is perpendicular to T. L has equation $x + by + c = 0$, where b and $c \in \mathbb{Z}$

(g) Find:

(i) the gradient of L **(ii)** the value of b and the value of c

10. Consider the function $f(x) = ax^3 - 3x + 5$, where $a \neq 0$.

(a) Find $f'(x)$

(b) Write down the value of $f'(0)$

The function has a local maximum at $x = -2$

(c) Calculate the value of a

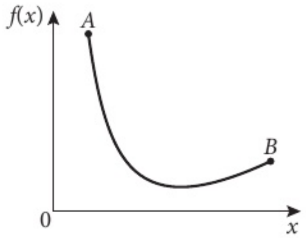


Figure 10.16 Graph of f for question 11

11. The graph of the function $f(x) = \dfrac{14}{x} + x - 6$, for $1 \leqslant x \leqslant 7$, is shown.

(a) Calculate $f(1)$

(b) Find $f'(x)$

(c) Use your answer to part **(b)** to show that the x-coordinate of the local minimum point of the graph of f is 3.7 correct to 2 significant figures.

(d) Find the range of f

Points A and B lie on the graph of f. The x-coordinates of A and B are 1 and 7 respectively.

(e) Write down the y-coordinate of B

(f) Find the gradient of the straight line passing through A and B

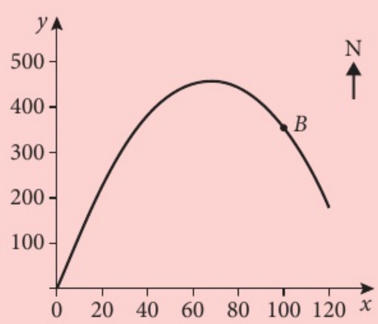
(g) M is the midpoint of the line segment AB. Write down the coordinates of M

L is the tangent to the graph of the function $y = f(x)$, at the point on the graph with the same x-coordinate as M

(h) Find the gradient of L

(i) Find the equation of L. Give your answer in the form $y = mx + c$

12. The diagram shows an aerial view of a bicycle track. The track can be modelled by the quadratic function $y = \dfrac{-x^2}{10} + \dfrac{27}{2}x$, where $x \geqslant 0$, $y \geqslant 0$, (x, y) are the coordinates of a point x metres east and y metres north of O, where O is the origin $(0, 0)$



B is a point on the bicycle track with coordinates $(100, 350)$.

(a) The coordinates of point A are $(75, 450)$. Determine whether point A is on the bicycle track. Give a reason for your answer.

(b) Find the derivative of $y = \dfrac{-x^2}{10} + \dfrac{27}{2}x$

(c) Use the answer in part **(b)** to determine if $A(75, 450)$ is the point furthest north on the track between O and B. Give a reason for your answer.

(d) (i) Write down the midpoint of the line segment OB.

(ii) Find the gradient of the line segment OB.

Scott starts from a point $C(0, 150)$. He hikes along a straight road towards the bicycle track, parallel to the line segment OB.

(e) Find the equation of Scott's road. Express your answer in the form $ax + by = c$, where a, b and $c \in \mathbb{R}$

(f) Use your GDC to find the coordinates of the point where Scott first crosses the bicycle track.

13. Given $f(x) = 5x^3 - 4x^2 + x$

(a) Find $f'(x)$

(b) Find, using your answer to part **(a)**, the x-coordinate of:

(i) the local maximum point

(ii) the local minimum point.

14. Consider the function $g(x) = bx - 3 + \frac{1}{x^2}, x \neq 0$

(a) Write down the equation of the vertical asymptote of the graph of $y = g(x)$

(b) Write down $g'(x)$

The line T is the tangent to the graph of $y = g(x)$ at the point where $x = 1$

The gradient of T is 3.

(c) Show that $b = 5$

(d) Find the equation of T

(e) Using your GDC find the coordinates of the point where the graph of $y = g(x)$ intersects the x-axis.

(f) (i) Sketch the graph of $y = g(x)$ for $-2 \leqslant x \leqslant 5$ and $-15 \leqslant y \leqslant 25$, indicating clearly your answer to part **(e)**.

(ii) Draw the line T on your sketch.

(g) Using your GDC find the coordinates of the local minimum point of $y = g(x)$

(h) Write down the interval for which $g(x)$ is increasing in the domain $0 < x < 5$

15. The equation of a curve is given as $y = 2x^2 - 5x + 4$

(a) Find $\frac{dy}{dx}$

The equation of the line L is $6x + 2y = -1$

(b) Find the x-coordinate of the point on the curve $y = 2x^2 - 5x + 4$ where the tangent is parallel to L

16. Consider the function $f(x) = -\frac{1}{3}x^3 + \frac{5}{3}x^2 - x - 3$

(a) Sketch the graph of $y = f(x)$ for $-3 \leqslant x \leqslant 6$ and $-10 \leqslant y \leqslant 10$ showing clearly the axis intercepts and local maximum and minimum points. Use a scale of 2 cm to represent 1 unit on the x-axis, and a scale of 1 cm to represent 1 unit on the y-axis.

(b) Find the value of $f(-1)$

(c) Write down the coordinates of the y-intercept of the graph of f.

(d) Find $f'(x)$

(e) Show that $f'(-1) = -\frac{16}{3}$

(f) Explain what $f'(-1)$ represents.

(g) Find the equation of the tangent to the graph of f at the point where $x = -1$

(h) Sketch the tangent to the graph of f at $x = -1$ on your diagram for **(a)**.

P and Q are points on the curve such that the tangents to the curve at these points are horizontal. The x-coordinate of P is a, and the x-coordinate of Q is b, $b > a$.

(i) Write down the value of:

(i) a **(ii)** b

(j) Describe the behaviour of $f(x)$ for $a < x < b$.

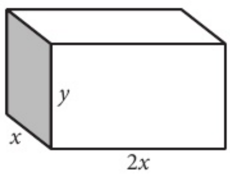


Figure 10.17 Diagram for question 17

17. A shipping container is to be made with six rectangular faces, as shown in Figure 10.17.

The dimensions of the container are length: $2x$, width: x, height: y

All the measurements are in metres. The total length of all twelve edges is 48 metres.

(a) Show that $y = 12 - 3x$

(b) Show that the volume V m^3 of the container is given by $V = 24x^2 - 6x^3$

(c) Find $\frac{dV}{dx}$

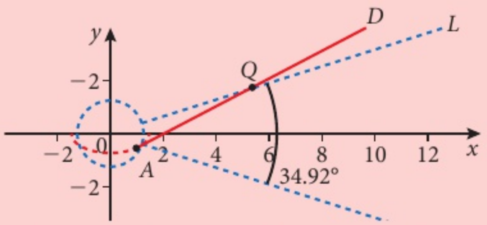
(d) Find the value of x for which V is a maximum.

(e) Find the maximum volume of the container.

(f) Find the length and height of the container for which the volume is a maximum.

(g) The shipping container is to be painted. One litre of paint covers an area of 15 m^2. Paint comes in tins containing four litres. Calculate the number of tins required to paint the shipping container.

18. In the sport of discus throwing, the athlete must throw the discus as far as possible from the centre of a 2.5 m diameter circle. The discus must land within a sector which has a central angle of 34.92°. Part of Kai's discus throw, before he releases the discus, can be modelled by the curve $y = -\sqrt{\frac{1}{2} - \frac{1}{4}x^2}$ where x and y are the coordinates in metres of the discus, with the origin placed at the centre of the throwing circle, as shown. On one throw, Kai releases the discus at the point A, and it travels according to the tangent at that point. The boundaries of the sector and throwing circle are shown in dashed lines.



(a) Given that point A has x-coordinate $x = 1$, find the equation of the tangent D in the form $y = ax + b$, where $a, b \in \mathbb{R}$

(b) The boundary line L passes through the centre of the throwing circle at the origin. Show that the equation of the line L, one boundary of the sector, is given by $y = 0.315x$

(c) Find the coordinates of point Q, the point at which the path of the discus leaves the sector.

(d) Kai's discus throw travels 20 m from the centre of the throwing circle before landing. For the throw to be valid, it must land within the sector. Is Kai's throw valid? Give a reason for your answer.

(e) The maximum distance Kai can throw (measured from the centre of the throwing circle) is 35 m. In order for Kai's throw to be valid, he must release the discus at point R on the curve $y = -\sqrt{\frac{1}{2} - \frac{1}{4}x\,2}$

(i) Find the coordinates of point R

(ii) Find the equation of the tangent at point R in the form $y = cx + d$, where $c, d \in \mathbb{R}$

19. A person in a rowing boat is 2 km from the nearest point on the shore. The house that she wishes to reach is 6 km from that point. She can row at a rate of 3 km h^{-1} and walk at a rate of 5 km h^{-1}.

(a) Show that the distance from the rowboat to the house is $2\sqrt{10}$ km.

To get to the house in the minimum amount of time, she should row for r km and walk for w km.

(b) Find a function r in terms of w.

(c) Hence, find a function for the total time to reach the house in terms of w.

(d) Hence or otherwise, find the values of w and r.

(e) Find the minimum amount of time required for her to reach the house.

20. A patient is given a hormone injection and the concentration of the hormone in their bloodstream is monitored. The rate of change of the bloodstream hormone concentration is given by

$\frac{dC}{dt} = (-0.75t^{1.2} + 40t^{0.2})0.97^t$ where $\frac{dC}{dt}$ is in nanograms $ml^{-1}\,min^{-1}$ and t is the number of minutes since the injection was given.

(a) Find the time at which the hormone concentration is increasing the fastest and the rate of increase at this point.

(b) Find the time at which the maximum hormone concentration is present.

(c) Find the time at which the hormone concentration is decreasing the fastest and the rate of decrease at this point.

(d) If the bloodstream concentration increases at more than 35 nanograms $ml^{-1}\,min^{-1}$, it can be dangerous for the patient. Find the interval where the patient is at risk.

21. In a small town, the number of new customers for a high-speed internet service provided by FastNet can be modelled by the function

$S'(t) = \frac{2000}{1200(0.65)^t + (1.25)^t + 250}$ where $S'(t)$ is the number of new subscribers in thousands, at time t, in years since 2000.

(a) Write down the units for $S'(t)$.

(b) In which year is the total number of subscribers increasing the fastest and at what rate?

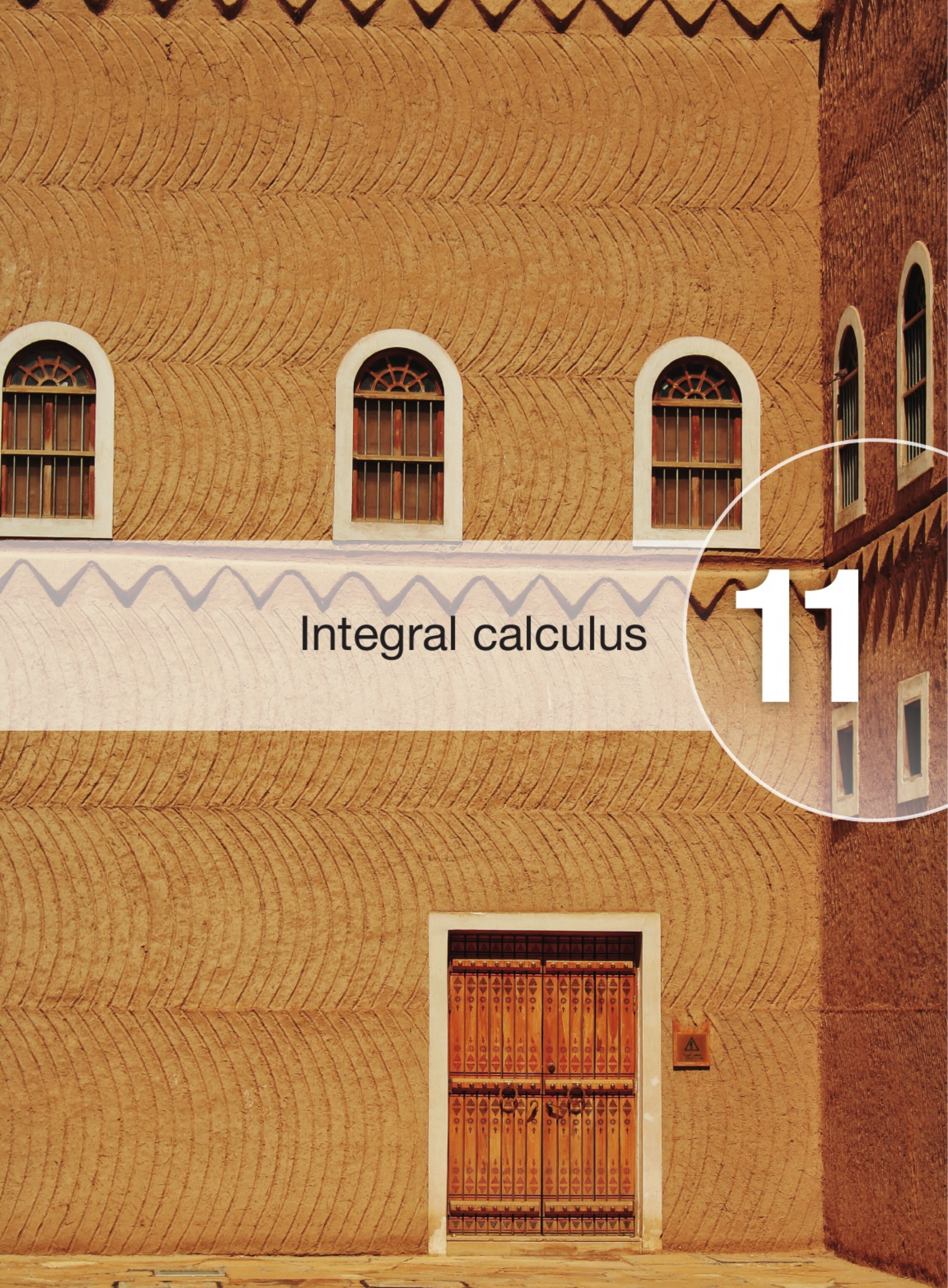
(c) FastNet has planned a peak growth rate of 15 new subscribers per day. During what time will this rate be exceeded? (Assume 365 days in a year.)

(d) Give a reason why the total number of subscribers will not decrease according to this model.

22. A barrel is to be made from strips of wood measuring 5 cm wide and 1 m long. A barrel maker has 40 strips with which to make a cylindrical barrel, including making the two ends of the barrel.

(a) Assuming that the wood can be entirely used (no wasted scraps), what are the dimensions of the barrel with maximum volume that can be made from these strips?

(b) What is the maximum volume, in litres?



11 Integral calculus

11 Integral calculus

Learning objectives

By the end of this chapter, you should be familiar with…

- integration as anti-differentiation of functions
- evaluating definite integrals
- performing anti-differentiation with a boundary value condition
- finding the area between a curve and the x axis with a definite integral
- working with kinematic problems involving position s, velocity v and acceleration a
- integrating polynomial and similar functions by inspection
- approximating definite integrals using the trapezoidal rule.

In Chapter 9, we used the derivative to find the rate of change of a function. In this chapter we learn how to reverse that process. If we know the derivative (rate of change) of an unknown function, can we find the original function? For example, can we find the distance that an object travels if we know its velocity?

11.1 Anti-derivatives

The idea of anti-differentiation may seem daunting, but you have probably done it without realising. For example, if we are given the velocity of an object over a specified interval of time, can we find the distance travelled? Imagine that an electric toy car moves in a straight line at a constant velocity of 8 centimetres per second (cm s^{-1}) for 12 seconds. What distance did the toy car travel during the 12 seconds?

In physics we refer to position as **displacement**.

distance = velocity × time

$\therefore$ distance travelled $= 8\,\text{cm s}^{-1} \times 12\,\text{s} = 96\,\text{cm}$

How is that anti-differentiation? Since the rate of change of position with respect to time is velocity, we have anti-differentiated to find the distance travelled.

Since the velocity function of the toy car is the constant function $v(t) = 8$, we can graph it as a horizontal line, as shown in Figure 11.1. Notice that the distance travelled over 12 seconds is exactly equal to the area between the curve and the x-axis. This makes sense because the units for the vertical axis are cm s^{-1} and the units for the horizontal axis are seconds. Therefore, if we multiply those quantities, we expect to obtain a result which has cm for units, which is exactly what we are looking for. We will return to this connection between area and the anti-derivative in Section 11.3.

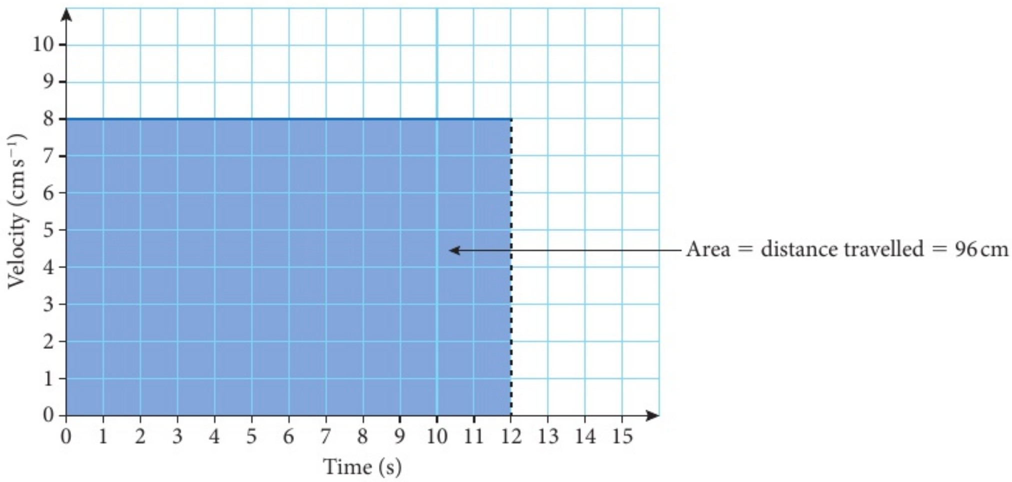


Figure 11.1 The shaded area represents the distance travelled in 12 seconds at 8 cm s^{-1}

Although we do not know the position function, $s(t)$, we do know that taking the derivative of the position function must give us the velocity function. That is, $\frac{\mathrm{d}}{\mathrm{d}t}[s(t)] = v(t)$ or $\frac{\mathrm{d}s}{\mathrm{d}t} = v$

So, if we could reverse the differentiation process, then we could find the position function from the velocity function. The 'reverse' of differentiation is, not surprisingly, often referred to as **anti-differentiation** (or finding an **anti-derivative**). The symbol for indicating the process of anti-differentiation is $\int$ and is referred to as an integral.

We are looking for the function $s(t)$ that has the derivative $s'(t) = v(t) = 8$, as shown in Figure 11.2.

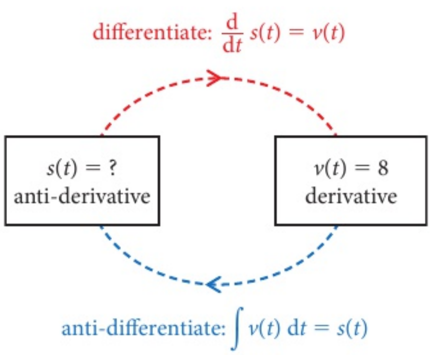


Figure 11.2 Derivative and anti-derivative

What function has a derivative with respect to t (time) that is 8? The function $s(t) = 8t$. However, $8t$ is not the only expression whose derivative is 8. For example, $8t + 2$, $8t - 3$ and $8t + 7.38$ all have 8 as a derivative:

Function	**Derivative**
$s(t) = 8t$	$s'(t) = v(t) = 8$
$s(t) = 8t + 2$	$s'(t) = v(t) = 8$
$s(t) = 8t - 3$	$s'(t) = v(t) = 8$
$s(t) = 8t + C$ where $C \in \mathbb{R}$	$s'(t) = v(t) = 8$

Notations used for differentiation (finding a derivative) include $\frac{\mathrm{d}}{\mathrm{d}x}(y)$, $\frac{\mathrm{d}y}{\mathrm{d}x}$, y' which all translate to 'the derivative of y' ('... with respect to x' for the first two examples). The notation $\int y\,\mathrm{d}x$ translates to 'the anti-derivative of y with respect to x', or 'the **integral** of y with respect to x'. As far as we are concerned, anti-derivative and integral mean the same thing.

It can be helpful to think about the meaning of the language we use. To find a derivative we differentiate, because we are finding a function that gives us the differential or difference or rate of change. When we anti-differentiate we integrate, because we are putting the differential pieces back together to find a function that gives us the total.

The last line is important: we add a constant C, known as the **constant of integration**, to form the **general anti-derivative**. The general anti-derivative represents all the functions that have $v(t)$ as their derivative. Since there are many anti-derivatives of $v(t) = 8$, which one should we use?

We will start with the general anti-derivative, but we will solve for the value of C to find a specific anti-derivative for our situation. We are interested in how far the toy car has travelled since time $t = 0$, and we know that the car has travelled 0 cm when $t = 0$. Therefore, we can write

$$s(0) = 0 \Rightarrow 8(0) + C = 0 \Rightarrow C = 0$$

Thus, the correct anti-derivative for this situation is $s(t) = 8t$. And, it works. We can check our previous answer by finding the position after 12 seconds: $s(12) = 8(12) = 96\,\text{cm}$

Using integral notation, we write:

$$s(t) = \int v(t)\,\mathrm{d}t = \int 8\,\mathrm{d}t = 8t + C$$

Notice also that the units for the anti-derivative are the product of the units of the range and the units of the domain:

cm s^{-1}	$\times$	s	$=$	cm
units of range of derivative		units of domain of derivative		units of anti-derivative

This might seem like a lot of work to confirm what we already know from the simple model distance = velocity × time. However, that simple model only works for constant velocities. When the velocity is changing, we do need anti-differentiation.

For example, what if the toy car is rolling down a 400 cm inclined ramp, so that the car's velocity is increasing, as in Figure 11.3?

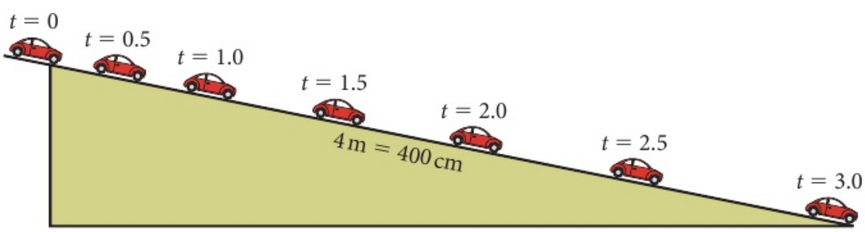


Figure 11.3 Velocity increasing as time increases

Suppose the velocity of the toy car is given by the function $v(t) = 90t$

How can we find the position function now?

We need to find the function $s(t)$ that has a derivative of $v(t) = 90t$

We know that when we differentiate, we decrease the exponent by 1. So to anti-differentiate, we will increase the exponent by 1. We can try $s(t) = 90t^2$ and use differentiation to check:

$$s(t) = 90t^2 \xrightarrow{\text{differentiate}} 90(2t^{2-1}) = 180t \neq v(t)$$

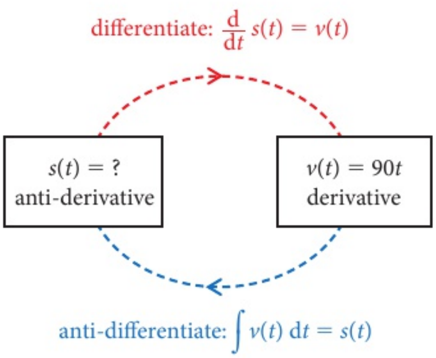


Figure 11.4 For what function is $v(t) = 90t$ the derivative?

This is not correct; when we differentiate $90t^2$ we get $180t$, which is twice the correct $v(t)$. This is because the power rule requires that we multiply the coefficient by the exponent. Since the exponent in $90t^2$ is 2, it doubled the expression for $v(t)$. Therefore, we can make the anti-derivative work by multiplying it by $\frac{1}{2}$. Let's check:

$$s(t) = 90\left(\frac{1}{2}\right)t^2 = 45t^2 \xrightarrow{\text{differentiate}} 45(2t^{2-1}) = 90t = v(t) \checkmark$$

Our expression for $s(t)$ now has a derivative equal to $v(t)$, so $s(t)$ is a correct anti-derivative. However, recall that there are many other correct anti-derivatives, because we can add a constant term C to $s(t)$ and it will still be a correct anti-derivative, as long as $C \in \mathbb{R}$:

$$s(t) = 45t^2 + C \underset{\text{anti-differentiate}}{\overset{\text{differentiate}}{\rightleftarrows}} v(t) = 90t$$

We will determine the value of C later, depending on the specific application.

The correct notation for anti-differentiation is the integral notation. For the toy car, we should write:

$$\begin{aligned} s(t) &= \int v(t)\,\mathrm{d}t \\ &= \int 90t\,\mathrm{d}t \\ &= 45t^2 + C \end{aligned}$$

This is the notation we will use in the rest of this chapter. By convention, we will always add the **constant of integration** C to indicate that many anti-derivatives are possible. We will usually not write it until our final answer.

So far, we have found anti-derivatives by inspection. That is, we look at the given function and use reasoning (and maybe some trial and error) to determine what function would have the given function as its derivative. Can we find a rule, like the power rule for differentiation, that can help us anti-differentiate?

The power rule for differentiation: for a function of the form $f(x) = ax^n$, the derivative is equal to $f'(x) = anx^{n-1}$. In words: To differentiate a power expression, multiply by the exponent, then reduce the exponent by one.

Example 11.1

Find an anti-derivative with respect to x of each of the following terms.

(a) 1 (b) x (c) x^2 (d) x^3 (e) x^4

Solution

(a) The derivative of x is 1, so an anti-derivative of 1 is x. We can write $\int 1\,\mathrm{d}x = x + C$

(b) At first, we might guess that the anti-derivative is x^2

However, $\frac{\mathrm{d}}{\mathrm{d}x}x^2 = 2x$ so we will need to fix the anti-derivative by multiplying it by $\frac{1}{2}$, which works since $\frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{1}{2}x^2\right) = x$

Therefore, $\int x\,\mathrm{d}x = \frac{1}{2}x^2 + C$

Since the question only asked for *an* anti-derivative, rather than the derivative in general form, we don't strictly need to add $+C$ since $\int 1\,\mathrm{d}x = x$ would represent the anti-derivative where $C = 0$. However, it is a good habit to do so.

(c) Again, we might guess the anti-derivative is x^3. But, as before, $\frac{d}{dx}x^3 = 3x^2$ so we will need to fix the anti-derivative by multiplying it by $\frac{1}{3}$ to obtain $\int x^2\,dx = \frac{1}{3}x^3 + C$

(d) We can start to see a pattern: increase the exponent by one, and multiply by the reciprocal of the new exponent. Therefore, $\int x^3\,dx = \frac{1}{4}x^4 + C$

(e) Increase the exponent by one, then multiply by the reciprocal of the new exponent: $\int x^4\,dx = \frac{1}{5}x^5 + C$

The anti-derivative of x^n

$\int x^n\,dx = \frac{1}{n+1}x^{n+1} + C$ where $n \in \mathbb{Z}, n \neq -1$

In words: to anti-differentiate, increase the exponent by one then multiply by the reciprocal of the new exponent.

From Example 11.1, we can detect the pattern $\int x^n\,dx = \frac{1}{n+1}x^{n+1} + C$

Does this work for all values of n? The answer is: almost. Remember, we cannot divide by zero, so we need $n + 1 \neq 0 \Rightarrow n \neq -1$

For this reason, we must exclude the exponent $n = -1$

Although this rule works for all other real-number values of n, in this course we will restrict n to the set of integers, excluding -1.

In **general form** means we add the constant of integration C to our final answer.

When we write $\int(8t^2 + 5t)\,dt$, it does not mean that dt is multiplied by $(8t^2 + 5t)$. You may also see it written without the parentheses, $\int 8t^2 + 5t\,dt$, and even sometimes within the expression itself: $\int \frac{dt}{t^2}$. Regardless, remember that dt is part of the notation.

We use the constant C to represent all the constants of integration; there is no need to use more than one.

Example 11.2

Find each anti-derivative in general form.

(a) $\int 4t\,dt$

(b) $\int 3x^2\,dx$

(c) $\int 14\,dt$

(d) $\int x^6\,dx$

(e) $\int(6x + 2)\,dx$

(f) $\int(8t^2 + 5t)\,dt$

(g) $\int ax^2\,dx$

(h) $\int(8n^{-3} + 5n^{-2} + 3)\,dn$

(i) $\int\left(\frac{5}{x^2} + 3x\right)dx$

(j) $\int\left(\frac{x + 200}{x^4}\right)dx$

Solution

(a) Increase the exponent by 1, and multiply by the reciprocal of the new exponent:

$\int 4t\,dt = 4\left(\frac{1}{2}t^{1+1}\right) = 2t^2 + C$

(b) Increase the exponent by 1, and multiply by the reciprocal of the new exponent:

$\int 3x^2\,dx = 3\left(\frac{1}{3}x^{2+1}\right) = x^3 + C$

(c) Notice we are integrating with respect to t. Therefore, $\int 14\,dt = 14t + C$

(d) $\int x^6\,dx = \frac{1}{7}x^7 + C$

(e) For a sum, we anti-differentiate each term:

$\int(6x + 2)\,dx = 6\left(\frac{1}{2}x^2\right) + 2x = 3x^2 + 2x + C$

(f) $\int(8t^2 + 5t)\,dt = 8\left(\frac{1}{3}t^3\right) + 5\left(\frac{1}{2}t^2\right) = \frac{8}{3}t^3 + \frac{5}{2}t^2 + C$

(g) We are integrating with respect to x, so we treat a as a constant coefficient. Thus,

$$\int ax^2\,dx = a\left(\frac{1}{3}x^3\right) = \frac{a}{3}x^3 + C$$

(h) We treat negative exponents in the same manner as positive exponents.

$$\begin{aligned}\int(8n^{-3} - 5n^{-2} + 3)\,dn &= 8\left(\frac{1}{-2}n^{-2}\right) - 5\left(\frac{1}{-1}n^{-1}\right) + 3n \\ &= -4n^{-2} + 5n^{-1} + 3n \\ &= -\frac{4}{n^2} + \frac{5}{n} + 3n + C\end{aligned}$$

(i) First, rewrite as a negative exponent, then anti-differentiate:

$$\begin{aligned}\int\left(\frac{5}{x^2} + 3x\right)dx &= \int(5x^{-2} + 3x)\,dx \\ &= 5\left(\frac{1}{-1}x^{-1}\right) + 3\left(\frac{1}{2}x^2\right) \\ &= -\frac{5}{x} + \frac{3}{2}x^2 + C\end{aligned}$$

(j) Rewrite into terms of the form ax^n, then anti-differentiate:

$$\begin{aligned}\int\left(\frac{x + 200}{x^4}\right)dx &= \int\left(\frac{x}{x^4} + \frac{200}{x^4}\right)dx \\ &= \int(x^{-3} + 200x^{-4})\,dx \\ &= \frac{1}{-2}x^{-2} + 200\left(\frac{1}{-3}x^{-3}\right) \\ &= -\frac{1}{2x^2} - \frac{200}{3x^3} + C\end{aligned}$$

Now we look at the constant of integration, C. In our toy car example, we found $v(t) = 90t$ and $s(t) = 45t^2 + C$

What value should we use for C? The answer is that it depends where we want to measure the position from. Since velocity is positive in this example, the toy car is moving in a positive direction. Therefore, we could make the top of the ramp position 0, so that $s(0) = 0$

This allows us to solve for C:

$s(0) = 0 \Rightarrow 45(0)^2 + C = 0 \Rightarrow C = 0$

Thus, the specific anti-derivative for this situation is $s(t) = 45t^2$

Example 11.3

A skydiver's vertical velocity is modelled by the function $v(t) = -9.8t$ where $v(t)$ is her vertical speed in $\mathrm{m\,s^{-1}}$ and t is time in seconds since she jumped out of an aeroplane, $0 \leqslant t \leqslant 10$

The aeroplane has an altitude of 3900 m at the instant the skydiver jumps out.

(a) Find a specific function for $s(t)$, the altitude of the skydiver at time t.

(b) Write down the units for $s(t)$.

(c) Hence, find the altitude of the skydiver at 1, 5, and 10 seconds, to 4 significant figures.

Solution

(a) Use anti-differentiation to find a function for $s(t)$:

$$\begin{aligned} s(t) &= \int v(t)\,dt \\ &= \int -9.8t\,dt \\ &= -9.8\left(\frac{1}{2}t^2\right) \\ &= -4.9t^2 + C \end{aligned}$$

To find C, recall that we want $s(t)$ to give us the altitude of the skydiver, and she was initially at 3900 m. With this, we can solve for C:

$s(0) = 3900 \Rightarrow -4.9(0)^2 + C = 3900 \Rightarrow C = 3900$

Thus, the anti-derivative that gives the correct position function is

$s(t) = -4.9t^2 + 3900$

(b) The units for the position function are:

$$\underbrace{\text{m s}^{-1}}_{\text{units of range of derivative}} \times \underbrace{\text{s}}_{\text{units of domain of derivative}} = \underbrace{\text{m}}_{\text{units of anti-derivative}}$$

(c) Substitute the values into the position function we have found:

$s(1) = -4.9(1)^2 + 3900 = 3895\text{ m}$

$s(5) = -4.9(5)^2 + 3900 = 3778\text{ m}$

$s(10) = -4.9(10)^2 + 3900 = 3410\text{ m}$

As we saw with the toy car and the skydiver, we will often want to find the specific anti-derivative for a given situation. This is finding the constant of integration using a **boundary condition**. A boundary condition is a known value that the anti-derivative must satisfy: for example, the altitude of 3900 m at time $t = 0$ for the skydiver.

Example 11.4

Find the specific anti-derivative for the given boundary condition.

(a) $f'(x) = 50x + 25$, $f(1) = 60$

(b) $\dfrac{dx}{dt} = 10t^3 + \dfrac{18}{t^2}$, when $t = 2$, $x = 46$

(c) $C'(n) = 100n^2 - 50n + 120$, $C(9) = 355$

Solution

(a) First find the general anti-derivative:

$$\int f(x)\,dx = \int 50x + 25\,dx$$
$$= 25x^2 + 25x + C$$

$f(1) = 60$, so substitute and solve for C:

$$f(1) = 60 \Rightarrow 25(1)^2 + 25(1) + C = 60$$
$$\Rightarrow 50 + C = 60$$
$$\Rightarrow C = 10$$

Therefore, the specific anti-derivative is $f(x) = 25x^2 + 25x + 10$

(b) First find the general anti-derivative:

$$x = \int \frac{dx}{dt}\,dt$$
$$= \int \left(10t^3 + \frac{18}{t^2}\right) dt$$
$$= \frac{5}{2}t^4 - \frac{18}{t} + C$$

When $t = 2$, $x = 46$, so substitute and solve for C:

$$\frac{5}{2}(2)^4 - \frac{18}{(2)} + C = 46$$
$$31 + C = 46$$
$$C = 15$$

Therefore, the specific anti-derivative is $x = \frac{5}{2}t^4 - \frac{18}{t} + 15$

(c) First find the general anti-derivative:

$$\int C'(n)\,dn = \int 100n^2 - 50n + 120\,dn$$
$$= \frac{100}{3}n^3 - 25n^2 + 120n + C$$

$C(9) = 355$, so substitute and solve for C:

$$C(9) = 355 \Rightarrow \frac{100}{3}(9)^3 - 25(9)^2 + 120(9) + C = 355$$
$$\Rightarrow 23\,355 + C = 355$$
$$\Rightarrow C = -23\,000$$

Therefore, the specific anti-derivative is

$$C(n) = \frac{100}{3}n^3 - 25n^2 + 120n - 23\,000$$

Of course, there are many applications of anti-derivatives. Whenever we have a rate of change and want to find the total or cumulative change, we can use anti-differentiation.

Example 11.5

A factory has modelled the marginal cost of producing sunglasses as $C'(n) = 45n^2 - 130n + 80$ where $C'(n)$ represents the additional cost in US dollars (USD) for an additional pair of sunglasses and n is the current production level in thousands of sunglasses per month. The factory is currently producing 1000 pairs of sunglasses per month at a cost of 50 000 USD.

(a) Write down the units for the anti-derivative function $C(n)$.

(b) Find a function $C(n)$ for the cost in thousands of USD for producing n thousand sunglasses per month.

(c) Find the minimum total production cost and the corresponding production level.

Solution

(a) Since $C'(n)$ is in USD per pair of sunglasses, and n is the current production level in thousands of sunglasses per month, we have:

$$\frac{\text{USD}}{\text{pair of sunglasses}} \times \text{thousands of sunglasses} = \text{USD (thousands)}$$

or thousands of USD.

(b) Find the anti-derivative $\int C'(n)\,dn = \int 45n^2 - 130n + 80\,dn$ to find $C(n)$:

$$\begin{aligned} C(n) &= \int 45n^2 - 130n + 80\,dn \\ &= 45\left(\frac{1}{3}n^3\right) - 130\left(\frac{1}{2}n^2\right) + 80(n) + D \\ &= 15n^3 - 65n^2 + 80n + D \end{aligned}$$

Since we know that the current production level is 1000 sunglasses per month, at a cost of 50 000 USD, we know the anti-derivative must satisfy $C(1) = 50$, so we can solve for D:

$$\begin{aligned} C(1) = 50 &\Rightarrow 15(1)^3 - 65(1)^2 + 80(1) + D = 50 \\ &\Rightarrow 30 + D = 50 \\ &\Rightarrow D = 20 \end{aligned}$$

Therefore the cost function is $C(n) = 15n^3 - 65n^2 + 80n + 20$

We usually use C as the constant of integration, but you can use another letter if C is already being used.

(c) Since the derivative is known (the marginal cost function), graph the derivative and apply the first derivative test.

The derivative changes from negative to positive at $n = 2$, so the minimum total cost must be when 2000 pairs of sunglasses per month are produced. The total cost will be

$C(2) = 15(2)^3 - 65(2)^2 + 80(2) + 20 = 40 \Rightarrow 40\,000$ USD

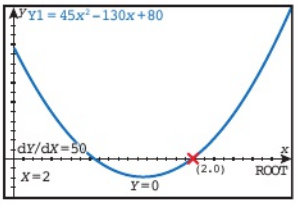


Figure 11.5 The function $C(n)$ changes from negative to positive at $n = 2$

Exercise 11.1

1. Find each anti-derivative in general form.

(a) $\int (x + 2)\,dx$ **(b)** $\int x^5\,dx$ **(c)** $\int 42\,dt$

(d) $\int a\,dn$ **(e)** $\int 5x^4\,dx$ **(f)** $\int 18t^2\,dt$

(g) $\int (-3x^2 + 5x - 8)\,dx$ **(h)** $\int (-9.8t + v_0)\,dt$

(i) $\int (2ax + b)\,dx$ **(j)** $\int (8x^3 + 27x^2 - 8x - 6)\,dx$

(k) $\int \left(14x^3 + 5x^2 - \frac{1}{2}\right) dx$ **(l)** $\int \frac{5}{h^2}\,dh$

(m) $\int \left(\frac{500 + w}{w^3}\right) dw$

2. For each function, explain why you cannot find an anti-derivative by using the rules in this section.

(a) $f(x) = \frac{1}{x}$ (b) $g(x) = \frac{400 + x^2}{x}$

3. Find the general anti-derivative of each function.

(a) $f(t) = 3t^2 - 2t + 1$

(b) $g(x) = \frac{1}{3} - \frac{2}{7}x^3$

(c) $f(t) = (t - 1)(2t + 3)$

(d) $f(x) = (3 + 2x)^2$

(e) $C(n) = (35n^2 - 110n + 70)n$

4. Find the specific anti-derivative for the given boundary condition.

(a) $f'(x) = 4x + 3, f(0) = 12$

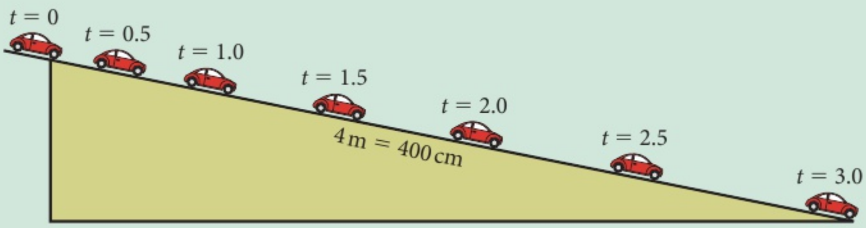
(b) $\frac{dy}{dx} = \frac{24}{x^3} + 5$, when $x = 2, y = 26$

(c) $f'(t) = 3t^2 - 2t + 1, f(10) = 1000$

(d) $\frac{dw}{dx} = \frac{x^2 - 350x}{x^4}$, when $x = 2, w = 44$

(e) $g'(x) = 500x^2 - 5500x + 12\,000, g(3) = 13\,000$

5. A toy car is rolling down a 400 cm long inclined ramp as shown in the diagram.



The velocity along the ramp at time t is given by $v(t) = -80t$, where $v(t)$ is the velocity in cm s^{-1} at time t in seconds.

(a) Find a function for $s(t)$, the distance of the car from the bottom end of the ramp.

(b) Write down the units for $s(t)$.

(c) Hence, find the distance of the toy car from the lower end of the ramp at 1 and 2 seconds.

(d) Use your function for $s(t)$ to predict and interpret the position of the toy car at 4 seconds.

(e) Find the time at which the toy car will reach the end of the ramp.

6. Water is draining from a container holding 150 litres of water.

   The change in water volume is given by $\frac{\mathrm{d}V}{\mathrm{d}t} = -2.5 + 0.02t$ where t is in minutes since the water started draining.

   (a) Write down the units of $\frac{\mathrm{d}V}{\mathrm{d}t}$

   (b) Find a function for the remaining water volume at time t.

   (c) Find the time at which half of the water is remaining.

   (d) Find the time when the water container will be empty.

7. Skydivers Antonia and Baxter are in a helicopter hovering at 2100 m. Antonia jumps and falls for 4 seconds before opening her parachute. The helicopter then ascends to 2300 m. Baxter jumps exactly 45 seconds after Antonia jumps; he falls for 13 seconds before opening his parachute. Assume that both skydivers descend at a constant rate of $5\,\mathrm{m\,s^{-1}}$ with their parachutes open and that their velocity when they are in free fall, before they open their parachutes, is described by the function $v(t) = -9t$

   (a) Find the altitude at which Antonia opens her parachute.

   (b) Find the altitude at which Baxter opens his parachute.

   (c) Baxter lands first. How long does he wait until Antonia lands?

8. A large sports-utility vehicle is slowing down from $108\,\mathrm{km\,h^{-1}}$ at a constant rate of $7.5\,\mathrm{m\,s^{-2}}$.

   (a) Write down the function for the acceleration $a(t)$ of the vehicle.

   (b) Find the velocity of the vehicle in $\mathrm{m\,s^{-1}}$.

   (c) Find a function for the velocity $v(t)$ of the vehicle where t is the time in seconds since the vehicle began slowing down.

   (d) It takes q seconds for the vehicle to stop. Find the value of q.

   (e) Using your function for $v(t)$, find a function for the distance $s(t)$ the vehicle has travelled since it began slowing down.

   (f) Find the distance travelled in the q seconds it takes to stop.

9. A small bicycle racing wheel company has modelled the marginal revenue from producing bicycle wheels as $R'(n) = -0.03n(n - 500)$ where $R'(n)$ represents the additional revenue in euros from producing additional wheels at the production level of n wheels per month. The company is currently earning a monthly revenue of €284 000 producing 200 wheels per month.

   (a) Write down the units for $R(n)$.

   (b) Find a function for the monthly revenue $R(n)$.

   (c) Find the production level that will maximise the monthly revenue.

   (d) Find the maximum monthly revenue.

11.2 Definite integrals

Consider the skydiver model we explored in Example 11.3. We used the velocity function $v(t) = -9.8t$, and the fact that the skydiver jumps from an altitude of 3900 metres, to find the function $s(t) = -4.9t^2 + 3900$ to model the skydiver's altitude (vertical position). With this function, we can find the altitude of the skydiver. However, we may also be interested in the change in altitude during some interval. For example, by how much did the skydiver's altitude change during the first 5 seconds? Or during the interval from 5 to 10 seconds?

To answer these questions, we could, of course, find the altitude at each of those times and subtract. For example:

- For the interval from 0 to 5 seconds, the skydiver's altitude changed by $s(5) - s(0) = [-4.9(5)^2 + 3900] - [-4.9(0)^2 + 3900] = -122.5$ metres.
- For the interval from 5 to 10 seconds, the skydiver's altitude changed by $s(10) - s(5) = [-4.9(10)^2 + 3900] - [-4.9(5)^2 + 3900] = -367.5$ metres.

This is the essence of **definite integration**.

We use a slightly different notation for definite integration: we write the **limits of integration** of the interval next to the integration symbol $\int$.

> Let $F(x)$ be any anti-derivative of $f(x)$. Then, the definite integral $\int_a^b f(x)\,dx = F(b) - F(a)$
>
> This simple rule is part of the **fundamental theorem of calculus**.

In our skydiver example, if we want to find the change in altitude from 5 to 10 seconds, given $v(t) = -9.8t$, we write

Change in altitude $= \int_5^{10} v(t)\,dt = \int_5^{10} -9.8t\,dt$

We know that the general anti-derivative for $-9.8t$ is $-4.9t^2 + C$

However, the rule above tells us that we can use any anti-derivative, so, for simplicity, we will let $C = 0$ and use $-4.9t^2$ for the anti-derivative.

Therefore, Change in altitude $= \int_5^{10} -9.8t\,dt = [-4.9(10)^2] - [-4.9(5)^2]$
$= -367.5$ metres

Mathematicians also use the following notation, to avoid writing the anti-derivative twice:
$\int_5^{10} -9.8t\,dt = [-4.9t^2]_5^{10}$
$= -367.5$
When we write $[-4.9t^2]_5^{10}$, it means $[-4.9(10)^2] - [-4.9(5)^2]$

Example 11.6

Find the value of each definite integral.

(a) $\int_0^2 3x^2\,dx$ (b) $\int_3^6 14\,dt$ (c) $\int_{-1}^1 x^3\,dx$

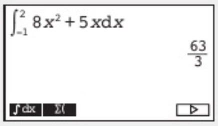
(d) $\int_{-4}^{-2} (6x + 2)\,dx$ (e) $\int_{-1}^2 (8t^2 + 5t)\,dt$

Solution

(a) $\int_0^2 3x^2\,dx = [x^3]_0^2 = 2^3 - 0^3 = 8$

(b) $\int_3^6 14\,dt = [14t]_3^6 = 14(6) - 14(3) = 42$

You can calculate the value of definite integrals directly using a GDC.



(c) $\int_{-1}^{1} x^3\,dx = \left[\frac{1}{4}x^4\right]_{-1}^{1} = \frac{1}{4}(1)^4 - \frac{1}{4}(-1)^4 = 0$

(d) $\int_{-4}^{-2}(6x+2)\,dx = [3x^2+2x]_{-4}^{-2} = -32$

(e) $\int_{-1}^{2}(8t^2+5t)\,dt = \left[\frac{8}{3}t^3 + \frac{5}{2}t^2\right]_{-1}^{2} = 31.5$

Since your GDC can calculate the value of definite integrals, we can now find the values of many anti-derivatives that do not match the rules we learned in Section 11.1.

Example 11.7

Use your GDC to calculate the value of the definite integral to 3 significant figures.

(a) $\int_{3}^{5}\left(\frac{450-w}{w^2}\right)dw$ (b) $\int_{1}^{5} e^{2x}\,dx$

Solution

Sample screen shots are shown in Figure 11.6.

(a) To 3 significant figures, $\int_{3}^{5}\left(\frac{450-w}{w^2}\right)dw = 59.5$

(b) To 3 significant figures, $\int_{1}^{5} e^{2x}\,dx = 11\,000$

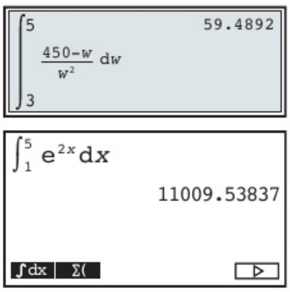
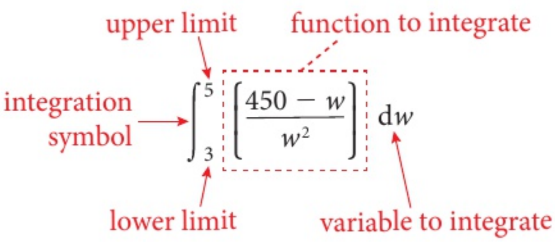


Figure 11.6 GDC screens for the solution to Example 11.7

In examinations, you will be expected to write the correct definite integral expression before using your GDC to calculate its value. For example, in Example 11.7 part (a) you would need to write all of the following:



Example 11.8

A tank containing $10\,m^3$ of water is draining. The change in water volume is given by $\frac{dV}{dt} = 0.0002t - 0.5$ where t is in minutes since the tank started draining.

(a) Find the change in the volume of water in the tank in the first 3 minutes.

(b) Find the amount of water that drains in the time between 4 and 8 minutes since the tank started draining.

(c) Find the time at which the tank will be empty.

Solution

(a) The definite integral required is $\int_0^3 \frac{dV}{dt}\,dt = \int_0^3 (0.0002t - 0.5)\,dt$

We can use a GDC or calculate by hand:

$$\begin{aligned}\int_0^3 (0.0002t - 0.5)\,dt &= [0.0001t^2 - 0.5t]_0^3\\ &= [0.0001(3)^2 - 0.5(3)] - [0.0001(0)^2 - 0.5(0)]\\ &= -1.50\,\text{m}^3\end{aligned}$$

(b) The definite integral required is $\int_4^8 \frac{dV}{dt}\,dt = \int_4^8 (0.0002t - 0.5)\,dt$

We can use a GDC or calculate by hand to obtain

$\int_4^8 0.0002t - 0.5\,dt = -2.00$ as shown.

Since the question asked for the amount, we take the absolute value: $|-2.00| = 2.00\,\text{m}^3$ drained from the tank.

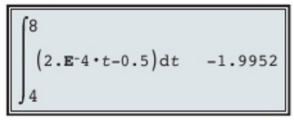


Figure 11.7 A GDC can quickly calculate the definite integral for Example 11.8 (b)

(c) There are two ways to approach this question.

Method 1: We want to know when the volume is zero, so find the specific anti-derivative and set it equal to zero:

$\int (0.0002t - 0.5)\,dt = 0.0001t^2 - 0.5t + C$

Since the volume is $10\,\text{m}^3$ at time $t = 0$, we can solve to find $C = 10$.

Hence, $0.0001t^2 - 0.5t + 10 = 0$

Use the quadratic formula or your GDC to find that volume is zero when $t = 20.1$ minutes.

Method 2: We want the change in the volume to represent the entire volume of the tank, that is, it must lose $10\,\text{m}^3$ of water. Hence, we are looking for $\int_0^x 0.0002t - 0.5\,dt = -10$, where x represents the time at which the tank is empty. Use the numerical solver on a GDC to solve this equation, as shown in Figure 11.8.

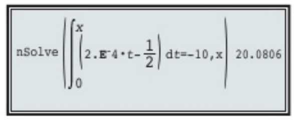


Figure 11.8 Using a GDC numerical solver in Example 11.8 (c), method 2. The tank will be empty after 20.1 minutes

Exercise 11.2

1. Evaluate each definite integral.

(a) $\int_2^7 8\,dx$ **(b)** $\int_{-2}^1 (3x^2 - 4x)\,dx$ **(c)** $\int_1^5 \frac{2}{t^3}\,dt$

(d) $\int_1^3 \frac{u^5 + 2}{u^2}\,du$ **(e)** $\int_1^3 (2 - \sqrt{x})^2\,dx$ **(f)** $\int_{-2}^2 |3w|\,dw$

(g) $\int_{-1}^1 \left(\frac{1}{x^2 + 1}\right)dx$

2. The flow rate in $\text{m}^3\,\text{h}^{-1}$ of a river recorded by a monitoring station can be modelled by the function $f(t) = 600t^3 - 7200t^2 + 21\,600t + 3600$ where t is in hours since 6:00, $0 \leqslant t \leqslant 8$

Find the total volume of water that flows past the monitoring station from 8:00 to 14:00.

3. The monthly production cost of a pencil factory changes according to the number of pencils produced. The marginal cost in thousands of USD per additional thousand pencils produced is known to be $\frac{dC}{dn} = 0.05n - 5$ where n is the current production level.
   (a) Find the additional cost when the factory increases production from 50 000 to 90 000 pencils.
   (b) The factory is currently spending 100 000 USD to produce 50 000 pencils. Find the total cost to produce 90 000 pencils.
   (c) Find the additional cost when the factory increases production from 90 000 to 140 000 pencils.
   (d) Hence, find the total cost to produce 140 000 pencils.

4. Cardiac output can be measured by injecting a dye into the right atrium of the heart and measuring the concentration of the dye as it leaves the heart. The dye concentration for a particular patient can be modelled by the function $c(t) = 0.25 + 3t - 0.27t^2 + 0.0061t^3$, where $c(t)$ is the concentration in mg l^{-1} at time t in seconds, $0 \leqslant t \leqslant 22$, after 8 mg of dye is injected. Since we know that all the dye will pass through the heart, we can use the formula $A = F\int_0^{22} c(t)\,dt$ where A is the amount of dye injected in mg and F is the cardiac output as a flow rate.
   (a) Find the units for F.
   (b) Find the cardiac output rate, F, for this patient.
   (c) Normal cardiac output is 3 to $4\,\text{l min}^{-1}$. Is this patient within normal range?

5. The change in water contained in a plant can be modelled by $\frac{dV}{dt} = -0.1t(24 - t) + k$ where $\frac{dV}{dt}$ is in ml h^{-1} at time t and k is the amount of water the plant receives from watering, in ml h^{-1}.
   (a) Find the total change in the water contained in the plant over 24 hours when $k = 5$
   (b) Find the value of k for the total change in water to be zero over 24 hours.

6. The growth rate of a small animal is given by $G(t) = \dfrac{9e^{-\frac{1}{2}(t-6)}}{2\left(e^{-\frac{1}{2}(t-6)} + 1\right)^2}$ where $G(t)$ is the change in mass in kg per year, at year t after birth.
   (a) Find the cumulative change in the mass of this animal:
      (i) in the first 5 years after birth
      (ii) in the first 15 years after birth
      (iii) during the second year after birth.
   (b) This animal had a mass of 0.5 kg at birth.
   Find the mass of the animal 5 years after birth.

7. A full swimming pool begins emptying at the rate of $\frac{\mathrm{d}V}{\mathrm{d}t} = 6 - \frac{2}{3}t$ where t is the time in hours and $\frac{\mathrm{d}V}{\mathrm{d}t}$ is in $\mathrm{m^3\,h^{-1}}$. The pool is completely empty after 9 hours. Find the volume of the pool when full.

8. The fuel consumption of an aircraft in $\mathrm{kg\,km^{-1}}$ on a certain 500 km flight can be modelled by the function $f(n) = (n + 17)^{1.25}\mathrm{e}^{-0.07(n+17)} - 0.001n + 3$ where n is the distance travelled in km.
   - **(a)** Use a GDC to find the distance where the rate of fuel consumption is greatest.
   - **(b)** Write down an integral to find the total fuel consumption for this 500 km flight.
   - **(c)** Find the total fuel consumption for this 500 km flight.

11.3 Area under a curve

Recall the electric toy car that moves in a straight line at a constant velocity of 8 centimetres per second ($\mathrm{cm\,s^{-1}}$) for 12 seconds. We wrote the velocity function $v(t) = 8$ and found the anti-derivative to find the total distance travelled in 12 seconds. Using the definite integral, we can write

$$\int_0^{12} 8\,\mathrm{d}t = [8t]_0^{12} = 96\,\mathrm{cm}$$

Notice that the value of the definite integral appears to be the area between the x-axis and the function we are integrating (Figure 11.9).

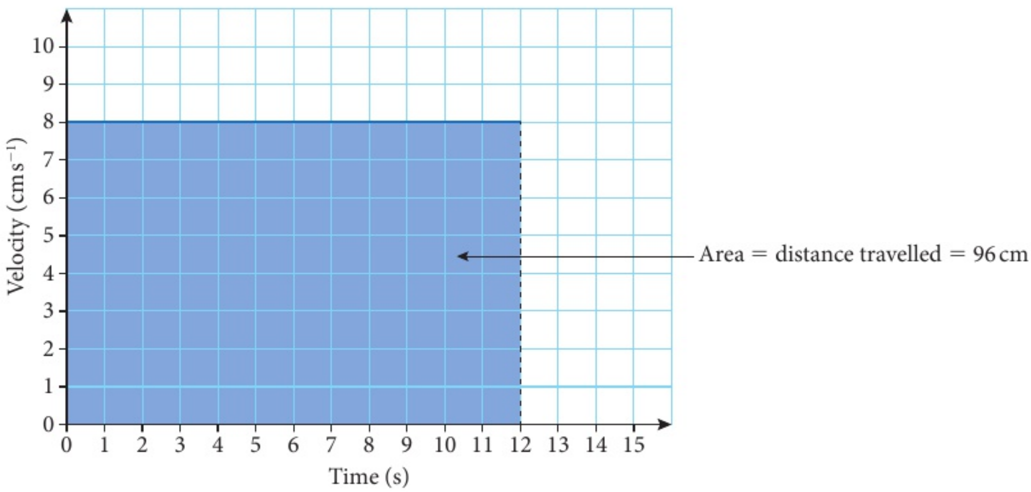


Figure 11.9 Area under the function

Is this always the case? We will test a few functions to see.

Example 11.9

For each function:

(i) draw a sketch of the function (use a GDC as needed) and shade the area between the curve and the x-axis for the interval given

(ii) use geometry to find the shaded area

(iii) calculate the value of the definite integral of the function for the same interval.

(a) $f(x) = 3, 1 \leqslant x \leqslant 5$

(b) $g(x) = 2x, 0 \leqslant x \leqslant 3$

(c) $h(x) = -x + 5, -1 \leqslant x \leqslant 1$

(d) $c(x) = \sqrt{4 - x^2}, -2 \leqslant x \leqslant 2$

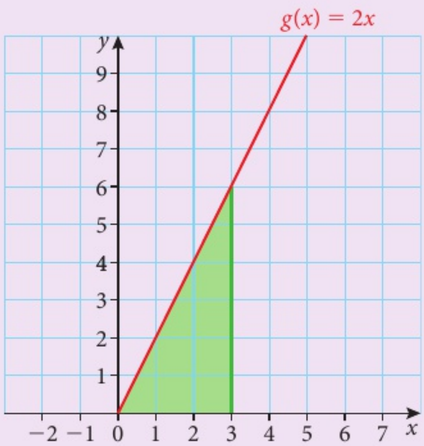
Solution

(a) (i)

(ii) This is a rectangle with sides of 4 and 3, so $A = bh = (4)(3) = 12$

(iii) $\int_1^5 3\,dx = [3x]_1^5 = 15 - 3 = 12$

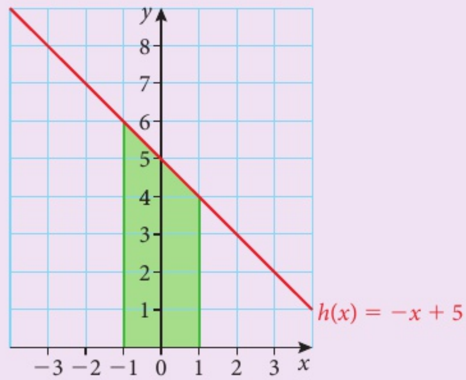
(b) (i)



(ii) This is a triangle with base 3 and height 6, so $A = \frac{1}{2}bh = \frac{1}{2}(3)(6) = 9$

(iii) $\int_0^3 2x\,dx = [x^2]_0^3 = 9 - 0 = 9$

(c) (i)

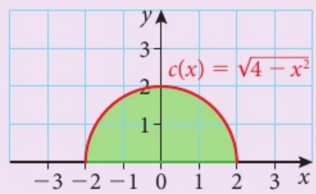


(ii) This is a trapezium with bases 4 and 6 and height 2, so

$A = \frac{1}{2}(b_1 + b_2)h = \frac{1}{2}(4 + 6)(2) = 10$

(iii) $\int_{-1}^{1}(-x + 5)\,\mathrm{d}x = \left[-\frac{1}{2}x^2 + 5x\right]_{-1}^{1} = \left[-\frac{1}{2} + 5\right] - \left[-\frac{1}{2} - 5\right] = 10$

(d) (i)



(ii) This is a semicircle with radius 2, so $A = \frac{1}{2}\pi r^2 = \frac{1}{2}\pi(2)^2 = 2\pi$
$= 6.28$

(iii) $\int_{-2}^{2}\sqrt{4 - x^2}\,\mathrm{d}x = 6.28$ (using a GDC)

When $f(x) < 0$ for some interval, the integral for that interval will give a negative value. Since a negative value doesn't make sense for an area, in this course we will not calculate areas for functions where $f(x) < 0$.

In Example 11.9 the value of the definite integral in each case is equal to the area. However, this is only true whenever the function is above (or touching) the x-axis, that is, $f(x) \geqslant 0$

The area between a function $f(x)$ and the x-axis for the interval $a \leqslant x \leqslant b$ is given by $\int_a^b f(x)\,\mathrm{d}x$ if and only if $f(x) \geqslant 0$ for the interval $a \leqslant x \leqslant b$

This is known as the area under a curve.

We may need to solve to find the bounds for the integral, depending on the situation and context.

Example 11.10

The graph of $f(x) = -x^2 + 2x + 3$ is given in Figure 11.10.

Find the area enclosed by the curve $f(x) = -x^2 + 2x + 3$ and the x-axis.

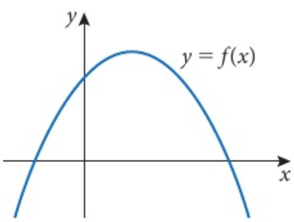


Figure 11.10 Graph of $f(x) = -x^2 + 2x + 3$

Solution

To find the area enclosed, we can use the definite integral. However, we need to know the bounds of integration. In this case, the bounds of integration are the x intercepts of the function, so we must solve $f(x) = 0$ by hand or by using a GDC.

$-x^2 + 2x + 3 = 0 \Rightarrow -(x + 1)(x - 3) = 0 \Rightarrow x = -1$ or $x = 3$

Then, we can set up and calculate the integral:

$$\int_{-1}^{3} f(x)\,\mathrm{d}x = \int_{-1}^{3} (-x^2 + 2x + 3)\,\mathrm{d}x$$

$$= \left[-\frac{1}{3}x^3 + x^2 + 3x\right]_{-1}^{3}$$

$$= \left[-\frac{1}{3}(3)^3 + (3)^2 + 3(3)\right] - \left[-\frac{1}{3}(-1)^3 + (-1)^2 + 3(-1)\right]$$

$$= \frac{32}{3}$$

Therefore the area enclosed is $\frac{32}{3}$ units^2.

Note that in examinations, you are expected to use technology to calculate the value, as shown in Figure 11.11.

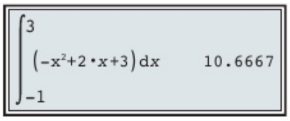


Figure 11.11 The definite integral gives the area under the curve in Example 11.10.

For more complex functions, we can rely on a GDC to find the value of the definite integral. We may also use two or more smaller regions to find the total area of a region.

Example 11.11

Find the area enclosed by curves $f(x) = \sqrt{x + 3}$, $g(x) = -2(x - 2)^3$, and the x-axis.

Solution

Begin by graphing the functions to see the areas we are looking for and finding relevant intersections.

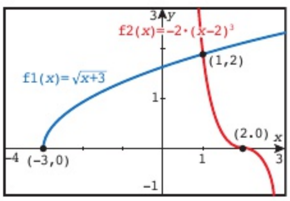


Figure 11.12 Using a GDC to find the intersection

We need to find the intersection of the curves in order to find the bounds of integration. Use a GDC to find that the intersection of $f(x)$ and $g(x)$ is $(1, 2)$. The curves appear to intersect the x-axis at -3 and 2, but it is a good idea to use a GDC to check, using the zero tool. Now, we can write down the definite integral required:

$$A = \int_{-3}^{1} f(x)\,\mathrm{d}x + \int_{1}^{2} g(x)\,\mathrm{d}x$$

$$= \int_{-3}^{1} (\sqrt{x + 3})\,\mathrm{d}x + \int_{1}^{2} (-2(x - 2)^3)\,\mathrm{d}x$$

Most GDCs can calculate integrals from the graphing page (Figure 11.13(a)).

From the graph, we see the separate areas; we then add the two areas to obtain a final answer of $5.33 + 0.5 = 5.83$ units^2. Alternatively, we can enter the integrals directly (Figure 11.13(b)).

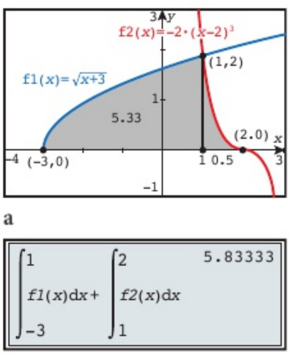


Figure 11.13 Definite integrals can be calculated on a graph screen (a) or directly (b)

In both cases, the total area enclosed by the curves $f(x) = \sqrt{x + 3}$, $g(x) = -2(x - 2)^3$, and the x-axis is 5.83 units^2.

Finding areas of irregular shapes is a powerful application of the definite integral.

Example 11.12

A ramp for skateboarding has a side piece that is modelled by the area enclosed by the y-axis, the x-axis, the function $f(x) = 1$, and the curve $g(x) = 3 - \sqrt{9 - (x - 3)^2}$, with all measurements in metres.

(a) Use your GDC to sketch the shape of the side piece on a coordinate plane using the equations given.

(b) Find the coordinates of the intersection of $f(x)$ and $g(x)$

(c) Find the area of the side piece in m^2.

Figure 11.14 Diagram for Example 11.12

Solution

(a) The four boundary lines are shown in the graph in Figure 11.15.

Figure 11.15 Graph showing the four boundary lines

(b) From a GDC, the intersection of $f(x)$ and $g(x)$ is $(0.764, 1)$.

(c) Integrate $f(x)$ from 0 to 0.764, and $g(x)$ from 0.764 to 3, and add the results:

$$A = \int_0^{0.764} f(x)\,\mathrm{d}x + \int_{0.764}^{3} g(x)\,\mathrm{d}x$$

$$= \int_0^{0.764} 1\,\mathrm{d}x + \int_{0.764}^{3} \left(3 - \sqrt{9 - (x - 3)^2}\right)\mathrm{d}x$$

Use a GDC to obtain the final value.

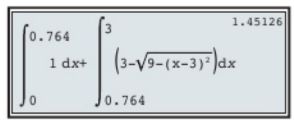


Figure 11.16 Using a GDC to find the final value

Hence $A = 1.45\,\text{m}^2$

Exercise 11.3

1. Find the area between the curve and the x-axis within the given bounds without using technology.

(a) $f(x) = 2x + 4, \quad 0 \leqslant x \leqslant 4$

(b) $y = -x^2 + 4, \quad 1 \leqslant x \leqslant 4$

(c) $f(x) = \dfrac{1}{x^2}, \quad 1 \leqslant x \leqslant 3$

2. Find the area between the curve and the x-axis within the given bounds.

(a) $y = 10 - x^2, \quad -2 \leqslant x \leqslant 2$

(b) $y = -x + 6, \quad 1 \leqslant x \leqslant 6$

(c) $g(x) = x^3 - 3x^2 - x + 8, \quad -1 \leqslant x \leqslant 3$

(d) $g(x) = 2^x, \quad 0 \leqslant x \leqslant 3$

(e) $y = 1 + \sqrt{x}, \quad 1 \leqslant x \leqslant 4$

(f) $f(x) = \dfrac{1}{\sqrt{x}} + 1, \quad 2 \leqslant x \leqslant 4$

3. Find the area enclosed between the curve and the x-axis.

(a) $y = 9 - x^2$

(b) $f(x) = -x^2 + 4x$

(c) $g(x) = x^3 - 3x^2 + 4$

(d) $y = |x - 4| - 3$

(e) $y = 1 - x^4$

4. Find the area enclosed by the given curves and the x-axis.

   (a) $f(x) = x^2, g(x) = -2x + 8$

   (b) $f(x) = \sqrt{x+3}, g(x) = \sqrt{5-x}$

   (c) $f(x) = \frac{1}{x}, y = x, x = 3$

5. Given the functions $f(x) = -\frac{1}{2}x(x-4)$ and $g(x) = \sqrt{4-(x-2)^2}$:

   (a) draw a graph of $f(x)$ and $g(x)$ on the domain $0 \leqslant x \leqslant 4$

   (b) find the area between $f(x)$ and the x-axis

   (c) find the area between $g(x)$ and the x-axis

   (d) hence, by subtracting areas, find the area between the curves $f(x)$ and $g(x)$ in the interval $0 \leqslant x \leqslant 4$

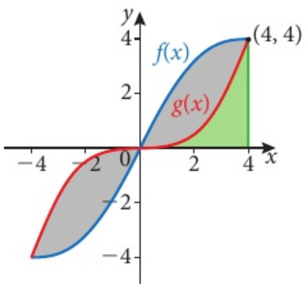


Figure 11.17 Toy propeller in question 6

6. A toy propeller can be modelled by the curves $f(x) = \frac{32x}{x^2+16}$ and $g(x) = \frac{1}{16}x^3$ as shown in Figure 11.17, where all measurements are in cm.

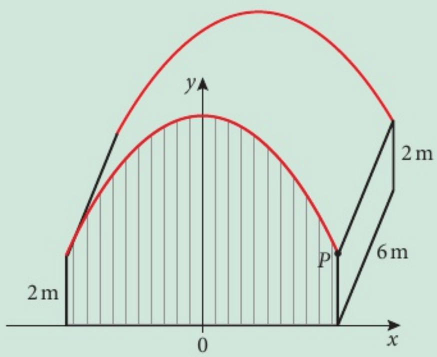
   (a) Find the area shaded in green, between $g(x)$ and the x-axis, for the interval $0 \leqslant x \leqslant 4$

   (b) Find the area between $f(x)$ and the x-axis for the interval $0 \leqslant x \leqslant 4$

   (c) Hence, by subtracting areas, find the total area of the toy propeller.

7. A large storage facility has a roof with a parabolic profile. The rectangular vertical walls are 6 m long and 2 m high, as shown in the diagram.

   The curve of the roof is given by the equation $y = 6 - x^2$ where x and y are in metres. The point P is the corner of the roof where it meets both vertical walls.



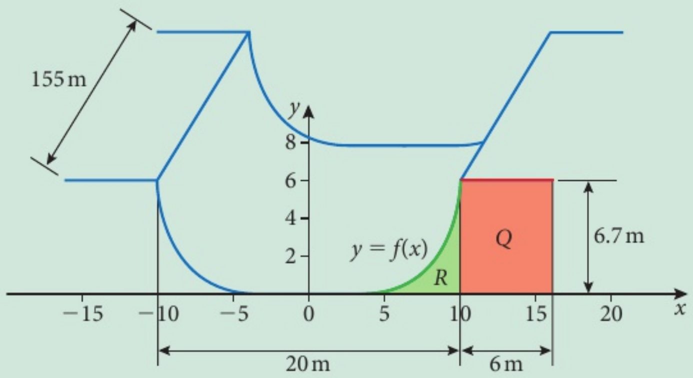
   (a) Find the coordinates of point P.

   (b) The shaded area represents the front wall of the storage facility. Find the area of the front wall.

   (c) Hence, find the total volume of the storage facility in m^3.

   (d) All four walls of the facility must be painted. Find the total area of all four walls in m^2.

8. A competition superpipe for snowboarding must have walls that are 6.7 m tall with top edges that are 20 m apart as shown in the diagram. The flat deck at the top of the wall must be 6 m wide and 6.7 m tall. The superpipe is 155 m long.



(a) The region R is modelled by the curve
$f(x) = \sqrt{6.7^2 + (x - 3.3)^2} - 6.7$, $3.3 \leqslant x \leqslant 10$ where all units are metres.

(i) Write down an integral expression for the area of region R.

(ii) Hence, find the area of region R.

(b) Find the area of region Q.

(c) Given that the superpipe is 155 m long, and the two sides are symmetric, find the volume of snow required to construct this superpipe to 3 significant figures.

11.4 Approximating area under a curve

So far we have seen that the area under a curve is exactly equal to the value of the definite integral. We can find the value of the definite integral by hand, using anti-differentiation (for some functions), or by using technology. However, some functions don't have an anti-derivative, and sometimes we may not even have a function – we may only have some data points.

For example, a researcher is monitoring the flow rate of a small river during a day. Table 11.1 gives a sample of the researcher's data.

Hour	0	2	4	6	8
Flow rate (thousands $m^3 h^{-1}$)	12	17	18	16	15

Table 11.1 Sample data

The researcher wishes to estimate the total volume of water that flowed during this time. How can this be done? If we had a function to model the flow rate, then we could use a definite integral to calculate the total volume of water, but we don't have a model so we must find a different approach. Start by generating a plot of the data.

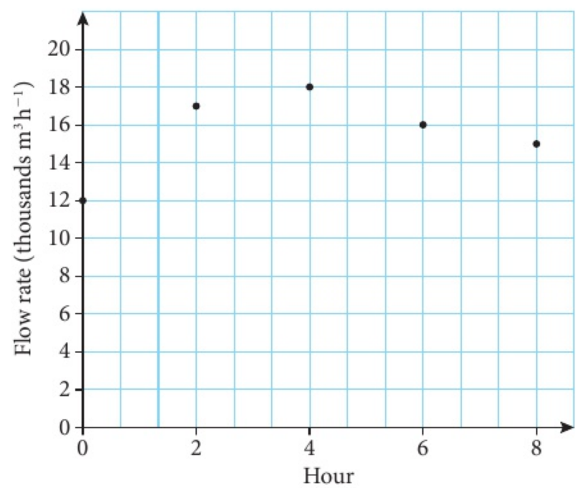


Figure 11.18 Plot of the flow rate data in Table 11.1

Since the definite integral is equal to area under the curve, we could use geometry to estimate the area. We can do this by drawing a series of trapeziums with bases parallel to the y-axis. Figure 11.19 shows these trapeziums, and the dimensions of the trapezium A_1 are labelled.

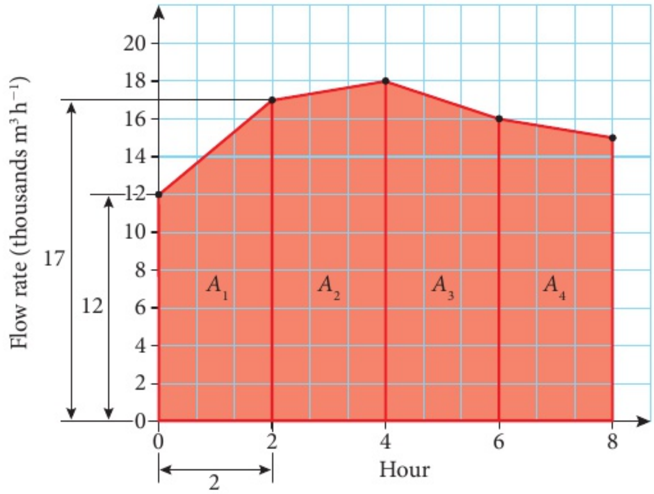


Figure 11.19 Using trapeziums to calculate the area under a curve based on data

Now we can calculate the area of each trapezium in turn. The lengths of the bases of each trapezium are the given data values (flow rates), while the height of the trapezium is the distance between the two bases. For example, trapezium A_1, which represents the interval between hours 0 and 2, has bases with lengths 12 and 17 and a height of 2, thus

The area of a trapezium is $A = \frac{1}{2}(b_1 + b_2)h$, where b_1 and b_2 are the lengths of the parallel bases and h is the 'height,' or distance between the bases.

$$A_1 = \frac{1}{2}(12 + 17)(2)$$

Repeating for each trapezium, the entire area is

$$\frac{1}{2}(12 + 17)(2) + \frac{1}{2}(17 + 18)(2) + \frac{1}{2}(18 + 16)(2) + \frac{1}{2}(16 + 15)(2)$$

$$= 129$$

Therefore, we estimate that a total of 129 000 m^3 of water flowed in this river during this 8 hour time period. Remember, this is an estimate since we don't know how the water flow rate changed between each reading – we are making an assumption that the flow rate between each pair of measurements can be estimated by the mean of the measurements. For example, we are estimating the flow rate between 0 and 2 hours as $\frac{1}{2}(12 + 17) = 14.5$ thousand $m^3\,h^{-1}$.

As with the anti-derivative and definite integral, the units for the area are the product of the units of the horizontal and vertical axes, that is, the product of the units of independent and dependent variables:

$$\underbrace{10^3\ m^3\ h^{-1}}_{\text{units of dependent variable}} \times \underbrace{h}_{\text{units of independent variable}} = \underbrace{10^3\ m^3}_{\text{units of area}}$$

Example 11.13

A medical test for diabetes measures glucose concentration in the patient's bloodstream to determine the patient's glucose tolerance by calculating the total area under the curve in mg h dl^{-1}. After a patient is given a glucose solution, the following data are collected.

Time since glucose given (h)	0.0	0.5	1.0	1.5	2.0
Glucose concentration (mg dl^{-1})	95	150	125	110	100

Estimate the area under the curve for this patient.

Solution

We will estimate the total glucose concentration by finding areas of trapeziums:

$$\begin{aligned}\text{Area under curve} &= \frac{1}{2}(95 + 150)(0.5) + \frac{1}{2}(150 + 125)(0.5)\\ &\quad + \frac{1}{2}(125 + 110)(0.5) + \frac{1}{2}(110 + 100)(0.5)\\ &= 241\ \text{mg h dl}^{-1}\end{aligned}$$

The method of dividing the area into trapeziums is known as the **trapezoidal rule**. The trapezoidal rule is one method that is used to estimate the value of definite integrals when the anti-derivative is unknown. While a GDC can quickly calculate the numerical value of a definite integral, it is still helpful to be able to use the trapezoidal rule.

Example 11.14

Use the trapezoidal rule to find the value of $\int_0^8 |2^x - 8|\,dx$, using trapeziums 0.8 units wide.

Solution

The trapeziums that we should use are 0.8 units wide. This means the x values we will use are $x = \{0, 0.8, 1.6, 2.4, 3.2, 4.0\}$. For each of these x values, we find the value of the function $f(x) = |2^x - 8|$ in order to find the length of the trapezium bases.

We can then build a table listing the x values, the value of the function, and the area calculation for each trapezium.

x	$f(x)$	Area
0	7	
		$A_1 = \frac{1}{2}(7 + 6.26)(0.8) = 5.30$
0.8	6.26	
		$A_2 = \frac{1}{2}(6.26 + 4.97)(0.8) = 4.49$
1.6	4.97	
		$A_3 = \frac{1}{2}(4.97 + 2.72)(0.8) = 3.08$
2.4	2.72	
		$A_4 = \frac{1}{2}(2.72 + 1.19)(0.8) = 1.56$
3.2	1.19	
		$A_5 = \frac{1}{2}(1.19 + 8)(0.8) = 3.68$
4.0	8	

Therefore, the total area is $5.30 + 4.49 + 3.08 + 1.56 + 3.68 = 18.1$ units^2.

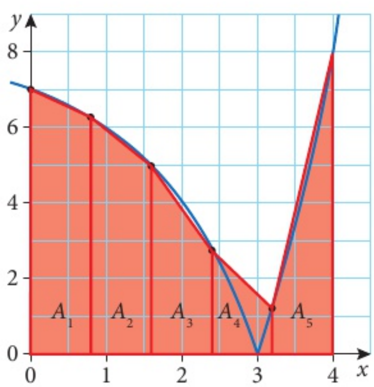


Figure 11.20 It helps to sketch a graph of the situation with the trapeziums shown

Notice that each area calculation involves multiplying by $\frac{1}{2}$ (from the trapezium formula) and 0.8 (the width of the trapeziums). Also, all bases except the first and last appear twice, so we can collect the 'inner' bases and multiply by two. Therefore, we can rearrange the calculation as shown:

$$A = \frac{1}{2}(7 + 6.26)(0.8) + \frac{1}{2}(6.26 + 4.97)(0.8) + \frac{1}{2}(4.97 + 2.72)(0.8)$$
$$+ \frac{1}{2}(2.72 + 1.19)(0.8) + \frac{1}{2}(1.19 + 8)(0.8)$$
$$= \frac{1}{2}(0.8)[(7 + 6.26) + (6.26 + 4.97) + (4.97 + 2.72) + (2.72 + 1.19) + (1.19 + 8)]$$
$$= \frac{1}{2}(0.8)\left[\underbrace{(7 + 8)}_{\substack{\text{first and}\\ \text{last base}}} + 2\underbrace{(6.26 + 4.97 + 2.72 + 1.19)}_{\text{all 'inner' bases}}\right]$$

This rearrangement shows that we can add the first and last base, plus twice the sum of all the 'inner' bases, then multiply the sum by half of the trapezium width. For a general formula, we can rewrite this as shown in the key fact box.

Suppose an interval $a \leqslant x \leqslant b$ is divided into n equal subintervals of width $h = \frac{b-a}{n}$, so that $x_0 = a, x_1 = a + h, x_2 = a + 2h, \dots, x_{n-1} = a + (n-1)h = b - h, x_n = a + nh = b$. Then, the area under a curve $y = f(x)$ for the interval $a \leqslant x \leqslant b$, where $f(x) \geqslant 0$, can be estimated by

$$\int_a^b y\,dx \approx \frac{1}{2}h\left[\underbrace{y_0 + y_n}_{\substack{\text{first and}\\ \text{last base}}} + 2\underbrace{(y_1 + y_2 + \dots + y_{n-1})}_{\text{all 'inner' bases}}\right] \text{ where } y_n = f(x_n)$$

This formula may look intimidating, but it is simply formalising what we have done above.

Example 11.15

You are given that $y = f(x)$.

(a) Write down an expression using the trapezoidal rule to estimate the area under the curve from $x = 1$ to $x = 3$ with intervals 0.5 units wide.

(b) The value A is equal to the area under the curve $f(x) = x^2 + 1$ when estimated using the trapezoidal rule. Find the value of A to 4 significant figures for the interval $1 \leqslant x \leqslant 3$.

(c) The value B is equal to the area under the curve $f(x) = x^2 + 1$.

(i) Write down the definite integral expression for B for the interval $1 \leqslant x \leqslant 3$.

(ii) Use a GDC to calculate the value of B to 4 significant figures.

(d) Find the percentage error of A relative to B.

Solution

(a) Since the interval is $3 - 1 = 2$ units wide, there are $n = \frac{2}{0.5} = 4$ subintervals. The trapezium bases are at $x_0, x_1, ..., x_4 = \{1, 1.5, 2, 2.5, 3\}$.

Therefore,

$$\int_a^b y\,\mathrm{d}x \approx \frac{1}{2}(0.5)\left[f(1) + f(3) + 2\sum_{k=1}^{3}(y_k)\right]$$
$$= 0.25\left[f(1) + f(3) + 2\left(f(1.5) + f(2) + f(2.5)\right)\right]$$

(b) We need to calculate the value of $f(x)$ for each x value, then evaluate the expression.

$$A = 0.25\left[f(1) + f(3) + 2\left(f(1.5) + f(2) + f(2.5)\right)\right]$$
$$= 0.25[2 + 10 + 2(3.25 + 5 + 7.25)]$$
$$= 10.75 \text{ units}^2$$

(c) (i) $B = \int_1^3 (x^2 + 1)\,\mathrm{d}x$

(ii) $B = 10.67$

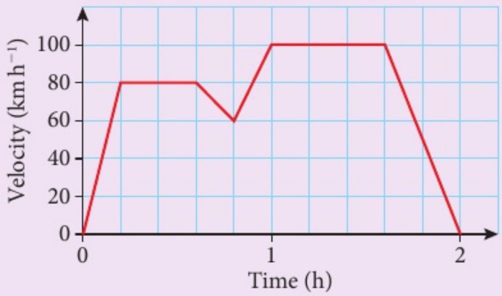
(d) Error $= \frac{A - B}{B} \times 100\% = \frac{10.75 - 10.67}{10.67} \times 100\% = 0.750\%$

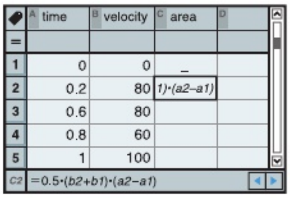
As you can see, the trapezoidal rule gives quite good approximations of the area under many functions, and can be made more accurate by using more trapeziums, i.e., making the width of each trapezium smaller.

Another way to make our work easier is to use a spreadsheet or a GDC to automate the calculation of the trapezium areas.

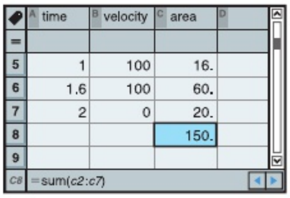
Example 11.16

The velocity–time graph shows a train journey. Use the trapezoidal rule to calculate the total distance travelled by the train.





a



b

Figure 11.21 Using a GDC spreadsheet to calculate area with the trapezium rule

Solution

The total distance travelled will be equal to the area under the curve. We can use a GDC to organise the coordinates representing the upper vertices of the trapeziums: (0, 0), (0.2, 80), (0.6, 80), (0.8, 60), (1, 100), (1.6, 100), and (2, 0). Then, by creating a general formula and filling it in adjacent cells, we can calculate the area of each trapezium. Figure 11.21(a) shows the first formula being entered. Notice that we want to add the lengths of the bases, but subtract the x-coordinates to find the height of the trapezium.

By filling the formula in adjacent cells and calculating the sum, we can find the area under the curve quickly and accurately (Figure 11.21(b)).

Therefore, the train has travelled 150 km in this two-hour journey.

Exercise 11.4

1. Find the area under the curve represented by the given data.

(a)

x	1	2	3	4
y	4	6	7	2

(b)

x	−3	0	5	6
y	5	2	1	1

(c)

x	1	1.5	3	3.5
y	1	3	5	11

2. For each function:

(i) sketch the curve and trapeziums for the indicated interval

(ii) estimate the area using the trapezoidal rule.

(a) $f(x) = -x(x - 5)$; $1 \leqslant x \leqslant 4$ with trapeziums every 1 unit

(b) $g(x) = x^3 - 4x^2 + 3x + 3$; $0 \leqslant x \leqslant 2$ with 4 equal-width trapeziums

(c) $y = |x^2 - 2x|$; $0 \leqslant x \leqslant 3$ with trapeziums every 1 unit

3. A yacht records the speeds shown in the table. A knot is a measure of speed used in maritime and aviation equal to one nautical mile per hour.

Time since leaving port (minutes)	0	30	45	60	72	90
Speed (knots)	5	5.5	6	6.2	5.8	4

(a) Write down the number of hours since the yacht left port.

(b) Estimate the total number of nautical miles this yacht has travelled using the trapezoidal rule.

4. A pilot of a small aeroplane records the effective airspeed data shown in Table 11.2.

Estimate the total distance this aeroplane has travelled using the trapezoidal rule.

Airspeed (km h^{-1})	Hours since departure
170	0
200	0.5
190	1
200	1.25
230	2
210	2.25

Table 11.2 Data for question 4

5. A sports car accelerates from stopped to $200\,\text{km h}^{-1}$. Table 1.3 gives the times that various speeds are reached.

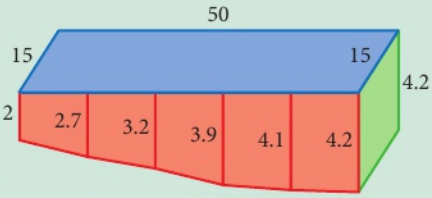
   (a) Convert the units of speed so that the trapezoidal rule gives a result in metres.

   (b) Hence, use the trapezoidal rule to find the total distance in metres travelled by the car to reach $200\,\text{km h}^{-1}$.

Speed (km h^{-1})	**Time (s)**
0	0
50	4.3
65	6.4
80	9.6
95	14.3
110	21.4
125	32.1
140	48
155	71.9
170	107.6
185	161.1
200	241.1

Table 11.3 Data for question 5

6. A rectangular swimming pool is 15 m wide and 50 m long. The diagram gives the depth at 10 m intervals.



   (a) Use the trapezoidal rule to find the area of the front wall (shaded in red)

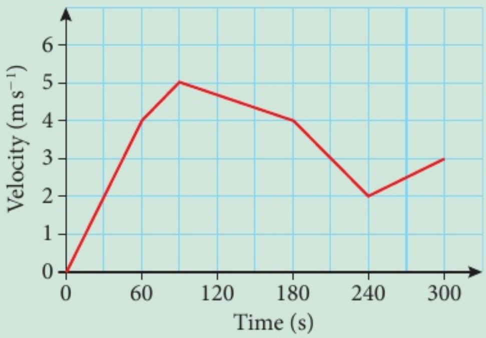
   (b) The four vertical sides of the pool are going to be repainted.

      (i) Find the total area of the four vertical sides.

      (ii) A tin of pool paint can cover $50\,\text{m}^2$. Find the number of tins of paint required.

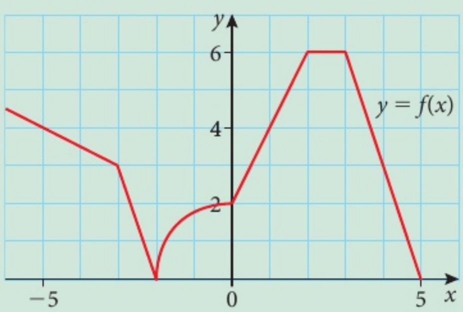
   (c) Find the total volume of the pool.

7. The velocity–time graph shows a 300-second run.



   Find the total distance travelled by the runner over the 300 seconds.

8. Use the graph of $y = f(x)$ to find each integral. The curved segment is a circular arc.



   (a) $\int_{-5}^{-3} f(x)\,dx$ **(b)** $\int_{-3}^{-2} f(x)\,dx$ **(c)** $\int_{-2}^{0} f(x)\,dx$

   (d) $\int_{0}^{5} f(x)\,dx$ **(e)** $\int_{-5}^{5} f(x)\,dx$

Chapter 11 practice questions

1. Find each anti-derivative in general form.

(a) $\int (x+8)\,\mathrm{d}x$ **(b)** $\int \left(\frac{1}{2}x^4\right)\mathrm{d}x$

(c) $\int 24\,\mathrm{d}p$ **(d)** $\int k\,\mathrm{d}x$

(e) $\int (10x^9 + 9x^8)\,\mathrm{d}x$ **(f)** $\int (-15t + v)\,\mathrm{d}t$

(g) $\int (2x^2 - 8x + 3)\,\mathrm{d}x$ **(h)** $\int (x^2+1)^2\,\mathrm{d}t$

(i) $\int (mx+b)\,\mathrm{d}x$ **(j)** $\int \left(2x^3 - 5x + \frac{2}{x^2}\right)\mathrm{d}x$

(k) $\int \left(x^3 + 2x^2 - \frac{1}{3}\right)\mathrm{d}x$ **(l)** $\int \frac{5}{n^3}\,\mathrm{d}n$

(m) $\int \left(\frac{300 - 2h}{h^3}\right)\mathrm{d}h$

2. Find the value of each definite integral.

(a) $\int_{-2}^{2} (3t^2 + 2t + 5)\,\mathrm{d}t$ **(b)** $\int_{1}^{2} e^x\,\mathrm{d}x$ **(c)** $\int_{-3}^{3} (x^3 - 9x)\,\mathrm{d}x$

(d) $\int_{-3}^{3} |x^3 - 9x|\,\mathrm{d}x$ **(e)** $\int_{e}^{e^2} \frac{1}{x}\,\mathrm{d}x$

3. Use the graph of $f(x)$ to find the given integrals. The curved segment is a circular arc.

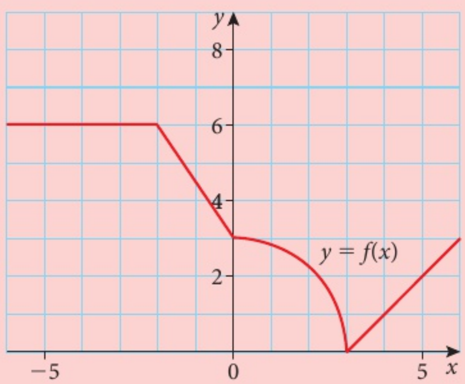
(a) $\int_{-5}^{-2} f(x)\,\mathrm{d}x$

(b) $\int_{-2}^{0} f(x)\,\mathrm{d}x$

(c) $\int_{0}^{3} f(x)\,\mathrm{d}x$

(d) $\int_{3}^{5} f(x)\,\mathrm{d}x$

(e) $\int_{-5}^{5} f(x)\,\mathrm{d}x$



4. A scale map of a pond is created from an aerial photo and measurements in metres are taken every 1.5 m. Figure 11.22 shows the measurements.

The evaporation rate of the pond can be estimated by calculating the surface area.

(a) Estimate the surface area of the pond.

(b) The evaporation rate for this pond is calculated to be $0.45\,\mathrm{l\,h^{-1}\,m^{-2}}$. Find the expected evaporation in $\mathrm{l\,h^{-1}}$.

(c) The pond has an average depth of 3 m.

(i) Calculate the volume of water in the pond in $\mathrm{m^3}$.

(ii) There are 1000 litres per $\mathrm{m^3}$. Write down the volume of the pond in litres.

(d) Assuming a uniform depth of 3 m, calculate the number of days until the pond has completely evaporated.

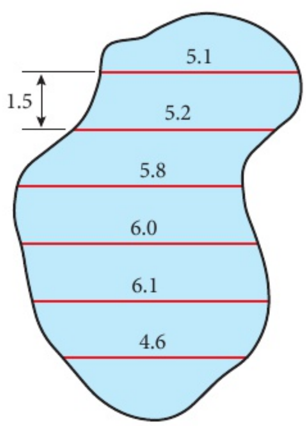


Figure 11.22 Diagram for question 4

5. Consider the quarter circle centred on the origin with radius 5 units as shown in Figure 11.23.
   (a) Calculate the area of the shaded region using 5 trapeziums of equal height.
   (b) Write down an appropriate integral for the area of the shaded region.
   (c) Hence, calculate the area of the shaded region using integration to 5 significant figures.
   (d) The area of the shaded region is A.
      (i) Calculate the value of A exactly using the formula for the area of the circle.
      (ii) Write down the value of A to 5 significant figures.
   (e) Comment on the accuracy of the methods in parts **(a)** and **(c)** in comparison to the true value in part **(d)**.

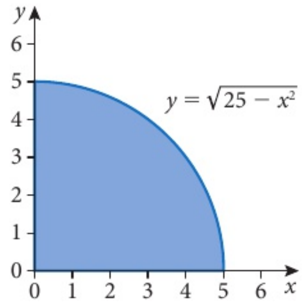


Figure 11.23 Diagram for question 5

6. Fuel tank A is filling from fuel tank B at a rate of $\frac{dV}{dt} = 3 - 0.1t$ where t is in seconds and $\frac{dV}{dt}$ is in $l\,s^{-1}$.
   (a) Write down an integral expression for the volume of fuel in fuel tank A after t seconds.
   (b) Find the volume of fuel transferred to fuel tank A from fuel tank B in the first 10 seconds of filling.
   (c) Fuel tank A has a capacity of 40 l. Find the time when fuel tank A will be full.
   (d) After fuel tank A is full, fuel tank B is then used to fill fuel tank C. When $\frac{dV}{dt} = 0$ there is no more fuel in tank B.
      (i) Find the time when this occurs.
      (ii) Hence find the number of litres in fuel tank B before filling tank A.

7. The Red Chair company sells n red chairs per month. The marginal profit in GBP can be modelled by the function $\frac{dP}{dn} = 300 - 3n$
   (a) Find the number of chairs that should be sold in order to maximise the profit.
   (b) The profit from the sale of 40 chairs is £5000. Find:
      (i) $P(n)$
      (ii) the profit from selling 80 chairs
      (iii) the least number of chairs that must be sold in order to make a profit
      (iv) the maximum profit.

Time (s)	Speed ($m\,s^{-1}$)
0	28.4
1	24
2	19.5
4	9.6
6	0

Table 11.4 Table for question 8

8. The velocity of a car braking to a stop on a wet road is recorded in Table 11.4.

(a) Use the trapezoidal rule to find the total distance travelled by the car.

Using a GDC, we can find a line of best fit for this data is

$$\frac{ds}{dt} = -4.76t + 28.7$$

(b) Write down an integral expression for the distance travelled by the car after 3 seconds.

(c) Hence, find the distance travelled by the car after 3 seconds.

(d) Find a function giving the distance travelled $s(t)$ of the car at time t.

(e) Use your function $s(t)$ to find the distance travelled by the car to come to a complete stop.

(f) The results from **(a)** and **(e)** agree to within 2 significant figures. Give a reason for this, relating to the shape of the data.

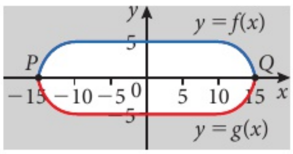


Figure 11.24 Diagram for question 9

9. A foam pillow is made by pressing liquid foam through an opening as shown in Figure 11.24. All measurements are in cm.

The equation of the upper curve is given by $f(x) = -(0.0783x)^{10} + 5$

(a) The x intercepts of $f(x)$ are at points P and Q. Find the coordinates of point P and point Q.

(b) Write down an integral expression for the area between $f(x)$ and the x-axis.

(c) Hence find the area between $f(x)$ and the x-axis.

(d) The curve $g(x)$ is symmetric to the curve $f(x)$ with respect to the x-axis. Write down the area between the curves $f(x)$ and $g(x)$.

(e) A certain pillow is 90 cm long. Find the volume of the pillow.

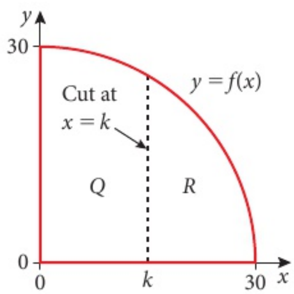


Figure 11.25 Diagram for question 10

10. One quarter of a pizza with radius 30 cm is shown in Figure 11.25.

Charles and Dominic are going to share the pizza. To divide it, they will make one cut, parallel to the y-axis. The curve of the pizza edge is given by the function $f(x) = \sqrt{30^2 - x^2}$

(a) Region Q is bounded by the y-axis, the x-axis, the line $x = k$, and the curve $y = f(x)$. Write down an integral expression for the area of region Q.

(b) Region R is bounded by the line $x = k$, the x-axis, and the curve $y = f(x)$. Write down an integral expression for the area of region R.

(c) Find the areas of regions Q and R when $k = 10$

(d) Charles and Dominic want to cut the pizza such that the regions Q and R have equal areas. Find the value of k that satisfies this requirement to 4 significant figures.

12 Probability distributions

12 Probability distributions

Learning objectives

By the end of this chapter, you should be familiar with…

- discrete random variables and their probability distributions
- expected value and the effect of linear transformations of X on its value
- the normal distribution: properties, normal probability calculations, and inverse normal calculations
- standardising normal variables (z-values)
- inverse normal calculations where mean and standard deviation are unknown
- the binomial distribution, including its mean and variance
- applications of all the above concepts.

Investing in securities, calculating premiums for insurance policies or overbooking policies used in the airline industry are only a few of the many applications of probability and statistics. Actuaries, for example, calculate the expected loss or gain that an insurance company will incur on a policy and decide on how high the premiums should be.

Chapter 8 discussed the concepts and rules of probability. This chapter extends the concept of probability to explain probability distributions. Any given statistical experiment has more than one outcome. It is impossible to predict which of the many possible outcomes will occur if an experiment is performed. Consequently, decisions are made under uncertain conditions.

> For example, a lottery player does not know in advance whether he is going to win that lottery. If the player knows that he is not going to win, he will definitely not play. It is the uncertainty about winning (some positive probability of winning) that makes him play. This chapter shows that if the outcomes and their probabilities for a statistical experiment are known, we can find out what will happen, on average, if that experiment is performed many times. For the lottery example, we can find out what a lottery player can expect to win (or lose), on average, if he continues playing this lottery again and again.

In this chapter, random variables and types of random variables are discussed. Then, the concept of a probability distribution and its mean and standard deviation for a discrete random variable are discussed. Finally, two special probability distributions for a random variable—the binomial probability distribution and the normal probability distribution—are developed.

12.1 Random variables

In Chapter 8, variables were defined as characteristics that change or vary over time and/or for different objects under consideration. A numerically valued variable X will vary or change depending on the outcome of the experiment we are performing. For example, one survey counts the number of mobile phones in each household in a certain town. Table 12.1 gives the frequency and relative frequency distributions of the number of mobile phones owned by all 4000 households in this town.

Number of phones	Frequency	Relative frequency
0	60	$\frac{60}{4000} = 0.015$
1	940	$\frac{940}{4000} = 0.235$
2	1700	$\frac{1700}{4000} = 0.425$
3	980	$\frac{980}{4000} = 0.245$
4	320	$\frac{320}{4000} = 0.080$
	$N = 4000$	**Sum = 1.000**

Table 12.1 Frequency and relative frequency distributions

Suppose one household is randomly selected from this population. The process of randomly selecting a household is called a random or chance experiment. Let X denote the number of phones owned by the selected household. Then X can assume any of the five possible values (0, 1, 2, 3, and 4) listed in the first column of Table 12.1. The value assumed by X depends on which household is selected. Thus, this value, denoted by x, depends on the outcome of a random experiment. Consequently, X is called a **random variable** or a **chance variable**.

A **random variable** is a variable whose value is determined by the outcome of a random experiment.

When a probability experiment is performed, often we are not interested in all the details of the outcomes, but rather in the value of some numerical quantity determined by the result. For instance, in rolling two dice (used in plenty of games), often we care about their sum and not the values on the individual dice. Consider this specific experiment. A sample space for which the points are equally likely is given in Table 12.2. It consists of 36 ordered pairs (a, b) where a is the number on the first die, and b is the number on the second die. For each sample point we can let the random variable X stand for the sum of the numbers. The resulting values of X are also presented in Table 12.2.

Random variables are customarily denoted by upper case letters, such as X and Y. Lower case letters are used to represent specific values of the random variable. That is, if X represents the sum of numbers resulting in the throw of a die, then $x = 2$ represents the case when the outcome is 2.

Sample point	X	f(x)	Sample point	X	f(x)	Sample point	X	f(x)
(1, 1)	$x = 2$	1	(2, 4), (1, 5), (4, 2), (5, 1), (3, 3)	$x = 6$	5	(5, 5), (4, 6), (6, 4)	$x = 10$	3
(1, 2), (2, 1)	$x = 3$	2	(2, 5), (5, 2), (1, 6), (6, 1) (3, 4), (4, 3)	$x = 7$	6	(5, 6), (6, 5)	$x = 11$	2
(1, 3), (2, 2), (3, 1)	$x = 4$	3	(2, 6), (6, 2), (3, 5), (5, 3) (4, 4)	$x = 8$	5	(6, 6)	$x = 12$	1
(1, 4), (2, 3), (3, 2), (4, 1)	$x = 5$	4	(4, 5), (5, 4), (3, 6), (6, 3)	$x = 9$	4			

Table 12.2 Sample space and the values of the random variable X in the two-dice experiment

Notice that events can be more accurately and concisely defined in terms of the random variable X; for example, the event of rolling a sum at least equal to 5 but less than 9 can be replaced by $5 \leqslant X < 9$

12 Probability distributions

There are many examples of random variables:

- X = the number of calls received by a household on a Friday night
- X = the number of free beds available at hotels in a large city
- X = the number of customers a sales person contacts on a working day
- X = the length of a metal bar produced by a certain machine
- X = the mass of newly born babies in a large hospital.

These variables are classified as **discrete** or **continuous**, according to the values that X can assume. In the examples above, the first three are discrete and the last two are continuous. The random variable is discrete if its set of possible values are isolated points on the number line, i.e. there are a countable number of possible values for the variable. The variable is continuous if its set of possible values is an entire interval on the number line, i.e. it can take any value in an interval. Consider the number of times you flip a coin until the head side appears. The possible values are $x = 1, 2, 3, \ldots$. This is a discrete variable, even though the number of times may be infinite! On the other hand, consider the time it takes a student at your school to eat lunch. This can be anywhere between zero and the length of the lunch period at your school.

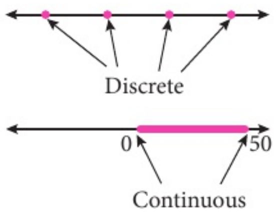


Figure 12.1 Discrete and continuous variables

Example 12.1

State whether each of the following is a discrete or a continuous random variable.

(a) The number of hairs on a Scottish terrier.

(b) The height of a building.

(c) The amount of fat in a steak.

(d) A high school student's grade on a maths test.

(e) The number of fish in the Atlantic Ocean.

(f) The temperature of an electric kettle.

Solution

(a) Even though the number of hairs is almost infinite, it is countable. So, it is a discrete random variable.

(b) This can be any real number. Even when you say this building is 15 m high, the number could be 15.1, or 15.02, etc. Hence, it is continuous.

(c) This is continuous, as the amount of fat could be zero, up to the maximum amount of fat that can be held in one piece.

(d) Grades are discrete. No matter how detailed a score the teacher gives, the grades are isolated points on a scale.

(e) This is almost infinite, but countable, hence discrete.

(f) This is continuous, as the temperature can take any value from room temperature to 100 °C.

Discrete probability distribution

In Chapter 7 you learned how to work with the frequency distribution and relative or percentage frequency distribution for a set of numerical measurements on a variable X. The distribution gave the following information about X:

- what value of X occurred
- how often each value occurred.

You also learned how to use the mean and standard deviation to measure the centre and variability of the data set.

Table 12.3 shows the relative frequency distribution of 4000 households and the number of mobile phones they own. One of the interpretations of probability is that it is understood to be the long-term relative frequency of the event.

Number of phones	Relative frequency
0	0.015
1	0.235
2	0.425
3	0.245
4	0.080

Table 12.3 Relative frequency distribution

A table like this, where we replace the relative frequency with probability, is called a **probability distribution** of the random variable.

The **probability distribution** for a discrete random variable is a table, graph, or formula that gives the possible values of x, and the probability $\mathrm{P}(x) = \mathrm{P}(X = x)$ associated with each value of x. This is also called the **probability mass function (PMF)** and in many sources it is called the **probability distribution function (PDF)**.

For every possible value x of the random variable X, the **probability mass function** specifies the probability of observing that value when the experiment is performed.

We write $\mathrm{P}(X = x)$ as $\mathrm{P}(x)$ for convenience.

Letting X be the number of phones in each household in Table 12.3, a new table showing the probability distribution of X can be completed as shown in Table 12.4.

X	0	1	2	3	4
$\mathrm{P}(x)$	0.015	0.235	0.425	0.245	0.080

Table 12.4 Probability distribution of X

The other way of representing the probability distribution is with a histogram as shown in Figure 12.2. Each column corresponds to the probability of the associated value of X. The values of X naturally represent mutually exclusive events. Summing $\mathrm{P}(x)$ over all values of X is equivalent to adding all probabilities of all simple events in the sample space, and hence the total is 1.

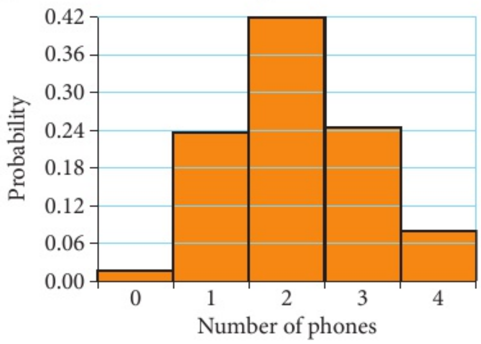


Figure 12.2 Histogram

Either the table or the histogram enable us to answer questions such as: if a household is chosen at random, what is the probability that it has at most one phone, or more than two phones?

Required properties of probability distribution functions of discrete random variables
Let X be a discrete random variable with probability distribution function, $P(x)$. Then

- $0 \leqslant P(x) \leqslant 1$ for any value x
- the individual probabilities sum to 1; that is $\sum_x P(x) = 1$ where the notation indicates summation over all possible values x.

Using notation, we can write $P(x \leqslant 1) = 0.015 + 0.235 = 0.250$ or $P(x > 2) = 0.245 + 0.080 = 0.325$

The result can be generalised for all probability distributions of discrete random variables.

Example 12.2

In the 2-dice experiment, find

(a) the probability distribution of x using a table and a graph

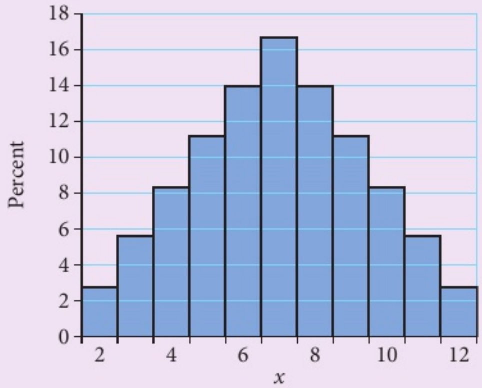
(b) $P(x \leqslant 4)$

(c) $P(x \geqslant 6)$

Solution

(a)

X	$P(x)$	X	$P(x)$	X	$P(x)$
$x = 2$	$\frac{1}{36}$	$x = 6$	$\frac{5}{36}$	$x = 10$	$\frac{3}{36}$
$x = 3$	$\frac{2}{36}$	$x = 7$	$\frac{6}{36}$	$x = 11$	$\frac{2}{36}$
$x = 4$	$\frac{3}{36}$	$x = 8$	$\frac{5}{36}$	$x = 12$	$\frac{1}{36}$
$x = 5$	$\frac{4}{36}$	$x = 9$	$\frac{4}{36}$		



(b) $P(x \leqslant 4) = \dfrac{1 + 2 + 3}{36} = \dfrac{1}{6}$

(c) $P(x \geqslant 6) = \dfrac{5 + 6 + 5 + 4 + 3 + 2 + 1}{36} = \dfrac{13}{18}$

Example 12.3

Radon is a major cause of lung cancer. It is a radioactive gas produced by the natural decay of radium in rocks that contain small amounts of uranium. Studies in areas with high levels of radon revealed that one third of houses in these areas have dangerous levels of this gas. Suppose that two houses are randomly selected, and we define the random variable X to be the number of houses with dangerous levels. Find the probability distribution of X by a table and a graph.

Solution

Since two houses are selected, then the possible values of X are 0, 1, or 2. The assumption here is that we are choosing the houses randomly and independently of each other.

$$\begin{aligned}
P(x = 2) = P(2) &= \text{P(1st house with dangerous levels and 2nd house with dangerous levels)} \\
&= \text{P(1st house with dangerous levels)} \times \text{P(2nd house with dangerous levels)} \\
&= \frac{1}{3} \times \frac{1}{3} = \frac{1}{9}
\end{aligned}$$

$$\begin{aligned}
P(x = 0) &= P(0) \\
&= \text{P(1st house without dangerous levels and 2nd house without dangerous levels)} \\
&= \text{P(1st house without dangerous levels)} \times \text{P(2nd house without dangerous levels)} \\
&= \frac{2}{3} \times \frac{2}{3} = \frac{4}{9}
\end{aligned}$$

$$P(x = 1) = 1 - [P(0) + P(2)] = 1 - \left[\frac{4}{9} + \frac{1}{9}\right] = \frac{4}{9}$$

x	0	1	2
$P(x)$	$\frac{4}{9}$	$\frac{4}{9}$	$\frac{1}{9}$

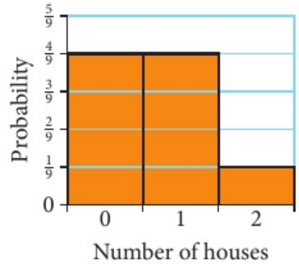


Figure 12.3 Probability is displayed as the height of a rectangle

Any type of graph can be used to give the probability distribution as long as it shows the possible values of X and the corresponding probabilities. In Example 12.3 probability is graphically displayed in Figure 12.3 as the height of a rectangle. Moreover, the rectangle corresponding to each value of X has an area equal to the probability $P(x)$. The histogram is the preferred tool due to its connection to the normal distribution discussed later in the chapter.

Formula/rule

Sometimes, the probability distribution of a random variable, X, can be given by a formula.

Don't be concerned now with how we came up with the formula as you will not be asked to use it.

For example, consider rolling a fair 6-sided dice and looking at the number of rolls it takes till the number on the top face is 3. Three can come up on the first roll, or the second, or the third and so on. That is, the possible set of values of X is $\{1, 2, 3, \ldots\}$. In order to show the frequency distribution, we can have a rule.

Recall from Chapter 8, that $P(X = 3) = \frac{1}{6}$ and $P(X \neq 3) = \frac{5}{6}$

So, to get a 3 on the second throw, it has to fail on the first throw and then succeed on the second throw. That is $P(X = 2) = \frac{5}{6} \cdot \frac{1}{6}$

To get a 3 on the 3rd throw, it has to fail on the first two throws and then succeed on the third throw. That is $P(X = 3) = \left(\frac{5}{6}\right)^2 \cdot \frac{1}{6}$

Thus, to get a 3 on the xth throw, it has to fail on $(x - 1)$ throws and then succeed on the xth throw. That is $P(X = x) = \left(\frac{5}{6}\right)^{x-1} \cdot \left(\frac{1}{6}\right)$

This is the rule or formula for this probability distribution.

We can verify this:

$P(X = 1) = \left(\frac{5}{6}\right)^{1-1} \cdot \left(\frac{1}{6}\right) = \frac{1}{6}$ as you expect, or

$P(X = 2) = \left(\frac{5}{6}\right)^{2-1} \cdot \left(\frac{1}{6}\right) = \frac{5}{6} \cdot \frac{1}{6}$

as we found earlier.

Cumulative distribution function

Cumulative distribution function (CDF)
The **cumulative distribution function** of a random variable X, also known as the cumulative probability function $F(x)$, expresses the probability that X does not exceed the value x, as a function of x. That is

$$F(x) = P(X \leqslant x) = \sum_{y: y \leqslant x} P(y)$$

The notation here indicates that summation is over all possible values of y that are less than or equal to x. The choice of the variable name to be y is arbitrary – you can use any letter.

x	$F(x)$
0	0.015
1	0.25
2	0.675
3	0.92
4	1.00

Table 12.5 CDF

For some value x of the random variable X, we often wish to compute the probability that the observed value of X is at most x. This gives rise to the **cumulative distribution function** (CDF).

For example, for the distribution in Table 12.4, the CDF is shown in Table 12.5.

So, $F(3) = 0.92$, stands for the probability of households that own up to 3 mobile phones. This result of course can be achieved by adding the probabilities corresponding to $x = 0, 1, 2$, and 3.

In many cases, as we will see later, we use the cumulative distribution to find individual probabilities,

$P(X = x) = P(X \leqslant x) - P(X < x)$.

For example, to find the probability that $x = 3$, we can use the cumulative distribution table.

$P(x = 3) = P(x \leqslant 3) - P(x < 3) = 0.92 - 0.675 = 0.245$

This property is of great value when studying the binomial distribution.

For example, many universities have the policy of posting the grade distributions for their courses. Several universities have a grade-point average that codes the grades as follows:

A = 4, B = 3, C = 2, D = 1, and F = 0

During the spring term at a certain large university 13% of the students in an introductory statistics course received As, 37% Bs, 45% Cs, 4% Ds and 1% Fs. The experiment here is to choose a student at random and mark down the grade. The student's grade on the 4-point scale is a random variable X.

Here is the probability distribution of X:

x	0	1	2	3	4
$P(x)$	0.01	0.04	0.45	0.37	0.13

Table 12.6 Probability distribution of X

Is this a probability distribution?

Yes, it is. Each probability is between 0 and 1, and the sum of all probabilities is 1.

What is the probability that a randomly chosen student receives a B or better?

$P(x \geqslant 3) = P(x = 3) + P(x = 4) = 0.37 + 0.13 = 0.50$

Example 12.4

The probability distribution for the first digit people choose for the codes for their mobile phones is shown in the table.

First digit	0	1	2	3	4	5	6	7	8	9
Probability	0.009	0.300	0.174	0.122	0.096	0.078	0.067	0.058	0.051	0.045

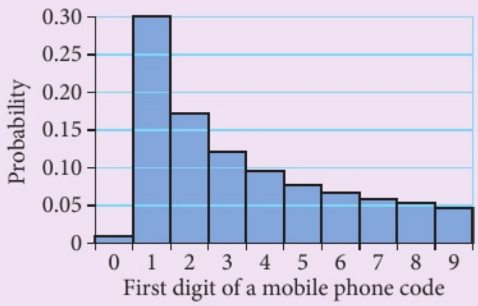
(a) What is the probability that you pick a first digit and it is more than 5?

(b) Show a probability histogram for the distribution.

Solution

(a) $P(x > 5) = P(x = 6) + P(x = 7) + P(x = 8) + P(x = 9) = 0.221$

(b)



Note that the height of each bar shows the probability of the outcome at its base. The heights add up to 1, of course. The bars in these histograms have the same width, namely 1. So, the areas also display the probability assignments of the outcomes. Think of such histograms (probability histograms) as idealised pictures of the results of very many repeated trials

Expected values

The probability distribution for a random variable looks very similar to the relative frequency distribution discussed in Chapter 7. The difference is that the relative frequency distribution describes a sample of measurements, whereas the probability distribution is constructed as a model for the entire population. Just as the mean and standard deviation gave you measures for the centre and spread of the sample data, you can calculate similar measures to describe the centre and spread of the population.

The population mean, which measures the average value of X in the population, is also called the **expected value** of the random variable X. It is the value that you would expect to observe on average if you repeated the experiment an infinite number of times. The formula we use to determine the expected value can be more easily understood with the help of an example.

Revisiting the household's mobile phones example
Let x be the number of phones owned. Here is the table of probabilities again

X	0	1	2	3	4
$\mathrm{P}(x)$	0.015	0.235	0.425	0.245	0.080

Suppose we choose a large number of households, say 100 000. Intuitively, using the relative frequency concept of probability you would expect to observe 1500 households with no mobile phone, 23 500 with one phone, and so on: 42 500, 24 500, and 8000.
The average (mean) value of X as defined in Chapter 7 would then be equal to:

$$\frac{\text{Sum of all measurements}}{n} = \frac{0 \times 1500 + 1 \times 23500 + 2 \times 42500 + 3 \times 24500 + 4 \times 8000}{100000}$$
$$= \frac{0 \times 1500}{100000} + \frac{1 \times 23500}{100000} + \frac{2 \times 42500}{100000} + \frac{3 \times 24500}{100000} + \frac{4 \times 8000}{100000}$$
$$= 0 \times 0.025 + 1 \times 0.235 + 2 \times 0.425 + 3 \times 0.245 + 4 \times 0.080$$
$$= 0 \times \mathrm{P}(0) + 1 \times \mathrm{P}(1) + 2 \times \mathrm{P}(2) + 3 \times \mathrm{P}(3) + 4 \times \mathrm{P}(4) = 2.14$$

That is, we expect to see households, on average, possessing 2.14 mobile phones! This does not mean that we know what a household will own, but we can say what we expect to happen.

Let X be a discrete random variable with probability distribution $\mathrm{P}(x)$. The mean or **expected value** of X is given by
$\mu = \mathrm{E}(X) = \sum x \cdot \mathrm{P}(x)$

Insurance companies make extensive use of expected value calculations. Here is a simplified example.

An insurance company offers a policy that pays you €10 000 when your car is damaged beyond repair or €5000 for major damages (50%). They charge you €50 per year for this service. Can they make a profit?

Suppose that in any year, 1 out of every 1000 cars is damaged beyond repair, and that another 2 out of 1000 will have serious damages. Then we can display the probability model for this policy in a table like this.

Type of accident	Amount paid x (€)	Probability $\mathrm{P}(X = x)$
Total damage	10 000	$\frac{1}{1000}$
Major damage	5000	$\frac{2}{1000}$
Minor or no damage	0	$\frac{997}{1000}$

Table 12.7 Probability model for a car insurance policy

The expected amount the insurance company pays is given by

$$\mu = \mathrm{E}(X) = \sum x\mathrm{P}(x) = €10000\left(\frac{1}{1000}\right) + €5000\left(\frac{2}{1000}\right) + €0\left(\frac{997}{1000}\right) = €20$$

This means that the insurance company expects to pay, on average, an amount of €20 per insured car. Since it is charging people €50 for the policy, the company expects to make a profit of €30 per car. Thinking about the problem in a different perspective, suppose they insure 1000 cars, then the company would expect to pay €10 000 for 1 car, and €5000 to each of two cars with major damage.

This is a total of €20 000 for all cars, or an average of $\frac{20000}{1000} = €20$ per car.

Of course, this expected value is not what actually happens to any particular policy. No individual policy actually costs the insurance company €20. We are dealing with random events, so a few car owners may require a payment of €10 000 or €5000, many others receive nothing. Because of the need to anticipate such variability, the insurance company needs to know a measure of this variability, which is the **standard deviation**.

Variance and standard deviation

For data in Chapter 7, we calculated the variance by computing the deviation from the mean, $x - \mu$, and then squaring it. We do that with random variables as well.

We can use similar arguments to justify the formulae for the population variance σ^2 and consequently the population standard deviation σ. These measures describe the spread of the values of the random variable around the centre. We similarly use the idea of the 'average' or 'expected' value of the squared deviations of the x-values from the mean μ or $E(X)$.

Let X be a discrete random variable with probability distribution $P(x)$ and mean μ.
The **variance** of X is given by

$$\begin{aligned}\sigma^2 &= E\left((X-\mu)^2\right)\\ &= \sum (x-\mu)^2 \cdot P(x)\end{aligned}$$

This is sometimes called $\text{Var}(X)$, or $V(X)$.

The standard deviation σ of a random variable x is equal to the positive square root of its variance.

It can also be shown that there is another formula for the variance:

$$\sigma^2 = \sum (x-\mu)^2 \cdot P(x) = \sum x^2 \cdot P(x) - (E(X))^2 = \sum x^2 \cdot P(x) - \left(\sum xP(x)\right)^2$$

This formula is useful in your explorations if you need to set up a spreadsheet to do the calculations.

Let us go back to the mobile phones example. We calculated the expected mean value to be 2.14 phones. To calculate the variance, we can tabulate our work to make the manual calculation easier.

x	$P(x)$	Deviation $(x-\mu)$	Squared deviation $(x-\mu)^2$	$(x-\mu)^2 \cdot P(x)$
0	0.015	−2.14	4.58	0.068694
1	0.235	−1.14	1.30	0.305406
2	0.425	−0.14	0.02	0.00833
3	0.245	0.86	0.74	0.181202
4	0.080	1.86	3.46	0.276768
		Total	$\sum (x-\mu)^2 \cdot P(x)$	**0.8404**

Table 12.8 Calculating the variance

So, the variance of the number of mobile phones per household is 0.8404 phones^2, or the standard deviation is 0.9167 phones.

The unit of the variance is the square of the unit of measurement of the random variable, which does not make much sense. Thus we calculate the standard deviation.

GDC notes

You can do these calculations using your GDC. Depending on which GDC you are using, some may require that you store your data in lists and perform the calculations as described by the formulas above, and some may give you the results after you enter your data in lists making sure that the probability is given as a frequency. In discrete random variable calculations take the σx values and not the sx values. Here is a sample of a GDC output.

```
1-Variable
x̄     =2.14
Σx    =2.14
Σx²   =5.42
σx    =0.91673333
```

You can also do the calculation using a spreadsheet.

x	$P(X = x)$
0	0.08
1	0.40
2	0.24
3	0.15
4	0.08
5	0.05

Table 12.9 Data for Example 12.5

Example 12.5

A computer store sells a specific type of laptop. The number of laptops sold each day is shown in Table 12.9. x is the potential number of laptops sold each day. The store would like to ensure that they have enough stock for all potential sales of laptops.

Calculate the expected value of the demand and the standard deviation.

Solution

$$E(X) = \sum xP(x) = 0 \times 0.08 + 1 \times 0.40 + 2 \times 0.24 + 3 \times 0.15 + 4 \times 0.08 + 5 \times 0.05 = 1.90$$

$$\begin{aligned}\text{Var}(X) = \sigma^2 &= \sum (x - \mu)^2 P(x) \\ &= (0 - 1.9)^2 \cdot 0.08 + (1 - 1.9)^2 \cdot 0.40 + (2 - 1.9)^2 \cdot 0.24 + \cdots \\ &\quad + (5 - 1.9)^2 \cdot 0.05 = 1.63\end{aligned}$$

Thus, $\sigma = 1.28$

A spreadsheet output and a sample of a GDC output are shown.

x	$P(x)$	$xP(x)$	$x - \mu$	$(x - \mu)^2$	$(x - \mu)^2P(x)$
0	0.08	0	−1.9	3.61	0.2888
1	0.4	0.4	−0.9	0.81	0.324
⋮	⋮	⋮	⋮	⋮	⋮
3	0.05	0.25	3.1	9.61	0.4805
Totals	**1**	**1.9**			**1.63**

Here is an example of a GDC output.

```
1-Variable
x̄     =1.9
Σx    =1.9
Σx²   =5.24
σx    =1.27671453
```

For the variance, we follow the same procedure as described earlier.

Example 12.6

The Statistical Abstract of countries is usually published annually. One of the questions asks the country's households to report the number of persons living in the household. Table 12.10 summarises the data for a certain country.

(a) Develop the probability distribution of the random variable defined as the number of persons per household.

(b) Find the probability that a randomly chosen household has 4 or more persons.

(c) Calculate the expected value and standard deviation. Consider the last class to be 7.

Number of persons	Number of households (Millions)
1	31.1
2	38.6
3	18.8
4	16.2
5	7.2
6	2.7
7 or more	1.4
Total	**116.0**

Table 12.10 Data for Example 12.6

Solution

(a) The probability of each value of X is computed as the relative frequency. Divide each frequency by 116, producing the probability distribution shown in Table 12.11.

(b) $P(X \geqslant 4) = P(4) + P(5) + P(6) + P(7 \text{ or more})$
$\quad = 0.140 + 0.062 + 0.023 + 0.012 = 0.237$

(c) We use a GDC. The expected value is 2.512 and the standard deviation is 1.398.

X	$P(x)$
1	0.268
2	0.333
3	0.162
4	0.140
5	0.062
6	0.023
7	0.012
Total	1.000

Table 12.11 Solution for Example 12.6 (a)

```
1-Variable
x̄    =2.512
Σx   =2512
Σx²  =8264
σx   =1.39780399
sx   =1.39850341
n    =1000
```

Figure 12.4 Solution for Example 12.6 (c)

The following example will introduce you to the ideas discussed in the next section.

Example 12.7

What is the probability distribution of the discrete random variable X that counts the number of heads in four flips of a coin?

Solution

We can derive this distribution if we make two reasonable assumptions:

- The coin is balanced, so it is fair and each flip is equally likely to give H or T.
- The coin has no memory, so flips are independent.

The outcome of four flips is a sequence of heads and tails such as HTTH. There are 16 possible outcomes in all. The outcomes are listed here along with the value of X for each outcome.

		HTTH		
		HTHT		
	HTTT	THTH	HHHT	
	THTT	HHTT	HHTH	
	TTHT	THHT	HTHH	
TTTT	TTTH	TTHH	THHH	HHHH
$X = 0$	$X = 1$	$X = 2$	$X = 3$	$X = 4$

x	P(x)
0	0.0625
1	0.25
2	0.375
3	0.25
4	0.0625

Table 12.12 Solution for Example 12.7

So, for example, $\mathrm{P}(X = 2) = \dfrac{\text{Number of ways } X = 2 \text{ occurs}}{16} = \dfrac{6}{16} = 0.375$

We can find the probability of each value of X from the outcomes we listed in the same way. Table 12.12 shows the results.

However, the multiplication rule for independent events tells us that, for example,

$\mathrm{P(HTTH)} = \mathrm{P(H)} \cdot \mathrm{P(T)} \cdot \mathrm{P(T)} \cdot \mathrm{P(H)} = \dfrac{1}{2} \cdot \dfrac{1}{2} \cdot \dfrac{1}{2} \cdot \dfrac{1}{2} = \dfrac{1}{16} = 0.0625$

This is one case of $X = 2$, but there are 6 such cases, so

$\mathrm{P}(X = 2) = 6 \cdot \dfrac{1}{16} = 0.375$

Exercise 12.1

1. Classify each of the following as discrete or continuous random variables.
   (a) The number of words spelled correctly by a student on a spelling test.
   (b) The amount of water flowing through the Niagara Falls per year.
   (c) The length of time by which a student is late to class.
   (d) The number of bacteria per ml of drinking water in Geneva.
   (e) The amount of carbon dioxide produced per litre of unleaded fuel.
   (f) The amount of a flu vaccine in a syringe.
   (g) The heart rate of a lab mouse.
   (h) The barometric pressure at the top of Mount Everest.
   (i) The distance travelled by a taxi driver per day.
   (j) Total score of football teams in national leagues.
   (k) Height of ocean tides on the shores of Portugal.
   (l) Tensile breaking strength (in newtons per square metre) of a 5 cm diameter steel cable.
   (m) Number of overdue books in a public library.

2. The amount of money students earn on their summer jobs is a random variable.
   (a) What are the possible values of this random variable?
   (b) Are the values countable? Explain.
   (c) Is there a finite number of values? Explain.
   (d) Is the random variable continuous? Explain.

3. The mark on a maths test of 100 multiple choice questions is a random variable.
   (a) What are the possible values of this random variable?
   (b) Are the values countable? Explain.
   (c) Is there a finite number of values? Explain.
   (d) Is the random variable continuous? Explain.

4. An internet pharmacy advertises that it will deliver products that customers purchase in 3 to 6 days. The manager of the company wanted to be more precise in its advertising and recorded the number of days it took to deliver to a large number of customers. The table shows the probability distribution.

No. of days	0	1	2	3	4	5	6	7	8
Probability	0	0	0.01	0.04	0.28	0.42	0.21	0.02	0.02

What is the probability that a delivery will be:

(a) made within the advertised period

(b) late

(c) early?

5. The manager of a bookstore recorded the number of customers who arrive at checkout counters every 2 minutes during late afternoon hours and set up a probability distribution, shown in Table 12.13.

(a) What is the probability that more than 2 customers arrive within 2 minutes?

(b) Calculate the mean and standard deviation of this random variable.

No. of customers	Probability
0	0.10
1	0.20
2	0.25
3	0.25
4	0.20

Table 12.13 Data for question 5

6. A random variable y has the probability distribution shown in Table 12.14.

(a) Find P(2).

(b) Construct a probability histogram for this distribution.

(c) Find μ and σ.

(d) Locate the interval $\mu \pm \sigma$ as well as $\mu \pm 2\sigma$ on the histogram.

(e) We create another random variable $Z = y + 1$. Find μ and σ of Z.

(f) Compare your results for **(c)** and **(e)** and generalise for $Z = Y + b$, where b is a constant.

y	$P(y)$
0	0.1
1	0.3
2	
3	0.1
4	0.05
5	0.05

Table 12.14 Data for question 6

7. A discrete random variable x can assume five possible values: 12, 13, 15, 18, and 20. Its probability distribution is shown in Table 12.15.

(a) What is P(15)?

(b) What is the probability that x equals 12 or 20?

(c) What is $P(x \leqslant 18)$?

(d) Find $E(x)$.

(e) Find $V(x)$.

(f) Let $Y = 0.5X - 4$. Find $E(Y)$ and $V(Y)$.

(g) Compare your results in **(d)**, **(e)** and **(f)** and generalise for $Y = aX + b$, where a and b are constants.

x	$P(x)$
12	0.14
13	0.11
15	
18	0.26
20	0.23

Table 12.15 Data for question 7

x	$P(x)$
0	0.002
1	0.029
2	0.132
3	0.309
4	0.360
5	0.168

Table 12.16 Data for question 8

8. Medical research has shown that a certain type of chemotherapy is successful 70% of the time when used to treat skin cancer. In a study to check the validity of such a claim, researchers chose different treatment centres and chose five of their patients at random. Table 12.16 shows the probability distribution of the number of successful treatments for groups of five.

(a) Find the probability that at least two patients would benefit from the treatment.

(b) Find the probability that the majority of the group does not benefit from the treatment.

(c) Find $E(x)$ and interpret the result.

(d) Show that $\sigma(X) = 1.02$.

9. The probability function of a discrete random variable X is given by $P(X = x) = \frac{kx}{2}$ for $x = 12, 14, 16, 18$

Set up the table showing the probability distribution and find the value of k.

X	$P(x)$
5	$\frac{3}{20}$
10	$\frac{7}{30}$
15	k
20	$\frac{3}{10}$
25	$\frac{13}{60}$

Table 12.17 Data for question 10

10. X has probability distribution as shown in Table 12.17.

(a) Find the value of k

(b) Find $P(x > 10)$

(c) Find $P(5 < x \leqslant 20)$

(d) Find the expected value and the standard deviation.

(e) Let $Y = \frac{1}{5}x - 1$. Find $E(Y)$ and $V(Y)$.

11. The discrete random variable x has probability function given by

$$P(x) = \begin{cases} \left(\frac{1}{4}\right)^{x-1} & x = 2, 3, 4, 5, 6 \\ k & x = 7 \\ 0 & \text{otherwise} \end{cases}$$

where k is a constant. Determine the value of k and the expected value of X.

y	$P(Y = y)$
0	0.1
1	0.11
2	k
3	$(k - 1)^2$

Table 12.18 Data for question 12

12. Table 12.18 shows a probability distribution for a random variable y.

(a) Find the value of k.

(b) Find the expected value.

13. A closed box contains eight red balls and four white balls. A ball is taken out at random, its colour noted, and then it is returned. This is done three times. Let X represent the number of red balls drawn.

(a) Set up a table to show the probability distribution of X.

(b) What is the expected number of red balls in this experiment?

14. Airlines sometimes overbook flights. Suppose for a 50-seat plane, 55 tickets were sold. Let X be the number of ticketed passengers that show up for the flight. Table 12.19 shows the PMF for this flight, taken from the airline's records.
    (a) Construct a CDF table for this distribution.
    (b) What is the probability that the flight will accommodate all ticketed passengers that show up?
    (c) What is the probability that not all ticketed passengers will have a seat on the flight?
    (d) Calculate the expected number of passengers who will show up.
    (e) Calculate the standard deviation of the passengers who will show up.

15. A small internet provider has 6 telephone service lines operating 24 hours daily. X is defined as the number of lines in use at any specific 10-minute period of the day. The PMF of X is given in Table 12.20.
    (a) Construct a CDF table.
    (b) Calculate the probability that at most three lines are in use.
    (c) Calculate the probability that a customer calling for service will have a free line.
    (d) Calculate the expected number of lines in use.
    (e) Calculate the standard deviation of the number of lines in use.

16. Some flashlights use one AA-type battery to work. The voltage in any new battery is considered acceptable if it is at least 1.3 volts. 90% of the AA batteries from a specific supplier have an acceptable voltage. Batteries are usually tested until an acceptable one is found. It is then installed in the flashlight. Let X be the number of batteries that must be tested.
    (a) What is P(1), i.e., $P(X = 1)$?
    (b) What is P(2)?
    (c) What is P(3)?
    (d) To have $X = 5$, what must be true of the
        (i) fourth battery tested
        (ii) fifth battery tested?
    (e) Use your observations above to obtain a general model for $P(x)$.

17. A biased dice with four faces is used in a game. A player pays 10 counters to roll the dice. Table 12.21 shows the possible scores on the dice, the probability of each score and the number of counters the player receives in return for each score.
    Find the value of n for the player to get an expected return of 9 counters per roll.

x	$P(x)$
45	0.05
46	0.08
47	0.12
48	0.15
49	0.25
50	0.20
51	0.05
52	0.04
53	0.03
54	0.02
55	0.01

Table 12.19 Data for question 14

x	$P(x)$
0	0.08
1	0.15
2	0.22
3	0.27
4	0.20
5	0.05
6	0.03

Table 12.20 Data for question 15

Score	Probability	Number of counters player receives
1	$\frac{1}{2}$	4
2	$\frac{1}{5}$	5
3	$\frac{1}{5}$	15
4	$\frac{1}{10}$	n

Table 12.21 Data for question 17

12.2 The binomial distribution

Examples of discrete random variables are abundant in everyday situations. However, there are a few discrete probability distributions that are widely applied and serve as models for a great number of the applications. One of them is the **binomial distribution**.

Example 12.8

A cereal company puts miniature figures in boxes of cornflakes to make them attractive for children and thus boost sales. The manufacturer claims that 20% of the boxes contain a figure. You buy three boxes of this cereal.

What is the probability that you will get:

(a) exactly three figures

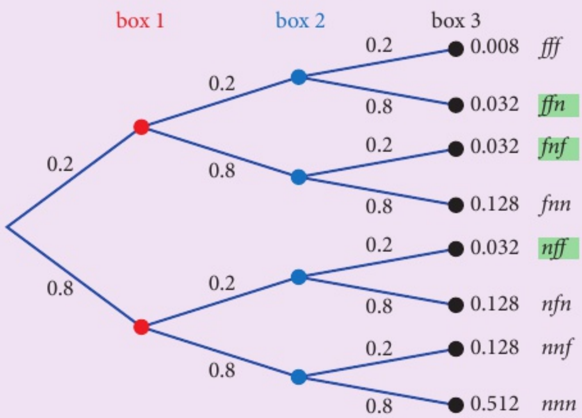
(b) exactly two figures

(c) exactly two figures in five boxes?

Solution

(a) To get three figures means that the first box contains a figure (0.20 chance), as does the second (also 0.20), and the third (0.20). You want three figures; therefore, this is the intersection of three events and the probability is $0.20^3 = 0.008$

(b) To get exactly two figures, the situation becomes more complicated. A tree diagram can help you visualise it better.



Let f stand for figure, and n for no figure. There are three events of interest to us. Since we are interested in two figures, we want to see *ffn*, which has a probability of

$0.2 \times 0.2 \times 0.8 = 0.2^2 \times 0.8 = 0.032$

The other events of interest are *fnf* and *nff* with probabilities

$0.2 \times 0.8 \times 0.2 = 0.032$ and $0.8 \times 0.2 \times 0.2 = 0.032$

Since the order of multiplication is not important, you see that the three probabilities are the same. These three events are disjoint, as can be seen from the tree diagram, and hence the probability of exactly two figures is the sum of the three numbers, $0.032 + 0.032 + 0.032 = 0.096$. Of course, you may realise by now that it would be much simpler if you wrote $3(0.032)$, since there are three events with the same probability.

(c) The situation is similar, of course. However, a tree diagram would not be useful in this case, as there is too much information to construct a sensible diagram. No matter how you succeed in finding a figure, whether it is in the first box, the second or the third, it has the same probability, 0.2. So, to have two successes (finding figures) in the five boxes, you need the other three to be failures (no figures) with a probability of 0.8 for each failure. Therefore, the chance of having a case like *ffnnn* is $0.2^2 \times 0.8^3$. However, this can happen in several mutually exclusive ways. There are 10 ways:

ffnnn, fnfnn, fnnfn, fnnnf, nffnn, nnffn, nnnff, nfnfn, nnfnf, nfnnf

The probability of having exactly two figures in five boxes is

$10 \times 0.2^2 \times 0.8^3 = 0.2048$

The number 10 is known as the binomial coefficient, which comes from Pascal's triangle . This is also the combination of two events out of five. You can use a GDC to find this number.

You can find experiments like this one in many situations. Flipping a coin is a simple example.

Recall the experiment with flipping four coins and counting the number of heads in Example 12.7.

The event of interest is the number of heads showing. Every coin has a probability of 0.50 to show heads and 0.50 to show tails. Take the case of $X = 1$. This means only one coin shows heads. The other three show tails. As you saw in Example 12.7, $X = 1$ corresponds to HTTT, THTT, TTHT, and TTTH.

These simple events have probabilities

$0.5 \times 0.5 \times 0.5 \times 0.5$; $0.5 \times 0.5 \times 0.5 \times 0.5$;

$0.5 \times 0.5 \times 0.5 \times 0.5$ and $0.5 \times 0.5 \times 0.5 \times 0.5$

As multiplication is commutative, then each can be written as 0.5×0.5^3

Thus $\mathrm{P}(X = 1) = 4 \times 0.5 \times 0.5^3$

Another very common example is opinion polls that are conducted before elections and used to predict voter preferences. Each sampled person can be compared to a coin – but a biased coin! A voter you sample in favour of your candidate can correspond to either a head or a tail on a coin. Such experiments all exhibit the typical characteristics of the binomial experiment.

A **binomial experiment** has the characteristics:

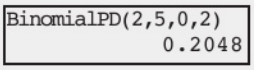
- The experiment consists of a fixed number of identical trials. We represent the number of trials by n.
- Each trial has one of two outcomes. We call one of them success, S, and the other failure, F.
- The probability of success on a single trial, p, is constant throughout the whole experiment. The probability of failure is $1 - p$ which is sometimes denoted by q. That is $p + q = 1$.
- The trials are independent, which means that the outcome of one trial does not affect the outcomes of any other trials.
  The random variable, X, of a binomial experiment is defined as the number of successes possible in the n trials. That is $X = 0, 1, 2, \ldots, n$.

In Example 12.8, we started with $n = 3$, $p = 0.2$, and asked for the probability of two successes, i.e., $x = 2$. In the second part we have $n = 5$.

Imagine repeating a binomial experiment n times. If the probability of success is p, then the probability of having x successes is $pppp\ldots$, x times (or p^x), because the order is not important, as we saw before. However, to have exactly x successes, the other $n - x$ trials must be failures, that is, with probability of $qqqq\ldots$, $(n - x)$ times; q^{n-x}

This is only one order (combination) where the successes happen the first x times and the rest are failures. We have to count the number of orders (combinations) possible. This is given by the binomial coefficient ${}^nC_x$.

You can use your GDC to do calculations with this formula.



Suppose that a random experiment can result in two possible mutually exclusive and collectively exhaustive outcomes, success and failure, and that p is the probability of a success in a single trial. When n independent trials are carried out, the distribution of the number of successes x is called the **binomial distribution**. Its probability distribution function for the binomial random variable x is:

P(x successes in n independent trials) $= P(x) = {}^nC_x p^x (1 - p)^{n-x} = {}^nC_x p^x q^{n-x}$, for $x = 0, 1, 2, \ldots n$.

(You GDC is capable of giving answers to such calculations and thus, there is no need to remember this formula.)

The notation used to indicate that a variable has a binomial probability distribution with n trials and probability of success p is: $X \sim B(n, p)$

Example 12.9

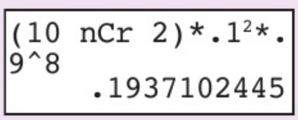
A computer shop orders its laptops from a supplier, which has a rate of defective items of 10%. The shop usually takes a sample of 10 laptops and checks them for defects. If they find two laptops are defective, they return the shipment. What is the probability that their random sample will contain two defective computers?

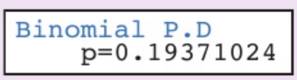
Solution

We will consider this to be a random sample and the shipment large enough to render the trials independent of each other. The probability of finding 2 defective computers in a sample of 10 is given by

$P(x = 2) = {}^{10}C_2 0.1^2 0.9^{10-2} = 45 \times 0.01 \times 0.43047 = 0.194$

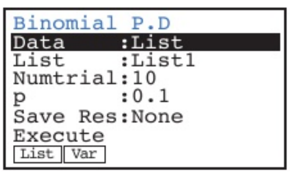
Of course, it is a daunting task to do all the calculations by hand. A GDC can do this calculation for you. You need to learn how your GDC performs such calculations. Samples from two GDCs are shown.





Using a spreadsheet, you can also produce this result or even a set of probabilities covering all the possible values.

Similarly, a GDC can also give you a list of the probabilities , as in Figure 12.5.



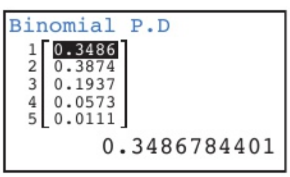


Figure 12.5 List of probabilities

Like other distributions, when you look at the binomial distribution, you want to look at its expected value and standard deviation.

Using the formula we developed for the expected value, $\sum xP(x)$, we can of course add $xP(x)$ for all the values involved in the experiment. The process would be long and tedious for something we can intuitively know. For example, in Example 12.9, if we know that the rate of defective laptops is 10%, then it is natural to expect to have $10 \times 0.1 = 1$ defective laptop.

If we have 100 computers with a defect rate of 10%, how many would you expect to be defective? Can you think of a reason why it would not be 10?

The expected value of the successes in the binomial is actually nothing but the number of trials n multiplied by the probability of success, i.e., np.

The binomial probability model
n = number of trials
p = probability of success, $q = 1 - p$ = probability of failure
x = number of successes in n trials
$P(x) = {}^nC_x p^x (1-p)^{n-x} = {}^nC_x p^x q^{n-x}$, for $x = 0, 1, 2, \dots, n$
Expected value $= \mu = np$
Variance $= \sigma^2 = npq$, $\sigma = \sqrt{npq}$

So, in the defective laptops case, the expected number of defective items in the sample of 10 is $np = 10 \times 0.1 = 1$

And the standard deviation is $\sigma = \sqrt{npq} = \sqrt{10 \times 0.1 \times 0.9} = 0.949$

Example 12.10

A study to examine the effectiveness of advertising on the internet reported that 4 out of 10 surfers remember advertisement banners after they have seen them.

(a) 20 surfers are chosen at random and shown an advertisement. What is the expected number of surfers that would remember the advertisement?

(b) What is the chance that 5 of those 20 will remember the advertisement?

(c) What is the probability that at most 1 surfer would remember the advertisement?

(d) What is the chance that at least 2 surfers would remember the advertisement?

Solution

(a) $X \sim B(20, 0.4)$. The expected number is $20 \times 0.4 = 8$
We expect 8 of the surfers to remember the advertisement. Notice on the histogram that the area of the tallest bar, in red, corresponds to the expected value 8.

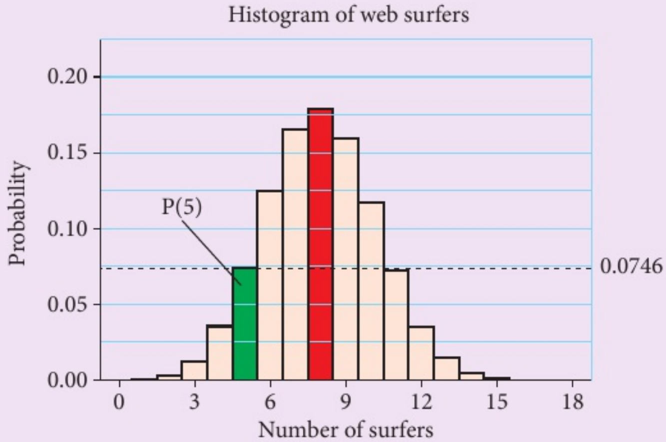
(b) $P(5) = {}^{20}C_5 0.4^5 (0.6)^{15} = 0.0746$
Or see the output from a GDC. This area is shown in the histogram as the green area.

```
Binomial P.D
     p=0.07464701
```

Figure 12.6 GDC output

(c) $P(x \leqslant 1) = P(x = 0) + P(x = 1) = 0.000524$

(d) $P(x \geqslant 2) = 1 - P(x \leqslant 1) = 1 - 0.000524 = 0.999475$



The cumulative binomial distribution function

The cumulative distribution function $F(x)$ of a random variable X expresses the probability that X does not exceed the value x. That is

$$F(x) = P(X \leqslant x) = \sum_{y:y \leqslant x} P(y)$$

So, for the binomial distribution, the cumulative distribution function (CDF) is given by

$$F(x) = P(X \leqslant x) = \sum_{y:y \leqslant x} P(y) = \sum_{y:y \leqslant x} {}^nC_y p^y q^{n-y}$$

The cumulative distribution is very helpful when we need to find the probability that a binomial variable assumes values over a certain interval.

You will not be required to perform calculations manually. A GDC can produce the values requested.

Example 12.11

In a large shipment of light bulbs, 4% of the bulbs are defective. In a sample of 20 randomly selected bulbs from the shipment, what is the probability that:

(a) there are at most three defective bulbs

(b) there are at least six defective bulbs.

Solution

(a) This can be considered as a binomial distribution with $n = 20$ and $p = 0.04$

We need $P(x \leqslant 3)$, which we can calculate either by finding the probabilities for $x = 0, 1, 2$, and 3, and then adding them, or by using the cumulative distribution function. In both cases, we will use a GDC.

As you can see in Figure 12.7, using the CDF is a much more straight forward procedure. $P(x \leqslant 3) = 0.993$

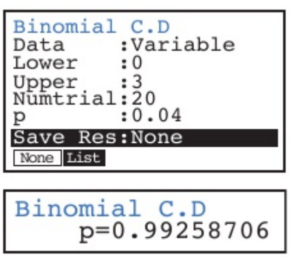


Figure 12.7 GDC solution to Example 12.11 (a)

(b) Here we need $P(X \geqslant 6)$. The first approach is not feasible at all as we would need to calculate 15 individual probabilities and add them. However, setting the problem as a complement and then using the cumulative distribution is much more efficient.

$P(X \geqslant 6) = 1 - P(X < 6) = 1 - P(X \leqslant 5) = 0.000098$

```
1-binomcdf(20,0.04,5)
         9.765401703E-5
```

OR

```
Binomial C.D
Data     :Variable
Lower    :6
Upper    :20
Numtrial:20
p        :0.04
Save Res:None
```

```
Binomial C.D
     p=9.7654E-05
```

Figure 12.8 GDC solution to Example 12.11 (b)

Example 12.12

Jim is a student taking a statistics course. Unfortunately, Jim is not a 'wise' student. He does not read the textbook before and after class, does not do homework, and regularly misses class. Jim is about to take a quiz. He decides to rely on mere luck to pass this quiz. The quiz consists of 10 multiple choice questions. Each question has four possible answers, only one of which is correct. He plans to guess the answer to each question.

(a) What is the probability that Jim gets all answers wrong?

(b) What is the probability that he gets 3 answers correct?

(c) What is the probability that he gets at most 3 answers correct?

(d) To pass the quiz, Jim must have at least 6 answers correct. What is the probability that he passes the quiz?

Solution

The situation can be modelled by a binomial distribution with success being guessing an answer correctly with a probability $p = \frac{1}{4} = 0.25$

Hence, the probability of getting it wrong is 0.75.

(a) To get all answers wrong is therefore $0.75^{10} = 0.0563$. There is no need to resort to the binomial commands. However, you will receive the same answer if you use a GDC.

(b) This is equivalent to asking for 3 successes out of 10 trials, $P(X = 3) = 0.250$

(c) This is a cumulative probability up to 3 successes, $P(X \leqslant 3) = 0.776$

(d) This is another cumulative probability with lower limit of 6.

We can get this probability by considering the event to be the complement of 'at most 5', $P(X \geqslant 6) = 1 - P(X \leqslant 5) = 0.0197$

```
BinomialPD(0,10,0.25
          0.05631351471
```

```
BinomialPD(3,10,0.25
           0.2502822876
```

```
BinomialCD(3,10,0.25
           0.7758750916
```

```
1-BinomialCD(5,10,0.
          0.01972770691
BinomialCD(6,10,10,.
          0.01972770691
```

Figure 12.9 GDC solutions to Example 12.12

Exercise 12.2

1. Consider the binomial distribution $X \sim B(5, 0.6)$

(a) Make a table for this distribution.

(b) Graph this distribution.

(c) Find the mean and standard deviation in two ways:

(i) using the formula

(ii) using the table of values you created in part **(a)**.

2. A poll of 20 adults is taken in a large city. The purpose is to determine whether they support banning smoking in restaurants. It is known that approximately 60% of the population supports the decision. Let x represent the number of respondents in favour of the decision.
   (a) What is the probability that 5 respondents support the decision?
   (b) What is the probability that none of the 20 supports the decision?
   (c) What is the probability that at least 1 respondent supports the decision?
   (d) What is the probability that at least two respondents support the decision?
   (e) Find the mean and standard deviation of the distribution.

k	$P(x \leqslant k)$
0	
1	
2	
3	
4	
5	
6	

Table 12.22 Table for question 3 (a)

3. Consider the binomial random variable with $n = 6$ and $p = 0.3$.
   (a) Copy Table 12.22 and fill in the probabilities.
   (b) Copy and complete the following table. Some cells have been filled in to guide you.

Number of successes x	List the values of x	Write the probability statement	Explain it, if needed	Find the required probability
At most 3				
At least 3				
More than 3	4, 5, 6	$P(x > 3)$	$1 - P(x \leqslant 3)$	0.07047
Fewer than 3				
Between 3 and 5 (inclusive)				
Exactly 3				

4. Repeat question 3 with $n = 7$ and $p = 0.4$
5. A box contains 8 balls: 5 are green and 3 are white, red and yellow. Three balls are chosen at random without replacement, and the number of green balls y is recorded.
   (a) Explain why y is not a binomial random variable.
   (b) Explain why, when we repeat the experiment with replacement, then y is a binomial random variable.
   (c) Give the values of n and p and display the probability distribution in tabular form.
   (d) What is the probability that at most 2 green balls are drawn?
   (e) What is the expected number of green balls drawn?
   (f) What is the variance of the number of balls drawn?
   (g) What is the probability that some green balls will be drawn?

6. On a multiple-choice test, there are 10 questions, each with 5 possible answers, one of which is correct. Nick is unaware of the content of the material so he guesses on all questions. Find:
   - **(a)** the probability that Nick does not answer any question correctly
   - **(b)** the probability that Nick answers at most half of the questions correctly
   - **(c)** the probability that Nick answers at least one question correctly.
   - **(d)** How many questions should Nick expect to answer correctly?

7. Houses in a large city are equipped with alarm systems to protect them from burglary. A company claims their system to be 98% reliable. That is it will trigger an alarm in 98% of the cases. In a certain neighbourhood, 10 houses equipped with this system experience an attempted burglary. Find the probability that:
   - **(a)** all the alarms work properly
   - **(b)** at least half of the houses trigger an alarm
   - **(c)** at most 8 alarms will work properly.

8. Harry Potter books are purchased by readers of all ages. 40% of Harry Potter books were purchased by readers 30 years of age or older. 15 readers are chosen at random. Find the probability that:
   - **(a)** at least 10 of them are 30 years or older
   - **(b)** 10 of them are 30 or older
   - **(c)** at most 10 of them are younger than 30.

9. A factory makes computer hard disks. Over a long period, 1.5% of them are found to be defective. A random sample of 50 hard disks is tested.
   - **(a)** Write down the expected number of defective hard disks in the sample.
   - **(b)** Find the probability that three hard disks are defective.
   - **(c)** Find the probability that more than one hard disk is defective.

10. Car colour preferences change over time and according to the area the customer lives in and the car model he/she is interested in. In a certain city a large dealer of one brand of cars noticed that 10% of the cars sold are metallic grey. Twenty of his customers are selected at random, and their car orders are checked for colour.
    - **(a)** Find the probability that:
      - **(i)** at least 5 cars are metallic grey
      - **(ii)** at most 6 cars are metallic grey
      - **(iii)** more than 5 are metallic grey
      - **(iv)** between 4 and 6 are metallic grey
      - **(v)** more than 15 are not metallic grey.

(b) In a sample of 100 customer records, find:
    (i) the expected number of metallic grey car orders
    (ii) the standard deviation of metallic grey car orders.

According to the empirical rule, 95% of the orders of metallic grey orders are between a and b.

(c) Find a and b.

11. Owners of dogs in many countries buy health insurance for their dogs. 3% of all dogs have health insurance. In a random sample of 100 dogs in a large city find:
    (a) the expected number of dogs with health insurance
    (b) the probability that five of the dogs have health insurance.
    (c) the probability that more than ten dogs have health insurance.

12. A balanced coin is flipped five times. Let x be the number of heads observed.
    (a) Using a table, construct the probability distribution of x.
    (b) What is the probability that no heads are observed?
    (c) What is the probability that all flips are heads?
    (d) What is the probability that at least one head is observed?
    (e) What is the probability that at least one tail is observed?
    (f) Another coin is unbalanced so that it shows 2 heads in every 10 flips. Repeat questions **(a)**–**(e)** for this coin.

13. When John throws a stone at a target, the probability that he hits the target is 0.4.

    He throws a stone 6 times.
    (a) Find the probability that he hits the target exactly 4 times.
    (b) Find the probability that he hits the target for the first time on his third throw.

14. On a television channel the news is shown at the same time each day. The probability that Alice watches the news on a given day is 0.4. Calculate the probability that on five consecutive days, she watches the news on at most three days.

15. A satellite relies on solar cells for its power and will operate provided that at least one of the cells is working. Cells fail independently of each other, and the probability that an individual cell fails within one year is 0.8
    (a) For a satellite with ten solar cells, find the probability that all ten cells fail within one year.
    (b) For a satellite with ten solar cells, find the probability that the satellite is still operating at the end of one year.
    (c) For a satellite with n solar cells, write down the probability that the satellite is still operating at the end of one year. Hence, find the smallest number of solar cells required so that the probability of the satellite still operating at the end of one year is at least 0.95

12.3 Continuous distributions – the normal distribution

When a random variable X is discrete, you assign a positive probability to each value that X can take and get the probability distribution for X. The sum of all the probabilities associated with the different values of X is 1.

A **continuous random variable** is a random variable whose values are not countable. A continuous random variable can assume any value over an interval or intervals.

Because the number of values contained in any interval is infinite, the possible number of values that a continuous random variable can assume is also infinite. Moreover, we cannot count these values.

Earlier, we mentioned that the life of a battery, heights of people, time taken to complete a task, amount of milk in a glass, and weights of babies are all examples of continuous random variables.

In a study of the growth of 12-month old children in a large city, the heights of 5000 children were recorded. X is the continuous random variable that represents the heights of these children in centimetres. Table 12.23 lists the frequency and relative frequency (Probability) distributions of X.

The relative frequencies given in Table 12.23 can be used as the probabilities of the respective classes.

Height of child X (cm)	Frequency	Probability
$60 \leqslant x < 61$	90	0.018
$61 \leqslant x < 62$	170	0.034
$62 \leqslant x < 63$	460	0.092
$63 \leqslant x < 64$	750	0.150
$64 \leqslant x < 65$	970	0.194
$65 \leqslant x < 66$	760	0.152
$66 \leqslant x < 67$	640	0.128
$67 \leqslant x < 68$	440	0.088
$68 \leqslant x < 69$	320	0.064
$69 \leqslant x < 70$	220	0.044
$70 \leqslant x < 71$	180	0.036
	$N = 5000$	Sum $= 1.0$

Table 12.23 Relative frequencies as probabilities

Figure 12.10 displays the histogram and polygon for the relative frequency distribution of Table 12.23.

Figure 12.11 shows the smoothed polygon for the data of Table 12.23.

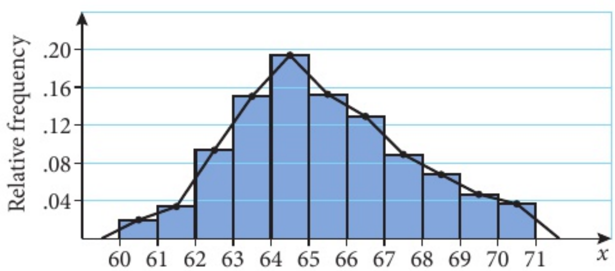


Figure 12.10 Histogram and polygon

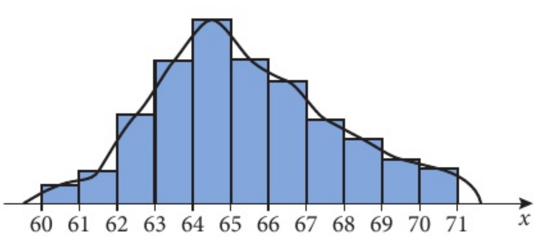


Figure 12.11 Smoothed polygon

Let X be a continuous random variable. The **probability density function**, $f(x)$, of the random variable is a function with the following characteristics:

- $f(x) > 0$ for all values of X
- The probability that x assumes a value in any interval lies in the range 0 to 1.
- The total probability of all the (mutually exclusive) intervals within which x can assume a value is 1.0.
- Suppose this density function is graphed. Let a and b be two possible values of the random variable X, with $a < b$. Then the probability that x lies between a and b, $P(a < x < b)$, is the area under the density function between these points.

The smoothed polygon approximates the probability distribution curve of the continuous random variable X. Note that each class in Table 12.23 has a width equal to 1 cm. If the width of classes is more than 1 unit, we first calculate the relative frequency densities and then graph these relative frequency densities to obtain the distribution curve. The relative frequency density of a class is obtained by dividing the relative frequency of that class by the class width. The relative frequency densities are calculated to make the sum of the areas of all rectangles in the histogram equal to 1.0. The probability distribution curve of a continuous random variable is also called its probability density function.

The second characteristic in the key point box states that the area under the probability distribution curve of a continuous random variable between any two points is between 0 and 1, as shown in Figure 12.12. The third characteristic indicates that the total area under the probability distribution curve of a continuous random variable is always 1.0, or 100%, as shown in Figure 12.13.

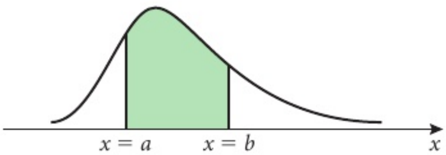


Figure 12.12 Shaded area is between 0 and 1

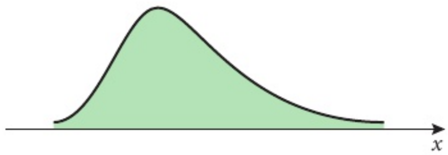


Figure 12.13 Shaded area is 1.0 or 100%

The fourth characteristic, the probability that a continuous random variable X assumes a value within a certain interval is given by the area under the curve between the two limits of the interval, is shown in Figure 12.14. The shaded area under the curve from a to b in this figure gives the probability that x falls in the interval a to b; that is,

$P(a \leqslant x \leqslant b)$ = area from a to b.

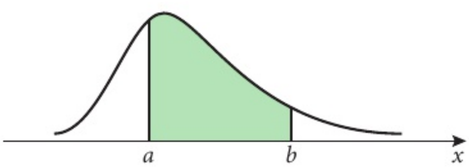


Figure 12.14 Shaded area gives the probability $P(a \leqslant x \leqslant b)$

Note that the interval $a \leqslant x \leqslant b$ states that x is greater than or equal to a but less than or equal to b.

Reconsider the example on children's heights. The probability that the height of a randomly selected child lies in the interval 65 cm to 68 cm is given by the area under the distribution curve of the heights of children from $x = 65$ to $x = 68$, as shown in Figure 12.15. This probability is written as

$P(65 \leqslant x \leqslant 68)$

which states that x is greater than or equal to 65 but less than or equal to 68.

For a continuous probability distribution, the probability is always calculated for an interval.

For example, in Figure 12.15, the interval representing the shaded area is from 65 to 68. Consequently, the shaded area in that figure gives the probability for the interval $65 \leqslant x \leqslant 68$

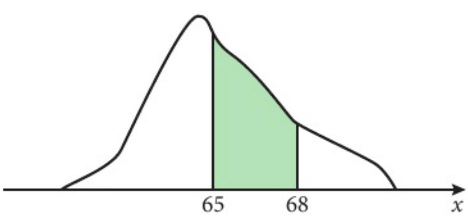


Figure 12.15 Shaded area gives the probability $P(65 \leqslant x \leqslant 68)$

The probability that a continuous random variable x assumes a single value is always zero. This is so because the area of a line, which represents a single point, is zero. For example, if x is the height of a randomly selected child, then the probability that this child is exactly 67 cm tall is zero; that is,

$$P(x = 67) = 0$$

This probability is shown in Figure 12.16. Similarly, the probability for X to assume any other single value is zero.

In general, if a and b are two of the values that X can assume, then

$$P(a) = 0 \text{ and } P(b) = 0$$

From this we can deduce that for a continuous random variable,

$$P(a \leqslant x \leqslant b) = P(a < x < b)$$

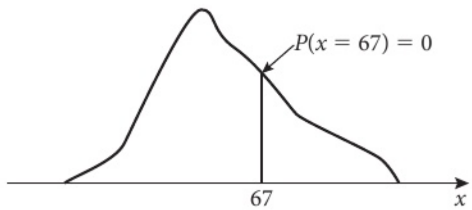


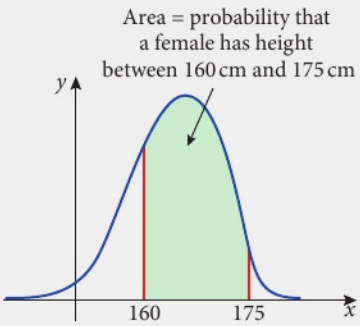
Figure 12.16 Probability of a single value is zero

In other words, the probability that x assumes a value in the interval a to b is the same whether or not the values a and b are included in the interval.

$$P(a < x < b) = P(a \leqslant x \leqslant b) = P(a \leqslant x < b) = P(a < x \leqslant b)$$

This example is only meant to clarify the concept and will not be on any examination for IB/SL.

The graph shows a model for the PDF f for a random variable X defined to be the height, in cm, of an adult female in Spain. The probability that the height of a female chosen at random from this population is between 160 cm and 175 cm is equal to the area under the curve between 160 and 175.



The function represented here is

$$f(x) = \int_{160}^{175} \frac{e^{-\frac{(x-165)^2}{50}}}{5\sqrt{2\pi}}\,dx$$

This is not an integral you can calculate exactly. So, we use a GDC to approximate it.

```
∫[160,175] (e^((−(x−165)²)÷50))÷(5/
                          0.8185946141
```

So, the chance of choosing a female at random with a height between 160 cm and 175 cm is approximately 81.9%.

The normal distribution

Continuous probability distributions can assume a variety of forms. Here we will focus on one distribution – the **normal probability distribution**.

The normal distribution is one of the many probability distributions that a continuous random variable can possess. The normal distribution is the most important and most widely used of all probability distributions. A large number of phenomena in the real world are normally distributed either exactly or approximately. The continuous random variables representing heights and masses of people, scores on an examination, masses of packages (e.g., cereal boxes, boxes of cookies), amount of milk in a bottle, life of an item (such as a light-bulb or a television set), and time taken to complete a certain job have all been observed to have a (approximate) normal distribution.

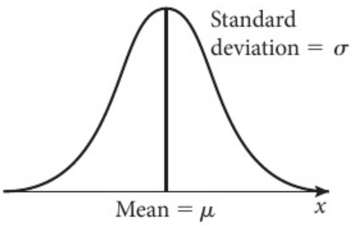


Figure 12.17 Bell-shaped curve

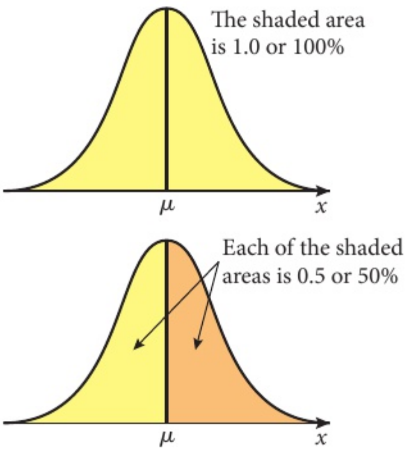


Figure 12.18 The curve is symmetric about the mean

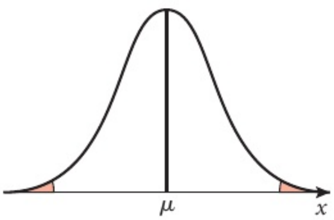


Figure 12.19 The shaded regions are virtually zero

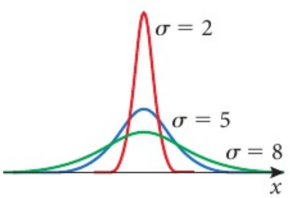


Figure 12.20 The curves have the same mean but different standard deviation

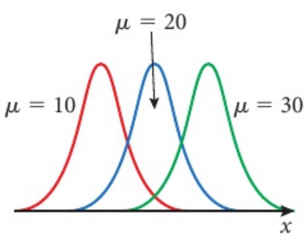


Figure 12.21 The curves have different means but the same standard deviation

The normal probability distribution or the normal curve is a bell-shaped (symmetric) curve. Such a curve is shown in Figure 12.17. Its mean is denoted by μ and its standard deviation by σ. A continuous random variable X that has a normal distribution is called a normal random variable. Note that not all bell-shaped curves represent a normal distribution curve. Only a specific kind of bell-shaped curve represents a normal curve.

The graph of a normal probability distribution is a bell-shaped curve with the following characteristics:

- The total area under the curve is 1.0.
- The curve is symmetric about the mean. These areas are shown in Figure 12.18.
- The two tails of the curve extend indefinitely in both directions without touching or crossing the horizontal axis. Although a normal distribution curve never meets the horizontal axis, beyond the points represented by $\mu - 3\sigma$ and $\mu + 3\sigma$ it becomes so close to this axis that the area under the curve beyond these points in both directions can be taken as virtually zero (Figure 12.19).

The mean, μ, and the standard deviation, σ, are the parameters of the normal distribution. Given the values of these two parameters, we can find the area under a normal distribution curve for any interval. Remember, there is not just one normal distribution curve but a family of normal distribution curves. Each separate set of values of μ and σ gives a different normal distribution.

The value of μ determines the centre of a normal distribution curve on the horizontal axis, and the value of σ gives the spread of the normal distribution curve. The three normal distribution curves in Figure 12.20 have the same mean but different standard deviations. By contrast, the three normal distribution curves in Figure 12.21 have different means but the same standard deviation.

Like the binomial probability distribution, the normal probability distribution can also be expressed by a mathematical equation. However, we will not use this equation to find the area under a normal distribution curve. Instead, we will use a GDC or software to do the calculations.

Historically tables were used, but they are becoming obsolete with the ability of calculators to produce more accurate values.

When a variable is normally distributed, we write: $X \sim \mathrm{N}(\mu, \sigma^2)$

Although there are many normal curves, they all have common properties. Here is one important property.

The **empirical rule** states that in the normal distribution with mean μ and standard deviation σ:

- approximately 68% of the observations fall within σ of the mean μ
- approximately 95% of the observations fall within 2σ of the mean μ
- approximately 99.7% of the observations fall within 3σ of the mean μ

Figure 12.22 illustrates this rule. Later in this section, you will learn how to find these areas from your GDC.

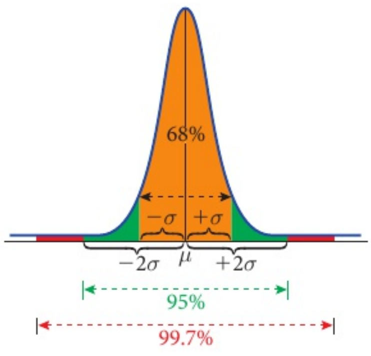


Figure 12.22 Empirical rule

Example 12.13

Heights of young German men between 18 and 19 years of age follow a distribution that is approximately normal with mean 181 cm with a standard deviation of approximately 8 cm. Describe this population of young men.

Solution

According to the empirical rule, we find that approximately 68% of those young men have a height between 173 cm and 189 cm, 95% of them between 165 cm and 197 cm, and 99.7% between 157 cm and 205 cm. Looking further, you can say that only 0.15% are taller than 205 cm or shorter than 157 cm.

Calculations using the normal distribution

To find the probability that a normal variable X lies in the interval a to b, we need to find the area under the normal curve $N(\mu, \sigma^2)$ between the points a and b.

In Example 12.13, if we are interested in the proportion of young German men whose height is between 173 cm and 189 cm, a GDC can give the answer in several ways, as shown in Figure 12.23.

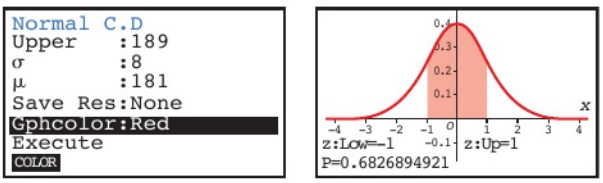


Figure 12.23 Proportion of young German men whose height is between 173 cm and 189 cm

That is $P(173 \leqslant x \leqslant 189) = 0.683$ or 68.3% which is a slightly more accurate value than the 68% we have with the empirical rule.

Figure 12.24 shows the other two probabilities given by the empirical rule:

$P(165 \leqslant x \leqslant 197) = 0.954$ or $P(157 \leqslant x \leqslant 205) = 0.997$

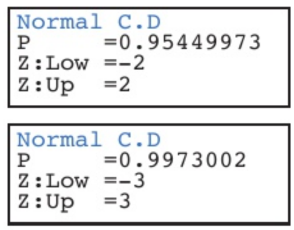


Figure 12.24 Calculations for the empirical rule

Furthermore, if we want to find the chance that a young German man is taller (shorter) than 175 cm, perform similar calculations putting the upper limit much higher than 3 standard deviations above (below) the mean.

For demonstration purposes we used 300 (alternatively 100). The use of 100 or 300 is arbitrary. You can choose other values.

That is $P(x \geqslant 175) = 0.773$ and $P(x \leqslant 175) = 0.227$, as in Figure 12.25.

Notice here that $P(x \leqslant 175) + P(x \geqslant 175) = 1.0$ as expected.

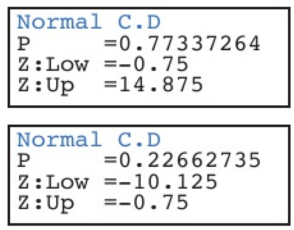


Figure 12.25 $P(x \geqslant 175)$ or $P(x \leqslant 175)$

You may notice that in most GDC outputs there are values corresponding to a variable z, which is the standard normal variable corresponding to any normal value we have. z expresses the deviation of any normal variable from its mean as a multiple of the standard deviation and is calculated using the formula $z = \dfrac{x - \mu}{\sigma}$

Example 12.14

The age of graduate students in engineering programs throughout the USA is normally distributed with mean $\mu = 24.5$ years and standard deviation $\sigma = 2.5$ years.

A student is chosen at random.

(a) What is the probability the student is younger than 26 years old?

(b) What proportion of students is older than 23.7 years?

(c) What percentage of students is between 22 and 28 years old?

Solution

If we let X = age of students, then $X \sim \mathrm{N}(\mu = 24.5, \sigma^2 = 6.25)$

(a) To answer this, we need $\mathrm{P}(x < 26)$

$\mathrm{P}(x < 26) = 0.726$

See the first two screens of Figure 12.26. Notice here that we put 0 as a lower limit. You can put a number as a lower limit far enough from the mean to make sure you are receiving the correct cumulative distribution. Remember to put the standard deviation of 2.5.

(b) This can be done similarly, we need $\mathrm{P}(x > 23.7)$,

$\mathrm{P}(x > 23.7) = 0.626$

See the middle two screens of Figure 12.26. Also notice here that we wrote 100 as an upper limit, which is an arbitrary number far enough to the right to be sure we include the entire population.

(c) $\mathrm{P}(22 < X < 28) = 0.761$ (last two screens of Figure 12.26)

```
Normal C.D
Data     :Variable
Lower    :0
Upper    :26
σ        :2.5
μ        :24.5
```

```
Normal C.D
P      =0.72574688
Z:Low  =-9.8
Z:Up   =0.6
```

```
Normal C.D
Data     :Variable
Lower    :23.7
Upper    :100
```

```
Normal C.D
P      =0.62551583
Z:Low  =-0.32
Z:Up   =30.2
```

```
Normal C.D
Data     :Variable
Lower    :22
Upper    :28
```

```
Normal C.D
P      =0.76058808
Z:Low  =-1
Z:Up   =1.4
```

Figure 12.26 GDC screens for the solution to Example 12.14

The inverse normal distribution

Another type of problem arises in situations similar to the one above when we are given a cumulative probability and would like to find the value in our data that has this cumulative probability. For example, what height marks the 95th percentile? That is, what height is taller than or equal to 95% of the population? To answer this question, we need to reverse our steps. So far, we have been given a value and then looked for the area corresponding to it. Now, we are given the area, and we have to look for the number. That is why this is called the inverse normal distribution.

That is, we need to find k such that $\mathrm{P}(x < k) = 0.95$

You can use a GDC. The process is identical to the normal calculation. Just choose invNorm instead.

From Figure 12.27, 95% of the young German men are shorter than 194.16 cm.

```
Inverse Normal
Data     :Variable
Tail     :Left
Area     :0.95
σ        :8
μ        :181
Save Res:None
```

```
Inverse Normal
 xInv=194.158829
```

Figure 12.27 Using invNorm on a GDC

Example 12.15

Since November 2007, the average time it takes fast trains (Eurostar) to travel between London and Paris is 2 hours 15 minutes with a standard deviation of 4 minutes. Assume a normal distribution.

(a) What is the probability that a randomly chosen trip will take longer than 2 hours and 20 minutes?

(b) What is the probability that a randomly chosen trip will be shorter than 2 hours and 10 minutes?

(c) What is the IQR of a trip on these trains?

(d) 5% of trips take longer than time t. Find t.

Solution

We will do each problem using a GDC. We will use hours as a unit. So, minutes are changed into decimal by dividing by 60. The mean $\mu = 2.25$ and $\sigma = 0.0667$

(a) 2 hours 20 minutes = 2.333

Using your GDC, $P(x > 2.333) = 0.1067$

(b) 2 hours 10 minutes = 2.167

$P(x < 2.167) = 0.1067$

(c) To find IQR, we need to find Q_1 and Q_3.

Q_1 is the number that leaves 25% of the data before it. Q_3 is the number that leaves 75% of the data before it. So, we need to find the inverse normal variable that has an area of 0.25 or 0.75 before it.

Using a GDC and the inverse normal, we find that $Q_1 = 2.205$ and $Q_3 = 2.295$

$IQR = 2.295 - 2.205 = 0.090$ of an hour, i.e., 5.4 minutes.

(d) 5% is the area under the normal curve in the tail to the right of t. Thus, t is the number that has an area of 0.95 before it. So, to find t, we need to use the inverse normal.

Using a GDC and the inverse normal, we find that $t = 2.3597$, that is 2 hours and 22 minutes (to the nearest minute).

Exercise 12.3

1. The time the batteries of your GDC last is approximately normal with mean 50 hours and standard deviation of 7.5 hours.

Find the probability that your newly equipped GDC will last:

(a) at least 50 hours

(b) between 50 and 75 hours

(c) less than 42.5 hours

(d) between 42.5 and 57.5 hours

(e) more than 65 hours

(f) 47.5 hours

2. Find each probability.
   (a) $P(x < 3.7)$, where $X \sim N(3, 3)$
   (b) $P(x > -3.7)$, where $X \sim N(3, 3)$
3. A car manufacturer introduces a new model that has a fuel consumption of 11.4 litres per 100 km in urban areas. Tests show that this model has a standard deviation of 1.26 litres per 100 km. The distribution is assumed to be normal.
   A car is chosen at random from this model.
   (a) What is the probability that it will have consumption less than 8.4 litres per 100 km?
   (b) What is the probability that the consumption is between 8.4 and 14.4 litres per 100 km?
4. The assembly times in a factory for a racing car toy follow a normal distribution with a mean of 55 minutes and a standard deviation of 4 minutes. The factory closes at 5 p.m. every day. If one worker starts to assemble a racing car at 4 p.m., what is the probability that she will finish this job before the company closes for the day? Shade the area corresponding to your answer on a normal distribution sketch.
5. A manufacturer produces apple juice. The filling machines are adjusted to pour 360 cc of juice into each 360 cc bottle. However, the actual amount of juice poured into each bottle is not exactly 360 cc; it varies from bottle to bottle. It has been observed that the net amount of juice in such a bottle has a normal distribution with a mean of 360 cc and a standard deviation of 0.45 cc.
   (a) What is the probability that a randomly selected bottle contains 359.1 cc to 359.7 cc of juice?
   (b) What percentage of the juice bottles contain 360.6 cc to 362.1 cc of juice?
6. The scores on a public schools' examination are normally distributed with a mean of 550 and a standard deviation of 100.
   (a) What is the probability that a randomly chosen student from this population scores below 400?
   (b) What is the probability that a student will score between 450 and 650?
   (c) What score should you have in order to be in the 90th percentile?
   (d) Find the IQR of this distribution.
7. A company producing and packaging sugar for home consumption put labels on their sugar bags noting the mass to be 500 g. Their machines are known to fill the bags with masses that are normally distributed with a standard deviation of 5.7 g. A bag that contains less than 500 g is considered to be underweight.
   (a) The company decides to set their machines to fill the bags with a mean of 512 g. What fraction will be underweight?
   (b) Bags that have a mass more than 525 g are considered a loss to the company. What percent of the bags are considered a loss?

8. In a large school, heights of students who are 13 years old are normally distributed with a mean of 151 cm and a standard deviation of 8 cm.

   Find the probability that a randomly chosen student is:

   (a) shorter than 166 cm

   (b) within 6 cm of the mean.

9. The time it takes Kevin to get to school every day is normally distributed with a mean of 12 minutes and 2 minutes standard deviation. Estimate the number of days when he takes:

   (a) longer than 17 minutes

   (b) less than 10 minutes

   (c) between 9 and 13 minutes.

   There are 180 school days in Kevin's school year.

10. The life span of a calculator manufactured by a major company has a normal distribution with a mean of 54 months and a standard deviation of 8 months. The company guarantees that any calculator that starts malfunctioning within 36 months of the purchase will be replaced by a new one. About what percentage of calculators made by this company are expected to be replaced?

11. Consider the normal distribution with mean 56 and standard deviation 2. Find the values of two numbers a and b in this distribution that are symmetric with respect to the mean and enclose an area of 0.90 between them as shown in Figure 12.28.

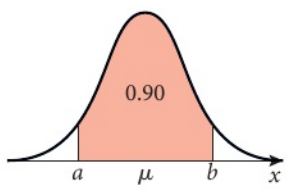


Figure 12.28 Diagram for question 11

12. Consider the normal distribution with mean 72 and standard deviation 3. Find the values of two numbers a and b in this distribution that are symmetric with respect to the mean and enclose an area of 0.80 between them.

13. Consider the normal distribution with mean 72 and standard deviation 3. Find the values of two numbers a and b in this distribution that are symmetric with respect to the mean so that the interval $[a, b]$ leaves an area of 0.025 in the two tails as shown in Figure 12.29.

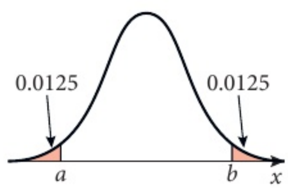


Figure 12.29 Diagram for question 13

14. Wooden poles produced for electricity networks in rural areas have lengths that are normally distributed with mean of 6.966 m and standard deviation of 0.324 m.

   (a) The acceptable range is between 6.3 m and 7.5 m. What percentage of the poles are:

   (i) too short **(ii)** too long?

   (b) In a randomly selected sample of 20 poles, find the probability of finding:

   (i) 2 poles that are considered too short

   (ii) 2 poles that are considered too long.

15. Bottles of mineral water sold by a company are advertised to contain 1 litre of water. The company adjusts its filling process to fill the bottles with an average of 1012 ml to ensure that there is a minimum of 1 litre. The process follows a normal distribution with standard deviation of 5 ml.

(a) Find the probability that a randomly chosen bottle contains more than 1010 ml.

(b) Find the probability that a bottle contains less than the advertised volume.

(c) In a shipment of 10 000 bottles, what is the expected number of underfilled bottles.

16. Cholesterol plays a major role in a person's heart health. High blood cholesterol is a major risk factor of coronary heart disease and stroke. The level of cholesterol in the blood can be measured in milligrams per decilitre of blood (mg/dL). According to the World Health Organisation, in general, less than 200 mg/dL is a desirable level, 200–239 mg/dL is borderline high, and above 240 mg/dL is a high risk level and a person with this level has more than twice the risk of heart disease as a person with less than 200 mg/dL.

In a certain country, it is known that the average cholesterol level of their adult population is 184 mg/dL with a standard deviation of 22 mg/dL. It can be modelled by a normal distribution.

(a) What percent do you expect to be borderline high?

(b) What percent do you consider are high risk?

(c) Estimate the interquartile range of the cholesterol levels in this country.

(d) Above what value are the highest 2% of adults' cholesterol levels in this country?

17. A manufacturer of car tyres claims that its winter tyres can be described by a normal model with an average life of 52 000 km and a standard deviation of 4000 km.

(a) You buy a set of tyres from this manufacturer. Is it reasonable for you to hope they last more than 64 000 km?

(b) What fraction of these tyres do you expect to last less than 48 000 km?

(c) What fraction of these tyres do you expect to last between 48 000 km and 56 000 km?

(d) What is the IQR of the life of this type of tyre?

(e) The company wants to guarantee a minimum life for these tyres. They will refund customers whose tyres last less than a specific distance. What should their minimum life guarantee be so that they do not end up refunding more than 2% of their customers?

18. A machine produces bearings with diameters that are normally distributed with mean 3.0005 cm and standard deviation 0.0010 cm. Specifications require the bearing diameters to lie in the interval 3.000 ± 0.0020 cm.

Those outside the interval are considered scrap and must be disposed of. What fraction of the production will be scrap?

19. A soft-drink machine can be regulated so that it discharges an average 216 cc per bottle. The amount of fill is normally distributed with standard deviation 9 cc. Give the maximum size of the bottles so that they will overflow only 1% of the time.

20. A soft-drink machine can be regulated so that it discharges an average 216 cc per bottle. The amount of fill is normally distributed with standard deviation 9 cc. The amount of drink discharged 95% of the time lies between a and b above and below the mean. Find a and b.

21. The speeds of cars on a main highway are approximately normal with mean 111.89 km h^{-1} and standard deviation 17.9 km h^{-1}. The speed limit on this highway is set to be 140 km h^{-1}. Cars travelling slower than 90 km h^{-1} are considered a hazard because they are too slow for the rest of cars.

(a) What percent of cars travel within the acceptable limits?

(b) Find the proportion of cars that travel at speeds exceeding 110 km h^{-1}.

Chapter 12 practice questions

1. Residents of a small town have savings which are normally distributed with a mean of \$3000 and a standard deviation of \$500.

(a) What percentage of townspeople have savings greater than \$3200?

(b) Two townspeople are chosen at random. What is the probability that both of them have savings between \$2300 and \$3300?

(c) The percentage of townspeople with savings less than d dollars is 74.22%. Find the value of d

2. The mass, W, of bags of rice follows a normal distribution with mean 1000 g and standard deviation 4 g.

(a) Find the probability that a bag of rice chosen at random has a mass between 990 g and 1004 g.

95% of the bags of rice have a mass less than k grams.

(b) Find the value of k .

For a bag of rice chosen at random, $P(1000 - a < W < 1000 + a) = 0.9$

(c) Find the value of a.

3. A fair coin is flipped eight times. Calculate the probability of obtaining:

(a) exactly 4 heads **(b)** exactly 3 heads **(c)** 3, 4 or 5 heads

4. The lifespan of a particular species of insect is normally distributed with a mean of 57 hours and a standard deviation of 4.4 hours.

(a) Find the probability that the lifespan of an insect of this species is:

(i) more than 55 hours

(ii) between 55 and 60 hours.

(b) 90% of the insects die after t hours. Find the value of t.

5. Intelligence quotient (IQ) in a certain population is normally distributed with a mean of 100 and a standard deviation of 15.

(a) What percentage of the population has an IQ between 90 and 125?

(b) If two people are chosen at random from the population, what is the probability that both have an IQ greater than 125?

6. The mass of packets of a breakfast cereal is normally distributed with a mean of 750 g and standard deviation of 25 g.

(a) Find the probability that a packet chosen at random has mass:

(i) less than 740 g **(ii)** at least 780 g **(iii)** between 740 g and 780 g

(b) Two packets are chosen at random. What is the probability that both packets have a mass which is less than 740 g?

(c) The mass of 70% of the packets is more than x grams. Find the value of x.

7. In a village in Wales, the height of adults is normally distributed with a mean of 187.5 cm and a standard deviation of 9.5 cm.

(a) What percentage of adults in the village have a height greater than 197 cm?

(b) A standard doorway in the village is designed so that 99% of adults have a space of at least 17 cm over their heads when going through a doorway. Find the height of a standard doorway in the village. Give your answer to the nearest cm.

8. It is claimed that the masses of a population of lions are normally distributed with a mean mass of 310 kg and a standard deviation of 30 kg.

(a) Calculate the probability that a lion selected at random will have a mass of 350 kg or more.

(b) The probability that the mass of a lion lies between a and b is 0.95, where a and b are symmetric about the mean. Find the value of a and of b.

9. Reaction times of human beings are normally distributed with a mean of 0.76 seconds and a standard deviation of 0.06 seconds.

(a) Calculate the probability that the reaction time of a person chosen at random is:

(i) greater than 0.70 seconds

(ii) between 0.70 and 0.79 seconds.

Three percent (3%) of the population have a reaction time less than c seconds.

(b) **(i)** Represent this information on a diagram, indicating the area representing 3%.

(ii) Find c

10. A factory makes calculators. Over a long period, 2% of them are found to be faulty. A random sample of 100 calculators is tested.

(a) Write down the expected number of faulty calculators in the sample.

(b) Find the probability that three calculators are faulty.

(c) Find the probability that more than one calculator is faulty.

11. Ball bearings are used in engines in large quantities. A car manufacturer buys these bearings from a factory. They agree on the following terms. The car company chooses a sample of 50 ball bearings from the shipment. If they find more than 2 defective bearings, the shipment is rejected. It is a fact that the factory produces 4% defective bearings.

(a) What is the probability that the sample is clear of defects?

(b) What is the probability that the shipment is accepted?

(c) What is the expected number of defective bearings in the sample of 50?

12. The table shows the probability distribution of a random variable X.

X	0	1	2	3
P(x)	$2k$	$2k^2$	$k^2 + k$	$2k^2 + k$

(a) Calculate the value of k.

(b) Find E(X).

13. It is estimated that 2.3% of the cherry tomatoes produced on a certain farm are considered to be small and cannot be sold for commercial purposes. The farmers have to separate such fruits and use them for domestic consumption instead.

(a) 12 tomatoes are randomly selected from the produce. Calculate the probability that:

(i) three are not fit to be sold **(ii)** at least 4 are not fit to be sold.

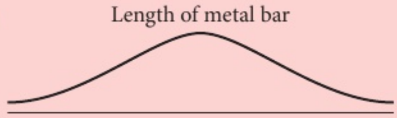
(b) It is known that the sizes of such tomatoes are normally distributed with a mean of 3 cm and a standard deviation of 0.5 cm. Tomatoes that are categorised as large have to be larger than 2.5 cm. What proportion of the produce is large?

14. A factory makes metal bars. Their lengths are assumed to be normally distributed with a mean of 180 cm and a standard deviation of 5 cm.

(a) On the provided diagram, shade the region representing the probability that a metal bar, chosen at random, will have a length less than 175 cm.

A metal bar is chosen at random.

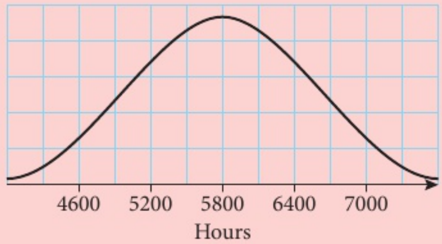
(b) **(i)** The probability that the length of the metal bar is less than 175 cm is equal to the probability that the length is greater than h cm. Write down the value of h.



(ii) Find the probability that the length of the metal bar is greater than one standard deviation above the mean.

15. The lifetime, L, of light bulbs made by a company follows a normal distribution.

L is measured in hours.
The normal distribution curve of L is shown in the diagram.



(a) Write down the mean lifetime of the light bulbs.

The standard deviation of the lifetime of the light bulbs is 850 hours.

(b) Find the probability that $5000 \leqslant L \leqslant 6000$, for a randomly chosen light bulb.

The company states that 90% of the light bulbs have a lifetime of at least k hours.

(c) Find the value of k. Give your answer correct to the nearest hundred.

16. A speed camera on Peterson Road records the speed of each passing vehicle. The speeds are found to be normally distributed with a mean of $67\,\text{km h}^{-1}$ and a standard deviation of $3.4\,\text{km h}^{-1}$.

(a) Sketch a diagram of this normal distribution and shade the region representing the probability that the speed of a vehicle is between 60 and $70\,\text{km h}^{-1}$.

A vehicle on Peterson Road is chosen at random.

(b) Find the probability that the speed of this vehicle is:

(i) more than $60\,\text{km h}^{-1}$

(ii) less than $70\,\text{km h}^{-1}$

(iii) between 60 and $70\,\text{km h}^{-1}$.

It is found that 19% of the vehicles are exceeding the speed limit of $s\,\text{km h}^{-1}$.

(c) Find the value of s, correct to the nearest integer.

There is a fine of US$65 for exceeding the speed limit on Peterson Road. On a particular day the total value of fines issued was US$14,820.

(d) **(i)** Calculate the number of fines that were issued on this day.

(ii) Estimate the total number of vehicles that passed the speed camera on Peterson Road on this day.

13 Statistical analysis

13 Statistical analysis

Learning objectives

By the end of this chapter, you should be familiar with…

- null and alternative hypotheses, H_0 and H_1
- significance levels and p-values
- expected and observed frequencies
- the χ^2 test for independence: contingency tables, degrees of freedom, critical values, and carrying out the test using technology
- the χ^2 goodness of fit test
- the t-test for means of populations – one-tailed and two-tailed tests
- the p-value to compare the means of two populations.

In several countries, court trials are familiar cases where a choice between two opposing arguments must be made. The person charged of an offense must be judged as innocent or at fault. Initially, in these countries, the accused is assumed innocent. Only strong evidence to the contrary will form a basis for the innocence claim to be rejected and the guilty claim to be assumed instead. The burden of proof is on the prosecution to prove the guilty claim.

Put differently, in court, there are two statements (hypotheses):

H_0: The defendant is innocent – this we call the **null hypothesis**.
H_0 is assumed unless there is evidence beyond doubt that it is false.
We cannot have evidence to prove a defendant innocent. We only fail to find evidence against the defendant.

H_1: The defendant is guilty – this we call the **alternative hypothesis**.
H_1 is claimed true if we have evidence beyond doubt against the null hypothesis.

13.1 Hypothesis testing in statistics

Calculations involved in the following two examples are for demonstration only. You will not be asked to perform them on exams. The only tests you will be examined on are 2-sample means t-test and the chi-squared test. Additionally, your GDC will provide you with the necessary values you can use to interpret the results!

Hypothesis testing is an area of inferential statistics. Several courses (e.g. biology and psychology) use hypothesis testing to perform the mathematics necessary to complete their Internal Assessment (IA) component.

The hypothesis tests that we will look at here are the chi-squared test and the t-test (for difference of means.) However, so that you will more fully appreciate these tests, a short introduction into the general concepts and methodologies of hypothesis testing is given on the following pages. These concepts will not be examined directly.

In an assembly plant for personal computers, the time technicians spend putting together systems is of extreme importance. Regular inspection of the time each process takes is a task that supervisors need to carry out. A supervisor believes that the average time it takes technicians to install a graphics card has changed. Previously, it took on average 2 minutes ($\mu = 2$) to complete this task with a standard deviation of $\sigma = 0.2$. Now she believes that the average has increased, i.e., $\mu > 2$

How might she decide, based on a sample of 100 observations, whether there is evidence that the mean is larger than 2 minutes?

Let us assume we are helping her out. We begin by calculating the mean of this sample and find that it is $\bar{x} = 2.2$ minutes.

Knowing something about the sampling process, we realise that, no matter how random our sample is, there will always be a chance for **sampling error**.

The question we need to answer is: assuming that the mean of the population is 2 minutes, how likely is it to have a sample with mean of 2.2 minutes or larger?

How significant/surprising is the error of 0.2 minutes?

The task is to calculate $P(\bar{x} > 2.2)$

A GDC output is shown in Figure 13.1. $P(\bar{x} > 2.2) \approx 7.6 \times 10^{-24} \approx 0$

```
Normal C.D
    p=7.6199E-24
```

Figure 13.1 GDC output

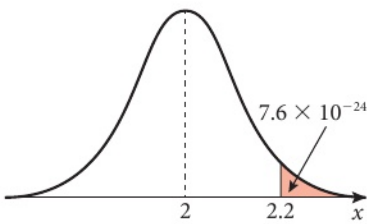


Figure 13.2 $P(\bar{x} > 2.2)$

That is, the likelihood that a process with a mean of 2 minutes would produce a sample with a mean of 2.2 minutes is extremely small. We are confident that we have evidence from the collected data to show that the mean time cannot be 2 minutes. This means that the claim that $\mu = 2$ is rejected on the basis of collected evidence.

What we just did here is a test of the hypothesis $\mu = 2$.

In hypothesis-testing problems, we have two contradictory hypotheses under consideration. One hypothesis in the example above is that $\mu = 2$, and the other is $\mu > 2$. If it were possible to collect data from the entire population, we would know which of the two hypotheses was true. Usually this is not possible, or very difficult and impractical in most cases. As a result, we must decide which hypothesis we think is true by using information from a well randomised sample.

A **statistical hypothesis** is a claim (assertion) about a population parameter. It is often a claim about the value of the parameter, in this case, the mean, μ.

In every hypothesis testing problem, we will have a **null hypothesis** H_0 (in this case $\mu = 2$) and an **alternative hypothesis** H_1 (in this case $\mu > 2$). The testing is a method for deciding which of the two hypotheses may be correct. We assume the null hypothesis, H_0, is correct. After carrying out the test, this hypothesis will only be rejected in favour of the alternative, H_1, if sample evidence is incompatible with H_0. If we fail to provide evidence against H_0, then the claim is that we fail to reject H_0.

The null hypothesis, H_0

- states the assumption to be tested – for example, the mean time technicians take to complete the task is 2 minutes ($H_0: \mu = 2$)
- begins with the assumption that the null hypothesis is true
- refers to the existing situation
- may or may not be rejected.

The alternative hypothesis, H_1

- is the opposite of the null hypothesis
- challenges the existing situation
- is generally the hypothesis that is believed (or needed to be proven) to be true by the researcher.

In the example above, the null hypothesis is that the mean time technicians take to complete the task is 2 minutes ($H_0: \mu = 2$). In fact this includes the case that $\mu < 2$ too, because if the mean time is less than 2 minutes, there is no reason for the manager to complain. This is why, when we say $H_0: \mu = 2$ we actually mean $H_0: \mu \leqslant 2$.

The alternative hypothesis is that the true mean time is larger than 2 minutes, $H_1: \mu > 2$.

Example 13.1

A factory produces light bulbs that are known to last 8000 hours on average. A new type has been developed by engineers and is claimed to last more than 8000 hours. The management is concerned that the cost of the process of producing the new type is more expensive than the traditional method. They can justify shifting to the new process only if they have evidence that the new light bulbs do actually have a longer life than 8000 hours. A random sample of 25 new bulbs was tested and gave an average life of 8600 hours with standard deviation 1000 hours. Is this enough evidence to justify the shift to the new process?

Solution

Since we are attempting to show that the population mean exceeds 8000 hours, it seems reasonable to base our decision on the value of the sample mean. If $\bar{x}$ is much larger than (or significantly away from) 8000 hours, then we have compelling evidence that the mean $m > 8000$. We need to determine whether the sample mean of 8600 exceeds 8000 by an amount that would be considered unlikely to occur by chance.

We first set up our hypotheses.

$H_0: \mu = 8000$

$H_1: \mu > 8000$

We can calculate the chance that a sample of mean 8600 or larger can happen by chance from a population of mean 8000. We call $\bar{x}$ the test statistic.

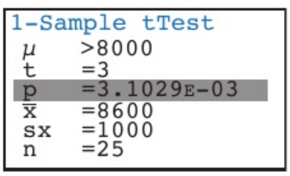


Figure 13.3 GDC output for solution to Example 13.1

Figure 13.3 shows the output of a GDC.

As you can see, $P(\bar{x} > 8600) \approx 3.1029 \times 10^{-3} \approx 0.0031029$, which is very close to zero. We can claim that there is a very small probability that a population whose mean is 8000 could ever yield a sample with 8600 by mere chance. We reject the null hypothesis in favour of the alternative. That is, we have enough evidence to reject the claim that the new light bulbs are the same as the old ones.

You are not expected to do the calculations involved in Example 13.1. Your task is to interpret the results your GDC gives you.

The hypothesis testing procedure for the population mean

In Example 13.1, how can we be confident about our decision that the new light bulbs are better than the old ones?

We start with the null hypothesis (H_0), specify a population parameter, μ, and suggest a value for that parameter, 8000 in this case. We usually write down a null hypothesis about a mean, for example, as

$H_0: \mu = \mu_0$ ($H_0: \mu = 8000$)

This is a short way of indicating the two items we need most: the nature of the parameter we hope to learn about (the true mean) and a particular assumed value for that parameter (8000 in this case). We need the particular value so we can judge our observed statistic against it.

The alternative hypothesis, H_1 (sometimes called H_a), contains the value(s) of the parameter that we regard as reasonable in case the null hypothesis is rejected. In the light bulbs example the alternative is the life of the bulbs being more than 8000 hours. We also write it as

$H_1: \mu > \mu_0$ ($H_1: \mu > 8000$)

In Example 13.1, we were interested in an alternative $H_1: \mu > \mu_0$, which is called an **upper tail test** But in other cases we could also be interested in $H_1: \mu < \mu_0$, which is called a **lower tail test**, or $H_1: \mu \neq \mu_0$, which is called a **two-tail test**.

What persuades us to believe that the light bulbs have a life more than 8000 hours?

We should not expect to have a sample mean exactly equal to 8000 hours as observations vary from one sample to the other. We base our decision on how significantly surprising our sample result is under the assumption that the true mean is 8000 in this example. That is, do we consider 8200 hours to be a surprising result? If not, is 8400 or 8600 hours surprising?

To answer these questions, we find $P(\bar{x} > 8200) = 0.164$

So, if the mean life of these bulbs is 8000 hours, the chance of randomly getting a sample mean of 8200 hours or above is about 16%. In other words, there is a good chance that a population with mean 8000 hours can possibly give a sample of average 8200 hours.

How surprising is 8400?

To answer this question, we find $P(\bar{x} > 8400) = 0.0285$

Thus, if the true mean life is 8000 hours, the chance that we can get a sample with average 8400 hours or more is less than 2.9%. You may think that this result is significantly surprising and you conclude that the mean lifetime has to be larger than 8000 in order to produce such a sample by chance. Someone else may not think so.

How surprising is 8600 hours?

As we have seen above, $P(\bar{x} > 8600) = 0.0031$

In this case, the event of finding a random sample with a mean of 8600 hours or above from a population with mean of 8000 hours is extremely unlikely, and we find ourselves convinced that the population must have a higher mean than 8000 hours in order to render a random sample with a mean of 8600 hours or more by mere chance.

As you notice from the previous discussion, the fundamental step in our analysis is the question, 'are the sample data unexpected, given the null hypothesis?' The key calculation is to determine how likely the sample data we observed would be if the null hypothesis were the true model of the world. That is why we need a probability. We would like to find the probability of observing sample data like these given the null hypothesis. This probability is the value we base our decision on. This probability is called the ***p*-value**.

A small p-value indicates that the sample data we see would be very unlikely had our null hypothesis been true. That is, we start with a model in mind, we collect the data, and then the model tells us that this data we have is unlikely to have happened. That is surprising. The model and data are not compatible and hence, we have to make a decision. Either the model, the null hypothesis, is true and we have been unlucky to get such a remarkably unexpected sample, or the null hypothesis is at fault – that is, we were not correct to use it as a basis for calculating our p-value. Given that the sample data is real, while the model (null hypothesis) is an assumption, we are tempted to reject the model.

When the p-value is large (or just not small enough), what do we conclude? In that case, we have not found anything unlikely or surprising or unexpected. So, we have no reason to reject the null hypothesis. In this case, it does not mean that we proved the null hypothesis. It only means that it does not appear that the hypothesis is false. Formally, we say that we **do not reject the null hypothesis**. All we were able to establish is that the sample data we have at hand is consistent with the model. We did not and could not collect all the evidence to support the null hypothesis. Unfortunately, the decision to reject it is more appealing, usually as we have a contradicting example that proves it wrong!

When performing the hypothesis test, we make our decision according to a **decision rule** (also called the **critical region**), which tells us when to reject the null hypothesis. We have a (100α)% error rate of making the incorrect decision of rejecting the null hypothesis when it is true. We call this the **level of significance of the test α**.

_p_-value

It is important that you find a definition of the p-value that makes sense to you. It is also very important to remember that we begin by assuming the null hypothesis is true to find this probability. Here are a few versions of its definition.

- The p-value is the probability of obtaining results like those of our data (or more extreme) given the null hypothesis is true. In probability notation we can write: P(Obtaining results like ours or more extreme | H_0 is True).
- We are asking, 'Assuming the null hypothesis is true, how rare is it to observe something as or more extreme than what I have found in my data?'
- We can also think of the p-value as the probability, assuming the null hypothesis is true, that the result we have seen is solely due to random error (or chance).

No matter how we look at it, a small probability gives us evidence that the null hypothesis is not true and thus our alternative is true – which is usually what we are hoping to show.

How small must the p-value be?

To answer this question, we need to investigate the ramifications of our decision. So, as we discussed earlier, our decision is to reject or not to reject the null hypothesis. Like any situation, where a decision has to be made, we are open to make a mistake. If we reject the null hypothesis based on sample data, it could well be that this data was so unrepresentative that we were misled to reject the hypothesis. If we fail to reject the hypothesis, it could be that the sample belongs to a population whose mean is close to 8000, for example 8200, but not

8000. To demonstrate this, see Figure 13.4. We receive a sample mean of 8150. The probability that a sample has mean 8150 or more when the population has a mean of 8000 is $P(\bar{x} > 8150) \approx 0.227$

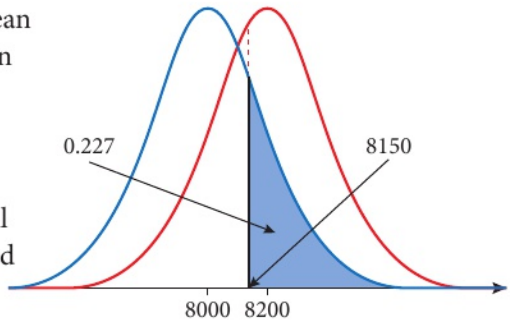


Figure 13.4 Area for type II error

Even though the sample belongs to a population whose mean is 8200, the chance of having a sample with this mean from a population having our hypothesised mean is 22.7%. This could well lead us to conclude that the sample is consistent with the model and we end up making the error of not rejecting the null hypothesis.

So, what types of errors may we end up committing?

When we perform a hypothesis test, we are faced with two types of error:

- if the null hypothesis is true, but we end up rejecting it – a type I error
- if the null hypothesis is false, but we do not reject it – a type II error.

Table 13.1 helps us keep track of our decision.

Our decision	The truth: H_0 True	The truth: H_1 True
Do not reject H_0	Correct decision	Type II Error
Reject H_0	Type I Error	Correct decision

Table 13.1 Type I and type II errors

So, in general, the decision of how small we want the p-value to be depends on how high a probability of type I error is acceptable. In the example of the light bulbs, committing a type I error means that the life of the bulbs is in fact 8000 hours but we end up saying that it is higher. The cost of our decision would be to cause the company to spend more money to produce a new line which is no better than the old one. Management, of course, wants to minimise the chance of this happening.

Type II error in this example is to conclude that you don't have evidence to say that the new light bulbs have a longer life, when they actually do. The consequence for this decision is to deprive the company of benefiting from the new innovation.

When the p-value is small, it indicates that our sample data are unusual given H_0. If our data are unusual enough, then we cannot assume that this could have occurred only by chance. Since the data were recorded, then something must be incorrect. All we can do is to reject the null hypothesis.

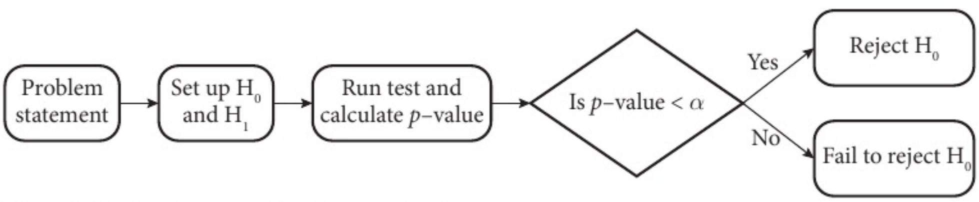


Figure 13.5 Steps for conducting a hypothesis test

As discussed earlier, there is a large resemblance between trial verdicts and hypothesis tests, as summarised in the table.

Hypothesis testing analogy to court verdicts

Verdict	The truth: Innocent	The truth: Guilty
Innocent	Correct decision	Error
Guilty	Error	Correct decision

Decision	The truth: H_0 True	The truth: H_1 True
Do not reject H_0	Correct decision	Type II Error
Reject H_0	Type I Error	Correct decision

But how unusual is 'unusual'? How small must the p-value be?

We can define unusual events arbitrarily by setting a limit for our p-value. If our p-value falls below that point, we will reject H_0. We call such results statistically significant.

The α-level is also called the **significance level** or **level of significance**. When we reject a hypothesis, we say that it was rejected at the k% level of significance, where $k = 1, 5, 10$, or any other number. The p-value is also called the **observed significance**.

A **statistically significant** result in hypothesis testing can be interpreted as a significantly rare event that will convince us to reject H_0.
The limit is called an **alpha level** (α-level). Common α-levels are 0.01, 0.05, and 0.10.
A statistician has to consider the alpha level carefully based on the situation. For example, if you are testing a hypothesis about the safety of a brake system in cars, you may want the α-level extremely low. If you are testing whether students use the school bus or not, you might be content with $\alpha = 0.10$. The level used mostly is $\alpha = 0.05$

Using hypothesis testing for comparing means of two populations

In this chapter, we will be using the t-test of hypothesis for the mean or difference of means of two populations. There are a few situations and different forms of the t-test, but we will focus only on one situation explained in Example 13.2 and beyond.

You will not be required to perform calculations manually, so learn how to use your GDC. Every GDC has its specific instructions.

Example 13.2

A comparison of the amount of nicotine that could be inhaled from cigarettes of the same brand differs between countries. There is some suspicion that local cigarette companies boosted their cigarettes' nicotine content to maintain or increase present addictive levels. A study of 40 randomly chosen cigarettes was conducted to investigate this issue between two European countries. The recorded average level of nicotine content in cigarettes in both countries is given below:

Country A: 0.85 mg/cigarette with a standard deviation of 0.2 mg/cigarette.

Country B: 1.30 mg/cigarette with a standard deviation of 0.18 mg/cigarette.

Does the nicotine content in the two countries differ? Test at the 5% level of significance.

There are a few tests for comparing means of populations, but in this book we limit our discussion to only one type of 2-sample test: independent samples with equal variances, also known as **pooled 2-sample *t*-test**.

Solution

Here we are testing

$H_0 : \mu_A = \mu_B$ against $H_1 : \mu_A \neq \mu_B$

An alternative formulation of the hypotheses is: $H_0 : \mu_A - \mu_B = 0$ against $H_1 : \mu_A - \mu_B \neq 0$.

This is a 2-sample t-test.

Carry out the calculations using a GDC.

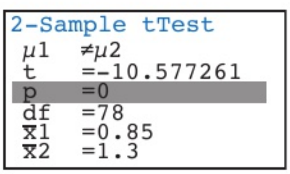


Figure 13.6 GDC solution to Example 13.2

The p-value for this test is zero, indicating that we have compelling evidence against H_0, and thus we reject and conclude that there is evidence that the nicotine content for the same brand of cigarettes differs among the two countries.

In Example 13.2, we have a 2-tail test since $H_1: \mu_A \neq \mu_B$

Example 13.3

Some investment funds can be purchased directly from banks whereas others must be purchased through brokers, who charge a fee for the service. There is suspicion that investors can do better by buying directly.

Researchers randomly sampled the annual returns from funds purchased directly and those purchased through brokers. The data given here are summaries of returns on investment after deducting all relevant fees:

Direct: $\bar{x}_1 = 6.63$, $s_1 = 6.12$; brokers: $\bar{x}_2 = 4.96$, $s_2 = 6.58$

Test, at the 5% level of significance, if the suspicion that investors can do better by buying directly is correct.

Solution

We are testing the following hypotheses:

$H_0: \mu_1 = \mu_2$ against $H_1: \mu_1 > \mu_2$

Alternatively: $H_0: \mu_1 - \mu_2 = 0$ against $H_1: \mu_1 - \mu_2 > 0$

This is a 2-sample t-test with pooled variance. It is also an upper-tail test.

The p-value of this test is 9.6%. That is, if we consider that buying direct or through brokers does not make a difference to return on investment (difference $= 0$), there is a 9.6% chance that we would get samples whose means are as far apart as our sample means. Thus, we do not have enough evidence against the null hypothesis and we do not reject it.

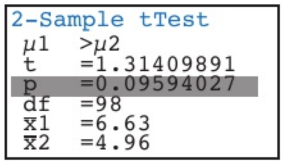


Figure 13.7 GDC solution to Example 13.3

Note that we do not say we accept the null hypothesis. We only fail to reject it.

An upper-tail test may as well be looked at as a lower-tail test. This happens if you change the order in which you write the hypothesis. That is, in Example 13.3, if we consider the purchase by brokers to be the first, then our hypotheses are as follows:

$H_0: \mu_1 = \mu_2$, against $H_1: \mu_1 < \mu_2$ and the output is given in the screenshot in Figure 13.8. Notice that the p-value is the same in both.

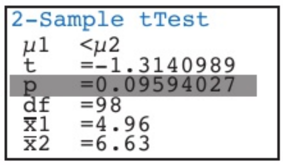


Figure 13.8 Note that the p-value is the same

Example 13.4

A company that specialises in job-placement services is interested in finding out if salaries of recent MA graduates in marketing are less than salaries of finance graduates. Random samples of 25 each from both subjects were collected and the data is as follows.

Marketing: Mean salary = US\$60,423, standard deviation = US\$16,193

Finance: Mean salary = US\$65,624, standard deviation = US\$18,985

Test at the 5% level of significance.

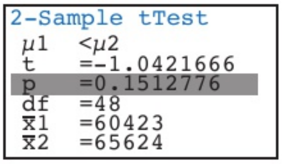


Figure 13.9 GDC solution to Example 13.4

Solution

The hypotheses are:

$H_0 : \mu_{marketing} = \mu_{finance}$ against $H_1 : \mu_{\text{marketing}} < \mu_{\text{finance}}$

Next, we compute the p-value $= 0.151$. The screenshot shows the GDC output.

Decision: Since $0.151 > 0.05$, i.e., the p-value is larger than the level of significance, α, we fail to reject the null hypothesis. Thus, we do not have enough evidence to conclude that marketing graduates earn less than finance graduates.

Hypothesis testing using critical values

In Example 13.4 we used the p-value approach to run the test. However, that is not the only way to run such a test.

We already know the level of significance, $\alpha = 5\%$, which can be used to set up what we call a **critical region**.

Our H_0 states that there is no difference between the means of the two populations. That is $H_0 : \mu_{\text{marketing}} - \mu_{\text{finance}} = 0$

However, $\bar{x}_{\text{marketing}} - \bar{x}_{\text{finance}} = 60423 - 65624 = -5201$

Our question now is: Is -5201 large enough to convince us that marketing graduates earn less?

Using the p-value approach, we know that the answer is no because it may be that, by chance, we could get two samples whose means differ by that amount. The chance such a difference can happen is 15.1%, which is larger than the acceptable level of 5%.

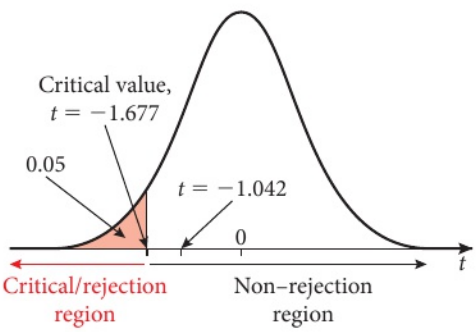


Figure 13.10 Hypothesis testing using critical values

Alternatively, the situation to start with can be shown on a diagram (Figure 13.10). The ***t*-value**, which we call the **critical number**, is a number on the t-distribution PDF that leaves an area equal to α below it in this case. (It can be to the right in an upper-tail test.)

We mark the region to the left of this point as the rejection region, and if the t-value of the difference of means lies in that region, this would mean that the p-value is less than α, and hence we reject the null hypothesis. The t-value of the difference is given in the GDC output to be -1.042 and lies in the non-rejection region. Thus, we do not reject the null hypothesis.

The calculations involved are beyond the scope of the course, but we are demonstrating it here because the concept is required in the next section.

You might notice that the results of the tests using the p-value approach or the critical values approach do not contradict each other. In most cases, it is advisable that you use the p-value approach since you can read the results directly from your GDC and then draw your conclusion. The other advantage is that when you reject or do not reject a hypothesis, you know your exact observed significance. If you apply a critical value approach, you must be prepared to settle for a flat reject/do not reject decision.

Example 13.5

A national magazine in the US claims that the average college student watches less television than the general public. The magazine collected data from a sample of 1200 people chosen at random from the non-college public. The average was 29.4 hours per week, with a standard deviation of 2.1 hours. A random sample of 900 college students has a mean of 28.9 hours with a standard deviation of 1.8 hours. Is there enough evidence to support the claim at the 1% level of significance?

Solution

To test the claim, we form the following hypotheses:

$H_0: \mu_{\text{college}} = \mu_{\text{public}}$ against $H_1: \mu_{\text{college}} < \mu_{\text{public}}$

Next, we run a 2-sample t-test with GDC output s hown in Figure 13.11.

Since the p-value $= 5.57 \times 10^{-9} < 0.01$, we have evidence to reject the null hypothesis. That is the magazine's claim is supported by our hypothesis testing at the 1% level of significance.

```
2-Sample tTest
 μ1   <μ2
 t    =-5.7353349
 p    =5.5709E-09
 df   =2098
 x̄1   =28.9
 x̄2   =29.4
```

Figure 13.11 GDC solution to Example 13.5

When we are running the 2-sample t-test, in this book we assume that the two populations in question are approximately normally distributed. If the populations are not normally distributed the t-test will not be valid.

Summary of hypothesis-testing terminology

- **Null hypothesis (H_0):** A maintained hypothesis that is held to be true unless sufficient evidence to the contrary is obtained. Should contain '=' in its statement.
- **Alternative hypothesis (H_1):** A hypothesis against which the null hypothesis is tested and which will be held to be true if the null is held false.
  - **One-tailed alternative:** An alternative hypothesis involving all possible values of a population parameter on either one side or the other of (that is, either greater than or less than) the value specified by a simple null hypothesis. Should contain either '<' or '>'.
  - **Two-tailed alternative:** An alternative hypothesis involving all possible values of a population parameter other than the value specified by a simple null hypothesis. Should contain '≠'.
- **Hypothesis test decisions:** A decision rule is formulated, leading the investigator to either reject or fail to reject the null hypothesis on the basis of sample evidence. (Decisions or decision rules are often called the critical region of the test and tell you when to reject a null hypothesis.) Whenever a decision is made, there will be a possibility of an error:
  - **Type I error:** The rejection of a true null hypothesis.
  - **Type II error:** The non-rejection of a false null hypothesis.
- **Significance level:** The probability of rejecting a null hypothesis that is true. This probability is sometimes expressed as a percentage, so a test of significance level a is referred to as a 100α%-level test.

Example 13.6

Atkins: Eat Right Not Less.

This is what this method of low-carbohydrate low-sugar diet advertises. Scientists at a university selected a random sample of 32 subjects from their local population of obese adults. They randomly assigned 17 to the Atkins diet and 15 to a conventional diet. Does this experiment show, at the 5% level of significance, that the Atkins diet leads to a loss in mass of 5 more kilograms in a year than a conventional diet?

Table 13.2 gives the mass loss (positive) or gain (negative) for each subject.

Atkins	Conventional
17.3	11.7
39.3	14.4
26.5	−7.2
10.3	1.9
27.3	10.7
10.2	−4.8
0.5	−1.3
−17.3	20.9
14.3	23.5
15.7	16
14	18.2
8.1	9.1
14.4	20.1
31.7	−7.6
19.4	0.6
25.2	
−1.2	

Table 13.2 Data for Example 13.6

Solution

We assume that the mass loss in both cases is normally distributed. Note here that the hypotheses deal with a difference more than 5 kg and not zero as in other examples. Also, the status quo would clearly include mass losses less than 5 kg and not only 5 kg. So, the hypotheses are:

$H_0: \mu_A - \mu_C \leqslant 5$; $H_0: \mu_A - \mu_C > 5$

We can use a GDC, where we enter the data into cells, rather than using summary statistics as before, or we can use a spreadsheet. Both are shown here. However, with your GDC, most probably, you need to add 5 to the list containing your second sample as has been done here. Note that

List 3 = List 2 + 5

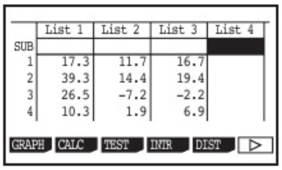


Figure 13.12 You will most probably need to add 5 to the list containing your second sample

After adjusting your lists, you can run the usual 2-sample t-test.

p-value $= 0.355 > 0.05$

We do not reject H_0. We do not have statistical evidence to reject the null hypothesis. We cannot claim that the Atkins leads to a greater loss in mass.

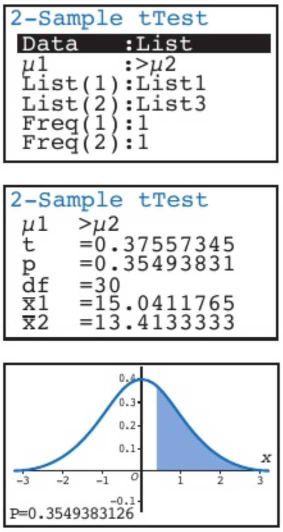


Figure 13.13 Running the usual 2-sample t-test

A spreadsheet will give you the same results, as shown below.

t-Test: Two-Sample Assuming Equal Variances		
	Atkins	Conventional
Mean	15.04117647	8.413333333
Variance	180.7900735	114.1712381
Observations	17	15
Pooled Variance	149.7012837	
Hypothesised Mean Difference	5	
df	30	
t Stat	0.375573457	
P(T,<=t) one-tail	0.354938313	
t Critical one-tail	1.697260887	
P(T,<=t) two-tail	0.709876625	
t Critical two-tail	2.042272456	

The methods in Example 13.6 may not be part of your examination but could be useful for your internal assessment work.

Exercise 13.1

1. In each case, test the hypothesis at the required level of significance. Assume that all the populations are normally distributed.

(a) $H_0: \mu_1 \leqslant \mu_2$ $\quad H_1: \mu_1 > \mu_2$ $\quad \alpha = 0.05$
Samples: $\bar{x}_1 = 412$, $s_1 = 128$, $n_1 = 150$; $\bar{x}_2 = 405$, $s_2 = 54$, $n_2 = 150$

(b) $H_0: \mu_1 = \mu_2$ $\quad H_1: \mu_1 \neq \mu_2$ $\quad \alpha = 0.05$
Samples: $\bar{x}_1 = 74$, $s_1 = 18$, $n_1 = 12$; $\bar{x}_2 = 71$, $s_2 = 16$, $n_2 = 12$

(c) $H_0: \mu_1 - \mu_2 = 0$ $\quad H_1: \mu_1 - \mu_2 < 0$ $\quad \alpha = 0.10$
Samples: $\bar{x}_1 = 51.3$, $s_1^2 = 52$, $n_1 = 31$; $\bar{x}_2 = 53.2$, $s_2^2 = 60$, $n_2 = 32$

2. In parts of southern Germany, rabbits are a major nuisance to farmers, frequently damaging crops, gardens, and landscaping. A local group arranges a test of two of the leading rabbit repellents. Fifty-six unfenced plots in areas having high concentrations of rabbits are used for the test. 29 plots are chosen at random to receive repellent A, and the other 27 receive repellent B. For each of the 56 plots, the time elapsed between application of the repellent and the appearance in the plot of the first rabbit is recorded.

   For repellent A, the mean time is 101 hours with a standard deviation of 15 hours. For repellent B, the mean time is 92 hours and standard deviation of 10 hours. Assume that the two populations of elapsed times have normal distributions.

   Test at the 2% significance level whether the mean elapsed times for repellents A and B are different.

   (a) Write down the hypotheses for the test

   (b) State two assumptions you make and the decision rule.

   (c) Carry out the test and write down your decision with reasons.

3. A large department store conducts an inventory check regularly and calculates losses from theft. They would like to reduce these losses and are considering two methods. The first is to hire a security guard, and the second is to install cameras. To help with the decision, they ran an experiment where in the first 6 months they hired a guard and the second 6 months they installed cameras. The monthly losses were recorded and are listed in Table 13.3. The management decided that since cameras are cheaper than hiring a guard, they would install cameras unless there was evidence to infer that the guard was better. What should they do? Use an α-level of 10%.

   (a) Write down the hypotheses for the test

   (b) State two assumptions you make and the decision rule.

   (c) Carry out the test and write down your decision with reasons.

Guard	Cameras
355	486
284	303
401	270
398	386
477	411
254	435

Table 13.3 Data for question 3

4. Although cruises are typically associated with elderly people, it appears that there is an increase in the number of younger people choosing cruises as their vacation. To determine their marketing strategy, a cruise line company sampled passengers last summer and determined their ages. They then compared the results to another survey that was done three years ago. Results are listed below:

   Old survey: $n = 125$, average age $= 59.8$ years,
   standard deviation $= 7.02$ years.

   New survey: $n = 159$, average age $= 57.4$ years,
   standard deviation $= 6.99$ years.

   What can the company conclude? Use a 10% level of significance.

   (a) Write down the hypotheses for the test.

   (b) State two assumptions you make and the decision rule.

   (c) Carry out the test and write down your decision with reasons.

13.2 Goodness-of-fit (GOF) test and test for independence

Goodness-of-fit

Between 2000 and 2012, there was a belief among some teachers that the distribution of IB grades for maths HL/SL was as follows:

25% of candidates earn grade 3 or below, distributed as 1%, 9%, and 15% for 1, 2, and 3 respectively; 75% for grades of 4 and above such that 20% had a grade of 4, 25% a grade of 5, 20% a grade of 6 and only 10% a grade of 7.

After a new syllabus was introduced in 2012, we wanted to check whether the distribution of grades was the same. After May 2016's session, a random sample of scores of 1800 candidates was collected. The results are shown in Table 13.4.

Grade	1	2	3	4	5	6	7
Observed number	13	146	295	382	434	382	148

Table 13.4 Random sample of maths scores

Does the observed data provide evidence of change in the distribution?

To be able to answer the question, we need to compare the observed data to what we would have expected if the model were true. That is, we compute the number of students we would expect to see in each grade level if the null model were true.

For example, in case of a score of 1, and if the 1% share were true, then we would expect $1800 \times 1\% = 1800 \times \frac{1}{100} = 18$

The 9% share for a score of 2 is $1800 \times 9\% = 1800 \times \frac{9}{100} = 162$

Table 13.5 shows the expected number of candidates for each grade level from applying the model percentages to the 1800 included in our sample.

Grade	1	2	3	4	5	6	7
Expected number	18	162	270	360	450	360	180

Table 13.5 Expected number of candidates for each grade

Of course, we should expect some variation. We should not expect the percentages for each grade level to match. After all, our sample is a random sample! However, are these deviations large enough to convince us that the number of students observed is different from those expected in each grade level?

To address this question, we test the table's **goodness-of-fit** to the proposed model. Here fit refers to the expected counts if we assume the belief is true.

The hypotheses for the test are:

H_0: The observed distribution is the same as the expected distribution.

H_1: The observed distribution is not the same as the expected distribution.

The goodness-of-fit test examines the discrepancy between what we observe and what we expect and generates a measure of how significant that difference is. In other words, to come to a decision on whether the null model is reasonable, we look at the differences between the expected values from the model, f_e and the counts we observe, f_o. We denote the differences between the observed and expected counts, $(f_0 - f_e)$. Are these differences so large that they call the model into question, or could they have occurred from usual sampling variability?

Adding these differences will create a problem as negative differences will cancel out the positive ones, so we square the differences.

Hence, we will look at $(f_0 - f_e)^2$ for each cell. This way we can focus our attention on cells with large differences.

Also, since samples of larger sizes will give results generally larger than small samples, we need to make these differences relative. So we divide every squared difference by the relative frequency of each cell, i.e., $\dfrac{(f_0 - f_e)^2}{f_e}$

The test statistic used for this purpose is called the **chi-squared** (or **chi-square**) **statistic**. It is found by adding all the squared differences divided by their respective expected frequencies

$$\chi^2_{\text{calc}} = \sum_{\text{all cells}} \frac{(f_o - f_e)^2}{f_e}$$

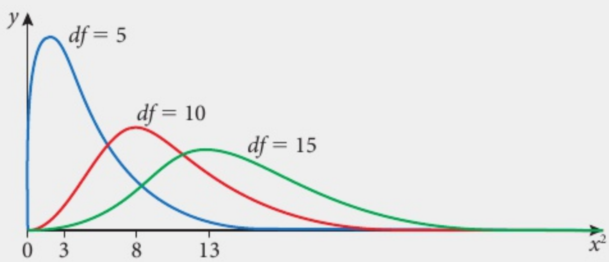
The decision will be to reject the null hypothesis if the chi-squared statistic is large. The critical value as always is determined by the level of significance and the number of degrees of freedom. The degrees of freedom for this test are $n - 1$, where n is the number of cells.

The reason for the square in deviations $(f_0 - f_e)^2$ is similar to the reason why we have $\sigma = \sqrt{\dfrac{\sum(x-\mu)^2}{n}}$ rather than simply $\dfrac{\sum(x-\mu)}{n}$

You are expected to use technology to find a p-value and the χ^2 statistic. The critical value will always be provided.

χ is the third last letter in the Greek alphabet and is pronounced chai or chi with a soft 'ch'.

The mathematics of the χ^2 distribution is beyond the IB syllabus and this book. It is worth mentioning that the chi-square distribution has only one parameter, called the degrees of freedom (df). The shape of a specific chi-square distribution depends on the number of degrees of freedom. The random variable χ^2 assumes non-negative values only. Hence, a chi-square distribution curve starts at the origin and lies entirely to the right of the vertical axis. The graph shows three chi-square distribution curves. They are for 5, 10, and 15 degrees of freedom, respectively.



The shape of a chi-square distribution curve is skewed for small degrees of freedom, and it changes drastically as the degrees of freedom increase. Eventually, for large degrees of freedom, the chi-square distribution curve looks like a normal distribution curve. The peak (or mode) of a chi-square distribution curve with 3 or more degrees of freedom is at $df - 2$. For instance, the peak of the chi-square distribution curve with $df = 5$ occurs at 3. The peak for the curve with $df = 10$ occurs at 8. Like all other continuous distribution curves, the total area under a chi-square distribution curve is 1.0.

To test a hypothesis using the chi-squared distribution:

1. Find the expected values. These come from the null hypothesis model. The expected frequency f_e for each cell is np, where n is the total number of observations and p is the proportion of the data you expect to be in this cell. For example, since we have 1800 students and the hypothesised proportion for a grade 7 is 10%, then our expected frequency is $1800 \times 0.10 = 180$ as you see in Table 13.5. (The result need not be an integer.)
2. Compute the differences (known as residuals): $(f_0 - f_e)$
3. Square the residuals: $(f_0 - f_e)^2$
4. Compute the cell contribution to the calculated chi-squared value: $\frac{(f_o - f_e)^2}{f_e}$ (also known as components)
5. Find the sum of the cell contributions: $\chi^2_{\text{calc}} = \sum_{\text{all cells}} \frac{(f_o - f_e)^2}{f_e}$
6. Find the number of degrees of freedom: $n - 1$, in this case.
7. Test the hypothesis. This can be done by either calculating the p-value or by constructing a critical region.

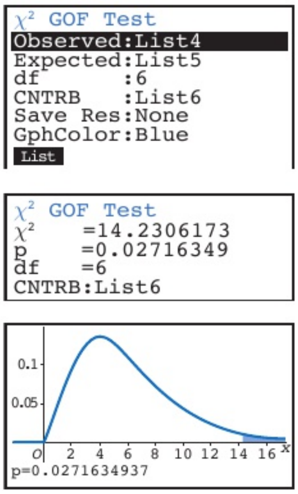


Figure 13.14 GDC output for chi squared test

You can carry out these calculations using a GDC, or a spreadsheet.

With p-value of 0.0272, we reject the null hypothesis. Thus, we can conclude that the distribution of grades after 2012 has changed.

The critical number for $df = 6$ and 5% significance is 12.59. Since the test statistic is 14.231, which is to the right of the critical number, the conclusion is, as expected, the same as before – reject H_0 (Figure 13.15).

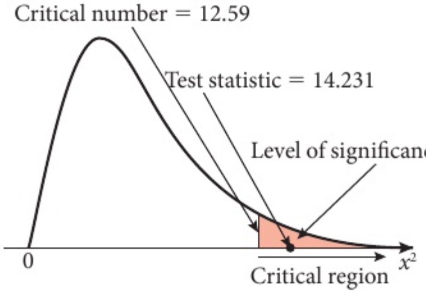


Figure 13.15 Test statistics in the critical region – reject H_0

The spreadsheet work is only meant for internal assessment work.

Grade	f_0	f_e	$(f_0 - f_e)$	$(f_0 - f_e)^2$	$\frac{(f_0 - f_e)^2}{f_e}$
1	13	18	−5	25	1.38888889
2	146	162	−16	256	1.58024691
⋮	⋮	⋮	⋮	⋮	⋮
4	148	180	−32	1024	5.68888889
	1800	1800	$x^2_{\text{calc}} = \sum_{\text{all cells}}$	$\frac{(f_0 - f_e)^2}{f_e}$	14.230617

Example 13.7

The population of the UK in 1996 had the following distribution.

Age group	0–15	16–64	65 and above
Percentage of population	20.7%	63.5%	15.8%

A sample of 1800 randomly chosen UK citizens in 2016 gave the following.

Age group	0–15	16–64	65 and above
Number of persons	340	1136	324

Did the population distribution change from 1996 to 2016? Test at the 5% level of significance.

Solution

Hypotheses:

H_0: The observed distribution of age in 2016 is the same as the expected distribution of 1996.

H_1: The observed distribution is not the same as the expected distribution.

Expected frequencies:

$1800 \times 0.207 = 372.6$, $1800 \times 0.635 = 1143$, and $1800 \times 0.158 = 284.4$

The GDC output is shown in Figure 13.16.

Since the p-value of $0.0149 < 0.05$, we reject H_0 and conclude that the evidence shows that the age distribution in 2016 is not the same as that in 1996.

Goodness of fit test – general considerations
Data for goodness of fit tests are organised in tables or lists. The cells of the tables represent counts that correspond to the different categories into which we divide our data set. In the example above, every grade level represented a category. So, data for the test must be counts (frequencies). The subjects counted should be a random sample.
There should be at least 5 observations in any individual cell. If this condition is not met, then you have to combine cells so that the condition is satisfied.
The chi-squared distribution used has, in general $(n - 1)$ degrees of freedom

```
χ² GOF Test
χ²     =8.40907496
p      =0.01492768
df     =2
CNTRB:List6
```

Figure 13.16 GDC output for Example 13.7

Test for independence: contingency table

Often, we would like to know whether two variables, each with several levels, are related. In such cases a chi-squared test is used. It is called a **test for independence**.

When we use the test, our hypotheses are in the following form:

H_0: The variables in question are independent – they are not related.

H_1: The variables in question are not independent – they are related.

Example 13.8

A researcher conjectures that seat belt usage, for drivers, is related to gender. She gathers data by interviewing a random sample of 216 car passengers and asking them about seat belt usage. The data has been recorded and organised in the table. Construct a chi-squared hypothesis test to determine if there is enough evidence to support the researcher's conjecture. Use a 5% level of significance.

	Seat belt use		
Gender	**Yes, all the time**	**Most of the time**	**Not at all**
Female	50	32	25
Male	40	26	43

Solution

Write down the hypotheses.

H_0: Seat belt usage is not related to gender.

- This agrees with the definition of the null hypothesis that states there is no significant difference between two parameters (numbers about a population).
- 'Independent of' is often used in place of 'not related to'.

H_1: Seat belt usage is related to gender.

- 'Related to' means that there is a relationship between seat belt usage and gender. If there is a relationship, one of the variables is dependent upon the other.
- 'Not independent of' is often used in place of 'is related to'.

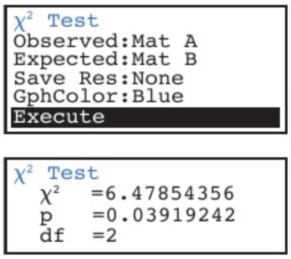


Figure 13.17 Entering the data in GDC matrix form

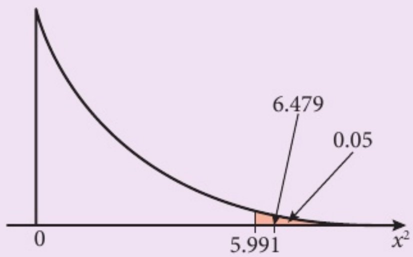
Like other tests, we either find the p-value or the test statistic, using a GDC. Unlike previous cases, the data should be entered into a GDC in matrix form.

Regardless of the GDC used, the output will mainly give the value of the test statistic, the p-value, and the degrees of freedom.

***p*-value approach:**
Since the p-value $= 0.0392 < 0.05$, we reject H_0 and conclude that there is evidence to suggest that seat belt usage may be related to gender.

Critical region approach:
We are given the critical number (How to find it is beyond the scope of your syllabus – it is given here for demonstration only!)

Since $\chi^2_{\text{statistic}} = 6.479 > \chi^2_{\text{critical}} = 5.991$, we also reject H_0.



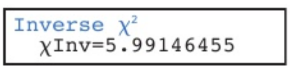


Figure 13.18 GDC output for finding critical values

In Example 13.9, we will carry out a test and show how the calculations are made. Again, you will not need to do the calculations manually.

Example 13.9

In a large school system, students in grades 9–12 were asked whether good grades, extracurricular activities, or esteem were most important to them. The students were divided in grade level groups. We are interested in whether there is a relationship between the preferred factor and grade level. Table 13.6 shows the results of a survey of 465 students chosen at random.

Test at the 1% level of significance, whether grade level and preferred factor are independent.

	Grade level			
Factor	9	10	11	12
Grades	35	30	54	55
Esteem	60	50	30	15
Activities	40	50	26	20

Table 13.6 Table for Example 13.9

This is known as a **contingency table**.

Solution

H_0: Grade level and preferred factor are not related.

H_1: Grade level and preferred factor are related.

Run a chi-squared test and calculate the p-value.

As mentioned earlier, the data should be entered into a GDC in matrix form.

Decision: As the p-value $= 3.5 \times 10^{-9} < 0.01$, we reject H_0 and can suggest that there is a connection between grade level and preferred factor among high school students.

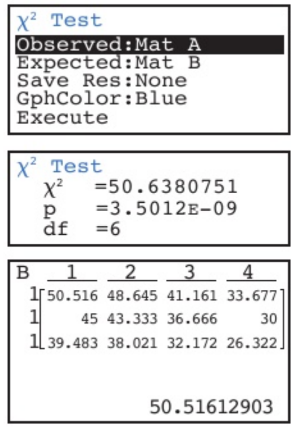


Figure 13.19 Entering the data in GDC matrix form

Calculations for the chi-squared test of independence.

Recall that the chi-squared test compares the observed frequencies, f_o with the expected frequencies, f_e for each cell using $\dfrac{(f_o - f_e)^2}{f_e}$

The collected data gives us the observed frequency f_o for each cell. Here is how we can calculate the expected frequency f_e for the cell.

	Grade level				
Factor	9	10	11	12	**Totals**
Grades	35	30	54	55	**174**
Esteem	60	50	30	15	**155**
Activities	40	50	26	20	**136**
Totals	**135**	**130**	**110**	**90**	**465**

Table 13.7 Calculating expected frequencies

Recall from Chapter 8, that if two events A and B are independent, then $P(A \cap B) = P(A)P(B)$

This is the key to calculating the expected frequencies. Assuming independence in H_0, every cell in the table is at the intersection of two levels, one from each. For example, the cell 35 in the grades row and 9 column is the observed frequency that a student in grade 9 cares more about grades. If grades were independent of grade level, then

$$P(\text{Grades} \cap 9) = P(\text{Grades})P(9) = \frac{174}{465} \cdot \frac{135}{465}$$

There are 174 among the whole group who say that grades are most important, and hence the probability of picking a grades student is $\dfrac{174}{465}$

Similarly there are 135 students in grade 9 in the group, so the probability of choosing a student in grade 9 is $\dfrac{135}{465}$

Now the expected frequency is np, and hence

$$f_{grades \cap 9} = \cancel{465} \cdot \frac{174}{\cancel{465}} \cdot \frac{135}{465} = \frac{174 \cdot 135}{465} = 50.516$$

Notice the first cell in the screen produced by a GDC in the screenshot in Figure 13.19. It is 50.516. This is the expected frequency for that cell. The calculation is done in a similar fashion for each of the other 11 cells.

This means, to find the expected frequency of a certain cell, multiply the row total by the column total and divide that product by the total number of observations in the whole sample. This is the matrix B produced by a GDC.

Calculating the expected frequency
The estimated expected number of observations in each cell, under H_0, is
$$f_e = \frac{RC}{n}$$
where R is the row total, and C is the column total. $R = 174$, $C = 135$, and $n = 465$ in Example 13.9.

Now, the chi-squared statistic is $\chi^2_{\text{calc}} = \sum_{\text{all cells}} \frac{(f_o - f_e)^2}{f_e}$ and is made up of the sum of all cells' contributions $\frac{(f_o - f_e)^2}{f_e}$ For example, to calculate the first cell's contribution to the statistic we have $\frac{(35 - 50.516)^2}{50.516} \approx 4.77$

Calculate the contribution from all the cells and add the resulting values to give $\chi^2_{\text{calc}} = 50.638$

The number of degrees of freedo m for this test is $(r - 1)(c - 1) = 2 \times 3 = 6$, where r is the number of rows and c is the number of columns. The critical number is 12.59 in this case. This is the number in the upper tail of the chi-squared distribution with 6 degrees of freedom which marks the 5% area.

Since 50.638 is to the right of 12.59, we reject the null hypothesis.

Alternatively, $\mathrm{P}\left(\chi^2_6 \geqslant 50.638\right) = 3.5 \cdot 10^{-9} \approx 0$, and we also reject the null hypothesis.

As mentioned earlier, this discussion is meant to help you understand what the chi-squared test is about. All calculations are done with a GDC, and critical numbers will be provided when needed.

Exercise 13.2

1. 180 spectators at a swimming championship were asked which, of four swimming styles, was the one they preferred to watch.

The results of their responses are shown in the table.

Swimming style	Male	Female
Freestyle	20	15
Butterfly	20	30
Backstroke	10	35
Breaststroke	10	40

A χ^2 test was conducted at the 5% significance level.

(a) Write down the null hypothesis for this test.

(b) Write down the number of degrees of freedom.

(c) Write down the value of χ^2_{calc}.

The critical value, at the 5% significance level, is 7.815.

(d) State, giving a reason, the conclusion to the test.

2. An agricultural cooperative uses three brands of fertiliser, *A*, *B* and *C*, on 120 different crops. The crop yields are classified as high, medium or low. The data collected are organised in the table.

	Fertiliser			Total
	A	*B*	*C*	
High yield	10	8	12	**30**
Medium yield	24	14	12	**50**
Low yield	16	8	16	**40**
Total	**50**	**30**	**40**	**120**

The agricultural cooperative decides to conduct a chi-squared test at the 1% significance level using the data.

(a) State the null hypothesis, H_0 for the test.

(b) Write down the number of degrees of freedom. (Hint: find the inverse chi-square.)

(c) Write down the critical value for the test.

(d) Show that the expected number of medium yield crops using fertiliser *C* is 17, correct to the nearest integer.

(e) Use your graphic display calculator to find for the data:

(i) the χ^2 calculated value **(ii)** the p-value.

(f) State the conclusion of the test. Give a reason for your decision.

3. In an effort to develop new software that replaces an existing office package, a new company ran a survey for 5958 randomly chosen potential users. One of the questions was 'How important is it for you that the new product has an email component similar to the one in the existing package?'

The options for the response were given on a scale of 1–4 as follows:
1 – not at all important, 2 – somewhat important,
3 – important, and 4 – extremely important.

The distribution of the respondents according to their age group are given in the table.

	age $<$ 20	20 $\leqslant$ age $<$ 40	40 $\leqslant$ age $<$ 60	age $\geqslant$ 60
4	595	526	485	442
3	427	459	447	381
2	408	392	407	391
1	143	154	148	153

The company decides to conduct a chi-squared test at the 5% significance level using the data.

(a) State the hypotheses for the test.

(b) Find the expected frequency in the cell corresponding to response 3 and $40 \leqslant \text{age} < 60$

(c) Find the number of degrees of freedom for this test.

(d) The critical value is 16.92. Write down the decision rule to reject the null hypothesis.

(e) Write down the value of the test statistic χ^2_{calc} and decide whether you reject H_0 giving your reasons.

(f) Write down the p-value and decide whether you reject H_0 giving your reasons.

4. The observed frequencies of an experiment with 300 trials and 5 cells is shown in the table. The model to be tested is:

$H_0: p_1 = 0.10, p_2 = 0.20, p_3 = 0.30, p_4 = 0.20, p_5 = 0.20$

Cell	1	2	3	4	5
Frequency	24	64	84	72	56

(a) Write down the alternative hypothesis.

(b) Find the expected frequencies.

(c) Test the goodness-of-fit of the data to the model at the 5% level of significance.

5. In a number of pharmaceutical studies, volunteers who take placebos (but are told they have taken a cold remedy) report the following side effects:

Side effect	**Headache**	**Drowsiness**	**Stomach upset**	**No side effect**
Cold (Percentage)	5%	7%	4%	84%
Anti-inflammatory	19	23	14	194

A random sample of 250 people were given a placebo (but thought they had taken an anti-inflammatory) and reported whether they had experienced each of the side effects. Their results in numbers are given in the third row of the table.

(a) Write down the hypotheses of the test.

(b) Find the expected frequencies.

(c) Is there evidence, at the 5% level of significance, that the reported side-effects of the anti-inflammatory group differ from those with a cold remedy?

6. To determine whether a single dice is fair, the dice was rolled 600 times. The table shows the number of times each of the faces showed.

Number	1	2	3	4	5	6
Frequency	102	107	101	84	92	114

(a) Write down the hypotheses for this test.

(b) Find the expected frequencies.

(c) Given that the critical value is 11.07, test whether the dice is fair at the 5% level of significance.

Chapter 13 practice questions

1. The producer of a TV dancing show asked a group of 150 viewers their age and the type of Latin dance they preferred. The types of Latin dances in the show were Argentine tango, samba, rumba and cha-cha-cha. The data obtained are shown in the table.

	Dance			
	Argentine tango	**Samba**	**Rumba**	**Cha-cha-cha**
20 years old and younger	35	23	12	10
Older than 20 years old	20	17	18	15

A χ^2 test was carried out, at the 5% significance level.

(a) Write down the null hypothesis for this test.

(b) Write down the observed number of viewers who preferred rumba and were older than 20 years old.

(c) Use your graphic display calculator to find the p-value for this test.

The producer claims that the type of Latin dance a viewer preferred is independent of their age.

(d) Decide whether this claim is justified. Give a reason for your decision.

2. Four parties contested an election held two years ago. Conservatives captured 29% of the vote, liberals 41% of the vote, while the greens and the extreme right parties split the rest among themselves equally. A survey of 1500 voters this year yielded the following: 410 conservatives, 590 liberals, 270 greens, and 230 went to the extreme right party.

Has the voter support changed since the elections? Test at the 10% level of significance.

(a) If the voter support has not changed, what is the expected number of voters for each party?

(b) Write down the hypotheses for this test.

(c) Perform the test in question.

3. In a debate on voting, a survey was conducted. The survey asked people's opinion on whether the minimum voting age should be reduced to 16 years of age. The results are shown as follows.

	Age 18–25	**Age 26–40**	**Age 41+**	**Total**
Oppose the reduction	12	20	48	80
Favour the reduction	18	15	17	50
Total	30	35	65	130

A χ^2 test at the 1% significance level was conducted. The χ^2 critical value of the test is 9.21

(a) State:

(i) H_0, the null hypothesis for the test

(ii) H_1, the alternative hypothesis for the test.

(b) Write down the number of degrees of freedom.

(c) Show that the expected frequency of those between the ages of 26 and 40 who oppose the reduction in the voting age is 21.5, correct to three significant figures.

(d) Find:

(i) the χ^2 statistic

(ii) the associated p-value for the test.

(e) Determine, giving a reason, whether H_0 should be rejected.

4. It has been estimated that employee absenteeism costs US companies more than $100 billion per year. To understand the trend of absenteeism among its employees, a large corporation recorded the weekdays during which individuals in a random sample of 724 absentees were away over the past 6 months, as shown in the table.

Do these data suggest that absenteeism is higher on some days of the week than on others? Test at the 10% level of significance.

Day of the week	Monday	Tuesday	Wednesday	Thursday	Friday
Number absent	174	124	142	136	148

5. Many restaurants use a device that allows credit card users to pay with their credit card at the table. The device also allows the user to specify an amount for a waiter tip. In an experiment to see how it works, a random sample of credit card users was drawn. Some paid with the device and some paid in cash. The percentage left as a tip was recorded and listed below.

Can we infer that users of the device leave larger tips? Test at the 10% level of significance.

Cash	10.3	15.2	13.0	9.9	12.1	13.4	12.2	14.9	13.2	12.0	
Device	13.6	15.7	12.9	13.2	12.9	13.4	12.1	13.9	15.7	15.4	17.4

(a) What type of test is this? Give two assumptions you are making.

(b) Write down your hypotheses for the test.

(c) Draw your conclusion and give a reason for it.

6. The manager of a travel agency surveyed 1200 travellers. She wanted to find out whether there was a relationship between a traveller's age and their preferred destination. The travellers were asked to complete the survey shown.

My age is:

25 or younger	26–40	41–60	61 or older

My preferred destination is:

New York	Tokyo	Melbourne	Dubai	Marrakech

A χ^2 test was carried out, at the 5% significance level, on the data collected.

(a) Write down the null hypothesis.

(b) Find the number of degrees of freedom.

The critical value of this χ^2 test is 21.026

(c) Use this information to write down the values of the χ^2 statistic for which the null hypothesis is rejected.

From the travellers taking part in the survey, 285 were 61 years or older and 420 preferred Tokyo.

(d) Calculate the expected number of travellers who preferred Tokyo and were 61 years or older.

7. A survey was conducted among a random sample of people about their favourite TV show. People were classified by gender and by category of favourite TV show (sports, documentary, news and reality TV).

The results are shown in the contingency table.

	Sports	Documentary	News	Reality TV	Total
Male	20	24	32	11	87
Female	18	30	20	25	93
Total	38	54	52	36	180

(a) Find the expected number of females whose favourite TV shows are documentaries.

A χ^2 test at the 5% significance level is used to determine whether TV show category is independent of gender.

(b) Write down the p-value for the test.

(c) State the conclusion of the test. Give a reason for your answer.

8. A psychologist examined whether the content of TV shows influenced the ability of viewers to recall brand names of items featured in commercials. The researcher randomly assigned volunteers to watch one of two programs, each containing the same nine commercials. One of the programs had violent content and the other had neutral content. After the show ended, the subjects were asked to recall the brands of products that were advertised. Results are listed below.

Violent: mean = 3.18, standard deviation = 1.86, $n = 108$

Neutral: mean = 4.16, standard deviation = 1.75, $n = 108$

Do these results indicate that viewer memory for commercials may differ depending on program content? Test at the 10% level of significance.

14 Bivariate analysis

14 Bivariate analysis

Learning objectives

By the end of this chapter, you should be familiar with...

- the key features of a scatter diagram and estimating a best-fit line by eye
- the difference between causation and correlation
- calculating (using your GDC) and interpreting Pearson's product-moment correlation coefficient for linear associations (r)
- calculating (using your GDC) and interpreting Spearman's rank-order correlation coefficient r_s
- the limitations of and differences between Pearson's r and Spearman's r_s
- how to find, use, and interpret least-square regression models and piecewise linear models.

We are often interested in whether one variable produces a change in another, and, if so, how much? For example:

- Do students who study more than others get better test grades?
- Do heavier cars have worse fuel economy?
- Do plants that are fertilised more grow bigger or faster?

All of these questions involve investigating the relationship between two variables, which is called **bivariate statistics**. We will start by examining data graphically, and then examine two quantitative methods to measure the relationship between two variables.

Deciding which variable is the explanatory or independent variable is largely up to us. Does the number of hours a student studies explain their test grade, or does the test grade explain the number of hours spent studying? The second statement seems rather silly. In experiments, the explanatory variable is often the variable that is directly manipulated, and then we observe a response in the response variable. If both variables are observed (neither is directly manipulated or controlled), then it is mostly about the question we want to ask or the model we want to create: does one variable explain the other? Do we want a model for a in terms of b or b in terms of a?

14.1 Scatter diagrams

Suppose we are interested in whether students who spend more time studying for a test get better scores on that test. We collect data on a class of 10 mathematics students and obtain the data shown in Table 14.1.

Student	Tim	Joon	Jim	Kevin	Steve	Niki	Henry	Anton	Cindy	Lukas
Hours	4	4.5	6	3.5	3	5	5.5	6.5	7	6.5
Grade	65	80	83	61	55	79	85	79	92	95

Table 14.1 Mathematics students' data

Since we are analysing whether more studying appears to produce higher test grades, we say

- 'hours' is the **explanatory** or **independent variable**
- 'grade' is the **response** or **dependent variable**
- the students whose time and grades are recorded are the **subjects** of this study/experiment.

A **response variable** measures an outcome of a study. An **explanatory variable** explains the changes in the response variable.
The response variable is traditionally known as the dependent variable; the explanatory variable is traditionally known as the independent variable.

To begin examining the relationship between hours spent studying and test grades, we draw a scatter diagram, where we plot each data point.

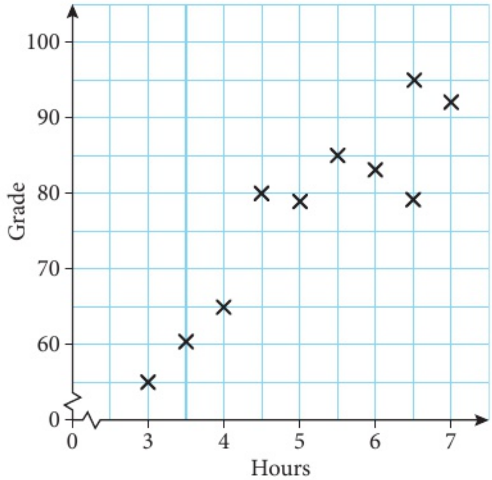


Figure 14.1 Scatter diagram

It appears that, in general, as the number of hours spent studying increases, so does the grade. Therefore, we say that the two variables are associated.

When we examine a scatter diagram, we look for form, direction, strength, and unusual features such as outliers or clusters.

Two variables measured on the same subjects are **associated** if specific values of one variable tend to occur in connection with particular values of the other variable.

Form

Form refers to the shape of the graph. This is often the first step to determining a likely model. In general we will start by deciding if the form is linear or non-linear.

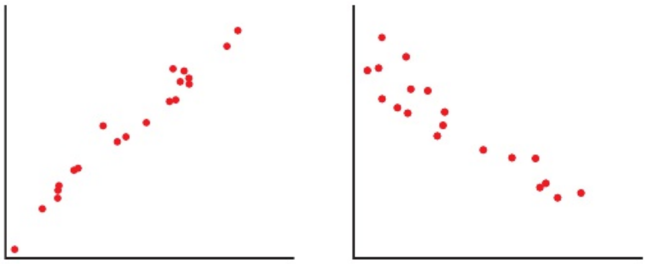


Figure 14.2 Linear forms

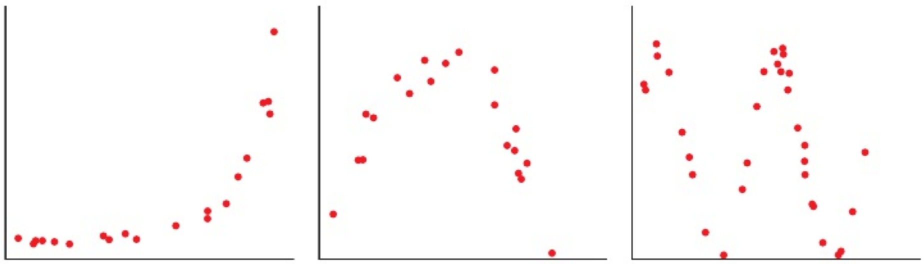


Figure 14.3 Non-linear forms

Direction

> You will see that many sources refer to correlation instead of association when describing direction and strength. We will use correlation primarily when discussing correlation coefficients, which are quantitative measures of specific types of association.

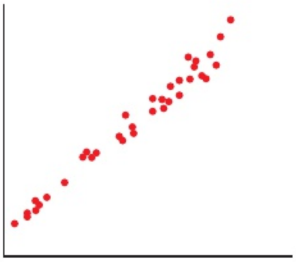


Figure 14.4 Positive association

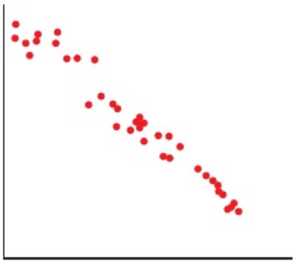


Figure 14.5 Negative association

Direction tells us how the response variable changes in relation to the explanatory variable. If the response variable increases along with the explanatory variable, then the association is positive (Figure 14.4). If the response variable decreases while the explanatory variable increases, then the association is negative (Figure 14.5). Graphically, this is analogous to the slope (gradient).

If the response variable does not change as the explanatory variable changes, or if the response variable changes in ways that appear random, then there is zero or no association (Figure 14.6).

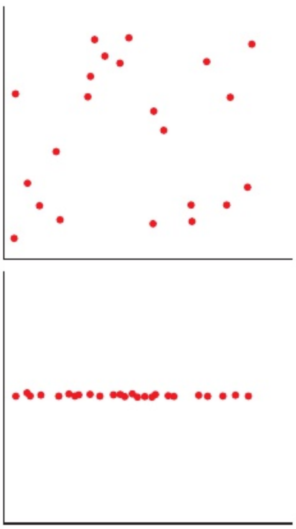
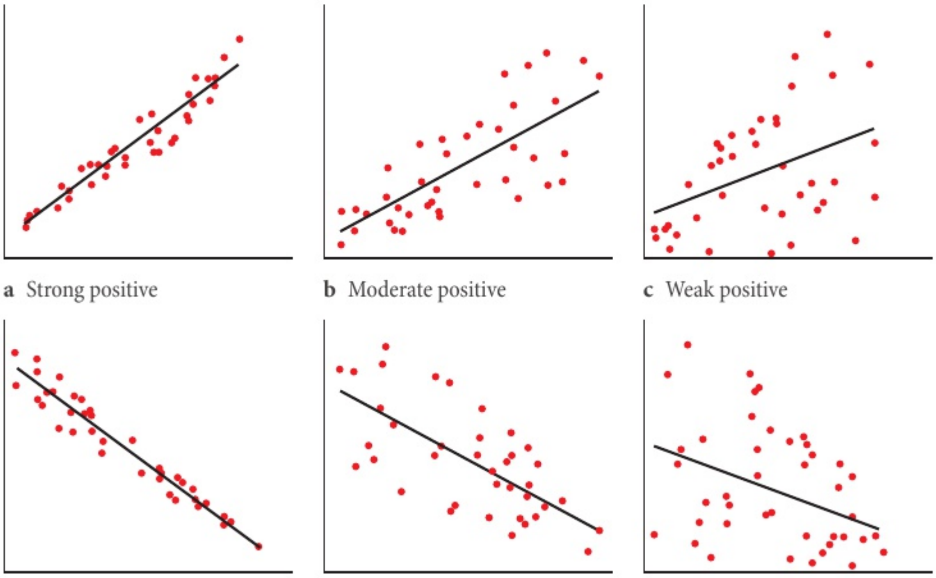


Figure 14.6 Two examples of zero association

Strength

Strength refers to how tightly the data appears to fit the form.


a Strong positive **b** Moderate positive **c** Weak positive

d Strong negative **e** Moderate negative **f** Weak negative

Figure 14.7 Strong, moderate, and weak associations

In Figure 14.7, the best-fitting linear model for each data set is also shown. As you can see, weak associations are widely scattered around the best-fit line, while strong associations show data tightly clustered around the best-fit line. Notice that both positive and negative associations can be strong and both can be weak.

Unusual features

We want to make note of any unusual features we see. **Outliers** are data points that do not fit the pattern or form of the other data points, as shown in Figure 14.8.

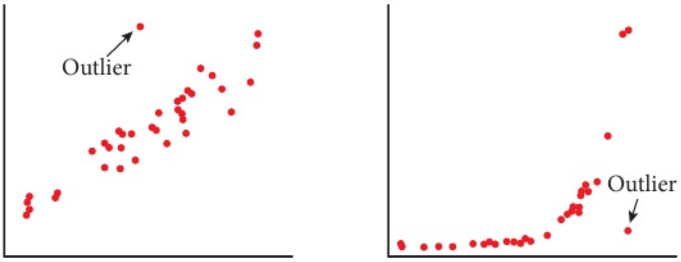
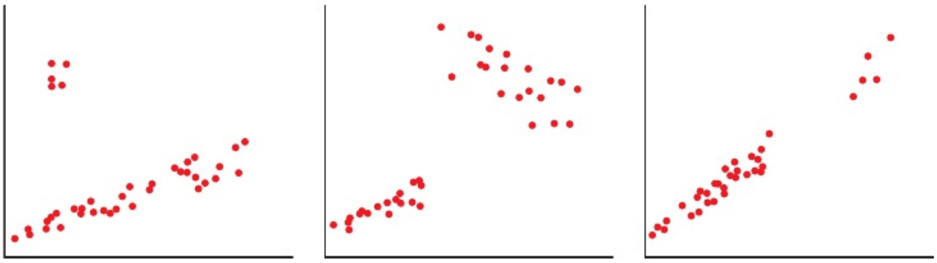


Figure 14.8 Outliers are a departure from the predominant pattern

You may also see clusters, gaps, or other patterns in the data that should be investigated, as shown in Figure 14.9.



a A cluster of outliers

b Two different directions

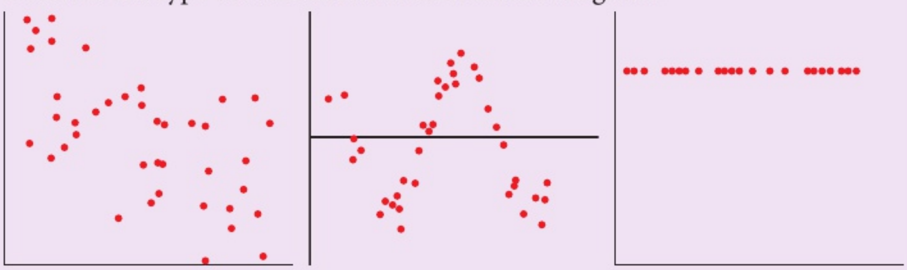
c This cluster fits the overall direction and form, but why the gap?

Figure 14.9 Clusters, gaps and other patterns

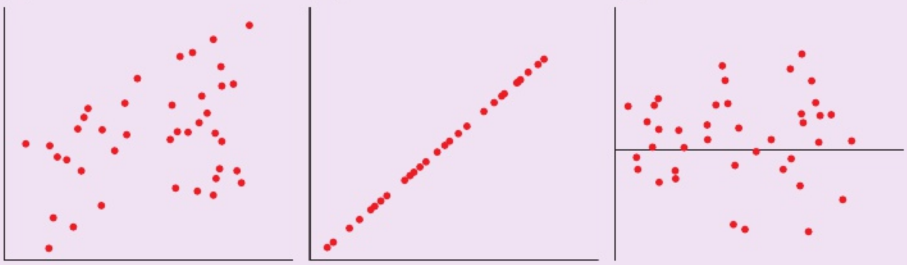
In general, unusual features in your data are always worth investigating. What caused the unusual feature? Is it a problem with the way the data was generated or collected? Is there an error in entering the data? Is there something in the phenomena you are investigating that is creating the feature?

Example 14.1

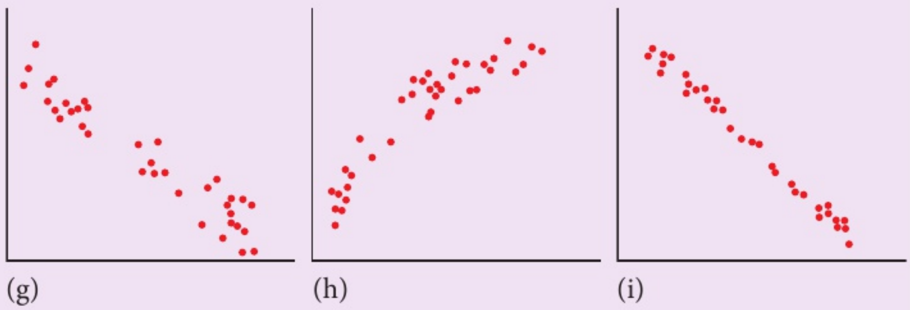
Describe the type of association in each scatter diagram.



(a) (b) (c)



(d) (e) (f)



Solution

(a) Moderate negative linear association.

(b) Nonlinear form, possibly sinusoidal.

(c) No association.

(d) Moderate positive linear association.

(e) Perfect positive linear association.

(f) No association.

(g) Strong negative linear association.

(h) Strong positive non-linear association.

(i) Very strong negative linear association.

Estimating the line of best fit

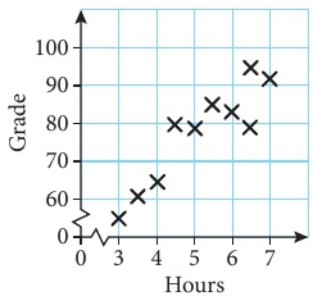


Figure 14.10 Mathematics students' scatter diagram

Consider our example from earlier, the class of mathematics students (Figure 14.10).

The association appears approximately linear, positive, and strong.

As a student in this class, we might ask questions such as:

- If I study for 4 hours, what grade can I predict?
- If I study for another hour, how much is my grade likely to increase?

To answer both of these questions, we would like to have a model of the relationship between hours and grade. In fact, we can obtain a reasonable model with very little effort by estimating a line of best fit by eye. Here is the procedure:

1. Calculate the mean of the x values, $\bar{x}$, and the mean of the y values, $\bar{y}$.
2. Plot the mean point $(\bar{x}, \bar{y})$ on the scatter diagram.
3. Draw a line through the mean point that fits the linear pattern. Aim to have approximately as many points above the line as below the line, but more important is that the points appear to be balanced above and below the line.

In this process we are very much relying on our visual sense to place the line. Why should the line go through the mean point $(\bar{x}, \bar{y})$? Intuitively, remember than our line of best fit is trying to describe the average behaviour, so it makes sense that the line should go through the mean point $(\bar{x}, \bar{y})$.

Example 14.2

The table shows some student test grade data.

Student	Tim	Joon	Jim	Kevin	Steve	Niki	Henry	Anton	Cindy	Lukas
Hours (x)	4	4.5	6	3.5	3	5	5.5	6.5	7	6.5
Grade (y)	65	80	83	61	55	79	85	79	92	95

(a) Draw a scatter diagram. Then draw the line of best fit by eye.

(b) Use your best-fit line to predict the score of a student who studies for 4 hours.

(c) Use your best-fit line to find the equation of a linear model.

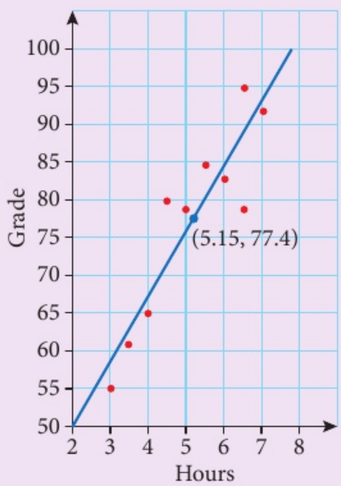
(d) Interpret your linear model to estimate how much a student's grade will improve for an additional hour of study.

Solution

(a) Draw the scatter diagram.
Then calculate the coordinates of the mean point:

$$\left.\begin{aligned} \bar{x} &= \frac{1}{n}\sum x_i = \frac{1}{10}(51.5) = 5.15 \\ \bar{y} &= \frac{1}{n}\sum y_i = \frac{1}{10}(774) = 77.4 \end{aligned}\right\} \Rightarrow (\bar{x}, \bar{y}) = (5.15, 77.4)$$

Plot the mean point on the scatter diagram and chose a line of best fit by eye. Here is one possible line:



Notice that although the line does not have the same number of data points above and below the line, it appears to be balanced.

Caution: Do not use existing data points to make predictions unless they happen to fall exactly on the best-fit line! Instead, use the best-fit line. For example, we see that Tim earned a grade of 65 with 4 hours of study. However, since Tim is not exactly on the best-fit line, he does not represent the average behaviour. Instead we predict that a student who studies 4 hours would, on average, receive a grade of 67. Likewise, do not use data points to estimate the slope of the best fit line, unless they happen to fall exactly on the best-fit line.

(b) To predict the grade of a student who studies for 4 hours, we simply read off the graph. The best-fit line appears to pass through the point (4, 67), so we predict that a student who studies for 4 hours would likely earn, on average, a grade of 67.

(c) To find the equation of our best-fit line, we will need to estimate the gradient of the best-fit line. To do this, we should pick two points on the best-fit line on the left and right sides of our data. On the left side, our best-fit line appears to pass through (2, 50), so that is an easy point to choose. On the right side, our best-fit line appears to pass through approximately (7.2, 95). Therefore, the gradient is

$$\frac{y_1 - y_2}{x_1 - x_2} = \frac{50 - 95}{2 - 7.2} \approx 8.7$$

Using the point (2, 50), we can find the linear model:

$$y - 50 = 8.7(x - 2) \Rightarrow y = 8.7x + 32.6$$

(d) In this context, the gradient of the linear model tells us the change in grade for each additional hour of study. Therefore, we predict that a student's grade will increase by 8.7 for each additional hour of study.

Association and correlation are not causation

It is important to remember than no matter how strong an association appears to be, changes in the explanatory variable do not necessarily cause changes in the response variable. It could be that the two variables are entirely unrelated. For example, consider the scatter diagram shown in Figure 14.11 (the axis labels are deliberately omitted).

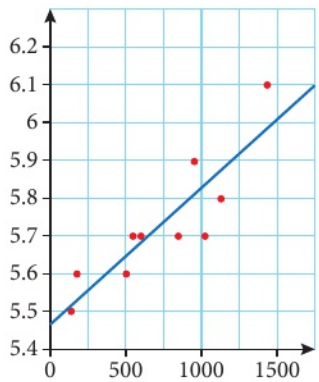


Figure 14.11 The data shows a strong, positive, linear correlation

It is perfectly correct to say that there is a strong, positive, linear correlation. However, do changes in the x variable cause changes in the y variable? No matter how strong the correlation appears to be, the data alone cannot prove that changes in x cause changes in y. To establish a causal relationship, we would need to design an experiment to manipulate the explanatory variable and look for changes in the response variable. In fact, it's not hard to find strong correlations between variables that have nothing to do with each other. For example, Figure 14.11 shows US consumption of milk vs European space launch revenue: the x variable is revenue from commercial space launches in Europe (in US$ millions); the y variable is per capita consumption of 1% skimmed milk in the USA (in US gallons). Clearly, increases in space launch revenue are not causing people to drink more milk in the USA, despite a strong correlation between the variables.

Exercise 14.1

1. Which point should every best-fit line pass through?

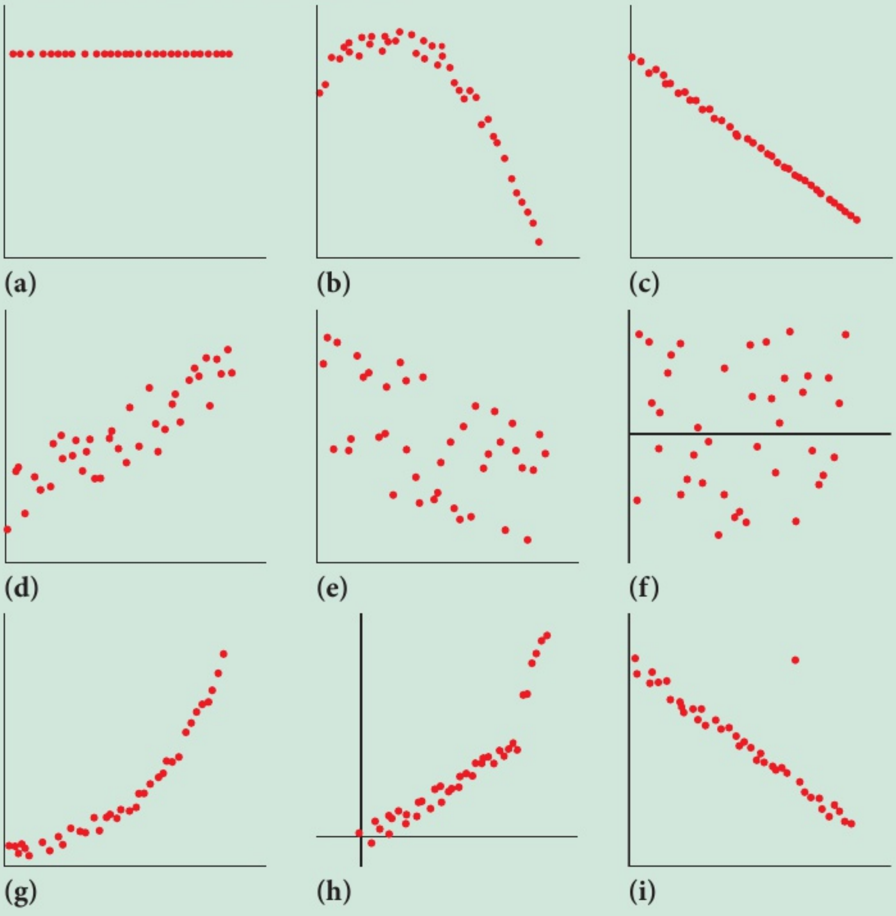
2. What four features should you look for when describing a scatter diagram?

3. In a 2008 study, researchers from Cornell University found that autism prevalence rates for school-aged children in California, Oregon and Washington in 2005 were positively related to the amount of precipitation these states received.

   (a) Identify the explanatory variable and response variable from this statement.

   (b) Does this mean that rainfall causes autism? Explain.

4. Describe each scatter diagram.



5. For each set of data, draw a scatter diagram, create a best-fit line by eye, and find the equation of your line.

   (a)

x	10	21	29	40	50	60
y	163	130	109	67	74	26

(b)

x	1.0	1.2	1.2	2.2	2.8	3.9	5.0	6.0	7.0	8.0	9.0	10.0
y	24.1	16.6	6.5	15.4	27.2	31.8	41.2	22.2	44.9	31.1	42.4	61.9

(c)

x	10	20	30	40	50	60
y	0	27	9	60	85	95

6. Lena is interested in whether the number of pages in a book determines how long it takes her to read it. She keeps track of the number of hours it takes her to read a book, and the number of pages in the book. The table shows the data for the last six books she read.

Pages	150	330	600	450	200	250
Hours	5	13	19	14	9	9

(a) Identify the explanatory and response variables in this context.

(b) Generate a scatter diagram and describe the relationship between the number of pages and the time required to finish a book.

(c) Create a best-fit line by eye and find the equation of your line for the number of hours t in terms of the number of pages n.

(d) Interpret the gradient and t-intercept in context.

(e) Predict the number of hours it would take Lena to read a 500-page book.

7. William is doing a study about students in his class concerning the mass of a random sample of 12 of his school mates against their metabolic rates. Metabolic rate is the rate at which the body consumes energy. The table shows the data he collected.

Student	Mass (kg)	Rate (kcal/day)
1	36.1	995
2	54.6	1425
3	48.5	1396
4	42	1418
5	50.6	1502
6	42	1256
7	40.3	1189
8	33.1	913
9	42.4	1124
10	34.5	1052
11	51.1	1347

(a) Identify the explanatory variable. Draw a scatter diagram of the data.

(b) Find the mean mass and the mean metabolic rate.

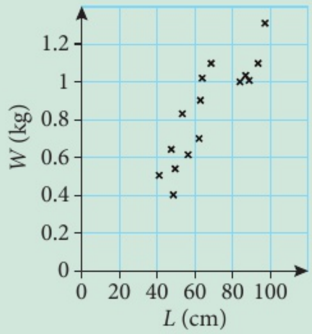
(c) Draw a line of best fit on your diagram.

(d) Describe the strength of the relationship and interpret the gradient of the line.

(e) Liz has a mass of 40 kg. What metabolic rate should she expect to see?

(f) Kevin has a mass of 70 kg. Can you use the model to predict his metabolic rate?

8. A sample of 15 fish was weighed. The weight, W, was plotted against length, L, as shown.



Exactly two of the following statements about the plot could be correct. Identify the two correct statements.

A The value of r, the correlation coefficient, is approximately 0.871.

B There is an exact linear relation between W and L.

C The line of best fit of W on L has equation $W = 0.012L + 0.008$

D There is negative correlation between the length and weight.

E The value of r, the correlation coefficient, is approximately 0.998

F The line of best fit of W on L has equation $W = 63.5L + 16.5$

9. The manager of a factory is curious about how production costs increase with production levels. She records the number of items produced and the production costs for 12 straight months. The table shows her data.

Number of items (thousands)	18	36	45	22	69	72	13	33	59	79	10	53
Production cost (£000)	37	54	63	42	84	91	33	49	79	98	32	71

(a) Identify the explanatory and response variable in this context.

(b) Generate a scatter diagram for production cost versus number of items.

(c) Describe the association between production cost and the number of items produced.

(d) Calculate the coordinates of the mean point.

(e) Draw a line of best fit by eye on your diagram.

(f) Use your line of best fit to predict the production cost when 70 000 items are produced.

(g) Find the gradient of your line of best fit and interpret in context.

(h) Estimate the y-intercept of your line of best fit and interpret in context. Give a reason why we should be cautious about this value.

(i) The manager would like to predict production costs when 100 000 items are produced. Give a reason why she should be cautious about this prediction.

Think back to extrapolation from Chapter 6.

10. When cars are driven, their tyres heat up. How does the driving speed relate to the temperature of the tyre? The table shows tyre temperature for a specific tyre model at the given speeds.

Speed (km h^{-1})	20	30	40	50	60	70	80	90
Temperature (°C)	45	52	64	66	91	86	98	105

(a) Identify the explanatory and response variables.

(b) Generate a scatter diagram for speed versus temperature and describe the association you see.

(c) Calculate the coordinates of the mean point.

(d) Draw a line of best fit by eye on your diagram.

(e) Use your best-fit line to predict the temperature when driving at 60 km h^{-1}.

(f) Find the gradient of your best-fit line and interpret in context.

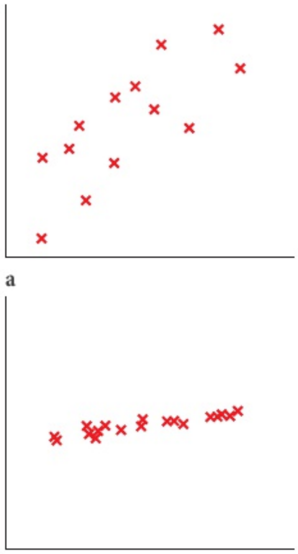
(g) Estimate the y-intercept of your best-fit line and interpret in context. Give a reason why we should be cautious about this value.

(h) Give a reason why we should not use your best-fit line to predict the tyre temperature when driving at 150 km h^{-1}.

14.2 Measures of correlation

So far, we have used our visual sense to describe the association we can see in a scatter diagram. Can we quantify that association? In this section, we will look at two methods for doing so: Pearson's product-moment correlation coefficient (r) and Spearman's rank correlation coefficient (r_s).

Pearson's product-moment correlation coefficient (r)



a

b

Figure 14.12 Which diagram shows a stronger correlation?

Consider the two scatter diagrams in Figure 14.12.

Which diagram shows a stronger correlation, diagram **a** or diagram **b**?

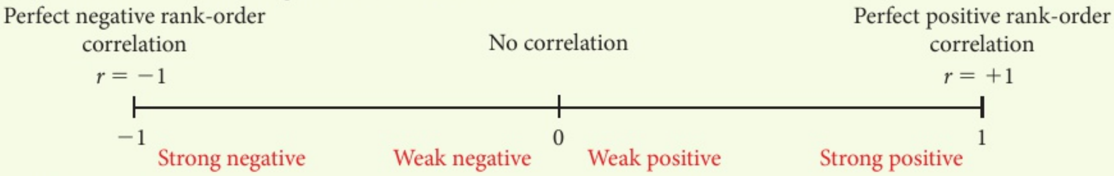
In fact, both diagrams show the same data. Sometimes, looks can be deceiving!

For this reason, we should try to quantify the strength and direction of an association. This is what Pearson's product-moment correlation coefficient aims to do. By convention, we refer to this statistic simply as r.

In this course, we will rely on technology to find the value of r. Our goal is to be able to interpret and use r. Consult your GDC manual to learn how to find r.

Facts about r

- The value of r is always in the interval $-1 \leqslant r \leqslant +1$
- The sign of r tells us the direction of the correlation: positive or negative or zero.
- The size of r tells us the strength of the linear correlation.



- When $r = +1$, there is a perfect positive linear correlation; all the points fall on a line with positive slope.
- When $r = 0$, there is zero correlation.
- When $r = -1$, there is a perfect negative linear correlation; all the points fall on a line with negative slope.
- r has no units, and is not a percentage.

There are no strict rules about what intervals of r indicate strong, moderate or weak correlations. Some authors claim $|r| > 0.5$ indicates a strong correlation, while others want $|r| > 0.87$. In reality, the interpretation of r depends on context: even a very low value of r may indicate a useful correlation in some contexts. Since in this course we are using only one explanatory variable, it is usually the case that r is not close to 1: in most (interesting) real-world situations there is more than one explanatory variable at work.

Remember, Pearson's correlation coefficient r applies only to linear relationships; that is, it measures the strength of a linear correlation, when it exists.

Here are some examples of r:

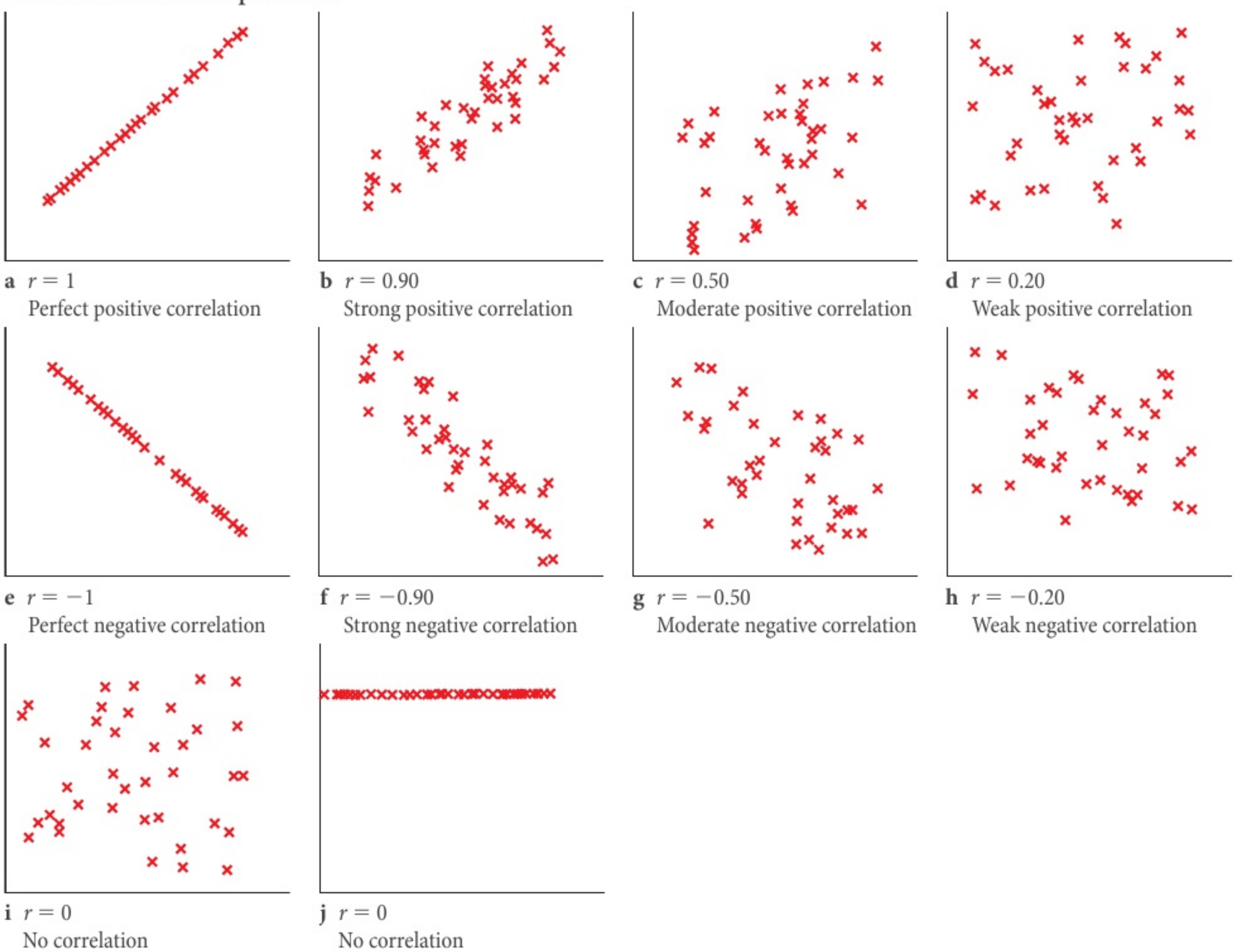
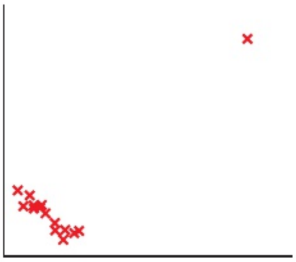
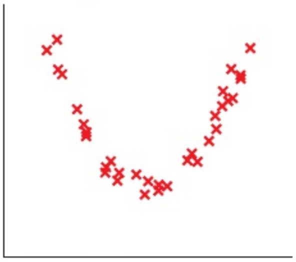


Figure 14.13 Examples of r



a $r = 0.75$
Strong positive correlation?



b $r = -0.02$
No correlation?

Figure 14.14 Some ways that r can mislead

It's always important to draw a scatter diagram. Do not rely on r to find an association, since there are several ways that r can show that there is a correlation when there is not, or that there is no linear correlation when a non-linear association exists. Figure 14.14 shows two examples.

In Figure 14.14a, there seems to be a negative association for most of the data, but a dramatic outlier fools r into indicating a positive linear correlation.

In Figure 14.14b, a non-linear pattern does not have a linear correlation, but there is clearly some sort of association between x and y.

Example 14.3

In an experiment, a metal bar is heated and its length is measured. The table shows measurements taken during the experiment.

Heat (°C)	40	45	50	55	60	65	70	75	80
Length (mm)	20	20.12	20.20	20.21	20.25	20.25	20.34	20.47	20.61

(a) Create a scatter diagram for this data.

(b) Describe the association, calculating and interpreting r if appropriate.

Solution

(a) The scatter diagram is shown in Figure 14.15.

(b) The association appears to be strong and positive. Since the association appears to be approximately linear, it is appropriate to proceed with calculating r. A GDC gives $r = 0.955$, which we interpret as a strong, positive linear correlation.

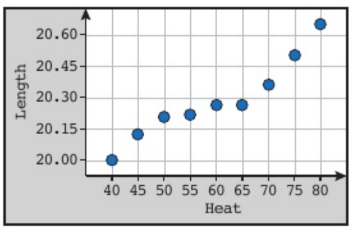


Figure 14.15 Scatter diagram for Example 14.3 (a)

Sometimes the data we obtain has unusual features such as an outlier. Depending on the context, we may decide to analyse the data without the outlier.

Example 14.4

A student is interested in whether the mass of a vehicle (kg) affects the fuel use (L per 100 km). The table contains data on 15 small cars.

Mass (kg)	1120	1170	1180	1180	1220	1250	1400	1400
Fuel consumption (L/100 km)	20.3	21.8	23.4	21	24.4	22.6	11.3	26.5

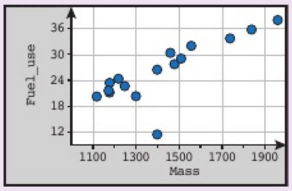
Mass (kg)	1460	1480	1510	1560	1740	1840	1960
Fuel consumption (L/100 km)	30.5	27.7	29	32.1	33.8	35.8	38.1

(a) Create a scatter diagram for this data.

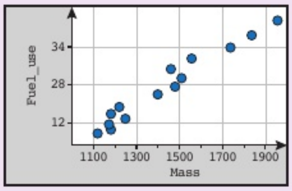
(b) Describe the association, calculating and interpreting r if appropriate.

Solution

(a) There is a clear outlier at (1400, 11.3). This vehicle is much more efficient than the others, for its mass. In fact, that data point represents a fuel-electric hybrid car. So, in this case, it's appropriate to remove that data point from our observation and focus only on the conventional fuel-powered vehicles. The new scatter diagram is shown below.



(b) The second scatter diagram suggests a strong, positive, and approximately linear association, so we will calculate r. A GDC shows that, with the outlier removed, $r = 0.980$, confirming a very strong, positive correlation.



Remember that r only has meaning for linear associations. If a non-linear association exists, r will not indicate the strength of the non-linear association and is therefore not useful.

Example 14.5

A student in a physics class collects data on the height of an object at different times after it is dropped from a height of 35 metres. The data are given in Table 14.2.

(a) Create a scatter diagram for this data.

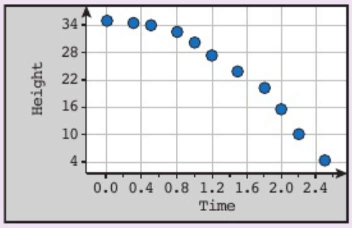
(b) Describe the association, calculating and interpreting r if appropriate.

Time	Height
0.00	34.83
0.25	34.62
0.50	34.11
0.75	32.58
1.00	30.21
1.25	27.25
1.50	23.85
1.75	20.27
2.00	15.52
2.25	10.1
2.50	4.38

Table 14.2 Data for Example 14.5

Solution

(a)



(b) The scatter diagram looks to have a strong, negative, and non-linear association. We know that falling objects can be modelled with quadratic functions and the scatter diagram suggests that a quadratic model could be appropriate. Since the data appears to have a non-linear association, it is not useful to calculate r.

Spearman's rank correlation coefficient (r_s)

Months since purchase	Change in value (%)
0	0
2	−29
4	−47
6	−58
8	−69
10	−72
12	−71
14	−77
16	−80
18	−82
20	−83
22	−85
24	−87

Table 14.3 Mobile phone value loss

Mobile phones often lose their resale value quickly after they are purchased. Data on the change in value for one particular model are given in Table 14.3.

How strong is the association between the age of the mobile phone and depreciation? Draw a scatter diagram to check for form before calculating Pearson's r. The scatter diagram is shown in Figure 14.16.

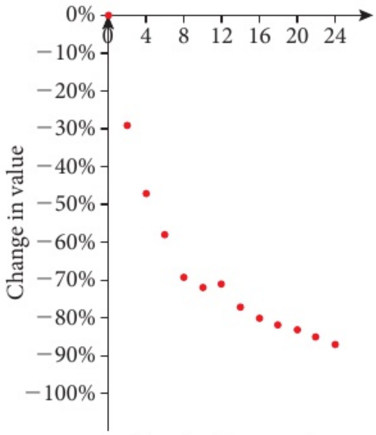


Figure 14.16 Mobile phone value loss scatter diagram

The association appears to be strong, negative, and non-linear. Since Pearson's r measures the strength of a linear correlation, it will likely under-estimate the strength of this association.

So how can we quantify the strength and direction of the association? Suppose we want to measure whether value consistently decreases as the age of the phone increases.

Consider this: if it is true that the value decreases consistently as time progresses, then we could rank the loss in value data and the ranks should be in strictly descending order. We will do the same with the month data to generate the rank order shown in Table 14.4.

Month rank	Loss in value rank
13	1
12	2
11	3
10	4
9	5
8	7
7	6
6	8
5	9
4	10
3	11
2	12
1	13

Table 14.4 Rank order

We observe that the rank-order is almost perfect. We can now quantify the association using Spearman's rank correlation coefficient, r_s. Using a GDC, we find that $r_s = -0.99$. The interpretation for Spearman's r_s is the same as Pearson's r.

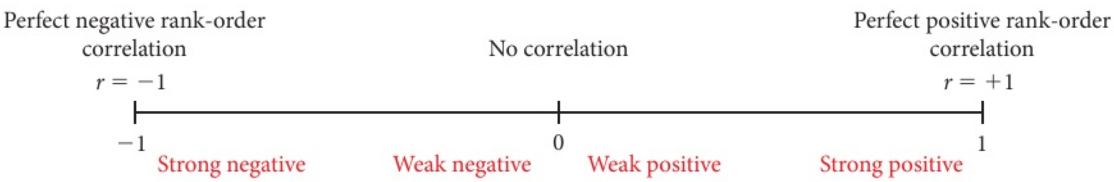


Figure 14.17 Interpration of Spearman's r_s.

Therefore, we conclude that there is a nearly perfect negative rank-order correlation between the age of this mobile phone and its loss in value.

Differences between Pearson's r and Spearman's r_s

Even though the calculations for Pearson's r and Spearman's r_s are identical, there are two differences to keep in mind. Both of these have to do with the fact that, for Spearman's r_s, we convert the data set into ranks before calculating the correlation.

Most GDCs do not include an option for calculating Spearman's r_s. Don't worry! The calculation for r_s is exactly the same as Pearson's r. Therefore, use your GDC's tool for finding r after you enter the rank order data. Make sure that you interpret the resulting r value as Spearman's r_s, not Pearson's r.

Linear and monotonic functions

As you have seen, both Pearson's r and Spearman's r_s attempt to measure the direction and strength of a correlation. However, while Pearson's r measures the strength of a linear correlation, Spearman's is measuring the strength of a rank-order correlation. Spearman's r_s measures whether the response variable consistently increases or decreases as the explanatory variable increases. This is called monotonicity. Monotonic functions only increase or decrease, as illustrated in Figure 14.18. Non-monotonic functions are shown in Figure 14.19.

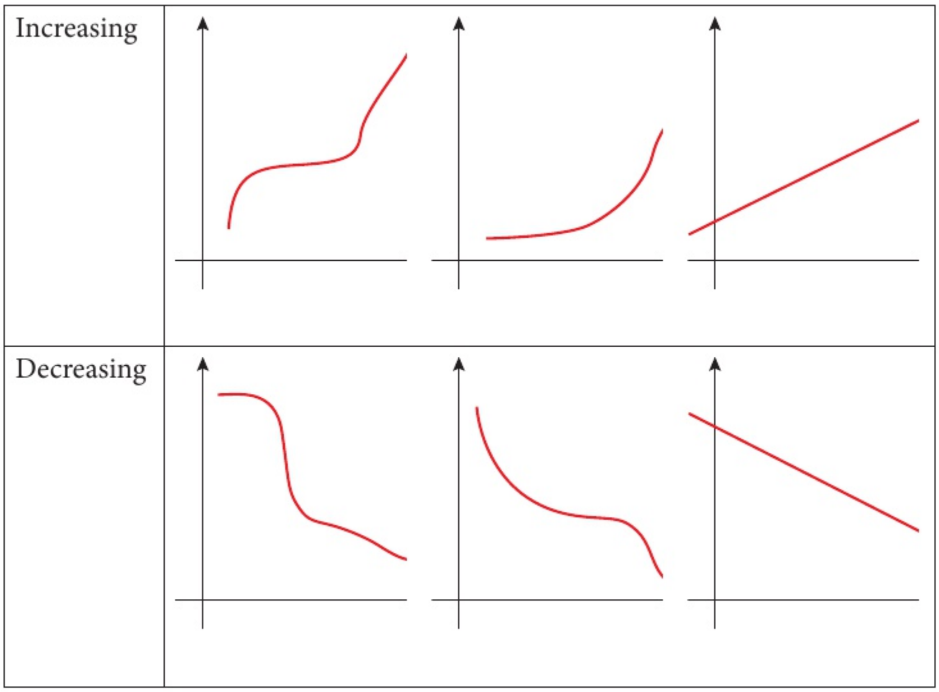


Figure 14.18 Monotonic functions

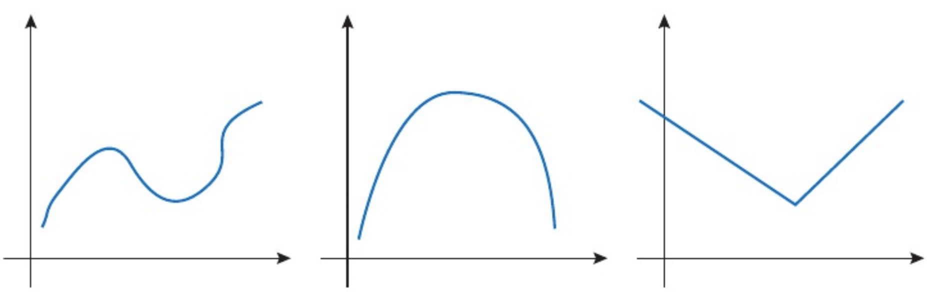


Figure 14.19 Non-monotonic functions

Effect of outliers

Recall that Pearson's r is sensitive to outliers. As you saw in Figure 14.14, the direction of Pearson's r can change completely when there are dramatic outliers. However, for Spearman's r_s, since we convert the data set into ranks before we calculate r_s, the effect of outliers is much reduced.

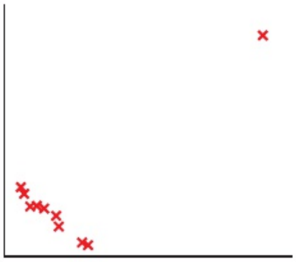


Figure 14.20 Dramatic outlier

For example, consider the data shown in Figure 14.20.

The value of Pearson's r is 0.77, indicating a strong positive correlation. From the scatter diagram, we see that is not the case. If we rank the data and calculate Spearman's r_s, we obtain $r_s = -0.48$. Spearman's r_s is a much better indicator of the direction and strength we can see in the scatter diagram. Of course, with a dramatic outlier like this, we should investigate the cause of that outlier and then decide what to do: remove the outlier or collect more or different data?

Facts about r_s

- The value of r_s is always in the interval $-1 \leqslant r \leqslant +1$
- The sign of r_s tells us the direction of the rank-order correlation: positive or negative or zero.
- The size of r_s tells us the strength of the rank-order correlation.
- If $r_s = +1$, there is a perfect positive rank-order correlation.
- If $r_s = 0$, there is zero correlation.
- If $r_s = -1$, there is a perfect negative rank-order correlation.
- r_s has no units, and is not a percentage.
- r_s is a measure of monotonicity.

When ranking data to calculate Spearman's r_s, we are not allowed to have tied **ranks**. But what do we do if there are tied ranks? We simply average the rank position for the tied data values.

Table 14.5 shows how much time a group of students spent studying for a mathematics test and their grades.

Student	Tim	Joon	Jim	Kevin	Steve	Niki	Henry	Anton	Cindy	Lukas	Julie
Hours (x)	4	4.5	6	3.5	3.5	5	5.5	6.5	7	6.5	6.5
Grade (y)	65	80	83	61	55	78	85	79	92	95	81

Table 14.5 How much time students spent studying and their grades

To calculate the rank order for this data set, we notice that there are two students (Kevin, Steve) who spent 3.5 hours studying and three students (Anton, Lukas, and Julie) who spent 6.5 hours studying.

Student	Tim	Joon	Jim	Kevin	Steve	Niki	Henry	Anton	Cindy	Lukas	Julie
Hours rank	9	8	5	10	10	7	6	2	1	2	2
Grade rank	9	6	4	10	11	8	3	7	2	1	5

Table 14.6 Rank order for the data set

Kevin and Steve are tied for rank 10, while Anton, Lukas, and Julie are tied for rank 2 (Table 14.6). Therefore, Anton, Lukas, and Julie are effectively occupying ranks 2, 3, and 4; the next rank (Joon) is rank 5. We assign Anton, Lukas, and Julie a rank equal to the mean rank they are occupying: Rank $= \dfrac{2+3+4}{3} = 3$.

Likewise, Kevin and Steve are taking up ranks 10 and 11, so we will assign the mean rank they are occupying as well: Rank $= \frac{10 + 11}{2} = 10.5$. Our new rank order is shown in Table 14.7.

Student	Tim	Joon	Jim	Kevin	Steve	Niki	Henry	Anton	Cindy	Lukas	Julie
Hours rank	9	8	5	10.5	10.5	7	6	3	1	3	3
Grade rank	9	6	4	10	11	8	3	7	2	1	5

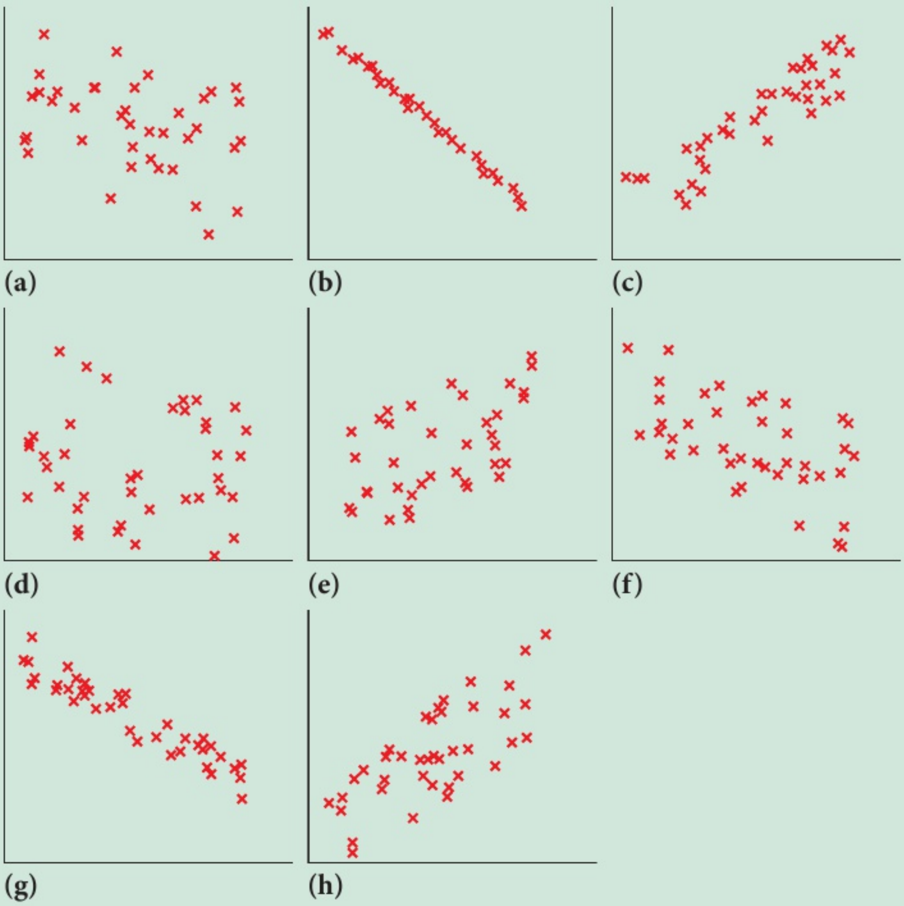
Table 14.7 New rank order

We can then calculate Spearman's r_s as usual.

Exercise 14.2

1. Explain the differences between Spearman's r_s and Pearson's r.
2. Assign each scatter diagram to one of the following r values:

$r = 0.90$, $r = 0.74$, $r = 0.51$, $r = -0.03$,
$r = -0.40$, $r = -0.58$, $r = -0.95$, $r = -0.99$



3. Six metal plates are immersed in an acid bath. After some time, they are removed and the mass is measured to determine the mass that was lost. Table 14.8 gives data from this experiment.
   (a) Generate a scatter diagram and calculate the value of r.
   (b) Describe the association between mass lost and hours.

Hours	Mass lost (%)
150	0.761
200	1.44
200	1.15
300	1.65
450	2.56
500	2.44

Table 14.8 Data for question 3

4. The table shows the cost in AUD of seven paperback books chosen at random, together with the number of pages in each book.

Book	**1**	**2**	**3**	**4**	**5**	**6**	**7**
Number of pages (x)	50	120	200	330	400	450	630
Cost (y AUD)	6.00	5.40	7.20	4.60	7.60	5.80	5.20

(a) Plot these pairs of values on a scatter diagram.

(b) Write down the linear correlation coefficient, r, for the data.

(c) Stephen wishes to sell a paperback book that has 350 pages in it. He plans to draw a line of best fit to determine the price. State whether or not this is an appropriate method in this case and justify your answer.

5. Does the acceleration ability of a car affect fuel consumption? Here are data from 15 cars.

Car	**0–60 mph time (s)**	**Fuel consumption (miles gal^{-1})**
Mazda MX-5 Miata Club	6.7	34
Honda Civic Si	7.3	34
Fiat 124 Spider Lusso	7.1	31
Mini Cooper S	7.0	30
Subaru BRZ Premium	7.2	30
Toyota 86	7.2	30
Volkswagen GTI Autobahn	6.6	29
Ford Fiesta ST	7.3	29
Fiat 500 Abarth	8.0	28
Porsche 718 Boxster (base)	4.4	26
Subaru Impreza WRX Premium	6.0	26
Audi TT 2.0T (AT)	6.3	26
Ford Focus ST	6.6	26
BMW M235i	5.2	25
Ford Mustang Premium (2.3T, AT)	6.4	25

(a) Generate a scatter diagram for this data and find the value of r.

(b) Describe the relationship.

(c) Rank the data. Hence, find the value of r_s and interpret in context.

(d) Comment on the difference between r and r_s. Is r or r_s a more appropriate measure of the strength of the association? Explain.

6. Here is the same data from question 5, but with fuel consumption in litres per 100 km, which is more frequently used outside the USA to measure fuel consumption.

Car	0–60 mph time (s)	Fuel consumption (L per 100 km)
Mazda MX-5 Miata Club	6.7	6.92
Honda Civic Si	7.3	6.92
Fiat 124 Spider Lusso	7.1	7.59
Mini Cooper S	7.0	7.84
Subaru BRZ Premium	7.2	7.84
Toyota 86	7.2	7.84
Volkswagen GTI Autobahn	6.6	8.11
Ford Fiesta ST	7.3	8.11
Fiat 500 Abarth	8.0	8.40
Porsche 718 Boxster (base)	4.4	9.05
Subaru Impreza WRX Premium	6.0	9.05
Audi TT 2.0T (AT)	6.3	9.05
Ford Focus ST	6.6	9.05
BMW M235i	5.2	9.41
Ford Mustang Premium (2.3T, AT)	6.4	9.41

(a) Generate a scatter plot and calculate the value of r.

(b) Describe the association and compare to question 5.

(c) Convert this data into ranked data.

(d) Find and interpret the value of r_s, then compare to question 5.

7. Fast food is often considered unhealthy because of high contents of fat and sodium. Are items that are high in fat also high in sodium? Here is some data on some popular menu items from fast-food restaurants.

Fat (g)	21	29	34	34	38	41	44
Sodium (mg)	900	1510	1320	830	1200	960	1210

(a) Generate a scatter diagram and calculate the value of r.

(b) Describe the association between fat and sodium for these menu items.

8. Fast food is often considered unhealthy because of high contents of fat and calories. Are items that are high in fat also high in calories? The table shows data on some popular menu items from fast-food restaurants.

Fat (g)	21	29	34	34	38	41	44
Calories (kcal)	410	580	570	560	660	680	670

(a) Generate a scatter diagram and calculate the value of r.

(b) Describe the association between fat and calories for these menu items.

9. Ian is doing a science experiment where he controls three explanatory variables (a, b, and c) and one response variable (S). After collecting the data and calculating Pearson's r, he obtains the following statistics:

Variable	a	b	c
r	0.86	0.32	0.55

(a) Can Ian conclude that a is the strongest explanatory variable?

(b) Do changes in a cause changes in S?

10. An anemometer measures wind speed by the rotation of a set of small cups called vanes. Anemometers may be calibrated by using a wind tunnel and measuring the rotational speed of the vanes. One such calibration generated the data shown in the table.

Wind speed ($\mathrm{m\,s^{-1}}$)	1	1.2	1.4	1.6	1.8	2
Rotations per minute (RPM)	29	42	65	91	110	121

(a) Generate a scatter diagram and calculate the value of r.

(b) Describe the association between wind speed and RPM.

(c) Give a reason why a linear model may not be the best for this situation.

11. Many universities calculate a measure of academic performance called a grade point average (GPA). Do students who work more hours in off-campus jobs have worse GPAs? Data from ten randomly sampled students at a certain university are shown in the table.

GPA	3.14	2.75	3.68	3.22	2.45	2.8	3	2.23	3.14	2.9
Hours	25	30	11	18	22	40	15	29	10	0

(a) Generate a scatter diagram for this data.

(b) Calculate the value of r_s and interpret in context.

(c) Calculate Pearson's r. Give a reason why r and r_s are not significantly different.

12. The specific weight of water varies with temperature according to the data shown in the table.

Temperature (°C)	0	0.1	1	4	10	15	20
Specific weight ($\mathrm{kN\,m^{-3}}$)	9.805	9.8052	9.8057	9.8064	9.804	9.798	9.789

(a) Generate a scatter diagram for this data.

(b) Calculate the value of r_s and interpret in context.

(c) Give a reason why r_s is not the best measure of association for this data.

14.3 Linear regression

So far we have found linear best-fit models by eye, using the mean point. Is there a more rigorous method? There is. In this course we use GDCs to find the linear models; we will focus on correct use and interpretation. The process of fitting a model to data is called regression.

Least-squares regression line (LSRL)

Example 14.6

In Zambia, a staple of the diet is ground maize, known as mealie meal. In one market, bags of mealie meal are priced as shown in the table. Prices are given in Zambian kwacha (ZMW).

Mass (kg)	5	10	25	50
Price (ZMW)	20	34	79	140

(a) Explain why it is appropriate to use a linear model for this data.

(b) Find the best-fit linear model for this data to predict the price given the mass. Use P to represent the price in ZMW and m to represent the mass in kg.

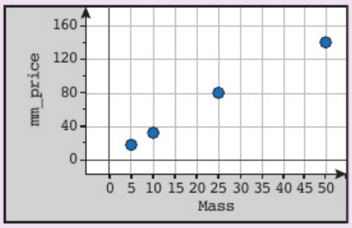
(c) Interpret the gradient and P-intercept of your model in context.

(d) Use your model to predict the price of a 50 kg bag. Does your model over-predict or under-predict the actual price? Give a reason why this may be the case.

(e) Write down a reasonable domain for your model and give a reason for your choice.

Solution

(a) To decide whether a linear model is appropriate, start by examining a scatter diagram.



We see that the data appears to have a strong, positive, linear association. To three significant figures, Pearson's $r = 0.999$

A linear model is appropriate.

(b) Use a GDC to find the least-squares regression line. To three significant figures, the best-fitting linear model is $P = 2.67m + 8.14$

There are two main methods for finding a best-fit linear model. In this course, we will use the least-squares regression line. This is the default in most GDCs, but you may also see a median-median line – we will not use this method in this course.

You may also see two different options for the least-squares regression line: $y = mx + b$ and $y = a + bx$

Mathematically, they are identical, but be careful to interpret them correctly: in the first, m represents the gradient and b represents the y intercept. In the second, a represents the y intercept and b represents the gradient.

Why is it called least-squares regression? The mathematical method for finding the least-squares line of best fit relies on minimising squared distances from each point to the regression line, hence, least-squares.

(c) The gradient of 2.67 suggests that, on average, the price for each additional kg of mealie meal is 2.67 ZMW. The P-intercept of 8.14 suggests that an empty bag of mealie meal costs 8.14 ZMW. This is a fixed cost representing distribution, packaging, and other costs.

(d) According to our model, the price of a 50 kg bag of mealie meal would be $P = 2.67(50) + 8.14 = 142$ ZMW (3 s.f.). This is a very slight over-prediction, which suggests that the 50 kg bag is a relatively good price compared to the other bags. This price is probably set to encourage consumers to purchase in greater volume.

(e) We can be confident predicting prices for bags with mass $5 \leqslant m \leqslant 50$ since any prediction in that interval would be an interpolation. Therefore, the domain for our model should be $5 \leqslant m \leqslant 50$

A prediction based on a mass outside that interval would be an extrapolation and is therefore unreliable.

What can go wrong?

When using linear regression models – or any other model – we need to be careful to avoid a few common pitfalls.

The least-squares regression line for y against x will be identical to the regression line for x against y if and only if $|r| = 1$

In other words, the difference between the two models is small when the linear correlation is strong, so predictions of x from y are more reliable. For moderate and weak correlations, however, the difference between the models can be significant, therefore predictions of x from y are unreliable.

Don't try to predict x from y using the model (especially if the correlation is not strong). Once we have a linear model, it's tempting to use the model in reverse. For example, in the model above, we could try to predict the mass of a bag priced at 50 ZMW. What's the problem? Algebraically, we can certainly solve the equation to express x in terms of y. However, because of the process used to find the least-squares regression line, if we swap x and y and re-calculate the regression on our GDC, we will get a model that is not algebraically equivalent to the previous model. Therefore, your least-squares regression line should only be used in one direction: to predict a value of the response variable (y) given a value of the explanatory variable (x). If you are asked to predict x from y using a model for y in terms of x, make sure you proceed with extreme caution.

Don't extrapolate. Extrapolating assumes that the pattern you've observed will continue. You may have good reason to believe that is the case, but it is always risky, even when the linear correlation is very strong.

Refer to Chapter 6 for more detail about the dangers of extrapolation.

The model is only as good as the strength of the correlation. Just because you can make a model, doesn't mean you should. You need to examine the scatter diagram and make sure a linear model is appropriate. Once you decide that the form is approximately linear, you may proceed to calculating the least-squares regression line. However, remember that the predictive power of your least-squares regression line is only as good as the strength of your correlation. If the linear correlation is weak to begin with, then you can't expect your model to make accurate predictions.

Piecewise linear models

For more information about piecewise models, refer to Chapter 6.

Sometimes we have data that very clearly has two linear trends visible. In this case, we can construct a piecewise linear model by partitioning the data set into two or more subsets.

Example 14.7

A research vessel is mapping the floor of the sea in a certain area. To do this, the boat travels directly out from the shore and measures the depth of the water continuously. The table shows a sample of the data collected.

Distance from shore (km)	0.2	0.4	0.6	0.8	1	1.2	1.4	1.6	1.8	2
Depth (m)	2	1.9	2.3	2.3	2.2	2.8	3.3	4	4.6	5.1

(a) Explain why a piecewise linear model is appropriate to model depth as a function of distance from shore.

(b) Generate a piecewise linear model, using d for distance from shore and h for depth.

(c) Interpret the gradients in your piecewise model.

(d) Use your model to predict the depth 500 m from the shore and 1.5 km from the shore.

Solution

(a) To determine if a piecewise linear model is appropriate, examine the data using a scatter diagram (see Figure 14.21). It appears that there are two linear forms; a piecewise linear model appears appropriate. Furthermore, it appears that the breakpoint between the linear models should be at a distance of 1.0 km.

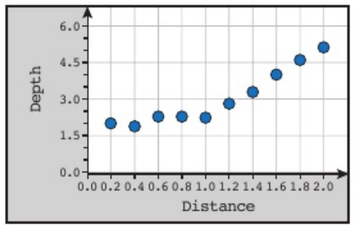


Figure 14.21 Scatter diagram for Example 14.7 (a)

(b) To generate the two pieces, we partition the data into two parts, before and after the breakpoint we have chosen visually. Note that the data point (1, 2.2) appears to fit in both parts, so we will include it in both least-squares regression line calculations.

Distance from shore (km)	0.2	0.4	0.6	0.8	1	1.2	1.4	1.6	1.8	2
Depth (m)	2	1.9	2.3	2.3	2.2	2.8	3.3	4	4.6	5.1

Then, enter the data as two separate data sets into a GDC and calculate a least-squares regression line for each. The GDC shows that the best-fit model for the first part of the data is $d = 0.400h + 1.90$, and the best-fit model for the second part of the data is $d = 2.94h - 0.747$

Now, we combine these two into a piecewise model:

$$d = \begin{cases} 0.400h + 1.90, & 0 \leqslant h \leqslant 1.0 \\ 2.94h - 0.747, & 1.0 < h \leqslant 2.0 \end{cases}$$

Graphically, our model appears to fit the data quite well, as shown in Figure 14.22.

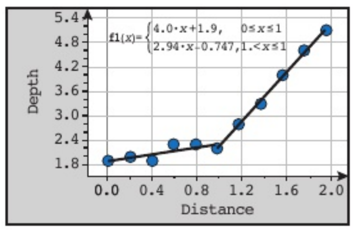


Figure 14.22 GDC screen for solution to Example 14.7 (b)

(c) The two pieces of the model suggest that the depth of the ocean is increasing at a rate of 0.4 m km^{-1} for the first km, then the rate increases to 2.94 m km^{-1}.

(d) To use the model to make predictions, proceed as with any piecewise model.
For 500 m from the shore: $h = 0.5 \Rightarrow d = 0.4(0.5) + 1.9 = 2.1$ m deep.
For 1.5 km from the shore: $h = 1.5 \Rightarrow d = 2.94(1.5) - 0.747 = 3.66$ m deep.

Exercise 14.3

Hours	Mass lost (%)
150	0.761
200	1.44
200	1.15
300	1.65
450	2.56
500	2.44

Table 14.9 Data for question 1

1. Six metal plates are immersed in an acid bath. After some time, they are removed, and the mass measured to determine the mass that was lost. Table 14.9 gives data from this experiment.
   (a) Find the equation of the least-squares regression line for mass lost, L, in terms of hours immersed, h. Interpret your model in context.
   (b) For each additional 100 hours a metal plate was in this acid bath, k percent of the mass was lost. Find the value of k.
   (c) Use your least-squares regression line to predict the mass loss of a metal plate immersed for 400 hours.
   (d) Is it appropriate to use your model to predict the number of hours a metal plate with 1.3% mass loss was immersed? Explain.

2. Does the acceleration ability of a car affect fuel efficiency? Here are data from 15 cars.

Car	0–60 mph time (s)	Fuel consumption (miles gal^{-1})
Mazda MX-5 Miata Club	6.7	34
Honda Civic Si	7.3	34
Fiat 124 Spider Lusso	7.1	31
Mini Cooper S	7.0	30
Subaru BRZ Premium	7.2	30
Toyota 86	7.2	30
Volkswagen GTI Autobahn	6.6	29
Ford Fiesta ST	7.3	29
Fiat 500 Abarth	8.0	28
Porsche 718 Boxster (base)	4.4	26
Subaru Impreza WRX Premium	6.0	26
Audi TT 2.0T (AT)	6.3	26
Ford Focus ST	6.6	26
BMW M235i	5.2	25
Ford Mustang Premium (2.3T, AT)	6.4	25

   (a) Find the equation of the least-squares regression line for fuel consumption, E, in terms of acceleration time, t. Interpret your model in context.
   (b) Use your least-squares regression line to predict the fuel consumption of a car with a 0–60 mph time of 5.7 s. Give a reason why this prediction may not be reliable.
   (c) Is it appropriate to use your model to predict the efficiency of a car with a 0–60 time of 9 seconds? Explain.

3. Do bigger aeroplanes use more fuel? The table contains data on the number of passenger seats and fuel consumption of aeroplanes currently in widespread use.

Seats	405	296	288	258	240	230	193	188	148
Fuel consumption (L min^{-1})	224	140	138	94	100	95	87	62	79

Seats	142	131	122	115	112	103	102	78
Fuel consumption (L min^{-1})	56	46	54	70	51	40	51	48

(a) Generate a scatter diagram and describe the association between the number of seats and fuel consumption.

(b) Find the equation of the least-squares regression line for fuel consumption, F, in terms of the number of seats, n. Interpret your model in context.

(c) Use your model to predict the fuel consumption of an aeroplane with 350 seats.

(d) Give a reason why a linear model may not be the most appropriate model for this data.

4. A computer-based workout app ranks its workouts according to the training stress score. Another training website calculates a separate relative effort value for each workout that is uploaded to the site. Are training stress score and relative effort related? The table shows a sample of one cyclist's last 12 workouts.

Training stress score	82	96	48	79	79	79	88	112	82	74	74	96
Relative effort	59	70	37	60	58	51	67	79	48	44	46	84

(a) Draw a scatter diagram and describe the association between training stress score and relative effort.

(b) Find the equation of the least-squares regression line for relative effort, E, in terms of the training stress score, S. Interpret your model in context.

(c) Use your model to predict the relative effort for a workout with a training stress score of 60.

(d) Write down a valid domain for your model.

(e) If you removed the possible outlier at (48, 37), how would your answers to (d) and (c) change?

(f) Remove the outlier at (48,37) and recalculate the least-squares regression line. Did the model change significantly?

Change in NEA (kcal)	Change in mass (kg)
−94	4.2
−57	3
−29	3.7
135	2.7
143	3.2
151	3.6
245	2.4
355	1.3
392	3.8
473	1.7
486	1.6
535	2.2
571	1
580	0.4
620	2.3
690	1.1

Table 14.10 Data for question 5

5. It is well known that many people who eat more don't gain weight as quickly as others do. Could non-exercise activity (fidgeting, maintaining posture, etc.) explain the difference? In a 1999 study, researchers deliberately overfed 16 non-obese volunteers and measured the change in their calorie expenditure through non-exercise activity (NEA) as well as their weight gain after 8 weeks. The data collected are shown in Table 14.10.

(a) Generate a scatter diagram and describe the association between change in non-exercise activity and change in mass.

(b) Find the equation of the least-squares regression line for change in mass, M, in terms of the change in non-exercise activity, N. Interpret your model in context.

(c) Use your model to predict the change in mass for an individual with a change in non-exercise activity of 200.

(d) Explain what a negative value for change in non-exercise activity means in this context.

(e) Write down a valid domain for your model.

(f) Making a prediction for change in mass given a change in non-exercise activity of 1100 is an extrapolation. Give another reason why it is particularly nonsensical in this context.

Defects	Production rate (units per hour)
20	400
30	450
10	350
20	375
30	400
25	400
30	450
20	300
10	300
40	300

Table 14.11 Data for question 6

6. A company believes that its workers make more mistakes when they work faster. To test this theory, they collect data on the production rate and the number of mistakes at that rate. The data are given in Table 14.11.

(a) Identify the explanatory and response variable in this context.

(b) Generate a scatter diagram and describe the association between defects and production rate.

(c) Find the equation of the least-squares regression line for defects, D, in terms of the production rate, P. Interpret your model in context.

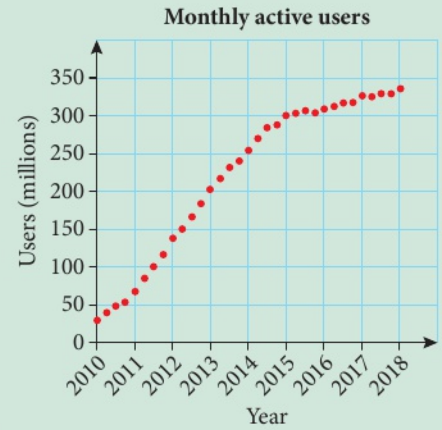
(d) Find the value of r and interpret in context.

(e) Would you advise this company to lower production rates in order to reduce defects?

(f) Remove the outlier from the data and re-do parts (b)–(e).

(g) Is it OK to remove this outlier in this case? Explain.

7. To track the number of active users for a website, it is common to count the number of users who have logged in at least once in a given month. This count is called monthly active users. The number of monthly active users for one social media service is shown in the scatter diagram.



(a) Describe the association shown in the scatter diagram.

(b) A piecewise linear model may be appropriate. Give the domain for each piece of a piecewise linear model, using years since 2010 (y) as the explanatory variable.

8. A dairy farmer would like to investigate whether temperature has an effect on the butterfat content of the milk his dairy cows produce. Over the course of a year, he records the butterfat content of the milk they produce and the air temperature. His data are shown in Table 14.12.

(a) Identify the explanatory and response variable in this context.

(b) Generate a scatter diagram and describe the association between temperature and butterfat content.

(c) Find the equation of the least-squares regression line for butterfat, F, in terms of the temperature, T. Interpret your model in context.

(d) Find the value of r and interpret.

(e) If the farmer wants to increase the butterfat content of the cows' milk, would investing in a climate-controlled barn for the cows be a good choice? Explain your answer.

Temperature (°C)	Butterfat (%)
3	4.87
3	5.09
4	4.97
5	4.52
7	4.83
8	4.85
13	4.77
13	4.48
13.5	4.23
13.5	4.85
14	4.51
14.5	4.74
15	4.45
15.5	4.7
16	4.65
16.5	4.45
18	4.63
18	4.65
18.5	4.65
18.5	4.59

Table 14.12 Data for question 8

Chapter 14 practice questions

1. What is wrong with each statement?

(a) The value of Pearson's correlation coefficient for the time it takes an athlete to run 5 km and the time it takes them cycle 30 km is $r = 1.21$

(b) The fuel consumption of cars decreases linearly with the mass of the car with Pearson's $r = 0.78$

(c) Among mammals, those with greater average body mass have longer life expectancy; the rank correlation coefficient is $r_s = -0.85$

(d) For a set of (x, y) data, the least-squares regression line is $y = 25 - 3.52x$ with $r = 0.64$

2. For each situation, would you expect the rank correlation coefficient r_s to be close to $+1$, 0, or -1?
   (a) The number of hours an IB Diploma student studies and the amount they spend on lunch each week.
   (b) The tax on cigarettes and the number of cartons sold.
   (c) The age of children and their mass.
   (d) The speed of a motor vehicle and the distance required to brake to a stop.
   (e) The amount spent on snack food and the average number of years of education in various countries.
   (f) The outside temperature and money spent on heating.

3. The intensity of radiation, I, from a certain source is measured at regular intervals. The measurements are given in the table.

Time (min)	0.2	0.4	0.6	0.8	1.0
ln(I)	0.508	0.212	−0.051	−0.387	−0.444

   (a) Show that the relationship between ln(I) and t is approximately linear.
   (b) Hence find the least-squares regression line for ln(I) in terms of t.

4. The time it takes for a runner to complete a marathon can be based on the time it takes to run 5 km according to the table.

5 km time (min)	15	17	19	21	23	25	27	29
Marathon time (min)	144	163	182	201	220	238	256	274

   (a) Generate a scatter diagram, find the value of r, and describe the relationship between 5 km time and marathon time.
   (b) Find the least-squares regression line for marathon time (M) based on 5 km time (T). Interpret the gradient and intercept in context.
   (c) Use your least-squares regression model to predict the marathon time for a runner with a 5 km time of 20 minutes.
   (d) The length of a marathon is 42.195 km. If runners ran at the same pace for marathons as they do for 5 km, the gradient of the least-squares model would be k.
      (i) Find the value of k.
      (ii) Find the percentage difference between k and the gradient from part (**b**).
      (iii) Explain the meaning of the percentage difference in context.

5. The width and length of several leaves from the same tree are measured. Data is given in the table.

Width (mm)	Length (mm)
44	102
43	99
45	103
48	111
52	119
43	100
46	106
50	114

(a) Generate a scatter diagram, find the value of r, and describe the relationship between width and length for these leaves.

(b) Find the least-squares regression line for length (L) based on width (W). Interpret the gradient and intercept in context.

(c) Write down a suitable domain for your model.

(d) Use your least-squares regression model to predict the length of a leaf with a width of 47 mm.

(e) Give a reason why we should not use the model to predict the length of a leaf that is 60 mm wide.

6. Some researchers drove a compact car at various speeds around a racetrack to see if fuel consumption relates to speed. Their data are shown in the table.

Speed (km h^{-1})	50	60	70	80	90	100	110	120
Fuel consumption (L per 100 km)	9.6	8.9	8.3	8	8.1	8.7	9.8	10.4

(a) Generate a scatter diagram for this data and describe the association.

(b) Give a reason why a linear model should not be used for this data.

(c) Give a reason why Spearman's r_s would not be a good measure of the strength and direction of the association.

7. In a 2017 study, researchers used data from US national health surveys from 1997 to 2009, involving 333 247 participants, and found that light and moderate consumption of alcohol correlated with fewer deaths from cardiovascular disease. Does this mean that drinking alcohol is beneficial?

8. Do faster aeroplanes use more fuel? The table contains data on the average speed and fuel consumption of different aeroplanes on mid-range journeys.

Speed (km h^{-1})	836	802	779	741	762	770	765	723	687
Fuel consumption (L min^{-1})	224	140	138	94	100	95	87	62	79

Speed (km h^{-1})	670	665	609	679	625	580	607	605
Fuel consumption (L min^{-1})	56	46	54	70	51	40	51	48

(a) Generate a scatter diagram and describe the association between speed and fuel consumption.

(b) Is Pearson's r or Spearman's r_s a better measure of strength and direction for this association? Give a reason for your answer, then calculate and interpret.

9. Is the flying speed of animals related to their overall body length? A sample of data for some flying species is given in the table.

Animal	**Length (cm)**	**Speed (cm s^{-1})**
Fruit fly	0.2	190
Horse fly	1.3	660
Hummingbird	8.1	1120
Dragonfly	8.5	1000
Bat	11	690
Common swift (bird)	17	2550
Flying fish	34	1560
Pintail duck	56	2280
Swan	120	1880
Pelican	160	2280

(a) Generate a scatter diagram of this data and describe it.

(b) Are there any outliers in the data? Assuming the data are correct, is it appropriate to remove them?

(c) Calculate the value of Pearson's r and interpret in context.

(d) Rank the data, then calculate the value of Spearman's r_s and interpret in context.

(e) Is one measure of association, r or r_s, more appropriate for this data? Give a reason for your answer.

10. The atmospheric concentration of carbon dioxide (CO_2) in parts per million (ppm) in the month of June at the Mauna Loa observatory in Hawaii, USA, is given in the table.

Year	**1960**	**1970**	**1980**	**1990**	**2000**	**2010**	**2017**	**2018**
CO_2 concentration (ppm)	320	328	341	356	372	392	409	411

(a) Generate a scatter diagram of this data and describe it.

(b) Calculate the value of Pearson's r and interpret in context.

(c) When constructing models, we often modify dates to be years since a certain date.

  (i) Modify the date value to be years since 1960.

  (ii) Recalculate the value of Pearson's r and interpret in context. Has it changed?

  (iii) Find a least-squares regression line for CO_2 concentration C based on y, years since 1960. Interpret the slope and intercept in context.

  (iv) Write down an appropriate domain for your model in context.

  (v) Use your model to predict the CO_2 concentration in the year 2005.

  (vi) Give a reason why this model should not be used to predict the CO_2 concentration in the year 2050.

11. In 1881, the Russian chemist D. Mendeleev investigated the solubility of sodium nitrate ($NaNO_3$) at different water temperatures. His data are shown in the table.

Temperature (°C)	0.0	4.0	10	15	21	29	36	51	68
Solubility (g per 100 ml)	66.7	71.0	76.3	80.6	85.7	92.9	99.4	114	125

(a) Generate a scatter diagram of this data and describe it.

(b) Calculate the value of Pearson's r and interpret in context.

(c) Find a least-squares regression line to predict solubility S based on temperature T and interpret in context.

(d) Use your least-squares regression line to predict the solubility of $NaNO_3$ at 25°C.

(e) Give a reason why your least-squares regression line should not be used to predict solubility of $NaNO_3$ at 95°C.

(f) A certain solution has a $NaNO_3$ concentration of 100 g per 100 ml. Predict the temperature of this solution, or give a reason why you cannot.

Time (s)	Velocity (cm s^{-1})
0	0
0.1	−91.3
0.2	−189
0.3	−316
0.4	−425
0.5	−500
0.6	−571
0.7	−740
0.8	−836
0.9	−940
1	−1020

Table 14.13 Data for question 12

12. In physics class, a student dropped an object and measured its velocity every 0.1 seconds for 1 second. Her data are shown in Table 14.13.

(a) Generate a scatter diagram of this data and describe it.

(b) Calculate the value of Pearson's r and interpret in context.

(c) Find a least-squares regression line to predict velocity $v(t)$ based on time t.

(d) Interpret the slope and intercept of your least-squares regression line, with appropriate units.

(e) Use your least-squares regression line to predict the velocity of this object at 0.15 s.

(f) The rate of change in velocity (acceleration) due to gravity on Earth should be 981 cm s^{-2} at the surface of the Earth. Does this experiment agree with that theoretical value? Explain.

Internal assessment

Internal assessment (IA) is an important component of the Applications and Interpretation SL course and contributes 20% to your final grade. It is a significant part of the overall assessment for the course and should be taken seriously. It should also be pointed out that your work in completing the IA component differs in important ways from the written exams (external assessment) for the course.

- Unlike written examinations, you do *not* perform IA work under strict time constraints.
- You have some freedom to decide which mathematical topic you wish to explore.
- Your IA work involves writing about mathematics, not just using mathematical procedures.
- Regular discussion with, and feedback from, your teacher will be essential.
- You should endeavour to explore a topic in which you have a genuine personal interest.
- You will be rewarded for evidence of creativity, curiosity, and independent thinking.

Mathematical exploration

To satisfy the IA component, you are required to complete a piece of written work on a mathematical topic that you choose in consultation with your teacher. This piece of written work is formally referred to as the mathematical exploration. It will be referred to simply as the 'exploration' throughout this chapter. Your primary objective is to *explore* a mathematical topic in which you are *genuinely interested* and that is at an *appropriate level* for the course. A fundamental aspect of your exploration must be the *use of mathematics* in a manner that clearly demonstrates your knowledge and understanding of the relevant mathematics. Your teacher may provide you with a list of ideas (or 'stimuli') to help you in the process of finding a suitable topic.

It is your responsibility to determine whether or not you are sufficiently interested in a particular topic – and it is your teacher's responsibility to determine if an exploration of the topic can be conducted at a mathematical level that is suitable for the course. Your teacher will help you determine if an exploration of a certain topic can potentially address the five assessment criteria satisfactorily. Your exploration should be approximately 12 to 20 pages long with double line spacing.

See the list of 200 ideas included in the eBook. You may find a suitable topic in the list, or the list may help you find or develop your own ideas for a mathematical topic to explore.

Internal assessment criteria

Your exploration will be assessed by your teacher according to the following five criteria.

A Presentation

This criterion assesses the organisation and coherence of the exploration. A well-organised exploration has an introduction, a rationale (which includes a brief explanation of why the topic was chosen), a description of the aim of the exploration, and a conclusion.

B Mathematical communication

This criterion assesses to what extent you are able to:

- use appropriate mathematical language (notation, symbols, terminology)
- clearly define key terms, variables, and parameters
- use multiple forms of mathematical representation, such as formulae, diagrams, tables, charts, graphs, and models
- apply a deductive approach in general, and present any proofs in a logical manner.

C Personal engagement

This criterion assesses the extent to which you engage with the exploration and present it in such a way that clearly shows *your own personal approach*. Personal engagement may be recognised in several different ways. These may include – but are not limited to – thinking independently and/or creatively, addressing personal interest, presenting mathematical ideas in your own words and diagrams, developing your own ideas and testing them, and creating your own examples to illustrate important results.

D Reflection

This criterion assesses how well you *review*, *analyse,* and *evaluate* the exploration. Although reflection may be seen in the conclusion to the exploration, you should also give evidence of reflective thought throughout the exploration. Reflection can be demonstrated by consideration of limitations and/or extensions, commenting on what you've learned, or comparing different mathematical methods and approaches.

E Use of mathematics

This criterion assesses to what extent and how well you use mathematics in your exploration. The mathematical working in your exploration needs to be *sufficiently sophisticated* and *rigorous*. The chosen topic should involve mathematics in the Applications and Interpretation SL syllabus or at a similar level. Sophistication and rigour can include understanding and use of challenging mathematical concepts, looking at a problem from different perspectives, mathematical arguments expressed clearly in a logical manner, or seeing underlying structures to link different areas of mathematics.

Your exploration will earn a score out of a total of 20 possible marks. The five criteria do not contribute equally to the overall score for your exploration. For example, criterion E (Use of mathematics) is 30% of the overall score, whereas crtieria C (Personal engagement) and D (Reflection) contribute 15% each.

It is very important that you familiarise yourself with the assessment criteria for the Applications and Interpretation SL exploration and refer to them while you are writing your exploration. The achievement levels for each criteria and associated descriptors are as follows:

A	**Presentation**
0	The exploration does not reach the standard described by the descriptors below.
1	The exploration has some coherence or some organisation.
2	The exploration has some coherence and shows some organisation.
3	The exploration is coherent and well organised.
4	The exploration is coherent, well organised, and concise.

B	**Mathematical communication**
0	The exploration does not reach the standard described by the descriptors below.
1	The exploration contains some relevant mathematical communication that is partially appropriate.
2	The exploration contains some relevant appropriate mathematical communication.
3	The mathematical communication is relevant, appropriate, and is mostly consistent.
4	The mathematical communication is relevant, appropriate, and consistent throughout.

C	**Personal engagement**
0	The exploration does not reach the standard described by the descriptors below.
1	There is evidence of some personal engagement.
2	There is evidence of significant personal engagement.
3	There is evidence of outstanding personal engagement.

D	**Reflection**
0	The exploration does not reach the standard described by the descriptors below.
1	There is evidence of limited reflection.
2	There is evidence of meaningful reflection.
3	There is substantial evidence of critical reflection.

E	**Use of mathematics**
0	The exploration does not reach the standard described by the descriptors below.
1	Some relevant mathematics is used.
2	Some relevant mathematics is used. Limited understanding is demonstrated.
3	Relevant mathematics commensurate with the level of the course is used. Limited understanding is demonstrated.

4	Relevant mathematics commensurate with the level of the course is used. The mathematics explored is correct. Some knowledge and understanding are demonstrated.
5	Relevant mathematics commensurate with the level of the course is used. The mathematics explored is correct. Good knowledge and understanding are demonstrated.
6	Relevant mathematics commensurate with the level of the course is used. The mathematics explored is correct. Thorough knowledge and understanding are demonstrated.

Guidance

Your teacher will provide oral and/or written advice on a draft of your exploration pertaining to how it can be improved. Your teacher will also write thorough and descriptive comments on the final version of your exploration to assist IB moderators in confirming the criteria scores they've awarded.

Conducting an in-depth individual exploration into the mathematics of a particular topic can be an interesting and very rewarding experience. It is important to take all stages of your work on the exploration seriously – not only because it is worth 20% of your final grade for the course but also because of the opportunity to pursue your own personal interests without the pressure of examination conditions. The exploration should *not* be approached as simply an extended homework assignment. The task of writing the exploration will require you to analyse, think, write, edit, and use mathematics in a readable and focused manner. Hopefully, it will also be enjoyable, thought-provoking, and satisfying, and it should give you the opportunity to gain a deeper appreciation for the beauty, power, and usefulness of mathematics.

Warning: Failure to properly cite any text, diagrams, images, mathematical working, or ideas that are not your own may result in your exploration being reviewed for malpractice, which could have serious consequences.

Although it is required that your exploration is completely your own work, you should consult with your teacher on a regular basis. You are allowed to work collaboratively with fellow students, but this should be limited to the following: selecting a topic, finding resources, understanding relevant mathematical knowledge and skills, and receiving peer feedback on your writing. While you are encouraged to *talk* through your ideas with others, it is not appropriate for you to *work* with others on your exploration. Your teacher should provide support and advice during the planning and writing stages of your exploration. Both you and your teacher will need to verify the authenticity of your exploration.

If you are uncertain about the formatting and style of citations and a bibliography (not the same thing), then you should consult with teacher(s) at your school who have expertise in this area – such as an English teacher or librarian. A bibliography is required but it does not replace the need for appropriate citations (inline or footnotes) at the pertinent location in the exploration.

Any text, diagrams, images, mathematical working, or ideas that are not your own must be cited where they appear in your exploration. Otherwise, all of the work connected with your exploration must be your own. Your exploration must provide the reader with the exact sources of quotations, ideas, and points of view with a complete and accurate bibliography. There are a number of acceptable bibliographic styles. Whichever style you choose, it must include all relevant source information and be applied consistently. Group work is not allowed. Also, if you are writing an extended essay for mathematics, you are not allowed to submit the same or similar piece of work for the exploration – and you should not write about the same mathematical topic for both.

In organising a successful exploration, consider the following suggestions:

1. Select a topic in which you are *genuinely interested*. Include a brief explanation in the early part of your exploration about why you chose your topic – including why you find it interesting.

2. Consult with your teacher to confirm that the topic is at the *appropriate level of mathematics* – namely, that it is at the same or similar level of the mathematics in the SL syllabus.
3. Find as much *information* about the topic as possible. Although information found on websites can be very helpful, try to also find information in books, journals, textbooks, and other printed material.
4. Although there is no requirement that you present your exploration to your classmates, it should be written so that they can follow it without trouble. Your exploration needs to be *logically organised* and use appropriate mathematical terminology and notation.
5. The most important aspects of your exploration should be about *mathematical communication and using mathematics*. Although other aspects of your topic – for example, historical, personal, cultural – can be discussed, be careful to keep focus on the mathematical features.
6. Two of the assessment criteria – Personal engagement and Reflection – are about *what you think about the topic* you are exploring. Don't hesitate to pose your own relevant and insightful questions as part of your exploration – and then to address these questions using mathematics at a suitably sophisticated level along with sufficient written commentary.
7. Although your teacher will expect and require you to work independently, you are allowed to *consult with your teacher* – and your teacher is allowed to give you advice and feedback to a certain extent while you are working on your exploration. It is especially important to check with your teacher that any *mathematics in your exploration is correct*. Your teacher will not give mathematical answers or corrections but can indicate where any errors have been made or where improvement is needed.

Mathematical exploration - SL student checklist

Warning: Although you will need to conduct some research, your exploration is *not* a research paper where you simply report what you've learned about a mathematical topic. You must discuss your thoughts about the mathematics and you must apply the mathematics in a way that clearly demonstrates your knowledge and understanding of the mathematics.

Keep in mind that you should write your exploration so that a student in your Applications and Interpretation SL class can understand it. That is, *your audience is your fellow students.*

- Is your exploration written entirely by yourself? Have you avoided simply replicating work and ideas from sources you found? ☐ Yes ☐ No
- Have you strived to apply your personal interest, develop your own ideas, and use critical thinking skills during your exploration? ☐ Yes ☐ No
- Did you refer to the five assessment criteria while writing your exploration? ☐ Yes ☐ No
- Does your exploration focus on good mathematical communication – and does it read like an article from a mathematical journal? ☐ Yes ☐ No
- Does your exploration have a clearly identified introduction and conclusion? ☐ Yes ☐ No
- Have you provided appropriate citation for any ideas, mathematical working, images, graphs, etc. that are not your own at the point they appear in your exploration? ☐ Yes ☐ No
- Not including the bibliography, is your exploration 12 to 20 pages? ☐ Yes ☐ No
- Are graphs, tables, and diagrams sufficiently described and labelled? ☐ Yes ☐ No
- To the best of your knowledge, have you used mathematics that is at the same level, or similar, to that studied in Applications and Interpretation SL? ☐ Yes ☐ No
- Have you attempted to discuss mathematical ideas, and use mathematics, with a sufficient level of sophistication and rigour? ☐ Yes ☐ No
- Are formulae, graphs, tables, and diagrams in the main body of text? (Preferably no full-page graphs, and no separate appendices.) ☐ Yes ☐ No
- Have you used technology – such as a GDC, spreadsheet, mathematics software, drawing and word-processing software – to enhance mathematical communication? ☐ Yes ☐ No
- Have you used appropriate mathematical language (notation, symbols, terminology) and defined key terms? ☐ Yes ☐ No
- Is the mathematics in your exploration performed precisely and accurately? ☐ Yes ☐ No
- Has calculator/computer notation and terminology been used? ($y = x^2$, not $y = x$^2; π, not <pi>; $\approx$, not 'approximately equal to'; $|x|$, not abs(x); etc) ☐ Yes ☐ No
- Have you included reflective and explanatory comments about the topic being explored throughout your exploration? ☐ Yes ☐ No

Finding, developing, and choosing a topic for your exploration

It is fair to say that the *most important stage* of completing your exploration is determining the mathematical topic you are going to investigate, write about, and apply. Your exploration is much more likely to be successful – and gratifying – if it focuses on a mathematical topic in which you have a genuine interest, is at a suitable level for the Applications and Interpretation SL course, and for which you are confident that you can discuss and use the relevant mathematics in a manner that demonstrates thorough knowledge and understanding. There is no single approach for determining an exploration topic that is guaranteed to be successful for all students. Your teacher will provide helpful advice and support. Your teacher may supply you with a short list of some broad stimuli to start the process of finding a much narrower topic. Many teachers have found that starting with a sufficiently narrow topic is often more successful than starting with a very broad topic that requires a significant effort to reduce to the extent that it can be explored in less than 20 pages (double spaced).

Avoid choosing a topic that is too broad and/or too complicated.

In the eBook for this textbook you will find a list of 200 mathematical topics. Some of the topics in the list are broad but many are already quite narrow in scope. It is possible that some of these 200 topics could be the focus of an exploration, while others will require you to investigate further to develop a narrower focus to explore. Do not restrict yourself to the topics in the list. This list is only the tip of the iceberg with regard to potential topics for your exploration. Reading through this list may stimulate you to think of some other topic(s) that you may find interesting to explore. Many of the items in the list may be unfamiliar to you. A quick search on the internet should give you a better idea what each is about and help you determine if you're interested enough to investigate further – and to see if it might be a suitable topic for your exploration.

Theory of knowledge

At the start of his wonderful book *Nature's Numbers,* the mathematician Ian Stewart writes:

> *'We live in a universe of patterns. Every night the stars move in circles across the sky. The seasons cycle at yearly intervals. No two snowflakes are ever exactly the same, but they all have sixfold symmetry. Tigers and zebras are covered in patterns of stripes, leopards and hyenas are covered in patterns of spots. Intricate trains of waves march across the oceans; very similar trains of sand dunes march across the desert. Coloured arcs of light adorn the sky in the form of rainbows, and a bright circular halo sometimes surrounds the moon on winter nights. Spherical drops of water fall from clouds.'*

We could add to Stewart's list. Wallpaper is patterned (there are surprisingly only 17 different distinct groups of possible patterns); buildings often exhibit mirror symmetry and their structure is carefully proportioned; the digital traces on memory sticks or hard drives are patterned in a way that makes them suitable for storing data; mechanical devices such as clocks and engines depend on symmetry and patterning for their smooth movement; the day is divided into equal parts that are represented using angles or digits; music possesses horizontal and vertical symmetries – and human behaviour is patterned.

It is no accident that the world is full of patterns. Symmetry in a building is not only easy on the eye but it ensures that the design is simple. Pattern is a labour-saving strategy. The same plan can be used for each window, or the plan for one side of a building can be used in reverse for the other side. These informational shortcuts can be found both in the man-made world and in nature. The same blueprint for generating twig patterns can be used for bigger branches, or one plan can be used for all the petals in a flower. It is a sort of design efficiency. The wealth of patterns in the world is a series of cost-effective solutions to problems – and that is why these patterns are worth studying.

Mathematics is one way in which human beings formally study patterns. While the natural sciences study patterns by going out into the world, collecting examples and analysing them, mathematics studies patterns in the abstract. Mathematics in its purest form is not fieldwork or experiment. Its raw materials are abstract structures specified by symbols, and mathematicians arrive at conclusions through their manipulation. In this sense, mathematics is a little 'other-wordly' – a characteristic that makes it interesting from a theory of knowledge (ToK) perspective. It means that in some sense, mathematics is more like an art than a science. There is in this suggestion more than a hint of a deep reliance on creativity and imagination. A comparison with the arts and the sciences is instructive and reveals the truly special place that mathematics occupies in human knowledge.

In this chapter, we will investigate mathematics using the basic structure of the knowledge framework: Perspectives, methods and tools, and the link to the individual.

Under 'Perspectives', we will look at the orientation of mathematics within the academy. There are a number of key questions to be answered here:

- What is mathematics about?
- How should we think of mathematics: as a human construction or something in the world?
- Why is mathematics useful?

Under 'Methods and Tools' we will discuss exactly what mathematicians do – how they arrive at mathematical knowledge and what counts as facts and truth in mathematics. This is where we unpack the key conceptual building blocks of mathematical thought.

The final section deals with mathematics and the individual. What is the link between mathematics and supposedly subjective phenomena such as beauty? How reliable are our mathematical intuitions? Is mathematics a personal journey or is it something that we collaborate on?

On the way, we will have fun with infinite numbers, self-similar patterns and security codes. While it might be removed from the physical world, the world of mathematics is just as fascinating, if not more so. Enjoy!

What role does mathematics play in your life?

Perspectives

Mathematics and number

As a first definition, let's say that mathematics is the formal study of patterns. In this section we will see how far this basic idea takes us.

Imagine a simple pattern in the world – a set of similar objects, for example, a field of animals. Let's say that the animals are of the same kind – they are cows. To recognise that a group of different things all belong to the same kind is already remarkable. It means ignoring all the things that mark out individual animals and focusing only on what they have in common. Grouping a set of things together by common characteristics is a powerful technique in the sciences. If such a classification is effective, it might yield understanding, generalisations and predictions. We call groups that have these properties **natural kinds** – it is something that might be expected to happen in biology. But mathematics goes one step further. Suppose that we make a mark 'I' on a clay tablet for every cow in the field. We end up with a mark 'IIIIIII'. What we have done now is to abstract away everything about the animals in the field: the fact that they are animals, that they are cows, that they are eating grass. What is left is their number.

So, the simplest pattern that we can deal with abstractly is number. In a somewhat magical way, the inscriptions of the tablet **represent** the cows in the field. They are a convenient stand-in for the real world. If we want to find out what happens when we remove 'III' cows from the field, we can either move them physically or simply separate the 'cow' symbols: 'IIIII III'. Manipulating the symbols is clearly easier to perform. Mathematics manipulates representations rather than the real world because it is easier.

We do not know if something like this story is accurate at the beginning of the long history of mathematics. But we do know that imprints on a Sumerian clay tablet led eventually to the astounding sophistication of the proof of Fermat's last theorem and to modern algebra, analysis, and geometry. Mathematics has been shaped by the job it is expected to perform and through countless quirks of culture. Improvised methods designed to deliver a temporary solution to an unforeseen problem become permanent. If they work well, they get passed on and take on a life of their own. Less good solutions eventually fall into disuse in a sort of Darwinian selection of competing ideas. We could call histories like this **cultural evolution**.

But has the counting of cows in a field really got anything to do with modern mathematics? Let's examine the example more closely. We add an 'I' on the tablet for each cow in the field, subject to two strict rules: no cow should be 'counted' more than once and all the cows in the field are counted. Although these rules are quite natural to us, they are mathematically sophisticated. Mathematically, we are establishing a mapping between the marks on the tablet and the cows in the field that is a **one-to-one correspondence**. This means a mapping links a mark to a unique cow (injective) and that all cows in the field are linked (surjective). While these early users of mathematics might not have understood it quite in these terms, they nonetheless needed to use these properties when counting. But there is something else at work here. The compound symbol 'IIIIIIII' stands for the whole field of cows. It is a property of the whole set. It expresses the size of the set or its **cardinality**. The counting of cows in a field has a lot to do with the deep nature of mathematics itself.

Indeed, there are three more ideas illustrated by this simple example. The first is the power of numbers to create ordering: I II III IIII is such an ordering. This is called the **ordinal** property of number. Second, it illustrates the special place of sets and mappings in mathematics. We focused on the set of cows and the set of marks on the tablet. Third, we counted the first set by establishing a one-to-one correspondence with the second. This is a technique that works with any sets, including those that have infinitely many members. Mathematics is truly about sets and the mappings between them.

By representing the real world by marks bearing a special relation to their targets, human beings initiated perhaps the most extraordinary technical advance in their history: the invention of symbolic representation. Manipulating symbols is easier than manipulating objects in the world. Moreover, symbols allow this information to be communicated over distance and time. But the most powerful feature of symbols is that they can be used to represent states of affairs that are not physically present. Symbols can represent past worlds, possible worlds, and desired future worlds. Symbols allow us to tell stories, write histories, and make plans. Symbols that do not actually correspond with the world are called **counterfactuals**. They describe 'what if' situations. What if the Allies had lost World War II? What if we add sulfuric acid to copper? What if we wake up one morning to discover that we have been transformed into a giant insect? What if parallel lines could actually meet? What if there was a solution to the equation $x^2 = -1$? The power of symbolic representation is that it allows us to build abstract worlds – virtual realities where the 'what if' conditions are true.

There is a sense in which the world of mathematics is one such virtual universe, containing all manner of exciting and weird things. Mathematicians discuss 11-dimensional hypercubes, infinite sets of numbers, infinite numbers, surfaces that turn you from being right-handed to left-handed as you traverse them, spaces where the angles of a triangle add up to more than 180 degrees, spaces where parallel lines diverge, systems where the order of the operation matters (where $A * B$ is not the same as $B * A$), vectors in infinite-dimensional space, series that go on forever, and geometric figures that are self-similar called fractals (where you can take a small piece of the original figure then enlarge it and it looks identical – truly identical – to the original). And all this started with the making of a simple mark on a clay tablet.

Mathematics uses symbols to describe these amazing structures in the basic language of sets and the mappings between them. Because symbols are abstract and not limited to representing things in the world, mathematicians can use their imaginations to create a virtual reality following its own rule system unhindered by what the world is really like, a **counterfactual world**. In this world, mathematicians can explore the patterns they encounter.

Yet mathematics is remarkably useful in this world. From building bridges to controlling strategy in football, mathematics lies at the heart of the modern world. If mathematics really is so other-worldly, how come it has so much to say about this one?

If symbolic representation is the most significant technical advance in history, what would you put in second place?

This is an important question that motivates much of what follows.

Purpose: mathematics for its own sake

ToK uses the map metaphor; knowledge is taken to be like a map that is used for a particular purpose, such as solving a particular problem or answering a question. The map is a simplified picture of the world and its simplicity is its strength. It ensures that we get the job done with the least cognitive cost. If this is right, then it is natural to ask about the purpose of this particular map. What problems does it solve or what questions does it answer? There seem to be two categories: those questions that occur strictly within the virtual reality of mathematics itself (mathematics for its own sake) and those that occur in the world outside (mathematics as a tool). These categories broadly correspond to two subdivisions of mathematics that are often two different departments within a university: pure mathematics and applied mathematics.

Let's start with pure mathematics. A typical example of a problem in this category is how to solve a particular type of equation.

An example of a problem in pure mathematics might be how to solve the equation

(1) $x^3 - 2x^2 - x + 2 = 0$

The task is to find a value for x that satisfies the equation. In books like this, there are many such equations and, in this context, they often have simple integer solutions. An initial strategy might be to try a value for x to see if it fits. If we try $x = 0$, then equation (1) gives us:

$0^3 - 2 \cdot 0^2 - 0 + 2 = 0$, i.e. $2 = 0$

which is clearly not true. So, we can say that $x = 0$ is not a solution to the equation.

But if we try $x = 1$, then equation (1) gives us:

$1^3 - 2 \cdot 1^2 - 1 + 2 = 0$

In other words, $1 - 2 - 1 + 2 = 0$ is true. So, $x = 1$ is a solution to the equation.

The trick now, as you know, is to factor out $(x - 1)$ from equation (1) to give:

(2) $(x - 1)(x^2 - x - 2) = 0$

We can now try to find values of x that make the second bracket in (2) equal to 0. This can be done either by trying out hopeful values of x (2 seems to be a good bet, for example) or using the quadratic formula. We end up with $x = 2$ or $x = -1$

The equation therefore has three solutions: $x = 1$ or $x = -1$ or $x = 2$

The history of these problems illustrates the great attraction of pure mathematics. Certainly, these problems were of interest from the 7th century in what is now the Middle East – the home of algebra. The great 11th century Persian mathematician and poet Omar Khayyam wrote a treatise about similar so-called cubic equations and realised they could have more than one solution. By the 16th century, cubic equations were of public interest. In Italy, contests were held to showcase the ability of mathematicians to solve cubic equations, often with a great deal of money at stake. One such contest took place in 1635 between Antonio Fior and Niccolò Tartaglia. Fior was a student of Scipione del Ferro, who had found a method for solving equations of the type $x^3 + ax = b$, which is known as the 'unknowns and cubes problem' (where a and b are given numbers).

Del Ferro kept his method secret until just before his death when he passed the method on to his student. Fior began to boast that he knew how to solve cubics. Tartaglia also announced that he had been able to solve a number of cubic equations successfully. Fior immediately challenged Tartaglia to a contest. Each was to give the other a set of 30 problems and put up a sum of money. The person who had solved the most after 30 days would take all the money.

Tartaglia had produced a method to solve a different type of cubic $x^3 + ax^2 = b$. Fior was confident that his ability to solve cubic equations would defeat Tartaglia and submitted 30 problems of the 'unknowns and cubes' type, but Tartaglia submitted a variety of different problems. Although Tartaglia could not initially solve the 'unknowns and cubes' type of equation, he worked hard and discovered a method to solve this type of problem. He then managed to solve all of Fior's problems in less than two hours. In the meantime, Fior had made little progress with Tartaglia's problems and it was obvious who was the winner. Tartaglia did not take Fior's money though; the honour of winning was enough.

Tartaglia represents the essence of the pure mathematician: someone who is intrigued by puzzles and has a deep desire to solve them. It is the problem itself that is the motivation, not possible real-world applications.

What other knowledge is worth pursuing for its own sake?

A modern example is the solution of Fermat's conjecture by Andrew Wiles. The French mathematician Pierre de Fermat wrote the conjecture in 1627 as a short observation in his copy of *The Arithmetics of Diophantus*.

The conjecture is that the equation

$$A^n + B^n = C^n$$

where A, B, C are positive integers and $n > 2$ has no solution. Despite a large number of attempts to prove it, the conjecture remained unproved for 358 years until Wiles published his successful proof in 1995. The proof is way beyond the scope of this book, but there have been a number of interesting books and TV programmes made about it, including Simon Singh *Fermat's Last Theorem* (1997) and the BBC Horizon programme *Fermat's Last Theorem* (1996). As mathematician Roger Penrose remarked, *'QED: how to solve the greatest mathematical puzzle of your age. Lock self in room. Emerge 7 years later'*.

Purpose: mathematical models

Unlike pure mathematics, which is about the solution of exclusively mathematical puzzles, applied mathematics is about solving real-world problems. The mathematics it produces can be just as interesting from an insider's viewpoint as the problems of pure mathematics (and often the two are inseparable), but a piece of applied mathematics is judged by whether it can be usefully applied in the world.

Here is an example of applied mathematics at work. This is a problem that could have been posed in this book or, indeed (and this is the point), in a physics course.

A stone is dropped down a 30 m well. How long will it take the stone to reach the bottom of the well, neglecting the effect of air resistance?

The typical way to solve this type of problem is to use what we call a **mathematical model**. The essence of mathematical modelling is to produce a description of the problem where the main physical features become variables in an equation which is then solved and translated back into the real world.

To model the situation above:

We know that the acceleration due to gravity is 9.8 m s^{-2}, and we also know that the distance travelled s is given by the equation:

$$s = \frac{1}{2}at^2, \text{ where } a = \text{acceleration and } t = \text{time}$$

So we substitute the known values into the equation and get:

$$30 = \frac{1}{2}(9.8)t^2$$

Rearranging the equation gives us:

$$\frac{60}{9.8} = t^2, \text{ so } t = \sqrt{\frac{60}{9.8}} = 2.47 \text{ seconds (3 s.f.)}$$

There are a number of points to make about the process here that are typical of mathematical models.

(1) The model neglects factors that are known to operate in the real-world situation. There are two big assumptions made: that the stone will not experience air resistance, which will act as a significant drag force, and that the acceleration due to gravity is constant.

(2) The model appeals to a law of nature. In this case, the law of acceleration due to gravity.

(3) The model uses values for constants that are established empirically. In this case, the acceleration due to gravity at the Earth's surface.

We know that neither of the assumptions in (1) is true. The effect of air resistance can be highly significant. We know that if you have the misfortune to fall from an airplane above 100 m or so, the height does not matter – the speed of impact with the ground will be the same, around 150 km h^{-1}, because of the effect of air resistance (of course, it matters how you fall). The changing strength of gravitational force is a less important factor for normal wells.

But if we are dealing with a well that is 4000 km deep, then this factor would be significant. The point is that the model is actually fictional (it even breaks a major law of physics). It could never be true in the sense of exactly corresponding to reality. However, it is a sort of idealisation that we accept because the model provides an approximation to the behaviour of the stone (although not such a good one for deeper wells) and more importantly it gives us understanding of the system. If we were to make the modelling assumptions more realistic, the mathematics in the model would become too complicated to solve easily. Points (2) and (3) show us that the actual content of the model depends on something outside mathematics – namely some well-established results in physics. The mathematics is only a tool, albeit an important one. A model is a mathematical map – a simplified picture of reality that is useful.

Another beautiful example is the Lotka–Volterra model of prey–predator population dynamics in biology. This model was proposed by Alfred Lotka in 1925 and independently by Vito Volterra in 1926.

The model assumes a closed environment where there are only two species, prey and predator, and no other factors. The rate of growth of prey is assumed to be a constant proportion A of the population. The rate at which predators eat prey is B, which is assumed to be a constant proportion of the product of predators and prey. The death rate of predators, C, is assumed to be a constant proportion of the population, and there is a rate of generation of new predators, D, dependent on the product of prey and predators.

These modelling assumptions give rise to a pair of coupled differential equations:

(1) $\dfrac{\mathrm{d}x}{\mathrm{d}t} = Ax - Bxy$

(2) $\dfrac{\mathrm{d}y}{\mathrm{d}t} = -Cy + Dxy$

A modern computer package gives the following evolution of prey and predators over time:

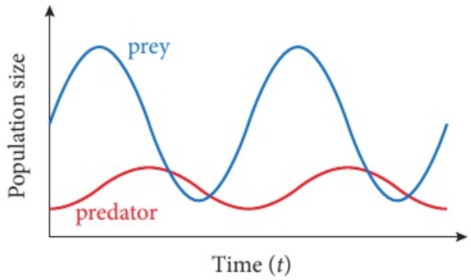


Figure 1 Evolution of prey and predator populations over time

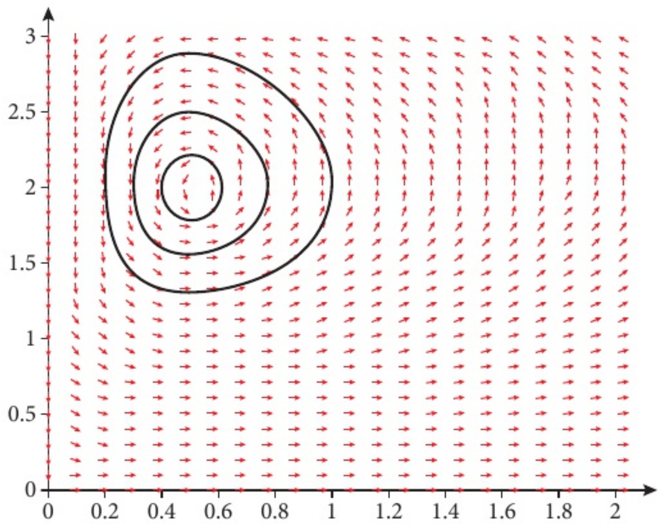


Figure 2 A phase space diagram. Number of prey (in units of 1000) on the *x*-axis, number of predators on the *y*-axis

It is interesting to look at a phase space diagram that represents each point (x, y) as a combination of numbers of prey and predators. Here the evolution of the system over time appears as a closed loop around the stationary point $\left(\frac{C}{D}, \frac{A}{B}\right)$, which is an 'attractor' of the dynamical system. (You could try to prove that this is a stationary point – it is not hard.) The position of an orbit around the attractor depends on the initial numbers of prey and predator. Notice that starting the model with too great a population of prey could end up with an extinction of predators (Figure 2) because the very high prey numbers leads to overpopulation of predators for whom there is not enough prey left to eat. The system itself is a nice example of circular causality.

As with the previous example, the modelling assumptions ensure that the mathematics of the model remains tractable, but the cost is that the model is not realistic. It is assumed that the prey do not die from natural causes or that the predators do not come into existence except through the provision of food. There is no competition between either prey or predators. Nonetheless, the model provides some important and powerful insights about the nature of population dynamics. As the model becomes more sophisticated and more factors are taken into consideration, not only does the mathematics become rapidly more difficult, but we lose sight of clear trends in the model (such as orbits around stationary points in phase space). We gain accuracy but lose understanding. This is a characteristic of both models and maps. A map that is as detailed as the territory it depicts is no use to anyone. It is precisely the simplification (literally what makes it false) that makes it useful. Virginia Woolf said about art, '*Art is not a copy of the world; one of the damn things is enough*', and the same could be said about models.

The distinction between pure and applied mathematics becomes blurred in the hands of someone like the great Carl Friedrich Gauss (1777–1855). He was perhaps happiest in the realm of number theory, which he called the 'queen of mathematics', and the idea that queens stay in their rarified towers and do not dirty their hands in the ways of the world was perhaps not so far from his thinking. He found great satisfaction in working with patterns and sequences of numbers. It is the same Gauss who, as a young man, enabled astronomers to rediscover the minor planet Ceres after they had lost it in the glare of the sun, by calculating its orbit from the scant data that had been collected on its initial discovery in 1801 and then predicting where in the sky it would be found more than a year later. This feat immediately brought Gauss to the attention of the scientific community. His skills as a number theorist presented him with the opportunity of solving a very real scientific problem.

Who would have guessed that recent work in prime number theory would give rise to a system of encoding data that is used by banks all over the world? The system is called 'dual key cryptography'. The key to the code is a very large number that is the product of two primes. The bank holds one of the primes and the client's computer the other. The key can be made public because in order for it to work it has to be split up into its component prime factors. This task is virtually impossible for large numbers. For example, present computer programs would take longer than the 13.8 billion years since the big bang to find the two prime factors of the number:

25 195 908 475 657 893 494 027 183 240 048 398 571 429 282 126 204 032
027 777 137 836 043 662 020 707 595 556 264 018 525 880 784 406 918 290
641 249 515 082 189 298 559 149 176 184 502 808 489 120 072 844 992 687
392 807 287 776 735 971 418 347 270 261 896 375 014 971 824 691 165 077
613 379 859 095 700 097 330 459 748 808 428 401 797 429 100 642 458 691
817 195 118 746 121 515 172 654 632 282 216 869 987 549 182 422 433 637
259 085 141 865 462 043 576 798 423 387 184 774 447 920 739 934 236 584
823 824 281 198 163 815 010 674 810 451 660 377 306 056 201 619 676 256
133 844 143 603 833 904 414 952 634 432 190 114 657 544 454 178 424 020
924 616 515 723 350 778 707 749 817 125 772 467 962 926 386 356 373 289
912 154 831 438 167 899 885 040 445 364 023 527 381 951 378 636 564 391
212 010 397 122 822 120 720 357

But this number is indeed of the form of the product of two large primes. If you know one of them, it takes an ordinary computer a fraction of a second to do the division and find the other.

Just as pure research in the natural sciences produced results that could also be used for technological or engineering applications, so in mathematics, problems motivated purely from within the most abstract recesses of the subject (pure mathematics) give rise to very useful techniques for solving problems with strong applications in the world outside of mathematics. Mathematicians often practise their art as art for its own sake. They are motivated by the internal beauty and elegance of their subject. Nevertheless, it often happens that pure mathematics created for no other purpose than solving internal mathematical problems turns out to have some extraordinary and very practical applications.

Can you think of an example of a model that does not represent the world well but is nonetheless useful?

What other examples are there of pure research that end up having immense practical benefit?

Constructivist view of mathematics

Having thought a little about what the purpose of mathematics could be, let's move on to the question of whether it is best thought of as an invention or as something out there in the world.

Broadly speaking, the **constructivist** views mathematics as a human invention. The vision we had of mathematics as a vast virtual reality limited only by the imagination and the rules that are installed there is a constructivist view. However, we are then bound to ask why mathematics has so many useful applications in the real world. Why is mathematics important when it comes to building bridges, doing science and medicine, economics and even playing basketball? Chess is also a game invented by humans, but it does not have very much use in the outside world. Constructivism cannot account for the success of mathematics in the outside world.

On this view, mathematics is what might be called a **social fact**. A social fact is true by virtue of the role that it plays in our social lives. Social facts do have real causal power in the world. That a particular piece of paper is money is a social fact that does make things happen. That piece of paper acquires its status ultimately from a whole set of social agreements. In the end, social facts are produced by **language acts** – performances that change the social world. A language act would be a registry officer saying 'I pronounce you married'. The use of language in a **performative** manner creates social facts. Social facts are no less real or definite than those about the natural world. The statement 'John is married' is definitely either true or false. One is reminded of the story about the little boy who, when asked by his grandmother what day it will be tomorrow, replies, 'Let's wait and see'. Social facts do not require us to wait and see. They rely on social agreements, not on empirical evidence.

The mathematician Reuben Hersh argues for a type of constructivism that he calls **Humanism**. For Hersh, numbers and other mathematical objects are social facts. Hersh defends this view on the Edge website:

'[Mathematics] … is neither physical nor mental, it's social. It's part of culture, it's part of history, it's like law, like religion, like money, like all those very real things, which are real only as part of collective human consciousness. Being part of society and culture, it's both internal and external: internal to society and culture as a whole, external to the individual, who has to learn it from books and in school. That's what math is.'

Hersh called his theory of mathematics humanism because it's saying that mathematics is something human. *'There's no math without people. Many people think that ellipses and numbers and so on are there whether or not any people know about them; I think that's a confusion.'*

Hersh points out that we do use numbers to describe physical reality and that this seems to contradict the idea that numbers are a social construction.

It is important to note here that we use numbers in two distinct ways: as nouns and as adjectives. When we say nine apples, nine is an adjective.

If it's an objective fact that there are nine apples on the table, that's just as objective as the fact that the apples are red, or that they're ripe or anything else about them; that's a fact. The problem occurs when we make a subconscious switch to 'nine' as an abstract noun in the sort of problems we deal with in Mathematics class. Hersh thinks that this is not really the same nine. They are connected, but the number nine is an abstract object as part of a number system. It is a result of our mathematics game – our deduction from axioms. It is a human creation.

Hersh sees a political and pedagogical dimension to his thinking about mathematics. He thinks that a humanistic vision of mathematics chimes in with more progressive politics. How can politics enter mathematics? As soon as we think of mathematics as a social construction then the exact arrangements by which this comes about – the institutions that build and maintain it – become important. These arrangements are political. Particularly interesting for us here is how a different view of mathematics can bring about changes in teaching and learning.

'Humanism sees mathematics as part of human culture and human history. It's hard to come to rigorous conclusions about this kind of thing, but I feel it's almost obvious that Platonism and Formalism are anti-educational, and interfere with understanding, and Humanism at least doesn't hurt and could be beneficial. Formalism is connected with rote, the traditional method, which is still common in many parts of the world. Here's an algorithm; practise it for a while; now here's another one. That's certainly what makes a lot of people hate mathematics (…) There are various kinds of Platonists. Some are good teachers, some are bad. But the Platonist idea, that, as my friend Phil Davis puts it, Pi is in the sky, helps to make mathematics intimidating and remote. It can be an excuse for a pupil's failure to learn, or for a teacher's saying "some people just don't get it". The humanistic philosophy brings mathematics down to earth, makes it accessible psychologically, and increases the likelihood that someone can learn it, because it's just one of the things that people do.'

There is a possibility that the arguments explored in this section might cast light on an aspect of mathematics learning that has seemed puzzling – why it is that mathematical ability is seen to be closely correlated with a certain type of intelligence. There is a widespread view that mathematics polarises society into two distinct groups: those who can do it and those who cannot. Those who cannot do it often feel the stigma of failure and that there is an exclusive club whose membership they have been denied. Those who can do it often find themselves labelled as 'nerds' or as people who are, in some sense, socially deficient. Is Hersh correct in attributing this to a formalistic or Platonic view? Is he right to suggest that if mathematics is just a meaningless set of formal exercises, then it will not be valued by society? If we deny that mathematics is out there to be discovered, it takes the stigma away from the particular individual who does not make the discovery. It is interesting to speculate on how consequences in the classroom flow from a humanist view of mathematics.

What are the strengths and weaknesses of mathematical humanism?

Platonic view of mathematics

One way to explain why mathematics applies so well to things like bridges and planets is simply to take mathematics as being out there in the world, independent of human beings. As with other things in the natural world, it is our task to discover it (literally to 'lift the cover'). This is called the **Platonic** view because the philosopher Plato (427–347 BC) took the view that mathematical objects belonged to the real world, underlying the world of appearances in which we lived. Mathematical objects such as perfect circles and numbers existed in this real world; circles on Earth were mere inferior shadows. Many mathematicians have at least some sympathy with this view. They talk about mathematical objects as though they had an existence independent of us and that we are accountable to mathematical truths in the same way as we are accountable to physical facts about the universe. They feel that there really is a mathematical world out there and that they are trying to discover truths about it, much like natural science discovers truth about the physical world.

This view is itself not entirely without problems. In ToK we might want to ask: 'If mathematics is out there in the world, where is it?' We do not see circles, triangles, $\sqrt{2}\,\pi$, i, e, and other mathematical objects obviously floating around in the world. We have to do a great deal of work to find them through inference and abstraction.

While this might be true, there is some evidence that mathematics is hidden not too far below the surface of our reality. Take prime numbers as an example. The Platonist might want to try to find them somewhere in nature. One place where she might start is in Tennessee. In the summer of 2016, the forests were alive with a cicada that exploits a property of prime numbers for its own survival. These cicadas have a curious life cycle. They stay in the ground for 13 years. Then they emerge and enjoy a relatively brief period courting and mating before laying eggs in the ground and dying. There is another species of cicada that has the same cycle and no fewer than 12 types that have a cycle of 17 years. There are, to add to the puzzle, none that have cycles of 12, 14, 15, 16 or 18 years. The clue is that 13 and 17 are prime numbers. There is a predator wasp that has evolved to have a similar life cycle. But if a predator had a life cycle of 6 years, the prey and the predator would only meet every $6 \times 17 = 112$ years. Whereas, if the cicada had a life cycle of 12 years, the prey and predator would meet every cicada cycle. Nature has discovered prime numbers through the cicada life cycles by evolutionary trial and error.

The relationship of nature to geometry was explored by the Scottish biologist D'Arcy Wentworth Thompson in his magnificent book of 1917, *On Growth and Form*. He explored the formation of shells and the wings of dragonflies, and examined the skeletons of dinosaurs through the eyes of a civil engineer constructing bridges and wondered about the formation of bee cells and the arrangement of sunflower seeds.

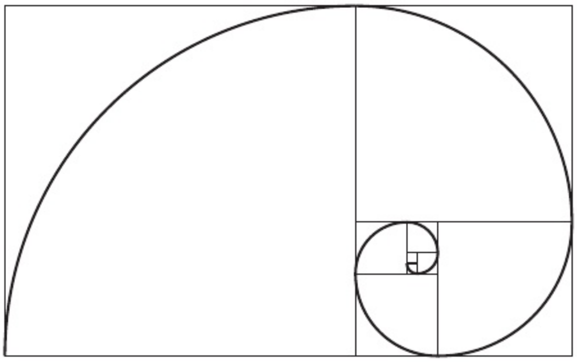


Figure 3 Spirals in nature

Many spirals in nature are formed, like the one in Figure 3, from the sequence:

1, 1, 2, 3, 5, 8, 13, 21, 34, …

This is called the Fibonacci sequence after the Italian mathematician Leonardo Pisano Bigolio (1170–1250), known as Fibonacci. The Fibonacci sequence is related to the golden number φ. The interested reader is referred to the many excellent sources on the internet.

Do you think that the mathematics teaching you have experienced reflects a Platonist or constructivist view of mathematics?

The methods and tools of mathematics

The language and concepts of mathematics

Knowledge in mathematics is like a map representing some aspect of the world. Like other areas of knowledge, it possesses a specialised vocabulary naming important concepts to build this map. Unlike some areas, this vocabulary is very precisely defined. This makes sense. If the world of mathematics is populated by some rather esoteric objects that are literally like nothing on Earth, then it is very important that these objects are precisely specified.

The other chapters in this book are all about establishing and using this very special vocabulary and becoming fluent in the methods that connect mathematical concepts into meaningful mathematical sentences. We will not spend too much time on these matters here, but there are a few aspects to highlight.

Notation

Since mathematical objects are abstract and we cannot point to them, we have to represent them with symbols. But the symbol and the idea are different things – there is a danger that we confuse them. Take representations of fractions. The symbols $\frac{1}{3}$, $\frac{2}{6}$, $\frac{3}{9}$, 0.3333… all represent the same number despite appearing to be quite different. (Perhaps the infinite number of ways of representing fractions is one of the reasons why some students have so much difficulty with them.) Some symbols such as $\frac{1}{0}$, $\sin^{-1}(1.2)$ or $\log(-2)$ have no meaning at all. More worrying is that an expression such as 'the smallest real number larger than 1' doesn't actually mean anything either. This is because there is no smallest real number larger than 1. (Think carefully about this.)

In a similar vein, the fact that there are different conventions for writing mathematics does not mean that the mathematics is different. Some conventions represent the number $\frac{3}{10}$ by the decimal 0.3, others by 0,3. Either way, the mathematics is the same and these do not really count as different mathematical cultures. Carl Friedrich Gauss, one of the greatest mathematicians of all time, said '*non notations, sed notions*' – not notations but notions.

Algebra

A staple method used in mathematics is the substitution of letters for numbers. In fact, mathematicians use letters for many sorts of mathematical objects, not just numbers. The reason is that they want to make generalised statements. By using a letter, they do not have to commit to making a statement about a specific number, but instead can make one about all numbers of a particular kind at once. This is a very powerful tool.

This is illustrated by a worked example. Imagine we want to prove that if we add an odd number to another odd number we get an even number. We hope to show that this is true for any choice of odd numbers. We could proceed by trying out different pairs of odd numbers and checking that the result is even:

$1 + 3 = 4$ even

$5 + 7 = 12$ even

$13 + 9 = 22$ even

$131 + 257 = 388$ even

You can see that this method will not serve as a proof because we would have to check every possible pair of odd numbers and, since this set is infinite, we would never finish. What we need is to define a general odd number without committing to a particular one. For example, we can define 'odd' by being 'one more than an even number'.

If k is an even number, then we can write $k = 2j$ for some whole number j.

If m is an odd number, then we can write $m = 2j + 1$ for some whole number j.

All we have to do now is to add two of these general odd numbers together.

So, we want to take two odd numbers, let's say $m = 2j + 1$ and $n = 2i + 1$ where j and i are whole numbers. There is a subtlety here because we use different letters j and i for the whole numbers in the expressions above because we want to allow m and n the possibility of being different odd numbers.

If we used the same letter, say j, in the expressions for m and n then we would be making our odd numbers equal and we would only have proved that if we add together two equal odd numbers, then the result is even.

Now we have to use some symbolic rules.

$$m + n = (2j + 1) + (2i + 1)$$

We can remove the brackets and rearrange to give:

$$m + n = 2j + 2i + 2$$

Finally, we can use the fact that 2 is a common factor of all terms in the expression to place it outside a bracket.

$$m + n = 2(j + i + 1)$$

But j, i and 1 are all whole numbers so $j + i + 1$ is also a whole number. Technically, this comes from the fact that the whole numbers are **closed under the operation of addition** because they form an important structure called a **group**. Let's call this whole number p.

So, we have that $m + n = 2p$. But this is precisely the definition of an even number that we started with. An even number is 2 times a whole number. This proves that any two odd numbers added together gives an even number.

The big chain of reasoning above is called a proof. It is immensely powerful because it covers an infinite number of situations. There is an infinite number of possible pairs of odd numbers to which the result applies. This is the power and beauty of using letters for numbers — a practice that was developed in Baghdad and Damascus about 1000 years ago. In one sense, mathematicians have a god-like ability when it comes to dealing with infinite sets.

Proof

Proof is the central concept in mathematics because it guarantees mathematical truth. When something is proved, we can say that it is true.

This type of truth is independent of place and time. In contrast to the science of the day, the mathematical truths of Pythagoras are just as true today as they were then – indeed his famous theorem is still taught today as can be seen in this book. But the science of the time has long been rejected. There were four chemical elements in the 4th century BC, and Aristotle thought that the heart was the organ for thinking. Actually, we do not have to go far back in time to find textbooks in the natural sciences that contain statements that we would dispute today. The truths of the natural sciences are always subject to revision, but mathematical truths are eternal.

But there is something even more striking about mathematical truths — that is, mathematical statements that have been proved. A statement such as 'odd + odd = even' has such power that we can say that it is certain. This is not just a matter of confidence – we are not talking about psychological certainty here.

It is certain because it cannot be otherwise. The negation of a mathematical truth (like 'odd + odd = odd') is to utter a self-contradiction or absurdity. Let's reflect on the power of this statement. This means that there is no possible world in which 'odd + odd = odd' (given the standard meanings of these terms). A story that makes this statement is describing a world that is self-contradictory – that is, an absurd and unintelligible world. Such a story is just not credible. But this means that mathematics is really radically different from other areas of knowledge, including the natural sciences. It is not a contradiction to say that the moon is composed of green cheese. There could be universes where this is true, but it just happens not to be in ours. Mathematics deals in what we call **necessary truths**, while the sciences deal mainly in **contingent truth**.

This is something that students of ToK should think about carefully. What is it about mathematical truth that makes it immune to revision and provides the basis for certainty and makes the negation of a mathematical truth a contradiction?

Recall that the constructivist sees mathematics as a big abstract game played by human beings according to invented rules. The hero of *The Glass Bead Game*, a novel by the German writer Hermann Hesse, must learn music, mathematics, and cultural history to play the game. On this view, mathematics is just like the glass bead game. There are parallels we can draw between a game like chess and mathematical proof. First, chess is played on a special board with pieces that can move in a particular way. The pieces must be set up on the board in a particular fashion before the game can begin. The same is true of mathematical proof. It starts with a collection of statements in mathematical language called **axioms**. They themselves cannot be proved. They are simply taken as self-evidently true and form the starting point for mathematical reasoning.

Once the game is set up, we can start playing. A move in chess means transforming the position of the pieces on the board by applying one of the game's rules that govern movement. Typically in chess, a move involves the movement of only one piece. (Can you think of an exception?) If the state of the pieces before the move was legitimate and the move was made according to the rules of the game, then the state of the pieces after the move is also legitimate. The same is true of a mathematical proof. One applies the rules (these are rules of algebra typically) to a line in the proof to get the next line. The whole proof is a chain of such moves.

Finally, the chess game ends. Either one of the players has achieved checkmate, or a stalemate (a draw) has been agreed. Similarly, a mathematical proof has an end. This is a point where the proof arrives at the required result at the end of the chain of reasoning. This result is called a **theorem**.

Once a proof of a mathematical statement is produced, we have a logical duty to believe the result, however unlikely. This is illustrated with a famous example.

Many people do not believe that $1 = 0.99999999\ldots$
(The three dots indicate that the 9's continue indefinitely).

The proof is straightforward.

Let	$x = 0.9999999\ldots$
Then	$10x = 9.9999999\ldots$
Subtract both equations	$10x - x = 9.99999999\ldots - 0.9999999\ldots$
This implies	$9x = 9$

Giving $x = 1$ as required.

0.999999… really does look very different to 1 but if the proof works then we are forced to believe that they are the same.

Are you happy with every stage of this proof?

Sets

A set is a collection of elements that can themselves be sets. They can be combined in various ways to produce new sets. The concepts of a set and membership of a set are **primitive**. This means that they cannot be explained in terms of more simple ideas. These seem to be rather modest beginnings on which to build the complexities of modern mathematics. Nevertheless, in the 20th century there were a number of projects that were designed to do just that: reduce the whole of mathematics to set theory. The most important work here was by Quine, von Neumann and Zermelo, and Bertrand Russell and Alfred North Whitehead in the three volumes of their *Principia Mathematica* of 1910–1913. Starting out with the notion of the empty set and the idea that no set can be a member of itself, we can construct the whole number system.

Mappings between sets

Once we have established sets in our mathematical universe, we want to do something useful with them. One of the most important ideas in the whole of mathematics is that of a mapping. A mapping is a rule that associates every member of a set with a member of a second set. This is what we were doing when we started this chapter by counting cows. We set up a one-to-one correspondence between a set of numbers and a set of cows.

Infinite sets

Consider the function $f(x) = 2x$ defined over the natural numbers.

Clearly it sets up a one-to-one correspondence between the set of natural numbers and the set of even numbers (check this yourself). So, this means that there are as many even numbers as there are natural numbers.

This is rather strange because we would think intuitively that there were more natural numbers than even numbers – they are after all the result of taking away an infinite number of odd numbers from the original set. But we are saying that the set that is left over has as many members as the original set. This strangeness is characteristic of infinite sets (indeed it can be used to define what we mean by infinite). Infinite sets can be put in a one-to-one correspondence with a proper subset of themselves.

But the story doesn't stop here. Using sets and mappings we can show that there are many different types of infinity. The set of natural numbers contains the smallest type of infinity, usually denoted by $\aleph_0$, which we call 'aleph nought'. In the 19th century, the German mathematician Georg Cantor showed by an ingenious argument that the number of numbers between 0 and 1 is a bigger type of infinity than aleph nought.

What is it about the difference between the methods of the natural sciences and mathematics that accounts for the radical difference in types of knowledge produced?

It turns out that there is an infinity of different types of infinity — a whole hierarchy of infinities, in fact — and this probably does not surprise you anymore, there are more infinities than finite cardinal numbers.

The methods and concepts of mathematics, therefore, are quite unlike anything to be found in the sciences, although they do seem to bear a strong resemblance to the arts in terms of the setting of the rules of the game and the use of the imagination. This is something we will explore in the next section.

Mathematics and the knower

English poet John Keats said, *"Beauty is truth, truth beauty – that is all / Ye know on earth, and all ye need to know."*

In this section we will see how mathematics impinges on our personal thinking about the world. One of the more surprising aspects of mathematics is the two-way link to the arts and beauty.

Beauty by the numbers

There is a long-held view that we find certain things beautiful because of their special proportions or some other intrinsic mathematical feature. This is the thinking that has inspired architects since the times of ancient Egypt and generations of painters, sculptors, musicians, and writers. Mathematics seems to endow beauty with a certain eternal objectivity. Things are beautiful because of the mathematical relationships between their parts. Moreover, this is a very public beauty because it can be dissected and discussed.

Let's take the example of the builders of the Parthenon. They were deeply interested in symmetry and proportion. In particular, they were interested in how to divide a line so that the proportion of the shorter part to the longer part is the same as that of the longer part to the whole. You can check that you get the quadratic equation $x^2 + x - 1 = 0$. One solution to this equation is the golden ratio $x = \frac{-1 + \sqrt{5}}{2} = 0.61803398875\ldots = \varphi$.

This proportion features significantly in the design of the Parthenon and many other buildings of the period. Since it is also related to the Fibonacci sequence, you will find φ turning up anywhere where there are spirals. It is used quite self-consciously in painting (Piet Mondrian, for example) and in music (particularly the music of Debussy). There are those who go as far as saying that it is present in the proportions of the perfect human figure and that we have a predisposition towards this ratio.

See if you can spot the connection between the golden ratio and the Fibonacci sequence. Hint: write down a difference equation for generating the sequence.

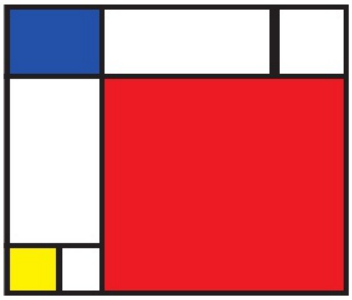


Figure 4 *Composition with Red, Blue and Yellow* (1926) Piet Mondrian. The proportions of some of the rectangles in this painting is φ

Beauty in numbers

Keats also put it the other way around: the beautiful is the true. Could we allow ourselves to be guided to truth in mathematics because of the beauty of the equations? This is a position taken by surprisingly many mathematicians. They look for beauty and elegance as an indicator of truth. Many mathematical physicists were guided in the 20th century by considerations of beauty and elegance.

Einstein suggested that the most incomprehensible thing about the universe was that it was comprehensible. From a ToK point of view, the most incomprehensible thing about the universe is that it is comprehensible in the language of mathematics. Galileo wrote, '*Philosophy is written in this grand book, the universe… It is written in the language of mathematics, and its characters are triangles, circles and other geometric figures…*'

Perhaps what is more puzzling is not just that we can describe the universe in mathematical terms, but that the mathematics we need to do this is mostly simple, elegant, and even beautiful.

To illustrate this, let's look at some of the famous equations of physics. Most people will be familiar with Einstein's field equations and Maxwell's equations.

EINSTEIN'S FIELD EQUATION
for general relativity

$$R_{\mu\nu} - Rg_{\mu\nu} = \frac{8\pi G}{C^4} T_{\mu\nu}$$

Figure 5 Einstein's field equation

1. $\nabla \cdot \mathbf{D} = \rho_V$
2. $\nabla \cdot \mathbf{B} = 0$
3. $\nabla \times \mathbf{E} = -\dfrac{\partial \mathbf{B}}{\partial t}$
4. $\nabla \times \mathbf{H} = \dfrac{\partial \mathbf{D}}{\partial t} + \mathbf{J}$

Figure 6 Maxwell's equations

It is perplexing that the whole crazy complex universe can be described by such simple, elegant, and even beautiful equations. It seems that our mathematics fits the universe rather well. It is difficult to believe that mathematics is just a mind game that we humans have invented.

But the argument from simplicity and beauty goes further. Symmetry in the underlying algebra led mathematical physicists to propose the existence of new fundamental particles, which were subsequently discovered. In some cases, beauty and elegance of the mathematical description have even been used as evidence of truth. The physicist Paul Dirac said, *'It seems that if one is working from the point of view of getting beauty in one's equations, and if one has really a sound insight, one is on a sure line of progress'.*

Dirac's own equation for the electron must rate as one of the most profoundly beautiful of all. Its beauty lies in the extraordinary neatness of the underlying mathematics – it all seems to fit so perfectly together:

$$(i\not{\partial} - m)\psi = 0$$

Figure 7 Dirac's equation of the electron

The physicist and mathematician Palle Jorgensen wrote:

> *'[Dirac] … liked to use his equation for the electron as an example stressing that he was led to it by paying attention to the beauty of the math, more than to the physics experiments.'*

It was because of the structure of the mathematics, in particular that there were two symmetrical parts to the equation — one representing a negatively charged particle (the electron) and the other a similar particle but with a positive charge — that scientists were led to the discovery of the positron. It seems fair to say that the mathematics did really come first here.

We will leave the last word on this subject to Dirac himself, writing in Scientific American in 1963:

> *'I think there is a moral to this story, namely that it is more important to have beauty in one's equations than to have them fit experiment.'*

By any standards this is an extraordinary statement for a mathematical physicist to make.

Mathematics and personal intuitions

Sometimes our intuition can let us down badly when it comes to making judgments of probability. Here is an example to illustrate how we might have to correct our intuition by careful mathematical reasoning.

Consider the following case. There is a rare genetic disease among the population. Very few people have the disease. As a precaution, a test has been developed to detect whether particular individuals have the disease. Although the test is quite good, it is not perfect — it is only 99% accurate. Person X takes the test and it shows positive. The question for your mathematical intuition is: 'What is the probability that X actually has the disease?' (You should recognise this as being a problem of conditional probability.)

Think about this for a moment before we continue.

Many of the students (and teachers) we have worked with in the past give the same answer: the probability that X actually has the disease given a positive test result is about 99%. Did you say the same? If you did, then your mathematical intuition let you down – very badly.

Let's put some numbers into the problem to illustrate this. For the sake of simplicity, assume that the country in which the test takes place has a population of 10 million. We are told that the disease is very rare. Assume that only 100 people in the whole country have the disease. We are told that the test is 99% accurate so, of the 100 cases of the disease the test would show positive in 99 cases and negative in one. So far so good.

Now consider the 9 999 900 people who don't have the disease. In 99% of these cases the test does its job and records a negative result. In 1% of the cases however it gets it wrong and produces a positive result. 1% of 9 999 900 is 99 999. So, of the whole population tested there would be a total of $99\,999 + 99 = 100\,098$ positive results. But of these only 99 have the disease. Therefore, the probability of having the disease given a positive result is $\frac{99}{100\,098} = 0.0989\%$ or about 1 in 1000. This is quite a big difference from the 990 in 1000 that we expected intuitively. That is out by a whopping 99 000%.

What went wrong with intuition here?

Mathematics and personal qualities

There are undoubtedly special qualities well-suited to doing mathematics. There are a host of great mathematicians from Archimedes, Euclid, Hypatia, through to Andrew Wiles, Grigori Perelman, and Maryam Mirzakhani, who contributed significantly to the area. Maryam was the first woman to receive the Fields medal (the equivalent of the Nobel prize in mathematics). Although mathematics is collaborative in the sense that mathematicians build on the work of others and take on the challenges that the area itself has recognised as being important, it is nevertheless largely a solitary pursuit. It requires great depth of thought, imaginative leaps, careful and sometimes laborious computations, innovative ways of solving very hard problems, and, most of all, great persistence. Mathematicians need to develop their intuition and their nose for a profitable strategy. They are guided by emotion and by hunches – they are a far cry from the stereotype of the coldly logical thinker who is closer to computer than human.

Conclusion

We have seen that mathematics is really one of the crowning achievements of human civilisation. Its ancient art has been responsible for some of the most extraordinary intellectual journeys taken by humankind, and its methods have allowed the building of great cities, and the production of great art, and it has been the language of great science.

From a ToK perspective, mathematics, with its absolute and unchanging notion of necessary truth, makes a good contrast to the natural sciences with their reliance on observation of the external world, experimental method, and provisional nature of its results.

Two countering arguments should be set against this view of mathematics. The idea that the axioms of mathematics (the rules of the game) are arbitrary both deprives mathematics of its status as something independent of human beings, and makes it vulnerable to the charge that its results cannot ever be entirely relevant to the world outside mathematics.

Platonists would certainly argue that mathematics is out there in the universe, with or without human beings. They would argue that it is built into the structure of the cosmos – a fact that explains why the laws of the natural sciences lend themselves so readily to mathematical expression.

Both views produce challenging questions in ToK. The constructivist is a victim of the success of mathematics in fields such as the natural sciences. She has to account for why mathematics is so supremely good at describing the outside world to which, according to this view, it should ultimately be blind. The Platonist, on the other hand, finds it hard to identify mathematical structures embedded in the world or has a hard time explaining why they are there.

We have seen how mathematics is closely integrated into artistic thinking; perhaps because both are abstract areas of knowledge indirectly linked to the world and not held to account through experiment and observation, but instead, open to thought experiment and leaps of imagination. Mathematics can challenge our intuitions and can push our cognitive resources as individual knowers. Infinity is not something that the human mind can fathom in its entirety. Instead, mathematics gives us the tools to deal with it in precisely this unfathomed state. We can be challenged by results that seem counter to our intuition, but ultimately, the nature of mathematical proof is that it forces us to accept them nonetheless. In turn, individuals can, through their insight and personal perspectives, make ground-breaking contributions that change the direction of mathematics forever. The history of mathematics is a history of great thinkers building on the work of previous generations to do ever more powerful things using ever more sophisticated tools.

The Greek thinkers of the 4th century BC thought that mathematics lay at the core of human knowledge. They thought that mathematics was one of the few areas in which humans could apprehend the eternal forms only accessible to pure unembodied intellect. They thought that in mathematics they could glimpse the very framework on which the world and its myriad processes rested. Maybe they were right.

Answers

Chapter 1

Exercise 1.1

1. (a) 208 (b) 200 (c) 207.63 (d) 208
   (e) 207.6 (f) 210
2. (a) (i) 5 (ii) 45800
   (b) (i) 3 (ii) 25300
   (c) (i) 4 (ii) 1.00×10^3
   (d) (i) 4 (ii) 0.0350
   (e) (i) 4 (ii) 2.35
   (f) (i) 5 (ii) 4.00
   (g) (i) 1 (ii) 20.0
   (i) (i) 1 (ii) 0.0100
   (i) 0.0 has no significant figures
3. (a) $22.5\,\text{kg} \leqslant M < 23.5\,\text{kg}$ (b) $2.65\,\text{s} \leqslant t < 2.75\,\text{s}$
   (c) $5.225\,\text{N} \leqslant F < 5.235\,\text{N}$ (d) $0.0195\,\text{W} \leqslant P < 0.0205\,\text{W}$
4. (a) 2.2% (b) 1.7% (c) 2.1% (d) 0.05%
5. (a) 1.62 (b) 1.1% (c) Requires research
6. (a) 210.25, 2.10×10^2 (b) 0.12%
   (c) 3.81. Irrational numbers cannot be written as exact decimals.
   (d) $3.805 \leqslant \sqrt{x} \leqslant 3.815$
7. (a) 20.6 cm, 21 cm correct to 2 significant figures
   (b) 1.4 cm (c) 2.9%
8. (a) $0.885\,\text{m} \leqslant L < 0.895\,\text{m}$
   $2.05\,\text{m} \leqslant W < 2.15\,\text{m}$
   $1.5\,\text{m} \leqslant H < 2.5\,\text{m}$
   (b) $2.72 \leqslant V \leqslant 4.81$
   (c) An estimate could be $0.895\,\text{m} \times 2.15\,\text{m} \times 2.5\,\text{m}$. The percentage error would be 29%
   (d) \$410.04 $\leqslant$ cost $<$ \$725.11
9. (a) 22
   (b) $74.5\,\text{kg} \leqslant M < 75.5\,\text{kg}$
   $1.845\,\text{m} \leqslant h < 1.855\,\text{m}$
   $\text{BMI}_{\text{min}} = 21.7$ $\text{BMI}_{\text{max}} = 22.2$
   (c) 1.4%
   (d) (i) If heights are equal, BMI decreases with mass
   (ii) If masses are equal, BMI decreases as height increases.
   (iii) Research question
10. (a) (i) 7.07 s (ii) 7.07 s
    (b) $7.053\,\text{s} \leqslant T < 7.082\,\text{s}$
    (c) Minimum time – 423 s or 7 min 3 s
    Maximum time – 425 s or 7 min 5 s
    (d) 3.8%
11. (a) scale 1 – 0.21% scale 2 – 2.1%
    (b) Increasing the precision from 1 kg to 0.1 kg reduced the percent error by a factor of ten.
    (c) For scale 1, the difference between maximum and minimum mass of the suitcases is 10 kg, whereas for scale 2, this difference is 100 kg. There is an increased chance of exceeding the maximum tolerated mass for the plane using scale 2.

Exercise 1.2

1. (a) x^7 (b) x (c) x^6 (d) $16x^4$
   (e) $\frac{1}{x^3}$ (f) $x^{\frac{1}{7}}$ (g) $x^{\frac{2}{3}}$ (h) 1
   (i) x^{14} (j) 8 (k) $\frac{1}{3x^{\frac{3}{2}}}$ (l) $\frac{x^{12}}{y^6}$
   (m) $\frac{x^{35}}{y^7}$ (n) $\frac{9}{4x^4y^6}$ (o) $\frac{1}{x^{\frac{7}{4}}}$
2. (a) 125 (b) $\frac{4}{9} \approx 0.44$ (c) $\frac{1}{64} \approx 0.015625$
   (d) 10 (e) 243 (f) 16
3. (a) $p = 6$ (b) $p = 0$ (c) $p = -1$ (d) $p = \frac{1}{5}$
   (e) $p = \frac{3}{2}$ (f) $p = 1$
4. 2^{-6}
5. (a) $49.0\,\text{cm}^2$ (b) $(\sqrt{2})^3 = 2^{\frac{3}{2}} \approx 2.83\,\text{cm}^3$
   (c) $\sqrt[3]{20} = 20^{\frac{1}{3}} \approx 2.71\,\text{cm}$
6. 1, 64, 729
7. (a) $512 = 2^9 = (2^3)^3 = 8^3$
   (b) The exponent is a prime number.
   (c) The exponent P is the product of two primes.
8. x can be any positive real number, any negative integer and any negative rational number whose fraction representation in lowest terms has an odd denominator.
9. $2^{\sqrt{2}} \approx 2.665$, $2^{\pi} \approx 8.825$. Since $\sqrt{2}$ and π are irrational numbers, they cannot be written as fractions of integers, and therefore do not align with any of our exponent laws. (These exponents are calculated using a process of representing irrational numbers as infinite series of rational numbers and using the existing exponent laws to obtain an approximation)
10. 0^0 gives an error if entered into a calculator because it is undefined. If we consider x^0, $x = 0.1, 0.01, 0.001...$ all values are 1. However, if we consider 0^x, $x = 0.1, 0.01, 0.001....$ all values are 0. Since both expressions approximate 0^0, they must give equal values, which they do not.
11. Answers may vary. Students should consider models of exponents of different types, positive integers, negative integers, zero and rational and compare with the repeated multiplication model.
12. For $m > n$, $a \neq 0$, $\frac{a^m}{a^n} = \frac{a^{m-n}a^n}{a^n} = a^{m-n}$ using Exponent Law 1.

Exercise 1.3

1. (a) 1.0807×10^4 (b) 9.83×10^{-3}
   (c) 3.45×10^4
2. (a) 9.45×10^7 (b) 1.024×10^{-3}
   (c) 5×10^{-2} (d) 2.9×10^4
3. $5.478 \times 10^{-4}\,g$
4. (a) 2.06×10^6 (b) 21%
5. $2.2 \times 10^{-30}\,\text{N}$
6. $9.47 \times 10^{12}\,\text{km}$

Exercise 1.4

1. (a) $3 = \log 1000$ (b) $3 = \log_4 64$
   (c) $\frac{3}{2} = \log_{100} 1000$ (d) $\frac{1}{2} = \log_9 3$
   (e) $\frac{1}{2} = \log_8 2\sqrt{2}$ (f) $0 = \log 1$
   (g) $0 = \ln 1$ (h) $-2 = \log_6 \frac{1}{36}$

(i) $-2 = \log_{\sqrt{2}}\frac{1}{2}$
(j) $-\frac{1}{2} = \log_3\frac{1}{\sqrt{3}}$
(k) $-3 = \log_{\frac{1}{2}}8$
(l) $-\frac{1}{2} = \log_8\frac{\sqrt{2}}{4}$
(m) not possible
(n) $-1 = \log_{0.01}100$
(o) $3 = \log_{\frac{\sqrt{2}}{2}}\frac{\sqrt{2}}{4}$

2. (a) $x = \log_2 y$ (b) $x = \log y$
(c) $x = \ln y$ (d) $x = \frac{1}{3}\log_2 y$
(e) $x = \log_2\frac{y}{3}$ (f) $x = \log_2(5 - y)$
(g) $x = \frac{1}{2}\log_3 y$ (h) $x = 2\log_3 y$
(i) $x = \frac{1}{2}\ln y$ (j) $x = \log_{2(y)} + 3$
(k) $x = 2\ln y$ (l) $x = \frac{1}{2}\ln 2y$

3. (a) $R = 6.2$ (b) $R = 5.5$

Chapter 1 Practice questions

1. (a) $p = 1.75$
(b) (i) $x = 2, y = 1, z = 50$ (ii) 1.98
(c) 13.1%
2. (a) $p = 9.3 \times 10^3$ cm (b) 5 280 000 cm^2
3. (a) 6900 km (b) 43354 km (c) 4.3354×10^4 km
4. (a) $\frac{224}{3}$ m (b) 20.5%
5. (a) $1.52 \times 10^6\ \text{m}^2$ (b) 5.26%
6. (a) 0.3% (b) 0.05%
(c) In some cases, rounding to 2 decimal places produces very small error.
7. (a) $1.885\text{ m} \leqslant h < 1.895\text{ m}$, $81.5\text{ kg} \leqslant m < 82.5\text{ kg}$
(b) $\text{BMI}_{\min} = 22.7$, $\text{BMI}_{\max} = 23.2$
(c) The largest percentage error in the BMI calculation is 2.2%
8. (a) City A is 2.2 times larger than city B, correct to 1 decimal place.
(b) 1.88×10^6
(c) 44.6%
9. (a) $1.93 \times 10^3\ s$ (b) 32 minutes
10. (a) 64 (b) $\frac{16}{9}$ (c) $\frac{1}{8}$ (d) 9
11. 2^{-11x}
12. (a) 2.23×10^3 (b) 5
(c) 1.22×10^{10} (d) 7.93×10^4
13. (a) x^{11} (b) x^7 (c) x^8 (d) $16x^4$
(e) $\frac{1}{x^2}$
14. (a) $p = 7$ (b) $p = 0$ (c) $p = -1$ (d) $p = \frac{1}{4}$
(e) $p = \frac{4}{3}$
15. (a) 2 (b) -1 (c) 0.5 (d) 2.11
16. $x \approx 1.40$
17. (a) 25 primes between 1 and 100
(b) $\frac{100}{\ln(100)} \approx 21.7$
(c) 13%
18. (a) For sulfuric acid, $[H^+] = 10^{-2}$ Mol/L and for hydrochloric acid $[H^+] = 10^{-5}$ Mol/L
(b) 1000 times more acidic.
19. 20 times more intense
20. $y = \log_a x$ or $y = \frac{\log x}{\log a}$

Chapter 2

Exercise 2.1

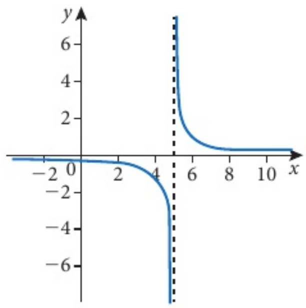
1. (a) $x = h - \frac{n}{m}$ (b) $a = \frac{v^2 + t}{b}$
(c) $b_1 = \frac{2A}{h} - b_2$ (d) $r = \sqrt{\frac{2A}{\theta}}$
(e) $k = \frac{gh}{f}$ (f) $t = \frac{x}{a + b}$
(g) $r = \sqrt[3]{\frac{3V}{\pi h}}$ (h) $k = \frac{g}{F(m_1 + m_2)}$
2. (a) $y = -\frac{2}{3}x - 5$ (b) $y = -4$
(c) $y = \frac{5}{4}x - 6$ (d) $x = \frac{7}{3}$
3. $y = -4x + 11$
4. $y = -\frac{5}{2}x - 7$
5. (a) $(5, 1)$ (b) $\left(4, \frac{1}{2}\right)$
6. (a) $\frac{5}{7}, \frac{3}{7}$ (b) $(-3, -8)$
7. (a) $(5, -3)$
(b) $(9.6, 7.6)$ (c) $\left(\frac{11}{19}, -\frac{18}{19}\right)$
8. (a) $0, -3$ (b) 2 (c) 2 (d) $-\frac{1}{2}$
(e) $x + 2y - 10 = 0$

Exercise 2.2

1. (a) G (b) L (c) H (d) K
(e) J (f) C (g) A (h) I
(i) F
2. $A = \frac{C^2}{4\pi}$
3. $A = \frac{l^2\sqrt{3}}{4}$
4. $A = 4x^2 + 60x$
5. $h = x\sqrt{2}$
6. (a) 9.4 (b) $V = \frac{3525}{P}$
7. (a) $F = kx$ (b) 6.25 (c) 37.5 N
8. (a) $\{-6.2, -1.5, 0.7, 3.2, 3.8\}$
(b) $r > 0$
(c) $\mathbb{R}$
(d) $\mathbb{R}$
(e) $t \leqslant 3$
(f) $\mathbb{R}$
(g) $x \neq \pm 3$
(h) $-1 \leqslant x \leqslant 1$ and $x \neq 0$
9. no, $x = c$ is a vertical line
10. (a) (i) $\sqrt{17}$
(ii) 7
(iii) 0
(b) $x < 4$
(c) domain: $x \geqslant 4$, range: $h(x) \geqslant 0$

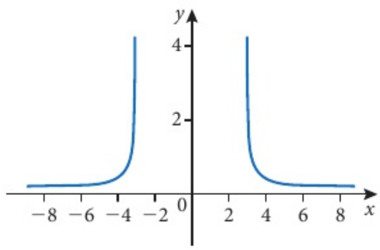
11. (a) (i) domain $\{x : x \in \mathbb{R}, x \neq 5\}$, range $\{y : y \in \mathbb{R}, y \neq 0\}$

(ii) y-intercept $\left(0, -\frac{1}{5}\right)$, vertical asymptote $x = 5$, horizontal asymptote $y = 0$



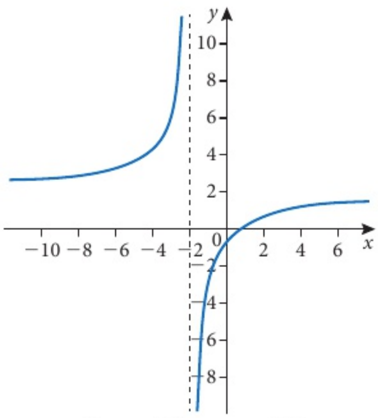
(b) (i) domain $\{x : x < -3, x > 3\}$, range $\{y : y > 0\}$

(ii) vertical asymptotes $x = -3$ and $x = 3$



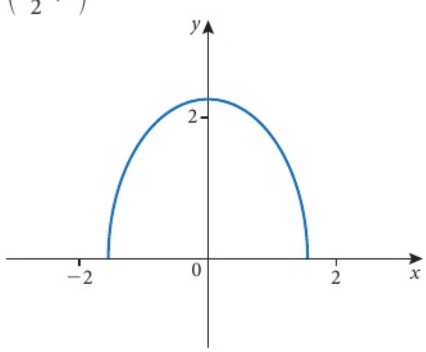
(c) (i) domain $\{x : x \in \mathbb{R}, x \neq -2\}$, range $\{y : y \in \mathbb{R}, y \neq 2\}$

(ii) y-intercept $\left(0, -\frac{1}{2}\right)$, vertical asymptote $x = -2$, horizontal asymptote $y = 2$



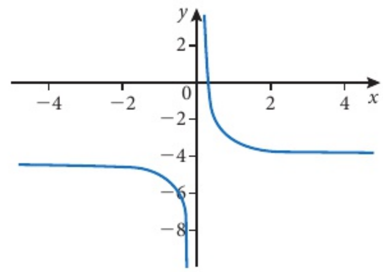
(d) (i) domain $\left\{x : -\frac{\sqrt{10}}{2} \leqslant x \leqslant \frac{\sqrt{10}}{2}\right\}$, range $\{y : 0 \leqslant y \leqslant \sqrt{5}\}$

(ii) y-intercept $(0, \sqrt{5})$, x-intercepts $\left(-\frac{\sqrt{10}}{2}, 0\right)$ and $\left(\frac{\sqrt{10}}{2}, 0\right)$



(e) (i) domain $\{x : x \in \mathbb{R}, x \neq 0\}$, range $\{y : y \in \mathbb{R}, y \neq -4\}$

(ii) vertical asymptote $x = 0$, horizontal asymptote $y = -4$



Exercise 2.3

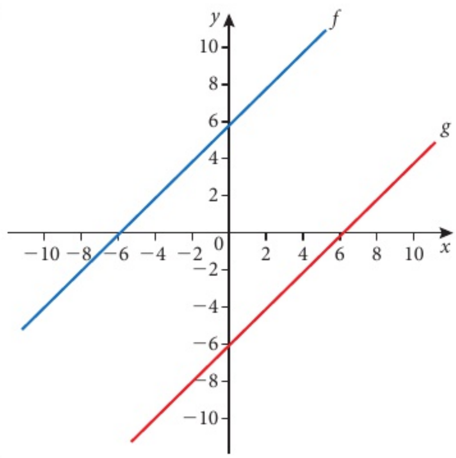
1. **(a)** 2 **(b)** 6

2. **(a)** -1 **(b)** b

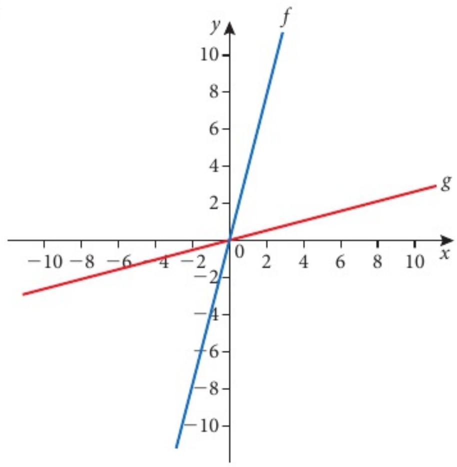
3. 4

4. 6

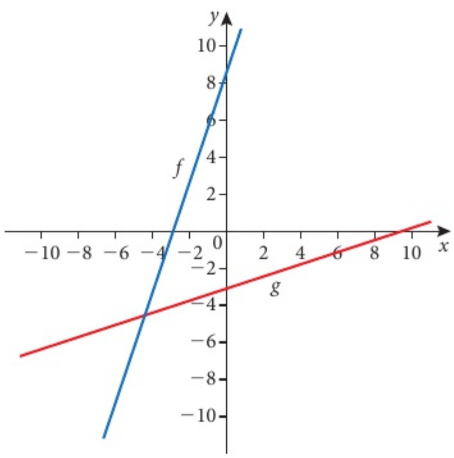
5. **(a)**



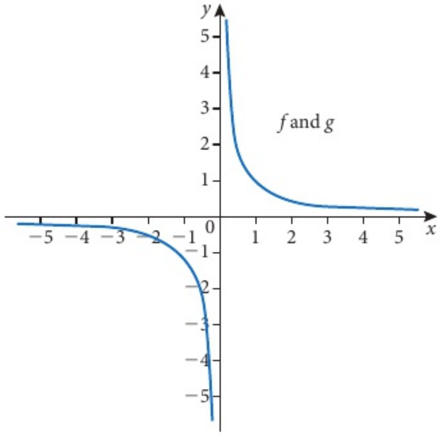
(b)



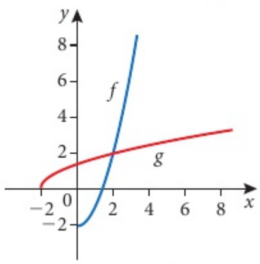
(c)



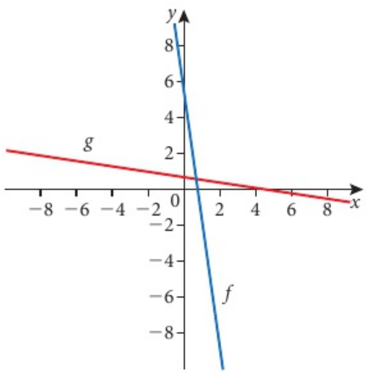
(d)



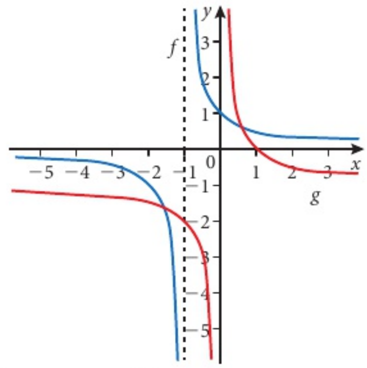
(e)



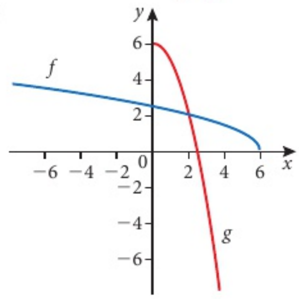
(f)



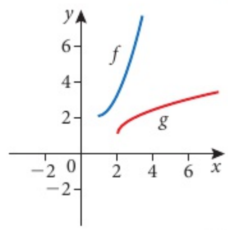
(g)



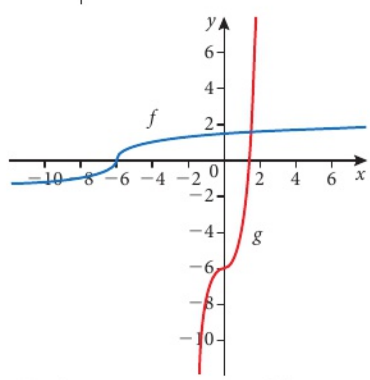
(h)



(i)



(j)



6. **(a)** P(5, 5) **(b)** -3
(c) (5, 5) and (-3, 1) **(d)** $h^{-1}(x) = \frac{1}{2}x + \frac{5}{2}$

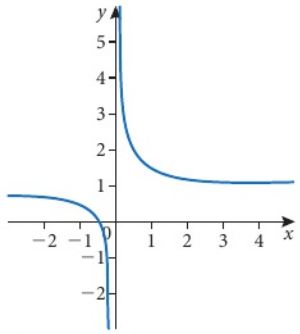
7. (2.30, 2.30)

Chapter 2 Practice questions

1. **(a)** $k = -9.5$
(b) $-\frac{2}{5}$ or -0.4
(c) $y = -\frac{2}{5}x - \frac{4}{5}$ or $y = -0.4x - 0.8$
(d) $2x + 5y + 4 = 0$

2. **(a)** **(i)** 1.6180
(ii) 1.62
(b) $xy = 1$
(c) $\frac{y}{x} = r$ or $\frac{y}{x} = \frac{1 + \sqrt{5}}{2}$
(d) $x \approx 0.786$

3. (a) (i) $x = -\frac{1}{2}$, or $x = -0.5$
   (ii) $x = 0$
   (b)



4. (a) $180 = 150m + c$ (b) $181.5 = 210m + c$
   (c) $L \approx 177$ mm
5. (a) 3600 m² (b) 180 m
   (c) (i) 100 m
   (ii) 10 000 m²
   (d) $m = 3600$, $n = 10000$
6. (a) (i) $p = 3$
   (ii) $q = 5$
   (b) $r = 8$ (c) $s = 6$
7. (a) $2 \times 0 - 3 \times 6 \neq 11$ (b) $\frac{3}{2}$
   (c) $m = -\frac{2}{3}$ (d) $c = 4$
8. (a) $-\frac{2}{5}$ (b) $\frac{5}{2}$ (c) $y = \frac{5}{2}x - \frac{19}{2}$
   (d) $(3, -2)$
9. (a) 100°C (b) 95.3°C (c) 8820 m

Chapter 3

Exercise 3.1

1. (a) $-1, 1, 3, 5, 7$ (b) $-1, 1, 5, 13, 29$
   (c) $\frac{3}{2}, \frac{3}{4}, \frac{3}{8}, \frac{3}{16}, \frac{3}{32}$ (d) $5, 8, 11, 14, 17$
   (e) $1, 7, -5, 19, -29$ (f) $3, 7, 13, 21, 31$
2. (a) $5, 7, 9, 11, 13, 103$
   (b) $6, 18, 54, 162, 486, 1.43 \times 10^{24}$
   (c) $\frac{2}{3}, -\frac{2}{3}, \frac{6}{11}, -\frac{4}{9}, \frac{10}{27}, \frac{50}{1251}$
   (d) $1, 2, 9, 64, 625, 1.776 \times 10^{83}$
   (e) $3, 11, 27, 59, 123, 4.50 \times 10^{15}$
   (f) $0, 3, \frac{3}{7}, \frac{21}{13}, \frac{39}{55}$, approx 1
   (g) $2, 6, 18, 54, 162, 4.786 \times 10^{23}$
   (h) $-1, 1, 3, 5, 7, 97$

Exercise 3.2

1. (a) arithmetic (b) not
   (c) not (d) arithmetic
2. (a) arithmetic, $d = 2$, $a_{50} = 97$
   (b) arithmetic, $d = 1$, $a_{50} = 52$
   (c) arithmetic, $d = 2$, $a_{50} = 97$
   (d) not arithmetic, *no common difference.*
   (e) not arithmetic, *no common difference.*
   (f) arithmetic, $d = -7$, $a_{50} = -341$
3. $13, 7, 1, -5, -11, -17, -23$
4. (a) (i) 26 (ii) $a_n = -2 + 4(n - 1)$
   (b) (i) 1 (ii) $a_n = 29 - 4(n - 1)$
   (c) (i) 57 (ii) $a_n = -6 + 9(n - 1)$
   (d) (i) 9.23 (ii) $a_n = 10.07 - 0.12(n - 1)$
   (e) (i) 79 (ii) $a_n = 100 - 3(n - 1)$
   (f) (i) $-\frac{27}{4}$ (ii) $a_n = 2 - \frac{5}{4}(n - 1)$
5. (a) -7 (b) 45 (c) -251 (d) 223
6. 99th
7. $a_n = 4n - 14$
8. $a_n = -\frac{142}{3} + \frac{11}{3}(n - 1) = -51 + \frac{11}{3}n$
9. (a) 88 (b) 36 (c) 11 (d) 16 (e) 11
10. $a_n = 4n + 27$
11. Yes, 3271th term
12. (a) proof (b) Yes, 1385th term
13. No
14. (a) 0 (b) $-\frac{9}{2}$
15. (a) 2700 (b) 3500 (c) 6200
16. 22

Exercise 3.3

1. (a) No (b) Yes (c) Yes (d) No
2. (a) Geometric, $r = 3^a$, $g_{10} = 3^{9a+1}$
   (b) Arithmetic, $d = 3$, $a_{10} = 27$
   (c) Geometric, $r = 2$, $b_{10} = 4096$
   (d) Neither
   (e) Geometric, $r = 3$, $u_{10} = 78732$
   (f) Geometric, $r = 2.5$, $a_{10} = 7629.39453125$
   (g) Geometric, $r = -2.5$, $a_{10} = -7629.39453125$
   (h) Arithmetic, $d = 0.75$, $a_{10} = 8.75$
   (i) Geometric, $r = -\frac{2}{3}$, $a_{10} = -\frac{1024}{2187}$
   (j) Arithmetic, $d = 3$, $a_{10} = 79$
   (k) Geometric, $r = -3$, $u_{10} = 19683$
   (l) Geometric, $r = 2$, $u_{10} = 51.2$
   (m) Neither
   (n) Neither
   (o) Arithmetic, $d = 1.3$, $a_{10} = 14.1$
3. 6,12,24,48
4. (a) (i) 32 (ii) $-3 + 5(n - 1)$
   (b) (i) -9 (ii) $19 - 4(n - 1)$
   (c) (i) 69 (ii) $-8 + 11(n - 1)$
   (d) (i) 9.35 (ii) $10.05 - 0.1(n - 1)$
   (e) (i) 93 (ii) $100 - (n - 1)$
   (f) (i) $-\frac{17}{2}$ (ii) $2 - 1.5(n - 1)$
   (g) (i) 384 (ii) $3 \times 2^{n-1}$
   (h) (i) 8748 (ii) $4 \times 3^{n-1}$
   (i) (i) -5 (ii) $5 \times (-1)^{n-1}$
   (j) (i) -384 (ii) $3 \times (-2)^{n-1}$
   (k) (i) $-\frac{4}{9}$ (ii) $972 \times \left(-\frac{1}{3}\right)^{n-1}$
   (l) (i) $\frac{2187}{64}$ (ii) $-2\left(-\frac{3}{2}\right)^{n-1}$
   (m) (i) $\frac{390625}{117649}$ (ii) $35\left(\frac{5}{7}\right)^{n-1}$
   (n) (i) $-\frac{3}{64}$ (ii) $-6\left(\frac{1}{2}\right)^{n-1}$
   (o) (i) 1216 (ii) $9.5 \times 2^{n-1}$
   (p) (i) 69.833 729 609 (ii) $100\left(\frac{19}{20}\right)^{n-1}$
   (q) (i) 0.002 085 686 (ii) $2\left(\frac{3}{8}\right)^{n-1}$
5. 35,175,875
6. 36

7. 21,63,189,567
8. $-24, 24$
9. $1.5, a_n = 24\left(\frac{1}{2}\right)^{n-1}$
10. £9033.86
11. **(a)** ¥11982.79 **(b)** ¥12216.10 **(c)** ¥12270.48
12. 10th term
13. Yes, 10th term
14. Yes , 10th term
15. €228.92
16. **(a)** 4.52% **(b)** 8 years and 5 months.
17. €2968.79
18. 8 690 000
19. **(a)** 5 **(b)** 1 **(c)** 4 **(d)** 1/2
20. £8615.24
21. €3714.87
22. £2921.16
23. 12 years
24. $14 171.76
25. 325 778.93 SAR
26. **(a)** $3906.78, $3795.96
    (b) $3000\left(1 + \frac{0.038}{2}\right)^{6\times2} = 3760.20$
    (c) 16 years

Exercise 3.4

1. **(a)** **(i)** 60 **(ii)** 11280
   (b) **(i)** 14 **(ii)** 0.7
   (c) **(i)** 20 **(ii)** 940
   (d) **(i)** 46 **(ii)** 6578
   (e) **(i)** 125 **(ii)** 42625
2. **(a)** **(i)** $-\frac{3}{2}$ **(ii)** 12 **(iii)** $-\frac{105469}{1024}$
   (b) **(i)** $-\frac{2}{5}$ **(ii)** 7 **(iii)** 1.431
   (c) **(i)** $\frac{\sqrt{3}}{4}$ **(ii)** 8 **(iii)** 0.587
   (d) **(i)** $-\frac{2}{3}$ **(ii)** 12 **(iii)** 1.191
3. 5929.92
4. 17 terms
5. 29 terms
6. **(a)** €36 677.46 **(b)** €16 502.18
7. **(a)** £150 677 **(b)** £200 903
8. $60.34
9. **(a)** **(i)** $S_1 = 7$, **(ii)** $S_2 = 16$
   (b) 9 **(c)** $d = 2$ **(d)** 25 **(e)** 498
   (f) 36
10. 49.2
11. **(a)** 1.945 **(b)** 131.945
12. **(a)** 11 866 **(b)** 763 517 **(c)** 14 348 906 **(d)** ~150
13. **(a)** $42 426.61 **(b)** 53 months
14. **(a)** 220 **(b)** 11
    (c) 2530 = 42 minutes and 10 seconds
    (d) 198.45 seconds **(e)** 775.82 **(f)** 7th
15. **(a)** £27 127.50 **(b)** £6947.50
    (c) £8285.02 **(d)** £24 686.03
    (e) $27\,127.50 \times (1 - 0.09)^4 \approx$ £18 600
    (f) €423.65
16. **(a)** $94 629.66 **(b)** $465 859.68

Chapter 3 Practice questions

1. $d = 5, n = 20$
2. €2098.63
3. **(a)** Nick:20
   Charlotte: 17.6
   (b) Nick: 390
   Charlotte: 381.3
   (c) Charlotte will exceed the 40 hours during week 13.
   (d) In week 11 Charlotte will catch up with Nick and exceed him.
4. **(a)** loss for the second month = 1060 g
   loss for the third month = 1123.6 g
   (b) Plan A loss = 1880 g
   Plan B loss = 1898.3 g
   (c) **(i)** Loss due to plan A in all 12 months = 17280 g
   (ii) Loss due to Plan B in all 12 months = 16 869.9 g
5. **(a)** €895.42
   (b) £6590.40
6. **(a)** 142.5 **(b)** 19003.5
7. **(a)** 3.3 km **(b)** 28.47km **(c)** 19 days
8. **(a)** On the 37th day **(b)** 407 km
9. **(a)** 1.5
   (b) **(i)** 207 595 **(ii)** 2019
   (c) 619 583 **(d)** Market saturation
10. 6, 2001
11. **(a)** **(i)** 6 **(ii)** 70
    (b) 3550 **(c)** 16
12. **(a)** 1220 **(b)** 36 920 m
13. **(a)** 729 **(b)** 1093 **(c)** 12:45
14. **(a)** **(i)** Kell: €18 400, €18 800; YBO: €18 190, €19 463.30
    (ii) Kell: €198 000; YBO: €234 897.62
    (iii) Kell: €21 600; YBO: €31 253.81
    (b) **(i)** After the second year
    (ii) 4th year
15. **(a)** 62 **(b)** 936
16. **(a)** $7000(1 + 0.0525)^t$
    (b) 7 years
    (c) Yes since 10 084.7 > 10 015.0
17. **(a)** 11 **(b)** 2 **(c)** 15
18. **(a)** 15 **(b)** -8
19. 10300
20. **(a)** $a_n = 8n - 3$ **(b)** 50
21. **(a)** 1160.75 **(b)** 5.25 years
22. 559
23. **(a)** 4 **(b)** $16(4^n - 1)$
24. **(a)** 80 **(b)** -88 **(c)** -3050
25. **(a)** $\frac{n(3n+1)}{2}$ **(b)** 30
26. **(a)** 30 082.74 (EUR) **(b)** 12 183.39 (EUR)
27. **(a)** **(i)** 10 m **(ii)** 35 m.
    (b) 145 m **(c)** 1125 m or 1.125 km.
    (d) 19 seconds **(e)** 18 700
    (f) 271 000 (2 70 936)
28. **(a)** **(i)** 37 200 **(ii)** 37 080;
    (b) **(i)** $36\,000 + 1200(n - 1)$ **(ii)** $36\,000(1.03)^{n-1}$
    (c) 9th year
    (d) **(i)** 669 561 **(ii)** 3561
29. **(a)** $N(t) = 300\,(1.05)^t$ **(b)** 489 **(c)** 39 weeks
30. $2, -3$

31. $-2, 4$

32. **(a)** 835.90 **(b)** 8135

33. **(a)** 6847.26 **(b)** $r = 7.18\%$

34. **(a)** $d = 6$ **(b)** $S = 5940$ **(c)** 9.41%

35. **(a)** $5000(1.063)^n$ **(b)** 6786.35
(c) 12 years

36. 7

37. $u_1 = 12, d = -1.5$

38. **(a)** 41 **(b)** $\sum_{n=1}^{41} 7 + 7n$ **(c)** 6314 **(d)** 287

39. **(a)** **(i)** $S_1 = 6$ **(ii)** $S_2 = 16$
(b) $u_2 = 16 - 6$ **(c)** $d = 4$
(d) 42 **(e)** $n = 500$ **(f)** $n = 38$

40. **(a)** £121521.60 **(b)** 56 000

Chapter 4

Exercise 4.1

1. **(a)** $\sqrt{5}$ cm ≈ 2.24 cm **(b)** 1 cm
(c) $\sqrt{3}$ cm ≈ 1.73 cm **(d)** $\sqrt{12}$ cm ≈ 3.46 cm

2. **(a)** $3^2 + 4^2 = 25, 5^2 = 25$
(b) $(3n)^2 + (4n)^2 = 25n^2, (5n)^2 = 25n^2$
(c) 1,4,9,16,25,36,49,64,81,100,121,144,169,196
$6^2 + 8^2 = 10^2$, $12^2 + 5^2 = 13^2$ are two Pythagorean Triples from the list
(d) Pythagorean triples correspond to right triangles with whole number side lengths. In general, at least one side is an irrational number.
(e) $(m^2 - n^2)^2 + (2mn)^2 = m^4 - 2m^2n^2 + n^4 + 4m^2n^2$
$= m^4 + 2m^2n^2 + n^4$
$(m^2 + n^2)^2 = m^4 + 2m^2n^2 + n^4$, as desired

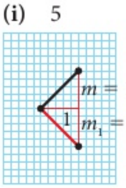
3. 233 km

4. 16 m

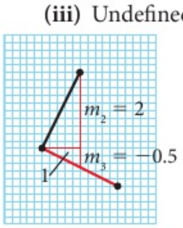
5. **(a)** [CD] = 5 m, [AB] = $\sqrt{34} \approx 5.83$
(b) 11.93 m

6. **(a)** **(i)** $\sqrt{98} \approx 9.9$ **(ii)** (5.5,8.5) **(iii)** 1
(b) **(i)** $\sqrt{90} \approx 9.5$ **(ii)** (5.5,4.5) **(iii)** -3
(c) **(i)** $\sqrt{50} \approx 7.07$ **(ii)** (1.5,3.5) **(iii)** $-\frac{1}{7}$
(d) **(i)** 4 **(ii)** (6,7) **(iii)** 0
(e) **(i)** 5 **(ii)** $(-3,2.5)$ **(iii)** Undefined

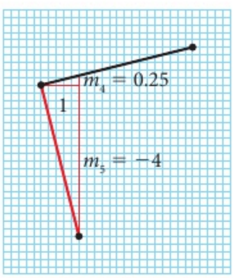
7. **(a)**



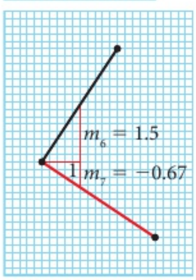
(b)



(c)



(d)

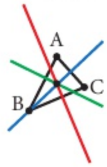


The gradients of each pair are negative reciprocals, and the line segments are perpendicular to each other.

8. **(a)** $y = 3x + 7$ **(b)** $y = -2x + 12$
(c) $y = 3x - 1$ **(d)** $y = -5x$
(e) $y = -3x + 15$

9. **(a)** $y = -\frac{a}{b}x - \frac{c}{b}$
(b) The gradient is $-\frac{a}{b}$ and the y intercept is $-\frac{c}{b}$

10. red $y = 2x + 1$ blue $y = -x + 4$ purple $y = \frac{1}{2}x + 1$
(a) $A(1, 3), B(0, 1), C(2, 2)$
(b) $AB = \sqrt{5}, AC = \sqrt{2}, BC = \sqrt{5}$ and the perimeter is $2\sqrt{5} + \sqrt{2} \approx 5.88$*units*
(c) $m = \frac{1}{2}, M(B, C) = (1, 1.5)$
(d) $y = -2x + 3.5$
(e) The perpendicular bisectors meet in a single point called the *orthocenter*.



Exercise 4.2

1. **(a)** $\frac{B}{C}$ **(b)** $\frac{A}{C}$ **(c)** $\frac{B}{A}$ **(d)** $\frac{A}{C}$
(e) $\frac{B}{C}$ **(f)** $\frac{A}{B}$ **(g)** $\beta = 90° - \theta$
(h) $\sin\theta = \cos(90 - \theta), \cos\theta = \sin(90 - \theta)$,
$\tan\theta = \dfrac{1}{\tan(90 - \theta)}$

2. **(a)** 13.02 cm **(b)** 9.5 cm **(c)** 6 cm **(d)** 10.15 cm

3. **(a)** 66° **(b)** 37° **(c)** 25° **(d)** 60°

4. **(a)** 127 km North, 272 km East
(b) 103 km South, 282 km East
(c) Since the plane flies on a bearing larger than 090, we use a right triangle that has an interior angle of $110° - 90° = 20°$ to calculate distance travelled. In addition, the plane flies south instead of north.

5. The height of the yellow building is 32 m and the height of the green building is 47 m.

6. 66.4°

7. **(a)** 24 cm^2 **(b)** 38 cm^2

8. **(a)** $\frac{3}{2\sqrt{3}}$ cm^2 or $3\frac{\sqrt{3}}{2}$ cm^2 **(b)** $\frac{5\cos 36}{4\sin 36}$cm^2
(c) The polygon is divided into n congruent isosceles triangles with interior angles of $\frac{360}{n}, \frac{180n - 360}{2n}$,and $\frac{180n - 360}{2n}$. One can now find the height and area of one of these triangles, and multiple this area by n to get the total area of the polygon.

9.

θ	$\sin(\theta)$	$\cos(\theta)$
10°	0.17365	0.98481
5°	0.08716	0.99619
2°	0.03490	0.99939
1°	0.01745	0.99985
0.5°	0.00873	0.99996

As θ approaches zero, the opposite side decreases and sin(θ) approaches zero. At the same time, the adjacent side increases and approaches the same length as the hypotenuse, and cos(θ) approaches one.

Exercise 4.3

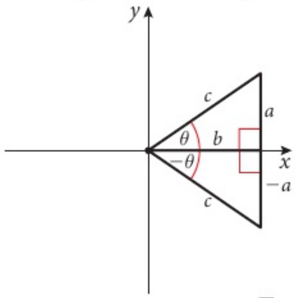
1.

	$\sin(\theta)$	$\cos(\theta)$	$\tan(\theta)$
$0° < \theta < 90°$	+	+	+
$90° < \theta < 180°$	+	−	−
$180° < \theta < 270°$	−	−	+
$270° < \theta < 360°$	−	+	−

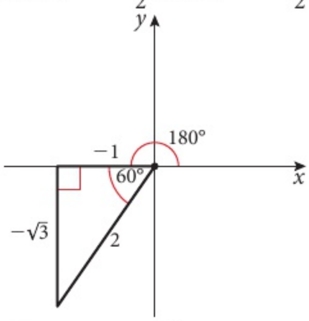
2. $\theta = 45°, 135°$

3. **(a)** $\cos\theta = \cos(-\theta) = \frac{b}{c}$

(b) $\sin\theta = \frac{a}{c}$, $\sin(-\theta) = -\frac{a}{c}$



4. $\cos 240^0 = -\frac{1}{2}$, $\sin 240^0 = -\frac{\sqrt{3}}{2}$, $\tan 240 = \sqrt{3}$



5. **(a)** 49.3 cm **(b)** 12.7 cm

6. **(a)** 43° **(b)** 156°

7. **(a)** $\sin b = \frac{h_1}{A}$, $\sin a = \frac{h_1}{B} \Rightarrow h_1 = A\sin b = B\sin a(*)$

(b) Dividing both sides of (*) by AB gives $\frac{\sin b}{B} = \frac{\sin a}{A}$

(c) $\sin(180 - c) = \frac{h_2}{B}$, but $\sin(180 - c) = \sin c$ and $h_2 = B\sin c$ using the large right triangle, $\sin b = \frac{h_2}{C}$

Therefore, $h_2 = B\sin c = C\sin b(**)$

(d) Dividing (**) by BC gives $\frac{\sin c}{C} = \frac{\sin b}{B}$

Combining with **(a)** gives equality with $\frac{\sin a}{A}$

8. **(a)** $x \approx 1.46$ m **(b)** $x \approx 17.4$ cm

9. **(a)** $\theta \approx 39°$ **(b)** $\theta \approx 126°$

10. The Sine law can be used to calculate missing side lengths or angles in a triangle when we know a side length and an opposite interior angle. The Cosine Law can be used to find a side length of a triangle when we have the opposite interior angle and the two enclosing sides. Alternatively, we can find an interior angle if we know all three sides of the triangle.

11. **(a)** $\cos b = \frac{x}{A} \Rightarrow x = A\cos b$

(b) $h^2 = A^2 - x^2 = B^2 - (C - x)^2$

(c) Expanding the expression in **(b)** and making B the subject gives:

$A^2 - x^2 = B^2 - (C^2 - 2Cx + x^2)$

$A^2 - x^2 = B^2 - C^2 - 2Cx + x^2$ we can now cancel x^2

$B^2 = A^2 + C^2 - 2Cx$

$B^2 = A^2 + C^2 - 2AC\cos b$

12. **(a)** $\theta = 50.2°$, $x = 32.4$ m

(b) $\theta = 122.9°$, $\beta = 23.1°$

(c) $\theta = 127.9°$

13. 472 cm^2

14. **(a)** 310 km

(b) 34°

(c) 2 hours, 49 minutes

Exercise 4.4

1. **(a)** (i) 84 cm^3 (ii) 122 cm^2

(b) (i) 10.4 cm^3 (ii) 26.4 cm^2

(c) (i) 904.8 cm^3 (ii) 452.4 cm^2

(d) (i) 11.18 cm^3 (ii) 34.5 cm^2

(e) (i) 180 m^3 (ii) 227 m^2

(f) (i) 134 cm^3 (ii) 151 cm^2

2. **(a)** $V = 13.5\ cm^2$, $S = 37.8\ cm^2$

(b) $V = 2.83\ cm^3$ $S = 12.3\ cm^2$

3. **(a)** $\theta \approx 83.6°$ **(b)** $V \approx 1500\ m^3$ **(c)** $S = 822\ m^2$

(d) The volume is important for calculating heating costs, surface area if important for calculating costs for painting and insulation.

4. **(a)** $V_{Earth} = 1.05 \times 10^{12}\ km^3$, $V_{Jupiter} = 1.43 \times 10^{15}\ km^3$

(b) Jupiter's volume is about 1366 times that of earth.

(c) $S_{Earth} = 4.99 \times 10^8\ km^2$, $S_{Jupiter} = 6.14 \times 10^{10}\ km^2$

(d) Jupiter's surface area is about 123 times that of Earth.

5. **(a)** The surface of a sphere can be covered by the surface of four of its equatorial circles.

(b) The volume increases by a factor of 8 when the radius is doubled.

(c) The surface area increases by a factor of 4 when the radius is doubled.

6. **(a)** All square pyramids with the same base and equal heights have the same volume.

(b) $V = 64\ cm^3$, $S = 127\ cm^2$

7. **(a)** $OC \approx 24$ m **(b)** $\theta \approx 42°$ **(c)** $B\widehat{A}C \approx 95°$

(d) Triangle BAC has an area of 266.4 m^2 and the pyramid has a surface area of 1065 m^2(without the base)

(e) $V = 8323\ m^3$

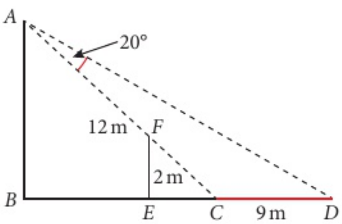
Chapter 4 Practice questions

1.

Equation	Diagram number
$y = c$	2
$y = -x + c$	3
$y = 3x + c$	4
$y = \frac{1}{3}x + c$	1

2.

Condition	Line
$m > 0$ and $c > 0$	5
$m < 0$ and $c > 0$	4
$m < 0$ and $c < 0$	1
$m < 0$ and $c = 0$	2

3. **(a)** $\frac{2}{3}$ **(b)** $y = \frac{2}{3}x + \frac{14}{3}$ **(c)** $y = 10$
4. **(a)** $P(0, 6), Q(12, 0)$ **(b)** $A(6, 3)$
5. **(a)** $A(0, 1), B(1, 2), C(0, 0)$
   (b) $AB = \sqrt{2}, AC = 1, BC = \sqrt{5}$
   (c) $A\widehat{C}B \approx 26.6°$ **(d)** 0.5 units^2
6. **(a)** L_1 has gradient 2, L_2 has gradient $-\frac{1}{4}$
   (b) L_2 is drawn incorrectly.
   (c) The gradients are not negative reciprocals of each other.
   (d) The blue line is the correctly drawn line.
7. 

   (a) In triangle ACD, using the law of sines:
   $\frac{\sin C\widehat{D}A}{12} = \frac{\sin 20°}{9} \Rightarrow C\widehat{D}A = \sin^{-1}\left(\frac{12 \sin 20°}{9}\right) \approx 27.13°$
   (b) Let the point at which the new wall to be placed be E.
   $B\widehat{C}A = 20 + 27.13 = 47.13$
   In triangle CEF
   $\tan E\widehat{C}F = \frac{2}{EC} \Rightarrow EC = \frac{2}{\tan E\widehat{C}F} \approx 3.90$
   Thus the furthest point from B is 10 – 3.90 = 6.10 m.
8. **(a)** $AC = 12$ m **(b)** $AG \approx 12.5$ m **(c)** 16.3°
9. **(a)** $A\widehat{C}B = 60°$ **(b)** 3.3%
10. **(a)** 1380 m **(b)** 743 m **(c)** 3.5%
11. **(a)** $\frac{\text{Length}_{\text{base}}}{\text{Length}_{\text{roof}}} = \frac{\text{Height}_{\text{total}}}{h} \Rightarrow \frac{55}{h} = \frac{50}{10} \Rightarrow h = 11$ m
    (b) 466.7 m^3 **(c)** 57.3°
12. **(a)** **(i)** $\tan 30 = \frac{1}{\sqrt{3}}, \frac{\sin 30}{\cos 30} = \frac{\frac{1}{2}}{\frac{\sqrt{3}}{2}} = \frac{1}{\sqrt{3}}$

    $\tan 45 = 1, \frac{\sin 45}{\cos 45} = \frac{\frac{1}{\sqrt{2}}}{\frac{1}{\sqrt{2}}} = 1$

    (ii) $\cos^2 45 + \sin^2 45 = \left(\frac{1}{\sqrt{2}}\right)^2 + \left(\frac{1}{\sqrt{2}}\right)^2 = \frac{1}{2} + \frac{1}{2} = 1$

    $\cos^2 30 + \sin^2 30 = \left(\frac{\sqrt{3}}{2}\right)^2 + \left(\frac{1}{2}\right)^2 = \frac{3}{4} + \frac{1}{4} = 1$

    (b) $\tan\theta = \frac{b}{a}, \frac{\sin\theta}{\cos\theta} = \frac{\frac{b}{c}}{\frac{a}{c}} = \frac{b}{a}, \sin^2\theta + \cos^2\theta = \frac{a^2}{c^2} + \frac{b^2}{c^2}$
    $= \frac{c^2}{c^2} = 1$

    (c) The key step in deriving the identity is the use of the Pythagorean Theorem, $a^2 + b^2 = c^2$
13. **(a)** **(i)** $-\frac{1}{3}$ **(ii)** $y = -\frac{1}{3}x + 2$
    (b) 4.5 units^2
14. **(a)** 39 270 cm^3 **(b)** 93.75 cm
15. **(a)** $h = 10$ cm **(b)** 196 cm^3 **(c)** 131 cm^3 **(d)** 33%
16. **(a)** $r = 2$ m **(b)** $V = 33.5$ m^3, $S = 50.3$ m
17. **(a)** 126° **(b)** 17.0 m
18. **(a)** **(i)** 8.5 cm **(ii)** 120° **(iii)** 30°
    (b) 14.7 cm
19. **(a)** 277° **(b)** 84°
20. **(a)** 2160 m^3 **(b)** 1 hour 30 minutes
    (c) 1176 m^2 **(d)** $h = 43.9$ cm.

Chapter 5

Exercise 5.1

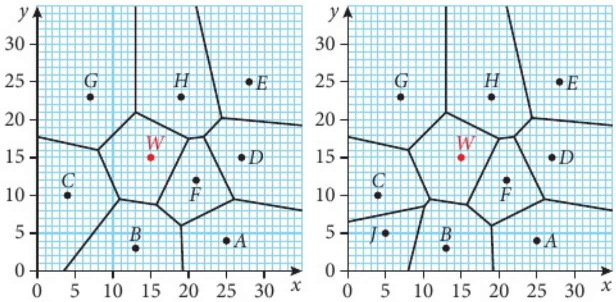
1. **(a)** arc length $= \frac{120}{360}(2\pi)(6) = 4\pi$ cm
   (b) arc length $= \frac{70}{360}(2\pi)(12) = \frac{14}{3}\pi$ cm
2. **(a)** area of sector $= \frac{30}{360}\pi(10)^2 = \frac{25}{3}\pi$ cm^2 arc length $= \frac{30}{360}(2\pi)(10) = \frac{5}{3}\pi$ cm
   (b) area of sector $= \frac{45}{360}\pi(8)^2 = 8\pi$ m^2 arc length $= \frac{45}{360}(2\pi)(8) = 2\pi$ m
   (c) area of sector $= \frac{52}{360}\pi(180)^2 = 4680\pi$ mm^2 arc length $= \frac{52}{360}(2\pi)(180) = 52\pi$ cm
   (d) area of sector $= \frac{n}{360}\pi(15)^2 = \frac{5\pi n}{8}$ cm^2 arc length $= \frac{n}{360}(2\pi)(15) = \frac{\pi n}{12}$ cm
3. $12 = \frac{\theta}{360}(2\pi)(8) \Rightarrow 270 = \pi\theta \Rightarrow \theta = \frac{270}{\pi} \approx 85.9°$
4. **(a)** $1.5 \times 360 = 540°\,\text{s}^{-1}$
   (b) Bicycle speed is equal to the speed of a point along the circumference of the wheel.
   Speed $= 1.5 \times 2\pi(35) \approx 330\,\text{cm s}^{-1}$
   $330\,\frac{\text{cm}}{\text{s}} \times \frac{60\text{ s}}{1\text{ min}} \times \frac{60\text{ min}}{1\text{ h}} \times \frac{1\text{ m}}{100\text{ cm}} \times \frac{1\text{ km}}{1000\text{ m}}$
   $= 11.9\,\text{km h}^{-1}$
5. $\frac{16}{5}\pi = \frac{\theta}{360}\pi(4)^2 \Rightarrow \frac{1}{5} = \frac{\theta}{360} \Rightarrow \theta = 72°$
6. $A = l \Rightarrow \frac{\theta}{360}\pi r^2 = \frac{\theta}{360}2\pi r \Rightarrow r^2 = 2r \Rightarrow r = 0$ or $r = 2$
7. Maria $= \frac{112}{360}(2\pi)(230) \approx 450$ m
   Norbert $= \frac{66}{360}(2\pi)(500) \approx 576$ m
   $\therefore$ Norbert walks $576 - 450 = 126$ m farther.
8. $\frac{1}{12}\pi(400)^2 \approx 41\,888 \approx 42\,000\,\text{m}^2\,\text{h}^{-1}$
9. **(a)** The watered area is 11.8 m^2, so the water requirement is
   $2.5\text{ mm} \times \frac{1\text{ m}}{1000\text{ mm}} \times 11.8 = 0.0295\text{ m}^3$.
   In cm^3, $0.0295\text{ m}^3 \times \frac{100^3\text{ cm}^3}{1\text{ m}^3} = 29\,500\text{ cm}^3$
   (b) $\frac{29\,500}{800} \approx 36.9$ min
   (c) New area $= \frac{100}{360}\pi(3)^2 \approx 7.85\text{ m}^2$
   Water requirement $= 7.85 \times \frac{2.5}{1000} \approx 0.020\text{ m}^3$
   $\approx 19\,600\text{ cm}^3$
   Required time $= \frac{19\,600}{800} \approx 24.5$ min

Exercise 5.2

1. **(a)** F

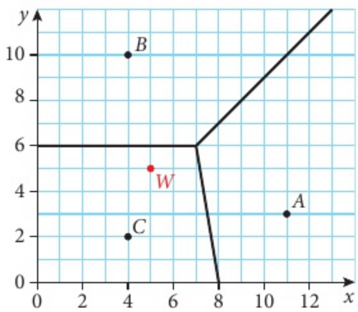
   (b) $y = x + 8$

   (c) The cells for sites F, B, C, G, and H will be changed.

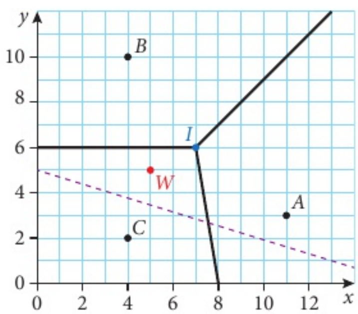


   (d) No, the new cell will have at least two rays as edges.

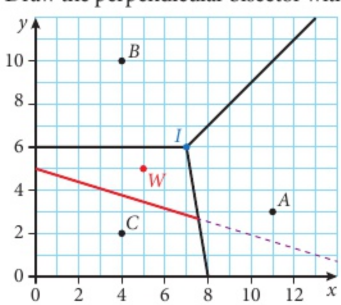
2. The process is shown below.



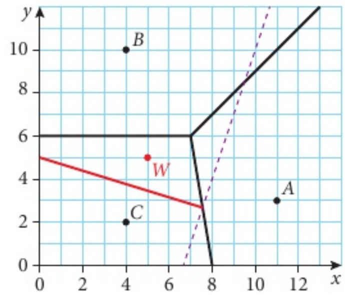
Plot the new site W.



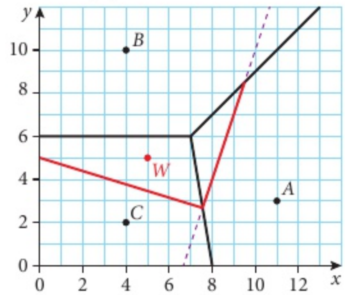
Draw the perpendicular bisector with containing cell's site (C).



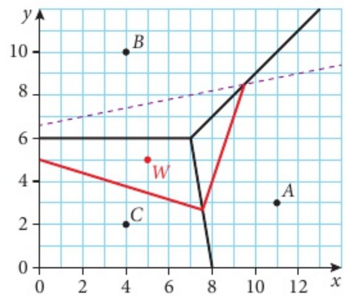
Add ray or line segment between W and C.



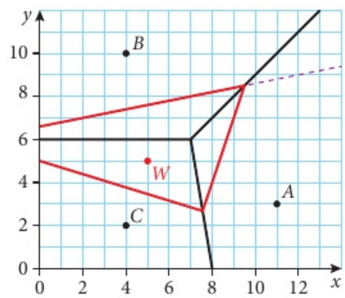
Find perpendicular biector between W and next adjacent cell site (A).



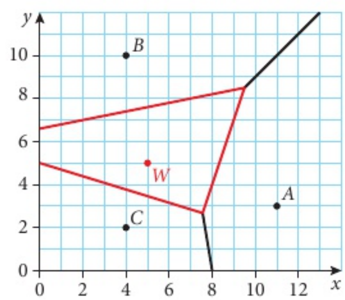
Add ray or line segment between W and A.



Find perpendicular biector between W and next adjacent cell site (B).



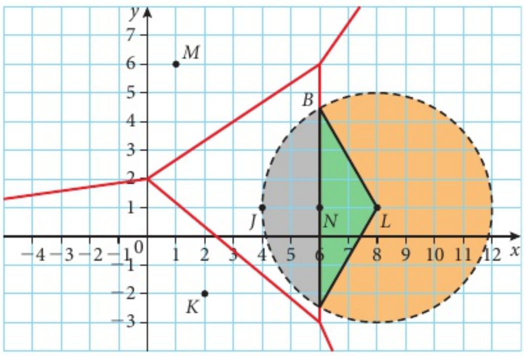
Add ray or line segment between W and B.



Remove edges inside W's new cell.

3. **(a)** We can find Q by reflecting P across the edges shared with the adjacent, empty cell, which gives $Q(4, 1)$.
(b) $A = \frac{1}{2}(4)(3) = 6\,\text{km}^2$
(c) The point $X(9, 2)$ is in the cell of site R. Therefore, we find the distance to point $R(5, 4)$using the distance formula: $d = \sqrt{(9-5)^2 + (2-4)^2}$
$= \sqrt{4^2 + (-2)^2}$
$= \sqrt{20} \approx 4.47\,\text{km}$.

4. **(a)** $A = \frac{1}{2}(9)(6) = 27\,\text{km}^2$
(b) $BLN = 56°$.
(c) The region can be seen as a triangle (shown in blue) and the sector of a circle (shown in orange).



Area = Area of triangle + Area of sector
$= \frac{1}{2}(2 \times 2\sqrt{2})(2) + \frac{360 - 112.6}{360}\pi(4^2)$
$\approx 41.9\,\text{km}^2$
(d) The three vertices to consider are at (0, 2), (6, 6), and (6, −3). Since all three are adjacent to cell J, we will calculate the distance of each vertex to site J:

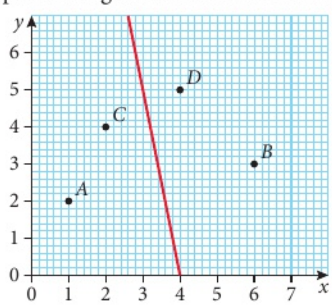
(0, 2): $d = \sqrt{(0-4)^2 + (2-1)^2}$
$= \sqrt{16+1} = \sqrt{17}$

(6, 6): $d = \sqrt{(6-4)^2 + (6-1)^2}$
$= \sqrt{4+25} = \sqrt{29}$

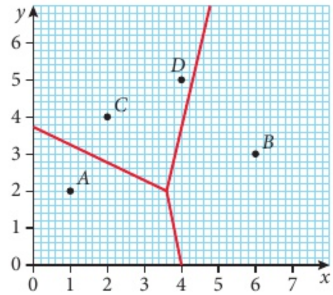
(6, −3): $d = \sqrt{(6-4)^2 + (-3-1)^2}$
$= \sqrt{4+16} = \sqrt{20}$

Since (6, 6) is furthest from site J is is also furthest from the other sites and therefore is the centre of the LEC on a vertex. This vertex will have the weakest signal of the three vertices.

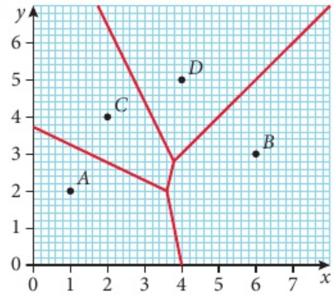
5. **(a)** The partial diagrams are shown below.
(i)



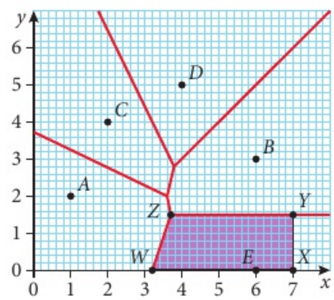
(ii)



(iii)



(b) Point $M(3.5, 2.5)$ is in the cell for site C. Therefore, it is likely to have soil type **silt**.
(c) $N(5, 4)$ is on the boundary between the cells for sites B and C. Since is it equidistant to sites B and C, it is not possible to determine the likely soil type.
(d) Cells for sites A and B would be divided. Cell B would be divided because $E(6, 0)$ is in the cell B. Cell A would be divided because the perpendicular bisector of segment EB intersects the boundary of cell A. There are no more adjacent cells, so no other cells would be divided.
(e) No, the answer to **(b)** does not change because the cell for site C, which contains m, does not change.
(f) The study area likely to be loam is contained within cell E. After we add site E to the Voronoi diagram, we have the following:



(The purple shaded area is the cell for site E.)
Cell E forms a trapezoid. To find the area we must determine the height and length of the two bases. The height = 1.5 (from the graph). The base lengths require more work: we must find the equation of the boundary between sites A and E to find exact coordinates of the points Z and W.

Midpoint of $AE = \left(\frac{6+1}{2}, \frac{0+2}{2}\right) = (3.5, 1)$.

Slope of $AE = \frac{2-0}{1-6} = -\frac{2}{5} \Rightarrow$ slope of $ZW = \frac{5}{2}$.

$\therefore$ equation of $ZW: y - 1 = \frac{5}{2}(x - 3.5) \Rightarrow y = \frac{5}{2}x - 7.75$

$\therefore$ when $y = 0 \Rightarrow 0 = \frac{5}{2}x - 7.75 \Rightarrow x = 3.1 \Rightarrow W(3.1, 0)$

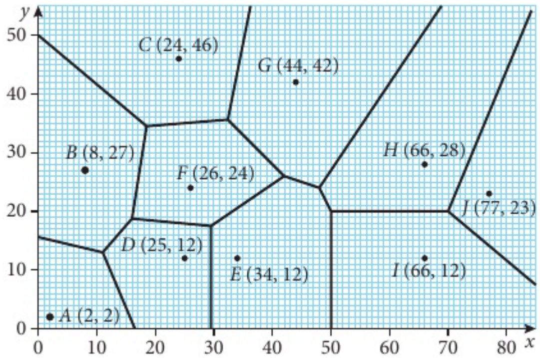
$\therefore$ when $y = 1.5 \Rightarrow 1.5 = \frac{5}{2}x - 7.75 \Rightarrow x = 3.7$
$\Rightarrow Z(3.7, 1.5)$

Therefore the bottom base is $7 - 3.1 = 3.9$, top base is $7 - 3.7 = 3.3$.

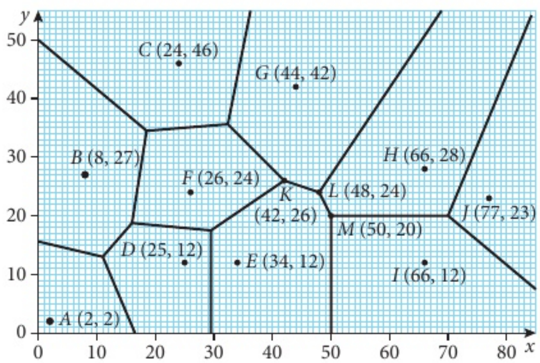
Area of cell $E = \frac{1}{2}(b_1 + b_2)(h) = \frac{1}{2}(3.9 + 3.3)(1.5)$
$= 5.4\ \text{m}^2$

6. Since $V(10, 16)$ and $W(15, 15)$ are both $\sqrt{65}$ units from their nearest site, there are two largest empty circles, centered at V or W, both with radius $\sqrt{65} \approx 8.06$.

7. (a) After adding site H, the Voronoi diagram should look like this:



The region bounded by points E, G, H, and I contains three vertices, labelled as K, L, and M on the diagram below:



(b) Since all three are adjacent to vertex E, we will find the distance of each to vertex E.

For K, $d_K = \sqrt{(42 - 34)^2 + (26 - 12)^2} = \sqrt{260}$;
for L, $d_L = \sqrt{(48 - 34)^2 + (24 - 12)^2} = \sqrt{340}$;
for m, $d_M = \sqrt{(50 - 34)^2 + (20 - 12)^2} = \sqrt{320}$.

Therefore, the best location for a new bank branch is at $L(48, 24)$.

Chapter 5 Practice questions

1. (a) 2.58 million km (b) 194 000 billion km^2
(c) 3.02 million km (d) 163 000 billion km^2

2. (a) 1.75 m^2, 573 lx (b) 7.00 m
(c) 21.9 m (d) 1690 lm

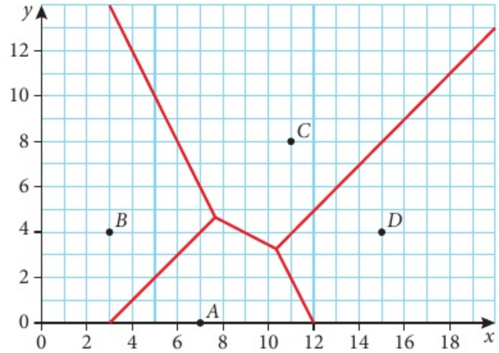
3. (a) $2\sqrt{269} = 32.8$ cm (b) 46.4 cm^2

4. (a) $r = \frac{15}{\sqrt{2}}$ (b) 255°

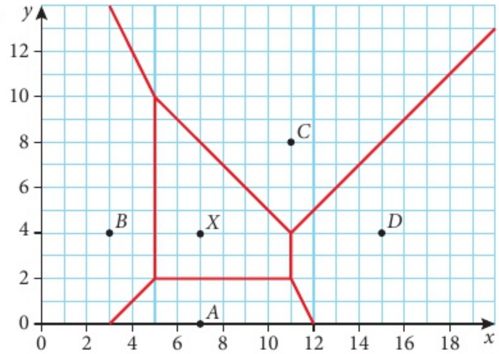
5. (a) 3890 km (b) 11 700 km
(c) 9300 km (d) 19 200 km

6. (a) 41.9 m (b) 227 m^2 (c) 117 panels

7. (a)

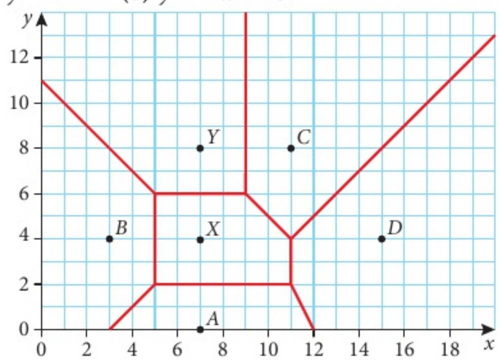


(b)



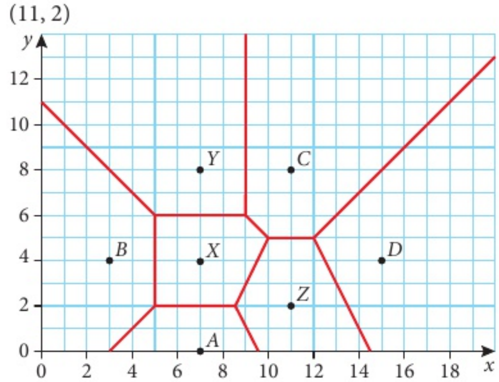
(c) $y = 2$ (d) $y = -x + 15$

(e)



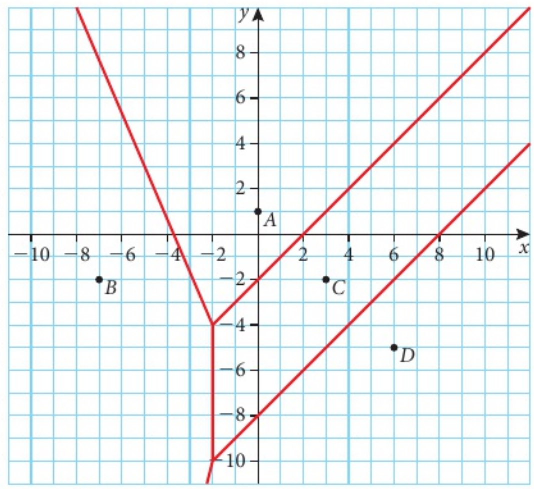
(f) (11, 2)

(g)

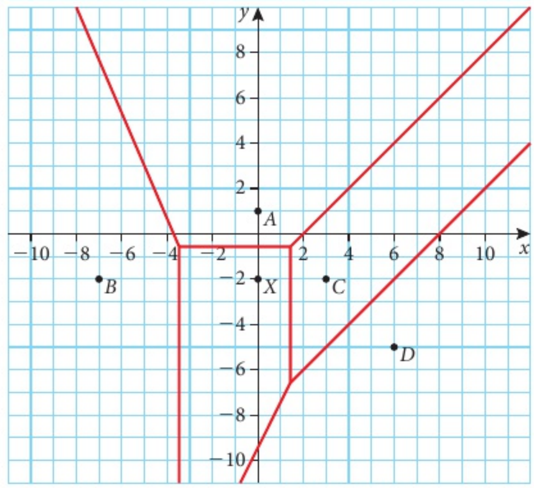


(h) $2x + y = 29$

8. (a)

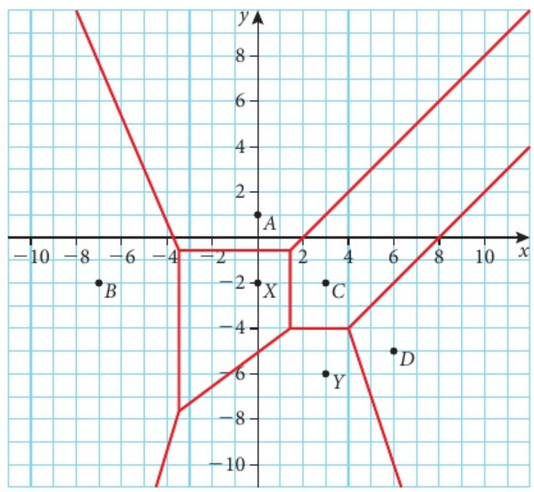


(b)



(c) $x = 1.5$

(d)



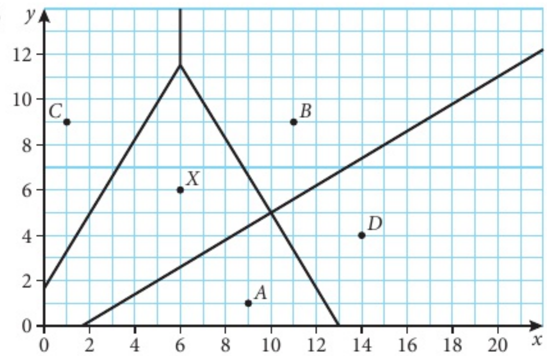
(e) $6x - 8y = 41$

9. (a) the largest empty circle (LEC) must be centred on a vertex

(b) All vertices in the diagram are adjacent to cell K and (15,8) is the vertex farthest from site K, therefore it must be the centre of the largest empty circle.

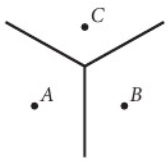
10. (a) The LEC must be at (6,6). The LEC must be on a vertex and (6,6) is the vertex farthest from any site.

(b)

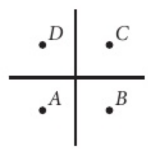


11. Since A, B, C, and D are collinear, the perpendicular bisectors between each pair of points must be parallel. Hence, the edges in the Voronoi diagram are parallel which implies there are no vertices.

12. (a) The Voronoi diagram has 3 edges that meet in a single vertex as shown.



(b) The Voronoi diagram has 4 edges that meet in a single vertex as shown.



(c) The Voronoi diagram for a regular n-gon will have n edges that meet at a single vertex in the centre of the n-gon. The angle between each pair of adjacent edges will be $\frac{360}{n}$

Chapter 6

Exercise 6.1

1. (a) $d = 3000 - 900t$

(b) (0, 3000); 3000 km represents the distance remaining at time $t = 0$

(c) $t = 3.33$; 3 hours and 20 minutes remaining until the plane reaches its destination.

(d) Domain: $0 \leqslant t \leqslant \frac{10}{3}$; Range: $0 \leqslant d \leqslant 3000$

2. (a) $d = -800t + 5000$

(b) The plane is travelling 800 km h^{-1}

(c) (0, 5000); 5000 km represents the distance remaining at time $t = 0$

(d) $t = 6.25$; 6 hours and 15 minutes remaining until the plane reaches its destination.

(e) Domain: $0 \leqslant t \leqslant 6.25$; Range: $0 \leqslant d \leqslant 5000$

3. (a) EU = USA + 33

(b) 45

(c) 11

(d) the EU size increases by 1 for every USA size increase by 1

(e) $6 \leqslant$ USA $\leqslant 16$, $39 \leqslant$ EU $\leqslant 49$

4. (a) $t = \frac{7}{5}n + 14$

(b) 1414 minutes

(c) The gradient of $\frac{7}{5}$ implies that it takes 7 minutes to read 5 pages $\Rightarrow$ 1.4 pages per minute; the t-intercept of 14 suggests that it takes an additional 14 minutes to start/finish a book

(d) $n > 0$; $t > 14$

5. (a) $F = \frac{9}{5}c + 32$

(b) Every degree Celsius is equal to 1.8 degrees Fahrenheit.

(c) 0 °C is equal to 32 °F

(d) 0 °F is equal to −17.8 °C

(e) 50 °F

(f) −40

(g) $C > -273$; $F > -459$

6. (a) *IB Cool*: $C = 150 + 75t$; *MC Numbers*: $D = 120 + 80t$; t is the length of the party in hours, C, D are the cost in dollars.

(b) 6 hours

7. (a) (i) 6000 m (ii) 8400 m (iii) 10000 m

(b) $$D(t) = \begin{cases} 300t & 0 \leqslant t < 20 \\ 150(t - 20) + 6000 & 20 \leqslant t < 36 \\ 400(t - 36) + 8400 & 36 \leqslant t \leqslant 40 \end{cases}$$

(c) (i) 7500 m (ii) 9200 m

(d) 33.3 minutes; 37.5 minutes

8. (a) 11.5 m per year

(b) 2540

(c) $d = 300 + 0.0115t$

(d) 24200 B.C.E.

9. (a) $C = 8.5n + 350$

(b) $n > 0$; $C > 350$

(c) (i) 1200 ZAR (ii) 2050 ZAR (iii) 3750 ZAR

(d) (i) 12 ZAR per cup

(ii) 10.25 ZAR per cup

(iii) 9.375 ZAR per cup

(e) The average cost per cup decreases when more cups are ordered.

(f) $D = 8n + 550$

(g) $n > 0$; $D > 550$

(h) It costs 8 ZAR for each additional cup ordered.

(i) 5350 ZAR

(j) $x = 400$

10. (a) $$C = \begin{cases} 8.5n + 350, & x < 500 \\ 8.5n, & x \geqslant 500 \end{cases}$$

(b) $a = 400$, $b = 499$, $k = 1100$

11. (a) $$C = \begin{cases} 2.6, & m \leqslant 234.8 \\ 2.6 + 0.2\left(\frac{m - 234.8}{117.4}\right), & 234.8 < m \leqslant 9656.1 \\ 18.65 + 0.2\left(\frac{m - 9656.1}{86.9}\right), & 9656.1 < m \end{cases}$$

(b) 2.6 GBP, 10.80 GBP,1.00 GBP

(c) $$D = \begin{cases} 2.6, & t \leqslant 50.4 \\ 2.6 + 0.2\left(\frac{t - 50.4}{25.2}\right), & 50.4 < t \end{cases}$$

(d) 2.6 GBP, 4.60 GBP, 9.40 GBP

(e) (i) 9.20 GBP

(ii) 9.20 GBP

Exercise 6.2

1. (a) $h(t) = -4.9t\ 2 + 5t + 60$

(b) (i) 61.3 m (ii) 4.05 s (iii) $0 \leqslant t < 2.03$

2. (a) Proof.

(b) $P(x) = R(x) - C(x) = (180x - 0.55x^2) - (0.35x^2 + 3200) = -0.9x^2 + 180x - 3200$

maximum profit when $x = \frac{-180}{2(-0.9)} = 100$machines

(c) $R(100) = 180(100) - 0.55\,(100)^2 = 12500$; selling price per machine $\frac{12500}{100} = \$125$

(d) $P(x) = 0 \Rightarrow x = 19.7$ or $x = 180$; smallest number of machines to make positive profit is 20 machines.

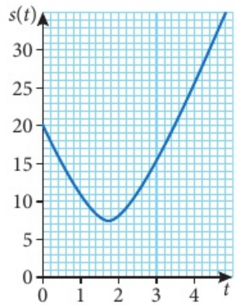
3. (a) 10.8 km

(b) 8 km

(c) The distance of A from the point where B starts at noon is $(20 - 10t)$. The distance B travels since noon is $4t$. Therefore the distance between them is

$$\begin{aligned} s(t) &= \sqrt{(20 - 10t)^2 + (4t)^2} \\ &= \sqrt{400 - 400t + 100t^2 + 16t^2} \\ &= \sqrt{116t^2 - 400t + 400} \end{aligned}$$

(d)



(e) (i) $1.17 \leqslant t \leqslant 2.3$ hours (ii) from 13:15 to 14:12.

4. (a) 320 kg per person (b) 348 kg per person, in 1988

(c) 311 kg per person (d) In the year 2030

(e) This model predicts that worldwide grain production will drop to zero in 2038.

5. (a) $D = (-4.87708 \times 10^{-6})(T - 4)^2 + 1$

(b) 0.99992 g mm^{-1}

(c) 94.6 °C

6. (a) $$\begin{aligned} h &= \frac{36}{105625}(d - 650)^2 + 6 \\ &= 0.000341d^2 - 0.443d + 150 \end{aligned}$$

(b) Domain: $0 \leqslant d \leqslant 1300$, Range: $6 \leqslant h \leqslant 150$

(c) 109 m

(d) $290 < d < 1010$

7. (a) $A = \left(600 - \frac{3}{2}x\right)x$

(b) 60 000 m^2 when $x = 200$ and $y = 150$

8. (a) $16 + x$ (b) $40 - 2x$

(c) 20 t-shirts (d) $R = (16 + x)(40 - 2x)$

(e) $x = 2 \Rightarrow$ €18 per t-shirt for €648 per day

9. (a) If we set the vertex of the parabola on the y axis, then the archway can be described by the function $y = -0.555(x + 3)(x - 3)$. In the truck is centred in the arch, a top corner of the truck will be at (1.3, 4.3), but the parabola passes through (1.3, 4.06), so the truck is 0.24 m too tall. (Or, at $y = 4.3$ m, the parabola passes through (1.12, 4.3) so the truck is 0.36 m too wide.)

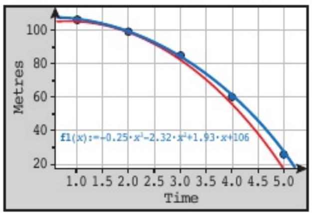
(b) 4.06 m

(c) 2.24 m

10. (a) 1 m; this is the initial height of the ball when it is hit.
(b) $k = 15.4$
(c) $a = 0.779$, $b = 2.36$
(d) $12.9\ \text{m s}^{-1}$
(e) 13.0 m at 1.57 s

Exercise 6.3

1. (a) $V = 4x^3 - 39x^2 + 93.5x$
(b) $0 \leqslant x < 4.25$
(c) Length = 7.83; width = 4.91; height = 1.59 inches
(d) $66.1\ \text{in}^2$
2. (a) $V = 4x^3 - 144x^2 + 1260x$
(b) $0 \leqslant x < 15$
(c) Length = 18.5; width = 5.51; height = 5.76 cm
(d) $3240\ \text{cm}^2$
3. (a) 106 m (b) $h = 44.9$ m
(c) $t = 4.91$ s (d) $t = 5.39$ s
(e) Jane's model appears to fit the data better and Kevin's model increases at the beginning which incorrectly suggests the rock is going up.
(f)



4. (a)

Number of layers in stack, n	1	2	3	4
Total number of balls, B	1	5	**14**	**30**

(b) $a = \frac{1}{3}, b = \frac{1}{2}, c = \frac{1}{6}$
(c) 385 balls (d) 13 layers
5. (a) 204 000 cases (b) 1994 (c) 2003
(d) The model predicts that the number of cases becomes zero/negative in 2013.
6. (a) 1200 bacteria (b) 1500 bacteria
(c) 17.6 minutes
(d) 1930 bacteria at 11.3 minutes
(e) 20 minutes
7. (a) Increase by 2.4 °C.
(b) $a = 4.81\ \text{mg kg}^{-1}$, $b = 8.23\ \text{mg kg}^{-1}$
(c) 3.7 °C for $6.67\ \text{mg kg}^{-1}$

Exercise 6.4

1. (a) $v = 56(0.84)^t$ (b) $16.5\ \text{m s}^{-1}$
(c) 13.9 s (d) 23.1 s
2. (a) 59 bacteria (b) 8.77 hours
(c) $k = \frac{\ln 2.5}{4} = 0.229$ (d) $n > 4.06 \Rightarrow n = 5$
3. (a) $9\ \text{mg l}^{-1}$ (b) $5.03\ \text{mg l}^{-1}$
(c) After $223.07 \approx 224$ minutes $\Rightarrow$ 17:44
4. (a) $L = 16(1.08)^t$ (b) 20.2 cm
(c) 5.80 years
5. (a) $k = -0.0462$ (b) 45.6 units
(c) 88.2 units (d) 23.8 years
6. (a) The initial quantity of glyphosate is 500 units.
(b) t represents the number of 45-day half-lives.
(c) $A(1) = 250$, there are 250 units of glyphosate left after 45 days.
7. (a) the temperature in the oven/the maximum possible temperature of the cake mix
(b) $a = 158$
(c) $t = 50.6$ minutes
8. (a) $a = 810$, $b = 1.92$ (b) 44 fish
(c) August (d) $p = 40$
9. (a) a is the location of the horizontal asymptote. The temperature of the water cannot fall below room temperature, 20 °C, so there must be a horizontal asymptote at $T = 20$, hence $a = 20$.
(b) $b = 80$; $85 = 20 + 80k^{-2} \Rightarrow \frac{65}{80} = k^{-1} \Rightarrow k = 1.11$
(c) $T = 67.6$ °C
(d) $m = 16.12 \Rightarrow$ 16 minutes and 7 seconds
10. (a) 10 °C (b) 123 °C
(c) $k = 5.64$ minutes
11. (a) 2.0 (b) $C = 3$
(c) after 4 hours
12. (a) $p + q = 47$, $4p + q = 53$
(b) $p = 2$, $q = 45$ (c) $C = 109$

Exercise 6.5

1. (a) always (b) sometimes
(c) sometimes (d) never
2. (a) never (b) never (c) never (d) never
3. (a) $y = 210$ (b) $x = 16$
4. (a) $y = 160$ (b) $x = 10$
5. (a) $y = 1024$ (b) $x = 4$
6. (a) $y = 2.5$ (b) $x = 100$
7. (a) $y = 0.625$ (b) $x = 10$
8. (a) $y = 2$ (b) $x = 3$
9. (a) $S = 0.75d$ (b) 5.25 m tall (c) 13.3 m
10. (a) 9.8 (b) 39.2 (c) 5.66 seconds
11. (a) 1.62 (b) 40.5 m (c) 11.1 s
12. (a) 4.00×10^{14} (b) $7560\ \text{m s}^{-1}$
(c) 6.25×10^6 m
13. (a) 7.664 (b) $3920\ \text{cm}^3$ (c) 2.35 m
14. (a) 0.613 (b) 1060 W (c) $14.8\ \text{m s}^{-1}$

Exercise 6.6

1. (a) 6.5 hours (6 hours 30 minutes)
(b) 3.6 m
(c) $p = 1.8$, $q = \frac{360}{13} = 27.7$, $r = 4$
(d) 06:00
(e) 145 minutes
2. (a) $a = 2.8$, $b = 28.8$
(b) 5.68 m
(c) 4.42 hours
3. (a) 1 m
(b) $k = 3.08$ minutes
(c) 9 minutes
4. (a) (i) 120 m (ii) 90 m
(b) $a = -60$, $b = 12$, $c = 60$
(c) 8 minutes

5. **(a)** $a = 8$
(b) $b = \frac{360}{90} = 4$
(c) 34.0 seconds

6. **(a)** $a = 2.4$, $c = 12$
(b) $b = \frac{360}{365} = \frac{72}{73}$
(c) 68 days

7. **(a)** 2 **(b)** 15
(c) -3 **(d)** $16 \leqslant t \leqslant 20$

8. **(a)** 35 cm **(b)** 5 cm
(c) 15 cm **(d)** $A = 15$, $C = 20$
(e) 4 seconds **(f)** $b = 90$
(g) $t = 0.535$ **(h)** 15

9. **(a)** 4 m **(b)** 11 m **(c)** 3 hours **(d)** 17:00

10. **(a)** $a = 19$, $b = 2$ **(b)** $c = 30$
(c) $3.5 < x < 8.5$

11. **(a)** 6.5 hours **(b)** 10 m
(c) The function is initially decreasing
(d) $p = -5$, $q = \frac{360}{13} = 27.7$, $r = 8$
(e) 23:00

Exercise 6.7

1. **(a)** inverse variation
(b) linear/direct linear variation
(c) quadratic/direct quadratic variation
(d) trigonometric
(e) cubic
(f) exponential (growth)
(g) trigonometric
(h) exponential (decay)
(i) inverse variation
(j) linear (probably NOT direct linear variation because there will be a fixed cost as well)

2. **(a)** $2 \leqslant t \leqslant 6$ **(b)** $11 \leqslant x \leqslant 100$
(c) $1.3 \leqslant n \leqslant 8.5$

3. **(a)** A constant rate of change; data in a linear pattern
(b) Data that appears parabolic
(c) Data that is cyclical/periodic
(d) Values that increases quickly, with bound, or decreases to approach a fixed value. The rate of change is a constant factor.
(e) As one value increases, the other decreases; data appears hyperbolic.

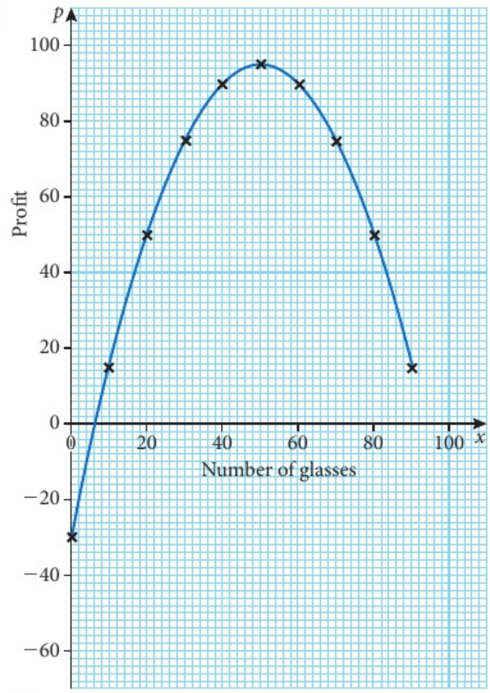
Chapter 6 Practice questions

1. **(a)** $T = 493$
(b) $n = 6.11 \Rightarrow$ in the year 2007
(c) $P = 39636$
(d) At the end of 7 years, $P = 46807 < 25600 \times 2$, therefore not doubled.
(e) $\frac{25600}{280} = \frac{640}{7} = 91.4$
(f) $n = 9.31 \Rightarrow$ after 10 years

2. **(a)**

x	0	10	20	30	40	50	60	70	80	90
P	**−30**	15	**50**	75	90	**95**	**90**	75	50	**15**

(b)



(c) **(i)** maximum profit = 95 swiss francs
(ii) 50 glasses
(iii) 67 ± 2 or 33 ± 2 (from graph)
(iv) 30 Swiss francs

(d) Fiona's share $= \frac{3}{6}$; Profit from 40 glasses $= 90$ swiss francs; Fiona's profit $= \frac{1}{2} \times 90$

3. **(a)** $Q(0) = 25$
(b) $Q(20) = 0.003(20)^2 - 0.625(20) + 25 = 13.7$
(c) $$\text{Average rate of change} = \frac{Q(20) - Q(10)}{20 - 10} = \frac{13.7 - 19.05}{10} = 0.535 \text{ units per minute}$$
(d) $0.003t^2 - 0.625t + 25 = 0 \Rightarrow$ Energy runs out after 154.0 minutes

4. **(a)** 400 USD **(b)** $t = 8.64 \Rightarrow$ 9 months
(c) $8500(0.95)^2 - (400 \times 2 + 2000) = 4871.25$ USD

5. **(a)** $3600\ \text{m}^2$ **(b)** 180 m
(c) **(i)** 100 m **(ii)** $10000\ \text{m}^2$
(d) $m = 3600$, $n = 10000$

6. **(a)** 100 °C **(b)** 95.3 °C **(c)** $h = 8820$ m

7. **(a)** $T(0) = 20 + 70 \times 2.72^{-0.4 \times 0} = 90$
(b) 21.3 °C
(c) 1.7 minutes
(d) 20 °C; that is the horizontal asymptote of the model/the room temperature.
(e) It indicates by how much the temperature increases per minute.
(f) $m = 3.96$; this is the time when the temperature of the soup and the temperature of the water are equal.
(g) $3.96 < m \leqslant 6$

8. **(a)** 25000 USD
(b) 19601.32 USD
(c) 8.55 years

9. **(a)** 1800 bacteria
   (b) 145800 bacteria
   (c) 33.5 hours

10. **(a)** 45 euros **(b)** 8 or 25 people
    (c) $n = 14$ **(d)** 28.29 euros

11. **(a)** 20.3 hours
    (b) The function starts at a minimum.
    (c) $p = -8.3, r = 12$
    (d) $q = \frac{360}{365} = \frac{72}{73}$
    (e) 89 days

Chapter 7

Exercise 7.1

1. **(a)** No experimental units. Total number of students. Students in one school. Qualitative.
   (b) No units. All 10th grade students. One school or class. Quantitative
   (c) Units of length (m, inches etc). All new born children. Children at one hospital or on one day. Quantitative.
   (d) No units. All children under 14. One school or one geographical location. Qualitative.
   (e) Units of time (hours, mins etc). All commuters. Workers in one city or a one company. Quantitative.
   (f) No units. Country's population. Subset of registered voters (age group, location etc). Qualitative.
   (g) No units. All students at international schools. One year group or from one school. Qualitative.

2. **(a) (i)** 1176 students in grades 10–12.
   (ii) Random sampling, appropriate stratified random sampling, quota sampling or convenience sampling of first 30 students at a location.
   (b) (i) All bolts produced.
   (ii) Random sampling from shift or container or testing each nth bolt from shift or container.

3. **(a)** Qualitative **(b)** Quantitative, discrete.
   (c) Quantitative, continuous **(d)** Quantitative, continuous
   (e) Quantitative, continuous **(f)** Quantitative, discrete
   (g) Quantitative, continuous **(h)** Qualitative
   (i) Qualitative

4. **(a)** Numerical, continuous **(b)** Numerical, discrete
   (c) Numerical, continuous **(d)** Numerical, continuous
   (e) Numerical, continuous **(f)** Numerical, continuous
   (g) Categorical

5. **(a)** Inferential. **(b)** Descriptive

6. **(a)** Inferential **(b)** Descriptive **(c)** Descriptive
   (d) Inferential **(e)** Inferential

7. **(a)** Non-random **(b)** Judgement sample
   (c) Non-sampling errors: there will be bias in the sample as no randomisation was used to obtain it. Not all members of the population have the same chance of being selected, this will limit generalisations of the results.
   (d) Random **(e)** Simple random
   (f) Sampling errors: differences between the sample and the population.

8. Stratified sampling

9. Any appropriate stratified sampling plan.

10. Randomly selected quota sampling. Equal quotas of men and women but the individuals are selected randomly.

Exercise 7.2

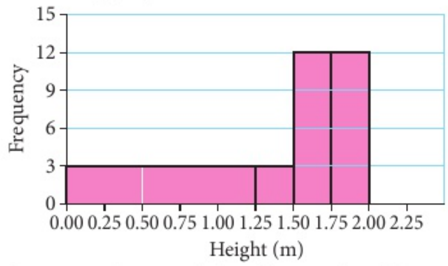
1. **(a)** $x = 168, y = 200$ **(b)** 168
   (c) 18

2. **(a)**

Height h (m)	Frequency
$h < 1.25$	3
$1.25 \leqslant h < 1.50$	3
$1.50 \leqslant h < 1.75$	12
$1.75 \leqslant h \leqslant 2.00$	12

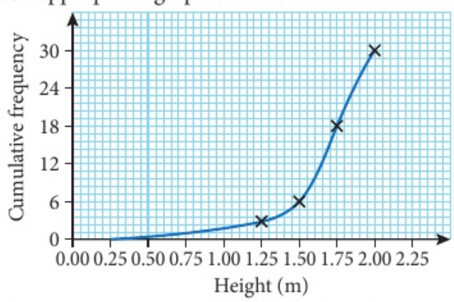
or other appropriate intervals.

(b)



Or appropriate graph based on students' intervals

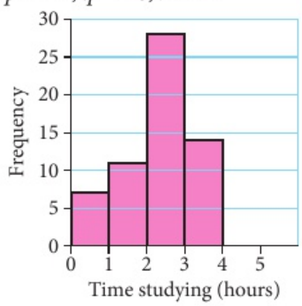
(c)

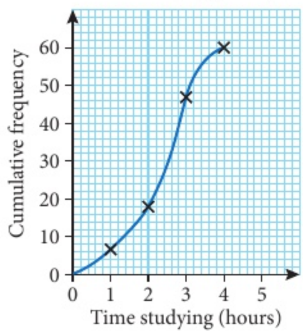


Or appropriate graph based on students' intervals

(d) The majority of students are over 1.50 m tall equally split between 1.50-1.75 and 1.75-2.00 m.

3. **(a)** $p = 11, q = 46, r = 14$
   (b)





(c) Data is left skewed with the majority of students studying more than 2 hours a night. Very few students studied less than 1 hour a night.

4. (a) 55 (b) 7.27%
(c) 8, 14.5% (d) 44 students

5. (a)

Mass of students M (kg)	Frequency
$15 \leqslant M < 30$	13
$30 \leqslant M < 45$	17
$45 \leqslant M < 60$	7

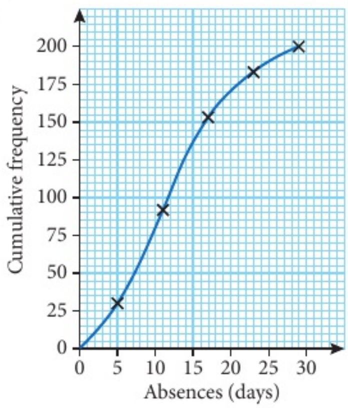
(b) 37 (c) $30 \leqslant M < 45$
(d) 27 (e) 27 kg

6. (a) 40.4 cm to 54 cm (b) 30
(c) 11 (d) 2%
(e) In the second sample fewer pike were shorter than 42 cm (4), more pike were longer than 47 cm (70) and a larger percentage are big (40.4%). There were also fewer pike in the sample.

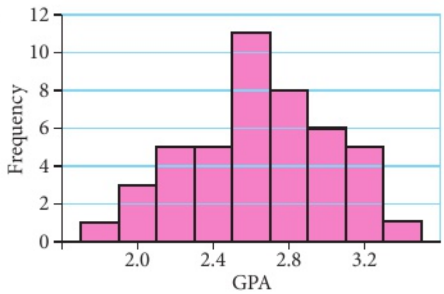
7. (a)

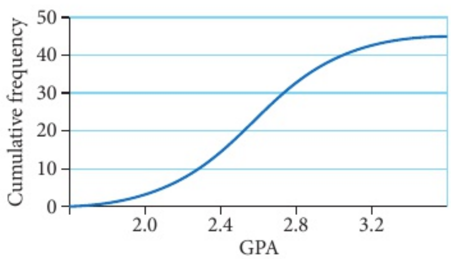
Absences (days)	Frequency	Cumulative Frequency
$0 \leqslant x \leqslant 5$	30	30
$6 \leqslant x \leqslant 11$	62	92
$12 \leqslant x \leqslant 17$	61	153
$18 \leqslant x \leqslant 23$	30	183
$19 \leqslant x \leqslant 29$	17	200

(b)



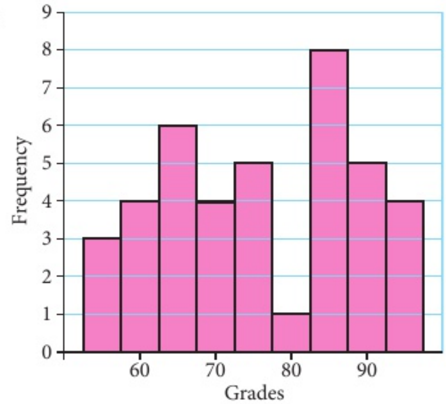
8.





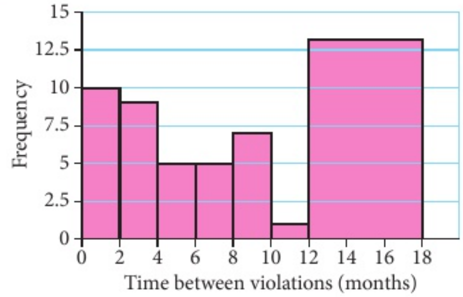
The majority of the results are grouped with a mode of 2.6 with a slight positive skew.

9.



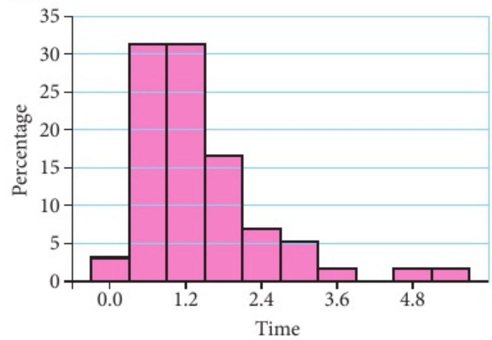
The grades appear to be divided into two groups, one with mode around 65 and the other around 85. No outliers are detected.

10. (a)

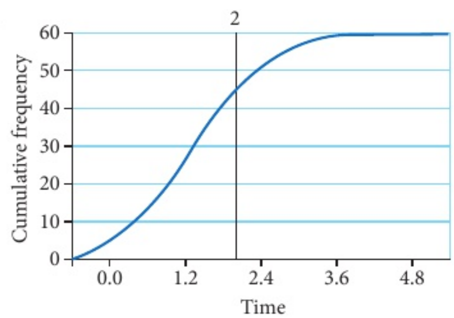


(b) Depends on how the intervals are selected.
(c) 72%

11. (a)



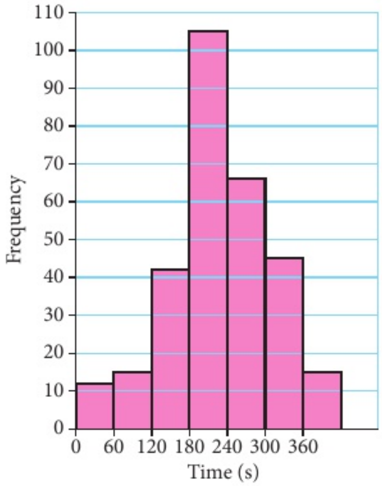
(b)



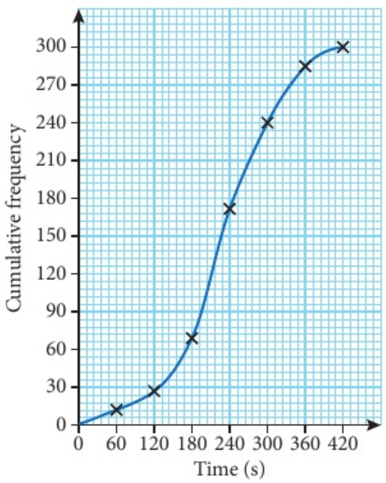
(c) Drawing a vertical line at 2 minutes, it intersects the cumulative frequency graph at approximately 45 out of 60. About 15 customers have to wait more than 2 minutes.

12. (a) 30 km h^{-1} (b) 13.3%
(c) $k = 45$ (d) 48 km h^{-1}

13. **(a)**



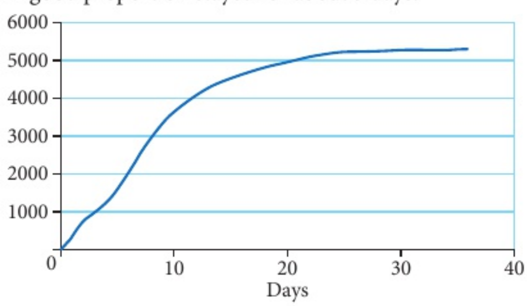
(b)



(c) 288 seconds **(d)** 288 seconds

14. **(a)** Skewed to the right, there is a mode at about 7 days stay, and a few extremes that stayed more than 20 days. A good proportion stayed for about 3 days.

(b)

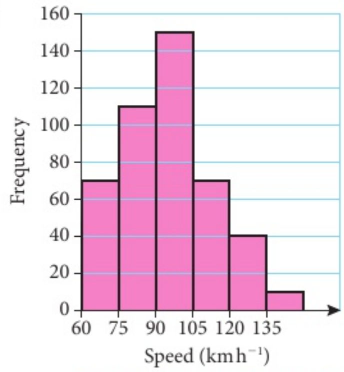


Approximately 35% of the patients.

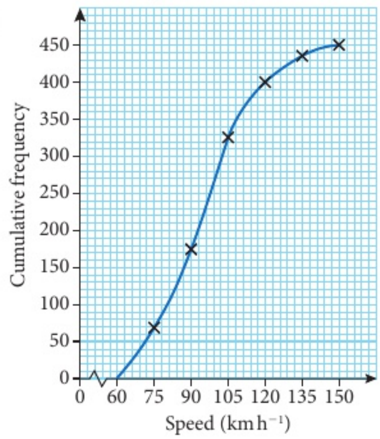
15. **(a)**

Speed, s (km h^{-1})	Frequency	Cumulative Frequency
$60 \leqslant s < 75$	70	70
$75 \leqslant s < 90$	110	180
$90 \leqslant s < 105$	150	330
$105 \leqslant s < 120$	70	400
$120 \leqslant s < 135$	40	440
$135 \leqslant s$	10	450

(b)



(c)



(d) 3 %

Exercise 7.3

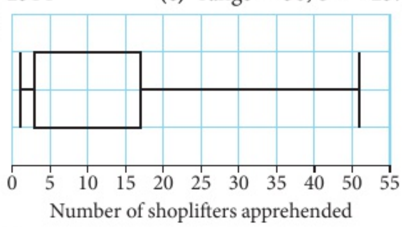
1. **(a)** $n = 4$ **(b)** $\mu = 9{,}080$, $\sigma = 2{,}040$

(c) 24% **(d)** 7,000

2. **(a)** mean = 11, median = 3, mode = 1

(b) 2314 **(c)** range = 50, $\sigma^2 = 257$, $\sigma = 16.0$

(d)



Lower limit is equal to Q_1.

(e) Outliers = median $\pm$ 1.5 IQR

IQR = 17 − 1 = 16

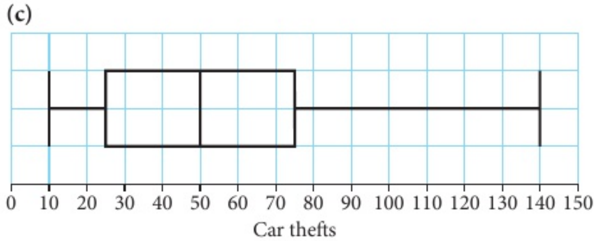
Outliers $< 3 - 16$ and $> 3 + 16$

$51 > 3 + 16$, therefore 51 is an outlier.

3. **(a)** $\mu = 55.8$, median = 60, mode = 10, 30, 60 and 70 (all occur twice)

(b) range = 130, $\sigma = 36.6$

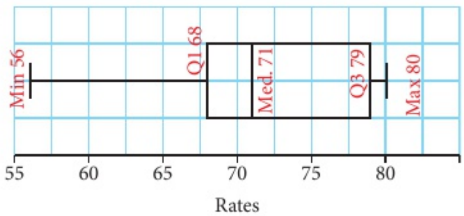
(c)



(d) Outliers = median ± 1.5 IQR
IQR = 75 − 25 = 50
Outliers < 60 − 50 and > 60 + 50
140 > 60 + 50, therefore 140 is an outlier

4. (a) $\mu = 71.5, \sigma = 7.04$

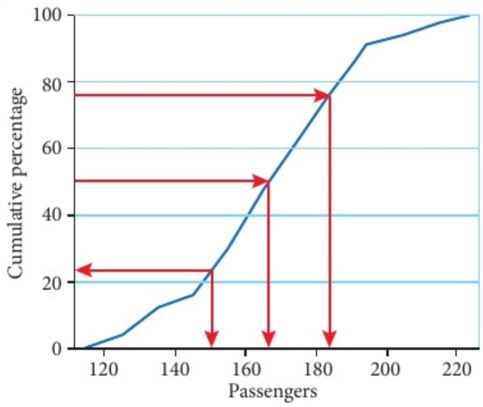
(b)



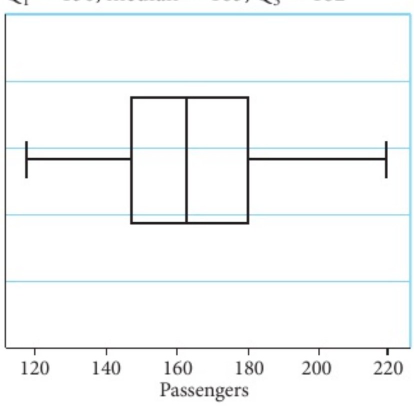
(c) IQR = 11. There are no outliers.

5. (a) $\mu = 162.6, \sigma = 23.11$

(b)



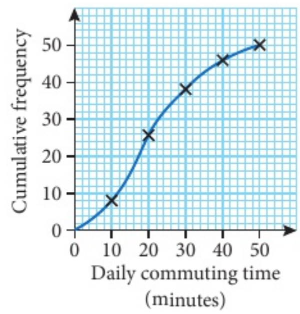
$Q_1 \approx 150$, median ≈ 165, $Q_3 \approx 182$



(c) Your GDC shows: $Q_1 = 146.75$, $Q_3 = 179.25$, IQR = 32.5 No outliers.

6. (a)

Daily commuting time (minutes)	Frequency	Cumulative Frequency
$0 \leqslant t < 10$	8	8
$10 \leqslant t < 20$	18	26
$20 \leqslant t < 30$	12	38
$30 \leqslant t < 40$	8	46
$40 \leqslant t < 50$	4	50



(b) $Q_1 \approx 14$, median ≈ 20, $Q_3 \approx 30$, IQR ≈ 16

(c) $\mu = 21.4, \sigma = 11.6$

7. (a) $\mu = €108.4$

(b) $\sigma = €35.6$

(c) $\mu = €84.4, \sigma = €35.6$

8. (a) $\mu = 106{,}500$, median = 76,000, $\sigma = 93{,}200$, IQR = 37, 000

(b) Yes 382,000 is an outlier
$\mu = 75{,}900$, median = 74,000, $\sigma = 16{,}600$, IQR = 21, 500
Standard deviation, as it uses the square of the difference from the mean so outliers make up a greater contribution.

(c) The second set, they are more representative of the data set.

9. (a) $x = 5$ **(b)** median = 5, $\sigma = 1.41$

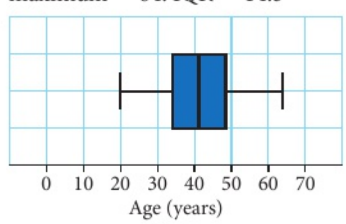
(c) 15 **(d)** 4.73

10. (a) $\mu = 31.4, \sigma = 9.35, Q_1 = 25.5$, median = 28.5, $Q_3 = 36$, IQR = 10.5

(b) More than 50% of customers spend less than the mean time on coffee and dessert, this means that a few customers are the main cause of losses.

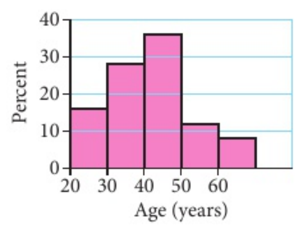
11. (a) minimum = 20, $Q_1 = 34$, median = 41, $Q_3 = 48.5$ and maximum = 64. IQR = 14.5

(b)

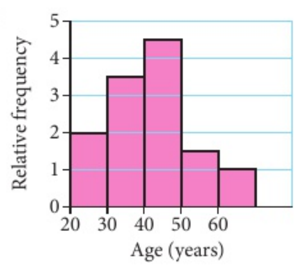


The data is very nearly symmetrical about the median.

(c)



(d)



The most frequent interval is 40–50 and the majority of the jockeys are under 50.

Chapter 7 Practice questions

1. **(a)** 30°C **(b)** 30°C, 24°C and 12°C
 (c) 22.5°C
2. **(a)** **(i)** $\mu = 13.7$ **(ii)** $\sigma = 2.52$
 (b) 13.1
3. **(a)** $a = 4$ **(b)** $b = 2$
4. **(a)** $x = 30$
 (b) **(i)** $\mu = 2.6$ **(ii)** 3 **(iii)** $\sigma = 1.06$
 (c) **(i)** $\mu = 4.6$ **(ii)** 5 **(iii)** $\sigma = 1.06$
5. **(a)** 1.75 **(b)** 2
6. **(a)** 61 **(b)** 14 **(c)** 20 **(d)** 53 kg
7. **(a)** **(i)** 13 seconds
 (ii) $Q_1 = 10$ seconds, $Q_3 = 16$ seconds
 (iii) 6 seconds
 (b) 1400 **(c)** 14 seconds
 (d) **(i)** 500 **(ii)** 150
 (e) **(i)** 13.25 seconds **(ii)** 4.41 seconds
 (f) No, required time is 8.84 seconds.
8. **(a)** $170 \leqslant h < 180$ **(b)** $\mu = 171$ cm, $\sigma = 11.1$
 (c) 171 cm **(d)** 13.3 cm
 (e) 28%
9. **(a)**

Examination score x (%)	$0 \leqslant x < 20$	$20 \leqslant x < 40$	$40 \leqslant x < 60$
Frequency	14	26	56

Examination score x (%)	$60 \leqslant x < 80$	$80 \leqslant x < 100$
Frequency	18	6

 (b) 50 **(c)** 46

Chapter 8

Exercise 8.1

Note: Some answers may differ from one person to the other due to different graph accuracies.

1. **(a)** {left handed, right handed}
 (b) all real numbers from (say) 50 cm to 210 cm.
 (c) all real numbers from 0 to 720 (say).
2. {(1,h), (2, h), …, (1, t), …, (6, t)}
3. **(a)** {(1, Hearts), …, (King, Hearts), (1, Spades), …}
 (b) {[(1, hearts), (King, Diamonds)], …,[(1, Spades), (10, Diamonds)],…}
 (c) **a**: 52, **b**: 1326
4. **(a)** 0.47
 (b) anywhere from 0 to 20!
 (c) 100 000
5. **(a)** {(1, 1), (1,2), …, (4, 4)}
 (b) {3, 4, …, 9}
6. **(a)** {(b, b), (b, g), (b, y), (g, b), (g, g), (g, y), (y, b), (y, g), (y, y)}
 (b) {(y, y), (y, b), (y, g)}
 (c) {(b, b), (g, g), (y, y)}
7. **(a)** {(b, g), (b, y), (g, b), (g, y), (y, b), (y, g)}
 (b) {(y, b), (y, g)}
 (c) φ
8. **(a)** {(t,t,t), (t,t,h), (t,h,t), (h,t,t), (h,t,h), (h,h,t), (t,h,h), (h,h,h)}
 (b) {(h,t,h), (h,h,t), (t,h,h), (h,h,h)}
9. {(I, fly), (I, dr), (I, tr), (H,dr), (H, b)}, {(I, fly)}
10. **(a)** {(1, g), (1, f), …, (0, c)}
 (b) {(0, c), (0, s)}
 (c) {(1, g), (1, f), (0, g), (0, f)}
 (d) {(1, g), (1, f), (1, s), (1, c)}
11. **(a)** $\{(K_1, K_1, K_1), (K_1, K_2, K_1), (K_1, K_1, K_2), \ldots\}$
 (b) A = all triplets containing K_2
 B = all triplets Not containing K_1; C = all triplets containing K_2.
 (c) **(i)** $A \cup B$ = All males or persons who drink;
 (ii) $A \cap C$ = All single males;
 (iii) C' = All non-single persons
 (iv) $A \cap B \cap C$ = All single males who drink
 (v) $A' \cap B$ = All females who drink.
12. **(a)** {(*R, L, L, S*), (*L, R, L, R*), …}, 81
 (b) {(*R,R,R,R*), (*L,L,L,L*) (*S,S,S,S*)}
 (c) {(*R,R,L,L*), (*R,L,R,S*), …}
 (d) {(*R,L,R,S*), (*S, S, R, L*), …}
13. **(a)** {(*T, SY, O*), (*C, SN, O*), …}
 (b) {(*T, SY, O*), (*T, SY, F*), (*B, SY, O*), …}
 (c) {(*C, SY, O*), (*C, SN, O*), (*C, SY, F*), …}
 (d) $C \cap SY = \{(C, SY, O), (C, SY, F)\}$
 $C' = \{(T, \ldots, \ldots), (B, \ldots, \ldots)\}$
 $C \cup SY$ = all triplets containing C or SY.
14. **(a)** {(1,1,1), (1,1,0), (0,1,0), …}
 (b) $X = \{(1,1,0), (1,0,1), (0,1,1)\}$
 (c) $Y = \{(1,1,1), (1,1,0), (1,0,1), (0,1,1)\}$
 (d) $Z = \{(1,1,1), (1,1,0), (1,0,1)\}$
 (e) **(i)** $Z' = \{(0,1,1), (0,1,0), (0,0,1), (0,0,0), (1,0,0)\}$
 (ii) $X \cup Z = \{(1,1,0), (1,0,1), (0,1,1), (1,1,1)\}$
 (iii) $X \cap Z = \{(1,1,0), (1,0,1)\}$
 (iv) $Y \cup Z = \{(1,1,0), (1,0,1), (0,1,1), (1,1,1)\}$
 (v) $Y \cap Z = \{(1,1,1), (1,1,0), (1,0,1)\}$
15. **(a)** {1, 2, 31, 32, 41, 42, 51, 52, 341, 342, …5432}
 (b) {31, 32, 41, 42, 51, 52}
 (c) all except {1, 2}
 (d) {1, 31, 41, 51, 341, 351, 431, 451, 531, 541, 3451, 4351, …}

Exercise 8.2

1. **(a)** $U = \{0, 1, 2, 3\}$; not equally likely.
 (b) $U = \{1, 2, 3, 4, 5, 6\}$; not equally likely.
 (c) $U = \{BBB, BBG, \ldots, GGG\}$; equally likely.
 (d) $U + \{0, 1, 2, \ldots, 20\}$; not equally likely.
2. **(a)** valid **(b)** valid **(c)** not valid **(d)** valid
 (e) not valid
3. **(a)** 0.59 **(b)** 0.64 **(c)** 0.99
4. **(a)** $\frac{3}{10}$ **(b)** $\frac{3}{4}$
5. **(a)** 0.63 **(b)** 1
6. **(a)** **(i)-(iv)** $\frac{1}{52}, \frac{7}{26}, \frac{4}{13}, \frac{10}{13}$ **(b)** **(i)-(ii)** $\frac{1}{51}, \frac{13}{17}$
 (c) **(i)-(ii)** $\frac{1}{52}, \frac{10}{13}$
7. **(a)** $\frac{4}{5}$ **(b)** $\frac{11}{30}$ **(c)** 1
8. **(a)** $\frac{1}{3}$ **(b)** $\frac{1}{12}$
9. **(a)** $\frac{1}{7}$ **(b)** $\frac{4}{7}$

10. (a) (i) $\{(1,1), (1,2) \ldots, (6,6)\}$
(ii) $\frac{1}{6}$ (iii) $\frac{2}{9}$ (iv) $\frac{5}{6}$
(b) (i) 0 (ii) $\frac{1}{9}$ (iii) $\frac{5}{36}$ (iv) 0

11. (a) 0.496 (b) 0.0834 (c) 0.0869

12. (a) Yes (b) No (c) No

13. (a) 0.06 (b) 0.42 (c) 0.36

14. (a) 0.183 (b) 0.835

15. (a) $\frac{5}{18}$ (b) $\frac{1}{3}$ (c) $\frac{5}{18}$

16. (a) $\frac{3}{28} \simeq 0.107$ (b) $\frac{17}{190} \simeq 0.0895$
(c) $\frac{68}{95} \simeq 0.716$

17. (a) 0.096 (b) 0.008 (c) 0.41

18. (a)

Cause of death	Male	Female	Total
Heart diseases	335 002	298 840	633 842
Cancer	313 818	282112	595 930
Accidents	92 919	53652	146 571
Diabetes	43 123	36412	79 535
Other causes	588 542	668 210	1 256 752
Total	1 373 404	1 339 226	2 712 630

(b) (i) 0.506
(ii) 0.123
(iii) 0.220
(c) 0.047
(d) 0.366

19. (a) 0.596
(b) lowest India, highest Mexico
(c) 0.05

Exercise 8.3

1. (a) (i) $\frac{1}{2}$ (ii) $\frac{3}{10}$
(b) $\frac{13}{20}$ (c) $\frac{7}{10}$

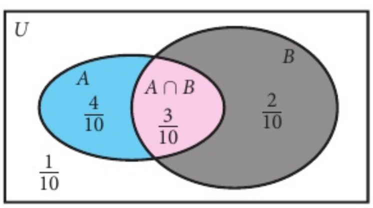
2. (a) (b) $\frac{7}{10}$

3. (a) (b) 1 (c) $\frac{3}{8}$

4. (a) $\{(1, y), (1, g), (3, y), (3, g), \ldots (6, y), (6, g), \ldots\}$
(b) (i) $\frac{1}{6}$ (ii) $\frac{2}{3}$ (iii) $\frac{1}{2}$

5. $\frac{7}{20}$

6. (a)



(b) (i) $\frac{5}{10}$ (ii) $\frac{4}{10}$ (iii) $\frac{2}{10}$ (iv) $\frac{1}{10}$ (v) $\frac{2}{3}$

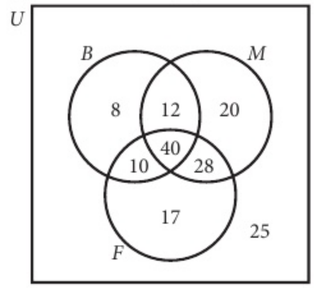
7. $P(A \cap B) = \frac{1}{9} \neq 0 \neq P(A)\,P(B)$

8. $\frac{29}{35}$

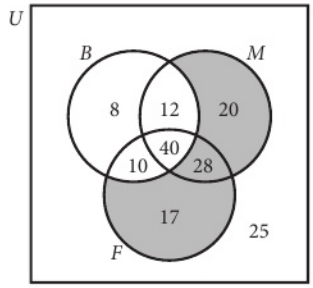
9. 0.90

10. (a) 92%
(b) (i) 0.64% (ii) 15.36% (iii) 14.72%

11 (a)

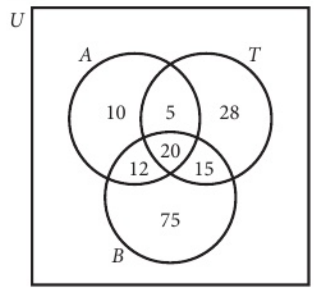


(b) 160
(c) (i)



(ii) 65
(d) (i) $\frac{100}{160} = 0.65$ (ii) $\frac{20}{65} = 0.308$ (iii) $\frac{90}{160} = 0.563$
(e) $\frac{100}{160} \times 2550 \approx 1594$

12 (a)



(b) (i)-(iv) 75, 12, 52, 15 (c) $\frac{113}{180}$ =0.628
(d) $\frac{10}{113} = 0.0885$

13. (a) $\{(1, 1), (1,2), \ldots, (6,6)\}$
(b)

x	2	3	4	5	6	7	8	9	10	11	12
P(x)	$\frac{1}{36}$	$\frac{1}{18}$	$\frac{1}{12}$	$\frac{1}{9}$	$\frac{5}{36}$	$\frac{1}{6}$	$\frac{5}{36}$	$\frac{1}{9}$	$\frac{1}{12}$	$\frac{1}{18}$	$\frac{1}{36}$

(c) $\frac{11}{36}$ (d) $\frac{11}{12}$ (e) $\frac{1}{3}$ (f) $\frac{2}{3}$

14. (a) $\frac{7}{15}$ (b) $\frac{11}{75}$ (c) $\frac{9}{35}$ (d) $\frac{46}{75}$
(e) $\frac{11}{20}$

15. (a) (i) 0.56 (ii) 0.15
(b) $\frac{15}{56}$
(c) no

16.

$P(A)$	$P(B)$	Conditions for events A and B	$P(A \cap B)$	$P(A \cup B)$	$P(A\|B)$
0.3	0.4	Mutually exclusive	0.00	0.70	0.00
0.3	0.4	Independent	0.12	0.58	0.30
0.1	0.5	Mutually exclusive	0.00	0.60	0.00
0.2	0.5	Independent	0.10	0.60	0.20

17 **(a)** 0.30 **(b)** yes

18. **(a)** 65% **(b)** 35% **(c)** 52%

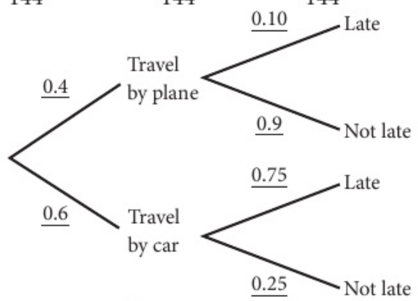
19. **(a)** 0.56 **(b)** 0.10

20. **(a)** $\frac{1}{216}$ **(b)** $\frac{91}{216}$ **(c)** $\frac{25}{72}$

21. **(a)** 0.3 **(b)** 0.441 **(c)** 0.657

22. **(a)** $\frac{23}{144}$ **(b)** $\frac{11}{144}$ **(c)** $\frac{15}{144}$ **(d)** $\frac{5}{12}$

23. **(a)**



(b) 0.51 **(c)** 0.706

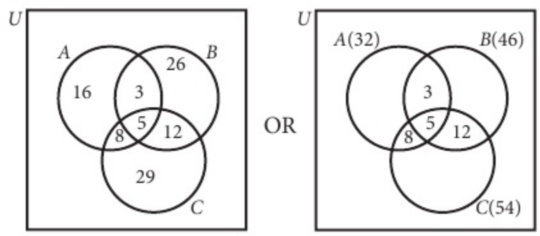
24. **(a)** 0.8805 **(b)** 0.0471

25. **(a)** 0.325 **(b)** 0.427 **(c)** 0.553

Chapter 8 Practice questions

1. **(a)** 290
(b) **(i)** 0.386
(ii) 0.434
(iii) 0.028
(iv) 0.586
(c) 0.614

2. **(a)**



(b) 1 **(c)** 71 **(d)** 45 **(e)** True
(f) 1.19%

3. **(a)** 0.30 **(b)** 0.72

4. **(a)** 0.0004 **(b)** 0.9996 **(c)** 0.0004

5. **(a)** **(i)** 0.85
(ii) 0.80
(iii) 0.15
(b) 0.083

6. **(a)** **(i)** 0.3405
(ii) 0.0108
(iii) 0.9622
(iv) 0.30
(b) Yes.

7. **(a)** 0.97 **(b)** 0.971

8. **(a)** 0.60 **(b)** Yes, $P(B|A) = P(B) = 0.60$

9. **(a)**

	Boys	Girls
Passed the ski test	32	16
Failed the ski test	12	12
Training, but did not take the test yet	22	16
Too young to take the test	4	6

(b) **(i)** 0.6
(ii) 0.56
(iii) 0.1463

10. **(a)** 0.4 **(b)** 0.6

11. **(a)** 0.38 **(b)** 0.283

12. **(a)** **(b)** $\frac{11}{36}$

13. **(a)**
(b) **(i)** 2
(ii) $\frac{1}{18}$
(c) No, $n(X \cap Y) \neq 0$

14. **(a)** **(b)** 35 **(c)** 0.35

15. **(a)** $\frac{7}{12}$ **(b)** $\frac{11}{36}$ **(c)** $\frac{1}{3}$

16. **(a)** $\frac{1}{11}$ **(b)** $\frac{12}{121}$

17. **(a)** Ind.
(b) Mutually excl.
(c) Neither

18. **(a)** **(b)** $\frac{47}{160}$ **(c)** $\frac{35}{47}$

19. **(a)** $\frac{1}{3}$ **(b)** $\frac{7}{12}$ **(c)** $\frac{3}{7}$

20. **(a)**
(b) **(i)** 0.36
(ii) 0.84
(iii) 0.429

21. **(a)** $\frac{1}{6}$ **(b)** $\frac{1}{12}$ **(c)** $\frac{2}{9}$

22. **(a)** **(i)** $\frac{8}{21}$
(ii) $\frac{1}{6}$
(iii) No $P(A \cap B) \neq P(A)P(B)$
(b) $\frac{10}{17}$
(c) $\frac{220}{441}$

Chapter 9

Exercise 9.1

1. Achilles can get arbitrarily close to the tortoise. Select any distance between the tortoise and Achilles and after some number of intervals Achilles must be that close. Therefore, the distance between the tortoise and Achilles approaches zero.

2. **(a)** 4 **(b)** -5 **(c)** 6 **(d)** 3
(e) 0.354

3. **(a)** **(i)** 31 cm s^{-1} **(ii)** 92.5 cm s^{-1} **(iii)** 153 cm s^{-1}
(b) Around 92 cm s^{-1} by drawing a line by eye and estimating two points on the line, then calculating slope, e.g. (5, 325) and (1.5, 0) yields 92.9 cm s^{-1}.

4. **(a)** **(i)** $3.75\,\text{km h}^{-1}$ **(ii)** about $12.4\,\text{km h}^{-1}$ **(iii)** No.
(b) **(i)** $8\,\text{km h}^{-1}$ **(ii)** about $12.4\,\text{km h}^{-1}$
(iii) The gradient appears to approach zero at Q, so the bicyclist is likely stopped at this point.
(c) **(i)** $7.33\,\text{km h}^{-1}$ **(ii)** $0\,\text{km h}^{-1}$
(iii) The gradient appears to approach zero at P, so the bicyclist is likely stopped at this point.
5. **(a)** $-11.3\,°\text{C min}^{-1}$
(b) $-27.8\,°\text{C min}^{-1}$
(c) $-1.73\,°\text{C min}^{-1}$
6. **(a)** The slope is very negative, then zero, then very positive.
(b) at $x = -1.5$, about $-1.13\,\text{mm}^{-1}$
at $x = -1$, about $-0.577\,\text{mm}^{-1}$
at $x = 0$, $0\,\text{mm}^{-1}$
at $x = 1$, about $5.77\,\text{mm}^{-1}$
at $x = 1.5$, about $1.13\,\text{mm}^{-1}$
The curve is symmetric over the y axis so the slope at x is the negative of the slope at $-x$.
(c) Vertical/undefined.
7. At A, $m = 0.368$; at B, $m = 1$; at C, $m = 2.72$. It appears that at A, $m = \frac{1}{e}$ and at B, $m = e$
This suggests that the value of the derivative at a point is equal to the value of the function at that point.

Exercise 9.2

1. **(a)** km hour^{-1} **(b)** $\text{cm}^3\,\text{cm}^{-1}$ **(c)** N m^{-1}
(d) thousands of individuals per year
(e) euros per shirt
(f) dollars per item
2. **(a)** 2.86s
(b) $140\,\text{cm s}^{-1}$
(c)

t	distance travelled	velocity
0	0	40
1	75	110
2	220	180
3	435	250

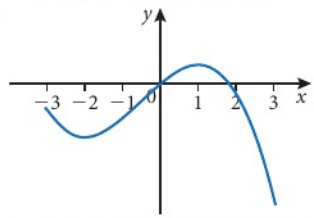
(d) The distance computed for $t = 3$ is longer than the length of the ramp.
(e) $d'(2.86) = 240\,\text{cm s}^{-1}$ **(f)** $40\,\text{cm s}^{-1}$
(g) 1.89 s **(h)** $d'(1.89) = 172\,\text{cm s}^{-1}$
(i) 1.29 s
3. **(a)** $100\pi = 314\,\text{cm}^2$ **(b)** $\text{cm}^2\,\text{cm}^{-1}$
(c) $20\,\text{cm}^2\,\text{cm}^{-1}$ **(d)** $40\pi = 126\,\text{cm}^2\,\text{cm}^{-1}$
4. **(a)** 17.7 km
(b) $d(5) = \frac{2}{3}(5)^3 - \frac{35}{6}(5)^2 + \frac{35}{2}(5) = 25\,\text{km}$
(c) $5\,\text{km h}^{-1}$
(d) km h^{-1}
(e) at $t = 0$, speed is $17.5\,\text{km h}^{-1}$; at $t = 5$, speed is approximately $9.17\,\text{km h}^{-1}$
(f) maximum speed of $17.5\,\text{km h}^{-1}$ at $t = 0$
(g) minimum speed of $0.486\,\text{km h}^{-1}$ at $t \approx 2.92$[1]
5. **(a)** m s^{-1}
(b) $2\,\text{m s}^{-1}$
(c) $1.73\,\text{m s}^{-1}$, $1.22\,\text{m s}^{-1}$, $1\,\text{m s}^{-1}$
(d) 3.46 m, 4.90 m, 6 m

Exercise 9.3

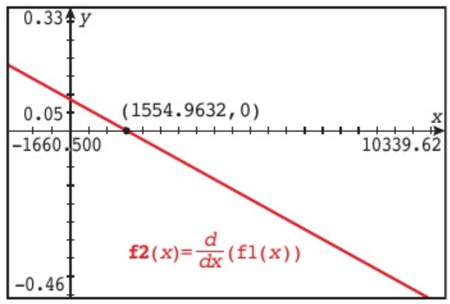
1. **(a)** Between A and B
(b) **(i)** A, B, and F **(ii)** D and E **(iii)** C
(c) Pair B and D, and pair E and F
2.

Function	Derivative diagram
f_1	d
f_2	e
f_3	b
f_4	a

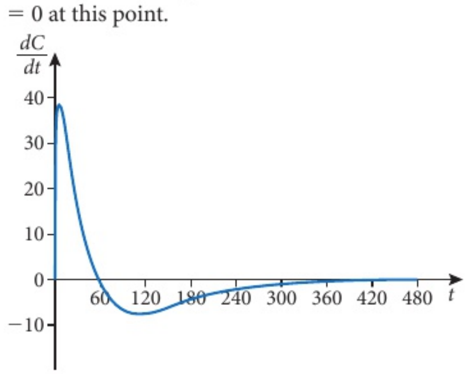
3. **(a)** Increasing for $1 < x < 5$; decreasing for $0 < x < 1$ or $5 < x < 6$
(b) Increasing for $0 < x < 1$, $3 < x < 5$; decreasing for $1 < x < 3$, $5 < x < 6$
4. **(a)** $-3 < x < -2$ or $1 < x < 3$
(b)



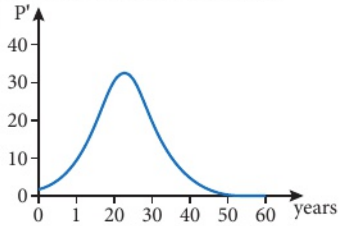
5. **(a)** If $T'(r) > 0$, then the torque is increasing as RPM increases.
(b)



Since $T'(r) > 0$ for $0 \leqslant r < 1555$, the torque is increasing on that interval. Since $T'(r) < 0$ for $r > 1555$, the torque is decreasing on that interval.
(c) The maximum torque is 505.8 Nm at 1555 RPM. $T'(r) = 0$ at this point.
6. **(a)**



(b) Increasing on $0 < t < 57.0$, decreasing on $t > 57.0$
(c) 110 min, $-7.51\,\text{nl min}^{-1}$
(d) 57.0 min

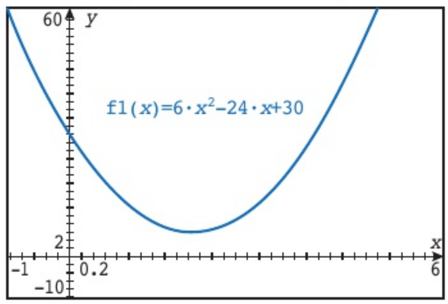
7. **(a)** thousands of people per year
(b)


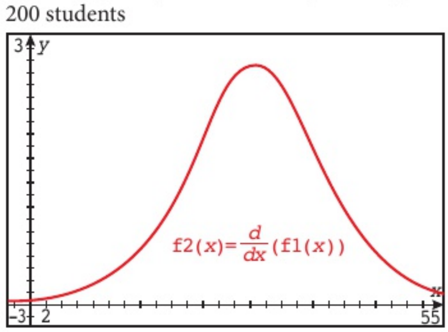
(c) The derivative is always positive
(d) There is a rapid increase ('boom') in population, then it levels off
(e) 8830 people per year
(f) 1995 to 2010
(g) in the year 2002, 32500 people per year

Exercise 9.4

1. **(a) (i)** $y' = 6x - 4$ **(ii)** -4
(b) (i) $y' = -2x - 6$ **(ii)** 0
(c) (i) $y' = -\dfrac{6}{x^4}$ **(ii)** -6
(d) (i) $y' = 5x^4 - 3x^2 - 1$ **(ii)** 1
(e) (i) $y' = 2x - 4$ **(ii)** 0
(f) (i) $y' = 2 - \dfrac{1}{x^2} + \dfrac{9}{x^4}$ **(ii)** 10
(g) (i) $y' = 1 - \dfrac{2}{x^3}$ **(ii)** 3
2. **(a)** $(0, 0)$ **(b)** $(2, 8)$ and $(-2, -8)$
(c) $\left(\dfrac{5}{2}, -\dfrac{21}{4}\right)$ **(d)** $(1, -2)$
3. $a = -5, b = 2$
4. The power rule gives us $y' = 1 \times mx^{1-1} + 0 = m$; therefore, the gradient of $y = mx + b$ is always m.
5. **(a)** cases year^{-1}
(b) $\dfrac{dC}{dt} = -666t^2 + 14520t - 12700$
(c) At $t = 7$, $\dfrac{dC}{dt} = 56300$ cases year^{-1}. At the beginning of 1990, the cumulative number of cases was increasing at a rate of 56300 cases per year.
(d) In 1990, there were 58700 new cases reported. This represents is the average rate of change during 1990, while the result in **(c)** is the instantaneous rate of change at the beginning of 1990.
(e) For $0 \leqslant t < 0.913$, $\dfrac{dC}{dt} < 0$. Therefore, during most of 1983 the cumulative number of cases decreased.
For $t > 0.913$, $\dfrac{dC}{dt} > 0$, so from late 1983 until 1998 the cumulative number of cases increased.
6. **(a)** $\dfrac{dP}{dt} = 34t - 3t^2$
(b) For $0 \leqslant t \leqslant 11.3$, $\dfrac{dP}{dt} > 0 \Rightarrow$ population is increasing.
For $11.3 < t \leqslant 20$, $\dfrac{dP}{dt} < 0 \Rightarrow$ population is decreasing.
(c) 40 bacteria min^{-1}
(d) 15 min
(e) 96.3 bacteria min^{-1} at 5.67 min
7. **(a)** $h'(t) = -9.8t + 16$, this function gives the velocity of the tennis ball in m s^{-1}
(b) 6.20 m s^{-1}
(c) 2.65 s
(d) 1.63 s, this when the tennis ball has reached its maximum height.
(e) Just before impact at 3.33 seconds, the speed is 16.6 m s^{-1}

Chapter 9 Practice questions

1. **(a)** $f'(x) = 50x$ **(b)** $f'(x) = 60(x - 2)$
(c) $f'(x) = -\dfrac{1}{x^2} + \dfrac{8}{x^3}$ **(d)** $f'(x) = \dfrac{1}{2} - \dfrac{1}{x^2}$
(e) $f'(x) - \dfrac{91}{3x^{14}}$
2. **(a)** $t = 3$ or $t = 5$ s **(b)** $v(t) = 3t^2 - 14t + 7$
(c) $t = 0.570$ or $t = 4.10$ s **(d)** $a(t) = 6t - 14$
(e) (i) -14 cm s^{-2} **(ii)** 19 cm s^{-2}
3. $a = 1$
4. **(a)** $\dfrac{dy}{dx} = \dfrac{2}{x^2}$ **(b)** Increasing for all $x \neq 0$
5. **(a)** 10 m s^{-1} **(b)** $t = 10$
(c) 50 m
6. **(a)** $c'(x) = 6x^2 - 24x + 30$
(b) $c'(1) = 12$, the marginal cost when producing 100 baseball caps is 12000 THB.
(c) Since $c'(x)$ is positive (it is a quadratic opening up, with no x-intercepts) for all $x > 0$, $c(x)$ is positive for all $x > 0$:


(d) $p(x) = -2x^3 + 12x^2 - 15x$
(e) $p'(x) = -6x^2 + 24x - 15$
(f) $p(x)$ increases for $0.775 < x < 3.22$, decreasing elsewhere
(g) Profit of 9350 THB for producing 322 baseball caps.
7. **(a)** -9000 l min^{-1}
(b) -6000 l min^{-1}
(c) Since the rate at which the water is draining is a negative quantity, it is increasing (slowing down).
8. **(a)** 200 students
(b)


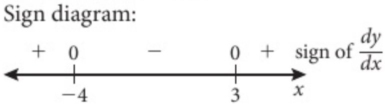
(c) 4.94 students day^{-1}
(d) 1.57 students day^{-1}
(e) 9.12 students day^{-1} on day 29
(f) day 48

9. (a) $M'(t) = -\dfrac{4500}{x^2}$
   (b) Since x^2 is always positive, $M'(t)$ is always negative $\Rightarrow$ $M(t)$ is a decreasing function.
   (c) After 2.12 years
   (d) At $t = 1$
   (e) After 9.49 years, \$1220
10. (a) $P'(n) = -6n^2 + 12n + 1$. $P'(n)$ is a quadratic function opening downward with zeros at -0.0801 and 2.08, hence $P'(n) > 0$ on $0 \leqslant n < 2.08$, so $P(n)$ is increasing on that interval.
   (b) 2080 t-shirts, \$10,000 profit
11. (a) $E'(t) = -0.0021t^2 + 0.0556t - 0.0843$
   (b) $a = 1972$, $b = 1995$
12. (a) CD sales increased sharply and then decreased.
   (b) During 1993
   (c) At the very end of 1998
   (d) During 2007, CD sales decreased by 158 million.
   (e) graph **A**
13. (a) 0.0347
   (b) 510 bacteria min^{-1}

Chapter 10

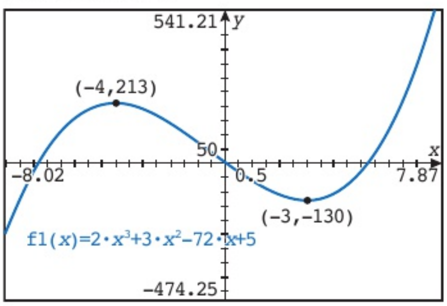
Exercise 10.1

1. (a) $\dfrac{dy}{dx} = 2x - 2 = 0 \Rightarrow x = 1, y = 1^2 - 2(1) - 6 = -7$
   $\therefore (1, -7)$
   (b) $y' = 8x + 12 = 0 \Rightarrow x = -1.5, y = 4(-1.5)^2 + 12(-1.5) + 17 = 8 \therefore (-1.5, 8)$
   (c) $\dfrac{dy}{dx} = -2x + 6 = 0 \Rightarrow x = 3, y = -(3)^2 + 6(3) - 7 = 2$
   $\therefore (3, 2)$
2. (a) (i) $\dfrac{dy}{dx} = 6x^2 + 6x - 72 = 0 \Rightarrow (-4, 213), (3, -130)$ are stationary points
   (ii) Sign diagram:

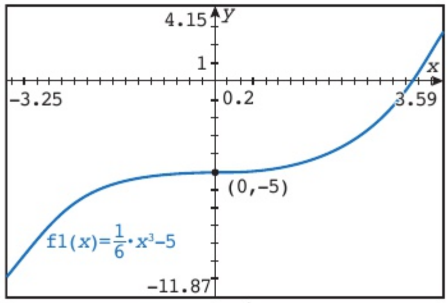


Therefore $(-4, 213)$ is local maximum, $(3, -130)$ is a local minimum.

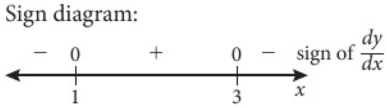
   (iii)



   (b) (i) $y' = \dfrac{1}{2}x^2 \Rightarrow (0, -5)$ is stationary
   (ii) Sign diagram:
   Therefore $(0, -5)$ is neither a minimum nor a maximum.
   (iii)

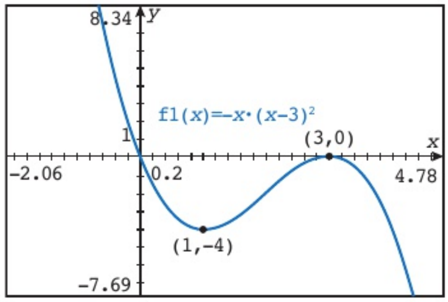


   (c) (i) $y = -x^3 + 6x^2 - 9x \Rightarrow \dfrac{dy}{dx} = -3x^2 + 12x - 9 \Rightarrow (1, -4), (3, 0)$
   (ii) Sign diagram:

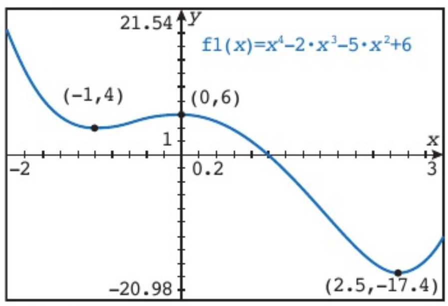


Therefore $(1, -4)$ is a local minimum and $(3, 0)$ is a local maximum.

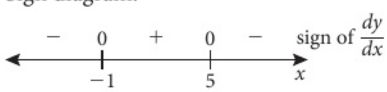
   (iii)



   (d) (i) $(-1, 4), (0, 6), \left(\dfrac{5}{2}, -\dfrac{279}{16}\right)$
   (ii) $(-1, 4)$ is a local minimum since $f'(x)$ changes from negative to positive at $x = -1$; $(0, 6)$ is a local maximum since $f'(x)$ changes from postive to negative at $x = 0$; $\left(\dfrac{5}{2}, -\dfrac{279}{16}\right)$ is a local minimum since $f'(x)$ changes from negative to positive at $x = \dfrac{5}{2}$.
   (iii)

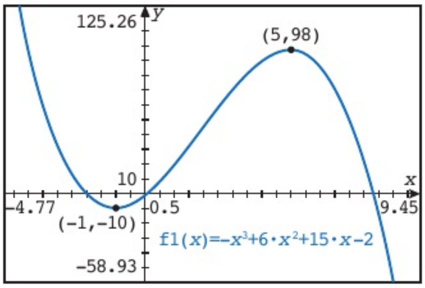


   (e) (i) $(-1, -10), (5, 98)$
   (ii) Sign diagram:



Therefore $(-1, -10)$ is a local maximum and $(5, 98)$ is a local minimum.

(iii)

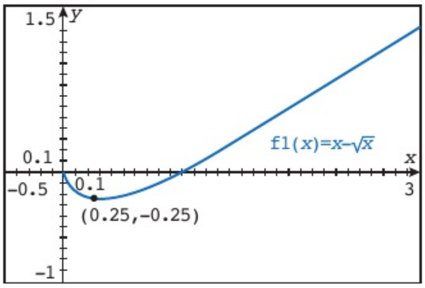


(f) (i) $y = x - x^{\frac{1}{2}} \Rightarrow \dfrac{\mathrm{d}y}{\mathrm{d}x} = 1 - \dfrac{1}{2}x^{-\frac{1}{2}} = 1 - \dfrac{1}{2\sqrt{x}} = 0$

$\Rightarrow x = \dfrac{1}{4}$, $\therefore$ stationary point at $\left(\dfrac{1}{4}, -\dfrac{1}{4}\right)$

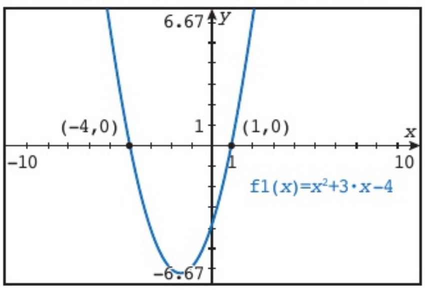
(ii) $\dfrac{\mathrm{d}y}{\mathrm{d}x}$ changes sign from negative to positive at $x = \dfrac{1}{4}$ so $\left(\dfrac{1}{4}, -\dfrac{1}{4}\right)$ is a local minimum.

(iii)



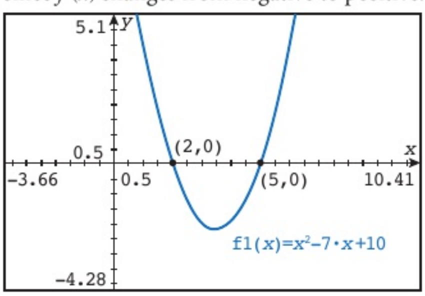
3. **(a)** Local max $(-2, 16)$; local min $(2, -16)$
(b) Local max $(0, 0)$, local minima $(-2, -4)$ and $(2, -4)$
(c) Local max $(2, 4)$, local min $(-2, -4)$
(d) Local max $(1, 2)$, local min $(-1, -2)$
(e) Local max $(0, 5)$, local minima $(-1, 0)$ and $(2, -27)$

4. **(a) (i)**



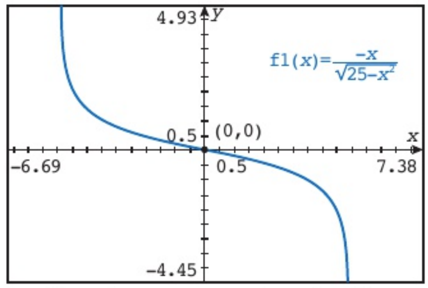
(ii) Local maximum at $x = -4$ since $f'(x)$ changes from positive to negative; local minimum at $x = 1$ since $f'(x)$ changes from negative to positive.

(b) (i)



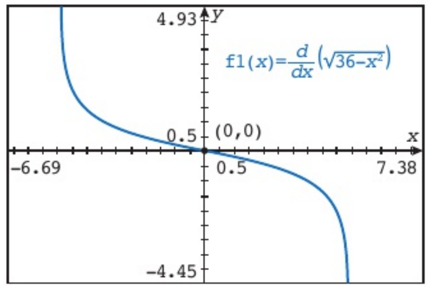
(ii) Local maximum at $x = 2$ since $f'(x)$ changes from positive to negative; local minimum at $x = 5$ since $f'(x)$ changes from negative to positive.

(c) (i)

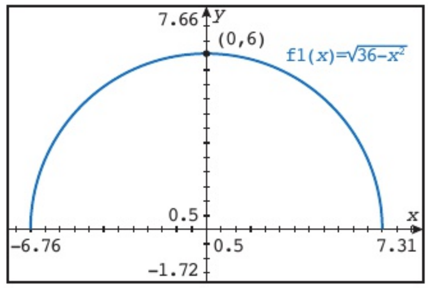


(ii) Local maximum at $x = 0$ since $f'(x)$ changes from positive to negative.

5. **(a) (i)**

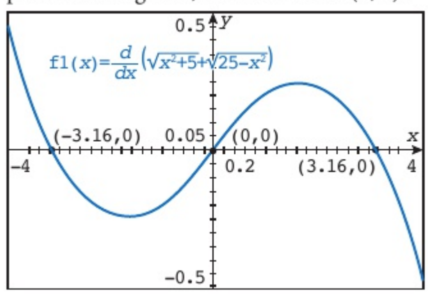


(ii)

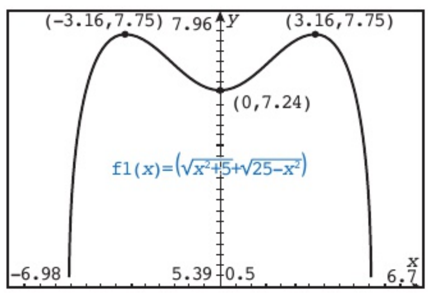


Local maximum at $x = 0$ since $f'(x)$ changes from positive to negative; maximum is at $(0, 6)$.

(b) (i)

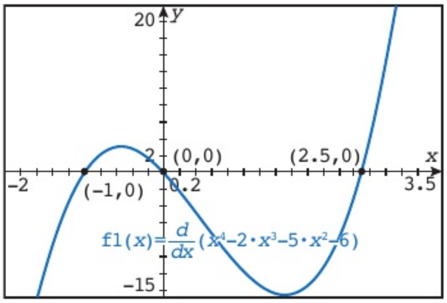


(ii)

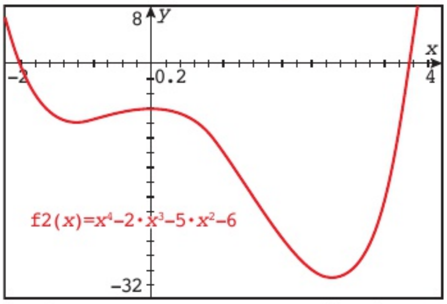


Local maximum at $x = -3.16$ since $f'(x)$ changes from positive to negative; local minimum at $x = 0$ since $f'(x)$ changes from negative to positive; local maximum at $x = 3.16$ since $f'(x)$ changes from positive to negative.

(c) (i)



(ii)



Local minimum at $x = -1$ since $f'(x)$ changes from negative to positive; local maximum at $x = 0$ since $f'(x)$ changes from positive to negative; local minimum at $x = 2.5$ since $f'(x)$ changes from negative to positive. Graph:

6. **(a)** **(i)** Increasing on $1 < x < 5$; decreasing elsewhere.
(ii) Local min at $x = 1$, local max at $x = 5$
(b) **(i)** Increasing on $0 < x < 1$ or $3 < x < 5$, decreasing on $1 < x < 3$ or $x > 5$
(ii) Local maxima at $x = 1$ and $x = 5$; local min at $x = 3$

7. **(a)** $v(t) = 3t^2 - 8t + 1$; $a(t) = 6t - 8$
(b) Max displacement is -6.88 m; $t = \dfrac{4 + \sqrt{13}}{3} = 2.54$ s
(c) Min velocity $= -4.33\ \text{m s}^{-1}$; $t = 1.33\ \text{s}^{-1}$
Starting from -6 m, the object moves toward the origin, passing the origin at $t = 0$ seconds, before reaching its maximum positive displacement at 0.131 s, then moving in the negative direction reaching minimum displacement of -6.88 m at 2.54 seconds, where it changes directions again.

8. **(a)** Max: $(-2, 18)$, $(2, 18)$; Min: $(0, 0)$
(b) Min: $(2, 12)$
(c) Min: $(-3, -18)$, $(3, 18)$
(d) Max: $(0, 0)$; Min: $(-1, -4)$, $(1, -4)$
(e) Max: $(0, 1)$, $(2.51, 1)$; Min: $(1.77, -1)$
(f) Min: $(-1, -0.368)$
(g) Max: $(2.03, 1.82)$; Min: $(0, 0)$, $(4.91, -4.81)$

9. **(a)** $v(0) = 27\ \text{m s}^{-1}$
(b) $v(3) = 45\ \text{m s}^{-1}$
(c) 0.5 s or 2.25 s. These are times when the displacement is at a relative maximum or minimum.

10. $x = 5.77$ tonnes; $D = 34.6 = 34641$ dollars. D' is negative to the left and positive to the right of $x = 5.77$.

11. **(a)** $10\ \text{m s}^{-1}$
(b) 10 s
(c) 50 m

12. **(a)** $v = 14 - 9.8t$
(b) Max height is 10 m at 1.43 s
(c) Velocity at max height is zero.

13. 467 spinners; 9650 RMB profit.

14. **(a)** 122 wheels, 16,600 EUR **(b)** 170 wheels, 112 EUR

15. 16.4 minutes

Exercise 10.2

1. **(a)** **(i)** $y' = 6x - 4$ **(ii)** -4
(iii) $y = -4x$ **(iv)** $y = \frac{1}{4}x$
(b) **(i)** $y' = -2x - 6$ **(ii)** 0
(iii) $y = 10$ **(iv)** $x = -3$
(c) **(i)** $y' = -\dfrac{6}{x^4}$ **(ii)** -6
(iii) $y = -6x - 4$ **(iv)** $y = \frac{1}{6}x + \frac{13}{6}$
(d) **(i)** $y' = 5x^4 - 3x^2 - 1$ **(ii)** 1
(iii) $y = x - 2$ **(iv)** $y = -x$
(e) **(i)** $y' = 2x - 4$ **(ii)** 0
(iii) $y = -16$ **(iv)** $x = 2$
(f) **(i)** $y' = 2 - \dfrac{1}{x^2} + \dfrac{9}{x^4}$ **(ii)** 10
(iii) $y = 10x - 10$ **(iv)** $y = -\frac{1}{10}x + \frac{1}{10}$
(g) **(i)** $y' = 1 - \dfrac{2}{x^3}$ **(ii)** 3
(iii) $y = 3x + 3$ **(iv)** $y = -\frac{1}{3}x - \frac{1}{3}$
(h) **(i)** $y' = 2x + 2$ **(ii)** -4
(iii) $y = -4x - 8$ **(iv)** $y = \frac{1}{4}x + \frac{19}{4}$
(i) **(i)** $y' = 3x^2 + 2x$ **(ii)** 0
(iii) $y = \dfrac{4}{27}$ **(iv)** $x = -\frac{2}{3}$
(j) **(i)** $y' = 6x - 1$ **(ii)** -1
(iii) $y = -x + 1$ **(iv)** $y = x + 1$
(k) **(i)** $y' = 2 - \dfrac{1}{x^2}$ **(ii)** -2
(iii) $y = -2x + 4$ **(iv)** $y = \frac{1}{2}x + \frac{11}{4}$

2. **(a)** $x = 0$ **(b)** $x = \pm 2$ **(c)** $x = \frac{5}{2}$ **(d)** $x = 1$

3. $y = 2x$, $y = -x + 1$, $y = 2x - 4$

4. $y = -2x$

5. $a = -2$, $b = -4$

6. $a = -3$, $b = 2$

7. $(3,6)$

8. $y = 7x - 16$, $y = 7x + 16$

9. $y = \frac{1}{2}x - \frac{7}{2}$; $\left(-\frac{1}{2}, -\frac{15}{4}\right)$

10. **(a)** At $x = 1$, $y' = -3$
(b) Tangent: $y = -3x + 3$; Normal: $y = \frac{1}{3}x - \frac{1}{3}$

11. **(a)** $y = 2x + \frac{5}{2}$ **(b)** $\left(\frac{2}{3}, \frac{41}{27}\right)$

12. Tangent: $y = -\frac{3}{4}x + 1$; Normal: $y = \frac{4}{3}x - \frac{22}{3}$

13. **(a)** $x = 1$
(b) Tangent to $y = x^2 - 6x + 20$ is $y = -4x + 1$;
Tangent to $y = x^3 - 3x^2 - x$ is $y = -4x + 19$

14. **(a)** $(-2, 4)$; $(2, -20)$ **(b)** $y = -2x$, $y = -10x$

15. $y = 11x - 25$, $y = -x - 1$

Exercise 10.3

1. **(a)** Maximum volume is 66.1 in^2.
(b) Box should be 5.33 by 7.83 by 1.59 inches.

2. **(a)** 10.9 m **(b)** 37.2 m

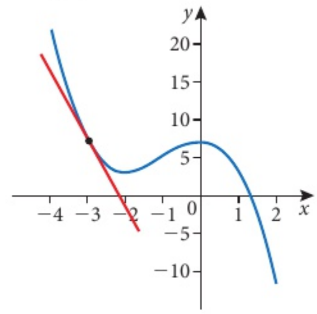
3. The lifeguard should run 283 m and swim the remaining distance; the total time will be 99.2 s.

4. $\sqrt{2}$ by $\frac{\sqrt{2}}{2}$
5. $13\frac{1}{3}$ cm by $6\frac{2}{3}$ cm
6. $3\frac{\sqrt{2}}{4}$
7. $x = 5\sqrt{2\pi} = 12.5$ cm
8. $x = 2.64$ km
9. $h = 10\frac{\sqrt{6}}{3}, r = 20\frac{\sqrt{3}}{3}$
10. (a) $h = \frac{27 - r^2}{r}$; $V = \pi r(27 - r^2)$ (b) $r = 3$ cm
11. $a = -2, b = 8, c = 10$
12. (a) Equipment maintenance $= \frac{4}{25}x^2$
 (b) $C(x) = \frac{4}{25}x^2 + 85x + 2150$
 (c) $C'(x) = \frac{8}{25}x + 85$, this is positive for all $x > 0$ hence production costs increase for all $x > 0$.
 (d) $\overline{C}(x) = \frac{0.16x^2 + 85x + 2150}{x}$
 (e) The minimum average manufacturing cost per device is \$122; 116 devices should be produced to minimize average cost per device.
13. $x = 11.5$ cm; maximum volume $= 403$ cm^3
14. 6 nautical miles

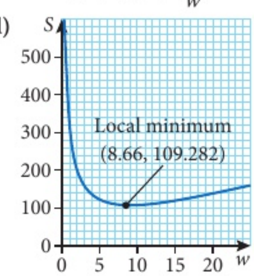
Chapter 10 Practice questions

1. (a) (i) The slant height is equal to the radius of the semicircle.
 Using $A_{\text{semicircle}} = \frac{1}{2}\pi r^2$ we get:
 $$39.27 = \frac{1}{2}\pi l^2$$
 $$25 = l^2$$
 $$5 = l$$
 The circumference of the base of the cone is equal to the arc length of the semicircle. Since $r = l = 5$, we have $C = \frac{1}{2}(2\pi r) = \pi(5) = 5\pi = 15.7$ m.
 (ii) For the distance C to be formed into a circle, it must satisfy $C = 2\pi r \Rightarrow 5\pi = 2\pi r \Rightarrow r = \frac{5}{2}$.
 (iii) $r^2 + h^2 = l^2 \Rightarrow \left(\frac{5}{2}\right)^2 + h^2 = 5^2 \Rightarrow h = \sqrt{5^2 - \left(\frac{5}{2}\right)^2}$
 $= \frac{5\sqrt{3}}{2} = 4.33.$
 (b) $h + 2r = 9.33 \Rightarrow h = 9.33 - 2r$
 (c) Using the general model for the volume of a cone, $V = \frac{1}{3}Bh$, we have $V = \frac{1}{3}\pi r^2 h$
 $\Rightarrow V = \frac{1}{3}\pi r^2(9.33 - 2r) = 3.11\pi r^2 - \frac{2}{3}\pi r^3.$
 (d) $\frac{dV}{dr} = 3.11\pi(2)r - \frac{2}{3}\pi(3)r^2 = 6.22\pi r - 2\pi r^2$
 (e) We need to find where $\frac{dV}{dr} = 0$ hence $6.22\pi r - 2\pi r^2 = 0$ $\Rightarrow (\pi r)(6.22 - 2r) = 0 \Rightarrow r = 0$ or $r = 3.11$. We discard $r = 0$. Sign analysis shows that $\frac{dV}{dr}$ is positive to the left of $r = 3.11$ and negative to the right, so the volume is maximized at $r = 3.11$ m
 (f) $V = 3.11\pi(3.11)^2 - \frac{2}{3}\pi(3.11)^3 = 31.5$ m^3

2. (a) $f'(x) = 5x^4$ (b) 80
 (c) $y - 32 = -\frac{1}{80}(x - 2) \Rightarrow 80y - 2560 = -x + 2$
 $\Rightarrow x + 80y - 2558 = 0$
3. (a) $\frac{dy}{dx} = 3x^2 + 2kx$ (b) $k = -\frac{9}{2}$
 (c) $y = (3)^3 - \frac{9}{2}(3)^2 = -\frac{27}{2}$
4. (a) $(0, 7)$ (b) $(-2, 3)$ (c) $-2 < x < 0$
 (d) $f'(x) = -3x^2 - 6x \Rightarrow f'(-3) = -9$
 $f(-3) = 7$
 $\therefore y - 7 = -9(x + 3)$ or $y = -9x - 20$
 (e) See graph.



5. (a) $f'(x) = 9x^2 - 7 - \frac{4}{x^2}$ (b) $(1.08, 9.92)$
 (c) $(-1.08, 10.1)$
6. (a) $T = 2\pi r + 4r + 4l$
 (b) $V = Bh = \frac{1}{2}\pi r^2 l \Rightarrow 0.75 = \frac{1}{2}\pi r^2 l \Rightarrow 1.5 = \pi r^2 l$
 (c) From (b), $1.5 = \pi r^2 l \Rightarrow l = \frac{1.5}{\pi r^2}$. From (a),
 $T = 2\pi r + 4r + 4l \Rightarrow T = 2\pi r + 4r + 4\left(\frac{1.5}{\pi r^2}\right)$
 $\Rightarrow T = 2\pi r + 4r + \frac{6}{\pi r^2}$
 $\Rightarrow T = (2\pi + 4)r + \frac{6}{\pi r^2}.$
 (d) $\frac{dT}{dr} = 2\pi + 4 - \frac{12}{\pi r^3}$
 (e) We solve $\frac{dT}{dr} = 2\pi + 4 - \frac{12}{\pi r^3} = 0 \Rightarrow r = \sqrt[3]{\frac{12}{\pi(2\pi + 4)}}$
 $= 0.719$ m
 (f) $l = \frac{1.5}{\pi r^2} = \frac{1.5}{\pi(0.719)^2} = 0.924$ m
 (g) $T = (2\pi + 4)(0.719) + \frac{6}{\pi(0.719)^2} = 11.1$ m
7. (a) $V = lwh \Rightarrow V = 20lw$
 (b) $V = 20lw \Rightarrow 3000 = 20lw \Rightarrow l = \frac{150}{w}$
 (c) $S = 2(20) + 4w + 2l = 40 + 4w + 2\left(\frac{150}{w}\right)$
 $= 40 + 4w + \frac{300}{w}$
 (d)



 (e) $\frac{dS}{dW} = 4 - \frac{300}{w^2}$

(f) $\dfrac{dS}{dW} = 4 - \dfrac{300}{w^2} = 0 \Rightarrow w = \sqrt{75} = 8.66$ cm

(g) $l = \dfrac{150}{w} = \dfrac{150}{\sqrt{75}} = 17.3$ cm

(h) $S = 40 + 4w + \dfrac{300}{w} = 40 + 4(\sqrt{75}) + \dfrac{300}{\sqrt{75}}$
$= 40 + 40\sqrt{3} = 110$ cm

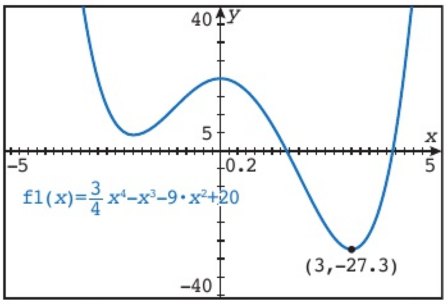
8. (a) $16x^4 - 27x$ (b) $f'(x) = 64x^3 - 27$

(c) $x = \dfrac{3}{4}$

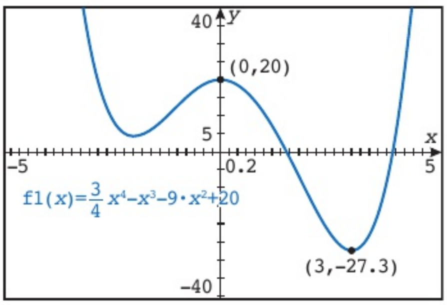
9. (a) $f(-2) = \dfrac{3}{4}(-2)^4 - (-2)^3 - 9(-2)^2 + 20 = 4$

(b) $f'(x) = 3x^3 - 3x^2 - 18x$

(c) $f'(3) = 3(3)^3 - 3(3)^2 - 18(3) = 0$ therefore $x = 3$ is a stationary point.
Now, we must show that this stationary point is in fact a local minimum. We can do this by using the first derivative test or sketching the graph on our GDC.
First derivative test:
$f'(2) < 0, f'(4) > 0$ therefore there is a local minimum at $x = 3$.
GDC graph:



(d)



(e) $(0, 20)$

(f) $f'(2) = 3(2)^3 - 3(2)^2 - 18(2) = -24$

(g) (i) $\dfrac{1}{24}$

(ii) $y + 12 = \dfrac{1}{24}(x - 2) \Rightarrow 24y + 288 = x - 2$
$\Rightarrow -x + 24y + 290 = 0 \Rightarrow x - 24y - 290 = 0$
$\therefore b = -24, c = -290$

10. (a) $f'(x) = 3ax^2 - 3$ (b) $f'(0) = -3$

(c) Local maximum implies
$f'(x) = 0 \Rightarrow f'(-2) = 3a(-2)^2 - 3 = 0 \Rightarrow a = \dfrac{1}{4}$.

11. (a) $f(1) = \dfrac{14}{(1)} + (1) - 6 = 9$

(b) $f'(x) = -\dfrac{14}{x^2} + 1$

(c) $f'(x) = -\dfrac{14}{x^2} + 1 = 0 \Rightarrow x^2 = 14 \Rightarrow x = \sqrt{14} = 3.7$
We see from the given graph that the stationary point at $x = 3.7$ must be a local minimum.

(d) From the graph, minimum occurs at $x = \sqrt{14}$
$\Rightarrow f(\sqrt{14}) = \dfrac{14}{\sqrt{14}} + \sqrt{14} - 6 = 1.48331$.
Therefore range is $1.48 \leqslant f(x) \leqslant 9$.

(e) $f(7) = \dfrac{14}{(7)} + (7) - 6 = 3$

(f) $m = \dfrac{9 - 3}{1 - 7} = -1$

(g) $M = \left(\dfrac{1 + 7}{2}, \dfrac{9 + 3}{2}\right) = (4, 6)$

(h) $f'(4) = -\dfrac{14}{4^2} + 1 = \dfrac{1}{8}$

(i) Point on function is $(4, f(4))$, $f(4) = \dfrac{14}{(4)} + (4) - 6 = 1.5$
$\Rightarrow (4, 1.5)$
Therefore equation of L is $y - 1.5 = \dfrac{1}{8}(x - 4)$
$\Rightarrow y = \dfrac{1}{8}x + 1$.

12. (a) $y = -\dfrac{75^2}{10} + \dfrac{27}{2} \times 75 = 450$ therefore A is on the track.

(b) $\dfrac{dy}{dx} = -\dfrac{2x}{10} + \dfrac{27}{2} = -0.2x + 13.5$

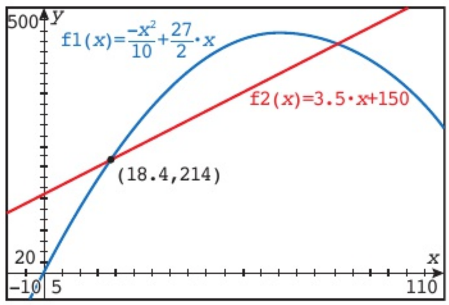
(c) Stationary point(s) occur at $\dfrac{dy}{dx} = -0.2x + 13.5 = 0$
$\Rightarrow x = 67.5$. Since A has x-coordinate 75, it cannot be the farthest point north.

(d) (i) $M = \left(\dfrac{0 + 100}{2}, \dfrac{0 + 350}{2}\right) = (50, 175)$

(ii) $m = \dfrac{350 - 0}{100 - 0} = 3.5$

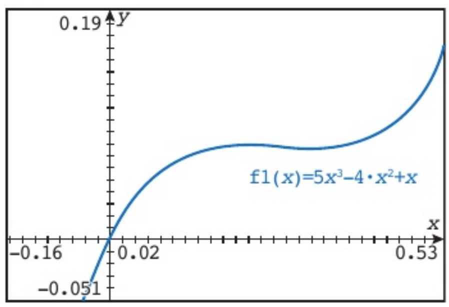
(e) $y - 150 = 3.5(x - 0) \Rightarrow 3.5x - y = -150$

(f) Use your GDCs Intersection feature to find the first point of intersection is at $(18.4, 214)$.

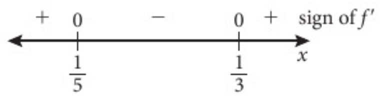


13. (a) $f'(x) = 15x^2 - 8x + 1$

(b) $f'(x) = 15x^2 - 8x + 1 = 0 \Rightarrow x = \dfrac{1}{5}, \dfrac{1}{3}$. Use GDC graph or first derivative test to find which stationary point is a local minimum and which is a local maximum. GDC graph:



Or, via first derivative test, using sign diagram:



Therefore,

(i) local maximum is at $x = \frac{1}{5}$

(ii) local minimum is at $x = \frac{1}{3}$

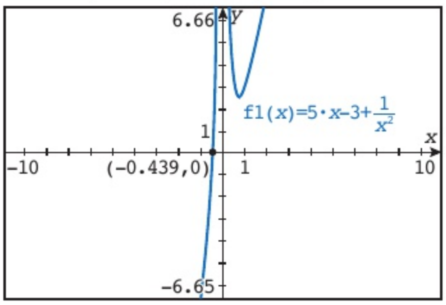
14. (a) Vertical asymptote occurs when denominator equals zero; therefore, $x^2 = 0 \Rightarrow x = 0$.

(b) $g'(x) = b - \frac{2}{x^3}$

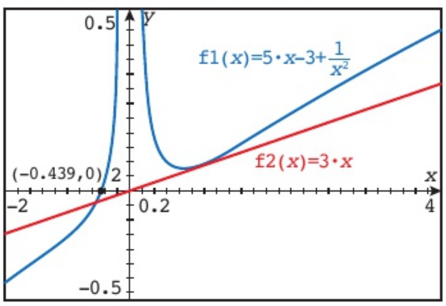
(c) $3 = b - \frac{2}{(1)^3} \Rightarrow b = 5$

(d) Point of tangency is at $g(1) = 3 \Rightarrow (1, 3)$, gradient is 3 (given), equation of T is $y - 3 = 3(x - 1) \Rightarrow y = 3x$.

(e) x-intercept is at $(-0.439, 0)$. Graph:



(f) **(i)** and **(ii)** shown on graph below:



(g) $(0.737, 2.53)$

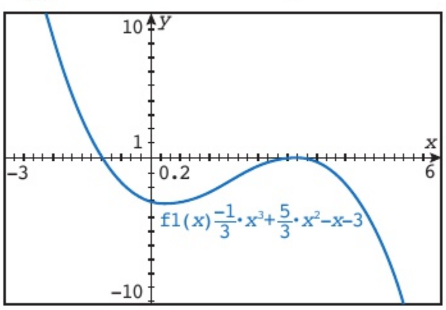
(h) $0.737 < x < 5$

15. (a) $\frac{dy}{dx} = 4x - 5$

(b) $L: 6x + 2y = -1 \Rightarrow y = -3x - \frac{1}{2} \Rightarrow$ Gradient of L is $m = -3$.

$\therefore \frac{dy}{dx} = 4x - 5 = -3 \Rightarrow x = \frac{1}{2}$.

16. (a)



(b) $f(-1) = -\frac{1}{3}(-1)^3 + \frac{5}{3}(-1)^2 - (-1) - 3 = 0$

(c) $(0, -3)$

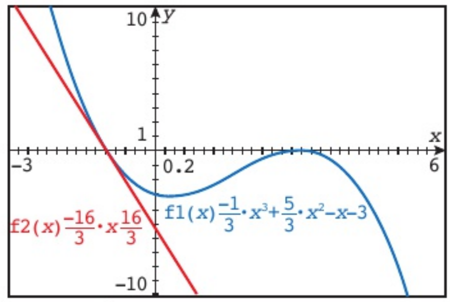
(d) $f'(x) = -x^2 + \frac{10}{3}x - 1$

(e) $f'(-1) = -(-1)^2 + \frac{10}{3}(-1) - 1 = -\frac{16}{3}$

(f) $f'(-1)$ gives the gradient of the tangent to the curve at the point where $x = -1$.

(g) $y - 0 = -\frac{16}{3}(x + 1) \Rightarrow y = -\frac{16}{3}x - \frac{16}{3}$

(h)



(i) **(i)** $a = \frac{1}{3}$ **(ii)** $b = 3$

(j) $f(x)$ is increasing on the interval $a < x < b$.

17. (a) $4(2x) + 4(x) + 4(y) = 48 \Rightarrow 3x + y = 12 \Rightarrow y = 12 - 3x$

(b) $V = lwh = (x)(2x)(y) = 2x^2y = 2x^2(12 - 3x) = 24x^2 - 6x^3$

(c) $\frac{dV}{dx} = 48x - 18x^2$

(d) $\frac{dV}{dx} = 48x - 18x^2 = 0 \Rightarrow x = 0, \frac{8}{3}$. $x = 0$ is nonsensical, so maximum volume occurs at $x = \frac{8}{3} = 2.67$ m.

(e) $V = 24\left(\frac{8}{3}\right)^2 - 6\left(\frac{8}{3}\right)^3 = 56.9\,\text{m}^3$.

(f) length $= 2x = 5.33$ m, height $= y = 12 - 3\left(\frac{8}{3}\right) = 4$ m

(g) $\text{SA} = 2 \times \frac{16}{3} \times 4 + 2 \times \frac{8}{3} \times 4 + 2 \times \frac{16}{3} \times \frac{8}{3} = 92.4\,\text{m}^2$.

Therefore $\frac{92.4}{15 \times 4} = 1.54 \Rightarrow 2$ tins are required.

18. (a) $y = 0.5x - 1$

(b) Equation is of the form $y = mx$ since it passes through the origin. Angle with x axis is $\frac{34.92°}{2} = 17.46°$.

Gradient $m = \tan(17.46°) = 0.315$; therefore equation is $y = 0.315x$.

(c) $(5.39, 1.70)$

(d) The throw is not valid, since the point Q is 5.65 m from the centre of the throwing circle, which implies that the discus will land outside the sector.

(e) **(i)** $R(0.795, -0.585)$

(ii) $y = 0.425x - 0.892$

19. (a) $\sqrt{2^2 + 6^2} = \sqrt{40} = 2\sqrt{10}$ km

(b) $r = \sqrt{(6 - w)^2 + 4}$

(c) $T = \frac{\sqrt{(6 - w)^2 + 4}}{3} + \frac{w}{5}$

(d) $w = 4.5$ km, $r = 2.5$ km

(e) 1.73 hours

20. (a) 3.94 min, 43.2 nanograms $\text{ml}^{-1}\,\text{min}^{-1}$

(b) 53.3 min

(c) 88.8 min, -4.36 nanograms $\text{ml}^{-1}\,\text{min}^{-1}$

(d) From 0.594 min to 12.3 min

21. (a) thousands of new subscribers per year

(b) When $t = 11.8 \Rightarrow$ late in 2011 at a rate of 7370 new subscribers per year.

(c) For $5.51 < t < 21.3 \Rightarrow$ mid-2005 to early 2021
(d) $S'(t)$ is positive for all $t \in \mathbb{R}$

22. (a) Radius = 32.6 cm, Height = 65.1 cm
(b) 217 litres

Chapter 11

Exercise 11.1

1. (a) $\frac{1}{2}x^2 + 2x + C$ (b) $\frac{1}{6}x^6 + C$
(c) $42t + C$ (d) $an + C$
(e) $x^5 + C$ (f) $6t^3 + C$
(g) $-x^3 + \frac{5}{2}x^2 - 8x + C$ (h) $-4.9t^2 + v_0t + C$
(i) $ax^2 + bx + C$ (j) $2x^4 + 9x^3 - 4x^2 - 6x + C$
(k) $\frac{7}{2}x^4 + \frac{5}{3}x^3 - \frac{1}{2}x + C$ (l) $-\frac{5}{h} + C$
(m) $-\frac{250}{w^2} - \frac{1}{w} = \frac{-250 - w}{w^2} + C$

2. (a) If we increase the power by one and multiply by the reciprocal of the new exponent, we obtain $\frac{1}{0}x^0$. Since division by zero is undefined, we cannot use this rule to find the antiderivative.
(b) We must first rewrite $g(x)$ as $g(x) = 400x^{-1} + x$. If we increase the power by one and multiply by the reciprocal of the new exponent, we obtain $400\left(\frac{1}{0}x^0\right) + \frac{1}{2}x^2$. Since division by zero is undefined, we cannot use this rule to find the antiderivative.

3. (a) $t^3 - t^2 + t + C$
(b) $-\frac{1}{14}x^4 + \frac{1}{3}x + C$
(c) $\frac{2}{3}t^3 + \frac{1}{2}t^2 - 3t + C$
(d) $\frac{4}{3}x^3 + 6x^2 + 9x + \frac{9}{2} + C$
(e) $\frac{35}{4}n^4 - \frac{110}{3}n^3 + 35n^2 + C$

4. (a) $f(x) = 2x^2 + 3x + 12$
(b) $y = -\frac{12}{x^2} + 5x + 19$
(c) $f(t) = t^3 - t^2 + t + 90$
(d) $w = \frac{175}{x^2} - \frac{1}{x} + \frac{3}{4}$
(e) $g(x) = \frac{500}{3}x^3 - 2750x^2 + 12000x - 2750$

5. (a) $s(t) = -40t^2 + 400$ (b) cm
(c) $s(1) = 360$ cm; $s(2) = 240$ cm
(d) $s(4) = -240$ cm; this means that the car has gone off the end of the ramp.
(e) $t = 3.16$ seconds.

6. (a) litres min^{-1} (b) $V = 0.01x^2 - 2.5x + 150$
(c) $t = 34.9$ minutes (d) $t = 100$ minutes

7. (a) $2028 \approx 2030$ m (b) $1539.5 \approx 1540$ m
(c) 43.7 s

8. (a) $a(t) = -7.5$ (b) 30 m s^{-1}
(c) $v(t) = -7.5t + 30$ (d) 4 s
(e) $s(t) = -3.75t^2 + 30t$ (f) 60 m

9. (a) monthly revenue in euros
(b) $R(n) = -0.01n^3 + 7.5n^2 + 64000$
(c) 500 wheels per month
(d) 689,000 euros

Exercise 11.2

1. (a) 40 (b) 15 (c) $\frac{24}{25}$ (d) $\frac{64}{3}$
(e) 0.810 (f) 12 (g) $\frac{\pi}{2} = 1.57$

2. 72 000 m^3

3. (a) US\$60,000 (b) US\$40,000
(c) US\$37,500 (d) US\$77,500

4. (a) $l\,s^{-1}$ (b) $0.0613\,l\,s^{-1}$
(c) Yes, $0.0613\,l\,s^{-1} = 3.68\,l\,min^{-1}$

5. (a) –110 ml (b) $k = 9.6$ ml h^{-1}

6. (a) (i) 2.97 kg (ii) 8.47 kg (iii) 0.390 kg
(b) (i) 3.47 kg

7. 27 m^3

8. (a) at 0.83 km
(b) $\int_0^{500} \left((n + 17)^{1.25} e^{-0.07(n+17)} - 0.001n + 3\right) dn$
(c) 1705 kg

Exercise 11.3

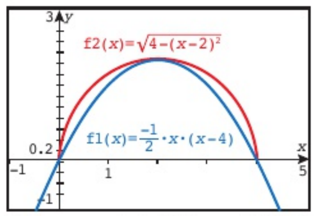
1. (a) 32 (b) $\frac{37}{3}$ (c) $\frac{2}{3}$

2. (a) 34.7 (b) $\frac{25}{2}$ (c) 20 (d) 10.1
(e) 7.67 (f) 3.17

3. (a) 36 (b) 10.7 (c) 6.75 (d) 9
(e) 1.60

4. (a) 6.67 (b) 10. 7 (c) 1.60

5. (a)



(b) 5.33 (c) 6.28 (d) 0.950

6. (a) 4 cm^2 (b) 11.1 cm^2 (c) 14.2 cm^2

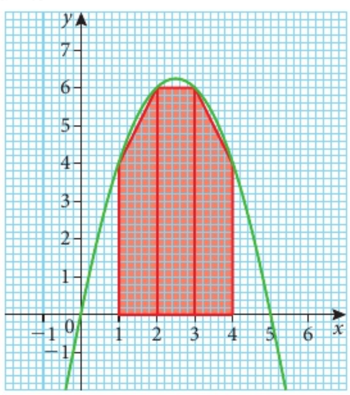
7. (a) (2, 2) (b) 18.7 m^2 (c) 112 m^3 (d) 61.3 m^2

8. (a) (i) $\int_{3.3}^{10} \left(\sqrt{6.7^2 + (x - 3.3)^2} - 6.7\right) dx$ (ii) 6.63 m^2
(b) 40.2 m^2 (c) 14 500 m^3

Exercise 11.4

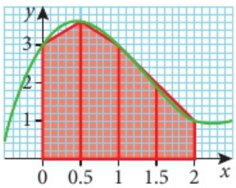
1. (a) 16 units^2
(b) 19 units^2
(c) 11 units^2

2. (a) (i)



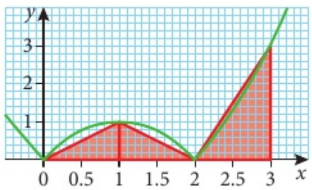
(ii) 16 units^2

(b) **(i)**



(ii) 5.25 units²

(c) **(i)**



(ii) 2.5 units²

3. **(a)** 1.5 hours **(b)** 8.26 nautical miles

4. 455 km

5. **(a)**

m s^{-1}	0	14	18.1	22.2	26.4	31	34.7	38.9	43.1	47.2	51.4	55.6
Time	0	4.3	6.4	9.6	14.3	21	32.1	48	71.9	107.6	161.1	241.1

(b) 2510 m

6. **(a)** 170 m²

(b) **(i)** 433 m² **(ii)** 9 tins

(c) 2550 m³

7. 990 m

8. **(a)** 7 units² **(b)** $\frac{3}{2}$ units² **(c)** π units² **(d)** 20 units²

(e) $\frac{57}{2} + \pi = 31.6$ units²

Chapter 11 Practice questions

1. **(a)** $x^2 + 8x + C$ **(b)** $\frac{1}{10}x^5 + C$

(c) $24p + C$ **(d)** $kx + C$

(e) $x^{10} + x^9 + C$ **(f)** $-7.5t^2 + vt + C$

(g) $\frac{2}{3}x^3 - 4x^2 + 3x + C$ **(h)** $t(x^2 + 1)^2 + C$

(i) $\frac{1}{2}mx^2 + bx + C$ **(j)** $\frac{1}{2}x^4 - \frac{5}{2}x^2 - \frac{2}{x} + C$

(k) $\frac{1}{4}x^4 + \frac{2}{3}x^3 - \frac{1}{3}x + C$ **(l)** $-\frac{5}{2n^2} + C$

(m) $-\frac{150}{h^2} + \frac{2}{h} = \frac{2h - 150}{h^2}$

2. **(a)** 36 **(b)** 4.67 **(c)** 0 **(d)** 40.5

(e) 1

3. **(a)** 18 units² **(b)** 9 units²

(c) $\frac{9\pi}{4}$ units² **(d)** 2 units²

(e) $29 + \frac{9\pi}{4} = 36.1$ units²

4. **(a)** 49.2 m², by modeling the end sections of the lake as triangles.

(b) 22.1 $l\,\text{h}^{-1}$

(c) **(i)** 147.6 m³ **(ii)** 147 600

(d) 278 days

5. **(a)** 19.0 units² **(b)** $\int_0^5 \sqrt{25 - x^2}\,\text{d}x$ **(c)** 19.635 units²

(d) **(i)** $\frac{1}{4}\pi r^2 = \frac{25\pi}{4}$ **(ii)** 19.635 units²

(e) The value from part **(c)** is accurate to at least 5 significant digits, while the value in part **(a)** is accurate only to 2 significant digits.

6. **(a)** $V = \int(3 - 0.1t)\,\text{d}t$ **(b)** 25 l **(c)** 20 s

(d) **(i)** 30 s **(ii)** 45 l

7. **(a)** 100 chairs

(b) **(i)** $P(n) = -\frac{3}{2}n^2 + 300n - 4600$

(ii) £9800 **(iii)** 17 chairs **(iv)** £10,400

8. **(a)** 86.7 m **(b)** $\int_0^3 (-4.76t + 28.7)\,\text{d}t$

(c) 64.7 m **(d)** $s(t) = -2.38t^2 + 28.7t + 28.4$

(e) 86.5 m

(f) The data is close to linear so the integral of the best-fit line provides a very good approximation.

9. **(a)** $P(-15, 0)$, $Q(15, 0)$ **(b)** $\int_{-15}^{15} (-(0.0783x)^{10} + 5)\,\text{d}x$

(c) 136 cm² **(d)** 273 cm² **(e)** 24 550 cm³

10. **(a)** $\int_0^k \sqrt{30^2 - x^2}\,\text{d}x$ **(b)** $\int_k^{30} \sqrt{30^2 - x^2}\,\text{d}x$

(c) $Q = 294$ cm², $R = 413$ cm² **(d)** $k = 12.12$ cm

Chapter 12

Exercise 12.1

1. **(a)** discrete **(b)** continuous **(c)** continuous

(d) discrete **(e)** continuous **(f)** continuous

(g) continuous **(h)** continuous **(i)** continuous

(j) discrete **(k)** continuous **(l)** continuous

(m) discrete

2. **(a)** The values in cents are 0 ,1 ,2, …

(b) Yes, because we can identify the first, second, etc.

(c) Yes, it is finite because students cannot earn an infinite amount of money.

(d) Technically, the variable is discrete.

3. **(a)** 0, 1, 2, …, 100

(b) Yes.

(c) Yes, there are 101 values.

(d) The variable is discrete because it is countable.

4. **(a)** $\text{P}(3 \leqslant X \leqslant 6) = \text{P}(3) + \text{P}(4) + \text{P}(5) + \text{P}(6) = 0.04 + 0.28 + 0.42 + 0.21 = 0.95$

(b) $\text{P}(X > 6) = \text{P}(X \geqslant 7) = \text{P}(7) + \text{P}(8) = 0.02 + 0.02 = 0.04$

(c) $\text{P}(X < 3) = \text{P}(X \leqslant 2) = \text{P}(0) + \text{P}(1) + \text{P}(2) = 0 + 0 + 0.01 = 0.01$

5. **(a)** 0.45

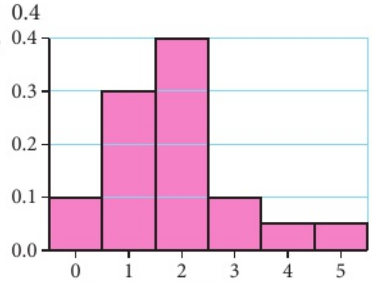
(b) $\mu = \text{E}(X) = \sum x\text{P}(x) = 0(.10) + 1(.20) + 2(.25) + 3(.25) + 4(.20) = 2.25$

$\sigma^2 = \text{V}(X) = \sum (x - \mu)^2\,\text{P}(x) = (0–2.25)^2(.10) + (1-2.25)^2(.20) + (2-2.25)^2(.25) + (3-2.25)^2(.13) + (4-2.25)^2(.20) = 1.59$

$\sigma = \sqrt{\sigma^2} = \sqrt{1.59} = 1.26$

6. **(a)** 0.4

(b)



(c) 1.85, 1.19 **(d)** **(e)** 2.85, 1.19

(f) $\text{E}(x + b) = \text{E}(x) + b$; $\text{V}(x + b) = \text{V}(x)$

7. **(a)** 0.26 **(b)** 0.37 **(c)** 0.77
(d) 16.29 **(e)** 8.126 **(f)** 4.125; 2.01325
(g) $E(ax + b) = aE(x) + b$; $V(ax + b) = a^2V(x)$

8. **(a)** 0.969 **(b)** 0.163 **(c)** 3.5
(d) $\int (x - 3.5)^2 \cdot P(x) = 1.048 \Rightarrow \sigma = \sqrt{1.048} \approx 1.02$

9. $k = \frac{1}{30}$

x	12	14	16	18
$P(X = x)$	$6k$	$7k$	$8k$	$9k$

10. **(a)** $k = \frac{1}{10}$ **(b)** $\frac{37}{60}$ **(c)** $\frac{19}{30}$
(d) $E(x) = 16$, SD$= 7$ **(e)** $E(y) = \frac{11}{5}$; $V(Y) = \frac{49}{25}$

11. $k = 0.667$, $E(x) = 5.44$

12. **(a)** $k = 0.3$ or 0.7
(b) for $k = 0.3$: $E(x) = 2.18$; for $k = 0.7$: $E(x) = 1.78$

13. **(a)**

y	0	1	2	3
$P(Y = y)$	$\frac{1}{27}$	$\frac{2}{9}$	$\frac{4}{9}$	$\frac{8}{27}$

(b) 2

14. **(a)**. See table below.

x	45	46	47	48	49	50	51	52	53	54	55
$P(x)$	0.05	0.13	0.25	0.40	0.65	0.85	0.90	0.94	0.97	0.99	1.00

(b) 0.85 **(c)** 0.15 **(d)** 48.87 **(e)** 2.057

15. **(a)** See table below

x	0	1	2	3	4	5	6
$P(x)$	0.08	0.23	0.45	0.72	0.92	0.97	1.00

(b) 0.72 **(c)** 0.97 **(d)** 2.63 **(e)** 1.44

16. **(a)** 0.90 **(b)** 0.09 **(c)** 0.009
(d) **(i)** unacceptable, **(ii)** acceptable
(e) $P(x) = (0.1^{x-1}) \times 0.9$

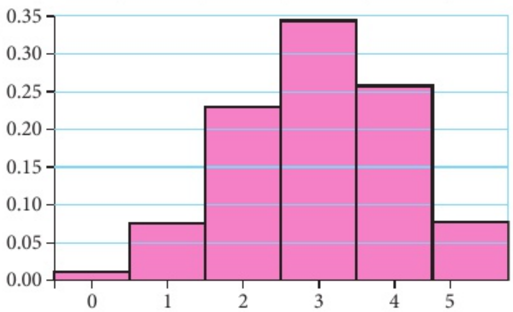
17. $n = 30$

Exercise 12.2

1. **(a)**

x	0	1	2	3	4	5
$P(X = x)$	0.01024	0.0768	0.2304	0.3456	0.2592	0.07776

(b)



(c) **(i)** Mean $= 3$, SD$= 1.095$
(ii) Mean $= 3$, SD$= 1.095$

2. **(a)** 0.001294494 **(b)** 0.000000011
(c) 0.99999999 **(d)** 0.99999966
(e) mean $= 12$, SD $= 2.19$

3. **(a)**

k	0	1	2	3	4	5	6
$P(x \leqslant k)$	0.11765	0.42017	0.74431	0.92953	0.98907	0.99927	1

(b)

Number of successes x	List the values of x	Write the probability statement	Explain it, if needed	Find the required probability
At most 3	0,1,2,3	$P(x \leqslant 3)$	$P(x \leqslant 3)$	0.92953
At least 3	3,4,5,6	$P(x \geqslant 3)$	$1 - P(x \leqslant 2)$	0.25569
More than 3	4, 5, 6	$P(x > 3)$	$1 - P(x \leqslant 3)$	0.07047
Fewer than 3	0,1,2	$P(x \leqslant 2)$	$P(x \leqslant 2)$	0.74431
Between 3 and 5 (inclusive)	3,4,5	$P(3 \leqslant x \leqslant 5)$	$P(x \leqslant 5) - P(x \leqslant 2)$	0.25496
Exactly 3	3	$P(x = 3)$	$P(x = 3)$	0.18522

4. **(a)**

k	0	1	2	3	4	5	6	7
$P(x \leqslant k)$	0.02799	0.15863	0.41990	0.71021	0.90374	0.98116	0.99836	1

(b)

Number of successes x	List the values of x	Write the probability statement	Explain it, if needed	Find the required probability
At most 3	0,1,2,3	$P(x \leqslant 3)$	$P(x \leqslant 3)$	0.71021
At least 3	3,4,5,6,7	$P(x \geqslant 3)$	$1 - P(x \leqslant 2)$	0.58010
More than 3	4, 5, 6,7	$P(x > 3)$	$1 - P(x \leqslant 3)$	0.28979
Fewer than 3	0,1,2	$P(x \leqslant 2)$	$P(x \leqslant 2)$	0.41990
Between 3 and 5 (inclusive)	3,4,5	$P(3 \leqslant x \leqslant 5)$	$P(x \leqslant 5) - P(x \leqslant 2)$	0.56126
Exactly 3	3	$P(x = 3)$	$P(x = 3)$	0.290304

5. **(a)** p is not constant, trials are not independent.
(b) p becomes constant.
(c) $n = 3$, $p = \frac{5}{8}$

y	0	1	2	3
$P(Y = y)$	0.05273	0.263672	0.439453	0.244141

(d) 0.75586 **(e)** 1.875 **(f)** 0.703125 **(g)** 0.6836

6. **(a)** 0.107374 **(b)** 0.99363 **(c)** 0.89263 **(d)** 2

7. **(a)** 0.817073 **(b)** 1 **(c)** 0.0161776

8. **(a)** 0.033833 **(b)** 0.024486 **(c)** 0.782722

9. **(a)** 0.75 **(b)** 0.0325112 **(c)** 0.172678

10. **(a)** **(i)** 0.0431745 **(ii)** 0.997614 **(iii)** 0.0112531
(iv) 0.130567 **(v)** 0.956826
(b) **(i)** 10 **(ii)** 3.13
(c) 4, 16.

11. **(a)** 3 **(b)** 0.101308 **(c)** 0.000214925

12. **(a)**

x	0	1	2	3	4	5
$P(x)$	0.03125	0.15625	0.31250	0.31250	0.15625	0.03125

(b) 0.03125 **(c)** 0.03125 **(d)** 0.96875 **(e)** 0.96875

(f) **a**

x	0	1	2	3	4	5
P(x)	0.32768	0.40960	0.20480	0.05120	0.00640	0.00032

b 0.32768 **c** 0.00032 **d** 0.67232 **e** 0.99968

13. **(a)** 0.1382 **(b)** 0.144

14. 0.91296

15. **(a)** 0.107 **(b)** 0.893 **(c)** $n = 14$

Exercise 12.3

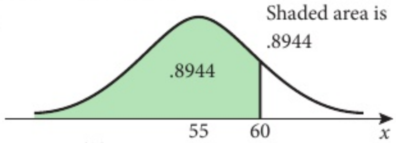
(some answers are rounded)

1. **(a)** 0.5 **(b)** 0.499571 **(c)** 0.158655
(d) 0.682690 **(e)** 0.022750 **(f)** 0

2. **(a)** 0.656947 **(b)** 0.999944

3. **(a)** 0.008634 **(b)** 0.982732

4. 0.**8944**,



5. **(a)** 0.2292 **(b)** 9.12%

6. **(a)** 0.066807 **(b)** 0.68269 **(c)** 678.16 **(d)** 134.898

7. **(a)** 1.8% **(b)** 1.13%

8. **(a)** 0.9696 **(b)** 0.546746

9. **(a)** 1 day **(b)** 29 days **(c)** 112 days

10. 1.22%

11. 52.71, 59.29

12. 68.16, 75.84

13. 65.28, 78.72

14. **(a)** **(i)** 2% **(ii)** 5%
(b) **(i)** 0.0524 **(ii)** 0.1875

15. **(a)** 0.655422 **(b)** 0.008198 **(c)** 82 bottles

16. **(a)** 22.7% **(b)** 0.55% **(c)** 29.678 **(d)** 229.2

17. **(a)** Not likely: chance is 0.14%
(b) 0.1587 **(c)** 0.6827 **(d)** 5396 **(e)** 43785

18. 0.073

19. 236.94

20. 198.4, 233.6

21. **(a)** 83% **(b)** 54.2%

Chapter 12 Practice questions

1. **(a)** 34.5% **(b)** 0.416 **(c)** \$3325

2. **(a)** 0.835 **(b)** 1010 **(c)** 6.58

3. **(a)** $\frac{35}{128}$ **(b)** $\frac{7}{32}$ **(c)** $\frac{91}{128}$

4. **(a)** **(i)** 0.675 **(ii)** 0.428
(b) $t = 62.6$

5. **(a)** 70.1% **(b)** 0.00226

6. **(a)** **(i)** 0.345 **(ii)** 0.115 **(iii)** 0.540
(b) 0.119
(c) 737

7. **(a)** 15.9% **(b)** 227

8. **(a)** 0.0912 **(b)** a = 251, b = 369.

9. **(a)** **(i)** 0.841 **(ii)** 0.533
(b) **(i)** **(ii)** 0.647

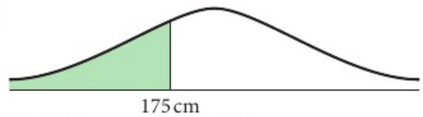
10. **(a)** 2 **(b)** 0.182 **(c)** 0.597

11. **(a)** 0.129886 **(b)** 0.676714 **(c)** 2

12. **(a)** $\frac{1}{5}$ **(b)** $\frac{7}{5}$

13. **(a)** **(i)** 0.217% **(ii)** 0.00045%
(b) 84.13%

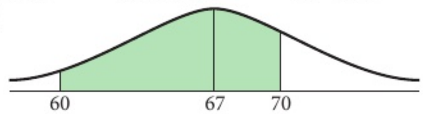
14. **(a)**



(b) **(i)** 185 **(ii)** 0.159

15. **(a)** 5800 **(b)** 0.420 **(c)** 4700

16. **(a)**



(b) **(i)** 0.980 **(ii)** 0.811 **(iii)** 0.792
(c) 70
(d) **(i)** 228 **(ii)** 1200

Chapter 13

Exercise 13.1

1. **(a)** p-value $= 0.269 > 0.05$, do not reject the null hypothesis.
(b) p-value $= 0.670 > 0.05$, do not reject the null hypothesis.
(c) p-value $= 0.159 > 0.10$, do not reject the null hypothesis.

2. **(a)** $H_0: \mu_A - \mu_B = 0$; $H_1: \mu_A - \mu_B \neq 0$
(b) Data are normally distributed and the distributions have equal variances.
If p-value < 0.02, reject H_0.
(c) p-value $= 0.0114 < 0.02$. Reject H_0. The two repellents are different.

3. **(a)** $H_0: \mu_{Guard} - \mu_{Cameras} = 0$; $H_1: \mu_{Gaurd} - \mu_{Cameras} < 0$
(b) Data are normally distributed and the distributions have equal variances.
If p-value < 0.10, reject H_0.
(c) p-value $= 0.338 > 0.10$. Do not reject H_0. There is no evidence that the guard is better.

4. **(a)** $H_0: \mu_{Old} - \mu_{New} = 0$; $H_1: \mu_{Old} - \mu_{New} > 0$
(b) Data are normally distributed and the distributions have equal variances.
If p-value < 0.10, reject H_0.
(c) p-value $= 0.00223 < 0.10$. Reject H_0. There is evidence that average age is lower now.

Exercise 13.2

1. **(a)** The (preferred) swimming style is independent of gender
(b) 3
(c) $\chi^2_{calc} = 16.4$
(d) Reject the Null Hypothesis, $16.4 > 7.815$, or p-value of $9.26148 \times 10^{-4} < 0.05$.

2. **(a)** The (crop) yield is **independent** of the (type of) fertiliser used.
(b) 4
(c) 13.277
(d) $\frac{50 \times 40}{120} \approx 17$
(e) **(i)** $\chi^2_{calc} = 3.86$ **(ii)** p-value $= 0.425$
(f) Since p-value > 0.01, do not reject the null hypothesis.

3. **(a)** H_0: Users' preferences are independent of age
H_1: Users' preferences are not independent of age
(b) 428.53
(c) 9
(d) If $\chi^2_{calc} > 16.92$, reject the null hypothesis.
(e) $\chi^2_{calc} = 18.23 > 16.92$. Reject the null hypothesis.
(f) p-value $= 0.0325 < 0.05$. Reject the null hypothesis.

4. **(a)** H_1: The data does not fit the proposed model.
   (b) 30, 60, 90, 60, 60
   (c) p-value $= 0.339 > 0.05$. Fail to reject the null hypothesis. Data may fit the model.
5. **(a)** H_0: reports of side effects are not different among the two groups
   H_1: reports of side effects are different among the two groups
   (b) 12.5, 17.5, 10, 210
   (c) This is a GOF test. p-value $= 0.0475 < 0.05$. Reject the null hypothesis. There is evidence that the two groups differ.
6. **(a)** H_0: The collected data fits a fair die distribution.
   H_1: The collected data does not fit a fair die distribution.
   (b) 100 in each cell
   (c) Either $\chi^2_{calc} = 5.7 < 11.07$, or p-value $= 0.337 > 0.05$. We fail to reject the null hypothesis. The die appears to be fair.

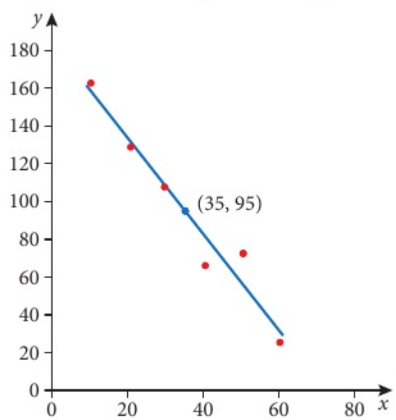
Chapter 13 Practice questions

1. **(a)** H_0: the type of Latin dance the viewer prefers is independent of their age.
   (b) 18
   (c) $p = 0.0876$
   (d) p-value > 0.05. The producer's claim may be justified.
2. **(a)** Conservatives 435, Liberals 615, Greens 225, rightists 225
   (b) H_0: voter support has not changed since election.
   H_1: voter support has changed since election.
   (c) GOF test with df $= 3$. P-value $= 0.009 \leqslant 0.10$. We reject the null hypothesis and conclude that voters support has changed.
3. **(a)** **(i)** H_0: age and opinion (about the reduction) are independent.
   (ii) H_1: age and opinion (about the reduction) are not independent.
   (b) 2
   (c) $\dfrac{80 \times 35}{130} = 21.5$
   (d) **(i)** 10.3
   (ii) 0.00573
   (e) Since p-value < 0.01, reject H_0, or χ^2statistic $> \chi^2$ critical, reject H_0.
4. This is a GOF test with 4 df. p-value $= 0.0230 < 0.10$. Reject H_0. Absences differ from one day to the other.
5. **(a)** This is a one tail t test of difference of means. The populations are approximately normal and with equal variances.
   (b) $H_0 : \mu_{device} = \mu_{usual;} H_1 : \mu_{device} > \mu.$
   (c) p-value $= 0.0213 < 0.10$. We reject the null hypothesis and conclude that users pay more with the device.
6. **(a)** Age and preferred destination are independent
   (b) $(4 - 1) \times (5 - 1) = 12.$
   (c) $\chi^2 > 21.026$
   (d) $\dfrac{285 \times 420}{1200} = 99.8$
7. **(a)** 27.9 **(b)** 0.0321
   (c) Reject the null hypothesis, TV show preference is not independent of gender.
8. This is a two-tail t test of difference of means. The populations are approximately normal and with equal variances.
   $H_0 : \mu_{violent} = \mu_{neutral}; H_1 : \mu_{violent} \neq \mu.$
   p-value $= 0.0000915 < 0.10$. We reject the null hypothesis and conclude that content of program affect viewers memory.

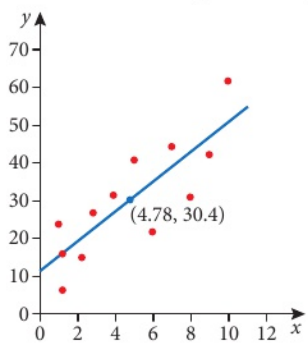
Chapter 14

Exercise 14.1

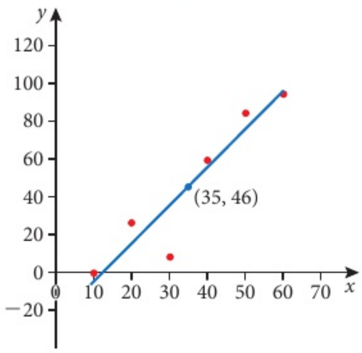
1. $(\bar{x}, \bar{y})$
2. form, direction, strength, unusual features
3. **(a)** Precipitation is the explanatory variable; autism prevalence rate is the response variable
   (b) No; no matter how strong the association is, it does not prove that precipitation causes autism, only that they are associated.
4. **(a)** No association.
   (b) Strong nonlinear association, possibly quadratic.
   (c) Nearly perfect negative linear association.
   (d) Strong positive linear association.
   (e) Moderate negative linear association.
   (f) No association.
   (g) Strong positive nonlinear association, possibly exponential.
   (h) Mostly strong positive linear association, but a cluster of outliers is a departure from the major pattern.
   (i) Very strong negative linear association with one outlier.
5. **(a)** The best fit line is approximately $y = -2.6x + 185$



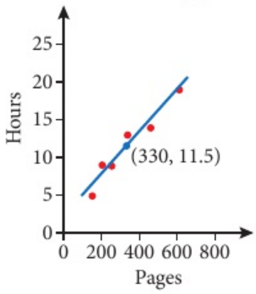
   (b) The best fit line is approximately $y = 4x + 12$



(c) The best fit line is approximately $y = 2x - 24$

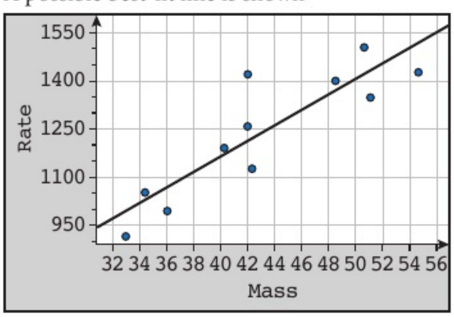


6. (a) Number of pages is the explanatory variable, hours to finish is the response variable.
(b) Scatter diagram is shown. The association is strong, positive, and approximately linear.
(c) The best fit line is approximately $t = 0.03n + 2.22$



(d) The gradient of 0.03 indicates that each additional page adds about 0.03 hours to the reading time, or an additional 100 pages adds 3 hours. The t-intercept of 2.22 indicates that a book with zero pages will take 2.22 hours to read: this indicates the extra time it takes to get the book and find the last page and any other task that is not related to the number of pages.
(e) About 16.3 hours.

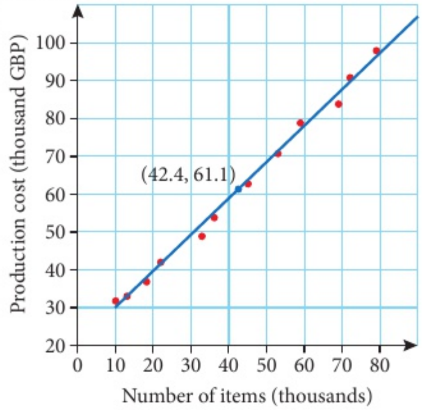
7. (a) The explanatory variable is mass. Scatter diagram shown.
(b) $\overline{M} = 43.2$, $\overline{R} = 1240$
(c) A possible best-fit line is shown



(d) There appears to be a strong, positive, linear correlation. The gradient of the line is about 24, which indicates that metaboloic rate increases by about 24 units for each additional kg in mass.
(e) About 1160.
(f) Kevin's mass is beyond the range of the observed data; we cannot predict his metabolic rate as it would be an extrapolation.

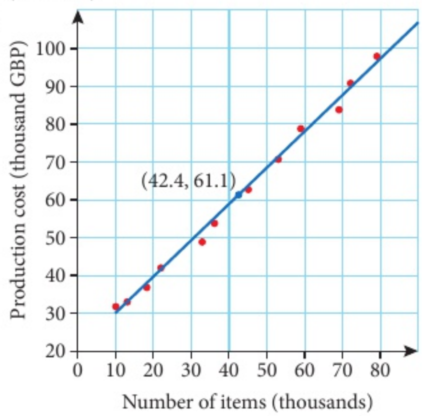
8. (a) and (c) are correct.

9. (a) Number of items produced is the explanatory variable; production cost is the response variable.
(b) Scatter diagram shown.
(c) Strong, positive, approximately linear, no outliers visible.
(d) (42.4, 61.1)
(e)



(f) 88 thousand GBP
(g) 1.0 (approximate); best fit is 0.966. This suggests that that the production cost increases by about 1000 GBP for each additional 1000 items produced.
(h) 20 (approximate); this could be interpreted as the fixed costs to keep the factory operating. This is an extrapolation.
(i) Predicting production cost for 100 thousand items would be an extrapolation.

10. (a) Speed is the explanatory variable; temperature is the response variable.
(b) The association is strong, positive, and approximately linear.
(c) (55, 75.9)
(d)

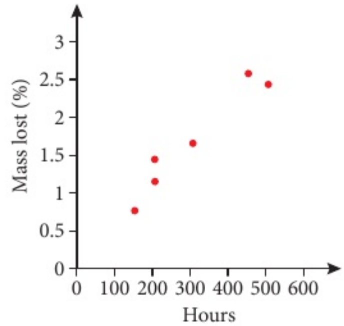


(e) About 80 km h^{-1}
(f) About 0.9 (best fit is 0.882). This indicates that tire temperatures increase by about 0.9 °C for each additional km h^{-1}.
(g) About 27 (best fit is 27.4). This suggests that the inital tire temperature before being driven was about 27 °C. This is an extrapolation.

(h) A prediction from a speed of 150 km h^{-1} would be an extrapolation.

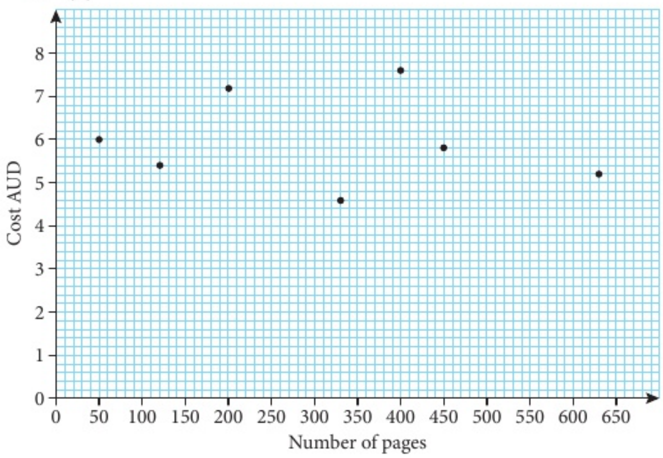
Exercise 14.2

1. Spearman's r_s measures the strength of a **monotonic** (continually increasing or decreasing) relationship. It is not sensitive to outliers. Pearson's r measures the strength of a **linear** relationship, and is sensitive to outliers.
2. **(a)** $r = -0.40$ **(b)** $r = -0.99$
   (c) $r = 0.90$ **(d)** $r = -0.03$
   (e) $r = 0.51$ **(f)** $r = -0.58$
   (g) $r = -0.95$ **(h)** $r = 0.74$
3. **(a)** $r = 0.967$



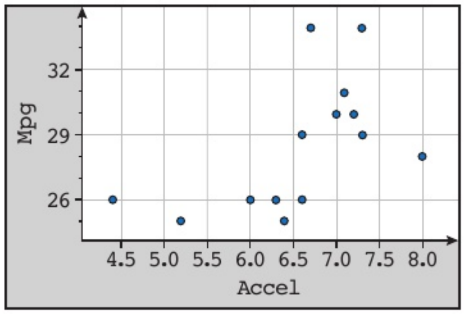
(b) There is a strong positive correlation between hours in the acid bath and mass of metal lost.

4. **(a)**



(b) $r = -0.141$
(c) No, the linear correlation is too weak for reliable predictions, as seen on the scatter plot and shown by the value of r.

5. **(a)** $r = 0.561$



(b) There is a moderate positive correlation between acceleration and fuel efficiency; as cars get slower the efficiency appears to increase. It appears to be only somewhat linear, which is supported by the value of r.

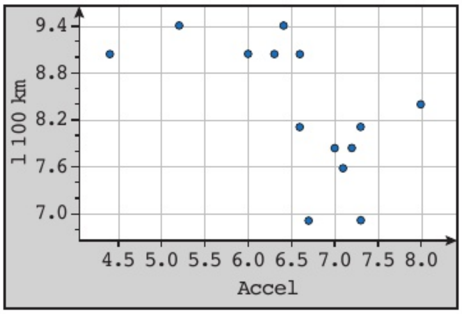
(c)

Cars	time to 60 mph rank	MPG rank
Mazda MX-5 Miata Club	8	1.5
Honda Civic Si	2	1.5
Fiat 124 Spider Lusso	6	3
Mini Cooper S	7	5
Subaru BRZ Premium	4	5
Toyota 86	4	5
Volkswagen GTI Autobahn	9	7.5
Ford Fiesta ST	2	7.5
Fiat 500 Abarth	1	9
Porsche 718 Boxster (base)	15	11.5
Subaru Impreza WRX Premium	13	11.5
Audi TT 2.0T (AT)	12	11.5
Ford Focus ST	9	11.5
BMW M235i	14	14.5
Ford Mustang Premium (2.3T, AT)	11	14.5

$r_s = 0.670$; in general there is a moderate positive rank correlation between acceleration and fuel efficiency in MPG; as acceleration times increase MPG increases as well. Cars that accelerate slower are more efficient in general.

(d) Person's r measures the strength of a linear correlation. Since this data appears to be somewhat nonlinear, Spearman's r_s is a more appropriate measure since it measures the strength of the monotonic association.

6. **(a)** $r = -0.598$



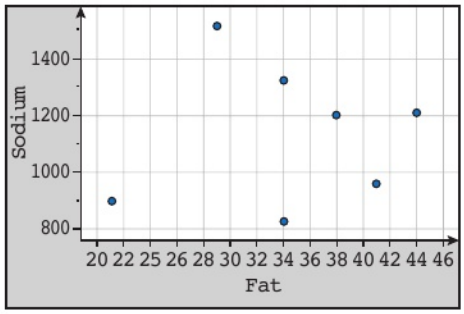
(b) There is a moderate negative linear correlation. As cars get slower, the L per 100km required decreases. The linear correlation is slightly stronger than in the previous exercise. By converting to L per 100km, we had to find the reciprocal of the units which may have made the association slightly more linear.

(c) The ranked data is shown below.

Cars	time to 60mph rank	L per 100 km rank
Mazda MX-5 Miata Club	8	14.5
Honda Civic Si	2	14.5
Fiat 124 Spider Lusso	6	13
Mini Cooper S	7	11
Subaru BRZ Premium	4	11
Toyota 86	4	11
Volkswagen GTI Autobahn	9	8.5
Ford Fiesta ST	2	8.5
Fiat 500 Abarth	1	7
Porsche 718 Boxster (base)	15	4.5
Subaru Impreza WRX Premium	13	4.5
Audi TT 2.0T (AT)	12	4.5
Ford Focus ST	9	4.5
BMW M235i	14	1.5
Ford Mustang Premium (2.3T, AT)	11	1.5

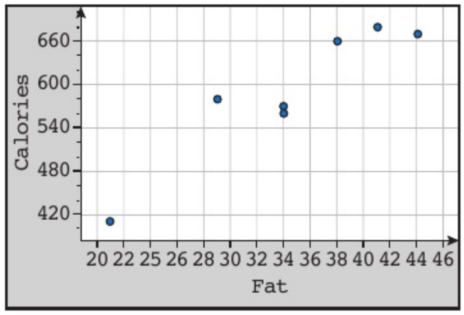
(d) $r_s = -0.670$; in general there is a moderate negative rank correlation between acceleration and fuel efficiency in L per 100 km; as acceleration times increase, L per 100 km decreases. Cars that accelerate slower are more efficient in general. The value of r_s has the opposite sign from the previous exercises as the direction of correlation has reversed.

7. **(a)** $r = 0.0852$



(b) There is almost no association between fat and sodium for these menu items.

8. **(a)** $r = 0.940$

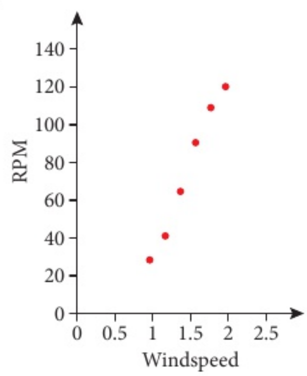


(b) There is a very strong positive linear correlation between fat and calorie content for these menu items.

9. **(a)** While a has the strongest linear correlation, we should not make any conclusions without looking at a scatter diagram. It could be that the design of the experiment causes outliers in variable a which are causing the relatively strong linear correlation.

(b) Without knowing more about the nature and design of the experiment, we cannot conclude that changes in a **cause** changes in S. However, if Ian deliberately manipulated the values of a and measured S, we may have strong evidence for causality (we would also need to suspect an underlying relationship between the two variables).

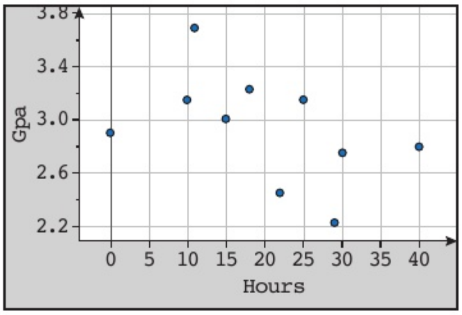
10. **(a)** $r = 0.993$



(b) There is a strong positive linear correlation between wind speed and RPM.

(c) The scatter diagram appears to have a slight curve to it and it is reasonable to suspect that the force applied by wind may be nonlinear. A logistic or power model may be more appropriate.

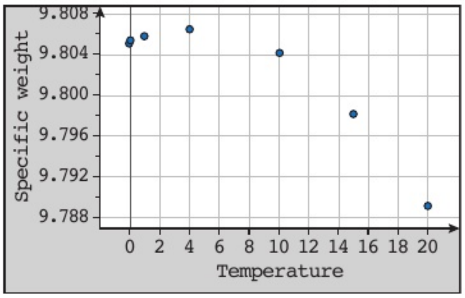
11. (a) The association appears moderate, negative, and approximately linear.



(b) $r_s = -0.533$; there is a moderate negative rank correlation between GPA and hours worked.

(c) $r = -0.461$; since the form of the association is approximately linear with no strong outliers, r is relatively close to r_s.

12. (a) The association is generally negative, but is nonlinear with a possible quadratic form.



(b) $r_s = -0.643$; there is a moderate negative rank correlation between temperature and the specific weight of water.

(c) The form does not appear monotonic.

Exercise 14.3

1. (a) $L = 0.00475h + 0.242$. On average, each additional hour increases mass loss by 0.00475%. The L intercept of 0.242 is not meaningful in this context (it is an extrapolation, and it suggests a mass loss of 0.242% for 0 hours).

(b) $k = 0.475\%$ (c) 2.14%

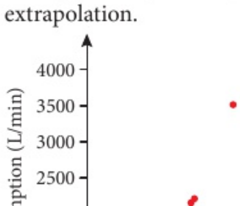
(d) Although the model appears quite good, the LSRL we generated should only be used to predict mass loss from hours in the acid bath. Using the model 'in reverse' lowers our confidence in the prediction significantly.

2. (a) $E = 1.85t + 16.4$. On average, for each additional second in 0–60 mph time, fuel efficiency increases my 1.85 miles gal^{-1}. The E-intercept of 16.4 is not meaningful in this context; it is an extrapolation and a car would have to have instantaneous acceleration to 60 mph!

(b) 26.9 miles gal^{-1}. Since the correlation appears moderate (not strong), there may be other factors.

(c) 9 seconds is beyond the range of the observed value of the explanatory variable; any prediction would be an extrapolation.

3. (a)



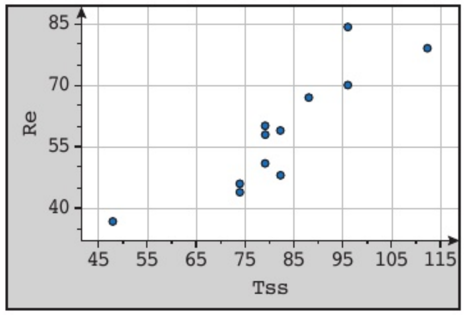
The association between the number of seats and fuel consumption is strong, positive, and approximately linear. There is one possible outlier but it appears to fit the general trend.

(b) $F = 0.462n + 3.65$. The fuel consumption increases by 0.462 L min^{-1} for each additional seat. The F-intercept of 3.95 could indcate fuel use beyond what is used to lift the weight of the seats.

(c) 166 L min^{-1}

(d) The data appears to have a slight nonlinear form in the scatter diagram. It would make sense that fuel consumption would increase in a faster-than-linear rate as the limits of current technology are reached, so a linear model may not be the most appropriate.

4. (a) The scatter diagram shows a strong, positive, linear association. There is a possible outlier at (48, 37).



(b) $E = 0.815(S) - 8.62$. For each additional unit increase in Training Stress Score, Relative Effort increases by 0.815 units. The E-intercept of -8.62 is not meaningful in this context.

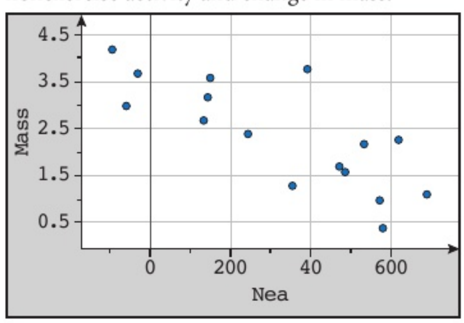
(c) 40.3

(d) $48 \leqslant S \leqslant 112$

(e) Since the low outlier is the minimum value in the domain, if we remove the outlier we would need to adjust the domain to $74 \leqslant S \leqslant 112$ and a prediction based on a Relative Effort of 60 would then become an extrapolation.

(f) Without (48,37), the LSRL is $E = 0.996(S) - 24.7$ with $R^2 = 0.756$. The LSRL has changed significantly.

5. (a) The scatter diagram shows a strong negative approximately linear association between change in nonexercise activity and change in mass.



(b) $M = -0.00344N + 3.51$. On average, for each additional 100 calories in change of NEA, change in mass decreases by 0.344 kg. The M-intercept suggests that when NEA does not change, mass will increase by 3.51 kg.

(c) 2.822 kg

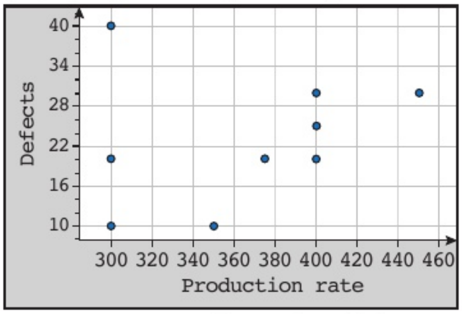
(d) A negative change in NEA suggests that the individual did less nonexercise activity after being overfed.

(e) $-94 \leqslant N \leqslant 690$

(f) The change in mass for $N = 1100$ would be negative, suggesting that the person was able to lose weight by eating more!

6. **(a)** Production Rate is the explanatory variable; Defects is the response variable.

(b) There appears to be a weak positive association between production rate and defects. However, an outlier at (300,40) may be affecting the strength of the association.

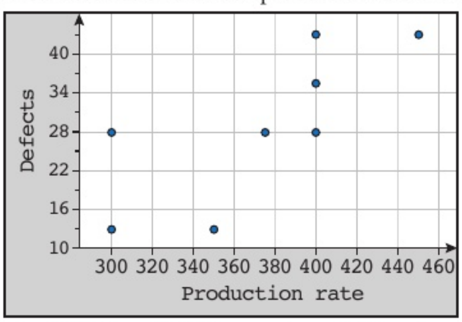


(c) $D = 0.0479R + 5.67$; the number of defects increases by about 5 for each 100 unit increase in production rate. The D-intercept of 5.67 is not meaningful in this context.

(d) $r = 0.295$; there is a weak positive linear correlation.

(e) The low value of r suggests that there are likely to be other factors determining the variation in the number of defects. It is better to try to find other factors instead.

(f) There appears to be a moderate positive approximately linear association between production rate and defects.



$D = 0.113R - 21.3$; The number of defects increases by about 11 for each 100 unit increase in production rate. The D-intercept of -21.3 is not meaningful in this context.

$r = 0.794$; there is a moderately strong positive linear correlation between defect and production rate.

The production rate appears to be correlated to the number of defects. Lower production rates should be attempted and further data can then be recorded.

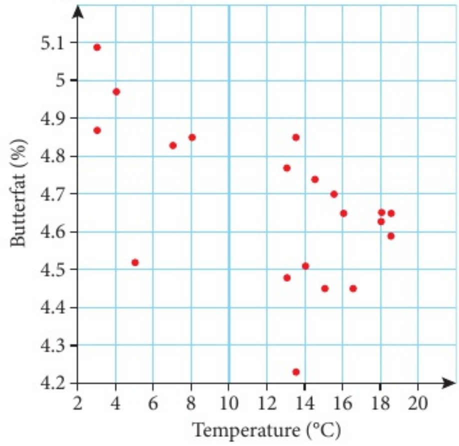
(g) It is OK to remove the outlier to investigate how influential it is. However, we must be careful to not put too much confidence in our predictions and recommendations using the revised data. Instead, we must investigate the outlier and see what caused it: was it a particularly unskilled worker? Mis-entered data? A power failure during production or other failure? Something else? If there are no unusual causes for the outlier, we should not remove it in our final analysis.

7. **(a)** There is a strong positive association between year and millions of active monthly users. The form appears nonlinear or piecewise linear.

(b) Suitable domains would be $0 \leqslant y \leqslant 5$ and $5 \leqslant y \leqslant 8$

8. **(a)** Temperature is the explanatory variable, butterfat is the response variable.

(b) The association appears to be negative and moderate and approximately linear.



There is a gap in the data; why are there no data values for temperatures in the interval $8 < T < 13$? We should proceed with caution.

(c) The LSRL is $F = -0.0216T + 4.94$. The butterfat content decreases by 0.02% for each increase of 1 °C. The F-intercept of 4.94 indicates that at a temperature of 0 °C, we predict that the butterfat content would be 4.94% – beware, this is an extrapolation.

(d) $r = -0.564$; there is a moderate negative linear correlation between butterfat and temperature.

(e) Correlation is not causation: while temperature and butterfat may have a moderate correlation, we cannot claim that lowering the temperature will cause increased butterfat content. Further research is needed; an investment in a climate-controlled barn may be an expensive experiment.

Chapter 14 Practice questions

1. **(a)** The value of r in the interval $-1 \leqslant r \leqslant 1$,

(b) If the association is negative, the value of r must be negative.

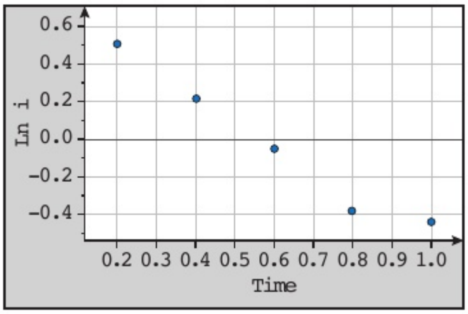
(c) If life expectancy increases as body mass increases, then r_s must be positive.

(d) The gradient in the LSRL is negative, but r is positive.

2. **(a)** 0 **(b)** -1 **(c)** $+1$ **(d)** $+1$

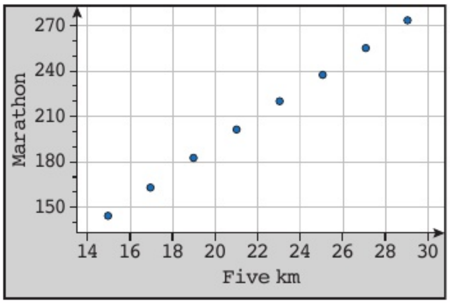
(e) 0 **(f)** -1

3. **(a)** The scatter diagram shows an strong, negative, approximately linear association, with $r = -0.984$.



(b) $\ln(I) = -1.25t + 0.719$

4. **(a)** The scatter diagram shows a nearly perfect, positive linear correlation with $r = 1.00$



(b) $M = 9.30T + 5.20$. For each additional minute in 5 km time, we estimate an additional 9.3 minutes in marathon time. The M intercept of 5.20 is not meaningful in this context.

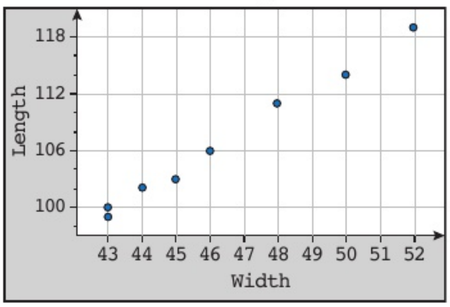
(c) 191 minutes

(d) **(i)** The slope gradient be $\dfrac{42.195}{5} = 8.44\ \text{min min}^{-1}$

(ii) The model gradient from part **(b)** is 10% more than this theoretical gradient.

(iii) We see that, on average, runners' pace is about 10% slower on a marathon than on a 5 km run.

5. **(a)** There is a very strong, positive, linear correlation between leaf width and length, with $r = 0.997$.



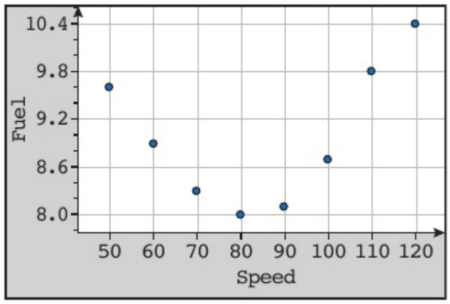
(b) $L = 2.15W + 6.85$. For each additional mm in width, we estimate an additional 2.15 mm in length. The L intercept of 6.85 is not meaningful in this context.

(c) $43 \leqslant W \leqslant 52$

(d) 108 mm

(e) 60 mm is outside the domain for our model and it would be an extrapolation.

6. **(a)** There appears to be a strong relationship, perhaps quadratic, it is initially negative and then positive.

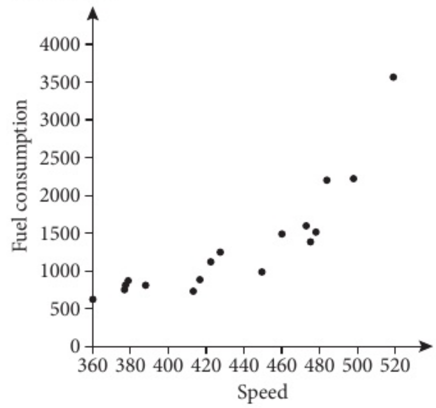


(b) The data does not appear to have a linear relationship.

(c) The data does not appear to be monotonic.

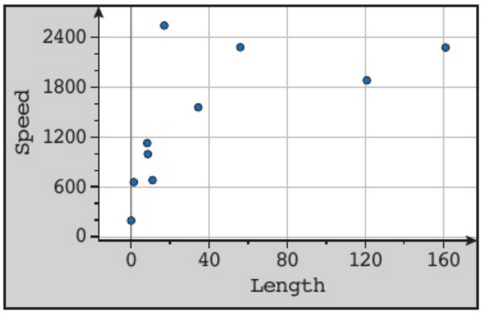
7. No, the study was observational and so did not control alcohol intake and then observe responses. There may be other factors that are linked with alcohol intake that reduce the risk of cardiovascular disease.

8. **(a)** There appears to be a strong, positive, nonlinear correlation.



(b) Spearman's r_s is a better measure since the data is nonlinear but approximately monotonic. Spearman's $r_s = 0.950$; there is a very strong positive correlation between speed and fuel consumption ranks.

9. **(a)** There is a positive association. It appears somewhat nonlinear. There appears to be an outlier at (17, 2550) – the common swift.



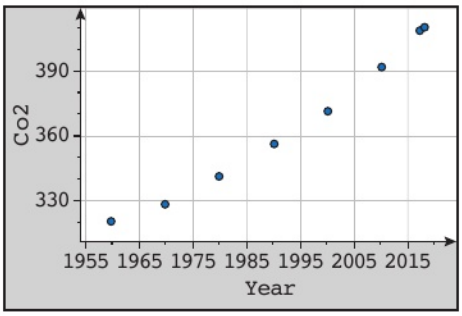
(b) There is one outlier (common swift). Assuming the data is correct, we cannot justify removing this data point simply because it doesn't 'fit.'

(c) $r = 0.620$ There is a moderate positive linear correlation between body length and flying speed.

(d) $r_s = 0.820$ There is a strong positive rank correlation between body length and flying speed.

(e) The data appears to have a slight nonlinear curve, so Spearman's r_s is more appropriate. However, the data is approximately linear so tis is also acceptable to measure the strength of the linear association with Pearson's r.

10. (a) There appears to be a strong, positive association, approximately linear but with some evidence of an nonlinear association, with no outliers.



(b) $r = 0.991$, there is a strong positive linear correlation between year and CO_2 concentration.

(c)

Year	0	10	20	30	40	50	57	58
CO_2 concentration (ppm)	320	328	341	356	372	392	409	411

(i) $r = 0.991$, there is a strong positive linear correlation between years since 1960 and CO_2 concentration. The correlation coefficient has not changed by subtracting 1960 from each date value.

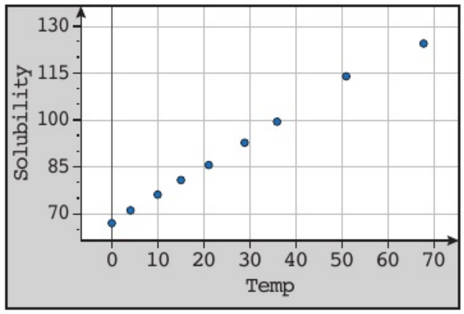
(ii) $C = 1.62y + 313$. The slope of 1.62 ppm year^{-1} indicates that CO_2 concentration is increasing by 1.62 ppm for each additional year. The intercept of 313 indicates the ppredicted CO_2 concentration in 1960.

(iii) Years from 1960 to 2018

(iv) 386 ppm

(v) The year 2050 is beyond the range of the observed data; it would be an extrapolation.

11. (a) There appears to be a nearly perfect positive linear association between temperature and solubility of $NaNO_3$.



(b) $r = 0.999$, There is a nearly perfect positive linear association between temperature and solubility of $NaNO_3$.

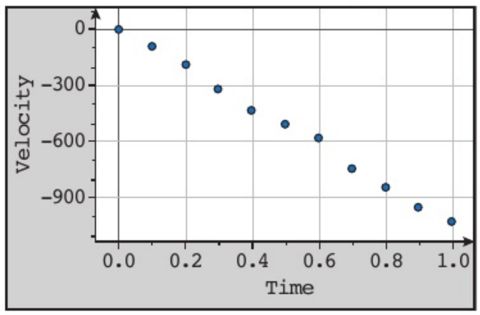
(c) $S = 0.872T + 67.5$. The gradient of 0.872 g per 100 ml per °C indicates that for each additional °C, we expect an additional 0.872 g per 100 ml can be dissolved. The S-intercept of 67.5 indicates that at 0 °C, we expect solubility of 67.5 g per 100 ml.

(d) At $T = 25$ °C, we expect solubility of $S = 89.3$ g per 100 ml.

(e) 95 °C is beyond the range of observed data; it would be an extrapolation.

(f) We should not use this model to predict temperature from solubility. (Also we do not know if the solution given has the maximum amount of $NaNO_3$ dissolved for that temperature.)

12. (a) The appears to be a very strong negative linear correlation between time and velocity.



(b) $r = -0.998$. There is an almost perfect negative linear correlation between time and velocity.

(c) $v(t) = -1040t + 7.87$

(d) The gradient of -1040 cm s^{-2} indicates that for each additional second, velocity changes by -1040 cm s^{-1}. The intercept of 7.87 suggests that the object was moving upward at 7.87 cm s^{-1} at the start of the experiment.

(e) -148 cm sec^{-1}

(f) The rate of change of velocity predicted by this experiment is -1040 cm s^{-2}. This is greater that the possible acceleration due to gravity; most likely there are errors in measurement.