MATHEMATICS
Common Core

Bill Blyth, Györgyi Bruder, Fabio Cirrito, Millicent Henry, Benedict Hung, William Larson, Rory McAuliffe, James Sanders.

6th Edition

For use with the I.B. Diploma Programme

MATHEMATICS COMMON CORE

Bill Blyth, Györgyi Bruder,

Fabio Cirrito, Millicent Henry,

Benedict Hung, William Larson,

Rory McAuliffe, James Sanders.

6th Edition

Copyright ©IBID Press, Victoria.

www.ibid.com.au

First published in 2019 by IBID Press, Victoria

Library Catalogue:

Bruder, Cirrito, Henry, Hung, Larson, McAuliffe, Sanders

1. Mathematics

2. International Baccalaureate.

Series Title: International Baccalaureate in Detail

ISBN 978-1-921784-86-6

IBID Press expresses its thanks to the International Baccalaureate Organization for permission to reproduce its intellectual property.

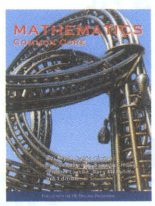
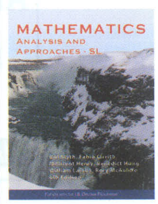
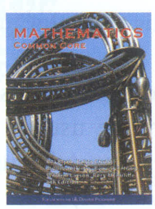
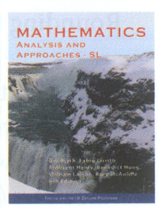
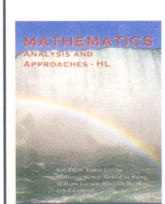
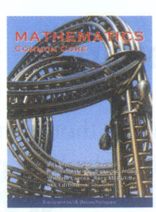
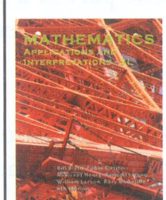
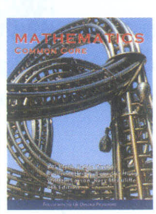
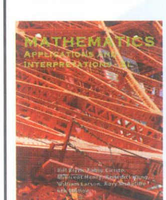
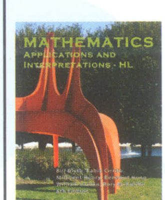
Cover design by Key-Strokes.

Published by IBID Press, www.ibid.com.au

Printed by Red Planet

Using the 6th Editions of IBID Press Mathematics Texts

This series of texts has been written for the IB Courses Mathematics: Analysis and Approaches and Mathematics: Applications and Interpretations that start teaching in August 2019.

Course Studied	Common Core	Mathematics: Analysis and Approaches (SL)	Mathematics: Analysis and Approaches (HL)	Mathematics: Applications and Interpretations (SL)	Mathematics: Applications and Interpretations (HL)	Discounted Package
Mathematics: Analysis and Approaches (SL)						Pure(SL)
Mathematics: Analysis and Approaches (HL)						Pure(HL)
Mathematics: Applications and Interpretations (SL)						Applied(SL)
Mathematics: Applications and Interpretations (HL)						Applied(HL)

PREFACE

This text for the Mathematics: Analysis and Approaches course has been prepared to closely align with the current course.

It has concise explanations, clear diagrams and calculator references.

Appropriate, graded exercises are provided throughout.

Also relating to International Perspectives and the Theory of Knowledge, it provides more than just the basics. It is an essential resource for those teachers and students who are looking for a reliable guide for their SL course.

This is a re-worked and revised edition of the Standard Level text first published by IBID Press in 1997.

2nd Edition published in 1999

3rd Edition published in 2004

4th Edition published 2012

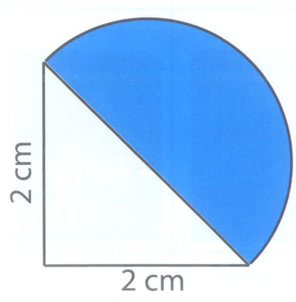
5th Edition published in 2018

6th Edition published in 2019

Rounding

When carrying an answer from one part of a calculation to a subsequent part, it is best to use unrounded values. For example:

A semicircle is constructed on the hypotenuse of a 2 cm right angled triangle. Find its area correct to 4 significant figures.



Stage 1:

Calculate the length of the hypotenuse using the theorem of Pythagoras:

$\sqrt{2^2+2^2} \approx 2.828 (4 \text{ s.f.})$

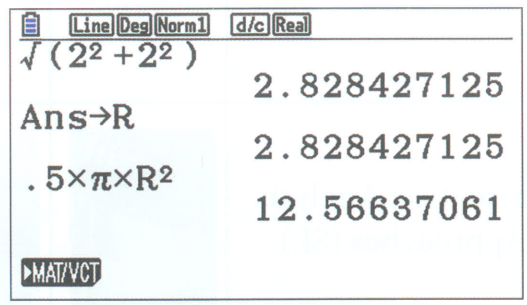
Stage 2

Find the area of the circle and halve it:

Area $= \frac{1}{2}\pi r^2 \approx \frac{1}{2}\pi \times 2.828^2 \approx 12.56$ cm (4 s.f.)

In this text, we will show calculations in this way as, we believe, students will be able to follow our explanations more easily if we do this.

However, if using a calculator, the best procedure is to use the memory to store a full accuracy version of the radius (second line).



Rounding this gives the answer 12.57 cm^2 (4 s.f.)

Note that the two answers are different. We understand that both answers are usually marked correct in examinations. However, we suggest that using the memory and the calculator value of π (not 3.14) is the better method.

Calculators

Students who are thoroughly familiar with the capabilities of their model of calculator place themselves at a considerable advantage over students who are not.

In preparing a text such as this, we cannot provide an exhaustive account of every place in which a calculator can help. Or, for that matter, an explanation of how each model works!

This text uses examples from Texas Instruments and Casio graphic calculators.

The manufacturers all provide extensive 'manuals'. These can be intimidating.

We suggest that a good strategy is to take each topic and, as you are learning it, take some time to discover your model's capability in that topic.

For example, Section 1.3 deals with counting principles. It is highly likely your calculator will be very helpful here. A good strategy can be to 'Google' or 'Bing' your model plus the topic.

There are now a number of training videos available on YouTube.

Answers

Answers to the Exercises are available as a free download from the publisher's website:

www.ibid.com.au

Also, there are QR codes embedded in the text that link directly to these.

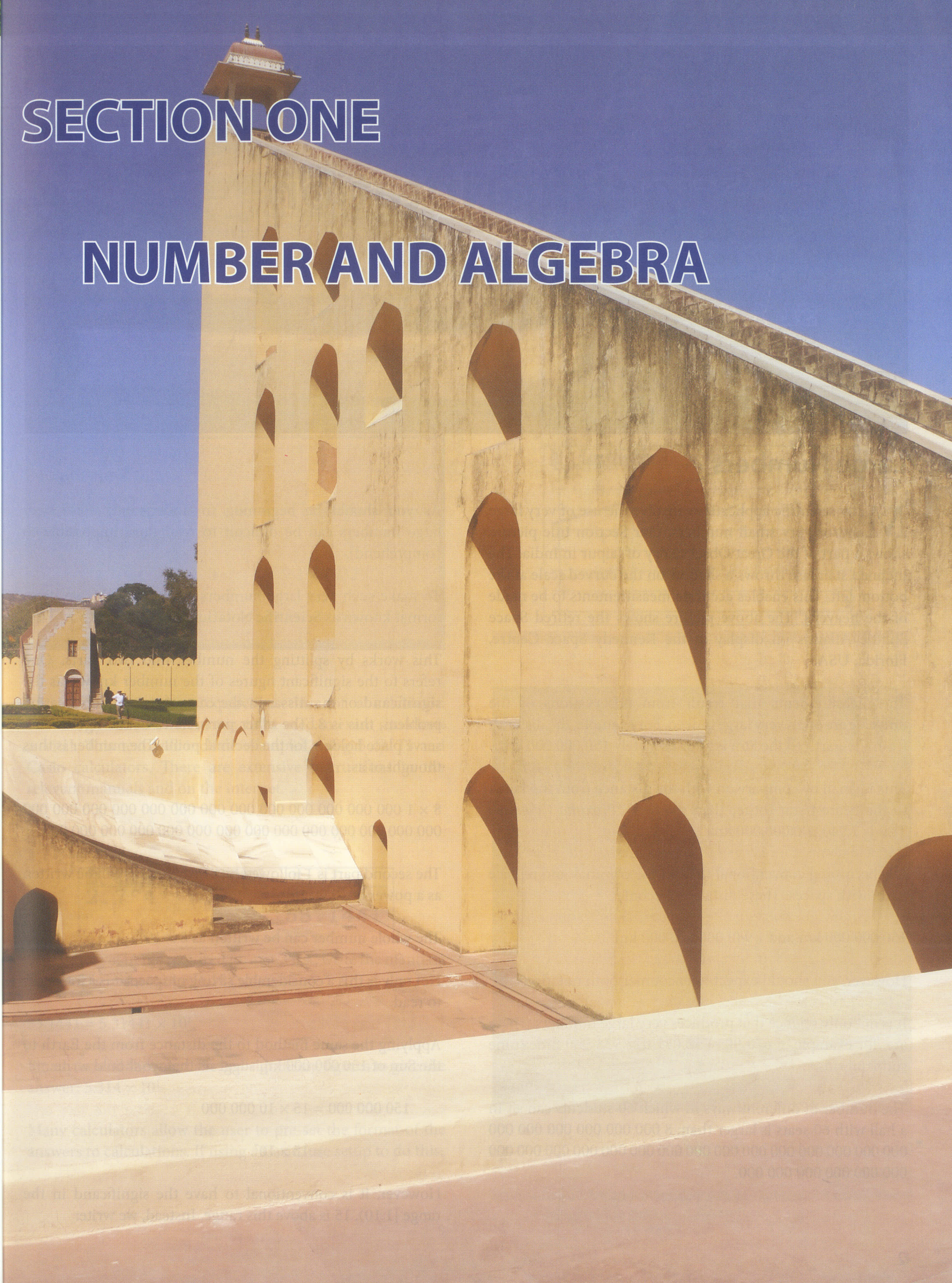
Online Errata

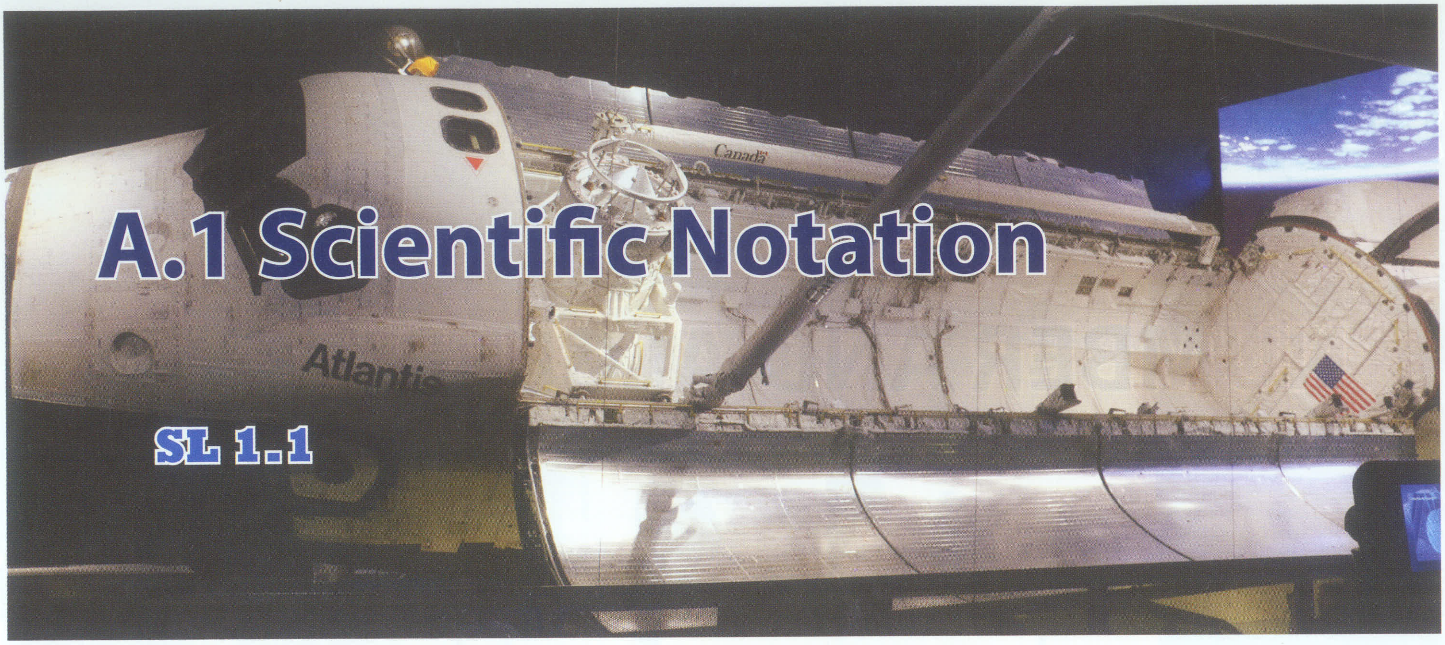
Supplementary Material

TABLE OF CONTENTS

SECTION ONE

NUMBER AND ALGEBRA





A.1 Scientific Notation

SL 1.1

Large Numbers

Many scientific applications involve the use of very large and/or very small numbers. Our Section title picture shows a part of the Great Observatory of Jaipur in India. The inclined staircase throws a shadow on the curved scale at the bottom left. This enables accurate measurements to be made of the heavens. The above picture shows the retired Space Shuttle Atlantis on display at the Kennedy Space Centre, Florida, USA.

The measurements that result from observations of the universe are often very large indeed. For example, the distance from the Earth to the Sun is approximately 150,000,000 km.

This is small by comparison with the distance from the Earth to the nearest star (other than the Sun), *Proxima Centauri*. This is 1,940,000,000,000 km.

It is becoming conventional to omit the commas and replace them with spaces. These distances become:

150 000 000 km and 1 940 000 000 000 km.

However, you should expect to encounter both notations.

It is not only science that produces very large numbers. There are, for example, 6 670 903 752 021 072 936 960 nine×nine sudoku grids.

The number of different ways in which 60 students can sit in a hall with 60 seats is larger than: 8 000 000 000 000 000 000 000 000 000 000 000 000 000 000 000 000 000 000 000 000 000 000 000 000 000 000 000 000.

As you should be beginning to understand, these very large numbers can be difficult to read (and impossible to comprehend).

To make such very large numbers more readable, we use a format known as **Scientific Notation**.

This works by splitting the number into two parts. One refers to the significant figures of the number known as the **significand** or **mantissa**. In the case of our student seating problem, this is 8. The other part relates to the figures that serve place holders for the decimal point. The number is thus thought of as:

8 × 1 000 000 000 000 000 000 000 000 000 000 000 000 000 000 000 000 000 000 000 000 000 000 000 000 000 000 000 000

The second part is 1 followed by 81 zeros. This can be written as a power of 10 ie. 10^{81}.

The whole number can be written as 8×10^{81}.

Not only is this a considerable saving of space, it is a lot easier to read.

Applying the same method to the distance from the Earth to the Sun of 150 000 000 km suggests that we should write:

$$150\,000\,000 = 15 \times 10\,000\,000$$

$$= 15 \times 10^7.$$

However, it is conventional to have the significand in the range [1,10). 15 is above this range. Instead, we write:

$150\,000\,000 = 1.5 \times 100\,000\,000$

$= 1.5 \times 10^8$.

Thus, in scientific notation, the distance from the Earth to the Sun $\approx 1.5 \times 10^8$ km.

In the same way, the distance from the Earth to *Proxima Centauri* $\approx 1\,940\,000\,000\,000$ km.

In scientific notation:

$$1\,940\,000\,000\,000 = 1.94 \times 1\,000\,000\,000\,000$$

$$= 1.94 \times 10^{12}.$$

Example A.1.1

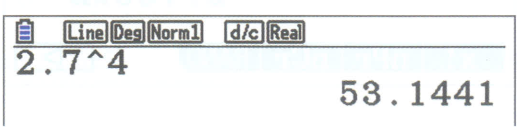
Evaluate, expressing your answers in scientific notation, correct to four significant figures:

a $(2.7)^4$ b $4\,562 \times 7\,432$

c 15! d $(3.05^{\,2})^9$

This sort of question requires the use of a calculator. Familiarity with the calculator to be used in exam situations is vital. This text will include examples using both TI and Casio calculators. There are extensive instructions in the relevant manuals and on the internet.

a

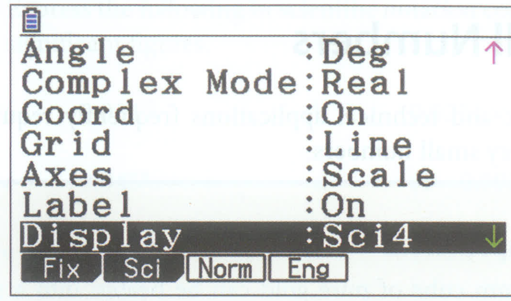


This is not yet in the required format. First, write the mantissa so that it is in the range [1,10).

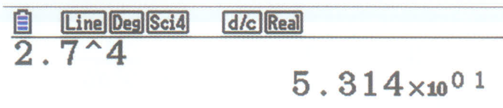
$53.1441 = 5.31441 \times 10^1$.

Finally, round the mantissa to 4 significant figures to give the answer: 5.314×10^1

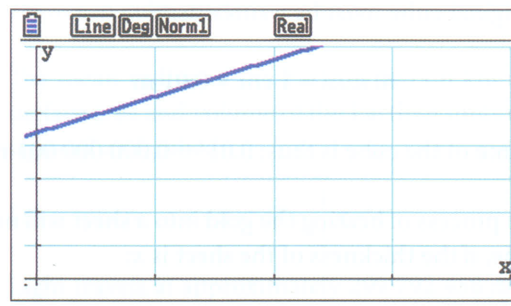
Many calculators allow the user to pre-set the format of the answers to calculations. If using a Casio, use setup to do this.



Repeat the calculation with this set-up and the result is:



b-d



International Perspective

The convention of digit grouping in threes (3 956 000) is now standard in scientific literature. It is not, however, universal.

Throughout the Indian Subcontinent (where our numeration system was invented) it is common, particularly in financial applications, to group digits otherwise.

For example 250 000 rupees is written as 2,50,000 rupees

or ₹2.5 *lakh*. Prices are often quoted in *lakh*.

There is a larger denomination known as the *crore*.

141 232 413 is written as 14,12,32,413 and described as fourteen *crore* twelve *lakh* thirty two thousand four hundred and thirteen.

Small Numbers

Scientific and technical applications frequently require the use of very small numbers.

Example A.1.2

A 10 mm cube of pure gold can be beaten into a sheet of 'gold leaf' of area 1 square metre. Approximately how thick is this sheet?

When solving these sorts of problems, it is important to work with consistent units. This means converting all the quantities into the same units. If we choose to use metres, the dimension of the original cube must be written as:

$$10 \text{ mm} = 1 \text{ cm} = 0.01 \text{ m}$$

The volume of the cube is thus: $0.01^3 = 0.000\,000\,001 \text{ m}^3$

Since the process of beating the gold into a sheet will not alter its volume, if the thickness of the sheet is x:

$$0.000\,000\,001 \times x = 1$$

Since the volume of the sheet is its surface area multiplied by its thickness. Thus the thickness is: 0.000 000 001 m.

This suffers from the same readability problems that we discussed with regard to very large numbers. To deal with this, we write:

$$\begin{aligned} 0.000\,000\,001 &= 1 \times \frac{1}{1000000000} \\ &= 1 \times \frac{1}{10^9} \\ &= 1 \times 10^{-9} \end{aligned}$$

Note the use of the negative index.

Example A.1.3

It is sometimes said that a roomful of monkeys typing aimlessly would eventually type the complete works of Shakespeare. Estimate the probability that one such monkey would type "To be or not to be, that is the question" at its first go?

Questions such as this require some assumptions. In this case, what chance does the monkey have of getting the first character (capital T) correct? There are approximately 45 keys on a QWERTY keyboard. The shift key doubles this to 90. This means that the probability of getting the first character right is $^1/_{90}$. The probability of getting the first two characters right is this $(^1/_{90})^2$. See chapter 4.6 for the background to this.

The entire quotation has 40 characters. The probability of getting all these right is $(^1/_{90})^{40}$.

This is such an extremely small number that it is effectively zero.

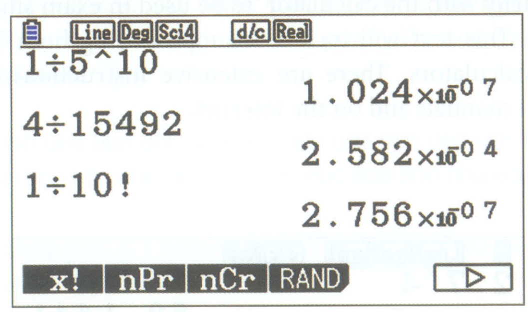
Even if every monkey that has ever lived spent its whole life typing, there is little chance that we would ever see even this quotation.

Example A.1.4

Evaluate, expressing your answers in scientific notation, correct to four significant figures:

a $\quad 1 \div 5^{10}$ $\qquad$ b $\quad 4 \div 15\,492$

c $\quad 1 \div 10!$



Estimation

Particularly when using technology, it is important to get into the habit of estimating answers before entering the calculation into the device. This is to counter the very real risks of keying errors. These errors (missing a decimal point, pressing + instead of × etc.) often produce very large errors in the answers.

Estimation usually means rounding the numbers to one significant figure and conducting 'mental arithmetic'.

For example, 43 516 × 3 927 can be estimated by writing each of these to one significant figure:

$$43\,516 \times 3\,927 \approx 40\,000 \times 4\,000$$

$$= 160\,000\,000$$

This compares with the correct answer of 170 887 332. This is not too far away from the estimated answer and we can be confident in accepting it.

A keying error (such as omitting a digit) will give a very different answer and alert the user to the error.

Scientific notation can help with estimation. If this sample calculation is written this way, it becomes:

$$43\,516 \times 3\,927 \approx 4 \times 10^4 \times 4 \times 10^3$$

$$= 16 \times 10^7 \text{ (using addition of indices)}$$

Beware, however, when using addition and subtraction.

$1.7 \times 10^{12} + 7.84 \times 10^5$ is very little different from 1.7×10^{12}. It is all a matter of scale!

As a final comment on estimation, when solving applications questions, you should have an idea of the magnitude of your answer. Answers such as 'the average speed of the car is 12 945 kph' ought to raise concerns in your mind.

Exercise A.1.1

1. Express the following in scientific notation correct to 4 significant figures.

a	525^8	b	$3\,634 \times 78\,900$
c	3.74^{11}	d	$(14.7 + 102)^4$
e	$(152.9 - 9.66)^9$	f	$7(223.9 - 4.82)^4$
g	$\frac{1}{0.03^9}$	h	$\frac{7.45^9}{0.35^4}$
i	$11!$	j	$5.5^{6.3}$
k	$\frac{1}{\sqrt[3]{0.05}}$	l	$4!^{4!}$
m	$(3.03 \times 10^7)^4$	n	$\left(\frac{400}{\sqrt{0.045}}\right)^3$

2. Express the following in scientific notation correct to 4 significant figures.

a	0.034^{12}	b	0.0214×0.0045
c	0.045^9	d	$(1 - 0.993)^7$
e	$(.56^{\,3})^9$	f	$(1.44 - 0.93)^{-5}$
g	$\frac{1}{3.7^9}$	h	$\frac{1.2^4}{9.7^{11}}$
i	$(9!)^{-6}$	j	$0.045^{5.7}$
k	$\frac{1}{\sqrt[3]{650}}$	l	$4!^{-4!}$
m	$(4.6 \times 10^{-7})^3$	n	$\left(\frac{3}{56}\right)^9$

3. Find the volume of a cube of side 500 mm. Give your answer in cubic metres.

4. The mean radius of the Earth is 6371.0 km. Find the surface area and volume of the Earth in units of metres squared and cubed.

   For a sphere $A = 4\pi r^2, V = \frac{4}{3}\pi r^3$.

5. Light travels at approximately 299 792 458 metres per second. Use the figure given earlier in this chapter for the distance from the Earth to the Sun to calculate the time it takes light to travel from the Sun to the Earth.

6. One gram of hydrogen contains approximately 6.022×10^{23} hydrogen atoms. The Sun weighs approximately 1.99×10^{30} kg and is 74.9% hydrogen. Estimate the number of hydrogen atoms in the Sun.

7. Find the smallest whole numbered value of x such that $x^x > 10^{17}$.

8. The yellow discharge lamps that are frequently used for street lighting work by passing an electric discharge through sodium vapour. The discharge produces light with a wavelength (λ) near 589 nanometres. A nanometre is 1×10^{-9} metre. Use the relationship $c = f\lambda$ where c is the velocity of light (3×10^8 ms^{-1}) and f is the frequency of the light. Find the frequency of sodium light.

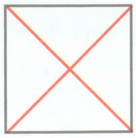
9. Knowledge of logarithms required. The Prime Number Theorem estimates that there are approximately:

   $\frac{N}{\log_e N}$ prime numbers less than N.

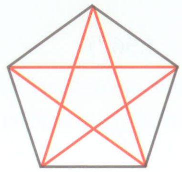
   Estimate the percentage of numbers between 10^{12} and 10^{13} that are prime.

10. The number of diagonals of a regular polygon depends upon the number of sides:

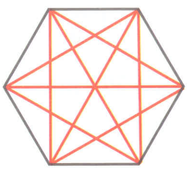
Sides: 4
Diagonals: 2



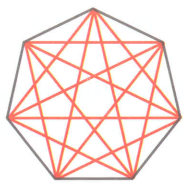
Sides: 5
Diagonals: 5



Sides: 6
Diagonals: 9



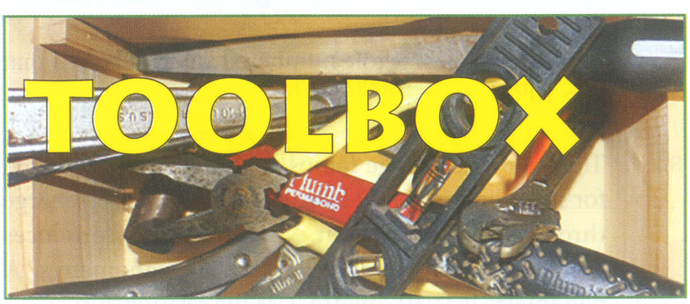
Sides: 7
Diagonals: 14



Look at the number of diagonals that meet at each vertex and find a rule for the number of diagonals of a *n* sided polygon.

How many diagonals can be drawn in a million sided regular polygon?

Answers



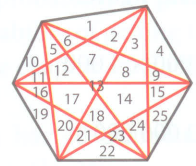
This section will provide some suggestions for further study of this topic.

1. Return to the diagrams of the diagonals of polygons. The diagonals divide the polygons (question 10 in the exercise).

In the case of the square, the diagonals divide it into 4 regions.

In the case of the pentagon, there are 11 regions.

The hexagon is a bit more complex as there is a point at which more than 2 diagonals intersect. If we displace one vertex, this need not be the case and the number of regions will be maximised.



There are 25 regions.

The diagram of the heptagon has no places where three diagonals intersect and needs no modification. There are 50 regions.

Continue adding sides in such a way that the number of regions is maximised (there are no places where more than two diagonals intersect). Can you find a formula that will allow you to predict the number of regions in a *n* sided polygon?

How many regions are produced by the diagonals on a million sided polygon?

2. Tom Stoppard's highly entertaining play *Rosencrantz and Guildenstern are Dead* takes a look at Shakespeare's *Hamlet* from the point of view of two minor characters.

The play opens with Rosencrantz and Guildenstern betting on coin tosses. Rosencrantz, who bets heads each time, wins ninety-two tosses in a row.

Earlier on in this chapter, we looked at the probability that this sort of thing might happen by chance.

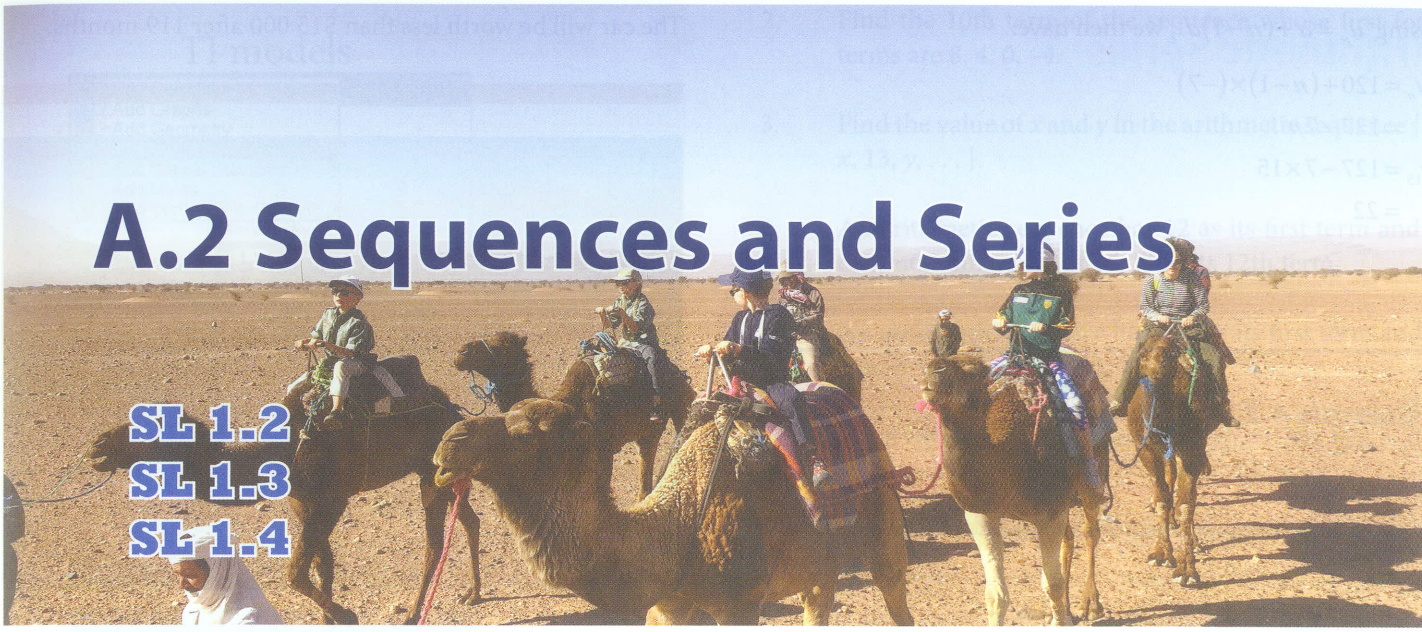
It is very small (2.019×10^{-28}).

However, if you actually observed this, would you really conclude that you have just seen a highly unlikely event?

Nine tails in a row. Less than 1 chance in a thousand. Do you believe it?

s

A.2 Sequences and Series



Arithmetic Sequences

A sequence is a set of quantities arranged in a definite order.

1, 2, 3, 4, 5, 6, ... −1, 2, −4, 8, −16, ... 1, 1, 2, 3, 5, 8, 13, ...

are all examples of sequences. When the terms of a sequence are added, we obtain a series. Sequences and series are used to solve a variety of practical problems in, for example, business.

There are two major types of sequences: arithmetic and geometric. This section will consider arithmetic sequences (also known as arithmetic progressions, or simply A.P.). The characteristic of such a sequence is that there is a common difference between successive terms. For example:

1, 3, 5, 7, 9, 11, . . . (the odd numbers) has a first term of 1 and a common difference of 2.

18, 15, 12, 9, 6, . . . has a first term of 18 and a common difference of −3 (sequence is decreasing).

The terms of a sequence are generally labelled:

$$u_1, u_2, u_3, u_4, \ldots, u_n .$$

The 'nth term' of a sequence is labelled u_n. In the case of an arithmetic sequence which starts with a and has a common difference of d, the nth term can be found using the formula:

$$u_n = a + (n-1)d \text{ where } d = u_2 - u_1 = u_3 - u_2 = \ldots \; n = 1, 2, 3\ldots$$

Example A.2.1
For the sequence 7, 11, 15, 19, ... , find the 20th term.

In this case, $a = 7$ and $d = 4$ because the sequence starts with a 7 and each term is 4 larger than the one before it, i.e. $d = 11 - 7 = 4$. Therefore the nth term is given by

$$u_n = 7 + (n-1)4$$

That is, $u_n = 4n + 3$

$$u_{20} = 4 \times 20 + 3 = 83$$

($n = 20$ corresponds to the 20th term)

Example A.2.2
An arithmetic sequence has a first term of 120 and a 10th term of 57. Find the 15th term.

The data is: $a = 120$ and when $n = 10$, $u_{10} = 57$ (i.e. 10th term is 57).

This gives, $u_{10} = 120 + (10-1)d$

$$120 + 9d = 57$$

$$d = -7$$

Using $u_n = a + (n-1)d$, we then have:

$$\begin{aligned} u_n &= 120 + (n-1) \times (-7) \\ &= 127 - 7n \\ u_{15} &= 127 - 7 \times 15 \\ &= 22 \end{aligned}$$

Example A.2.3
An arithmetic sequence has a 7th term of 16.5 and a 12th term of 24. Find the 24th term.

In this instance we know neither the first term nor the common difference and so we need to set up equations to be solved simultaneously.

The data is: $u_7 = a + 6d = 16.5$ – (1)

$u_{12} = a + 11d = 24$ – (2)

We first solve for 'd':(2) – (1) : $5d = 7.5 \Leftrightarrow d = 1.5$

Substituting into (1): $a + 6 \times 1.5 = 16.5 \Leftrightarrow a = 7.5$

To find the 24th term we use the general term: $u_n = a + (n-1)d$ with n = 24:

$u_{24} = 7.5 + (24-1) \times 1.5 = 42$

Example A.2.4
A car whose original value was \$25 600 decreases in value by \$90 per month. How long will it take before the car's value falls below \$15 000?

The values can be seen as a sequence: \$25 600, \$25 510, \$25 420 etc.

In this case a = 25 600 and d =25 510 – 25 600 = –90 so that:

$$\begin{aligned} u_n &= 25690 + (n-1) \times (-90) \\ &= 25690 - 90n \\ 15000 &= 25690 - 90n \\ 90n &= 25690 - 15000 \\ n &= 118.77 \end{aligned}$$

The car will be worth less than \$15 000 after 119 months.

On 'final approach' a pilot aims to hold the airspeed constant and to descend in a straight line. As a result, if measured at regular intervals, range to the 'piano keys' and altitude form arithmetic sequences.

Using a graphics calculator

Most graphic calculators have an automatic memory facility (often called **Ans**) that stores the result of the last calculation as well as an ability to remember the actual calculation. This can be very useful in listing a sequence.

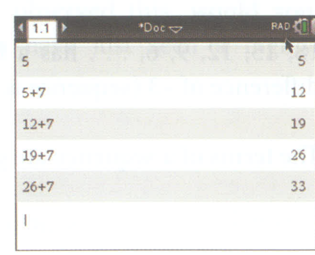
Example A.2.5

List the arithmetic sequence 5, 12, 19, 26, ...

The sequence has a first term of 5. Enter this and press ENTER or EXE.

The common difference of the sequence is 7 so enter + 7.

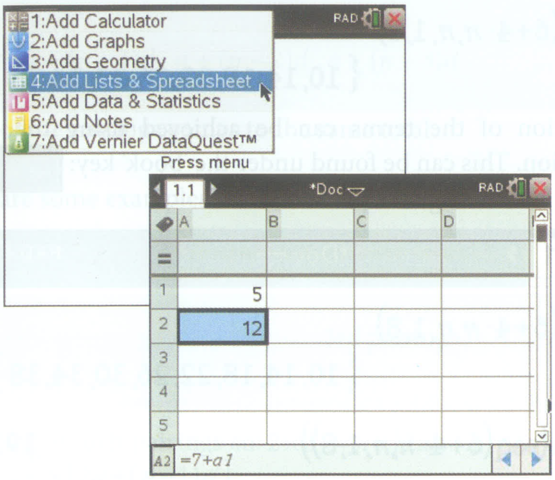
The display will show Ans + 7 which means 'add 7 to the previous answer'.



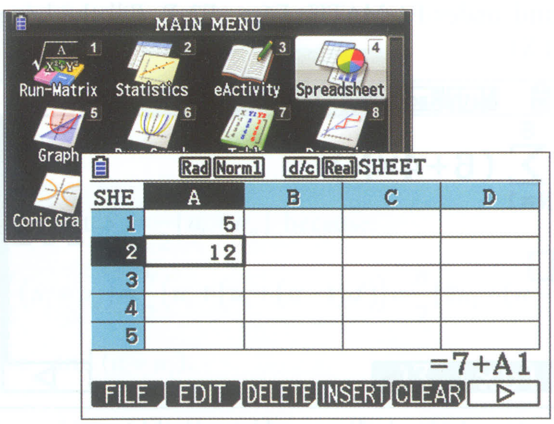
From here, every time you press ENTER (or EXE), you will repeat the calculation, generating successive terms of the sequence.

However most calculators are more sophisticated than this. It is possible to set up a spreadsheet type file. These can be used much as one might use Excel to make repetitive calculations.

TI models



Casio models



We now consider Example A.2.2, where we obtained the sequence $u_n = 127 - 7n$ and wished to determine the 15th term.

Most calculators have many features that can be used with sequences. Become familiar with all of them for your model.

Exercise A.2.1

1.

a Show that the following sequences are arithmetic.

b Find the common difference.

c Define the rule that gives the nth term of the sequence.

i $\{2, 6, 10, 14, \ldots\}$

ii $\{20, 17, 14, 11, \ldots\}$

iii $\{1, -4, -9, \ldots\}$

iv $\{0.5, 1.0, 1.5, 2.0, \ldots\}$

v $\{y+1, y+3, y+5, \ldots\}$

vi $\{x+2, x, x-2, \ldots\}$

2. Find the 10th term of the sequence whose first four terms are 8, 4, 0, –4.

3. Find the value of x and y in the arithmetic sequence $\{5, x, 13, y, \ldots\}$.

4. An arithmetic sequence has 12 as its first term and a common difference of –5. Find its 12th term.

5. An arithmetic sequence has –20 as its first term and a common difference of 3. Find its 10th term.

6. The 14th term of an arithmetic sequence is 100. If the first term is 9, find the common difference.

7. The 10th term of an arithmetic sequence is –40. If the first term is 5, find the common difference.

8. If $n + 5$, $2n + 1$ and $4n - 3$ are three consecutive terms of an arithmetic sequence, find n.

9. The first three terms of an arithmetic sequence are 1, 6, 11.

   a. Find the 9th term.

   b. Which term will equal 151?

10. Find x and y given that $4-\sqrt{3}, x, y$ and $2-\sqrt{3}$ are the first four terms of an arithmetic sequence.

11. For each of the following sequences, determine:

   i. its common difference

   ii. its first term

   a. $u_n = -5 + 2n, n \geq 1$

   b. $u_n = 3 + 4(n+1), n \geq 1$

12. The third and fifth terms of an A.P. are $(x + y)$ and $(x - y)$ respectively. Find the twelfth term.

13. The sum of the fifth term and twice the third of an arithmetic sequence equals the twelfth term. If the seventh term is 25 find an expression for the general term, u_n.

14. For a given arithmetic sequence, $u_n = m$ and $u_m = n$. Find:

   a. the common difference.

   b. u_{n+m}.

Arithmetic Series

If the terms of a sequence are added, the result is known as a series.

The **sequence**: 1, 2, 3, 4, 5, 6, . . .

gives the **series**: 1 + 2 + 3 + 4 + 5 + 6 + . . .

and the **sequence**: –1, –2, –4, –8, –16 . . .

gives the **series**: (–1) + (–2) + (–4) + (–8) + (–16) + . . .

(or – 1 – 2 – 4 – 8 – 16 – . . .)

The sum of the terms of a series is referred to as S_n, the **sum of *n* terms of a series**.

For an arithmetic series, we have:

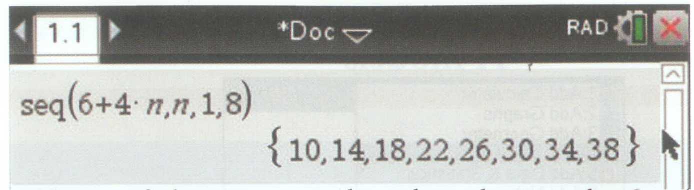
$$S_n = u_1 + u_2 + u_3 + ... u_n$$
$$= a + (a+d) + (a+2d) + (a+3d) + ... + (a+(n-1)d)$$

For example, if we have a sequence defined by $u_n = 6 + 4n, n \geq 1$ then the sum of the first 8 terms is given by:

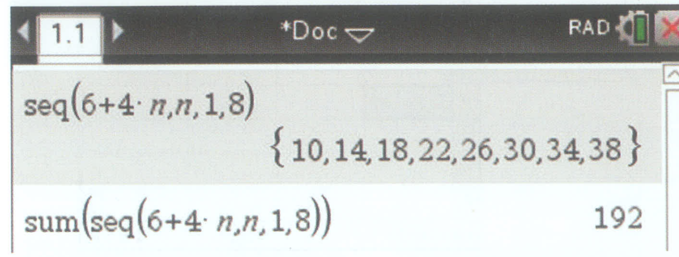
$$S_8 = u_1 + u_2 + u_3 + ... u_8$$
$$= 10 + 14 + 18 + ... + 38$$
$$= 192$$

Most calculators have several ways of handling sequences and series. We have already referred to the LIST (spreadsheet) feature.

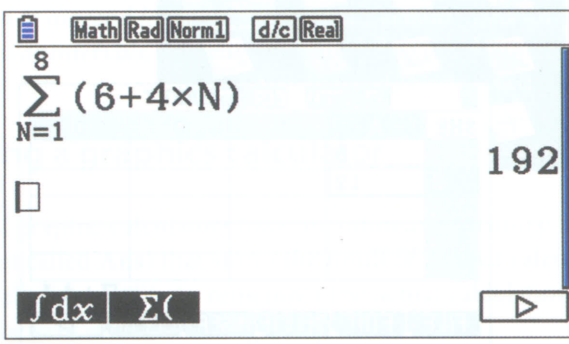
An alternative (TI) is to use MENU, 6 (STATISTICS), 4 (LIST OPERATIONS),5 (SEQUENCE) - the exact calculator method will vary - consult the manual!:



Addition of the terms can be achieved using the **Sum(** function. This can be found under the 'book' key:



Casio provide a summation function in run mode. This can be found under F4-MATH, F6-▷, F2-Σ. Fill the boxes and press EXE:



The sigma notation is discussed later in this section.

There will be many cases in which we can add the terms of a series in this way. If, however, there are a large number of terms to add, a formula is more appropriate.

There is a story that, when the mathematician Gauss was a child, his teacher was having problems with him because he always finished all his work long before the other students. In an attempt to keep Gauss occupied for a period, the teacher asked him to add all the whole numbers from 1 to 100. '5050' Gauss replied immediately.

It is probable that Gauss used a method similar to this:

1	2	3	4	5	6	......,	96	97	98	99	100
100	99	98	97	96	95	......,	5	4	3	2	1
101	**101**	**101**	**101**	**101**	**101**	**......,**	**101**	**101**	**101**	**101**	**101**

Adding each of the pairings gives 100 totals of 101 each. This gives a total of 10100. This is the sum of two sets of the numbers 1 + 2 + 3 + ... + 98 + 99 +100 and so dividing the full answer by 2 gives the answer 5050, as the young Gauss said.

It is then possible to apply the same approach to such a sequence, bearing in mind that the sequence of numbers must be arithmetic.

Applying this process to the general arithmetic series we have:

a	$a + d$	$a + 2d$	...	$a + (n-3)d$	$a + (n-2)d$	$a + (n-1)d$
$a + (n-1)d$	$a + (n-2)d$	$a + (n-3)d$	...	$a + 2d$	$a + d$	a

Each of the pairings comes to the same total.

Here are some examples:

1st pairing: $a + a + (n-1)d = 2a + (n-1)d$

2nd pairing: $a + d + a + (n-2)d = 2a + (n-1)d$

3rd pairing: $a + 2d + a + (n-3)d = 2a + (n-1)d$

There are n such pairings so: $2\times S_n = n\times[2a+(n-1)d]$ That is, $S_n = \frac{n}{2}\times[2a+(n-1)d]$.

Giving the formula, for the sum of n terms of a sequence:

$$S_n = \frac{n}{2}\times[2a+(n-1)d]$$

Note also that $S_n = \frac{n}{2}(u_1+u_n)$ because:

$$S_n = \frac{n}{2}(u_1+u_n) = \frac{n}{2}(u_1+(u_1+(n-1)d)) = \frac{n}{2}(2u_1+(n-1)d)$$

as above - and $(a = u_1)$.

This formula can now be used to sum large arithmetic series:

Example A.2.6

Find the sum of 20 terms of the series $-2 + 1 + 4 + 7 + 10 + \ldots$

We have the following information: $a = u_1 = -2$

and $d = u_2 - u_1 = 1 - (-2) = 3$.

Then, the sum to n terms is given by: $S_n = \frac{n}{2}\times[2a+(n-1)d]$

So that the sum to 20 terms is given by

$$S_{20} = \frac{20}{2}\times[2\times(-2)+(20-1)\times 3]$$
$$= 10[-4+19\times 3]$$
$$= 530$$

Example A.2.7

Find the sum of 35 terms of the series: $-\frac{3}{8}-\frac{1}{8}+\frac{1}{8}+\frac{3}{8}+\frac{5}{8}+\ldots$

We have the following information: $a = u_1 = -\frac{3}{8}$

and $d = u_2 - u_1 = -\frac{1}{8}-\left(-\frac{3}{8}\right)=\frac{1}{4}$.

Then, with $n = 35$ we have

$$S_{35} = \frac{35}{2}\left[2\times\left(-\frac{3}{8}\right)+(35-1)\times\frac{1}{4}\right]$$
$$= 17.5\left[-\frac{3}{4}+34\times\frac{1}{4}\right]$$
$$= 135\frac{5}{8}$$

Example A.2.8

An arithmetic series has a third term of 0. The sum of the first 15 terms is –300. What is the first term and the sum of the first ten terms?

From the given information we have: $u_3 = a + 2d = 0$ – (1)

and: $S_{15} = \frac{15}{2}[2a+14d] = -300$

i.e. $15a + 105d = -300$

$\therefore\ a + 7d = -20$ – (2)

The pair of equations can now be solved simultaneously:

(2) – (1): $5d = -20 \Leftrightarrow d = -4$

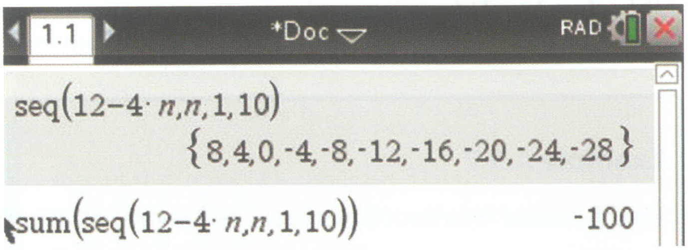
Substituting into (1) we have: $a + 2\times -4 = 0 \Leftrightarrow a = 8$

This establishes that the series is $8 + 4 + 0 + (-4) + (-8) + \ldots$

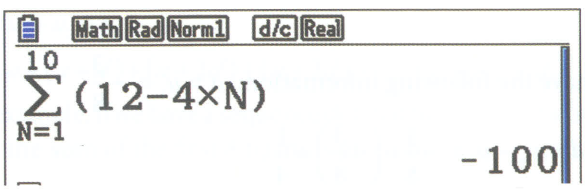
So the first term is 8 and the sum of the first ten terms is:

$$S_{10} = \frac{10}{2}[16 + 9\times -4] = -100$$

Using the TI_NSpire we have, with the general term $12 - 4n$:



Using a Casio model, this can be evaluated:



Example A.2.9

A new business is selling home computers. They predict that they will sell 20 computers in their first month, 23 in the second month, 26 in the third and so on, in arithmetic sequence. How many months will pass before the company expects to sell their thousandth computer?

The series is: 20 + 23 + 26 + ...

The question implies that the company is looking at the total number of computers sold, so we are looking at a series, not a sequence.

The question asks how many terms (months) will be needed before the total sales reach more than 1000. From the given information we have: $a = 20$, $d = 23 - 20 = 3$.

Therefore, we have the sum to n terms given by:

$$S_n = \frac{n}{2}[2\times 20 + (n-1)\times 3]$$

$$= \frac{n}{2}[3n + 37]$$

Next, we determine when $S_n = 1000$:

$$\frac{n}{2}[3n + 37] = 1000$$

$$3n^2 + 37n = 2000$$

$$3n^2 + 37n - 2000 = 0$$

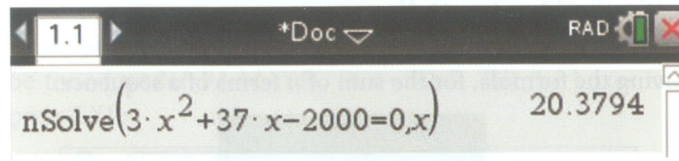
Solving for n can be done using several methods:

Method 1: Quadratic formula

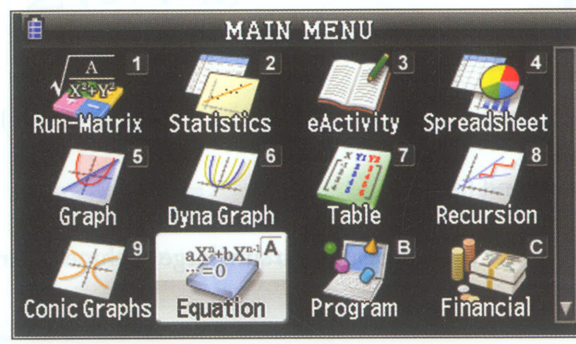
$$n = \frac{-37 \pm \sqrt{37^2 - 4\times 3\times -2000}}{2\times 3}$$

$$= 20.38\ldots\ -32.7$$

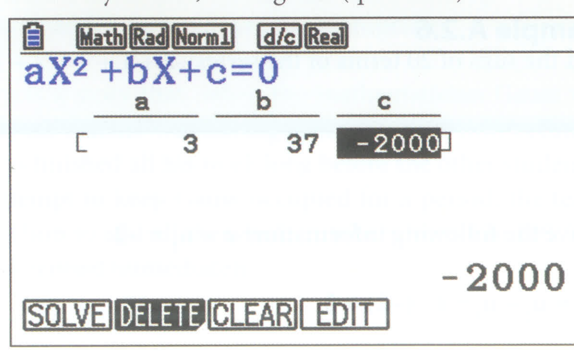
Method 2: Graphics Calculator Solve function



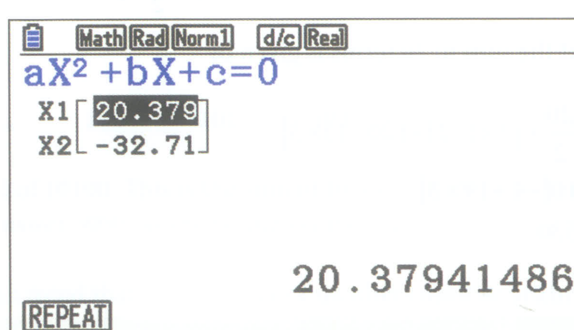
If using Casio, select the Equations module:



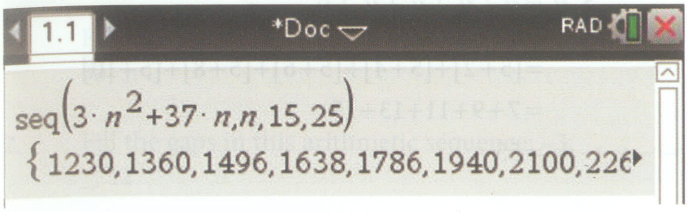
Select F2-Polynomial, F1-degree2 (quadratic)



Then F1 will initiate solve:

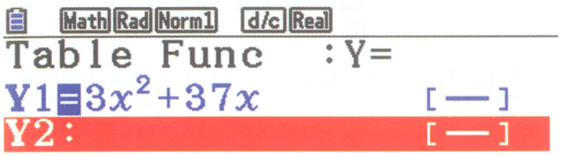


Method 3: Table of values

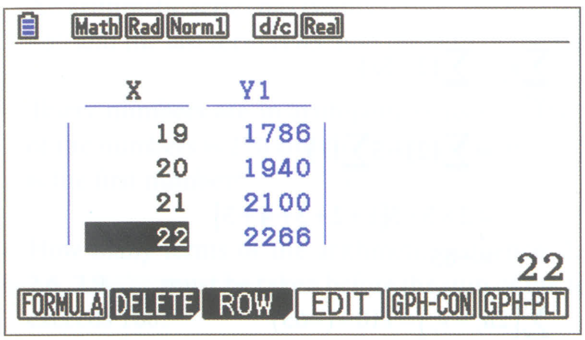


Casio: Using the Table Module (7)

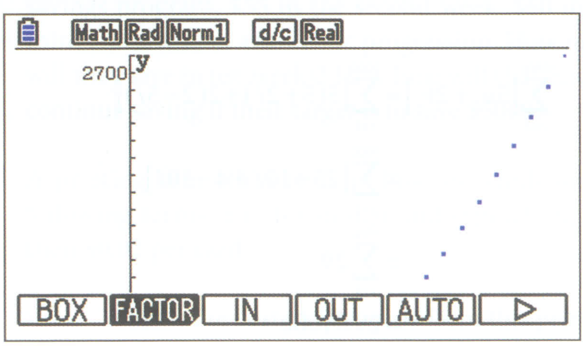
Enter the rule (using x as the variable can save time):



Use F5-SET to set the values of the variable (15 to 25) and then EXE F6-Table:



A nice feature of the Casio is that it will show you a graph. Press F6 and Shift F2-ZOOM, F5-AUTO (to find the points),



Notice that we have entered the expression for S_n as the sequence rule. In fact, the series itself is made up of terms in a sequence of so-called partial sums, often called a sum sequence.

That is, we have that $\{S_1, S_2, S_3, ...\} = \{15, 33, 54, ...\}$ forms a sequence.

The answer then, is that the company will sell its thousandth computer during the 20th month.

Exercise A.2.2

1. Find the sum of the first ten terms in the arithmetic sequences

   a $\{1, 4, 7, 10, \ldots\}$

   b $\{3, 9, 15, 21, \ldots\}$

   c $\{10, 4, -2, \ldots.\}$.

2. For the given arithmetic sequences, find the sum, S_n, to the requested number of terms.

   a $\{4, 3, 2, \ldots\}$ for $n = 12$

   b $\{4, 10, 16, \ldots\}$ for $n = 15$

   c $\{2.9, 3.6, 4.3, \ldots\}$ for $n = 11$

3. Find the sum of the following sequences:

   a $\{5, 4, 3, \ldots, -15\}$

   b $\{3, 9, 15, \ldots, 75\}$

   c $\{3, 5, 7, \ldots, 29\}$

4. The weekly sales of washing machines from a retail store that has just opened in a new housing complex increase by 2 machines per week. In the first week of January 2015, 24 machines were sold.

   a How many were sold in the last week of December 2015?

   b How many machines did the retailer sell in 2015?

   c When was the 500th machine sold?

5. The fourth term of an arithmetic sequence is 5 while the sum of the first 6 terms is 10. Find the sum of the first nineteen terms.

6. Find the sum of the first 10 terms for the sequences defined by:

   a $u_n = -2 + 8n$

   b $u_n = 1 - 4n$

7. The sum of the first eight terms of the sequence $\{\ln x, \ln x^2 y, \ln x^3 y^2, \ldots\}$ is given by $4(a\ln x + b\ln y)$. Find a and b.

Sigma notation

There is a second notation to denote the sum of terms. This other notation makes use of the Greek letter Σ as the symbol to inform us that we are carrying out a summation.

In short, Σ stands for 'The sum of . . .'.

This means that the expression

$$\sum_{i=1}^{n} u_i = u_1 + u_2 + u_3 + \ldots + u_{n-1} + u_n$$

For example, if $u_i = 2+5(i-1)$, i.e. an A.P. with first term a = 2 and common difference d = 5, the expression:

$$S_n = \sum_{i=1}^{n}[2+5(i-1)]$$

would represent the sum of the first n terms of the sequence.

So, the sum of the first 3 terms would be given by:

$$\sum_{i=1}^{3}[2+5(i-1)]$$

$$=[2+5(1-1)]+[2+5(2-1)]+[2+5(3-1)]$$

Term 1 Term2 Term 3

= 2 + 7 + 12 = 21

Properties of Σ

1. Σ is distributive. That is, $\sum_{i=1}^{n}[u_i+v_i]=\sum_{i=1}^{n}u_i+\sum_{i=1}^{n}v_i$.

2. $\sum_{i=1}^{n}ku_i = k\sum_{i=1}^{n}u_i$, for some constant value k.

3. $\sum_{i=1}^{n}k = kn$, i.e. adding a constant term, k, n times is the same as multiplying k by n.

Example A.2.10

Given that $u_i = 5 + 2i$ and $v_i = 2 - 5i$, find:

a $\sum_{i=1}^{5}u_i$ b $\sum_{i=1}^{5}[2u_i - v_i]$

c $\sum_{i=1}^{1000}[5u_i + 2v_i]$

a
$$\sum_{i=1}^{5}u_i = u_1+u_2+u_3+u_4+u_5$$
$$=[5+2]+[5+4]+[5+6]+[5+8]+[5+10]$$
$$=7+9+11+13+15$$
$$=55$$

b
$$\sum_{i=1}^{5}[2u_i - v_i]=\sum_{i=1}^{5}[2u_i]+\sum_{i=1}^{5}[-v_i]$$
$$=2\sum_{i=1}^{5}[u_i]-\sum_{i=1}^{5}[v_i]$$

Now,
$$2\sum_{i=1}^{5}[u_i]=2\times 55=110$$

and (Using properties)

$$\sum_{i=1}^{5}v_i=\sum_{i=1}^{5}(2-5i)$$
$$=\sum_{i=1}^{5}(2)-5\sum_{i=1}^{5}(i)$$
$$=2\times 5-5[1+2+3+4+5]$$
$$=-65$$
$$\sum_{i=1}^{5}[2u_i - v_i]=110-(-65)$$
$$=175$$

c
$$\sum_{i=1}^{1000}[5u_i+2v_i]=\sum_{i=1}^{1000}[5(5+2i)+2(2-5i)]$$
$$=\sum_{i=1}^{1000}[25+10i+4-10i]$$
$$=\sum_{i=1}^{1000}29$$
$$=29\times 1000$$
$$=29000$$

In this example we have tried to show that there are a number of ways to obtain a sum. It is not always necessary to enumerate every term and then add them. Often, an expression can first be simplified.

Exercise A.2.3

1. Find the twentieth term in the sequence 9, 15, 21, 27, 33, . . .

2. Fill the gaps in this arithmetic sequence: –3, _, _, _, _, _, 12.

3. An arithmetic sequence has a tenth term of 17 and a fourteenth term of 30. Find the common difference.

4. If $u_{59} = \frac{1}{10}$ and $u_{100} = -1\frac{19}{20}$ for an arithmetic sequence, find the first term and the common difference.

5. Find the sum of the first one hundred odd numbers.

6. An arithmetic series has twenty terms. The first term is –50 and the last term is 83, find the sum of the series.

7. Thirty numbers are in arithmetic sequence. The sum of the numbers is 270 and the last number is 38. What is the first number?

8. How many terms of the arithmetic sequence: 2, 2.3, 2.6, 2.9, . . . must be taken before the sum of the terms exceeds 100?

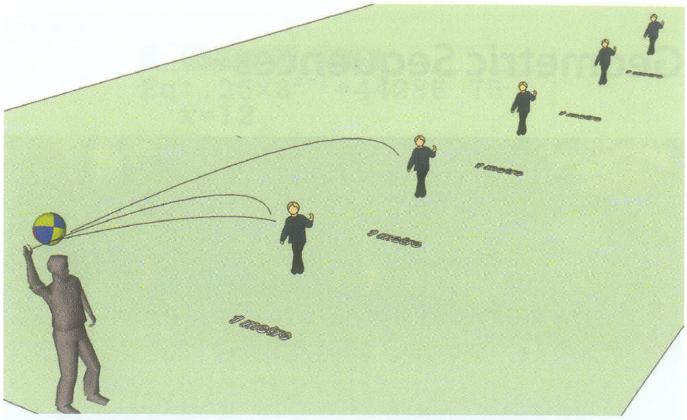
9. Sandip and Melissa save $50 in the first week of a savings program, $55 in the second week, $60 in the third and so on, in arithmetic progression. How much will they save in ten weeks? How long will they have to continue saving if their target is to save $5000?

10. A printing firm offers to print business cards on the following terms: $45 for design and typesetting and then $0.02 per card.

    a What is the cost of 500 cards from this printer?

    b How many cards can a customer with $100 afford to order?

11. A children's game consists of the players standing in a line with a gap of 2 metres between each. The adult at the left-hand end of the line has a ball which s/he throws to the first child in the line, a distance of 2 metres. The ball is then thrown back to the adult who then throws the ball to the second child in the line, a distance of 4 metres. The ball is then returned to the adult, and so on until all the children have touched the ball at least once.



    a If a total of five children play and they make the least number of throws so that only the leftmost child touches the ball more than once:

        i What is the largest single throw?

        ii What is the total distance travelled by the ball?

    b If seven children play, what is the total distance travelled by the ball?

    c If n children play, derive a formula for the total distance travelled by the ball.

    d Find the least number of children who need to play the game before the total distance travelled by the ball exceeds 100 metres.

    e The children can all throw the ball 50 metres at most.

        i What is the largest number of children that can play the game?

        ii What is the total distance travelled by the ball?

12. Find each sum.

    a $\sum_{i=1}^{100} k$ b $\sum_{i=1}^{100} (2k+1)$ c $\sum_{i=1}^{51} (3k+5)$

13. If $u_i = -3 + 4i$ and $v_i = 12 - 3i$ find:

    a $\sum_{i=1}^{10} [u_i + v_i]$ b $\sum_{i=1}^{10} [3u_i + 4v_i]$ c $\sum_{i=1}^{10} [u_i v_i]$

14.a Show that for an arithmetic sequence, $u_n = S_n - S_{n-1}$, where u_n is the nth term and S_n is the sum of the first n terms.

    b Find the general term, u_n, of the A.P given that $\sum_{i=1}^{10} u_i = \frac{n}{2}(3n-1)$.

Geometric Sequences

Attempts to understand the sizes of animal populations have often used sequences and series.

Sequences such as 2, 6, 18, 54, 162, ... and 200, 20, 2, 0.2, ... in which each term is obtained by multiplying the previous one by a fixed quantity are known as **geometric sequences**.

The sequence: 2, 6, 18, 54, 162, ... is formed by starting with 2 and then multiplying by 3 to get the second term, by 3 again to get the third term, and so on.

For the sequence 200, 20, 2, 0.2, ..., begin with 20 and multiply by 0.1 to get the second term, by 0.1 again to get the third term and so on.

The constant multiplier of such a sequence is known as the **common ratio**.

The common ratio of 2, 6, 18, 54, 162,... is 3 and of 200, 20, 2, 0.2,... it is 0.1.

The *n*th term of a geometric sequence is obtained from the first term by multiplying by $n-1$ common ratios.

This leads to the formula for the *n*th term of a geometric sequence:

$$u_n = a \times r^{n-1}$$

where $r = \frac{u_2}{u_1} = \ldots = \frac{u_n}{u_{n-1}}$ and *n* is the term number, *a* the first term and *r* is the common ratio.

Example A.2.11
Find the tenth term in the sequence 2, 6, 18, 54, 162, ...

The first term is $a = 2$. The common ratio $r = 3 = {}^6/_2 = {}^{18}/_6$ and *n*, the required term, is 10.

Use the formula to solve the problem:

$$u_n = a \times r^{n-1}$$
$$u_{10} = 2 \times 3^{10-1}$$
$$= 2 \times 3^9$$
$$= 39366$$

Example A.2.12
Find the fifteenth term in the sequence 200, 20, 2, 0.2, ...

In this case, $a = 200$, $r = \frac{20}{200} = \frac{1}{10} = 0.1$ and $n = 15$.

Using the general term $u_n = a \times r^{n-1}$, the 15th term is given by:

$$u_{15} = 200 \times 0.1^{(15-1)}$$
$$= 200 \times 0.1^{14}$$
$$= 2 \times 10^{-12}$$

Example A.2.13
Find the eleventh term in the sequence: $1, -\frac{1}{2}, \frac{1}{4}, -\frac{1}{8}, \frac{1}{16}, \ldots$

The sequence $1, -\frac{1}{2}, \frac{1}{4}, -\frac{1}{8}, \frac{1}{16}, \ldots$, has a common ratio of $r = -\frac{1}{2}$.

Using the general term , we have:

$$u_{11} = 1 \times \left(-\frac{1}{2}\right)^{11-1}$$
$$= \left(-\frac{1}{2}\right)^{10}$$
$$\approx 0.000977$$

Many questions will be more demanding in terms of the way in which you use this formula. You should also recognise that the formula can be applied to a range of practical problems.

Many of these involve growth and decay.

Example A.2.14

A geometric sequence has a fifth term of 3 and a seventh term of 0.75. Find the first term, the common ratio and the tenth term.

From the given information we can set up the following equations:

$$u_5 = a \times r^4 = 3 \qquad -(1)$$

and $$u_7 = a \times r^6 = 0.75 \qquad -(2)$$

As with similar problems involving arithmetic sequences, the result is a pair of simultaneous equations. In this case these can best be solved by dividing (2) by (1) to get:

$$\frac{a \times r^6}{a \times r^4} = \frac{0.75}{3} \Leftrightarrow r^2 = 0.25 \Leftrightarrow r = \pm 0.5$$

Substituting results into (1) we have: $a\left(\pm\frac{1}{2}\right)^4 = 3 \Leftrightarrow a = 48$

Therefore, the 10th term is given by: $u_{10} = 48 \times (\pm 0.5)^9 = \pm\frac{3}{32}$

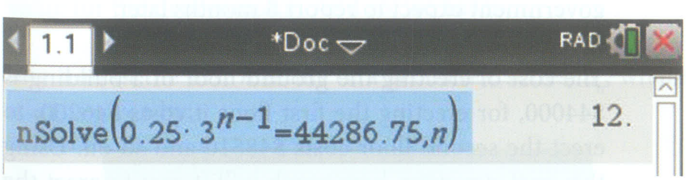
There are two solutions: 48, 24, 12, 6, . . . (for the case $r = 0.5$) & 48, –24, 12, –6, . . . ($r = -0.5$).

Example A.2.15

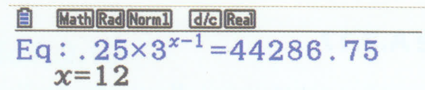
Find the number of terms in the geometric sequence: 0.25, 0.75, 2.25, . . . , 44286.75.

The sequence 0.25, 0.75, 2.25, ..., 44286.75 has a first term $a = 0.25$ and a common ratio $r = \frac{0.75}{0.25} = 3$. In this problem it is n that is unknown. Substitution of the data into the formula gives: $u_n = 0.25 \times 3^{n-1} = 44286.75$

The equation that results can be solved using a calculator or logarithms (see Sec. 1.2)



If using a Casio calculator, select module A-Equation. Then choose F3-Solver. Enter the equation and press EXE, F6-SOLVE.



The analytic solution is: $3^{(n-1)} = \left(\frac{44286.75}{0.25}\right)$

$$= 177147$$

$$\log_{10}\left(3^{(n-1)}\right) = \log_{10}(177147)$$

$$(n-1)\log_{10}(3) = \log_{10}(177147)$$

$$n-1 = \frac{\log_{10}(177147)}{\log_{10}(3)}$$

$$n-1 = 11$$

$$n = 12$$

Example A.2.16

A car originally worth \$34 000 loses 15% of its value each year.

a Write a geometric sequence that gives the year-by-year value of the car.

b Find the value of the car after 6 years.

c When will the value of the car fall below \$10 000?

a If the car loses 15% of its value each year, its value will fall to 85% (100% – 15%) of its value in the previous year. This means that the common ratio is 0.85 (the fractional equivalent of 85%). Using the formula, the sequence is: $u_n = 34000 \times 0.85^{(n-1)}$, i.e. \$34000, \$28900, \$24565, \$20880.25, . . .

b The value after six years have passed is the seventh term of the sequence. This is because the first term of the sequence is the value after no years have passed. $u_7 = 34000 \times 0.85^6 \approx 12823$ or \$12 823.

c
$$10000 = 34000 \times 0.85^n$$

$$0.85^n = 0.2941$$

$$\log_{10}(0.85^n) = \log_{10}(0.2941)$$

$$n\log_{10}(0.85) = \log_{10}(0.2941)$$

$$n = \frac{\log_{10}(0.2941)}{\log_{10}(0.85)}$$

$$n \approx 7.53$$

The car's value will fall to \$10000 after about 7 years 6 months.

Example A.2.17

The number of people in a small country town increases by 2% per year. If the population at the start of 1970 was 12500, what was the population at the start of 2010?

A quantity can be increased by 2% by multiplying by 1.02. Note that this is different from finding 2% of a quantity which is done by multiplying by 0.02. The sequence is 12500, 12500 $\times 1.02$, 12500×1.02^2 etc. with $a = 12500$, $r = 1.02$.

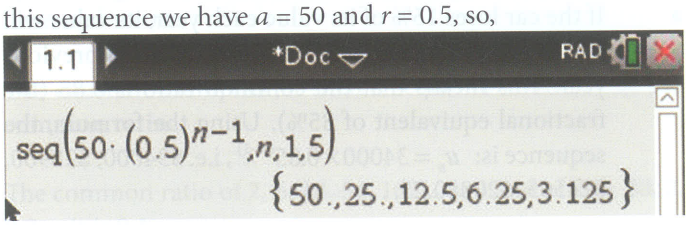
It is also necessary to be careful about which term is required. In this case, the population at the start of 1970 is the first term, the population at the start of 1971 the second term, and so on. The population at the start of 1980 is the eleventh term and at the start of 2010 we need the forty-first term:

$$u_{41} = 12500 \times 1.02^{40}$$
$$\approx 27600$$

In all such cases, you should round your answer to the level given in the question or, if no such direction is given, round the answer to a reasonable level of accuracy.

Using a graphics calculator

As with arithmetic sequences, geometric sequences such as 50, 25, 12.5, . . . can be listed using a graphics calculator. For this sequence we have $a = 50$ and $r = 0.5$, so,



Exercise A.2.4

1. Find the common ratio, the 5th term and the general term of the following sequences.

   a $3, 6, 12, 24, \ldots$ b $3, 1, \frac{1}{3}, \frac{1}{9}$

   c $2, \frac{2}{5}, \frac{2}{25}, \frac{2}{125}, \ldots$ d $-1, 4, -16, 64, \ldots$

   e $ab, a, \frac{a}{b}, \frac{a}{b^2}, \ldots$ f $a^2, ab, b^2, \ldots$

2. Find the value(s) of x if each of the following are in geometric sequence.

   a $3, x, 48$ b $\frac{5}{2}, x, \frac{1}{2}$

3. The third and seventh terms of a geometric sequence are 0.75 and 12 respectively.

   a Find the 10th term.

   b What term is equal to 3072?

4. A rubber ball is dropped from a height of 10 metres and bounces to reach $^5/_6$ of its previous height after each rebound. Let u_n be the ball's maximum height before its nth rebound.

   a Find an expression for u_n.

   b How high will the ball bounce after its 5th rebound.

   c How many times has the ball bounced by the time it reaches a maximum height of $^{6250}/_{1296}$ m.

5. The terms $k+4$, $5k+4$, $k+20$ are in a geometric sequence. Find the value(s) of k.

6. A computer depreciates each year to 80% of its value from the previous year. When bought the computer was worth $8000.

   a Find its value after: i 3 years ii 6 years.

   b How long does it take for the computer to depreciate to a quarter of its purchase price?

7. The sum of the first and third terms of a geometric sequence is 40 while the sum of its second and fourth terms is 96. Find the sixth term of the sequence.

8. The sum of three successive terms of a geometric sequence is $^{35}/_2$, while their product is 125. Find the three terms.

9. The population in a town of 40000 increases at 3% per annum. Estimate the town's population after 10 years.

10. Following new government funding it is expected that the unemployed workforce will decrease by 1.2% per month. Initially there are 120000 people unemployed. How large an unemployed workforce can the government expect to report 8 months later.

11. The cost of erecting the ground floor of a building is $44000, for erecting the first floor it costs $46200, to erect the second floor costs $48510 and so on. Using this cost structure, how much will it cost to erect the 5th floor? What will be the total cost of erecting a building with six floors?

Geometric Series

Wall Street in New York - where interest becomes a geometric series

When the terms of a geometric sequence are added, the result is a geometric series.

For example:

The sequence 3, 6, 12, 24, 48, . . . gives rise to the series: 3 + 6 + 12 + 24 + 48 + . . .

and, the sequence 24, –16, $10^2/_3$, $-7^1/_9$... leads to the series 24 –16 + $10^2/_3$ $-7^1/_9$+...

Geometric series can be summed using the formula that is derived by first multiplying the series by r:

$$S_n = a + ar + ar^2 + ar^3 ... + ar^{n-3} + ar^{n-2} + ar^{n-1}$$
$$rS_n = \quad ar + ar^2 + ar^3 ... + ar^{n-3} + ar^{n-2} + ar^{n-1} + ar^n$$

Subtracting the second equation from the first:

$$S_n - rS_n = a - ar^n$$
$$S_n(1-r) = a(1-r^n)$$
$$S_n = \frac{a(1-r^n)}{1-r}$$

This formula can also be written as: $S_n = \frac{a(r^n-1)}{r-1}, r \neq 1$.

It is usual to use the version of the formula that gives a positive value for the denominator. And so, we have:

The sum of the first n terms of a geometric series, S_n, where $r \neq 1$ is given by:

$$S_n = \frac{a(1-r^n)}{1-r}, |r| < 1 \text{ or } S_n = \frac{a(r^n-1)}{r-1}, |r| > 1$$

Example A.2.18

Sum the following series to the number of terms indicated.

a $2 + 4 + 8 + 16 + ...$ 9 terms

b $5 - 15 + 45 - 135 + ...$ 7 terms

c $24 + 18 + 27 + {}^{27}/_2 + {}^{81}/_8$ 12 terms

d $20 - 30 + 45 - 67.5 + ...$ 10 terms

a In this case $a = 2$, $r = 2$ and $n = 9$.

Because $r = 2$ it is convenient to use: $S_n = \frac{a(r^n-1)}{r-1}$

$$S_9 = \frac{2(2^9-1)}{2-1} = 1022$$

Using this version of the formula gives positive values for the numerator and denominator. The other version is correct but gives negative numerator and denominator and hence the same answer.

b $a = 5$, $r = -3$ and $n = 7$.

$$S_n = \frac{a(1-r^n)}{1-r} \quad \text{or} \quad S_n = \frac{a(r^n-1)}{r-1}$$

$$S_7 = \frac{5(1-(-3)^7)}{1-(-3)} = 2735 \qquad S_7 = \frac{5((-3)^7-1)}{(-3)-1} = 2735$$

c $a = 24$, $r = 0.75$ and $n = 12$.

$$S_n = \frac{a(1-r^n)}{1-r}$$

$$S_{12} = \frac{24\left(1-\left(\frac{3}{4}\right)^{12}\right)}{1-\left(\frac{3}{4}\right)} = 92.95907$$

This version gives the positive values.

d $a = 20$, $r = -1.5$ and $n = 10$.

$$S_n = \frac{a(1-r^n)}{1-r}$$

$$S_{10} = \frac{20(1-(-1.5)^{10})}{1-(-1.5)} = -453.32031$$

When using a calculator to evaluate such expressions, it is advisable to use brackets to ensure that correct answers are obtained. For both the graphic and scientific calculators, the negative common ratio must be entered using the +/– or (–) key.

Other questions that may be asked in examinations could involve using both formulae. A second possibility is that you may be asked to apply sequence and series theory to some simple problems.

Example A.2.19

The second term of a geometric series is –30 and the sum of the first two terms is –15. Find the first term and the common ratio.

From the given information we have:

$u_2 = -30 \therefore ar = -30$ – (1)

$S_2 = -15 \therefore \dfrac{a(r^2-1)}{r-1} = -15$ – (2)

The result is a pair of simultaneous equations in the two unknowns. The best method of solution is substitution.

From (1): $a = \dfrac{-30}{r}$.

Substituting into (2): $\dfrac{\frac{-30}{r}(r^2-1)}{r-1} = -15$

$$\frac{-30(r^2-1)}{r(r-1)} = -15$$

$$\frac{-30(r+1)(r-1)}{r(r-1)} = -15$$

$$-30(r+1) = 15r$$

$$-30r - 30 = -15r$$

$$r = -2$$

$$\therefore a = \frac{-30}{r} = \frac{-30}{-2} = 15$$

The series is 15 – 30 + 60 – 120 + 240 – ... which meets the conditions set out in the question.

Example A.2.20

A family decide to save some money in an account that pays 9% annual compound interest calculated at the end of each year. They put \$2500 into the account at the beginning of each year. All interest is added to the account and no withdrawals are made. How much money will they have in the account on the day after they have made their tenth payment?

The problem is best looked at from the last payment of \$2500 which has just been made and which has not earned any interest.

The previous payment has earned one lot of 9% interest and so is now worth 2500×1.09.

The previous payment has earned two years' worth of compound interest and is worth 2500×1.09^2.

This process can be continued for all the other payments and the various amounts of interest that each has earned. They form a geometric series:

Last payment   First payment

$2500 + 2500 \times 1.09 + 2500 \times 1.09^2 + \ \ldots \ + \ldots + 2500 \times 1.09^9.$

The total amount saved can be calculated using the series formula:

$$S_n = \frac{a(r^n - 1)}{r-1}$$

$$S_{10} = \frac{2500(1.09^{10}-1)}{1.09-1}$$

$$= 37982.32$$

The family will save about \$37 982.

Exercise A.2.5

1. Find the common ratios of these geometric sequences:

   a 7, 21, 63, 189, ... b 12, 4, $^4/_3$, $^4/_9$, ...

   c 1, –1, 1, –1, 1, ... d 9, –3, 1, $-^1/_3$, $^1/_9$, ...

   e 64, 80, 100, 125, ... f 27, –18, 12, –8, ...

2. Find the term indicated for each of these geometric sequences.

   a 11, 33, 99, 297, ... 10th term.

b $1, 0.2, 0.04, 0.008, \ldots$ 5th term.

c $9, -6, 4, -{}^{8}/_{3}, \ldots$ 9th term.

d $21, 9, {}^{27}/_{7}, {}^{81}/_{49}, \ldots$ 6th term.

e $-{}^{1}/_{3}, -{}^{1}/_{4}, -{}^{3}/_{16}, -{}^{9}/_{64}, \ldots$ 6th term.

3. Find the number of terms in each of these geometric sequences and the sum of the numbers in each sequence:

a $4, 12, 36, \ldots, 236196$ b $11, -22, 44, \ldots, 704$

c $100, -10, 1, \ldots, -10^{-5}$ d $48, 36, 27, \ldots, \frac{6561}{1024}$

e $\frac{1}{8}, -\frac{9}{32}, \frac{81}{128}, \ldots \frac{6561}{2048}$ f $100, 10, 1, \ldots, 10^{-10}$

4. Write the following in expanded form and evaluate.

a $\sum_{k=1}^{7}\left(\frac{1}{2}\right)^k$ b $\sum_{i=1}^{6}2^{i-4}$ c $\sum_{j=1}^{4}\left(\frac{2}{3}\right)^j$

d $\sum_{s=1}^{4}(-3)^s$ e $\sum_{t=1}^{6}2^{-t}$

5. The third term of a geometric sequence is 36 and the tenth term is 78 732. Find the first term in the sequence and the sum of these terms.

6. A bank account offers 9% interest compounded annually. If $750 is invested in this account, find the amount in the account at the end of the twelfth year.

7. When a ball is dropped onto a flat floor, it bounces to 65% of the height from which it was dropped. If the ball is dropped from 80 cm, find the height of the fifth bounce.

8. A computer loses 30% of its value each year.

a Write a formula for the value of the computer after n years.

b How many years will it be before the value of the computer falls below 10% of its original value?

9. A geometric sequence has a first term of 7 and a common ratio of 1.1. How many terms must be taken before the value of the term exceeds 1000?

10. A colony of algae increases in size by 15% per week. If 10 grams of the algae are placed in a lake, find the weight of algae that will be present in the lake after 12 weeks. The lake will be considered 'seriously polluted' when there is in excess of 10 000 grams of algae in the lake. How long will it be before the lake becomes seriously polluted?

11. A geometric series has nine terms, a common ratio of 2 and a sum of 3577. Find the first term.

12. A geometric series has a third term of 12, a common ratio of $-{}^{1}/_{2}$ and a sum of $32{}^{1}/_{16}$. Find the number of terms in the series.

13. A geometric series has a first term of 1000, seven terms and a sum of $671{}^{7}/_{8}$. Find the common ratio.

14. A geometric series has a third term of 300, and a sixth term of 37500. Find the common ratio and the sum of the first fourteen terms (in scientific form correct to two significant figures).

15. A $10 000 loan is offered on the following terms: 12% annual interest on the outstanding debt calculated monthly. The required monthly repayment is $270. How much will still be owing after nine months.

16. As a prize for inventing the game of chess, its originator is said to have asked for one grain of wheat to be placed on the first square of the board, 2 on the second, 4 on the third, 8 on the fourth and so on until each of the 64 squares had been covered. How much wheat would have been the prize?

Combined arithmetic and geometric sequences and series

There will be occasions on which questions will be asked that relate to both arithmetic and geometric sequences and series.

Example A.2.21

A geometric sequence has the same first term as an arithmetic sequence. The third term of the geometric sequence is the same as the tenth term of the arithmetic sequence with both being 48. The tenth term of the arithmetic sequence is four times the second term of the geometric sequence. Find the common difference of the arithmetic sequence and the common ratio of the geometric sequence.

When solving these sorts of questions, write the data as equations, noting that a is the same for both sequences. Let u_n denote the general term of the arithmetic sequence and v_n the general term of the geometric sequence.

We then have:

$u_{10} = a + 9d$ $v_3 = ar^2 = 48$

i.e. $a + 9d = ar^2 = 48$ – (1)

$u_{10} = 4v_2 \Rightarrow a + 9d = 4ar$ – (2)

(1) represents the information 'The third term of the geometric sequence is the same as the tenth term of the arithmetic sequence with both being 48'.

(2) represents 'The tenth term of the arithmetic sequence is four times the second term of the geometric sequence'.

There are three equations here and more than one way of solving them. One of the simplest is:

From (1) $a + 9d = 48$ and so substituting into (2):

$48 = 4ar \Leftrightarrow ar = 12$ – (3)

Also from (1) we have: $ar^2 = 48 \Leftrightarrow (ar)r = 48$ – (4)

Substituting (3) into (4): $12r = 48 \Leftrightarrow r = 4$

Substituting result into (1): $a \times 16 = 48$ $a = 3$

Substituting result into (1): $3 + 9d = 48 \Leftrightarrow d = 5$

The common ratio is 4 and the common difference is 5.

It is worth checking that the sequences are as specified:

Geometric sequence: 3, 12, 48

Arithmetic sequence: 3, 8, 13, 18, 23, 28, 33, 38, 43, 48

Exercise A.2.6

1. Consider the following sequences:

Arithmetic: 100, 110, 120, 130, . . .

Geometric: 1, 2, 4, 8, 16, . . .

Prove that:

a The terms of the geometric sequence will exceed the terms of the arithmetic sequence after the 8th term.

b The sum of the terms of the geometric sequence will exceed the sum of the terms of the arithmetic after the 10th term.

2. An arithmetic series has a first term of 2 and a fifth term of 30. A geometric series has a common ratio of –0.5. The sum of the first two terms of the geometric series is the same as the second term of the arithmetic series. What is the first term of the geometric series?

3. An arithmetic series has a first term of –4 and a common difference of 1. A geometric series has a first term of 8 and a common ratio of 0.5. After how many terms does the sum of the arithmetic series exceed the sum of the geometric series?

4. The first and second terms of an arithmetic and a geometric series are the same and are equal to 12. The sum of the first two terms of the arithmetic series is four times the first term of the geometric series. Find the first term of each series, if the A.P. has $d = 4$.

5. Bo-Youn and Ken are to begin a savings program. Bo-Youn saves \$1 in the first week \$2 in the second week, \$4 in the third and so on, in geometric progression. Ken saves \$10 in the first week, \$15 in the second week, \$20 in the third and so on, in arithmetic progression. After how many weeks will Bo-Youn have saved more than Ken?

6. Ari and Chai begin a training program. In the first week Chai will run 10km, in the second he will run 11km and in the third 12km, and so on, in arithmetic progression. Ari will run 5km in the first week and will increase his distance by 20% in each succeeding week.

a When does Ari's weekly distance first exceed Chai's?

b When does Ari's total distance first exceed Chai's?

7. The Fibonacci sequence: 1, 1, 2, 3, 5, 8, 13, 21, . . . in which each term is the sum of the previous two terms is neither arithmetic nor geometric. However, after the eighth term (21) the sequence becomes approximately geometric. If we assume that the sequence is geometric:

a What is the common ratio of the sequence (to four significant figures)?

b Assuming that the Fibonacci sequence can be approximated by the geometric sequence after the eighth term, what is the approximate sum of the first 24 terms of the Fibonacci sequence?

Convergent Series

If a geometric series has a common ratio between −1 and 1, the terms get smaller and smaller as n increases.

The sum of these terms is still given by the formula

$$S_n = \frac{a\left(1-r^n\right)}{1-r}, r \neq 1$$

For $-1 < r < 1$, $r^n \to 0$ as $n \to \infty$, $S_n \to \dfrac{a}{1-r}$

If $|r| < 1$, the infinite sequence has a sum given by: $S_\infty = \dfrac{a}{1-r}$

This means that if the common ratio of a geometric series is between −1 and 1, the sum of the series will approach a value of $\dfrac{a}{1-r}$ as the number of terms of the series becomes large, i.e. the series is convergent.

Example A.2.22

Find the sum to infinity of the series:

a 16 + 8 + 4 + 2 + 1 + ...

b $9 - 6 + 4 - {}^{8}/_{3} + {}^{16}/_{9} + \ldots$

a 16 + 8 + 4 + 2 + 1 + ...

In this case: a $a = 16, r = \dfrac{1}{2} \Rightarrow S_\infty = \dfrac{a}{1-r} = \dfrac{16}{1-\dfrac{1}{2}} = 32$

b $9 - 6 + 4 - {}^{8}/_{3} + {}^{16}/_{9}$

$$a = 9, r = -\frac{2}{3} \Rightarrow S_\infty = \frac{a}{1-r} = \frac{9}{1-\left(-\dfrac{2}{3}\right)} = 5.4$$

There are many applications for convergent geometric series. The following examples illustrate two of these.

Example A.2.23

Use an infinite series to express the recurring decimal $0.4\dot{6}\dot{2}$ as a rational number.

$0.4\dot{6}\dot{2}$ can be expressed as the series:

$$0.462 + 0.000462 + 0.000000462 + \ldots$$

or $\dfrac{462}{1000} + \dfrac{462}{1000000} + \dfrac{462}{1000000000} + \ldots$

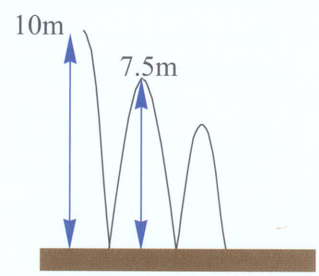
This is a geometric series with: $a = \dfrac{462}{1000}, r = \dfrac{1}{1000}$.

It follows that: $S_\infty = \dfrac{a}{1-r} = \dfrac{\dfrac{462}{1000}}{1-\dfrac{1}{1000}} = \dfrac{\dfrac{462}{1000}}{\dfrac{999}{1000}} = \dfrac{462}{999}$

Example A.2.24

A ball is dropped from a height of 10 metres. On each bounce the ball bounces to three quarters of the height of the previous bounce. Find the distance travelled by the ball before it comes to rest (if it does not move sideways).

The ball bounces in a vertical line and does not move sideways. On each bounce after the drop, the ball moves both up and down and so travels twice the distance of the height of the bounce.



Distance = $10 + 15 + 15 \times \dfrac{3}{4} + 15 \times \left(\dfrac{3}{4}\right)^2 + \ldots$

All terms, except the first, are geometric.

Distance $10 + S_\infty = 10 + \dfrac{15}{1-\dfrac{3}{4}} = 70$ m

Exercise A.2.7

1. Evaluate:

   a $27 + 9 + 3 + \dfrac{1}{3} + \ldots$

   b $1 - \dfrac{3}{10} + \dfrac{9}{100} - \dfrac{27}{1000} + \ldots$

   c 500 + 450 + 405 + 364.5 + …

   d 3 – 0.3 + 0.03 – 0.003 + 0.0003 –…

2. Use geometric series to express the recurring decimal 23.232323… as a mixed number.

3. Biologists estimate that there are 1000 trout in a lake. If none are caught, the population will increase at 10% per year. If more than 10% are caught, the population will fall. As an approximation, assume that if 25% of the fish are caught per year, the population will fall by 15% per year. Estimate the total catch before the lake is 'fished out'. If the catch rate is reduced to 15%, what is the total catch in this case? Comment on these results.

4. Find the sum to infinity of the sequence 45, –30, 20, ...

5. The second term of a geometric sequence is 12 while the sum to infinity is 64. Find the first three terms of this sequence.

6. Express the following as rational numbers:

   a $0.3\dot{6}$ b $0.\dot{3}\dot{7}$ c $2.1\dot{2}$

7. A swinging pendulum covers 32 centimetres in its first swing, 24 cm on its second swing, 18 cm on its third swing and so on. What is the total distance this pendulum swings before coming to rest?

8. The sum to infinity of a geometric sequence is $^{27}/_{2}$ while the sum of the first three terms is 13. Find the sum of the first 5 terms.

9. Find the sum to infinity of the sequence: $1+\sqrt{3},1,\frac{1}{\sqrt{3}+1},...$.

10. a Find: i $\sum_{i=0}^{n}(-t)^i, |t|<1$, ii $\sum_{i=0}^{\infty}(-t)^i, |t|<1$.

    b i Hence, show that ,

    $$\ln(1+x)=x-\frac{1}{2}x^2+\frac{1}{3}x^3-\frac{1}{4}x^4+..., |x|<1$$

    ii Using the above result, show that:

    $$\ln 2=1-\frac{1}{2}+\frac{1}{3}-\frac{1}{4}+...$$

11. a Find: i $\sum_{i=0}^{n}\left(-t^2\right)^i, |t|<1$, ii $\sum_{i=0}^{\infty}\left(-t^2\right)^i, |t|<1$

    i Hence, show that:

    $$\arctan x=x-\frac{1}{3}x^3+\frac{1}{5}x^5-\frac{1}{7}x^7+..., |x|<1$$

    Using the above result, show that $\frac{\pi}{4}=1-\frac{1}{3}+\frac{1}{5}-\frac{1}{7}+...$.

Exercise A.2.8

1. $2k + 2$, $5k + 1$ and $10k + 2$ are three successive terms of a geometric sequence. Find the value(s) of k.

2. Evaluate $\dfrac{1+2+3+...+10}{1+\frac{1}{2}+\frac{1}{4}+...+\frac{1}{512}}$.

3. Find a number which, when added to each of 2, 6 and 13, gives three numbers in geometric sequence.

4. Find the fractional equivalent of:

   a $2.3\dot{8}$ b $4.\dot{6}\dot{2}$ c 0.41717...

5. Find the sum of all integers between 200 and 400 that are divisible by 6.

6. Find the sum of the first 50 terms of an arithmetic progression given that the 15th term is 34 and the sum of the first 8 terms is 20.

7. Find the value of p so that $p + 5$, $4p + 3$ and $8p - 2$ will form successive terms of an arithmetic progression.

8. For the series defined by $S_n = 3n^2 - 11n$, find t_n and hence show that the sequence is arithmetic.

9. How many terms of the series $6 + 3 + {}^3/_2 + ...$ must be taken to give a sum of $11^{13}/_{16}$?

10. If $1 + 2x + 4x^2 + ... = {}^3/_4$, find the value of x.

11. Logs of wood are stacked in a pile so that there are 15 logs on the top row, 16 on the next row, 17 on the next and so on. If there are 246 logs in total,

    a how many rows are there?

    b how many logs are there in the bottom row?

12. The lengths of the sides of a right-angled triangle form the terms of an arithmetic sequence. If the hypotenuse is 15 cm in length, what is the length of the other two sides?

13. The sum of the first 8 terms of a geometric series is 17 times the sum of its first four terms. Find the common ratio.

14. Three numbers a, b and c whose sum is 15 are successive terms of a G.P, and b, a, c are successive terms of an A.P. Find a, b and c.

15. The sum of the first n terms of an arithmetic series is given by:

$$S_n = \frac{n(3n+1)}{2}.$$

a Calculate S_1 & S_2.

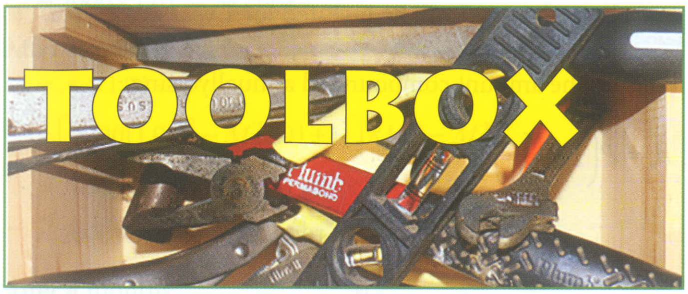
b Find the first three terms of this series.

c Find an expression for the nth term.

16. An ant walks along a straight path. After travelling 1 metre it stops, turns through an angle of 90° in an anticlockwise direction and sets off in a straight line covering a distance of half a metre. Again, the ant turns through an angle of 90° in an anticlockwise direction and sets off in a straight line covering a quarter of a metre. The ant continues in this manner indefinitely.

a How many turns will the ant have made after covering a distance of $^{63}/_{32}$ metres?

b How far will the ant eventually travel?

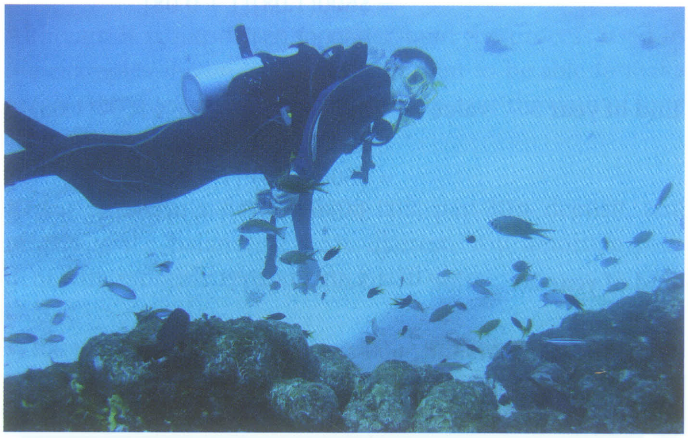


Sequences and series are used in a range of areas. The financial centres of all cities use series to plan loans.

Founded as a refuge from lawlessness, Venice was a great trade and financial centre in the 13th Century.

Financial District, Dubai, UAE.

This diver is taking extra nitrogen into his blood because of the elevated pressure of the air he is breathing - one extra atmosphere for every 10 metres of depth. This nitrogen is released when he surfaces. If he surfaces too rapidly, this may cause bubbles to form in his blood. This results in the painful and sometimes fatal condition known as 'the bends'.



To prevent this, he carries a 'dive computer' that records his depth as a time sequence and uses a mathematical model (a time series, as nitrogen is progressively added) to predict the nitrogen uptake and advise as to the safe ascent procedure.

Compound Interest

We have already come across some practical examples of the use of G.P.s in the area of finance. In this section we further develop these ideas and look at the area of compound interest and superannuation.

Example A.2.25

Find what $600 amounts to in 20 years if it is invested at 8% p.a. compounding annually.

End of year 1 value = \$600 + 8% × \$600
= \$600(1.08)

End of year 2 value = \$600(1.08) + 8% × \$600(1.08)
= \$600(1.08) + 0.08 × \$600(1.08)
= \$600(1.08)[1 + 0.08]
= $\$600(1.08)^2$

End of year 3 value = $\$600(1.08)^2 + 8\% \times \$600(1.08)^2$
= $\$600(1.08)^2 + 0.08 \times \$600(1.08)^2$
= $\$600(1.08)^2[1 + 0.08]$
= $\$600(1.08)^3$
⇩

End of year 20 value = $\$\,600(1.08)^{20}$

Thus, after 20 years the \$600 amounts to \$2796.57.

Looking closely at the terms of the sequence, they form a G.P.:

$$600(1.08),\ 600(1.08)^2,\ 600(1.08)^3,\ \ldots,\ 600(1.08)^{20}$$

where $a = 600$ and $r = 1.08$.

Developing a formula for compound interest

In general, if \$$P$ is invested at r% p.a. compound interest, it grows according to the sequence:

$$P\left(1+\frac{r}{100}\right), P\left(1+\frac{r}{100}\right)^2, P\left(1+\frac{r}{100}\right)^3, \ldots, P\left(1+\frac{r}{100}\right)^n$$

where $a = P\left(1+\frac{r}{100}\right)$ and $r = \left(1+\frac{r}{100}\right)$ so that

$$A_n = P\left(1+\frac{r}{100}\right)^n$$

where A_n is the amount after n time periods.

Superannuation

Superannuation is a common way in which working people attempt to provide for themselves in retirement. In many cases, workers save a fixed amount from each pay-packet into an interest bearing account.

Example A.2.26

A woman invests \$1000 at the beginning of each year in a superannuation scheme. If the interest is paid at the rate of 12% p.a. on the investment (compounding annually), how much will her investment be worth after 20 years?

t_1 = the 1st \$1000 will be invested for 20 years at 12% p.a.

t_2 = the 2nd \$1000 will be invested for 19 years at 12% p.a.

t_3 = the 3rd \$1000 will be invested for 18 years at 12% p.a.

⇩

t_{20} = the 20th \$1000 will be invested for 1 year at 12% p.a.

Finding the amount compounded annually using :

$A_n = P\left(1+\frac{r}{100}\right)^n$, we have:

$$t_1 = 1000\left(1+\frac{12}{100}\right)^{20} = 1000(1.12)^{20}$$

$$t_2 = 1000\left(1+\frac{12}{100}\right)^{19} = 1000(1.12)^{19}$$

$$t_3 = 1000\left(1+\frac{12}{100}\right)^{18} = 1000(1.12)^{18}$$

⇩

$$t_{20} = 1000\left(1+\frac{12}{100}\right)^{1} = 1000(1.12)^{1}$$

To find the total of her investment after 20 years, we need to add the separate amounts:

$$1000(1.12)^{20} + 1000(1.12)^{19} + 1000(1.12)^{18} + \ldots 1000(1.12)^{1}$$

= \$80 698.74

Thus her total investment amounts to \$80 698.74

Example A.2.27

Linda borrows $2000 at 1% per month reducible interest. If she repays the loan in equal monthly instalments over 4 years, how much is each instalment?

Amount borrowed = $2000, r = 1% per month = 0.01 and $n = 4 \times 12 = 48$ months.

Let the monthly instalment be = $ M and the amount owing after n months = $$A_n$.

Our aim is to find $$M$ i.e. the amount of each instalment.

After 1 month (after paying the 1st instalment), we have:

A_1 = 2000 + interest – M = 2000 + 2000 × 0.01 – M

After 2 months,

$$\begin{aligned} A_2 &= A_1 \times 1.01 - M \\ &= [2000(1.01) - M] \times 1.01 - M \\ &= 2000 \times 1.01^2 - 1.01 \times M - M \\ &= 2000 \times 1.01^2 - M(1.01+1) \end{aligned}$$

After 3 months,

$$\begin{aligned} A_3 &= A_2 \times 1.01 - M \\ &= \left[2000(1.01)^2 - M(1.01+1)\right] \times 1.01 - M \\ &= 2000 \times 1.01^3 - M(1.01+1) \times 1.01 - M \\ &= 2000 \times 1.01^3 - M\left[1.01^2 + 1.01 + 1\right] \end{aligned}$$

After 4 months,

$$\begin{aligned} A_4 &= A_3 \times 1.01 - M \\ &= \left[2000(1.01)^3 - M\left(1.01^2 + 1.01 + 1\right)\right] \times 1.01 - M \\ &= 2000 \times 1.01^4 - M\left(1.01^3 + 1.01^2 + 1.01\right) \times 1.01 - M \\ &= 2000 \times 1.01^4 - M\left[1.01^3 + 1.01^2 + 1.01 + 1\right] \end{aligned}$$

After n months, we then have

$A_n = 2000(1.01)^n - M[1+1.01+1.01^2+1.01^3 \ldots 1.01^{n-1}]$

thus, $A_{48} = 2000(1.01)^{48} - M[1+1.01+1.01^2+1.01^3 \ldots 1.01^{47}]$

Now, the loan is repaid after 48 months, meaning that $A_{48} = 0$, therefore, solving for M, we have

$2000(1.01)^{48} - M[1+1.01+1.01^2+1.01^3 \ldots 1.01^{47} = 0$

$2000(1.01)^{48} = M[1+1.01+1.01^2+1.01^3 \ldots 1.01^{47}$

$$M = \frac{2000 \times 1.01^{48}}{1+1.01+1.01^2+1.01^3+\ldots+1.01^{47}}$$

The denominator is a G.P. with $a = 1$, $r = 1.01$ and $n = 48$, so that

$$\begin{aligned} 1+1.01+1.01^2+1.01^3+\ldots+1.01^{47} &= S_{48} \\ &= \frac{1\left(1-1.01^{48}\right)}{1-1.01} \\ &= 61.22261 \end{aligned}$$

$$\text{Therefore, } M = \frac{2000(1.01)^{48}}{61.22261} = 52.67$$

That is, each instalment must be $52.67.

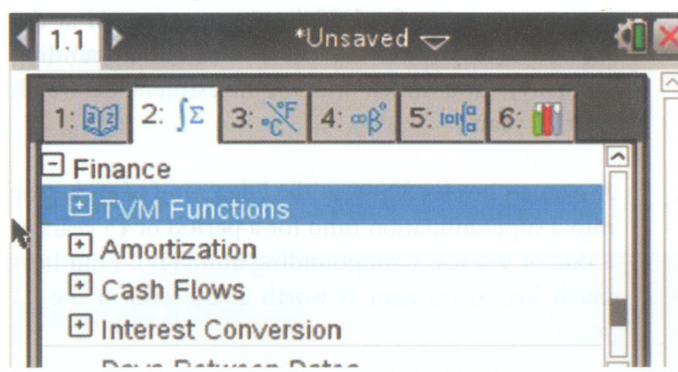
The total paid = 52.67 × 48 = 2528.16 so that the interest paid = 2528.16 – 2000 = 528.05

That is, she ends up paying $528.05 in interest.

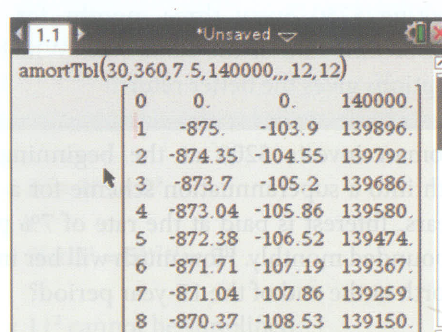
Although it is important to understand the process used in the examples shown, it is also important to be able to make use of technology. Calculators can help ease the pain of long calculations.

If you purchase a flat for $200 000, pay 30% deposit, and mortgage the balance at 7.5% interest. You amortize your debt with monthly repayments for 30 years.

What is the repayment?



and the data given:



In the first month, the payment will be $875 interest + $103.90 principal making a total of (to the nearest $), $979.

If using Casio, select the Financial Module (C), F4 -Amortization.

Exercise A.2.9

1. To how much will $1000 grow to if it is invested at 12% p.a. for 9 years, compounding annually?

2. A bank advertises an annual interest rate of 13.5% p.a. but adds interest to the account monthly, giving a monthly interest rate of 1.125%. Scott deposits $3500 with the bank. How much will be in the account in 20 months time?

3. To what amount will $900 grow to if it is invested at 10% p.a. for 7 years, compounding every 6 months?

4. A man borrows $5000 at 18% p.a. over a period of 5 years, with the interest compounding every month. Find to the nearest dollar the amount owing after 5 years.

5. Find the total amount required to pay off a loan of $20,000 plus interest at the end of 5 years if the interest is compounded half yearly and the rate is 12%.

6. A man invests $500 at the beginning of each year in a superannuation fund. If the interest is paid at the rate of 12% p.a. on the investment (compounding annually), how much will his investment be worth after 20 years?

7. A woman invests $2000 at the beginning of each year into a superannuation fund for a period of 15 years at a rate of 9% p.a. (compounding annually). Find how much her investment is worth at the end of the 15 years.

8. A man deposits $3000 annually to accumulate at 9% p.a. compound interest. How much will he have to his credit at the end of 25 years? Compare this to depositing $750 every three months for the same length of time and at the same rate. Which of these two options gives the better return?

9. A woman invests $200 at the beginning of each month into a superannuation scheme for a period of 15 years. Interest is paid at the rate of 7% p.a. and is compounded monthly. How much will her investment be worth at the end of the 15-year period?

10. Peter borrows $5000 at 1.5% per month reducible interest. If he repays the loan in equal monthly instalments over 8 years, how much is each instalment, and what is the total interest charged on the loan? Compare this to taking the same loan, but at a rate of 15% p.a. flat rate.

11. Kevin borrows $7500 to be paid back at 12.5% p.a. monthly reducible over a period of 7 years. What is the amount of each monthly instalment and what is the total interest charged on the loan. Find the equivalent flat rate of interest.

12. Find the possible values of x if $x + 1$, $3x + 2$, and $2x^2$ are three consecutive terms of an arithmetic sequence.

13. Find k, given that $\sum_{i=1}^{k}(4i-29)=45$.

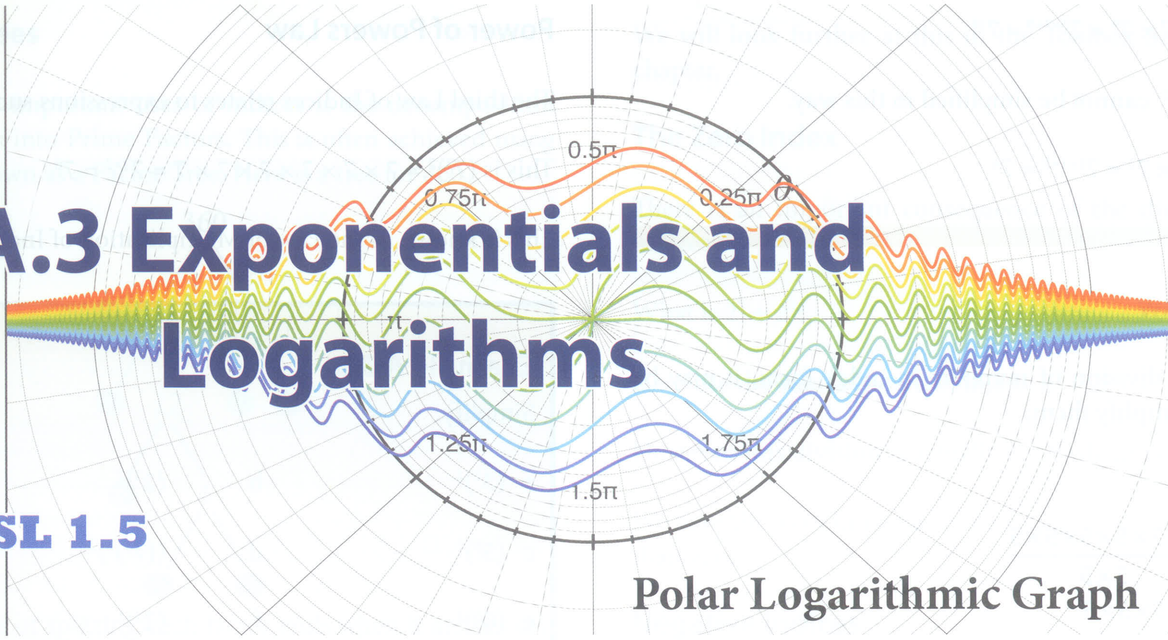
14. Show that $\sum_{k=3}^{8}(2k-2)=\sum_{k=1}^{6}(3(k+2)-1)$.

15. Given four consecutive terms in a progression, 4, m, n, 49. Find the possible values of m and n, if the first three terms form an arithmetic sequence and the last three terms form a geometric sequence.

Answers

A.3 Exponentials and Logarithms

SL 1.5



Polar Logarithmic Graph

Exponents

Basic rules of indices

We start by looking at the notation involved when dealing with indices (or exponents).

The expression:

$$a \times a \times a \times a \times a \times \ldots \times a$$

$$\longleftarrow \quad n \text{ times} \quad \longrightarrow$$

can be written in index form, a^n, where n is the index (or power or exponent) and a is the base.

This section deals with integer exponents and the rules of arithmetic that apply when using them.

A number in index form looks like this: 2^7. The 2 is known as the **base** and the 7 is called the **index**.

2^7 is read as 'two to the power 7'. It means two multiplied by itself seven times:

$$2^7 = 2 \times 2 \times 2 \times 2 \times 2 \times 2 \times 2 \ (= 128)$$

Multiplication Law

If we have numbers in index form, it may be possible to simplify them:

$$2^3 \times 2^6 = 2 \times 2 \times 2 \times 2 \times 2 \times 2 \times 2 \times 2 \times 2 = 2^{3+6} = 2^9$$

This is known as the Law of Addition of Indices.

The most common errors in using this rule are:

Applying it to mixed bases:

$$2^3 \times 5^6$$

This cannot be simplified using this rule.

Applying it to operations other than multiplication:

$$2^4 + 2^6$$

Also cannot be simplified using this rule.

Example A.3.1

Simplify:

a	$2^3 \times 2^2$	b	$5^3 \times 5^7 \times 5^1$
c	$5^3 \times 11^2$	d	$7^2 \times 7^7 \times 7^9$
e	$2^3 - 2^2$	f	$7 \times 7 \times 7^2$

a $\quad 2^3 \times 2^2 = 2^{3+2} = 2^5$

b $\quad 5^3 \times 5^7 \times 5^1 = 5^{3+7+1} = 5^{11}$

c $\quad 5^3 \times 11^2$ cannot be simplified.

d $7^2 \times 7^7 \times 7^9 = 7^{2+7+9} = 7^{18}$

e $2^3 - 2^2$ cannot be simplified in this way.

f $7 \times 7 \times 7^2 = 7^{1+1+2} = 7^4$

Division Law

If we have a division of numbers in index form, it may be possible to simplify them:

$$5^6 \div 5^2 = \frac{5^6}{5^2}$$
$$= \frac{5\times5\times5\times5\times5\times5}{5\times5}$$
$$= 5\times5\times5\times5$$
$$= 5^{6-2}$$
$$= 5^4$$

The red lines indicate a process known as 'cancellation'. The 5s are removed in pairs because 5 ÷ 5 = 1 and multiplication by one has no effect on the answer. The 5s can be cancelled in pairs. The indices are thus subtracted.

This is known as the Law of Subtraction of Indices. The same warnings as before about misapplying this rule should be borne in mind.

Example A.3.2

Simplify:

a	$2^3 \div 2^2$	b	$5^{14} \div 5^7$
c	$5^3 - 11^2$	d	$7^2 \div 3^7$
e	$2^3 \div 2$	f	$7^3 \times 7^5 \div 7^4$

a $2^3 \div 2^2 = 2^{3-2} = 2^1 = 2$

b $5^{14} \div 5^7 = 5^{14-7} = 5^7$

c $5^3 - 11^2$ cannot be simplified.

d $7^2 \div 3^7$ cannot be simplified.

e $2^3 \div 2 = 2^3 \div 2^1 = 2^{3-1} = 2^2$

f $7^3 \times 7^5 \div 7^4 = 7^{3+5-4} = 7^4$

Power of Powers Law

The third Law of Indices relates to expressions such as $(7^3)^2$.

This is $(7^3)^2 = 7 \times 7 \times 7 \times 7 \times 7 \times 7 = 7^{3\times2} = 7^6$

This is known as the Law of Multiplication of Indices.

Example A.3.3

Simplify:

a	$(3^3)^2$	b	$(2^3)^2$
c	$(5^6)^2$	d	$((7^3)^5)^2$
e	$(0^3)^2$	f	$(2^3)^1$

a $(3^3)^2 = 3^{3\times2} = 3^6$

b $(2^3)^2 = 2^{3\times2} = 2^6$

c $(5^6)^2 = 5^{6\times2} = 5^{12}$

d $((7^3)^5)^2 = 7^{3\times5\times2} = 7^{30}$

e $(0^3)^2 = 0$

f $(2^3)^1 = 2^3$

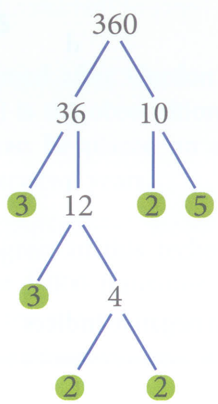
Exercise A.3.1

Simplify where possible:

1. $2^4 \times 2^3 \times 2^2$
2. $5^7 \times 5^3$
3. $2^4 \times 2^3 \div 2^2$
4. $7^6 \times 7^2 \div 7^5$
5. $((3^3)^2)^2 \div 3^3$
6. $((11^4)^3)^2 \div ((11^3)^2)^2$
7. $(3^2)^2 \times (3^3)^2$
8. $(3^4)^2 \times (3^3)^2 \div 3$
9. $(9^4)^3 \times (9^2)^2 \div 9^3$

Factor Trees

One of the important uses of indices is the decomposition of numbers into Prime Factors. This is often achieved using what is known as a Factor Tree. Here is a tree for 360.



The process of splitting each factor ends when it arrives at a prime number (green dots).

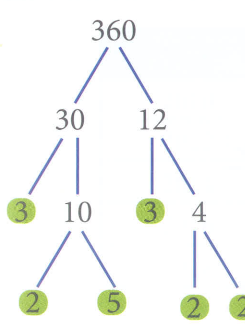
The numbers at the end of each branch are the prime factors of 360. They are (reading from left to right): 3, 3, 2, 2, 2, 5.

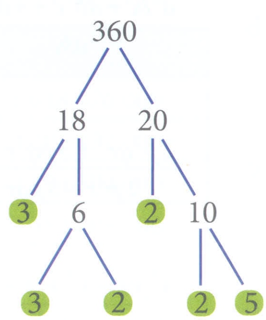
Because of the way the tree is constructed, it follows that:

$$360 = 3 \times 3 \times 2 \times 2 \times 2 \times 5 = 2^3 \times 3^2 \times 5.$$

It is an important fact of arithmetic, known as the **Fundamental Theorem of Arithmetic**, that this decomposition is unique for each number and does not depend on the way in which the tree is constructed.

For 360, which is rich in factors, there are several trees.





However, they all produce the same prime factors.

The prime factor decomposition of a number is like a fingerprint. Each one is unique to its number and no two different numbers have the same prime factors.

These are almost always given in index form. They are also indicative of the characteristics of the number. 360 is very rich in factors. This makes it a good choice for the number of degrees in a circle. 361 is entirely different. Can you see why?

We will look further at this in the Toolbox Section of this chapter.

The Zero Index

There is an important consequence of the division law of indices.

What is $2^3 \div 2^3$?

If we use the subtraction law, this is: $2^3 \div 2^3 = 2^{3-3} = 2^0$

However, this division is actually $8 \div 8$ which is, of course, 1.

We have this very important conclusion: $a^0 = 1$ for all a (other than zero).

Negative Indices

Another consequence of the subtraction law is to give meaning to negative indices.

$3^3 \div 3^5 = 3^{3-5} = 3^{-2}$ - but also:

$3^3 \div 3^5 = \dfrac{3\times3\times3}{3\times3\times3\times3\times3} = \dfrac{1}{3\times3} = \dfrac{1}{3^2}$ after cancelling.

A negative index is, therefore one over a positive index and vice versa.

$$a^{-n} = \frac{1}{a^n} \text{ and } \frac{1}{a^{-n}} = a^n$$

Example A.3.4

Simplify:

a $\dfrac{x^2 \times x^3}{x^7}$ b $\dfrac{(2x^2)^3 \times x}{4x \times x^9}$

c $\dfrac{3y^2}{(3y)^2}$ d $\dfrac{2xy^2}{(3x^2y)^3}$

a $\dfrac{x^2 \times x^3}{x^7} = x^{-2} = \dfrac{1}{x^2}$

b $\dfrac{(2x^2)^3 \times x}{4x \times x^9} = \dfrac{2^3 \times x^{2\times3} \times x}{4x^{1+9}} = \dfrac{8x^7}{4x^{10}} = 2x^{-3} = \dfrac{2}{x^3}$

c $\dfrac{3y^2}{(3y)^2}=\dfrac{3y^2}{3^2\times y^2}=\dfrac{1}{3}$

d $\dfrac{2xy^2}{(3x^2y)^3}=\dfrac{2xy^2}{3^3x^6y^3}=\dfrac{2}{27x^5y}=\dfrac{2}{27}x^{-5}y^{-1}$

This is also a moment to review factorisation involving exponents.

Example A.3.5

Factorise:

a 1568 – 112

b $(6ab)^3-27a^2b$

c $\dfrac{7a^2b+28ab^2}{a+4b}$

d $\dfrac{6x^3-12xy+18x}{3x^2y-6y^2+9y}$

a Using factor trees or otherwise:

$$\begin{aligned}1568-112 &= 2^5\times 7^2-2^4\times 7\\ &= 2^4\times 7(2\times 7-1)\\ &= 2^4\times 7\times 13\end{aligned}$$

There are many problems whose solutions are facilitated by being able to think of numbers as their prime factor decompositions in this way.

b
$$\begin{aligned}(6ab)^3-27a^2b&=2^3\times 3^3a^3b^3-3^3a^2b\\ &=3^3a^2b\left(2^3\times ab^2-1\right)\end{aligned}$$

c
$$\begin{aligned}\frac{7a^2b+28ab^2}{a+4b}&=\frac{7a^2b+4\times 7ab^2}{a+4b}\\ &=\frac{7ab(a+4b)}{a+4b}\\ &=7ab\end{aligned}$$

d
$$\begin{aligned}\frac{6x^3-12xy+18x}{3x^2y-6y^2+9y}&=\frac{6x(x^2-2y+3)}{3y(x^2-2y+3)}\\ &=\frac{6x}{3y}\\ &=\frac{2x}{y}\end{aligned}$$

Exercise A.3.2

1. Express as a product of prime numbers:

a 36 b 1 225

c 120 d 3 528

e 78 f 5 250

g 210 h 18 750

2. Express without negative indices

a $a^{-4}b^3$ b $\dfrac{2x^2y}{4x^{-3}y^{-2}}$

c $\dfrac{32x}{(2x^2)^2}$ d $\dfrac{x}{(2x^3y)^2}$

e $\dfrac{abc}{(2a)^3}$ f $\dfrac{12p^3q}{3(2pq)^3}$

3. Simplify:

a $\dfrac{x^2y+xy^2+xy}{2x^3+2x^2+2x^2y}$

b $\dfrac{a^2bc+ab^2c++abc^2}{abc}$

c $\dfrac{2p^2qr^2+3pq^2r^2+pqr^2}{2p^2r+3pqr+pr}$

d $\dfrac{x^3yz-x^2y^2z+3x^2yz}{x^2y^2z-xy^3z+3xy^2z}$

4. A ream (500 sheets) of paper is 6 cm thick. If it were possible to fold one sheet 15 times, approximately how thick would it be?

5. A helium party balloon contains 5 litres of gas. Its volume is halved each day. Find its volume after 15 days.

Modelling

As hinted at in the last two questions, exponents feature prominently in mathematical modelling.

Example A.3.6

Moore's law (named after Gordon Moore, one of the founders of Intel) is the observation that the number of transistors that can be placed on an integrated circuit doubles about every two years.

Tabulate the progress of this technology starting from 1970, when about 1 000 transistors could be placed on one circuit, up to the present.

The early predictions are:

Year	Prediction
1970	1 000
1972	2 000
1974	4 000
1976	8 000
1978	16 000
1980	32 000
1982	64 000
1984	128 000
1986	256 000

At the other end, these become:

Year	Prediction
2010	1 048 576 000
2012	2 097 152 000
2014	4 194 304 000
2016	8 388 608 000
2018	16 777 216 000

This is not far from the true figure.

Exponential models also work well for situations in which quantities are decreasing.

The next example will look at a 'decay' model.

Example A.3.7

The half life of carbon 14 is about 5 730 years. Carbon 14 is an unstable isotope of carbon. Most of the carbon on Earth is carbon 12. Only about 1 in every 10^{12} carbon atoms is carbon 14 - it is rare!

The proportion of carbon 14 in the atmosphere, and hence in living plants, has remained approximately constant over the last 60 000 years. This is because it is created at a steady rate by natural processes in the upper atmosphere.

When a plant dies, the proportion of carbon 14 is fixed at the atmospheric proportion of $1:10^{12}$. This then halves every 5 730 years. Tabulate these proportions that could be expected for 25 000 years working in multiples of the half life.

Year	Proportion
0	1 in 10^{12}
5 730	0.5 in 10^{12} or 1 in 2×10^{12}
11 460	0.25 in 10^{12} or 1 in 4×10^{12}
17 190	0.125 in 10^{12} or 1 in 8×10^{12}
22 920	0.0625 in 10^{12} or 1 in 1.6×10^{13}
28 650	0.03125 in 10^{12} or 1 in 3.2×10^{13}

The proportion of carbon 12 is measured by taking accurate measurements of the radioactive decay of the sample, not by counting atoms.

You will also notice that this table only gives values at times that are very widely spread. The full story will have to wait until (and if) you learn about fractional indices.

The Case of the Generous Banker

This modelling exercise requires you to believe that there are banks that will offer 100% annual interest on deposits.

So, we have found our generous banker.

What will happen if we invest $1 for 1 year? Well, our money will double to $2. A 100% interest rate implies multiplying the principal (original investment) by 2.

But, we can do better than that without actually asking the bank to increase its interest rate. Suppose we ask the bank to calculate the interest six monthly?

We obviously cannot expect two lots of 100% interest. Two lots of 50% interest would seem to be fair. It seems that this will produce the same result. However, it does not.

To increase an amount by 50%, we multiply by $1+\frac{50}{100}=1.5$.

But this is done twice so the new amount is: $1 \times 1.5^2 = \$2.25$.

This looks so good, we ask the banker if we can have 10 lots of 10% interest, resulting in a return of:

$$1\times\left(1+\frac{10}{100}\right)^{10}=1.1^{10}\approx 2.5937$$

This implies a return of \$2.59. Even better!

Can this process go on for ever?

Here are some more results:

Number of periods	Factor	Amount
1	2	\$2.00
10	$1 \times 1.1^{10} \approx 2.59374246$	\$2.59
100	$1 \times 1.01^{100} \approx 2.704813829$	\$2.70
1000	$1 \times 1.001^{1000} \approx 2.716923932$	\$2.72

Whilst this appears to be the 'gift that goes on giving', it seems that the gifts of extra money, whilst generous at first, are beginning to slow down.

When we get to 1 000 000 periods, the factor is:

$1 \times 1.000001^{1000000} \approx 2.71828$

Use a calculator or spreadsheet to investigate this further and you will find that we cannot do much better than this.

It is true that every time we take more interest periods, the return gets larger. However, it never reaches, for example 2.71829.

This is an example of limiting behaviour that is considered in more detail elsewhere.

This number is the limit as n becomes very large of:

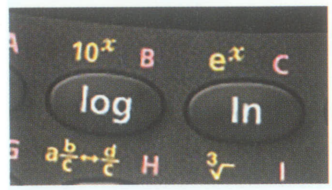
$$\left(1+\frac{1}{n}\right)^n$$

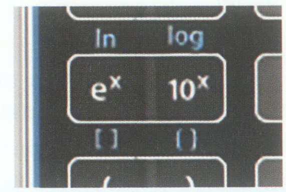
is known as 'Euler's Number' (after Leonhard Euler 1707-1783) and is designated e.

This number is the natural number of growth and decay, rather like π is the key number for circles.

The number is so important that it features on all scientific calculators as the function e^x. This is usually paired on the same button with $\ln x$. More of this later.

For base 10, the same pairing is almost always used with a key labelled $\log x/10^x$.





These two examples (Casio at left, TI-NSPIRE at right) have the functions reversed (ie. log is the principal function on Casio and the Alt function on TI.

You should know how to evaluate powers of 10 and e using you calculator.

Note that there is a difference between the subtraction key and the key used to enter negative numbers. Usually this is labelled (–). Also, some calculators only return a numerical answer if you press CTRL ENTER. Otherwise they return fractions, surds etc. Know your own model!

Here is a short exercise. It gives you a format for your answers. Do you know how to instruct your calculator to answer in the specified format?

Exercise A.3.3

1. Evaluate, giving your answer in scientific notation, correct to 3 decimal places:

   a 10^3 b 10^{-3}

   c 10^7 d 10^{-6}

   e $10^4 \times 10^{-4}$

2. Evaluate, giving your answer in scientific notation, correct to 3 decimal places:

   a e^3 b e^{-3}

   c e^7 d e^{-6}

   e $e^4 \times e^{-4}$

Modelling

The exponential function is extensively used in modelling growth and decay. Here are two examples which we will tackle with a calculator.

Example A.3.8

A population of bacteria is modelled by the expression:

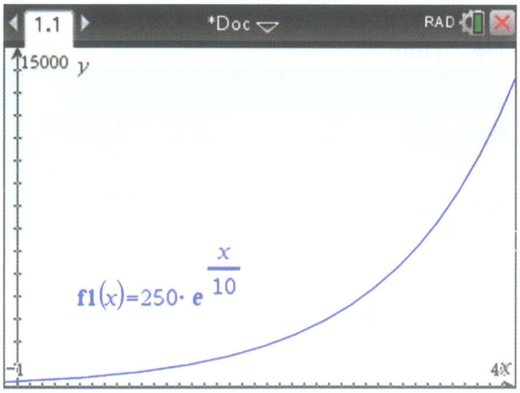
Population at time t is $250 \times e^{0.1 \times t}$.

Tabulate the population for $t = 0, 5, 10, ..., 40$.

Show this as a graph.

Using a calculator:

t	Population
0	250
5	412
10	680
15	1120
20	1847
25	3046
30	5021
35	8279
40	13650



Example A.3.9

Glass is made in furnaces at a temperature of 1 500°C.

When a glass statuette has been manufactured, it is important that it is cooled slowly. This is to prevent it cracking or otherwise distorting.

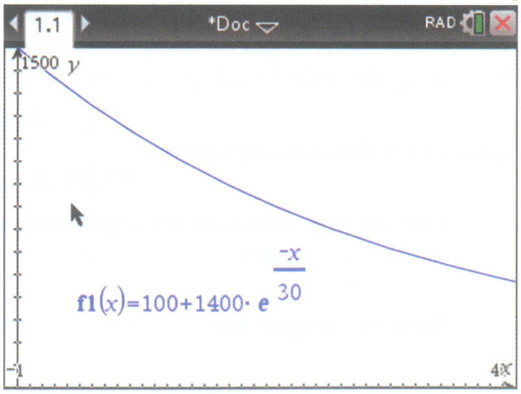
The temperature (T) at time t (hours) is modelled by:

$$T = 100 + 1400 \times e^{-t/30}.$$

Use a calculator to tabulate and graph these temperatures.

Using a calculator:

t	Temp (°C)
0	1500
5	1285
10	1103
15	949
20	819
25	708
30	615



Exercise A.3.4

1. A group of surgeon fish has populated a coral outcrop. At first the population grows rapidly.

   It is modelled by $P = 650 - 600e^{-t/10}$ where P is the population and t is the time in weeks.

   Graph this and explain its general shape.

2. The concentration C (mg/litre) of a drug in a patients blood t hours after it is taken is modelled by:

$$C = 12 \times e^{-t/30}$$

If the drug is only effective if present in concentrations greater than 8 mg/litre, how often should the drug be administered.

3. The annual sales S (\$ 000s) of a small business n years after its formation are modelled by:

$$S = 10 - 6e^{-n/10}$$

What does this model predict for the sales that the business might expect?

4. The effectiveness E (%) of an industrial catalyst t hours after installation is modelled by:

$$E = 100 \times e^{-t/400}$$

The catalyst is replaced when it reaches 50% effectiveness.

Use the graphic capabilities of your calculator to determine when the catalyst needs to be replaced.

5. Consider the function expression: $\dfrac{2}{1+3e^{-n}}$.

   a Tabulate the values of the expression for n = −5,−4, −3, ... 3, 4, 5.

   b Show this graphically.

   c Suggest some physical situations which expressions of this sort might model.

6. Consider the function expression: $2e^{-n^2}$

   a Tabulate the values of the expression for n = −5,−4, −3, ... 3, 4, 5.

   b Show this graphically.

   c Suggest some physical situations which expressions of this sort might model.

Logarithms

Many mathematical processes occur in pairs that are inverse to one another.

Doubling - Halving

Squaring - Square rooting

and so on.

The reverse of exponentiation (taking exponents) is taking logarithms.

This can be confusing so it is as well to have a specific case in mind to aid the memory.

$10^2 = 100$ is a statement in exponentials. In this statement, the 'question' contains a base (10 blue) and an exponent (^{2} green). The 'answer', 100, is in orange.

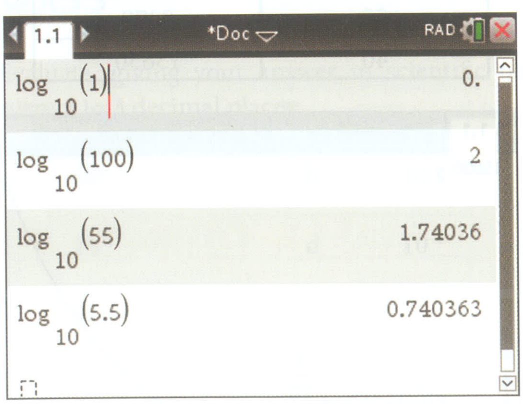
The corresponding statement in logarithms works in reverse with the 'question' becoming the 'answer' and vice-versa.

$$\log_{10} 100 = 2$$

The same numbers as appeared in the exponential statement appear in the logarithm.

At this stage, we will use a calculator to enumerate logarithms.

For example: $\log_{10}1$, $\log_{10}100$, $\log_{10}55$, $\log_{10}5.5$



There are several important results on this screen.

The first one: $\log_{10}1 = 0$ is a reflection of the fact that $10^0 = 1$.

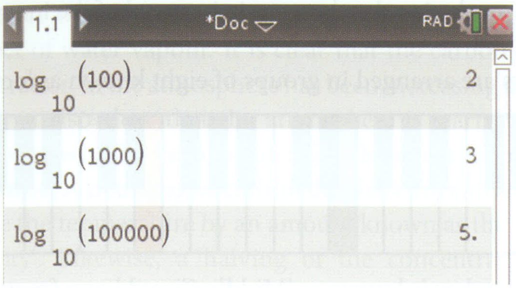
The second result is our original example.

The third and fourth results show an interesting property of logarithms: $55 = 10 \times 5.5$ and the logarithm of 55 is 1 bigger than the logarithm of 5.5.

This is a general property that results from the Law of Addition of Indices.

In indices: $10^2 \times 10^3 = 10^5$.

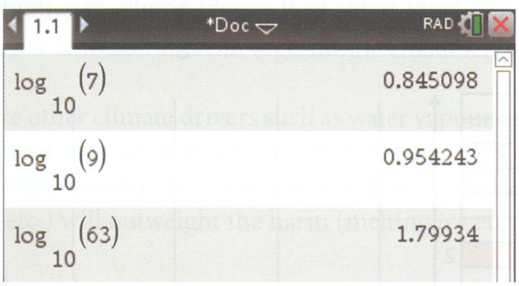
In logarithms:



Note that 2 + 3 = 5. Thus a multiplication problem in exponents becomes an addition problem in logarithms.

The principle works for numbers other than powers of 10.

For example: $7 \times 9 = 63$

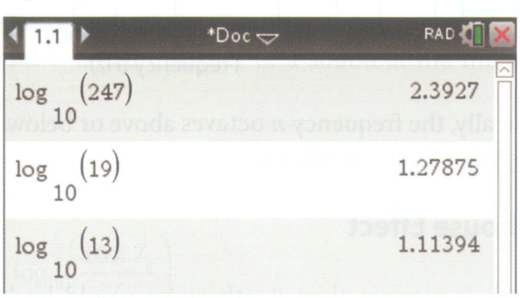


Note also that 0.845098 + 0.954243 = 1.799341 - to within calculator accuracy, this is the correct result:

$$\log_{10}7 + \log_{10}9 = \log_{10}63$$

The same principle works for division/subtraction.

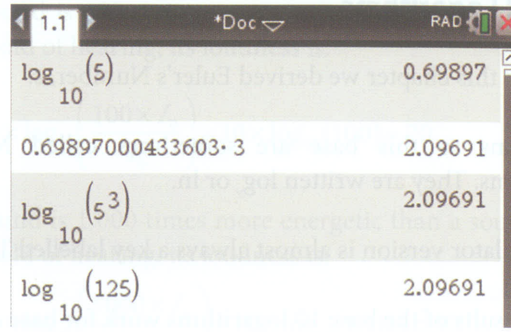
For example: $247 \div 19 = 13$



$$\log_{10}247 - \log_{10}19 = \log_{10}13$$

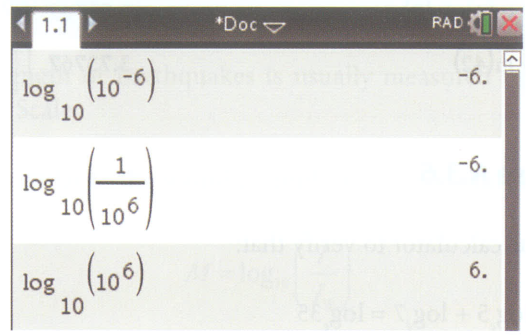
The third rule of exponents (power of powers) has its logarithmic version:

For example: $3 \times \log_{10}5 = \log_{10}5^3$.



There is also an implication related to negative indices.

Recall that: $10^{-6} = \frac{1}{10^6}$. In logarithms:



You should use your calculator to familiarise yourself with these rather strange properties.

In the days before the invention of calculators, logarithms were used to reduce multiplications and divisions to additions and subtractions. A great saving in time. The calculator was first marketed after the Apollo 11 Moon landings - not in the distant past.

Exercise A.3.5

Use your calculator to verify that:

1. $\log_{10}6 + \log_{10}9 = \log_{10}54$
2. $\log_{10}12 + \log_{10}5 = \log_{10}60$
3. $\log_{10}12 - \log_{10}2 = \log_{10}6$
4. $\log_{10}100 - \log_{10}5 = \log_{10}20$
5. $2\times\log_{10}7 = \log_{10}49$
6. $2\times\log_{10}6 - \log_{10}18 = \log_{10}2$
7. $2\times\log_{10}4 - \log_{10}8 = \log_{10}2$
8. $2\times\log_{10}12 - \log_{10}6 = \log_{10}24$

Natural Logarithms

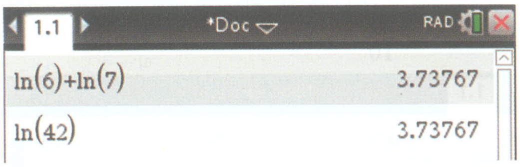
Earlier in this chapter we derived Euler's Number, *e*.

Logarithms to this base are sometimes called Natural Logarithms. They are written $\log_e$ or ln.

The calculator version is almost always a key labelled 'ln'.

All the results of the base 10 logarithms work for base *e*.

For example: $\log_e 6 + \log_e 7 = \log_e 42$



Exercise A.3.6

Use your calculator to verify that:

1. $\log_e 5 + \log_e 7 = \log_e 35$
2. $\log_e 6 + \log_e 5 = \log_e 30$
3. $\log_e 14 - \log_e 2 = \log_e 7$
4. $\log_e 120 - \log_e 5 = \log_e 24$
5. $2\times\log_e 3 = \log_e 9$
6. $2\times\log_e 6 - \log_e 2 = \log_e 18$
7. $3\times\log_e 2 - \log_e 4 = \log_e 2$
8. $2\times\log_e 12 - \log_e 3 = \log_e 48$

Modelling

Logarithms, as we have seen, have some curious properties. These make them suitable in the modelling of a number of physical phenomena. We will look at a few of these. There are many others.

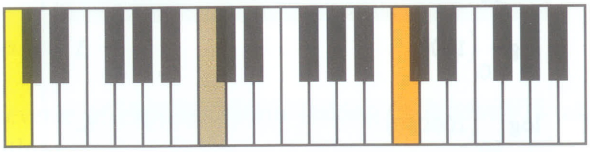
Musical Scales

Music was studied by the Ancient Greeks who made a number of discoveries that led them to believe that there was a direct connection between beauty in music and the order of numbers.

We will look at one aspect of this - musical scales.

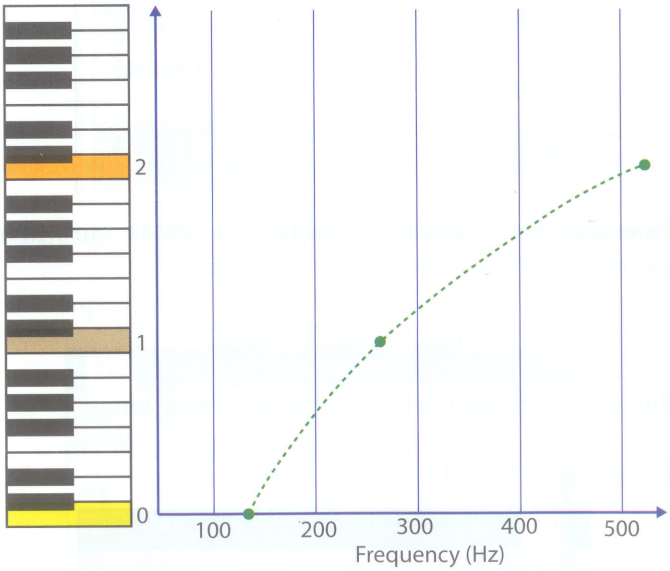
The piano keyboard is a row of keys. When pressed, these strike strings. These strings are usually arranged in groups of three that are struck simultaneously by felt hammers. The frame and body of the piano amplify the vibration of the strings, producing the characteristic sound of the instrument.

The keys are arranged in groups of eight known as 'octaves'.



The brown key is known as 'Middle C' and has a frequency of 262 cycles per sec (Hertz - Hz). The yellow key an octave below has a frequency of half this (131 Hz) and the orange key has a frequency of twice this (524 Hz). The notes sound similar and, if sounded together (a chord) they sound harmonious. It was this that the Ancient Greeks found so satisfying.

Thus, every eight notes to the right results in a doubling of frequency. This is the behaviour of logarithms:



Algebraically, the frequency n octaves above or below middle C is $F = 262 \times 2^n$, $n = 0, \pm1, \pm2, \ldots$

Greenhouse Effect

The Earth is warmer than it otherwise would be because, like a greenhouse, the atmosphere traps a proportion of the Sun's radiant energy. Without this, it is unlikely that the Earth could sustain life.

The process is complex, but there is little doubt that water vapour and carbon dioxide are important factors in the greenhouse effect. The water vapour content varies from time to time and from place to place. So do the clouds that result from the water vapour content. Clouds tend to reflect sunlight, reducing the greenhouse effect. It is this sort of thing, combined with the turbulence of the atmosphere that makes climate modelling such a complex enterprise.

The effect of carbon dioxide on the greenhouse effect is still the subject of some debate, but is better understood than the effect of water vapour. It is clear that the carbon dioxide concentration in the atmosphere has been increasing in recent years. It is also clear that the atmosphere is warmer that it would otherwise be because of this. The effect is, however, logarithmic. A doubling of carbon dioxide concentration will increase the temperature by an amount known as the 'climate sensitivity'. Likewise, a halving of the concentration will reduce the temperature by the same amount.

The size of this climate sensitivity is very uncertain and is currently thought to be 1.5-4.5˚C. This big range underlines the fact that this issue is very far from being settled. Currently, the carbon dioxide concentration is about 400 ppm. If it doubles to 800 ppm, the prediction is an increase of 1.5-4.5˚C. If it halves to 200 ppm, the prediction is for a decrease of the same amount. This has logarithmic characteristics.

There are other climate drivers such as water vapour and solar activity and it is unclear which will dominate. It is unclear whether the benefits of more carbon dioxide (better plant growth etc.) will outweight the harm (melting ice etc.).

Sound

Sound intensity is measured in decibels. This is a logaritmic scale.

$$L = 10 \times \log_{10}\left(\frac{I}{I_0}\right)$$

L is loudness in decibels (db), I is the intensity (power) of a sound and I_0 is the intensity of a sound at the threshold of hearing.

It follows that, if a sound is 10 times more energetic than a sound at the threshold of hearing, its loudness is:

$$\begin{aligned} L &= 10 \times \log_{10}\left(\frac{10 \times I_0}{I_0}\right) \\ &= 10 \times \log_{10}(10) \\ &= 10 \end{aligned}$$

The loudness is 10 dB.

If a sound is 100 times more energetic than a sound at the threshold of hearing, its loudness is:

$$L = 10 \times \log_{10}\left(\frac{100 \times I_0}{I_0}\right) = 10 \times \log_{10}(100) = 20$$

If a sound is 1 000 times more energetic than a sound at the threshold of hearing, its loudness is:

$$L = 10 \times \log_{10}\left(\frac{1000 \times I_0}{I_0}\right) = 10 \times \log_{10}(1000) = 30$$

Every 10 dB increase reflects at 10 TIMES more energetic sound.

Earthquakes

The strength of Earthquakes is usually measured using the Richter Scale.

This is similar to the sound example and is logarithmic.

$$M = \log_{10}\left(\frac{I}{I_0}\right)$$

M is the Richter Magnitude of an earthquake. The I values have a technical definition but can be thought of as intensities or energy dissipated in a standard (small earthquake I_0) and that dissipated in a new event.

The consequence of this is that every increase of 1 in the Richter Magnitude means that the energy dissipation is 10 times greater.

Many people have trouble in understanding these logaritmic scales despite the fact that they are quite common. A bit of practice with your calculator will help you avoid the more common misunderstandings.

Exercise A.3.7

1. The frequencies audible by humans are often given as 20 Hz to 20 000 Hz.

   How many octaves is this?

2. The effectiveness of a chemical catalyst (E) depends on it concentration (%) in the reaction medium (C). This is modelled by:

   $$E = 30 \times \log_e\left(\frac{C}{10}\right), C > 10$$

   a Find the effectiveness of the catalyst when its concentration is 16%.

b Find the effectiveness of the catalyst when its concentration is 35%.

c Find the smallest concentration that will provide an effectiveness of about 50.

3. The *Rock Blaster* disco has 100 watt amplifiers and the *Ground Shaker* disco has 1 000 watt amplifiers.

   Approximately how much louder (dB) is the sound in *Ground Shaker* than the sound in *Rock Blaster*.

4. How much stronger is a magnitude 6 earthquake than a magnitude 4 earthquake?

5. The concentration of a drug in the blood (C mg/litre) t hours after administration is modelled by:

$$C = 10 - \ln(t+1)$$

   a What is the initial concentration?

   b What is the concentration after 4 hours?

   c Approximately when does the concentration first fall below 7.44 mg/L?

6. The profit to earnings ratio (P/E ratio) P of a company t years after its formation is modelled by:

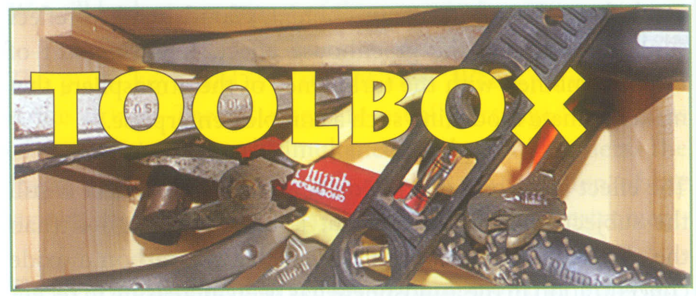
$$P = 10 + 5 \times \ln(t+1)$$

   a What is the initial P/E ratio?

   b What is the P/E ratio after 4 years?

   c Approximately when does the P/E ratio first exceed 18?

Answers



Public Key Encryption

Return to the issue of prime factors that we discussed at the start of the chapter.

In 1999, Irish teenager Sarah Flannery caused a media frenzy when she published her idea for an encryption system to protect electronic financial transactions.

Her book, *In Code* ISBN: 0-7611-2384-9 explains her system in admirably clear terms. If you are interested in the issue of how the banks protect card transactions and how your PIN is linked to your card, this is a good place to start.

The important factor here is that the card must be quickly verified when it is presented with the PIN. However, having the card and discovering the PIN must be very difficult.

What we need is a process that proceeds easily in one direction but is almost impossible to reverse.

Prime factors present such a process. When we asked you to construct factor trees, we were careful to choose numbers such as 48 that have a lot of small prime factors.

What if we had given you 371 081?

It is likely that this will take you quite a bit of time even with a calculator. The reason is that 371 081 = 433 × 857 and both these factors are prime.

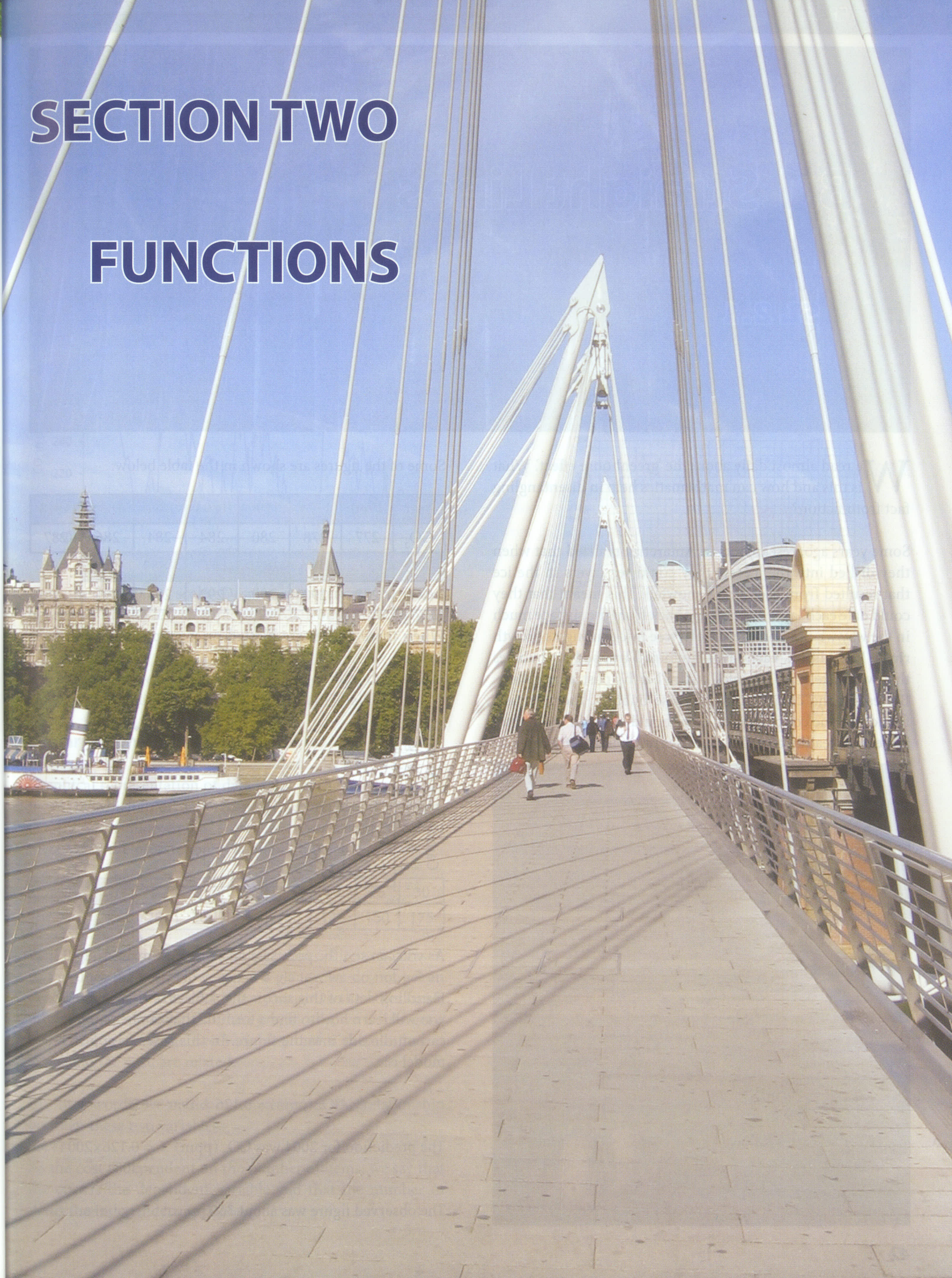
Going from 433 × 857 to 371 081 is easy. The reverse is difficult.

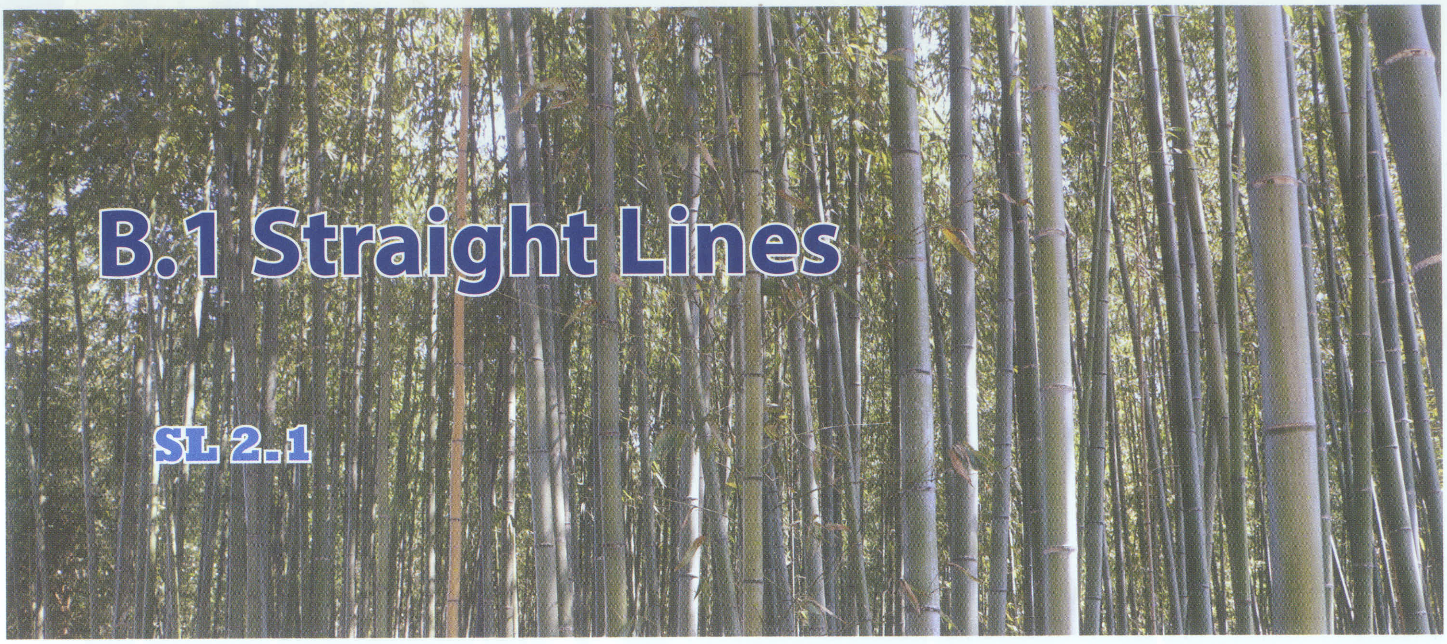
Suppose that, instead of two primes each with 3 digits, we can find a pair of primes with thousands of digits, we can produce a number with millions of digits that has exactly two factors that are virtually impossible for someone else to discover.

This is the basis of one encryption system.

SECTION TWO

FUNCTIONS



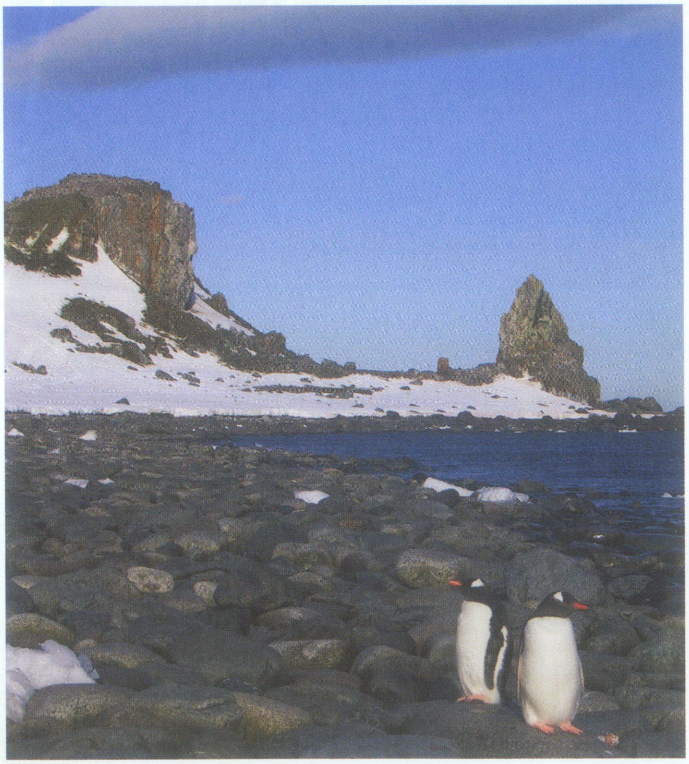


B.1 Straight Lines

SL 2.1

We read almost daily about the 'greenhouse effect'. What is this and how can mathematics help in disentangling fact from fiction?

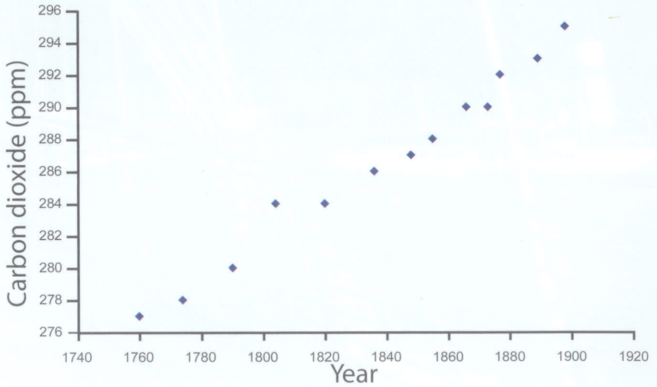
Some years ago, researchers in Antarctica realized that when they drilled into the icecap, they could see layers in the ice that resulted from each year's snowfall. This meant that they could date each layer by counting back from the surface much in the way we can find the age of a tree by counting its growth rings. In addition, each layer contained trapped air bubbles that were like time capsules of the Earth's past atmosphere. By testing the composition of these bubbles, the scientists found that the concentration of carbon dioxide in the air had increased steadily over the centuries preceding 1900.



Some of the figures are shown in the table below:

Year	1760	1774	1790	1804	1820	1836	1848
CO_2	277	278	280	284	284	286	287
Year	1855	1866	1873	1877	1889	1898	
CO_2	288	290	290	292	293	295	

When these figures are plotted on a graph, the result is as shown.



As you can see, the points lie on an approximate straight line.

Handling data of this sort is the subject of this chapter. In it you will learn how to find a formula that can help predict the carbon dioxide into the future. In this case, an approximate model is:

$$CO_2 \text{ (ppm)} = 0.126 \times \text{Year} + 55.$$

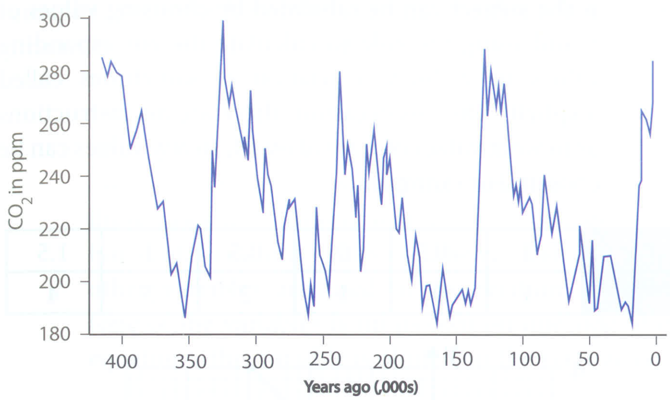
The prediction for 2004 was CO_2 (ppm) $= 0.126 \times 2004 + 55$
$= 308.$

The observed figure was about 360 ppm.

This illustrates that, whilst mathematical modelling is useful, it is not infallible. There is, however, a lot of evidence that CO_2 levels in the atmosphere are currently increasing.

What, if anything, should be done about this is another matter. Both positive and negative effects of this are predicted. Amongst the negative effects are damaging weather and rising sea levels. Amongst the positive are improved green plant growth - the CO_2 level for optimal plant growth is often reported as 800 ppm.

It is also untrue to assume that this trend has always been linear. On a larger time scale, the measured CO_2 levels were:



Note the horizontal scale. Despite appearances, the oldest figures are to the left so time proceeds from left to right as usual.

Over long time periods, the data is more cyclical than linear. Don't assume that mathematical models are infallible!

Linear Graphs

Graphical representation of data is one of the most powerful means by which information can be interpreted and analysed. Consider the following set of data, representing the cost involved in the production of an item (in batches of five):

No. of items	0	5	10	15	20	25	30
Cost ($)	25	50	75	100	125	150	175

From the table it appears that as the number of items increases, so too does the cost. However, at a glance, we do not get an overall feel for how the cost is increasing. On the other hand, if we were to represent this data graphically, we would get a better feel for how the two variables, that is the number of items and the cost, are related.

In other words, we could more easily 'see' the relationship between these two variables.

As the cost is dependent on the number of items, we say that the cost is the dependent variable and that the number of items the independent variable.

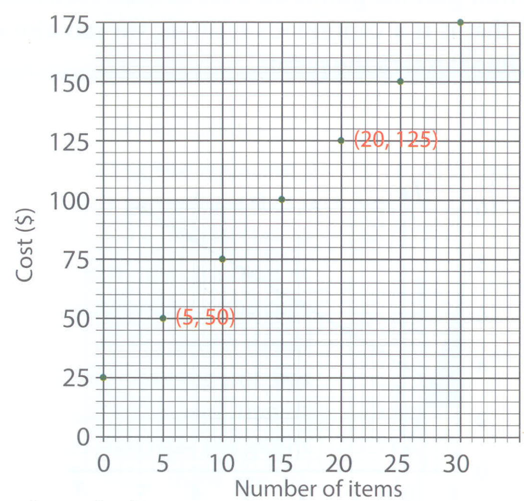
Next we plot a graph of the above data with the independent variable along the horizontal axis and the dependent variable along the vertical axis. To plot the relationship between these two variables we use a set of ordered pairs:

When 5 items are sold, we have an associated cost of $50 and so this would correspond to the ordered pair (5, 50). When 15 items are sold, we have a cost of $100, giving the ordered pair (15,100).

We continue in this manner until we have a complete set of ordered pairs:

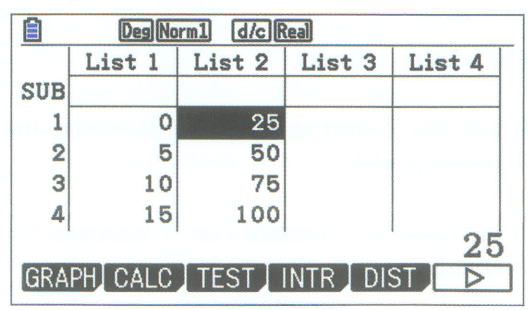
{(0, 25), (5, 50), (10, 75), (15, 100), (20, 125), (25, 150), (30, 175)}.

These points can now be plotted, either using graph paper:

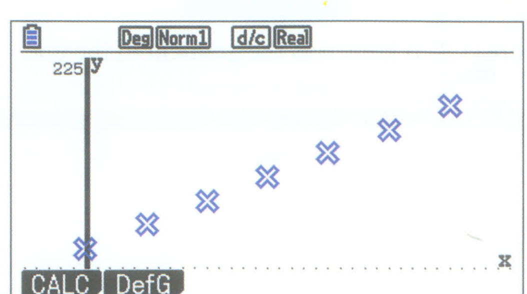


or, a graphic calculator:

One way is to use the statistical mode. The data is entered as a pair of lists:



and using scattergraph:



Example B.1.1

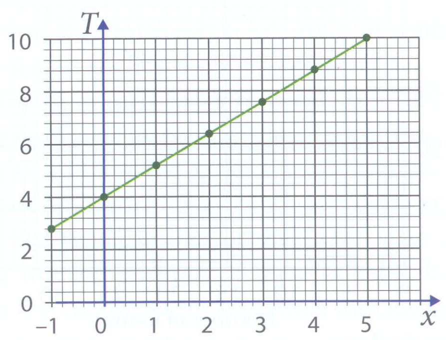
For the continuous relation, $T = 1.2x + 4$, construct a table of values for values of x from –1 to 5 and use it to sketch the graph of this relation. Give a brief description of this graph.

Use each x value in turn in the formula. For example,

if $x = 3$, $T = 1.2 \times 3 + 4 = 7.6$, giving the point (3,7.6).

x	–1	0	1	2	3	4	5
T	2.8	4	5.2	6.4	7.6	8.8	10

We can now plot the points on a set of axes. However, as we are told that it is a continuous relation, we can also draw a straight line through the points.



The following observations can be made:

1. The relation is **linear**.
2. It cuts the T–axis at the point (0, 4), i.e. line has a T-intercept of 4.
3. It increases at a constant rate. T increases by 1.2 for every increase of unit in x.

$$\frac{\text{rise in } T}{\text{run in } x} = 1.2$$

This figure (1.2) is known as the **gradient** of the line.

Example B.1.2

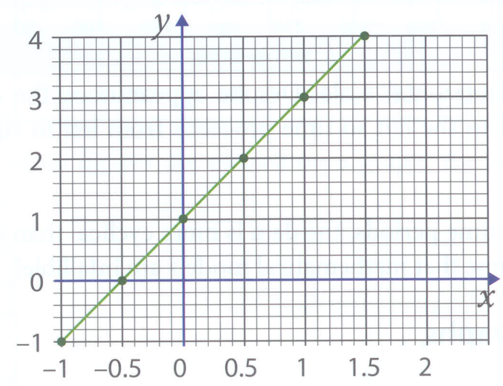
Using tables of values for the following relations, draw their graphs.

a $y = 2x + 1$ b $x + y = 2$

c $\frac{2t}{3} - s = 4$

a An equation of this type in which one of the variables is the subject can be tabulated by choosing values of x and using the rule to calculate the corresponding values of y. Such a relation is sometimes called '**explicit**'. On this occasion, there are no instructions as to what values of x are allowed, so any values can be chosen: For example:

x	–1	–0.5	0	0.5	1	1.5
y	–1	0	1	2	3	4

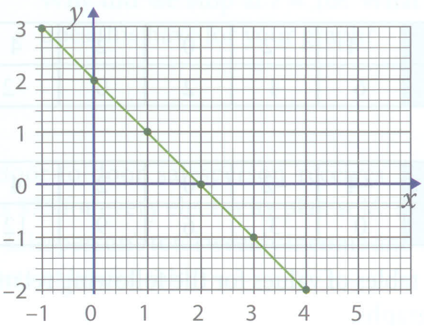


The points of the graph lie on a straight line, hence the description, 'linear'. It should also be noted that only a sample of the points has been plotted on the graph. If all the points are added they will form a continuous straight line.

This line passes through the y-axis at 1 and has a slope (gradient) of 2 (the line rises by 2 units for each unit moved to the right).

b $x + y = 2$ is known as an implicit relation. It is possible to make y the subject of the relation (to get $y = 2 - x$), but this is not necessary. All that is required is to find pairs of numbers that add up to 2.

x	–1	0	1	2	3	4
y	3	2	1	0	–1	–2



As before, the entries in this table are only examples. If the complete set of such points is plotted, the result is a straight line with y-intercept 2 and gradient -1.

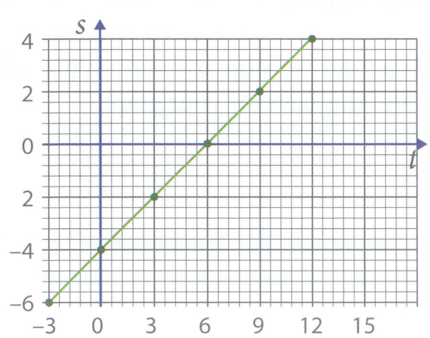
c The previous examples used the variables x and y. In these cases it is conventional to plot x values on the horizontal axis and y values on the vertical axis. When other variables are used, if the relation is explicit (one of the variables is on its own on one side of the equation), this variable is plotted on the vertical axis.

In this case, it is not clear which variable should be plotted on each axis. Also, it is probably easiest to make one of the variables the subject of the equation:

$$\frac{2t}{3}-s=4 \Rightarrow \frac{2t}{3}=s+4 \Rightarrow s=\frac{2t}{3}-4$$

This can now be used as in part a to calculate values of s after choosing values for t. The table could be:

t	–3	0	3	6	9	12
s	–6	–4	–2	0	2	4



There are two commonly used standard terms that allow us to avoid the construction of tables.

Gradient/Intercept form

This relies on the equation being of the form $y = mx + c$ (ie. explicit).

1. The gradient (or slope) of the line is given by m.

2. The y-intercept for the line is c, i.e. the straight line cuts the vertical axis at $(0, c)$.

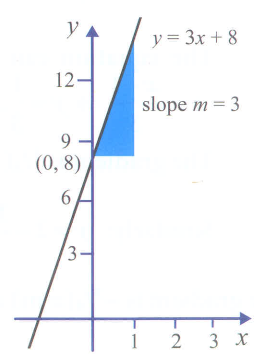
Example B.1.3

For each of the following linear relations, identify its gradient and y-intercept and then, sketch its graph.

a $y = 3x + 8$ b $y = x - 2$

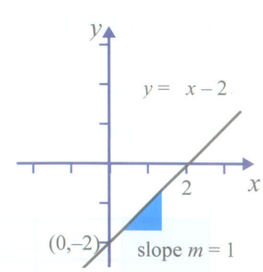
c $4x - 8y = 24$

a Comparing the equation $y = 3x + 8$ with that of $y = mx + c$, we have that $m = 3$ and $c = 8$.

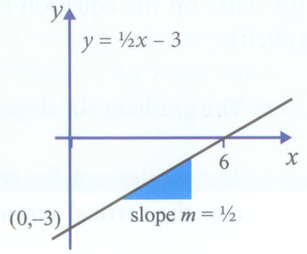


Therefore, the gradient is 3 (= m) and the y-intercept is 8 (= c). The graph needs to 'show' that the line has a slope of 3 and that it passes through the point (0, 8). Also, there are no restrictions on the x-values that can be used, and so we can assume that we can use the set of real values for x. The set of values of x that can be used is called the **domain**. Notice that although the graph seems to 'stop', it does in fact continue. To indicate that a graph does not go beyond a particular point, we place a closed circle at the end point(s).

b Comparing the equation $y = x - 2$ with that of $y = mx + c$, we have that $m = 1$ and $c = -2$. Therefore, the gradient is 1 (= m) and the y-intercept is –2 (= c). The graph should display a line passing through the point (0,–2) and with a gradient of 1.



c The equation $4x - 8y = 24$ is not in the standard form and so, this time, we need to first carry out a little algebra. We need to rearrange the equation $4x - 8y = 24$ and make y the subject.



$4x - 8y = 24 \Rightarrow -8y = -4x + 24 \Rightarrow y = \frac{1}{2}x - 3$

Example B.1.4

State the gradient and y-intercepts of the graphs related to these equations.

a $y = 2x - 3$ b $y = \dfrac{x-1}{3}$

c $y = 2 - \dfrac{3x}{4}$

a By comparing the equation to the standard form , the gradient $(m) = 2$ and the y-intercept $(c) = -3$ can be read directly from the equation.

b The equation can be written in the standard form: $y = \dfrac{x-1}{3} \Rightarrow y = \dfrac{1}{3}x - \dfrac{1}{3}$.
The gradient is $^1/_3(=m)$ and the y-intercept is $-^1/_3(=c)$.

c Similarly: $y = 2 - \dfrac{3x}{4} \Rightarrow y = -\dfrac{3}{4}x + 2$.

The gradient is $-^3/_4(= m)$ and the y-intercept is 2 (= c).

Exercise B.1.1

1. For each table of values (you may wish to use a graphics calculator):

   i plot the set of points on a set of axes.

   ii draw a continuous line through the points in part i.

a.

x	0	1	2	3	4	5
y	3	5	7	9	11	13

b.

x	–3	–1	1	3	5	7
y	1	3	5	7	9	11

c.

x	–4	–2	0	2	4	6
y	6	4	2	0	–2	–4

d.

x	–4	–2	0	2	4	6
y	0	3	6	9	12	15

2 Using tables of values for the following relations, draw their graphs.

   a $y = 3x - 1$ b $y = 4x + 2$

   c $y = -2x + 4$ d $y = {}^2/_3x - \frac{1}{2}\ 2$

3. Sketch the graphs of the following.

   a $x - 3y = 6$ b $2x = 3y + 9$

   c $6x = -2y + 12$

4. A computer is purchased for $900 at the start of 2002 and losses $150 of its value every year.

   a Construct a table showing the value of the computer, $$V$, t years since the start of 2002.

t (years)	0	1	2	3	4	5	6
V ($)	900						

   b Graph the relationship between the computer's value $$V$ and the number of years, t, since the start of 2002.

   c Why did we stop at $t = 6$? What implications does this have for your graph?

5. A tank containing 250 litres of water has its tap left open and water is leaking at a constant rate of 25 litres per hour.

   a Construct a table of values for the amount of water, V litres, left in the tank every hour.

t (hours)	0	1	2	3	4	5	...	9	10
V (Litres)	250								

   b Graph the relationship.

c Why did we stop at $t = 10$? What implications does this have for your graph?

6. A door-to-door salesperson receives \$750 per week plus \$75 for every electricity plan sold. Let \$$I$ represent the salesperson's total weekly income and n the number of sets of books sold.

a Construct a table of values for the salesperson's total weekly income.

b Plot your results on a set of axes.

c Graph the relationship between the total weekly income \$$I$ and the number of electricity plans sold.

7. The relationship between a father's height, H cm, and the average height of his sons, h cm, can be approximated by the linear relation $h = ½H + 87$.

a Construct a realistic table of values for the relationship between h and H.

b From your graph, determines the son's height if the father is 187 cm tall.

Intercept Form

If the equation is in the form $ax + by = c$ it can be easiest to look at intercepts when sketching graphs. This also applies to equations in the form: $ax + by + c = 0$.

Example B.1.5

Find the axial intercepts and sketch the graphs of these linear relations.

a $2x-3y=12$ b $\frac{2}{3}x-\frac{y}{4}=4$

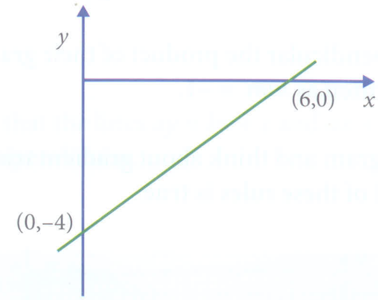
c $2x+3y-4=0$

a For y-intercept set $x = 0$:

$$2\times0-3y=12 \Rightarrow -3y=12 \Rightarrow y=-4$$

For x-intercept set $y = 0$:

$$2x-3\times0=12 \Rightarrow 2x=12 \Rightarrow x=6$$

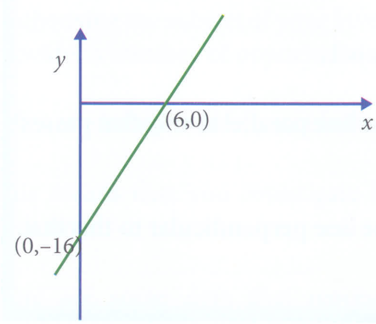


b For y-intercept set $x = 0$:

$$\frac{2}{3}\times0-\frac{y}{4}=4 \Rightarrow -\frac{y}{4}=4 \Rightarrow y=-16$$

For x-intercept set $y = 0$:

$$\frac{2}{3}x-\frac{y}{4}\times0=4 \Rightarrow \frac{2}{3}x=4 \Rightarrow x=6$$



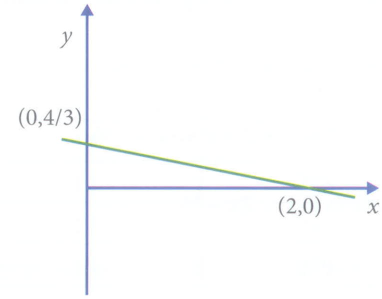
c For y-intercept set $x = 0$:

$$2x+3y-4=0 \Rightarrow 2\times0+3y-4=0 \Rightarrow 3y=4 \Rightarrow y=\frac{4}{3}$$

For x-intercept set $y = 0$:

$$2x+3y-4=0 \Rightarrow 2x+3\times0-4=0 \Rightarrow 2x-4=0$$

$$\Rightarrow 2x=4 \Rightarrow x=2$$

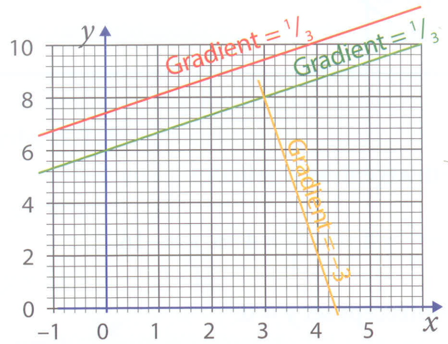


Perpendicular and Parallel Lines

If lines are parallel, they have the same gradient.

If lines are perpendicular the product of their gradients is –1. This is often written as $mm' = -1$.

Look at the diagram and think about gradient triangles to see why the second of these rules is true.



Example B.1.6

A line has equation $x + 2y = 1$.

a Find the equation of the line parallel to this that passes through (5,2).

b Find the equation of the line perpendicular to this that passes through (2,3.5).

a $x + 2y = 1$ rearranges to $y = -\frac{1}{2}x + \frac{1}{2}$. This has gradient $-\frac{1}{2}$. We are looking for the line with gradient $-\frac{1}{2}$ that passes through (5,2).

This has equation $y = -\frac{1}{2}x + c$.

At (5,2) $2 = -\frac{1}{2} \times 5 + c$ so $c = 4\frac{1}{2}$

The equation is $y = -\frac{1}{2}x + 4\frac{1}{2}$ or $x + 2y = 9$.

b Perpendicular lines have gradient = 2 using $mm' = -1$.

They have equations of the form $y = 2x + c$.

At (2,3.5) $3.5 = 2 \times 2 + c$ so $c = -\frac{1}{2}$

The equation is $y = 2x - \frac{1}{2}$.

Exercise B.1.2

1. Sketch graphs of these functions:

a $2x+3y=6$

b $2x-y=4$

c $\frac{x}{2}+y=2$

d $3x+2y+6=0$

e $\frac{2x}{3}-\frac{y}{4}=1$

f $3x-2y+7=0$

g $\frac{x-1}{3}+\frac{y+2}{4}=1$

h $2(x-4)-3(y+6)=1$

2. Find the equation of the straight line that:

a has a gradient of 2 and passes through the point (0, 5)

b has a gradient of –1 and passes through the point (0, 3)

c has a gradient of –2 and passes through the point (4, 0)

d has a gradient of –0.5 and passes through the point (4, 1)

e has a gradient of $^2/_3$ and passes through the point (6, 4).

3. It costs \$4 to set the type to print business cards. Each card costs \$0.02 to print.

a Write an equation that could be used to calculate the cost (\$$C$) of printing n business cards.

b Sketch the graph of the relationship between the \$$C$ and the number of cards, n.

c What cost is involved in printing 2000 cards?

4. Joe's telephone account is made up of a fixed rental

charge of $44 per three months, plus $0.22 per mobile call. Assume that Joe has not made any non-mobile calls.

a Write down a relationship between Joe's telephone bill, $\$B$, and the number of calls, n, he makes in a three month period.

b Sketch the relationship in part a.

c What was the amount shown on Joe's last telephone bill if he made 168 calls during the last three-month period?

d How many calls would Joe have?

5. It costs 35 cents per kilometre to run a car on petrol. If the car is converted to liquefied petroleum gas, the cost becomes 27 cents per kilometre.

a Write equations that give the cost of running the car on:

i petrol (P) for k kilometres.

ii gas (G) for k kilometres.

b On the same set of axes, sketch both graphs in part a.

c If it costs $1000 to convert a car from petrol to gas, how far would the owner have to drive before recovering the conversion costs?

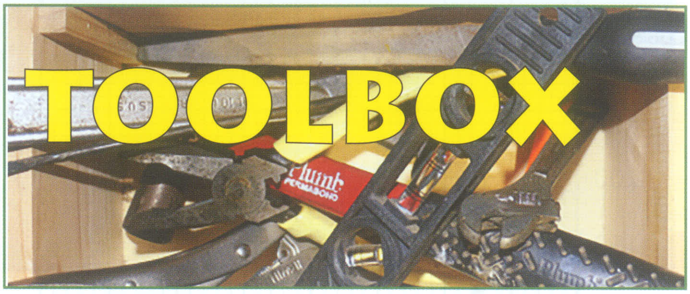
6. Find the equation of the line parallel to $y = x + 9$ that passes through (4,5).

7. Find the equation of the line parallel to $y + x = 9$ that passes through (1,2).

8. A line joins the points (4,3) and (6,7). Find the equation of the line that is perpendicular to this and which passes through (5,5)

9. A line segment joins the points (–2,4) & (6,–4). Find the equation of the perpendicular bisector of this line segment.

10. Prove that the lines $ay = bx + c$ and $ax + by = c$ $(a,b\neq 0)$ are perpendicular.



We began this chapter discussing carbon dioxide and the hazards of assuming linear models.

In choosing the subject of your investigation, linear modelling provides a number of opportunities for relevant study.

Often the independent variable (ie. x) will be time.

This means that you investigate how some quantity varies over time.

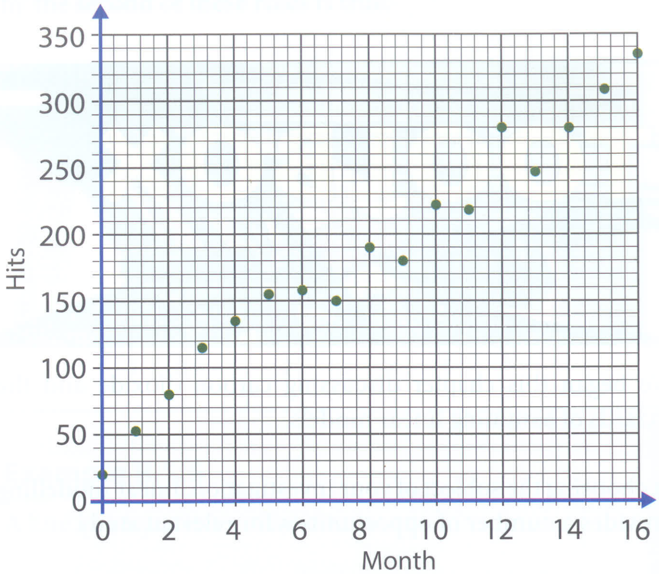
Here are some data that represent the number of 'hits' registered by a new website in the months after its launch.

Month	Hits
0	21
1	53
2	79
3	115
4	134
5	155
6	159
7	151
8	191
9	180
10	223
11	217
12	278
13	246
14	279
15	310
16	336

What conclusions can we draw about this?

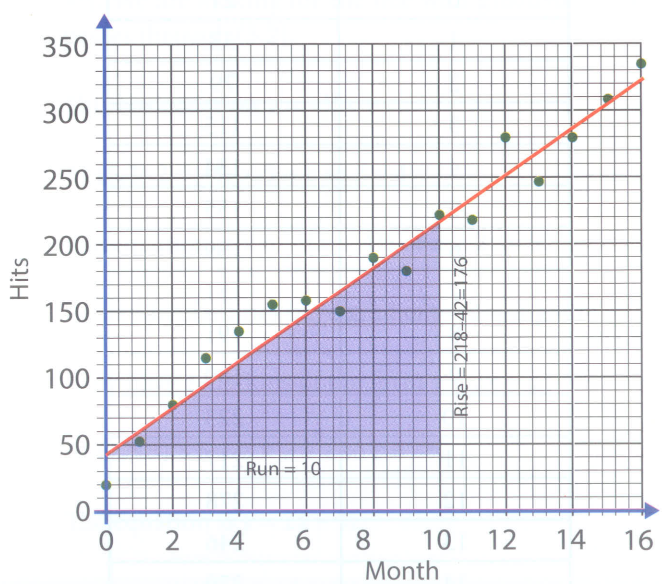
The first step will almost certainly be to choose an appropriate way of visualising the data - in this case a graph of the type discussed in this chapter.

Whilst it is possible to do all the graphing using technology, a pencil and paper graph works very well.



There appears to be a linear trend.

To test this, we try to fit a straight line. This will not pass through all the points. However, the aim is to draw the line that passes as close to the points as possible. This can be done with a ruler, but a stretched length of string will enable you to see the points above and below the line. Once you have judged the best position for the line, remove the string and draw it with a ruler.



Reading values from the (red) trendline gives:

Gradient = rise/run = 176/10 = 17.6.

Intercept = 42

Using hits = h and month = m, this gives us the linear model:

$$h = 17.6m + 42$$

Before using this model, it is a good idea to evaluate it. Evaluate can mean several things. In the first instance, we can look at how closely the model fits the established data. Here are a few test values:

Month	Hits	Model	Error
0	21	42	-21
2	79	77.2	2
4	134	112.4	22
6	159	147.6	11
8	191	182.8	8
10	223	218	5
12	278	253.2	25
14	279	288.4	-9

Some of the deviations are positive and some negative. Both looking at the graph and these deviations, the linear model works well.

It can now be used to predict the number of hits that can be expected for a limited period into the future.

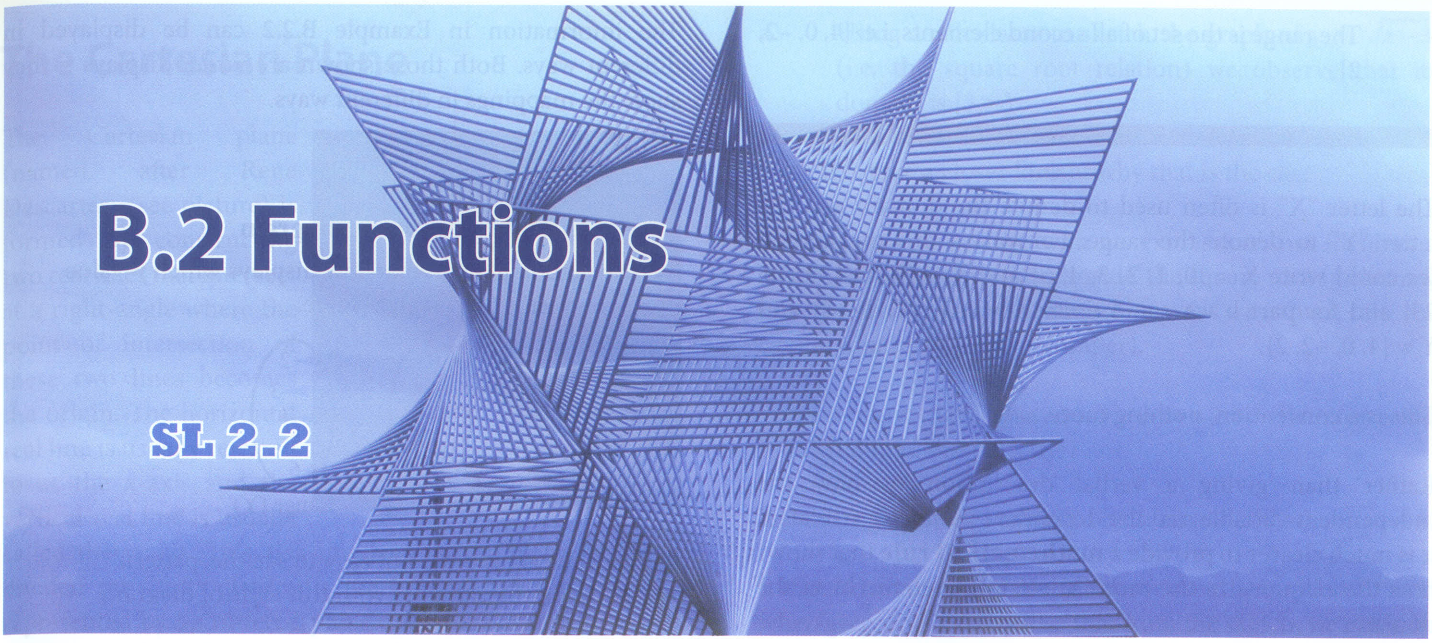
For example, the predicted number of hits after 20 months is 394. Such evidence could be used in trying to attract advertisers to the site.

However, it would be very unwise to predict into even the medium term future - such as 5 years.

Alternative

If your interests are very large, you might look at Hubble's Law and the fascinating story of the discovery of the Expanding Universe and the Big Bang.

Answers



B.2 Functions

SL 2.2

Relations

Consider the relationship between the weight of five students and their ages as shown below.

Age (years)	Weight (kg)
10	31
12	36
14	48
16	53
18	65

We can represent this information as a **set of ordered pairs**. An age of 10 years would correspond to a weight of 31 kg. An age of 16 years would correspond to a weight of 53 kg and so on.

This type of information represents a **relation** between two sets of data. This information could then be represented as a set of ordered pairs,

$$\{(10, 31), (12, 36), (14, 48), (16, 53), (18, 65)\}$$

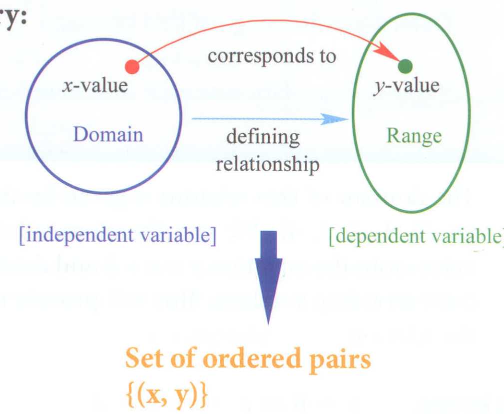
The **set of all first elements** of the ordered pair is called the **domain** of the relation and is referred to as the **independent variable**. The **set of all second elements** is called the **range** and is referred to as the **dependent variable**.

For the above example, the domain= $\{10, 12, 14, 16, 18\}$

and the range = $\{31, 36, 48, 53, 65\}$.

Notice that (10, 31) and (31,10) are not the same! This is because the ordered pair (10, 31) provides the correct relation between age and weight, i.e. at age 10 years the weight of the student is 31 kg. On the other hand, the ordered pair (31,10) would be informing us that at age 31 years the weight of the student is 10 kg!

Summary:



Example B.2.1

Determine the domain and range for each of the following relations:

a $\{(0, 0) (1, 1), (2, 4), (3, 9), (4, 16), (5, 25)\}$

b $\{(-3, 4), (-1, 0), (2, -2), (-2, 2)\}$.

a The domain is the set of all first elements, i.e. $\{0, 1, 2, 3, 4, 5\}$.

The range is the set of all second elements, i.e. $\{0, 1, 4, 9, 16, 25\}$.

b The domain is the set of all first elements, i.e. $\{-3, -1, 2, -2\}$.

The range is the set of all second elements, i.e. {4, 0, –2, 2}

The letter "X" is often used to denote the domain and the letter "Y" to denote the range. For part a this means that we could write X = {0, 1, 2, 3, 4, 5} and Y = {0, 1, 4, 9, 16, 25} and for part b we could write X = {–3, –1, 2, –2} and Y = {4, 0, –2, 2}.

This is a convention, nothing more.

Rather than giving a verbal description of how the independent variable and the dependent variable are related, it is much clearer to provide a **mathematical rule** that shows how the elements in the range relate to the elements in the domain.

Example B.2.2

A relation is defined by the rule 'add 2'.

a Determine the range of this relation.

b Express this relation as a set of ordered pairs.

a The domain of this relation is given by the *x*-values, i.e. {0, 1, 2, 3, 4}. We can therefore substitute these values into the equation $y = x + 2$ and determine their corresponding *y*-values. This will provide the range of the relation.

Substituting: $x = 0 \Rightarrow y = 0 + 2 = 2$

$x = 1 \Rightarrow y = 1 + 2 = 3$

$x = 2 \Rightarrow y = 2 + 2 = 4$, and so on.

This produces a set of *y*-values {2, 3, 4, 5, 6} that defines the range.

b The set of ordered pairs would be {(0, 2), (1, 3), (2, 4), (3, 5), (4, 6)}.

Notice that we can describe the set of ordered pairs more formally as:

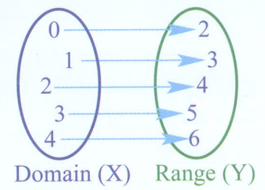
$\{(x,y) : y = x + 2, x \in \{0, 1, 2, 3, 4\}\}$ which is read as:

"The set of ordered pairs *x* and *y*, such that $y = x + 2$, where *x* is an element of the set of values {0, 1, 2, 3, 4}."

The information in Example B.2.2 can be displayed in different ways. Both those shown are *visual* displays – they show the mappings in different ways.

Mapping diagram

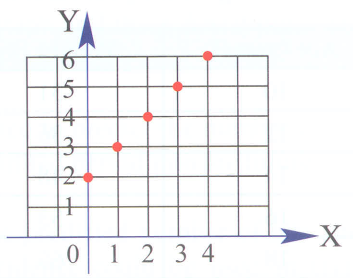
The mapping diagram below displays which *y*-value corresponds to a given *x*-value.



However it is often not easy to see the 'pattern' between the variables with this style of diagram.

Cartesian plane

The Cartesian plane is made up of a horizontal axis (independent variable, X) and a vertical axis (dependent variable, Y).



We plot the points on the grid, so that (3, 5) is 3 units to the right and 5 units up.

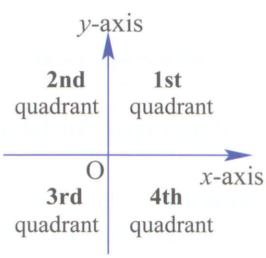
Notice that in the mapping diagram that uses the Cartesian plane, we have not joined the points together in a straight line. This is because the domain specifies that the only values of *x* that can be used must be from the set {0, 1, 2, 3, 4}, and so a value such as $x = 2.4$ cannot be used.

Both these visual representations are useful in displaying which values in the domain generate a given value in the range. However, the Cartesian plane more readily gives a quick overview of what the underlying relationship between the two variables is. It is very easy (and quick) to see that as the *x*-values increase, so too do the *y*-values. We can do this by simply looking at the points on the graph and observing the 'trend' without really concerning ourselves with what the actual values are.

We now provide a formal definition of the Cartesian plane and a relation.

The Cartesian Plane

The Cartesian plane (named after Rene Descartes - see picture) is formed by constructing two real lines that intersect at a right-angle where the point of intersection of these two lines becomes the origin. The horizontal real line is usually referred to as the x-axis and the vertical real line is usually called the y-axis. This also implies that the plane has been divided into four quadrants. Each point on this plane is represented by an ordered pair (x,y) where x and y are real numbers and are the coordinates of the point.



The set of all ordered pairs (x, y), where $x \in X$ and $y \in Y$ can also be defined by making use of the Cartesian product,

$$X \times Y = \{(x,y): x \in X, y \in Y\}$$

Implied domain

So far we have looked at examples for which a domain has been specified. Suppose we were asked to find the range of the relation $y = 1 + x^2, x \geq 3$? After sketching its graph, we would determine its range to be $[10, \infty)$. However, what if we wanted to know the range of the relation $y = 1 + x^2$? In this case, because we have not been provided with any restriction on the x-values, we will need to assume that we can **use the largest possible set of x-values for which the relation is defined** – this domain is known as the **implied domain** (or **maximal domain**) – in this case that would be the real number set, $\mathbb{R}$. Then, after sketching the graph of $y = 1 + x^2$ for all real values of x we would have a range defined by $[1,\infty)$.

Example B.2.3

Determine the domain and range of the following relations:

a $y = \sqrt{x-3}$ b $y = \dfrac{2}{\sqrt{x-3}}$ c $y = \dfrac{3}{2-x}$

a Using a calculator to sketch the graph of $y = \sqrt{x-3}$ (i.e. the square root relation) we observe that its domain is $[3,\infty)$.

Now, let's take a closer look at why that is the case.

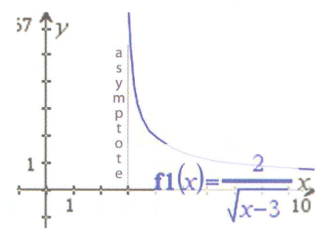
Because we are dealing with an expression that involves a square root, then, the term 'inside' the square root (radicand) must be greater than or equal to zero (as we cannot take the square root of a negative number).

So, we must have that $x - 3 \geq 0 \Leftrightarrow x \geq 3$. Therefore, the implied domain is $\{x: x \geq 3\}$.

From the graph, the range can be seen to be $[0,\infty)$.

It should be noted that the calculator uses the implied domain when graphing. Also realise that from the sketch, we could be misled into thinking that there is a 'gap' at the point (3, 0). Be careful with this – use the graphics calculator as an aid, then, double check to make sure.

b The equation: $y = \dfrac{2}{\sqrt{x-3}}$ represents the reciprocal of a square root relation.

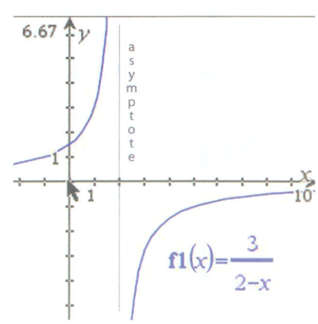


As in part a, we must have that $x - 3 \geq 0 \Leftrightarrow x \geq 3$.

However, this time we have another restriction – we cannot divide by zero and so we cannot include $x = 3$ in our domain. So, at $x = 3$, we draw an **asymptote**.

We then have $x - 3 > 0 \Leftrightarrow x > 3$. This leads to a range of $(0,\infty)$ (or $]0,\infty[$).

c The only restriction that can be readily seen for the relation $y = \dfrac{3}{2-x}$ is that we cannot divide by zero and so, we must have that $2 - x \neq 0$. That is, $x \neq 2$.



As it is a reciprocal relation, we have an asymptote at $x = 2$. So, the domain is given by $]-\infty,2[\cup]2,\infty$ or simply, $\mathbb{R}\backslash\{2\}$.

The range can then be seen to be $\mathbb{R}\backslash\{0\}$.

Exercise B.2.1

1 State the domain and range of the following relations.

a $\{(2,4), (3,-9), (-2,4), (3,9)\}$

b $\{(1,2), (2,3), (3,4), (5,6), (7,8), (9,10)\}$

c $\{(0,1), (0, 2), (1,1), (1,2)\}$

2 Find the range for each of the following.

a $\{(x,y): y = x + 1, x \in \mathbb{R}^+\}$

b $\{(x,y): y \geq x, x \geq 0\}$

c $y = x^2 + 2x + 1, x > 2$

d $y = 2x - x^2, x \in \mathbb{R}$

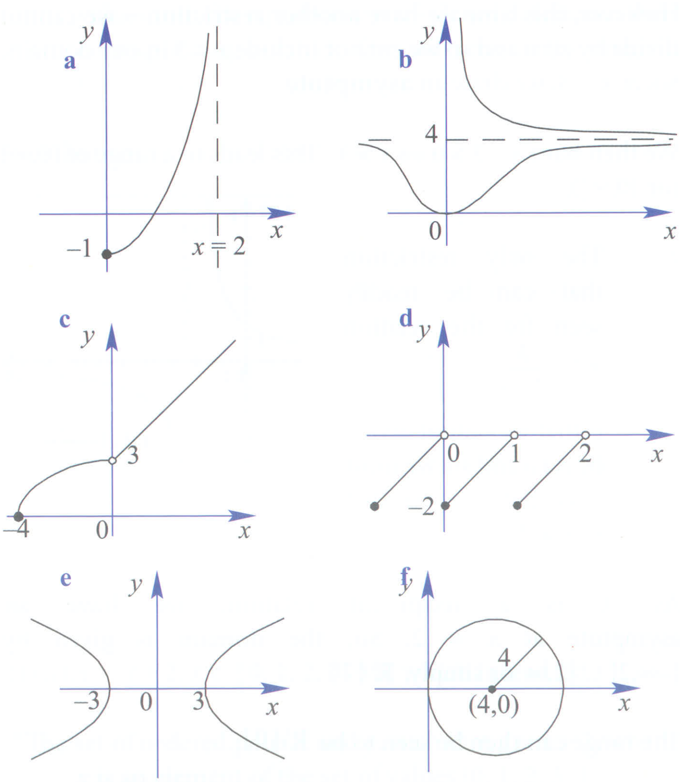
e $x^2 + y^2 = 9, -3 \leq x \leq 3$

f $x^2 - y^2 = 9, x \geq 3$

g $y = x - 1, 0 < x \leq 1$

h $y = 4 - x^2, -2 \leq x < 1$

3. State the range and domain for each of the following relations.



4. Determine the implied domain for each of the following relations.

a $y = \dfrac{2x}{x+2}$

b $y = \dfrac{3}{\sqrt{9-x}}$

c $y = \sqrt{16 - x^2}$

d $y = \sqrt{x^2 - 4}$

e $xy - x = 3$

f $y = \dfrac{2}{x^2 + 1}$

g $y = \dfrac{2}{x^3 + 1}$

5. Find the range of the following relations.

a $y = x - a, x < 0, a > 0$

b $y = \dfrac{ab}{x+1}, x \geq 0, ab > 0$

c $y = a^2x - ax^2, x \geq \frac{1}{2}a, a > 0$

d $y = a^2x - ax^2, x \geq \frac{1}{2}a, a < 0$

e $y = \dfrac{a}{x} + a, a > 0$

Extra questions

Not all mappings are mathematical. You may have noticed that the term 'mapping' is suggestive of maps in the everyday sense.

A map is, in every sense, a mathematical mapping. Every place on the map corresponds to a place in reality. It is one of our expectations when consulting a map that places that are close in reality will also be close on the map.

This is true when we are looking at locality maps such as the street directory of a city. It is not, however, true of maps of the entire World.

The history of map projections and the mathematics behind them makes fascinating reading and a potential topic for your investigation.

Functions

There is a special group of relations which are known as **functions**. This means that every set of ordered pairs is a relation, but **every relation is not a function**. Functions then make up a subset of all relations.

A function is defined as a relation such that each domain element has a unique image in the codomain. That is a function is a relation for which no two ordered pairs have the same first element.

For example, the function 'Take a real number, double it and add one' is commonly expressed in mathematical notation as:

$$f(x)=2x+1, x\in\mathbb{R} \text{ or } f:\mathbb{R}\mapsto\mathbb{R}, f(x)=2x+1$$

When you use a function on a calculator (such as x^2) you get a single answer - not a choice. This is why mathematicians like functions - they remove doubt! There are two ways to determine if a relation is a function.

Method 1: Algebraic approach

For Method 1 we use the given equation and determine the number of y-values that can be generated from one x-value.

Example B.2.4

Determine which (if any) of the following are functions.

a $y^3 - x = 2$ b $y^2 + x = 2$

a From $y^3 - x = 2$, we have $y = \sqrt[3]{2+x}$, then for any given value of x, say $x = a$, we have that $y = \sqrt[3]{2+a}$ which will only ever produce one unique y-value.

Therefore, the relation $y^3 - x = 2$ is a function. In fact, it is a one-to-one function.

b From $y^2 + x = 2 : y^2 = 2 - x \Leftrightarrow y = \pm\sqrt{2-x}$.

Then, for any given value of x, say $x = a$ (where $a \le 2$), we have that $y = \pm\sqrt{2-a}$, meaning that we have two different y-values; $y_1 = \sqrt{2-a}$ and $y_2 = -\sqrt{2-a}$, for the same x-value. Therefore, this relation is not a function.

Method 2: Vertical line test

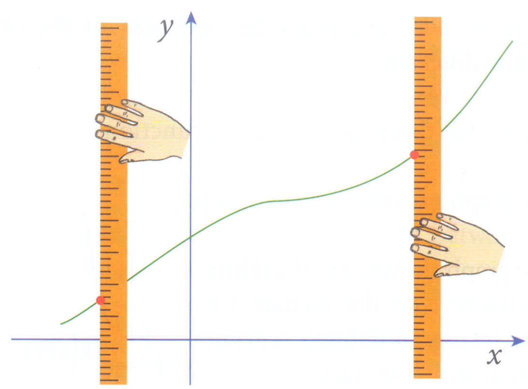
Step 1: Sketch the graph of the relation.

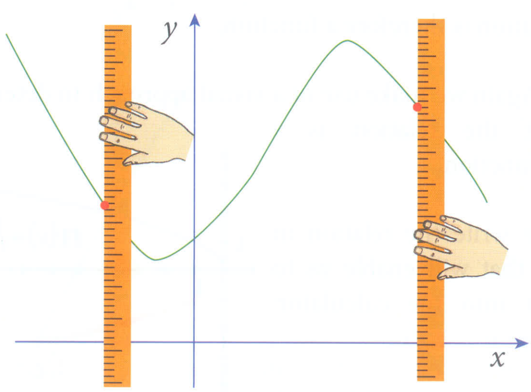
Step 2: Make a visual check of the number of times a vertical line would cut the graph.

Step 3: If the vertical line only ever cuts at one place for every value in the domain the relation is a function.

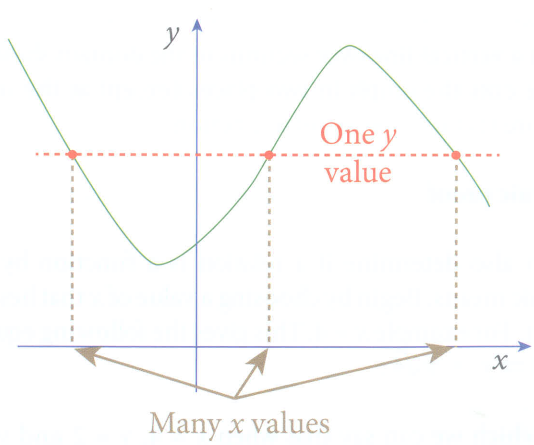
In both these examples, any vertical line (ruler) cuts the graph in at most one place.

Vertical Line Test





The first example is known as a 'one to one' function. The second is known as 'a many to one function' as many (in maths this 'means more than one') x values produce one y value.



Example B.2.5

Which of the following defines a function?

a $\{(0,2), (1,2), (2,1)\}$

b $\{(x, y): y = x^3 + 1, x \in \mathbb{R}\}$

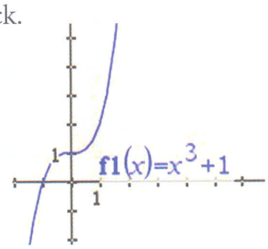
c $y^2 = x, x \geq 0$

d $\{(x, y): x^2 + y^2 = 16\}$

a Clearly, we have every first element of the ordered pairs different.

This means that this relation is also a function:

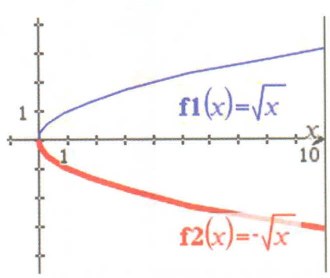
b A graph provides a visual check.



From the graph shown, a vertical line drawn anywhere on the domain for which the relation is defined, will cut the graph at only one place.

This relation is therefore a function.

c Again we make use of a visual approach to determine if the relation is a function.



First we write the relation in a form that will enable us to enter it into the calculator: $y^2 = x \Rightarrow y = \pm\sqrt{x}$

We can therefore define the relation $Y_1 = \sqrt{X}$ and $Y_2 = -\sqrt{X}$ and sketch both on the same set of axes.

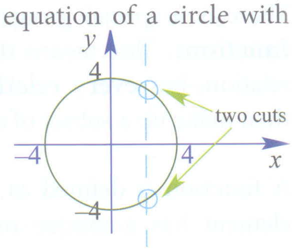
Placing a vertical line over sections of the domain shows that the line cuts the graph in two places (except at the origin). Therefore this relation is not a function.

Algebraic proof

We can also determine if a relation is a function by using algebraic means. Begin by choosing a value of x that lies in the domain. For example $x = 4$. This gives the following equation: $y^2 = 4 \Rightarrow y = \pm\sqrt{4}$.

From which we can say that when $x = 4$, $y = 2$ **and** $y = -2$, so that there are two ordered pairs, (4, 2) and (4, –2). As we have two different y-values for one x-value this relation is not a function.

d This relation describes the equation of a circle with radius 4 units and centre at the origin. The graph of this relation is shown alongside. The graph fails the vertical line test, and so is not a function.



Exercise B.2.2

1. A function is defined as follows, $f: x \mapsto 2x + 3, x \geq 0$.

a Find the value of $f(0), f(1)$.

b Evaluate the expressions: i $f(x + a)$
ii $f(x + a) - f(x)$

c Find $\{x: f(x) = 9\}$.

2. If $f(x) = \frac{x}{x+1}, x \in [0, 10]$, find:

a $f(0), f(10)$ b $\{x: f(x) = 5\}$

c the range of $f(x) = \frac{x}{x+1}, x \in [0, 10]$.

3. For the mapping $x \mapsto 2 - \frac{1}{2}x^2, x \in \mathbb{R}$, find:

a $f(x+1)$, $f(x-1)$ b a, given that $f(a) = 1$

c b, given that $f(b) = 10$.

4. A function is defined as, $y = x^3 - x^2, x \in [-2, 2]$

a Find the value(s) of x such that $y = 0$.

b Sketch the graph of $y = x^3 - x^2, x \in [-2, 2]$ and determine its range.

5. The function f is defined as $f{:}]-\infty,\infty[\ \mapsto \mathbb{R}$, where $f(x) = x^2 - 4$.

a Sketch the graph of:

i f ii $y = x + 2, x \in\ ]-\infty,\infty[$

b Find:

i $\{x{:}f(x) = 4\}$ ii $\{x{:}f(x) = x + 2\}$

Extra questions

Modelling

One of the most powerful applications of mathematics is in modelling.

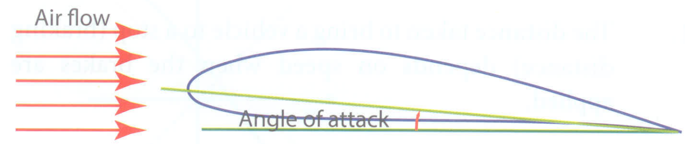
This is the process of taking a physical phenomenon (such as a chemical reaction or a design problem) and building a mathematical structure that behaves in a way similar to the phenomenon.

If this can be done, it becomes possible to work out how a particular design will work before it is constructed.

The 'Wright Flier' (Smithsonian Institution, Washington) was built after much experimentation with kites, gliders etc. because there was no base of experience that could be used to model powered flight.

Since then, an extensive theory of flight has developed. The designers of modern aeroplanes use functions, equations etc. to work out how big the wings need to be, how much fuel is needed and so on long before a single sheet of metal is ordered.

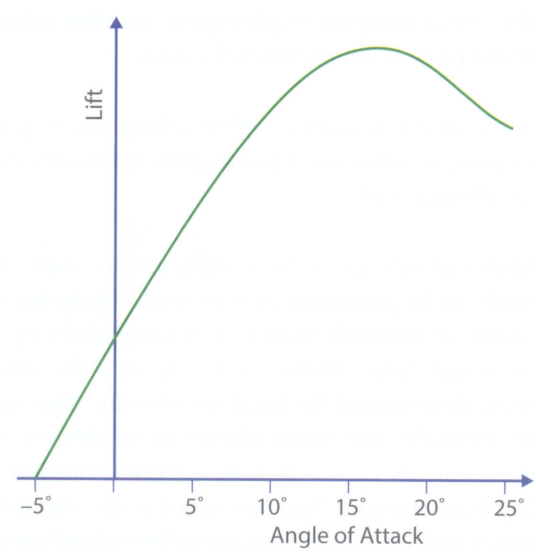
Functions play a considerable role in this design. For example, the amount of lift generated by a wing depends on a variety of factors such as its size, profile, thickness etc. Also important is the 'angle of attack'. This is illustrated:



The amount of lift generated by the wing depends on (is a function of) the angle of attack.

Note that there are other factors affecting lift as well. These vary from the very important (eg. ambient air pressure) to the minor (eg. humidity). We will look only at angle of attack.

A typical graph depicting the relationship between lift and angle of attack looks like this.



What does this graph tell us about the relationship between these two quantities? There are two main points. At low angles of attack, lift increases almost linearly as the aeroplane 'noses up'. At around 15˚ this reverses and lift falls quite suddenly. This is known as a 'stall' and is very dangerous. It results from a break-up of the airflow around the wing. Predicting when this will happen before it actually does is very useful!

Exercise B.2.3

Research these situations and sketch graphs that depict approximate models of the variables mentioned.

1. The distance taken to bring a vehicle to a stop (braking distance) depends on speed when the brakes are applied.

2. The concentration of a drug administered to a patient changes with time after it was administered.

3. The pressure of a fixed amount of gas varies with volume (Boyle's Law).

4. The amount of stretching (extension) produced in a wire when it is placed under tension.

5. Enzymes are biological substances that aid the chemical reactions needed for life - digestion, cell division etc. Their effectiveness often depends on temperature.

6. The sales a company might expect are often influenced by the price at which they sell goods.

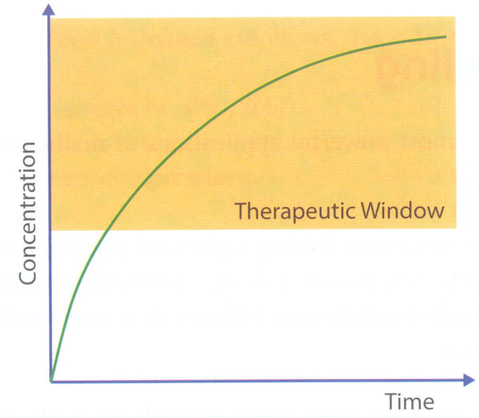
7. In the above scenario, higher selling prices generate more income per sale. How might total income depend on selling price?

8. Electrical energy is very difficult to store and it needs to be generated as it is used. There are now a variety of methods used to generate electricity: Coal, oil, wind, tides, hydro electricity etc. The choice is often determined by local conditions. You cannot, for example, use hydro electricity in deserts. Power requirements vary with time - more is needed in the day than at night. Another factor is the variability of power generated over time caused by natural variability in conditions such as the wind. Our photograph shows a part of the 300 MW Hellisheiði Geothermal Power Station in Iceland. Use a graphical method to demonstrate how supply might be organised to meet typical patterns of demand.

Inverses

Having considered some mathematical models of physical situations, we look at the idea of an inverse. As its name implies, this is the reverse of a process.

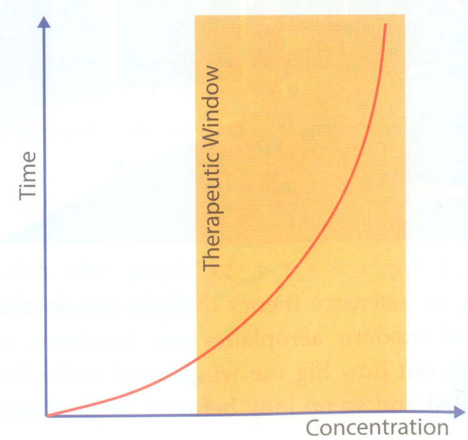
Consider this example. A drug is to be administered by an implant. The drug is taken up rapidly at first with the uptake gradually slowing as the implant is depleted. The concentration in the blood is important as most drugs have a 'therapeutic window' - a band of concentrations within which the drug is effective. Too little and there is not enough drug to be effective. Too much and there may be toxic side effects. Suppose the uptake and the therapeutic window are as shown on this graph:



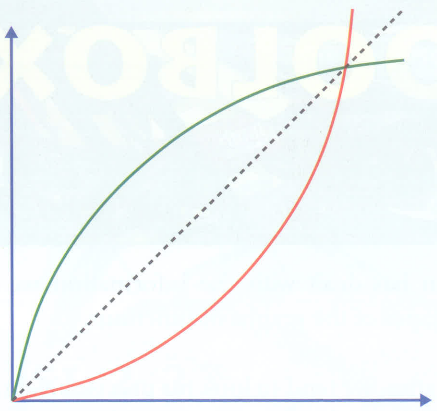
This graph is arranged so that the independent variable is time and the dependent variable is drug concentration. If we think about a time and want to know the concentration, this is the graph we would use.

However, suppose we ask a 'reverse' question, we might choose the inverse graph. A reverse question is one of the type 'when does the drug have a particular concentration'? We might ask this question if we want to know when the drug is within the therapeutic window.

The inverse relation has the variables reversed. It looks like this:



The two graphs are symmetric, as shown in this diagram:



The symmetry is a reflection in a line at 45° to the axes. It has the generic equation '$y = x$'. Note that, as the axes are reversed, the labels are reversed too.

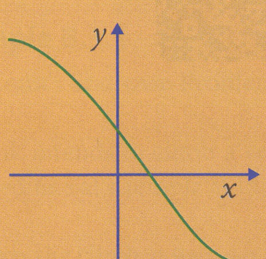
Depending on the course you have chosen to follow, the detailed algebra relating to inverses will need to wait until later in your studies.

At this stage, we will look only at the graphical implications of functions and their inverses.

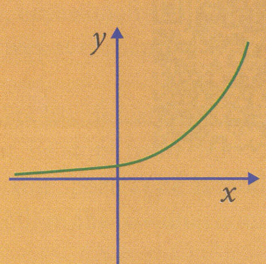
Example B.2.6

Show the inverse of each of these functions. Which of the inverses are also functions?

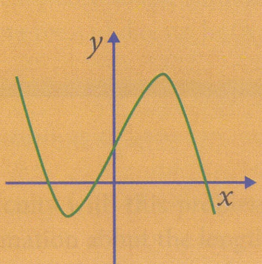
a



b

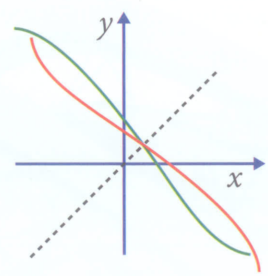


c



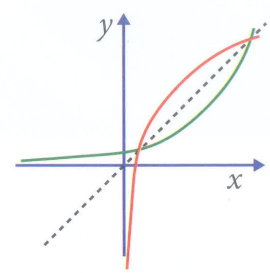
We will answer all these questions by using reflection in the line $y = x$ (black dotted). The inverses are shown in red.

a



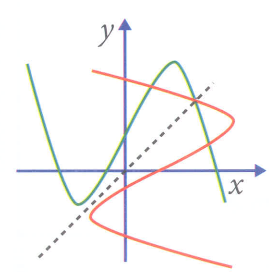
The inverse passes the vertical line test and so is a function.

b



The inverse passes the vertical line test and so is a function. Notice also that the function and its inverse intersect on the line of symmetry.

c

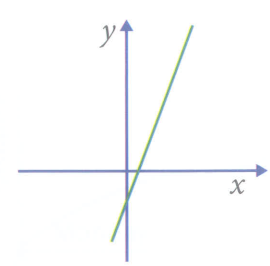


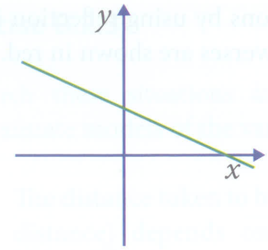
The inverse does not pass the vertical line test and so is not a function.

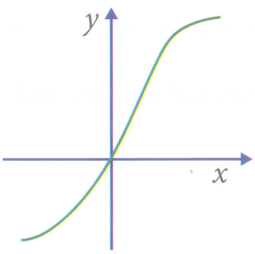
Exercise B.2.4

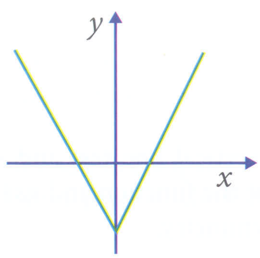
Show the inverse of each of these functions. Which of the inverses are also functions?

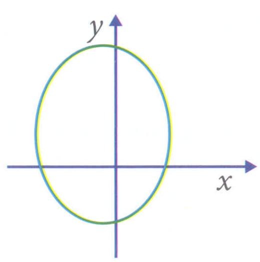
1.

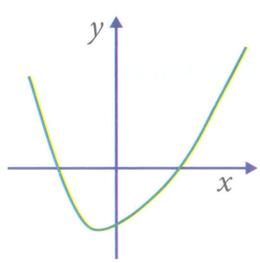


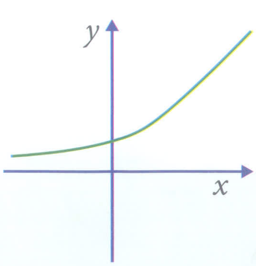
2. 

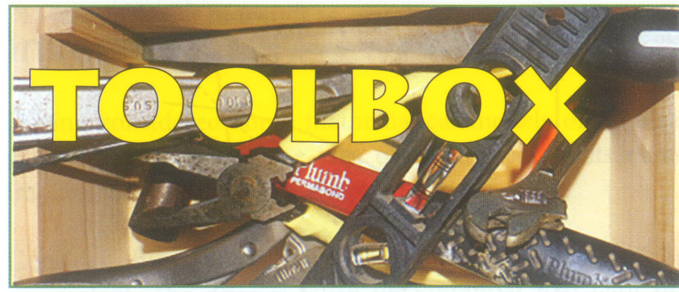
3. 

4. 

5. 

6. 

7. 



This section has dealt with the information we can derive from the shapes of the graphs of functions.

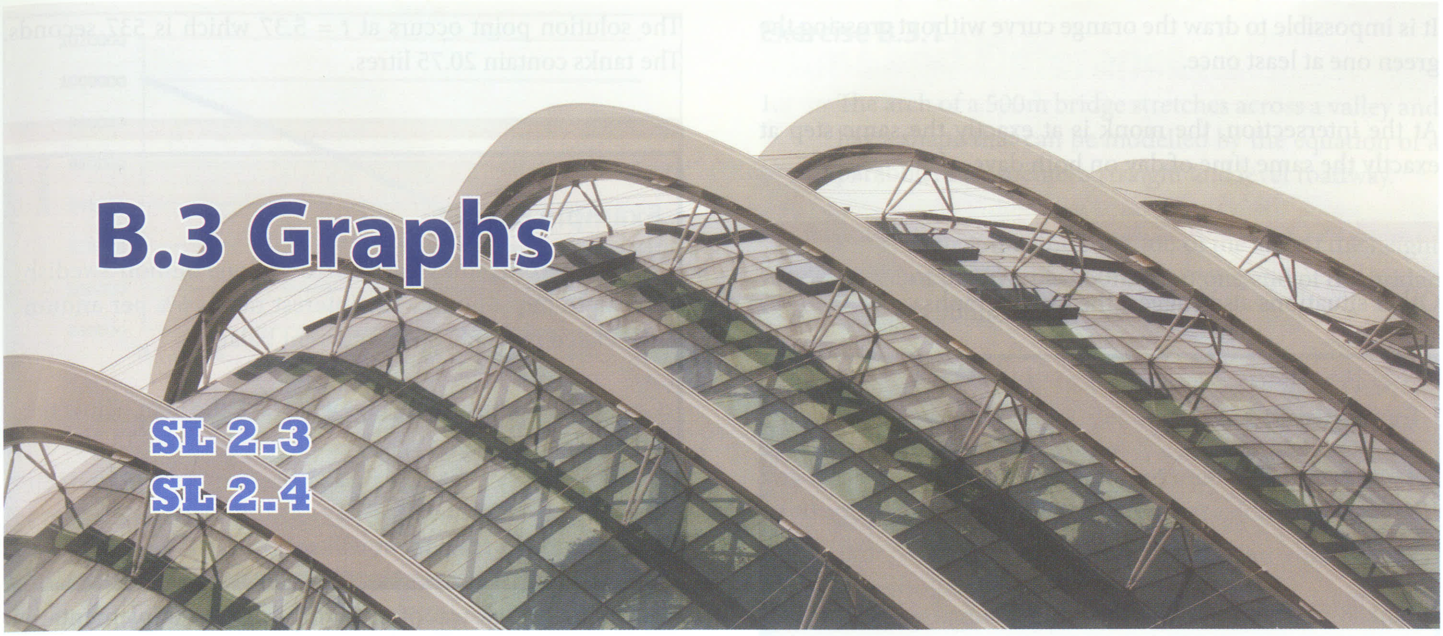
In mathematics, we tend to look for precise algebraic rules to describe functions. However, we have already met a number of situations in which exact rules are not possible. However, useful information can be gathered by looking at the forms of graphs derived from experimental data.

Exercise B.2.3 question 8 suggests that you look at the issues surrounding the generation and consumption of mains electricity.

One good source of information about the 'real time' power supply position of the UK grid is the website:

http://www.gridwatch.templar.co.uk

Answers



B.3 Graphs

SL 2.3
SL 2.4

In the previous section, we dealt with the graphing of functions that are mathematically defined. In this section we will deal with the use of graphs in other contexts such as the processing of experimental data. It is also important to be able to construct graphs that illustrate real situations and to interpret graphs that you encounter in books, magazines, scientific literature etc.

We will begin with a famous old problem.

Example B.3.1

The Monk's Staircase Problem.

A monk intends to climb a very long staircase so that he can meditate at a temple at the top of a hill.

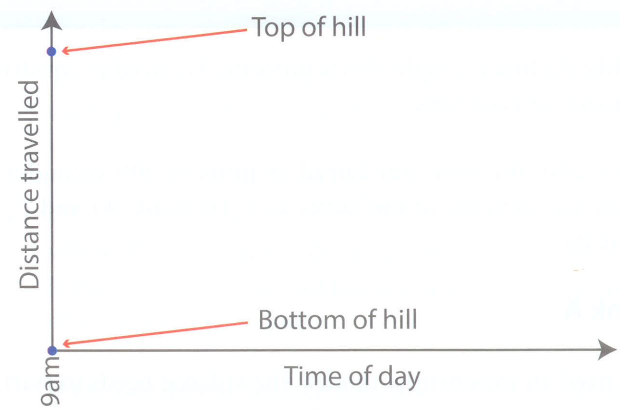
He starts the climb at 9am on Monday. During the day, he climbs, occasionally stopping to rest, admire the view etc. He arrives at the temple at around sunset.

He starts his descent on Tuesday at 9am. Again, on his descent, he stops from time to time. He does not descend at the same speed as he ascended. He arrives at the bottom of the hill in the late afternoon.

Is there a point on the staircase that he reaches at exactly the same time of day on both Monday and Tuesday?

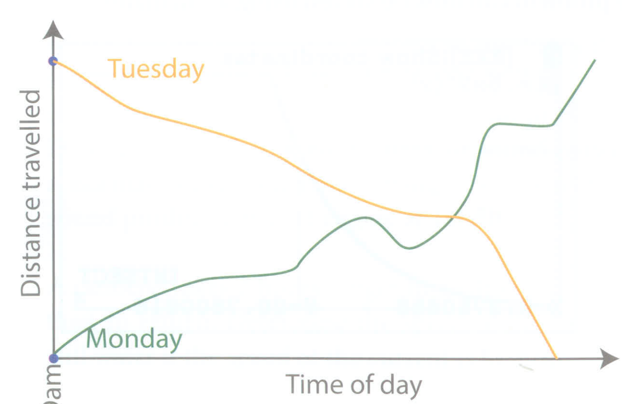
The difficulty with this problem is the lack of data. There is no information about the length of the staircase, the speed at which the monk moves or when and where he rests.

A well chosen graph, however, will clarify things. The axes represent the time of day (horizontal) and the distance travelled along the staircase (vertical).



On Monday (green curve) the monk goes from the bottom to the top across the course of the day. The curve shows variable speed and even one occasion on which the monk went back down a bit.

Tuesday's curve (orange) starts at the top and finishes at the bottom - not necessarily at the same time as Monday's curve.



It is impossible to draw the orange curve without crossing the green one at least once.

At the intersection, the monk is at exactly the same step at exactly the same time of day on both days.

Other situations also suggest the use of graphs.

Example B.3.2

Tank A initially contains 0.5 litres. Water is poured in at a rate that doubles the amount in the container every 100 seconds.

Tank B initially contains 10 litres. Water is added at the rate of 2 litres every 100 seconds.

When will the two tanks contain the same amount of water?

Unlike the first example, this is numeric. We need to construct a model for each tank.

If we label the time (measured in units of 100 seconds) as t and the contents of the tanks as $C_A(t)$ (tank A) and $C_B(t)$ (tank B).

Tank A

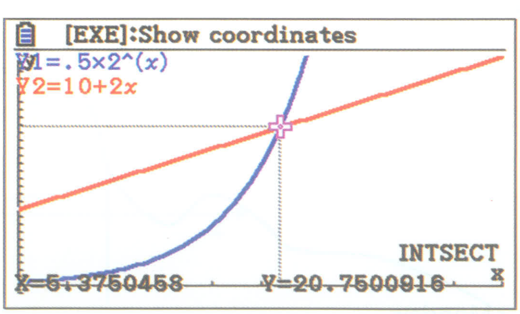
We need an exponential model. The volume needs to start at 0.5, become 1 at $t = 1$, 2 at $t = 2$, 4 at $t = 3$ etc.

The function $C_A(t) = 0.5 \times 2^t$ will do this.

Tank B

This is a linear model $C_B(t) = 10 + 2t$.

The problem can now be solved using a calculator.



The solution point occurs at $t = 5.37$ which is 537 seconds. The tanks contain 20.75 litres.

Example B.3.3

The Ångstrom family have borrowed one million Swedish Krona to buy a house. The interest rate is 6% per annum calculated monthly on the current balance.

Construct a spreadsheet that will track the loan. Show the results on a graph.

The Ångstroms intend to pay off the loan in ten years. What monthly repayment will secure this?

There is not a single way of achieving this. One implementation is:

	A	B	C	D
1	Loan Repayment			
2		Monthly payment		5600
3				
4	Month	Balance	Interest	
5	0	1000000	5000	
6	1	999400	4997	

Enter =A5+1

Enter =B5+C5–D$2

Enter =B5*0.005 and copy to cell C6

Column A records the month, column B the balance (the current level of the debt) and column C the monthly interest (which depends on the balance at the start of the month.

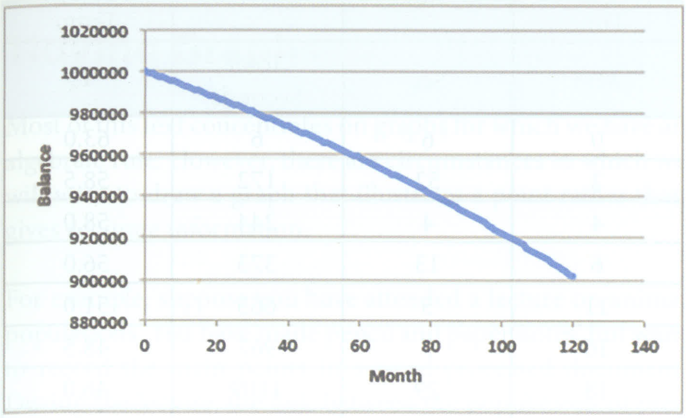
The formulas are:

In cell A6: =A5+1. This increases the month by 1 for each row.

In cell B6: =B5+C5–D$2. This takes the last balance (B5), adds the interest (C5) and subtracts the payment (D$2). The $ sign allows the formula to be copied down the rows with the D$2 element unchanged.

Use 'Fill Down' to copy A6 to C6 down another 120 rows. Look at the way the cells are copied. You should find that the formula in B7 has 'copied' as =B6+C6–D$2. The first two elements have been updated, but the last one (payment) has not. This is known as relative cell addressing.

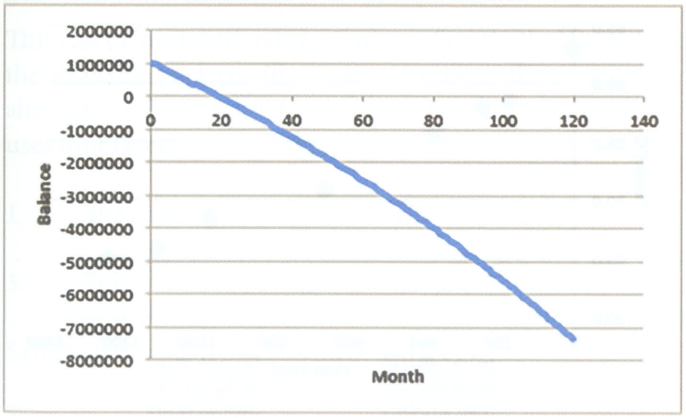
Next, select the relevant data that you want to graph. The question talks about 10 years so you need to go from 0 to 120 months. You need the balances in column B as well. We have chosen an XY scatter diagram with the line only option.



Note that the result is not linear. The debt is also not going to get paid off in ten years because the montly payment of 5600 is too low.

We next alter the value in D2 to try to get the balance at 120 months to be zero. Unless you have exceptional eyesight or a giant screen, you will not be able to see the two relevant cells at once. A way around this is to enter =C125 in cell B3. This enables you to see the payment and final balance next to each other.

Now use 'guess and check' to home in on the solution. As our first guess was way too small, it is a good idea to make the next guess a lot bigger. If you make the payment 10 time bigger, you will get a negative final balance. If you leave the graph active, you will also see the effect graphically.



If you are too timid and make a small increase to, for example, 5700, you will take a long time, even with the technology.

We have bracketed the answer between 5 600 and 56 000 with the likely answer closer to the lower figure. We made our next guess 12 000 and got a final balance of –736. You are probably thinking 'lucky'. Well, yes, but not entirely. A good final answer to the correct payment is 11 100.

Exercise B.3.1

1. The arch of a 500m bridge stretches across a valley and has a shape that can be modelled by the equation of a parabola. It suspends a straight and level roadway.

   a Determine the equation of the arch if the height of the overpass 50m from one end of the bridge is 36m.

   b Find the distance from the top of the arch to the roadway 45m from one end of the bridge.

2. In chemistry, the pH measure of acidity is based on a logarithmic scale and it is given by the formula $\mathrm{pH}=-\log_{10}\left[\mathrm{H}^+\right]$, where $\left[\mathrm{H}^+\right]$ is the concentration of hydrogen ions in moles per litre.

   a Determine the concentration of hydrogen ions for household bleach if its pH is 12.6.

   b The pH for milk is about 6.5. How much stronger in acidity is milk than household bleach?

3. The Richter scale is used to compare the intensity of earthquakes and it is given by:

$$M=\log_{10}\left[\frac{I}{S}\right]$$

   where M is the magnitude reader on the Richter scale, I is the intensity of the earthquake, and S is the intensity of a standard earthquake.

   If an average earthquake measures 4.5 on the Richter scale, how much more intense is an earthquake that measures 6?

4. The formula that is used to measure sound is:

$$L=10\times\log_{10}\left[\frac{I}{I_0}\right]$$

   where L is the loudness measured in decibels (dB), I is the intensity of the sound being measured, and I_0 is the intensity of sound at the threshold of hearing. Determine the change in intensity of sound between a normal conversation measuring at 60dB with the sound produced by a jet engine at 140dB.

5. A canoeist takes 2 hours longer to go 38km up a river than to return. Determine the speed of the canoe in still water if the speed of the current is 5km/hr.

6. A swimming pool can be filled in 4 hours by two pipes working simultaneously. The smaller pipe takes 3 hours longer than the bigger pipe, if each pipe is to pump the water into the swimming pool alone. Determine the exact amount of time required by the bigger pipe to pump the water into the swimming pool.

7. A rope footbridge across a river is modelled by the function:

$$h(x)=6\left(e^{-0.3x}+e^{0.3x}\right),-3\leq x\leq 3$$

where h is the height of the bridge above the mean river level at a point x across the bridge. Both are measured in metres. Find:

i the maximum height of the bridge.

ii the minimum height of the bridge.

8. Use the data in Example B.3.3 to answer these questions:

a How much interest do the family pay over the course of the loan?

b If the family receive a legacy at the start and reduce the starting balance to 800 000, what is the new payment and how much interest will they save?

Graphs from Experimental Data

Frequently, you will use graphs to analyse the results of experiments.

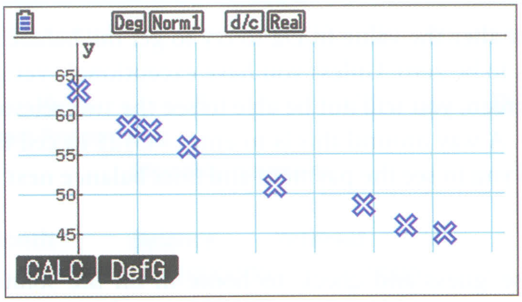
This video shows an experiment on cooling using very rudimentary equipment. The video has been very considerably speeded up. You should pause it periodically in order to take readings of time and temperature.

Cooling video:

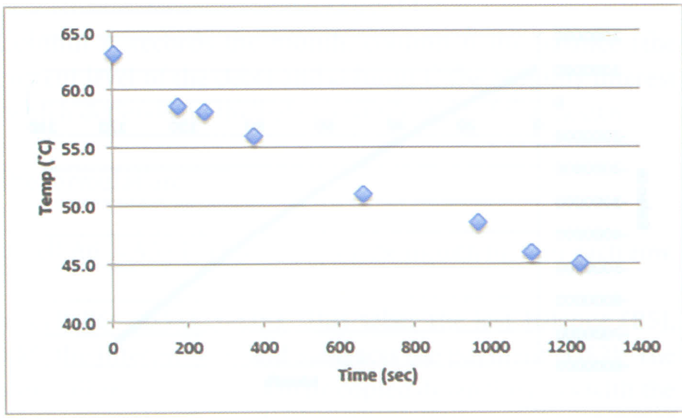
Data can be collected from the video (or preferably your own data) using freeze frame. Here are some.

Time			Temp
Min	Sec	Total Seconds	°C
0	6	6	63.0
2	52	172	58.5
4	4	244	58.0
6	13	373	56.0
11	5	665	51.0
16	7	967	48.5
18	29	1109	46.0
20	40	1240	45.0

These can now be transferred to a graph using graph paper, a calculator,



or a spreadsheet.



From here on, it is possible to make qualitative comments that relate to the shape of the graph and hence to the experiment.

It is, for example, possible to make comments such as:

"The coffee cools most quickly when it is hot"

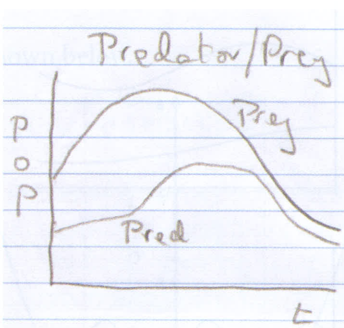
and feel that the graphs supports it (just!).

Tidying Graphs

Most of this text concentrates on graphs for which we have an algebraic rule. However, there are circumstances in which we will want to draw a graph that illustrates a point rather than gives accurate information.

For example, suppose you have attended a lecture on animal populations. You have made pencil and paper notes, but want to record the main points in a word processed document. During the lecture, the link between the populations of prey animals and the predators that feed off them was discussed. The point was made that an increase in the population of prey is often followed, after a delay, by an increase in the population of predators. This is because of the increased food supply.

You have made the following hurried sketch and want to convert it into an electronic document. The first step is to scan it.

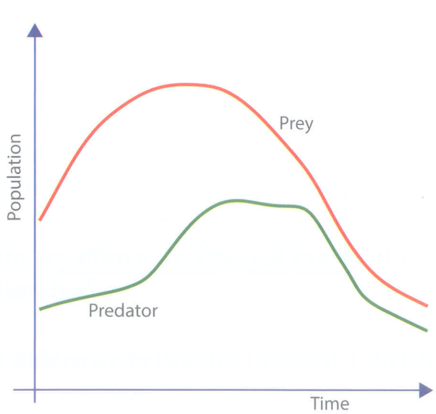


This can be neatened using a variety of methods. We will use the drawing package *Illustrator*. However, there are many alternatives. We find it helps to use layers. Our sample file uses four layers:

1. The scan
2. The axes
3. Text
4. Graphs

Sample file:

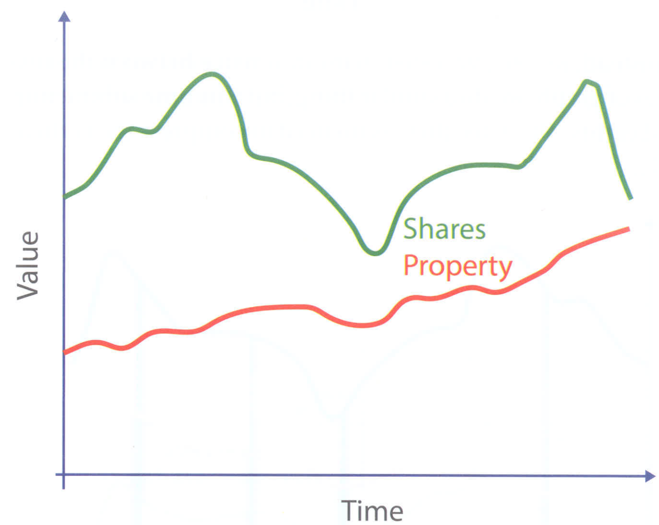
Video:



Sums and Differences

Sometimes we are presented with graphical data without the underlying data. If we are asked to plot the sum and or differences of the data

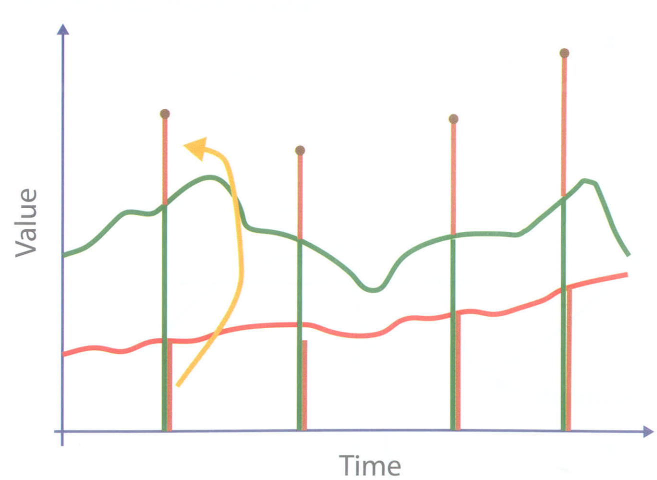
The diagram is typical of the sort of graph you will see in magazines. In this case, the performance of share market investments is being compared with property investments. The complete absence of any numbers makes arithmetic impossible.



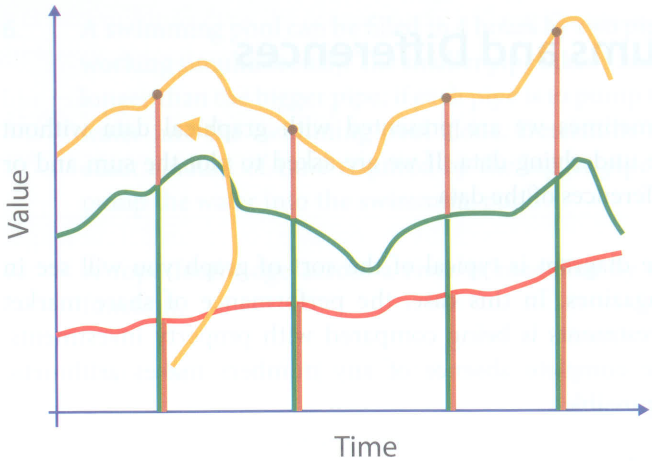
Suppose we are interested in the performance of a portfolio that contains both types of investment. We need to 'add' the two graphs.

Before doing that, we will compress the vertical scale. As we are working in a drawing package (Illustrator), we can use select and compress this graphs.

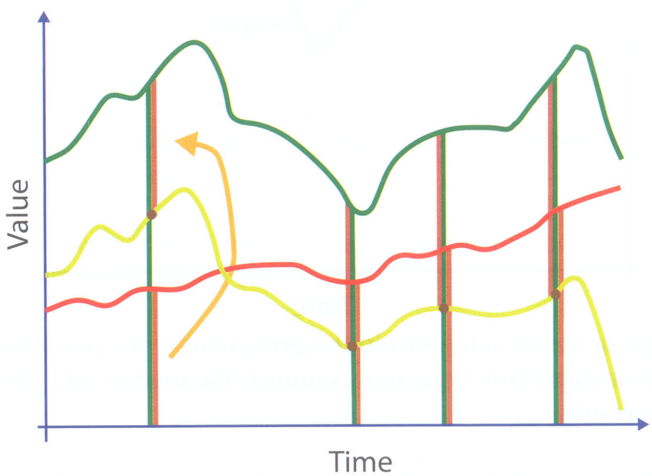
Following this, we proceed point by point across the graph. Use the line tool to measure the two heights and then stack them. We show four examples:



The result (orange) has some of the features of both graphs.



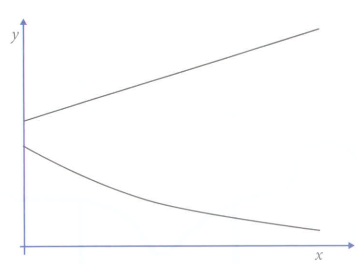
If, instead, we are interested in the difference between the two types of graph, we do a similar thing, but this time subtracting the heights. This time there is no need to compress the vertical scale.



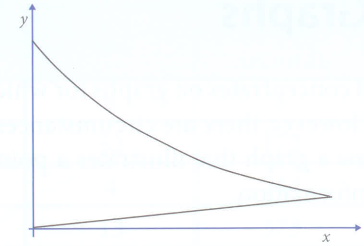
Exercise B.3.2

1. Sketch these graphs and show the sums of the functions and their differences.

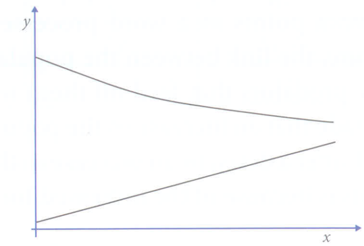
a



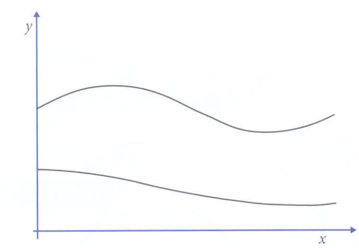
b



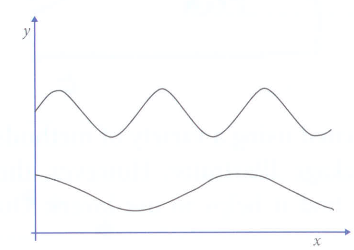
c



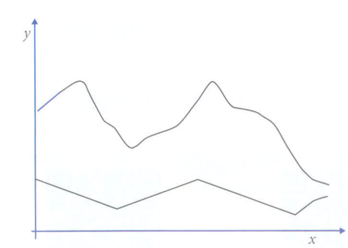
d



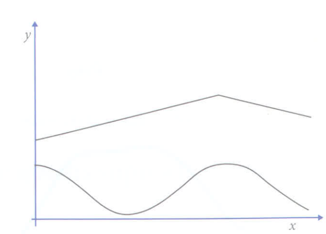
e



f

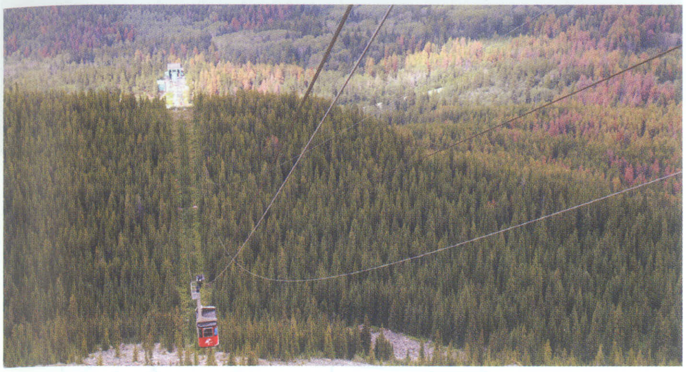


g



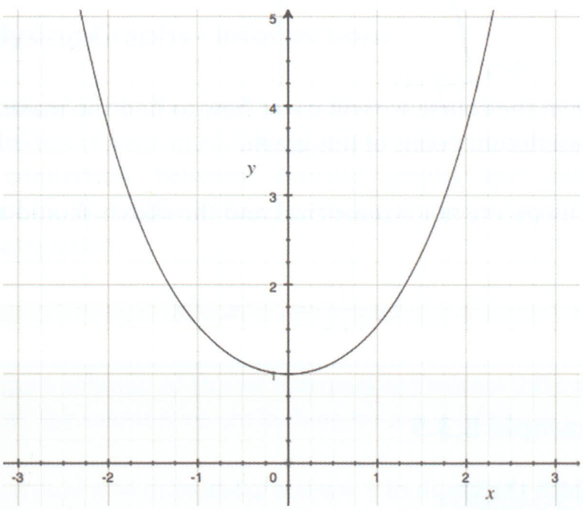
Features of Graphs

A curve followed by a freely hanging wire is known as a catenary.



The function that describes this curve is known as the hyperbolic cosine (or cosh for short). This function is not included in this course, but you will find it on most calculators and graphing software.

The graph is shown below.



Note that this is not all engineers need to know about the catenary if they are designing a cable lift of the type shown in our title photograph. The passenger car will deform the curve significantly. There is a lot more work to do to ensure that the cable lift can operate safely.

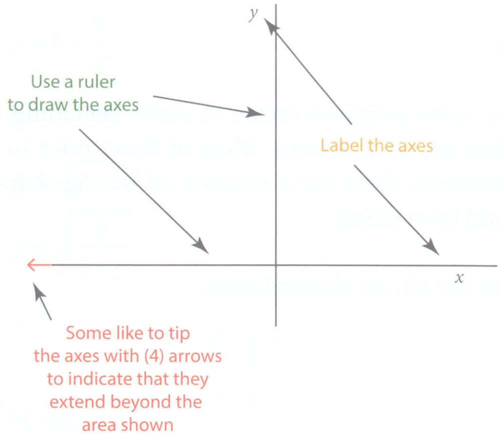
This means remaining clear of the ground, trees etc. under all operating circumstances.

The diagram on the previous page is often known as a 'plot'. The points have been put on the graph with as much accuracy as the computer can muster.

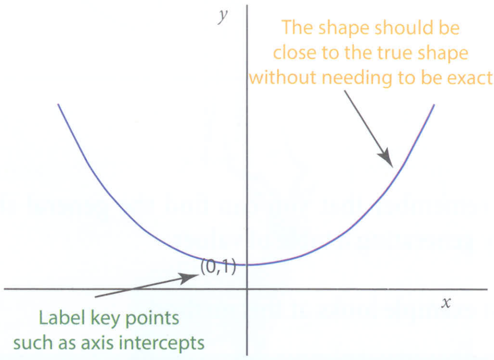
When graphs are drawn in tests and exams, they will almost always be 'sketches'.

What is the difference between a plot and a sketch?

Here is a sketch of the catenary:



Note that the axes labels will not always be x & y. They may be v (for volume) and C (for cost).

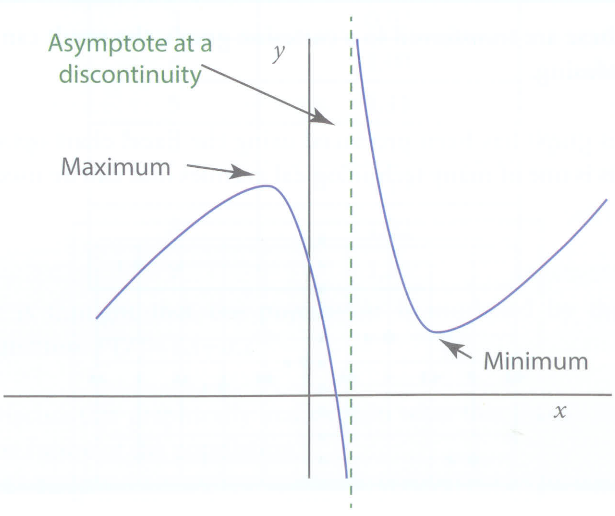


When labelling points on a sketch, it is a good idea to get used to using coordinates. In the present case, the label is (0,1) rather than just 1.

Exam questions often use the phrase 'showing the coordinates of all...' in which case the coordinate pair becomes obligatory.

This example also shows a minimum point at (0,1). Be aware that such points can be important in applications.

Other features that often appear on graphs are maxima/ minima and discontinuities.

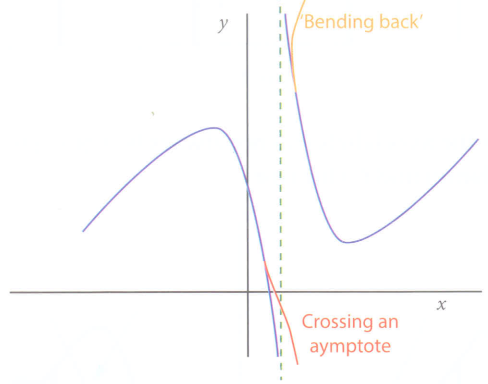


We will see a number of these in the examples.

Pitfalls

There are some common errors in curve sketching that can lead to lost marks in exams. Most of these refer to missing labels. However, there are a number of 'wrong shape errors' that should be avoided.

Some, but not all, are shown below.



Always remember that you can find the general shape of a graph by generating a table of values.

Our first example looks at this method.

Example B.3.4

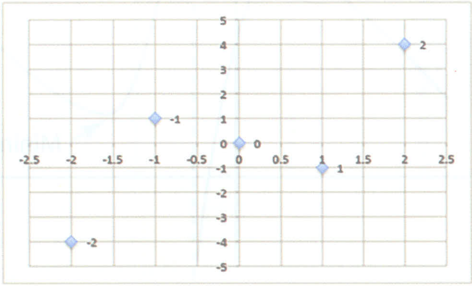
Sketch the graph of $y = x^3 - 2x$, $\quad -2<x\leq 2$.

Note that there is a specified domain ($-2<x\leq 2$). This needs to be reflected in the table:

x	−2	−1	0	1	2
x^3	−8	−1	0	1	8
$2x$	−4	−2	0	2	4
y	−4	1	0	−1	4

If these are transferred to a cartesian graph, the result can be confusing.

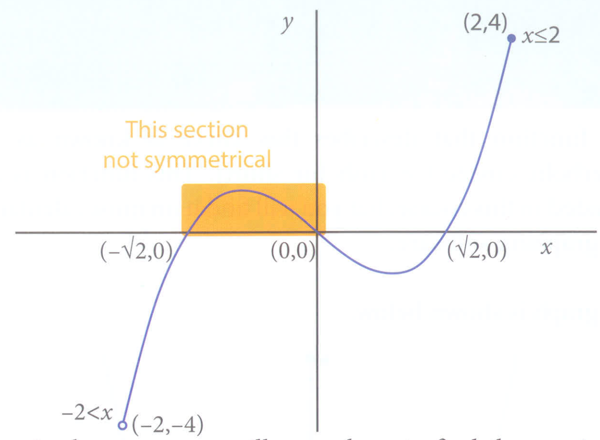
This graph has been produced using the Excel chart option. This is one of many technological avenues that can be used.



The three points on a straight line are a bit of a mystery. If you are using the table of values method, you will need to clarify matters by choosing some extra points in the region of concern.

x	−1	−0.5	0	0.5	1
x^3	−1	−0.125	0	0.125	1
$2x$	−2	−1	0	1	2
y	1	0.875	0	−0.875	−1

The sketch should look like this:



Later in the course we will cover how to find the maximum and minimum points of this graph.

The 'loops' are not symmetrical and the sketch should show this.

Example B.3.5

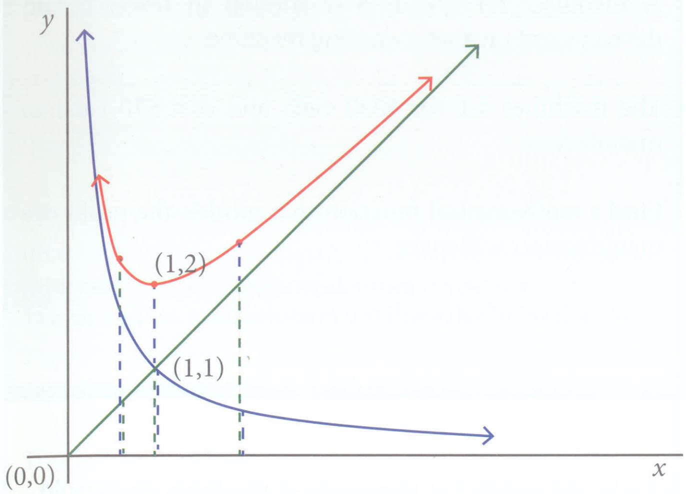
Sketch the graph of $y = x + x^{-1}$, $0<x$.

We will approach this solution from a slightly different viewpoint. The function is the sum of two parts. The first part is $y = x$, a straight line through the origin with gradient 1.

The other function: $y = x^{-1} = \dfrac{1}{x}$ is a hyperbola.

The full function (red) can be found by adding the heights of the blue curve to those of the green line, point by point (as discussed earlier). We have shown three examples of this - the dotted lines show the stacking of the two heights of the graph to produce the sum of the functions. This is not as time consuming as it seems as the green line has small y values for small x. This means that the red function behaves like the blue curve for small x. For large x, the position is reversed because the reciprocal of a large number is small. The red curve approaches the green line for large x.

We have used arrows on the graph to show that the graph extends beyond what is shown on the diagram.



This is an example of a curve (red) being asymptotic to another curve (blue).

Analysing Graphs - Intersections

You should be thoroughly familiar with the graphing capabilities of your model of calculator. This video discusses the connection between drawing graphs and solving equations. We will be looking at finding the positive solution of the equation:

$$x^5 - x^2 - 2 = \frac{1}{x+1} + 5$$

Finding a solution of this by rearranging presents difficulties - try it! The solution we are looking for is $x \approx 1.581$

The method is to draw the graphs of

$y = x^5 - x^2 - 2$ and $y = \frac{1}{x+1} + 5$

Exercise B.3.3

Sketch the graphs of the following functions showing the coordinates of all axis intercepts, maxima and minima.

1. $y = 2 - x$
2. $y = x^2 - x - 2$
3. $y = \frac{1}{x^2}$
4. $y = x^2 - 4$
5. $y = 2^{-x}$
6. $y = 1 - \frac{1}{x}$
7. $y = x^x$
8. $y = \sqrt{x-1}$
9. $y = \sqrt{\frac{1}{2-x}}$
10. $y = 2^{\frac{1}{x}}$

Applications

Mathematical models and their graphs are crucial to many scientific and technical activities.

The first deals with the modelling of animal populations.

Example B.3.6

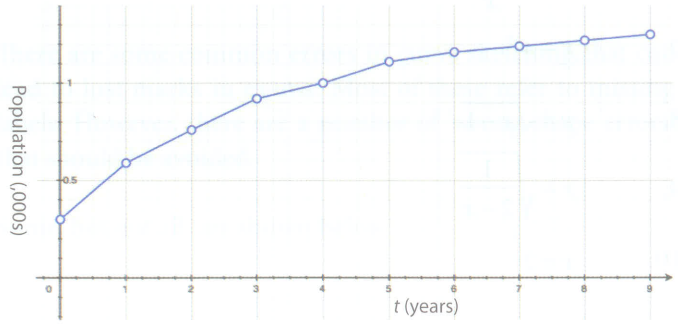
A new game park is stocked with a small population of gazelles. The population of animals is counted annually with the results shown in this table.

Year	Population (1,000s)
0	0.30
1	0.59
2	0.76
3	0.92
4	1.00
5	1.11
6	1.16
7	1.19
8	1.22
9	1.25

It is thought that the population is modelled by the function: $P(t) = 1.3 - 0.2^{0.2t}$

Discuss this graphically and explain what this means for the future of the population.

Graphing the raw data gives:



The straight line segments between the data points are not that useful and are produced by the graphing package used.

The general shape of the graph shows (amongst other things):

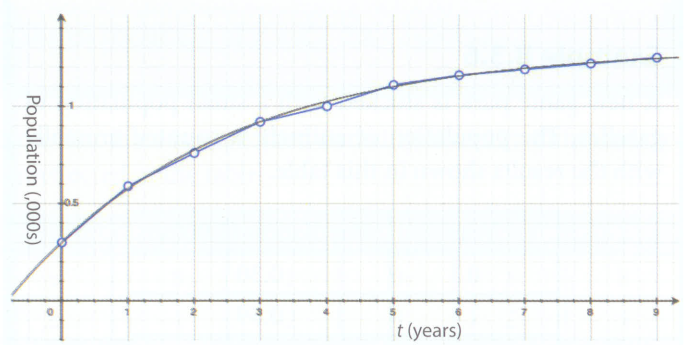
1. The initial 'seeding' population was 300 animals.

2. In the early years, the population rises quite quickly.

3. As time passes, the rise in population slows down.

The third point is probably due to the limited supply of food suitable for the gazelles.

Adding the given function to the graph gives:



The graph of the function fits the data points closely. This means that the modelling function works quite well.

One feature of the mathematical function is that, as t gets large, P approaches 1.3. This is asymptotic behaviour and suggests that the population will stabilise at around 1,300.

Of course, we always have to be careful about predicting the future.

Example B.3.7

A manufacturer spends $1.2 million in research and development on a new washing machine.

The machines sell for $350 each and cost $70 each to manufacture

Find a mathematical function that models the profit the manufacturer will make.

At what level of sales will the manufacturer make a profit?

Let n be the number of thousands of machines made/sold.

Let P be the profit in $1,000s

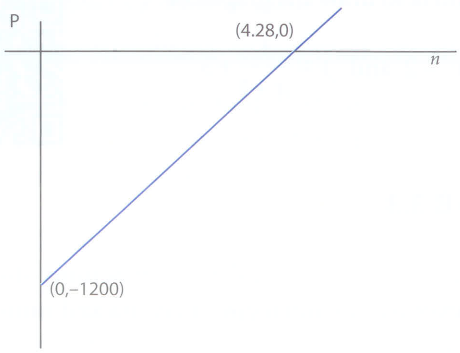
The costs are fixed and, as we are working in thousands of dollars, this is 1,200.

The income is $350 per machine but there is also a cost of $70 per machine. This means that the income is $280 per machine. This is 0.28 thousands. However, we are counting our machines in thousands so, for each thousand machines, we make $280 thousand.

The profit function is:

$$P = -1200 + 280n$$

This is linear:



The point at which the profit changes from negative (ie. a loss) is at approximately $n = 4.28$.

This means that the company will make a profit after it has sold about 4,280 machines.

Example B.3.8

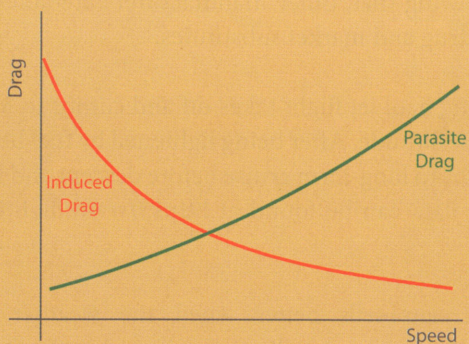
Aeroplanes experience drag due to air resistance. This drag resists the forward motion and is overcome by the engine(s).

There are two types of drag:

Lift induced drag that results from the need to provide lift to keep the aircraft flying. This decreases as the speed increases and the wings work more efficiently.

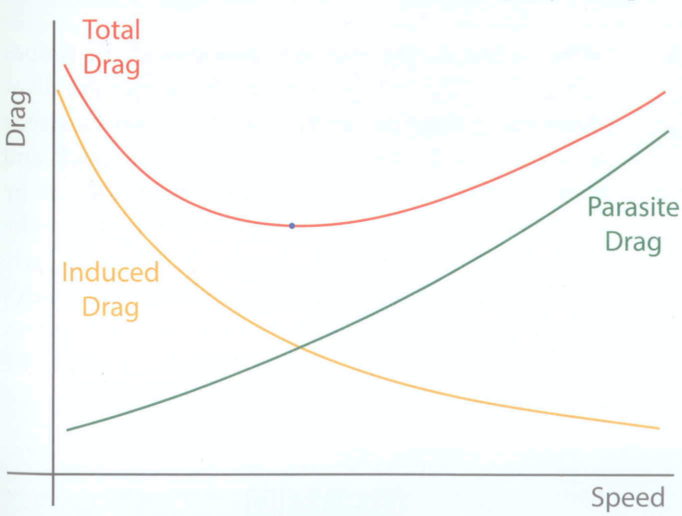
Parasite drag caused by friction as the air passes over the aeroplane. This increases with speed.

This is shown below:



Show the graph of the total drag and comment on its shape.

The total drag is obtained by adding the two functions in the same manner as we did in the shares and property example.



The total drag is at a minimum at the point marked by the blue dot. This is all that is required for comment in an examination. The following is for information.

This minimum is termed the 'glide speed' of the aeroplane.

If pilots experience an engine failure, they first adjust the airspeed to the correct glide speed for the aeroplane type. This maximises the range and the choice of sites for a forced landing.

Example B.3.9

An enzyme catalysed reaction proceeds at a rate that is affected by the temperature.

The table shows some measured rates at a range of temperatures:

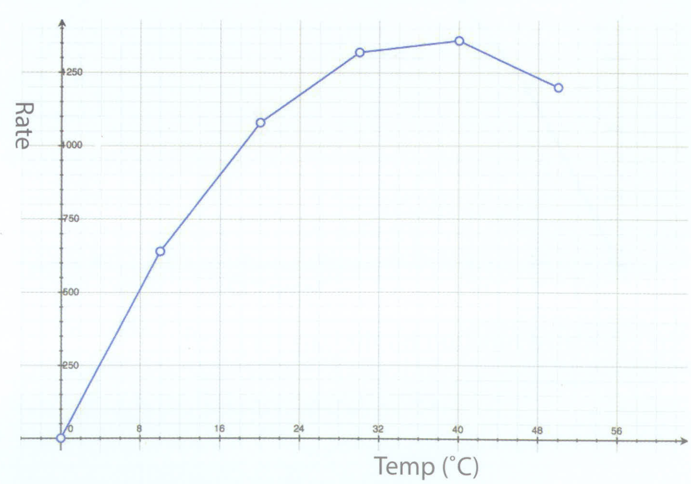
Temperature (˚C)	Rate
0	0
10	640
20	1080
30	1320
40	1360
50	1200

Show these data graphically.

It is hypothesised that the data can be modelled by a function of the form:

$$R(t) = at^2 + bt + c$$

where R is the rate, t is the temperature and a, b & c are constants. Find this function and test its fit to the data.



Note that the straight line segments joining the points are not very informative about this example.

As the graph passes through the origin, we can conclude immediately that $c = 0$.

Choosing two other points: (10,640) & (20,1080):

$10^2a+10b=640$

$20^2a+20b=1080$

$100a+10b=640$ [1]

$400a+20b=1080$ [2]

Halving [2] and subtracting from [1] gives:

$$100a-200a=640-540$$
$$-100a=100$$
$$a=-1$$

Substituting this in [1] gives:

$$-100+10b=640$$
$$10b=740$$
$$b=74$$

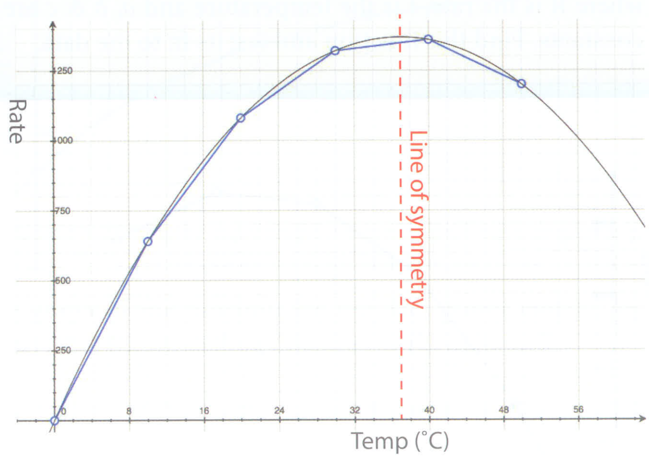
The modelling function is:

$$R(t)=-t^2+74t$$

Since this is a quadratic, it is symmetrical (it is a parabola).

The line of symmetry is temperature = 37°C

This means that the optimum temperature for conducting this reaction is 37°C.



The calculated function fits the data very well.

Exercise B.3.4

Use appropriate technology to draw graphs to illustrate these situations.

1. A population of antelope is small initially. Following good rains, the population grows rapidly until it reaches a new stable level.

2. Heat is lost from a cup of coffee most rapidly when it is freshly made.

3. Daylight hours over a year in Berlin.

4. A drug company has just developed a new treatment for asthma. Legislation guarantees the a period of time during which they alone can manufacture the drug. Following that, other manufacturers are able to copy the drug and market substitutes.

5. A long haul jet flight takes off and climbs to an initial cruise altitude. After being lightened by fuel burnoff, it climbs to long-term cruise before descent and landing. Live data can be found at https://www.flightradar24.com

6. Ice-cream sells better when the weather is hot. A multinational sells ice-cream in the southern hemisphere, the northern hemisphere and in the tropics. Show sales in these three regions and worldwide.

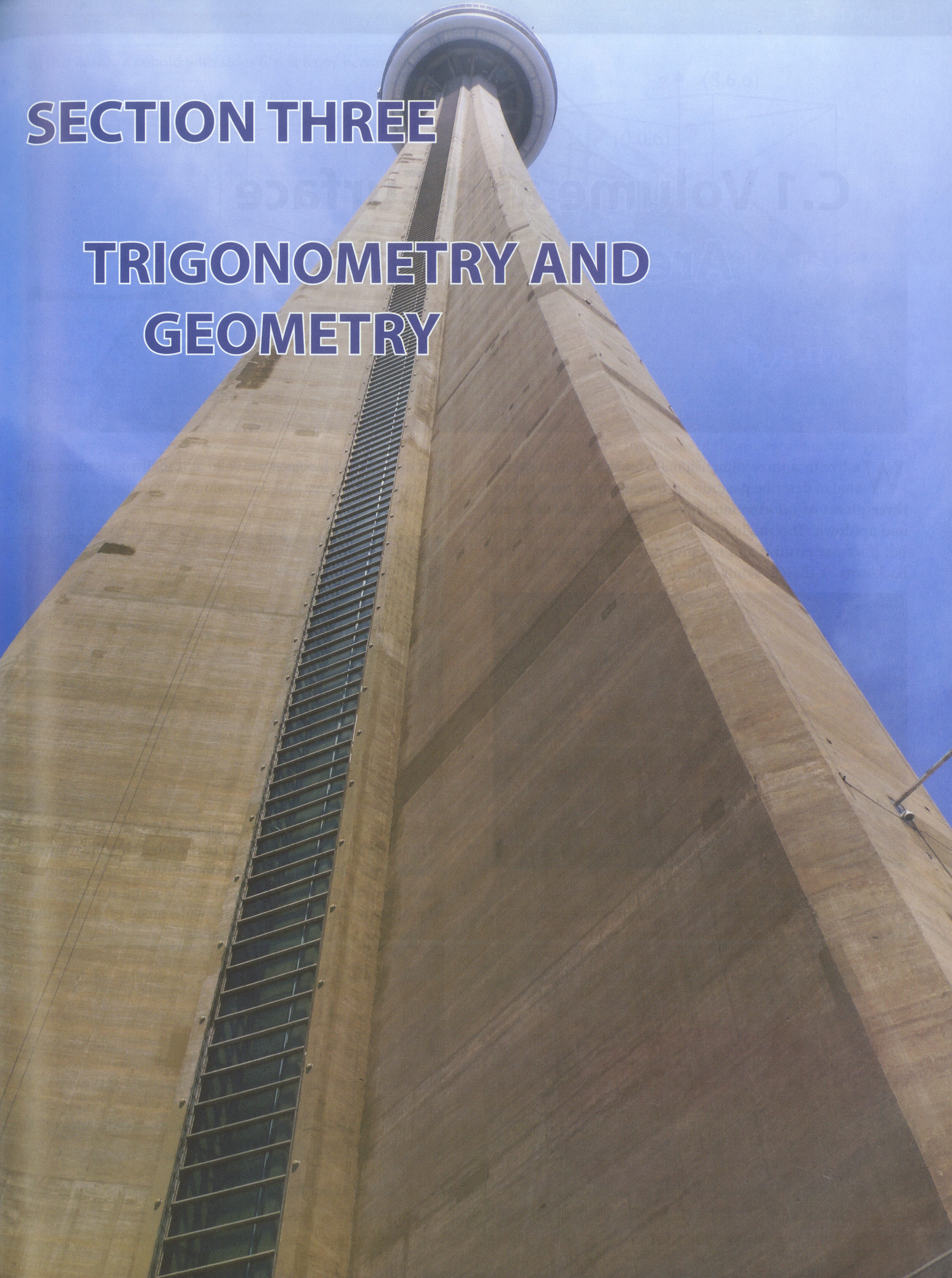
7. A sky-diver jumps from an aircraft and free falls before opening her parachute and descending to a safe landing. Show her vertical speed.

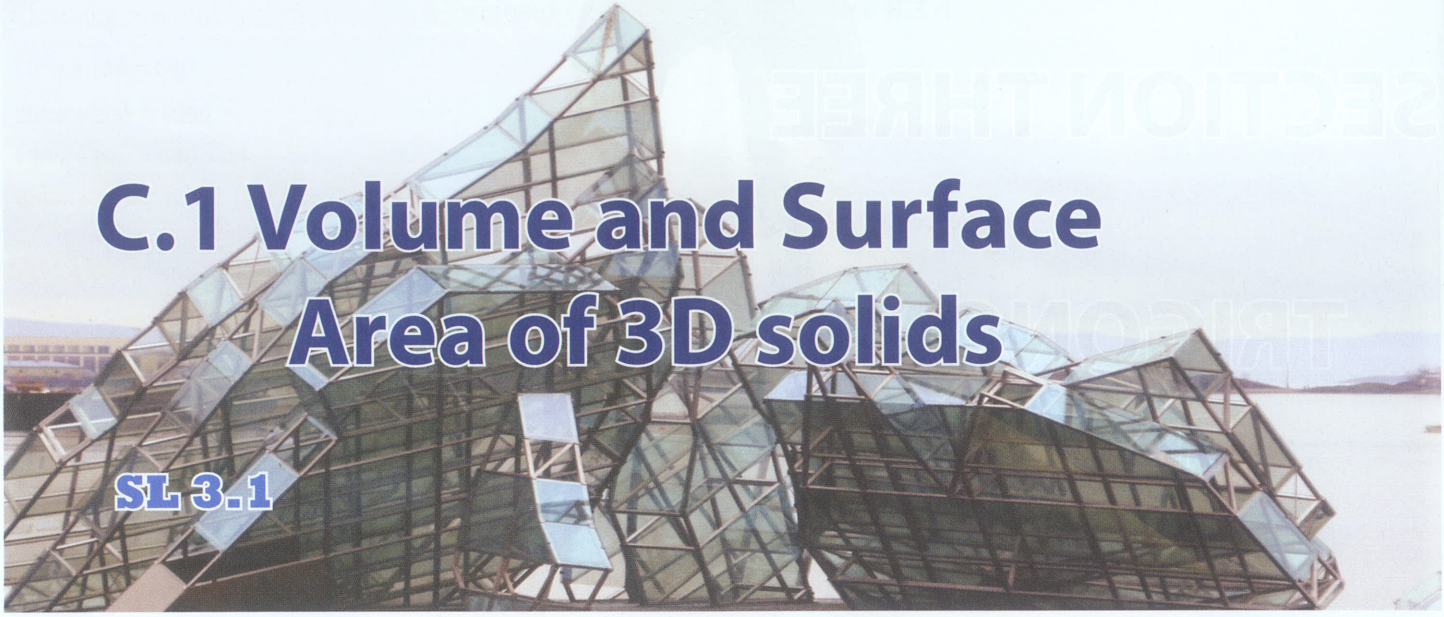
8. When scuba diving, nitrogen dissolves in the tissues of the diver. This rate gets bigger the deeper the diver descends. A diver descends rapidly to explore a wreck at a depth of 35 metres. He explores the wreck and then ascends slowly to a depth of 10 metres where he performs a decompression stop for 10 minutes. The purpose of this is to allow nitrogen to disperse slowly and hence avoid decompression sickness ('the bends') Finally he performs a 3 minute safety stop at 5 metres before returning to the surface. Show on a graph the depth profile and the nitrogen concentration.

Answers

SECTION THREE

TRIGONOMETRY AND GEOMETRY

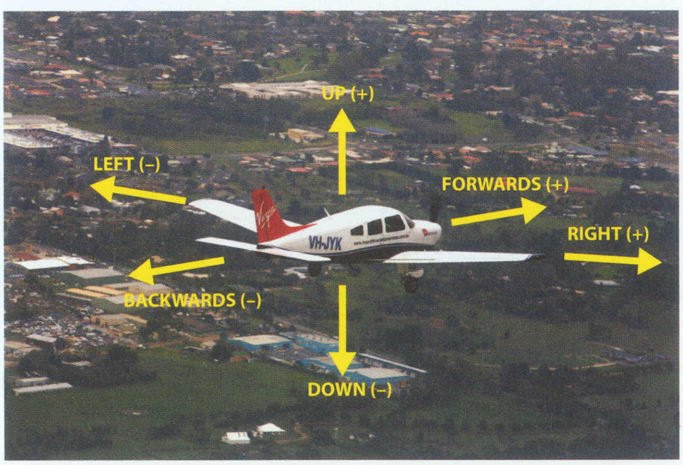




C.1 Volume and Surface Area of 3D solids

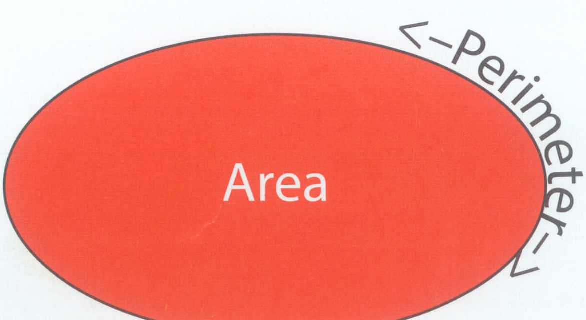
SL 3.1

We live in a three dimensional world (3D). This means that to describe position or movement we need to give three pieces of information: 'forwards/backwards, left/right and up/down. This might look like six pieces of information, but if we use positive and negative numbers (forward + and backward –), then it is only three.



By contrast, Flatlanders (imaginary creatures who slide across flat table tops) live in a two dimensional world (2D).

For Flatlanders, size is measured by area and perimeter.



For the 3D world, objects are solid and their size is measured by their surface area and their volume.

Anish Kapoor's *Cloud Gate* in Chicago, USA, the surface area of the object measures the extent of the job facing the workers who keep it polished. The volume measures the amount of water we would need were we to decide to fill it with water. The other buildings illustrate the wide range of shapes that are possible when constructing solids.

We will begin by looking at some of the simpler solids.

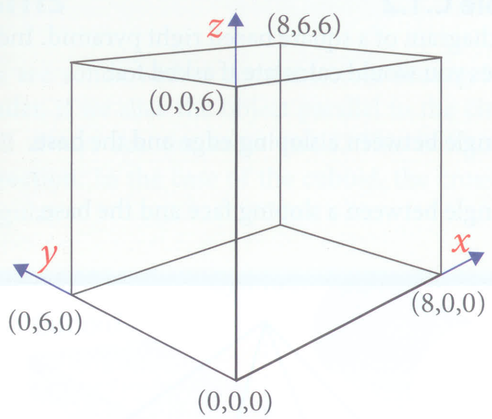
3D Space

There are two ways of describing objects in 3D space.

1. A verbal description such as 'a cuboid with edges 6, 6 & 8 cm'.

2. Using a coordinate system.

In this case, , 'a cuboid with sides 6, 6 & 8 cm' becomes:



Now that we have a coordinate system, we can find mid-points and distances.

Mid-point

The mid-point will be mid-way between two points in each of the *x*, *y* and *z* directions.

The mid-point between (1,3,–2) and (–3, 7,2) is:

$$\left(\frac{1+[-3]}{2},\frac{3+7}{2},\frac{[-2]+2}{2}\right) = (-1,5,0)$$

Distance

Distance is calculated using the Theorem of Pythagoras. In 3D, this becomes:

$$\text{Distance} = \sqrt{x^2+y^2+z^2}$$

The distance between (1,3,–2) and (–3, 7,2) is:

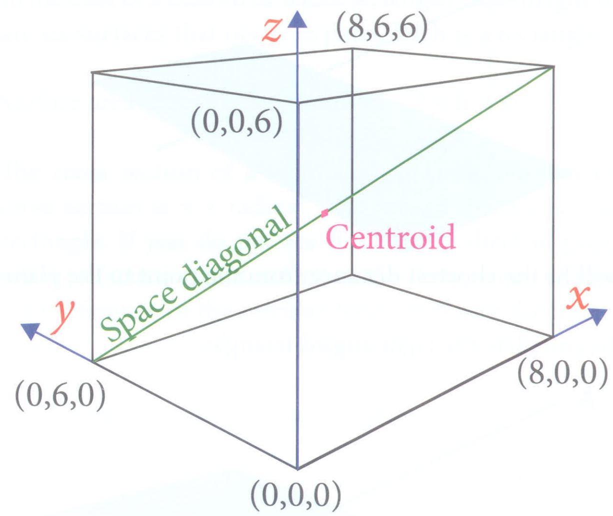
$$D=\sqrt{(1-[-3])^2+(3-7)^2+([-2]-2)^2}$$
$$=\sqrt{4^2+(-4)^2+(-4)^2}$$
$$=\sqrt{3\times 16}$$
$$=4\sqrt{3}$$

Example C.1.1

Find the distance from the centroid of a cuboid with sides 6, 6 & 8 cm to one of the corners.

The centroid is the point at which the space diagonals of the cuboid intersect. The space diagonals are straight lines joining opposite corners of the cuboid.

Begin with a diagram:



The space diagonal shown in green joins the points (0,6,0) and (8,0,6). Taking the means of these coordinates, we get the mid-point or centroid (4,3,3).

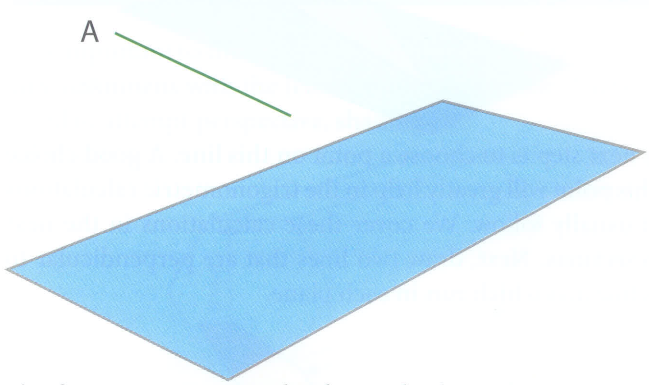
Using Pythagoras to find the distance between the centroid (4,3,3) and the corner (0,6,0):

$$D=\sqrt{(4-0)^2+(3-6)^2+(3-0)^2}$$
$$=\sqrt{4^2+(-3)^2+3^2}$$
$$=\sqrt{34}$$

Angle between a Line and a Plane

The purpose of this section is to define what is meant by the angle between a line and a plane. Methods of calculating such angles will be covered in the next two sections.

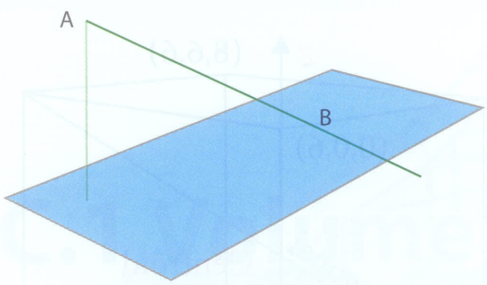
A line segment will not necessarily intersect with a plane. Even if this is the case, there is still an angle between the two.



The first step is to extend either or both until the line passes through the plane.

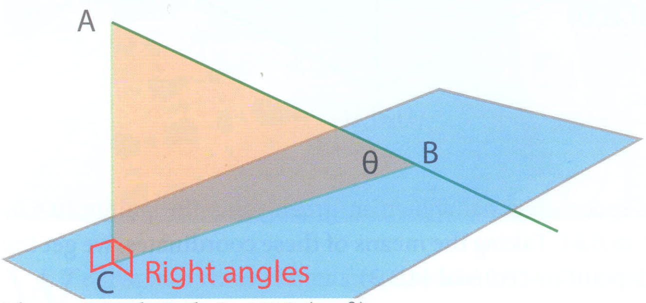
The only circumstance in which this will not be possible is if the line and the plane are parallel.

It is also necessary to drop a perpendicular from a point on the line (A) to the plane.



This will be the shortest distance from the point to the plane.

Finally, complete the right angled triangle:

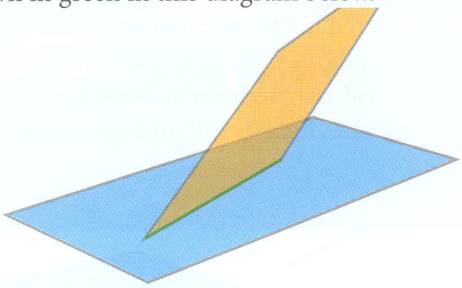


The required angle is ABC (or θ).

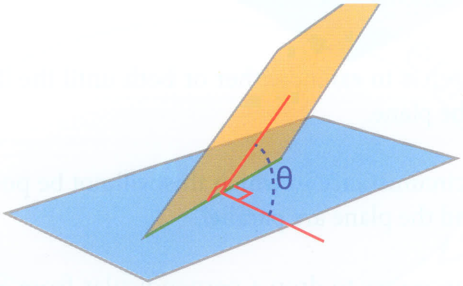
Angle between Two Planes

As with the previous section, the point here is to correctly identify the angle that is meant.

It is possible for two planes to be parallel - they never intersect even if they are indefinitely extended. A table top and a floor are, in theory, like this. Usually, planes will intersect if extended as far as necessary. The intersection will be a straight line (shown in green in this diagram below.



The next step is to choose a point on this line. A good choice of this point will greatly help in the trigonometric calculations that usually follow. We cover these calculations in the next two sections. Next, draw two lines that are perpendicular to this line and which run in each plane.



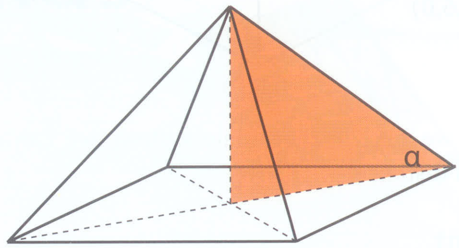
The required angle is θ.

Example C.1.2

Draw a diagram of a square based right pyramid. Indicate the angles you would calculate if asked for:

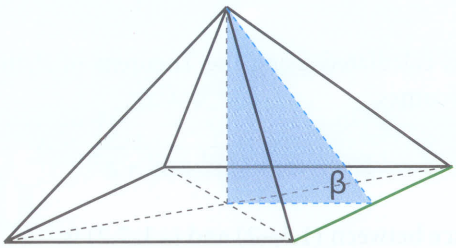
a The angle between a sloping edge and the base.

b The angle between a sloping face and the base.



a It makes sense to drop the perpendicular from the apex of the pyramid to the centre of the square base. Complete the red triangle and identify the angle α.

b We look for an edge where the base and the sloping face intersect (green). The best choice of a point on the edge is its mid-point. Then complete the blue right angled triangle and identify angle β.



Exercise C.1.1

Draw diagrams of the following solids. Indicate angles that you would use to find:

1. The angle between adjacent faces of a regular tetrahedron.

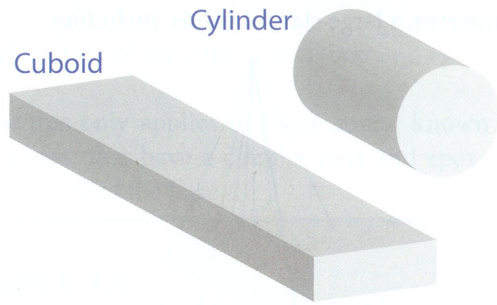
2. The angle between an edge and a face of a regular tetrahedron.

3. The angle between adjacent faces of a regular octahedron.

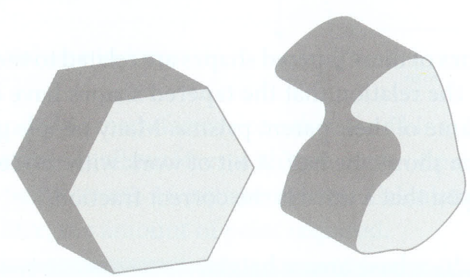
4. A polyhedron has a regular hexagonal base. The apex is vertically above one of the base vertices. You are asked to find the smallest angle between a sloping edge and the base.

Prisms

Prisms are solids with a constant cross section.In the case a cylinder, if we slice the object parallel to the circular face, we will always see the same circle. This is known as the **cross section**. In the case of the cuboid, the cross section is rectangular.



Not all prisms are this simple. Any shape with a constant cross section counts as a prism.



Volume

The volume all prisms is found by multiplying the area of the constant cross section (often called the **base**) by the length (often called **height**).

This leads to two frequently used formulas:

Volume of a cuboid = length × width × height

Volume of a cylinder = π × radius of base² × height.

Both of these conform to the ' base area × height' formula already referred to.

Other prisms, such as the hexagonal prism and the other example above may need special techniques to find the base area.

Surface Area

The various prisms require different techniques to calculate their surface areas. In each case, the areas of the various faces are calculated separately.

In the case of a cuboid of width w, length l and height h, there are six surfaces that occur in pairs. Each is a rectangle.

Surface area $= 2 \times l \times w + 2 \times l \times h + 2 \times h \times w$

The cross section of a cylinder is a circle. The area of the cross section is π × radius². The curved surface is actually a rectangle. If you do not see this, take a sheet of paper and roll it to form a cylinder. One of the dimensions of this is the circumference of the circular base ($2\pi r$) and the other is the height.

If a cylinder has radius r and height h:

Surface area of the closed cylinder: $2\pi r^2 + 2\pi rh$.

When cylinders appear in questions, you need to read the wording carefully. Sometimes you are asked about closed cylinders (with a top and a bottom). At other times the cylinder may be open at one end, as in rubbish bins. In this case, use $\pi r^2 + 2\pi rh$. Pipes are, of course, open at both ends.

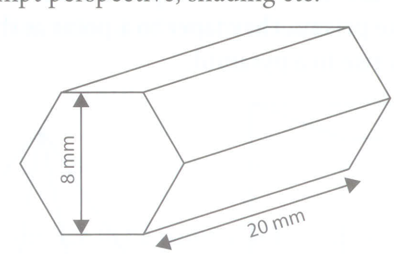
The hexagonal prism is covered in the following example. For other cases, we need to identify each face, use an appropriate method to find its area and then add. In the case of the curvy prism, this will present a very difficult problem!

Example C.1.3

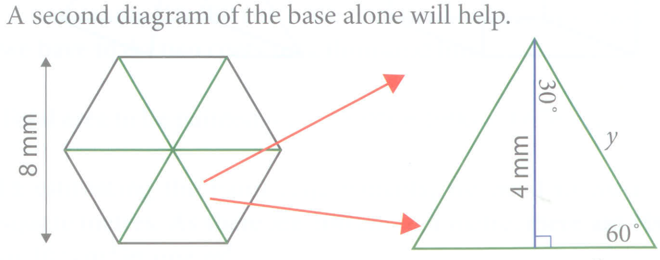
A tool consists of a hexagonal prism made of solid steel. The width of the tool as measured between opposite parallel faces is 8 mm. The length is 20 mm.

Find the volume and surface area of the tool, correct to 4 significant figures.

It is important to draw a good diagram, add the data and label any unknowns with the letters you intend to use. There is no need to attempt perspective, shading etc.



A second diagram of the base alone will help.



Note that $y = 2x$, as the triangle is equilateral.

Using trigonometry in the right angled triangle:

$$\frac{4}{y} = \sin 60^\circ$$

$$\frac{4}{y} = \frac{\sqrt{3}}{2}$$

$$y = \frac{8}{\sqrt{3}} = \frac{8\sqrt{3}}{3}$$

As already observed, $y = 2x$, so $x = \frac{4\sqrt{3}}{3}$ mm.

The area of one of the green triangles

(using area = ½base×height) = $\frac{16\sqrt{3}}{3}$

The base of the hexagon consists of six of these triangles.

$$\text{Base area } = 6 \times \frac{16\sqrt{3}}{3} \simeq 55.43 \text{ mm}^2$$

$$\text{Volume } = 6 \times \frac{16\sqrt{3}}{3} \times 20 \simeq 1109 \text{ mm}^3$$

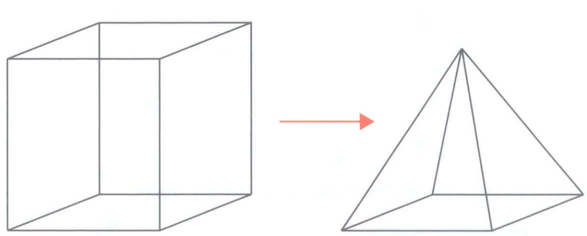
Since the tool is solid, we must find the surface area of each of the six rectangular faces and the two ends (already done).

The rectangular faces are each $20 \times \frac{8\sqrt{3}}{3}$ mm^2

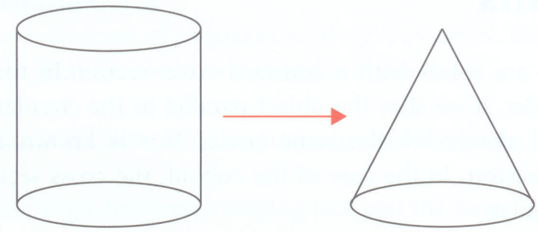
$$\text{Total surface area } = 2 \times 6 \times \frac{16\sqrt{3}}{3} + 6 \times 20 \times \frac{8\sqrt{3}}{3} \simeq 665.1 \text{ mm}^2$$

Cones and Pyramids

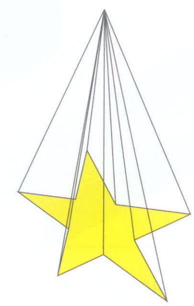
These shapes are derived from prisms. Instead of 'going straight up' like prisms, they taper to a point at the top. Thus, a cuboid gives rise to a pyramid.



The cylinder gives rise to the cone:



This is, however, a large class of solids, including:



Volume

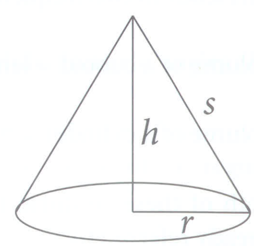
The volumes of these tapered shapes are related to their parent prisms by the relation that the tapered shapes have one third of the volume of their parent prisms. Many people guess that the fraction should be half. A bit of work with models should convince you that a third is the correct fraction.

Volume = $^1/_3$ × base area × height.

Surface Area

As with prisms, surface areas are found by treating each face separately. The pyramid, for example, is made from a rectangle or square and four triangles.

The cone is handled as follows.



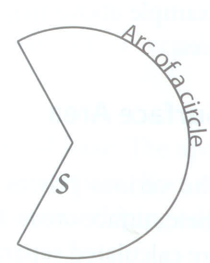
r = radius of the base

h = height of cone

s = slant height of cone.

By Pythagoras: $s^2 = h^2 + r^2$.

The base is a circle ($A = \pi r^2$).



If the curved surface is opened up, it forms a fan shape. Cut one out for yourself and you will see that it rolls up to form a cone.

The arc rolls up to become the circumference of the base of the cone ($2\pi r$).

The area of this fan shape is a fraction of the complete circle ($A = \pi s^2$ and $C = 2\pi s$). Note that the radius of this circle is s, not r. The fan is only a fraction of this circle. But what fraction is it? The arc rolls up to form the circumference of the base. This is $2\pi r$.

The fraction is $\dfrac{2\pi r}{2\pi s} = \dfrac{r}{s}$

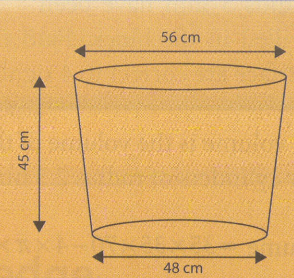
Area of fan $= \dfrac{r}{s} \times \pi s^2 = \pi r s$

Total surface area of the cone $= \pi r^2 + \pi r s$.

Note that this only applies to those cones, known as **right circular cones**, that have a circular base and apex vertically above the centre of the base.

Example C.1.4

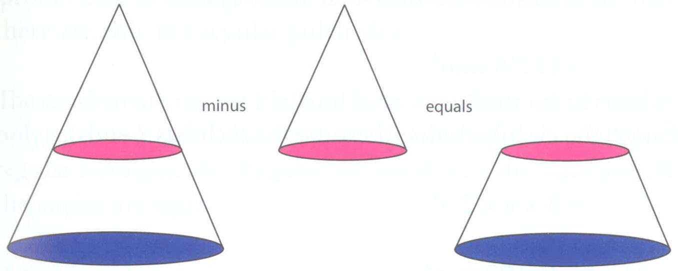
The diagram shows a design for an office recycle bin.



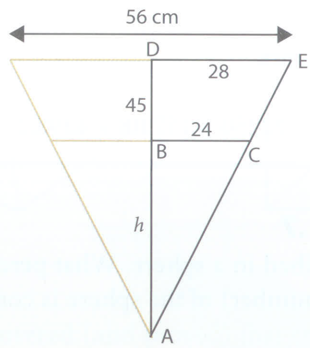
Find its volume in litres.

It is intended to manufacture one thousand bins. They are to be treated on all surfaces with two coats of paint. One litre of paint covers 16 square metres. Calculate, to the nearest litre, the amount of paint required.

The bin is essentially two cones, one of which is removed.



At the moment, we do not know the height of either cone. This is a problem of similar triangles. The diagram, not drawn to scale is:



Triangles ABC and ADE are similar because they share angle BAC and they both have right angles. It follows that corresponding sides are in proportion.

$$\frac{AB}{AD} = \frac{BC}{DE} \Rightarrow \frac{h}{h+45} = \frac{24}{28} \Rightarrow h = 270$$

The smaller cone thus has a base radius of 24 cm and a height of 270 cm. The figures for the larger cone are $r = 28$ & $h = 315$.

The volumes are:

Smaller cone: $V_s = \frac{1}{3}\pi r^2 h = \frac{1}{3}\pi \times 24^2 \times 270 \approx 162860 \text{ cm}^3$.

Larger cone: $V_l = \frac{1}{3}\pi r^2 h = \frac{1}{3}\pi \times 28^2 \times 315 \approx 258616 \text{ cm}^3$.

The volume of the bin is the difference between these figures or 95 756 cm^3. Remember that full calculator accuracy should be carried through this calculation. Situations in which two large and similar numbers are subtracted contain dangers that rounding errors can be exaggerated.

There are 1 000 cm^3 in a litre and we need to note that no accuracy has been specified, so we should round the final answer to a level similar to that of the data. As all the data is given to 2 significant figures, the bin will hold 96 litres.

Surface area. This calculation is treated similarly to the volume problem. The slant heights are found using Pythagoras.

Large cone: $s_L = \sqrt{28^2 + 315^2} \approx 316.24$ cm

Small cone: $s_S = \sqrt{24^2 + 270^2} \approx 271.06$ cm

We are, at this stage, only interested in the curved surface which can be calculated by subtracting, as with the volume.

Large cone: $A_L = \pi r s = \pi \times 28 \times 316.24 \approx 27818 \text{ cm}^2$

Small cone: $A_S = \pi r s = \pi \times 24 \times 271.06 \approx 20438 \text{ cm}^2$

Area of curved surface $\approx 27818 - 20438 \approx 7380 \text{ cm}^2$

The bin also has a circular bottom of area:

$A = \pi r^2 = \pi \times 24^2 \approx 1810$

The total area is 7380 + 1810 ≈ 9190 cm^2

However the surface to be painted is both inside and out and is double this (18 380 cm^2). We also need to bear in mind that we have to do two coats on a thousand bins.

Total area to be painted = 2 × 18380 × 1000 ≈ $3.676 \times 10^7 \text{cm}^2$.

In estimating the paint requirements, we need to work in square metres. As there are 100 cm in a metre, there are 100^2 or 10^4 cm^2 in one m^2.

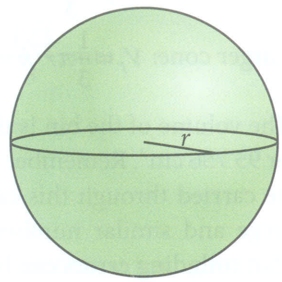
The area to be painted is 3.676×10^3 m². As each litre covers 16 m², we will need $3.676 \times 10^3 \div 16 \approx 230$ litres. As the data is only 2 s.f., the answer should be given to a similar accuracy.

Other Important Solids

Sphere

$$V = \frac{4}{3}\pi r^3$$

$$A = 4\pi r^2$$



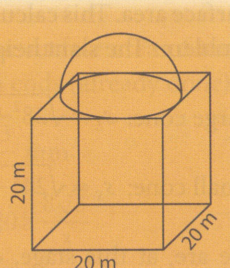
Composite solids

These are combinations of solids made by sticking parts together or by drilling holes etc.

Example C.1.5

A design for a new building consists of a 20 m cube topped by a hemispherical dome.

Find its volume and surface area.



The volume of the cube is $20^3 = 8\,000$ m³.

The volume of the hemisphere (half a sphere) is:

$$V = \frac{1}{2}\left(\frac{4}{3}\pi r^3\right)$$

$$= \frac{1}{2}\left(\frac{4}{3}\pi 10^3\right)$$

$$\approx 2094 \text{ m}^3$$

Total volume $\approx 8\,000 + 2\,094 = 10\,094$ m³.

Surface area.

Some assumptions need to be made here such as that the face of the cube that is on the ground is not included. Making that assumption, there are four 20 by 20 square faces.

Remember also that the dome is half a sphere of radius 10 m.

The flat part of the roof is a square minus a circle.

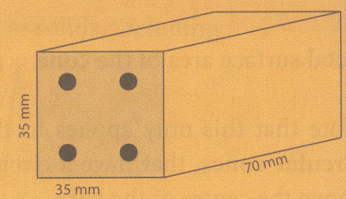
$$\text{Surface Area} = 4\times 20\times 20 + \frac{1}{2}\times 4\pi\times 10^2 + \left(20\times 20 - \pi\times 10^2\right)$$

$$\approx 2314 \text{ m}^2$$

Example C.1.6

An engine component consists of a steel cuboid 35 by 35 by 70 mm.

Four circular holes of radius 5 mm have been drilled in the block. Find the component's volume and surface are correct to 4 significant figures.



The volume is the volume of the cuboid minus the volume of four cylinders of radius 2.5 mm and length 70 mm.

$$\text{Volume} = 35\times 35\times 70 - 4\times\pi\times 2.5^2\times 70$$

$$\approx 80252 \text{mm}^3$$

The surface area is made up of three parts:

Part 1, the surface of the cuboid

$= 2 \times (35 \times 35 + 35 \times 70 + 35 \times 70)$

$= 12250$ mm².

Part 2, the circular holes (there are 8 - each has 2 ends).

$= 8 \times \pi \times 2.5^2$

≈ 157.08 mm²

Part 3, the interior surfaces of the four holes (added).

$= 4 \times \pi \times 5 \times 70$

≈ 4398.2 mm²

Surface area $= 12250 - 157.08 + 4398.2$

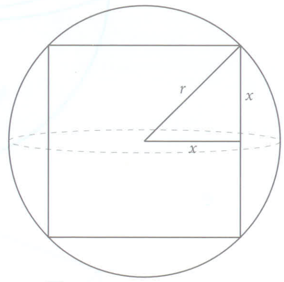
Example C.1.7

A cube is inscribed in a sphere. What percentage (to the nearest whole number) of the sphere is contained within the cube?

The question gives no numerical values. Rather than assuming that it does not matter, we will let the radius be r.

Let the side of the cube be $2x$.

By the Theorem of Pythagoras:



$$x^2 + x^2 = r^2$$

$$x = \frac{r}{\sqrt{2}}$$

The cube has side length $\frac{2r}{\sqrt{2}} = r\sqrt{2}$

The volume of the cube is: $\left(r\sqrt{2}\right)^3 = 2\sqrt{2}r^3$.

The volume of the sphere is: $\frac{4}{3}\pi r^3$

The fraction of the sphere occupied by the cube $\frac{2\sqrt{2}r^3}{\frac{4}{3}\pi r^3} = \frac{3\sqrt{2}}{2\pi}$

The percentage $= \frac{3\sqrt{2}}{2\pi} \times \frac{100}{1} \approx 68\,\%$

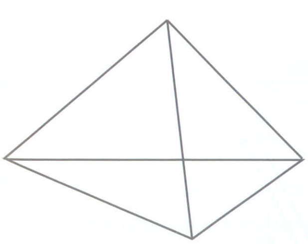
Theory of Knowledge

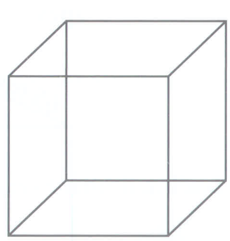
The Ancient Greeks

Greek Mathematics flourished from around 600 to 200BC. They were as much interested in form (shape) and pattern as they were in numbers. The school developed the notion of 'proof'. One of their greatest achievements was the proof that there are only five regular polyhedra.

The mathematicians of Flatland have an infinite set of regular polygons: equilateral triangle, square, regular pentagon, regular hexagon, etc. 'Regular' means that all the sides and all the angles are equal.

If mathematicians look for analogues in the 3D world, they find the regular polyhedra. The first two of these are the tetrahedron (4 equilateral triangle faces) and the cube (6 square faces).





The Greeks discovered (and proved) that there are only five such solids. Can you see why this is so?

Platonic Solids Video

https://www.youtube.com/watch?v=RbbaGGmaO6U

Biology - On Being the Right Size

Mathematics affects the forms of animals. At a very approximate level, the weight of an animal is proportional to its volume. Its strength is proportional to its surface area. This is because the strength of muscles is proportional to their cross sectional area. Since volume depends on the cube of linear dimensions and surface area on the square of these dimensions, a number of important generalisations follow.

If you were to double in size, your strength would become four times greater, but your weight would become eight times greater. In effect, you would become half as good at moving your body about. You might even have trouble standing up.

We see this in animals. Very large animals overcome this difficulty by having more substantial legs. The elephant and the gazelle can both run fast. However, the elephant is not simply a scaled up gazelle.

There are other consequences of this that are skilfully described in the short essay *On Being the Right Size* by J.B.S. Haldane (1928). Aside from describing these phenomena in more detail than we have, the essay is remarkable for the clarity of its expression. Reading it will help you understand that long complex sentences are not necessary to make one's point!

The essay can be found on the internet.

http://irl.cs.ucla.edu/papers/right-size.html

Exercise C.1.2

1. Find the volumes and surface areas of these cuboids:

   a 25 cm by 15 cm by 20 cm

   b 17 cm by 19 cm by 31 cm

   c 2.3 cm by 5.1 cm by 6.2 cm

   d 0.04 cm by 1.65 cm by 0.54 cm

2. Find the volumes and surface areas of these solids:

   a Cylinder of radius 3.5 cm and height 9.2 cm.

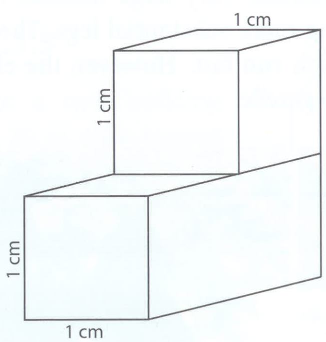
   b Cone of radius 9.1 cm and height 15.2 cm.

   c A sphere of radius 7.7 cm

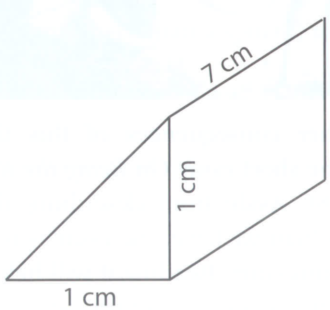
   d A hemisphere of radius 4.2 cm

3. Find the volumes and surface areas of these solids:

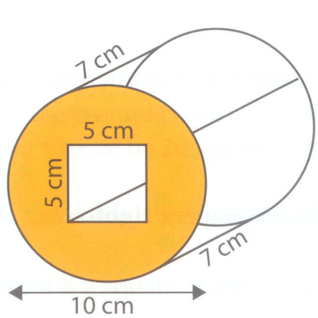
   a



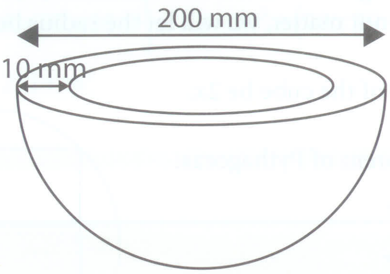
   b



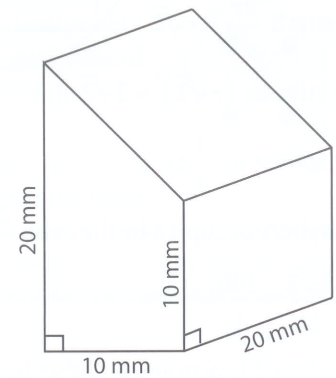
   c



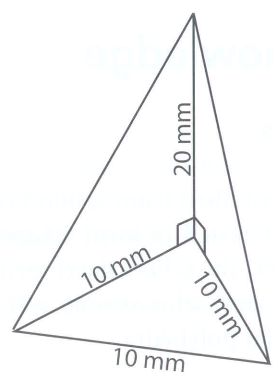
   d



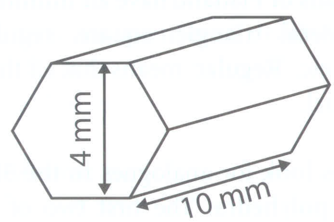
   e



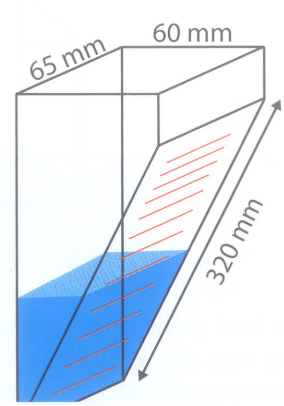
   f



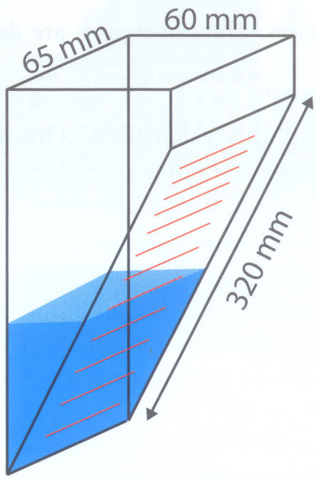
   g



   h Two regular hexagonal cones placed apex to apex.



4. The diagram shows a rain gauge that is made from transparent moulded plastic. The internal dimensions are shown.



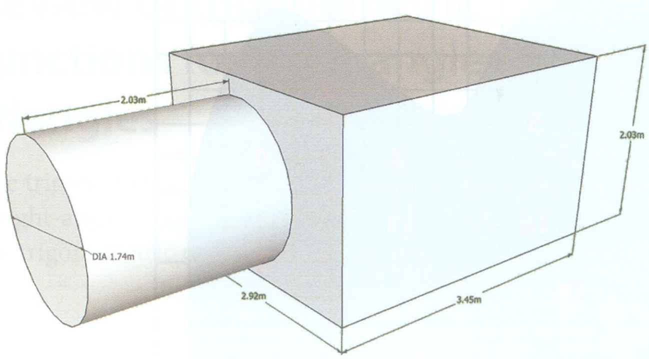
The rainfall scale is marked on the sloping edge.

a What is the highest rain reading that the gauge can measure?

b Find the distance up the scale that 10 mm of rain should be marked.

c Find the distance up the scale that 55 mm of rain should be marked.

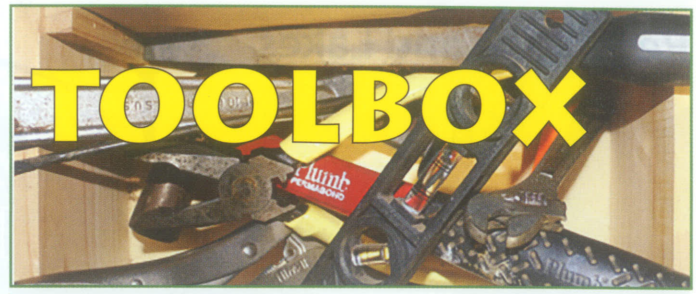
5. The object shown is part of a street art display. Find its volume.



6. A regular tetrahedron has edges of length a. Find a formula for its volume.

7. A regular dodecahedron has edges of length a. Find a formula for its volume.

8. A garden hose is made from flexible plastic with density 1.63 gm cm^{-3}. It has an inner diameter of 12 mm, an outer diameter of 16 mm and a length of 50 m. Find its weight in kilograms to the nearest tenth of a kilogram.



Ratios

How does an object's surface area relate to its volume? This is a somewhat meaningless question as the two quantities are not even measured in the same units. However, we will persist and assume that there is sense in comparing a surface area in m^2 with a volume in m^3 (or cm^2/cm^3).

Begin by looking at a cube and a sphere. The surface area divided by volume ratios are:

$$\text{Cube } \frac{6a^2}{a^3} = \frac{6}{a}. \quad \text{Sphere } \frac{4\pi r^2}{\frac{4}{3}\pi r^3} = \frac{3}{r}.$$

If we view a solid as a means of containing 'matter', this result means that the sphere uses less surface than a cube to contain the same amount of volume. The sphere is a more efficient container. But is it the best possible container?

Look at the world around you to see some examples of this. Why are soap bubbles spherical? Why do animals tend to 'curl up' when they sleep?

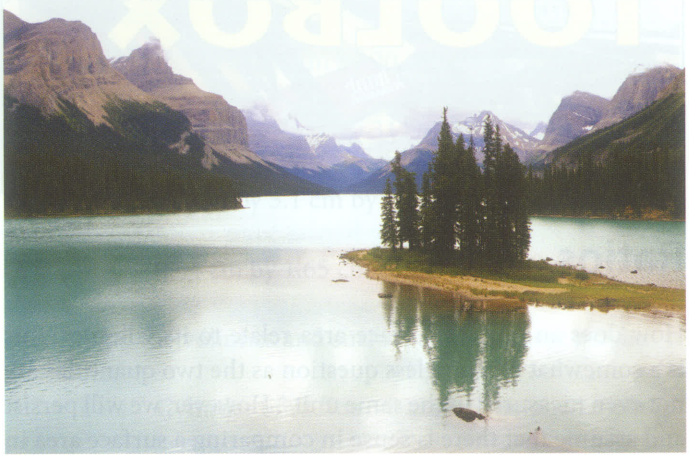
A second consequence of the fact that there is not a fixed ratio between a solid's volume and its surface area is very important in the sciences.

Imagine a unit cube. Next cut it in half. The volume remains the same, but the surface area is increased (we have not lost any surface and the cut has created two new surfaces). Repeat the process and you will get more area for the same volume. In principle (though not in practice) it is possible to continue the process indefinitely and arrive at a pile of stuff with infinite surface area. Practical applications of this process of finely dividing substances to increase surface area include:

- chemical catalysts
- gills and lungs
- gas masks

Can you find any more?

Reservoirs and Lakes



Measuring the amount of water in lakes and reservoirs is of interest to ecologists, water engineers and a whole range of other people.

The trouble is that lakes and even man-made reservoirs seldom have 'nice' geometric shapes of the sort we have been considering.

You are asked to look at the map of a (fictional) lake at the bottom of this page and arrive at an estimate for the amount of water it contains.

The grid is 10 metre squares. The red figures are depths in metres.

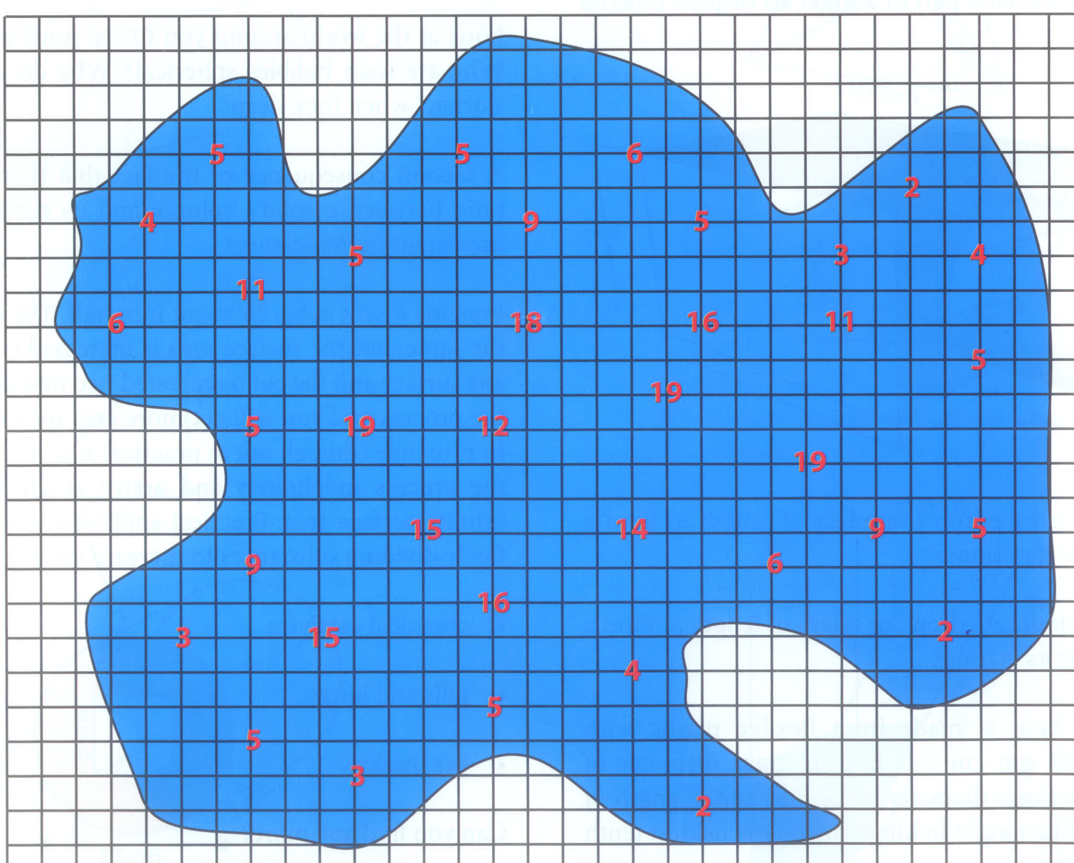
How would one set about this calculation? One thing is certain, there is no exact answer waiting to be discovered.

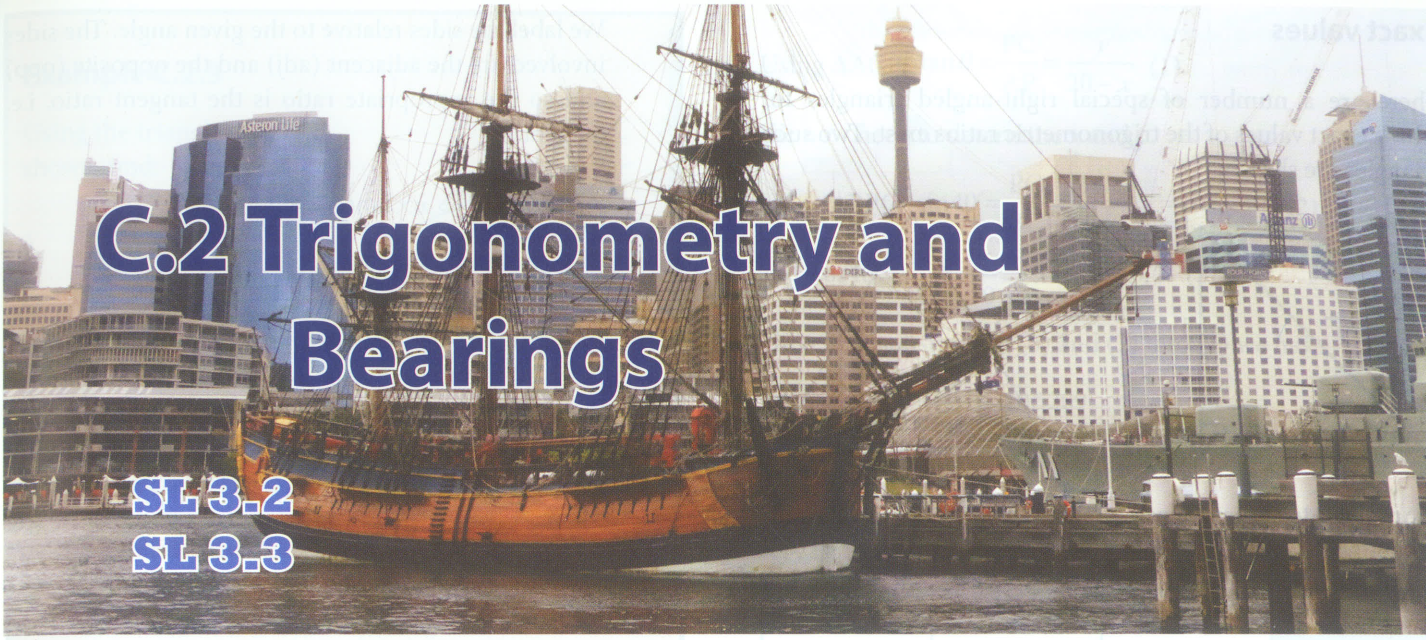
We suggest a couple of possibilities.

1. A geometric approach. Can you divide the lake into sections, each with a depth and estimate the amount of water in each section? Is it necessary to view these sections as prisms?

2. A probabilistic approach. How about taking random samples of the surface of the lake? If the sample does not hit a depth reading, use nearby figures to get a depth estimate. Can you get an average depth and use it to estimate the volume?

Answers





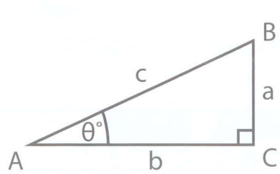
C.2 Trigonometry and Bearings

SL 3.2
SL 3.3

Our cover picture shows the replica of James Cook's *Endeavour* in modern Sydney. The reason why Cook's 1770 expedition, during which he circumnavigated the World, is considered to be one of the great voyages of discovery is not that he was the first person to visit the Pacific. He was not! However, he was to first person to know exactly where he was while he did it. Cook's charts of New Zealand, Eastern Australia and parts of Papua New Guinea were still in use until recently. His absolute error in position seldom exceeds 10 nautical miles. How did Cook manage this feat? He was a very capable and largely self-taught mathematician.

Review of trigonometric functions for right-angled triangles

The trigonometric functions are defined as ratio functions in a right-angled triangle. As such they are often referred to as the trigonometric ratios.



The trigonometric ratios are based on the right-angled triangle shown above. Such right-angled triangles are defined in reference to a nominated angle. In the right-angled triangle ABC the longest side [AB] (opposite the right-angle) is the hypotenuse. Relative to the angle ∠BAC of size θ° , the side BC is called the opposite side while the side AC is called the adjacent side.

The trigonometric ratios are defined as

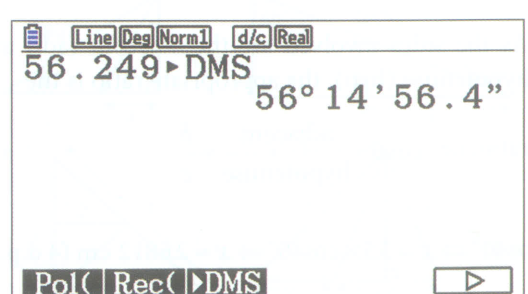
$$\sin\theta=\frac{\text{opposite}}{\text{hypotenuse}}=\frac{a}{c}$$
$$\cos\theta=\frac{\text{adjacent}}{\text{hypotenuse}}=\frac{b}{c}$$
$$\tan\theta=\frac{\text{opposite}}{\text{adjacent}}=\frac{a}{b}$$

This is often memorised using the mnemonic SOHCAHTOA.

Note also, that $\dfrac{\sin\theta}{\cos\theta}=\dfrac{\frac{a}{c}}{\frac{b}{c}}=\dfrac{a}{b}=\tan\theta$

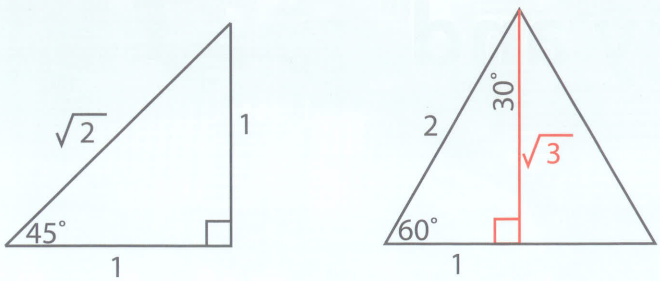
There also exists another important relation between the side lengths of a right-angled triangle. This relationship, using Pythagoras' Theorem is: $a^2+b^2=c^2$.

Do not forget to adjust the mode of your calculator to degree mode when necessary. On the TI–83, this is done by pressing MODE and then selecting the Degree mode. As angles can be quoted in degrees ' ° ', minutes ' ' ' and seconds ' " '. Most calculators can convert an angle quoted as a decimal into one quoted in degrees, minutes and seconds.



Exact values

There are a number of special right-angled triangles for which exact values of the trigonometric ratios exist. Two such triangles are shown:

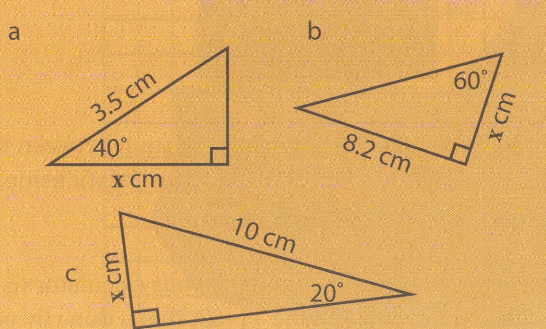


From these triangles we can tabulate the trigonometric ratios as follows:

	sin	cos	tan
30°	$\frac{1}{2}$	$\frac{\sqrt{3}}{2}$	$\frac{1}{\sqrt{3}}=\frac{\sqrt{3}}{3}$
45°	$\frac{1}{\sqrt{2}}=\frac{\sqrt{2}}{2}$	$\frac{1}{\sqrt{2}}=\frac{\sqrt{2}}{2}$	1
60°	$\frac{\sqrt{3}}{2}$	$\frac{1}{2}$	$\sqrt{3}$

Example C.2.1

Find x in each of the following triangles (correct to 4 d.p.).



a We label the sides relative to the given angle:

As the sides involved are the adjacent (adj) and the hypotenuse (hyp), the appropriate ratio is the

cosine ratio, i.e. $\cos\theta=\frac{\text{adjacent}}{\text{hypotenuse}}=\frac{b}{c}$

$$\frac{x}{3.5}=\cos40^\circ \Rightarrow x=3.5\times\cos40^\circ \Rightarrow x\approx 2.6812 \text{ cm (4 d.p.)}$$

b We label the sides relative to the given angle. The sides involved are the adjacent (adj) and the opposite (opp) and so the appropriate ratio is the tangent ratio, i.e. tangent.

$$\frac{8.2}{x}=\tan60^\circ \Rightarrow x\times\tan60^\circ=8.2$$

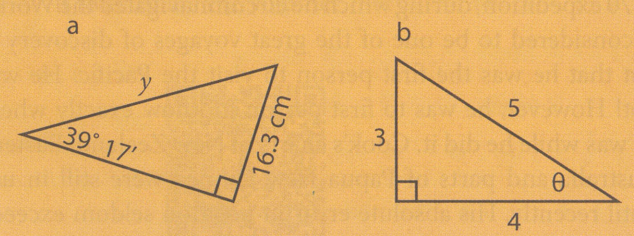
$$\Rightarrow x=\frac{8.2}{\tan60^\circ}\approx 4.7343 \text{ cm (4 d.p.)}$$

c We label the sides relative to the given angle. The sides involved are the opposite (opp) and the hypotenuse (hyp). The appropriate ratio is the sine ratio.

$$\frac{x}{10}=\sin20^\circ \Rightarrow x=10\times\approx 3.4202 \text{ cm (4 d.p.)}$$

Example C.2.2

Find y and θ in the following triangles.



a The important sides are the opposite and hypotenuse.

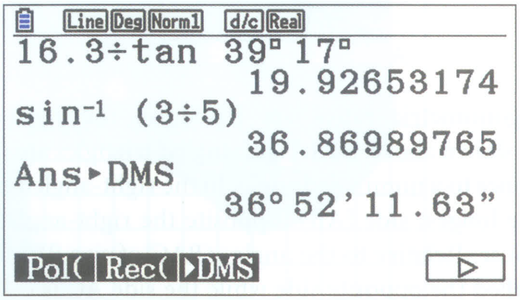
$$\frac{16.3}{y}=\sin39^\circ17' \Rightarrow y\times\sin39^\circ17'=16.3$$

$$\Rightarrow y=\frac{16.3}{\sin39^\circ17'}\approx 25.7 \text{ cm}$$

Most calculators accept angle inputs of angles in degrees minutes and seconds.

b Any one of the three ratios will do. Note that we must use the inverse trigonometric ratio (e.g. $\sin^{-1}$).

$$\sin\theta=\frac{3}{5}\Rightarrow\theta=\sin^{-1}\left(\frac{3}{5}\right)\approx 36^\circ52'12''$$



Example C.2.3

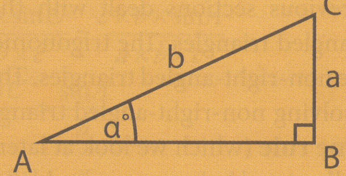
Using the triangle shown, find:

a i AB

ii cosα

iii sinα

b If: cosα = 0.2, find sin(90˚ – α)



a i Using Pythagoras' Theorem we have:

$$AC^2 = AB^2 + BC^2 \therefore b^2 = AB^2 + a^2$$

$$\Rightarrow AB^2 = b^2 - a^2$$

$$\Rightarrow AB = \sqrt{b^2 - a^2}$$

ii $\cos\alpha = \dfrac{AB}{AC} = \dfrac{\sqrt{b^2 - a^2}}{b}$

iii $\tan\alpha = \dfrac{BC}{AB} = \dfrac{a}{\sqrt{b^2 - a^2}}$

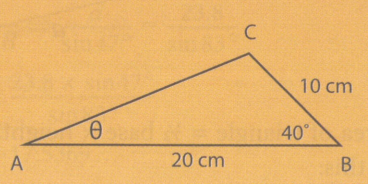
b $\sin(90° - \alpha) = \dfrac{AB}{AC} = \cos\alpha$

We often have to deal with non-right-angled triangles. However, these can be 'broken up' into at least two right-angled triangles, which then involves solving simultaneous equations. This is illustrated in the next example.

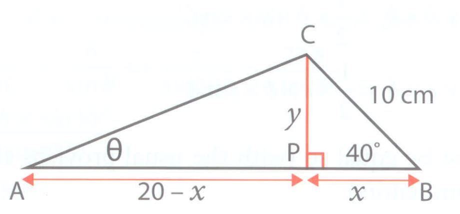
Example C.2.4

Find the angle θ in the diagram shown.

Note that $\angle ACB \neq 90^\circ$



We start by 'breaking-up' the triangle into two right-angled triangles as follows:



Using ΔACP: $\tan\theta = \dfrac{PC}{AP} = \dfrac{y}{20 - x}$ (1)

We now need to determine x and y.

Using ΔBPC: $\sin 40° = \dfrac{PC}{BC} = \dfrac{y}{10} \Rightarrow y = 10\sin 40°$ (2)

and $\cos 40° = \dfrac{BP}{BC} = \dfrac{x}{10} \Rightarrow x = 10\cos 40°$ (3)

Therefore, substituting (3) and (2) into (1) we have:

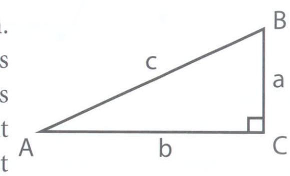
$$\tan\theta = \frac{10\sin 40°}{20 - 10\cos 40°} \approx 0.5209$$

$$\therefore \theta \approx \tan^{-1}(0.5209) \approx 27.5157° \approx 27°31'$$

Note that we have not rounded down our answer until the very last step.

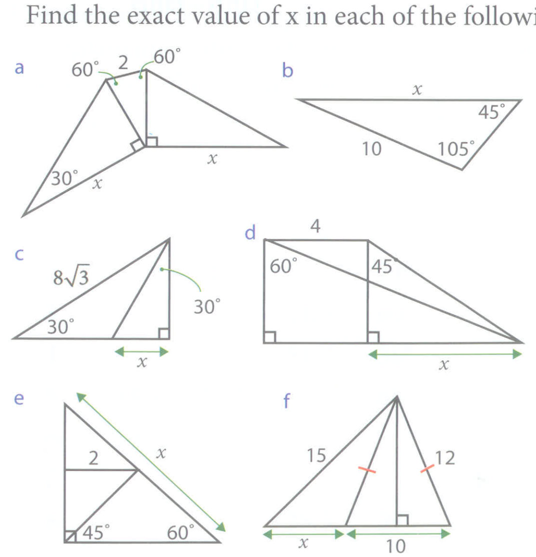
Exercise C.2.1

1 The parts of this question refer to the triangle shown. Complete the blank spaces in this table, giving lengths correct to three significant figures and angles correct to the nearest degree.

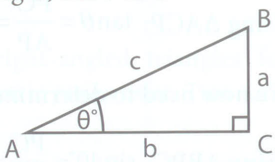


	a cm	b cm	c cm	A	B	C
a			1.6		90°	23°
b		98.3			90°	34°
c			33.9		90°	46°
d	2.3	30.7			90°	87°
e					90°	33°
f	44.4	77			90°	51°
g			68.4		90°	57°

2 Find the exact value of x in each of the following.



3. Using the triangle on the right, show that



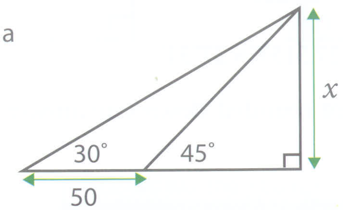
a $\sin(90°-\theta)=\cos\theta$

b $\cos(90°-\theta)=\sin\theta$

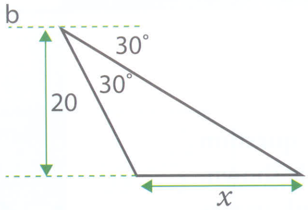
c $\tan(90°-\theta)=\dfrac{1}{\tan\theta}$

4. Find the exact value of x in each of the following.

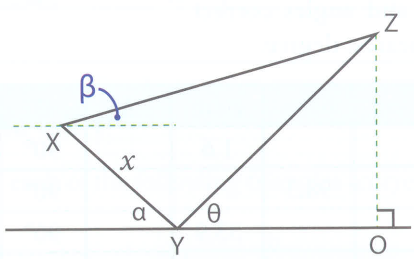
a



b



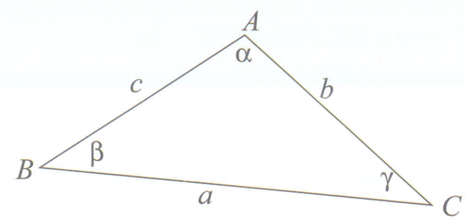
5.



Show that $OZ=\dfrac{x\tan\theta(\sin\alpha+\cos\alpha\tan\beta)}{\tan\theta-\tan\beta}$

The Sine Rule

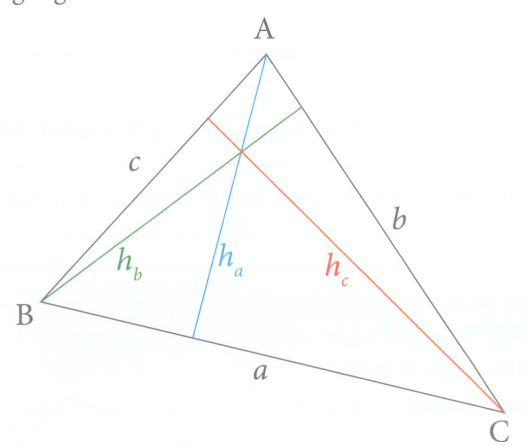
Previous sections dealt with the trigonometry of right-angled triangles. The trigonometric ratios can be used to solve non-right-angled triangles. There are two main methods for solving non-right-angled triangles, the **sine rule** and the **cosine rule** (which we look at later in this section). Both are usually stated using a standard labelling of the triangle. This uses capital letters to label the vertices and the corresponding small letters to label the sides opposite these vertices.



$\sin A = \sin\alpha$
$\sin B = \sin\beta$
$\sin C = \sin\gamma$

If we add the three altitudes to the triangle, we can calculate the area of the triangle in three ways.

The altitude from vertex A to side a is labelled h_a. These will meet in a single point, though this is not crucial to the following argument.



Area of triangle = ½ base × height. There are three versions of this:

$$\text{Area}=\frac{1}{2}\times a\times h_a=\frac{1}{2}\times a\times c\times\sin B$$

$$\text{Area}=\frac{1}{2}\times b\times h_b=\frac{1}{2}\times b\times a\times\sin C$$

$$\text{Area}=\frac{1}{2}\times c\times h_c=\frac{1}{2}\times c\times b\times\sin A$$

These must be equal so (with the usual provisos about non-zero denominators):

$$\frac{1}{2}\times a\times c\times \sin B=\frac{1}{2}\times b\times a\times \sin C=\frac{1}{2}\times c\times b\times \sin A$$

$$a\times c\times \sin B=b\times a\times \sin C=c\times b\times \sin A$$

$$\frac{a\times c\times \sin B}{a\times b\times c}=\frac{b\times a\times \sin C}{a\times b\times c}=\frac{c\times b\times \sin A}{a\times b\times c}$$

$$\frac{\sin B}{b}=\frac{\sin C}{c}=\frac{\sin A}{a}$$

The reciprocal version of this is the more usual version of the **sine rule**:

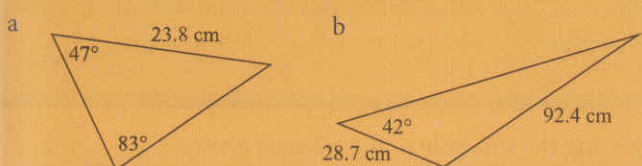
$$\frac{a}{\sin A}=\frac{b}{\sin B}=\frac{c}{\sin C} \quad \text{or} \quad \frac{\sin A}{a}=\frac{\sin B}{b}=\frac{\sin C}{c}$$

So, when should/can we make use of the sine rule?

Although the sine rule can be used for right-angled triangles, it is more often used for situations when we do not have a right-angled triangle, and when the given triangle has a known side and the angle opposite it is also known:

Example C.2.5

Solve the following triangles giving the lengths of the sides in centimetres, correct to one decimal place and angles correct to the nearest degree.



a Firstly, label the triangle using the standard method of lettering. 'Solve the triangle' means find all the angles and the lengths of all the sides. Since two of the angles are known, the third is C = 180° – 47° – 83° = 50°.

The lengths of the remaining sides can be found using the known pairing of side and angle, b and B.

$$\frac{a}{\sin A}=\frac{b}{\sin B}\Leftrightarrow\frac{a}{\sin 47^\circ}=\frac{23.8}{\sin 83^\circ}$$

$$a=\frac{23.8\times \sin 47^\circ}{\sin 83^\circ}$$

$$=17.5369\ldots$$

That is, BC is 17.5 cm (correct to one d.p.).

Similarly, the remaining side can be calculated:

$$\frac{c}{\sin C}=\frac{b}{\sin B}\Leftrightarrow\frac{c}{\sin 50^\circ}=\frac{23.8}{\sin 83^\circ}$$

$$\therefore c=\frac{23.8\times \sin 50^\circ}{\sin 83^\circ}$$

$$=18.3687\ldots$$

That is, AB is 18.4 cm (correct to one d.p.).

b This triangle is different from the previous example in that only one angle is known. It remains the case that a pair of angles and an opposite side are known and that the sine rule can be used. The angle A must be found first.

$$\frac{\sin A}{a}=\frac{\sin B}{b}\Leftrightarrow\frac{\sin A}{28.7}=\frac{\sin 42^\circ}{92.4}$$

$$\Leftrightarrow \sin A=\frac{28.7\times \sin 42^\circ}{92.4}$$

$$=0.207836$$

$$\therefore A=\sin^{-1}0.207836$$

$$=11.9956^\circ$$

$$=11^\circ 59' 44''$$

The answer to the first part of the question is 12° correct to the nearest degree. It is important, however, to carry a much more accurate version of this angle through to subsequent parts of the calculation. This is best done using the calculator memory.

The third angle can be found because the sum of the three angles is 180°. So, C = 180° – 12° – 42° = 126°.

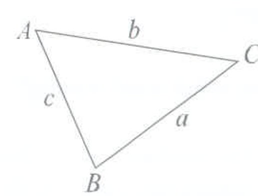
An accurate version of this angle must also be carried to the next part of the calculation. Graphics calculators have multiple memories labelled A, B, C etc. and students are advised to use these in such calculations.

The remaining side is: $\frac{c}{\sin 126^\circ}=\frac{28.7}{\sin 12^\circ}\Leftrightarrow c=\frac{28.7\sin 126^\circ}{\sin 12^\circ}$

$$\therefore c=111.6762\ldots$$

That is, AB is 111.7 cm (correct to one d.p.)

Exercise C.2.2



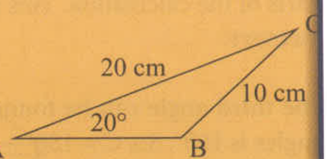
1. Use the sine rule to complete the following table, which refers to the standard labelling of a triangle.

	a cm	*b* cm	*c* cm	*A*	*B*	*C*
1			48.2		29°	141°
2		1.2		74°	25°	
3			11.3	60°		117°
4			51.7	38°		93°
5	18.5	11.4		68°		
6	14.6	15.0			84°	
7		7.3			16°	85°
8			28.5	39°		124°

	a cm	b cm	c cm	A	B	C
9	0.8		0.8	82°		
10			33.3	36°		135°
11	16.4			52°	84°	
12			64.3		24°	145°
13	30.9	27.7		75°		
14			59.1	29°		102°
15		9.8	7.9		67°	
16			54.2	16°		136°
17	14.8		27.2			67°
18			10.9		3°	125°
19			17.0		15°	140°
20			40.1	30°		129°

Example C.2.6

For the triangle shown, find the angle ABC.



Making use of the sine rule we have:

$$\frac{\sin A}{a} = \frac{\sin B}{b} \Leftrightarrow \frac{\sin 20°}{10} = \frac{\sin B}{20}$$

$$\Leftrightarrow \sin B = \frac{20 \sin 20°}{10}$$

$$\therefore B = \sin^{-1}(2 \sin 20°)$$

$$= 43.1601\ldots$$

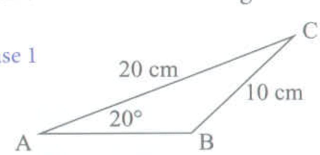
That is, $B = 43°10'$

However, from our diagram, the angle ABC should have been greater than 90°! That is, we should have obtained an **obtuse angle** (90° < B < 180°) rather than an **acute angle** (0° < B < 90°).

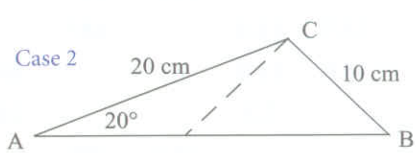
So, what went wrong?

This example is a classic case of what is known as the **ambiguous case**, in that, from the given information it is possible to draw two different diagrams, both having the same data. We show both these triangles:

Case 1



Case 2



Notice that the side BC can be pivoted about the point C and therefore two different triangles can be formed with BC = 10. This is why there are two possible triangles based on the same information. In the solution above, B = 43° 10' – representing Case 2. However, our diagram is represented by Case 1! Therefore, the correct answer is 180° – 43° 10' = 136° 50'.

The Ambiguous Case (SL Topic 3.5 & AHL 3.8)

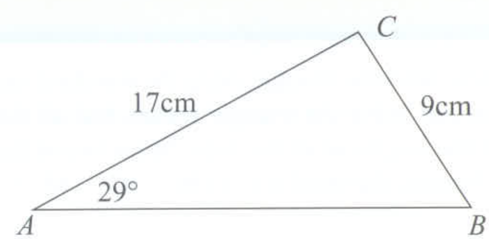
In understanding this rather troublesome difficulty, students are encouraged to relate this to the section on the graph of the sine function.

From Example C.2.6, it can be seen that an ambiguous (having a 'double meaning') case can arise when using the sine rule. In the given situation we see that the side CB can be pivoted about its vertex, forming two possible triangles. We consider another such triangle in the next example.

Example C.2.7

Draw diagrams showing the triangles in which AC = 17 cm, BC = 9 cm and A = 29° and solve these triangles.

Applying the sine rule to the triangle gives:



$$\frac{\sin B}{17} = \frac{\sin 29°}{9} \Leftrightarrow \sin B = \frac{17 \times \sin 29°}{9}$$

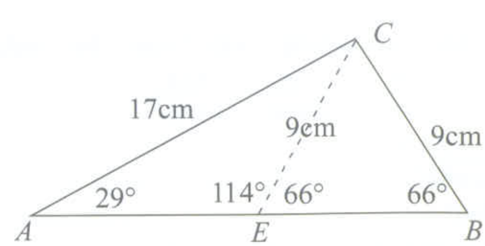
$$= 0.91575$$

$$\therefore B = 66°$$

Next, we have, $C = 180° - 29° - 66° = 85°$ and

$$\frac{c}{\sin 85°} = \frac{9}{\sin 29°} \Leftrightarrow c = 18.5$$

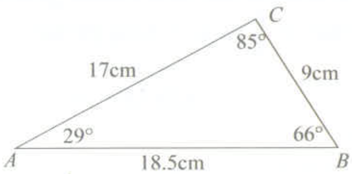
There is, however, a second solution that results from drawing an isosceles triangle BCE.

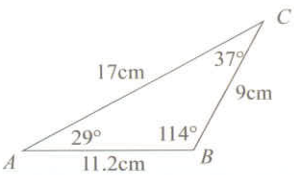


This creates the triangle AEC which also fits the data. The third angle of this triangle is 37° and the third side is:

$$\frac{AE}{\sin 37^\circ} = \frac{9}{\sin 29^\circ} \Leftrightarrow AE = 11.2$$

The original data is ambiguous in the sense that there are two triangles that are consistent with it.





You should also notice that the two angles in the solution are 66° and 114° and that sin66° = sin114°. (That is, sin 66° = sin(180° – 66°) = sin 114°.

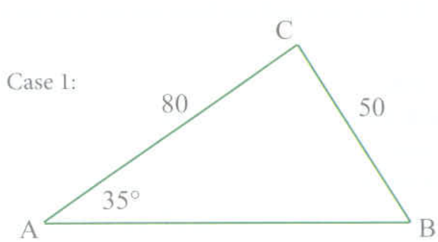
Example C.2.8

Find ∠ABC for the triangle ABC given that $a = 50$, $b = 80$ and A = 35°.

We first determine the value of $b\sin\alpha$ and compare it with the value a:

Now, $b\sin\alpha = 80\sin 35^\circ = 45.89$

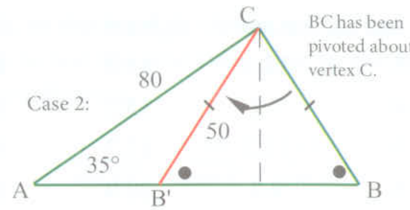
Therefore we have that $b\sin\alpha$ (= 45.89) < a (= 50) < b (= 80) meaning that we have an ambiguous case.



Using the sine rule, $\frac{\sin A}{a} = \frac{\sin B}{b}$, we have

$$\frac{\sin 35^\circ}{50} = \frac{\sin B}{80} \Leftrightarrow \sin B = \frac{80\sin 35^\circ}{50}$$

$$\therefore B = 66^\circ 35'$$



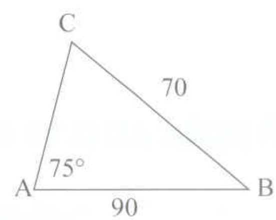
From case 1, the obtuse angle B' is given by 180° – 66°35' = 113°25'.

This is because $\Delta B'CB$ is an **isosceles** triangle, so that
$\angle AB'C = 180^\circ - \angle CB'B$

Example C.2.9

Find $\angle ACB$ for the triangle ABC given that $a = 70$, $c = 90$ and $A = 75°$.

We start by drawing the triangle with the given information:



Using the sine rule we have:

$$\frac{\sin C}{90} = \frac{\sin 75}{70} \Leftrightarrow \sin C = \frac{90\sin 75}{70}$$

$$\therefore \sin C = 1.241\ldots$$

Which is impossible to solve for as the sine of an angle can never be greater than one.

Therefore no such triangle exists.

Exercise C.2.3

Find the two solutions to these triangles which are defined using the standard labelling:

	a cm	b cm	A
1	7.4	18.1	20°
2	13.3	19.5	14°
3	13.5	17	28°
4	10.2	17	15°
5	7.4	15.2	20°
6	10.7	14.1	26°
7	11.5	12.6	17°
8	8.3	13.7	24°
9	13.7	17.8	14°
10	13.4	17.8	28°
11	12.1	16.8	23°
12	12	14.5	21°

	a cm	b cm	A
13	12.1	19.2	16°
14	7.2	13.1	15°
15	12.2	17.7	30°
16	9.2	20.9	14°
17	10.5	13.3	20°
18	9.2	19.2	15°
19	7.2	13.3	19°
20	13.5	20.4	31°

2. Solve the following triangles.

a $\alpha = 75°, a = 35, c = 45$

b $\alpha = 35°, a = 30, b = 80$

c $\beta = 40°, a = 22, b = 8$

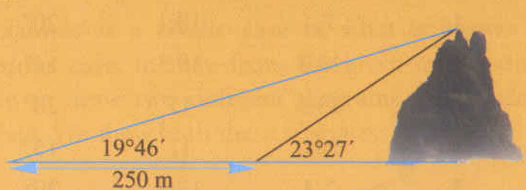
d $\gamma = 50°, a = 112, c = 80$

Applications of the sine rule

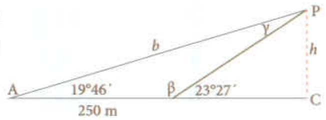
Just as in the case of right-angled triangles, the sine rule becomes very useful. In particular, it means that previous problems that required the partitioning of a non-right-angled triangle into two (or more) right-angled triangles can be solved using the sine rule.

Example C.2.10

A surveying team are trying to find the height of a hill. They take a 'sight' on the top of the hill and find that the angle of elevation is 23°27′. They move a distance of 250 metres on level ground directly away from the hill and take a second 'sight'. From this point, the angle of elevation is 19°46′. Find the height of the hill, correct to the nearest metre.



Labelling the given diagram using the standard notation we have:



With $\beta = 180 - 23°27' = 156°33'$

and $\gamma = 180 - 19°46' - 156°33' = 3°41'$

Then, using the sine rule,

$$\frac{b}{\sin 156°33'} = \frac{250}{\sin 3°41'}$$

$$\Leftrightarrow b = \frac{250\sin 156°33'}{\sin 3°41'}$$

$$= 1548.63\ldots$$

Then, using ΔACP we have,

$$\sin 19°46' = \frac{h}{b} \Leftrightarrow h = b\sin 19°46'$$

$$= 523.73$$

So, the hill is 524 m high (to nearest metre).

Exercise C.2.4

1. A short course biathlon meet requires the competitors to run in the direction S60°W to their bikes and then ride S40°E to the finish line, situated 20 km due South of the starting point. What is the distance of this course?

2. A pole is slanting towards the sun and is making an angle of 10° to the vertical. It casts a shadow 7 metres long along the horizontal ground. The angle of elevation of the top of the pole to the tip of its shadow is 30°. Find the length of the pole, giving your answer to 2 d.p.

3. A statue A, is observed from two other statues B and C which are 330 m apart. The angle between the lines of sight AB and BC is 63° and the angle between the lines of sight AC and CB is 75°. How far is statue A from statue B?

4. Town A is 12 km from town B and its bearing is 132° from B. Town C is 17 km from A and its bearing is 063°T from B. Find the bearing of A from C.

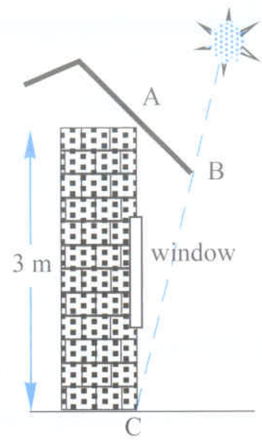
5. The angle of elevation of the top of a building from a park bench on level ground is 18°. The angle of elevation from a second park bench, 300 m closer to the base of the building is 30°. Assuming that the two benches and the building all lie on the same vertical plane, find the height of the building.

a A man standing 6 m away from a lamp post casts a shadow 10 m long on a horizontal ground. The angle of elevation from the tip of the shadow to the lamp light is 12°. How high is the lamp light?

b If the shadow is cast onto a road sloping at 30° upwards, how long would the shadow be if the man is standing at the foot of the sloping road and 6 metres from the lamp post?

6. At noon the angle of elevation of the sun is 72° and is such that a three metre wall AC, facing the sun, is just in the shadow due to the overhang AB. The angle that the overhang makes with the vertical wall is 50°.

Copy and illustrate this information on the diagram shown.



a Find the length of the overhang.

b At 4 p.m. the angle of elevation of the sun is 40° and the shadow due to the overhang just reaches the base of the window.

c How far from the ground is the window?

7. The lookout on a ship sailing due East at 25 km/h observes a reef N62°E at a distance of 30 km.

a How long will it be before the ship is 15 km from the reef, assuming that it continues on its easterly course.

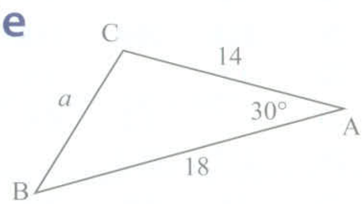
b How long is it before it is again 15 km from the reef?

c What is the closest that the ship will get to the reef?

Extra question

The Cosine Rule

Sometimes the sine rule is not enough to help us solve for a non right-angled triangle.



For example, in the triangle shown, we do not have enough information to use the sine rule. That is, the sine rule only provided the following:

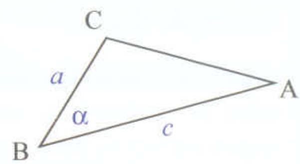
$$\frac{a}{\sin 30^\circ} = \frac{14}{\sin B} = \frac{18}{\sin C}$$

where there are too many unknowns.

For this reason we derive another useful result, known as the cosine rule. The cosine rule may be used when

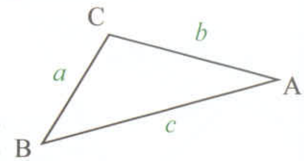
1. **two sides and an included angle are given**:

   This means that the third side can be determined and then we can make use of the sine rule (or the cosine rule again).



2. **three sides are given**:

   This means we could then determine any of the angles.

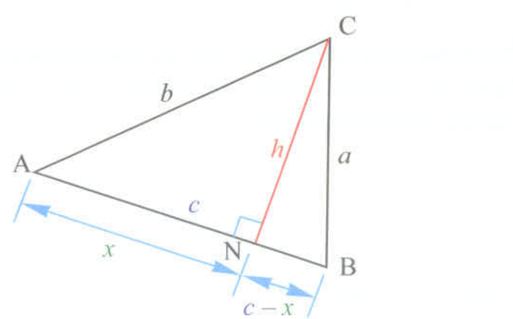


The cosine rule, with the standard labelling of the triangle has three versions:

$$a^2 = b^2 + c^2 - 2bc\cos A$$
$$b^2 = a^2 + c^2 - 2ac\cos B$$
$$c^2 = a^2 + b^2 - 2ab\cos C$$

The cosine rule can be remembered as a version of Pythagoras' Theorem with a correction factor. We now show why this works.

Consider the case where there is an acute angle at A. Draw a perpendicular from C to N as shown in the diagram.



In ΔANC we have $b^2 = h^2 + x^2$

$\Leftrightarrow h^2 = b^2 - x^2$ – (1)

In ΔBNC we have $a^2 = h^2 + (c-x)^2$

$\Leftrightarrow h^2 = a^2 - (c-x)^2$ – (2)

Equating (1) and (2) we have,

$$a^2-(c-x)^2 = b^2-x^2$$
$$\Leftrightarrow a^2-c^2+2cx-x^2 = b^2-x^2$$
$$\Leftrightarrow a^2 = b^2+c^2-2cx$$

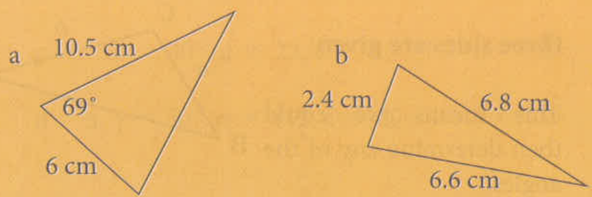
However, from ΔANC: $\cos A = \frac{x}{b} \Leftrightarrow x = b\cos A$

Substituting this result for x: $a^2=b^2+c^2-2bc\cos A$

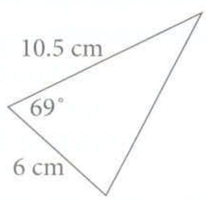
Although we have shown the result for an acute angle at A, the same rule applies if A is obtuse.

Example C.2.11

Solve the following triangles giving the lengths of the sides in centimetres, correct to one decimal place and angles correct to the nearest degree.



The data does not include an angle and the opposite side so the sine rule cannot be used. The first step, as with the sine rule, is to label the sides of the triangle. Once the triangle has been labelled, the correct version of the cosine rule must be chosen. In this case, the solution is:



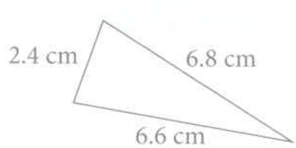
$$c^2 = a^2+b^2-2ab\cos C$$
$$c^2 = 10.5^2+6^2-2\times 10.5\times 6\times \cos 69°$$
$$= 101.0956$$
$$a = 10.1$$

The remaining angles can be calculated using the sine rule. Again, it is important to carry a high accuracy for the value of c to the remaining problem:

$$\frac{\sin B}{b} = \frac{\sin C}{c} \Leftrightarrow \sin B = \frac{6\times \sin 69°}{10.0546} \quad \therefore B = 34°$$

Finally, $A = 180°-34°-69° = 77°$

b In this case, there are no angles given. The cosine rule can be used to solve this problem as follows:



$$a^2 = b^2+c^2-2bc\cos A$$
$$6.6^2 = 2.4^2+6.8^2-2\times 2.4\times 6.8\times \cos A$$
$$2\times 2.4\times 6.8\times \cos A = 2.4^2+6.8^2-6.6^2$$
$$\cos A = \frac{2.4^2+6.8^2-6.6^2}{2\times 2.4\times 6.8}$$
$$= 0.25858$$
$$A = 75.014°$$
$$= 75°1'$$

Next, use the sine rule:

$$\frac{\sin B}{b} = \frac{\sin A}{a} \Leftrightarrow \sin B = \frac{2.4\times \sin 75}{6.6} \quad \therefore B = 20°34'$$

So that $C = 180° - 75° - 21° = 84°$

The three angles, correct to the nearest degree are $A = 75°$, $B = 21°$ & $C = 84°$.

Exercise C.2.5

Solve the following triangles.

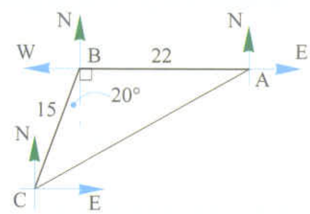
	a cm	*b* cm	*c* cm	*A*	*B*	*C*
1	13.5		16.7		36°	
2	8.9	10.8				101°
3	22.8		12.8		87°	
4	21.1	4.4				83°
5		10.6	15.1	74°		
6		13.6	20.3	20°		
7	9.2		13.2		46°	
8	23.4	62.5				69°
9		9.6	15.7	41°		
10	21.7	36.0	36.2			
11	7.6	3.4	9.4			
12	7.2	15.2	14.3			
13	9.1		15.8		52°	
14	14.9	11.2	16.2	63°	42°	75°
15	2.0	0.7	2.5			
16	7.6	3.7	9.0			
17	18.5	9.8	24.1			
18	20.7	16.3	13.6			
19		22.4	29.9	28°		
20	7.0		9.9		42°	
21	21.8	20.8	23.8			
22	1.1		1.3		89°	
23		1.2	0.4	85°		
24	23.7	27.2				71°
25	3.4	4.6	5.2			

Applications of the Cosine Rule

Example C.2.12

A cyclist rode her bike for 22 km on a straight road heading in a westerly direction towards a junction. Upon reaching the junction, she headed down another straight road bearing 200°T for a distance of 15 km. How far is the cyclist from her starting position?

We start with a diagram:



Note that:

$\angle ABC = 90° + 20° = 110°$

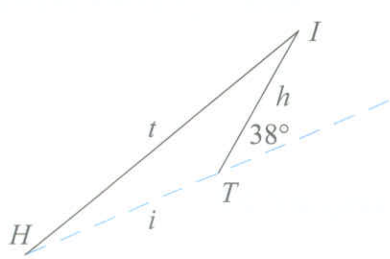
Using the cosine rule we have,

$$AC^2 = 15^2 + 22^2 - 2 \times 15 \times 22 \cos 110°$$
$$\Rightarrow AC = \sqrt{225 + 484 - 660 \times (-0.3420\ldots)}$$
$$\therefore AC = 30.5734\ldots$$

That is, she is (approximately) 30.57 km from her starting point.

Example C.2.13

A yacht starts from a harbour and sails for a distance of 11 km in a straight line. The yacht then makes a turn to port (left) of 38° and sails for 7 km in a straight line in this new direction until it arrives at a small island. Draw a diagram that shows the path taken by the yacht and calculate the distance from the harbour to the island.



The question does not give the bearing of the first leg of the trip so the diagram can show this in any direction. H is the harbour, I the island and T the point where the yacht makes its turn.

The angle in the triangle at T is 180° – 38° = 142°.

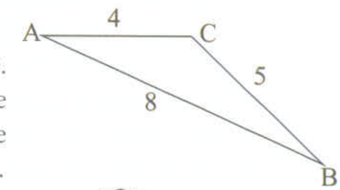
The problem does not contain an angle and the opposite side and so must be solved using the cosine rule.

$$t^2 = h^2 + i^2 - 2hi \cos T$$
$$= 7^2 + 11^2 - 2 \times 7 \times 11 \times \cos 142°$$
$$= 291.354$$
$$\therefore t = 17.1$$

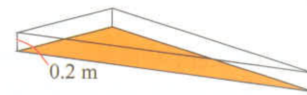
That is, distance from the harbour to the port is 17.1 km (to one d.p)

Example C.2.14

A triangular sandpit having side lengths 5 m, 4 m and 8 m is to be constructed to a depth of 20 cm. Find the volume of sand required to fill this sandpit.



We will need to find an angle. In this case we determine the largest angle, which will be the angle opposite the longest side.



From our diagram we have

$$8^2 = 4^2 + 5^2 - 2 \times 4 \times 5 \cos C$$
$$\therefore 64 = 16 + 25 - 40 \cos C$$
$$\Leftrightarrow \cos C = \frac{16 + 25 - 64}{40}$$
$$= -\frac{23}{40}$$
$$\therefore C = 125°6'$$

To find the volume of sand we first need to find the surface area of the sandpit.

Area $= \frac{1}{2}ab\sin C = \frac{1}{2} \times 4 \times 5 \times \sin(125°6') = 8.1815 \text{ m}^2$.

The volume of sand required is $0.2 \times 8.1815 = 1.64 \text{ m}^3$.

Exercise C.2.6

1. Thomas has just walked 5 km in a direction N70°E when he realises that he needs to walk a further 8 km in a direction E60°S.

   a How far from the starting point will Thomas have travelled?

   b What is his final bearing from his starting point?

2. Two poles, 8 m apart, are facing a rugby player who is 45 m from the left pole and 50 m from the right one. Find the angle that the player makes with the goal mouth.

3. The lengths of the adjacent sides of a parallelogram are 4.80 cm and 6.40 cm. If these sides have an inclusive angle of 40˚, find the length of the shorter diagonal.

4. During an orienteering venture, Patricia notices two rabbit holes and estimates them to be 50 m and 70 m away from her. She measures the angle between the line of sight of the two holes as 54°. How far apart are the two rabbit holes?

5. To measure the length of a lake, a surveyor chooses three points. Starting at one end of the lake she walks in a straight line for 223.25 m to some point X, away from the lake. She then heads towards the other end of the lake in a straight line and measures the distance covered to be 254.35 m. If the angle between the paths she takes is 82°25', find the length of the lake.

6. A light aeroplane flying N87°W for a distance of 155 km, suddenly needs to alter its course and heads S34°E for 82 km to land on an empty field.

   a How far from its starting point did the plane land.

   b What was the plane's final bearing from its starting point?

Area of a Triangle

Given **any** triangle with sides a and b, height h and included angle θ, the area, A, is given by:

$$A = \frac{1}{2}bh$$

However, $\sin\theta = \frac{h}{a} \Leftrightarrow h = a \times \sin\theta$ and so, we have that:

Area = ½ $a \times b$ sinθ

where θ is the angle between sides a and b.

Note that the triangle need not be a right-angled triangle.

Because of the standard labelling system for triangles, the term sinθ is often replaced by sin$C$, giving the expression Area = ½ $a \times b$ sin$C$.

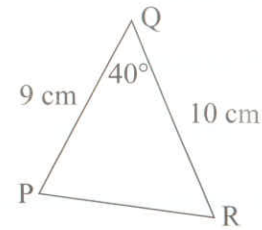
A similar argument can be used to generate the formulae:

Area = ½ $b \times c$ sin$A$ = ½ $a \times c$ sin$B$

Example C.2.15

Find the area of the triangle PQR given that PQ = 9 cm, QR = 10 cm and ∠PQR = 40˚.

Based on the given information we can construct the following triangle:



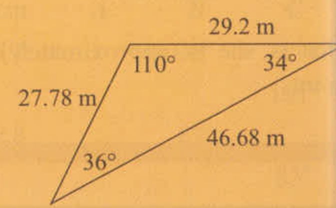
The required area, A, is given by:

$$A = \frac{1}{2}ab\sin\theta = \frac{1}{2}\times 9\times 10\times\sin 40° = 28.9$$

That is, the area is 28.9 cm^2.

Example C.2.16

The diagram shows a triangular children's playground. Find the area of the playground.



Since all the measurements of the triangle are known, any one of the three formulae could be used. Many people remember the formula as 'Area equals half the product of the lengths of two sides times the sine of the angle between them'.

$$\text{Area} = \frac{1}{2}\times 27.78\times 46.68\times\sin 36° = 381\text{m}^2$$

$$\text{Area} = \frac{1}{2}\times 27.78\times 29.2\times\sin 110° = 381\text{m}^2$$

$$\text{Area} = \frac{1}{2}\times 29.2\times 46.68\times\sin 34° = 381\text{m}^2$$

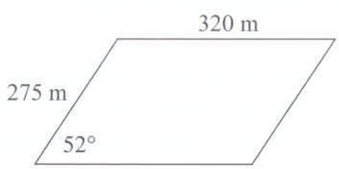
Exercise C.2.7

1. Find the areas of these triangles that are labelled using standard notation.

	a cm	*b* cm	*c* cm	A	B	C
a	35.94	128.46	149.70	12°	48°	120°
b	35.21	54.55	81.12	20°	32°	128°

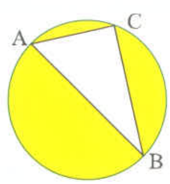
c	46.35	170.71	186.68	14°	63°	103°
d	33.91	159.53	163.10	12°	78°	90°
e	42.98	25.07	48.61	62°	31°	87°
f	39.88	24.69	34.01	84°	38°	58°
g	43.30	30.26	64.94	34°	23°	123°
h	12.44	2.33	13.12	68°	10°	102°
i	43.17	46.44	24.15	67°	82°	31°
j	23.16	32.71	24.34	45°	87°	48°
k	50.00	52.91	38.64	64°	72°	44°
l	44.31	17.52	48.77	65°	21°	94°
m	12.68	23.49	22.34	32°	79°	69°
n	42.37	42.37	68.56	36°	36°	108°
o	40.70	15.65	41.26	77°	22°	81°

2. A car park is in the shape of a parallelogram. The lengths of the sides of the car park are given in metres.



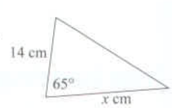
What is the area of the car park?

3. The diagram shows a circle of radius 10 cm. AB is a diameter of the circle. AC = 6 cm.

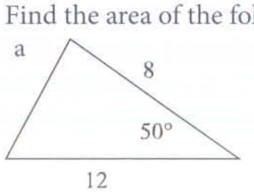


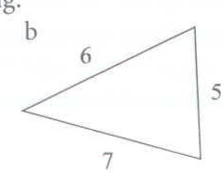
Find the area of the shaded region, giving an exact answer.

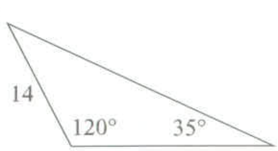
4. The triangle shown has an area of 110 cm². Find x.



5. Find the area of the following.

a


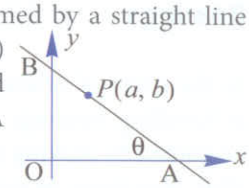
b


c


6. A napkin is in the shape of a quadrilateral with diagonals 9 cm and 12 cm long. The angle between the diagonals is 75°. What area does the napkin cover when laid out flat?

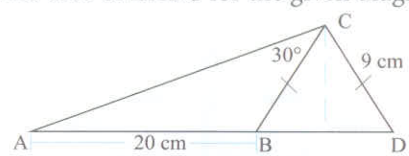
7. A triangle of area 50 cm² has side lengths 10 cm and 22 cm. What is the magnitude of the included angle?

8. A variable triangle OAB is formed by a straight line passing through the point $P(a,b)$ on the Cartesian plane and cutting the x-axis and y-axis at A and B respectively.



If $\angle OAB = \theta$, find the area of ΔOAB in terms of a, b and θ.

9. Find the area of ΔABC for the given diagram.

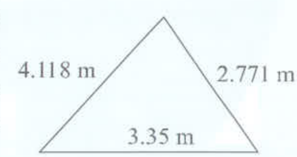


Exercise C.2.8

1. The diagram shows a triangular building plot. The distances are given in metres. Find the length of the two remaining sides of the plot giving your answers correct to the nearest hundredth of a metre.

2. Xiang is standing on level ground. Directly in front of him and 32 metres away is a flagpole. If Xiang turns 61° to his right, he sees a post box 26.8 metres in front of him. Find the distance between the flagpole and the post box.

3. A triangular metal brace is part of the structure of a bridge. The lengths of the three parts are shown in metres.



Find the angles of the brace.

4. Find the smallest angle in the triangle whose sides have length 35.6 cm, 58.43 cm and 52.23 cm.

5. Ayton is directly North of Byford. A third town, Canfield, is 9.93km from Ayton on a bearing of 128°T. The distance from Byford to Canfield is 16.49km. Find the bearing of Canfield from Byford.

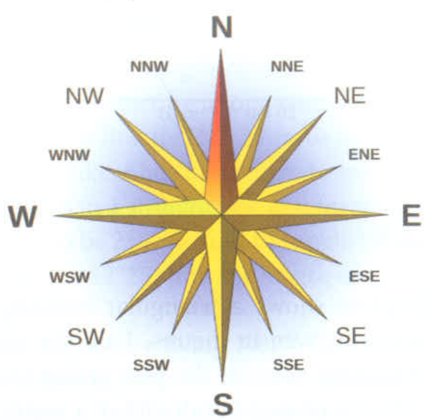
Extra Questions

APPLICATIONS

Bearings

In the sport of orienteering, participants need to be skilled at handling bearings and reading a compass. Bearings can be quoted by making reference to the North, South, East and West directions or using true bearings. We look at each of these.

Points of the Compass

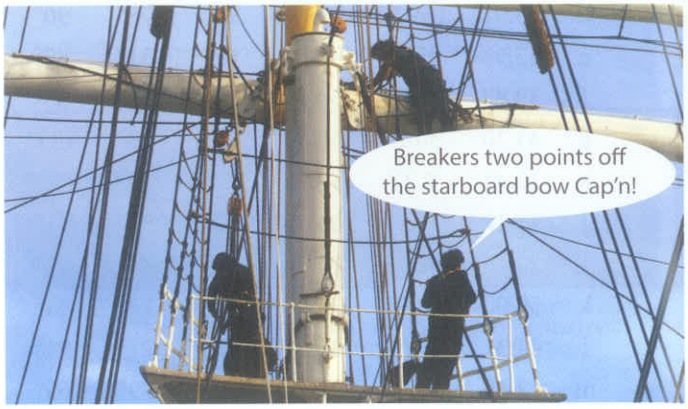


Viking ship, Oslo, Norway

The Vikings were expert sailors and crossed the Atlantic Ocean in open boats of the type show. They navigated using naturally occurring magnetised rock which, when suspended on a string aligned itself North/South. They probably only recognised the four main points of the compass.

As technology improved, more points were recognised and named. On our 'compass rose', there are 16 point. NE means 'North East' - mid-way between North and East (ie. 45° from North). NNE means North North East mid-way between North and North East.

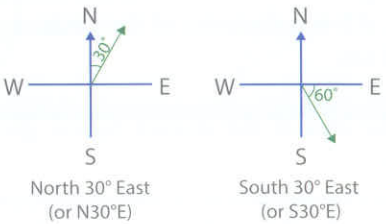
A further subdivision into 32 'points' followed further improvements in compass design.



Our lookout's warning means that he has spotted breaking waves, $^1/_{16}$th of a turn to the right. This probably means a reef and danger. Note that the reference point here is the bow of the ship. From here on the reference direction will be North.

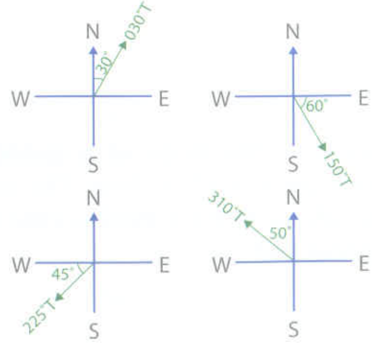
Compass Bearings

These are quoted in terms of an angle measured East, West, North or South, or somewhere in-between. For example, North 30° East, expressed as N30°E, informs us that from the North direction we rotate 30° towards the east and then follow that line of direction. The following diagrams display this for a number of bearings.



True bearings

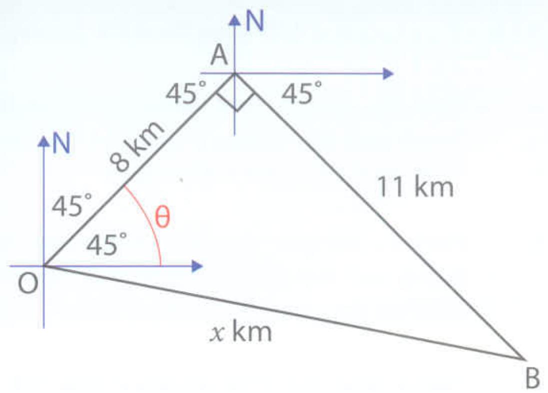
These are quoted in terms of an angle measured in a clockwise direction from North (and sometimes a capital T is attached to the angle to highlight this fact). So, for example, a bearing of 030°T would represent a bearing of 30° in a clockwise direction from the North – this corresponds to a compass bearing of N30°E. Using the above compass bearings we quote the equivalent true bearings:



Example C.2.17

Janette walks for 8 km north-East and then 11 km South-East. Find the distance and bearing from her starting point.

First step is to draw a diagram:



As $\angle OAB = 90^\circ$ we can make use of Pythagoras' Theorem:

$x^2 = 8^2 + 11^2 \Rightarrow x \approx 13.60$ (taking +ve square root)

Let $\theta = \angle AOB$ so that $\tan\theta = \frac{11}{8} \Rightarrow \theta = \tan^{-1}\left(\frac{11}{8}\right) \approx 53.97^\circ$

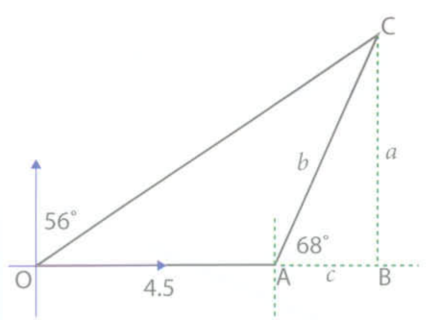
Therefore, bearing is $45^\circ + \theta = 45^\circ + 53.97^\circ = 98.97^\circ$

That is, B has a bearing of 98.97°T from O and is (approx.) 13.6 km away.

Example C.2.18

The lookout, on a ship sailing due East, observes a light on a bearing of 056°T. After the ship has travelled 4.5 km, the lookout now observes the light to be on a bearing of 022°T. How far is the light source from the ship at its second sighting?

As always, we start with a diagram.



Using ΔOBC we have: $\tan 34^\circ = \frac{BC}{OB} = \frac{a}{4.5+c}$

$\therefore a = (4.5+c)\tan 34^\circ \quad (1)$

Using ΔABC we have: $\tan 68^\circ = \frac{BC}{AB} = \frac{a}{c}$

$\therefore a = c\tan 68^\circ \quad (2)$

Equating (1) and (2) we have,

$$c\tan 68^\circ = (4.5+c)\tan 34^\circ$$
$$= 4.5\tan 34^\circ + c\tan 34^\circ$$
$$c\tan 68^\circ - c\tan 34^\circ = 4.5\tan 34^\circ$$
$$c(\tan 68^\circ - \tan 34^\circ) = 4.5\tan 34^\circ$$
$$c = \frac{\tan 34^\circ}{(\tan 68^\circ - \tan 34^\circ)}$$
$$\approx 1.6857$$

Substituting this result into (2) we have,

$$a = \frac{4.5\tan 34^\circ}{(\tan 68^\circ - \tan 34^\circ)} \times \tan 68^\circ \approx 4.1723$$

Then, using ΔABC and Pythagoras' Theorem, we have:

$$b^2 = a^2 + c^2$$
$$= 4.1723^2 + 1.6857^2$$
$$b \approx 4.49999$$

That is, the light is 4.5 km from the ship (at the second sighting). Can you see a much quicker solution? Hint: think isosceles triangle.

Exercise C.2.9

1 a Change the following compass bearings into true bearings.

i	N30°E	ii	N30°W
iii	S15°W	iv	W70°S

b Change the following true bearings into compass bearings.

i	025°T	ii	180°T
iii	220°T	iv	350°T

2 The angle of depression from the top of a building 60 m high to a swing in the local playground is 58°. How far is the swing from the foot of the building?

3 From a point A on the ground, the angle of elevation to the top of a tree is 52°. If the tree is 14.8 m away from point A, find the height of the tree.

4 Find the angle of elevation from a bench to the top of an 80 m high building if the bench is 105 m from the foot of the building.

5 Patrick runs in a direction N60°E and after 45 minutes finds himself 3900 m North of his starting point. What is Patrick's average speed in ms^{-1}?

6 A ship leaves Oldport and heads NW. After covering a distance of 16 km it heads in a direction of N68°22'W travelling a distance of 22 km where it drops anchor. Find the ship's distance and bearing from Oldport after dropping anchor.

7 From two positions 400 m apart on a straight road, running in a Northerly direction, the bearings of a tree are N36°40'E and E33°22'S. What is the shortest distance from the tree to the road?

8 A lamp post leaning away from the sun and at 6° from the vertical, casts a shadow 12 m long when the sun's angle of elevation is 44°. Assuming that the ground where the pole is situated is horizontal, find the length of the pole.

9 From a window, 29.6 m above the ground, the angle of elevation of the top of a building is 42°, while the angle of depression to the foot of the building is 32°. Find the height of the building.

10 Two towns P and Q are 50 km apart, with P due West of Q. The bearing of a station from town P is 040°T while the bearing of the station from town Q is 300°T. How far is the station from town P?

11 When the sun is 74° above the horizon, a vertical flagpole casts a shadow 8.5 m onto a horizontal ground. Find the length of the shadow cast by the sun when it falls to 62° above the horizontal.

12 A hiker walks for 5 km on a bearing of 053° true (North 53° East). She then turns and walks for another 3 km on a bearing of 107° true (East 17° South).

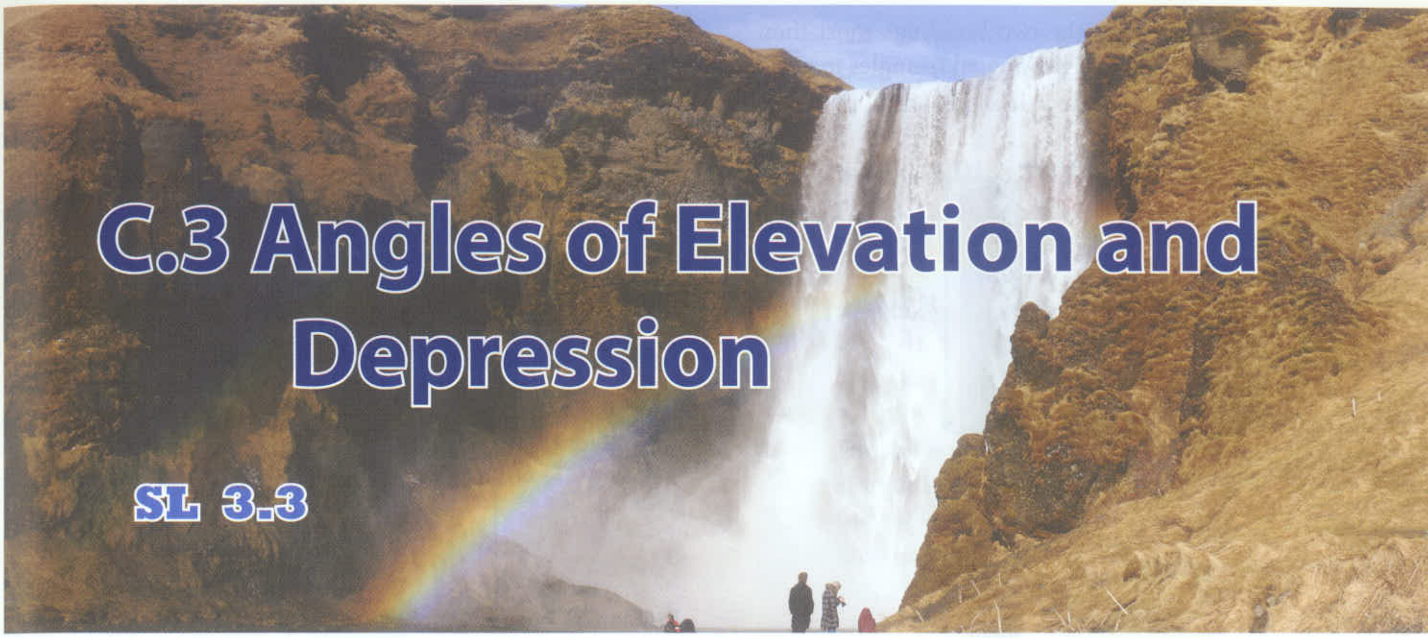
a Find the distance that the hiker travels North/South and the distance that she travels East/West on the first part of her hike.

b Find the distance that the hiker travels North/South and the distance that she travels East/West on the second part of her hike.

c Hence find the total distance that the hiker travels North/South and the distance that she travels East/West on her hike.

d If the hiker intends to return directly to the point at which she started her hike, on what bearing should she walk and how far will she have to walk?

Answers



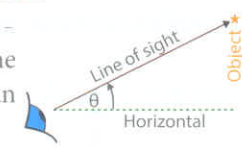
C.3 Angles of Elevation and Depression

SL 3.3

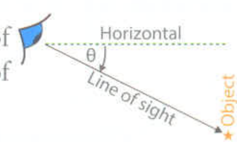
Applications that require the use of trigonometric ratios and right-angled triangles are many and varied. In this section we consider a number of standard applications to highlight this.

Angle of Elevation and Depression

The **angle of elevation** is the angle of the line of sight above the horizontal of an object seen above the horizontal.



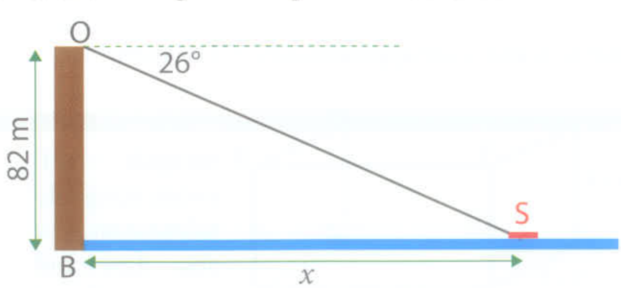
The **angle of depression** is the angle of the line of sight below the horizontal of an object seen below the horizontal.



Note that the angle of depression and elevation for the same line of sight are alternate angles.

> **Example C.3.1**
>
> An observer standing on the edge of a cliff 82 m above sea level sees a ship at an angle of depression of 26°.
>
> How far from the base of the cliff is the ship situated?

We first draw a diagram to represent this situation:



Let the ship be at point S, x metres from the base of the cliff, B, and let O be where the observer is standing.

Using the right-angled triangle OBS we have:

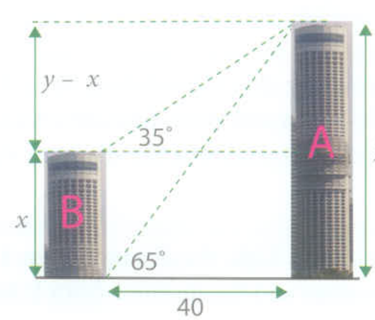
$$\tan 26^\circ = \frac{82}{x} \Rightarrow x\tan 26^\circ = 82$$

$$x = \frac{82}{\tan 26^\circ} \approx 168.12$$

Therefore, the ship is 168 m from the base of the cliff.

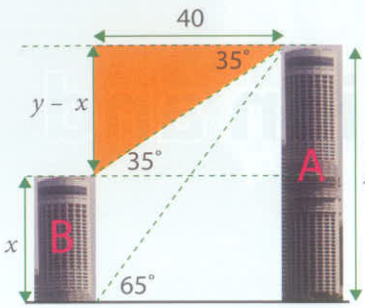
> **Example C.3.2**
>
> The angle of depression from the roof of building A to the foot of a second building, B, across the same street and 40 metres away is 65°. The angle of elevation of the roof of building B to the roof of building A is 35°. How tall is building B?

Let the height of building B be x m and that of building A be y m.

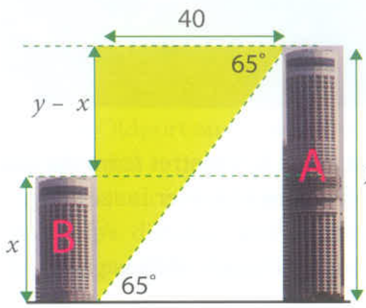


Note that we are using the fact that for the same line of sight, the angle of depression and elevation is equal.

The height difference between the two buildings must then be $(y - x)$ m. We now have two right-angled triangles to work with:



$$\tan 35^\circ = \frac{y-x}{40} \Rightarrow y - x = 40\tan 35^\circ \ (1)$$



$$\tan 65^\circ = \frac{y}{40} \Rightarrow y = 40\tan 65^\circ \ (2)$$

Substituting (2) into (1) we have:

$$40\tan 65^\circ - x = 40\tan 35^\circ \Rightarrow x \approx 57.77$$

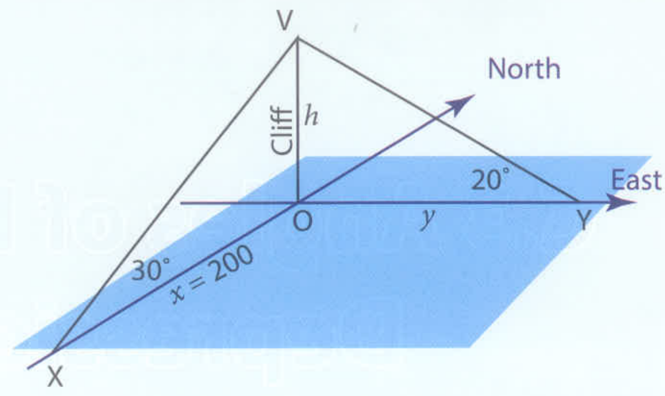
That is, building B is 57.77 m tall.

Example C.3.3

From a point X, 200 m due South of a cliff, the angle of elevation of the top of the cliff is 30°. From a point Y, due East of the cliff, the angle of elevation of the top of the cliff is 20°. How far apart are the points X and Y?

We start by illustrating this information on a 3-D diagram (Note that North–South and West–East are drawn on a plane. It is necessary to do this otherwise the diagram will not make sense).

Let the cliff be h metres high, the distance from X to the base of the cliff be x metres and the distance from Y to the base of the cliff be y metres.



As $\angle XOY = 90^\circ$, then $XY^2 = x^2 + y^2 = 200^2 + y^2$

But $\tan 20^\circ = \frac{h}{y}$ of which we know neither h nor y.

However, using triangle XOV, we have that:

$$\tan 30^\circ = \frac{h}{200} \Rightarrow h = 200\tan 30^\circ$$

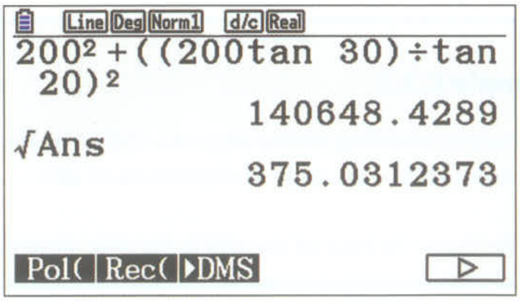
Therefore, we have: $\tan 20^\circ = \frac{200\tan 30^\circ}{y} \Rightarrow y = \frac{200\tan 30^\circ}{\tan 20^\circ}$

That is, $y = 317.25$

$$\text{Therefore, } XY^2 = x^2 + y^2 = 200^2 + \left(\frac{200\tan 30^\circ}{\tan 20^\circ}\right)^2$$

$$= 140648.4289$$

$$XY = 375.0312$$



That is, X and Y are approximately 375 m apart.

Exercise C.3.1

1. For the diagrams shown, determine the angle of inclination between:

Figure 1

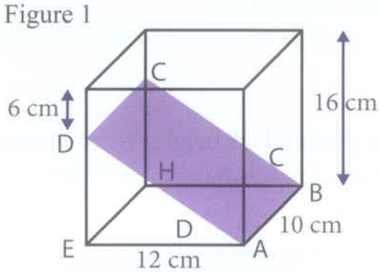
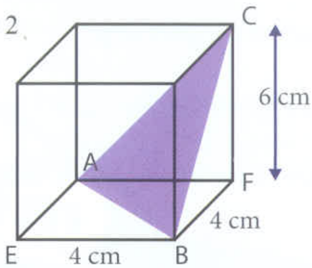


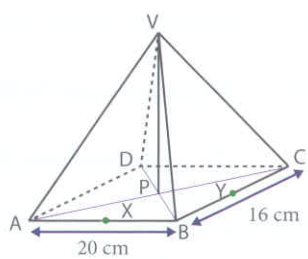
Figure 2



a ABCD and the base, EABH (Figure 1).

b ABC and the base EBFA (Figure 2).

2. A right pyramid with a rectangular base and a vertical height of 60 cm is shown in the diagram alongside. The points X and Y are the midpoints of the sides [AB] and [BC] respectively.



Find:

a the length, AP.

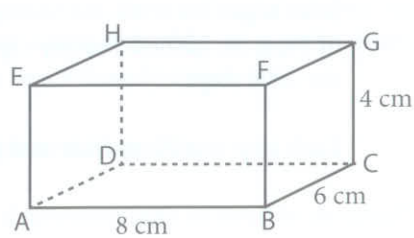
b the length of the edge [AV].

c the angle that the edge AV makes with the base ABCD.

d the length, YV.

e The angle that the plane BCV makes with the base.

3. The diagram alongside shows a rectangular box with side lengths AB = 8 cm, BC = 6 cm and CG = 4 cm.



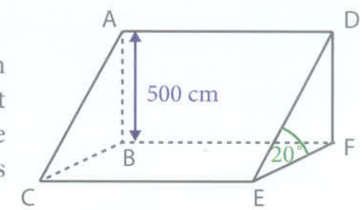
Find the angle between:

a the line [BH] and the plane ABCD.

b the lines [BH and [BA].

c the planes ADGF and ABCD.

4. For the wedge shown alongside, given that the angle between the lines EA and ED is 50°, find:

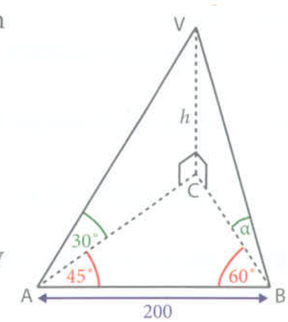


a the length of [AE].

b ∠AEB

5. From a point A, 100 m due south of a tower, the angle of elevation of the top of the tower is 40°. From a point B, due east of the tower, the angle of elevation of the top of the tower is 20°. How far apart are the points A and B?

6. For the triangular prism shown alongside, find:



a the value of h.

b the value of α.

c the angle that the plane ABV makes with the base ABC.

7. The angle of depression from the top of a tower to a point X south of the tower, on the ground and 120 m from the foot of the tower, is 24°. From point Y, due west of X, the angle of elevation to the top of the tower is 19°.

a Illustrate this information on a diagram.

b Find the height of the tower.

c How far is Y from the foot of the tower?

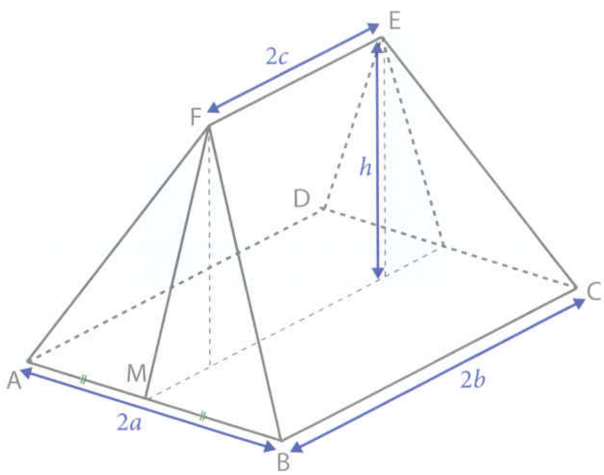
d How far apart are the points X and Y?

8. A mast is held in a vertical position by four ropes of length 60 metres. All four ropes are attached at the same point at the top of the mast so that their other ends form the vertices of a square when pegged into the (level) ground. Each piece of rope makes an angle of 54° with the ground.

a Illustrate this information on a diagram.

b How tall is the mast?

9. A symmetrical sloping roof has dimensions as shown in the diagram.



Find:

a the length of [FM].

b the angle between the plane BCEF and the ground.

c the angle between the plane ABF and the ground.

d the total surface area of the roof.

10. The angle of elevation of the top of a tower from a point A due south of it is 68°. From a point B, due east of A, the angle of elevation of the top is 54°. If A is 50 m from B, find the height of the tower.

11. A tower has been constructed on the bank of a long straight river. From a bench on the opposite bank and 50 m downstream from the tower, the angle of elevation of the top of the tower is 30°. From a second bench on the same side as the tower and 100 m upstream from the tower, the angle of elevation of the top of the tower is 20°.

Find:

a the height of the tower.

b the width of the river.

12. A right pyramid of height 10 m stands on a square base of side lengths 5 m.

Find:

a the length of the slant edge.

b the angle between a sloping face and the base.

c the angle between two sloping faces.

13. A camera sits on a tripod with legs 1.5 m long. The feet rest on a horizontal flat surface and form an equilateral triangle of side lengths 0.75 m.

Find:

a the height of the camera above the ground.

b the angles made by the legs with the ground.

c the angle between the sloping faces formed by the tripod legs.

14. From a point A due south of a mountain, the angle of elevation of the mountaintop is α. When viewed from a point B, x m due east of A, the angle of elevation of the mountaintop is β. Show that the height, h m, of the mountain is given by:

$$h = \frac{x.\sin\alpha.\sin\beta}{\sqrt{\sin^2\alpha - \sin^2\beta}}$$

15. The 1000 Steps

John began his work-out facing East. After an easy jog of 7km, he turned sharply right and began climbing the 1000 steps.

Each step was 10cm wide and 11cm high.

Resting at the top and admiring the view, he could see the Information Kiosk next to the bottom step. The Kiosk was 2.4m high. How far below was the top of the Kiosk in a direct line and at what angle of depression?

Draw a scale diagram to show this.

16. The Helicopter

A helicopter leaves its pad and heads NNE. Beth, the pilot, wants to clear a 1600m high mountain 725m away by 10m

Draw a scale diagram to show this.

At what angle of elevation should she fly?

She then turns through 015° clockwise and flies level for a further 5600m and drops her load without losing height. She now needs to return to the pad. There are no obstacles in the way so she can fly smoothly at a consistent angle of depression.

Add this information to your diagram.

At what angle of depression should she fly?

17. The Paraglider

A paraglider leaves the beach and ascends smoothly at an angle of elevation of 50° to a height of 35m. He wants to descend, again smoothly, to a point 350m further along the beach from his start.

Draw a scale diagram to show this.

At what angle of depression does he need to descend?

18. The Window Cleaner

A window cleaner started her day 4m from the top of a high-rise hotel. Her angle of depression to the top of the Food Shack below was 55°.

After 3 hours' work she is ready for a break. At the level she is now at, her angle of depression to The Food Shack below is 35°.

She lowers herself to the ground and walks the 40m directly to the 10m high Food Shack.

Draw a scaled diagram to show this.

How high is the hotel?

How far did she travel down the side of the hotel during her cleaning.

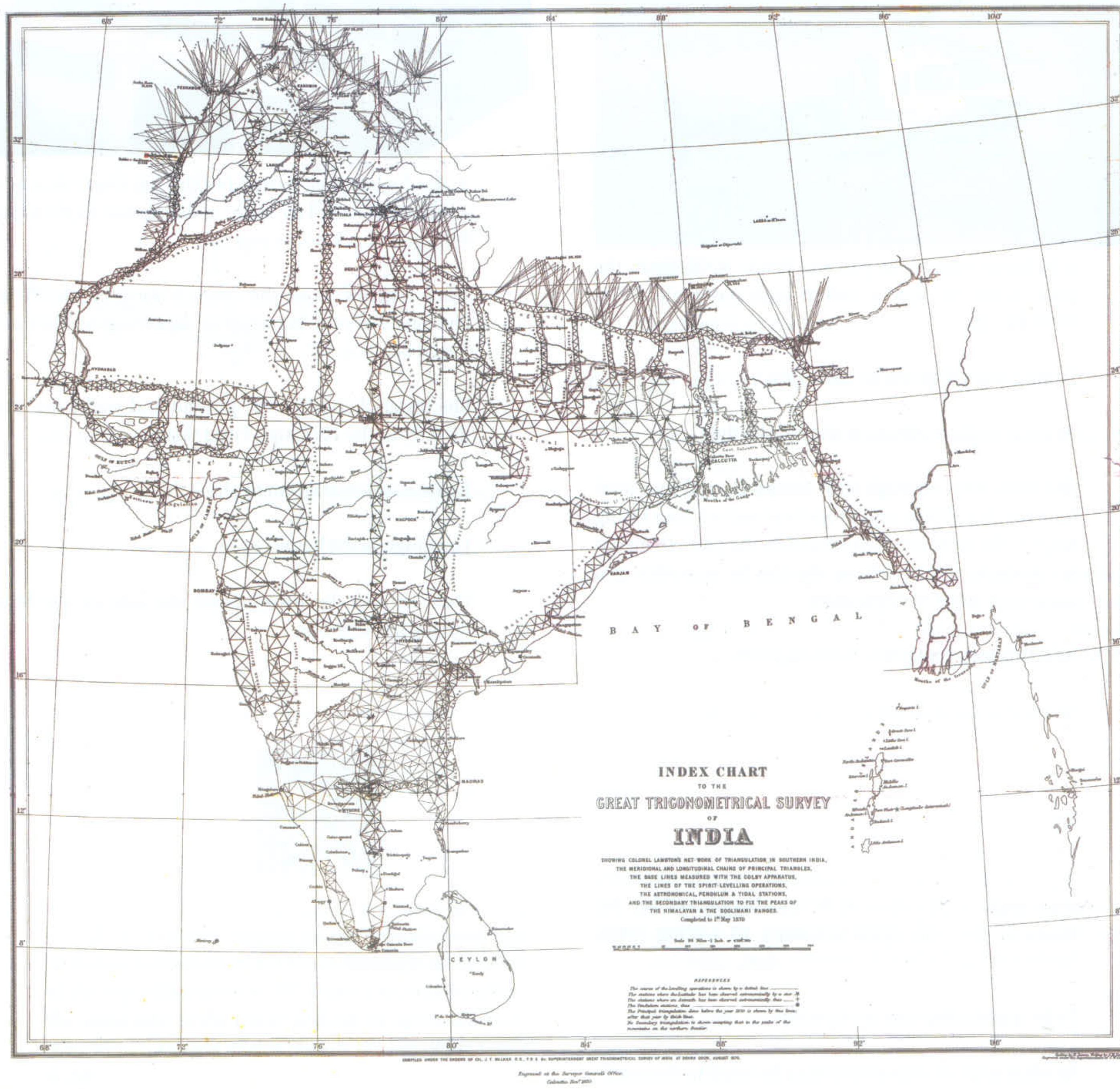
Answers

The Great Trigonometrical Survey

This immense mapping project, using the techniques discussed in this and the previous chapter, was conducted between 1802 and 1871. One of its leaders was George Everest and amongst its great achievements was finding the height of the mountain that bears his name.

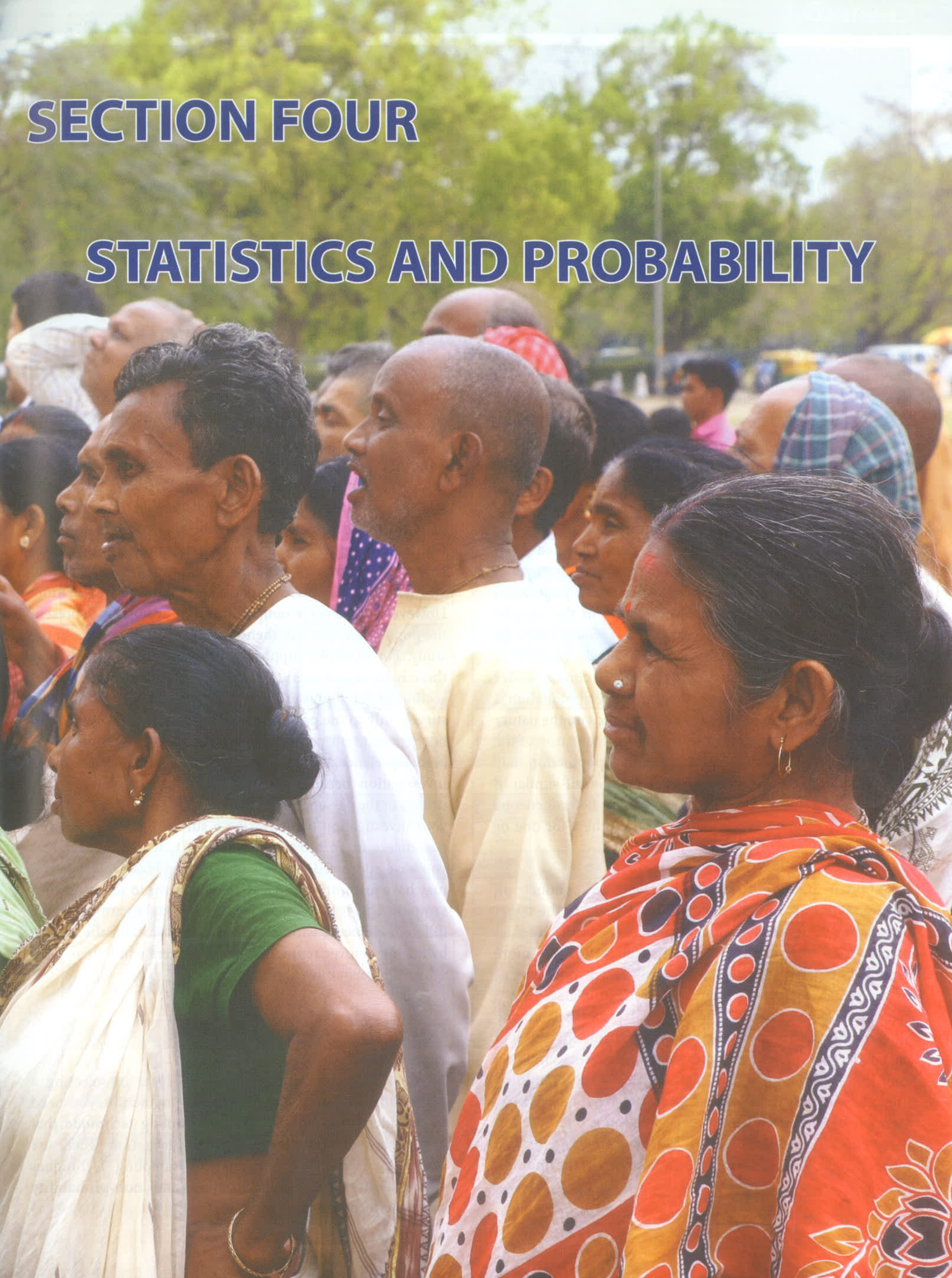
The immense precision achieved across all these triangles proved that the Earth is not a perfect sphere and that the Himalayas are the highest mountains on the planet. As well as producing a very accurate map of 'The Jewel in the Crown'.

A very readable account of this astonishing achievement can be found in *The Great Arc* by John Keay (ISBN: 0-00-257062-9).

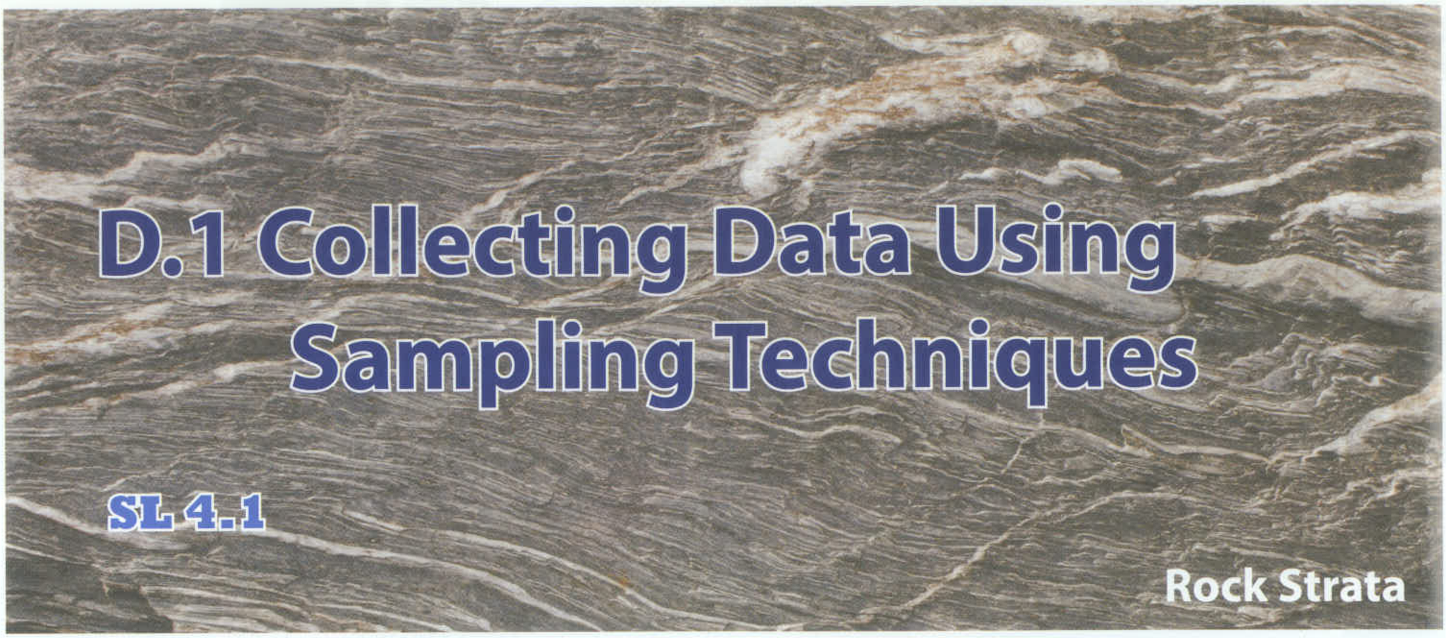


SECTION FOUR

STATISTICS AND PROBABILITY



D.1 Collecting Data Using Sampling Techniques



Rock Strata

Population and Sample

Regardless of the actual context of any statistical study - whether it is an opinion poll, observations from scientific experiments, manufactured units, or data collected from a wide geographical region - one of the main objectives is to draw some generalizations about the actual population inferred directly from the observed results in the sample. By definition, the **target population** in a study refers to all potential members, each of which needs to meet a particular set of parameters, in the predefined context under investigation. Any single member within a given population is known as an **element**. A given element depends on the nature of its population, which in turn depends on the nature of the investigation. The general composition of a population may either be **homogeneous** where every element is self-similar in all aspects as defined by the parameter set, or **heterogeneous** where every element is not similar to the others for one or more variables.

It is often inefficient and impractical to conduct an investigation to study every individual element in the target population because of limited resources. There is also an incorrect assumption that the statistical results are more accurate when an investigation has studied every element in the target population than an investigation based on a carefully selected sample. The potential errors in studying all elements in a population may come from inconsistent data collection methods, inaccurate data analysis techniques and, more importantly, the impracticality of re-examining the entire investigation process to ensure reliability. Hence, a sample derived from the target population is often used to complete the same investigation. This is more cost-effective, more efficient, and even improves the accuracy of the findings.

The **sample** is a subset of the target population that should share the same characteristics. Therefore, if a sample is selected carefully through a well-defined methodology then it should inevitably represent the entire population with a high degree of accuracy. Furthermore, the statistical results that are drawn from the sample can also be inferred and generalized with confidence to the entire population. However, whenever a sample does not accurately represent the population, you will then introduce sampling bias in the study. For example, suppose you are investigating whether there is an urgent need to install speed cameras in a certain section of the highway in a particular city. If you deliver a survey only to the passengers of sports cars, the sample does not represent all potential users of the same section of the highway. Hence, there is a sampling bias in the design of this investigation because there is no representative from the drivers or the passengers of other vehicle types. If you redesign your investigation by delivering a survey to the drivers and passengers in an arbitrary ratio of 4:1, it may appear that you have addressed the issue of sampling bias. However, you may have just introduced a systematic error in selecting your samples. This is because there is an unequal treatment of the different elements with similar characteristics in the sample by having an over representation of drivers while there is an under representation of passengers.

Sampling Techniques

In order to minimize inaccuracy and bias in selecting a sample that best represents the target population, you need to consider using an appropriate sampling technique that is suitable for the context of your investigation. There are essentially two main categories of sampling techniques - **probability sampling techniques** and **non-probability sampling techniques**.

Probability Sampling Techniques

In probability sampling, every element in the population has a known probability that it might be included in the sample. When you perform probability sampling, you will reduce systemic errors because elements with similar and different characteristics have their respective chances of being represented in your sample. This in turn also minimizes sampling bias. The design of an investigation based on probability sampling can become more laborious and expensive but will ensure inferences and results drawn from such a sample can be generalized to the target population.

There are several probability sampling techniques, but for the purpose of this text, the discussion on the theoretical differences and the actual implementation procedures will be limited to (1) simple random sampling, (2) systematic sampling, (3) stratified random sampling, and (4) cluster sampling. Each of these probability sampling techniques is designed and implemented for a different purpose, but more importantly and particularly, for the actual population to be investigated. In order to better illustrate the various probability sampling techniques, the following table with 100 data pieces, each corresponding to the length in millimetres of an opening made in a piece of metal, will be considered.

Table 1 Measurement of an opening in millimetres made by four drills on a piece of metal

Data No.	Drill	Data Value	Data No.	Drill	Data Value	Data No.	Drill	Data Value	Data No.	Drill	Data Value
1	A	3.23	26	B	3.37	51	C	3.11	76	C	3.38
2	B	3.44	27	D	3.56	52	A	3.17	77	C	3.06
3	C	3.26	28	A	3.04	53	B	3.05	78	D	3.08
4	C	3.21	29	C	3.45	54	B	3.31	79	A	3.46
5	D	3.12	30	C	3.14	55	A	3.15	80	A	3.27
6	D	3.06	31	A	3.57	56	C	3.05	81	C	3.25
7	B	3.28	32	D	3.51	57	C	3.24	82	C	3.14
8	D	3.42	33	C	3.50	58	C	3.59	83	D	3.14
9	B	3.27	34	C	3.25	59	D	3.35	84	D	3.28
10	D	3.12	35	C	3.03	60	A	3.47	85	A	3.26
11	C	3.35	36	D	3.28	61	D	3.38	86	D	3.40
12	B	3.08	37	B	3.38	62	B	3.19	87	C	3.30
13	D	3.49	38	A	3.13	63	D	3.04	88	B	3.17
14	D	3.35	39	B	3.52	64	D	3.38	89	C	3.06
15	C	3.19	40	C	3.16	65	A	3.05	90	B	3.13
16	D	3.35	41	C	3.59	66	D	3.09	91	C	3.29
17	C	3.30	42	C	3.07	67	C	3.47	92	D	3.27
18	C	3.55	43	B	3.30	68	B	3.56	93	C	3.04
19	A	3.05	44	C	3.20	69	C	3.40	94	C	3.11
20	B	3.26	45	C	3.14	70	A	3.46	95	C	3.23
21	B	3.59	46	A	3.26	71	C	3.19	96	B	3.00
22	C	3.12	47	B	3.33	72	D	3.24	97	B	3.11
23	B	3.35	48	B	3.55	73	C	3.00	98	B	3.28
24	D	3.21	49	D	3.35	74	C	3.41	99	A	3.07
25	D	3.03	50	A	3.01	75	B	3.19	100	C	3.43

Simple Random Sampling

For simple random sampling, each element in the population has an equal chance of being selected in the sample. This is the simplest probability sampling technique and you will find it to be most useful for sampling from a relatively small homogeneous population and for generating a small sample. Until your entire sample is formed, each potential element from your target population to be included in the sample is selected one at a time either by drawing from a lottery or referring to a random number table.

Suppose that $y_1, y_2, y_3, \ldots, y_n$ are the actual values included in your sample of n units from a population of N elements. Given that simple random sampling is sampling without replacement, therefore the corresponding sample mean is:

$$\bar{y} = \frac{\sum_{i=1}^{n} y_i}{n}$$

and the sample variance is:

$$s_y^2 = \frac{\sum_{i=1}^{n} (y_i - \bar{y})^2}{n-1}.$$

It must be noted that the unbiased estimate of the population mean ($\hat{\bar{Y}}$) for the sample is equivalent to the sample mean if you have implemented simple random sampling (i.e. $\hat{\bar{Y}} = \bar{y}$), however, the unbiased estimate of the sample variance is not the same as the sample variance itself.

Example D.1.1

Suppose that you have implemented simple random sampling by selecting the following 10 data pieces from Table 1 to form your sample.

Data No.	Drill	Data Value
9	B	3.27
36	D	3.28
37	B	3.38
41	C	3.59
54	B	3.31
62	B	3.19
70	A	3.46
79	A	3.46
92	D	3.27
98	B	3.28

Find the corresponding sample mean and sample variance.

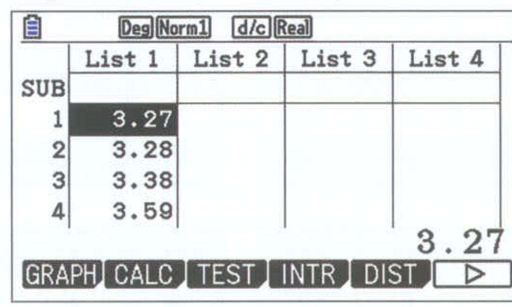
Sample Mean:

$$\bar{y} = \frac{3.27+3.28+3.38+3.59+3.31+3.19+3.46+3.46+3.27+3.28}{10}$$
$$= \frac{33.49}{10}$$
$$= 3.35 (3 \text{ s.f.})$$

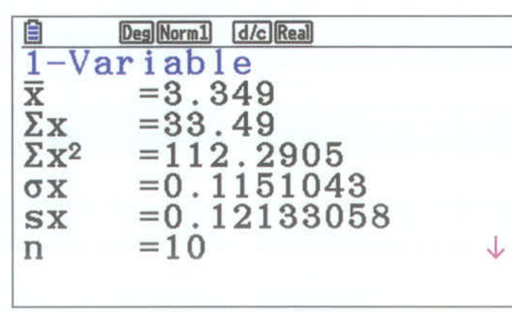
Sample Variance:

$$s_y^2 = \frac{\sum_{i=1}^{10} (y_i - 3.349)^2}{10-1}$$
$$= \frac{(3.27-3.349)^2 + (3.28-3.349)^2 + \ldots + (3.28-3.349)^2}{9}$$
$$= \frac{0.13249}{9}$$
$$= 0.0147 (3 \text{ s.f.})$$

If using a calculator, you will probably need to begin by selecting statistics mode. The data is usually entered as a list (which may not be visible all at once).



You may also need to select the correct list before calculating the 1-VAR statistics.



The sample mean is $3.349 \approx 3.35$ mm with the sample variance of 0.0147. It should be noted that the sample deviation is the principal root of its corresponding sample variance. Hence, the sample standard deviation is 0.121 mm. In other words, based on your sample, you can conclude that on average an opening made on a piece of metal is about 3.35 mm in length with a standard deviation of 0.121 mm.

You will find that one of the biggest challenges in implementing simple random sampling is to identify all elements in the population before sampling. The preparation needed to exhaustively list all elements may be impractical as the size of the population increases. Moreover, if the elements in the population are heterogeneous for some critical variables, then you may be introducing some systematic errors into your sample.

Systematic Sampling

For simple random sampling, every element in the population has an equal chance of being selected in the sample. In systematic sampling, all elements in the population are selected at regular intervals beginning from a randomly chosen first element until the entire sample is formed. The sampling interval is determined either by a particular duration of time or the sequential order of the elements. Due to the regularity in selecting the elements to form the sample, the biggest difference (also the greatest improvement) in the initial stage of implementation for this sampling technique is that it is not necessary to have an exhaustive list of all elements in the population prior to the selection process. Therefore, this probability sampling technique is particularly useful when it is not possible to predict exactly how many elements are to be included in the population or because the exact number of elements varies over a period of time. For example, the number of secondary school students taking the subway on a given day is unpredictable or the number of blueberries in a crate may vary. However, you should still only use this sampling technique when the elements in the population are homogeneous.

Example D.1.2

Suppose that you have implemented systematic sampling by selecting the elements in intervals of 10 starting from the 5th element from Table 1 to form your sample.

Data No.	Drill	Data Value
5	D	3.12
15	C	3.19
25	D	3.03
35	C	3.03
45	C	3.14
55	A	3.15
65	A	3.05
75	B	3.19
85	A	3.26
95	C	3.23

Find the corresponding sample mean and sample variance.

$$\text{Sample Mean: } \overline{y} = \frac{\sum_{i=1}^{n} y_i}{n} = \frac{31.39}{10} \approx 3.14(3 \text{ s.f.})$$

$$\text{Sample Variance: } s_y^2 = \frac{\sum_{i=1}^{10}(y_i - 3.139)^2}{10-1} = \frac{0.06029}{9} \approx 0.00670(3 \text{ s.f.})$$

The sample mean is 3.14 mm with a sample variance of 0.00670 (and a standard deviation of 0.00818 mm). In other words, on average, an opening made in a piece of metal is about 3.14 mm in length with a standard deviation of 0.00818 mm.

It is important to note that it is not possible to comment on whether the results drawn from systematic sampling are more accurate than the results drawn from simple random sampling. This is because, technically, you do not know the population mean and the population variance. In fact, it is possible that in the selection of every 10th element starting from the 5th element may introduce sampling bias if those particular elements have much wider openings or much smaller openings than the rest.

One method to reduce the possibility of introducing sampling bias when implementing systematic sampling is to draw more than one set of samples systematically at varying sampling intervals.

When two mutually exclusive sets of samples are combined together the resulting sample mean is the weighted arithmetic mean of the two corresponding sample means. Let $\overline{y}_1$ be the sample mean for sample set 1 with n_1 elements in the sample and $\overline{y}_2$ be the sample mean for sample set 2 with n_2 elements in the sample. The combined sample mean between sample set 1 and sample set 2 is:

$$\overline{y} = \frac{n_1(\overline{y_1}) + n_2(\overline{y_2})}{n_1 + n_2}$$

In general, the combined sample mean for k sets of samples, each of which is drawn by systematic sampling, is defined by.

$$\overline{y_{sys}} = \frac{\sum_{i=1}^{k} n_i(\overline{y_i})}{\sum_{i=1}^{k} n_i}$$

The combined sample variance between sample Set 1 and sample Set 2 is:

$$s_{y_{sys}}^2 = \frac{n_1\left(s_{y_1}^2 + \left(y_1 - \overline{y}\right)^2\right) + n_2\left(s_{y_2}^2 + \left(y_2 - \overline{y}\right)^2\right)}{n_1 + n_2}$$

where $s_{y_1}^2$ is the sample variance for Set 1 and $s_{y_2}^2$ is the sample variance for Set 2. Hence, in general, the combined sample variance for k sets of samples, each of which is drawn by systematic sampling, is defined by:

$$s_{y_{sys}}^2 = \frac{\sum_{i=1}^{k} n_i\left(s_{y_i}^2 + \left(y_i - \overline{y}\right)^2\right)}{\sum_{i=1}^{k} n_i}$$

Example D.1.3

Continuing from Example 4.1.2, suppose that you have now selected a second set of elements in intervals of 10 starting from the 2nd element from Table 1 to form your second sample set.

Data No.	Drill	Data Value
2	B	3.44
12	B	3.08
22	C	3.12
32	D	3.51
42	C	3.07
52	A	3.17
62	B	3.19
72	D	3.24
82	C	3.14
92	D	3.27

Find the corresponding sample mean and sample variance for this new sample set. Hence, find the combined sample mean and sample variance.

Sample Mean for Set 2: $\overline{y}_2 = \frac{\sum_{i=1}^{10} y_i}{n}$

$$= \frac{32.23}{10}$$

$$= 3.22(3\text{ s.f.})$$

Sample Variance for Set 2: $s_{y_2}^2 = \frac{\sum_{i=1}^{10}\left(y_i - 3.223\right)^2}{10-1}$

$$= \frac{0.19721}{9}$$

$$\approx 0.0219(3\text{ s.f.})$$

For this second set of samples drawn from systematic sampling, the sample mean is 3.223 mm with the sample variance of 0.0219.

The combined sample mean between Set 1 and Set 2:

$$\begin{aligned}\overline{y}_{sys} &= \frac{n_1\left(\overline{y}_1\right) + n_2\left(\overline{y}_2\right)}{n_1 + n_2} \\ &= \frac{10(3.139) + 10(3.223)}{10+10} \\ &= \frac{63.62}{20} \\ &\approx 3.18(3\text{ s.f.})\end{aligned}$$

The combined sample variance between Set 1 and Set 2:

$$\begin{aligned}s_{y_{sys}}^2 &= \frac{n_1\left(s_{y_1}^2 + \left(y_1 - \overline{y}\right)^2\right) + n_2\left(s_{y_2}^2 + \left(y_2 - \overline{y}\right)^2\right)}{n_1 + n_2} \\ &= \frac{10\left(0.006699 + (3.139 - 3.180)^2\right) + 10\left(0.02191 + (3.223 - 3.181)^2\right)}{10+10} \\ &\simeq \frac{0.3214}{20} \\ &\approx 0.0161(3\text{ s.f.})\end{aligned}$$

Therefore, the combined sample mean between Set 1 and Set 2 is 3.18 mm with the combined sample variance of 0.0161(and a standard deviation of 0.127 mm).

It is important to note that even though the combined sample mean amongst several mutually exclusive sample sets each of which has an equal number of elements is essentially the arithmetic mean of their respective individual sample means. In Example 4.1.2 and Example 4.1.3, the two individual sample means are 3.14 and 3.22. The arithmetic mean is 3.18 which is exactly the same as their combined sample mean. It is, however, incorrect to take the arithmetic mean amongst the sample variances of all sample sets. For example, if you take the sample variances from Example 2 and Example 3 as 0.00670 and 0.0219 respectively. The arithmetic mean between these two sample variances is 0.0143 but the actual combined variance is 0.0161.

Stratified Random Sampling

If you take a closer look at Table 1, you will notice that not only are those 100 data pieces taken from four different drills, there are more data pieces from drill C than the other three drills. In particular, there are 16 data pieces from drill A, 23 data pieces from drill B, 37 data pieces from drill C, and 24 data pieces from drill D. Now, if you consider the 10 randomly chosen elements in Example 4.1.1, you will notice that drill C is under-represented with only 1 element while drill B is over-represented with 5 elements. In other words, the number of elements for each of the drills in the sample is not in proportion to how they are distributed respectively in the target population. Hence, you may question whether there is now an obvious systematic error in those originally chosen 10 elements in the sample in Example 4.1.1. The issue is also true in Example 4.1.2 and Example 4.1.3 when the selected elements in the sample do not necessarily follow the same proportion as they appear in the target population. This particular issue is more apparent when you begin to work with a much wider population and especially when the population is no longer homogenous.

Stratified random sampling is particularly useful when the entire collection of the elements in the target population is heterogeneous. Often, the target population has natural divisions, known as **strata**, within which the elements are homogenous. For example, in selecting respondents to an election poll, the strata may be defined by each voting district. In counting the number of trees in a field, the strata may be defined by plots of land.

Although the strata are essentially determined by the topic and the nature of the investigation, it is often the practice to use predefined characteristics readily available for stratification prior to the design of an investigation, instead of arbitrary criteria, in order to avoid sampling bias. The stratification criteria are often based on gender, age, ethnicity, time, postal codes, land divisions, etc. Since the elements in the target population only differ by one unique stratification characteristic, the strata are mutually exclusive.

Once your strata are defined you will then perform simple random sampling within each stratum to form the final sample. There are two methods of selecting the elements that go into the strata. The first method is **proportional allocation** in which the sample size for each stratum is made proportional to the actual number of elements in the stratum. In order to use this method, you must have an exhaustive list of elements in the target population otherwise the correct proportions cannot be determined. The second method is equal allocation in which the sample size for each stratum is the same regardless of the actual number of elements in the stratum. The second method is more practical when the exact number of elements in the population is not known.

Given the complexity in the preliminary stage in designing an investigation with stratified random sampling, it is often more time consuming and requires more resources than simple random sampling and systematic sampling. However, the sample drawn from this sampling technique is usually a better representation of the target population as it reflects the true diversity of the elements.

In order to determine the sample mean and sample variance through stratified random sampling, you will begin by stratifying the target population with N elements into k strata. The number of elements in stratum i for $1 \le i \le k$ is denoted by N_i, and since the strata are mutually exclusive, you will have $N_1 + N_2 + ... + N_i$. It is important to note that the number of elements in each stratum may vary from one stratum to another. If you are following the proportional allocation method, then depending on the final size (n) of your sample, you will need to determine the actual number of elements to be included in each stratum. This should be proportional to the number of elements of the same stratum in the population. In other words, if n_i is the number of elements to be chosen in the sample for stratum i, then:

$$\frac{n_i}{N_i} \approx \frac{n}{N} \text{ and } n_1 + n_2 + \ldots + n_k = n$$

Notice that the proportion is only an approximated equality because sometimes the ratio may result in a fraction and needs to be rounded to ensure that a whole number of elements is selected in each stratum.

The mean of a given stratum is then determined by:

$$\overline{y}_i = \frac{\sum_{j=1}^{n_i} y_{ij}}{n_i}$$

where y_{ij} is the j th element in stratum i.

The mean of a sample with k strata is therefore determined by:

$$\overline{y}_{st} = \frac{\sum_{i=1}^{k} N_i \overline{y}_i}{N}$$

The variance of stratum i is: $s_{y_i}^2 = \dfrac{\sum_{j=1}^{n_i} \left(y_{ij} - \overline{y}_i\right)^2}{n_i - 1}$ where $\overline{y}_i$ is the mean of that stratum.

The variance of a sample with k strata is therefore determined by:

$$s_{y_{st}}^2 = \frac{\sum_{i=1}^{k} N_i \left(s_{y_i}^2 + \left(y_i - \overline{y}\right)^2\right)}{\sum_{i=1}^{k} N_i}$$

Example D.1.4

Suppose you have to implement stratified random sampling by selecting a sample of size 20 with proportional allocation of the elements from Table 1. Find the corresponding sample mean and sample variance.

The first step is to determine the number of elements per stratum that is proportional to the number of elements in the target population.

Given the different drill types (i.e. A, B, C, and D), there are a total of four strata. For the first stratum, drill A, there are 16 elements in the target population. Hence, the proportional number of elements to be chosen in this stratum is 3 (ie. $\frac{16}{100}\times 20 = 3.2$). The table below illustrates the exact number of elements to be chosen randomly for each stratum.

Drill	Stratum Number	Number of Elements in the Population	Number of Elements in the Stratum
A	1	16	3
B	2	23	5
C	3	37	7
D	4	24	5
Total		100	20

The second step is to select the elements per stratum randomly.

The table below lists the selected elements to form the sample.

Drill	Stratum Number	Number of Elements in the Stratum	Data No.	Data Value
A	1	3	28	3.04
			50	3.01
			80	3.27
B	2	5	9	3.27
			21	3.59
			54	3.31
			75	3.19
			88	3.17
C	3	7	3	3.26
			18	3.55
			33	3.50
			45	3.14
			57	3.24
			82	3.14
			91	3.29
D	4	5	5	3.12
			25	3.03
			59	3.35
			61	3.38
			92	3.27

In order to find the sample mean, you need to calculate the mean for each stratum and then combine them together. Hence, $\overline{y}_1 = 3.106, \overline{y}_2 = 3.306, \overline{y}_3 = 3.303$ and $\overline{y}_4 = 3.230$.

Therefore the sample mean is:

$$\overline{y}_{st} = \frac{16(3.106)+23(3.306)+37(3.303)+24(3.230)}{16+23+37} = \frac{325.4}{100} \approx 3.25$$

The variances for each stratum are $s^2_{y_1} = 0.0202$, $s^2_{y_2} = 0.0285$, $s^2_{y_3} = 0.0265$, and $s^2_{y_4} = 0.0227$.

Therefore the sample variance is:

$$s^2_{y_{st}} = \frac{\sum_{i=1}^{k} N_i\left(s^2_{y_i} + (y_i - \overline{y})^2\right)}{\sum_{i=1}^{k} N_i}$$

$$= \frac{16\left[0.02023+(3.107-3.255)^2\right]+23\left[0.02848+(3.306-3.255)^2\right]+\ldots}{100}$$

$$\frac{\ldots+37\left[0.02649+(3.303-3.255)^2\right]+24\left[0.02265+(3.230-3.255)^2\right]}{100}$$

$$= \frac{3.012985}{100} \approx 0.0301$$

If the same 20 elements in the sample are chosen through simple random sampling instead of stratified random sampling the mean becomes 3.26 mm and the variance becomes 0.0264.

Example D.1.5

Suppose that you have to implement stratified random sampling by selecting a sample of size 20 with equal allocation of the elements from Table 1. Find the corresponding sample mean and the sample variance.

The first step is to determine the number elements in each stratum to form a sample with a total of 20 elements. In other words, there needs to be 5 elements in each stratum.

The second step is to randomly select 5 elements in each stratum and the table (at left), lists the selected elements in each stratum.

In order to find the sample mean, you need to calculate the mean for each stratum and then combine them together. Hence, $\overline{y}_1 = 3.106, \overline{y}_2 = 3.306, \overline{y}_3 = 3.303$ and $\overline{y}_4 = 3.230$.

Therefore the sample mean is:

$$\bar{y}=\frac{16(3.136)+23(3.306)+37(3.266)+24(3.23)}{16+23+37+24}$$
$$=\frac{324.576}{100}$$
$$\approx 3.25$$

The variances for each of the strata are $s_{y_1}^2=0.0130$, $s_{y_2}^2=0.0285$, $s_{y_3}^2=0.0218$, and $s_{y_4}^2=0.0227$.

Therefore the sample variance is:

$$s_{y_{st}}^2=\frac{\sum_{i=1}^{4}N_i\left(s_{y_i}^2+(y_i-\bar{y})^2\right)}{\sum_{i=1}^{4}N_i}$$
$$=\frac{16\left[0.01298+(3.136-3.246)^2\right]+23\left[0.02848+(3.306-3.246)^2\right]+\dots}{100}$$
$$\frac{\dots+37\left[0.02178+(3.266-3.246)^2\right]+24\left[0.02265+(3.230-3.246)^2\right]}{100}$$
$$=\frac{2.509524}{100}$$
$$\approx 0.0251$$

If the same 20 elements in the sample are chosen through simple random sampling instead of stratified random sampling the mean becomes 3.24 mm and the variance becomes 0.222.

It is interesting to note that the sample mean and the sample variance for stratified random sampling with proportional allocation are the same as the sample mean and the sample variance for stratified random sampling with equal allocation. This is not always the case, especially when the number of elements in each stratum varies greatly from one to another.

Cluster Sampling

Cluster sampling shares some of the features of stratified random sampling. The elements in the population have one particular inherited defining characteristic that makes them heterogeneous as the entire collection but homogeneous in their individual subgroups. Those subgroups, in this case the clusters, are formed by that particular characteristic. The unique characteristic is often based on gender, age, ethnicity, time, postal codes, land divisions, etc. Since the elements in the target population only differ by one unique characteristic, the clusters are also mutually exclusive. Once the clusters are formed, the next step is to randomly select one of the clusters and all elements in that chosen cluster to be used in the sample. Since the number of elements in one cluster is often known in the process, the calculation for the sample mean and the sample variance is similar to the calculation performed for simple random sampling.

Cluster sampling is useful when the population is spread over a wide geographical region, or when the total number of elements in the population is large (e.g. manufactured items packaged in boxes). Hence, the actual total number of elements in the population is irrelevant for this technique. This probability sampling method helps to reduce costs and is more time-efficient when compared with the other probability sampling techniques. However, due to sampling on all the elements in a single cluster, the technique may sometimes lead to unnecessary sampling bias and systematic error if the original cluster is chosen poorly. For example, if you decide to sample the first tray of cookies made from a large batter mixture the result may be very different from sampling the last tray.

Example D.1.6

Suppose the four columns in Table 1 represent four distinct collections of data readings. You are implementing cluster sampling by selecting one of the four columns as your cluster. Find the corresponding sample mean and sample variance.

The first step is to randomly select a column from Table 1 and you have selected column 2. Next you will need to calculate the sample mean amongst all 25 elements in column 2.

Sample Mean: $$\bar{y}=\frac{\sum_{i=1}^{25}y_i}{25}=\frac{82.69}{25}\approx 3.31(3\text{s.f.})$$

Sample Variance: $$s_y^2=\frac{\sum_{i=1}^{25}(y_i-3.307)^2}{25-1}=\frac{0.84706}{24}\approx 0.0353(3.\text{s.f.})$$

The sample mean is 3.31 mm with the sample variance of 0.0353.

Multistage Sampling and Multiphase Sampling

There are two other more advanced probability sampling techniques. These are a combination of the aforementioned techniques. In multistage sampling, you will implement more than one probability sampling technique in a particular order. For example, you may implement cluster sampling in stage 1 before you continue with systematic sampling using that particular cluster. Therefore, multistage sampling is simply sampling from a sample. The most obvious advantage in multistage sampling is to reduce the amount of resources required to complete the sampling process. The drawback, however, is directly related to the way in which a cluster or stratum is selected in any stage throughout the process. The complexity in the sampling process increases as the number of stages increases, hence, there is also an increase to potential systematic errors.

Multiphase sampling involves **iterative sampling** but at each phase the sample becomes smaller because of the additional information you have collected from the previous phase. This sampling technique is useful when there is insufficient funding to analyze a large sample from the original target population or when the investigation on a larger sample increases the burden or puts stress on the elements. Multiphase sampling helps to reduce systematic error. However, given that samples in subsequent phases are chosen based on information previously collected, the sampling technique may lead to potential sampling bias when the investigator may have purposefully selected elements with predetermined decisions.

Although multistage and multiphase sampling techniques are similar in name they are very different in their structures, especially when you take a closer look at the composition of the elements in the sample. In multiphase sampling, the samples at any phase are taken from the same sample frame and the elements in those samples are structurally the same. However, in multistage sampling, the elements in any given stage may be very different from its previous stage because of the chosen sampling technique.

Compare and Contrast the Different Sampling Techniques

In order to better illustrate the differences between the four main types of probability sampling techniques, the following table summarizes their unique features, brief descriptions, their benefits, and limitations.

Technique	Requirements	Descriptions	Benefits	Limitations
Simple Random Sampling	• an exhaustive list of all elements in the population • homogeneous elements in the population	• every element in the population has an equal chance to be sampled • each element is sampled randomly through a lottery or by a random number table	• reduces systematic errors and sampling bias • easiest to implement	• impractical for a large population • not applicable if not all of the elements in the population are known in advance
Systematic Sampling	• homogeneous elements in the population	• elements in the population are chosen at regular sampling intervals • first element is selected at random, then all subsequent elements are selected based on the interval	• draws a sample from the population when there is an unknown number of elements	• impractical for a large population • introduces sampling bias if the first element is incorrectly chosen or the sampling interval does not match the context

Technique	Requirements	Descriptions	Benefits	Limitations
Stratified Random Sampling	• heterogeneous elements in the population	• divide the population into homogenous strata • randomly select elements in each strata by either proportional allocation or equal allocation	• ensures elements in the sample are drawn proportionally • works well with a large population	• requires more resources to implement
Cluster Sampling	• heterogeneous elements in the population	• divide the population into homogenous clusters • randomly select a cluster and sample all elements in that chosen cluster	• works well with a large population • reduces the resources to select a sample	• introduces systematic error if a cluster is chosen incorrectly • elements in the chosen cluster may not reflect the diversity in the population

Non-Probability Sampling Techniques

The second main category of sampling techniques is non-probability sampling in which the selection of the sample is not random. Hence it is impossible to determine the probability of selecting a particular element from the population. Due to the non-probabilistic nature in selecting a sample, the sampling procedure relies heavily on your subjective judgement. The statistical results gathered from these non-probability samples can rarely be generalized to the population, but rather they help to inform you in developing an argument relating to an issue. The actual implementation of these non-probability techniques requires many fewer resources than probability sampling techniques and the procedures are often less sophisticated. However, with the reduction in the complexity in sampling, there is obviously a much higher chance of introducing systematic errors and sampling bias.

There are several non-probability sampling techniques. For the purpose of this syllabus, the discussion on the theoretical differences and the actual implementation procedures will be limited to:

1. volunteer sampling,
2. convenience sampling,
3. purposive sampling,
4. quota sampling,
5. snowball sampling, and
6. matched sampling.

Given that the probability for a particular element in the population to be selected for the sample cannot be determined in any of these non-probability sampling techniques, it is not possible to determine the sample mean and sample variance.

Volunteer Sampling

For volunteer sampling, the elements in the target population self-select themselves into the sample. These elements are often human participants when they individually make the deliberate choice to become part of the sample. The benefit in this sampling technique is that it gathers the largest possible sample in the shortest amount of time if the researcher has provided sufficient incentive for potential participants to join. The major limitation is that the actual participants in the sample may already share a similar interest with each other and with the researcher. Hence, it is known for producing one of the greatest sampling biases.

Internet users often participate in volunteer sampling researches when they submit their responses to survey questionnaires for a company or for a service. For example, if you click on an advertisement on your social-media platform to complete a short survey for entering into a lucky draw, or if you place your business card in a box beside the cashier in a coffee shop to win their monthly grand prize, you have self-selected yourself into that sample.

Convenience Sampling

Convenience sampling, also known as **accidental sampling** or **haphazard sampling**, is when, after defining the population, the researcher selects the sample from those elements that are readily available at a particular moment. This technique is prone to have systematic errors, but it is one of the techniques that require minimum effort and resources.

If you are studying the preferences of the potential buyers of the latest smart-phone models, you can use the convenience sampling technique by surveying the customers queuing outside a phone store. Similarly, if you would like to understand passenger behaviours on public buses, you may choose to ride a few bus routes and make your observations.

Purposive Sampling

Purposive sampling, sometimes also referred to as **judgemental sampling**, is similar to convenience sampling. The researcher defines a limited scope in the study to serve a particular purpose or to have a predefined result. Hence the sample selection is based on their favourability towards the same issue. Samples collected through this technique are often homogeneous.

Some lobbying groups may use purposive sampling to craft their discussion papers in the direction they want. Likewise, some marketing groups may select a sample to specifically highlight the success of an advertising campaign. For example, if you want to understand the effectiveness of your leadership style in student council, you may want to sample your supporters only. Notice that you are only interested in those who support you.

Quota Sampling

As the name suggests, for quota sampling, the researcher selects readily available elements to form a sample until the quota is reached. The elements in the sample are often heterogeneous since they may represent a widely diverse population. However, depending on the actual procedure in sampling, it is likely that there will be sampling bias.

You may have been stopped by a research analyst on the street to collect your opinion on responses to their survey or have been given free gifts to rate a particular product. Those researchers are carrying out quota sampling, because it is likely that they will stop surveying after their quota is met or when their free gifts run out. For example, if you are researching the average spending per shopper at a nearby grocery store, you may want to survey the first 100 shoppers. The limitation in the design of this sampling procedure is the fact that you have not considered when the store is busiest and with higher spending shoppers.

Snowball Sampling

Snowball sampling is often used when the entire target population is difficult to define or access. The researcher will begin selecting the first sample, and from the first sample the researcher gathers information and suggestions for the second sample, and so on. This sampling technique is also known as **network sampling**, **chain-referred sampling**, or **reputational sampling**.

When completing your extended essay, you begin to read the first journal article you have found. That article has several references to another article, so you begin to research that second article. From the second article, you find quotes from and important thoughts from an expert in the field, so you continue to research about this expert. This is an example of snowball sampling in which your subsequent sample is suggested by a previous sample.

Matched Sampling

This sampling technique is often used in experimental research. A pair of two matched elements is sampled from the target population. This sampling technique allows the researchers to study the effect on the control group and vice versa. The challenge in selecting a pair of elements with this technique is that both elements must match in every aspect as defined by the researcher.

In order to illustrate the process in matched sampling, suppose that you want to investigate whether there is a correlation between a student taking IB SL Mathematics and IB SL Physics. From the class registers, you sample just those students who take both of these two courses. If a student only takes SL Mathematics or only takes SL Physics, then that student is not sampled.

Choice of Sampling Technique

Amongst the choices between probability and non-probability techniques, there is actually no single rule in defining which technique is the most suitable for sampling in general. The selection of a particular technique depends on the nature and the context of the research and the availability of resources. As a researcher, you need to consider all the factors to ensure you have minimized potential systematic errors and sampling bias.

Another question that also has an indefinite answer relates to the actual sample size. Similar to the selection of a sampling technique, the sample size also depends on the same factors mentioned above. As long as the sample represents the population, then its size is suitable.

Exercise D.1.1

1. Generate samples by selecting pages from this mathematics textbook using:

    a. simple random sampling.

    b. systematic random sampling.

    c. stratified random sampling, with proportional allocation of elements.

    d. stratified random sampling, with equal allocation of elements.

    e. cluster random sampling.

    f. convenience sampling.

    g. snowball sampling.

2. A small apartment complex has twenty 2-bedroom units and four 3-bedroom units. The 2-bedroom units are

Floor	Unit Number
1	101
	102
	103
	104
	105
2	201
	202
	203
	204
	205
3	301
	302
	303
	304
	305
4	401
	402
	403
	404
	405

and the 3-bedroom units are

Floor	Unit Number
1	106
3	306
2	206
4	406

The management decides to select six 2-bedroom units and one 3-bedroom unit by random sampling to measure the approval rating of the newly hired cleaning company. Give the unit numbers of those to form one possible sample, if the company uses stratified random sampling.

3. Give the steps necessary in drawing a 13-card hand from a standard deck of playing cards using:

    a. simple random sampling.

    b. systematic random sampling.

    c. stratified random sampling.

    d. cluster random sampling.

    e. multistage random sampling.

    f. purposive sampling.

4. The following table lists the postal codes for four districts in a city.

City	Postal Code
1	L2201
	L2202
	L2203
	L2204
	L2205
2	L2401
	L2402
	L2403
	L2404
3	L3511
	L3512
	L3513
	L3514
	L3515
	L3516
4	L3941
	L3942
	L3943
	L3944
	L3945
	L3946

Identify an appropriate sampling method that has been used to select each of the following samples:

a. L3941, L3942, L3943, L3944, L3945, L3946

b. L2201, L2203, L2205

c. L2201, L2204, L2401, L2404, L3511, L3514, L3941, L3944

d. L2203, L2401, L3512, L3946

e. L2204, L2402, L3513, L3516, L3942, L3943

5. Identify, and explain, an appropriate sampling method for each of the following scenarios:

a. a quality control procedure to select a sample of 100 rulers, from a crate of boxes, each box has bundles, and each bundle has rulers.

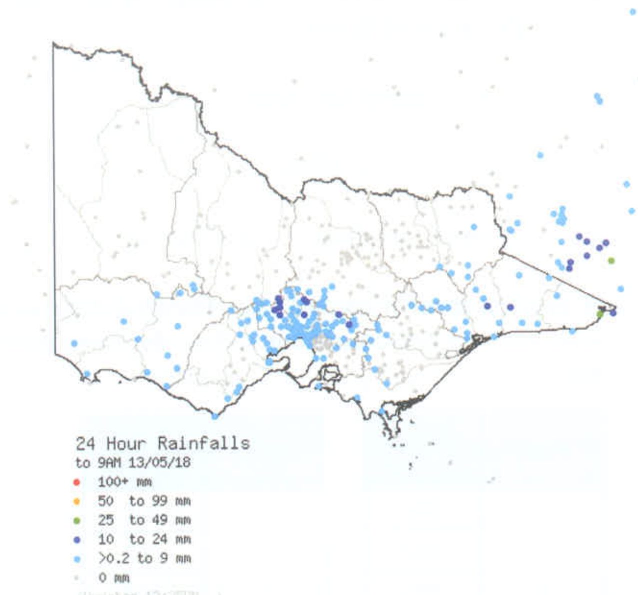
b. a sample of cars from a parking lot with indoor parking spots and outdoor parking spots.

c. a group panel of experts to present in a forum.

d. a sample of students from a school of students to try the new lunch menu for the following semester.

Answers will vary.

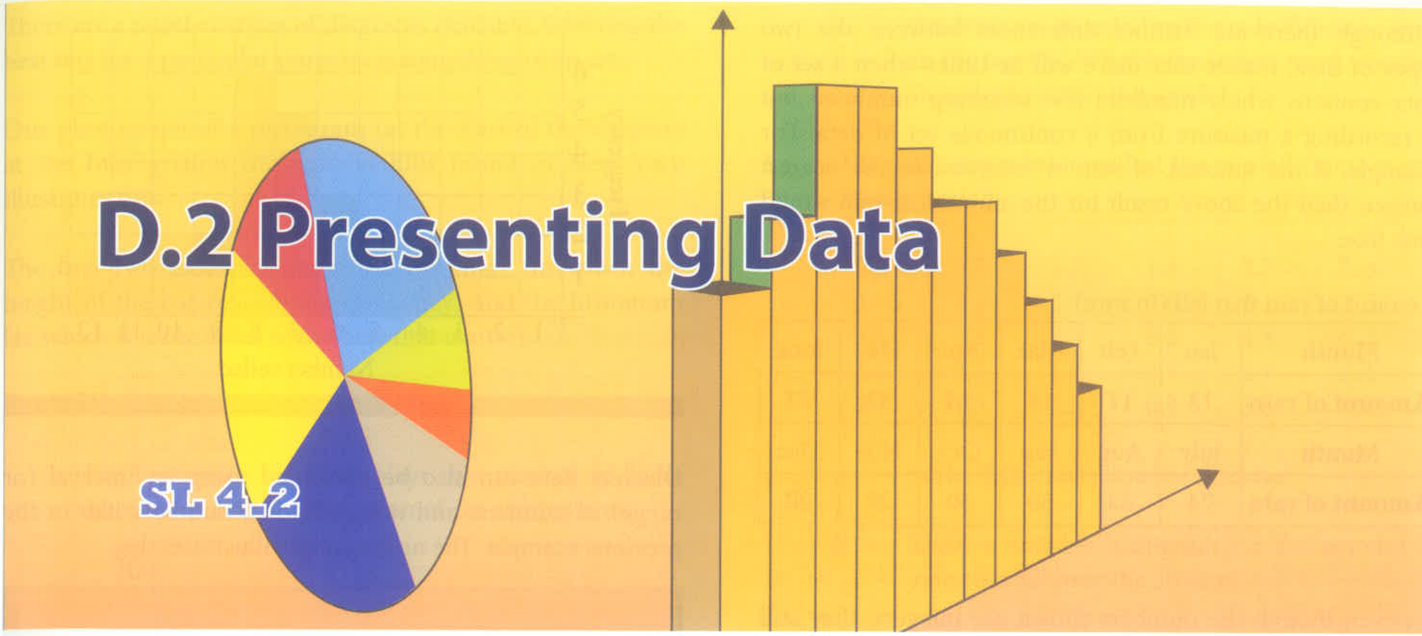
The diagram shows the location of rain gauges in the State of Victoria in Australia.



The current data can be found at: http://www.bom.gov.au/vic/flood/

What sort of sampling is evident here?

If you were asked to recommend a redistribution of the gauges, what would you say?



D.2 Presenting Data

SL 4.2

Displaying Data

There are a number of ways that we may display sets of data. The point of 'displaying data' is to make its meaning clear to others. Long lists of figures convey very little. The data display is a good example of how a well chosen graphic can make patterns more obvious. In this case, the aim of the display is to explain patterns in migration to the USA to tourists.

Statistical Display, Ellis Island, New York.

Our first task is to deal with the types of data we collect, that is, words or numbers, discrete or continuous etc.

If the sets of numeric data are counted, they are considered discrete and if they are measured, they are considered continuous. So, for example, if we were counting the number of fish that were caught over a period of 365 days, this would be considered as a discrete measure (as we are carrying out a counting process). However, if we were looking to carry out some analysis about the length of these fish, then that would be considered as a continuous measure (as we are carrying out a measuring process).

Similarly, if we were looking at rainfall, we could measure the number of days on which it rained – which would be considered to be a counting process or, the amount of rain that fell on each day – which would be considered as a measuring process. So, in the case of the rainfall example, we could have the following tables of data:

Discrete (counting the number of days):

Number of days on which it rained

Month	Jan	Feb	Mar	Apr	May	June
No. of days	4	4	5	9	14	18
Month	July	Aug	Sep	Oct	Nov	Dec
No. of days	19	19	14	13	7	5

Continuous (measuring the amount of rain):

Amount of rain that fell (in mm)

Month	Jan	Feb	Mar	Apr	May	June
Amount of rain	12.5	11.0	14.3	13.7	31.5	53.2
Month	July	Aug	Sep	Oct	Nov	Dec
Amount of rain	73.5	82.9	50.4	30.1	28.7	20.2

From the data, we also observe the different forms that the data takes on. For the discrete data (the number of days) we have whole numbers, i.e. counting numbers. For the continuous data (the amount of rain) we have rational numbers.

Although there are distinct differences between the two types of data, realize that there will be times when a set of data contains whole numbers (i.e. counting numbers) but is recording a measure from a continuous set of data. For example, if the amount of rain is recorded to the nearest integer, then the above result for the amount of rain would look like:

Amount of rain that fell (in mm)

Month	Jan	Feb	Mar	Apr	May	June
Amount of rain	13	11	14	14	32	53
Month	July	Aug	Sep	Oct	Nov	Dec
Amount of rain	74	83	50	30	29	20

So, even though the numbers shown are integers, they still reflect a measuring process and so are still considered as continuous.

You have already been exposed to some statistical work in the past, such as representing data in table form or graphical form. Here, we will review this by way of examples so that you may be reminded of the elements involved when dealing with these types of data.

Example D.2.1

Navneet rolls two dice 26 times, and records the sum showing uppermost on each throw:

5	6	11	6	10	2
3	7	8	4	7	9
8	7	8	12	8	4
6	5	7	8	7	8
9	10				

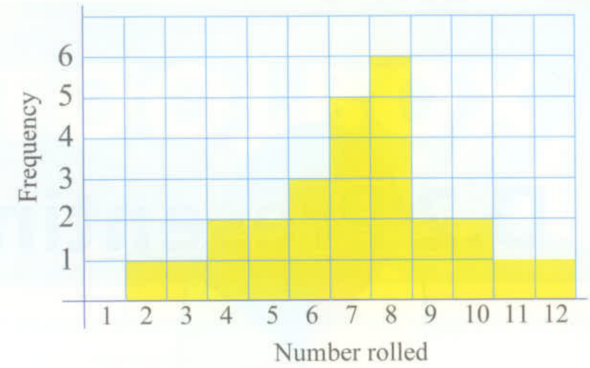
Construct a frequency table for his results and represent the data accordingly.

In this instance, we note that we have a counting process, meaning that we are dealing with a discrete data set. So, we can set up a frequency table.

Sum from rolling two dice

Score (x)	2	3	4	5	6	7	8	9	10	11	12
Frequency (f)	1	1	2	2	3	5	6	2	2	1	1

Based on these results, we can now draw our graph, with the axes labelled appropriately, frequency on the vertical axis and score (number rolled) on the horizontal axis:



Discrete data can also be presented using an interval (or range) of numbers and not individual numbers like in the previous example. The next example illustrates this.

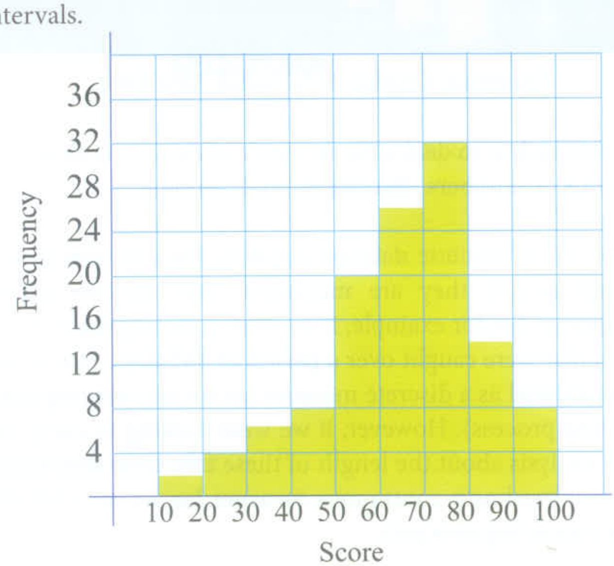
Example D.2.2

One hundred students sat a test, marked out of 120. Students were only awarded whole marks for their test. The results are shown in the table below. Draw a histogram for the table below.

Student test results

Score range (x)	[0, 10[	[10,20[	[20, 30[	[30, 40[	[40, 50[
Frequency (f)	0	2	4	6	8
Score range (x)	[50, 60[	[60, 70[	[70, 80[	[80, 90[	[90, 100]
Frequency (f)	20	26	32	14	8

Again, we have a discrete data set with well defined intervals. As such, we may proceed with our histogram using the given intervals.



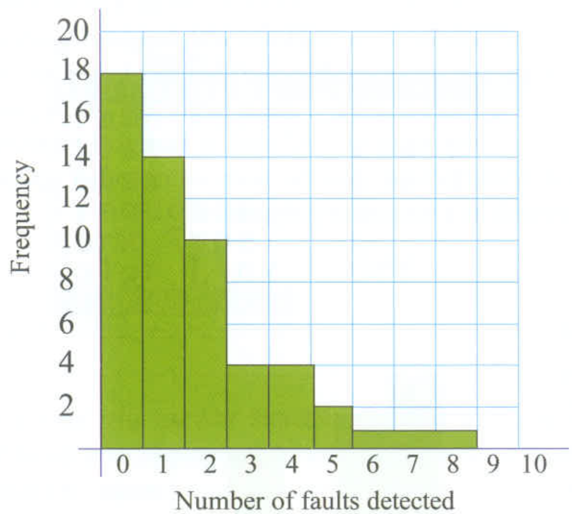
There are a number styles of diagrams available. Selecting the best one for a particular purpose is something of an art.

Our photograph of a pictogram (at the start of the chapter) at the Immigration Museum at Ellis Island in New York illustrates this.

The first two examples show the bar-chart (in which the height of the bar reflects the frequency) and the histogram (in which it is the areas of the bars that matter).

Exercise D.2.1

1. The histogram below displays the number of faults detected during an inspection of car components.



a How many components were inspected?

b What percentage of components had no faults?

c What percentage of components had at least 5 faults?

2. The table below shows the frequency distribution of marks in a science test by 500 students.

Test scores

Score range (x)	[0, 20[	[10, 20[	[20,30[	[30, 40[	[40, 50[
Frequency (f)	0	45	85	145	105
Score range (x)	[50, 60[	[60, 70[	[70, 80[	[80, 90[	[90, 100[
Frequency (f)	60	25	15	20	0

a Draw a histogram for the students' marks.

b If the pass mark is 30, what proportion of students passed?

c Students are awarded an A-grade if their score is in the top 4%. What is the lowest mark possible in order to attain this grade?

3. The data shown below reflect the rainfall (mm) over a 30-day period:

2.0	3.7	3.2	1.5	2.7	7.5
10.5	8.7	2.2	4.6	3.1	2.5
1.7	7.3	2.2	5.2	4.8	6.2
2.1	7.2	1.2	4.7	2.7	2.1
8.1	1.3	2.5	0.9	5.6	12.2

a Is the data continuous or discrete?

b Draw a suitable histogram for the amount of rain that fell over the 30 days.

Choice of Diagram

We have looked at bar charts and histograms as a way of displaying data. Other choices are discussed below. In making your choice in a particular context, remember that the point of using a graphic is to make the meaning of your data clearer than the data does on its own.

Histogram

The terms 'bar chart' and 'histogram' are sometimes used as if they are interchangeable. They are not. In a bar chart, it is the height of the bars that reflect the frequency of the data. In a histogram it is the area of the bar that is the key quantity. This makes the histogram a bit more flexible in advanced applocations.

The following figures are the heights (in centimetres) of a group of students:

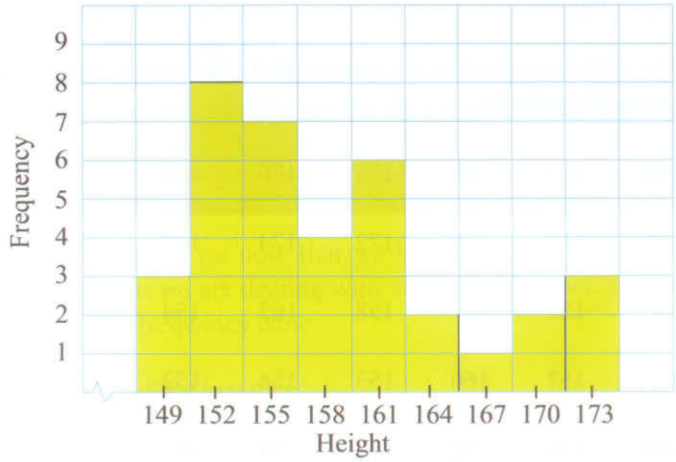
156	172	168	153	170	160	170
156	160	160	172	174	150	160
163	152	157	158	162	154	159
163	157	160	153	154	152	155
150	150	152	152	154	151	151
154						

These figures alone do not give us much information about the heights of this group of people. One of the first things that is usually done in undertaking an analysis is to make a frequency table. In this case, as there are a large number of different heights, it is a good idea to group the height data into the categories (or classes) 148–150, 151–153, 154–156, etc. before making a tally.

Height	Tally	Frequency
148–150	///	3
151–153	////////	8
154–156	///////	7
157–159	////	4
160–162	//////	6
163–165	//	2
166–168	/	1
169–171	//	2
172–174	///	3

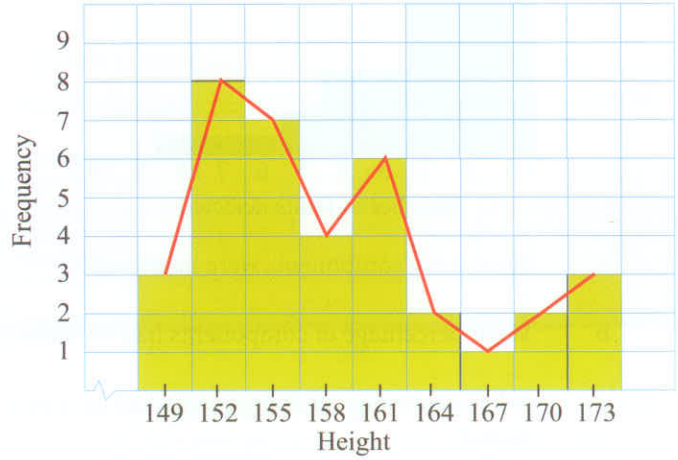
Each height is recorded in the appropriate row of the tally column. Finally, the frequency is the number of tally marks in each row. As a check, the total of the frequency column should equal the count of the number of data items. In this case there are 36 heights.

The choice of class interval in making such a frequency table is generally made so that there are about ten classes. This is not inevitably the case and it is true to say that this choice is an art rather than a science. The objective is to show the distribution of the data as clearly as possible. This can best be seen when the data is shown graphically. There are a number of ways in which this can be done. In the present example, we are dealing with heights. Since heights vary continuously, we would usually use a histogram to display the distribution.



There are two details connected with the construction of histograms that you should not ignore. Firstly, as far as the horizontal scales are concerned, we are representing the continuous variable 'height'. The first class interval represents all the people with heights in the range 148 to 150 cm. Since these have been rounded to the nearest whole centimetre, anyone with a height from 147.5 to 150.5 cm, or [147.5, 150.5], will have been placed in this class. Similarly, anyone with a height in the range [150.5, 153.5) will be categorized in the class 151–153 cm. If you want to label the divisions between the blocks on the histogram, technically these should be 147.5, 150.5 etc. Secondly, in a histogram, it is the area of the bars and not their height that represents the number of data items in each class. To be completely correct, we should give the area as a measure of the vertical scale. This definition allows us to draw histograms with class intervals of varying widths. This is sometimes done to smooth out the variations at the extremes of a distribution where the data is sparse.

Once we have drawn a histogram, it should be possible to see any patterns that exist in the data. In this case, there is a big group of students with heights from about 150 to 160 cm. There are also quite a few students with heights significantly larger than this and very few with heights below the main group. The distribution has a much larger 'tail' at the positive end than at the negative end and is said to be positively skewed. Patterns can also be seen using other graphical devices such as a frequency polygon:

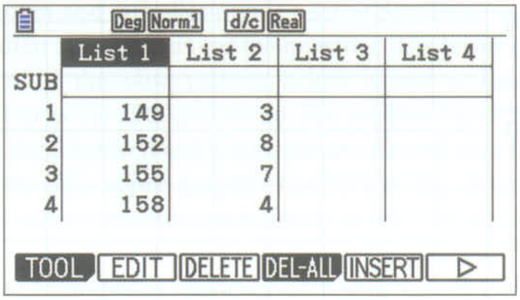


The same patterns are evident from this diagram as were seen from the histogram.

Many of these diagrams are available on graphic calculators. The details vary (consult your manual), but mostly, the calculator needs to be in Statistics mode.

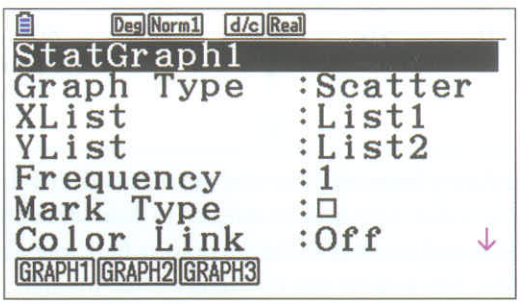
This is how it looks on Casio calculators:

If working with a single set of data, it can be entered as a single list. This example has grouped data. The data values are entered in list 1 and the frequencies are entered as list 2.

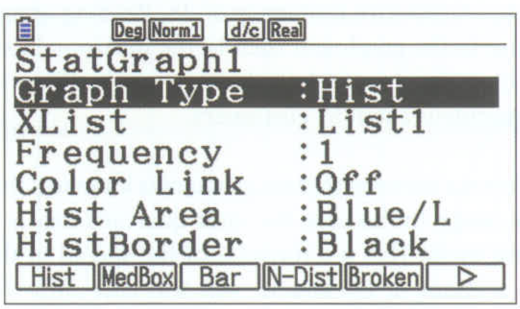


(not all data shown)

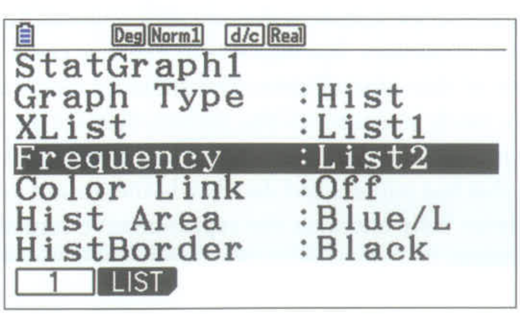
The next step is to set the type of graph (casio keying is F1 - GRAPH, F6 - SET.



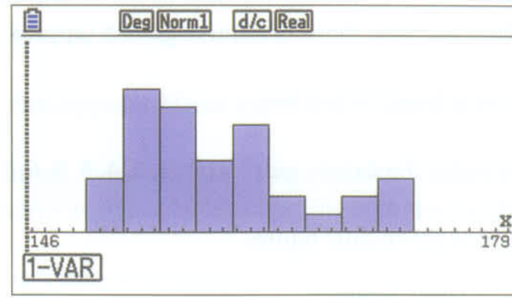
This screen tells us that StatGraph1 is set to 'Scatter'. We want to set it to 'Histogram' by editing the second line of the settings F6 - more, :



The frequency must also be set to List2 (as that is where the frequencies are stored)



The histogram can then be displayed using EXIT, F1 - GRAPH1, EXE



Cumulative Frequency Curve

Bar Charts, Histograms and Frequency Polygons are useful for displaying the general characteristics of data. However, data often has two important characteristics that we want to emphasise. The first is **central tendency** - the tendency of data to 'clump' around a central value. There are three commonly used measures of this:

Mode: the most commonly occurring value.

Mean: obtained by adding the data and dividing by the number of data items.

Median: the value with as many data below the median as above.

The other characteristic is **spread** - the tendency of the data to differ from the central tendency.

The point of this can be a bit difficult to understand. At this stage, consider these two examples:

1. Ball bearing manufacture: the bearings need to be as identical as possible. The manufacturer will want to achieve a small spread when the diameters of the products are checked.

2. Ability testing. An ability test is only useful if the competent students score more than the less competent. The setters are looking for a reasonably big spread in the results.

In this section we will look at **inter-quartile range** as a measure of this. This is considered below. Other measures of spread are covered elsewhere.

The following data represent the number of employees absent from work over a nine day period:

2, 6, 5, 4, 7, 1, 0, 5, 2.

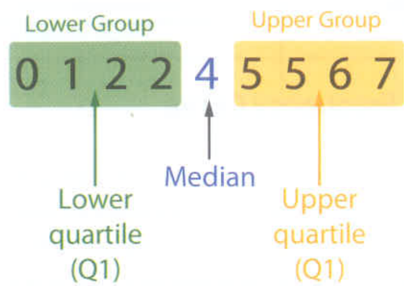
The mode is 2 as it occurs more often than the others.

The mean is found by adding the data to get 32 and dividing by 9, as there are nine items of data, to get 3.6 (approx.).

The median is found as follows:

Firstly, we order the data to get 0, 1, 2, 2, 4, 5, 5, 6, 7.

The median is the middle figure:



The median divides the distribution into an upper and lower group. The lower quartile is the middle figure of the lower group and the upper quartile is the middle figure of the upper group. As with finding the median, it may be necessary when dealing with a group with an even number of data items to take the mid point between two numbers. This is the case with the current data set. The lower quartile is 1.5 and the upper quartile is 5.5.

When dealing with large data sets or grouped data, there is an alternative method of finding the median and quartiles based on **cumulative frequency**.

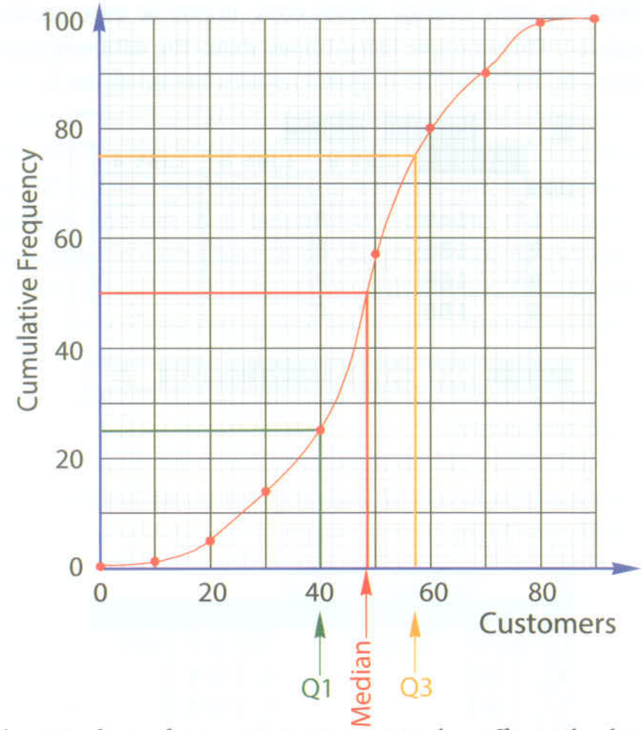
This is calculated as follows:

These figures represent the numbers of customers in a small cinema:

Customers	Frequency	Cumulative Frequency
0-9	1	1
10-19	4	5
20-29	9	14
30-39	11	25
40-49	32	57
50-59	23	80
60-69	10	90
70-79	9	99
80-89	1	100

The cumulative frequency is calculated by 'accumulating' the frequencies as we move down the table. Thus the figure 25 in the shaded box means that on 25 occasions there were fewer than 40 customers.

Cumulative frequencies can now be used to produce a **cumulative frequency curve**:



The cumulative frequency curve or **ogive** has effectively placed the data in order. This now enables us to read off estimates of the median and quartiles. The median is half way along the list of data. Since there are 100 figures, the median point is at 50. Technically this should be figure number 51, however, this method only produces an approximate figure and we seldom worry about this distinction. Reading across from 50 and down to the 'customers' scale gives a figure of about 48 customers as the median. Similarly, the lower quartile can be found at a cumulative frequency of 25. Reading across from this figure to the graph and then to the horizontal axis gives a lower quartile of approximately 40 customers. Similarly, the upper quartile is about 57 customers.

The difference between the two quartiles is known as the **inter-quartile range**. In this case, the inter-quartile range is 57 – 40 = 17 customers. This is a measure of the spread of the data. In later sections of the course, you will encounter variance and standard deviation. These are alternative measures of spread which may differ from inter-quartile range quite considerably. In choosing which measure of spread to use, we generally use the quartiles and the median for a data set that contains a few numbers that are very unusual. Such data are known as **outliers**. The data sets in Exercise D.2.3 contain examples of this type of data. The standard deviation and mean are much more sensitive to outliers than are the median and inter-quartile range. Of course, you will need to look at a data set that has outliers and decide whether or not you want to minimise their effect on the representative statistics that you calculate.

Do not expect anything useful to flow from comparing the inter-quartile range of one data set with the variance of another!

Percentiles

The median and quartiles divide a set of data into halves (for the median and quarters for the quartiles. The **lower quartile** has 25% of the observations below it and is sometimes referred to as the 25th percentile. The **median** has 50% of the observations below it and is sometimes referred to as the 50th percentile. The **upper quartile** has 75% of the observations below it and is sometimes referred to as the 75th percentile.

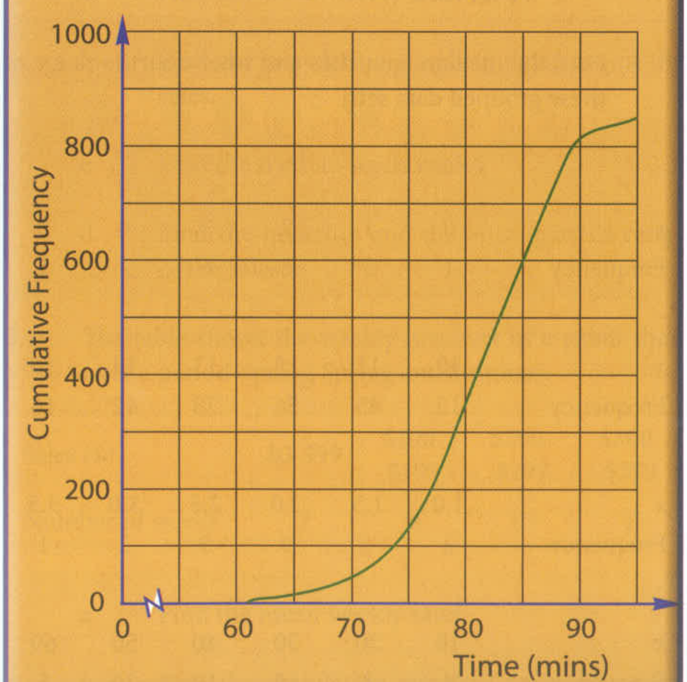
There are a number of applications in which other percentiles are useful. It may be that a 'Distinction' in a course is awarded to the top 10% of the students - in which case, the cut off mark would be the 90th percentile (90% 'less than' and 10% 'more than').

The term 'decile' is also used to divide a distribution into tenths. Thus the 3rd decile is $^{3}/_{10}{}^{\text{ths}}$ of the way from the bottom to the top.

Example D.2.3

The organisers of a 'fun run' award 'Pheidippides Medals' to the fastest 10% of the runners and 'Participation Awards' to the slowest 20%. The remainder receive 'Commemorative Medallions'.

The results are shown on this graph:



From the graph, there were 850 runners. The 10th percentile is at the 85th runner

This is because 10% of 850 is 85.

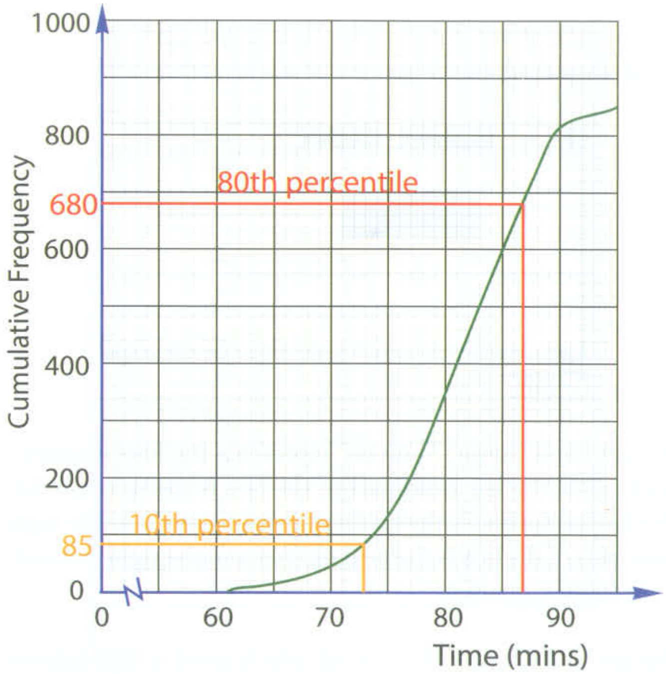
So the Pheidippides Medal cutoff is at 73 minutes.

The cut-off between the Participation Award and the Commemorative Medallions is at the 80th percentile.

Since there were 850 runners:

80% of 850 is $0.8 \times 850 = 680$.

These two points are added to the graph:

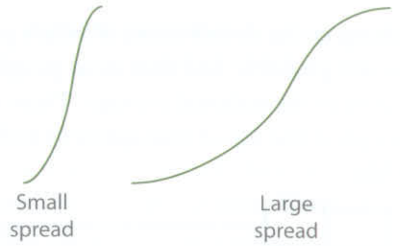


Runners who completed the course in 73 minutes or fewer receive Pheidippides Medals.

Runners who took longer than 87 minutes receive Participation Awards.

The remainder receive Commemorative Medallions.

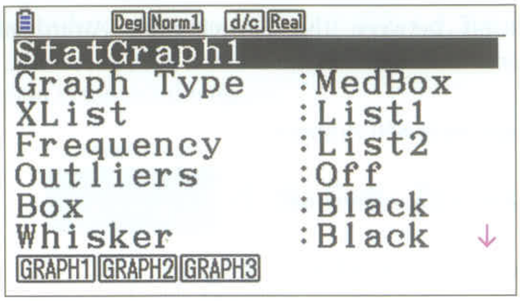
The general shape of the cumulative frequency curve does give some indication of the spread of data:



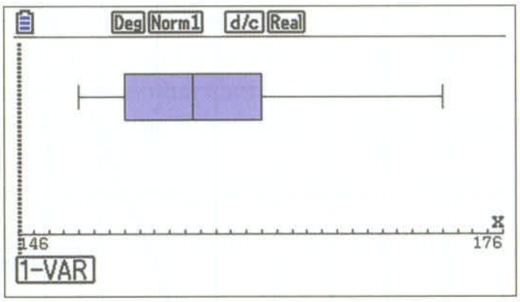
Box and Whisker Plot

Another display that is often used to is known as a 'box and whisker plot'.

The Casio calls this a MedBox graph:



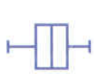
When displayed, the graph is:



The left hand (lower) 'whisker' represents the lowest quarter of the data, the left hand part of the box is the next quarter. The line in the box is the half way point or median. The right hand part of the box is the next quarter and the right 'whisker' is the highest quarter.

The box and whisker plot is particularly good at highlighting patterns in data.

For example, if the data is quite clumped and symmetrically distributed (as in the case of ball bearing manufacture) the box and whisker plot will look a bit like this:

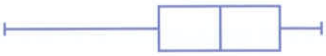


If the data is symmetric but well spread (as our ability testers will want), the diagram will be more like:

If you are looking at the distribution of salary packages in a company, you will probably find that most people are paid at a 'modest' level near the national average. There will be a few highly paid people at the top of management and the plot will be somewhat like:

In extreme cases, such as a small company with one highly paid owner, the extreme figure can be displayed as an **outlier** - see below:

Distributions can be asymmetric in the reverse sense. Discounted fuel prices can look like this if a small number of fuel (gas) stations try to generate business and ambush their competitors with a short price cutting blitz.



Exercise D.2.2

1. Find the median, quartiles and inter-quartile range of these data sets:

   a 5, 6, 2, 8, 9, 2, 7, 0, 5, 3

   b 2.8, 4.9, 2.8, 0.9, 3.3, 5.8, 2.9, 3.7, 6.9, 3.3, 5.1

   c 142, 167, 143, 126, 182, 199, 172, 164, 144, 163, 192, 101, 183, 153

   d 0.02, 0.25, 1.72, 0.93, 0.99, 1.62, 0.67, 1.42, 1.75, 0.04, 1.12, 1.93

   e 1200, 2046, 5035, 4512, 7242, 6252, 5252, 8352, 6242, 1232

2. Find the median, quartiles and inter-quartile range of these grouped data sets:

a

x	0	1	2	3	4	5
Frequency	1	3	6	6	7	1

b

x	10	11	12	13	14	15
Frequency	12	45	56	78	42	16

c

x	1.0	1.5	2.0	2.5	3.0	3.5
Frequency	2	4	9	9	2	1

d

x	10	20	30	40	50	60
Frequency	4	8	15	19	20	5

e

x	0	5	10	15	20	25
Frequency	0	3	0	6	7	5

3. The weekly expenses paid to a group of employees of a small company were

$25	$0	$10	$10
$55	$0	$12	$375
$75	$445	$7	$2

a Find the mean weekly expense.

b Find the population standard deviation of the expenses.

c Find the median weekly expense.

d Find the quartiles and the inter-quartile range of the expenses.

e Which of these statistics are the best representatives of the expenses data?

4. The table shows the numbers of cars per week sold by a dealership over a year.

Cars sold	0	1	2	3	4	5
Number of weeks	2	13	15	12	7	3

a Find the mean weekly sales.

b Find the population standard deviation of the sales.

c Find the median weekly sales.

d Find the quartiles and the inter-quartile range of the sales.

5. The table shows the weekly turnover of a small shop over a period during Spring and Summer.

Sales ($)	$0-$99	$100-$199	$200-$299	$300-$399
Number of weeks	2	9	15	7

a Find the mean weekly sales.

b Find the population standard deviation of the sales.

c Construct a cumulative frequency table and draw the cumulative frequency curve.

d Find the median weekly sales from your graph.

e Find the quartiles and the inter-quartile range of the sales from your graph.

6. Plot the cumulative frequency curves for these data and hence estimate the median,

quartiles and inter-quartile range of the data.

x	0-4	5-9	10-14	15-19	20-24	25-29
Frequency	2	5	11	9	7	2

7. A rehabilitation hospital conducts manual dexterity test on its patients. The number of errors made in the test is shown in the table:

Number of Errors	Number of Patients
0	1
1	3
2	11
3	28
4	34
5	14
6	15
7	23
8	11
9	7
10	2
11	1

Find:

a The percentage of patients who made 5 errors or fewer.

b The percentage of patients who made 7 errors.

c The 10th percentile.

d The 30th percentile.

e The 60th percentile.

f The 90th percentile.

Outliers

Occasionally a coherent set of data will have a few very unusual figures known as outliers. There are a number of reasons for this. Here are three examples.

1. It is common practice in science to repeat experiments. Thus, if a scientist is measuring the density of a new polymer, s/he will measure a number of samples. This should lead to data that are very closely the same. If, amongst these, there is a result that differs substantially from the others, it is common to assume that an error was made in reading an instrument or recording a number. In this case the outlier is deleted from the data set.

2. Real Estate Agents commonly publish data on property values. These vary from area to area within a city, but not that much within a particular suburb. However, there may be a few derelict properties with values much smaller than the norm. To prevent these from skewing the general data downwards, Agents commonly quote the median value for a suburb rather than the mean. We will discuss these measures in more detail in the next chapter.

3. If you are applying for a job at a small company, you might be interested in what the general levels of pay that are offered to the existing workers. Typically, most of the employees will be paid at a similar level. However, it is common for the top manager to be paid much more (an outlier) than the general level. You may choose to discard this number when making your assessment of your likely remuneration.

An outlier is often defined as a value that is more than 1.5 times the interquartile range from the nearest quartile. Thus, if a data set has $Q_1 = 15$ and $Q_2 = 21$, the IQR is 6 ... 1.5×6=9 so any datum smaller than 15–9 =6 or bigger than 21+9=30 is classified as an outlier.

Exercise D.2.3

1. The following figures contain an outlier. Use a calculator to show the box and whisker plot with the outlier shown as a part of the data and, separately, as an outlier.

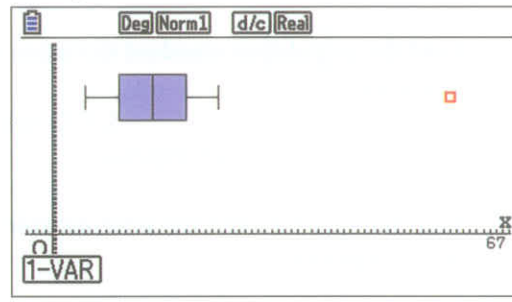
124	126	122	123	120	121
125	124	124	125	120	122
123	125	147	125	123	121

2. The following figures are all measurements of the density (gm/cc) of an experimental alloy:

2.738	2.735	2.730	2.735
2.735	2.733	2.737	2.735
2.737	2.731	2.735	2.732
2.735	2.736	27.36	2.737
2.730	2.734	2.732	2.730

What should the technicians record as the density?

3. The screen below shows the donations received by a charity.



What level of donation is 'normal'?

Answers

As we observed in the previous chapter, Statistics provides a number of good topics for relevant investigations. We have already mentioned a few. Here are two more suggestions.

1. **House prices** in your city/state/county/country. These distributions are often skewed. Comparisons between parts of a city can lead to important conclusions relevant to house buyers, investors etc. The raw data is often freely available.

2. The distribution of **wealth** in your city etc. Wealth is not the same as income. The **Gini Index** is a measure of wealth distribution. The web carries data on this and the discussions of 'fairness' that follow.

D.3 Statistical Measures

SL 4.3

Further Measures

We have already considered mode, mean and median as measures of central tendency. Also, we have looked at spread.

There are, however other more complex measures of spread. In this section, we look at **variance**.

Variance

The following sets of data are test results obtained by a group of students in two tests in which the maximum mark was 20.

Test 1:

4	12	11	10	5	10	12	12
6	8	19	13	3	7	11	13
4	9	12	10	6	13	19	11
3	12	14	11	6	13	16	11
5	10	12	13	7	8	13	14
6	10	12	10	7	10	12	10

Test 2:

9	8	10	10	8	9	10	11
8	8	11	10	9	8	11	10
9	8	10	11	8	9	11	10
9	8	11	11	9	9	11	10
8	9	11	10	8	9	11	11
8	8	11	10	8	9	10	10

The means of the two data sets are fairly close to one another (Test 1, 10.1, Test 2, 9.5). However, there is a substantial difference between the two sets which can be seen from the frequency tables.

Test 1:

Mark	3	4	5	6	7	8	9	10	11
Frequency	2	2	2	4	3	2	1	8	5
Mark	12	13	14	15	16	17	18	19	
Frequency	8	6	2	0	1	0	0	2	

Test 2:

Mark	3	4	5	6	7	8	9	10	11
Frequency	0	0	0	0	0	13	11	12	12
Mark	12	13	14	15	16	17	18	19	
Frequency	0	0	0	0	0	0	0	0	

The marks for Test 1 are quite spread out across the available scores whereas those for Test 2 are concentrated around 9, 10 and 11. This may be important as the usual reason for setting tests is to rank students in order of their performance. Test 2 is less effective at this than Test 1 because the marks have a very small spread. In fact, when teachers and examiners set a test, they are more interested in getting a good spread of marks than they are in getting a particular value for the mean. By contrast, manufacturers of precision engineering products want a small spread on the dimensions of the articles they make. Either way, it is necessary to have a way of calculating a numerical measure of the spread of data. The most commonly used measures are variance, standard deviation and interquartile range.

Variance and Standard Deviation

Although statistical computations will usually be carried out using a calculator or computer, we start with a few examples showing the 'background calculations' that are actually carried out. Thereafter, make use of available technology to do the number crunching. We continue with the situation described in Test 1

To calculate the variance of a set of data, the frequency table can be extended as follows:

Mark (M)	Frequency	$M-\mu$	$f(M-\mu)^2$
3	2	−7.10	100.82
4	2	−6.10	74.42
5	2	−5.10	52.02
6	4	−4.10	67.24
7	3	−3.10	28.83
8	2	−2.10	8.82
9	1	−1.10	1.21
10	8	−0.10	0.08
11	5	0.90	4.05
12	8	1.90	28.88
13	6	2.90	50.46
14	2	3.90	30.42
15	0	4.90	0.00
16	1	5.90	34.81
17	0	6.90	0.00
18	0	7.90	0.00
19	2	8.90	158.42
		Total:	640.48

Test 1:

The third column in this table measures the amount that each mark **deviates** from the mean mark of 10.10. Because some of these marks are larger than the mean and some are smaller, some of these deviations are positive and some are negative. If we try to calculate an average deviation using these results, the negative deviations will cancel out the positive deviations. To correct this problem, one method is to square the deviations. Finally, this result is multiplied by the frequency to produce the results in the fourth column.

The last row is calculated:

$$2 \times (3 - 10.10)^2 = 2 \times 50.41 = 100.82 .$$

The total of the fourth column is divided by the number of data items (48) to obtain the variance of the marks:

$$\text{Variance} = \frac{640.48}{48} = 13.34$$

The measure most commonly used is the square root of the variance (remember that we squared the deviations). This is a measure known as the standard deviation of the marks. In the previous case: Standard deviation = $\sqrt{13.34} \approx 3.65$

Repeating this calculation for the second set of marks:

Mark (M)	Frequency	$M-m$	$f(M-m)^2$
8	13	−1.48	28.475
9	11	−0.48	2.534
10	12	0.52	3.245
11	12	1.52	27.725
		Total:	61.979

$$\text{Variance} = \frac{61.979}{48} = 1.291$$

$$\text{Standard deviation} = \sqrt{1.291} = 1.136$$

This figure is about one-third of the figure calculated for Test 1. This reflects the fact that Test 2 has not spread the students very well.

In summary, the variance and population standard deviation are calculated using the formulae:

For individual observations:

$$\sigma^2 = \frac{\sum_{i=1}^{n}(x_i-\mu)^2}{n} = \frac{1}{n}\sum_{i=1}^{n}(x_i-\mu)^2$$

For grouped data:

$$\sigma^2 = \frac{\sum_{i=1}^{N}(x_i-\mu)^2}{N} = \frac{1}{N}\sum_{i=1}^{k}f_i(x_i-\mu)^2 = \frac{1}{N}\sum_{i=1}^{k}f_i x_i^2-\mu^2$$

where $N = \sum_{i=1}^{k} f_i$.

Then, the standard deviation is calculated as the square root of σ^2.

That is, the standard deviation of a population,

$$\sigma = \sqrt{\frac{1}{N}\sum_{i=1}^{k}f_i(x_i-\mu)^2}.$$

In the same way that we had a sample mean, $\bar{x}$, and a population mean, μ, when calculating 'the mean', we have a similar situation with the variance. Calculators will have two variance functions; the population variance, σ^2 and the sample variance, s^2. In this course, you will always use the population variance σ^2 (as all data sets will be considered as the population).

The relationship between the sample variance, s^2, and the population variance, σ^2 is $(n-1)s^2 = n\sigma^2$ with $\mu = \bar{x}$.

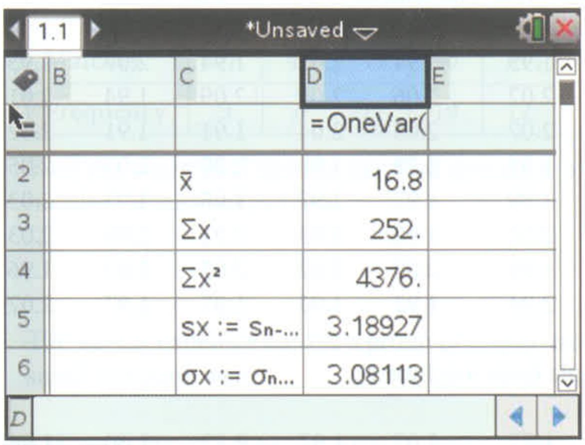
Having done some number-crunching with the illustrated example for test results, we now consider a couple more examples.

Example D.3.1

Calculate the standard deviation for the following data set.

12	15	11	17	14	16	20
22	15	21	16	17	19	20
17						

With all statistics data, we have the option of making use of a graphics calculator – which we will do in this instance. We start by entering the data as a list and then use the appropriate options to allow the calculator to number crunch,



So, in this case we have that the standard deviation is 3.08.

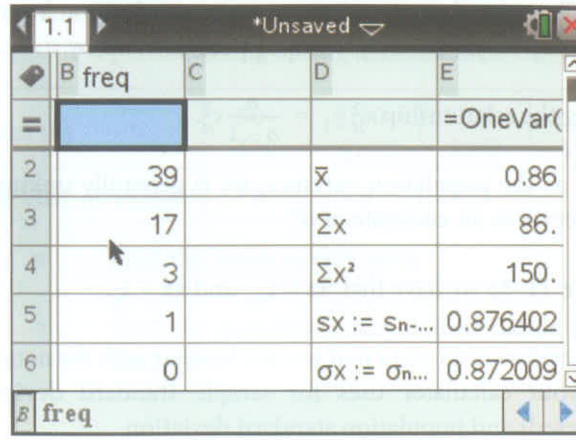
Example D.3.2

An experiment consists of rolling 5 dice 100 times and recording the number of sixes observed after they land. The result is tabulated below. Calculate the standard deviation of the number of sixes observed.

Results of rolling 5 dice

Number of sixes	0	1	2	3	4	5
Frequency	40	39	17	3	1	0

Again, we use a graphics calculator, which will enable us to quickly work out the standard deviation. We enter the observed values (number of sixes) as the first list, and the corresponding frequencies as the second list:



This gives us the standard deviation as $\sigma = 0.8720$ (the population standard deviation).

Sx is known as the **sample standard deviation**. This is the same as the standard deviation discussed above but with one less than the number of data items in the denominator (47 in this case).

σx is the **population standard deviation** discussed above.

Sample standard deviation? Population standard deviation? What's it all about? Unfortunately there are regional variations (as well as in textbooks) in the notation and the language that is used to define these terms.

When we refer to the **sample variance**, it suggests that we are finding the variance of a sample and, by default, the sample is a subset of a population and so we are in fact finding an estimate of the population variance. This estimate is known as the **unbiased estimate** of the population variance.

The **unbiased estimate** of the population variance, σ^2, is given by:

$$s_{n-1}^2 = \frac{1}{n-1}\sum_{i=1}^{k} f_i(x_i - \bar{x})^2.$$

The standard deviation of the sample is given by the square root of s_{n-1}^2, i.e. $\sqrt{s_{n-1}^2}$, which

corresponds to the value **Sx** that is produced by the TI-83.

The **variance of a population**, σ^2, is given by:

$$\sigma^2 = \frac{1}{n}\sum_{i=1}^{k} f_i(x_i - \mu)^2$$

The standard deviation then is $\sigma = \sqrt{\sigma^2}$.

To differentiate between division by n and division by $n - 1$ we use s_n^2 for division by n and s_{n-1}^2 for division by $n - 1$.

Giving the relationship $s_{n-1}^2 = \frac{n}{n-1}s_n^2$.

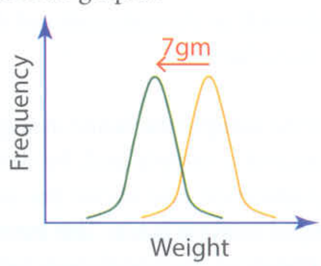
Then, as the population variance, σ^2, is generally unknown, s_{n-1}^2 serves as an estimate of σ^2.

On the TI–83 we have that **Sx** $= s_{n-1}$ and **sx** $= s_n$.

It is therefore important that you are familiar with the notation that your calculator uses for sample standard deviation (unbiased) and population standard deviation.

Transformations

If we have calculated statistics on a set of data and that data is transformed, it may not be necessary to recalculate all the results. For example, if 7 is subtracted from a set of data items, then the mean is decreased by 7, but the standard deviation is unchanged. Thus, if we have a set of weights and we have calculated the mean and standard deviation only to discover that the weighing machine is weighing things 7gms too much, the mean weight needs to be reduced by 7. Measures of spread remain unchanged. This is similar to results used in the transformations of graphs.



If all the data items are doubled, the mean is doubled, but the variance is increased by a factor of 4 and the standard deviation is doubled.

Example D.3.3

The mean daily high temperature for the month of March in Buenos Aires is 26.8°C with a standard deviation of 2.3°C.

What are these statistics in °Fahrenheit?

The conversion between Celsius (C) and Fahrenheit (F) is:

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

The mean temperature needs to be converted to Fahrenheit:

$$F = \frac{9}{5} \times 26.8 + 32 \approx 80\ ^\circ\text{F}$$

The standard deviation is multiplied by $\frac{9}{5}$ to get 4.14°F.

Exercise D.3.1

1. The weights (kg) of two samples of bagged sugar taken from a production line.

Sample from machine A:

1.95	1.94	2.02	1.94	2.07	1.95
2.02	2.06	2.09	2.09	1.94	2.01
2.07	2.05	2.04	1.91	1.91	2.02
1.92	1.99	1.98	2.09	2.05	2.05
1.99	1.97	1.97	1.95	1.93	2.03
2.02	1.90	1.93	1.91	2.00	2.03
1.94	2.00	2.02	2.02	2.03	1.96
2.04	1.92	1.95	1.97	1.97	2.07

Sample from machine B:

1.77	2.07	1.97	2.22	1.60	1.96
1.95	2.23	1.79	1.98	2.07	2.32
1.66	1.96	2.05	2.32	1.80	1.96
2.06	1.80	1.93	1.91	1.93	2.25
1.63	1.97	2.08	2.32	1.94	1.93
1.94	2.22	1.76	2.06	1.91	2.39
1.98	2.06	2.02	2.23	1.75	1.95
1.96	1.80	1.95	2.09	2.08	2.29

a Find the mean weights of the bags in each sample.

b Use the formula: $S_x = \sqrt{\frac{\sum f_i(x_i - \bar{x})^2}{\sum f_i - 1}}$

to calculate the sample standard deviations of each sample.

c Use the formula: $\sigma_x = \sqrt{\frac{\sum f_i(x_i - \mu)^2}{\sum f_i}}$

to calculate the population standard deviations of each sample.

2. The following frequency table gives the numbers of passengers using a bus service over a week-long period.

Passengers	0–4	5–9	10–14	15–19	20–24	25–29
Frequency	3	5	11	15	10	7

a Find the mean number of passengers carried per trip.

b Find the population standard deviation of the number of passengers carried per trip.

3. The number of matches per box in a sample of boxes taken from a packing machine was:

Matches	47	48	49	50	51	52
Frequency	3	6	11	19	12	9

Find the mean and sample standard deviation of the number of matches per box.

4. The weekly expenses paid to a group of employees of a small company were

\$25	\$0	\$10	\$10	\$55	\$0
\$12	\$375	\$75	\$445	\$7	\$2

a Find the mean weekly expense.

b Find the population standard deviation of the expenses.

5. The table shows the numbers of cars per week sold by a dealership over a year.

Cars sold	0	1	2	3	4	5
Number of weeks	2	13	15	12	7	3

a Find the mean weekly sales.

b Find the population standard deviation of the sales.

6. The table shows the weekly turnover of a small shop over a period during Spring and Summer.

Sales (\$)	\$0–\$99	\$100–\$199	\$200–\$299	\$300–\$399
Number of weeks	2	9	15	7

a Find the mean weekly sales.

b Find the population standard deviation of the sales.

7. The frequency distribution of Mathematics and English test results at a local secondary school are shown in the table below:

Test Scores

Mark	Mathematics	English
[0, 10[	7	2
[10, 20[	11	5
[20, 30[	13	11
[30, 40[	17	18
[40, 50[	22	34
[50, 60[	22	34
[60, 70[	16	20
[70, 80[	14	12
[80, 90[	11	2
[90, 100]	7	2

a Draw a histogram showing the test scores for:
i Mathematics ii English.

b i Draw a table of the cumulative frequencies for each of Mathematics and English.

ii Draw the cumulative frequency graphs for Mathematics and English scores.

c The pass mark for Mathematics is the lowest score obtained by 78% of the students. What is the minimum score required for a student to pass Mathematics?

8. Doctors are under pressure to diagnose as many patients as possible, meaning that the session time they allocate to each patient is closely monitored. The table below shows the times that two doctors have spent with their patients.

Session time distribution

Time (minutes)	Number of sessions	
	Doctor A	Doctor B
5–9	5	3
10–14	10	7
15–19	23	x
20–24	16	27
25–29	12	9
30–34	5	6
35–39	3	2

a For doctor A, calculate the:

i mean session time.

ii standard deviation of session times.

b Doctor B has misplaced the tally for the number of times she has seen patients for a period of 15–19 minutes. She decides that she should have the same mean as her colleague for the time spent seeing patients. What value of x should she use (to the nearest integer)?

c A third doctor at the clinic recorded an average of 22 minutes after seeing 60 patients. What is the overall mean time these three doctors spend with patients?

9. The mean weight of a sample of chocolate boxes is 502.7gm with a standard deviation of 1.5gm.

The USA version of the product is sold in pounds (lbs). There are 2.20 lbs to the kilo. What quality statistics should the US distributors expect?

10. In a cycling trial, the competitors average 56 kph with a standard deviation of 2 kph. What are these figures in miles per hour? A kilometre is about $^5/_8$th of a mile.

11. A weighing scale reads 4% too heavy. A set of weights recorded using the scale has a median of 25.6kg and a range of 3.6kg. Estimate the true median and range.

Answers

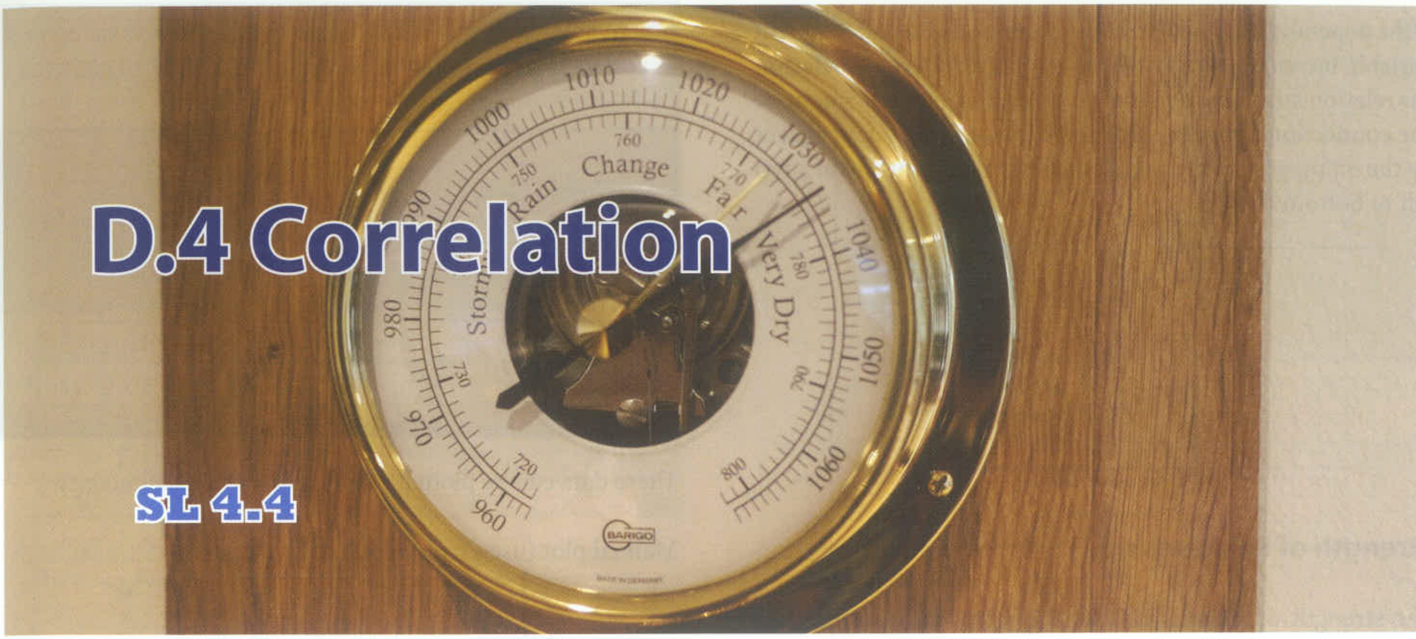
It can be an interesting exercise to compare the various measures of spread to see how well each performs in different circumstances.

As a starting point, look at a data set that is highly skewed and has outliers (such as pay scales).

Inter-quartile range ignores the effect of outliers whereas variance does not.

In what circumstances is this sort of evasion desirable when presenting data.

When considering this, it might help to think from the point of view of a politician writing a speech, a union official preparing for a wage negotiation etc.



D.4 Correlation

SL 4.4

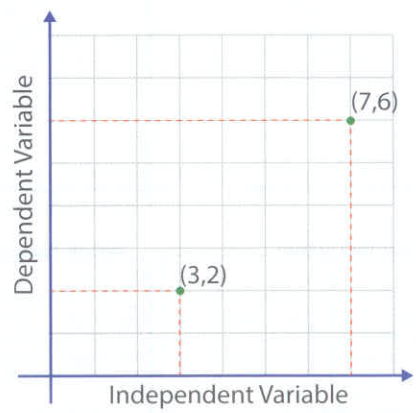
Sailors have long believed that there is a connection between the pressure of the atmosphere (which varies) and the weather that they can expect (rainfall, wind speed etc.). Most ships are still equipped with barometers (see photo) which are used in addition to radio weather forecasts. These technological services are, of course comparatively recent developments. Barometric pressure reports from ships to land based forecasters also contribute to our daily weather reports.

This sort of connection between quantities that are seemingly unrelated forms the subject of this chapter.

Such information is known as **bivariate data**. In particular we will be looking at the use of **scatter diagrams** as an initial visual aid in describing any relationship that may exist between two variables. Next we look at measuring the strength of such a relationship (if one exists) and finally we consider the issue of **regression analysis.** This will help in obtaining equations that can be used to predict (or explain) the value of one variable (the **dependent variable**) based on the value of a second observation (the **independent variable**). In particular, we will only be considering linear relationships, as such, the regression lines will take on the form $y = a + bx$.

Scatter Diagrams

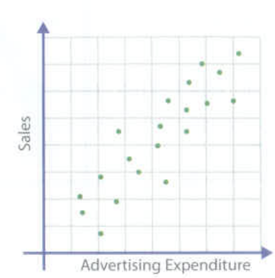
A scatter diagram (or scattergraph) is a method by which we can obtain a very quick visual appreciation of how two variables are related. Such diagrams are obtained by plotting a set of points that correspond to the bivariate data. Usually the independent variable runs along the horizontal axis, while the dependent variable runs along the vertical axis. Once the data has been plotted we are interested in finding some indication of the association between the two variables. One such measure is the correlation. Qualitative descriptors that are useful include: direction, form and strength of relationship.



The diagram shows a simple scatter diagram with two points. The one at lower left shows a data point in which the independent variable takes a value of 3 and the dependent variable has value 2.

Direction of relationship

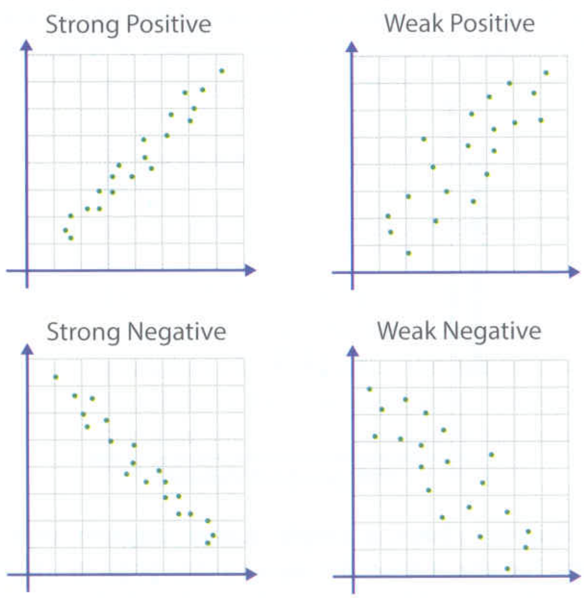
If the dependent variable tends to increase as the independent variable increases, we say that there is a **positive association** (or relationship) between the variables. Our example shows the possible relationship between advertising expenditure and sales for a small business. The points lie in a bottom left to top right band.



If the dependent variable tends to decrease as the independent variable increases, we say that there is a **negative association** (or relationship) between the variables. An example might be the connection between sales and the number of days absent by the employees at the same business. The points lie in a top left to bottom right band.

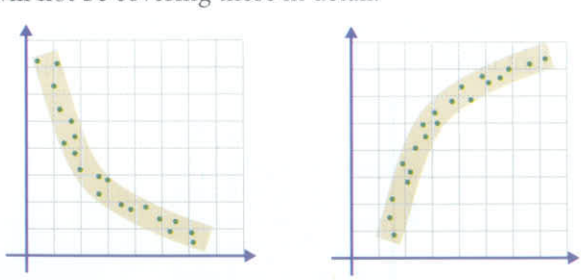
Strength of Relationship

The strength of a linear relationship gives an indication of how closely the points in the scatter diagram fit a straight line.



There is a continuum between very positive through the weak positive to the weak negative to the strong negative. In the middle there is the random scattering pattern which shows no correlation at all.

There are also cases where a non-linear correlation occurs but we will not be covering these in detail.



Calculators usually work by requiring the data to be entered as paired lists. Consider this example.

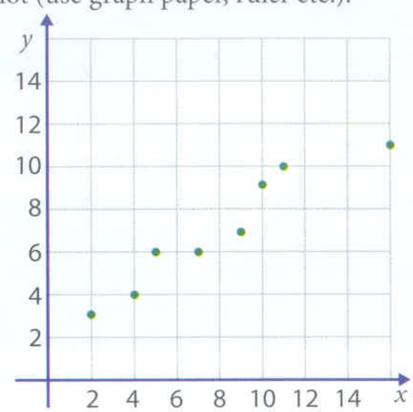
Example D.4.1

Determine if these data have a linear relationship, stating the direction and strength.

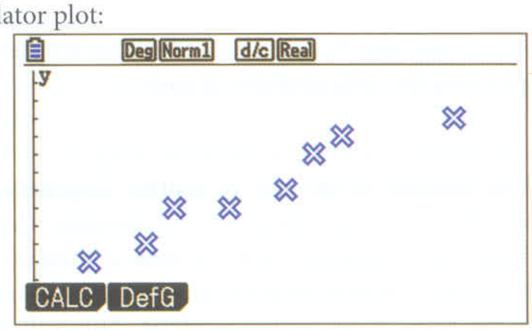
x	2	4	5	7	9	10	11	15
y	3	4	6	6	7	9	10	11

These data can be plotted manually or using technology.

Manual plot (use graph paper, ruler etc.):



Calculator plot:



The calculator plot is described on this short video:

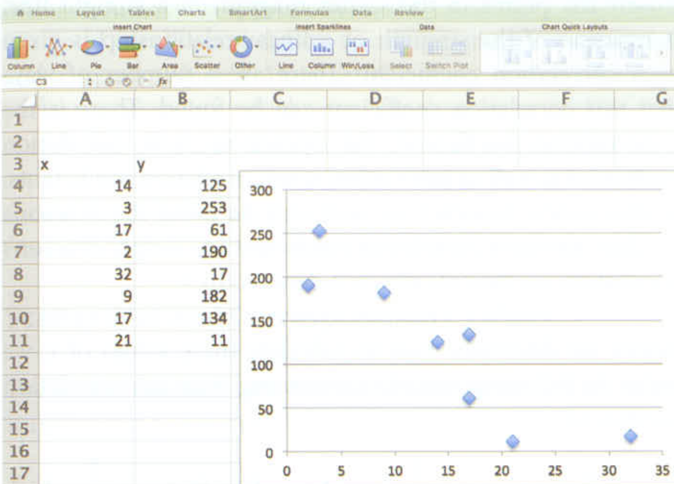
The diagram shows a strong positive correlation. This is because the points lie in a fairly narrow upward sloping band.

Example D.4.2

Determine if the data has a linear relationship, stating the direction and strength.

x	14	3	17	2	32	9	17	21
y	125	253	61	190	17	182	134	11

To use a spreadsheet, the data needs to be entered (cells A5 - B11). It is then selected and then Chart/Scatter is selected.



The Excel plot is described on this short video:

The diagram shows a strong(ish) negative correlation.

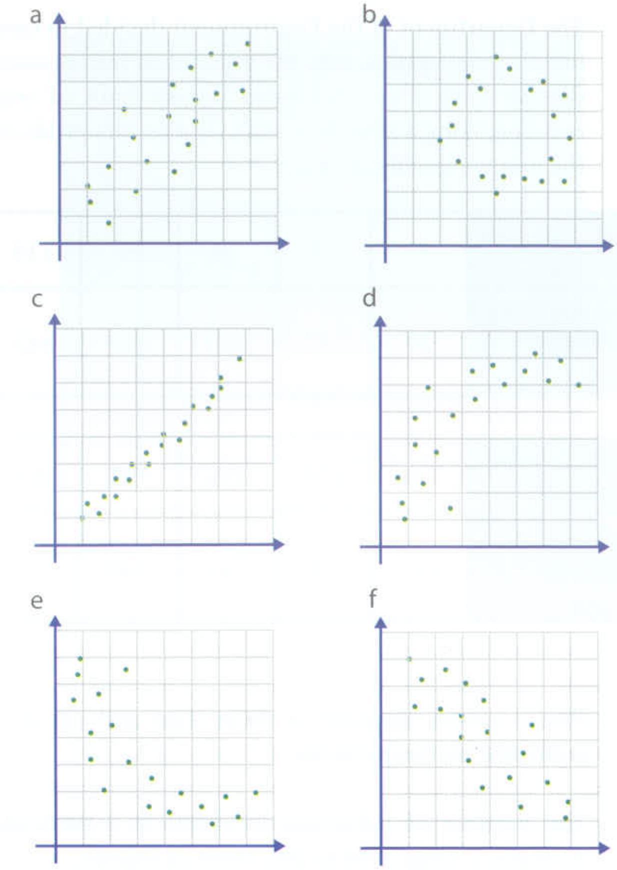
Exercise D.4.1

1. For each of the following, give a statement about:

   i the direction

   ii the form

   iii the strength of the relationship.



2. A group of students had their Science and Maths results tabulated.

Student	1	2	3	4	5	6	7	8	9	10
Science	55	70	40	67	80	80	55	60	20	84
Maths	60	78	39	65	82	90	50	71	18	79

   a Plot these data on a scatter diagram.

   b Describe the direction, form and strength of the relationship between Science marks and Maths marks.

3. The data in the table below shows students' reading test scores and their corresponding I.Q. scores

Student	1	2	3	4	5	6	7	8	9	10
Reading score	50	73	74	62	70	57	60	62	70	65
I.Q. scores	99	118	131	111	113	101	106	113	121	118

   a Plot these data on a scatter diagram.

   b Describe the direction, form and strength of the relationship between reading scores and IQ scores.

4. The Department of the Environment decided to carry out an investigation into the amount of lead content, due to traffic flow, deposited on the bark of trees running along a stretch of road. The results produced the following table of values.

Traffic flow (in thousands)	32	35	70	73	119
Lead content (mg/g dry weight)	29	110	164	349	442

Traffic flow (in thousands)	121	125	194	193	204
Lead content (mg/g dry weight)	337	530	743	540	557

Plot a scatter diagram of these data and use it to comment on these results.

5. The number of industrial accidents in a particular workplace, from 1994 to 2003 were as follows.

Year	1994	1995	1996	1997	1998
Number of accidents	166	131	123	162	160

Year	1999	2000	2001	2002	2003
Number of accidents	130	91	82	65	53

a Plot a scatter diagram and use it to comment on the data.

b How would you rate the work safety policy that the company implemented since 1994?

Strength of a Linear Relationship

So far we have given qualitative measures of the strength of a linear relationship, i.e. we have used expressions such as strong and weak, etc. However, as in all aspects of good statistical analysis, it is important that we provide a quantitative measure to describe our observations. Such measures are crucial when comparing sets of data.

The strength of a linear relationship is an indication of how closely the points in the scatter diagram fit a straight line. A measure of the strength of a linear relationship is given by a **correlation coefficient**. There are a number of ways that this correlation coefficient can be found. There is the q–correlation coefficient, the Spearman rank correlation coefficient as well as some other rarer correlation coefficients. In this course however, we will only be using the Pearson's product-moment correlation coefficient (or simply Pearson's correlation coefficient) which is denoted by r.

Before we determine how to calculate these values of r, we make the following remarks concerning Pearson's correlation coefficient.

1. The value of r does not depend on the units or which variable is chosen as x or y.

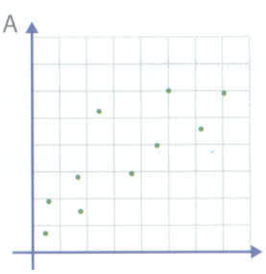
2. The value of r always lies in the range $-1 \le r \le 1$. A positive r indicates a positive association between the variables while a negative r indicates a negative association.

3. Perfect linear association when scatter plot points lie on a straight line, occurs if $r = \pm 1$.

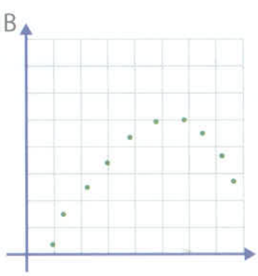
4. r measures only the strength of a **linear** association between two variables.

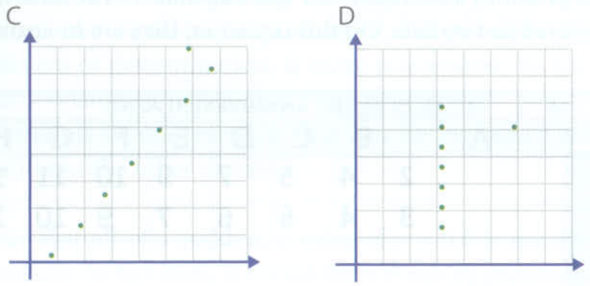
The last point is particularly important – that is, it is only of use when there is strong evidence that a linear relationship does indeed exist. A famous example was provided by Anscombe in 1973 where four radically different scatter plots were produced with a contrived value of $r = 0.82$ in each case. These are shown below:

A



B

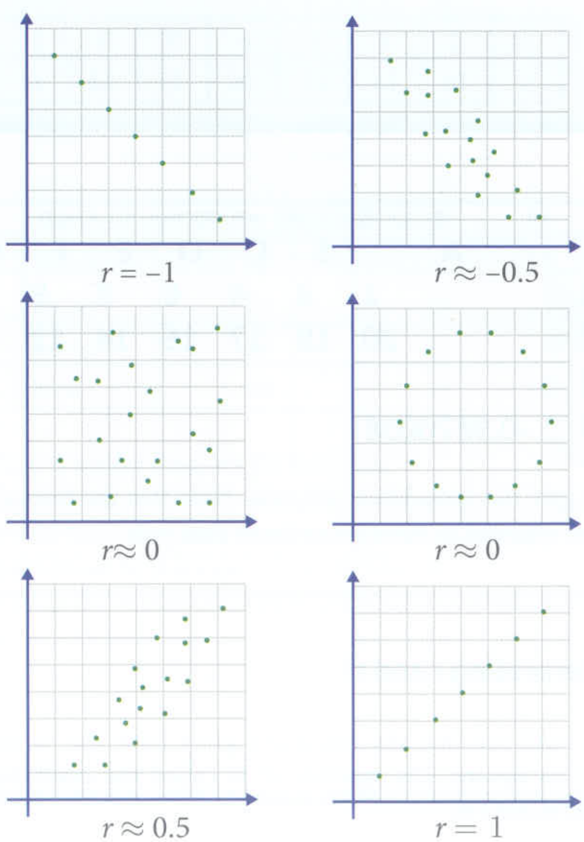




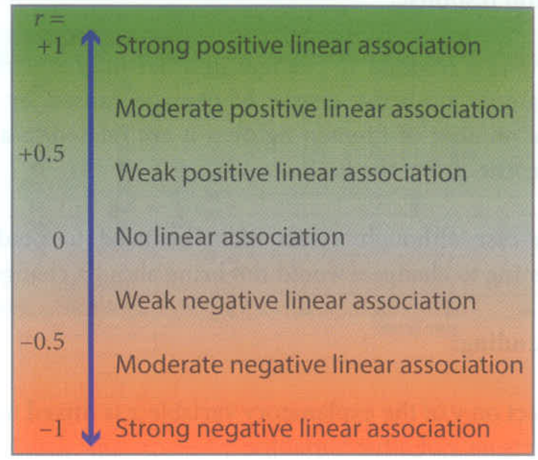
Data set A is a reasonably standard bivariate set for which the linear model seems appropriate and r meaningful. Data set B is a good curvilinear relation so a linear model is unsuitable and r is not valuable. Data set C has an outlier ruining a perfect relationship which may signal an error in data. The correlation probably understates the true relationship. Data set D is anomalous in the extreme – all points are the same except one and this exerts a large influence on the value of r. No meaningful conclusion could be drawn from this set.

Scatter Plot and Corresponding r Values

The properties of r and corresponding scatter plots can be summarized as follows.



The following diagram provides an indication of the qualitative description of the strength of the linear relationship and the quantitative value of r.



Cause and Association

It can be risky to confuse a relationship with a cause. Just because two variables are highly correlated, it does not mean that one necessarily causes the other. For example, the degree of fatigue you experience during summer may be influenced by the temperature of the day (i.e. your fatigue depends on the temperature level), however, if you happen to reach a certain level of fatigue on some other day, this will not in turn indicate the temperature level of the day!

That is, the temperature is independent of your fatigue level, and so, a rise in your fatigue level will not cause the temperature to rise.

Some 'relationships' suggested by correlation are spurious. For example, we have:

- Damage caused by fire, and the number of firemen fighting it.
- Weight and height in individuals.
- Smoking rates and lung cancer deaths.

Some of these may represent causal relationships, others may seem ridiculous, but too often people jump to unjustified conclusions on the basis of high correlations alone.

Two variables x and y may exhibit a strong link for a number of reasons. These include:

Causation:

Changes in x cause changes in y – for example, a change in outdoor temperature causes change in ice cream consumption. In cases where we have control over one variable, if we can change x, we can bring about a change in y. If smoking causes lung cancer, then reducing the prevalence of smoking should reduce the incidence of lung cancer.

Common response:

Both x and y respond to changes in other hidden variables. For example, both the degree of damage caused by a fire and the number of firemen fighting it are related to a third variable, the size of the fire!

In this case although x can often be used to predict y, intervening to change x would not bring about a change in y.

Confounding:

The effect on y of the explanatory variable x is mixed up with the effects on y of other variables.

When experiments are not possible, good evidence for causation is less direct and requires a combination of several factors – where each of the above is adequately addressed.

Determining the Value of *r* from a Data Set

This is best done by calculator or spreadsheet.

We concluded the solution to Example D.4.1 with a screen showing the scatter diagram. EXIT twice to the data lists, F2-CALC and then press F3-REG (for regression). Choose F1-X for a simple linear model. There are two types of model. Pick the F1 option to get this screen:

```
Deg Norm1 d/c Real
LinearReg(ax+b)
    a =0.64864864
    b =1.89189189
    r =0.9686196
    r²=0.93822393
   MSe=0.57657657
y=ax+b
                          COPY
```

The information is that the r value is ≈ 0.9686...

The other data will be covered in subsequent sections. At this stage, we observe that the r value is close to +1 indicating a strong positive correlation.

The Casio calculator procedure is described on this short video:

Excel provides a function that will calculate r. The data must be entered as two lists. On this occasion, they are horizontal.

A4 fx =PEARSON(A1:H1,A2:H2)

	A	B	C	D	E	F	G	H
1	2	4	5	7	9	10	11	15
2	3	4	6	6	7	9	10	11
3								
4	0.9686196							
5								

Note that the syntax of the function is a bit complex. The range of cells for the first variable (blue) must be entered and separated from the second set (green) with a comma.

=PEARSON(A1:H1,A2:H2)

The result is as produced by the calculator.

Example D.4.3

Find the Pearson correlation coefficient for the following set of data.

x	2	3	4	6	8	9	10
y	20	18	17	16	14	12	11

A4 fx =PEARSON(A1:G1,A2:G2)

	A	B	C	D	E	F	G
1	2	3	4	6	8	9	10
2	20	18	17	16	14	12	11
3							
4	-0.9870038						

We give the spreadsheet solution. Note that the formula has been adjusted as there is one fewer data pair.

Interpreting *r* and r^2

You will see that the calculator gives both the value of r and its square. The quantity r^2 is called the **coefficient of determination**.

This quantity gives us an approximate value of the extent to which there is a real connection between the two variables.

In the case of Example D.4.3 $r \approx -0.987$ and $r^2 \approx 0.974$. This can be interpreted as meaning that 97% of the connection between x and y is 'real'. The rest is random.

In general, when interpreting the magnitude of the relation between two variables, regardless of directionality, r^2, the coefficient of determination, is more informative. So for two linearly related variables, this value provides the proportion of variation in one variable that can be explained by the variation in the other variable.

Notice that all of a sudden, a value of $r = 0.6$ is not all that impressive! Why? Well, if $r = 0.6$ then $r^2 = 0.36$, meaning that only 36% of the variation in one variable is explained by the variation in the other variable.

One way of visualizing the meaning of the coefficient of determination, r^2, is to consider a perfect positive correlation and then observe what happens for a small value of r.

Example D.4.4

As part of a Mathematics course, a teacher is required to submit marks to the IBO for the internal assessment part of the course. The internal marks and the final awarded marks for 10 students in a particular class were as follows:

Student	Internal mark	Final mark
Peter	65	72
Mark	80	87
John	78	65
Lois	60	73
Jane	92	95
Tom	81	90
Fiona	50	45
Becky	55	65
Louie	50	51
Sam	90	93

a Draw a scatter diagram for this data.

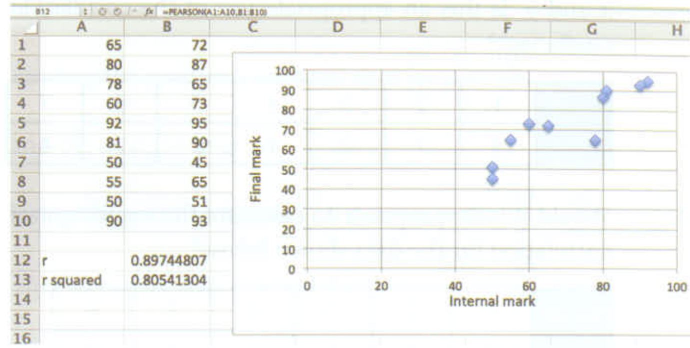
b Would the use of the Pearson correlation coefficient be appropriate in this case? If so, calculate it.

c What conclusions can you make about the relationship between internally assessed marks and the final marks awarded to students?

We will use a spreadsheet on this occasion.

The data will first be entered as columns.

a



b From the spreadsheet, we have a strong positive correlation between the two amounts.

c There exists a positive linear relationship between internal marks and the final mark awarded to the students in this class. As the value of $r = 0.8974$ we can say that there is a strong positive correlation between the internal mark awarded and the final mark awarded.

Also, the value of $r^2 = 0.8054$, is telling us that 80.54% of the variation in the final mark awarded can be accounted for by the variation in the internal mark awarded.

That is, only 20% (approximately) of the variation is attributed to other factors. The internal assessment appears to be in substantial agreement with the final assessment.

Exercise D.4.2

1. a Assuming that the data has a linear relationship, find the coefficient of correlation for this set.

x	4	6	7	9	11	12	13	17
y	8	9	11	11	12	14	15	16

b Draw a scatter diagram for the given data.

2. a Draw a scatter diagram for the given data.

x	1	5	6	6	2	3	4	4
y	2	4	5	3	1	2	5	4

b Find the coefficient of correlation for this set of data. What assumption have you made in determining this value?

3. For the set of paired data, find the correlation between x and y. Is this an appropriate use of the correlation coefficient?

x	1	2	3	4	5	6	7
y	4	3	2	1	2	3	4

4. Would it be appropriate to calculate the coefficient of correlation for the data shown below?

x	1	2	3	4	5	6
y	3	2	1	1	2	3

5. Calculate the proportion of the variance of Y which:

a can be predicted from (or explained by) the variance of X if:

i $r = 0.8$. ii $r = -0.9$

b cannot be predicted from (or explained by) the variance of X if:

i $r = 0.7$ ii $r = -0.6$.

6. The data below represents entrance examination marks (x) and first-year average test marks (y) for a group of ten students.

x	55	59	62	80	92
y	61	69	52	61	90

x	63	69	84	62	55
y	85	70	67	72	60

a Draw a scatter diagram for the data.

b Determine the correlation coefficient between x and y.

7. How many times is the difference in predictive capacity between correlations of 0.70 and 0.80 greater than between correlations of 0.20 and 0.30?

8. What correlation between X and Y is required in order to assert that 85% of the variance of X depends on the variance of Y?

9. For the data below, calculate the proportion of the variance of y which can be explained by the variance of x

x	3	4	6	7	9	12
y	20	14	12	10	9	7

Line of Best Fit

Having established that a linear relation exists between two variables x and y (for example), we can then search for a line of best fit. That is, a line that will best represent the data on the scatter diagram. There are a number of ways this can be done. Some possibilities are:

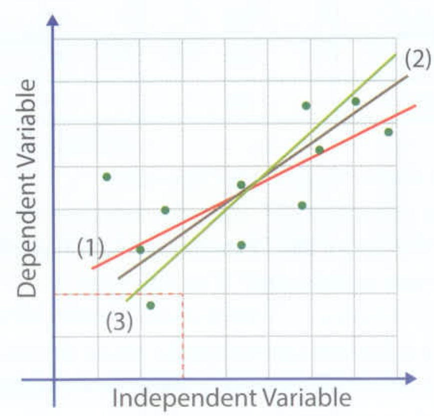
1. Drawing a line 'by eye'.
2. Using the locus of means.
3. Using the median-median line of best fit.
4. Using the least squares regression equation found using technology.

We will consider options 1 and 4.

Drawing a line 'by eye'

If the scatter plot signals that a linear model is reasonable, we may attempt to model the data using a line of best fit. The choice of a suitable linear model may be done informally or by other formal methods which are usually easier to produce by hand. The problem with drawing a line 'by eye' is that many lines may seem equally suitable.

Consider the scatter plot shown below. Which of the lines, (1), (2) or (3) seems best and are there better lines we could fit?



If you are using nothing more than your eyes to decide on the best line, using a stretched length of string can be better than a ruler. This is because you can see both sides of the string at once whereas a ruler tends to obscure half the points.

Once a line is decided on, we can then use the methods of Chapter B1 to determine the equation of this line. When we have an equation for the linear model, we can use it for predicting the values of y we may expect for a given value of x.

However, this method relies too much on individual preferences. So, to help in this endeavour, we have two methods:

1. Balancing the number of points above and below the line.

2. Calculating the mean value of x, and y and plotting this mean point. This can be used as a pivot and only leaves you to decide the best gradient.

3. Balancing the errors from the line to the data points.

The third method uses a method that balances the error on either side of the straight line. However, this method is more suitable when the data points are scattered and fitting the straight line is not as obvious. Errors are based on the vertical distances between the data points and the fitted straight line, with those above the straight line being positive while those below the straight line being negative.

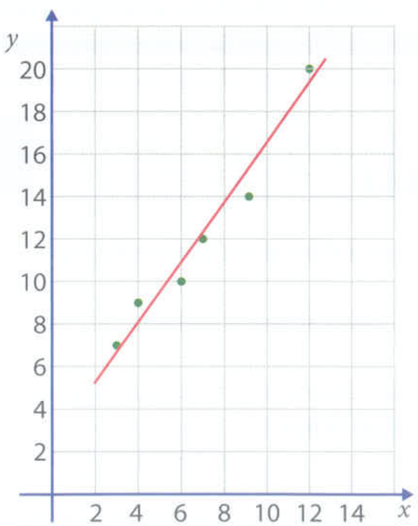
In this instance we will only consider method 1.

Example D.4.5

For the data shown below, find the equation of the line of best fit by eye.

x	3	4	6	7	9	12
y	7	9	10	12	14	20

We start by drawing a scatter plot of the data:



Having fitted a line so that there are three points 'above' the line and three points 'below' the line, we proceed to determine the equation of the straight line.

The line passes through the points with coordinates (4, 8) and (11, 18).

Finding the gradient, we have $m=\frac{18-8}{11-4}=\frac{10}{7}$

and the straight line: $y=\frac{10}{7}x+c$.

Using the point (4,8): $8=\frac{10}{7}\times 4+c \Rightarrow c=8-\frac{40}{7}=\frac{16}{7}$

This 'line of best fit' has equation: $y=\frac{10}{7}x+\frac{16}{7}$ [Eq. 1].

We can now compare the original data with the results based on this equation.

x	3	4	6	7	9	12
y	7	9	10	12	14	20
Eq. 1	6.571	8.000	10.857	12.286	15.143	19.429
Error	0.429	1.000	–0.857	–0.286	–1.143	0.571

The results seem quite good (errors quite small with some positive and others negative.). There is an element of confidence then in using this equation to predict the y-value when, for example, $x = 8$.

The errors in the bottom row give a measure of the 'goodness of fit' of the line that we drew to the actual data. We cannot, however, just add these errors as some are positive and others negative. Adding will result in cancellation and a useless result. The usual way around this is to square the errors and try to minimise the result:

x	3	4	6	7	9	12
y	7	9	10	12	14	20
Eq. 1	6.571	8.000	10.857	12.286	15.143	19.429
Error	0.429	1.000	–0.857	–0.286	–1.143	0.571
Error^2	0.184	1.000	0.735	0.082	1.306	0.327

The total Error^2 = 3.633

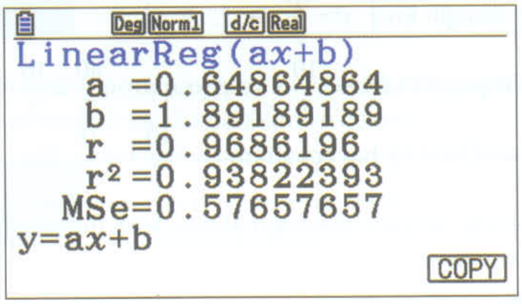
The calculations needed to minimise this error are complex and best left to technology. From here on, we will use calculators and Excel to 'grind the numbers'.

Example D.4.6

Assuming that the data has a linear relationship, find the regression line of y on x.

x	2	4	5	7	9	10	11	15
y	3	4	6	6	7	9	10	11

To use a calculator, first enter the data as two lists. This was described earlier. The choose REG (regression), X (linear model). The select the $ax+b$ model to get:



This suggests that the best model is $y = 0.6486x + 1.8919$

The value of r^2 suggests that 94% of the variability in y is attributable to x and 6% to random factors.

This video shows the same calculation using Excel.

Example D.4.7

The data shown represents observations made on the rate of cricket sounds by the number y chirps per 15 seconds at different temperatures x in degrees Fahrenheit.

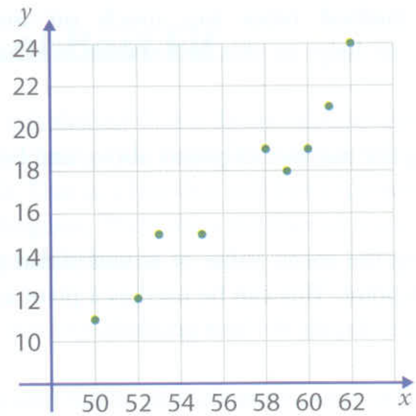
x	62	61	60	59	58	55	53	52	50
y	24	21	19	18	19	15	15	12	11

a Plot these points on a scatter diagram.

b Determine the line of regression of y on x and draw it on your scatter diagram.

c Estimate the cricket's rate when the temperature is 65 degrees Fahrenheit.

a



b Using Excel, $y = 0.9482x - 36.62$ with $r^2 \approx 0.93$

c This means that there is a strong positive correlation between chirp rate and temperature. For every degree (F) rise in temperature, the chirp rate increases by 1 per 15 seconds.

If the temperature is 65°F, the predicted chirp rate is:

$y = 0.9482 \times 65 - 36.62 \approx 25$ chirps per 15 seconds.

Because this value lies outside our data set, we are **extrapolating** and so we are relying on the assumption that the linear trend based on our regression equation will continue.

Extrapolation is risky. For a temperature of 33°F (just above freezing) the predicted chirp rate is negative. Whilst it would almost certainly be zero, a negative result makes no sense.

Exercise D.4.3

1. For each of the following sets of data:

   i draw a scatter diagram.

   ii make use of the 'by eye' approach to find the line of best fit.

   a

x	3	4	6	7	9	12
y	20	14	12	10	9	7

b

x	15	13	17	14	18	12	20	16	18	17	19
y	18	16	18	15	19	16	18	15	21	17	18

2. For the sets of data shown below:

   i draw a scatter diagram.

   ii determine the least squares regression line.

   iii draw the regression line on your scatter diagram.

a

x	3	4	6	7	9	12
y	20	14	12	10	9	7

b

x	2	1	4	5	3
y	4	2	6	5	3

c

x	11	5	4	5	2	3
y	52	31	30	34	20	25

d

x	1	2	3	4	5
y	2	1	3	5	4

3. The following table shows the income (in thousands of dollars) and the annual expenditure, in hundreds of dollars for ten single working people aged 20–24 years.

Income	22	14	16	18	20	19	16	18	19	18
Expenditure	75	59	67	69	75	73	62	64	70	71

   a Plot the data on a scatter diagram.

   b Find the correlation coefficient.

   c Calculate the proportion of the variance of Expenditure which can be explained by the variance of the Income.

   d Find the least squares equation of the regression line.

   e On the scatter diagram from part a, sketch the regression line.

4. The result of the first two tests given to a group of Mathematics students is shown in the table below.

Test 1 (x)	60	50	80	80	70
Test 2 (y)	80	70	70	100	50

Test 1 (x)	60	100	40	90	70
Test 2 (y)	80	100	60	80	60

   a Draw a scatter diagram for these data.

   b Find the coefficient of correlation.

   c Find the least squares regression line of:

      i y on x ii x on y.

5. A cafe owner wishes to improve the efficiency of her cafe. One aspect that needs to be looked into is that of the rate at which customers are being served by the staff.

   The table below shows the number of weeks that eight employees have been working at the cafe and the average number of customers that each served per hour.

Weeks at cafe	8	5	15	3
Customers served	18.4	12.2	32.3	10

Weeks at cafe	10	2	13	6
Customers served	21.0	8.2	28.1	16.5

   a Draw a scatter diagram for the given set of data. Define the variable C to represent the average number of customers an employee served per hour and the variable w to represent the number of weeks that employee has been working at the cafe.

   b The owner decides to use a straight line to model the data. Is the owner justified? Give a reason for your answer.

   c i Calculate the correlation coefficient for the given data set.

   ii Use the method of least squares to determine the line of best fit.

   iii Graph the regression line on the scatter diagram in part a.

d Estimate how many customers employees should be able to serve in one hour if they have been working at the cafe for:

i 9 weeks. ii 50 weeks.

iii What constraints can you see this model having?

6. The table below shows the results of measurements taken for systolic blood pressures (y) of 8 women and their respective ages (x).

Age (x)	60	42	68	72
Blood pressure (y)	155	140	152	160

Age (x)	42	36	55	49
Blood pressure (y)	125	118	155	145

a Draw a scatter diagram for the given set of data.

b Calculate the correlation coefficient for the given data set. Is this an appropriate statistic to calculate for this data set?

c i Use the method of least squares to determine the line of best fit.

ii Graph the regression line on the scatter diagram in part a.

d Based on your line of best fit, determine the level of systolic blood pressure for a woman aged:

i 45 ii 85

iii What is the difference in using the line of best fit when answering parts i and ii?

7. The yield, y kilograms, of a vegetable, obtained by using x kilograms of a new fertilizer, produced these results.

x	1.4	3.3	5.9	8.8	7.3	5.1
y	5.0	7.5	7.7	10	9	8.3

a Draw a scatter diagram for the given set of data

b Calculate the correlation coefficient for the given data set. Is this an appropriate statistic to calculate for this data?

c i Use the method of least squares to determine the line of best fit.

ii Graph the regression line on the scatter diagram in part a.

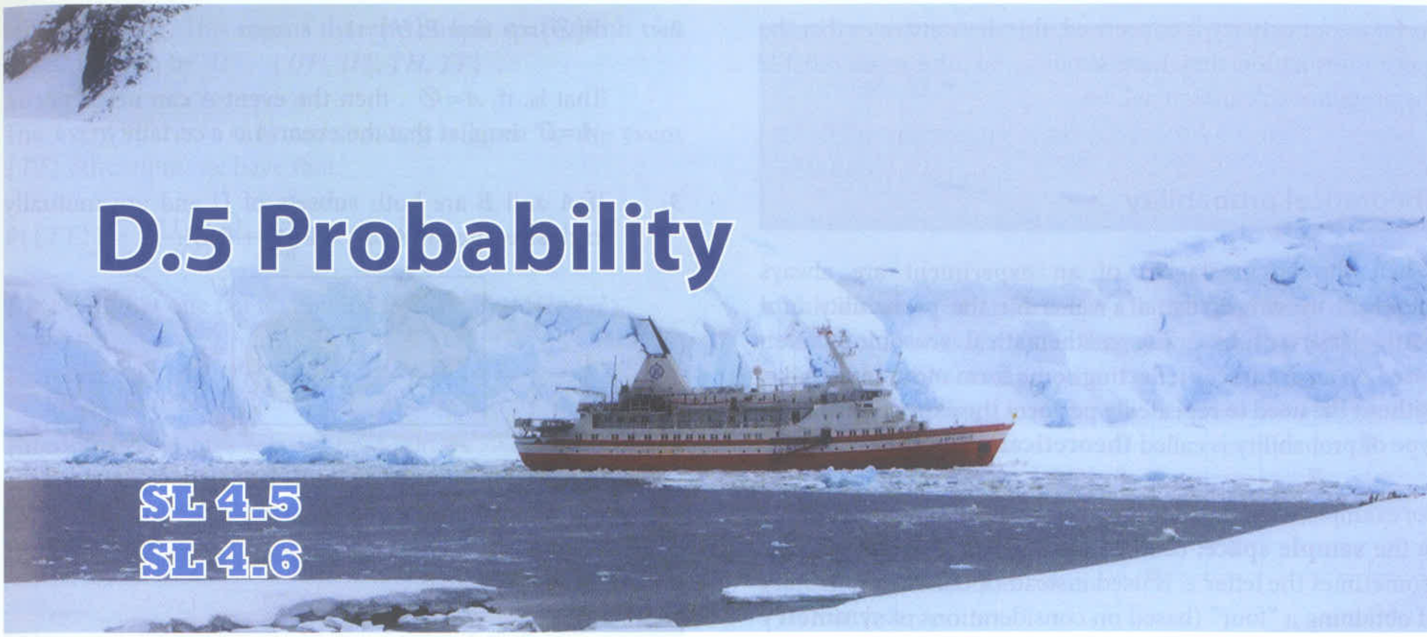
We began this chapter with the sailors' forecaster, the barometer. Does it work?

Detailed weather data are available online. For example Weather Underground will enable you to search for data from a wide range of locations. As we have observed elsewhere investigations tend to work best if you pick a problem relevant to you. So, for example, choose your own location or a place where you may be planning a holiday.

Remember to; define the problem; collect appropriate data analyse the data using appropriate mathematical techniques form appropriate conclusions; evaluate these conclusions.

Weather Underground:

Answers



D.5 Probability

SL 4.5
SL 4.6

Probability

We are often faced with statements that reflect an element of likelihood, For example, "It is likely to rain later in the day" or "What are the chances that I roll a six?". Such statements relate to a level of uncertainty (or indeed, a level of certainty). It is this element of likelihood in which we are interested. In particular, we need to find a measure of this likelihood — i.e. the associated probability of particular events.

Our title picture is of the *MV Explorer* in Antarctica in December 2001. This voyage was completed safely. Several years later, the ship was holed and sank in Antarctic waters (with no injuries). How did the insurers calculate the premium due for the voyage? The answer is to be found in this and subsequent sections - and in experience.

Roughly, the insurance industry argues:

$$\text{Probability of accident} = \frac{\text{Number of accidents}}{\text{Number of voyages}}$$

The premium is decided by multiplying this probability by the amount insured and adding a profit margin. This is, of course, a considerable simplification of the very complex work of the insurance actuary.

Probability as a long-term relative frequency

An experiment is repeated in such a way that a series of independent and identical trials are produced, so that a particular event A is observed to either occur or not occur. We let N be the total number of trials carried out and $n(A)$ (or $|A|$) be the number of times that the event A was observed.

We then call the ratio $\frac{n(A)}{n}$ $\left(\text{or } \frac{|A|}{N}\right)$ the **relative frequency** of the event A.

This value provides some indication of the likelihood of the event A occurring.

In particular, for large values of N we find that the ratio $\frac{n(A)}{N}$ tends to a number called the **probability** of the event A, which we denote by $p(A)$ or $P(A)$.

As $0 \le n(A) \le N$, this number, $P(A)$, must lie between 0 and 1 (inclusive), i.e. $0 \le P(A) \le 1$.

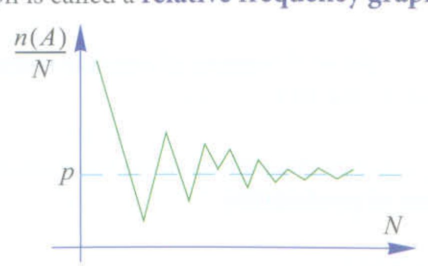
A more formal definition is as follows:

If a random experiment is repeated N times, in such a way that each of the trials is identical and independent, where $n(A)$ is the number of times event A has occurred after N trials, then:

$$\text{As } N \to \infty, \frac{n(A)}{N} \to P(A)$$

It is possible to provide a graph of such a situation, which shows that as N increases, the ratio $\frac{n(A)}{N}$ tends towards some value p, where in fact, $p = P(A)$.

Such a graph is called a **relative frequency graph**.



As far as our actuary is concerned, this demonstrates that the more information they have about a risk, the more reliable the premium calculation will be.

Theoretical probability

When the circumstances of an experiment are always identical, we can arrive at a value for the probability of a particular event by using mathematical reasoning, often based on an argument reflecting some form of symmetry (i.e. without the need to repeatedly perform the experiment). This type of probability is called **theoretical probability**.

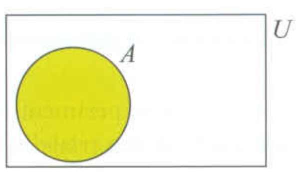
For example, when we roll a die, every possible outcome, known as the **sample space**, can be listed as $U = \{1, 2, 3, 4, 5, 6\}$ (sometimes the letter ε is used instead of U). The probability of obtaining a "four" (based on considerations of **symmetry of equal likelihood**) is given by $^1/_6$. Such a probability seems obvious, as we would argue that:

"Given there are six possible outcomes and each outcome is equally likely to occur (assuming a fair die), then the chances that a 'four' occurs must be one in six, i.e. $^1/_6$."

Laws of probability

We will restrict our arguments to **finite sample spaces**. Recall, that a **sample space** is the set of every possible outcome of an experiment, and that an **event** is any subset of the sample space. This relationship is often represented with a Venn diagram:

The Venn diagram shows the sample A space U, with the event A, as a subset.



Definition of probability

If an experiment has equally likely outcomes and of these the event A is defined, then the **theoretical probability of event A** occurring is given by:

$$P(A) = \frac{n(A)}{n(U)} = \frac{\text{Number of outcomes in which A occurs}}{\text{Total number of outcomes in the sample space}}$$

Where $n(U)$ is the total number of possible outcomes in the sample space, U, (i.e. $n(U) = N$).

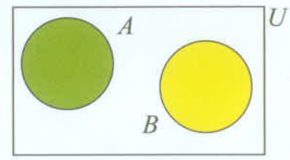
As a consequence of this definition we have what are known as the **axioms of probability**:

1. $0 \le P(A) \le 1$

2. $P(\varnothing) = 0$ and $P(U) = 1$

   That is, if $A = \varnothing$, then the event A can never occur
   $A = U$ implies that the event A is a certainty.

3. If A and B are both subsets of U and are mutually exclusive, then $P(A \cup B) = P(A) + P(B)$.



Note: Two events A and B are said to be **mutually exclusive** (or disjoint) if they have no elements in common, i.e. if $A \cap B = \varnothing$.

Example D.5.1

A fair die is thrown. List the sample space of the experiment and hence find the probability of observing:

a a multiple of 3 b an odd number.

Are these events mutually exclusive?

The sample space is $U = \{1, 2, 3, 4, 5, 6\}$. Let A be the even 'obtaining a multiple of 3'.

We then have that $A = \{3, 6\}$.

Therefore, $P(A) = \frac{n(A)}{n(U)} = \frac{2}{6} = \frac{1}{3}$

Let B be the event 'obtaining an odd number'.

Here $B = \{1, 3, 5\}$ and so $P(B) = \frac{n(B)}{n(U)} = \frac{3}{6} = \frac{1}{2}$.

In this case, $A = \{3, 6\}$ and $B = \{1, 3, 5\}$, so that $A \cap B = \{3\}$
Therefore, as $A \cap B \neq \varnothing$ A and B are not mutually exclusive.

Example D.5.2

Two coins are tossed. Find the probability that:

a two tails are showing

b a tail is showing.

Let H denote the event a head is showing and T the event tail is showing. This means that the sample space (with two coins) is given by $U = \{HH, HT, TH, TT\}$.

The event that two tails are showing is given by the event $\{TT\}$, therefore, we have that:

$P(\{TT\}) = \frac{n(\{TT\})}{n(U)} = \frac{1}{4}$.

The event that one tail is showing is given by $\{HT, TH\}$.

Therefore,: $P(\{HT, TH\}) = \frac{n(\{HT, TH\})}{n(U)} = \frac{2}{4} = \frac{1}{2}$.

Example D.5.3

A card is drawn from a standard deck of 52 playing cards. What is the probability that a diamond card is showing?

Let D denote the event 'a diamond card is selected'.

This means that $n(D) = 13$ as there are 13 diamond cards in a standard deck of cards.

Therefore, $P(D) = \frac{n(D)}{n(U)} = \frac{13}{52} = \frac{1}{4}$.

Problem-solving Strategies in Probability

When dealing with probability problems it is often useful to draw some form of diagram to help 'visualize' the situation. Diagrams can be in the form of:

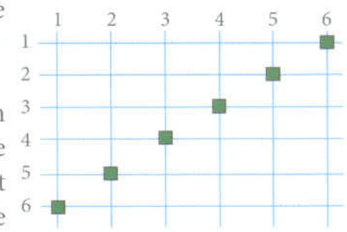
1. Venn diagrams.
2. Tree diagrams.
3. Lattice diagrams.
4. Karnaugh maps (probability tables).
5. As a last resort, any form of diagram that clearly displays the process under discussion (e.g. flow chart).

It is fair to say that some types of diagrams lend themselves well to particular types of problems. These will be considered in due course.

Example D.5.4

Find the probability of getting a sum of 7 on two throws of a die.

In this instance, we make use of a lattice diagram to display all possible outcomes. From the diagram, we can list the required event (and hence find the required probability):



Let S denote the event 'A sum of seven is observed'. From the lattice diagram, we see that there are 6 possibilities where a sum of seven occurs.

In this case: $S = \{(1, 6), (2, 5), (3, 4), (4, 3), (5, 2), (6, 1)\}$.

Therefore: $P(S) = \frac{n(S)}{n(U)} = \frac{6}{36} = \frac{1}{6}$

Exercise D.5.1

1. From a bag containing 6 white and 4 red balls, a ball is drawn at random. What is the probability that the ball selected is:

   a red. b white. c not white.

2. From an urn containing 14 marbles of which 4 are blue and 10 are red, a marble is selected at random. What is the probability that:

   a the marble is blue. b the marble is red.

3. A letter is chosen at random from the letters of the alphabet. What is the probability that:

   a the letter is a vowel.

   b the letter is a consonant.

4. A coin is tossed twice. List the sample space and find the probability of observing:

   a two heads.

   b at least one head.

5. A coin is tossed three times. List the sample space and find the probability that:

   a two heads show uppermost.

   b at least two heads show uppermost.

   c three heads or three tails are showing.

6. A letter is chosen at random from the word FERTILITY. Find the probability that the letter chosen is:

   a a T. b an I.

   c a consonant. d a vowel.

7. A bag has 20 coins numbered from 1 to 20. A coin is drawn at random and its number is noted. What is the probability that the coin drawn has:

   a an even number on it?

   b has a number that is divisible by 3?

   c has a number that is divisible by 3 or 5?

8. A die is rolled twice. Use a lattice diagram to illustrate the sample space. What is the probability of observing:

   a at least one five. b a four and a three.

   c a pair. d a sum of eight.

9. A family has three children. List the sample space and hence find the probability that:

   a there are 3 boys.

   b there are 2 boys and 1 girl.

   c there are at least two girls.

10. A card is selected from a pack of 52 cards. Find the probability that the card is:

   a red b a heart

   c red and a heart.

11. A cube is drawn at random from an urn containing 16 cubes of which 6 are red, 4 are white and 6 are black. Find the probability that the cube is:

   a red b white

   c black d red or black.

12. A coin is tossed and a die is rolled simultaneously. Draw a lattice diagram to depict this situation.

   a Using your lattice diagram, list the sample space.

   b What is the probability of observing a tail and an even number?

13. A die is rolled three times. Find the probability of observing:

   a three sixes.

   b three even numbers.

   c two odd numbers.

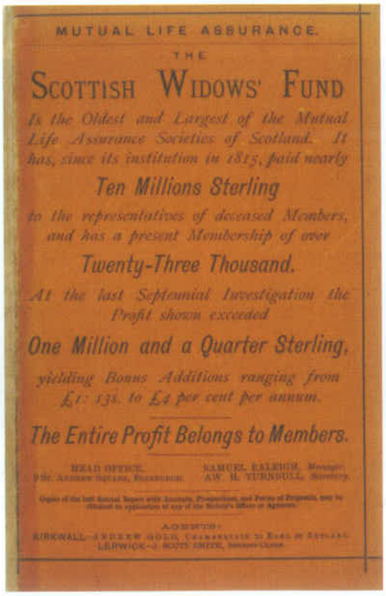
   (Hint: You may need to draw a three-dimensional lattice diagram.)

Application

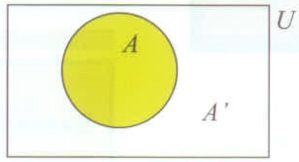
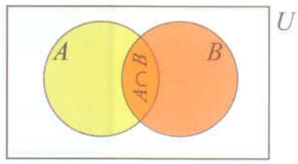
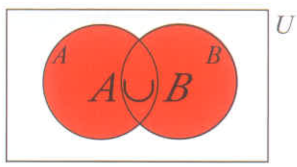
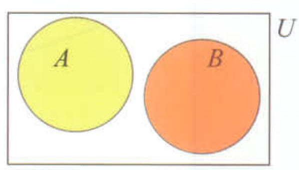
Scottish Widows One of the first life insurance companies in the world was set up in 1812 to care for widows and orphans. Their plan was to collect premiums to create a capital fund. The pensions would be paid from the interest generated by the fund. The capital was intended to be preserved.

They got their calculations right and are trading to this day.

This advertisement dates from 1878.



From the axioms of probability we can develop further rules to help solve problems that involve chance. We illustrate these rules with the aid of Venn diagrams.

Event	Set language	Venn diagram	Probability result
The **complement** of A is denoted by A'.	A' is the complement to the set A, i.e. the set of elements that do not belong to the set A.		$\text{P}(A')=1-\text{P}(A)$ $\text{P}(A')$ is the probability that event A does not occur.
The **intersection** of A **and** B: $A\cap B$	$A\cap B$ is the intersection of the sets A and B, i.e. the set of elements that belong to **both** the set A **and** the set B.		$\text{P}(A\cap B)$ is the probability that both A and B occur.
The **union** of events A and B: $A\cup B$	$A\cup B$ is the union of the sets A and B, i.e. the set of elements that belong to A **or** B or both A **and** B.		$\text{P}(A\cup B)$ is the probability that either event A or event B (or both) occur. From this we have what is known as the '**Addition rule**' for probability: $\text{P}(A\cup B)=\text{P}(A)+\text{P}(B)-\text{P}(A\cap B)$
If $A\cap B=\varnothing$ the events A and B are said to be **disjoint**. That is, they have no elements in common.	If $A\cap B=\varnothing$ the sets A and B are **mutually exclusive**.		If A and B are mutually exclusive events then event A and event B cannot occur simultaneously, i.e. $\text{n}(A\cap B)=0$ $\Rightarrow \text{P}(A\cap B)=0$ Therefore $\text{P}(A\cup B) = \text{P}(A) + \text{P}(B)$.

Although we now have a number of 'formulae' to help us solve problems that involve probability, using other forms of diagrams to clarify situations and procedures should not be overlooked.

Example D.5.5

A card is randomly selected from an ordinary pack of 52 playing cards. Find the probability that it is either a 'black card' or a 'king'.

Let B be the event 'A black card is selected.' and K the event 'A king is selected'.

We first note that event B has as its elements the Jack of spades (J♠), the Jack of clubs (J♣), the Queen of spades (Q♠), the Queen of clubs (Q♣) and so on. This means that:

B = {K♠,K♣,Q♠,Q♣,J♠,J♣,10♠,10♣,9♠,9♣,8♠,8♣,7♠,7♣, 6♠,6♣,5♠,5♣,4♠,4♣,3♠,3♣,2♠,2♣,A♠,A♣} and

K = {K♠, K♥, K♣, K♦}, so that $B\cap K$ = {K♠, K♣}.

Using the addition rule, $\text{P}(B\cup K) = \text{P}(B)+\text{P}(K)-\text{P}(B\cap K)$

we have $\text{P}(B\cup K) = \frac{26}{52}+\frac{4}{52}-\frac{2}{52} = \frac{7}{13}$.

Note the importance of subtracting $\frac{2}{52}$ as this represents the fact that we have included the event {K♠, K♣} twice when finding B and K.

Example D.5.6

A bag has 20 coins numbered from 1 to 20. A coin is drawn at random and its number is noted. What is the probability that the coin has a number that is divisible by 3 or by 5?

Let T denote the event "The number is divisible by 3" and S, the event "The number is divisible by 5".

Using the addition rule we have:

$$P(T \cup S) = P(T) + P(S) - P(T \cap S)$$

Now, $T = \{3, 6, 9, 12, 15, 18\}$ and $S = \{5, 10, 15, 20\}$ so that $T \cap S = \{15\}$.

Therefore, we have $P(T) = \frac{6}{20}$ and $P(S) = \frac{4}{20}$

and $P(T \cap S) = \frac{1}{20}$.

This means that $P(T \cup S) = \frac{6}{20} + \frac{4}{20} - \frac{1}{20} = \frac{9}{20}$.

Example D.5.7

If $P(A) = 0.6, P(B) = 0.3$ and $P(A \cap B) = 0.2$, find:

a $P(A \cup B)$ b $P(B')$

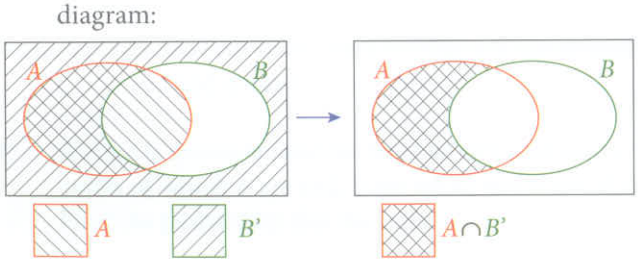
c $P(A \cap B')$

a Using the addition formula, we have:

$$\begin{aligned} P(A \cup B) &= P(A) + P(B) - P(A \cap B) \\ &= 0.6 + 0.3 - 0.2 \\ &= 0.7 \end{aligned}$$

b Using the complementary formula: $P(B') = 1 - P(B)$
$= 1 - 0.3$
$= 0.7$

c To determine $P(A \cap B')$, we need to use a Venn diagram:



Using the second Venn diagram we are now in a position to form a new formula:

$$\begin{aligned} P(A \cap B') &= P(A) - P(A \cap B) \\ &= 0.6 - 0.2 \\ &= 0.4 \end{aligned}$$

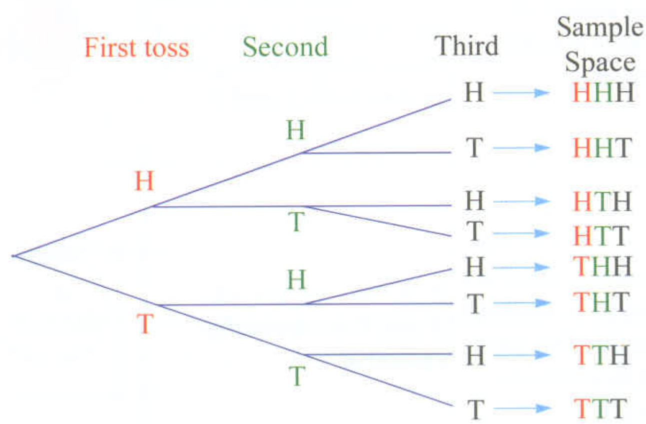
Example D.5.8

A coin is tossed three times. Find the probability of:

a obtaining three tails

b obtaining at least one head.

We begin by drawing a tree diagram to describe the situation:



From the tree diagram we have a sample space made up of eight possible outcomes:

{HHH, HHT, HTH, HTT, THH, THT, TTH, TTT}

Let X be the event "Obtaining three tails", so $X = \{TTT\}$.

Therefore $P(X) = \frac{1}{8}$.

Although we can answer this question by using the tree diagram, we make use of complementary events to solve this problem.

Note - complementary and complimentary are two different words with quite specific different meanings.

Notice that 'At least one head' is the complement of no heads.

So, $P(\text{At least one head}) = P(X') = 1 - P(X) = 1 - \frac{1}{8} = \frac{7}{8}$.

Exercise D.5.2

1. A letter is chosen at random from the letters of the word TOGETHER.

   a Find the probability of selecting a T.

   b Find the probability of selecting a consonant.

   c Find the probability of not selecting an E.

2. A card is drawn at random from a standard deck.

   a Find the probability that the card is an ace.

   b Find the probability that the card is black.

   c Find the probability that the card is an ace and black.

   d Find the probability that the card is an ace or black.

3. A letter is selected at random from the alphabet. Find the probability that the letter is a vowel or comes from the word 'helpful'.

4. The events A and B are such that $P(A) = 0.5$, $P(B) = 0.7$ and $P(A \cap B) = 0.2$.

   Find:

   a $P(A \cup B)$.

   b $P(B')$.

   c $P(A' \cap B)$.

5. The events A and B are such that $p(A) = 0.35$, $p(B) = 0.5$ and $p(A \cap B) = 0.15$.

   Using a Venn diagram (where appropriate), find:

   a $p(A')$.

   b $p(A \cup B)$.

   c $p(A \cup B')$.

6. The events A and B are such that $p(A) = 0.45$, $p(B) = 0.7$ and $p(A \cap B) = 0.20$.

   Using a Venn diagram (where appropriate), find:

   a $p(A \cup B)$.

   b $p(A' \cap B')$.

   c $p((A \cap B)')$.

7. A coin is tossed three times.

   a Draw a tree diagram and from it write down the sample space.

   b Use the results from part a to find the probability of obtaining:

   i only one tail.

   ii at least 2 tails.

   iii 2 tails in succession.

   iv 2 tails.

8. In a class of 25 students it is found that 6 of the students play both tennis and chess, 10 play tennis only and 3 play neither. A student is selected at random from this group.

   Using a Venn diagram, find the probability that the student:

   a plays both tennis and chess.

   b plays chess only.

   c does not play chess.

9. A blue and a red die are rolled together (both numbered one to six).

   a Draw a lattice diagram that best represents this experiment.

   b Find the probability of observing an odd number.

   c Find the probability of observing an even number with the red die.

   d Find the probability of observing a sum of 7.

   e Find the probability of observing a sum of 7 or an odd number on the red die.

10. A card is drawn at random from a standard deck of 52 playing cards. Find the probability that the card drawn is:

a a diamond.

b a club or spade.

c a black card or a picture card.

d a red card or a queen.

11. A and B are two events such that $P(A) = p$, $P(B) = 2p$ and $P(A \cap B) = p^2$.

a Given that $P(A \cup B) = 0.4$, find p.

b Use a Venn Diagram to help you find the following:

i $P(A' \cup B)$.

ii $P(A' \cap B')$.

12. In a group of 30 students 20 hold an Australian passport, 10 hold a Malaysian passport and 8 hold both passports. The other students hold only one passport (that is neither Australian nor Malaysian). A student is selected at random.

a Draw a Venn diagram which describes this situation.

b Find the probability that the student has both passports.

c Find the probability that the student holds neither passport.

d Find the probability that the student holds only one passport.

13. This aircraft has one piston engine.

The engine has four horizontally opposed cylinders. Each cylinder has two spark plugs.

The two sets of spark plugs are connected to separate ignition systems (magnetos).

The fuel is stored in wing tanks. As the tanks are below the engine, there is no gravity feed. There are, however, two fuel pumps. One is mechanical (driven by the engine) and the other is electrical and can be turned on and off by the pilot.

Discuss how these two features and the principles of probability, make this aircraft safer than it would be if it was powered by a car engine.

Using Permutations and Combinations in Probability

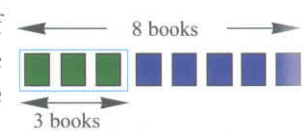
Because enumeration is such an important part of finding probabilities, a sound knowledge of permutations and combinations can help to ease the workload involved.

Example D.5.9

Three maths books, three chemistry books and two physics books are to be arranged on a shelf. What is the probability that the three maths books are together?

The total number of arrangements of all 8 books is 8! = 40320

To determine the number of arrangements that contain the three maths books together we make use of the box method:



We now have 6 boxes to arrange, giving a total of 6! arrangements.

However, the three maths books (within the blue box) can also be arranged in 3! ways.

Therefore, there are $6! \times 3! = 4320$ ways this can be done.

So, P(maths books are together) $= \frac{6! \times 3!}{8!} = \frac{6! \times 3!}{8 \times 7 \times 6!}$

$= \frac{6}{8 \times 7}$

$= \frac{3}{28}$

Example D.5.10

A committee of 5 is randomly chosen from 8 boys and 6 girls. Find the probability that the committee consists of at least 3 boys.

The possibilities are:

Boys	Girls	No. of Selections
3	2	$\binom{8}{3}\times\binom{6}{2} = 840$
4	1	$\binom{8}{4}\times\binom{6}{1} = 420$
5	0	$\binom{8}{5}\times\binom{6}{0} = 56$

The total number of committees with at least 3 boys is 840 + 420 + 56 = 1316

However, the total number of committees of 5 from 14 is $\binom{14}{5} = 2002$.

If X denotes the number of boys on the committee, then $p(X \geq 3) = \frac{1316}{2002} = \frac{94}{143}$.

Exercise D.5.3

1. Five red cubes and 4 blue cubes are placed at random in a row. Find the probability that:
   - a the red cubes are together.
   - b both end cubes are red.
   - c the cubes alternate in colour.
2. Five books of different heights are arranged in a row. Find the probability that:
   - a the tallest book is at the right-hand end.
   - b the tallest and shortest books occupy the end positions.
   - c the tallest and shortest books are together.
   - d the tallest and shortest books are never next to each other.
3. Three cards are dealt from a pack of 52 playing cards. Find the probability that:
   - a two of the cards are kings.
   - b all three cards are aces.
   - c all three cards are aces given that at least one card is an ace.
4. The letters of the word LOTTO are arranged in a row. What is the probability that the Ts are together?
5. A committee of 4 is to be selected from 7 men and 6 women. Find the probability that:
   - a there are 2 women on the committee.
   - b there is at least one of each sex on the committee.
6. A basketball team of 5 is to be selected from 12 players. Find the probability that:
   - a the tallest player is selected.
   - b the captain and vice-captain are selected.
   - c either one, but not both of the captain or vice-captain are selected.
7. Find the probability of selecting one orange, one apple and one pear at random without replacement from a bag of fruit containing five oranges, four apples and three pears.
8. Three red cubes, four blue cubes and six yellow cubes are arranged in a row. Find the probability that:
   - a the cubes at each end are the same colour.
   - b the cubes at each end are of a different colour.
9. A sample of three light bulbs is selected from a box containing 15 light bulbs. It is known that five of the light bulbs in the box are defective.
   - a Find the probability that the sample contains a defective.
   - b Find the probability that the sample contains at least two defectives.

10. Eight people of different heights are to be seated in a row. What is the probability that:

a the tallest and shortest persons are sitting next to each other?

b the tallest and shortest occupy the end positions?

c there are at least three people sitting between the tallest and shortest?

Extra questions

Conditional Probability

Every living creature owes its existence to the fact that its parent(s) existed. We are 'here' because our parents were - their existence is a **condition** for ours.

Informal Definition of Conditional Probability

Conditional probability works in the same way as simple probability. The only difference is that we are provided with some prior knowledge (or some extra condition about the outcome). So, rather than considering the whole sample space, ε, given some extra information about the outcome of the experiment, we only need to concentrate on part of the whole sample space, ε^*. This means that the sample space is reduced from ε to ε^*. Before formalizing this section, we use an example to highlight the basic idea.

Example D.5.11

a In the roll of a die, find the probability of obtaining a '2'.

b After rolling a die, it is noted that an even number appeared. What is the probability that it is a '2'?

a This part is solved using the methods of the previous section: $U = \{1, 2, 3, 4, 5, 6\}$, and so P('2') = $^1/_6$.

b This time, because *we know that an even number has occurred*, we have a new sample space, namely $U^* = \{2, 4, 6\}$. The new sample size is $n(U^*)$.

P('2' *given that an even number showed up*) = $^1/_3$.

Formal Definition of Conditional Probability

If A and B are two events, then **the conditional probability of event A given event B** is found using:

$$P(A|B)=\frac{P(A\cap B)}{P(B)}, P(B)\neq 0$$

Note:

1. If A and B are mutually exclusive then: $P(A|B)=0$

2. From the above rule, we also have the general Multiplication rule:

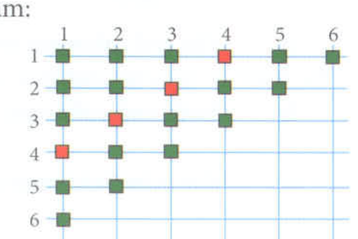
$$P(A\cap B)=P(A|B)\times P(B).$$

It should also be noted that usually $P(A|B)\neq P(B|A)$.

Example D.5.12

Two dice numbered one to six are rolled onto a table. Find the probability of obtaining a sum of five given that the sum is seven or less.

We first draw a lattice diagram:



From the diagram we see that the new sample space is made up of 21 outcomes (boxes) and the event we want (red boxes) consists of 4 outcomes.

Then, $P((X=5)\cap(X\leq 7)) = \frac{4}{36}$ and $P(X\leq 7) = \frac{21}{36}$.

Therefore, $P(X=5|X\leq 7) = \frac{\frac{4}{36}}{\frac{21}{36}} = \frac{4}{21}$.

Example D.5.13

A box contains 2 red cubes and 4 black cubes. If two cubes are chosen at random, find the probability that both cubes are red given that:

a the first cube is not replaced before the second cube is selected

b the first cube is replaced before the second cube is selected.

Let A be the event 'the first cube is red' and B be the event 'the second cube is red'. This means that the event $A \cap B$ must be 'both cubes are red'.

Now, $P(A) = {}^2/_6 = {}^1/_3$ (as there are 2 red cubes from a total of 6 cubes in the box). The value of $P(B)$ depends on whether the selection is carried out with or without replacement.

a If the first cube selected is red and it is not replaced, then we only have 1 red cube left in the box out of a total of five cubes.

So, the probability that the second cube is red given that the first is red is ${}^1/_5$.

That is:

$$p(B|A) = \frac{1}{5} \Rightarrow P(A \cap B) = P(B|A) \times P(A) = \frac{1}{5} \times \frac{1}{3} = \frac{1}{15}.$$

b This time, because the cube is replaced, the probability that the second cube is red given that the first one is red is still ${}^1/_3$.

So that:

$$P(B|A) = \frac{1}{3} \Rightarrow P(A \cap B) = P(B|A) \times P(A) = \frac{1}{3} \times \frac{1}{3} = \frac{1}{9}.$$

Example D.5.14

Two events A and B are such that $P(A) = 0.5$, $P(B) = 0.3$ and $P(A \cup B) = 0.6$. Find:

a $P(A|B)$ b $P(B|A)$ c $P(A'|B)$.

$P(A|B) = \frac{P(A \cap B)}{P(B)}$, therefore we need to find $P(A \cap B)$.

Using the addition rule we have:

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

$$0.6 = 0.5 + 0.3 - P(A \cap B)$$

$$\therefore P(A \cap B) = 0.2$$

So,

$$P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{0.2}{0.3} = \frac{2}{3}$$

$$P(B|A) = \frac{P(B \cap A)}{P(A)} = \frac{0.2}{0.5} = 0.4$$

$$P(A'|B) = \frac{P(A' \cap B)}{P(B)} = \frac{P(B) - P(A \cap B)}{P(B)} = \frac{0.3 - 0.2}{0.3} = \frac{1}{3}$$

Independence

The events A and B are said to be statistically independent if the probability of event B occurring is not influenced by event A occurring.

Therefore we have the mathematical definition:

Two events **A and B are independent if, and only if,**

$$P(A|B) = P(A) \text{ and } P(B|A) = P(B)$$

A more convenient definition for independence is:

Two events **A and B are independent if, and only if,**

$$P(A \cap B) = P(A) \times P(B)$$

This definition can be used as a test to decide if two events are independent. However, as a rule of thumb, if two events are 'physically independent' then they will also be statistically independent.

There are a few points that should always be considered when dealing with independence:

1. Never assume that two events are independent unless you are absolutely certain that they are independent.

2. How can you tell if two events are independent? A good rule of thumb is: If they are physically independent, they are mathematically independent.

3. Make sure that you understand the difference between mutually exclusive events and independent events.

Mutually exclusive means that the events A and B have nothing in common and so there is no intersection, i.e. $A \cap B = \varnothing \Rightarrow P(A \cap B) = 0$.

Independent means that the outcome of event A will not influence the outcome of event B, i.e. $P(A \cap B) = P(A) \times P(B)$

4. Independence need not be for only two events. It can be extended, i.e. if the events A, B and C are each independent of each other then:

$$P(A \cap B \cap C) = P(A) \times P(B) \times P(C)$$

5. Showing that two events, A and B, are independent, requires three steps:

Step 1 Evaluate the product $P(A) \times P(B)$.

Step 2 Determine the value of $P(A \cap B)$ using any means (other than step 1), i.e. use grids, tables, Venn diagrams, . . . i.e. you must not assume anything about A and B.

Step 3 If the answer using Step 1 is equal to the answer obtained in Step 2, then and only then will the events be independent. Otherwise, they are not independent.

Notice that not being independent does not therefore mean that they are mutually exclusive. They simply aren't independent. That's all.

6. Do not confuse the multiplication principle with the rule for independence:

Multiplication principle is $P(A \cap B) = P(A|B) \times P(B)$.

Independence is given by $P(A \cap B) = P(A) \times P(B)$.

Example D.5.15

Two fair dice are rolled. Find the probability that two even numbers will show up.

Let the E_1 and E_2 denote the events 'An even number on the first die.' and 'An even number on the second die.' respectively. In this case, the events are physically independent, i.e. the outcome on one die will not influence the outcome on the other die, and so we can confidently say that E_1 and E_2 are independent events.

Therefore, we have:

$$P(E_1 \text{ and } E_2) = P(E_1 \cap E_2) = P(E_1) \times P(E_2) = \frac{1}{2} \times \frac{1}{2} = \frac{1}{4}$$

Example D.5.16

Debra has a chance of 0.7 of winning the 100 m race and a 60% chance of winning the 200 m race.

a Find the probability that she only wins one race.

b Find the probability that she wins both races.

a Let W_1 denote the event 'Debra wins the 100 m race' and W_2, the event 'Debra wins the 200 m race'.

If Debra wins only one race she must either:

win the 100 m **and** lose the 200 m **or**

win the 200 m **and** lose the 100 m.

That is, we want:

$$P(W_1 \cap W_2') = P(W_1) \times P(W_2') = 0.7 \times 0.4 = 0.28$$

or we can multiply the probabilities because the events are independent **(why?)**:

$$P(W_2 \cap W_1') = P(W_2) \times P(W_1') = 0.6 \times 0.3 = 0.18.$$

Therefore, the required probability is $0.28 + 0.18 = 0.46$

Notice that if W_1 and W_2 are independent, then so too are their complements.

b Winning both races means that Debra will win the 100 m **and** 200 m race.

Therefore, we have:

$$P(W_1 \cap W_2) = P(W_1) \times P(W_2) = 0.7 \times 0.6 = 0.42.$$

Notice how we have made repeated use of the word '**and**'. This emphasizes the fact that we are talking about the intersection of events.

Example D.5.17

Four seeds are planted, each one having an 80% chance of germinating. Find the probability that:

a all four seeds will germinate

b at least one seed will germinate.

a Let G_i denote the event that the i th seed germinates.

This means that $P(G_1) = P(G_2) = P(G_3) = P(G_4) = 0.8$

It is reasonable to assume that each seed will germinate independently of the other.

Therefore, P(All four seeds germinate) =

$$P(G_1 \cap G_2 \cap G_3 \cap G_4) = P(G_1) \times P(G_2) \times P(G_3) \times P(G_4)$$

$$= (0.8)^4$$

$$= 0.4096$$

b Now, P(At least one seed will germinate) = 1 – p(No seeds germinate).

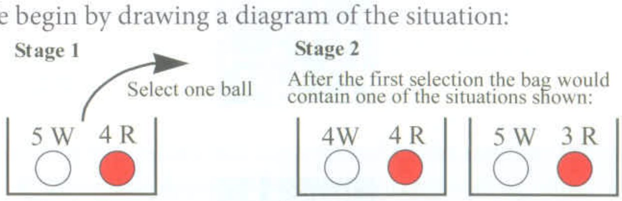
P(Any **one** seed does not germinate) $= P(G_i') = 0.2$

Therefore, P(At least one seed will germinate) $= 1 - (P(G_i'))^4 = 1 - (0.2)^4 = 0.9984$.

Example D.5.18

A bag contains 5 white balls and 4 red balls. Two balls are selected in such a way that the first ball drawn is not replaced before the next ball is drawn. Find the probability of selecting exactly one white ball.

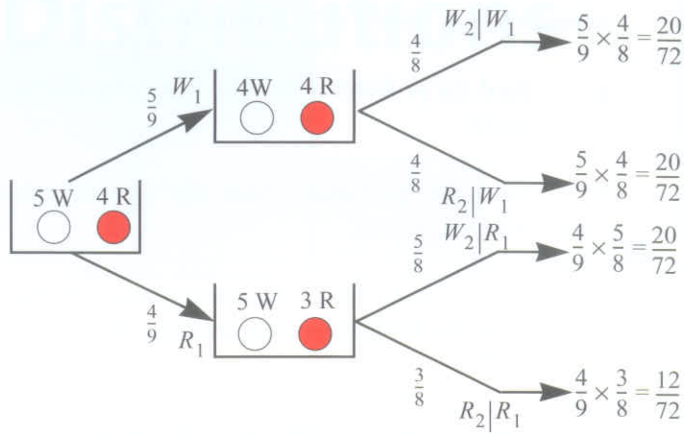
We begin by drawing a diagram of the situation:



From our diagram we notice that there are two possible sample spaces for the second selection.

As an aid, we make use of a tree diagram, where W_i denotes the event 'A white ball is selected on the ith trial' and R_i denotes the event 'A red ball is selected on the ith trial'.

The event 'Only one white' occurs if the first ball is white **and** the second ball is red, **or** the first ball is red **and** the second ball is white.



$$P(\text{One White ball}) = P(W_1 \cap R_2) + P(R_1 \cap W_2)$$

$$= P(R_2|W_1) \times P(W_1) + P(W_2|R_1) \times P(R_1)$$

$$= \frac{4}{8} \times \frac{5}{9} + \frac{5}{8} \times \frac{4}{9}$$

$$= \frac{5}{9}$$

Exercise D.5.4

1. Two events A and B are such that $p(A) = 0.6$, $p(B) = 0.4$ and $p(A \cap B) = 0.3$. Find the probability of the following events.

   a $A \cup B$ b $A|B$

   c $B|A$ d $A|B'$

2. A and B are two events such that $p(A) = 0.3$ $p(B) = 0.5$ and $p(A \cup B) = 0.55$. Find the probability of the following events:

   a $A|B$ b $B|A$

   c $A|B'$ d $A'|B'$

3. Urn A contains 9 cubes of which 4 are red. Urn B contains 5 cubes of which 2 are red. A cube is drawn at random and in succession from each urn.

   a Draw a tree diagram representing this process.

   b Find the probability that both cubes are red.

   c Find the probability that only 1 cube is red.

d If only 1 cube is red, find the probability that it came from urn A.

4. A box contains 5 red, 3 black, and 2 white cubes. A cube is randomly drawn and has its colour noted. The cube is then replaced, together with 2 more of the same colour. A second cube is then drawn.

a Find the probability that the first cube selected is red.

b Find the probability that the second cube selected is black.

c Given that the first cube selected was red, what is the probability that the second cube selected is black?

5. A fair coin, a double-headed coin and a double-tailed coin are placed in a bag. A coin is randomly selected. The coin is then tossed.

a Draw a tree diagram showing the possible outcomes.

b Find the probability that the coin lands with a tail showing uppermost.

c In fact, the coin falls 'heads', find the probability that it is the 'double-headed' coin.

6. Two unbiased coins are tossed together. Find the probability that they both display heads given that at least one is showing a head.

7. A money box contains 10 discs, 5 of which are yellow, 3 of which are black and 2 green. Two discs are selected in succession, with the first disc not replaced before the second is selected.

a Draw a tree diagram representing this process.

b Hence find the probability that the discs will be of a different colour.

c Given that the second disc was black, what is the probability that both were black?

8. Two dice are rolled. Find the probability that the faces are different given that the dice show a sum of 10.

9. Given that $p(A) = 0.6$, $p(B) = 0.7$ and that A and B are independent events.

Find the probability of the events:

a	$A \cup B$	b	$A \cap B$
c	$A \mid B'$	d	$A' \cap B$

10. The probability that an animal will still be alive in 12 years is 0.55 and the probability that its mate will still be alive in 12 years is 0.60. Find the probability that:

a both will still be alive in 12 years.

b only the mate will still be alive in 12 years.

c at least one of them will still be alive in 12 years.

d the mate is still alive in 12 years given that only one is still alive in 12 years.

11. Tony has a 90% chance of passing his maths test, whilst Tanya has an 85% chance of passing the same test. If they both sit for the test, find the probability that:

a only one of them passes.

b at least one passes the test.

c Tanya passed given that at least one passed.

12. The probability that Roger finishes a race is 0.55 and the probability that Melissa finishes the same race is 0.6. Because of team spirit, there is an 80% chance that Melissa will finish the race if Roger finishes the race. Find the probability that:

a both will finish the race.

b Roger finishes the race given that Melissa finishes.

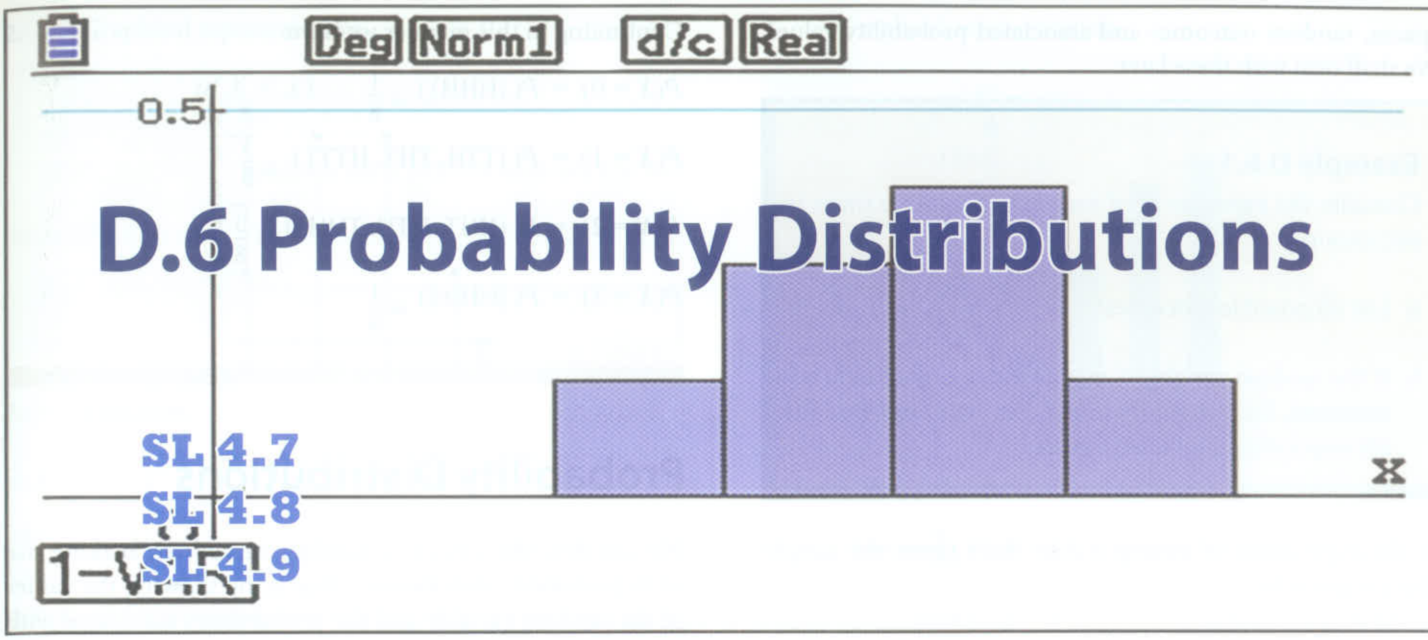
13. If A and B are independent events, show that their complementary events are also independent events.

Extra questions

Answers

D.6 Probability Distributions

SL 4.7
SL 4.8
SL 4.9



Discrete Random Variables

Concept of a random variable

Consider the experiment of tossing a coin twice. The sample space, S, (i.e. the list of all possible outcomes) of this experiment can be written as S = {HH, HT, TH, TT}.

We can also assign a numerical value to these outcomes. For example, we can assign the number

0 to the outcome {HH},

1 to the outcomes {HT, TH} and

2 to the outcome {TT}.

These numerical values are used to represent the number of times that a tail was observed **after** the coin was tossed.

The numbers 0, 1 and 2 are **random** in nature, that is, until the coins are tossed we have no idea as to which one of the outcomes will occur. We define a random variable as follows:

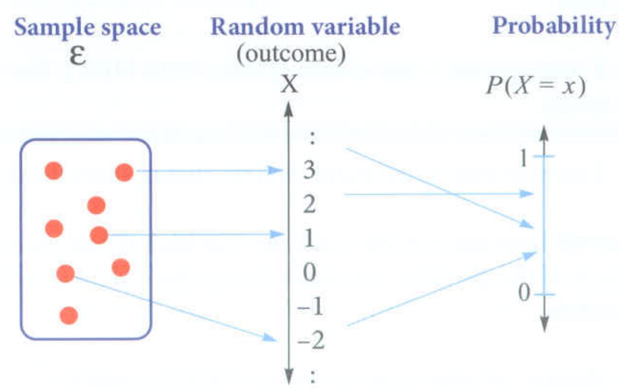
A **random variable, X** (random variables are usually denoted by capital letters), which can take on exactly n numerical values, each of which corresponds to one and only one of the events in the sample space is called a **discrete random variable**.

Note that the values that correspond to the random variables X, Y, Z ... are denoted by their corresponding lower case letters, x, y, z ... For the example above, $X = \{x: x = 0, x = 1, x = 2\}$.

Discrete random variable

A **discrete random variable** is one in which we can produce a countable number of finite or infinite outcomes. X of these discrete random variables are usually associated with a counting process. The number of plants that will flower, the number of defective items in a box or the number of items purchased at a supermarket store are examples with a finite set of values. On the other hand, the number of times to roll a pair of dice until we get a double is an example of a discrete random variable with an infinite set of values, because the possible values for becomes $X = 0, 1, 2,...$

We can display this concept using a simple diagram such as the one below:



We write
'The **probability** that the random variable
X = x is p' as:
P(X = x) = p.

Note that the probability of any event must always lie between 0 and 1, inclusive.

To obtain the sample space, we may need to carry out an experiment. However, as we shall see, there are many types of random variables that already possess their own sample spaces, random outcomes and associated probability values. We shall deal with these later.

Example D.6.1

Consider the experiment of tossing a coin three times in succession.

a List all possible outcomes.

b If the random variable X denotes the number of heads observed, list the values that X can have and find the corresponding probability values.

In the experiment of tossing a coin three times the sample space is given by

S = {HHH, HHT, HTH, THH, TTH, THT, HTT, TTT},

where the event {HTH} represents the observation, head, tail, head, in that order.

This means that on any one trial of this experiment, we could have obtained 0 heads, 1 head, 2 heads or 3 heads.

Therefore, the random variable X has as its possible values the numbers 0, 1, 2, 3.

That is:

$X = 0$ corresponds to the event {TTT}, that is, no heads.

$X = 1$ corresponds to the events {TTH, THT, HTT}, that is, one head.

$X = 2$ corresponds to the events {THH, HTH, HHT}, that is, two heads.

$X = 3$ corresponds to the event {HHH}, that is, three heads.

Once we have our sample space, we can look at the chances of each of the possible outcomes. In all there are 8 possible outcomes.

The chances of observing the event {HHH} would be $\frac{1}{8}$, i.e. $P(X = 3) = \frac{1}{8}$.

To find $P(X = 2)$, we observe that the outcome '$X = 2$' corresponds to {THH, HTH, HHT}.

In this case there is a chance of 3 in 8 of observing the event where '$X = 2$'.

Continuing in this manner we have:

$$P(X = 0) = P(\{\text{HHH}\}) = \frac{1}{8}$$

$$P(X = 1) = P(\{\text{TTH, THT, HTT}\}) = \frac{3}{8}$$

$$P(X = 2) = P(\{\text{HHT, HTH, THH}\}) = \frac{3}{8}$$

$$P(X = 3) = P(\{\text{HHH}\}) = \frac{1}{8}$$

Probability Distributions

We can describe a discrete random variable by making use of its probability distribution. That is, by showing the values of the random variable and the probabilities associated with each of its values.

A probability distribution can be displayed in any one of the following formats:

1. Tabular form

2. Graphical representation

(*With the probability value on the vertical axis, and the values of the random variable on the horizontal axis.*)

3. Function

(*A formula that can be used to determine the probability values.*)

Example D.6.2

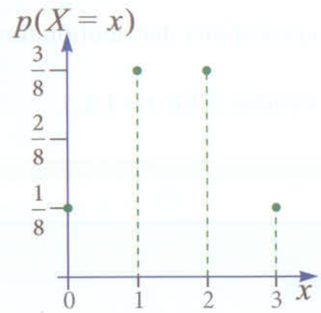
Use each of the probability distribution representations discussed to display the results of the experiment where a coin is tossed three times in succession.

Let the random variable X denote the number of heads observed in three tosses of a coin.

1. Tabular form:

x	0	1	2	3
$p(X = x)$	$\frac{1}{8}$	$\frac{3}{8}$	$\frac{3}{8}$	$\frac{1}{8}$

2. Graphical representation:



3. Function:

$p(X=x) = \binom{3}{x}\left(\frac{1}{2}\right)^3, x = 0, 1, 2, 3$, where:

$\binom{3}{x} = \dfrac{3!}{(3-x)!x!}$.

A probability distribution function (pdf) allows us to quickly identify the probability for a single event in our entire sample space. However, sometimes it may be more convenient to consider the sum of certain events. For example, we may be interested in finding the probability for seeing at least two heads by considering:

$$P(X=2)+P(X=3)=\frac{4}{8}.$$

For such interests, we can consider using the **cumulative distribution function** (cdf) of a random variable, which is defined by for $F(x)=P(X \le x)$ for $-\infty < x < \infty$. Similar to a probability distribution function, the corresponding cumulative distribution function also be expressed in function notation, in tabular form and/or with a graph.

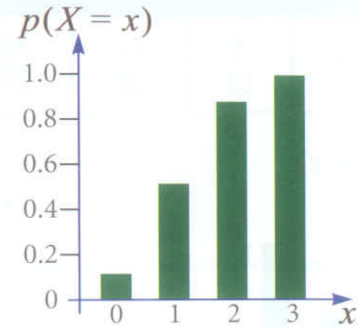
Example D.6.3

Referring to Example D.6.2, use each of the probability distribution representations discussed to display the cumulative results of the experiment where a coin is tossed three times in succession.

1. Tabular form:

x	0	1	2	3
$p(X=x)$	$\frac{1}{8}$	$\frac{3}{8}$	$\frac{3}{8}$	$\frac{1}{8}$
$F(X)$	$\frac{1}{8}$	$\frac{4}{8}$	$\frac{7}{8}$	$\frac{8}{8}$

2. Graphical representation:



3. Function:

$$F(x)=\sum P(X=x), x=0,1,2,3$$

It is important to note that not all probability distributions are symmetrical.

Example D.6.4

A stretch of road between point A and point B has three sets of traffic lights and each set of traffic light has three colours (i.e. red, yellow, and green).

a. List the sample space for all possible permutations of the traffic lights as one travels from point A to point B.

b. Let X be the random variable for the number of red lights. Determine the corresponding probabilities for this random variable.

c. Display the probability distribution with on a graph.

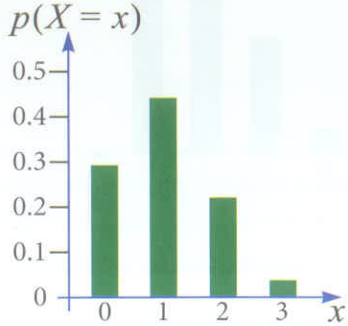
a.

RRR	RRG	RRY	RGR	RYR	RGG
RYY	RGY	RYG	YYY	YYR	YYG
YRY	YGY	YRR	YGG	YRG	YGR
GGG	GGR	GGY	GRG	GYG	GRR
GYY	GYR	GRY			

b.

x	0	1	2	3
$p(X=x)$	$\frac{8}{27}$	$\frac{12}{27}$	$\frac{6}{27}$	$\frac{1}{27}$

c.



Properties of the probability function

We can summarise the features of any discrete probability function as follows:

1. The probability for any value of X must **always lie between** 0 and 1 (inclusive).

   That is, $0 \le P(X = x_i) \le 1$ for all values of x_i.

2. For the n mutually exclusive and exhaustive events, $A_1, A_2, \ldots, A_n$ that make up the sample space ε, then, the sum of the corresponding probabilities must be 1.

That is,

$$\sum_{i=1}^{i=n} P(X = x_i) = P(X = x_1) + P(X = x_2) + \ldots + P(X = x_n) = 1$$

Where $P(X = x_i)$ is the probability of event A_i occurring.

Any function that does not abide by these two rules cannot be a probability function.

Example D.6.5

Show that $f(x) = \frac{x+2}{12}$ for $x = 1,2,3$ is a probability distribution function for a discrete random variable X.

We will begin by substituting $x = 1,2,3$ $f(x) = \frac{x+2}{12}$ into and get:

$f(1) = \frac{3}{12}, f(2) = \frac{4}{12}$ and $f(3) = \frac{5}{12}$

Since $0 \le f(x) \le 1$ for $x = 1,2,3$ and $f(1) + f(2) + f(3) = 1$

$f(x) = \frac{x+2}{12}$ is a probability distribution function for a discrete random variable X for $x = 1,2,3$.

Example D.6.6

Consider the random variable X with probability function defined by

$P(X = 0) = \alpha$, $P(X = 1) = 2\alpha$ and $P(X = 2) = 3\alpha$

Determine the value of α.

Because we are given that this is a probability function, then summing all the probabilities must give a result of 1.

Therefore we have that:

$$P(X = 0) + P(X = 1) + P(X = 2) = 1$$
$$\therefore \alpha + 2\alpha + 3\alpha = 1$$
$$\Leftrightarrow 6\alpha = 1$$
$$\Leftrightarrow \alpha = \frac{1}{6}$$

Example D.6.7

The probability distribution of the random variable X is represented by the function

$$P(X = x) = \frac{k}{x}, x = 1, 2, 3, 4, 5, 6.$$

Find:

a the value of k b $P(3 \le X \le 5)$.

Using the fact that the sum of all the probabilities must be 1, we have:

$$P(X = 1) + P(X = 2) + \ldots + P(X = 6)$$
$$= \frac{k}{1} + \frac{k}{2} + \frac{k}{3} + \frac{k}{4} + \frac{k}{5} + \frac{k}{6} = 1$$
$$\Leftrightarrow \frac{147k}{60} = 1$$

Therefore, $k = \frac{60}{147} = \frac{20}{49}$.

Now, $P(3 \le X \le 5) = P(X = 3) + P(X = 4) + P(X = 5)$

$$= \frac{k}{3} + \frac{k}{4} + \frac{k}{5}$$

$$= \frac{47k}{60}$$

However, we know that $k = \frac{60}{147}$.

Therefore, $P(3 \le X \le 5) = \frac{47}{60} \times \frac{60}{147} = \frac{47}{147}$.

Example D.6.8

A discrete random variable X has a probability distribution defined by the function:

$P(X = x) = \binom{4}{x}\left(\frac{2}{5}\right)^x\left(\frac{3}{5}\right)^{4-x}$ where $x = 0, 1, 2, 3, 4$

Display this distribution using:

a i a table form ii a graphical form.

b Find: i $P(X = 2)$ ii $P(1 < X \le 3)$

a i We begin by evaluating the probability for each value of x: (Using the notation $\binom{4}{x} = {}^4C_x$)

$$p(X = 0) = {}^4C_0\left(\frac{2}{5}\right)^0\left(\frac{3}{5}\right)^{4-0} = \frac{81}{625},$$

$$p(X = 1) = {}^4C_1\left(\frac{2}{5}\right)^1\left(\frac{3}{5}\right)^3 = \frac{216}{625},$$

$$p(X = 2) = {}^4C_2\left(\frac{2}{5}\right)^2\left(\frac{3}{5}\right)^2 = \frac{216}{625},$$

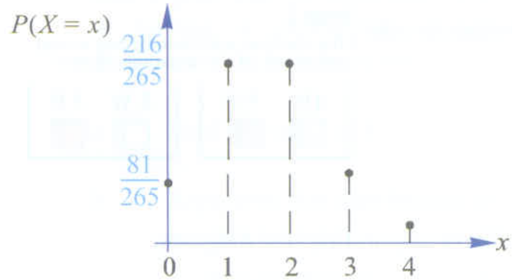
$$p(X = 3) = {}^4C_3\left(\frac{2}{5}\right)^3\left(\frac{3}{5}\right)^1 = \frac{96}{625},$$

$$p(X = 4) = {}^4C_4\left(\frac{2}{5}\right)^4\left(\frac{3}{5}\right)^0 = \frac{16}{625}$$

We can now set up this information in a table:

x	0	1	2	3	4
$P(X = x)$	$\frac{81}{625}$	$\frac{216}{625}$	$\frac{216}{625}$	$\frac{96}{625}$	$\frac{16}{625}$

ii Using the table found in part i, we can construct the following graph:



i From our probability table: $P(X = 2) = \frac{216}{625}$.

ii The statement $P(1<X\le3)$ requires that we find the probability of the random variable X taking on the values from 1 to 3, but excluding 1. Since X is a discrete random variable, the next available value for X when $x = 1$ is excluded is $x = 2$. Therefore, $P(1<X\le3)$ is essentially $P(2\le X\le3)$ and this amounts to evaluating the sum of the corresponding probabilities.

Hence, $P(1<X\le3) = P(2\le X\le3) = \frac{312}{625}$.

Constructing Probability Functions

When we are given the probability distribution, we can determine the probabilities of events. However, there is still one issue that we must deal with:

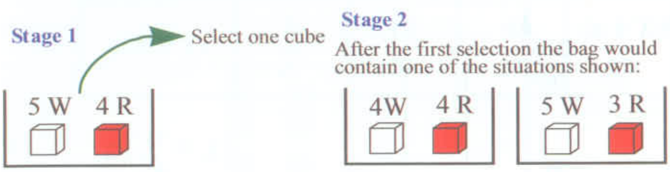
How do we obtain the probabilities in the first place?

Sometimes we recognise a particular problem and know of an existing model that can be used. However, resolving this question is not always an easy task, as this often requires the use of problem-solving skills and modelling skills as well as interpretive skills.

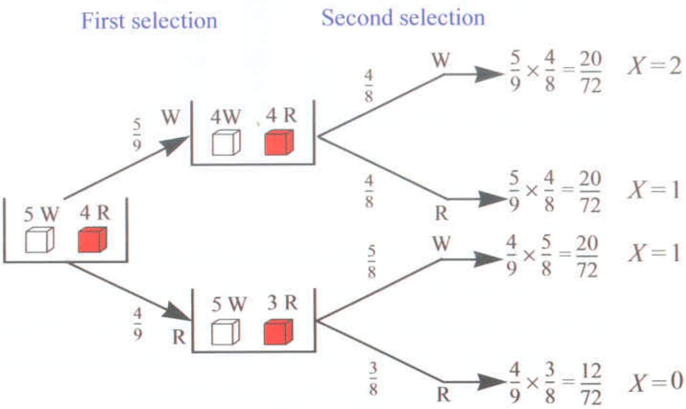
Example D.6.9

A bag contains 5 white cubes and 4 red cubes. Two cubes are selected in such a way that the first cube drawn is not replaced before the second cube is drawn. Find the probability distribution of X, where X denotes the number of white cubes selected from the bag.

We start by drawing a diagram that will help us visualise the situation:



Next, we set up the corresponding tree diagram:



We are now in a position to complete the probability table.

x	0	1	2
$P(X=x)$	$\frac{12}{72}$	$\frac{40}{72}$	$\frac{20}{72}$

Example D.6.10

Two friends, Kirsty and Bridget, independently applied for different jobs. The chance that Kirsty is successful is 0.8 and the chance that Bridget is successful is 0.75.

If X denotes the number of successful applications between the two friends, find the probability distribution of X.

Hence find the probability that:

i both are successful

ii that if one is successful, it is Kirsty.

a Let K denote the event that Kirsty is successful, so that $P(K) = 0.8$ and let B denote the event that Bridget is successful, so that $P(B) = 0.75$.

Now, the event '$X = 0$' translates to 'nobody is successful':

That is, $P(X=0) = P(K' \cap B') = P(K') \times P(B')$.

$= 0.2 \times 0.25 = 0.05$

Similarly, the event '$X = 1$' translates to 'only one is successful':

That is, $P(X=1) = P(K \cap B') + P(K' \cap B)$

$= 0.8 \times 0.25 + 0.2 \times 0.75 = 0.35$

Lastly, the event '$X = 2$' translates to 'both are successful':

That is,

$$P(X=2) = P(K \cap B) = P(K) \times P(B) = 0.8 \times 0.75 = 0.6$$

We can now construct a probability distribution for the random variable X:

x	0	1	2
$P(X=x)$	0.05	0.35	0.60

b i $P(\text{Both successful}) = P(X=2) = 0.60$

$$P(K|\text{Only one is successful}) = \frac{P(K \cap \{X=1\})}{P(\{X=1\})} = \frac{P(K \cap B')}{P(\{X=1\})}$$

$$= \frac{0.20}{0.35} (= 0.5714)$$

Exercise D.6.1

1. The following table show the cumulative distribution function of a discrete random variable X. Find the probability distribution of X.

x	0	1	2	3	4
$F(x)$	0	0.1	0.3	0.8	1

2. If X is a discrete random variable, show that the corresponding probability distribution is related to its cumulative probability distribution by $P(x) = F(x+1) - F(x)$.

3. Find the value of k, so that the random variable X describes a probability distribution.

x	1	2	3	4	5
$P(X=x)$	0.25	0.20	0.15	k	0.10

4. The discrete random variable Y has the following probability distribution:

y	1	2	3	4
$P(Y=y)$	β	2β	3β	4β

Find the value of β.

Find: i $P(Y=2)$ ii $P(Y>2)$

5. A delivery of six television sets contains 2 defective sets. A customer makes a random purchase of two sets from this delivery. The random variable X denotes the number of defective sets purchased by the customer.

a Find the probability distribution table for X.

b Represent this distribution as a graph.

c Find $P(X \le 1)$.

6. A pair of dice is rolled. Let Y denote the sum showing uppermost.

a Determine the possible values that the random variable Y can have.

b Display the probability distribution of Y in tabular form.

c Find $P(Y=8)$.

d Sketch the probability distribution of Y.

7. A fair coin is tossed 3 times.

a Draw a tree diagram representing this experiment.

b Display this information using both graphical and tabular representations.

c If the random variable Y denotes the number of heads that appear uppermost, find $P(Y \ge 2 | Y \ge 1)$.

8. The number of customers that enter a small corner newsagency between the hours of 8 p.m. and 9 p.m. can be modelled by a random variable X having a probability distribution given by $P(X=x) = k(3x+1)$, where $x = 0, 1, 2, 3, 4$.

a Find the value of k.

b Represent this distribution in:
i tabular form ii graphical form.

c What are the chances that at least 2 people will enter the newsagency between 8 p.m. and 9 p.m. on any one given day?

9. The number of cars passing an intersection during the hours of 4 p.m. and 6 p.m. follows a probability distribution that is modelled by the function

$$P(X=x) = \frac{(0.1)^x}{x!}e^{-0.1}, x = 0, 1, 2, 3, \ldots,$$

where the random variable X denotes the number of cars that pass this intersection between 4 p.m. and 6 p.m.

a Find: i $P(X=0)$ ii $P(X=1)$.

b Find the probability of observing at least three cars passing through this intersection between the hours of 4 p.m. and 6 p.m.

10. The number of particles emitted during a one-hour period is given by the random variable X, having a probability distribution

$$P(X=x) = \frac{(4)^x}{x!}e^{-4}, x = 0, 1, 2, 3, \ldots.$$

Find $P(X>4)$.

11. A random variable X has the following probability distribution

x	0	1	2	3
$P(X=x)$	$\frac{1}{6}$	$\frac{1}{2}$	$\frac{1}{5}$	$\frac{2}{15}$

a Find the probability distribution of $Y = X^2 - 2X$

b Find: i $P(Y=0)$ ii $P(Y<3)$.

12. A bakery has six indistinguishable muffins on display. However, two of them have been filled with strawberry jam and the others with apricot jam. Claire, who hates strawberry jam, purchases two muffins at random. Let N denote the number of strawberry jam muffins Claire buys. Find the probability distribution of the random variable N.

Extra questions

Mean and Variance

Central tendency and expectation

For a discrete random variable X with a probability distribution defined by $P(X = x)$, we define the **expectation of the random variable X** as

$$E(X) = \sum_{i=1}^{i=n} x_i P(X = x_i)$$
$$= x_1 \times P(X = x_1) + x_2 \times P(X = x_2) + \ldots + x_n \times P(X = x_n)$$

Where $E(X)$ is read as "The expected value of X". $E(X)$ is interpreted as the **mean value** of X and is often written as μ_X (or simply μ). Often we write the expected value of X as $\Sigma x.P(X = x)$. This is in contrast to the **mode** which is the most common value(s) and the **median** which is the value with half the probabilities below and half above the median value.

So what exactly does *E*(*X*) measure?

The expected value of the random variable is a **measure of the central tendency** of X. That is, it is an indication of its 'central position' – based not only on the values of X, but also the **probability weighting** associated with each value of X. That is, it is the probability-weighted average of its possible values.

To find the value of $E(X)$ using the formula:

$$E(X) = \sum_{i=1}^{i=n} x_i P(X = x_i)$$

we take each possible value of x_i, multiply it by its associated probability $P(X = x_i)$ (i.e. its 'weight') and then add the results. The number that we obtain can be interpreted in two ways:

As a **probability-weighted average**, it is a summary number that takes into account the relative probabilities of each x_i value.

As a **long-run average**, it is a measure of what one could expect to observe if the experiment were repeated a large number of times.

For example, when tossing a fair coin a large number of times (say 500 times), and the random variable X denotes the number of tails observed, we would expect to observe 250 tails, i.e. $E(X) = 250$.

Note: *Although we would expect 250 tails after tossing a coin 500 times, it may be that we do not observe this outcome! For example, if the average number of children per 'family' in Australia is 2.4, does this mean we expect to see 2.4 children per 'family'?*

In short, we may not be able to observe the value $E(X)$ that we obtain.

Example D.6.11

For the random variable X with probability distribution defined by:

x	1	2	3	4
$P(X = x)$	$\frac{1}{10}$	$\frac{2}{10}$	$\frac{3}{10}$	$\frac{4}{10}$

Find the mode, median and mean values of X.

The mode is $X = 4$ (the most probable) and the median 3 (half are above and half below 3 – this is probably best done by sketching the cdf of X).

To find the mean of X we use the formula:

$$E(X) = \sum_{i=1}^{i=n} x_i P(X = x_i)$$

$$E(X) = \left(1 \times \frac{1}{10}\right) + \left(2 \times \frac{2}{10}\right) + \left(3 \times \frac{3}{10}\right) + \left(4 \times \frac{4}{10}\right)$$
$$= \frac{1}{10} + \frac{4}{10} + \frac{9}{10} + \frac{16}{10}$$
$$= 3$$

Therefore, X has a mean value of 3.

Example D.6.12

A fair die is rolled once. If the random variable X denotes the number showing, find the expected value of X.

Because the die is fair we have the following probability distribution:

x	1	2	3	4	5	6
$P(X = x)$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$

$$E(X) = \left(1 \times \frac{1}{6}\right) + \left(2 \times \frac{1}{6}\right) + \left(3 \times \frac{1}{6}\right) + \left(4 \times \frac{1}{6}\right) + \left(5 \times \frac{1}{6}\right) + \left(6 \times \frac{1}{6}\right)$$
$$= 3.5$$

Example D.6.13

Suppose you play a game by tossing a biased coin once. The likelihoods for each side of the coin $P(\text{heads}) = {}^3/_5$ and $P(\text{tails}) = {}^2/_5$. If you toss a head, you will lose \$7.00, but if you toss a tail, you will win \$10.00. Determine the expected value for this game if you play it many times.

$$E(x) = 10\left(\frac{2}{5}\right) - 7\left(\frac{3}{5}\right)$$
$$= \frac{20}{5} - \frac{21}{5}$$
$$= -\frac{1}{5}$$

On average, you would expect to lose \$0.20 each time you play the game.

Variance

Although we now have a means by which we can calculate a measure of the central tendency of a random variable, an equally important aspect of a random variable is its spread. For example, the mean of the three numbers, 100, 110 and 120 is 110. Similarly, the mean of the three numbers 10, 100, and 220 is also 110. Yet clearly, the values in the second set of data have a wider spread than those in the first set of data. The **variance** (or more so, the **standard deviation**) provides a measure of this spread.

The **variance** of a discrete random variable may be considered as the average of the squared deviations about the mean. This provides a measure of the variability of the random variable or the probability dispersion. The variance associated with the random variable X is given by:

$$Var(X) = E((X-\mu)^2) = \sum_{i=1}^{i=n} (x-\mu)^2 P(X=x)$$

However, for computational purposes, it is often better to use the alternative definition:

$$Var(X) = E(X^2) - (E(X))^2 = E(X^2) - \mu^2$$

The variance is also denoted by σ^2 (read as "**sigma squared**"), i.e. $Var(X) = \sigma^2$.

We also have the **standard deviation**, given by

$$Sd(X) = \sigma = \sqrt{Var(X)},$$

which also provides a measure of the spread of the distribution of X.

What is the difference between the $Var(X)$ and the Sd(X)?

Because of the squared factor in the equation for $Var(X)$ (i.e. $Var(X) = E(X^2) - \mu^2$), the units of $Var(X)$ are not the same as those for X. However, because the $Sd(X)$ is the square root of the $Var(X)$, we have "adjusted" the units of $Var(X)$ so that they now have the same units as the random variable X.

The reason for using the $Sd(X)$ rather than the $Var(X)$ is that we can make clearer statistical statements about the random variable X (in particular, statements that relate to an overview of the distribution).

Example D.6.14

The probability distribution of the random variable X is shown below:

x	–2	–1	0	1	2
$P(X = x)$	$\frac{1}{64}$	$\frac{12}{64}$	$\frac{38}{64}$	$\frac{12}{64}$	$\frac{1}{64}$

Find:

a the variance of X

b the standard deviation of X.

a First we need to find $E(X)$ and $E(X^2)$:

$$E(X) = \sum xP(X = x)$$
$$= -2 \times \frac{1}{64} + (-1) \times \frac{12}{64} + 0 \times \frac{38}{64} + 1 \times \frac{12}{64} + 2 \times \frac{1}{64}$$
$$= 0$$

$$E(X^2) = \sum x^2 P(X = x)$$
$$= (-2)^2 \times \frac{1}{64} + (-1)^2 \times \frac{12}{64} + 0^2 \times \frac{38}{64} + 1^2 \times \frac{12}{64} + 2^2 \times \frac{1}{64}$$
$$= \frac{1}{2}$$

b Therefore, $Var(X) = E(X^2) - \mu^2 = \frac{1}{2} - 0^2 = \frac{1}{2}$.

Using $Sd(X) = \sigma = \sqrt{Var(X)}$, we have that:

$$Sd(X) = \sigma = \sqrt{\frac{1}{2}} \approx 0.707 \cdot$$

Exercise D.6.2

1. A discrete random variable X has a probability distribution given by

x	1	2	3	4	5
$P(X = x)$	0.25	0.20	0.15	0.3	0.10

   a Find the mean value of X.

   b Find the variance of X.

2. The discrete random variable Y has the following probability distribution:

y	1	2	3	4
$P(Y = y)$	0.1	0.2	0.3	0.4

   a Find the mean value of Y.

   b Find: i $Var(Y)$ i $Sd(Y)$.

   c Find: i $E(2Y)$ ii $E\left(\frac{1}{Y}\right)$.

3. A random variable X has the following probability distribution:

x	0	1	2	3
$P(X = x)$	$\frac{1}{6}$	$\frac{1}{2}$	$\frac{1}{5}$	$\frac{2}{15}$

   a Find: i $E(X)$ ii $E(X^2)$

   iii $E(X^2 - 2X)$.

   b Find: i $Sd(X)$

   ii $Var(3X + 1)$

   c If $Y = \frac{1}{X+1}$, find : i $E(Y)$ ii $E(Y^2)$

4. A delivery of six television sets contains 2 defective sets. A customer makes a random purchase of two sets from this delivery. The random variable X denotes the number of defective sets the customer purchased. Find the mean and variance of X.

5. A pair of dice are rolled. Let Y denote the sum showing uppermost.

   a Find $E(Y)$.

   b Find $Var(Y)$.

6. How many tails would you expect to observe when a fair coin is tossed 3 times?

7. The number of customers that enter a small corner newsagency between the hours of 8 p.m. and 9 p.m. can be modelled by a random variable X having a probability distribution given by $P(X = x) = k(2x + 1)$, where $x = 0, 1, 2, 3, 4$.

   a Find the value of k.

   b How many customers can be expected to enter the newsagency between 8 p.m. and 9 p.m.?

c Find the standard deviation of X.

8. A discrete random variable Y has its probability distribution function defined as

y	-2	-1	0	1
$P(Y=y)$	k	0.2	$3k$	0.4

a Find k.

b Given that the function, F, is defined by $F(y) = P(Y \le y)$, find:

i $F(-1)$ ii $F(1)$.

c Find:

i the expected value of Y.

ii the variance of Y.

iii the expected value of $(Y+1)^2$.

9. A dart board consisting of concentric circles of radius 1, 2 and 3 units is placed against a wall. A player throws darts at the board, each dart landing at some random location on the board. The player will receive \$9.00 if the smaller circle is hit, \$7.00 if the middle annular region is hit and \$4.00 if the outer annular region is hit. Should players miss the board altogether, they would lose \$$k$ each time. The probability that the player misses the dart board is 0.5. Find the value of k if the game is to be fair.

10. A box contains 7 black cubes and 3 red cubes. Debra selects three cubes from the box without replacement. Let the random variable N denote the number of red cubes selected by Debra.

a Find the probability distribution for N.

b Find: i $E(N)$ ii $Var(N)$.

Debra will win \$2.00 for every red cube selected and lose \$1.00 for every black cube selected. Let the random variable W denote Debra's winnings.

c If $W = aN + b$, find a and b. Hence, find $E(W)$.

11.a A new gambling game has been introduced in a casino: A player stakes \$8.00 in return for the throw of two dice, where the player wins as many dollars as the sum of the two numbers showing uppermost.

How much money can the player expect to walk away with?

b At a second casino, a different gambling game has been set up: A player stakes \$8.00 in return for the throw of two dice, if two sixes come up, the player wins \$252.

Which game would be more profitable for the casino in the long run?

12. Given that $Var(X) = 2$, find:

a $Var(5X)$ b $Var(-3X)$

c $Var(1-X)$.

13. Given that $Var(X) = 3$ and $\mu = 2$, find:

a $E(2X^2 - 4X + 5)$ b $Sd\left(4 - \frac{1}{3}X\right)$

c $E(X^2) + 1 - E((X+1)^2)$.

14. A store has eight toasters left in its storeroom. Three of the toasters are defective and should not be sold. A salesperson, unaware of the defective toasters, selects two toasters for a customer. Let the random variable N denote the number of defective toasters the customer purchases.

Find:

a $E(N)$ b $Sd(N)$.

15. a The random variable Y is defined by:

y	-1	1
$P(Y=y)$	p	$1-p$

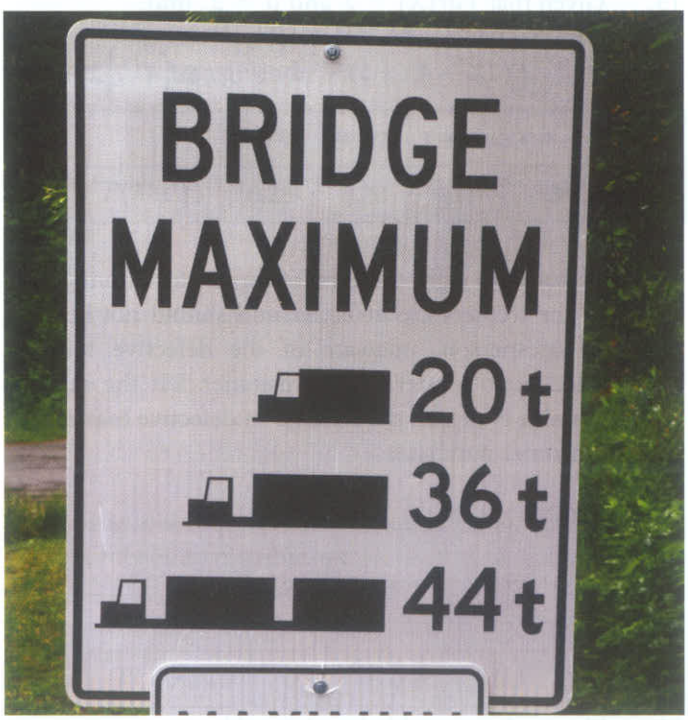
Find the mean and variance of Y.

b The random variable X is defined as $X = Y_1 + Y_2 + Y_3 + \ldots + Y_n$ where each Y_i $i = 1, 2, 3, \ldots, n$ is independent and has the distribution defined in part **a**.

Find: i $E(X)$ ii $Var(X)$

Extra questions

What is the explanation for this rather strange set of weight restrictions for a bridge in the Rockies, Canada.



Is this a probability issue or is the explanation connected to the geometry of the bridge?

The Binomial Distribution

The Binomial Experiment

The binomial distribution is a special type of discrete distribution which finds applications in many settings of everyday life. In this section we summarise the important features of this probability distribution.

Bernoulli Trials

Certain experiments consist of repeated trials where each trial has only two mutually exclusive, possible outcomes. Such trials are referred to as **Bernoulli trials** (after Jacob Bernoulli - pictured). The outcomes of a Bernoulli trial are often referred to as 'a success' or 'a failure'. The terms 'success' and 'failure' in this context do not necessarily refer to the everyday usage of the word success and failure. For example, a 'success' could very well be referring to the outcome of selecting a defective transistor from a large batch of transistors.

We often denote $P(\text{Success})$ by p and $P(\text{Failure})$ by q, where $p + q = 1$ (or $q = 1 - p$).

Properties of the binomial experiment

1. There are a fixed number of trials. We usually say that there are n trials.

2. On each one of the n trials there is only one of two possible outcomes, labelled '*success*' and '*failure*'.

3. Each trial is identical and independent.

4. On each of the trials, the probability of a success, p, is always the same, and the probability of a failure, $q = 1 - p$, is also always the same.

A Bernoulli random variable can be expressed using the following function:

$$p(x) = \begin{cases} p^x(1-p)^{1-x} & \text{if } x = 0 \text{ or } x = 1 \\ 0 & \text{otherwise} \end{cases}$$

The Binomial Distribution

If a (discrete) random variable X has all of the above mentioned properties, we say that X has a **binomial distribution**. The probability distribution function is given by

$$p(X=x)=\binom{n}{x}p^x q^{n-x}=\binom{n}{x}p^x(1-p)^{n-x}, x=0,1,2,\ldots n$$

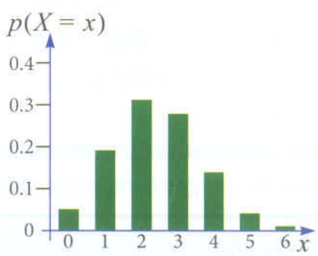
Where X denotes the **number of successes** in n trials such that the **probability of a success on any one trial** is p, $0 \le p \le 1$ and $p+q=1$ (or $q=1-p$)).

We can also express the binomial distribution in a compact form, written as $X \sim B(n,p)$, read as "**X is distributed binomially with parameters n and p**", where n is the number of trials and $p = P(success)$ [it is also common to use $X \sim \text{Bin}(n, p)$].

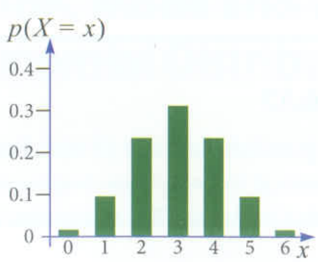
For example, the probability function for $X \sim B(6,0.4)$ (i.e. 6 trials and $p = 0.4$) would be

$$P(X=x)=\binom{6}{x}(0.4)^x(0.6)^{6-x}, x=0,1,2,\ldots,6.$$

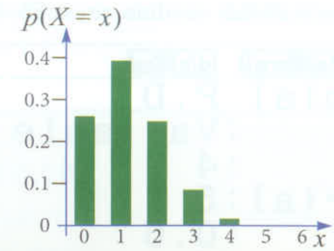
The corresponding graphical representation for this binomial distribution $X \sim B(6,0.4)$ is:



In order to further illustrate the concept of a probability distribution, let us consider the following two graphs:



This graph represents $X \sim B(6,0.5)$, and you will notice that it is a much more symmetrical distribution than the previous one because the chance for success and failure as illustrated by this binomial distribution is 0.5. Hence, there is an equal chance for a single event to be either successful or not successful. When compared to the previous graph, we notice that the distribution is positively skewed because with only 0.4 probability of success, it is more likely to have an unsuccessful outcome than a successful outcome. If the chance for success is further reduced to 0.2, the asymmetry becomes more obvious, as the chance for an unsuccessful outcome is much more likely than for a successful outcome.



Example D.6.15

If $X \sim B(5, 0.6)$, find $P(X=4)$.

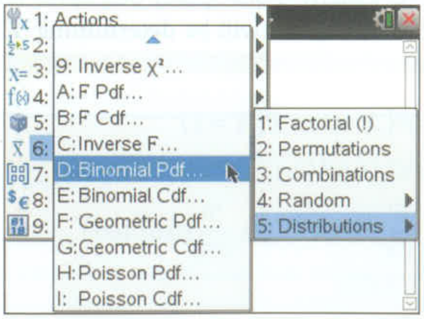
$X \sim B(5, 0.6)$ means that $P(X = 4)$ is the probability of 4 successes in five trials, where each trial has a 0.6 chance of being a success, that is, $n = 5$, $p = 0.6$ and $x = 4$.

$$\begin{aligned}\therefore P(X=4) &= {}^5C_4(0.6)^4(0.4)^{5-4}\\ &= \frac{5!}{1!4!}(0.6)^4(0.4)^1\\ &= 0.2592\end{aligned}$$

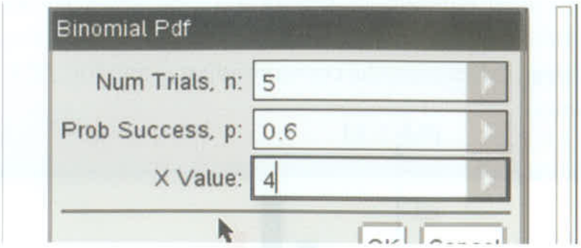
Most modern calculators can perform these calculations. TI NSpire calculators:

MENU / 5. Probability / 5. Distributions / D. Binomial PDF

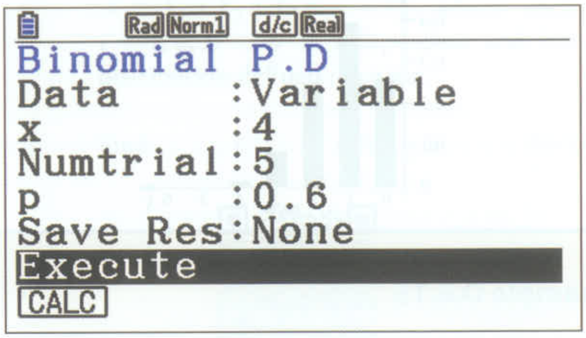
PDF stands for 'Probability Density function'. Make sure you are clear about this and CDF 'Cumulative Density function'.



The previous example is solved:

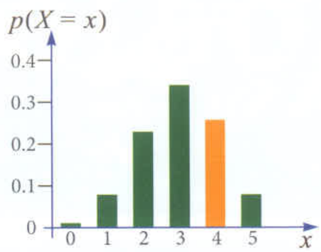


If using the Casio Statistics module, press F5, F5, F1 and set:



F1-CALC completes the calculation.

In Example D.6.15, we have determined the probability for a single value of the random variable X, where $X \sim B(5,0.6)$. Graphically, it can be represented as the following:

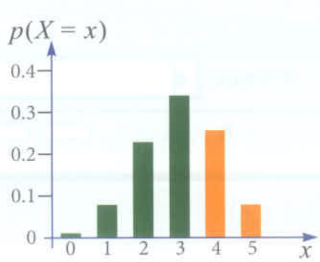


The bar corresponding to the probability for $x = 4$ is highlighted, because we are interested in determining the probability for exactly 4 successful outcomes out of a total of 5 trials.

Sometimes, it may be useful for us to determine the total probability when several exact outcomes are combined together. For example, if we are interested in determining the probability of at least 4 successful outcomes out of a total of 5 trials (i.e. exactly 4 successful outcomes and exactly 5 successful outcomes), we will be determining $P(X \geq 4)$ for $X \sim B(5,0.6)$.

$$P(X \geq 4) = P(X=4) + P(X=5)$$
$$= {}^5C_4(0.6)^4(0.4)^{5-4} + {}^5C_5(0.6)^5(0.4)^{5-5}$$
$$= \frac{5!}{1!4!}(0.6)^4(0.1)^1 + \frac{5!}{0!5!}(0.6)^5(0.1)^0$$
$$= 0.3370$$

When using the calculator, we will be selecting **binomialcdf** for binomial cumulative density function when we are not interested in an exact number of outcomes but a range of outcomes. Moreover, the corresponding graph for $P(X \geq 4)$ becomes:



Example D.6.16

A manufacturer finds that 30% of the items produced from one of the assembly lines are defective. During a floor inspection, the manufacturer selects 6 items from this assembly line. Find the probability that the manufacturer finds:

a two defectives b at least two defectives.

a Let X denote the number of defectives in the sample of six. Therefore, we have that $n = 6$, $p(success) = p = 0.30$ ($\Rightarrow q = 1 - p = 0.70$), so that $X \sim B(6, 0.3)$.

Note that in this case, a 'success' refers to a defective.

i $P(X = 2) = {}^6C_2(0.3)^2(0.7)^4 = 0.3241$.

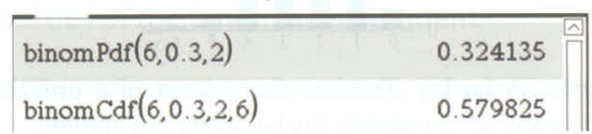
ii $P(X \geq 2) = P(X = 2) + P(X = 3) + \ldots + P(X = 6)$

$= {}^6C_2(0.3)^2(0.7)^4 + {}^6C_3(0.3)^3(0.7)^3 + \ldots + {}^6C_6(0.3)^6(0.7)^0$

A second method makes use of the complementary event:

$$P(X \geq 2) = 1 - P(X < 2) = 1 - P(X \leq 1)$$
$$= 1 - [P(X = 1) + P(X = 0)]$$
$$= 1 - [0.1176 + 0.3025]$$
$$= 0.5798$$

Note: Using the cumulative binomial distribution on a calculator, we enter the range of values as 2 to 6.



Example D.6.17

Sophie has 10 pots labelled one to ten. Each pot, and its contents, is identical in every way. Sophie plants a seed in each pot such that each seed has a germinating probability of 0.8.

a What is the probability that all the seeds will germinate?

b What is the probability that only three seeds will not germinate?

c What is the probability that more than eight seeds do germinate?

d How many pots must Sophie use to be 99.99% sure that at least one seed germinates?

a Let X denote the number of seeds germinating. Therefore we have that $X \sim B(10, 0.8)$,

i.e. X is binomially distributed with parameters $n = 10$ and $p = 0.8$ (and $q = 1 - p = 0.20$).

$$P(X = 10) = \binom{10}{10}(0.8)^{10}(0.2)^{0} = 0.1074.$$

If only three seeds will **not** germinate, then only seven seeds must germinate!

We want, $P(X = 7) = \binom{10}{7}(0.8)^{7}(0.2)^{3} = 0.2013$.

Now, $P(X > 8) = P(X = 9) + P(X = 10)$

$= 0.2684 + 0.1074$

$= 0.3758$

At least one flower means, $X \geq 1$, therefore we need to find a value of n such that $P(X \geq 1) \geq 0.9999$.

$$P(X \geq 1) = 1 - P(X = 0) = 1 - \binom{10}{0}(0.8)^{0}(0.2)^{n} = 1 - (0.2)^{n}$$

Solving for n we have: $1 - (0.2)^{n} \geq 0.9999 \Leftrightarrow (0.2)^{n} \leq 0.0001$

This inequality can be solved by trial and error, algebraically, or: $(0.2)^{n} \leq 0.0001$, $n = 5.72$

Therefore, Sophie would need 6 pots to be 99.99% that at least one seed germinates

Expectation, Mode and Variance for the Binomial Distribution

If the random variable X, is such that $X \sim B(n,p)$, we have:

1. the **expected value of X** is $\mu = E(X) = np$.

2. the **mode of X** is that value of x which has the largest probability

3. the **variance of X** is $\sigma^2 = Var(X) = npq = np(1 - p)$

Notes:

1. Although we can use our earlier definitions of the expected value and the variance of a random variable, the formulae above are in a nice compact form and can only be used when dealing with the binomial distribution.

2. The standard deviation, $Sd(X)$, is still given by $\sigma = \sqrt{Var(X)} = \sqrt{npq}$.

Example D.6.18

A fair die is rolled six times. If X denotes the number of fours obtained, find:

a $E(X)$ b the mode of X c Sd(X)

a In this case we have that $X \sim B\left(6, \frac{1}{6}\right)$, therefore $q = \frac{5}{6}$

$$\mu = E(X) = 6 \times \frac{1}{6} = 1.$$

b To find the mode, we need to know the probability of each outcome. We do this by constructing a table of values:

x	0	1	2	3	4	5	6
$P(X=x)$	$\frac{15625}{46656}$	$\frac{18750}{46656}$	$\frac{9375}{46656}$	$\frac{2500}{46656}$	$\frac{375}{46656}$	$\frac{30}{46656}$	$\frac{1}{46656}$

So that the mode of X is 1 (as it has the highest probability value). Notice in this case, the mode of X = expected value of X. Will this always be true?

c $\sigma = \sqrt{Var(X)} = \sqrt{npq} = \sqrt{6 \times \frac{1}{6} \times \frac{5}{6}} \approx 0.9129$.

Example D.6.19

An urn contains 7 marbles of which 2 are blue. A marble is selected, its colour noted and then replaced in the urn. This process is carried out 50 times. Find;

a The mean number of blue marbles selected

b The standard deviation of the number of blue marbles selected.

Because we replace the marble before the next selection, each trial is identical and independent. Therefore, if we let X

denote the number of blue marbles selected, we have that:

$p = \frac{2}{7}, n = 50$ and $q = \frac{5}{7}$.

a $E(X) = np = 50 \times \frac{2}{7} = 14.29$.

b $Var(X) = npq = 50 \times \frac{2}{7} \times \frac{5}{7} = \frac{500}{49} \therefore \sigma = \sqrt{\frac{500}{49}} \approx 3.19$.

Example D.6.20

The random variable X is such that $E(X) = 8$ and $Var(X) = 4.8$. Find $p(X = 3)$.

This time we are given that $np = 8$ and $npq = np(1 - p) = 4.8$.

Therefore, after substituting $\boldsymbol{np} = 8$ into $\boldsymbol{np}(1 - p) = 4.8$, we have that

$8(1 - p) = 4.8 \quad \therefore (1 - p) = 0.6 \Rightarrow p = 0.4$

Substituting $p = 0.4$ back into $np = 8$, we have that $n = 20$.

Therefore, $P(X = 3) = \binom{20}{3}(0.4)^3(0.6)^{17} = 0.0123$.

A binomial distribution is constructed from a finite number of independent Bernoulli trials. When an infinite number of independent Bernoulli trials are placed together until the first successful outcome, we would have constructed the **geometric distribution**. In other words, if the first success is denoted by $X = k$, then there must be $k - 1$ failures leading to the first success.

$$P(X=k)=(1-p)^{k-1}p, k=1,2,3,...$$

In theory, the number of trials could go on forever, because the first successful outcome may not ever come. However, there must be at least one trial in a geometric distribution. If a random variable follows a geometric distribution with a probability of success, p, then $X \sim G(p)$ (sometimes denoted as $X \sim \text{Geo}(p)$).

Example D.6.21

Given that there is a 12% chance of making a successful call when telemarketing, find the probability for the first successful call after 9 unsuccessful attempts.

We will begin by defining the distribution as $X \sim G(0.12)$.

Therefore: $P(X=10)=0.88^9 \times 0.12$

≈ 0.0380

Notice that the probability of success diminishes quickly as the number of failures increases. For example, if there are 19 unsuccessful telemarketing calls, then $P(X=20) \approx 0.0106$ and if there are 29 unsuccessful telemarketing calls, then $P(X=30) \approx 0.00295$. These results can also be seen graphically.

Exercise D.6.3

1. At an election 40% of the voters favoured the Environment Party. Eight voters were interviewed at random. Find the probability that:

   a exactly 4 voters favoured the Environment Party.

   b a majority of those interviewed favoured the Environment Party.

   c at most 3 of the people interviewed favoured the Environment Party.

2. In the long run, Thomas wins 2 out of every 3 games. If Thomas plays 5 games, find the probability that he will win:

   a exactly 4 games. b at most 4 games.

   c no more than 2 games.

   d all 5 games.

3. A bag consists of 6 white cubes and 10 black cubes. Cubes are withdrawn one at a time with replacement. Find the probability that after 4 draws:

a all the cubes are black.

b there are at least 2 white cubes.

c there are at least 2 white cubes given that there was at least one white cube.

4. An X-ray has a probability of 0.95 of showing a fracture in the leg. If 5 different X-rays are taken of a particular leg, find the probability that:

a all five X-rays identify the fracture.

b the fracture does not show up.

c at least 3 X-rays show the fracture.

d only one X-ray shows the fracture.

5. A biased die, in which the probability of a '2' turning up is 0.4, is rolled 8 times.

Find the probability that:

a a '2' turns up 3 times.

b a '2' turns up on at least 4 occasions.

6. During an election campaign, 66% of a population of voters are in favour of a food quality control proposal. A sample of 7 voters was chosen at random from this population.

Find the probability that:

a there will be 4 voters that were in favour.

b there will be at least 2 voters who were in favour.

7. During an election 35% of the people in a town favoured the fishing restrictions at Lake Watanaki. Eight people were randomly selected from the town. Find the probability that:

a 3 people favoured fishing restrictions.

b at most 3 of the 8 favoured fishing restrictions.

c there was a majority in favour of fishing restrictions.

8. A bag containing 3 white balls and 5 black balls has 4 balls withdrawn one at a time, in such a way that the first ball is replaced before the next one is drawn. Find the probability of:

a selecting 3 white balls.

b selecting at most 2 white balls.

c selecting a white ball, two black balls and a white ball in that order.

d selecting two white balls and two black balls.

9. A tennis player finds that he wins 3 out of 7 games he plays. If he plays 7 games straight, find the probability that he will win:

a exactly 3 games.

b at most 3 games.

c all 7 games.

d no more than 5 games.

e After playing 30 games, how many of these would he expect to win?

10. A true–false test consists of 8 questions. A student will sit for the test, but will only be able to guess at each of the answers. Find the probability that the student answers:

a all 8 questions correctly.

b 4 questions correctly.

c at most 4 of the questions correctly.

The following week, the same student will sit another true–false test, this time there will be 12 questions on the test, of which he knows the answer to 4.

e What are the chances of passing this test (assuming that 50% is a pass)?

11. The births of males and females are assumed to be equally likely. Find the probability that in a family of 6 children:

a there are exactly 3 girls.

b there are no girls.

c the girls are in the majority.

d How many girls would you expect to see in a family of 6 children?

12. During any one production cycle it is found that 12% of items produced by a manufacturer are defective. A sample of 10 items is selected at random and inspected. Find the probability that:

a no defectives will be found.

b at least two defectives will be found.

c A batch of 1000 such items are now inspected. How many of these items would you expect to be defective?

13. Ten per cent of washers produced by a machine are considered to be either oversized or undersized. A sample of 8 washers is randomly selected for inspection.

a What is the probability that there are 3 defective washers?

b What is the probability that there is at least one defective washer?

14. Over a long period of time, an archer finds that she is successful on 90% of her attempts. In the final round of a competition she has 8 attempts at a target.

a Find the probability that she is successful on all 8 attempts.

b Find the probability that she is successful on at least 6 attempts.

The prize that is awarded is directly proportional to the number of times she is successful, earning 100 fold, in dollars, the number of times she is successful.

c What can she expect her winnings to be after one round?

She draws with another competitor. However, as there can be only one winner, a second challenge is put into place – they must participate in another 3 rounds, with 5 attempts in each round.

d Find the probability that she manages 3 perfect rounds.

15. For each of the random variables:

a $X \sim B(7, 0.2)$

b $X \sim B(8, 0.38)$

Find: i the mean ii the mode

iii the standard deviation

iv $P(X \geq 6 | X > 4)$ v $P(X > 4 | X \leq 6)$

Extra questions

The Normal Distribution?

The examples considered in previous sections mainly dealt with data that was discrete. Discrete data is generally counted and can be found exactly. Discrete data is often made up of whole numbers. For example, we may have counted the number of occupants in each of the cars passing a particular point over a period of two hours. In this case the data is made up of whole numbers. If we collect information on the European standard shoe sizes of a group of people, we will also be collecting discrete data even though some of the data will be fractional: shoe size nine and a half.

Alternatively, sometimes we collect data using measurement. For example, we may collect the birth weights of all the babies delivered at a maternity hospital over a year. Because weight is a continuous quantity (all weights are possible, not just whole numbers or certain fractions), the data collected is continuous. This remains the case even though we usually round continuous data to certain values. In the case of weight, we may round the data to the nearest tenth of a kilogram. In this case, if a baby's weight is given as 3.7 kg it means that the weight has been rounded to this figure and lies in the interval [3.65,3.75). If we are looking at data such as these weights it may seem as if the data is discrete even in cases where it is in fact continuous.

When dealing with continuous data, we use different methods. The most important distinction is that we can never give the number of babies that weigh *exactly* 3.7kg as there may be *none* of these. All that we can give is the number of babies born that have weights in the range [3.65,3.75).

One of the ways in which we can handle continuous data is to use the normal distribution. This distribution is only a model for real data. This means that its predictions are only approximate. The normal distribution generally works best in a situation in which the data clusters about a particular mean and varies from this as a result of random factors. The birth weights of babies cluster about a mean with variations from this mean resulting from a range of chance factors such as genetics, nutrition etc. The variation from the mean is measured by the standard deviation of the data. In examples such as this, the normal distribution is often a fairly good model. The basis of all normal distribution studies is the standard normal curve.

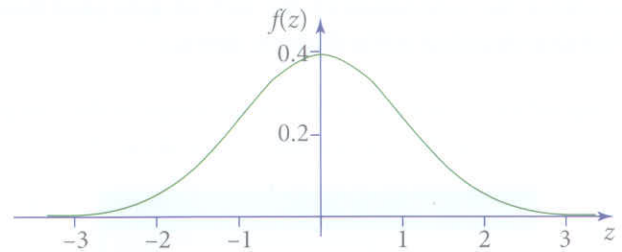
The Standard Normal Curve

The standard normal curve models data that has a mean of zero and a standard deviation of one. The equation of the standard normal curve is:

$$f(z) = \frac{1}{\sqrt{2\pi}}e^{\left(-\frac{z^2}{2}\right)}, -\infty < z < \infty$$

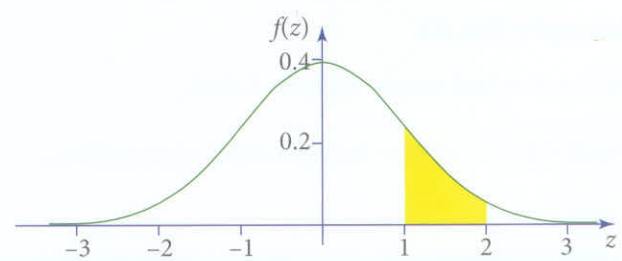
The equation of this distribution is complex and does not directly give us any information about the distribution. The shape of the curve does, however, indicate the general shape of the distribution.

The shape of this curve is often referred to as the 'bell curve'. On the next page we see how this function behaves.



As a result of the fact that the variable z is continuous, it is not the height of the curve but the areas underneath the curve that represent the proportions of the variable that lie between various values. The total area under the curve is 1 (even though the curve extends to infinity in both directions without actually reaching the axis).

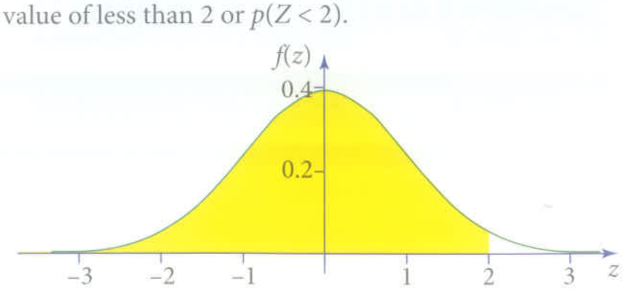
For example, the proportion of the standard normal data that lies between 1 and 2 is represented by the area shown.



Areas under curves are usually found using a method covered in Integral Calculus. In the case of the normal curve, the complexity of the equation of the graph makes this impossible at least at this level. Instead, we rely on graphics calculators.

Using a Calculator

The diagram shows the area that represents the proportion of values for which $z < 2$. This proportion can also be interpreted as the probability that a randomly chosen value of z will have a value of less than 2 or $p(Z < 2)$.



The area to the left is of infinite extent and yet the area is finite.

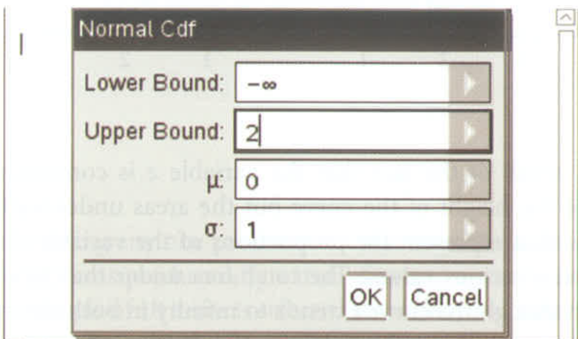
The area is found using a calculator (much as we find the trigonometric ratios etc.

The entries to solve this problem follow the same pattern used with the other probability distributions. If using at TI-NSpire:

MENU / 5. Probability / 5. Distributions / 2. Normal CDF

This allows the calculation of the sort of area (and hence probability) depicted in the diagram above.

The dialog box allows you to fill out the range of the variable $(-\infty,2]$ and the mean (0) and standard deviation (1).



normCdf(-∞,2,0,1)	0.97725

Example D.6.22

For the standard normal variable Z, find:

a $p(Z < 1)$ b $p(Z < 0.96)$

c $p(Z < 0.03)$.

All these examples can be solved by direct use of a calculator:

$p(Z < 1) = 0.8413$ $p(Z < 0.96)$ $(= 0.8315)$

$p(Z < 0.03)$ $(= 0.5120)$

normCdf(-∞,1,0,1)	0.841345
normCdf(-∞,0.96,0,1)	0.831472
normCdf(-∞,0.03,0,1)	0.511967

Example D.6.23

For the standard normal variable Z, find:

a $p(Z > 1.7)$ b $p(Z > -0.88)$

c $p(Z > -1.53)$.

normCdf(1.7,∞,0,1)	0.044565
normCdf(-0.88,∞,0,1)	0.81057
normCdf(-1.53,∞,0,1)	0.936992

Example D.6.24

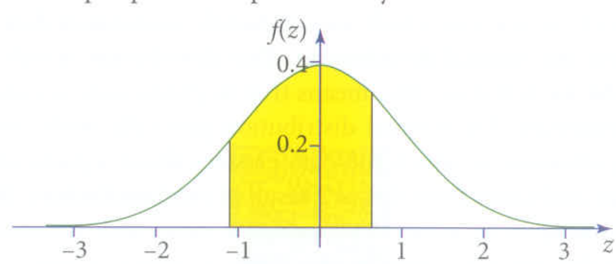
For the standard normal variable Z, find:

a $p(1.7 < Z < 2.5)$ b $(-1.12 < Z < 0.67)$

c $p(-2.45 < Z < -0.08)$.

It is as well when using technology to answer questions of this sort to have an estimate of the correct answer in mind.

For example, part b is represented by this area:



The whole are under the curve is 1, the shaded area looks to be a bit more than half this, so an answer a bit over 0.5 is to be expected. This is confirmed by the calculator.

normCdf(1.7,2.5,0,1)	0.038356
normCdf(-1.12,0.67,0,1)	0.617214
normCdf(-2.45,-0.08,0,1)	0.460976

Exercise D.6.4

1. For the standard normal variable Z, find:

a	$p(Z < 0.5)$	b	$p(Z < 1.84)$
c	$p(Z < 1.62)$	d	$p(-2.7 < Z)$
e	$p(-1.97 < Z)$	f	$p(Z < -2.55)$

2. For the standard normal variable Z, find:

   a $p(1.75 < Z < 2.65)$

   b $p(0.3 < Z < 2.5)$

   c $p(1.35 < Z < 1.94)$

   d $p(-1.92 < Z < -1.38)$

   e $p(2.23 < Z < 2.92)$

The Normal Distribution

Standardizing any normal distribution

Very few practical applications will have data whose mean is 0 and whose standard deviation is 1. The standard normal curve is, therefore, not directly usable in most cases. We overcome this difficulty by relating every problem to the standard normal curve.

As we have already seen, a general variable, X, is related to the standard normal variable, Z, using the relation:

$$Z = \frac{X-\mu}{\sigma}$$

where μ = the mean of the data and σ is the standard deviation. We use an example to illustrate this.

Example D.6.25

A assembly line produces bags of sugar with a mean weight of 1.01 kg and a standard deviation of 0.02 kg:

a Find the proportion of the bags that weigh less than 1.03 kg.

b Find the proportion of the bags that weigh more than 1.02 kg.

c Find the percentage of the bags that weigh between 1.00 kg and 1.05 kg.

When approaching these problems, it is important to estimate answers. In the case of non-standard normal distributions, it is best to think graphically.

This involves relating the distribution you have been given to the standard normal curve. The former is centred on the mean and spreads three standard deviations either side of this.

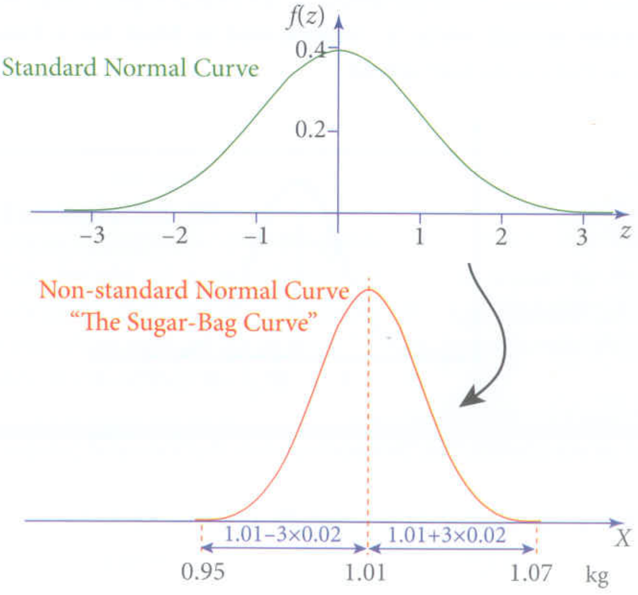
The weight variable, X, is related to the standard normal variable, Z, using the relation:

$$Z = \frac{X-\mu}{\sigma}$$

where μ = the mean of the data and σ is the standard deviation.

In this case: $Z = \dfrac{X-1.01}{0.02}$.

The curve centres on 1.01 kg and spreads $3\times0.02 = 0.06$ ether side of this mean (i.e 0.95 to 1.07 kg).



The answers to the three parts are:

normCdf(-∞,1.03,1.01,0.02)	0.841345
normCdf(1.02,∞,1.01,0.02)	0.308538
normCdf(1,1.05,1.01,0.02)	0.668712

Rule of Thumb

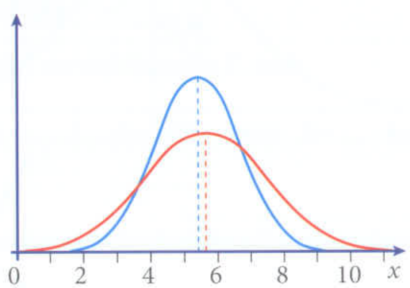
For normally distributed variables, about two thirds (68.27%) of the values lie within one standard deviation of the mean, roughly 95% (95.45%) lie within two standard deviations, and more than 99% (99.73) lie within three standard deviations.

Another important use of the standard normal distribution is to compare two sets of related continuous data, each of which has a different mean and standard deviation. In theory, each set of continuous data with its unique mean and standard deviation gives a different normal distribution. In order to analyze two sets of data and to draw meaningful conclusions, we need to standardize both sets of data.

Example D.6.26

Given that $X_1 \sim N(5.64, 1.85^2)$ and $X_2 \sim N(5.35, 1.27^2)$ compare the distributions and determine which has a higher likelihood of producing a value of x equal to or higher than 6.

When we examine the means and the standard deviations, it may not be obvious which distribution yields the bigger result. The corresponding graph is not particularly useful either, because X_1 (represented in red) has a higher mean but a wider spread, while X_1 (represented in blue) has a lower mean but a narrower spread.

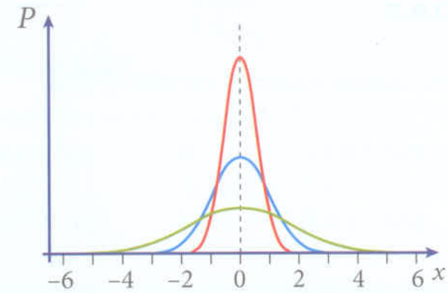


Let us standardize the two normal distributions and compare:

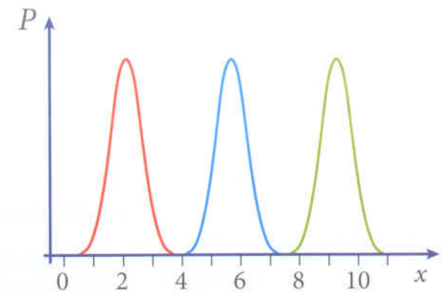
$$P(X_1 \ge 6) = P(Z_1 \ge 0.195) = 0.423 \quad \text{and} \quad P(X_2 \ge 6) = P(Z_2 \ge 0.512) = 0.304$$

Hence, the first distribution gives the higher chance.

We will now consider the effects on the distribution caused by changing the parameters. If we keep the mean constant while changing the standard deviation, we will notice that the location where the symmetry appears will not move, but the width of the distribution changes accordingly.



Conversely, if we keep the standard deviation constant but instead we change the mean, we will get the same distribution but with a horizontal translation corresponding to the new mean.



The normal distribution, as a continuous distribution, indeed provides a close approximation to the binomial distribution when n, the number of trials, is very large and p, the probability of success, is close to 0.5.

Example D.6.27

a. Find the probability of getting less than 20 heads in 30 flips of a fair coin.

b. Use a normal distribution to approximate the binomial distribution in part [a].

a. Let us begin by defining the binomial distribution, $X \sim B(30, 0.5)$, $P(X \le 20) \approx 0.979$

b. We need to determine μ and σ.

$\mu = 30 \times 0.5 = 15$ and $\sigma = \sqrt{30 \times 0.5 \times 0.5} \approx 2.739$

Therefore, $X \sim N(15, 2.739)$

$P(X \le 20) \approx 0.966$

Note, the error of the approximation is about 0.013.

Inverse Problems

There are occasions when we are told the proportion of the data that we are to consider and asked questions about the data conditions that are appropriate to these proportions.

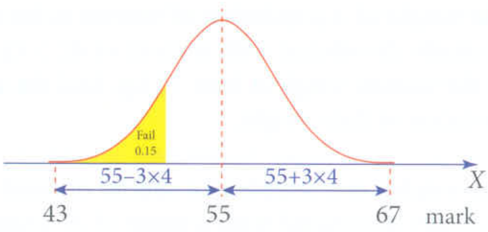
Example D.6.28

The Board of Examiners have decided that 85% of all candidates sitting Mathematical Methods will obtain a pass grade in the examination. The actual examination marks are found to be normally distributed with a mean of 55 and a variance of 16. What is the lowest score a student can get on the exam to be awarded a pass grade?

This requires used of the inverse normal (option 3 on the TI) distribution.

Again, an estimate is useful. The mean is 55 and the standard deviation (the square root of the variance) is 4.

Thus almost all the students can be expected to score between 55–3×4=43 and 55+3×4=67.



On the basis of the graph, an answer of around 50 marks would seem reasonable:

```
invNorm(0.15,55,4)          50.8543
```

Therefore a student needs to score at least 51 marks to pass the exam.

Example D.6.29

The lifetime of a particular make of television tube is normally distributed with a mean of 8 years, and a standard deviation of σ years. The chances that the tube will not last 5 years is 0.05. What is the value of the standard deviation?

Let X denote the life-time of the television tubes, so that $X \sim N(8, \sigma^2)$.

Given that $p(X < 5) = 0.05 \Rightarrow p\left(Z < \frac{5-8}{\sigma}\right) = 0.05$.

That is, we have that:

$$p\left(Z < -\frac{3}{\sigma}\right) = 0.05 \quad \Leftrightarrow -\frac{3}{\sigma} = -1.6449$$
$$\Leftrightarrow \sigma = 1.8238$$

And so the standard deviation is approximately one year and 10 months.

Example D.6.30

The weight of a population of men is found to be normally distributed with mean 69.5 kg. Thirteen per cent of the men weigh at least 72.1 kg, find the standard deviation of their weight.

Let the random variable X denote the weight of the men, so that $X \sim N(69.5, \sigma^2)$.

We then have that $p(X \geq 72.1) = 0.13$ or $p(X \leq 72.1) = 0.87$.

$$\therefore p\left(Z \leq \frac{72.1 - 69.5}{\sigma}\right) = 0.87 \Leftrightarrow \frac{72.1 - 69.5}{\sigma} = 1.1264$$

Exercise D.6.5

1. If Z is a standard normal random variable, find:

   a $p(Z > 2)$ b $p(Z < 1.5)$

2. If Z is a standard normal random variable, find:

   a $p(Z > -2)$ b $p(Z < -1.5)$

3. If Z is a standard normal random variable, find:

   a $p(0 \le Z \le 1)$ b $p(1 \le Z \le 2)$

4. If Z is a standard normal random variable, find:

   a $p(-1 \le Z \le 1)$ b $p(-2 \le Z \le -1)$

5. If X is a normal random variable with mean $\mu = 8$ and variance $\sigma^2 = 4$, find:

   a $p(X \ge 6)$ b $p(5 < X \le 8)$

6. If X is a normal random variable with mean $\mu = 100$ and variance $\sigma^2 = 25$, find:

   a $p(X \ge 106)$ b $p(105 < X \le 108)$

7. If X is a normal random variable with mean $\mu = 60$ and standard deviation $\sigma = 5$, find:

   a $p(X \ge 65)$ b $p(55 < X \ge 65)$

8. Scores on a test are normally distributed with a mean of 68 and a standard deviation of 8. Find the probability that a student scored:

   a at least 75 on the test

   b at least 75 on the test given that the student scored at least 70

   c In a group of 50 students, how many students would you expect to score between 65 and 72 on the test.

9. If X is a normally distributed variable with a mean of 24 and standard deviation of 2, find:

   a $p(X > 28 | X \ge 26)$

   b $p(26 < X < 28 | X \ge 27)$

10. The heights of men are normally distributed with a mean of 174 cm and a standard deviation of 6 cm. Find the probability that a man selected at random:

    a is at least 170 cm tall

    b is no taller than 180 cm

    c is at least 178 cm given that he is at least 174 cm.

11. If X is a normal random variable with a mean of 8 and a standard deviation of 1, find the value of c, such that:

    a $p(X > c) = 0.90$ b $p(X \le c) = 0.60$

12. If X is a normal random variable with a mean of 50 and a standard deviation of 5, find the value of c, such that:

    a $p(X \le c) = 0.95$ b $p(X \ge c) = 0.95$

    c $p(-c \le X \le c) = 0.95$

13. The Board of Examiners has decided that 80% of all candidates sitting the Mathematical Methods Exam will obtain a pass grade. The actual examination marks are found to be normally distributed with a mean of 45 and a standard deviation of 7. What is the lowest score a student can get on the exam to be awarded a pass grade?

14. The weight of a population of women is found to be normally distributed with a mean of 62.5 kg. If 15% of the women weigh at least 72 kg, find the standard deviation of their weight.

15. The weights of a sample of a species of small fish are normally distributed with a mean of 37 grams and a standard deviation of 3.8 grams. Find the percentage of fish that weigh between 34.73 and 38.93 grams. Give your answer to the nearest whole number.

16. The weights of the bars of chocolate produced by a machine are normally distributed with a mean of 232 grams and a standard deviation of 3.6 grams. Find the proportion of the bars that could be expected to weigh less than 233.91 grams.

17. For a normal variable, X, $\mu = 196$ and $\sigma = 4.2$. Find:

    a $p(X < 193.68)$ b $p(X > 196.44)$

18. The circumferences of a sample of drive belts produced by a machine are normally distributed with a mean of 292 cm and a standard deviation of 3.3 cm. Find the

percentage of the belts that have diameters between 291.69 cm and 293.67 cm.

19. A normally distributed variable, X, has a mean of 52. $p(X < 51.15) = 0.0446$. Find the standard deviation of X.

20. The lengths of the drive rods produced by a small engineering company are normally distributed with a mean of 118 cm and a standard deviation of 0.3 cm. Rods that have a length of more than 118.37 cm are rejected. Find the percentage of the rods that are rejected. Give your answer to the nearest whole number.

21. After their manufacture, the engines produced for a make of lawn mower are filled with oil by a machine that delivers an average of 270 mL of oil with a standard deviation of 0.7 mL.

    Assuming that the amounts of oil delivered are normally distributed, find the percentage of the engines that receive more than 271.12 mL of oil. Give your answer to the nearest whole number.

22. A sample of detergent boxes have a mean contents of 234 grams with a standard deviation of 4.6 grams. Find the percentage of the boxes that could be expected to contain between 232.22 and 233.87 grams. Give your answer to the nearest whole number.

23. A normally distributed variable, X, has a mean of 259. $p(X < 261.51) = 0.9184$. Find the standard deviation of X.

24. A normally distributed variable, X, has a standard deviation of 3.9. Also, 71.37% of the data are larger than 249.8. Find the mean of X.

25. The times taken by Maisie on her way to work are normally distributed with a mean of 26 minutes and a standard deviation of 2.3 minutes. Find the proportion of the days on which Maisie's trip takes longer than 28 minutes and 22 seconds.

26. In an experiment to determine the value of a physical constant, 100 measurements of the constant were made. The mean of these results was 138 and the standard deviation was 0.1. What is the probability that a final measurement of the constant will lie in the range 138.03 to 139.05?

27. In an experiment to determine the times that production workers take to assemble an electronic testing unit, the times had a mean of 322 minutes and a standard deviation of 2.6 minutes. Find the proportion of units that will take longer than 324 minutes to assemble. Answer to two significant figures.

28. A normally distributed variable, X, has a standard deviation of 2.6. $p(X < 322.68) = 0.6032$. Find the mean of X.

29. The errors in an experiment to determine the temperature at which a chemical catalyst is at its most effective, were normally distributed with a mean of 274°C and a standard deviation of 1.2°C. If the experiment is repeated what is the probability that the result will be between 275°C and 276°C?

30. The weights of ball bearings produced by an engineering process have a mean of 215 g with a standard deviation of 0.1 g. Any bearing with a weight of 215.32 g or more is rejected. The bearings are shipped in crates of 10 000. Find the number of bearings that may be expected to be rejected per crate.

31. If $X \sim N(\mu, 12.96)$ and $p(85.30 < X) = 0.6816$, find μ to the nearest integer.

32. At a Junior track and field meet it is found that the times taken for children aged 14 to sprint the 100 metres race are normally distributed with a mean of 15.6 seconds and standard deviation of 0.24 seconds. Find the probability that the time taken for a 14 year old at the meet to sprint the 100 metres is:

    i less than 15 seconds.

    ii at least 16 seconds.

    iii between 15 and 16 seconds.

    On one of the qualifying events, eight children are racing. What is the probability that six of them will take between 15 and 16 seconds to sprint the 100 metres?

33. Rods are manufactured to measure 8 cm. Experience shows that these rods are normally distributed with a mean length of 8.02 cm and a standard deviation of 0.04 cm.

    Each rod costs $5.00 to make and is sold immediately if its length lies between 8.00 cm and 8.04 cm. If its length exceeds 8.04 cm it costs an extra $1.50 to reduce its length to 8.02 cm. If its length is less than 8.00 cm it is sold as scrap metal for $1.00.

a What is the average cost per rod?

b What is the average cost per usable rod?

34. The resistance of heating elements produced is normally distributed with a mean of 50 ohms and standard deviation 4 ohms.

a Find the probability that a randomly selected element has resistance less than 40 ohms.

bi If specifications require that acceptable elements have a resistance between 45 and 55 ohms, find the probability that a randomly selected element satisfies these specifications.

ii A batch containing 10 such elements is tested. What is the probability that exactly 5 of the elements satisfy the specifications?

c The profit on an acceptable element, i.e. one that satisfies the specifications, is \$2.00, while unacceptable elements result in a loss of \$0.50 per element. If \$ P is the profit on a randomly selected element, find the profit made after producing 1000 elements.

35. Given that $X \sim N\left(\mu, \sigma^2\right)$ and $Y = aX + b$, where $a < 0$.

Show that $Y \sim N\left(a\mu + b, a^2\sigma^2\right)$.

36. If $X \sim N\left(\mu, \sigma^2\right)$, show that:

$P\left(|X - \mu| \le 0.675\sigma\right) = 0.5$.

37. Find the mean and the standard deviation of a normally distributed random variable X if $P(X \le 20) = 0.3$ and $P(X \ge 50) = 0.2$.

38. Given that $X \sim N(10, 25)$. Find Q_1 and Q_3.

39. Given that $\mu = 6.62$ and $Q_3 = 13.38$ and is a normal distribution. Find the standard deviation and Q_1.

40. Kelly scored 76% in Biology where the class mean was 71% and the standard deviation was 11.2%. In Mathematics she scored 69%, the class mean 65% was and the standard deviation was 7.8%. Assume that the scores for both subjects were normally distributed. In which subject did Kelly perform better when compared with the rest of her class?

41. The amount of time an IB student spends on SNS per day is a normally distributed random variable. It is known that 65% of the students spend more than 4.9 hours per day, and 25% of the students spend less than 4.2 hours. Find the mean and the standard deviation of the amount of time an IB student spends on SNS per day.

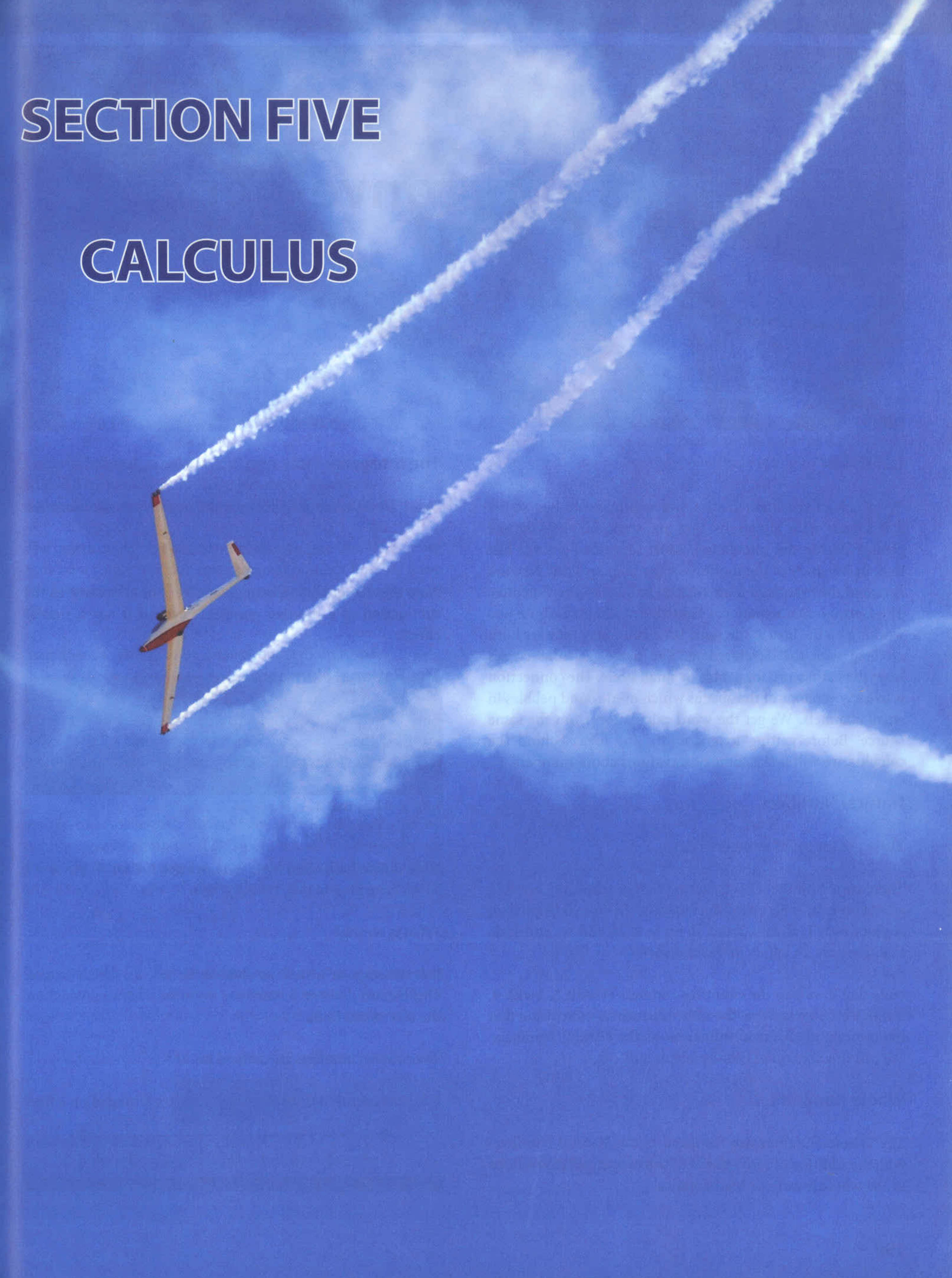
42. An airline knows from experience that only 85% of the travellers who make reservations actually show up for the flight. The airline decides for one flight to overbook and accept reservations when only 90 seats are available. What is the probability that more than 90 travellers will show up?

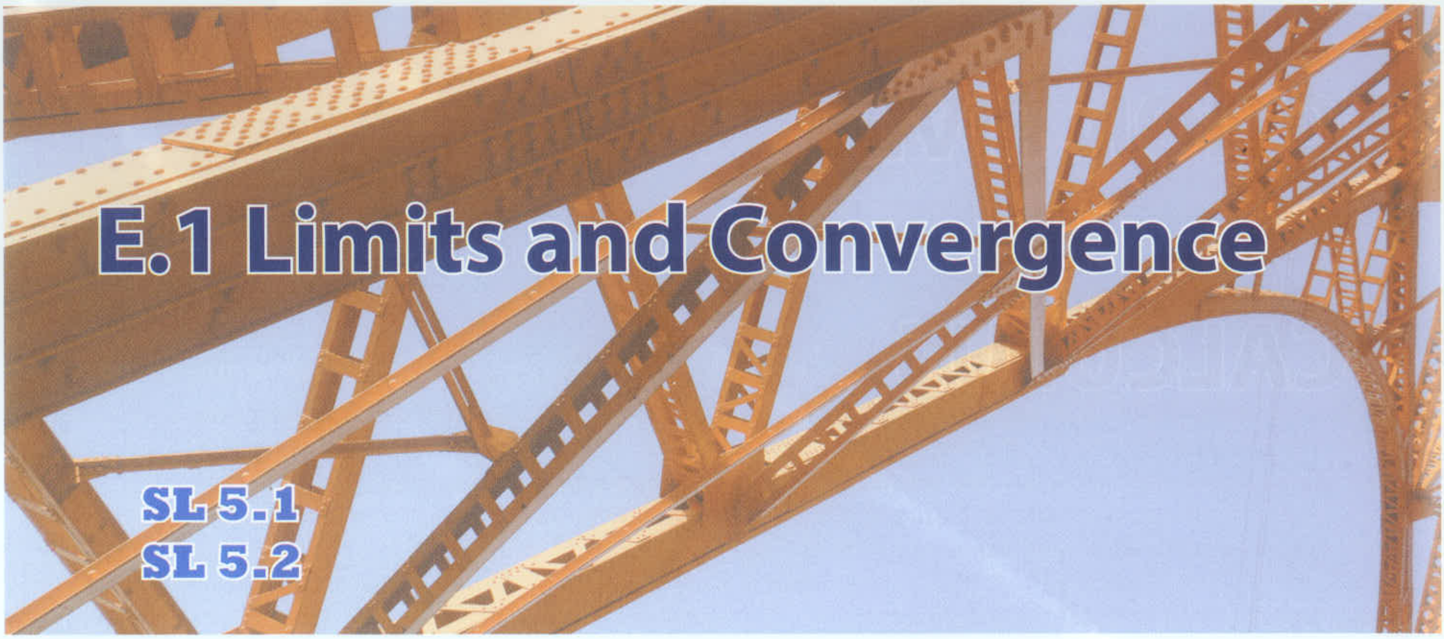
Extra questions

Answers

SECTION FIVE

CALCULUS





E.1 Limits and Convergence

SL 5.1
SL 5.2

Limits

Much of the mathematics you will have studied so far deals with situations that are static ie. they do not change during the calculation. Most situations are not like that. In his study of the motions of the planets, Isaac Newton invented the branch of Mathematics known as **The Calculus**. The German mathematician Gottfried Wilhelm Leibniz was working on the idea at the same time and today they are both credited with the development. The name Calculus comes from the Latin word for pebble or small stone. The connection is to early versions of the abacus which used round pebbles in small channels. We get the word calculation from the same source. Before getting to grips with the ideas behind the calculus, we will make some observations about numbers.

Natural Numbers

These are the set of numbers 1,2,3,4,...

There are an infinite set of these and they are used for counting sets of objects. If we count the letters of the Roman Alphabet, A pairs with 1, B with 2 and so on to Z 26 and we conclude that there are 26 letters in the alphabet.

Note that if we pair the counting numbers (1 with 2, 2 with 4, 3 with 6 etc.) we arrive at the rather disturbing conclusion that the even numbers are as numerous as the Natural Numbers. We call this version of infinity 'aleph null - $\aleph_0$ '.

Whole numbers

The Whole Numbers are the same as the Natural Numbers with the addition of zero - one of the most original innovations in the whole history of Mathematics.

The Integers

This ($\mathbb{Z}$) is the set of all positive and negative whole numbers:

$$...-4,-3,-2,-1,0,1,2,3,4...$$

Note that the quality of being negative is not the same as the subtraction of a positive number - even if it has a similar effect.

The Rational Numbers

This is the set of all ratios (ie. fractions) made from the integers. There is an exception that we do not allow division by zero as this leads to this sort of unacceptable contradiction.

Let: $x=1$ so that $x-1=x^2-1$

$$x-1=(x+1)(x-1)$$

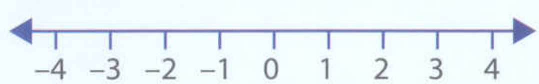
If we divide both sides by $(x - 1)$, we get $1 = x + 1$. But $x = 1$ so we have proved that $1 = 2$.

What is wrong?

It is the stage at which we divided by $(x - 1)$, which is zero. This ban on division by zero has important consequences for the rest of this book.

The rational numbers are defined as: $\mathbb{Q}=\left\{\frac{p}{q}, q\neq 0, p,q\in\mathbb{Z}\right\}$

It is now useful to think of these numbers arranged on a line.



Are there any gaps in this line?

It looks like the only answer to this is that there are no gaps as the finite space [0,1] contains an infinite number of fractions of the form $^1/_2, ^1/_3, ^1/_4, \ldots$.

Note that all decimals both terminating and recurring can be written as fractions.

However, there are gaps even allowing for the fact that we have an infinite number of points squashed into a finite space.

When these gaps are filled we get the Real Numbers.

The Real Numbers

Depending on which course you follow, you may see it proved that $\sqrt{2}$ cannot be written as a fraction. It is said to be **irrational**. The decimal version of $\sqrt{2}$ neither terminates nor recurs - and this is a characteristic of irrational numbers.

Other irrational numbers are even more arcane than the surds. An example is π. They are more numerous than the rationals and are termed uncountable. There are $\aleph_1$ of them.

The purpose of this discussion is to point out that, if we remove a single number from the real number line, we leave a hole of no 'size' and which we can approach as closely as we like!

So is there anything to be done about zero divisors?

Example E.1.1

Investigate the behaviour of $f(x)=\frac{x^2-4}{x-2}, x\neq 2$ in the region of $x = 2$.

As we have observed, division by zero is forbidden so we cannot simply compute $f(2)$. Because the real numbers are infinitely densely packed, we can take values that get as close to 2 as we like. As bit of algebra helps:

$$f(x)=\frac{x^2-4}{x-2}=\frac{(x+2)(x-2)}{x-2}=x+2, x\neq 2$$

It is a bit easier to say what happens to $x + 2$ as x approaches 2 than it would be using the original function. The closer x approaches 2, the closer $x + 2$ approaches 4. We say:

"The limit as $x\to 2$ of $f(x) = 4$" or

$$\lim_{x\to 2} f(x)=4$$

Remember, we cannot actually use 2!

Example E.1.2

Express 0.9 recurring as a rational number.

This involves another limiting process. Every time we add another 9 to the number, it gets bigger. As we are adding an infinite number of 9s, we might expect an infinite answer.

$$0.9 < 0.99 < 0.999 < 0.999 < \ldots$$

This is, however not so as we will now prove:

Let: $x = 0.99999999\ldots$ [1]

$10x = 9.9999999\ldots$ [2]

Now subtract [1] from [2]. On the left hand side we get $9x$. On the right hand side, all the places to the right of the decimal point are $9 - 9 = 0$. The right hand side is 9. It follows that 0.9 recurring is EQUAL TO 1 (not nearly equal to).

Exercise E.1.1

1. Find: $\lim_{x\to -2} f(x)=x^2$

2. Find: $\lim_{x\to 1} f(x)=\frac{x^2-1}{x-1}, x\neq 1$

3. Find: $\lim_{x\to 1} f(x)=\frac{x^4-1}{x-1}, x\neq 1$

4. Find: $\lim_{x\to 3} f(x)=\frac{x^2-9}{x-3}, x\neq 3$

5. Find: $\lim_{x\to 0} f(x)=\frac{x^2}{x}, x\neq 0$

6. Express as rational numbers:

   i $0.\dot{5}$ ii $0.\dot{7}$

   iii $0.\dot{4}2857\dot{1}$ iv $0.\dot{4}\dot{5}$

7. Find: $\lim_{x\to 0} f(x)=x^x, x\neq 0$

8. Find: $\lim_{x\to \infty} f(x)=\left(1+\frac{1}{x}\right)^x$

We now move on to how limits can help us with the problem of finding the slope of a curve.

Rates of Change

The divers in our picture are breathing compressed air. As a result, their bodies are taking up nitrogen.

The rate at which the nitrogen enters and leaves their tissues determines whether or not they are at risk from 'the bends'. In this and many other applications, it is the rate at which things happen that matters. This is the subject of this Chapter.

Functional dependence

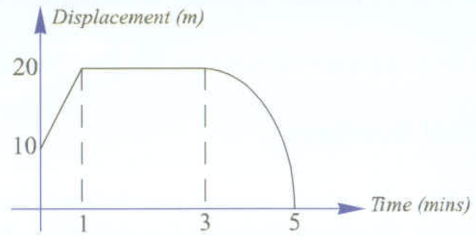
The notion of functional dependence of a function $f(x)$ on the variable x has been dealt with earlier. However, apart from this algebraic representation, sometimes it is desirable to create a graphical representation using a qualitative rather than quantitative description. In doing so, there are a number of key words that are often used.

Words to be kept in mind are:

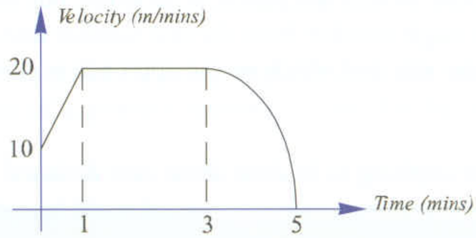
Rate of change (slow, fast, zero)	Increasing, decreasing
Positive, negative	Maximum, minimum
Average	Instantaneous
Stationary	Initial, final
Continuous, discontinuous	Range, domain

Such terms enable us to describe many situations that are presented in graphical form. There is one crucial point to be careful of when describing the graphical representation of a given situation. Graphs that look identical could very well be describing completely different scenarios. Not only must you consider the behaviour (shape) of the graph itself, but also take into account the variables involved.

Consider the two graphs below. Although identical in form, they tell two completely different stories. We describe what happens in the first five minutes of motion:



An object is moving in such a way that its displacement is increasing at a constant rate, that is, the object maintains a constant velocity (or zero acceleration) for the first minute. During the next two minutes the object remains stationary, that is, it maintains its displacement of 20 metres (meaning that it doesn't move any further from its starting position). Finally the particle returns to the origin.



An object is moving at 10 m/min and keeps increasing its velocity at a constant rate until it reaches a velocity of 20 m/min, that is, it maintains a constant acceleration for the first minute. During the next two minutes the object is moving at a constant velocity of 20 m/mins (meaning that it is moving further away from its starting position). Finally, the particle slows to rest, far from the origin.

Although the shape of both graphs is identical, two completely different situations have been described!

Quantitative aspects of change

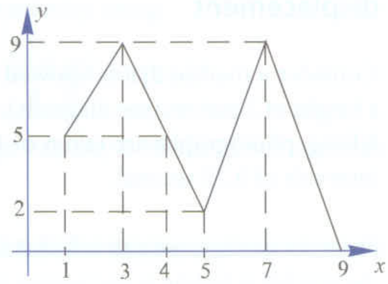
When dealing with the issue of rates of change, we need to consider two types of rates:

1. the average rate of change and
2. the instantaneous rate of change.

We start by considering the first of these terms, the average rate, and then see how the second, the instantaneous rate, is related to the first.

Average rate of change

The average rate of change can be best described as an 'overview' of the effect that one variable (the independent variable) has on a second variable (the dependent variable). Consider this graph.



We can describe the change in the y-values (relative to the change in the x-values) as follows:

For $x \in [1,3]$:

There is a **constant** increase from $y = 5$ to $y = 9$ as x increases from 1 to 3.

An increase of 2 units in x has produced an increase of 4 units in y.

We say that the **average rate** of change of y with respect to x is $\frac{4}{2} = 2$.

For $x \in [1,4]$:

This time, the overall change in y is 0. That is, although y increases from 5 to 9, it then decreases back to 5. So from its initial value of 5, because it is still at 5 as x increases from 1 to 4, the overall change in y is 0. This time the average rate of change is $\frac{0}{3} = 0$.

For: $x \in [1,5]$:

As x now increases from 1 to 5 we observe that there is an overall decrease in the value of y, i.e. there is an overall decrease of 3 units (y: $5 \to 9 \to 5 \to 2$).

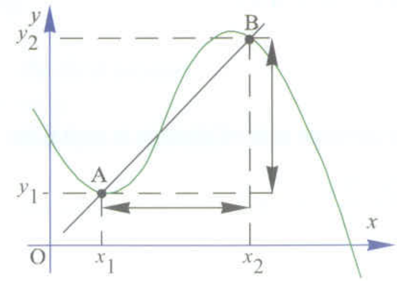
In this instance we say that the average rate of change is $-\frac{3}{4} = -0.75$.

Notice that we have included a negative sign to indicate that there was an overall decrease in the y-values (as x has increased by 4). Similarly for the rest of the graph. Note that we need not start at $x = 1$. We could just as easily have found the change in y for $x \in [3, 5]$. Here, the average rate of change is $-\frac{7}{2} = -3.5$.

The question then remains, is there a simple way to find these average rates of change and will it work for the case where we have non-linear sections? As we shall see in the next sections, the answer is 'yes'.

Determining the average rate of change

To find the average rate of change in y it is necessary to have an initial point and an end point, as x increases from x_1 to x_2.



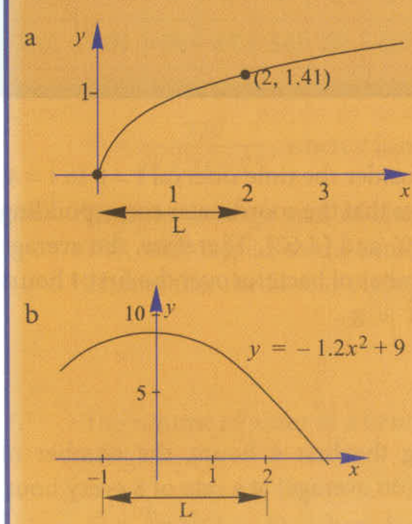
At A $x = x_1, y = y_1$ and at B $x = x_2, y = y_2$.

To obtain a numerical value, we find the gradient of the straight line joining these two points.
Average rate of change from A to B = gradient from A to B

$$= \frac{y_2 - y_1}{x_2 - x_1}$$

Example E.1.3

For each of the graphs below, find the average rate of change of y with respect to x over the interval specified (i.e. over the domain L).



a For this case we have the 'starting point' at the origin (with coordinates (0, 0)) and the 'end' point with coordinates (2, 1.41).

This means that the average rate of change of y with respect to x, over the domain L is given by:

$$\frac{y_2 - y_1}{x_2 - x_1} = \frac{1.41 - 0}{2 - 0} = 0.705.$$

b This time we will need to first determine the coordinates of the extreme points:

For $x = -1$, $y = -1.2\times(-1)^2+9 = 7.8$ and for $x = 2$, $y = -1.2\times(2)^2+9 = 4.2$.

Therefore, the average rate of change is equal to:

$\frac{y_2-y_1}{x_2-x_1} = \frac{4.2-7.8}{2-(-1)} = -1.2$.

It is not always necessary to have a graph in order to find the average rate of change. Often we are given information in the form of a table.

Example E.1.4

The table below shows the number of bacteria, *N*, present in an enclosed environment. Find the average growth rate of the population size over the first 4 hours.

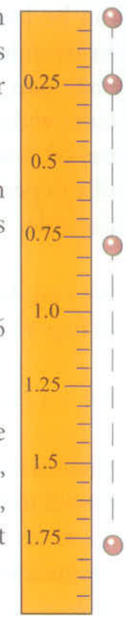
Time (hrs)	0	1	2	3	4	5	6	7	9
N	30	36	43	52	62	75	90	107	129

This time we need to consider the time interval $t = 0$ to $t = 4$. From the table we observe that the coordinates corresponding to these values are; (0,30) and (4,62). Therefore, the average rate of growth of the number of bacteria over the first 4 hours is equal to $\frac{62-30}{4-0} = \frac{32}{4} = 8$.

This means that during the first 4 hours, the number of bacteria was increasing (on average) at a rate of 8 every hour.

Notice that in the 1st hour, the average rate was $\frac{36-30}{1-0} = \frac{6}{1} = 6$ (< 8), whereas in the 4th hour the average rate of increase was $\frac{62-52}{4-3} = \frac{10}{1} = 10$ (> 8).

Velocity as a measure of the rate of change of displacement



Consider a marble that is allowed to free fall from a height of 2 metres (see diagram). As the marble is falling, photographs are taken of its fall at regular intervals of 0.25 second.

From its motion, we can tell that the rate at which the marble is falling is increasing (i.e. its velocity is increasing).

What is its average velocity over the first 0.6 second?

Reading from the diagram, we see that the marble has fallen a total distance of 1.75 (approximately), therefore, the average velocity v_{ave} of the marble, given by the rate at which its displacement increases (or decreases), is given by

$$v_{ave} = \frac{1.75-0}{0.6-0} \approx 2.92 \text{ m/sec}$$

Video discussion of average and instantaneous rates of change.

Example E.1.5

The displacement, x m, of an object, t seconds after it is dropped from the roof of a building is given by $x = 4.9t^2$ metres.

a What is the object's displacement after 4 seconds?

b What is the average velocity of the object over the first 4 seconds of its motion?

a After 4 seconds of free fall, the object's displacement will be $4.9(4)^2 = 78.4$ m.

We obtained this result by substituting the value of $t = 4$ into the equation for the displacement $x = 4.9t^2$.

b The average velocity is given by the average rate of change of displacement, x m, with respect to the time t seconds.

Once we have the starting position and the end position we can determine the average velocity using:

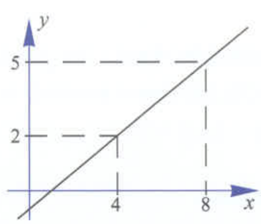
$$v_{ave} = \frac{x_2 - x_1}{t_2 - t_1} = \frac{78.4 - 0}{4 - 0} = 19.6\,.$$

That is, the object's average velocity over the first 4 seconds is 19.6 m/s.

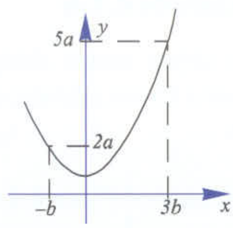
Exercise E.1.2

1. For each of the following graphs determine the average rate of change over the specified domain.

a $x \in [4, 8]$



b $x \in [-b, 3b]$



2. For each of the following functions, find the average rate of change over the given domain.

a $x \mapsto x^2 + 2x - 1, x \in [0, 2]$

b $x \mapsto \sqrt{x+1}, x \in [3, 8]$

c $x \mapsto 10 - \frac{1}{\sqrt{x}}, x \in [2, 20]$

d $x \mapsto \frac{x}{x+1}, x \in [0.1, 1.1]$

e $x \mapsto \frac{1}{1+x^2} - 1, x \in [0, 100]$

f $x \mapsto x\sqrt{400 - x}, x \in [300, 400]$

g $x \mapsto 2^x, x \in [0, 5]$

h $x \mapsto (x-1)(x+3), x \in [-3, 2]$

3. The displacement of an object, t seconds into its motion, is given by the equation, $s(t) = t^3 + 3t^2 + 2t, t \geq 0$.

Find the average rate of change of displacement during:

a the first second.

b the first 4 seconds.

c the interval when $t = 1$ to $t = 1 + h$.

4. The distance s metres that a particle has moved in t seconds is given by the function $s = 4t + 2t^2, t \geq 0$ Find the particle's average speed over the first 4 seconds.

5. The distance s metres that a particle has moved in t seconds is given by the function $s = 4t + 2t^2, t \geq 0$ Find the particle's average speed during the time interval from $t = 1$ to $t = 1 + h$.

6. The temperature T °C of food placed inside cold storage is modelled by the equation:

$$T = \frac{720}{t^2 + 2t + 25}$$, where t is measured in hours.

Find the average rate of change of the temperature, T°C, with respect to the time, t hours, during the first 2 hours that the food is placed in the cold storage.

7. The volume of water in a hemispherical bowl of radius r is given by:

$V = \frac{1}{3}\pi h^2(3r - h)$, where h is the height of the water surface inside the bowl.

Extra questions

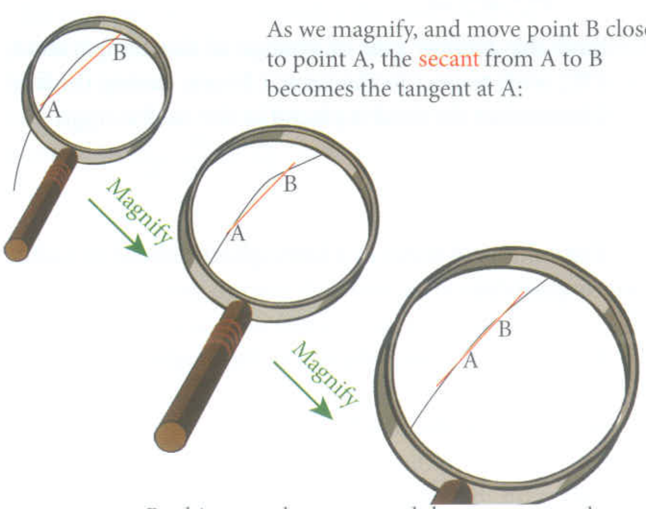
Instantaneous Rate of Change

Informal idea of limits

As already discussed, the average rate of change between two points on a curve is determined by finding the gradient of the straight line joining these two points. However, we often need to find the rate of change at a particular instant, and so the method used for finding the average rate of change is no longer appropriate. However, it does provide the foundation that leads to obtaining the instantaneous rate of change. We refine our definition of the average rate of change to incorporate the notion of the instantaneous rate of change. The basic argument revolves around the notion of magnifying near the point where we wish to find the instantaneous rate of change, that is, by repeatedly 'closing in' on a section of a curve. This will give the impression that over a very small section, the curve can be approximated by a straight line. Finding the gradient of that straight line will provide us with a very good approximation of the rate of change of the curve (over the small region under investigation). To obtain the exact rate of change at a particular point on the curve we will then need to use a limiting approach.

The process used to determine the rate of change at A is carried out as follows:

Start by drawing a secant from A to B, where B is chosen to be close to A. This will provide a reasonable first approximation for the rate of change at A. Then, to obtain a better approximation we move point B closer to point A.



As we magnify, and move point B closer to point A, the secant from A to B becomes the tangent at A:

By this stage the secant and the tangent are almost the same at point A.
Therefore the gradient of the tangent, the secant and that of the curve at A are almost the same.

Next, zoom-in towards point A, again. We move point B closer to point A, whereby a better measure for the rate of change at point A is now obtained.

We then repeat step 2, i.e. move B closer to A and zoom in, move point B closer to A and zoom in, and so on.

Finally, the zooming-in process has reached the stage whereby the secant is now virtually lying on the curve at A. In fact the secant is now the tangent to the curve at the point A.

Using the process of repeatedly zooming in to converge on a particular region lies at the heart of the limiting process. Once we have understood the concepts behind the limiting process, we can move on to the more formal aspect of limits. We now provide a 'visual' representation of steps 1 to 3 described above.

Example E.1.6

An object moves along a straight line. Its position, x metres (from a fixed point O), at time t seconds is given by:

$x(t) = t - \frac{1}{4}t^2$, $t \geq 0$. Determine:

a its average velocity over the interval from $t = 1$ to $t = 2$

b its average velocity over the interval $t = 1$ to $t = 1.5$

c its average velocity over the interval $t = 1$ to $t = 1.1$

d its average velocity over the interval $t = 1$ to $t = 1 + h$, where h is small.

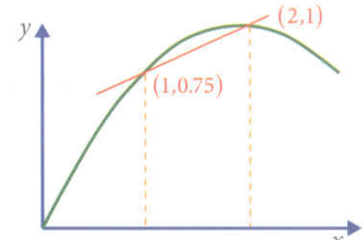
How can the last result help us determine the object's velocity at $t = 1$?

a The average velocity over the required second (from $t = 1$ to $t = 2$) is found by looking at the slope of the secant joining those two points on the graph of $x(t)$.

At $t = 2$, we have $x(2) = 2 - \frac{1}{4}(2)^2 = 1$, and at $t = 1$, $x(1) = 1 - \frac{1}{4}(1)^2 = \frac{3}{4}$

Therefore, we have that:

$$v_{ave} = \frac{x(2) - x(1)}{2 - 1} = \frac{1 - 0.75}{1} = 0.25$$



Therefore, the average velocity over the second is 0.75m/s.

b For $t = 1$ to $t = 1.5$ we have,

$$v_{ave} = \frac{x(1.5)-x(1)}{1.5-1} = \frac{(1.5-0.25\times 1.5^2)-0.75}{0.5} = 0.375$$

c Similarly, for $t = 1$ to $t = 1.1$, we have,

$$v_{ave} = \frac{x(1.1)-x(1)}{1.1-1} = 0.475$$

d We are now in a position to determine the average rate over the interval $t = 1$ to $t = 1 + h$.

The average velocity is given by $v_{ave} = \frac{x(1+h)-x(1)}{1+h-1}$

Now, $x(1+h) = (1+h)-0.25(1+h)^2$
$= 1+h-0.25(1+2h+h^2)$
$= 0.75+0.5h-0.25h^2$

Therefore,

$$v_{ave} = \frac{0.75+0.5h-0.25h^2-0.75}{1+h-1} = \frac{0.5h-0.25h^2}{h}$$

$$= \frac{h(0.5-0.25h)}{h}$$

$$= 0.5-0.25h,\ h \neq 0$$

Notice that for part b, (i.e. $t = 1$ to $t = 1.5$) $h = 0.5$, so that substituting $h = 0.5$ into this equation we have, $v_{ave} = 0.5-0.25(0.5) = 0.375$, providing the same result as before.

We can set up a table of values and from it determine what happens as we decrease the time difference.

We notice that, as h becomes very small, the average rate of change from $t = 1$ to $t = 1 + h$ becomes the instantaneous rate of change at t = 1! This is because we are zooming in onto the point where $t = 1$.

h	v_{ave}
0.1	0.475
0.01	0.4975
0.001	0.4999

This means that the rate of change at $t = 1$ (h '= 0') would therefore be 0.5 m/s. This means that the particle would have a velocity of 0.5 m/s after 1 second of motion.

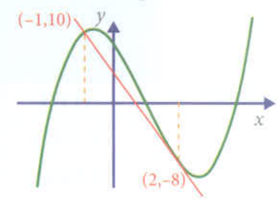
Example E.1.7

For the graph with equation $f(x) = (x+2)(x-1)(x-4)$,

a Find the average rate of change of f over the interval [–1,2].

b Find the rate of change of f, where $x = 4$.

a We first find the coordinates of the end points for the interval [–1,2]:



$x = -1,\ y = f(-1) = ...$

$= (-1+2)(-1-1)(-1-4) = 10.$

$x = 2,\ y = f(2) = (2+2)(2-1)(2-4) = -8.$

Therefore, the average rate of change in y with respect to x over the interval [–1,2] is given by

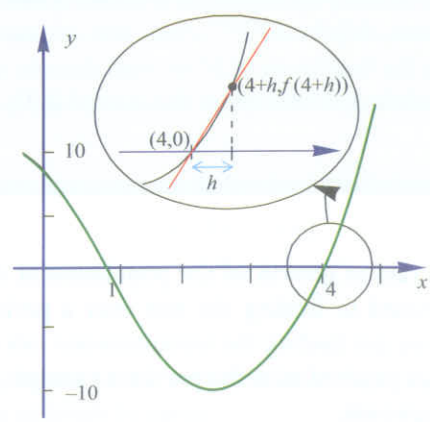
$$\frac{f(2)-f(-1)}{2-(-1)} = \frac{-8-10}{3} = -6$$

b To determine the rate of change at $x = 4$, we choose a second point close to $x = 4$. In this case, we use the point $x = 4 + h$, where h can be considered to be a very small number.

We will look at what happens to the gradient of the secant joining the points (4, 0) and $(4+h, f(4+h))$ as h approaches zero.

The gradient of the secant is given by:

$$\frac{f(4+h)-f(4)}{(4+h)-4} = \frac{f(4+h)-f(4)}{h}$$



We now need to determine the value of $f(4+h)$ and $f(4)$. However, we already know that $f(4) = 0$.

We can now find values for $f(4+h)$ as h approaches zero.

For $h = 0.1$, $f(4+0.1) = f(4.1) = (4.1+2)(4.1-1)(4.1-4)$
$= 6.1 \times 3.1 \times 0.1 = 1.891$

Therefore, $\frac{f(4+h)-f(4)}{h} = \frac{1.891-0}{0.1} = 18.91$, for $h = 0.1$.

We can continue in this same manner by making the value of h smaller still.

We do this by setting up a table of values:

h	$\frac{f(4+h)-f(4)}{h}$
0.01	18.09010000
0.001	18.00900100
0.0001	18.00090001

From the table, it appears that as h approaches zero, the gradient of the secant (which becomes the gradient of the tangent at (4,0)) approaches a value of 18.

Therefore, we have that the rate of change of f at (4,0) is 18.

More formally we write this result as,

$$\lim_{h \to 0} \frac{f(4+h)-f(4)}{h} = 18 \text{ which is read as}$$

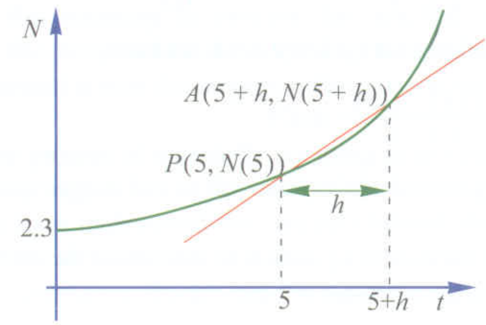
"The limit as h tends to zero of $\frac{f(4+h)-f(4)}{h}$ is equal to 18."

Example E.1.8

The population of a city at the start of 2010 was 2.3 million, and its projected population, N million, is modelled by the equation $N(t) = 2.3^{0.0142t}, t \geq 0$ and is measured in years since the beginning of 2010. Find the rate of growth of the population in this city at the start of 2015.

Finding the rate of growth of the population at the start of 2005 as opposed to finding the rate over a period of time means that we are finding the instantaneous rate of change. To do this, we proceed as in the previous example, i.e. we use a limiting approach.

Consider the two points, $P(5, N(5))$ (start of 2015) and $A(5+h, N(5+h))$ on the curve representing the population size:



The gradient of the secant passing through P and A is given by:

$$\frac{N(5+h)-N(5)}{(5+h)-5} = \frac{N(5+h)-N(5)}{h}$$

Now, $N(5) = 2.3e^{0.0142 \times 5} = 2.3e^{0.071}$

and $N(5+h) = 2.3e^{0.0142(5+h)}$

Therefore, the gradient of the secant is given by

$$\frac{2.3e^{0.0142(5+h)} - 2.3e^{0.071}}{h} = \frac{2.3e^{0.071+0.0142h} - 2.3e^{0.071}}{h}$$
$$= \frac{2.3e^{0.071}(e^{0.0142h}-1)}{h}$$

Again we set up a table of values:

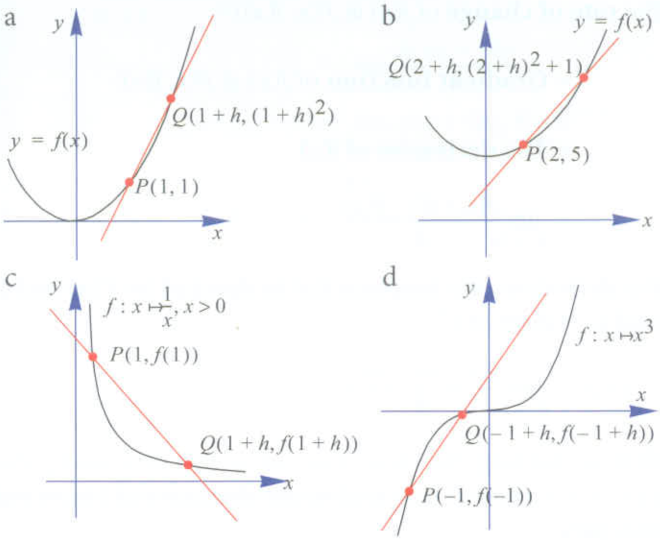
h	$\frac{2.3e^{0.071}(e^{0.0142h}-1)}{h}$
0.1	$\frac{2.3e^{0.071}(e^{0.0142 \times 0.1}-1)}{0.1} = 0.035088$
0.01	$\frac{2.3e^{0.071}(e^{0.0142 \times 0.01}-1)}{0.01} = 0.035066$
0.001	$\frac{2.3e^{0.071}(e^{0.0142 \times 0.001}-1)}{0.001} = 0.035063$
0.0001	$\frac{2.3e^{0.071}(e^{0.0142 \times 0.0001}-1)}{0.0001} = 0.035063$

Using limit notation we have: $\lim_{h \to 0} \frac{N(5+h)-N(5)}{h} = 0.035063$

That is, the growth rate at the start of 2015 is 35 063 people per year.

Exercise E.1.3

1. For each of the graphs shown, find the gradient of the secant joining the points P and Q.



2. For each of the graphs in Question 1, use a limiting argument to deduce the instantaneous rate of change of the given function at the point P.

3. For each of the functions, f, given below, find the gradient of the secant joining the points $P(a, f(a))$ and $Q(a+h, f(a+h))$.

   a $f(x) = 3 + x^2$ b $f(x) = 1 - x^2$

   c $f(x) = (x+1)^2 - 2$ d $f(x) = x^3 + x$

   e $f(x) = 2 - x^3$ f $f(x) = x^3 - x^2$

   g $f(x) = \frac{2}{x}$ h $f(x) = \frac{1}{x-1}$

   i $f(x) = \sqrt{x}$

4. An object moves along a straight line. Its position, x metres (from a fixed point O), at time t seconds is given by $x(t) = 2t^2 - 3t + 1$, $t \geq 0$.

   a Sketch the graph of its displacement function.

   b Determine :

   i its average velocity over the interval from $t = 1$ to $t = 2$

   ii its average velocity over the interval $t = 1$ to $t = 1.5$

   iii its average velocity over the interval $t = 1$ to $t = 1.1$

   c Show that its average velocity over the interval $t = 1$ to $t = 1 + h$, where h is small, is given by $1 + 2h$.

   d How can the last result help us determine the object's velocity at $t = 1$?

   e Show that its average velocity over any time interval of length h is given by $4t + 2h - 3$. Hence deduce the object's velocity at any time t during its motion

Extra questions

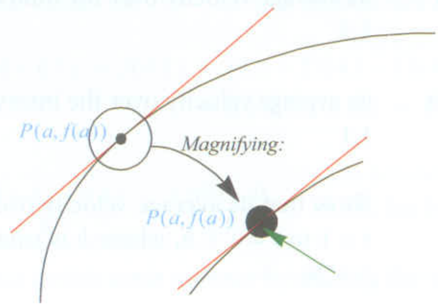
The Derivative and the Gradient Function

In the previous section we concentrated on determining the average rate of change for a function over some fixed interval. We then proceeded to find the instantaneous rate of change at a particular point (on the curve). We now consider the same process, with the exception that we will discuss the instantaneous rate at any point $P(x, f(x))$. The result will be an expression that will enable us to determine the instantaneous rate of change of the function at any point on the curve. Because the instantaneous rate of change at a point on a curve is a measure of the gradient of the curve at that point, our newly found result will be known as the gradient function (otherwise known as the derivative of the function).

For a continuous function, $y = f(x)$, we deduced that the instantaneous rate of change at the point $P(a, f(a))$ is given by:

$$\frac{f(a+h) - f(a)}{h}$$

where h is taken to be very small (in fact we say that h approaches or tends to zero).

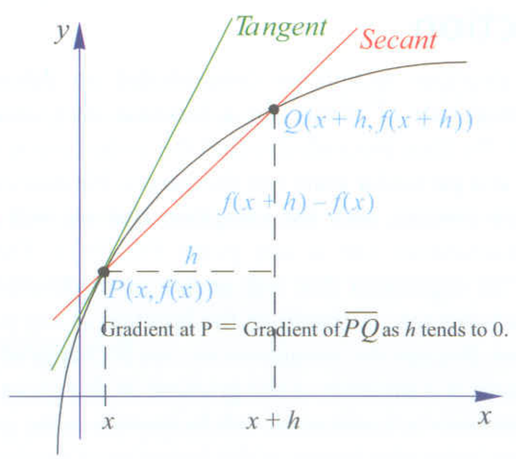


At the point P, the tangent and the line are one and the same.
Therefore, finding the gradient of the tangent at P is the same as finding the rate of change of the function at P.

So, to determine the rate at which a graph changes at a single point, we need to find the slope of the tangent line at that point.

This becomes obvious if we look back at our 'zooming in process—where the tangent line to the function at the point $P(a, f(a))$ is the line that best approximates the graph at that point.

Rather than considering a fixed point $P(a, f(a))$, we now consider any point $P(x, f(x))$ on the curve with equation $y = f(x)$:



The rate of change of the function f at $P(x, f(x))$ is therefore given by the gradient of the tangent to the curve at P.

If point Q comes as close as possible to the point P, so that h approaches zero, then, the gradient of the tangent at P is given by the gradient of the secant joining the points $P(x, f(x))$ and $Q(x+h, f(x+h))$ as $h \to 0$.

In mathematical notation we have:

$$\text{Rate of change at P} = \lim_{h \to 0} \frac{f(x+h) - f(x)}{h}$$

Notation and language

We now introduce the term derivative of a function:

The **rate of change** of $f(x)$ at $P(x, f(x))$

= **Gradient function** of $f(x)$ at $P(x, f(x))$

= **The derivative** of $f(x)$

$$= \lim_{h \to 0} \frac{f(x+h) - f(x)}{h}$$

The derivative of a function $f(x)$ is denoted by $f'(x)$ and is read as "f dash of x".

$$f'(x) = \lim_{h \to 0} \frac{f(x+h) - f(x)}{h}$$

That is, finding the derivative of a function using this approach is referred to as *finding the derivative of f from first principles*.

It is important to realise that in finding $f'(x)$ we have a new function – called the gradient function, because the expression $f'(x)$ will give the gradient anywhere on the curve of $f(x)$. If we want the gradient of the function $f(x)$ at $x = 5$, we first determine $f'(x)$ and then substitute the value of $x = 5$ into the equation of $f'(x)$.

Example E.1.9

Using the first principles method, find the derivative (*or the gradient function*) of the function $f(x) = 3x^2 + 4$.

Hence, find the gradient of the function at $x = 3$.

Using the first principles method means that we must make use of the expression

$$f'(x) = \lim_{h \to 0} \frac{f(x+h) - f(x)}{h} \quad - (1)$$

We start by first evaluating the expression $f(x+h) - f(x)$:

That is: $f(x+h) - f(x) = 3(x+h)^2 + 4 - [3x^2 + 4]$
$= 3(x^2 + 2xh + h^2) + 4 - 3x^2 - 4$
$= 3x^2 + 6xh + 3h^2 - 3x^2$
$= 6xh + 3h^2$

Substituting this result into (1):

$$\begin{aligned}\lim_{h \to 0} \frac{f(x+h)-f(x)}{h} &= \lim_{h \to 0} \frac{6xh+3h^2}{h} \\ &= \lim_{h \to 0} \frac{h(6x+3h)}{h} \\ &= \lim_{h \to 0} (6x+3h), h \neq 0 \\ &= 6x\end{aligned}$$

That is, we now have the gradient function $f'(x) = 6x$.

To determine the gradient of the function at $x = 3$, we need to substitute the value $x = 3$ into the gradient function. That is, $f'(3) = 6 \times 3 = 18$.

Exercise E.1.4 (* indicates extension questions)

1. Use a limiting process to find the gradients of these curves at the points indicated.

   a $x \mapsto x^3$ at $x = 1$

   b $v = 2t^2 - 1$ at $t = 2$

   c* $f(x) = \frac{1}{x}$ at $x = 3$

   d* $x \mapsto 2^x$ at $x = 1$

   e $f = t^2 - 2t + 3$ at $t = 0.5$

   f* $t \mapsto \frac{t^2-1}{t}$ at $t = 4$

2. An object is dropped from a high building. The distance, d metres, that the object has fallen, t seconds after it is released, is given by the formula $d = 4.9t^2, 0 \le t \le 3$.

   a Find the distance fallen during the first second.

   b Find the distance fallen between $t = 1$ and $t = h + 1$ seconds.

   c Hence, find the speed of the object 1 second after it is released.

3. Find, from first principles, the gradient function, $f'(x)$, of the following.

   a $f{:}x \mapsto 4x^2$ b $f{:}x \mapsto 5x^2$

   c $f{:}x \mapsto 4x^3$ d $f{:}x \mapsto 5x^3$

   e $f{:}x \mapsto 4x^4$ f $f{:}x \mapsto 5x^4$

   Can you see a pattern in your results?

4. Find, from first principles, the derivatives of the following functions.

   a $f(x) = 2x^2 - 5$ b $g(x) = 2 - x$

   c $g(x) = 2 - x + x^3$ d $f(x) = \frac{1}{x}$

   e $f(x) = \frac{2}{x+1}$ f $f(x) = \sqrt{x}$

5. A particle moving along a straight line has its position at time t seconds governed by the equation $x(t) = 2t - 0.5t^2, t \ge 0$, where $x(t)$ is its position in metres from the origin O.

   a Find the particle's velocity after it has been in motion for 1 second.

   b Find the particle's velocity at time $t = a, a > 0$.

6. A particle moving along a straight line has its position at time t seconds governed by the equation $x(t) = 4t^2 - t^3, t \ge 0$, where $x(t)$ is its position in metres from the origin O.

   a Sketch the displacement-time graph of the motion over the first five seconds

   b Find the particle's velocity at time:

   i $t = 1$ ii $t = 2$

   c Find the particle's velocity at any time $t, t \ge 0$.

   d When will the particle first come to rest?

Answers

At various time in this course you will encounter rules that will enable you to find exact expressions for the gradients of functions.

There are two situations in which exact answers are not possible.

1. When the function is too complicated to differentiate.

2. When there is only experimental data.

All is not lost as you can always fall back on the limiting processes discussed in this chapter.

In the first case, suppose you were asked to find the gradient of the function:

$$f(x)=x^{x^x}, x>0$$

at the point where $x = 1$. Note that there is no problem evaluating the function at 1. $f(x)=1^{1^1}=1$.

Can you use technology to implement the limiting process and obtain an answer?

One solution is illustrated:

	A	B	C	D
1	Gradient Finder			
2				
3	Mid-point	1		
4	Increment	0.01		
5				
6	x	f(x)	Difference	Approx Gradient
7	0.97	0.97174771		
8	0.98	0.98078434	0.00903663	0.90366328
9	0.99	0.99019802	0.00941368	0.94136811
10	1	1	0.00980198	0.98019785
11	1.01	1.01020202	0.01020202	1.02020219
12	1.02	1.02081635	0.01061433	1.06143313
13	1.03	1.0318558	0.01103945	1.10394517

The red cell shows the gradient triangle below (1,1) and the green cell the gradient triangle above (1,1). The limiting process can be achieved by reducing the value in cell B4.

The gradients appear to be homing in on 1 from both above and below. Is this the answer?

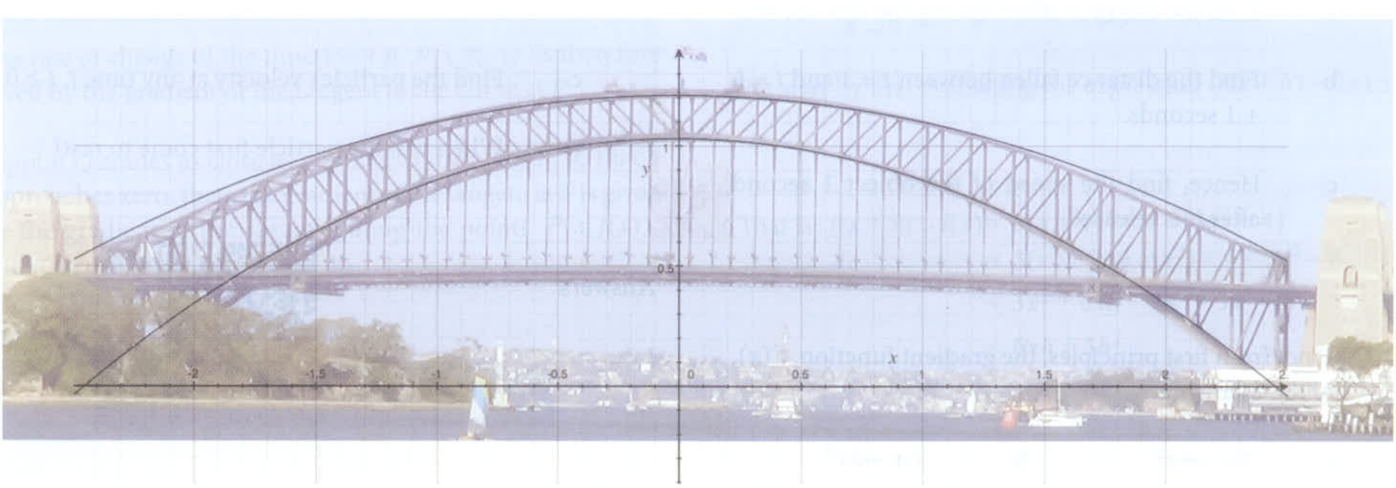
The spreadsheet can be downloaded here:

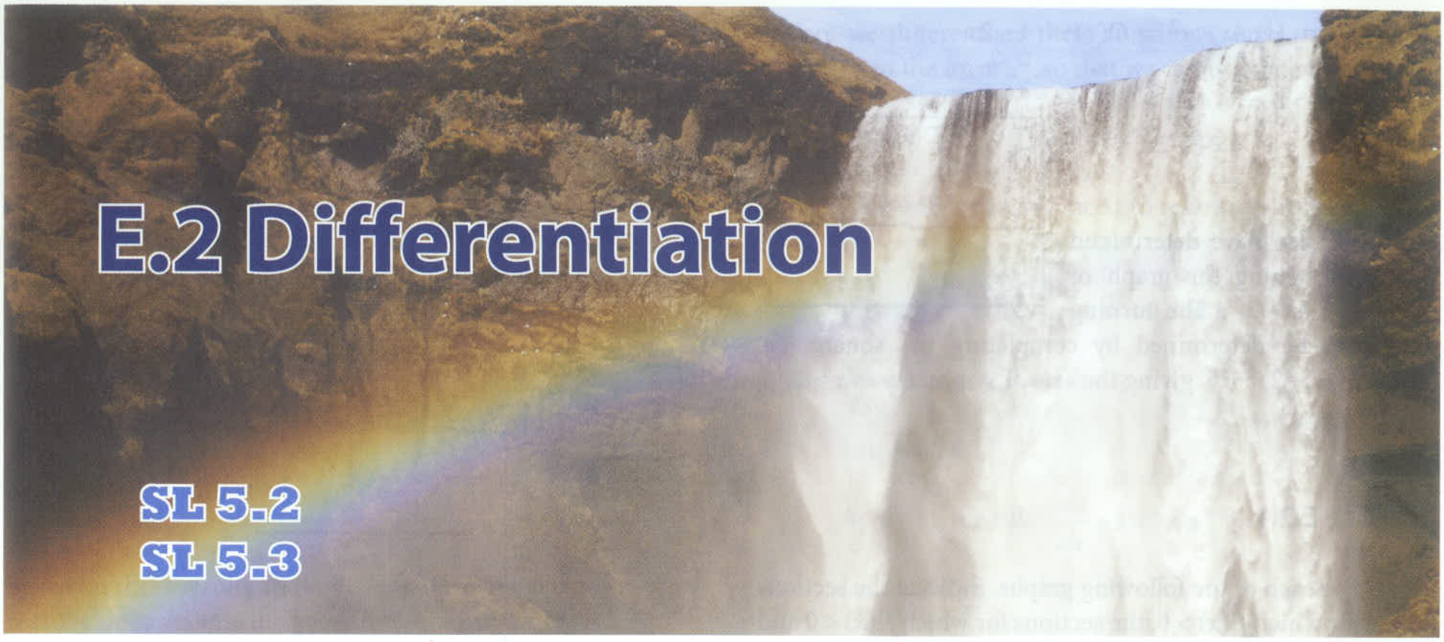
Here is an example of the second situation in which we may need to use numerical methods.

The graphic below is a picture of the Sydney Harbour bridge. The file may be downloaded here:

Axes and a (very approximate) curve have been added which you may use or ignore.

How steep is the upper arch (where the famous 'Bridge Walk' occurs)?





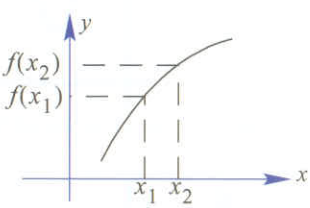
E.2 Differentiation

SL 5.2
SL 5.3

Increasing and Decreasing Functions

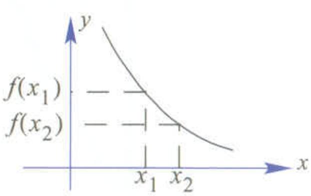
In the previous section, we looked at the idea of the gradient of a curve and a few ways in which this can be found. In Section B we looked at graphs and the ways these can illustrate complex situations.

A function f is said to be increasing if its graph rises as it is sketched from left to right.



That is, if :

$x_2 > x_1 \Rightarrow f(x_2) > f(x_1)$ (i.e. the y-values increase as the x-values increase).



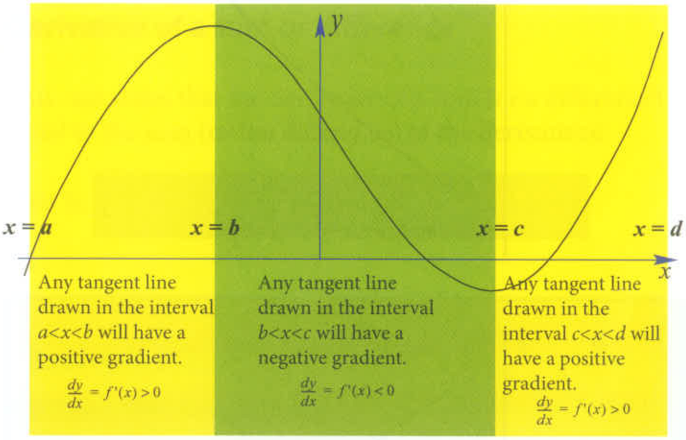
Similarly,

A function f is said to be decreasing if its graph falls as it is sketched from left to right.

That is, if $x_2 > x_1 \Rightarrow f(x_2) < f(x_1)$ (i.e. the y-values decrease as the x-values increase).

... A calculus point of view

The derivative can be used to determine whether a function is increasing or decreasing and so it can be used to help find those values of x for which the function is increasing or decreasing.



This means that, to determine where a function is increasing or decreasing, the values of x for which $f'(x) > 0$ and $f'(x) < 0$ respectively need to be found.

Example E.2.1

Find the values of x for which the function $f(x) = 1 + 4x - x^2$ is increasing.

By definition, a function is increasing for those values of x for which $f'(x) > 0$.

Therefore find:

1. $f'(x)$

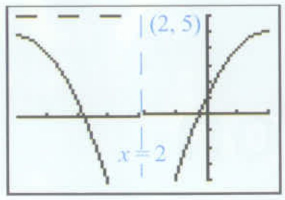
2. the values of x such that $f'(x) > 0$

Now, $f(x) = 1 + 4x - x^2 \Rightarrow f'(x) = 4 - 2x$

Then, $f'(x) > 0 \Leftrightarrow 4 - 2x > 0$

$\Leftrightarrow 4 > 2x$

$\Leftrightarrow x < 2$

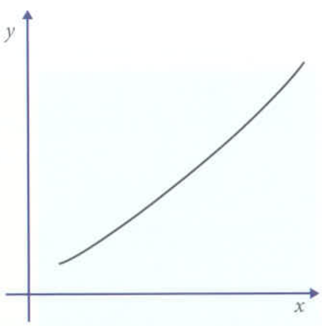


We could also have determined this by sketching the graph of $f(x) = 1 + 4x - x^2$. The turning point can be determined by completing the square, i.e. $f(x) = -(x-2)^2 + 5$ giving the axis of symmetry as $x = 2$.

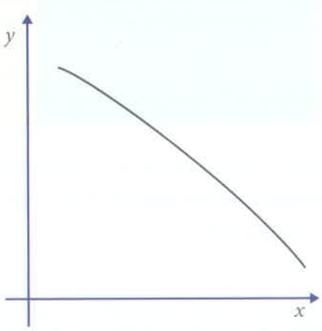
Exercise E.2.1

1. For each of the following graphs, indicate the sections for which $f'(x) > 0$, the sections for which $f'(x) < 0$ and any points where $f'(x) = 0$.

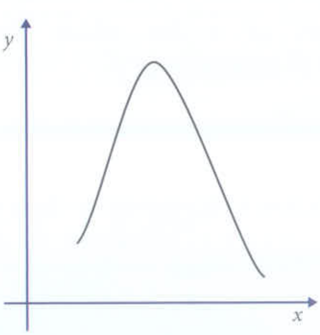
a



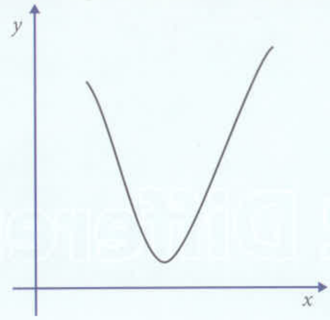
b



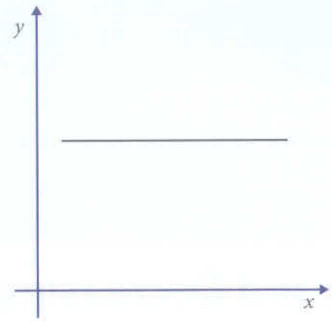
c



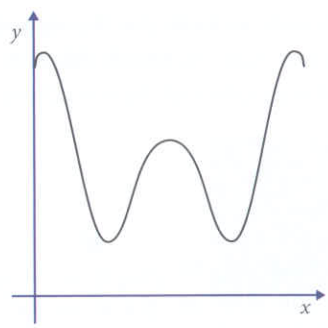
d



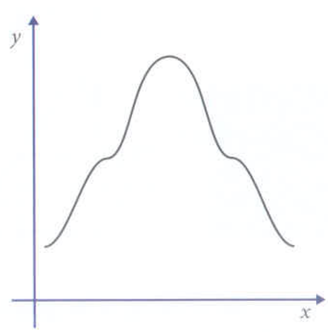
e



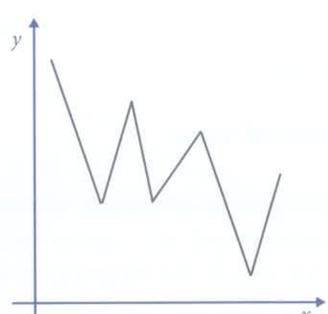
f



g



h



Differential Calculus

Power rule for differentiation

Finding the derivative from first principles can be tedious. The previous two examples clearly show this. However, using the first principles approach produces the results shown in the table below:

Function $y = f(x)$	x^4	x^3	x^2	x^1	x^{-1}	x^{-2}
Derivative $\frac{dy}{dx} = f(x)$	$4x^3$	$3x^2$	$2x^1$	$1x^0$	$-1x^{-2}$	$-2x^{-3}$

Based on these results and following the general pattern, it is reasonable to assume the general result that if:

$$y = x^n, n \in Z, \text{ then } \frac{dy}{dx} = nx^{n-1}$$

In fact this rule is true for any exponent $n \in \mathbb{R}$, i.e. for any real number n. However, at this level, all that is required is examples using $n \in \mathbb{Z}$.

For example, if we look at the reciprocal function, then we have that:

$$y = \frac{1}{x} = x^{-1}, x \neq 0.$$

When using the rule, be careful to identify n. In this case $n = -1$.

Following that, calculate $n - 1$. This is where most mistakes get made, particularly when n is negative.

In this case, $n - 1 = -2$.

This result is known as the power rule for differentiation.

Notice that for the case $n = 0$, then $y = x^0$ and so we have that:

$$\frac{dy}{dx} = 0x^{0-1} = 0.$$

All constants differentiate to zero.

Example E.2.2

Use the power rule to differentiate the following functions.

a x^6 b 5 c $7x^3$ d $\frac{1}{x^2}$

Before we differentiate these functions, each part must be rewritten in the form x^n so that we can use the power rule.

a Let $f(x) = x^6 \Rightarrow f'(x) = 6x^{6-1} = 6x^5$

b Remember, we first need to rewrite it in the form x^n: Let $y = 5x^0$.

Identify n. In this case, $n = 0$ and $n - 1 = -1$.

$$\frac{dy}{dx} = 5 \times 0 \times x^{-1} = 0$$

c Let $y = 7x^3$.

Applying the rule: $\frac{dy}{dx} = 3 \times 7 \times x^{3-1} = 21x^2$

d Let $f(x) = \frac{1}{x^2}$ so that $f(x) = x^{-2}$

$\therefore f'(x) = -2x^{-2-1} = -2x^{-3}$, that is, $f'(x) = -\frac{2}{x^3}$.

Derivative of a sum or difference

This rule states that the derivative of a sum (or a difference) is equal to the sum (or the difference) of the derivatives.

That is, If $y = f(x) \pm g(x)$ then $\frac{dy}{dx} = f'(x) \pm g'(x)$

Example E.2.3

Differentiate these functions.

a $y = 2x^3 + 5x - 9$

b $f(x) = (3x+1)^2$

c $f(x) = \frac{3x^2 - 2x + 3}{x^2}$

a
$$y = 2x^3 + 5x - 9 \Rightarrow \frac{dy}{dx} = \frac{d}{dx}(2x^3 + 5x - 9)$$
$$= \frac{d}{dx}(2x^3) + \frac{d}{dx}(5x) - \frac{d}{dx}(9)$$
$$= 6x^2 + 5$$

Notice we have used a slightly different notation, namely that:

$$f'(x) = \frac{d}{dx}(f(x)).$$

We can think of $\frac{d}{dx}$ as the differentiation operator, so that $\frac{d}{dx}(f(x))$ or $\frac{d}{dx}(y)$ is an operation of differentiation done on

$f(x)$ or y respectively.

b $f(x)=(3x+1)^2$

At this stage, it is necessary to simplify (multiply out) the expression. Remember to use the Distributive Law correctly - there are four multiplications to complete:

$$f(x)=(3x+1)^2$$
$$=3x\times3x+3x\times1+1\times3x+1\times1$$
$$=9x^2+6x+1$$

Next, apply the rule to each term:

$$f'(x)=9\times2x^{2-1}+6\times1x^0+0$$
$$=18x+6$$

c As with the previous example, simplify to the point where you have a sum and difference of powers of x.

$$f(x)=\frac{3x^2-2x+3}{x^2}$$
$$=\frac{3x^2}{x^2}-\frac{2x}{x^2}+\frac{3}{x^2}$$
$$=3-2x^{-1}+3x^{-2}$$

Next, use the rule:

$$f'(x)=0-2\times(-1)\times x^{-2}+3\times(-2)\times x^{-3}$$
$$=2x^{-2}-6x^{-3}$$
$$=\frac{2}{x^2}-\frac{6}{x^3}$$

Exercise E.2.2

1. Find the derivative of each of the following.

a	x^5	b	x^9	c	x^{25}
d	$9x^3$	e	$-4x^7$	f	$\frac{1}{4}x^8$
g	x^2+8	h	$5x^4+2x-1$		
i	$-3x^5+6x^3-x$	j	$20-\frac{1}{3}x^4+10x$		
k	$3x^3-6x^2+8$	l	$3x-1+\frac{x^2}{5}+x^4$		

2. Find the derivative of each of the following.

a $y=(2x-1)^2$

b $y=\frac{1}{x}-2x$

c $f(x)=(3-x)^2$

d $f(x)=(x^{-2})^2$

e $f(x)=\left(1+\frac{1}{x}\right)^2$

f $y=(x+1)(x-2)$

g $y=(x^2+3)(x-2)$

h $f(x)=(x^2-2)^2$

i $f(x)=\left(x-\frac{1}{x}\right)^3$

3. If: $f(x)=x^3, g(x)=x+1, h(x)=\frac{1}{x}, x\neq0$

find the derivative, with respect to x, of:

a $2.f(x)$

b $f(x)+g(x)$

c $f(x)\times g(x)$

d $f(x)\times h(x)$

e $g(x)\div h(x)$

f $(g(x))^3$

g $f(x)[g(x)+h(x)]$

h $f(x)\times g(x)\times h(x)$

4.a Show that if $f(x)=x^2-x$, then $f'(x)=1+\frac{2f(x)}{x}$.

b Show that if $f(x)=\sqrt{2x}-2\sqrt{x}, x\geq0$,
then $\sqrt{2x}f'(x)=1-\sqrt{2}, x>0$

c Show that if $y=ax^n$ where a is real and $n\in N$, then $\frac{dy}{dx}=\frac{ny}{x}, x\neq0$

Differentiating with variables other than *x* and *y*

Although it was convenient to establish the underlying theory of differentiation based on the use of the variables x and y, it must be pointed out that not all expressions are written in terms of x and y. In fact, many of the formulae that we use are written in terms of variables other than y and x, e.g. volume, V, of a sphere is given by:

$V = \frac{4}{3}\pi r^3$, where r is its radius.

The displacement of a particle moving with constant acceleration is given by

$s = ut + \frac{1}{2}at^2$.

However, it is reassuring to know that the rules are the same regardless of the variables involved. Thus, if we have that y is a function of x, we can differentiate y *with respect to* (w.r.t.) x to find $\frac{dy}{dx}$.

On the other hand, if we have that y is a function of t, we would differentiate y w.r.t. t and write $\frac{dy}{dt}$.

Similarly, if W was a function of θ, we would differentiate W w.r.t. θ and write $\frac{dW}{d\theta}$.

Example E.2.4

Differentiate the following with respect to the appropriate variable.

a $V = \frac{4}{3}\pi r^3$ b $p = 3w^3 - 2w + 20$

c $s = 10t + 4t^2$

a As V is a function of r, we need to differentiate V with respect to r:

$$V = \frac{4}{3}\pi r^3 \Rightarrow \frac{dV}{dr} = \frac{4}{3}\pi(3r^2) = 4\pi r^2.$$

b This time p is a function of w, and so we would differentiate p with respect to w:

$$p = 3w^3 - 2w + 20 \Rightarrow \frac{dp}{dw} = 9w^2 - 2.$$

c In this expression we have that s is a function of t and so we differentiate s w.r.t t:

$$s = 10t + 4t^2 \Rightarrow \frac{ds}{dt} = 10 + 8t.$$

Exercise E.2.3

1. Differentiate the following functions with respect to the appropriate variable.

a $s = 12t^4$ b $Q = \left(n + \frac{1}{n^2}\right)^2$

c $P = \frac{(2r-3)^2}{r^2}$ d $T = \left(\theta + \frac{1}{\theta}\right)^3$

e $A = 40L - L^3$ f $F = \frac{50}{v^2} - v$

g $V = 2l^3 + 5l$ h $A = 2\pi h + 4h^2$

2. Differentiate the following with respect to the independent variable:

a $v = \frac{2}{3}\left(5 - \frac{2}{t^2}\right)$

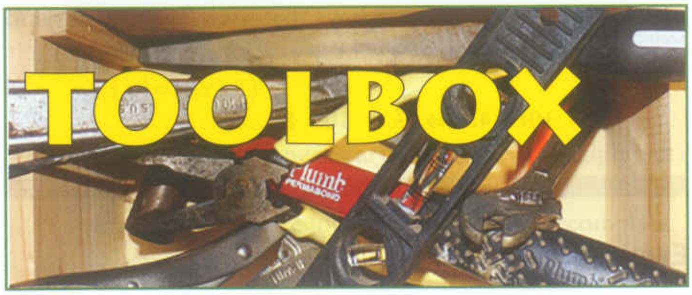
b $S = \pi r^2 + \frac{20}{r}$

c $q = s^4 - \frac{3}{s}$

d $h = \frac{2 - t + t^2}{t^3}$

e $L = \frac{(4-b)^2}{b}$

f $W = (m-2)^2(m+2)$



Here are some events described in terms of the rate at which things are happening.

1. A population of microbes grows under the following rules:

- In any minute, there is a fixed probability that an individual will divide to produce two new individuals.
- In any minute, there is a fixed probability that an individual will die.

Draw a graph of the population through time and investigate how changing the two probabilities will alter the graph. Use of technology is suggested (e.g. a spreadsheet).

2. As water cascades over a cliff, some of it disperses into droplets.

Early on, the droplets have a small vertical velocity. Under the influence of gravity, this increases. As the velocity increases, so does air resistance. This tends to slow the droplets down.

Draw a graph showing how this velocity varies with time.

3. These two photographs show some aeroplane instruments during a flight.

Photo 1: Time signature 11.31.49

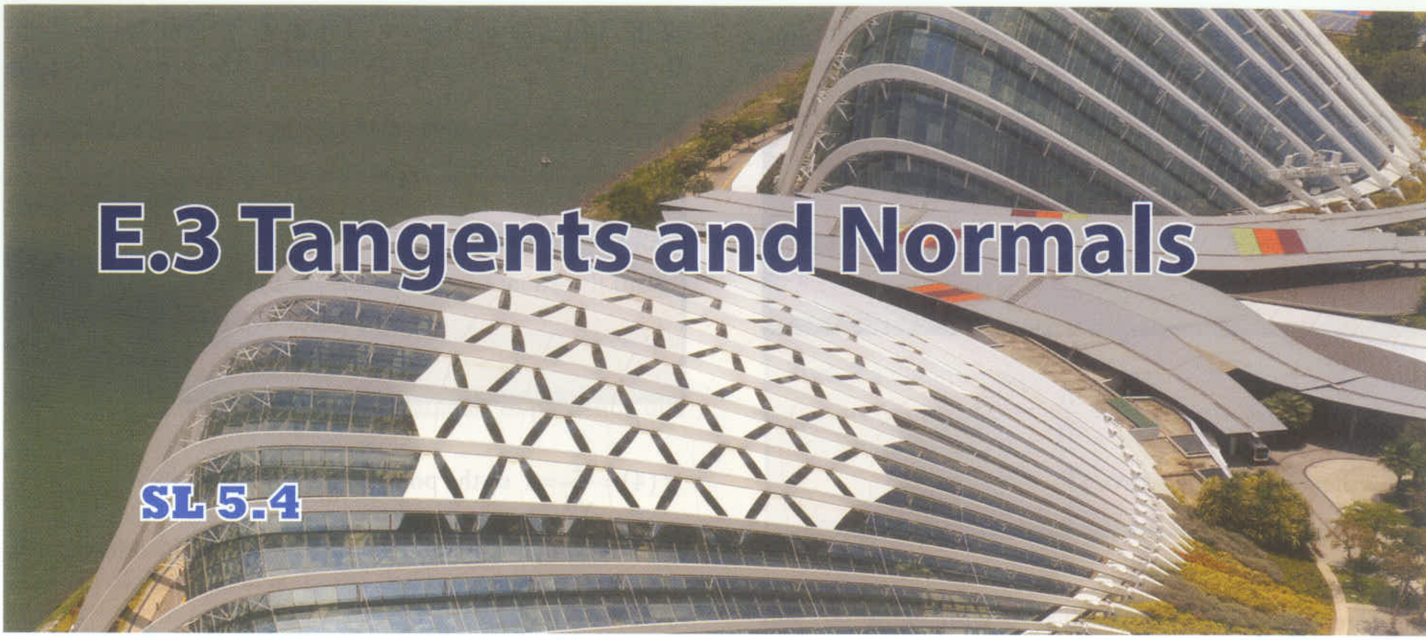
The instrument second from the right of the group of eight circular instruments is the altimeter (measures altitude). It works like an analogue clock and currently reads 2 530 feet. Yes, aviation uses feet, not metres.

The instrument directly beneath it measures rate of climb/descent in hundreds of feet per minute.

Photo 2: Time signature 11.31.59

Use the techniques developed in this chapter to explain how these two pieces of information are linked and can be used to describe part of the flight path of the aeroplane.

Answers



E.3 Tangents and Normals

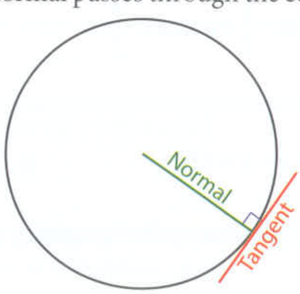
SL 5.4

Equation of Tangent

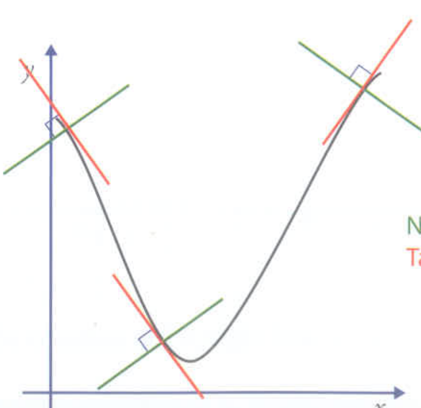
The gradient of a curve $y = f(x)$ at any point (x_1, y_1) is equal to the gradient of the tangent to the curve at that point.

A tangent is a straight line that just touches a curve. A normal is the straight line at right angles to the tangent that passes through the point of contact.

You are probably familiar with the idea of a tangent to a circle. In this case, the normal passes through the centre of the circle.



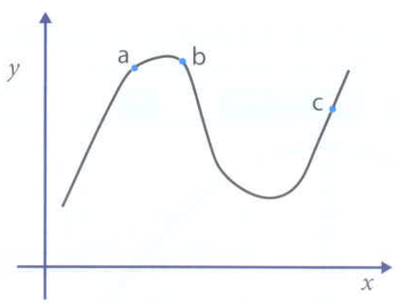
If we generalise this idea to other curves, this becomes:



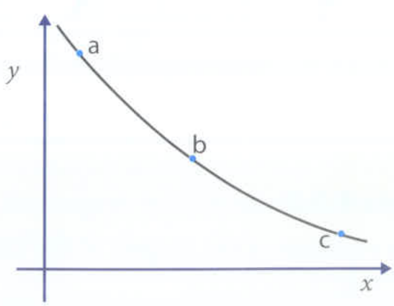
Exercise E.3.1

Estimate the gradients of the tangents and normals at the points indicated on each of these curves:

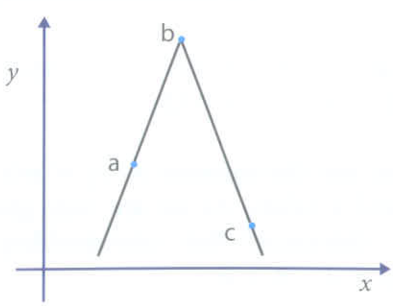
1.



2.



3.



We now move on to finding the equation of a straight line through a given point - as in this example:

Example E.3.1

Find the equation of the tangent to the curve $y = 5 - x^2$ at the point (1, 4).

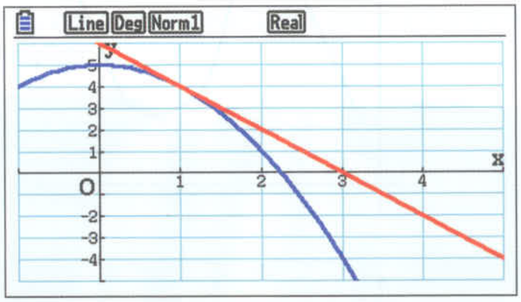
1. Given that $y = 5 - x^2 \Rightarrow \frac{dy}{dx} = -2x$.
2. Then, for $x = 1$, we have $\frac{dy}{dx} = -2(1) = -2$.

Therefore, using $y - y_1 = m(x - x_1)$, with $m = -2$ and $(x_1, y_1) \equiv (1, 4)$, we have the equation of the tangent given by:

$y - 4 = (-2)(x - 1) \Leftrightarrow y - 4 = -2x + 2$

That is, $y = -2x + 6$

Entering both these functions on a calculator suggests (does not prove!) that this answer is correct.



Example E.3.2

Find the equation of the tangent to the curve $y = x^3 - 8$ where $x = 2$.

Given that $y = x^3 - 8 \Rightarrow y' = 3x^2$. Then, for $x = 2$, $y' = 3 \times 2^2 = 12$, i.e. $m = 12$.

In order to use the equation $y - y_1 = m(x - x_1)$ we need both x- and y-values. As we are only given the x-value, we now determine the corresponding y-value, i.e. $x = 2 \Rightarrow y = 2^3 - 8 = 0$.

With $(x_1, y_1) \equiv (2, 0)$ the equation of the tangent is: $(y - 0) = 12(x - 2) \Leftrightarrow y = 12x - 24$.

Example E.3.3

Find the equation of the tangent to the curve:

$f(x) = \frac{2}{\sqrt{x}}, x > 0$ where $x = 4$.

$f(4) = \frac{2}{\sqrt{4}} = 1$ so the point is (4,1).

Before trying to differentiate the function, write it in x^n form:

$f(x) = \frac{2}{\sqrt{x}} = 2x^{-1/2}$

Next, use the standard derivative:

$f'(x) = 2 \times -\frac{1}{2}x^{-3/2} = -x^{-3/2} = -\frac{1}{x\sqrt{x}}$

It follows that: $f'(4) = -\frac{1}{4\sqrt{4}} = -\frac{1}{8}$

Next, using $y = mx + c$.

$y = -\frac{1}{8}x + c$

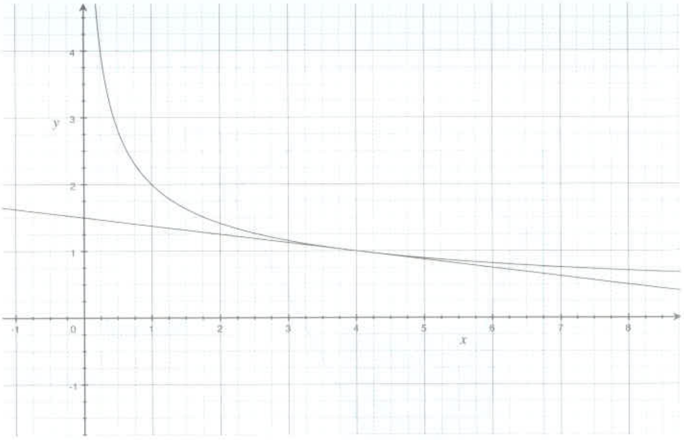
The tangent passes through (4,1) so that:

$1 = -\frac{1}{8} \times 4 + c \Rightarrow 1 = -\frac{1}{2} + c \Rightarrow c = \frac{3}{2}$

The tangent is:

$y = -\frac{1}{8}x + \frac{3}{2} \Rightarrow 8y + x = 12$

This time we will use a computer graphing package to check the answer.



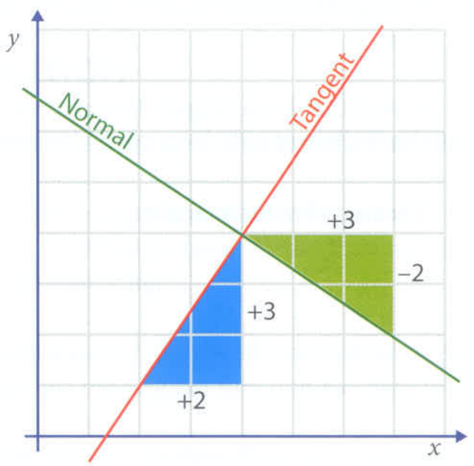
Equation of Normal

To find the equation of the normal at the point (x_1, y_1) we first need to determine the gradient of the tangent, m_t, and then use the relationship between the gradients of two perpendicular lines (given that the normal is perpendicular to the tangent).

To find the equation of the normal we need to repeat the gradient calculation but with a gradient determined by the condition that the product of the gradients of perpendicular lines is –1:

$$m \times m' = -1$$

This relationship is illustrated below for the case in which the gradient of the tangent is $^3/_2$ and that of the normal is $-^2/_3$:



You should draw your own diagram to satisfy yourself that this relationship is always true.

Example E.3.4

Find the equation of the normal to the curve $y = 2x^3 - x^2 + 1$ at the point $(1, 2)$.

First determine the gradient of the tangent: $\frac{dy}{dx} = 6x^2 - 2x$.

At $x = 1$, we have $\frac{dy}{dx} = 6(1)^2 - 2(1) = 4$. That is, $m_t = 4$.

We can now determine the gradient of the normal:

using $m_N = -\frac{1}{m_t}$ we have $m_N = -\frac{1}{4}$.

Using the equation of a straight line, $y - y_1 = m(x - x_1)$

where $(x_1, y_1) \equiv (1, 2)$ and $m = -\frac{1}{4}$

we have that $y - 2 = -\frac{1}{4}(x - 1) \Leftrightarrow 4y - 8 = -x + 1$

Hence the equation of the normal is given by $4y + x = 9$.

Example E.3.5

Find the equation of the normal to the curve:

$f(x) = \frac{x-3}{x\sqrt{x}}, x > 0$ at $x = 9$.

The coordinates of the point are found as follows:

$f(9) = \frac{9-3}{9\sqrt{9}} = \frac{6}{9 \times 3} = \frac{2}{9}$ so the point is: $\left(9, \frac{2}{9}\right)$.

Before finding the derivative, write the function in the form x^n.

$$f(x) = \frac{x-3}{x\sqrt{x}} = \frac{x-3}{x^{3/2}} = x^{-1/2} - 3x^{-3/2}$$

So that: $f'(x) = -\frac{1}{2}x^{-3/2} + \frac{9}{2}x^{-5/2}$

$$= -\frac{1}{2x\sqrt{x}} + \frac{9}{2x^2\sqrt{x}}$$

The gradient of the tangent at $x = 9$ is:

$$f'(9) = -\frac{1}{2 \times 9\sqrt{9}} + \frac{9}{2 \times 9^2\sqrt{9}}$$

$$= -\frac{1}{54} + \frac{1}{54}$$

$$= 0$$

It follows that the tangent is horizontal and the normal is vertical with equation $x = 9$.

Exercise E.3.2

1. Find the equations of the tangents to the following curves at the points indicated:

   a $y = x^3 - x^2 - x + 2$ at (2, 4)

   b $y = x^4 - 4x^2 + 3$ at (1, 0)

   c $y = \sqrt{x+1}$ at (3, 2)

   d $y = \sqrt{x}$ at (9, 3)

e $y=\frac{x+2}{x}, x>0$ at (1,3).

f $y=\frac{x+1}{\sqrt{x}}, x>0$ at (1,2).

g $y=(2x+1)^2$ at (−1,1).

h $y=(x+1)^3$ at (1,8).

i $y=\left(x+\frac{1}{x}\right)^2, x\neq 0$

2. Find the equation of the normal for each of the curves in Question 1.

3. Find the equation of the tangent to the curve $y = x^2(x^2-1)$ at the point A(2, 12).

The tangent at a second point, B(−2,12), intersects the tangent at A at the point C. Determine the type of triangle enclosed by the points A, B and C. Show that the tangents drawn at the points X and Y, where $x = a$ and $x = -a$ respectively will always meet at a third point Z which will lie on the y-axis.

4. Find the equation of the tangent and the normal to the curve $x \mapsto x + \frac{1}{x}, x \neq 0$ at the point (1, 2).

Find the coordinates of the points where the tangent and the normal cross the x- and y-axes, and hence determine the area enclosed by the x-axis, the y-axis, the tangent and the normal.

5. Find the equation of the tangent to the curve $y = x^2 - 2x$ that is parallel to the line with equation $y = 4x + 2$.

6. The straight line $y = -x + 4$ cuts the parabola with equation $y = 16 - x^2$ at the points A and B.

a Find the coordinates of A and B.

b Find the equation of the tangents at A and B, and hence determine where the two tangents meet.

7. The line L and the curve C are defined as follows:

$L: y = 4x - 2$ and $C: y = mx^3 + nx^2 - 1$

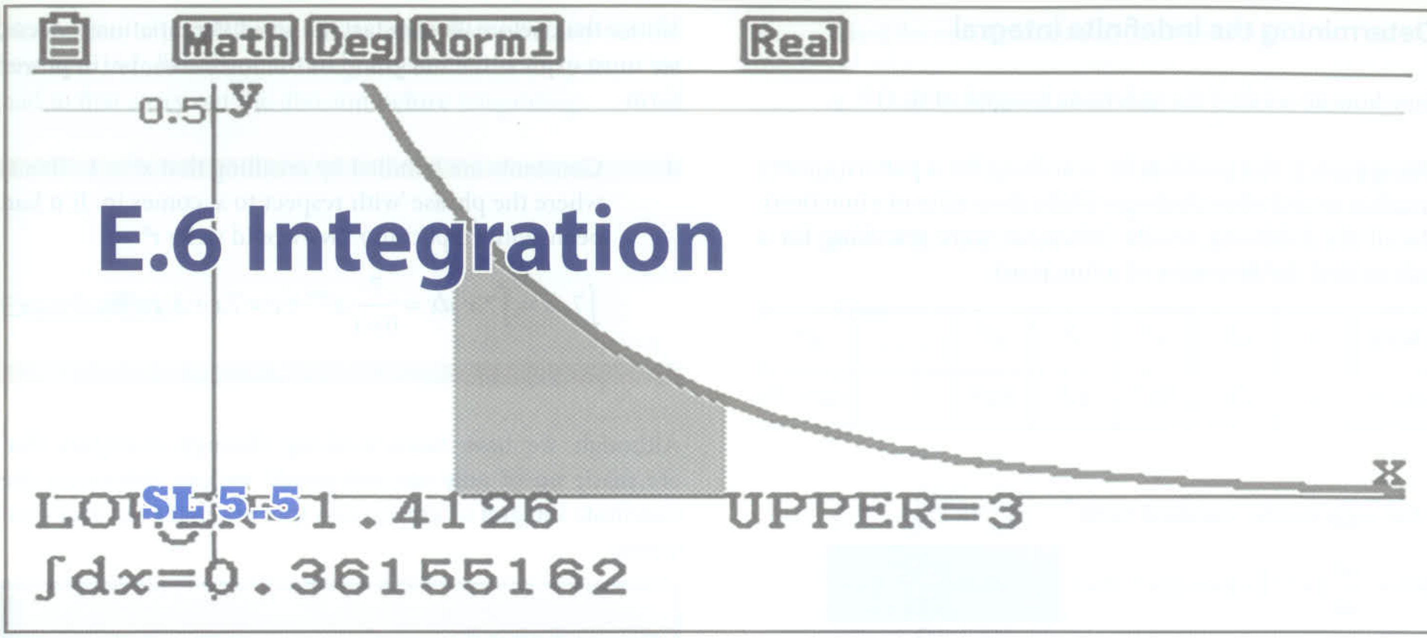
The line L is a tangent to the curve C at $x = 1$.

a Using the fact that L and C meet at $x = 1$, show that $m + n = 3$.

b Given that L is a tangent to C at $x = 1$, show that $3m + 2n = 4$.

c Hence, solve for m and n.

Answers



E.6 Integration

SL 5.5

Integration

Antidifferentiation and the indefinite integral

As the name suggests, ***anti*differentiation** is the reverse process of differentiation. We are then searching for the answer to the following:

Given an expression $f'(x)$ (i.e. the derivative of the function $f(x)$), what must the original function $f(x)$ have been?

For example, if $f'(x) = 2x$ then $f(x) = x^2$ is a *possible expression* for the original function. Why do we say ' *... is a possible expression for the original function*.'?

Consider the following results:

$f(x) = x^2 + 3, \Rightarrow f'(x) = 2x$ [1] & $f(x) = x^2 - 5, \Rightarrow f'(x) = 2x$ [2]

From equations 1 and 2 we see that given an expression $f'(x)$, there are a number of possible different original functions, $f(x)$. This is due to the fact that the derivative of a constant is zero and so when we are given an expression for $f'(x)$, there is no real way of knowing if there was a constant in the original function or what that constant might have been (unless we are given some extra information).

The best that we can do at this stage is to write the following:

Given that $f'(x) = 2x$, then $f(x) = x^2 + c$, where c is some real number that is yet to be determined (it could very well be that $c = 0$).

The antidifferentiation process described above can be summarized as follows:

Given that $\frac{dy}{dx} = f'(x)$, then (after antidifferentiating):

$$y = f(x) + c \text{ where } c \in \mathbb{R}$$

We say that $y = f(x) + c$ where $c \in \mathbb{R}$ is the **antiderivative** of $f'(x)$.

Language and notation

The set of all antiderivatives of a function $h(x)$ is called the **indefinite integral** of $h(x)$, and is denoted by $\int h(x)dx$.

The symbol $\int$ is called the **integral sign**, the function $h(x)$ is the **integrand** of the integral and x is the **variable of integration**.

Once we have found an antiderivative (or indefinite integral) of $h(x)$, $H(x)$ (say) we can then write:

$$\int h(x)dx = H(x) + c, \text{ where } c \in \mathbb{R}$$

The constant c is called the **constant of integration**. The above result is read as:

'The **antiderivative** of $h(x)$ with respect to x is $H(x) + c$, where $c \in \mathbb{R}$'.

or

'The **indefinite integral** of $h(x)$ with respect to x is $H(x) + c$, where $c \in \mathbb{R}$'.

Determining the indefinite integral

So—how do we find the indefinite integral of $h(x)$?

We approach this problem by searching for a pattern (pretty much as we did when dealing with the derivative of a function). Recall the following results (when we were searching for a rule to find the derivative of a function):

$h(x)$	x	x^2	x^3	x^4	x^5	...	x^n
$h'(x)$	1	$2x$	$3x^2$	$4x^3$	$5x^4$	...	nx^{n-1}

This suggests the 'standard form':

Since: $\frac{d}{dx}(x^{n+1})=(n+1)x^n$ then $\int x^n\,dx=\frac{x^{n+1}}{n+1}...(+c)$

Note that, since we cannot have a zero denominator, n cannot be -1.

The case $\int x^{-1}\,dx=\int\frac{1}{x}dx$ will be dealt with later.

A slightly more general result is one where we have ax^n rather than simply x^n. In this case we have that:

$$\int ax^n\,dx=\frac{ax^{n+1}}{n+1}...(+c),n\neq-1$$

Example E.6.1

Find the indefinite integral with respect to x of:

a $4x^2$ b x^{-3}

c $\frac{-3}{x^2}$ d 7

In each case we use the standard form. That is, we first increase the power by one and then divide by the new power.

a $\int 4x^2dx = \frac{4}{2+1}x^{2+1}+c, c\in\mathbb{R} = \frac{4}{3}x^3+c, c\in\mathbb{R}$

b $\int x^{-3}dx = \frac{1}{-3+1}x^{-3+1}+c, c\in\mathbb{R} = -\frac{1}{2}x^{-2}+c, c\in\mathbb{R}$

$= -\frac{1}{2x^2}+c, c\in\mathbb{R}$

c $\int\frac{-3}{x^2}dx=\int-3x^{-2}=\frac{-3}{-2+1}x^{-2+1}+c=\frac{3}{x}+c,c\in\mathbb{R}$

Notice that, before we can start the antidifferentiation process, we must express the integrand in the form ax^n, i.e. in power form.

d Constants are handled by recalling that $x^0 = 1$. This is where the phrase 'with respect to x' comes in. If it had been 'with respect to r', we would write r^0.

$$\int 7\,dx=\int 7x^0dx=\frac{7}{0+1}x^{0+1}+c=7x+c,c\in\mathbb{R}$$

Although we have been working through examples that are made up of only one integrand, we can determine the indefinite integral of expressions that are made up of several terms.

Example E.6.2

Find:

a $\int(2x^2+x^3-4)dx$

b $\int(x-1)(x^4+3x)dx$

c $\int\frac{z^4-2z^2+3}{z^2}dz$

a $\int(2x^2+x^3-4)dx = \int 2x^2dx+\int x^3dx-\int 4dx$

$= \frac{2}{2+1}x^{2+1}+\frac{1}{3+1}x^{3+1}-4x+c$

$= \frac{2}{3}x^3+\frac{1}{4}x^4-4x+c, c\in\mathbb{R}$

When determining the indefinite integral of 4, we have actually thought of '4' as '$4x^0$'.

So that $\int 4dx = \int 4x^0dx = \frac{4}{0+1}x^{0+1} = 4x$.

b $\int(x-1)(x^4+3x)dx = \int(x^5-x^4+3x^2-3x)dx$

$= \frac{1}{6}x^6-\frac{1}{5}x^5+x^3-\frac{3}{2}x^2+c$

c $\int\frac{z^4-2z^2+3}{z^2}dz = \int\left(\frac{z^4}{z^2}-\frac{2z^2}{z^2}+\frac{3}{z^2}\right)dz$

$= \int(z^2-2+3z^{-2})dz$

$= \frac{1}{3}z^3-2z+\frac{3}{-1}z^{-1}+c$

$= \frac{1}{3}z^3-2z-\frac{3}{z}+c, c\in\mathbb{R}$

Notice that in part b it was necessary to first multiply out the brackets **before** we could integrate. Similarly, for part c we had to first carry out the division **before** integrating.

Exercise E.6.1

1. Find the indefinite integral of the following.

   a x^3 b x^7 c x^5

   d x^8 e $4x^2$ f $7x^5$

   g $9x^8$ h $\frac{1}{2}x^3$

2. Find:

   a $\int 5dx$ b $\int 3dx$

   c $\int 10dx$ d $\int \frac{2}{3}dx$

   e $\int -4dx$ f $\int -6dx$

   g $\int -\frac{3}{2}dx$ h $\int -dx$

3. Find:

   a $\int (1-x)dx$ b $\int (2+x^2)dx$

   c $\int (x^3-9)dx$ d $\int \left(\frac{2}{5}+\frac{1}{3}x^2\right)dx$

   e $\int \frac{2x-3}{x^3}dx$ f $\int \frac{(2x-3)^2}{x^4}dx$

   g $\int x(x+2)dx$ h $\int x^2\left(3-\frac{2}{x}\right)dx$

   i $\int (x+1)(1-x)dx$

4. Find the antiderivative of the following.

   a $(x+2)(x-3)$ b $(x^2-3x)(x+1)$

   c $(x-3)^3$ d $(x+2x^3)(x+1)$

   e $(1-x)(1+x)$ f $\left(\frac{x}{4}-3\right)^3$

5. Find:

   a $\int \frac{x^2-3x}{x}dx$

   b $\int \frac{4u^3+5u^2-1}{u^2}du$

   c $\int \frac{(x+2)^2}{x^4}dx$

   d $\int \frac{x^2+5x+6}{x+2}dx$

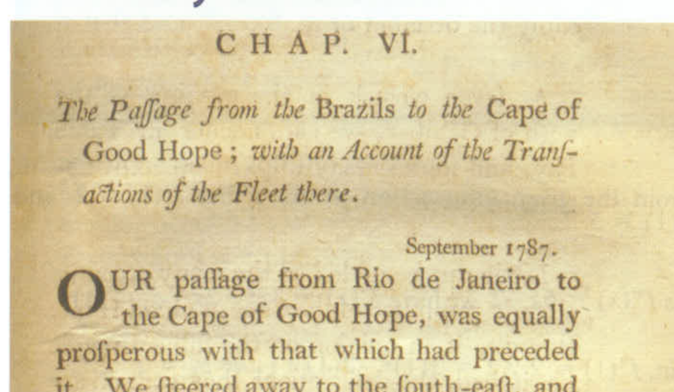
   e $\int \frac{x^2-6x+8}{x-2}dx$

   f $\int \left(\frac{t^2+1}{t}\right)^2 dt$

6. Given that $\frac{d}{dx}\left(\sqrt{x^2-1}\right)=\frac{x}{\sqrt{x^2-1}}$

   Find: $\int \frac{2x}{\sqrt{x^2-1}}dx$

Boundary Conditions


CHAP. VI.

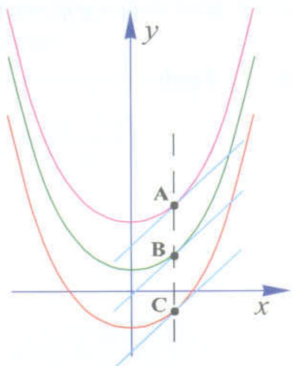
The Paſſage from the Brazils to the Cape of Good Hope; with an Account of the Tranſactions of the Fleet there.

September 1787.

OUR paſſage from Rio de Janeiro to the Cape of Good Hope, was equally proſperous with that which had preceded it. We ſteered away to the ſouth-eaſt, and

Our picture is from a book printed in 1789. Note the elongated 's's in the word 'passage'. These were common in both handwriting and printing at the time. The integral sign ∫ was chosen as an 'S' for 'summation'. The reason for this is part of the subject of this section.

Although we have already discussed the reason for adding a constant, c, when finding the indefinite integral, it is important that we can also determine the value of c.



We show the family of curves resulting from:

$$\frac{dy}{dx} = 2x \Rightarrow y = x^2 + c$$

To determine which of these curves is the one that we actually require, we must be provided with some extra information. In this case we would need to be given the coordinates of a point on the curve.

Example E.6.3

Find $f(x)$ given that $f'(x) = 2x$ and that the curve passes through the point (2, 9).

As $f'(x) = 2x \Rightarrow f(x) = x^2 + c$.

Using the fact that at $x = 2$, $y = 9$, or that $f(2) = 9$, we have $9 = 2^2 + c \Leftrightarrow c = 5$.

Therefore, of all possible solutions of the form $y = x^2 + c$, the function satisfying the given information is $f(x) = x^2 + 5$.

Example E.6.4

Find $f(x)$ given that $f''(x) = 6x - 2$ and that the gradient at the point (1, 5) is 2.

From the given information we have that $f'(1) = 2$ and $f(1) = 5$.

As $f''(x) = 6x - 2$ we have $f'(x) = 3x^2 - 2x + c_1$ – (1)

But, $f'(1) = 2 \therefore 2 = 3(1)^2 - 2(1) + c_1 \Leftrightarrow c_1 = 1$
(i.e. substituting into (1))

Therefore, we have $f'(x) = 3x^2 - 2x + 1$.

Next, from $f'(x) = 3x^2 - 2x + 1$ we have:

$f(x) = x^3 - x^2 + x + c_2$ – (2)

But, $f(1) = 5 \therefore 5 = (1)^3 - (1)^2 + (1) + c_2 \Leftrightarrow c_2 = 4$
(i.e. substituting into (2))

Therefore, $f(x) = x^3 - x^2 + x + 4$

Sometimes, information is not given in the form of a set of coordinates. Information can also be 'hidden' in the context of the problem.

Example E.6.5

The rate of change in pressure, p units, at a depth x cm from the surface of a liquid is given by $p'(x) = 0.03x^2$. If the pressure at the surface is 10 units, find the pressure at a depth of 5 cm.

Antidifferentiating both sides with respect to x, we have $\int p'(x)dx = \int 0.03x^2 dx$.

$\therefore p(x) = 0.01x^3 + c$ – (1)

At $x = 0$, $p = 10$. Substituting into (1) we have:

$10 = 0.01(0)^3 + c \Leftrightarrow c = 10$

Therefore, the equation for the pressure at a depth of x cm is $p(x) = 0.01x^3 + 10$.

At $x = 5$, we have $p(5) = 0.01(5)^3 + 10 = 11.25$. That is, the pressure is 11.25 units.

Exercise E.6.2

1. Find the equation of the function in each of the following.

   a $f'(x) = 2x + 1$, given that the curve passes through (1, 5).

   b $f'(x) = 2 - x^2$ and $f(2) = \frac{7}{3}$.

c $\frac{dy}{dx} = 2x - 4$, given that the curve passes through (4, 0).

d $f'(x) = x - \frac{1}{x^2} + 2$, and $f(1) = 2$

e $\frac{dy}{dx} = 3(x+2)^2$, given that the curve passes through (0, 8).

f $\frac{dy}{dx} = \frac{1}{x^2} + x - 3$, given that the curve passes through (1, 0).

g $f'(x) = (x+1)(x-1) + 1$, and $f(0) = 1$

h $f'(x) = 4x^3 - 3x^2 + 2$, and $f(-1) = 3$

2. Find the equation of the function $f(x)$ given that it passes through the point (–1,2) and is such that $f'(x) = ax + \frac{b}{x^2}$, where $f(1) = 4$ and $f(1) = 0$.

3. The marginal cost for producing x units of a product is modelled by the equation $C'(x) = 30 - 0.06x$. The cost per unit is \$40. How much will it cost to produce 150 units?

4. If $\frac{dA}{dr} = 6 - \frac{1}{r^2}$, and $A = 4$ when $r = 1$,

find A when $r = 2$.

5. The rate, in cm^3/sec, at which the volume of a sphere is increasing is given by the relation $\frac{dV}{dt} = 4\pi(2t+1)^2, 0 \le t \le 10$.

If initially the volume is π cm^3, find the volume of the sphere when $t = 2$.

6. The rate of change of the number of deer, N, in a controlled experiment, is modelled by the equation $\frac{dN}{dt} = 3\sqrt{t^3} + 2t, 0 \le t \le 5$.

There are initially 200 deer in the experiment. How many deer will there be at the end of the experiment?

7. If $\frac{dy}{dx} \propto x$, find an expression for y, given that $y = 7$ when $x = 1$ and $y = 16$ when $x = 2$.

8. A function with gradient defined by $\frac{dy}{dx} = 4x - m$ at any point P on its curve passes through the point (2, –6) with a gradient of 4. Find the coordinates of its turning point.

9. The marginal revenue is given by $\frac{dR}{dx} = 25 - 10x + x^2, x \ge 0$, where R is the total revenue and x is the number of units demanded.

Find the equation for the price per unit, $P(x)$.

Extension

10. The rate of growth of a culture of bacteria is modelled by the equation $200t^{1.01}$, $t \ge 0$, t hours after the culture begins to grow. Find the number of bacteria present in the culture at time t hours if initially there were 500 bacteria.

Extra questions

Why the definite integral?

Unlike the previous section where the indefinite integral of an expression resulted in a new expression, when finding the definite integral we produce a numerical value.

Definite integrals are important because they can be used to find different types of measures, for example, areas, volumes, lengths and so on. It is, in essence, an extension of the work we have done in the previous sections.

Language and notation

If the function $f(x)$ is continuous at every point on the interval [a,b] and $F(x)$ is any antiderivative of $f(x)$ on [a,b], then:

$\int_a^b f(x)dx$ is called the definite integral and is equal to $F(b)-F(a)$.

That is, $\int_a^b f(x)dx = F(b)-F(a)$.

Which is read as:

"the integral of $f(x)$ with respect to x from a to b is equal to $F(b) - F(a)$."

Usually we have an intermediate step to aid in the evaluation of the definite integral. This provides a somewhat 'compact recipe' for the evaluation process. This intermediate step is written as $[F(x)]_a^b$.

We therefore write

$$\int_a^b f(x)dx = [F(x)]_a^b = F(b)-F(a)$$

The process is carried out in four steps:

1. Find an indefinite integral of $f(x)$, $F(x)$ (say).
2. Write your result as $[F(x)]_a^b$.
3. Substitute a and b into $F(x)$.
4. Subtract: $F(b)-F(a)$ to obtain the numerical value.

Notice that the constant of integration c is omitted. This is because it would cancel itself out upon carrying out the subtraction: $(F(b)+c)-(F(a)+c) = F(b)-F(a)$.

In the expression $\int_a^b f(x)dx$, x is called the variable of integration, and a and b are called the lower limit and upper limit respectively. It should also be noted that there is no reason why the number b need be greater than the number a when finding the definite integral.

That is, it is just as reasonable to write $\int_{-3}^{2} f(x)dx$ as it is to write $-\int_{2}^{-3} f(x)dx$, both expressions are valid.

Example E.6.6

Evaluate the following:

a $\int_{-1}^{2}(x^2-2x+1)dx$ b $C \in$.

c $\int_5^2 (3x-4)^4 dx$

a $\int_{-1}^{2}(x^2-2x+1)dx = \left[\frac{x^3}{3}-x^2+x\right]_{-1}^{2}$

$= \left(\frac{2^3}{3}-2^2+2\right)-\frac{(-1)^3}{3}+(-1)^2-(-1)$

$= \frac{8}{3}-4+2+\frac{1}{3}+1+1$

$= 3$

b $\int_2^4 \left(x+\frac{1}{x}\right)^2 dx = \int_2^4 \left(x^2+2+\frac{1}{x^2}\right)dx = \left[\frac{1}{3}x^3+2x-\frac{1}{x}\right]_2^4$

$= \left(\frac{1}{3}(4)^3+2(4)-\frac{1}{4}\right)-\left(\frac{1}{3}(2)^3+2(2)-\frac{1}{2}\right)$

$= \left(\frac{64}{3}+8-\frac{1}{4}\right)-\left(\frac{8}{3}+4-\frac{1}{2}\right)$

$= \frac{275}{12} (\approx 22.92)$

c $\int_5^2 (3x-4)^4 dx = \left[\frac{1}{3(5)}(3x-4)^5\right]_5^2 = \frac{1}{15}((2)^5-(11)^5)$

$= -\frac{161019}{15} = -10734.6$

Example E.6.7

The production rate for radios by the average worker at Bat-Rad Pty Ltd t hours after starting work at 7:00 a.m., is given by $N'(t) = -2t^2+8t+10, 0 \le t \le 4$.

How many units can the average worker assemble in the second hour of production?

The second hour starts at $t = 1$ and ends at $t = 2$. Therefore, the number of radios assembled by the average worker in the second hour of production is given by:

$$N = \int_{t=1}^{t=2} N'(t)dt.$$

That is, $N = \int_1^2 (-2t^2 + 8t + 10)dt = \left[-\frac{2}{3}t^3 + 4t^2 + 10t\right]_1^2$

$$= \left(-\frac{2}{3}(2)^3 + 4(2)^2 + 10(2)\right) - \left(-\frac{2}{3}(1)^3 + 4(1)^2 + 10(1)\right)$$

$$= \frac{52}{3}$$

Beware of situations in which an integral may have values missing from its domain.

$\int_{-2}^{1} \frac{1}{x^2}dx$ is undefined because $x = 0$ gives a zero denominator.

Exercise E.6.3

1. Evaluate the following.

   a $\int_1^4 x\,dx$ b $\int_{-2}^{2} (2x+1)^2 dx$

   c $\int_2^3 \frac{2}{x^3}dx$ d $\int_{-2}^{-1} \frac{x+2}{x^3}dx$

2. Evaluate the following definite integrals (giving exact answers).

   a $\int_1^2 \left(x^2 - \frac{3}{x^4}\right)dx$ b $\int_{-2}^{-1} \frac{(3x+2)^2}{x^4}dx$

   c $\int_0^2 (1 + 2x - 3x^2)dx$ d $\int_{-2}^{0} (x+1)dx$

   e $\int_0^{-1} x^3(x+1)dx$ f $\int_{-1}^{1} (x+1)(x^2-1)dx$

   g $\int_{-2}^{3} \frac{1}{x^3}dx$ h $\int_1^2 \left(x - \frac{1}{x}\right)^2 dx$

   i $\int_1^3 \left(\frac{x^3 - x^2 + x}{x}\right)dx$ j $\int_{-1}^{1} (x - x^3)dx$

   k $\int_{-2}^{3} \frac{1}{x-1}dx$

3. Evaluate the following definite integrals (giving exact values).

   a $\int_0^1 (x+1)^4 dx$ b $\int_0^1 \frac{1}{x^2}dx$

   c $\int_{-1}^{2} (1-2x)^3 dx$ d $\int_0^1 (3x-2)^4 dx$

4. Given that $\int_a^b f(x)dx = m$ and $\int_a^b g(x)dx = n$, find:

   a $\int_a^b 2f(x)dx - \int_a^b g(x)dx$ b $\int_a^b (f(x) - 1)dx$

Extension Questions

5. Given that $\frac{d}{dx}\left(\frac{1}{\sqrt{x^2-4}}\right) = \frac{-x}{(x^2-4)\sqrt{x^2-4}}$

   find the exact value of: $\int_3^5 \frac{x}{(x^2-4)\sqrt{x^2-4}}dx$.

6. Given that $\frac{d}{dx}\left(\sqrt{\sqrt{x}+1}\right) = \frac{1}{4\sqrt{x}\left(\sqrt{\sqrt{x}+1}\right)}$

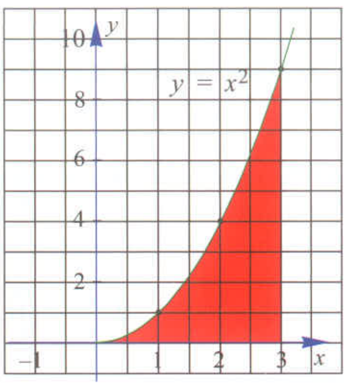
   find the exact value of: $\int_2^3 \frac{1}{\sqrt{x}\sqrt{\sqrt{x}+1}}dx$.

Areas

In Chapter E1 we saw that differentiation had a geometric meaning, that is, it provided a measure of the gradient of the curve at a particular point. We have also seen applications of the definite integral throughout the previous sections in this chapter. In this section we will investigate the geometric significance of the integral.

Introduction to the area beneath a curve

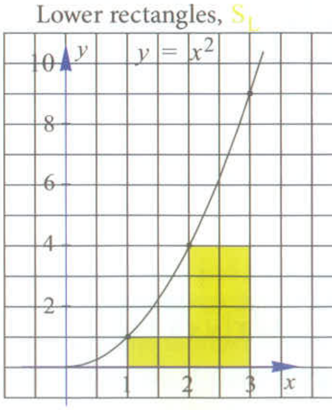
Consider the problem of finding the exact value of the red shaded area, A sq. units, in the diagram shown.



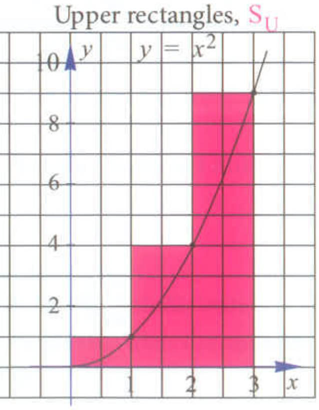
As a first step we make use of rectangular strips as shown below to obtain an approximation of the shaded area. We can set up a table of values, use it to find the area of each strip and then sum these areas.

In the figure, the green rectangles lie below the curve, and so we call these the lower rectangles. In the figure, magenta rectangles lie above the curve, and so we call these the upper rectangles. The red figure above shows that the true area (or exact area) lies somewhere between the sum of the areas of the lower rectangles, S_L, and the sum of the areas of the upper rectangles, S_U.

Lower rectangles, S_L



Upper rectangles, S_U



That is, we have that:

Lower Sum = S_L < Exact Area, A < S_U = Upper Sum

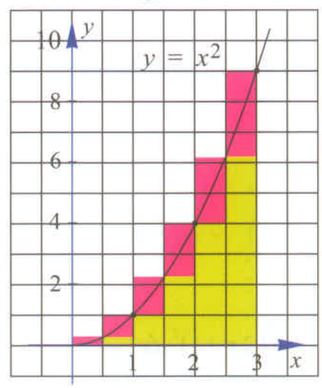
We have that $S_L = 1\times1 + 1\times4 = 5$ and
$S_U = 1\times1 + 1\times4 + 1\times9 = 14$.

Therefore, we can write $5 < A < 14$. However, this does seem to be a poor approximation as there is a difference of 9 sq. units between the lower approximation and the upper approximation. The problem lies in the fact that we have only used two rectangles for the lower sum and three rectangles for the upper sum. We can improve on our approximation by increasing the number of rectangles that are used. For example, we could use 5 lower rectangles and 6 upper rectangles, or 10 lower rectangles and 12 upper rectangles and so on.

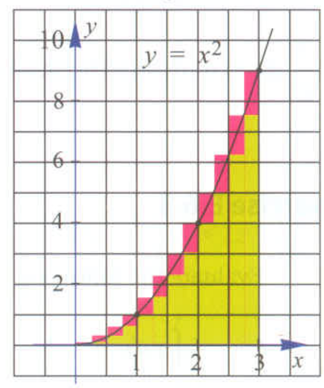
In search of a better approximation

As shown in the diagrams below, as we increase the number of rectangular strips (or decrease the width of each strip) we obtain better approximations to the exact value of the area.

5 lower & 6 upper rectangles



11 lower & 12 upper rectangles



For intervals of width 0.5 we have:

$$S_L = \frac{1}{2}\times[0.25 + 1 + 2.25 + 4 + 6.25] = 6.875$$

$$S_U = \frac{1}{2}\times[0.25 + 1 + 2.25 + 4 + 6.25 + 9] = 11.375$$

So: 6.875 < True area < 11.375

For intervals of width 0.25 we have:

$$S_L = \frac{1}{4}\times[0.25 + 0.5625 + 1 + 1.5625 + 2.25 + \ldots + 7.5625]$$
$$\approx 7.89$$

$$S_U = \frac{1}{4}\times[0.0625 + 0.25 + 0.5625 + \ldots + 7.5625 + 9] \approx 10.16$$

So: 7.56 < True area < 10.16

By continuing in this manner, the value of A will become *sandwiched* between a lower value and an upper value. Of course the more intervals we have the '*tighter*' the sandwich will be! What we can say is that if we partition the interval [0,3] into n equal subintervals, then, as the number of rectangles increases, S_L increases towards the exact value A while S_U decreases towards the exact value A.

That is, $\lim_{n \to \infty} S_L = A = \lim_{n \to \infty} S_U$

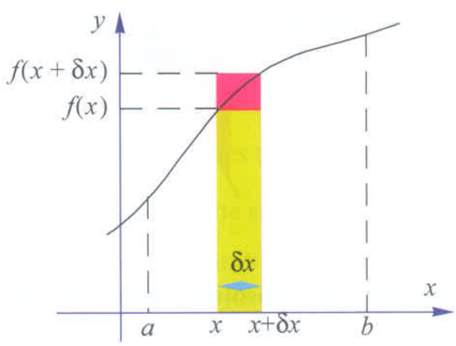
As we have seen, even for a simple case such as $y = x^2$, this process is rather tedious. And as yet, we still have not found the exact area of the shaded region under the curve $y = x^2$ over the interval [0,3].

Towards an exact area

We can produce an algebraic expression to determine the exact area enclosed by a curve. We shall also find that the definite integral plays a large part in determining the area enclosed by a curve.

As a starting point we consider a single rectangular strip.

Consider the function $y = f(x)$ as shown:



Divide the interval from $x = a$ to $x = b$ into n equal parts: $a = x_0, x_1, x_2, \ldots, x_n = b$.

This means that each strip is of width $\frac{b-a}{n}$.

We denote this width by δx so that $\delta x = \frac{b-a}{n}$.

The area of the lower rectangle is $f(x) \times \delta x$ and that of the upper rectangle is $f(x + \delta x) \times \delta x$.

Then, the sum of the areas of the lower rectangles for $a \le x \le b$ is

$$S_L = \sum_{x=a}^{b-\delta x} f(x)\delta x$$

and the sum of the areas of the upper rectangles for $a \le x \le b$ is

$$S_U = \sum_{x=a}^{b} f(x+\delta x)\delta x$$

Then, if A sq units is the area under the curve $y = f(x)$ over the interval $[a, b]$ we have that:

$$\sum_{x=a}^{b-\delta x} f(x)\delta x < A < \sum_{x=a}^{b} f(x+\delta x)\delta x$$

As the number of strips increases, that is, as $n \to \infty$ and therefore $\delta x \to 0$ the area, A sq units, approaches a common limit, i.e. S_L from below, and S_U from above. We write this result as:

$$A = \lim_{\delta x \to 0} \sum_{x=a}^{b} f(x)\delta x$$

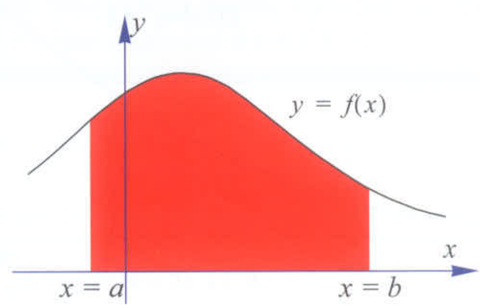
In fact this result leads to the use of the integral sign as a means whereby we can find the required area.

That is, $\lim_{\delta x \to 0} \sum_{x=a}^{b} f(x)\delta x = \int_a^b f(x)dx$

Notice that we've only developed an appropriate notation and a 'recognition' that the definite integral provides a numerical value whose geometrical interpretation is connected to the area enclosed by a curve, the x-axis and the lines $x = a$ and $x = b$. We leave out a formal proof of this result in preference to having developed an intuitive idea behind the concept and relationship between area and the definite integral.

We can now combine our results of the definite integral with its geometrical significance in relation to curves on a Cartesian set of axes.

The definite integral and areas

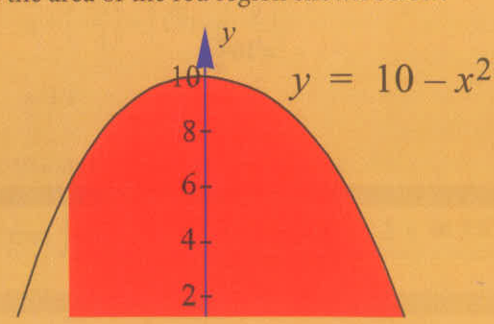


If $y = f(x)$ is **positive** and **continuous** on the interval $[a,b]$, the area, A sq units, bounded by $y = f(x)$, the x-axis and the lines $x = a$ and $x = b$ is given by

$$\text{Area} = A = \int_a^b f(x)dx = \int_a^b y\,dx$$

Example E.6.8

Find the area of the red region shown below:



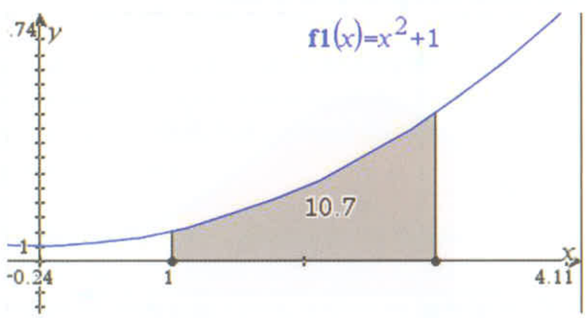
$$\text{Area} = \int_{-2}^{3}(10-x^2)dx = \left[10x-\frac{1}{3}x^3\right]_{-2}^{3}$$
$$= \left((30-9)-\left(-20+\frac{8}{3}\right)\right)$$
$$= \frac{115}{3}$$

Therefore, the red area measures $\frac{115}{3}$ square units.

Example E.6.9

Find the area enclosed by the curve with equation $f(x)=x^2+1$, the x-axis and the lines $x = 1$ and $x = 3$.

Some calculators will produce numeric solutions to these area problems. Make sure you are familiar with the capabilities of your model.



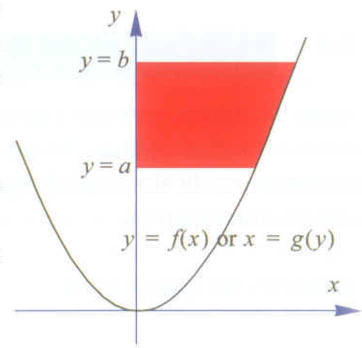
The analytic solution is: $\text{Area} = \int_1^3 (x^2+1)dx$

$$= \left[\frac{x^3}{3}+x\right]_1^3$$
$$= \left(\frac{3^3}{3}+3\right)-\left(\frac{1^3}{3}+1\right)$$
$$= 9+3-\frac{1}{3}-1$$
$$= 10\frac{2}{3}$$

Further observations about areas

To find the area bounded by $y = f(x)$, the y-axis and the lines $y = a$ and $y = b$ we carry out the following process:

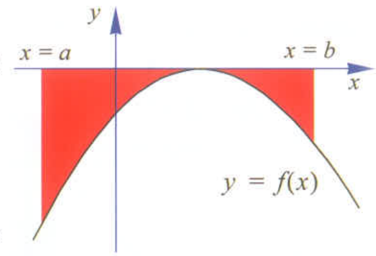
First you need to make x the subject, i.e. from $y = f(x)$ obtain the new equation $x = g(y)$.



Then find the definite integral, $\int_a^b x\,dy = \int_a^b g(y)dy$ sq units, which will give the red area.

If f is **negative** over the interval [a,b] (i.e $f(x) < 0$ for $a \le x \le b$, then the integral:

$\int_a^b f(x)dx$ is a negative number.



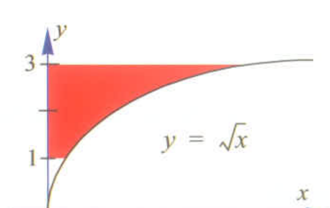
We therefore need to write the area, A, as:

$-\int_a^b f(x)dx$ or, use the **absolute value** of the integral:

$$A = \left|\int_a^b f(x)dx\right|$$

Example E.6.10

Find the area enclosed by the curve $y = \sqrt{x}$, the y-axis and the lines $y = 1$ and $y = 3$.



We need an expression for x in terms of y:

That is, $y = \sqrt{x} \Rightarrow y^2 = x, x > 0$

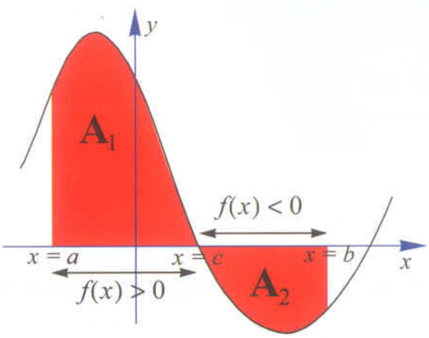
Therefore, the area of the shaded region is:

$$A = \int_1^3 x\,dy = \int_1^3 y^2 dy = \left[\frac{1}{3}y^3\right]_1^3 = \frac{1}{3}(3^3 - 1^3) = \frac{26}{3}.$$

The required area is $\frac{26}{3}$ sq units.

The Signed Area

It is possible for $y = f(x)$ to alternate between negative and positive values over the interval $x = a$ and $x = b$. That is, there is at least one point $x = c$ where the graph crosses the x-axis, and so $y = f(x)$ changes sign when it crosses the point $x = c$.



The integral $\int_a^b f(x)dx$ gives the **algebraic sum** of A_1 and A_2, that is, it gives the **signed area**.

For example, if $A_1 = 12$ and $A_2 = 4$, then the definite integral $\int_a^b f(x)dx = 12 - 4 = 8$.

This is because $\int_a^c f(x)dx = 12$, $\int_c^b f(x)dx = -4$ and so

$$\int_a^b f(x)dx = \int_a^c f(x)dx + \int_c^b f(x)dx = 12 + (-4) = 8$$

As $\int_c^b f(x)dx$ is a negative value, finding the negative of $\int_c^b f(x)dx$, that is $\left(-\int_c^b f(x)dx\right)$, would provide a positive value and therefore be a measure of the area of the region that is shaded below the x-axis. The red area would then be given by:

$$\int_a^c f(x)dx + \left(-\int_c^b f(x)dx\right)$$

This would provide the sum of two positive numbers.

Steps for finding areas

It follows, that in order to find the area bounded by the curve $y = f(x)$, the x-axis and the lines $x = a$ and $x = b$, we first need to find where (and if) the curve crosses the x-axis at some point $x = c$ in the interval $a \le x \le b$. If it does, we must evaluate the area of the regions above and below the x-axis **separately**.

Otherwise, evaluating $\int_a^b f(x)dx$ will provide the signed area (which only gives the correct area if the function lies above the x-axis over the interval $a \le x \le b$).

Therefore, we need to:

1. Sketch the graph of the curve $y = f(x)$ over the interval $a \le x \le b$. (In doing so you will also determine any x-intercepts).
2. Integrate $y = f(x)$ over each region separately (if necessary). That is, regions above the x-axis and regions below the x-axis.
3. Add the required (positive terms).

Example E.6.11

Find the area of the region enclosed by the curve $y = x^3 - 1$, the x-axis and $0 \le x \le 2$.

First sketch the graph of the given curve:

x-intercepts (when $y = 0$): $x^3 - 1 = 0 \Leftrightarrow x = 1$

y-intercepts (when $x = 0$): $y = 0 - 1 = -1$.

From the graph we see that y is negative in the region [0,1] and positive in the region [1,2], therefore the area of the region enclosed is given by:

$$A = \left(-\int_0^1 (x^3 - 1)dx\right) + \int_1^2 (x^3 - 1)dx$$

$$= \left(-\left[\frac{x^4}{4} - x\right]_0^1\right) + \left[\frac{x^4}{4} - x\right]_1^2$$

$$= -\left(-\frac{3}{4}\right) + \left((2) - \left(-\frac{3}{4}\right)\right)$$

$$= 3.5$$

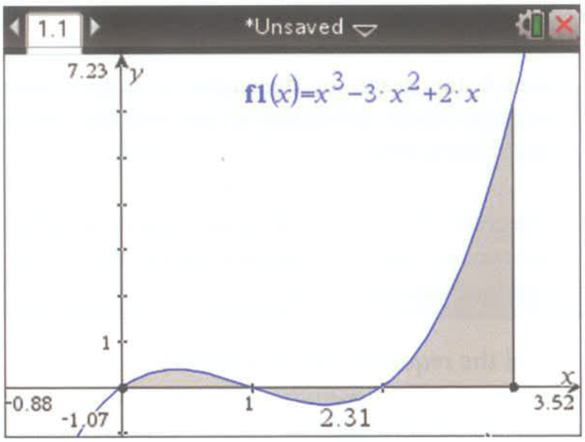
That is, the area measures 3.5 sq units.

Notice that $\int_0^2 (x^3 - 1)dx = \left[\frac{x^4}{4} - x\right]_0^2 = 2 \ (\neq 3.5)$

Example E.6.12

Find the area enclosed by the curve $y = x^3 - 3x^2 + 2x$, the x-axis and the lines $x = 0$ and $x = 3$.

First sketch the graph of the given curve:



x-intercepts (when $y = 0$):

$x^3 - 3x^2 + 2x = 0 \Leftrightarrow x(x-2)(x-1) = 0 \therefore x = 0, 2, 1$

y-intercepts (when $x = 0$): $y = 0 - 0 + 0 = 0$.

From the diagram we have, Area = $A_1 - A_2 + A_3$.

$$A_1 = \int_0^1 (x^3 - 3x^2 + 2x)dx = \left[\frac{x^4}{4} - x^3 + x^2\right]_0^1 = \frac{1}{4}$$

$$A_2 = \int_1^2 (x^3 - 3x^2 + 2x)dx = \left[\frac{x^4}{4} - x^3 + x^2\right]_1^2 = -\frac{1}{4}$$

$$A_3 = \int_2^3 (x^3 - 3x^2 + 2x)dx = \left[\frac{x^4}{4} - x^3 + x^2\right]_2^3 = \frac{9}{4}$$

Therefore, the required area is $\frac{1}{4} - \left(-\frac{1}{4}\right) + \frac{9}{4} = \frac{11}{4}$ sq units.

Exercise E.6.4

1. Find the area of the region bounded by:

   a $y = x^3$, the x-axis, and the line $x = 2$.

   b $y = 4 - x^2$, and the x-axis.

   c $y = x^3 - 4x$, the x-axis, and the lines $x = -2$ and $x = 0$.

   d $y = x^3 - 4x$, the x-axis, the line $x = 2$ and the line $x = 4$.

   e $y = x^2 + x$, the x-axis, and the lines $x = 0$ and $x = 1$.

2. Verify your answers to Question 1, using a graphics calculator.

3. Find the area of the region enclosed by the curve $y = 8 - x^3$, the y-axis and the x-axis.

4. Find the area of the region enclosed by the curve $y = x^2 + 1$, and the lines $y = 2$ and $y = 4$.

5. Find the area of the region enclosed by the curve $y = x^2 - 1$, the x-axis, the line $x = 0$ and $x = 2$.

6. Find the area of the region enclosed by the curve $y = x(x+1)(x-2)$ and the x-axis.

7. Find the area of the region enclosed by the curve $f(x) = 1 - \frac{1}{x^2}$

   a the x-axis, the line $x = 1$ and $x = 2$.

   b the x-axis, the line $x = \frac{1}{2}$ and $x = 2$.

   c and the lines $y = -\frac{1}{2}$ and $y = \frac{1}{2}$.

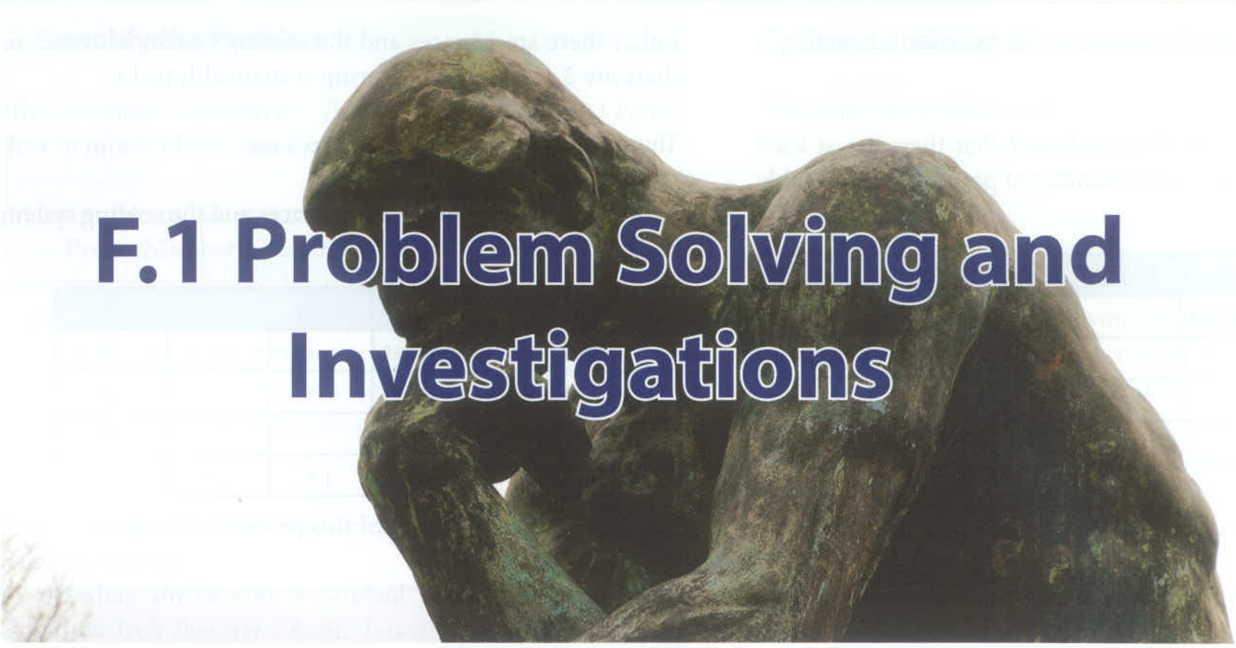
8. The area of the region enclosed by the curve $y^2 = 4ax$ and the line $x = a$ is ka^2 sq units. Find the value of k.

9. a Find the area of the region enclosed by the curve $y = |2x - 1|$, the x-axis, the line $x = -1$ and the line $x = 2$.

   b Find the area of the region enclosed by the curve $y = |2x| - 1$, the x-axis, the line $x = -1$ and the line $x = 2$.

Answers

F.1 Problem Solving and Investigations



This Chapter has been included to help students and teachers complete the 'Toolbox' component of each course. As this comprises 20% of the assessment (externally moderated) and 20% for SL & 12.5% for HL of the teaching time, it is a component that needs to be seriously addressed.

We divide this Chapter into two main sections.

1. **Problem Solving.** We will look at the meaning of this rather puzzling term. We will cover some of the common strategies used to solve problems and offer examples and practice questions.

2. **Investigations.** These require a written report that is externally moderated by IBO staff. We offer advice as to choice of topic, collection and analysis of data and preparation of a report.

Throughout this series of texts, students will find 'Toolbox' sections that provide further suggestions.

Problem Solving

In everyday life we encounter problems that we try to solve as best we can. The kettle has stopped working - use a saucepan on the cook-top. And so on, seemingly endlessly.

Mathematical problem solving is a bit different.

When you are learning the techniques of mathematics, it is usual to practise each technique by completing related examples. In the chapter on sequences and series, there was a section on summing arithmetic series. It was followed, surprise-surprise, by an exercise on summing arithmetic series.

Problem solving exercises tend to offer you a problem with no hints as to the techniques needed to solve it. This is probably the most challenging aspect of the task.

1. Don't Panic!

Just because you cannot see how to do a problem, don't just give up. Beware of staring at the ceiling hoping for the solution to drop into your head.

So what can you do instead?

Try to identify the information in the question. Once you have done that, attempt to get a version of it on paper, a table, a diagram, a graph, etc.

Identify the information that you are asked to find (ie. the question at the end).

Look for any information that you might like to have been told but were not. Above all, put it on paper.

Example F.1.1

Alice, Bella and Candice hold a series of races. They have a whole numbered scoring system with first place scoring more than second which scores more than third which does not score zero. There were no ties.

The first race was 100m. At the end of the competition, Alice had 20 points, Bella with 10 and Candice with 9. If Alice did not win the 200m, who did?

Try to solve this before turning the page.

There is not a lot of information in the question, so getting it on paper is vital.

We do know there are three girls and that there are at least two races. We also know the number of points gained. A table suggests...

	100m	200m	??	??	Total
Alice	Not 1st				20
Bella					10
Candice					9

That is all we know.

Next, what are we asked? - 'Who won the 200m?'

This means we have to work out all the results. At this stage, it seems impossible.

The next question is what don't we know? This is a really important question as it may tip us off as to an intermediate step.

Whilst we do know that there are three girls, we don't know how many races they ran. Hence the question marks in the table.

Importantly also, we don't know the scoring system. A table beckons here too:

	100m	200m	??	??	Total
1st					
2nd					
3rd					

We don't know any of the entries. Neither the number of races nor the scoring system.

So we now focus on what we do not know (the number of races and the scoring system) and compare it with what we do know. This is the points scored by the three girls (20, 10, 9).

What information is hidden here that will lead to our missing data?

If you have not already solved the problem, break off here and think again.

Welcome back!

The total points scored is 20 + 10 + 9 = 39. This has exactly one factorisation: 3 × 13 (other than 1 × 13). Prime Numbers!

What does this mean for the problem?

Either there are 13 races and the scoring system adds to 3 or there are 3 races and the scoring system adds to 13.

The scoring system cannot add to 3 so:

Major Breakthrough: there are 3 races and the scoring system adds to 13.

	100m	200m	??	Total
Alice	Not 1st			20
Bella				10
Candice				9
Total	13	13	13	39

Again, if you have not solved this problem, try again.

From here on, we are looking at one of the stalwarts of problem solving 'guess and check'. We will deal with this method in more detail later.

In essence, we are looking for three numbers that add to 13. Rather like a Sudoku, they must be filled into the above table to give the right totals for each row and column.

The only constraint is that Alice did not win every race.

The key with guess and check is to be systematic. This means that you need to check the possibilities in an ordered manner.

This is one possibility: 3rd: 1 pt, 2nd 2 pt 1st 10 pt. Having chosen the points for 3rd and 2nd (1+2) the points for 1st has to be 10 as the scoring system totals 13. Can this work? No. Just look at Alice who must come 1st twice and in another place in the third race.

How about 3rd: 1 pt, 2nd 3 pt 1st 9 pt? You cannot make 20 out of a set of 3 of these.

Work systematically through all of these (there are not as many as you may first have thought!) and you should arrive at the unique solution:

	100m	200m	??	Total
Alice	8	4	8	20
Bella	1	8	1	10
Candice	4	1	4	9
Total	13	13	13	39

See the Proof Chapter for the meaning of **Exhaustion**. There are not that many cases to check to establish that Bella won the 200m race. After you have solved the problem, it is wise to check back to the question. Have you answered the actual question asked?

2. Simplify the Problem

This is a close cousin to the 'Prime Directive' of 'Don't Panic'. Just to mix Sci Fi metaphors!

Many problems ask you to look at 'A Right Angled Triangle' or a 'Prime Number'. They then ask you to prove that some condition always holds. It can help to, in the example of the right angled triangle, look at a particular right angled triangle such as a 3,4,5. Can you prove it for that? If you can, does your proof generalise to every right angled triangle? This can be an easier proof as you have a strategy that worked for the particular case.

Our example works in a slightly different way. This technique is multi-faceted.

> **Example F.1.2**
>
> Find seven different positive integers a to f such that:
>
> $$\frac{1}{a}+\frac{1}{b}+\frac{1}{c}+\frac{1}{d}+\frac{1}{e}+\frac{1}{f}=1$$

As before, try to solve this without looking at what follows.

As we have said, one of the challenges of this part of Mathematics is that the appropriate techniques may not be that obvious. In this case, it seems that addition of fractions using LCMs is unavoidable. But the LCM here is horrid.

One way of doing this is to look at a simpler version of the same problem.

For example: $\frac{1}{a}+\frac{1}{b}+\frac{1}{c}=1$

Suppose we look at choosing $a = 2$

It follows that $\frac{1}{2}+\frac{1}{b}+\frac{1}{c}=1 \Rightarrow \frac{1}{b}+\frac{1}{c}=\frac{1}{2}$

It does not seem so difficult to find two fractions in the 'one-over' form that add to one half.

For example since $\frac{1}{2}-\frac{1}{3}=\frac{3}{6}-\frac{2}{6}=\frac{1}{6}$ it follows that: $\frac{1}{6}+\frac{1}{3}=\frac{1}{2}$

and so we have the next level of complication of the problem (three one-over fractions that add to one):

$$\frac{1}{2}+\frac{1}{3}+\frac{1}{6}=1$$

Continue to decompose the fractions until you have a (non-unique) solution to the problem.

3. Guess and Check

There are two golden rules here:

1. Be systematic
2. Learn from your mistakes.

The first of these means that you should choose your guesses sensibly, keep a record of the results and not find yourself checking solutions that fail more than once.

The second means that, when a guess fails, as it almost certainly will, why does it fail? Is it too big? Too small? Is it close or way off?

Our example illustrates this.

> **Example F.1.3**
>
> Five loaves of bread are weighed two at a time in all possible combinations. The weights in grams are 310, 312, 313, 314, 315, 316, 317, 318, 320 and 321. How much does each loaf weigh?

The first step is to make sure you understand the problem and the data you have been given. The key is to understand that the five loaves are weighed in pairs. Why does this lead to the 10 results?

Suppose the loaves are called A, B, C, D & E with the corresponding weights being a, b, c, d & e.

The weighing pairs are AB, AC, AD, AE, BC, BD, BE, CD, CE & DE. There are ten of these, corresponding to the data.

Can we simplify the problem? Since all the weights are 'three hundred and something', we could subtract 300 from each and then add 150 back in at the end. Remember that the loaves are weighed in pairs and all the answers will be 'one hundred and fifty something'.

This means that the pair weights become: 10, 12, 13, 14, 15, 16, 17, 18, 20 and 21.

Now ask what we do and do not know. There are 5 loaves, but are there two that weigh the same? Think about this and the consequences. Suppose C & D weigh the same. AC and AD would also weigh the same (as well as CE & DE etc.). As there are no repeated numbers in the pair weights, we conclude that all the weights are different.

Since all the weights are different, it may help to think of them arranged from lightest to heaviest:

$$a < b < c < d < e$$

From here on, it looks like guess and check will need to be employed. As we have said, it is vital to be systematic. Don't just have a wild guess, find it fails and follow up with another unrelated guess. Keep a record of your work so you don't end up going round in circles. For example, record your work in a table.

Guess	*a*	*b*	*c*	*d*	*e*	OK?
1						
2						
3						
4						
5						
6						

Suppose we begin with $a = 1$ (ie. a loaf with weight 151). We now know that $b = 9$ because the two smallest numbers add to 10. Allowing the smallest possible numbers for the other loaves we get the first row as:

Guess	*a*	*b*	*c*	*d*	*e*	OK?
1	1	9	10	11	12	No

The answer 'No' means that the given solution is not correct. This can be verified by looking at the two largest numbers 11 + 12 = 23 (not 21). Rather like we know the smallest numbers add to 10, we also know that the two largest numbers add to 21.

This means that there is no need to consider any more solutions with $a = 1$. What about $a = 2$, $b = 8$?

Guess	*a*	*b*	*c*	*d*	*e*	OK?
1	1	9	10	11	12	No
2	2	8	9	10	11	

This does not fail the largest pair test as 10 + 11 = 21 which is correct. However, it is not possible to make 12 so this is another 'No'. Also, there is no need to check any more $a = 2$ solutions. Next, $a = 3$.

Guess	*a*	*b*	*c*	*d*	*e*	OK?
1	1	9	10	11	12	No
2	2	8	9	10	11	No
3	3	7	8	9	12	

Note that the last pair is 9, 12 in order to make 21. However, this is not correct as we still cannot make 12. Another 'No'.

Guess	*a*	*b*	*c*	*d*	*e*	OK?
1	1	9	10	11	12	No
2	2	8	9	10	11	No
3	3	7	8	9	12	No

However, we are not yet finished with $a = 3$.

Guess	*a*	*b*	*c*	*d*	*e*	OK?
1	1	9	10	11	12	No
2	2	8	9	10	11	No
3	3	7	8	9	12	No
4	3	7	9	10	11	
5						
6						

This is another 'No' as there is still no way of making 12. Proceeding in this manner:

Guess	*a*	*b*	*c*	*d*	*e*	OK?
1	1	9	10	11	12	No
2	2	8	9	10	11	No
3	3	7	8	9	12	No
4	3	7	9	10	11	No
5	4	6	7	8	13	No
6	4	6	8	9	12	Yes

The sixth guess is correct: 4+6=10, 4+8=12, 4+9=13, 6+8=14, 6+9=15, 4+12=16, 8+9=17, 6+12=18, 8+12=20 & 9+12=21.

There is more work needed (not much) to establish that this solution is unique.

The weights of the five loaves are 154, 156, 158, 159 & 162 g.

4. Look for the Pattern

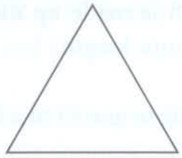
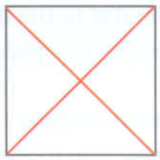
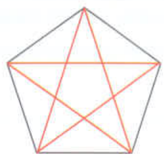
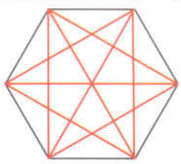
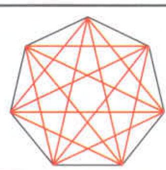
Pattern finding is possibly the most potent of all generally applied mathematical techniques. Some Mathematicians actually describe their job as 'pattern searching'.

Example F.1.4

Find the number of diagonals in a 100 sided convex polygon.

Actually drawing this is, of course possible.

However, a pattern search is easier.

Sides	Shape	Diagonals
3		0
4		2
5		5
6		9
7		14

Do we have enough data to look for a pattern?

One technique is to make a differences table:

Sides	Diagonals	1st diff	2nd diff
3	0		
		2–0=2	
4	2		3–2=1
		5–2=3	
5	5		4–3=1
		9–5=4	
6	9		5–4=1
		14–9=5	
7	14		

The difference in the third column is calculated by subtracting the upper brown cell from the lower orange cell. The same pattern is used throughout the table.

We have a constant second difference. What does this suggest?

If the first differences had been constant we would look for a linear relationship between the number of diagonals (d) and the number of sides (s).

This would mean a relationship of the form $d = a{\times}s+b$ where a and b are constants that we would need to determine.

Because we have a constant second difference, we should look for a quadratic relationship:

$$d = a{\times}s^2 + b{\times}s + c$$

There are three parameters a, b, c to determine. There are a number of ways of going about this. One is simultaneous equations

Triangle:	$0 = 9a + 3b + c$	[1]
Square:	$2 = 16a + 4b + c$	[2]
Pentagon:	$5 = 25a + 5b + c$	[3]
[2] – [1]	$2 = 7a + b$	[4]
[3] – [2]	$3 = 9a + b$	[5]
[5] – [4]	$1 = 2a$	

Thus $a = ½$

Substitute in [4] $2 = 7{\times}½ + b$

Thus $b = -1½$

Finally, using [2] $2 = 8 - 6 + c$

Thus $c = 0$

The rule is: $d = ½s^2 - {}^3/_2\, s$

It is a good idea to check the rule for a couple of cases:

Triangle:	$d = ½{\times}9 - {}^3/_2{\times}3 = 0$ correct
Pentagon:	$d = ½{\times}25 - {}^3/_2{\times}5 = 5$ correct
Hexagon:	$d = ½{\times}36 - {}^3/_2{\times}6 = 9$ correct

An alternative is to look for patterns in the diagrams.

You may have noticed that the number of vertices is the same as the number of sides (s). Each vertex 'sprouts' $s - 3$ diagonals. The product of these is $s(s - 3)$. But this counts each diagonal twice so:

$$d = ½s(s - 3) = ½s^2 - {}^3/_2\, s$$

Finally, to answer our question, if $s = 100$, $d = 50(100 - 3) = 4850$. Note that we have not had to work out all the answers up to the 100 side case but have gone straight to the answer.

These are not the only techniques used to solve problems, however, they are frequently useful.

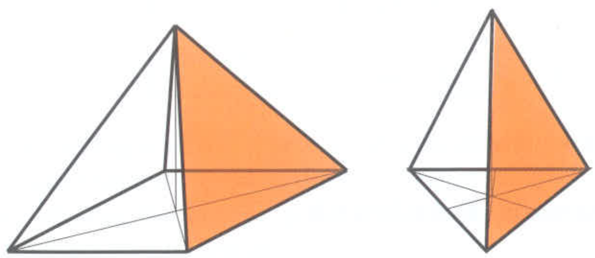
Now try some of these problems. We have made suggestions as to necessary prior knowledge, but these are only suggestions. As our last example showed, there can be more than one solution to a problem and some may be more 'elegant' than others.

Problem 1

A Sticky One

Assumed Knowledge: Core Chapter C1.

A square based pyramid and a regular tetrahedron are such that their triangular faces are congruent equilateral triangles.

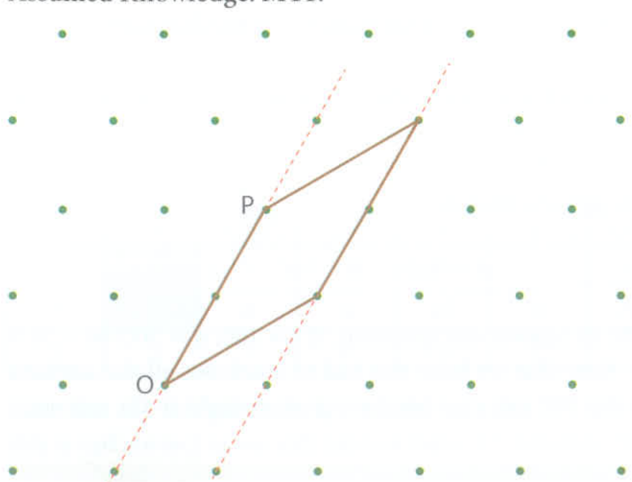


The red shaded faces are identical. If they are glued together, how many faces does the resulting solid have? The answer is not 7. Paul Halmos (one of the great writers on Mathematics - Bing/Google him) relates that this problem was set in a College Entrance Exam with the answer 7. A student with an alternative answer took the matter to court and won. This becomes (probably) the only Mathematics Problem whose solution has legal backing!

Problem 2

Lattice Parallelograms

Assumed Knowledge: MYP.



The diagram shows lattice points from a portion of a triangular lattice which is made up of successive equilateral triangles with sides of unit length.

A line is drawn from the origin O to a lattice point P.

Another line is drawn parallel to the line OP, so that there are no lattice points between the two parallel lines.

A parallelogram is formed so that the four corners of the parallelogram are at lattice points lying on the two parallel lines, and no lattice points, other than those at the corners, are on the boundary of the parallelogram.

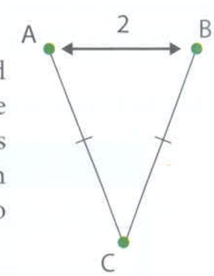
Calculate the area of any parallelogram which can be drawn in this way.

Problem 3

Fibre Optics

Assumed Knowledge: Core Chapters C1-C3.

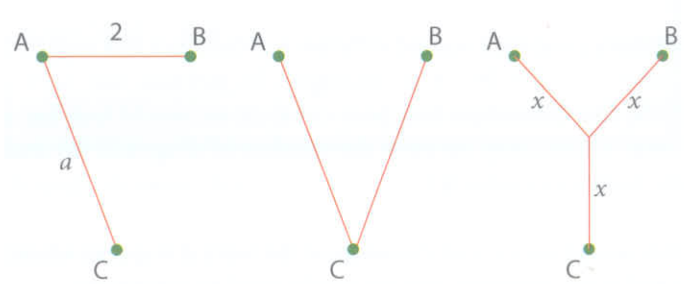
Part A



Engineers from a data link company need to provide a fibre optics link between three towns A, B and C which form an isosceles triangle with AB = BC. The engineers wish to use the shortest possible cable length to join the three towns.

Instead of just solving the problem for the particular towns being considered, the engineers would like to obtain a more general result which could be used if a similar situation arises in the future. To do this they propose the following model: fix the distance between A and C at 2 units and change the isosceles triangle by moving B along a line perpendicular to AC and through the mid-point of AC.

The engineers have put forward three plans which they believe will be useful in solving the problem. These plans, showing the cable links, are shown below.

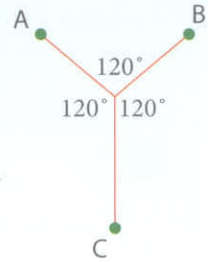


In the 'C-plan', cable is run from B to A to C. In the 'V–plan', cable is run from A to B to C. In the 'Y-plan', cables of equal length are run to A, B and C from a point between the towns.

For different allowable positions of B, when should a given plan be used to obtain the shortest cable length?

Part B

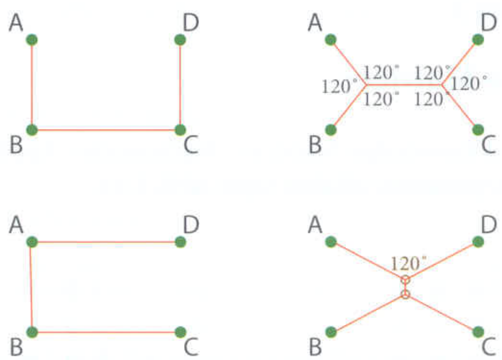
A consulting engineer suggests another option: to link the three towns using a modified Y-plan in which the cables are joined so that the angles between them are each 120°. Is this plan better than any of the three plans in part A?



Explain your answer fully.

Part C

The consulting engineer is working on the problem of connecting four towns A, B, C and D with cable. She is considering four plans which are shown below:



Suppose the distance between A and B is fixed at 2 units and we change the rectangle by adjusting the distance between B and C.

If the engineers want to obtain the shortest cable length, which plan should be used (consider all possible values for part A?

Problem 4

Buried Treasure

Assumed Knowledge: Chapter C2.

You have found a map and instructions to find treasure buried on an island on which are located a water tower, a large oak tree and a peppercorn tree. The instructions include the following directions.

'When you stand at the water tower you should be able to see the large oak tree ahead of you and somewhat to the left, and the peppercorn tree ahead of you and somewhat to the right.

Starting from the water tower, walk directly to the large oak tree, counting the number of steps taken. Turn left through an angle of 90° and walk the same distance away from the oak tree. Mark this spot.

Do this again by starting at the water tower, walking in a straight line to the peppercorn tree and counting the number of steps taken, turning right through an angle of 90°, walking the same distance away from the peppercorn tree and marking the spot.

Half-way along a straight line joining the two marked spots lies the treasure.'

Unfortunately, when you get to the island, the oak tree and the peppercorn tree are plainly visible, but the water tower is nowhere to be seen. It appears that the tower has been demolished and all traces removed. Where is the treasure buried?

You calculate the spot where you think the treasure is buried and have been digging unsuccessfully for some time when you remember that your compass was out of alignment and was not measuring angles accurately.

The angle at the oak tree through which you had turned left was only 85° and the angle at the peppercorn tree through which you had turned right was really 95°. Your compass has now broken down completely and you have no way of accurately determining a right-angle. What can you now do in your attempt to find the treasure?

In your concern to find the treasure you become exhausted and confused and you lose the map in a sudden gust of wind. You think that the instructions were as follows.

'Starting from the water tower, walk directly to the large oak tree, counting the number of steps taken. Turn right through an angle of 90° and walk twice the same distance away from the oak tree. Mark this spot. Do this again by starting at the water tower, walking in a straight line to the peppercorn tree and counting the number of steps taken, turning left through an angle of 90°, walking twice the same distance away from the peppercorn tree and marking the spot.

Halfway along a straight line joining the two marked spots lies the treasure.'

If you were able to find this mid-point, how far from the treasure would you be if the oak tree and the peppercorn tree are 50 metres apart?

Problem 5

Easter Sunday

Assumed Knowledge: MYP.

In theory, Easter Sunday occurs on the first Sunday after the Paschal full moon, which is the first full moon in Jerusalem after 21 March. In practice, the scheduled date of Easter Sunday in each year is determined by a formula specified in Christian literature.

A simpler formula was derived by the mathematician C. F. Gauss (1777–1855), which gives the same date as the scheduled date for every year this century except for 1954 and 1981. The formula derived by Gauss involves the use of the symbol a mod b which means the remainder when a is divided by b. For example, 18 mod 7 is equal to 4 since 18 divided by 7 gives 2 with a remainder of 4.

Gauss's calculation of the date of Easter Sunday is as follows.

For the year which is x years after 1900, for example the year 1931 has $x = 31$, the first full moon occurs c days after 22 March where:

$$c = [19\,(x \bmod 19) + 24] \bmod 30$$

The following Sunday, Easter Sunday, occurs d days after the full moon where:

$$d = (2a + 4b + 6c + 3) \bmod 7$$

with $a = x \bmod 4$, $b = x \bmod 7$, and c defined as before.

You will need to make use of the following information to answer the questions below.

1. In 1900 the full moon occurred 24.07 days after 22 March. This was a Sunday.

2. The time between two full moons is 29.53059 days.

Use the Gauss formula to calculate the date of Easter Sunday for each year in the period 1990–1999.

Explain the reasoning underlying the formula for c relating it to the full moon cycle.

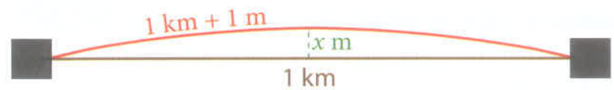
Now let c be any number of days, from 0 to 29 inclusive, after 22 March. Show that the first Sunday after this date occurs in a further d days, as given by the Gauss formula. In the cases for which this date is already a Sunday, show that the Gauss formula gives $d = 0$.

Problem 6

Pipes

Assumed Knowledge: Estimation and Geometry

A straight length of pipeline is fixed at both ends and is 1 km long. On a very hot day, the pipe expands by 0.1% (1 metre). As a result it bows into the arc of a circle. See the diagram below, which is not drawn to scale.



1. Guess the size of x.

2. Use a straight line approximation and Pythagoras to estimate the size of x.

3. Calculate the size of x.

Problem 7

Sparse Matrices

Assumed Knowledge: Matrices - Mathematics: Applications and Interpretation syllabus topic AHL 1.14.

Find:

1. $\sqrt{\begin{matrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{matrix}}$

2. $\sqrt{\begin{matrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{matrix}}$

Problem 8

Watermelons

Assumed Knowledge: MYP.

Watermelons are 99% water by weight.

500 kg of watermelons are stored over the weekend. On Monday, they are 98% water.

How much do they weigh on Monday?

Problem 9

Tilings

Assumed Knowledge: MYP.

Roman mosaic floor, Morocco

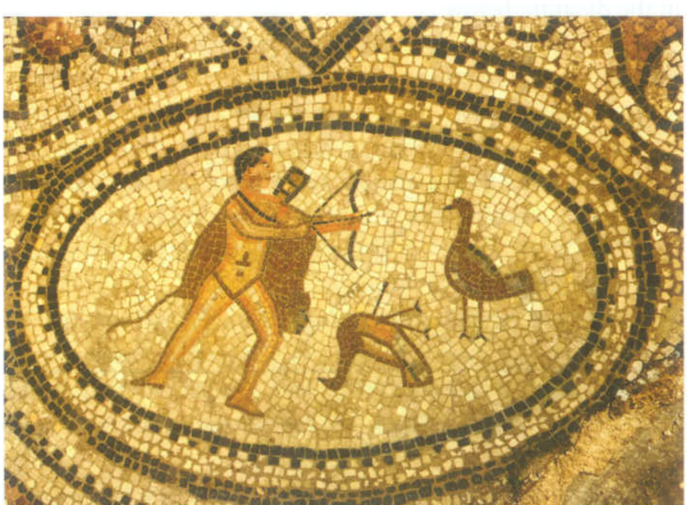
The tiling pattern on this mosaic floor contains triangles, squares, hexagons and rhombuses.

Consider the set of regular shapes:

{equilateral triangles, squares, regular pentagons, regular hexagons, rhombuses}

The set {equilateral triangles, regular hexagons, rhombuses} is said to be a subset of the original set. There are $2^5 = 32$ subsets (including the empty set and the complete set).

Investigate these and classify them according to whether they will or will not form a tessellation if all shapes in the subset are used in the pattern.



Another floor from the same site, Morocco

Problem 10

Tablets

Assumed Knowledge: Probability.

Carol has a bottle containing 18 tablets. She has been told by her doctor that, each day before breakfast, she should take one and a half tablets. On the first day she takes two tablets from the bottle, breaks one in half and puts the unused half back in the bottle. On a subsequent day, perhaps she might take two whole tablets and break one of them, or she might take a whole and a half or three halves so as not to have to break any tablets that day.

Suppose that on day d, before Carol takes her daily dose, there are W whole tablets and H half tablets in the bottle.

The full bottle corresponds to day zero.

a i On day zero there is only one possibility for the mixture of whole and half tablets, namely $W = 18$, $H = 0$.

How many possibilities are there on day d?

ii More generally, if there were N tablets in the bottle initially, find a rule or set of rules which will tell how many possibilities for the mixture of whole and half tablets there are on day d. The case $N = 18$, $d = 10$ is a good one to test your rule. Give the answer in this case, and similarly give several more cases which for one reason or another you think are good ones to test your rule.

b In going from the starting situation of N whole tablets to the final situation where a daily dose cannot be taken, the sequence of possible values for (W, H) shall be called a 'course'.

For example, starting with 14 tablets, one possible course is: $(14, 0) \rightarrow (12, 1) \rightarrow (10, 2) \rightarrow (9, 1) \rightarrow (7, 2) \rightarrow (5, 3) \rightarrow (5, 0) \rightarrow (3, 1) \rightarrow (2, 0) \rightarrow (0, 1)$.

How many possible courses are there starting with:

i 12 tablets?

ii 13 tablets?

iii 14 tablets?

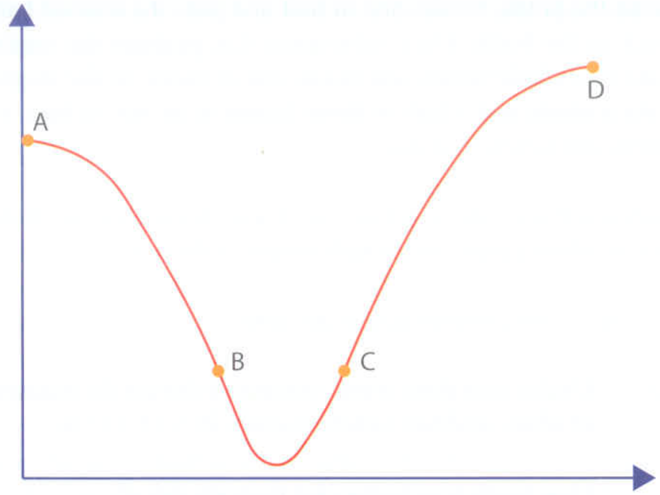
iv 15 tablets?

Problem 11

Wild Ride

Assumed Knowledge: Functions and Calculus.

The track of a section of a switchback railway at an amusement park is in disrepair and needs replacing. Viewed from the side, the track has the shape shown in the following diagram.



The profile of each of the segments AB, BC and CD is a parabola. The gradients of the profiles at the points B and C, where the segments meet, match so that the track is smooth.

The track at point A is 30 metres above ground level and inclined at an angle of 10° below the horizontal, while the track at point B is 10 metres above ground level and 8 metres horizontally across from A. The point C is at the same height as B and 11 metres across horizontally from A. The track at point D is 35 metres high and has zero slope.

The new track costs $100 per metre. Devise a method to estimate the cost of the new track.

Draw diagrams to scale.

Consider where the curved ramps meet the ground or other obstacles. How do the designers of these circuits ensure a 'smooth' ride for the riders?

The real problem is 3 dimensional:

Problem 12

Hockey

Assumed Knowledge: Geometry, Functions.

Hockey is played on a rectangular playing field with a goal area at each of the shorter sides of the rectangle. A player on the long boundary, as shown in the diagram below, wants to shoot a ball through the goal.

What position on the boundary will give the player the biggest shooting angle through the goal?

If the field was circular instead of rectangular, what position on the boundary would give the player the best shooting angle through the goal?

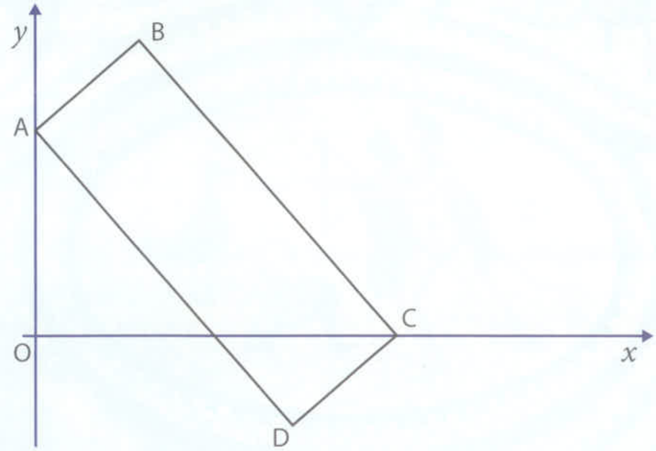
What is the best position on the boundary to shoot from if the field is elliptical?

Problem 13

Sliding Rectangles

Assumed Knowledge: Coordinate Geometry.

A rectangle ABCD is placed on the Cartesian plane as shown in the diagram below.



The rectangle is free to slide so that A is always somewhere on the positive y-axis and C is always somewhere on the positive x-axis.

As the rectangle moves, what path will the point B follow? Explain your results.

What will be the path of the point D?

If the point A can move over both the positive and negative parts of the y-axis, and the point C can move over both the positive and negative parts of the x-axis, what paths will the points B and D follow?

What effect will changing the dimensions of the rectangle have on the paths of points B and D?

Problem 14

Epidemic

Assumed Knowledge: Functions, Sequences & Series.

One Sunday evening, five people with infectious influenza arrive in a large city with a population of about two million. They go to different locations in the city and thus the disease begins to spread throughout the population.

At first, when a person becomes infected he or she shows no sign of the disease and cannot spread it. About one week after first catching the disease, this person becomes infective and can spread the disease to other people.

This infective phase also lasts for about one week. After this time the person is free from the influenza, although he or she may catch it again at some later time.

Epidemiologists are trying to model the spread of the influenza. They make a simplifying assumption that the infection progresses in one week units. That is, they assume that everybody who becomes infected does so on a Sunday evening, becomes infective exactly one week later, and is free of infection exactly one week after that. People free of the disease are called susceptibles.

The epidemiologists also assume that the city population is large enough so that the population size is constant for the duration of the disease. That is, they ignore births, deaths and other migrations into or out of the population.

Finally, they assume that each infective person infects a fixed fraction f of the number of susceptibles, so that:

1. the number of infectives at week $n+1$ is equal to $f \times$ (the number of susceptibles at week n) $\times$ (the number of infectives at week n).

and

2. the number of susceptibles at week $n+1$ is equal to (the number of susceptibles at week n) + (the number of infectives at week n) − (the number of infectives at week $n+1$).

Why must the number of infectives, plus the number of susceptibles, be constant from week to week?

Choose values of f between $\frac{1}{10^6}$ and $\frac{2}{10^6}$ and use this model to show how the number of infectives changes from week to week from the Sunday when the five infective people arrived.

What limiting values does the model predict for the number of infectives?

Will there always be a limiting value for the number of infectives? If so, how are the limiting values related to the population size and to f?

For what values of f will it be possible to have a situation where the number of infectives eventually oscillates between two values?

Problem 15

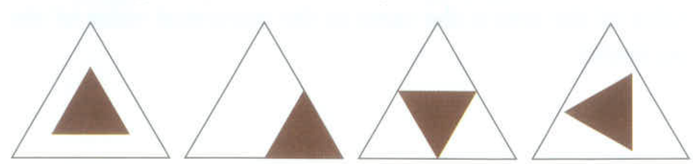
Chocolate Wrapping

Assumed Knowledge: Geometry.

A triangular slice of chocolate has to be wrapped in a triangular piece of paper, which cannot be cut or torn. The chocolate is sliced so thinly that you can ignore its thickness.

Both the chocolate and paper are equilateral triangles. The paper can be folded along an edge of the chocolate. The edges of the wrapping do not have to overlap, they can just meet.

The chocolate can be positioned in various ways on the paper. Some examples appear in the diagram below:



When the chocolate has a side length of 4 centimetres, what is the smallest side length of the paper that will allow the chocolate to be wrapped?

How does the solution differ if the chocolate and paper are both rectangular (not necessarily similar rectangles)?

What optimization issues arise when dealing with such a situation?

Problem 16

Area and Perimeter

Assumed Knowledge: Arithmetic, Geometry, Functions.

The following question was posed to a group of mathematics students.

> 'Are there shapes for which the numerical value of the length of the perimeter is the same as the numerical value of the area?'

One student quickly saw that a square is a shape with this property because a square which has a side length of 4 units has a perimeter of 16 units and an area of 16 square units. The student could also easily show that there could only be one square with this property.

After looking at families of shapes like triangles, circles, rectangles and other polygons the students made the following conjecture.

> 'For every family of shapes there is at least one of these shapes for which the numerical value of the area and the numerical value of the perimeter are the same.'

By a 'family of shapes' the students meant all shapes which are similar to a given shape. For example there is only one family of squares, but there is an infinite number of families of rectangles.

You are required to find the following.

For which shapes does the conjecture hold?

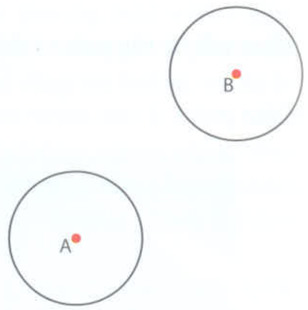
For each class of shapes for which the conjecture holds, give a method for finding an actual shape for which the numerical value of the area is the same as the numerical value of the perimeter.

Problem 17

Circles upon Circles

Assumed Knowledge: Geometry.

Here are two circles of radius 1 with centres A and B, which are 4 units apart.



For a given pair of circles, any circle that intersects each of the original circles twice, with every intersection forming a right angle, we will call perpendicular to that pair.

For the pair of circles pictured:

a

i give a precise description of a circle, perpendicular to the pair, with centre on the line joining A and B.

ii give a precise description of any other circles perpendicular to the pair.

iii comment on any noteworthy features of the family of all circles perpendicular to the pair.

b Consider the circles, perpendicular to the pair, with centres above the line joining A and B.

Each such circle has a point furthest from the line joining A and B, which we will call 'the top'. Suppose that the spacing between two such circles is defined to be the distance between their tops. Find and describe a family of 'equispaced' circles, all of which are perpendicular to the original pair of circles.

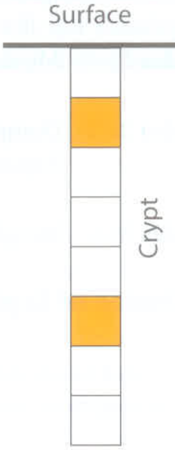
c. Now consider the following. Each of the circles considered in part b has a counterpart obtained by reflection about the line joining A and B. If we now take one of the circles from part b and its reflected counterpart to form a pair of circles, describe the family of circles perpendicular to this pair and comment on any noteworthy features.

Generalize your results in part a to the case where the original pair of circles has unequal radii.

Problem 18

Crypts

Assumed Knowledge: Statistics.



A crypt is found in the human body. It is like a tube of cells which descends from the surface of the colon. The figure opposite shows a highly simplified picture of a crypt which in this case is 10 cells long.

In this picture the shaded boxes show cells which are undergoing cell division. Such cells are found by a staining technique which makes them readily apparent under a microscope.

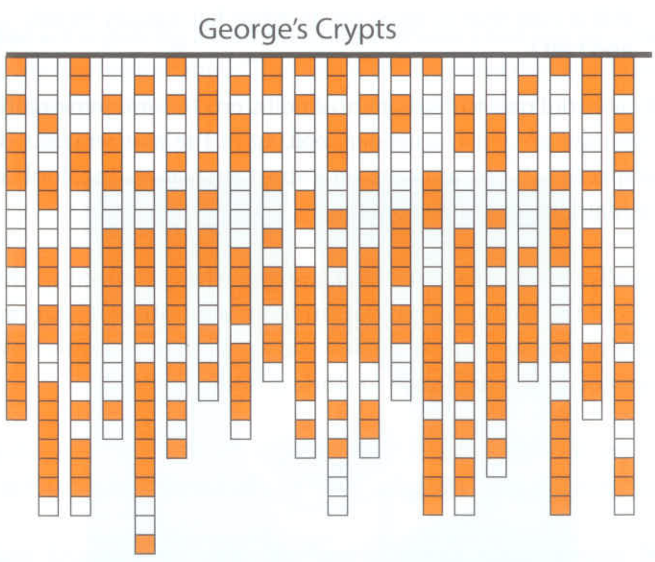
Each clear box shows a cell which is not stained and thus is not undergoing cell division.

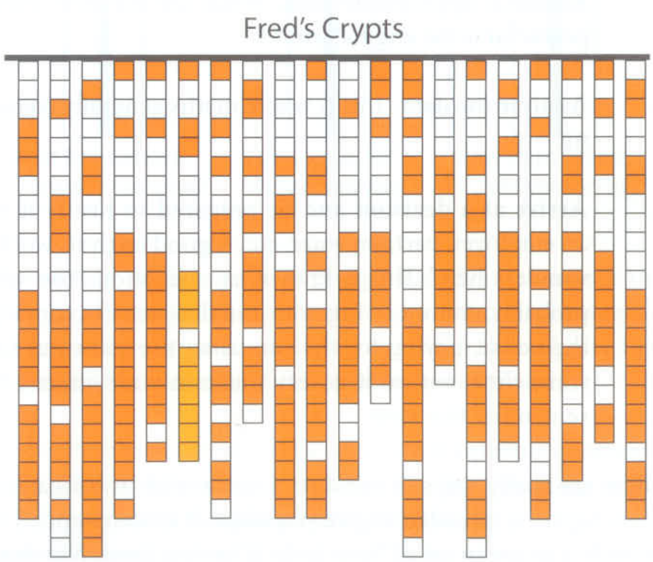
It is thought that a high-fibre diet may alter the way in which cell division occurs within a crypt.

The diagrams below show 20 crypts selected randomly from the colon of George, who has a normal diet.

They also show 20 crypts selected randomly from the colon of Fred, who has a high-fibre diet.

Find a good descriptive variable which indicates that statistically George's crypts are in some way rather different from Fred's crypts. Show also your analysis of other descriptive variables which do not indicate any significant difference between George and Fred.





Investigations

We begin this section by observing that the capabilities we are trying to develop are important life skills. We hope that you will use them as you face challenges in your personal and working lives.

Questions that face many of us at some stage in our lives: which job to apply for, which car to buy, which course to take, whether to buy or rent a home etc. may seem simple. Some are, but most are not. Some of these carry penalties if you make a poor decision.

We are suggesting that, when faced with an important decision, you should follow the process that we hope you will learn in this part of the course. That said, many of history's most valuable investigations (into, for example, the causes of disease, navigation, atomic structure etc.) have been undertaken more out of interest than in the expectation of personal gain.

The main steps in the process are:

1. **Define the problem**. Should I buy or rent a flat? How should I finance it? Where should my company buy its paper, cars, iron ore, legal advice …?
2. **Identify the parameters** that define a 'good decision'. These may be all sorts of things such as price, status value, fast service etc.
3. **Research appropriate data**. What are the fees? How powerful is the engine? etc.
4. **Analyse the data.** This is where maths generally comes in!
5. **Arrive at a decision** and be prepared to justify it to an employer, partner, voter, etc. A good analysis works wonders here! This is known as evaluation. You will probably end up living with the decision (occupying the house, paying for the car, using the software) for some time, and so it is very much in your interest to remember this step!

There are many ways to conduct a successful investigation. This chapter is intended to give you ideas. It is not intended to provide a recipe or set of 'how to do it' instructions. Nor does it provide a complete investigation!

We will look at investigation ideas in the areas of probability, functions and statistics.

To understand these you will need to have covered the relevant chapters:

Idea 1: Monte Carlo methods (Probability).

Idea 2: Damped oscillations (Trigonometric and Exponential Functions, Calculus).

Idea 3: Weights in lifts and transport (Statistics).

Choice of Topic (defining the problem)

Try not to spend too much time on this! We have provided investigations suggestions in two forms:

1. A few suggestions at the end of each chapter, but you should not feel bound by these.
2. A list of investigation 'themes' at the end of this chapter. We have adopted a 'themes' approach so that the investigations would not be prescriptive but, rather, allow you to use them as a springboard to topics of rich mathematical content and of interest to you.

However, some of the best investigations that we have seen have arisen from quite simple questions such as:

> Is it true that petrol prices go up at the weekend?

OR

> Will a coin that is spun on its edge fall equally heads and tails?

Try to avoid questions that are of no interest to you personally. If you are an artist, consider looking at the mathematics of perspective. If you are keen on a sport such as sailing, look at the mathematics of navigation.

Equally, avoid questions that are pointless. Collecting the sizes of the feet of a group of students is pointless. However, imagining that you run a shoe shop and want to know how many of each size of shoe to stock is not!

So, try to pick a simple, relevant question that is in an area that interests you. Then discuss your choice with your teacher.

Of course, your investigation will also be assessed and graded. So, it is important that you are familiar with the 'Internal Assessment Criteria' that will be used to grade your investigation work.

What will define a good decision?

This will depend heavily on your choice of problem and may be as simple as 'the ship will not run aground', but will probably be complicated by a requirement such as 'will be cheaper than my competitor's ship whilst still not running aground'. Do not get over-complicated here as you may make the next step too difficult!

Collecting Data

If you want real data (and we suggest that you do!), the internet is an excellent starting point. Our three case studies give examples in which data came from this source as well as other possibilities. There is nothing wrong with collecting your own data. Remember, however, that the data that you collect must be relevant to the problem you are working on.

Also, there are some dangers to the collection of data. Here are a few:

1. It is a general principle that the process of collecting data affects the results that you get.

A simple question such as

"How many people are in the room at the moment?"

may seem to be easy to answer exactly. However, do you know if there is a person hiding in a cupboard?

2. Measurement also affects the quantity that it is trying to measure.

The act of putting a cold thermometer into a hot cup of coffee lowers its temperature. The only way to prevent this is to know the temperature of the coffee and to heat the thermometer to that before putting it into the cup.

But, if we know the temperature to start with, why are we wasting time measuring it? The sharp eyed amongst you may also see the small bubble in the fluid. This makes this thermometer read about 1°C too much all the time. This is known as a **systematic error**.

The size of these errors depend on a variety of factors.

It will be very small if we use a thermometer to measure the temperature of a swimming pool, particularly if we guess that it is likely to be about 25°C and warm the thermometer up to this temperature before using it.

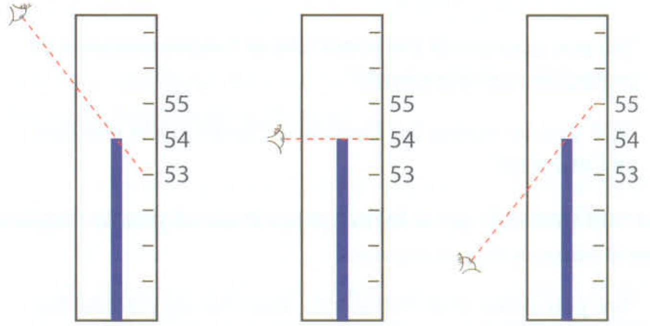
Of course, in the case of the swimming pool the 'cooling error' will be much less than the other main source of error: instrument inaccuracy. Thermometers cannot measure temperature exactly. Some are better than others. The thermometer shown cannot be relied on to measure to better than 1°C. More accurate instruments are more expensive and none are EXACT.

3. Users can also introduce errors by misreading an instrument.

The photograph in the previous column shows a view of the thermometer from above. It suggests that the coffee is at about 78°C.

However, the scale lies behind the liquid column which contains the possibility of a PARALLAX error.

Parallax error is illustrated next (the diagram exaggerates the problem).



View from above, 53°, answer too small. Level view, 54°, correct answer View from below, 55°, answer too big.

The second of our three investigation ideas measures the position of an object using a fairly crude method that is subject to the parallax error. It would be important to recognize that this limits the accuracy of the measurements made.

4. Problems can also arise when we ask people questions. There are two main reasons for this. The first is the possibility that people may lie to us.

Suppose you are approached by a researcher in the street and asked,

"Did you watch the 9 o'clock news on ABC last night?",

you will probably tell the truth.

If, however, this stranger asks:

"How much do you weigh?" or "How many friends do you have?",

you may well be less likely to tell the truth by 'massaging' the answer to be closer to what you would like it to be.

The other problem is that your questions may inadvertently affect the answers people give.

The question:

"Will you be going to cheer on the school basketball team in their final on Saturday?"

may well get a different answer from the question:

"Will you be going to the Rock-Concert or the basketball on Saturday?"

5. Sequences of questions can also affect answers:

Suppose you are leading up to the question:

"Will you be voting for President Clover in the election on Saturday?"

and you lead up to this question with:

"Are you aware that President Clover has increased the health care budget by 8%?"

"Do you approve of President Clover's new measures to protect the environment?"

"Will you be voting for President Clover in the election on Saturday?"

You will probably get a higher proportion of positive replies than if you use this sequence:

"Are you aware that President Clover has decreased the police budget by 8%?"

"Do you approve of President Clover's frequent trips to resorts in the Caribbean?"

"Will you be voting for President Clover in the election on Saturday?"

If you intend to use a questionnaire, you should be aware of these problems and your report should underline the steps that you have taken to minimize their effects.

During the Episode *The Ministerial Broadcast* of the BBC comedy show *Yes Prime Minister*, Sir Humphrey illustrates the issue of leading questions in a brilliant sequence.

Many questionnaires include 'lie detector questions'. One of the best of these is: "Have you ever told a lie?" Would you trust anyone who replied "no" to this question?

Many investigators make use of information derived from the internet. Whilst much of this has been collected by professional researchers who are skilled at minimizing the problems described above, the information should not be treated as completely reliable.

Analysing the data

Following this, you should use mathematics to analyse your data. Analyse means 'pull apart and find how it works'.

This is a mathematics project and you should use mathematics contained in the course. Your teacher should help you determine if your problem is beyond the scope of the course.

Remember to relate the mathematics that you use to your original question. If you are looking at the question, "Do petrol prices go up at the weekend?", you will have already collected information on 'petrol prices' and you will have noted that since you, presumably, have not visited every petrol retailer in the World, there are limits to the validity of your data. What mathematics is it appropriate to use here? There is a temptation to, for example, work out the average price as this is reasonably easy. But does this help you address your question? Far more relevant would be a price/time graph with the weekends clearly identified. Then, you can start to look at whether you have a pattern of steady increase (a trend) or whether the price is stable and 'spikes' at the weekend (or has both a trend and spiking).

It is also important to know what your mathematical results are telling you. If you, for example, calculate a trend line for petrol price data, make sure that you have understood the implications of the calculation before drawing any conclusions from it.

Conclusion

Finally, you should reach a conclusion, related to the question that forms the subject of your investigation. This should be evaluated. In some respects, this is the hardest part of the project.

What do we mean by 'evaluation'? Some decisions have a right answer. If you are playing 'noughts and crosses' there is always a best move that you should make, whatever the game situation. Chess, and most of life, is different. Some moves are better than others but we do not currently (and probably never will) know the best move in every situation. Your project will probably be like this. Your conclusion may be wrong for a variety of reasons including incomplete information. Evaluation means that you have looked at these possibilities!

Try also to avoid looking for the data to confirm your previous opinion. If you ask a motorist "Do petrol prices go up at the weekend?", you may well get the answer "Sure, everyone knows that!". This means that, if your evidence points to the truth being that prices do not go up at the weekend, you may feel under pressure to go along with what everyone thinks. That is not the scientific method! Whilst you will not be held up to public ridicule because you have written something in a school-based investigation that goes against public perceptions, many scientists have been vilified and occasionally imprisoned for insisting on the truth of their (unpopular) findings – the fate of Galileo Galilei (1564–1642) is the best-known case in point.

We will now look at three case studies to expand these ideas. We have taken some ideas further than others and none can be considered as even approaching complete. They are to be taken as skeletons on which a good report might be built!

Idea 1: Monte Carlo methods

(Background theory: Probability).

If you type 'Monte Carlo' into an internet search engine, you will get a lot of information about the small and glamorous Mediterranean state that is almost synonymous with 'high rolling'.

The third site listed by my search engine was a mathematical paper on 'Monte Carlo Methods'.

The name is no accident as it is a mathematical method based on gambling.

1. The question

The first question was raised by looking at this aerial photograph. It emphasizes that most real objects are not bounded by straight lines, circles etc. as they are in mathematics texts. This coastline is very 'crinkly'.

This coastline is smoother.

This suggests the problem: how do we find the areas of irregular shapes?

2. A good answer?

This will be a numerical answer to the problem. Given that there is no 'correct' answer with which we can compare ours, we will be looking at an approximation. Try to set a realistic target such as 'to the nearest 100 km^2'.

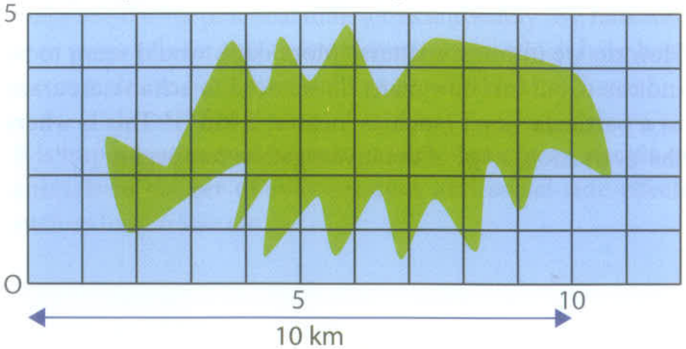
3. Collect information

Even though Monte Carlo methods are not specifically a part of the course, they depend on the concepts of probability that you study. They have many variants. At its simplest, the Monte Carlo method is like throwing darts at a map whilst wearing a blindfold. You will obviously need to research what this might have to do with area!

It would also be a good idea to set yourself the task of finding the area of a real island such as Hokkaido. You will need to obtain an accurate map. As part of your report, you should note the accuracy and scale of the map. A mariner's chart is made to very different standards from those you see on tourist brochures that encourage you to 'Holiday in Happy Hokkaido' (the ones that have cartoon fish leaping from cartoon lakes and cartoon waves rolling onto cartoon beaches).

4. Analysis

The simplest version of the method involves copying a scale version of your island into a rectangle of known dimensions.



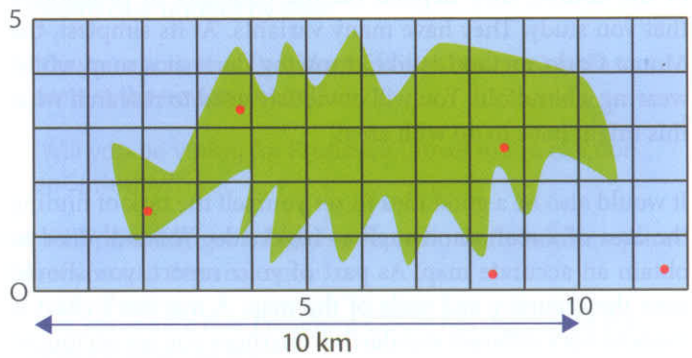
This rectangle measures 12 km by 5 km and so has an area of 60 km^2. But what proportion of this is inside the island?

This is where we don our mathematical blindfold and start throwing darts!

To blindfold ourselves, we can use the random number generator of a computer or calculator. These vary but, commonly, produce a random (what does this mean?) number between 0 and 1. If we multiply a random number in this range by 12, we will get a random number in between 0 and 12. If we take a second random number and multiply it by 5, we will get a random number in the range 0 to 5. If we now view these two numbers as a coordinate pair on the map, we will have, effectively, thrown a blind dart at the map.

Here are ten random numbers and the five 'darts' that they throw at the map:

Random 1	Random 2	12×Rand 1	5×Rand 2
0.723018492	0.529733397	8.68	2.65
0.310504197	0.673043069	3.73	3.37
0.167302964	0.293733395	2.01	1.47
0.96699594	0.078452756	11.60	0.39
0.695841987	0.075482986	8.35	0.38



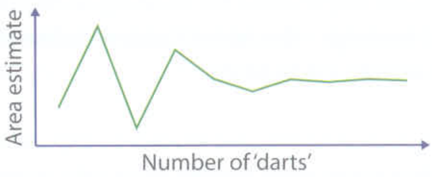
As things stand, 3 out of the 5 'darts' have landed in the island. This suggests that its area is three-fifths of that of the rectangle or $^3/_5 \times 60 = 36$ km^2.

However, if one of the 'darts' that 'hit' was recorded as a 'miss', the area estimate becomes $^2/_5 \times 60 = 24$ km^2. As this is quite different from the previous answer, we cannot be very confident about our current measure of the area.

How do we improve matters? More darts would seem to be indicated, but just how many do we need to achieve accuracy to a particular level (such as 'nearest 1 km^2')? This is where the 'evaluation' stage of the investigation can begin.

5. Evaluation and conclusion

There are several ways to go about this. If you read further into this subject, you should discover that it is possible to estimate the accuracy of an answer obtained by a Monte Carlo method by using a formula. However, if you calculate your area every 10 'darts' and plot a graph of area estimate against number of darts, you should see that the result starts to stabilize after early fluctuations. The extent of the fluctuations in the estimate can give you a measure of accuracy:



Whatever method you use to assess the accuracy of your method, you should comment on this in your report.

Idea 2: Damped Oscillations

(Background theory: Trigonometric and Exponential Functions and Calculus).

1. The question

This idea came from a very rough ride on the back of a truck along a dirt road! Why do we get a much smoother ride in a modern saloon car than on the back of a truck or tractor?

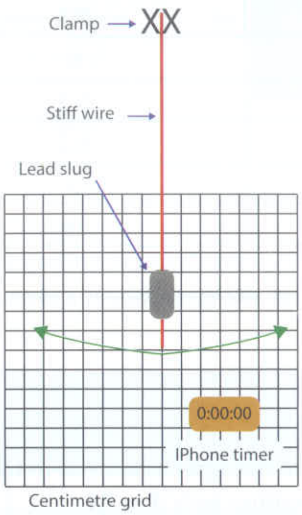
So why do we 'bounce around' more in some vehicles than others?

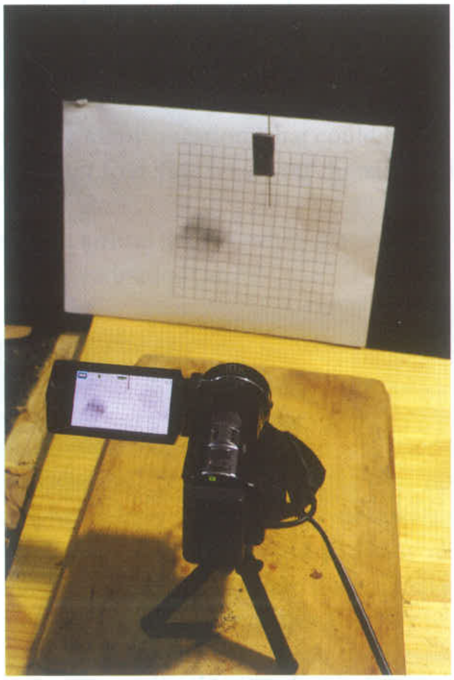
There are several answers to this question, not least the different surfaces over which they usually travel! One of the keys is the quality of the wheel springs and what are known as shock absorbers. The idea also came from the slightly worrying feelings that you get in tall buildings when they 'sway' and in aircraft when the wings 'flap'.

Real vehicle suspensions, buildings and aircraft wings are complex structures that are difficult to observe. What we decided to do was make our measurements on a 'home-built' structure that had similarities to the 'real thing'.

In setting up an experiment to investigate this, we have decided to simplify the problem to look at a simple damped vibrating system using a piece of stiff garden wire, a lead slug (weight) and a digital video camera. The structure is shown in the photograph. Apart from the camera, we used items that were to hand in a domestic garage.

The intention is to move the slug sideways and use pictures taken from the video recording and the grid to measure the position of the slug as it varies over time. To assist in measuring this position, we made and printed a 1 cm square grid using a computer. Note that this setup introduces some parallax error. There is also an error introduced by measuring the horizontal rather than the actual deflection of the tip of the wire.





2. Define the objectives and criteria for success

It is our hope that, working from measurements from the 'structure', we can set up a mathematical model that can be used to predict how it will perform under conditions different from the original experiment (a different initial deflection).

3. Collect the data

The complete data set is a digital video.

Use the pause button to collect data pairs: (time, deflection).

Here are some of ours - obtained by pausing the video. We chose the stopwatch reading 13.15 sec as the zero of time. The second row of the table was taken at the clock reading of 13.15 + 0.06 = 13.21 sec.

Time (sec)	Deflection (cm)
0	1
0.06	5.8
0.19	8.5
0.39	-0.7
0.68	-1.2
0.8	3.1
0.87	6.5
0.99	7
1.05	4.4

It is an important part of the data collection section of an investigation that you should report on the limitations of your method. In this case, we have used a 1 cm grid and have estimated to one decimal place. It is unlikely that we have done better than an accuracy level of ±0.2 cm, particularly in the frames where there is 'motion blur'. Also, there is the effect of parallax, although we have tried to minimize this. Note also that we are ignoring the fact that the tip moves vertically as well as horizontally (it moves in the approximate arc of a circle).

If you choose to do a similar modelling exercise, you will find it easier if you choose a slower moving event. Use a heavier weight or a less stiff wire. Or a different setup altogether.

We have included an example of a practical investigation to illustrate the fact that these may not be as difficult as they at first appear. However, we strongly advise that you discuss a proposal of this sort with your teacher before undertaking any such experiment. There are many reasons for this, the main ones being practical difficulties and safety. We have also seen proposals for this type of investigation which have been too complex to be carried out in a reasonable time frame.

It is important to avoid investigations that expose people or animals to danger or may produce anti-social side-effects such as loud noise etc.

4. Analysis

As a first step, we have entered the data given in the table above onto a spreadsheet. This enables us to generate a graph of the data set for the early part of the motion.

We will look at the two parts to modelling the motion separately.

1. Periodicity (trigonometric)

2. Decay (negative exponential)

Periodicity

If D = deflection from the zero point (the vertical gridline 2 cm to the left of the phone) and t is the time in seconds after our arbitrary start time of 13.15 on the clock, then we are looking for a model of the form:

$$D = A \times \sin(n \times t + c)$$

See the Chapter on the Trigonometric Functions to recall that the three parameters A, n, c control different features of the function:

A determines the amplitude. The video shows that 8 is a reasonable estimate for this.

n determines the period. Looking at the data table it looks like two successive maximum deflections to the right occur at about $t = 0.19$ & 0.87. This suggests a period of about 0.68.

Using $\tau = \dfrac{2\pi}{n} \Rightarrow n = \dfrac{2\pi}{\tau} \approx \dfrac{2\pi}{0.68} \approx 9.2$

c represents the left/right displacement or phase of the model. For no particular reason, we start with a value of 1.

Video discussion of the use of a spreadsheet to model the periodic component of these data.

Link to the spreadsheet.

The periodic component is modelled by:

$$D = 8 \times \sin(8.6t)$$

To model the decay aspect, we use a similar method. Here are some data points selected to be the maximum deflection to the right.

Time	Maximum Deflection
0	7.3
20	6
30	5.2
50	4.2
80	3
100	2.6

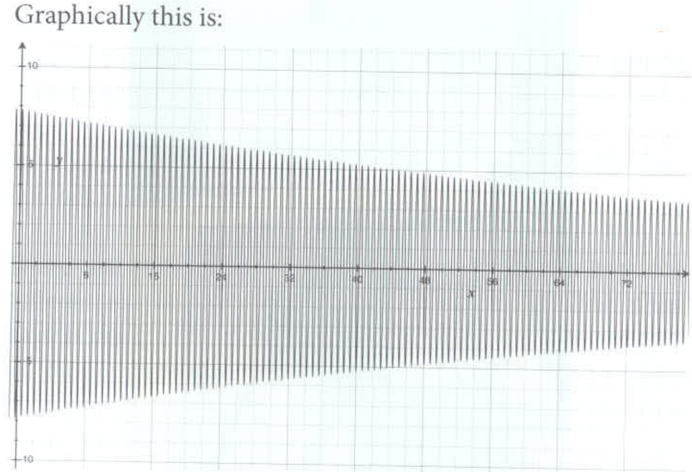
The second page of the spreadsheet shows how we arrive at the model:

$$D_{max} = 7.8e^{-t/100}$$

The two models do not agree exactly as to the correct value of A. We will compromise with $A = 7.9$

The final model is: $D = 7.9 \times e^{-t/100} \times \sin(8.6t)$

Graphically this is:



5. Evaluation and conclusion

We have succeeded in taking measurements from a real structure that was set in motion. We have used mathematics to find a model that can be used to predict the behaviour of the structure in circumstances different from those that we actually tried. The model performed much better in the short term than later in the motion. A good investigation will explain that this is much more likely to be due to errors in the trigonometric part of the model than in the exponential part.

Idea 3: Statistics (weights in lifts and transport)

The question here was raised by this safety plaque in an elevator (lift).

The manufacturers seem to be assuming an average weight of

$\frac{2030}{30} \approx 68$ kg per person.

Further research on the internet indicated that airlines assume standard weights for passengers. These are under review (people are getting heavier) and vary from country to country (as do people) but in the 1990s approximate values were:

Seating capacity	Adult Male (kg)	Adult Female (kg)	Teenage Male (kg)	Teenage Female (kg)	Child (kg)	Infant (kg)
7–9	86	71	65	56	44	17
40–59	83	68	63	57	42	16
100–149	82	66.9	61.1	55.2	41	16
300–499	81.4	66.3	60.6	54.8	41	16

There seem to be many questions that could be followed here and it is important to define what you choose to investigate.

If you accept the airline figures, what chance is there that the elevator will be overloaded?

If you accept the implied elevator figures, what chance is there that an aeroplane will be overloaded?

If you accept none of these figures, you will need to get data on the actual weights of people. This can be a real problem as 'body weight' is very definitely in the area of 'personal sensitivity'. Ask someone their weight and they may well tell you their preferred weight rather than their actual weight. Ask them to get on a set of bathroom scales and they will probably refuse. This seems to be a line of enquiry best avoided!

2. Define the problem

After a variety of internet searches on 'body weight', we failed to come up with any definitive data on this distribution. The outcome of a search for useful data on human weight distributions was to produce a flood of offers for diet programs! You may, of course, be more successful in forming your search questions than were the authors!

This became an investigation in which 'definition' emerged as a major feature.

We have airline data on the 'standard weights' of various classes of humans. We also have a safety plaque that indicates the safe loading of elevators.

One option is to accept the most conservative weights assumed by airlines (the row for 7–9 seat aeroplanes). Why is this the best assumption and why do the airlines assume that everyone gets lighter when they get onto large aeroplanes? There is a good investigation in this question alone!

However, we will accept the most pessimistic data on weights assumed by the airlines and ask ourselves:

What are the chances (mathematicians call this probability) that the elevator whose plaque was shown earlier will accept a load of 30 persons or fewer and yet still be overloaded?

This is now a clearly defined problem. We were not able to arrive at this definition until we looked at the available data. This may well be the case with the investigation that you choose and underlines the fact that the stages identified earlier in the chapter may well over-lap.

3. Collect data

Our problem definition has now assumed that we accept these data on human body weight:

Adult Male (kg)	Adult Female (kg)	Teenage Male (kg)	Teenage Female (kg)	Child (kg)	Infant (kg)
86	71	65	56	44	17

The elevator is restricted to 30 persons and a weight of 2030 kg.

If 30 adult males get into the elevator they will weigh: $30 \times 86 = 2580$ kg and the elevator will be overloaded. It appears that the manufacturers are assuming that the normal load will seldom consist of 30 adult males. So what is a normal load?

There are two questions to investigate here:

1. What is the distribution of the numbers of people who enter elevators?
2. What is the distribution of the types of people who enter elevators?

Both these questions can be answered by collecting real data. Observe an elevator and make a frequency distribution of:

The number of people riding in the elevator (not the number who get into it)

AND

The distribution of the types of people (adult females, infants etc. as defined by the weight data) who ride in elevators.

4. Analyse the data

Having collected the data described above, you now need to display it mathematically. This can mean graphs, tables etc.

However, remember that the issue is to look at the probability that the elevator will be overloaded and your analysis should reflect this.

For example, if you observed a family consisting of a mother, father, two female teenagers and an infant riding in the elevator, you can assume that they weigh: $86 + 71 + 2 \times 56 + 17 = 286$ kg which is well inside the weight limit. If, however, you observe a full load of adult males, you will record an overload.

Taking into account both these variables (number of occupants and composition of the load) what is the probability that the elevator will be overloaded?

By the way, you do not need to flee in panic if you find yourself in an elevator full of heavy people. Many have weight sensors that prevent them moving if overloaded. In common with all structures, the designers will have built in a safety factor that they are not keen to advertise. Imagine if the plaque had said '30 persons, but will take 35'. Everyone would work on 35 and it is 'goodbye to the safety factor'. Also, though this will not be a part of a mathematics project, most elevators include a truly ingenious 'fail safe' device that ensures that, even if the cable snaps, the 'cage' will not plummet to the bottom of the shaft.

5. Conclusions

This topic has made many assumptions with its data. Your conclusions should reflect on the difficulties that were experienced in acquiring reliable data. These, of course, follow through into the reliance that can be placed on the conclusions.

Index

MATHEMATICS
Common Core

IBID Press has published for the IB courses since 1998. Our books are specifically written for the courses they serve and are not adaptations from other sources.

For the new mathematics courses we have produced 5 texts which cover the Common Core, Analysis and Approaches SL & Applications and Interpretations SL and Analysis and Approaches HL & Applications and Interpretations HL. This format will help schools who need to teach as many joint classes as possible while providing the specific coverage necessary for the individual courses and the distinct levels.

Each book is in full colour, contains clear explanations, associated diagrams, cultural and international references, extensive exercises and Toolboxes.

ISBN: 978-1-921784-86-6

ISBN: 978-1-921784-86-6
9 781921 784866

For use with the I.B. Diploma Programme