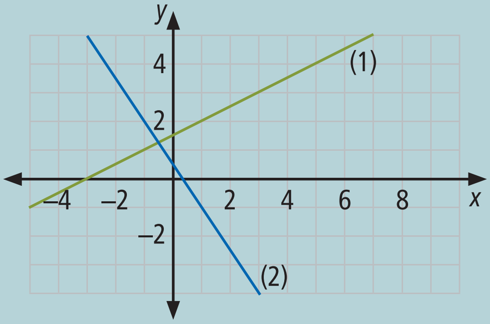
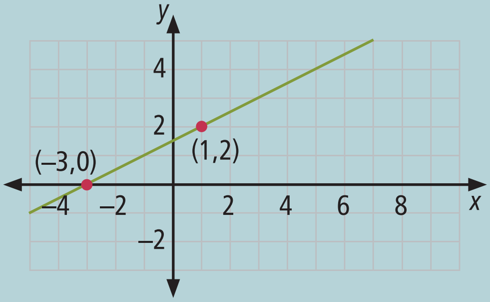
If the gradient of a line is required and no points are provided, identify two points on the line whose coordinates are integers (for ease of working).

**Example**

**Q.** Find the gradient of each line on the graph below.



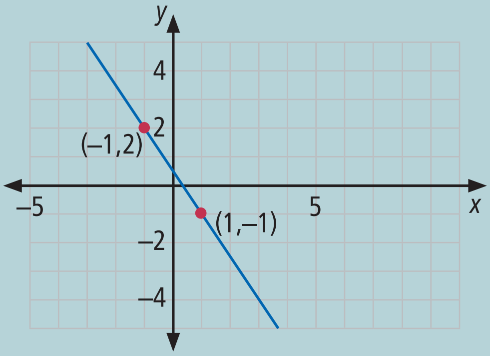

**A.** 1. The points (–3,0) and (1,2) lie on line (1).




$$m = \frac{2-0}{1--3} \quad \text{substituting in } \frac{y_2-y_1}{x_2-x_1}$$
$$= \frac{2}{4}$$
$$= \frac{1}{2}$$

*Note:* Any pair of points which lie on the line may be used. The gradient will always be $\frac{1}{2}$.

2. The points (–1,2) and (1,–1) lie on line (2).

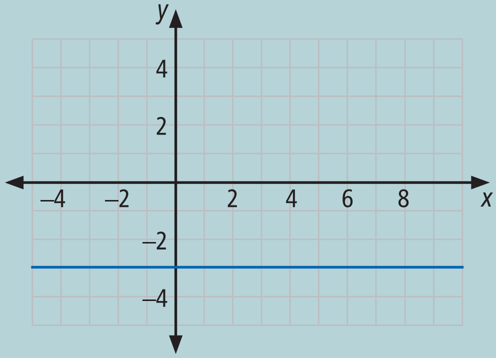



$$m = \frac{-1-2}{1--1} \quad \text{substituting in } \frac{y_2-y_1}{x_2-x_1}$$
$$= -\frac{3}{2}$$

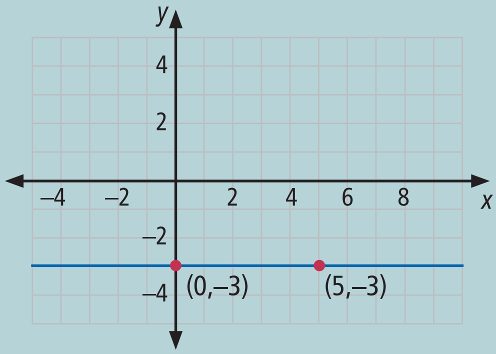
A line that is completely flat has a gradient of zero.

**Example**

**Q.** Find the gradient of the line shown below.




**A.** The points (0,–3) and (5,–3) lie on the line (any points on the line can be used).




$$m = \frac{-3--3}{5-0} \quad \text{substituting in } \frac{y_2-y_1}{x_2-x_1}$$
$$= \frac{0}{5} \quad \text{simplifying}$$
$$= 0$$

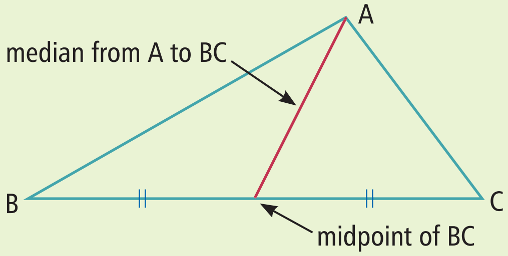
*Note:* The gradient of a vertical line is undefined.

 ISBN 978-0-947504-98-4 

## Medians, perpendicular bisectors and altitudes

### *Median of a triangle*

A **median** of a triangle is a line segment joining a vertex of a triangle to the midpoint of the opposite side.




The median divides the area of the triangle into two parts of equal area.

**Example**

A triangle ABC has vertices A (–2,–7), B (5,1) and C (6,–5).

Find the equation of the median from the point B to the side AC.

**Solution**

The median joins B (5,1) to the midpoint of AC

which is $\left(\frac{-2+6}{2},\frac{-7+-5}{2}\right)$ $\quad\left(\frac{x_1+x_2}{2},\frac{y_1+y_2}{2}\right)$

$= (2,-6)$

Gradient of median is $\frac{-6-1}{2-5}$ $\quad m = \frac{y_2-y_1}{x_2-x_1}$

$= \frac{7}{3}$

Equation of median is

$y - 1 = \frac{7}{3}(x-5)$ $\quad y - y_1 = m(x - x_1)$

$3y - 3 = 7(x-5)$ multiplying by 3

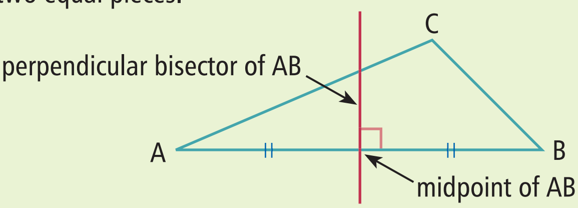
$3y - 3 = 7x - 35$ expanding brackets

$7x - 3y - 32 = 0$ rearranging

Alternatively, the equation is $y = \frac{7}{3}x - 10\frac{2}{3}$.

### *Perpendicular bisector of a line segment*

The **perpendicular bisector** or **mediator** of a line segment AB is a line which cuts AB at right angles into two equal pieces.




The perpendicular bisector of AB is the **locus** of points equidistant from A and B.

**Example**

On a map, town P is located at (12,2) and town Q is located at (18,6). A road is planned so that it is equidistant from P and Q. Find the equation of the road.

**Solution**

Require the equation of the perpendicular bisector of the line segment PQ, which is a line at right angles to the line segment PQ and passing through its midpoint.

The midpoint of PQ is (15,4) $\quad\left(\frac{x_1+x_2}{2},\frac{y_1+y_2}{2}\right)$

The gradient of PQ is $\frac{6-2}{18-12}$ $\quad m = \frac{y_2-y_1}{x_2-x_1}$

$= \frac{2}{3}$ simplifying

So the gradient of the perpendicular is $-\frac{3}{2}$

$m_2 = -\frac{1}{m_1}$

The equation of the perpendicular bisector is

$y - 4 = -\frac{3}{2}(x - 15)$ substituting $m = -\frac{3}{2}$ and $(x_1,y_1) = (15,4)$ in $y - y_1 = m(x - x_1)$

$2y - 8 = -3x + 45$ multiplying by 2 and expanding bracket

$3x + 2y - 53 = 0$ rearranging

Alternatively, the equation is $y = -\frac{3}{2}x + 26\frac{1}{2}$.

 ISBN 978-0-947504-98-4 

## The general term of a geometric sequence

A rule can be found for the general term of a **geometric sequence** by working out how each term is calculated.

**Example**

In the geometric sequence 2, 6, 18, 54, 162, ... the first term is $a = 2$, and the common ratio is $r = 3$. $6 \div 2 = 3$ $18 \div 6 = 3$ etc.

The first six terms are formed as follows:

| **Term** | $t_1 = 2$ | $t_2 = 6$ | $t_3 = 18$ |
|---|---|---|---|
| **Rule** | $t_1 = 2 \times 3^0$ | $t_2 = 2 \times 3^1$ | $t_3 = 2 \times 3^2$ |

| **Term** | $t_4 = 54$ | $t_5 = 162$ | $t_6 = 486$ |
|---|---|---|---|
| **Rule** | $t_4 = 2 \times 3^3$ | $t_5 = 2 \times 3^4$ | $t_6 = 2 \times 3^5$ |

In the rule the power of 3 is always 1 less than the term number, so the rule for the *n*th term is

$$t_n = 2 \times 3^{n-1}$$

Using this rule the 8th term, say, can be worked out to be:

$$t_8 = 2 \times 3^7 \text{ which is } 4\,374$$

In general, a geometric sequence is of the form $a, ar, ar^2, ar^3, ar^4, \ldots$

where $a$ is the first term and $r$ is the common ratio.

The **general term** for a geometric sequence is:

$$t_n = ar^{(n-1)}$$

- $n$ = the position of the term within the sequence
- $a$ = the first term of the sequence
- $r$ = the common ratio
- $t_n$ = the *n*th term

**Example**

**Q.** Calculate the 10th term of the geometric sequence: 1, 3, 9, 27, ...

**A.** $a = 1$ $r = 3$ $n = 10$

the first term is 1, the common ratio is 3 and the 10th term is required

Substituting these values into the formula $t_n = ar^{(n-1)}$ gives:

$$t_{10} = 1 \times 3^{(10-1)}$$
$$= 3^9$$
$$t_{10} = 19\,683$$

Many practical problems involve geometric sequences.

**Example**

**Q.** A student records the number of minutes she spends each day on a homework project. For the first three days the number of minutes spent are 200, 160 and 128. If this pattern follows a geometric sequence, how many minutes will she spend on the project on the 7th day?

**A.** $n = 7$ $a = 200$ $r = \frac{160}{200} = 0.8$

Substituting into $t_n = ar^{(n-1)}$ gives:

$$t_7 = 200 \times 0.8^{(7-1)}$$
$$= 200 \times 0.8^6$$
$$t_7 = 52.43$$

She will spend about 53 minutes on her project on the 7th day.

## *The general term of a geometric sequence*

Answers p. 307

1. Calculate the required term for the following geometric sequences:

   a. 1, 4, 16, 64, ..., $t_9$

   b. 1, 2, 4, 8, 16, ..., $t_{20}$

 ISBN 978-0-947504-98-4 

## Conditional probability

Conditional probabilities are calculated using a restricted set of outcomes.

**Example**

Customers in a shop buy items for themselves or as a gift. A record is kept for a week.

| | Self | Gift | Total |
|---|---|---|---|
| Male | 23 | 72 | 95 |
| Female | 78 | 143 | 221 |
| Total | 101 | 215 | 316 |

1. From the table, there are 95 male customers.
   Of these 95 male customers 72 purchased a gift.
   P(male customer buys gift) $= \frac{72}{95}$
2. There are 215 customers who bought gifts.
   Of these 215 customers, 143 were female.
   P(gift-buyer is female) $= \frac{143}{215}$
3. A female is chosen at random.
   P(female is gift buyer) $= \frac{143}{221}$

*Note:* Compare the wording of the conditional probabilities in part 2. and 3.

## *Basic probability*

Answers p. 327

1. Two fair dice are rolled together and the numbers added.

a. Complete the table showing the set of all 36 possible outcomes.

| | 1 | 2 | 3 | 4 | 5 | 6 |
|---|---|---|---|---|---|---|
| **1** | 2 | 3 | 4 | 5 | 6 | 7 |
| **2** | 3 | 4 | 5 | | | |
| **3** | 4 | | | | | |
| **4** | 5 | | | | | |
| **5** | 6 | | | | | |
| **6** | | | | | | |

b. When two dice are rolled and the numbers added, use the table to work out the probability that the sum is:

i. equal to 6

ii. equal to 13

iii. an odd number

iv. less than 5

v. more than 7

vi. less than 13

c. Ben said that getting a sum of 4 has the same chance as getting the sum of another number. What is this other number?

2. Amy carried out a week-long survey of customers who bought items in her shop. She had a total of 304 customers during the week, of which 115 were male.

a. What proportion of customers were male?

b. Estimate the probability that the next customer in her shop is female.

ESA PUBLICATIONS ISBN 978-0-947504-98-4 

## Expected numbers

An **expected number** is the number in a group which would be expected to have a particular feature or characteristic.

Suppose a sample of size $n$ is taken from a population in which the probability of having a certain characteristic is $p$. The expected number in the sample with this characteristic is:

Expected number $= np$ or $n \times p$

For example, suppose a darts player hits the bulls-eye on 5% of his throws.

If he throws 60 times, he would expect:

5% of 60 = 3 bulls-eyes.

**Example**

In a certain population, the probability of a person working in a manufacturing job is 0.23. How many people in a random sample of 270 people from this population would be expected to work in manufacturing?

**Solution**

Expected number $= 270 \times 0.23$ — multiplying number in sample by the probability

$= 62.1$

Therefore about 62 or 63 people in such a sample would be expected to work in manufacturing.

*Note:* The 'exact' answer of 62.1 results from long-run averaging, and is interpreted practically by rounding to the adjacent whole number(s).

In order to work out predictions of expected numbers, the proportions present in a representative sample from a population are used.

**Example**

**Q.** The table shows the grades received in a general knowledge test given to a random sample of 80 students aged 10–18.

| | A | B | C | Totals |
|---|---|---|---|---|
| 10–12 years | 6 | 11 | 9 | 26 |
| 13–15 years | 9 | 14 | 5 | 28 |
| 16–18 years | 10 | 12 | 4 | 26 |
| Totals | 25 | 37 | 18 | 80 |

If this group of students is representative of all students in a city of 3 682 students aged 10–18, find the expected number of students in the city who would score a C on this test.

**A.** This group of students is representative of the population of all students in the city, so sample proportions can be used as estimates of population probabilities.

The proportion of students in the sample who scored a C is $\frac{18}{80}$ or 0.225.

So the expected number of students in the city who would score a C on this test is

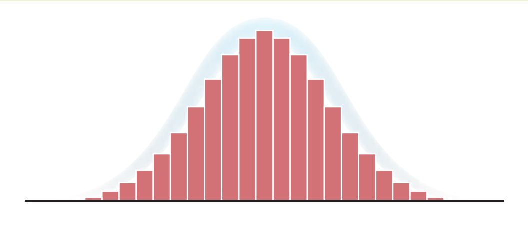
$3\,682 \times 0.225 = 828.45$ — expected number $= n \times p$

Thus about 830 students in the city would be expected to score a C.

 ISBN 978-0-947504-98-4 

## The normal distribution

Many continuous data sets have a distribution which is bell-shaped when graphed.

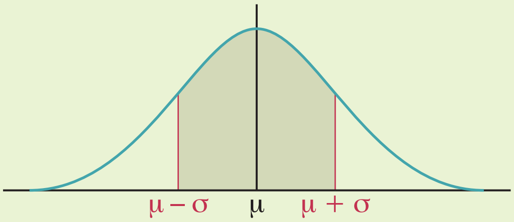


**Example**

Some examples of data sets with bell-shaped distributions are:

1. The heights of students in secondary school.
2. Weights of adult male German Shepherd dogs.
3. Fuel consumption per 100 L of a particular model of car.
4. The length of each advertisement break on television.

Variables with symmetrical bell-shaped graphs are said to have a **normal distribution**. A typical normal curve graph is shown below.




Each normal distribution is described by two measures: the **mean**, $\mu$ (the middle value) and the **standard deviation**, $\sigma$ (a measure of spread). There is no upper or lower bound on the values the variable can take.

Two other important features of the normal curve are:

- the graph of the normal curve is symmetrical about $\mu$
- the total area under the normal curve is 1.

### Areas under the normal curve

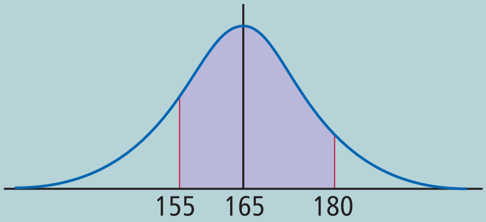
Each **event** corresponds to an interval of values on the horizontal axis. The **probability of an event** is given by the area under the normal curve over this interval.

*Note:* The probability is the same whether or not the end points of the interval are included.

**Example**

The heights of students in Year 12 are normally distributed with mean 165 cm and standard deviation 10 cm.

The probability that the height of a Year 12 student chosen at random is between 155 cm and 180 cm is shown shaded in the diagram below.

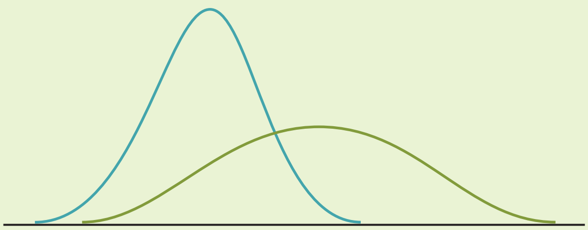



The shaded area is 0.775. This represents the probability that a Year 12 student chosen at random measures between 155 cm and 180 cm.

### The standard normal distribution

The shape of a normal curve depends on the values of the mean $\mu$ and standard deviation $\sigma$.

**Normal curves with different $\mu$ and $\sigma$**



To avoid having to calculate probabilities for each different combination of $\mu$ and $\sigma$, each normal curve is **standardised** so that the mean becomes 0 ($\mu = 0$) and the standard deviation becomes 1 ($\sigma = 1$).

The normal distribution which has $\mu = 0$ and $\sigma = 1$ is is called the **standard normal distribution**.

Special tables exist which give the probabilities between two values for a Standard normal curve.

### Standard normal tables

Before solving problems involving the normal distribution, it is important to know how to find probabilities using the Standard normal table.

The table gives the area under the Standard normal curve between 0 and a value *z*. This area is equal to the probability of a value lying between 0 and *z*.

ESA PUBLICATIONS ISBN 978-0-947504-98-4 

# MATHEMATICS AND STATISTICS 2.14

**Internally assessed**
**2 credits**

## Apply systems of equations in solving problems

## Introduction

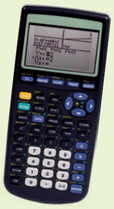
The solving of a pair of linear **simultaneous equations** was covered in NCEA Level 1 Mathematics and Statistics (these methods are revised below).

The focus for this standard is on systems of equations in which at least one equation is **non-linear**.

The ability to sketch the graphs of **linear**, **polynomial**, **circular** and **hyperbolic** functions is expected knowledge in this standard.

## Sketching graphs of lines and curves

One way of sketching graphs is to calculate a table of points $(x,y)$ (select values for $x$ then substitute them into the equation of the curve to find the associated $y$-values). Plot and join points to obtain the graph.



*Note:* Graphics calculators are very useful tools for drawing graphs. Consult your manual to find out how to do this on your calculator.

**Example**

The linear function $y = mx + c$ has gradient $m$ and $y$-intercept $c$.

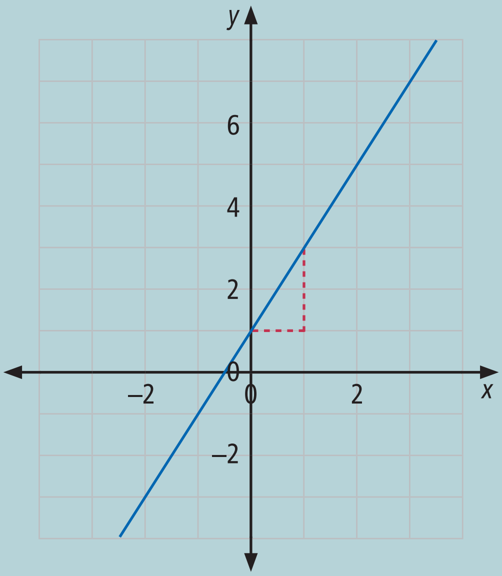
For example, the graph of $y = 2x + 1$ is a line, with gradient $= 2$ and $y$-intercept $= 1$.

Plot the $y$-intercept $(0,1)$, and from this point move 1 unit right and 2 units up (gradient is 2) to the point $(1,3)$.

Joining these points gives the required line, as shown. Alternatively, calculate points using the rule.

For example, when $x = 0$, $y = 2 \times 0 + 1 = 1$ giving the point $(0,1)$.

Plot and join these points with a straight line.




The graph of a **quadratic function** $y = ax^2 + bx + c$ is a symmetrical curve called a **parabola.**

**Example**

The graph of $y = x^2 - 2x + 2$ is a parabola.

This graph can be sketched by drawing up a table of values, as shown below.

For example,
if $x = -1$, then $y = (-1)^2 - 2(-1) + 2 = 5$

if $x = 0$, then $y = 0^2 - 2(0) + 2 = 2$

| $x$ | –1 | 0 | 1 | 2 | 3 |
|---|---|---|---|---|---|
| $y$ | 5 | 2 | 1 | 2 | 5 |

Plot the points $(-1,5)$, $(0,2)$ etc., and join them with a smooth curve, as shown. Make sure the turning point is rounded, not sharp.

ESA PUBLICATIONS ISBN 978-0-947504-98-4 

In the **substitution method** one variable is replaced with an expression in terms of the other variable.

**Example**

**Q.** John's age now is one more than double his brother's age now. In four years' time John will be three times older than his brother is now. How old is John now?

**A.** Define the variables:

let $J$ = John's age now,

$B$ = his brother's age now

Form equations from the given information:

$J = 2B + 1$ ...(1) — $J$ is 1 more than double $B$

$J + 4 = 3B$ ...(2) — when John's age is $J + 4$, it equals $3B$

Now solve simultaneously by replacing $J$ with $2B + 1$ in the second equation:

$2B + 1 + 4 = 3B$ — substitute (1) into (2)

$2B + 5 = 3B$

$B = 5$ — subtract $2B$ both sides

Substitute $B = 5$ in (1):

$J = 2 \times 5 + 1 = 11$

John's age now is 11 years.

## *Graphical interpretation of a system of linear equations*

When solving a system of two linear equations with two variables, the pair of lines representing the equations can relate to each other in three possible ways.

- A **unique solution** is found (as in the examples above) when the two lines intersect at a single point.
- **Infinitely many solutions** are found when the equations are in fact different representations of the same line. An attempt to solve equations like these algebraically will lead to a trivial but true statement (like 0 = 0).
- **No solution** can be found when the two lines are parallel. An attempt to solve equations like these algebraically will reduce to a false statement (such as 0 = 4).

If at least one solution can be found for a system of equations, then the system is referred to as **consistent**; if no solution can be found the system is **inconsistent**.

**Example**

Attempt to solve the following systems of equations.

**Q.** $3x + y = 1$ and $2y = -6x + 2$

**A.** $3x + y = 1$ ...(1)

$2y = -6x + 2$ ...(2)

$y = -3x + 1$ — dividing (2) by 2

Substitute $y = -3x + 1$ into (1)

$3x + (-3x + 1) = 1$

$1 = 1$

This is always true no matter what values of $x$ and $y$ are chosen so there are infinitely many solutions to this system.

Graphically both equations represent the same line.

**Q.** $6x - 2y = 3$ and $y = 3x - 4$

**A.** $6x - 2y = 3$ ...(1)

$y = 3x - 4$ ...(2)

Substitute (2) into (1)

$6x - 2(3x - 4) = 3$

$6x - 6x + 8 = 3$ — expanding

$8 = 3$

This is impossible hence this system is inconsistent and has no solution (no values of $x$ and $y$ will result in a correct solution). Graphically the two equations represent parallel lines.

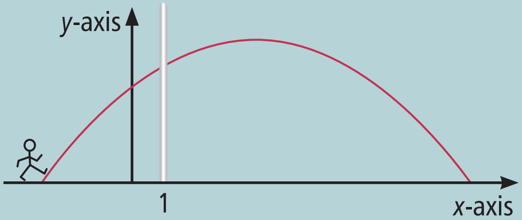
 ISBN 978-0-947504-98-4 

## Applications of systems of linear/non-linear equations

Many practical situations can be modelled using linear and non-linear functions.

**Example**

Daniel is attempting to kick a goal in a game of rugby. The ball travels from left to right along a path that can be described by the equation $y = 0.1(12 - x)(x + 3)$, where $x$ is the horizontal position of the ball in metres, relative to the axes shown below. The $x$-axis is at ground level, and the goal posts lie on the line $x = 1$.




**Q.** 1. If the crossbar on the goal posts is 3 metres high, does Daniel's kick make it over?

2. How far is the ball from Daniel when it touches the ground again?

**A.** 1. Substituting $x = 1$ into $y = 0.1(12 - x)(x + 3)$ gives

$y = 0.1(12 - 1)(1 + 3)$ — substituting $x = 1$ in $y = 0.1(12 - x)(x + 3)$

$= 4.4$

The height of the ball as it went between the posts was 4.4 m, so Daniel's kick did pass over the 3-metre high crossbar.

2. Substituting $y = 0$ into $y = 0.1(12 - x)(x + 3)$ gives $0.1(12 - x)(x + 3) = 0$, which has solutions $x = -3$ and $x = 12$.

This means that Daniel is 3 m left of the $y$-axis as he kicks the ball, and the ball lands 12 m to the right of the $y$-axis, a distance of $3 + 12 = 15$ m from Daniel.

In some problems you may need to form equations from the information given.

**Example**

**Q.** Fred and his younger friend Rebecca notice that when they multiply their ages together they get 72, and when they add their ages they get 17. Form and solve a pair of equations to find the ages of the two friends.

**A.** Let $F$ = Fred's age and $R$ = Rebecca's age

From the information given:

$FR = 72$ ...(1) — product of ages is 72

$F + R = 17$ ...(2) — sum of ages is 17

Rearranging (2) gives $F = 17 - R$.

Substituting this into (1) gives:

$(17 - R)R = 72$

$17R - R^2 = 72$ — expanding

$R^2 - 17R + 72 = 0$ — rearranging

$(R - 9)(R - 8) = 0$ — factorising

$R = 9$ or $R = 8$ — setting factors to 0

Substituting $R = 9$ or $R = 8$ into (2) gives $F = 8$ or $F = 9$, respectively.

But Rebecca is younger, so Fred is 9 and Rebecca is 8.

 ISBN 978-0-947504-98-4 

# Answers

## *Achievement Standard 91256 (Mathematics and Statistics 2.1)*

### Gradients of lines (page 4)

1. a. $\frac{3}{4}$ b. –1 c. 0 d. $\frac{1}{4}$
2. a. $\frac{5}{6}$ b. $2\frac{2}{5}$ c. 1 d. 0 e. $-\frac{3}{5}$ f. 1
3. a. $\frac{1}{3}$ b. –1 c. 2 d. 1
4. a. 1 b. –2 c. $-\frac{2}{3}$ d. $-\frac{2}{3}$
5. a. A (0,3), B (2,–2) b. $-\frac{5}{2}$ c. $\frac{5}{2}$
   d. i. AC is steeper (shorter ladder reaching same height)
      ii. $\frac{5}{3}$ (which is less than $\frac{5}{2}$)
6. a. $\frac{1}{2}$ b. –2 c. BC

### Distance between two points (page 8)

1. a. i. 4.24 ii. 5.10
   b. i. 5 ii. 5
   c. i. 9.22 ii. 6
2. a. 3.61 b. 10.44 c. 19.80 d. 10 e. 7.07 f. 5*a*
3. 549 mm
4. 11.7 km
5. a. 7 m b. i. 15 m ii. 20 m c. 12.82 m (2 d.p.)
6. 3.464 m (3 d.p.)
7. 8.93 or –4.93 (2 d.p.)

### Midpoint between two points (page 11)

1. a. (3,6) b. (0,2) c. (3,6) d. (–8,3)
   e. (4.5,5) f. (1,–2)
2. a. (1,0) b. (7,–4) c. (5,0)
3. a. (–1,0) b. 5

### Gradient-intercept form of the equation of a straight line (page 14)

1. a. $y = -2x + 4$ b. $y = 3x + 1$
   c. $y = x - 2$ d. $y = \frac{1}{2}x$
   e. $y = -2x + 4$
2. a. $y = -2x + 1$ b. $y = x + 2$
   c. $y = -\frac{1}{2}x$ d. $y = -\frac{4}{3}x + 4$
3. a. Line passing through (0,1) and (1,3)
   b. Line passing through (0,–2) and (1,–4)
   c. Line passing through (0,3) and (2,4)
   d. Line passing through (0,0) and (4,–3)
4. a. i. $y = -x + 4$ ii. $y = 2x + 3$
      iii. $y = \frac{1}{4}x + 2$ iv. $y = -2x - 4$
   b. i. Line passing through (0,4) and (1,3)
      ii. Line passing through (0,3) and (1,5)
      iii. Line passing through (0,2) and (4,3)
      iv. Line passing through (0,–4) and (–1,–2)
5. a. $y = \frac{2}{5}x + 2$ (or $2x - 5y = -10$)
   b. $y = \frac{2}{3}x - 4$ (or $2x - 3y = 12$)
   c. $y = -\frac{3}{4}x - 3$ (or $3x + 4y = -12$)
   d. $y = -\frac{2}{3}x + 10$ (or $2x + 3y = 30$)

### Equation of line when gradient and a point are known (page 16)

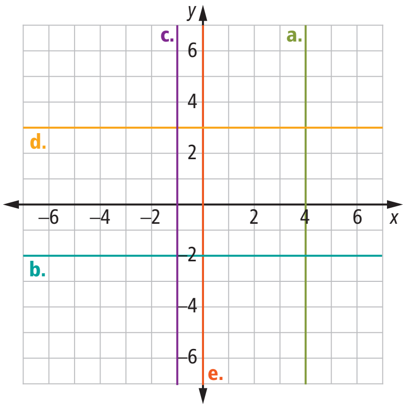
1. $y = 2x - 1$ 2. $y = -x + 7$ 3. $y = -4x - 1$
4. $y = \frac{1}{2}x + 4$ 5. $y = \frac{3}{2}x - 4$ 6. $y = -\frac{3}{4}x - 6$
7. $y = 0.5x + 1$
8. a. $-\frac{13}{14}$
   b. $y = -\frac{13}{14}x + 16\frac{15}{28}$ (or $26x + 28y = 463$)
   c. Both gradients are equal to $-\frac{13}{14}$

### Equation of a line when two points are known (page 18)

1. $y = -\frac{1}{3}x + \frac{2}{3}$ 2. $y = \frac{7}{4}x + \frac{1}{4}$
3. $y = -x + 10$ 4. $y = -x + 2$
5. $y = x - 3$ 6. $y = \frac{2}{3}x - 2$ 7. $y = 2$
8. a. –2 b. $y = -2x + 110$ c. $y = x + 20$ d. 113.14 m (2 d.p.)

### Equations of special lines (page 22)

1.



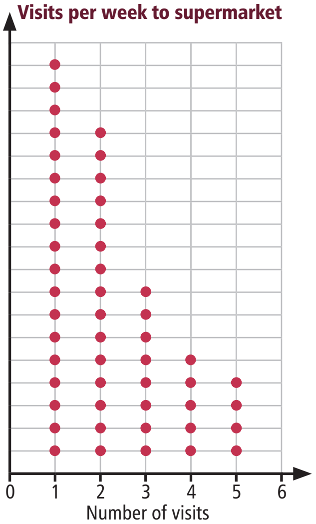

2. a. $x = -4$ b. $x = 6$ c. $y = 5$ d. $y = -6$ e. $y = 0$
3. a. $y = -2x$ b. $y = x$
   c. $y = 3x - 15$ d. $y = -\frac{2}{3}x + \frac{17}{3}$ (or $2x + 3y = 17$)
   e. $y = \frac{3}{2}x - \frac{7}{2}$ (or $3x - 2y = 7$)

 ISBN 978-0-947504-98-4 

b. i. 5.7 sec
ii. *Answers will vary*, e.g. different means or quartiles
iii. Sampling variability, each student's sample will be different.
c. i. Samples are likely to be more similar.
ii. Large sample sizes give less variability.

## Displaying sample distributions (page 132)

1. a.

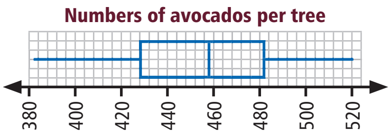



b. *Answers will vary*, e.g. Most shoppers went at most twice in the previous week, with 1 visit the most common number. The distribution is skewed right.

2. *Explanations will vary*.

a.

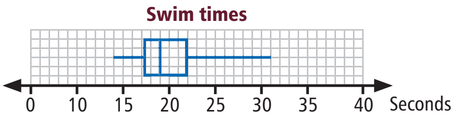
| Min | LQ | Med | UQ | Max |
|---|---|---|---|---|
| 382 | 427 | 457.5 | 482 | 521 |




The data in the sample is slightly negatively skewed (median of 457.5 avocados per tree is placed a little right of centre in the plot) with an interquartile range of 55 avocados per tree. So the numbers of fruit per tree is fairly consistent. This would be expected if these numbers came from a commercial orchard business, in which the correct species of tree for the conditions would be selected and similar conditions (watering, sunshine, etc.) provided for all trees in the orchard.

b. All measurements in seconds.

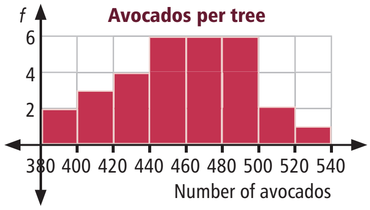
| Min | LQ | Med | UQ | Max |
|---|---|---|---|---|
| 14 | 17 | 19 | 22 | 31 |




The data in the sample is positively skewed (the median of 19 seconds is left of centre in the plot) so the faster times were spread over a smaller range than the slower times which were more spread out. This could be because of a group in the sample who train for this event. The maximum time of 31 sec is over double the fastest time, so there are obviously also some less able swimmers in the sample. It is therefore more likely that the sample is from a population of mixed ability, such as a Year group, and less likely to be from a population of members of a competitive swim club.

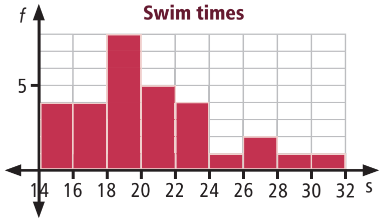
3. *Comments will vary*.

a. i.



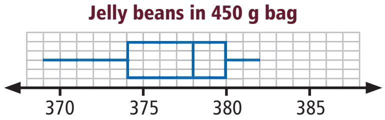

ii. The histogram shows that the distribution of the data is fairly symmetrical in shape. This could mean that the numbers of avocados per tree is actually symmetrically distributed back in the population. However, with a sample of only 30, the symmetry of the sample distribution could have resulted from the variability of sampling (other samples may be less symmetrically distributed), so that the population distribution is in fact more skewed than this sample would indicate.

b. i.




ii. The histogram clearly shows the positive skew of the sample. A feature not evident from the box plot is the large number of times (8 out of 30) in the range 18–19 seconds. This could be due to sampling variability (without this peak the sample is fairly uniformly distributed between 14 seconds and 23 seconds) and another sample may not show this feature. Alternatively it could mean that there is a similar peak back in the population from which this sample was selected. Further sampling would be necessary to be more definite about the distribution of the population.

4.




*Answers will vary*.

For this sample of thirty 450 g bags, the numbers of jellybeans per bag lies between 369 and 382 (a range of 13 jellybeans). This range seems quite large for bags of jellybeans being produced commercially by precision machines, which would be expected to produce uniformly sized jelly beans, and weigh the bags very accurately. The distribution of the numbers of jellybeans per bag is negatively skewed, with 50% of the bags having as many as 378–382 jellybeans per bag. The negative skew is likely to occur in the population, because there is a limit to how many jellybeans can fit in a bag before the bag bursts. There is no such restriction on bags with low numbers of jellybeans.

The middle 50% of bags had 374–380 jellybeans per bag, giving the sample an interquartile range of 6 jellybeans per bag (this seems a somewhat large interquartile range, for the same reasons as given above).

Despite the size of the sample being appropriately large (50 bags), this may be an unusual sample for some reason (e.g. the sample may not have been selected using a random sampling method). However, if the sample is representative of the population, then it would be inferred that the population distribution also has quite a lot of variability. This may suggest that the machines producing and weighing the jellybeans need servicing, so that the variability in the weights and numbers of jellybeans per bag is reduced.

The distribution is negatively skewed with an interquartile range of 6 jelly beans. The range of 13 jelly beans seems quite large for bags being produced commercially by precision machines. This sample could be unusual, however, and a different sample could give a different distribution.

5. The times taken to solve the memory game lay between 21 seconds and 66 seconds for this sample of students. The distribution of memory game times is positively skewed; the middle 50% of times is also positively skewed with an interquartile range of 10.5 seconds.

 ISBN 978-0-947504-98-4 

# Index



 ISBN 978-0-947504-98-4 