



BRITISH LEARNING CENTRE

CAMBRIDGE IGCSE · 0580

Mathematics

Probability & Statistics

Topic 8 Probability and Topic 9 Statistics of the Cambridge IGCSE Mathematics 0580 syllabus, Core and Extended. Each unit has key knowledge, worked examples, a practice sheet and exam-style questions with full mark schemes.

Name:

Key knowledge · Practice · Exam-style questions · Mark schemes

How to use this workbook

1

Key knowledge

Read the notes, worked examples and common mistakes.

2

Practice

Build the skill on Sheet A.

3

Exam-style

Answer exam-style questions on Sheet B. Marks are shown in [].

4

Check & reflect

Mark your work with the mark scheme, then colour the tracker.

My progress tracker

Colour each row: **Red** = not yet · **Amber** = nearly · **Green** = secure.

#	Syllabus ref	Topic · I can...	R	A	G
1	C8.1, E8.1 Core & Extended	The probability scale and single events I can find the probability of a single event and the probability of an event not happening.			
2	C8.2, E8.2 Core & Extended	Relative frequency and expected frequency I can estimate probabilities from experiments and calculate how many times an event is expected to happen.			
3	C8.3, E8.3 Core & Extended	Sample spaces and combined events I can list outcomes in a sample space and find probabilities of two events happening together.			
4	C8.3, E8.3 Core & Extended	Tree diagrams I can draw and use tree diagrams for two or more events, with or without replacement.			
5	C8.3, E8.3 Core & Extended	Venn diagrams and probability I can use Venn diagrams and set notation to find probabilities.			
6	E8.4 Extended	Conditional probability I can find conditional probabilities from tables, Venn diagrams and tree diagrams.			
7	C9.1, E9.1, C9.2 Core & Extended	Types of data and tables I can classify data and organise it in frequency tables and two-way tables.			
8	C9.2, C9.4, E9.4 Core & Extended	Bar charts, pictograms and pie charts I can draw and interpret bar charts, pictograms and pie charts.			
9	C9.3, E9.3, C9.4, E9.4 Core & Extended	Averages, range and stem-and-leaf I can find the mean, median, mode and range from lists, tables and stem-and-leaf diagrams.			
10	E9.3 Extended	Grouped data and the estimated mean I can estimate the mean of grouped data and find the modal class and the class containing the median.			
11	C9.5, E9.5 Core & Extended	Scatter diagrams and correlation I can draw and interpret scatter diagrams, describe correlation and use a line of best fit.			
12	E9.3, E9.6 Extended	Cumulative frequency, quartiles and box plots I can draw cumulative frequency diagrams and find medians, quartiles, percentiles and box plots.			
13	E9.7 Extended	Histograms and frequency density I can draw and interpret histograms with unequal class widths using frequency density.			



CHECK

Starting point check

Starting point check

Name _____ Date _____ Marks _____ / 13

Answer as many questions as you can. This check shows what you already know — do not worry if some topics are new.

- 1 A fair spinner has 8 equal sections numbered 1 to 8. Find the probability that it lands on a number less than 3. [1]

Answer

- 2 The probability of an event is 0.45. Find the probability that it does not happen. [1]

Answer

- 3 A coin lands on heads 27 times in 60 throws. Find the relative frequency of heads. [1]

Answer

- 4 The probability that a spinner lands on red is 0.2. It is spun 150 times. How many reds would you expect? [1]

Answer

- 5 Two fair coins are thrown. Find the probability of two tails. [1]

Answer

- 6 In a group of 25 people, 12 are in set A, 9 are in set B and 4 are in both. How many are in neither set? [2]

Answer

- 7 Find the mean of 4, 8, 6, 10, 7. [1]

Answer

- 8 Find the median of 3, 9, 5, 12, 8, 6. [1]

Answer

- 9 In a survey of 90 people, 15 chose blue. Find the angle for blue in a pie chart. [1]

Answer

- 10 Is the mass of a dog qualitative, discrete or continuous data? [1]

Answer

- 11 What type of correlation would you expect between hours of sunshine and the number of umbrellas sold? [1]

Answer

- 12 A class $20 < x \leq 30$ has frequency 25. Find its frequency density. [1]

Answer

1

UNIT 1 · KEY KNOWLEDGE

The probability scale and single events

I can... find the probability of a single event and the probability of an event not happening.

Syllabus ref: C8.1, E8.1 · Core & Extended

Probability measures how likely an event is, on a scale from 0 (impossible) to 1 (certain).

When all outcomes are **equally likely**: $P(\text{event}) = \text{number of favourable outcomes} \div \text{total number of outcomes}$.

$P(A')$ is the probability that A does **not** happen. The probabilities of all possible outcomes add up to 1.

Key words

probability

outcome

equally likely

event

complement

$P(A')$

Key points

- $P(A)$ means 'the probability of event A'; $0 \leq P(A) \leq 1$
- $P(A') = 1 - P(A)$
- The probabilities of all the different outcomes add up to 1

Worked example 1

A bag has 5 red, 3 blue and 4 green counters. One is taken at random. Find (a) $P(\text{blue})$ (b) $P(\text{not green})$.

Total = $5 + 3 + 4 = 12$ counters

(a) $P(\text{blue}) = 3/12 = \mathbf{1/4}$

(b) Not green = 8 counters, so $P(\text{not green}) = 8/12 = \mathbf{2/3}$

Worked example 2

A spinner lands on red with probability 0.16, blue 0.24, green x and yellow $2x$. Find x .

$0.16 + 0.24 + x + 2x = 1$

$3x = 0.6$, so $x = \mathbf{0.2}$ (and $P(\text{yellow}) = 0.4$)

Common mistakes

- Writing '1 : 4' or '1 in 4' instead of $1/4$.
- Not simplifying, e.g. leaving $3/12$.

Exam tip: When a table of probabilities has a missing value, start with: all the probabilities add up to 1.

Answer each question. Give probabilities as fractions in their simplest form unless told otherwise.

1 A fair six-sided dice is rolled. Find $P(5)$.

2 A fair six-sided dice is rolled. Find $P(\text{even number})$.

3 A fair six-sided dice is rolled. Find $P(\text{number greater than 4})$.

4 A fair six-sided dice is rolled. Find $P(7)$.

5 A letter is chosen at random from the word STATISTICS. Find $P(S)$.

6 A letter is chosen at random from the word STATISTICS. Find $P(I)$.

7 A letter is chosen at random from STATISTICS. Find $P(\text{the letter is a vowel})$.

8 A letter is chosen at random from STATISTICS. Find $P(\text{the letter is not T})$.

9 A card is chosen at random from cards numbered 1 to 20. Find $P(\text{multiple of 3})$.

10 A card is chosen at random from cards numbered 1 to 20. Find $P(\text{prime number})$.

11 A card is chosen at random from cards numbered 1 to 20. Find $P(\text{square number})$.

12 $P(A) = 0.37$. Find $P(A')$.

13 The probability that Lina wins a game is $\frac{2}{9}$. Find the probability that she does not win.

14 What word describes an event with probability 0?

15 A spinner can land on red, blue or green only. $P(\text{red}) = 0.35$ and $P(\text{blue}) = 0.25$. Find $P(\text{green})$.

16 A box has 6 milk and some dark chocolates. $P(\text{milk}) = \frac{1}{4}$. How many dark chocolates are there?

17 A bag holds red and white beads in the ratio 2 : 5. Find $P(\text{white})$.

18 Which is more likely: an event with probability 0.7 or one with probability $\frac{7}{11}$?

Reflect: ☐ Red — not yet ☐ Amber — nearly ☐ Green — secure

Answer all questions. Show your working clearly.

- 1 A spinner can land on red, green, blue or yellow. The table shows some of the probabilities.

Colour	Red	Green	Blue	Yellow
Probability	0.35	0.2	x	$x + 0.05$

- (a) Find the value of x .

[3]

Answer $x = \dots\dots$

- (b) Find the probability that the spinner lands on yellow.

[1]

Answer $\dots\dots$

- (c) Find the probability that the spinner does **not** land on red.

[1]

Answer $\dots\dots$

- 2 Ten discs are numbered 1, 2, 2, 3, 5, 5, 5, 8, 9, 10. One disc is chosen at random.

- (a) Find the probability that the number is odd.

[1]

Answer $\dots\dots$

- (b) Find the probability that the number is not 5.

[1]

Answer $\dots\dots$

- (c) Find the probability that the number is greater than 10.

[1]

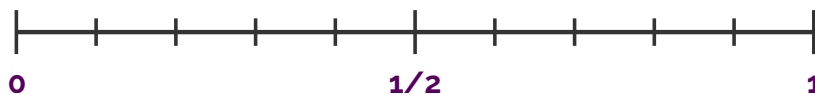
Answer $\dots\dots$

- (d) Some more discs, all numbered 5, are added. The probability of choosing a 5 is now $\frac{1}{2}$. How many discs are added?

[3]

Answer $\dots\dots$

- 3 The probability scale is shown below.



- (a) Mark with an arrow labelled H the probability that a fair coin lands on heads. [1]
- (b) Mark with an arrow labelled C the probability that a card taken at random from a normal pack of 52 is a club. [1]
- (c) $P(R)$ is the probability that it rains tomorrow. $P(R) = 0.62$. Write down, in words, what $P(R)$ means and find its value. [2]

Answer $P(R) = \dots\dots$

- 4 In a class of 30 students, 12 have brown eyes, 9 have blue eyes and the rest have green eyes. A student is chosen at random.

- (a) Find the probability that the student has green eyes. [2]

Answer

- (b) Find the probability that the student does not have blue eyes. [1]

Answer

- (c) Three new students with brown eyes join the class. Find the new probability that a student chosen at random has brown eyes. [2]

Answer

- 5 A box contains 30 pens. Each pen is red, black or blue. The probability of choosing a blue pen is $\frac{1}{6}$. The ratio red pens : black pens is 2 : 3.

(a) Find the number of black pens.

[3]

Answer

(b) A pen is taken at random. Find the probability that it is red. Give your answer in its simplest form.

[2]

Answer

Reflect: ☐ Red — not yet ☐ Amber — nearly ☐ Green — secure

Relative frequency and expected frequency

I can... estimate probabilities from experiments and calculate how many times an event is expected to happen.

Syllabus ref: C8.2, E8.2 · Core & Extended

When outcomes are not equally likely, we estimate probability by experiment. **Relative frequency** = number of times the event happens $\div$ number of trials.

The more trials there are, the more reliable the estimate. A **fair** dice or coin has equally likely outcomes; a **biased** one does not.

Expected frequency = probability $\times$ number of trials. It does not have to be a whole number.

Key words

relative frequency

trial

estimate

expected frequency

biased

fair

Key points

- Relative frequency = frequency $\div$ total trials
- Expected frequency = $P(\text{event}) \times$ number of trials

Worked example 1

A dice is rolled 200 times and a six appears 46 times. Find the relative frequency of a six. Is the dice fair?

Relative frequency = $46/200 = 0.23$

A fair dice would give about $200 \times 1/6 \approx 33$ sixes.

46 is a lot more, so the dice **may be biased** — more trials would make us more sure.

Worked example 2

$P(\text{a train is late}) = 0.15$. How many of the next 60 trains are expected to be late?

Expected frequency = $0.15 \times 60 = 9$ trains

Common mistakes

- Dividing by the wrong total — always divide by the total number of trials.
- Thinking an expected frequency is what will definitely happen.

Exam tip: If asked which estimate is best, choose the one from the largest number of trials and say why.

Answer each question. Give probabilities as fractions in their simplest form unless told otherwise.

- 1 A drawing pin is dropped 50 times. It lands point up 18 times. Find the relative frequency of point up.

- 2 A coin is spun 80 times and lands on heads 44 times. Find the relative frequency of heads.

- 3 A dice is rolled 300 times. A six appears 62 times. Find the relative frequency of a six.

- 4 The probability that a bus is late is 0.15. How many of the next 60 buses are expected to be late?

- 5 A fair dice is rolled 240 times. How many fours would you expect?

- 6 A fair coin is thrown 75 times. How many heads would you expect?

- 7 $P(\text{win}) = \frac{3}{8}$. A game is played 200 times. Expected number of wins?

- 8 A spinner has $P(\text{red}) = 0.28$. It is spun 450 times. Expected number of reds?

- 9 Out of 400 light bulbs, 6 are faulty. Estimate $P(\text{a bulb is faulty})$.

- 10 Using your answer to question 9, how many faulty bulbs would you expect in 5000?

- 11 A biased coin lands on tails with relative frequency 0.42 after 500 throws. How many tails were there?

- 12 Ali throws a dice 12 times; Ben throws it 600 times. Whose results give the better estimate of probability?

- 13 In 120 spins, red came up 30 times. Estimate the probability of **not** red.

- 14 The expected number of goals scored from 40 penalties is 34. Estimate $P(\text{scoring a penalty})$.

Reflect: ☐ Red — not yet ☐ Amber — nearly ☐ Green — secure

Answer all questions. Show your working clearly.

- 1 Sara spins a four-coloured spinner. She records the total number of times it has landed on red after different numbers of spins.

Number of spins	20	50	100	200	500
Number of reds	7	14	31	58	147
Relative frequency	0.35				

(a) Complete the table by working out the relative frequencies of red. Give each answer as a decimal. [2]

(b) Which relative frequency gives the best estimate of the probability of red? Give a reason. [2]

(c) If the spinner were fair, the probability of red would be 0.25. Is there any evidence that the spinner is biased? Explain. [1]

(d) Sara spins the spinner another 1500 times. Estimate how many times it lands on red. [2]

Answer

- 2** A six-sided dice is rolled 200 times. The results are shown.

Score	1	2	3	4	5	6
Frequency	32	35	29	30	34	40

- (a)** Find the relative frequency of rolling a 6. **[1]**

Answer

- (b)** How many times would you expect each number to appear if the dice were fair? **[2]**

Answer

- (c)** Kai says, 'The dice is biased because the 6 came up more often than the other numbers.' Comment on his statement. **[1]**

- 3** A factory tests its light bulbs. In a sample of 800 bulbs, 12 are faulty.

- (a)** Estimate the probability that a bulb chosen at random is faulty. **[1]**

Answer

- (b)** The factory makes 24 000 bulbs in a week. Estimate how many are faulty. **[2]**

Answer

- (c)** Give one reason why your answer to (b) is only an estimate. **[1]**

4 The probability that a seed grows into a plant is 0.86.

(a) Find the expected number of plants from 350 seeds. [2]

Answer

(b) Mia wants at least 430 plants. Find the smallest number of seeds she should plant. [3]

Answer

5 A game at a school fair costs 50 cents to play. The probability of winning is 0.12. A winner receives \$3. The game is played 250 times.

(a) Find the expected number of winners. [2]

Answer

(b) Find the expected profit for the school, in dollars. [3]

Answer \$

Reflect: ☐ Red — not yet ☐ Amber — nearly ☐ Green — secure

Sample spaces and combined events

I can... list outcomes in a sample space and find probabilities of two events happening together.

Syllabus ref: C8.3, E8.3 · Core & Extended

A **sample space** is a list or table of every possible outcome. Work **systematically** so that nothing is missed. For two dice there are $6 \times 6 = 36$ equally likely outcomes.

Two events are **independent** if one does not affect the other. Then $P(A \text{ and } B) = P(A) \times P(B)$.

Key words

sample space

systematic

independent

combined event

outcome

Key points

- Total outcomes for two events = (outcomes of first) $\times$ (outcomes of second)
- Independent: $P(A \text{ and } B) = P(A) \times P(B)$
- $P(\text{at least one}) = 1 - P(\text{none})$

Worked example 1

Spinner A shows 1–4 and spinner B shows 1–3. Both are spun and the scores added. Find $P(\text{total} = 5)$.

There are $4 \times 3 = 12$ outcomes.

Totals of 5: (2, 3), (3, 2), (4, 1) — 3 outcomes.

$P(\text{total} = 5) = 3/12 = \mathbf{1/4}$

Worked example 2

A fair coin is thrown and a fair dice is rolled. Find $P(\text{head and a 6})$.

The events are independent.

$P = 1/2 \times 1/6 = \mathbf{1/12}$

Common mistakes

- Missing outcomes, e.g. counting (1, 2) but not (2, 1).
- Adding probabilities when you should multiply for 'and'.

Exam tip: Draw a full sample space table for two dice — it takes one minute and stops counting mistakes.

Answer each question. Give probabilities as fractions in their simplest form unless told otherwise.

1 Two fair coins are thrown. List all the possible outcomes.

2 Two fair coins are thrown. Find $P(\text{two heads})$.

3 Two fair coins are thrown. Find $P(\text{one head and one tail})$.

4 Two fair dice are rolled and the scores added. How many outcomes are there?

5 Two fair dice are rolled and the scores added. Find $P(\text{total} = 7)$.

6 Two fair dice are rolled and the scores added. Find $P(\text{total} = 11)$.

7 Two fair dice are rolled and the scores added. Find $P(\text{total is more than 9})$.

8 Two fair dice are rolled. Find $P(\text{a double})$.

9 Two fair dice are rolled. Find $P(\text{the product of the scores is 12})$.

10 A fair coin is thrown and a fair dice is rolled. Find $P(\text{tail and 3})$.

11 A fair coin is thrown and a fair dice is rolled. Find $P(\text{head and even number})$.

12 $P(A) = 0.3$ and $P(B) = 0.5$. A and B are independent. Find $P(A \text{ and } B)$.

13 A fair coin is thrown three times. Find $P(\text{three tails})$.

14 Starters: soup or salad; mains: fish, beef or pasta. How many different two-course meals?

Reflect: ☐ Red — not yet ☐ Amber — nearly ☐ Green — secure

Answer all questions. Show your working clearly.

- 1 Spinner A has sides numbered 1, 2, 3, 4. Spinner B has sides numbered 2, 4, 6. Both spinners are fair. They are spun and the two numbers are **multiplied**.

×	2	4	6
1	2		
2		8	
3			
4			24

(a) Complete the sample space diagram. [2]

(b) Find the probability that the product is 12. [1]

Answer

(c) Find the probability that the product is greater than 10. [2]

Answer

(d) Find the probability that the product is an odd number. [1]

Answer

- 2 Two fair dice are rolled. The score is the **difference** between the two numbers.

(a) Find the probability that the score is 0. [1]

Answer

(b) Find the probability that the score is 2. [2]

Answer

(c) Which score is most likely? Give its probability. [2]

Answer score probability

- 3** The probability that it rains on Saturday is 0.3. The probability that it rains on Sunday is 0.4. These events are independent.

(a) Find the probability that it rains on both days.

[2]

Answer

(b) Find the probability that it rains on neither day.

[2]

Answer

(c) Find the probability that it rains on at least one of the two days.

[2]

Answer

-
- 4** The digits 1, 2 and 3 are each used once to make a three-digit number, chosen at random.

(a) List all the possible numbers.

[2]

(b) Find the probability that the number is greater than 300.

[1]

Answer

(c) Find the probability that the number is even.

[1]

Answer

- 5 A fair red dice is numbered 1 to 6. A fair blue dice has faces 1, 1, 2, 2, 3, 3. Both are rolled and the scores are added.

(a) Find the probability that the total is 4.

[2]

Answer

(b) Find the probability that the total is a prime number.

[2]

Answer

(c) Find the probability that the red dice shows an even number and the blue dice shows 3.

[1]

Answer

Reflect: ☐ Red — not yet ☐ Amber — nearly ☐ Green — secure

4

UNIT 4 · KEY KNOWLEDGE

Tree diagrams

I can... draw and use tree diagrams for two or more events, with or without replacement.

Syllabus ref: C8.3, E8.3 · Core & Extended

A **tree diagram** shows two or more events in order. The branches from one point add up to 1.

Multiply along the branches to find the probability of a route. **Add** the results for different routes that give the outcome you want.

Without replacement (Extended), the second set of probabilities changes because there is one item fewer.

Key words

tree diagram

branch

with replacement

without replacement

independent

Key points

- Multiply along a route ('and'); add between routes ('or')
- Without replacement: reduce the total, and the count of the colour taken, by 1

Worked example 1

A bag has 3 red and 5 blue beads. A bead is taken and replaced, then a second is taken. Find P(both the same colour).

$$P(R, R) = 3/8 \times 3/8 = 9/64 \text{ and } P(B, B) = 5/8 \times 5/8 = 25/64$$

$$P(\text{same}) = 9/64 + 25/64 = 34/64 = \mathbf{17/32}$$

Worked example 2 (Extended)

From the same bag, two beads are taken without replacement. Find P(both red).

After one red is taken: 2 red out of 7 remain.

$$P(R, R) = 3/8 \times 2/7 = 6/56 = \mathbf{3/28}$$

Common mistakes

- Adding along branches instead of multiplying.
- Without replacement: changing the denominator but not the numerator.
- Forgetting that 'one of each' has two routes (R then B, B then R).

Exam tip: Check your tree: every pair of branches must add to 1, and all the final outcomes must add to 1.

Answer each question. Give probabilities as fractions in their simplest form unless told otherwise.

- 1 A biased coin has $P(\text{head}) = 0.6$. It is thrown twice. Find $P(\text{two heads})$.

- 2 For the same coin, find $P(\text{two tails})$.

- 3 For the same coin, find $P(\text{one head and one tail})$.

- 4 A bag has 4 green and 6 yellow balls. One is taken and replaced, then another. Find $P(\text{both green})$.

- 5 Same bag, with replacement. Find $P(\text{both yellow})$.

- 6 Same bag, with replacement. Find $P(\text{one of each colour})$.

- 7 Same bag, with replacement. Find $P(\text{at least one green})$.

- 8 $P(\text{pass driving test}) = 0.7$ at each attempt. Find $P(\text{fail the first, pass the second})$.

- 9 **(Extended)** A bag has 4 green and 6 yellow balls. Two are taken **without** replacement. Find $P(\text{both green})$.

- 10 **(Extended)** Same bag, without replacement. Find $P(\text{both yellow})$.

- 11 **(Extended)** Same bag, without replacement. Find $P(\text{one of each colour})$.

- 12 **(Extended)** Two cards are taken from 5 cards numbered 1–5 without replacement. Find $P(\text{both odd})$.

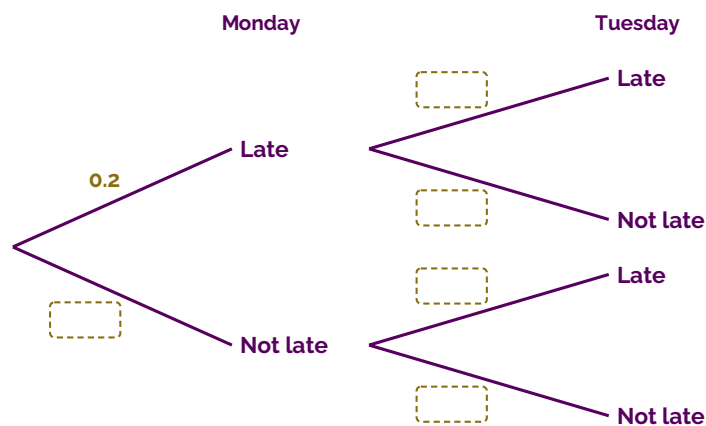
- 13 The branches from one point of a tree diagram show 0.35 and .

- 14 On a tree diagram, do you multiply or add along the branches?

Reflect: ☐ Red — not yet ☐ Amber — nearly ☐ Green — secure

Answer all questions. Show your working clearly.

- 1 The probability that Ravi's bus is late on any day is 0.2. The days are independent.



- (a) Complete the tree diagram for Monday and Tuesday. [2]
- (b) Find the probability that the bus is late on both days. [2]

Answer

- (c) Find the probability that the bus is late on exactly one of the two days. [3]

Answer

- 2 A bag contains 7 red and 3 white counters. Jo takes a counter at random, notes its colour and **replaces** it. She then takes a second counter.

(a) Draw a tree diagram to show this information.

[3]

(b) Find the probability that both counters are white.

[2]

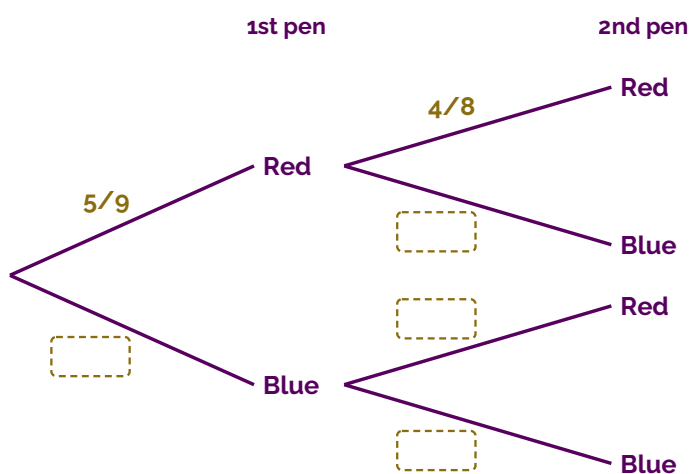
Answer

(c) Find the probability that at least one counter is red.

[2]

Answer

- 3 (Extended) A box has 5 red pens and 4 blue pens. Two pens are taken at random **without replacement**.



(a) Complete the tree diagram.

[2]

(b) Find the probability that both pens are the same colour.

[3]

Answer

(c) Find the probability that the pens are different colours.

[1]

Answer

- 4 **(Extended)** A packet has 6 lemon sweets and 4 orange sweets. Ana takes three sweets at random without replacement.

(a) Show that the probability that all three sweets are lemon is $\frac{1}{6}$. **[2]**

(b) Find the probability that at least one sweet is orange. **[1]**

Answer

(c) Find the probability that exactly two of the sweets are lemon. **[3]**

Answer

Reflect: ☐ Red — not yet ☐ Amber — nearly ☐ Green — secure

Venn diagrams and probability

I can... use Venn diagrams and set notation to find probabilities.

Syllabus ref: C8.3, E8.3 · Core & Extended

A **Venn diagram** shows how many items (or what probability) are in each region. The rectangle is the universal set $\mathcal{E}$.

$A \cap B$ (**intersection**) means in A **and** B. $A \cup B$ (**union**) means in A **or** B or both. A' means **not** in A.

$P(\text{region}) = \text{number in the region} \div \text{total number in } \mathcal{E}$. Always fill the middle ($\cap$) region first.

Key words

Venn diagram

intersection $\cap$

union $\cup$

complement A'

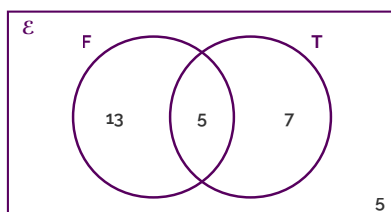
universal set $\mathcal{E}$

Key points

- $n(A \cup B) = n(A) + n(B) - n(A \cap B)$
- $P(A') = 1 - P(A)$
- All regions, including the outside, add up to the total

Worked example

Of 30 students, 18 play football (F), 12 play tennis (T) and 5 play both. Find $P(\text{T only})$ and $P(F \cup T)$.



Fill the Venn: both = 5, F only = 13, T only = 7, neither = $30 - 25 = 5$

$P(\text{T only}) = \mathbf{7/30}$

$P(F \cup T) = 25/30 = \mathbf{5/6}$

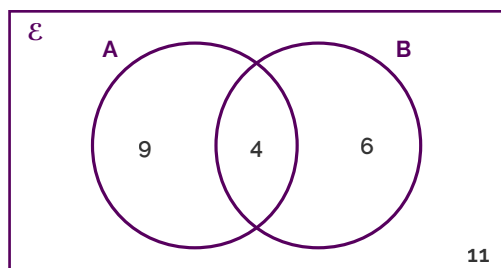
Common mistakes

- Putting the total of A (e.g. 18) in the 'A only' region.
- Forgetting the numbers outside the circles.
- Mixing up $\cap$ (and) with $\cup$ (or).

Exam tip: Start with the intersection, then subtract it from each set total, then find the outside region last.

Answer each question. Give probabilities as fractions in their simplest form unless told otherwise.

The Venn diagram shows the numbers of students in sets A and B. How many students are there altogether?



Use the Venn diagram in question 1.

2 Find $n(A)$.

Answer

Answer

3 Using the Venn diagram in question 1, find $P(A \cap B)$.

4 Using question 1, find $P(A \cup B)$.

5 Using question 1, find $P(A')$.

6 Using question 1, find $P(B \text{ only})$.

7 Using question 1, find $P(A' \cap B')$.

8 Using question 1, find $P(A \cap B)$.

9 $\mathcal{E} = \{1, 2, \dots, 10\}$, $E = \{\text{even numbers}\}$,
 $M = \{\text{multiples of 3}\}$. List $E \cap M$.

10 Same sets. Find $P(E \cup M)$ for a number chosen at random from $\mathcal{E}$.

11 Same sets. Find $P(M')$.

12 In a group of 40, 25 like tea, 20 like coffee and 8 like both. How many like neither?

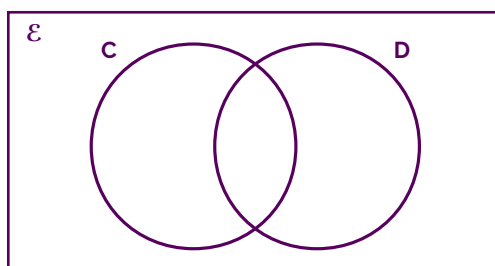
13 In question 12, find $P(\text{likes tea but not coffee})$.

14 What does $A \cup B$ mean in words?

Reflect: ☐ Red — not yet ☐ Amber — nearly ☐ Green — secure

Answer all questions. Show your working clearly.

- 1 50 people were asked whether they own a cat (C) or a dog (D). 22 own a cat, 26 own a dog and 9 own neither.



- (a) Find how many people own both a cat and a dog. [2]

Answer

- (b) Complete the Venn diagram. [2]

- (c) A person is chosen at random. Find $P(C \cup D)$. [1]

Answer

- (d) Find $P(D \cap C')$. [1]

Answer

2 $\mathcal{E} = \{1, 2, 3, \dots, 15\}$, $X = \{\text{prime numbers}\}$, $Y = \{\text{odd numbers}\}$.

(a) Draw a Venn diagram to show this information, putting every number in the correct region. [3]

(b) A number is chosen at random from $\mathcal{E}$. Find $P(X \cap Y)$. [1]

Answer

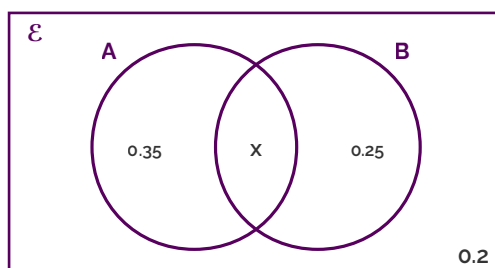
(c) Find $P(X \cup Y)$. [1]

Answer

(d) Find $P(X' \cap Y')$. [1]

Answer

3 The Venn diagram shows the probabilities of events A and B.



(a) Find the value of x . [2]

Answer $x =$

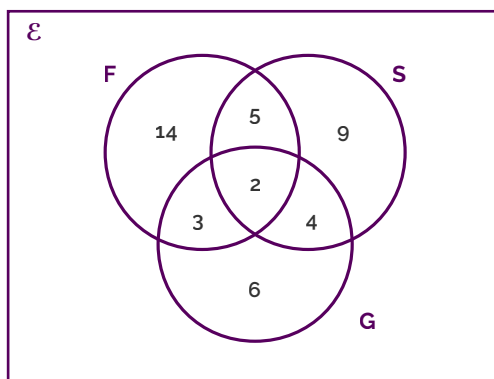
(b) Find $P(A)$. [1]

Answer

(c) Find $P(A \cup B)$. [1]

Answer

- 4 **(Extended)** 60 students were asked which languages they study: French (F), Spanish (S) or German (G). The Venn diagram shows the results. One student is chosen at random.



- (a) Find the number of students who study none of these languages. [1]

Answer

- (b) Find the probability that the student studies exactly two of the languages. [2]

Answer

- (c) Find the probability that the student studies French but not German. [2]

Answer

- 5 **(Extended)** In a class of 32 students, 14 study Art (A) and 12 study Music (M). x students study both and $2x$ study neither.

- (a) Show that $x = 6$. [2]

- (b) A student is chosen at random. Find $P(A \cap M')$. [1]

Answer

- (c) Find $P((A \cap M)')$. [1]

Answer

Reflect: ☐ Red — not yet ☐ Amber — nearly ☐ Green — secure

Conditional probability

I can... find conditional probabilities from tables, Venn diagrams and tree diagrams.

Syllabus ref: E8.4 · Extended

Conditional probability is the probability of an event **given that** another event has happened. $P(A | B)$ means 'the probability of A given B'.

'Given that B has happened' makes B the new total. From a table or Venn diagram: $P(A | B) = n(A \cap B) \div n(B)$. For example, if 18 students play football and 5 of them also play tennis, $P(\text{tennis} | \text{football}) = 5/18$.

With probabilities: $P(A | B) = P(A \cap B) \div P(B)$. Without replacement problems are conditional too.

Key words

conditional probability

given that

$P(A | B)$

dependent

independent

Key points

- $P(A | B) = n(A \cap B) \div n(B) = P(A \cap B) \div P(B)$
- On a tree, second-stage branches are conditional probabilities
- A and B are independent if $P(A | B) = P(A)$

Worked example

A bag has 3 red and 5 blue beads. Two are taken without replacement. Given that they are the same colour, find $P(\text{both red})$.

$$P(RR) = 3/8 \times 2/7 = 6/56; P(BB) = 5/8 \times 4/7 = 20/56$$

$$P(\text{same}) = 26/56$$

$$P(RR | \text{same}) = 6/56 \div 26/56 = \mathbf{3/13}$$

Common mistakes

- Dividing by the grand total instead of the 'given' total.
- Mixing up $P(A | B)$ and $P(B | A)$.

Exam tip: Underline the words 'given that': whatever follows is your new denominator.

Answer each question. Give probabilities as fractions in their simplest form unless told otherwise.

The table shows how 80 students travel to school. A student is chosen at random. Find $P(\text{Bus})$.

	Walk	Bus	Car	Total
Boys	14	12	10	36
Girls	18	16	10	44
Total	32	28	20	80

1

2

Use the table in question 1. Find $P(\text{Bus} \mid \text{Girl})$.

Answer

3

Use the table in question 1. Find $P(\text{Girl} \mid \text{Bus})$.

4

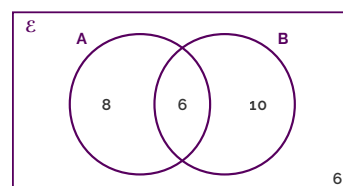
Use the table in question 1. Find $P(\text{Boy} \mid \text{Car})$.

5

Use the table in question 1. Find $P(\text{Walk} \mid \text{Boy})$.

6

The Venn diagram shows numbers of people. Find $P(A \mid B)$.



Answer

7

Using the Venn diagram in question 6, find $P(B \mid A)$.

8

A bag has 4 red and 6 blue beads. Two are taken without replacement. Find $P(\text{2nd red} \mid \text{1st red})$.

9

Same bag. Find $P(\text{2nd red} \mid \text{1st blue})$.

10

$P(A \cap B) = 0.12$ and $P(B) = 0.4$. Find $P(A \mid B)$.

11

$P(A \mid B) = P(A)$. What does this tell you about A and B?

12

A fair dice is rolled. Given that the score is even, find $P(\text{score is 6})$.

Reflect: ☐ Red — not yet ☐ Amber — nearly ☐ Green — secure

Answer all questions. Show your working clearly.

- 1 A café records the drinks bought by 200 customers.

	Tea	Coffee
Morning	46	74
Afternoon	38	42

- (a) A customer is chosen at random. Find the probability that they bought coffee. [1]

Answer

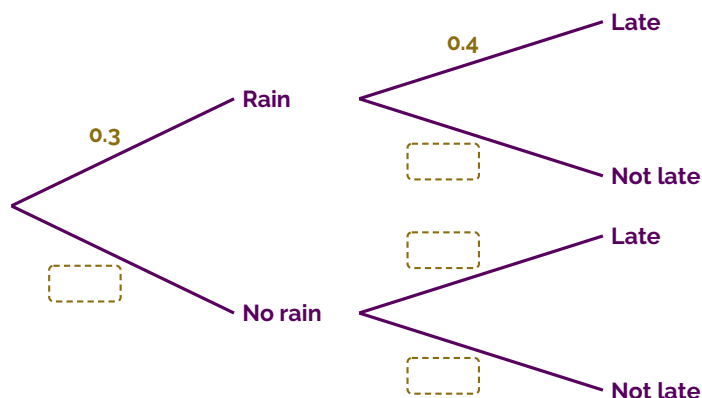
- (b) Given that the customer came in the morning, find the probability that they bought tea. [2]

Answer

- (c) Given that the customer bought coffee, find the probability that they came in the afternoon. [2]

Answer

- 2 The probability that it rains on a school day is 0.3. If it rains, the probability that Leo is late is 0.4. If it does not rain, the probability that Leo is late is 0.1.



(a) Complete the tree diagram. [2]

(b) Find the probability that Leo is late. [3]

Answer

(c) Given that Leo is late, find the probability that it rained. [2]

Answer

- 3 A bag contains 6 green and 4 yellow balls. Two balls are taken at random without replacement.

(a) Find the probability that both balls are the same colour. [3]

Answer

(b) Find the probability that at least one ball is yellow. [2]

Answer

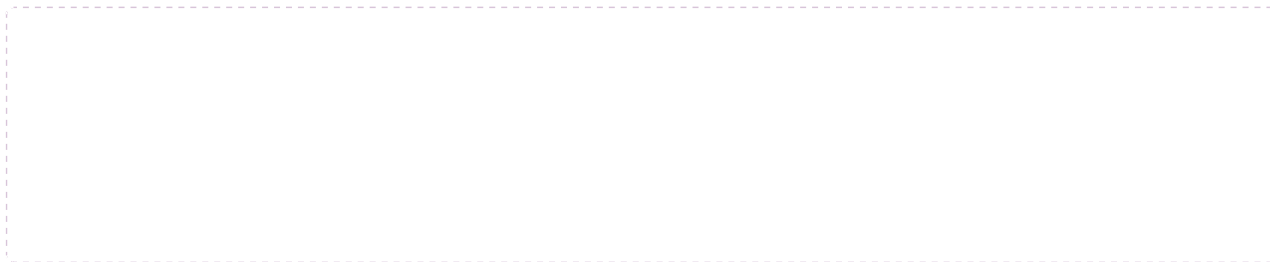
(c) Given that both balls are the same colour, find the probability that they are both green. [2]

Answer

4 For two events A and B, $P(A) = 0.5$, $P(B) = 0.4$ and $P(A \cap B) = 0.15$.

(a) Draw a Venn diagram showing the probability in each region.

[2]



(b) Find $P(A | B)$.

[2]



Answer

(c) Find $P(A | B')$.

[2]



Answer

(d) Are A and B independent? Give a reason.

[1]

Reflect: ☐ Red — not yet ☐ Amber — nearly ☐ Green — secure

Types of data and tables

I can... classify data and organise it in frequency tables and two-way tables.

Syllabus ref: C9.1, E9.1, C9.2 · Core & Extended

Qualitative data are words or categories (e.g. colour).

Quantitative data are numbers: **discrete** data are counted (e.g. number of pets); **continuous** data are measured (e.g. time, mass).

A **frequency table** shows how many times each value occurs. Continuous data are grouped into **classes** such as $10 < t \leq 20$, so that every value fits in exactly one class.

In a **two-way table**, each row and each column adds up to its total.

Key words

qualitative

discrete

continuous

frequency table

class interval

two-way table

Key points

- Discrete = counted; continuous = measured
- Classes must not overlap: use $<$ and $\leq$
- Two-way tables: rows and columns both add to the grand total

Worked example

Complete the two-way table: 40 adults chose tea or coffee. 21 are men; 12 men chose tea; 26 people chose tea.

Men, coffee: $21 - 12 = 9$

Women: $40 - 21 = 19$; women, tea: $26 - 12 = 14$

Women, coffee: $19 - 14 = 5$; coffee total: $9 + 5 = 14$

Common mistakes

- Calling heights or masses discrete because they are rounded — they are still continuous.
- Using overlapping classes such as 10–20, 20–30.
- Confusing a value with its frequency.

Exam tip: When you complete a two-way table, check the last row and column add up to the grand total.

Answer each question. Show your working where it helps.

- 1 Number of students in a class:
qualitative, discrete or continuous?

- 2 Mass of an apple: qualitative,
discrete or continuous?

- 3 Favourite sport: qualitative, discrete
or continuous?

- 4 Time taken to solve a puzzle:
qualitative, discrete or continuous?

- 5 Country of birth: qualitative,
discrete or continuous?

- Numbers of pets owned by 20
students:
0, 1, 2, 1, 0, 3, 2, 1, 1, 0, 2, 1, 4, 1, 2, 0,
1, 3, 1, 2

How many students have exactly 1
pet?

- 7 Using the pet data in question 6,
how many students have more
than 2 pets?

Complete the two-way table. How
many Year 10 students chose Music?

	Art	Music	Drama	Total
Year 10	12		9	30
Year 11		11		
Total	21			60

8

Answer

- 9 Using the table in question 8, how
many Year 11 students chose Art?

- 10 Using the table in question 8, how
many students altogether chose
Drama?

- 11 Using the table in question 8, what
fraction of all the students are in
Year 11 and chose Music?

- 12 Write 'more than 20 cm, up to and
including 30 cm' as an inequality in
h.

Reflect: ☐ Red — not yet ☐ Amber — nearly ☐ Green — secure

Answer all questions. Show your working clearly.

- 1 120 students were asked to choose a lunch. The two-way table shows some of the results.

	Hot meal	Sandwich	Salad	Total
Boys	28			64
Girls		18		
Total	50		21	120

(a) Complete the two-way table. [3]

(b) What fraction of the girls chose salad? Give your answer in its simplest form. [1]

Answer

(c) A student is chosen at random. Find the probability that the student is a boy who chose a sandwich. [1]

Answer

(d) What percentage of the students chose a hot meal? [2]

Answer %

- 2** The masses, in kg, of 24 parcels are:
1.2, 3.5, 4.8, 0.6, 2.0, 5.5, 6.3, 3.1, 2.2, 7.4, 4.1, 1.8, 3.9, 5.0, 2.6, 0.9, 4.4, 3.3, 6.8, 2.4, 3.7, 1.5, 4.6, 5.9

Mass, m kg	$0 < m \leq 2$	$2 < m \leq 4$	$4 < m \leq 6$	$6 < m \leq 8$
Frequency				

(a) Complete the frequency table. [3]

(b) Write down the modal class. [1]

Answer

(c) One parcel has a mass of 2.0 kg. Which class does it go in? Explain. [1]

(d) What fraction of the parcels have a mass of more than 4 kg? [1]

Answer

(e) Give one disadvantage of grouping the data. [1]

- 3** The table shows the monthly rainfall, in mm, in two cities.

Month	Jan	Feb	Mar	Apr	May	Jun
City A (mm)	12	18	35	60	142	210
City B (mm)	45	40	52	71	98	131

(a) In which month is the difference in rainfall between the two cities greatest? [1]

Answer

(b) Find the total rainfall in each city for the six months. [2]

Answer A mm B mm

(c) Make one comparison between the rainfall in the two cities. [1]

- 4 The table shows the number of goals scored in 30 football matches.

Goals	0	1	2	3	4	5
Frequency	6	9	8	4	2	1

- (a) Find the total number of goals scored.

[2]

Answer

- (b) In what fraction of the matches were more than 2 goals scored?

[1]

Answer

- (c) Ben says, 'The most common number of goals is 9.' Explain his mistake.

[1]

- (d) Is the number of goals discrete or continuous data?

[1]

Answer

- 5 For each of these, state whether the data are qualitative, discrete or continuous.

- (a) The heights of sunflowers.

[1]

Answer

- (b) The number of cars in a car park.

[1]

Answer

- (c) The flavours of ice cream sold in a shop.

[1]

Answer

Reflect: ☐ Red — not yet ☐ Amber — nearly ☐ Green — secure

Bar charts, pictograms and pie charts

I can... draw and interpret bar charts, pictograms and pie charts.

Syllabus ref: C9.2, C9.4, E9.4 · Core & Extended

A **bar chart** has bars of equal width with gaps between them. Label both axes and use an even scale. A **dual bar chart** puts two sets of data side by side.

A **pictogram** uses symbols. Always read the **key**; part of a symbol is part of the key value.

In a **pie chart** the whole circle (360°) is the total. Angle = frequency $\div$ total $\times 360^\circ$.

Key words

bar chart

pictogram

key

pie chart

sector

dual bar chart

Key points

- Pie chart angle = frequency $\div$ total $\times 360^\circ$
- Frequency from a pie chart = angle $\div 360^\circ \times$ total
- The angles in a pie chart add up to 360°

Worked example 1

60 students travel: bus 22, car 15, walk 18, cycle 5. Find the pie chart angles.

$$360^\circ \div 60 = 6^\circ \text{ per student}$$

$$\text{Bus } 22 \times 6 = \mathbf{132^\circ}, \text{ car } \mathbf{90^\circ}, \text{ walk } \mathbf{108^\circ}, \text{ cycle } \mathbf{30^\circ}$$

$$\text{Check: } 132 + 90 + 108 + 30 = 360^\circ$$

Worked example 2

In a pie chart, 18 people are shown by a 72° sector. How many people are there in total?

$$72^\circ \div 18 = 4^\circ \text{ per person}$$

$$\text{Total} = 360 \div 4 = \mathbf{90 \text{ people}}$$

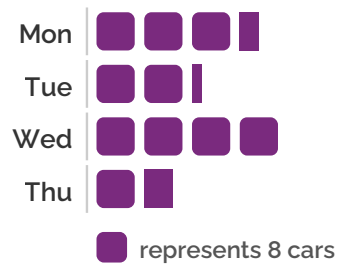
Common mistakes

- Forgetting to check that the angles add up to 360° .
- Ignoring the key in a pictogram.
- Comparing pie charts of different totals as if they showed numbers.

Exam tip: Draw pie charts with a sharp pencil and protractor: you are usually allowed $\pm 2^\circ$ for each sector.

Answer each question. Show your working where it helps.

- 1 The pictogram shows the number of cars sold by a garage.



- (a) How many cars were sold on Monday?

Answer

- (b) How many cars were sold on Tuesday?

Answer

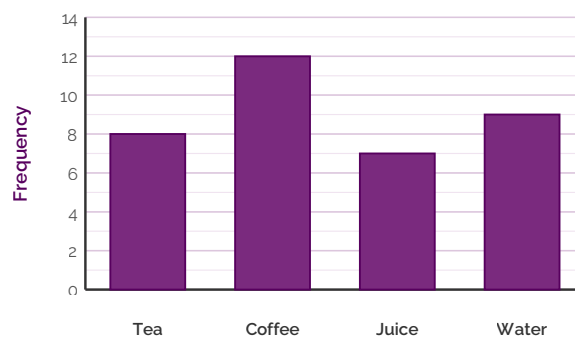
- (c) How many cars were sold altogether?

Answer

- (d) How many symbols would show 20 cars?

Answer

- 2 The bar chart shows the favourite drinks of a group of people.



- (a) How many people chose coffee?

Answer

- (b) How many people were asked?

Answer

- (c) What fraction of the people chose juice?

Answer

- (d) The data are shown in a pie chart. Find the angle for water.

Answer°

- 3 120 people chose their favourite sport. Work out the pie chart angle for each sport.

Sport	Football	Tennis	Swimming	Other
Frequency	45	30	25	20

(a) Football

Answer°

(b) Tennis

Answer°

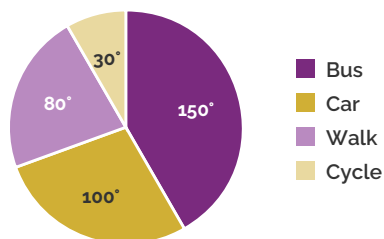
(c) Swimming

Answer°

(d) Other

Answer°

- 4 The pie chart shows how 72 students travel to school.



(a) How many students come by bus?

Answer

(b) How many students walk?

Answer

(c) What fraction of the students come by car?

Answer

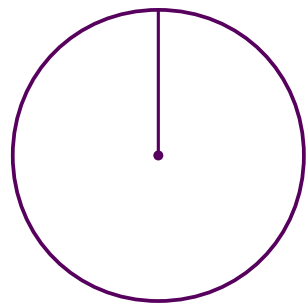
- 5 In a pie chart for 40 people, the sector for red is 81° . How many people chose red?

Answer

Reflect: ☐ Red — not yet ☐ Amber — nearly ☐ Green — secure

Answer all questions. Show your working clearly.

- 1 45 students were asked to name their favourite subject.



Subject	Maths	Science	English	Art	Other
Frequency	12	9	15	6	3

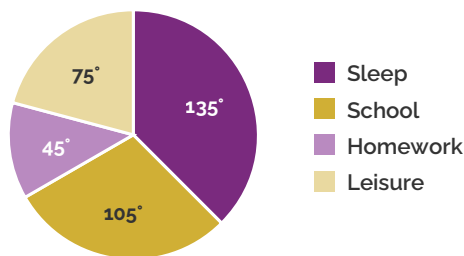
- (a) Work out the angle for each subject in a pie chart. [2]

- (b) Draw the pie chart in the circle. Label each sector. [2]

- (c) What percentage of the students chose English? [1]

Answer %

- 2 The pie chart shows how Sam spends a 24-hour day.



- (a) How many hours does Sam sleep?

[2]

Answer hours

- (b) How many hours does Sam spend on leisure?

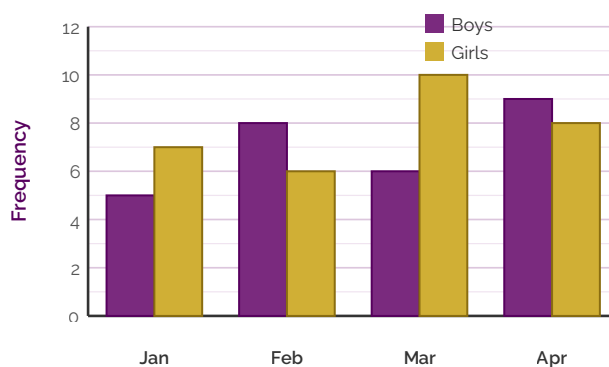
[1]

Answer hours

- (c) Sam says he spends more than a quarter of the day at school. Is he correct? Show how you decide.

[2]

- 3 The dual bar chart shows the number of books read by a group of boys and a group of girls.



- (a) In which month did the girls read the most books?

[1]

Answer

- (b) Find the total number of books read by the boys.

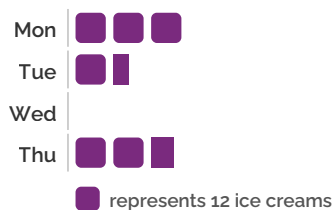
[1]

Answer

- (c) Compare the numbers of books read by the boys and the girls.

[2]

- 4 The pictogram shows the number of ice creams sold. The row for Wednesday is not finished.



(a) How many ice creams were sold on Monday? [1]

Answer

(b) How many ice creams were sold on Thursday? [1]

Answer

(c) 42 ice creams were sold on Wednesday. Complete the pictogram. [2]

(d) Find the total number of ice creams sold. [1]

Answer

- 5 School A has 600 students and School B has 300 students. In pie charts showing how students travel, the angle for 'walk' is 90° for School A and 120° for School B. Tom says more students walk at School B.

(a) Find how many students walk at each school. [2]

Answer A B

(b) Explain why Tom is wrong. [1]

- 6 The table shows the numbers of goals scored by a team in 20 matches. Draw a bar chart on the grid to show this information.



Goals	0	1	2	3	4
Frequency	3	6	5	4	2

[3]

Reflect: ☐ Red — not yet ☐ Amber — nearly ☐ Green — secure

Averages, range and stem-and-leaf

I can... find the mean, median, mode and range from lists, tables and stem-and-leaf diagrams.

Syllabus ref: C9.3, E9.3, C9.4, E9.4 · Core & Extended

Mean = total $\div$ number of values. **Median** = middle value when the data are in order (position $(n + 1)/2$). **Mode** = most common value. **Range** = largest – smallest.

For a frequency table: mean = $\Sigma fx \div \Sigma f$ (multiply each value by its frequency).

A **stem-and-leaf diagram** shows ordered data and must have a **key**. The range measures **spread**.

Key words

mean

median

mode

range

stem-and-leaf

key

Key points

- Mean from a table = $\Sigma fx \div \Sigma f$
- Median position = $(n + 1)/2$
- Compare data sets using an average and a measure of spread

Worked example 1

Stem-and-leaf (key 1 | 2 = 12): 1 | 2 5 7; 2 | 0 3 3 8; 3 | 1 4. Find the median, mode and range.

$n = 9$, so the median is the 5th value: **23**

Mode = **23**; range = $34 - 12 = 22$

Worked example 2

Goals 0, 1, 2, 3 have frequencies 4, 7, 6, 3. Find the mean.

$\Sigma f = 20$; $\Sigma fx = 0 + 7 + 12 + 9 = 28$

Mean = $28 \div 20 = 1.4$ goals

Common mistakes

- Finding the median without putting the data in order first.
- In a frequency table, dividing by the number of rows instead of Σf .

Exam tip: When comparing two data sets, write one sentence about the average and one about the spread, in context.

Answer each question. Show your working where it helps.

- 1** For the numbers 7, 3, 9, 4, 7, 10, 2, find:

(a) the mean

Answer

(b) the median

Answer

(c) the mode

Answer

(d) the range

Answer

- 2** For the numbers 12, 15, 9, 20, 14, 20, find:

(a) the median

Answer

(b) the mean

Answer

- 3** The mean of five numbers is 8. Four of them are 6, 9, 10 and 5. Find the fifth number.

Answer

- 4** The mean of four numbers is 12. The number 22 is added. Find the new mean.

Answer

- 5** The stem-and-leaf diagram shows some journey times.

3		1 4 8
4		0 2 2 5 7
5		1 3 6
6		0

Key: 3 | 1 means 31 minutes

(a) Find the median.

Answer minutes

(b) Write down the mode.

Answer minutes

(c) Find the range.

Answer minutes

(d) How many journeys took more than 50 minutes?

6 The table shows the number of siblings of 25 students.

Siblings	0	1	2	3	4
Frequency	5	9	7	3	1

(a) Calculate the mean.

Answer

(b) Find the median.

Answer

(c) Write down the mode.

Answer

7 Which average is most useful for the shoe sizes sold by a shop?

Answer

8 Which average is affected most by one very large value?

Answer

Reflect: ☐ Red — not yet ☐ Amber — nearly ☐ Green — secure

Answer all questions. Show your working clearly.

- 1 The back-to-back stem-and-leaf diagram shows the test marks of two classes.

Class A		Class B
8	1	
7 5 2	2	4 9
8 4 1 1	3	0 3 5 6 6 9
5 2 0	4	1 3 4

Key: 2 | 4 means 24 marks for Class B; 7 | 2 means 27 marks for Class A

- (a) Find the median and the range for Class B. [2]

Answer median range

- (b) For Class A, the median is 31 and the range is 27. Make two comparisons between the classes. [2]

- (c) Find the mode for Class A. [1]

Answer

- 2 The heights, in cm, of 15 plants are:
23, 35, 41, 28, 37, 52, 44, 31, 26, 38, 47, 35, 29, 40, 33

- (a) Draw an ordered stem-and-leaf diagram for these heights. Include a key. [3]

- (b) Find the median height. [1]

Answer cm

- (c) Find the range. [1]

Answer cm

- 3 The table shows the number of people in each of 50 cars.

Number of people	1	2	3	4	5
Frequency	23	15	7	4	1

- (a) Calculate the mean number of people per car.

[3]

Answer

- (b) Find the median.

[1]

Answer

- (c) Write down the mode.

[1]

Answer

- 4 Solve these problems about the mean.

- (a) The mean age of 6 people is 24. One person leaves and the mean age of the other 5 is 25. Find the age of the person who left.

[3]

Answer years

- (b) The mean mass of 12 boys is 52 kg. The mean mass of 18 girls is 47 kg. Find the mean mass of all 30 students.

[3]

Answer kg

- 5 The salaries, in thousands of dollars, of the 7 people in a small company are:
21, 22, 22, 24, 25, 26, 95

(a) Calculate the mean salary.

[2]

Answer

(b) Find the median salary.

[1]

Answer

(c) Which average better represents a typical salary? Give a reason.

[2]

Reflect: ☐ Red — not yet ☐ Amber — nearly ☐ Green — secure

Grouped data and the estimated mean

I can... estimate the mean of grouped data and find the modal class and the class containing the median.

Syllabus ref: E9.3 · Extended

In grouped data we do not know the exact values, so we use the **mid-point** of each class to **estimate** the mean.

Estimated mean = $\Sigma fx \div \Sigma f$, where x is the mid-point. The **modal class** is the class with the highest frequency.

The median is the $((n + 1)/2)$ th value: add up the frequencies to find which class it is in.

Key words

grouped data

mid-point

estimate

modal class

Σfx

Key points

- Mid-point = (lower bound + upper bound) $\div$ 2
- Estimated mean = $\Sigma fx \div \Sigma f$
- For discrete classes such as 11–20, the mid-point is 15.5

Worked example

Times: $0 < t \leq 10$: 5, $10 < t \leq 20$: 12, $20 < t \leq 30$: 15, $30 < t \leq 40$: 8. Estimate the mean.

Mid-points: 5, 15, 25, 35

$\Sigma fx = 25 + 180 + 375 + 280 = 860$; $\Sigma f = 40$

Estimated mean = $860 \div 40 = \mathbf{21.5 \text{ minutes}}$

Modal class: $20 < t \leq 30$; median (20.5th value) is in $20 < t \leq 30$

Common mistakes

- Dividing by the number of classes instead of Σf .
- Using the class width or upper bound instead of the mid-point.
- Giving the frequency as the modal class.

Exam tip: Add two columns to the table — mid-point and $f \times x$ — so your working is clear and earns method marks.

Answer each question. Show your working where it helps.

- 1 The table shows the heights of 40 students.

Height h cm	$150 < h \leq 155$	$155 < h \leq 160$	$160 < h \leq 165$	$165 < h \leq 170$	$170 < h \leq 175$
Frequency	4	9	14	8	5

- (a) Write down the mid-point of the class $155 < h \leq 160$.

Answer

- (b) Write down the modal class.

Answer

- (c) How many students are taller than 165 cm?

Answer

- (d) Which class contains the median?

Answer

- (e) Find Σfx , using the mid-points.

Answer

- (f) Calculate an estimate of the mean height.

Answer cm

- (g) Why is your answer to (f) only an estimate?

- 2 The table shows the marks of 30 students.

Mark	1–10	11–20	21–30	31–40
Frequency	3	7	12	8

- (a) Write down the mid-point of the class 21–30.

Answer

- (b) Calculate an estimate of the mean mark.

Answer

- (c) Write down the modal class.

Answer

- 3 The class $0 < t \leq 5$ has frequency 8. What does this class add to Σfx ?

Answer

Reflect: ☐ Red — not yet ☐ Amber — nearly ☐ Green — secure

Answer all questions. Show your working clearly.

- 1 The table shows the time, t minutes, that 60 students spent on homework one evening.

Time (t min)	$0 < t \leq 30$	$30 < t \leq 60$	$60 < t \leq 90$	$90 < t \leq 120$	$120 < t \leq 150$
Frequency	8	17	21	10	4

- (a) Write down the modal class. [1]

Answer

- (b) Write down the class that contains the median. [1]

Answer

- (c) Calculate an estimate of the mean time. [4]

Answer min

- (d) Explain why your answer to (c) is an estimate. [1]

- (e) Estimate the number of students who spent more than 100 minutes on homework. [2]

Answer

- 2 The table shows the masses of some packages. An estimate of the mean mass is 60 kg.

Mass (m kg)	$40 < m \leq 50$	$50 < m \leq 60$	$60 < m \leq 70$	$70 < m \leq 80$
Frequency	6	x	14	5

- (a) Find the value of x . [4]

Answer $x = \dots\dots$

- (b) What percentage of the packages have a mass of more than 60 kg? [2]

Answer $\dots\dots\%$

- (c) Write down the class that contains the median. [1]

Answer $\dots\dots$

- 3 The table shows the times of 25 runners in Group B. For Group A, an estimate of the mean time is 24.5 minutes.

Time (t min)	$15 < t \leq 20$	$20 < t \leq 25$	$25 < t \leq 30$	$30 < t \leq 35$
Frequency	4	10	9	2

- (a) Calculate an estimate of the mean time for Group B. [3]

Answer $\dots\dots$ min

- (b) Which group was faster on average? Give a reason. [1]

- (c) Explain why you cannot find the exact range of the times in Group B. [1]

- 4 The table shows the number of emails received by 30 people in one day.

Emails	0–4	5–9	10–14	15–19
Frequency	7	12	9	2

- (a) Calculate an estimate of the mean number of emails. [3]

Answer

- (b) Write down the modal class. [1]

Answer

- (c) Write down the class that contains the median. [1]

Answer

Reflect: ☐ Red — not yet ☐ Amber — nearly ☐ Green — secure

Scatter diagrams and correlation

I can... draw and interpret scatter diagrams, describe correlation and use a line of best fit.

Syllabus ref: C9.5, E9.5 · Core & Extended

A **scatter diagram** plots pairs of data to show whether two variables are related. **Positive correlation**: both increase together. **Negative correlation**: one increases as the other decreases. Points close to a line show **strong** correlation.

A **line of best fit** is a ruled straight line that follows the trend, with roughly equal numbers of points on each side. It should pass through the **mean point** (mean of x , mean of y).

Estimates **inside** the data range are reliable; estimates **outside** it (extrapolation) are not.

Key words

scatter diagram

correlation

positive

negative

line of best fit

outlier

Key points

- Correlation: positive, negative or none; strong or weak
- Mean point = $(\bar{x}, \bar{y})$ lies on the line of best fit
- Correlation does not prove that one thing causes another

Worked example

The points (2, 5), (4, 9), (6, 10), (8, 16) are plotted. Find the mean point and describe the correlation.

Mean of $x = 20 \div 4 = 5$; mean of $y = 40 \div 4 = 10$

Mean point = (5, 10)

As x increases, y increases: **positive correlation**

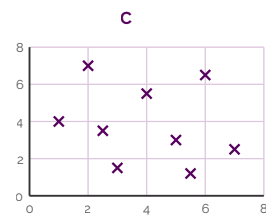
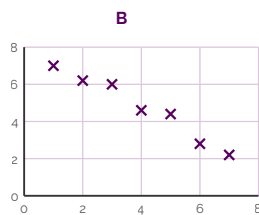
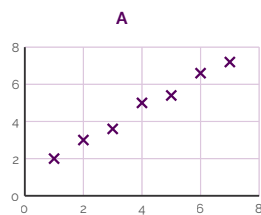
Common mistakes

- Joining the points dot-to-dot instead of drawing one straight line.
- Forcing the line through the origin.
- Using the line to predict far outside the data.

Exam tip: Use a ruler and a sharp pencil; mark the mean point with a small cross and draw your line through it.

Answer each question.

- 1 Describe the correlation shown in each scatter diagram.



(a) Diagram A

Answer

(b) Diagram B

Answer

(c) Diagram C

Answer

- 2 What correlation would you expect between temperature and the number of ice creams sold?

Answer

- 3 What correlation would you expect between the age of a car and its value?

Answer

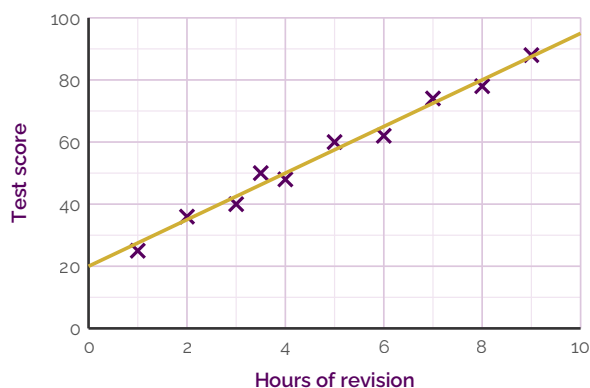
- 4 What correlation would you expect between shoe size and maths test score?

Answer

- 5 Find the mean point of (2, 5), (4, 9), (6, 10) and (8, 16).

Answer

- 6 The scatter diagram shows the hours of revision and test scores of 10 students, with a line of best fit.



- (a) Estimate the score of a student who revised for 5 hours.

Answer

- (b) Estimate the hours of revision for a score of 65.

Answer hours

- (c) Why would an estimate for 15 hours of revision be unreliable?

- 7 What is an outlier on a scatter diagram?

Answer

- 8 Does a strong correlation prove that one variable causes the other?

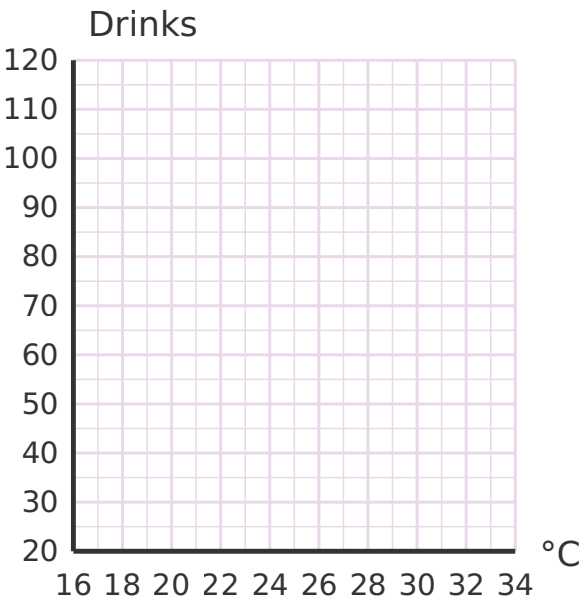
Answer

Reflect: ☐ Red — not yet ☐ Amber — nearly ☐ Green — secure

Answer all questions. Show your working clearly.

- 1 The table shows the midday temperature and the number of cold drinks sold at a café on 10 days.

Temperature (°C)	18	20	21	23	24	26	28	29	31	33
Drinks sold	38	45	52	58	66	70	82	85	96	104



- (a) On the grid, plot the points to draw a scatter diagram.

[3]
- (b) Describe the correlation.

[1]

Answer
- (c) Calculate the mean temperature and the mean number of drinks sold.

[2]

Answer °C drinks
- (d) Draw a line of best fit through the mean point.

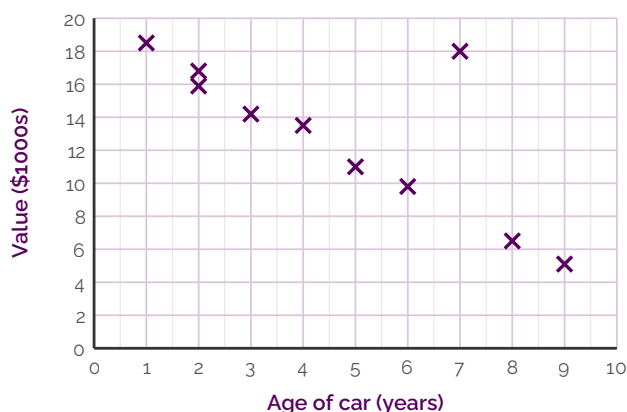
[1]
- (e) Use your line to estimate the number of drinks sold when the temperature is 27 °C.

[1]

Answer
- (f) Explain why it would not be sensible to use your line to estimate the sales at 40 °C.

[1]

- 2 The scatter diagram shows the ages and values of 10 cars of the same model.



- (a) Describe the correlation. [1]

Answer

- (b) One point is an outlier. Write down its coordinates and suggest a reason for it. [2]

- (c) Ignoring the outlier, draw a line of best fit on the diagram. [1]

- (d) Use your line to estimate the value of a car that is $5\frac{1}{2}$ years old. [1]

Answer \$

- (e) Should your line be used to estimate the value of a 20-year-old car? Explain. [1]

- 3 Ahmed measures the arm span, a cm, and the height, h cm, of some students. The equation of his line of best fit is $h = 0.95a + 9$.

- (a) What type of correlation does this show? [1]

Answer

- (b) Estimate the height of a student with an arm span of 160 cm. [2]

Answer cm

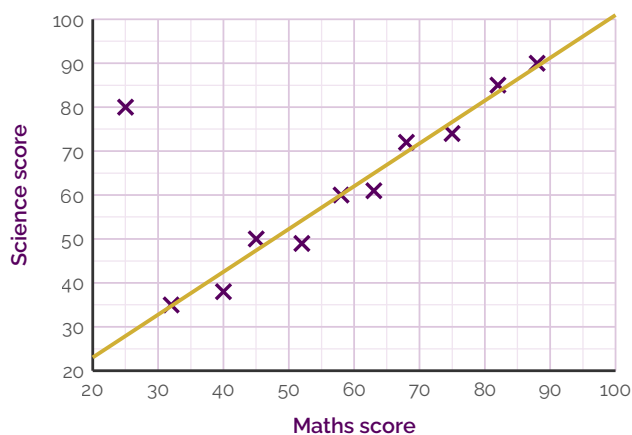
- (c) The mean arm span is 158 cm. Find the mean height. [1]

Answer cm

- (d) Estimate the arm span of a student who is 170 cm tall. [2]

Answer cm

- 4 The scatter diagram shows the maths and science test scores of 11 students, with a line of best fit.



- (a) Describe the relationship between the maths and science scores. [1]

- (b) One student scored 70 in maths but missed the science test. Estimate their science score. [2]

Answer

- (c) Describe the point (25, 80). [1]

- (d) Lina says, 'Good maths scores cause good science scores.' Comment on her statement. [1]

Reflect: ☐ Red — not yet ☐ Amber — nearly ☐ Green — secure

Cumulative frequency, quartiles and box plots

I can... draw cumulative frequency diagrams and find medians, quartiles, percentiles and box plots.

Syllabus ref: E9.3, E9.6 · Extended

Quartiles split ordered data into four equal parts. For a list of n values: lower quartile = $((n + 1)/4)$ th value, upper quartile = $(3(n + 1)/4)$ th value. **IQR** = UQ – LQ measures the spread of the middle half.

Cumulative frequency is a running total. Plot each value at the **upper bound** of its class and join with a smooth curve. Read the median at $n/2$, LQ at $n/4$, UQ at $3n/4$ and the p th percentile at $p\%$ of n .

A **box-and-whisker plot** shows the minimum, LQ, median, UQ and maximum.

Key words

cumulative frequency

quartile

interquartile range

percentile

box-and-whisker plot

Key points

- IQR = upper quartile – lower quartile
- Plot cumulative frequency at upper class bounds
- Compare distributions using the median and the IQR

Worked example

Find the quartiles and IQR of 3, 5, 6, 8, 9, 11, 12, 14, 15, 18, 20.

$n = 11$: median = 6th = **11**

LQ = 3rd value = **6**; UQ = 9th value = **15**

IQR = $15 - 6 = 9$

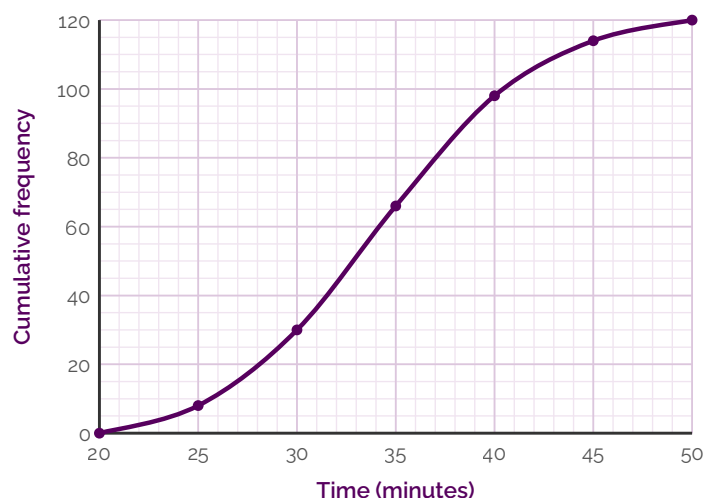
Common mistakes

- Plotting at mid-points instead of upper bounds.
- Giving the cumulative frequency (e.g. 40) as the median instead of reading across to the horizontal axis.

Exam tip: Draw your reading lines on the graph — they show the examiner your method.

Answer each question. Draw lines on the diagram to show your readings.

- 1 The cumulative frequency diagram shows the times of 120 runners.



- (a) Estimate the median.

Answer min

- (b) Estimate the lower quartile.

Answer min

- (c) Estimate the upper quartile.

Answer min

- (d) Estimate the interquartile range.

Answer min

- (e) Estimate how many runners took 33 minutes or less.

Answer

- (f) Estimate how many runners took more than 42 minutes.

Answer

- (g) Estimate the 90th percentile.

Answer min

- 2 Find the median of: 4, 7, 8, 10, 12, 13, 15, 16, 18, 19, 21, 22, 25, 27, 30

Answer

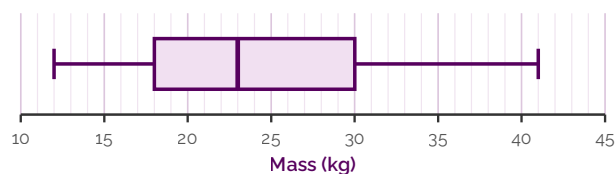
- 3 For the data in question 2, find the lower quartile.

Answer

- 4 For the data in question 2, find the interquartile range.

Answer

- 5 The box-and-whisker plot shows the masses of some parcels.



- (a) Write down the median.

Answer kg

- (b) Find the interquartile range.

Answer kg

- (c) Find the range.

Answer kg

- (d) What fraction of the parcels have a mass of more than 30 kg?

Answer

Reflect: ☐ Red — not yet ☐ Amber — nearly ☐ Green — secure

Answer all questions. Show your working clearly.

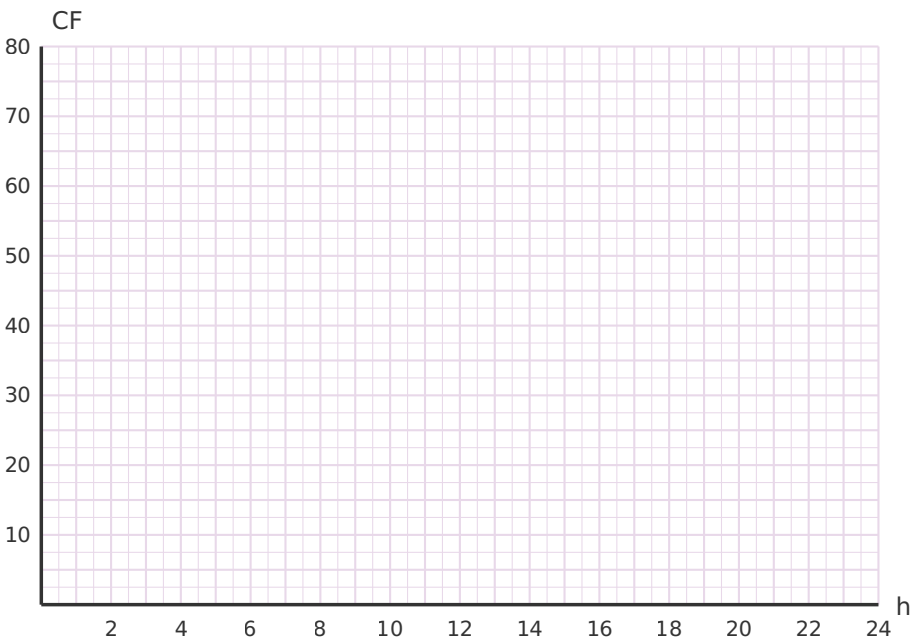
1 The table shows the battery life, h hours, of 80 phones.

Battery life (h hours)	$0 < h \leq 4$	$4 < h \leq 8$	$8 < h \leq 12$	$12 < h \leq 16$	$16 < h \leq 20$	$20 < h \leq 24$
Frequency	3	9	20	26	15	7

(a) Complete the cumulative frequency table. [1]

Battery life (h hours)	$h \leq 4$	$h \leq 8$	$h \leq 12$	$h \leq 16$	$h \leq 20$	$h \leq 24$
Cumulative frequency						

(b) On the grid, draw a cumulative frequency diagram. [3]



(c) Use your diagram to estimate the median. [1]

Answer hours

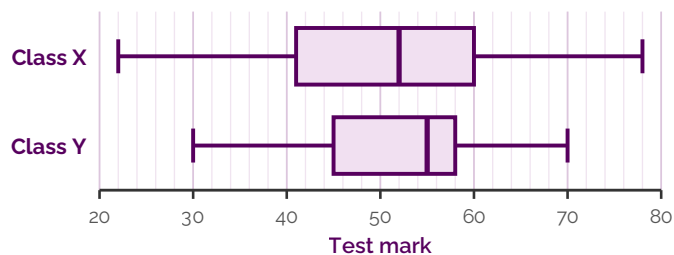
(d) Use your diagram to estimate the interquartile range. [2]

Answer hours

(e) Estimate the number of phones with a battery life of more than 18 hours. [2]

Answer

- 2 The box-and-whisker plots show the test marks of two classes.



- (a) Write down the median mark for Class X. [1]

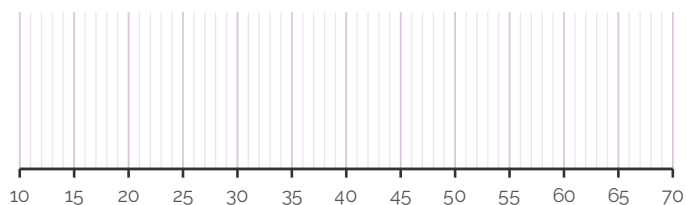
Answer

- (b) Find the interquartile range for each class. [2]

Answer X Y

- (c) Make two comparisons between the marks of the two classes. [2]

- 3 For a set of data, the smallest value is 15, the largest value is 62, the median is 38, the lower quartile is 27 and the interquartile range is 19.

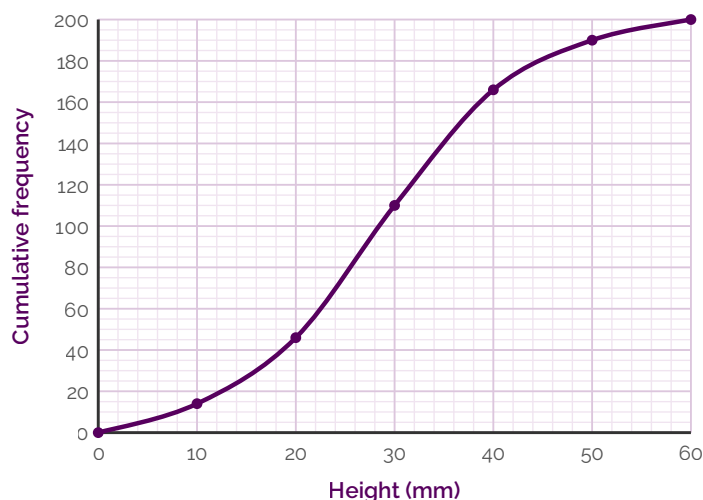


- (a) Find the upper quartile. [1]

Answer

- (b) Draw a box-and-whisker plot for the data. [2]

- 4 The cumulative frequency diagram shows the heights of 200 seedlings.



- (a) Estimate the median height. [1]

Answer mm

- (b) Estimate the 80th percentile. [2]

Answer mm

- (c) Estimate the number of seedlings taller than 35 mm. [2]

Answer

- 5 The waiting times, in minutes, of 11 patients are:

3, 5, 6, 8, 9, 11, 12, 14, 16, 19, 25

- (a) Find the median, the lower quartile and the upper quartile. [3]

Answer median LQ UQ

- (b) Find the interquartile range. [1]

Answer

Reflect: ☐ Red — not yet ☐ Amber — nearly ☐ Green — secure

Histograms and frequency density

I can... draw and interpret histograms with unequal class widths using frequency density.

Syllabus ref: E9.7 · Extended

A **histogram** is used for grouped continuous data. There are no gaps between the bars, and classes can have **different widths**.

The vertical axis shows **frequency density**, so the **area** of each bar shows its frequency.

Key words

histogram

frequency density

class width

area

unequal widths

Key points

- Frequency density = frequency $\div$ class width
- Frequency = frequency density $\times$ class width
- The tallest bar is not always the class with the greatest frequency

Worked example 1

Find the frequency densities: $0 < t \leq 10$: 8; $10 < t \leq 15$: 12; $15 < t \leq 30$: 18.

$$8 \div 10 = \mathbf{0.8}$$

$$12 \div 5 = \mathbf{2.4}$$

$$18 \div 15 = \mathbf{1.2}$$

Worked example 2

A bar for $20 < x \leq 35$ has height 1.6. How many values does it show? Estimate how many are between 20 and 25.

$$\text{Frequency} = 1.6 \times 15 = \mathbf{24}$$

$$20 \text{ to } 25 \text{ is } 5/15 \text{ of the class: } 1.6 \times 5 = \mathbf{8}$$

Common mistakes

- Plotting frequency instead of frequency density.
- Using the wrong class width (e.g. 10 for $20 < x \leq 25$).

Exam tip: If the axis has no scale, use one bar with a known frequency to find what one unit of area represents.

Answer each question. Show your working where it helps.

1

A class $10 < x \leq 20$ has frequency 15. Find the frequency density.

2

A class $20 < x \leq 25$ has frequency 18. Find the frequency density.

3

A class $0 < t \leq 40$ has frequency 12. Find the frequency density.

4

A class $50 < m \leq 80$ has frequency density 0.8. Find the frequency.

5

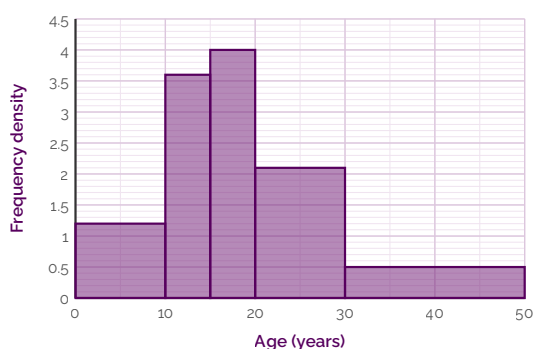
A class $5 < h \leq 6.5$ has frequency 9. Find the frequency density.

6

What does the area of a bar in a histogram represent?

The histogram shows the ages of people at a cinema. How many people were aged 0 to 10?

7



8

Using the histogram in question 7, how many were aged 10 to 15?

Answer

9

Using the histogram in question 7, how many were aged 20 to 30?

10

Using the histogram in question 7, how many were aged 30 to 50?

11

Using the histogram in question 7, how many people were there in total?

12

Using the histogram in question 7, estimate the number aged between 25 and 40.

13

Which class in question 7 has the highest frequency?

14

Why is the tallest bar in question 7 not the class with the highest frequency?

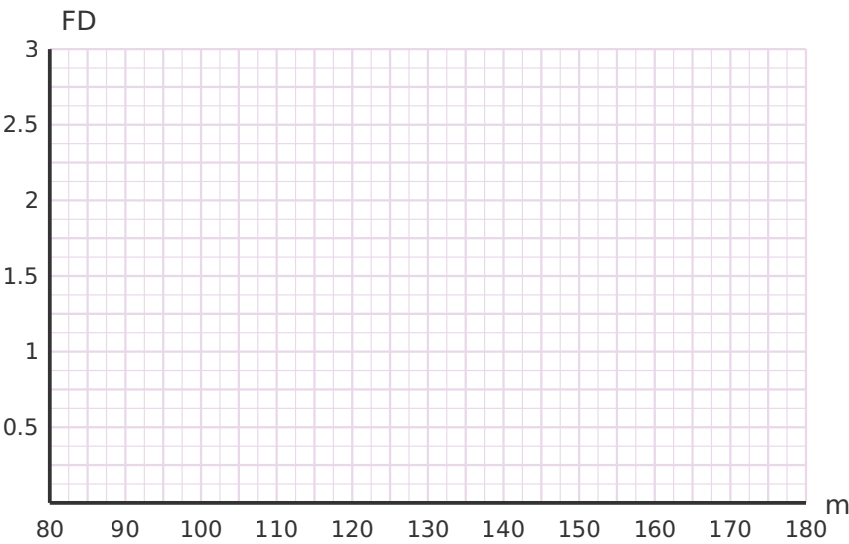
Reflect: ☐ Red — not yet ☐ Amber — nearly ☐ Green — secure

Answer all questions. Show your working clearly.

1 The table shows the masses of 100 apples.

Mass (m g)	$80 < m \leq 100$	$100 < m \leq 110$	$110 < m \leq 120$	$120 < m \leq 140$	$140 < m \leq 180$
Frequency	10	18	26	32	14
Frequency density					

- (a) Complete the frequency density row of the table. [2]
- (b) On the grid, draw a histogram to show this information. [3]



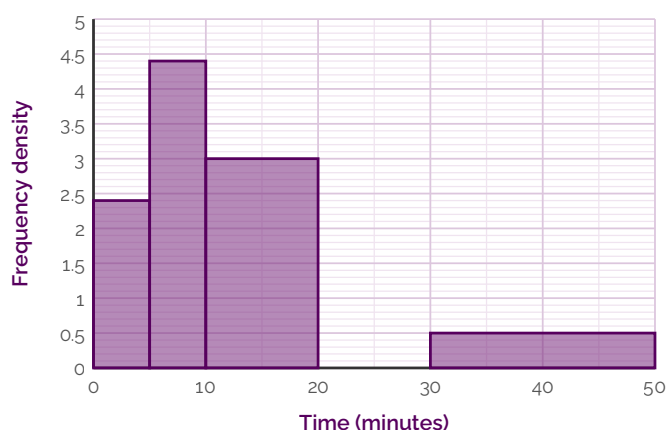
- (c) Write down the modal class. [1]

Answer

- (d) Estimate the number of apples with a mass of more than 130 g. [2]

Answer

- 2 The histogram and the table show the times some customers spent in a shop. The histogram is not complete.



Time (t min)	$0 < t \leq 5$	$5 < t \leq 10$	$10 < t \leq 20$	$20 < t \leq 30$	$30 < t \leq 50$
Frequency	12			16	

- (a) Use the histogram to complete the table. [2]
 (b) Use the table to complete the histogram. [1]
 (c) Find the total number of customers. [1]

Answer

- (d) Estimate the number of customers who spent more than 25 minutes in the shop. [2]

Answer

- 3 Use the completed table from question 2.

- (a) Calculate an estimate of the mean time. [4]

Answer min

- (b) Estimate the median time. [3]

Answer min

- 4 In a histogram, the bar for the class $10 < x \leq 20$ is 3 cm high and represents 45 people. The bar for $20 < x \leq 25$ is 4 cm high.

How many people does the bar for $20 < x \leq 25$ represent?

[3]

Answer

- 5 A histogram has classes of different widths.

Explain why the vertical axis must show frequency density and not frequency.

[1]

Reflect: ☐ Red — not yet ☐ Amber — nearly ☐ Green — secure



CHECK

End-of-book check

End-of-book check

Name _____ Date _____ Marks _____ / 19

Answer all questions. Show your working. Compare your score with your starting point check.

- 1 A bag has 4 red, 7 blue and 9 white beads. One is taken at random. Find $P(\text{not blue})$. [1]

Answer

- 2 A spinner lands on A, B, C or D. $P(A) = 0.15$, $P(B) = 0.4$ and $P(C) = P(D)$. Find $P(D)$. [2]

Answer

- 3 A dice is rolled 150 times and shows a six 36 times. (a) Find the relative frequency of a six.
(b) Estimate how many sixes there would be in 600 rolls. [2]

Answer (a) (b)

- 4 Two fair dice are rolled and the scores added. Find $P(\text{total} = 8)$. [1]

Answer

- 5 A biased coin has $P(\text{head}) = 0.7$. It is thrown twice. Find $P(\text{two heads})$. [1]

Answer

- 6 **(Extended)** A bag has 5 red and 3 green balls. Two are taken without replacement. Find $P(\text{both green})$. [2]

Answer

- 7 **(Extended)** 24 students in a club play chess. 9 of the chess players are girls. A chess player is chosen at random. Find $P(\text{girl} \mid \text{plays chess})$. [1]

Answer

- 8 Find the mean of this data. [2]

Value	1	2	3	4
Frequency	3	5	8	4

Answer

- 9 **(Extended)** Calculate an estimate of the mean. [3]

x	$0 < x \leq 10$	$10 < x \leq 20$	$20 < x \leq 30$
Frequency	6	10	4

Answer

- 10 **(Extended)** Find the interquartile range of 2, 4, 5, 7, 9, 12, 15. [2]

Answer

- 11 **(Extended)** A cumulative frequency diagram shows 60 values. At what cumulative frequency do you read the median? [1]

Answer

- 12 **(Extended)** A histogram bar for $15 < x \leq 30$ has frequency density 1.8. Find the frequency. [1]

Answer



Practice paper

Name _____ Date _____ Marks _____ / 60

Time: 1 hour 15 minutes
Calculator allowed
Total: 60 marks
Questions marked (Extended) are for Extended candidates.

Answer all questions. Show all your working. Give probabilities as fractions or decimals.

- 1 A sports club has 80 members. Each member chooses swimming or tennis.

	Swimming	Tennis	Total
Adults		12	35
Children	27		
Total		30	80

(a) Complete the two-way table. [3]

(b) A member is chosen at random. Find the probability that the member is an adult who chose swimming. [1]

Answer

(c) (Extended) A member who chose tennis is chosen at random. Find the probability that this member is a child. [2]

Answer

- 2 The stem-and-leaf diagram shows the scores of 19 players in a game.

2	1 4 5 8
3	0 1 1 3 5 6 8
4	0 1 1 1 4 7
5	0 3

Key: 2 | 1 means a score of 21

(a) Write down the mode.

[1]

Answer

(b) Find the median.

[1]

Answer

(c) Find the range.

[1]

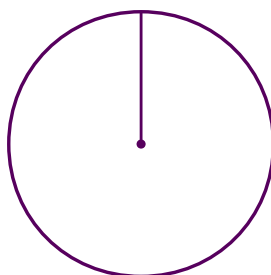
Answer

(d) (Extended) Find the interquartile range.

[2]

Answer

- 3 72 people were asked to name their favourite fruit.



Fruit	Apple	Banana	Mango	Other
Frequency	20	14	26	12

(a) Work out the pie chart angle for each fruit.

[2]

(b) Draw and label the pie chart.

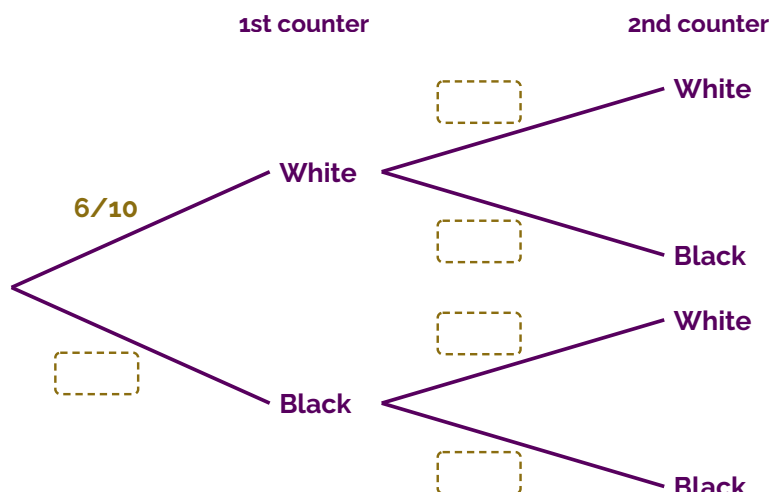
[2]

(c) What fraction of the people chose mango? Give your answer in its simplest form.

[1]

Answer

- 4 **(Extended)** A box contains 6 white counters and 4 black counters. Two counters are taken at random without replacement.



(a) Complete the tree diagram. [2]

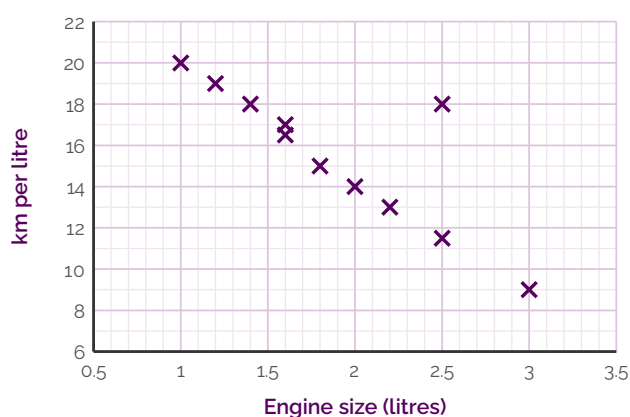
(b) Find the probability that both counters are the same colour. [3]

Answer

(c) Find the probability that at least one counter is black. [2]

Answer

- 5 The scatter diagram shows the engine size and fuel economy of 11 cars.



(a) Describe the correlation. [1]

Answer

(b) One car does not fit the pattern. Circle it on the diagram. [1]

(c) Draw a line of best fit, ignoring this car. [1]

(d) Use your line to estimate the fuel economy of a car with a 2.4 litre engine. [1]

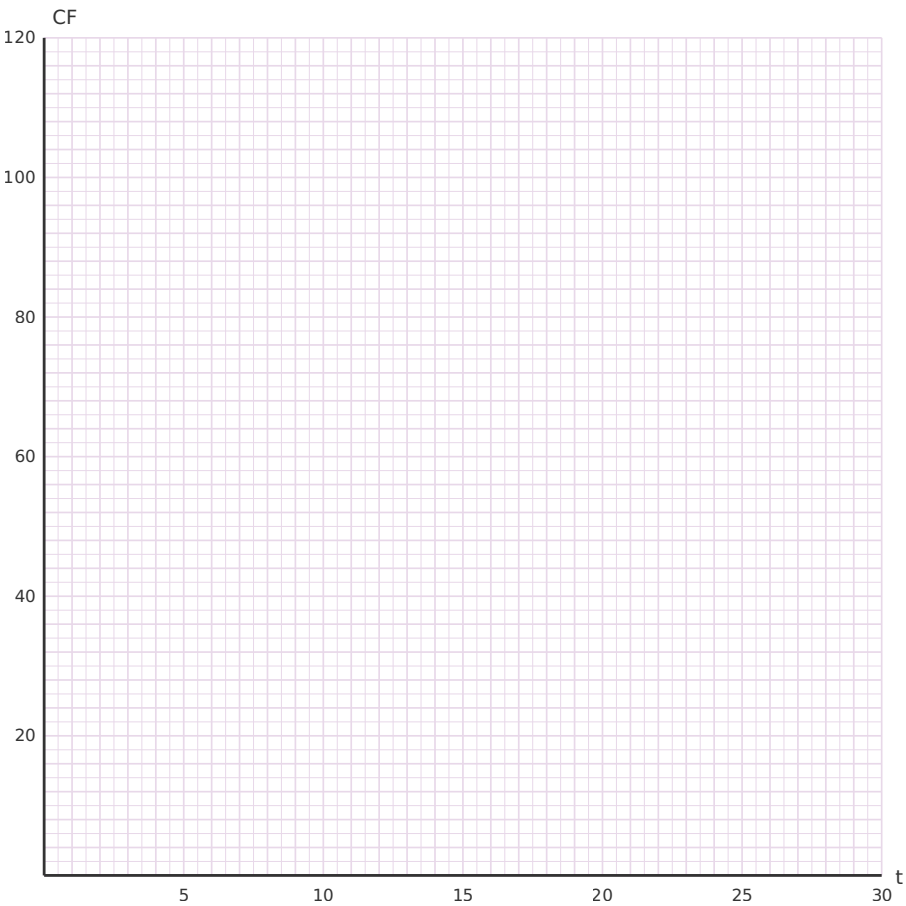
6 (Extended) The table shows the times, t minutes, that 120 students took to solve a puzzle.

Time (t min)	$0 < t \leq 5$	$5 < t \leq 10$	$10 < t \leq 15$	$15 < t \leq 20$	$20 < t \leq 25$	$25 < t \leq 30$
Frequency	6	18	37	32	19	8

(a) Complete the cumulative frequency table. [1]

Time (t min)	$t \leq 5$	$t \leq 10$	$t \leq 15$	$t \leq 20$	$t \leq 25$	$t \leq 30$
Cumulative frequency						

(b) Draw the cumulative frequency diagram. [3]



(c) Estimate the median time. [1]

Answer min

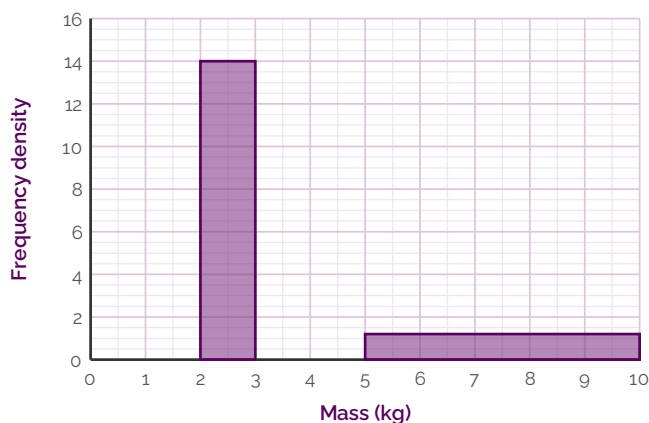
(d) Estimate the interquartile range. [2]

Answer min

(e) Estimate the number of students who took more than 22 minutes. [2]

Answer

- 7 **(Extended)** The table and the histogram show the masses of some parcels. Both are incomplete.



Mass (w kg)	$0 < w \leq 2$	$2 < w \leq 3$	$3 < w \leq 5$	$5 < w \leq 10$
Frequency	8		18	

- (a) Complete the table. [2]
 (b) Complete the histogram. [2]
 (c) Estimate the number of parcels heavier than 4 kg. [2]

Answer

- 8 **(Extended)** The table shows the ages of 40 people at a concert.

Age (a years)	$10 < a \leq 20$	$20 < a \leq 30$	$30 < a \leq 40$	$40 < a \leq 60$
Frequency	7	15	10	8

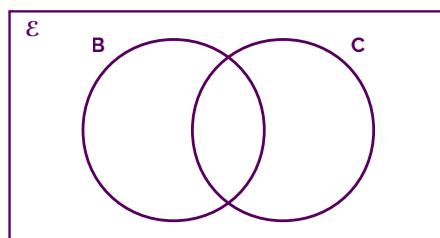
- (a) Calculate an estimate of the mean age. [4]

Answer years

- (b) Write down the modal class. [1]

Answer

- 9 In a class of 40 students, 22 study Biology (B), 18 study Chemistry (C) and 6 study neither.



(a) Complete the Venn diagram. [2]

(b) A student is chosen at random. Find $P((B \cup C)')$. [1]

Answer

(c) (Extended) A Biology student is chosen at random. Find the probability that they also study Chemistry. [2]

Answer

- 10 A dice is rolled 300 times. The results are shown in the table.

Score	1	2	3	4	5	6
Frequency	42	51	48	39	45	75

(a) Find the relative frequency of a 6. [1]

Answer

(b) The dice is rolled 1200 more times. Estimate how many times it shows a 6. [2]

Answer

(c) Is the dice fair? Give a reason. [1]

- 11 Two fair dice are rolled. The score is the **larger** of the two numbers (if they are equal, the score is that number).

(a) Find the probability that the score is 6. [2]

Answer

(b) Find the probability that the score is 3 or less. [2]

Answer



FOR TEACHERS & PARENTS

Answers & mark schemes

Mark schemes list the points that earn marks. Accept equivalent answers and follow-through where working is correct. M = method mark, A = accuracy mark, B = independent mark.

Starting point check

- 1 $2/8 = 1/4$ B1 [1]
- 2 0.55 B1 [1]
- 3 $27/60 = 0.45$ B1 [1]
- 4 30 B1 [1]
- 5 $1/4$ B1 [1]
- 6 $12 + 9 - 4 = 17$ M1; 8 A1 [2]
- 7 7 B1 [1]
- 8 7 B1 [1]
- 9 60° B1 [1]
- 10 continuous B1 [1]
- 11 negative B1 [1]
- 12 2.5 B1 [1]

Unit 1 · Sheet A — The probability scale and single events

- 1 $1/6$
- 2 $3/6 = 1/2$
- 3 $2/6 = 1/3$
- 4 0
- 5 $3/10$
- 6 $2/10 = 1/5$
- 7 $3/10$
- 8 $7/10$
- 9 $3/10$
- 10 $2/5$
- 11 $1/5$
- 12 0.63
- 13 $7/9$
- 14 impossible
- 15 0.4
- 16 18
- 17 $5/7$
- 18 $0.7 (7/11 = 0.636\dots)$

Unit 1 · Sheet B — The probability scale and single events

- 1 (a) $0.35 + 0.2 + x + x + 0.05 = 1$ M1; $2x = 0.4$ M1; $x = 0.2$ A1 [3] (b) 0.25 B1 ft [1] (c) 0.65 B1 [1]
- 2 (a) $6/10 = 3/5$ B1 [1] (b) $7/10$ B1 [1] (c) 0 B1 [1] (d) $(3 + n)/(10 + n) = 1/2$ M1; $6 + 2n = 10 + n$ M1; $n = 4$ A1 [3]
- 3 (a) Arrow at $1/2$ B1 [1] (b) Arrow at $1/4$ B1 [1] (c) probability that it does not rain tomorrow B1; 0.38 B1 [2]
- 4 (a) $30 - 12 - 9 = 9$ M1; $9/30 = 3/10$ A1 [2] (b) $21/30 = 7/10$ B1 [1] (c) $15/33$ M1; $5/11$ A1 [2]
- 5 (a) blue = $30 \times 1/6 = 5$ M1; red + black = 25 M1; black = $25 \times 3/5 = 15$ A1 [3] (b) red = 10 M1; $10/30 = 1/3$ A1 ft [2]

Unit 2 · Sheet A — Relative frequency and expected frequency

- 1 $18/50 = 0.36$
- 2 $44/80 = 0.55$
- 3 $62/300 = 31/150$
- 4 9
- 5 40
- 6 37.5
- 7 75
- 8 126
- 9 $6/400 = 3/200$
- 10 75
- 11 210
- 12 Ben (more trials)
- 13 $90/120 = 3/4$
- 14 $34/40 = 0.85$

Unit 2 · Sheet B — Relative frequency and expected frequency

- 1 (a) 0.28, 0.31, 0.29, 0.294 B2 (B1 for 2 or 3 correct) [2] (b) 0.294 B1; it is from the greatest number of spins B1 [2] (c) Yes: after 500 spins the relative frequency (0.294) is still well above 0.25 (expected about 125 reds, got 147) B1 [1] (d) 0.294×1500 M1; 441 A1 (ft their (b)) [2]
- 2 (a) $40/200 = 1/5$ (0.2) B1 [1] (b) $200 \times 1/6$ M1; 33.3 (or $33 \frac{1}{3}$) A1 [2] (c) e.g. 40 is not far from 33.3 for only 200 rolls — more rolls are needed before deciding B1 [1]
- 3 (a) $12/800 = 3/200$ (0.015) B1 [1] (b) $0.015 \times 24\,000$ M1; 360 A1 [2] (c) relative frequency is only an estimate of the probability / results vary B1 [1]
- 4 (a) 0.86×350 M1; 301 A1 [2] (b) $430 \div 0.86$ M1; = 500 A1; 500 seeds B1 (accept 500) [3]
- 5 (a) 0.12×250 M1; 30 A1 [2] (b) income $250 \times 0.5 = \$125$ M1; prizes $30 \times 3 = \$90$ M1; profit \$35 A1 [3]

Unit 3 · Sheet A — Sample spaces and combined events

- 1 HH, HT, TH, TT
- 2 $1/4$
- 3 $2/4 = 1/2$
- 4 36
- 5 $1/6$
- 6 $1/18$
- 7 $1/6$
- 8 $1/6$
- 9 $1/9$
- 10 $1/12$
- 11 $3/12 = 1/4$
- 12 0.15
- 13 $1/8$
- 14 6

Unit 3 · Sheet B — Sample spaces and combined events

- 1 (a) All 12 products correct B2 (B1 for at least 8 correct): row 1: 2, 4, 6; row 2: 4, 8, 12; row 3: 6, 12, 18; row 4: 8, 16, 24 [2] (b) $2/12 = 1/6$ B1 ft [1] (c) 5 outcomes seen M1; $5/12$ A1 [2] (d) 0 B1 [1]
- 2 (a) $6/36 = 1/6$ B1 [1] (b) 8 outcomes M1; $8/36 = 2/9$ A1 [2] (c) 1 B1; $10/36 = 5/18$ B1 [2]
- 3 (a) 0.3×0.4 M1; 0.12 A1 [2] (b) 0.7×0.6 M1; 0.42 A1 [2] (c) $1 - 0.42$ M1; 0.58 A1 ft [2]
- 4 (a) 123, 132, 213, 231, 312, 321 B2 (B1 for 4 or 5 correct) [2] (b) $2/6 = 1/3$ B1 ft [1] (c) $2/6 = 1/3$ B1 ft [1]
- 5 (a) outcomes (3, 1), (2, 2), (1, 3) seen (6 of 36, or 3 of 18) M1; $1/6$ A1 [2] (b) 9 prime totals out of 18 equally likely pairs M1; $1/2$ A1 [2] (c) $1/2 \times 1/3 = 1/6$ B1 [1]

Unit 4 · Sheet A — Tree diagrams

- 1 0.36
- 2 0.16
- 3 0.48
- 4 $4/25$
- 5 $9/25$
- 6 $12/25$
- 7 $16/25$
- 8 0.21
- 9 $2/15$
- 10 $1/3$
- 11 $8/15$
- 12 $3/10$
- 13 0.65
- 14 multiply

Unit 4 · Sheet B — Tree diagrams

- 1 (a) 0.8 on first 'not late' branch B1; 0.2, 0.8, 0.2, 0.8 on the second set B1 [2] (b) 0.2×0.2 M1; 0.04 A1 [2] (c) 0.2×0.8 M1; $\times 2$ or $+ 0.8 \times 0.2$ M1; 0.32 A1 [3]
- 2 (a) Two-stage tree B1; $7/10$, $3/10$ on first B1; same on second set B1 [3] (b) $3/10 \times 3/10$ M1; $9/100$ A1 [2] (c) $1 - 9/100$ M1; $91/100$ A1 [2]
- 3 (a) First: $5/9$, $4/9$ B1; second: $4/8$, $4/8$ after red and $5/8$, $3/8$ after blue B1 [2] (b) $5/9 \times 4/8$ or $4/9 \times 3/8$ M1; sum of both products M1; $32/72 = 4/9$ A1 [3] (c) $1 - 4/9 = 5/9$ B1 ft [1]
- 4 (a) $6/10 \times 5/9 \times 4/8$ M1; $= 120/720 = 1/6$ A1 (must see working) [2] (b) $1 - 1/6 = 5/6$ B1 [1] (c) $6/10 \times 5/9 \times 4/8$ for one order M1; $\times 3$ M1; $1/2$ A1 [3]

Unit 5 · Sheet A — Venn diagrams and probability

- 1 30
- 2 13
- 3 $2/15$
- 4 $19/30$
- 5 $17/30$
- 6 $1/5$
- 7 $11/30$
- 8 $3/10$
- 9 [6]
- 10 $7/10$
- 11 $7/10$
- 12 3
- 13 $17/40$
- 14 in A or in B or in both

Unit 5 · Sheet B — Venn diagrams and probability

- 1 (a) $22 + 26 + 9 - 50$ M1; 7 A1 [2] (b) 15, 7, 19, 9 in correct regions B2 (B1 for 2 or 3 correct) ft [2] (c) $41/50$ B1 ft [1] (d) $19/50$ B1 ft [1]
- 2 (a) $X \cap Y = \{3, 5, 7, 11, 13\}$ B1; X only = {2}, Y only = {1, 9, 15} B1; outside = {4, 6, 8, 10, 12, 14} B1 [3] (b) $5/15 = 1/3$ B1 ft [1] (c) $9/15 = 3/5$ B1 ft [1] (d) $6/15 = 2/5$ B1 ft [1]
- 3 (a) $0.35 + x + 0.25 + 0.2 = 1$ M1; $x = 0.2$ A1 [2] (b) 0.55 B1 ft [1] (c) 0.2 B1 [1]
- 4 (a) $60 - 43 = 17$ B1 [1] (b) $5 + 3 + 4 = 12$ M1; $12/60 = 1/5$ A1 [2] (c) $14 + 5 = 19$ M1; $19/60$ A1 [2]
- 5 (a) $14 - x + x + 12 - x + 2x = 32$ M1; $26 + x = 32$ so $x = 6$ A1 [2] (b) $8/32 = 1/4$ B1 [1] (c) $26/32 = 13/16$ B1 [1]

Unit 6 · Sheet A — Conditional probability

- 1 $7/20$
- 2 $4/11$
- 3 $4/7$
- 4 $1/2$
- 5 $7/18$
- 6 $3/8$
- 7 $3/7$
- 8 $1/3$
- 9 $4/9$
- 10 0.3
- 11 they are independent
- 12 $1/3$

Unit 6 · Sheet B — Conditional probability

- 1 (a) $116/200 = 29/50$ B1 [1] (b) $46/120$ M1; $23/60$ A1 [2] (c) $42/116$ M1; $21/58$ A1 [2]
- 2 (a) 0.7 B1; 0.6, 0.9 (and 0.1) B1 [2] (b) 0.3×0.4 or 0.7×0.1 M1; $0.12 + 0.07$ M1; 0.19 A1 [3] (c) $0.12 \div 0.19$ M1; $12/19$ A1 ft [2]
- 3 (a) $6/10 \times 5/9$ M1; $+ 4/10 \times 3/9$ M1; $42/90 = 7/15$ A1 [3] (b) $1 - 6/10 \times 5/9$ M1; $2/3$ A1 [2]
(c) $(30/90) \div (42/90)$ M1; $5/7$ A1 ft [2]
- 4 (a) 0.35, 0.15, 0.25 B1; 0.25 outside B1 [2] (b) $0.15 \div 0.4$ M1; 0.375 (3/8) A1 [2] (c) $0.35 \div 0.6$ M1; $7/12$ A1 [2] (d) No, $P(A | B) = 0.375 \neq P(A) = 0.5$ (or $0.5 \times 0.4 = 0.2 \neq 0.15$) B1 [1]

Unit 7 · Sheet A — Types of data and tables

- 1 discrete
2 continuous
3 qualitative
4 continuous
5 qualitative
6 8
7 3
8 9
9 9
10 19
11 $11/60$
12 $20 < h \leq 30$

Unit 7 · Sheet B — Types of data and tables

- 1 (a) Boys: sandwich 31, salad 5; Girls: hot meal 22, salad 16, total 56; sandwich total 49 B3 (B2 for 4 or 5 correct, B1 for 2 or 3 correct) [3] (b) $16/56 = 2/7$ B1 ft [1] (c) $31/120$ B1 ft [1] (d) $50/120 \times 100$ M1; 41.7% A1 [2]
- 2 (a) 6, 8, 7, 3 B3 (B2 for 3 correct, B1 for 2 correct) [3] (b) $2 < m \leq 4$ B1 ft [1] (c) $0 < m \leq 2$, because ≤ 2 includes 2 B1 [1] (d) $10/24 = 5/12$ B1 ft [1] (e) the exact values are lost (cannot find exact mean/range) B1 [1]
- 3 (a) Jun (difference 79 mm) B1 [1] (b) City A 477 mm B1; City B 437 mm B1 [2] (c) e.g. City A has more rain overall but drier early months; City A's rainfall varies much more B1 [1]
- 4 (a) $0 + 9 + 16 + 12 + 8 + 5$ M1; 50 A1 [2] (b) $7/30$ B1 [1] (c) 9 is the frequency; the most common number of goals is 1 B1 [1] (d) discrete B1 [1]
- 5 (a) continuous B1 [1] (b) discrete B1 [1]
(c) qualitative B1 [1]

Unit 8 · Sheet A — Bar charts, pictograms and pie charts

- 1 (a) 28 (b) 18 (c) 92 (d) $2\frac{1}{2}$
2 (a) 12 (b) 36 (c) $7/36$ (d) 90°
3 (a) 135° (b) 90° (c) 75° (d) 60°
4 (a) 30 (b) 16 (c) $100/360 = 5/18$
5 9

Unit 8 · Sheet B — Bar charts, pictograms and pie charts

- 1 (a) $360 \div 45 = 8^\circ$ per student M1; $96^\circ, 72^\circ, 120^\circ, 48^\circ, 24^\circ$ A1 [2]
(b) all sectors correct $\pm 2^\circ$ B1; labelled B1 ft [2] (c) $15/45 = 33.3\%$ ($33\frac{1}{3}\%$) B1 [1]
- 2 (a) $135/360 \times 24$ M1; 9 hours A1 [2] (b) 5 hours B1 [1] (c) $105 > 90$ or $105/360 = 0.29 > 0.25$ M1; Yes A1 [2]
- 3 (a) March B1 [1] (b) 28 B1 [1] (c) girls read more in total (31 v 28) B1; e.g. boys read more in February and April B1 [2]
- 4 (a) 36 B1 [1] (b) 33 B1 [1] (c) $42 \div 12 = 3.5$ M1; $3\frac{1}{2}$ symbols drawn A1 [2] (d) 129 B1 ft [1]
- 5 (a) A: $90/360 \times 600 = 150$ B1; B: $120/360 \times 300 = 100$ B1 [2]
(b) pie charts show proportions, not numbers; School A has more walkers ($150 > 100$) B1 [1]
- 6 axes labelled and scale B1; 5 bars correct heights (3, 6, 5, 4, 2) B1; equal widths with gaps B1 [3]

Unit 9 · Sheet A — Averages, range and stem-and-leaf

- 1 (a) 6 (b) 7 (c) 7 (d) 8
2 (a) 14.5 (b) 15
3 10
4 14
5 (a) 43.5 minutes (b) 42 minutes (c) 29 minutes (d) 4
6 (a) $36 \div 25 = 1.44$ (b) 1 (c) 1
7 mode
8 mean

Unit 9 · Sheet B — Averages, range and stem-and-leaf

- 1 (a) median 36 B1; range 20 B1 [2] (b) Class B did better on average (higher median) B1; Class B's marks were more consistent (smaller range) B1 [2] (c) 31 B1 [1]
- 2 (a) stems 2, 3, 4, 5 B1; leaves ordered: 2 | 3 6 8 9; 3 | 1 3 5 5 7 8; 4 | 0 1 4 7; 5 | 2 B1; key B1 [3] (b) 35 cm B1 [1] (c) 29 cm B1 [1]
- 3 (a) $\Sigma fx = 23 + 30 + 21 + 16 + 5$ M1; = 95 M1 (dep); $95 \div 50 = 1.9$ A1 [3] (b) 2 B1 [1] (c) 1 B1 [1]
- 4 (a) $6 \times 24 = 144$ M1; $5 \times 25 = 125$ M1; 19 A1 [3] (b) $12 \times 52 + 18 \times 47$ M1; = 1470 M1; $\div 30 = 49$ kg A1 [3]
- 5 (a) $235 \div 7$ M1; 33.6 (thousand dollars) A1 [2] (b) 24 B1 [1] (c) median B1; the mean is raised by the extreme value 95 B1 [2]

Unit 10 · Sheet A — Grouped data and the estimated mean

- 1 (a) 157.5 (b) $160 < h \leq 165$ (c) 13 (d) $160 < h \leq 165$ (e) 6505 (f) 162.6 cm (162.625) (g) the exact heights are not known; mid-points are used
2 (a) 25.5 (b) $715 \div 30 = 23.8$ (c) 21–30
3 $8 \times 2.5 = 20$

Unit 10 · Sheet B — Grouped data and the estimated mean

1 (a) $60 < t \leq 90$ B1 [1] (b) $60 < t \leq 90$ B1 [1] (c) mid-points 15, 45, 75, 105, 135 M1; $\Sigma fx = 120 + 765 + 1575 + 1050 + 540$ M1; $4050 \div 60$ M1; 67.5 min A1 [4] (d) mid-points are used / actual times not known B1 [1] (e) $20/30 \times 10 (= 6.67)$ M1; $+ 4 = 10.7$, so about 11 A1 [2]

2 (a) $6 \times 45 + 55x + 14 \times 65 + 5 \times 75 = 1555 + 55x$ M1; $\div (25 + x) = 60$ M1; $1555 + 55x = 1500 + 60x$ M1; $x = 11$ A1 [4] (b) $19/36$ M1; 52.8% A1 ft [2] (c) $60 < m \leq 70$ B1 ft [1]

3

(a) $\Sigma fx = 70 + 225 + 247.5 + 65$ M1; $607.5 \div 25$ M1; 24.3 min A1 [3] (b) Group B — lower estimated mean ($24.3 < 24.5$) B1 ft [1] (c) the exact fastest and slowest times are not known B1 [1]

4 (a) mid-points 2, 7, 12, 17 M1; $\Sigma fx = 240$ M1; 8 A1 [3] (b) 5–9 B1 [1] (c) 5–9 B1 [1]

Unit 11 · Sheet A — Scatter diagrams and correlation

- 1** (a) strong positive (b) negative (c) no correlation
2 positive
3 negative
4 none
5 (5, 10)
6 (a) 57–58 (57.5) (b) 6 hours (c) 15 is outside the range of the data (extrapolation)
7 a point that does not follow the pattern of the rest
8 No

Unit 11 · Sheet B — Scatter diagrams and correlation

- 1** (a) all 10 points correct ($\pm \frac{1}{2}$ small square) B3 (B2 for 8 or 9, B1 for 6 or 7) [3] (b) (strong) positive B1 [1] (c) 25.3 °C B1; 69.6 B1 [2] (d) ruled line through (25.3, 69.6) following the trend B1 [1] (e) answer from their line, typically 74–80 (regression gives 77) B1 ft [1] (f) 40 °C is outside the range of the data (extrapolation) B1 [1]
2 (a) negative B1 [1] (b) (7, 18) B1; e.g. car in excellent condition / rare / low mileage B1 [2] (c) ruled line following trend, ignoring (7, 18) B1 [1] (d) answer from their line, about 9.6 to 11.6 thousand dollars B1 ft [1] (e) No — 20 years is far outside the data; the line would give a negative value B1 [1]
3 (a) positive B1 [1] (b) $0.95 \times 160 + 9$ M1; 161 cm A1 [2] (c) $0.95 \times 158 + 9 = 159.1$ cm (line passes through mean point) B1 [1] (d) $(170 - 9) \div 0.95$ M1; 169 cm (169.5) A1 [2]
4 (a) positive correlation — higher maths scores go with higher science scores B1 [1] (b) reading from line at 70 M1; 71–73 A1 [2] (c) outlier — low maths score but high science score B1 [1] (d) correlation does not prove cause; another factor (e.g. revision) may affect both B1 [1]

Unit 12 · Sheet A — Cumulative frequency, quartiles and box plots

- 1** (a) 34.2 min (accept ± 0.5) (b) 30.0 min (c) 38.5 min (d) 8.5 min (e) 51 (f) 14 (g) 42.6 min
2 16
3 10
4 $22 - 10 = 12$
5 (a) 23 kg (b) 12 kg (c) 29 kg (d) $1/4$

Unit 12 · Sheet B — Cumulative frequency, quartiles and box plots

- 1** (a) 3, 12, 32, 58, 73, 80 B1 [1] (b) points plotted at upper bounds (4, 3) ... (24, 80) B2 (B1 for 4 or 5 correct); smooth curve or polygon through points B1 [3] (c) 13.2 hours (accept 13–14) B1 ft [1] (d) $UQ \approx 16.4$ and $LQ \approx 9.8$ M1; $IQR \approx 6.6$ (accept 6–7) A1 ft [2] (e) reading at 18 ≈ 67 M1; $80 - 67 \approx 13$ (accept 13–17) A1 ft [2]
2 (a) 52 B1 [1] (b) X: 19 B1; Y: 13 B1 [2] (c) Class Y had a higher median, so did better on average B1; Class Y had a smaller IQR, so marks were more consistent B1 [2]
3 (a) 46 B1 [1] (b) box from 27 to 46 with median at 38 B1; whiskers to 15 and 62 B1 ft [2]
4 (a) 28.4 mm (accept ± 1) B1 [1] (b) reading at 160 M1; 38.6 mm (accept ± 1) A1 [2] (c) reading at 35 ≈ 140 M1; 60 (accept ± 4) A1 [2]
5 (a) median 11 B1; LQ 6 B1; UQ 16 B1 [3] (b) 10 B1 ft [1]

Unit 13 · Sheet A — Histograms and frequency density

- 1** 1.5
2 3.6
3 0.3
4 24
5 6
6 the frequency
7 12
8 18
9 21
10 10
11 81
12 $10.5 + 5 = 15.5$, about 15 or 16
13 $20 < a \leq 30$
14 bars have different widths; frequency = area

Unit 13 · Sheet B — Histograms and frequency density

- 1** (a) 0.5, 1.8, 2.6, 1.6, 0.35 B2 (B1 for 3 or 4 correct) [2] (b) 5 bars with correct widths B1; heights correct ft B2 (B1 for 3 or 4 correct) [3] (c) $110 < m \leq 120$ B1 [1] (d) $\frac{1}{2} \times 32 (= 16)$ M1; $16 + 14 = 30$ A1 [2]
2 (a) 22 and 30 B1; 10 B1 [2] (b) bar $20 < t \leq 30$ at height 1.6 B1 [1] (c) 90 B1 ft [1] (d) $\frac{1}{2} \times 16 = 8$ M1; $8 + 10 = 18$ A1 [2]
3 (a) mid-points 2.5, 7.5, 15, 25, 40 M1; $\Sigma fx = 1445$ M1; $\div 90$ M1; 16.1 min A1 [4] (b) 45th value M1; 11 more needed in $10 < t \leq 20$ M1; $10 + 11/30 \times 10 = 13.7$ min A1 [3]
4 $45 \div (3 \times 10) = 1.5$ people per cm^2 M1; $4 \times 5 = 20 \text{ cm}^2$ M1; 30 A1 [3]
5 so that the area of each bar is proportional to the frequency (wider classes would look too big) B1 [1]

End-of-book check

- 1 13/20 B1 [1]
- 2 $(1 - 0.55) \div 2$ M1; 0.225 A1 [2]
- 3 (a) 0.24 B1 (b) 144 B1 ft [2]
- 4 5/36 B1 [1]
- 5 0.49 B1 [1]
- 6 $3/8 \times 2/7$ M1; $3/28$ A1 [2]
- 7 $9/24 = 3/8$ B1 [1]
- 8 $\Sigma fx = 53$ M1; 2.65 A1 [2]
- 9 mid-points 5, 15, 25 M1; $\Sigma fx = 280$ M1; 14 A1 [3]
- 10 LQ 4 and UQ 12 M1; 8 A1 [2]
- 11 30 B1 [1]
- 12 27 B1 [1]

Practice paper

- 1 (a) adults swimming 23; children tennis 18; children total 45; swimming total 50 B3 (B2 for 3 correct, B1 for 2) [3] (b) 23/80 B1 ft [1] (c) 18/30 M1; $3/5$ A1 ft [2]
- 2 (a) 41 B1 [1] (b) 36 B1 [1] (c) 32 B1 [1] (d) LQ = 30 and UQ = 41 M1; 11 A1 [2]
- 3 (a) $360 \div 72 = 5$ M1; 100° , 70° , 130° , 60° A1 [2] (b) sectors correct $\pm 2^\circ$ B1; labels B1 [2] (c) $26/72 = 13/36$ B1 [1]
- 4 (a) first: $6/10$, $4/10$ B1; second: $5/9$, $4/9$ and $6/9$, $3/9$ B1 [2] (b) $6/10 \times 5/9$ or $4/10 \times 3/9$ M1; add both M1; $42/90 = 7/15$ A1 [3] (c) $1 - 6/10 \times 5/9$ M1; $2/3$ A1 [2]
- 5 (a) negative B1 [1] (b) (2.5, 18) circled B1 [1] (c) ruled line following the trend B1 [1] (d) from their line, about 11 to 13 km per litre B1 ft [1]
- 6 (a) 6, 24, 61, 93, 112, 120 B1 [1] (b) points at upper bounds B2 (B1 for 4 or 5 correct); smooth curve B1 [3] (c) 14.9 min (accept ± 0.5) B1 ft [1] (d) UQ ≈ 19.4 , LQ ≈ 11.0 M1; ≈ 8.5 min A1 ft [2] (e) reading at 22 ≈ 102 M1; ≈ 18 (accept ± 3) A1 ft [2]
- 7 (a) 14 B1; 6 B1 [2] (b) $0 < w \leq 2$ at height 4 B1; $3 < w \leq 5$ at height 9 B1 [2] (c) $\frac{1}{2} \times 18 = 9$ M1; $9 + 6 = 15$ A1 [2]
- 8 (a) mid-points 15, 25, 35, 50 M1; $\Sigma fx = 1230$ M1; $\div 40$ M1; 30.75 (30.8) A1 [4] (b) $20 < a \leq 30$ B1 [1]
- 9 (a) both = 6 M1; 16, 6, 12, 6 A1 [2] (b) $6/40 = 3/20$ B1 [1] (c) $6/22$ M1; $3/11$ A1 ft [2]
- 10 (a) $75/300 = 0.25$ B1 [1] (b) 0.25×1200 M1; 300 A1 [2] (c) No (probably biased): 0.25 is well above $1/6 \approx 0.17$ after many rolls B1 [1]
- 11 (a) 11 outcomes M1; $11/36$ A1 [2] (b) 9 outcomes M1; $9/36 = 1/4$ A1 [2]