

Sampling and Populations

Solutions

For AQA GCSE Mathematics (8300) • Higher tier

Each question is shown in grey, followed by the working, one step per line. The final answer is in a box. *Marks* shows how the marks are usually given in the exam, and **Tip** points out a common mistake.

1 (5 marks)

A headteacher wants to find out what the 1200 students at her school think about the school lunches.

- (a) She asks 60 students who are in the lunch queue.
Give one reason why this sample may be biased.

Students in the lunch queue have chosen to buy school lunches.

So they are more likely to like them.

Students who bring a packed lunch are left out of the sample.

Answer: Students in the queue are more likely to like school lunches

Tip: Link your reason to this survey. “The sample is too small” does not get the mark: explain **who** is missed out or over-represented.

- (b) She then takes a random sample of 60 students.
Her deputy says, “38% of the students in this sample like the school lunches.”
Give a reason why the deputy must be wrong.

Work out 38% of the 60 students in the sample.

$$0.38 \times 60 = 22.8$$

The number of students must be a whole number.

So 38% of this sample is not possible.

Answer: 38% of 60 is 22.8, which is not a whole number of students

Tip: Many students work out 22.8 but do not say why it is a problem. Finish with the reason: a number of students must be a whole number.

- (c) In fact, 24 of the 60 students in the random sample like the school lunches.
Estimate how many of the 1200 students like the school lunches.
Assume that the sample is representative of the whole school.

Fraction of the sample who like the lunches: $\frac{24}{60} = 0.4$

The sample is representative, so the same fraction of all 1200 students like them.

$$0.4 \times 1200 = 480$$

Answer: 480

Marks: 1 mark for $\frac{24}{60} \times 1200$. 1 mark for 480.

- (d) Later, she finds out that 45 of the 60 students in the random sample have a school lunch every day, but only half of the 1200 students do.
What does this suggest about how many students at the school like the lunches?
Tick one box.

In the sample, $\frac{45}{60} = 75\%$ have a school lunch every day.

In the whole school, only 50% do.

So students who have a school lunch every day are over-represented in the sample.

These students are more likely to like the school lunches.

So the sample is likely to give too high a proportion who like the lunches.

The actual number is probably less than 480.

Answer: It is probably less than the answer to part (c)

2 (3 marks)

Amir and Beth each want to estimate the proportion of cars in a car park that are electric.

Amir chooses 20 cars at random. 3 of them are electric.

Beth chooses 150 cars at random. 27 of them are electric.

(a) Amir says,

“My results show that exactly 15% of the cars in the car park are electric.”

Is Amir correct? Tick a box.

Give a reason for your answer.

Amir's results: $\frac{3}{20} = 15\%$ of his sample are electric.

But he only looked at a sample of 20 cars, not every car in the car park.

A sample only gives an **estimate**.

A different sample could give a different percentage: Beth's sample gives

$$\frac{27}{150} = 18\%.$$

So Amir is not correct.

Answer: No. 15% is only an estimate from a small sample (another sample could give a different answer)

Tip: Say why a sample cannot give an **exact** answer for the whole car park. Just writing “No” does not get the mark.

(b) There are 900 cars in the car park.

Use Beth's results to estimate the number of electric cars in the car park.

Beth's sample is much larger than Amir's, so it should give the more reliable estimate.

Fraction of Beth's sample that are electric: $\frac{27}{150}$

Use the same fraction of all 900 cars:

$$\frac{27}{150} \times 900 = 162$$

Answer: 162

Marks: 1 mark for $\frac{27}{150} \times 900$. 1 mark for 162.

3 (1 mark)

A factory makes 5000 bolts a day. A quality inspector tests the first 50 bolts made each day.

Give one reason why this sample may not be representative.

The sample is not random: it is always the first 50 bolts.

The first bolts made each day may be different from the rest, e.g. while the machines warm up.

Answer: Not random / first bolts may differ

4 (1 mark)

Maya claims that most of the flats in her 10-floor building have a view of the park. She samples 40 flats from floors 1 to 3.

Give a reason why this sample may not be useful in testing Maya's claim.

The sample only includes flats on floors 1 to 3.

Lower flats are less likely to have a view of the park.

Flats on floors 4 to 10 are not included at all.

Answer: Only the lower floors are sampled, so it is not representative

Tip: Be specific about who is left out (floors 4 to 10). A general answer such as "she did not ask every flat" is not enough.

5 (4 marks)

Owen states a hypothesis.

"Most adults in my town spend more than £20 a week on takeaway food."

- (a) Owen plans to ask the first 10 customers who leave a pizza takeaway on a Friday evening.

Give **two** criticisms of Owen's sampling method.

Criticism 1: a sample of 10 is too small to represent a whole town.

Criticism 2: people leaving a takeaway have just bought takeaway food, so they are likely to spend more on takeaways than most adults. The sample is biased.

Other valid criticisms: it only samples one day and one time of the week; the first 10 customers are not a random sample; some of the customers may not live in the town or may not be adults.

Answer: Any two of: sample too small; biased (takeaway customers spend more on takeaways); only one day/time; not random; customers may not be adults from the town

Marks: 1 mark for each different valid criticism.

(b) Owen changes his method.

He takes a random sample of 120 adults from the town.

47 of them spend more than £20 a week on takeaway food.

Does this sample support Owen's hypothesis?

You must show your working.

"Most" means more than half.

Half of 120 is 60.

47 is less than 60.

Or: $\frac{47}{120} \times 100 = 39.2\%$ (1 dp), which is less than 50%.

So the sample does **not** support his hypothesis.

Answer: No, $\frac{47}{120} \approx 39\%$, which is less than half

Tip: "Most" means more than 50%. Compare the proportion in the sample with one half.

Marks: 1 mark for comparing 47 with 60, or for 39.2% (or 0.39). 1 mark for No with correct working.

6 (4 marks)

A theatre asks a sample of the people who went to a show one week how they heard about the show.

The table shows the results.

How they heard	Social media	Radio	A friend	Poster
Frequency	63	21	45	21

- (a) 4200 people went to the show that week.

Estimate how many more people heard about the show from social media than from a friend.

You may assume the sample represents everyone who went to the show.

Social media – friend in the sample: $63 - 45 = 18$

Sample size: $63 + 21 + 45 + 21 = 150$

Fraction of the sample: $\frac{18}{150}$

Estimate = $\frac{18}{150} \times 4200 = 504$

(Or: social media $\frac{63}{150} \times 4200 = 1764$, friend $\frac{45}{150} \times 4200 = 1260$, and $1764 - 1260 = 504$.)

Answer: 504

Tip: Find the total of the sample first. The estimate is a fraction of the 4200 people, not of the 150.

Marks: 1 mark for 150 or for $\frac{63}{150}$, $\frac{45}{150}$ or $\frac{18}{150}$. 1 mark for a complete method. 1 mark for 504.

- (b) The sample was taken from people who booked their tickets online.

Give one reason why the sample may not be representative of everyone who went to the show.

People who book online are more likely to use social media.

So social media is likely to be over-represented in the sample.

People who bought tickets at the box office (for example, older people) were not asked.

Answer: People who book online are more likely to use social media, so they are not typical of everyone who went

7 (4 marks)

A factory tests batches of light bulbs for faults.
The table shows the results.

Batch	A	B	C	D
Number of bulbs tested	80	150	250	120
Number of faulty bulbs	6	9	12	9

- (a) Which batch gives the most reliable estimate of the probability that a bulb is faulty?

Give a reason for your answer.

Batch C.

It has the most bulbs tested (250), and a bigger sample gives a more reliable estimate.

Answer: C, because it is the largest sample

Tip: The most reliable estimate comes from the **largest** sample, not from the batch with the most (or fewest) faulty bulbs.

- (b) Use all the results to work out the best estimate of the probability that a bulb is faulty.

Combine all four batches into one big sample.

Faulty: $6 + 9 + 12 + 9 = 36$

Tested: $80 + 150 + 250 + 120 = 600$

$$\text{Best estimate} = \frac{36}{600} = 0.06$$

Answer: 0.06 (or $\frac{3}{50}$)

Tip: Add up all the faulty bulbs and all the bulbs tested. Do not find the mean of the four relative frequencies: the batches are different sizes.

Marks: 1 mark for $\frac{36}{600}$ (or both totals). 1 mark for 0.06.

- (c) The factory makes 15 000 bulbs a week.
Estimate how many of these bulbs will be faulty.

Expected number = probability $\times$ number of bulbs

$$0.06 \times 15\,000 = 900$$

Answer: 900