Worksheet - Sampling and scatter graphs

Year: Year 11 · 20

Show your working.

Question 1

Question 1 – A stratified sample *(Section A · Skills practice)*

A school has 800 pupils. A stratified sample of 80 pupils is to be chosen by year group.

Year	7	8	9	10	11
Number of pupils	180	170	160	150	140

a. Explain why a stratified sample is better than asking the first 80 pupils to arrive. [2 marks]
b. Work out the number to choose from each year. [4 marks]
c. Check that your numbers add to 80. [1 mark]

Question 2

Question 2 – Estimating from a sample *(Section A · Skills practice)*

A random sample of 50 pupils is asked how they travel to school. 18 of them walk. There are 1 200 pupils in the school.

a. Estimate the number of pupils in the school who walk. [2 marks]
b. Explain why this is only an estimate. [1 mark]
c. A second random sample of 100 pupils has 41 who walk. Use both samples together to improve the estimate. [3 marks]

Question 3

Question 3 – Spotting bias *(Section A · Skills practice)*

For each survey, say why the sample may be biased and suggest an improvement.

a. Asking customers outside a gym about how often people exercise. [2 marks]
b. Asking only the pupils in the school choir which music the school should buy. [2 marks]
c. Phoning people at 11 am on a weekday about their jobs. [2 marks]

Question 4

Question 4 – Capture-recapture (Higher) *(Section A · Skills practice)*

A biologist wants to estimate the number of frogs in a pond. She catches 60 frogs,

marks them and releases them. Later she catches 40 frogs, of which 8 are marked.

a. Write a proportion for the second sample. [1 mark]
b. Estimate the number of frogs in the pond. [3 marks]
c. State two assumptions you have made. [2 marks]

Question 5

Question 5 – Describing a scatter graph *(Section A · Skills practice)*

Eight pupils revised for a test. The table shows hours of revision, x, and test mark, y (out of 50).

x	1	2	3	4	5	6	7	8
y	12	15	22	24	30	31	38	41

a. Describe the correlation. [1 mark]
b. Which points would you use to draw a line of best fit? Write the equation of the line through $(2, 15)$ and $(8, 41)$ to 1 decimal place in the gradient. [4 marks]
c. Use the line to estimate the mark for a pupil who revised for 5.5 hours. [2 marks]

Question 6

Question 6 – The equation of a line of best fit *(Section A · Skills practice)*

A line of best fit passes through $(2, 40)$ and $(10, 80)$.

a. Work out the gradient. [2 marks]
b. Find the equation in the form $y = mx + c$. [2 marks]
c. Use it to predict y when $x = 7$. [1 mark]
d. Explain what the gradient means if x is hours of revision and y is mark. [1 mark]

Question 7

Question 7 – Fish in a lake *(Section B · Real-life practice)*

A park ranger catches 80 fish in a lake and puts a small tag on each one. A week later she catches 60 fish and 12 of them have tags.

a. Estimate the number of fish in the lake. [3 marks]
b. The ranger catches the second sample using a different net that cannot catch small fish. Explain how this affects the estimate. [2 marks]
c. State why it is important to wait a week between the samples. [1 mark]

Question 8

Question 8 – A stratified sample in a company *(Section B · Real-life practice)*

A company has 150 office workers, 60 factory workers and 30 drivers. A stratified

sample of 40 employees is asked about travel to work.

a. Work out how many from each group should be in the sample. [4 marks]
b. A random sample of 40 from the whole list could contain no drivers. Why might that matter? [2 marks]
c. Describe how to choose the 10 factory workers fairly. [2 marks]

Question 9

Question 9 – Counting deer (Higher) *(Section B · Real-life practice)*

45 deer in a forest are marked. Later, 30 deer are seen and 6 of them have marks.

a. Estimate the number of deer in the forest. [3 marks]
b. Give one reason why the true number may be different. [1 mark]
c. If 5 more marked deer had been seen in the second sample, how would the estimate change? [2 marks]

Question 10

Question 10 – Arm span and height *(Section B · Real-life practice)*

The table shows the height and arm span (in cm) of 8 pupils.

Height	150	155	160	165	170	175	180	185
Arm span	148	157	158	168	169	178	179	190

a. Describe the relationship. [2 marks]
b. One pupil has height 168 cm. Estimate their arm span. [2 marks]
c. Explain why it is unwise to use the graph to predict the arm span of a person 120 cm tall. [2 marks]
d. Is it correct to say that a longer arm span causes greater height? [1 mark]

Question 11

Question 11 – Ice cream sales *(Section B · Real-life practice)*

An ice cream van records the temperature and sales. A line of best fit passes through $(15, 80)$ and $(30, 260)$.

a. Work out the gradient and explain what it means. [3 marks]
b. Write the equation of the line. [2 marks]
c. Predict the sales on a day of 27°C. [2 marks]
d. Why is a prediction for 40°C less reliable? [1 mark]

Question 12

Question 12 – A systematic sample *(Section B · Real-life practice)*

A school register has 600 names. A teacher takes a systematic sample by choosing every 15th name, starting with the 7th.

a. How many pupils are in the sample? [2 marks]
b. Write the position of the first four names chosen. [2 marks]
c. Give one advantage and one disadvantage of this method. [2 marks]

Question 13

Question 13 – Value of a car *(Section B · Real-life practice)*

The value of a car, y thousand pounds, falls with age, x years. A line of best fit has the equation $y = 15 - 1.2x$.

a. What type of correlation is shown? [1 mark]
b. Write down the meaning of the gradient and the intercept. [2 marks]
c. Estimate the value of a 5-year-old car. [2 marks]
d. After how many years does the equation predict the value to be zero? Comment. [2 marks]

Question 14

Question 14 – Choosing a random sample *(Section B · Real-life practice)*

The 480 pupils in Years 10 and 11 are numbered 1 to 480. A calculator random number generator gives: 47, 312, 479, 505, 47, 188, 0, 260.

a. Which of these numbers cannot be used? Explain. [2 marks]
b. Which number is repeated and what should you do? [2 marks]
c. Write the numbers of the first five pupils chosen. [2 marks]
d. Explain why this sample is random. [1 mark]

Question 15

Question 15 – Comparing two samples *(Section C · Problem solving)*

Two researchers estimate the mean time pupils spend on homework each week. Sample A has 20 pupils with mean 3.5 hours. Sample B has 100 pupils with mean 4.1 hours.

a. Which sample is more likely to give a reliable estimate? Explain. [2 marks]
b. Work out the mean of the two samples combined. [3 marks]
c. The school has 900 pupils. Estimate the total hours of homework per week using your combined mean. [2 marks]

Question 16

Question 16 – Weaknesses of capture-recapture (Higher) *(Section C · Problem*

solving)

50 fish are tagged. A second sample of 80 contains 10 tagged fish.

a. Estimate the population. [3 marks]
b. Some tags fall off between the two samples. Does this make the estimate too high or too low? Explain. [3 marks]
c. After the second sample, 20 new fish are born and 10 die. What other issue arises? [2 marks]

Question 17

Question 17 – An outlier on a scatter graph *(Section C · Problem solving)*

Seven pupils have these values of (x, y): $(1, 12)$, $(2, 15)$, $(3, 18)$, $(4, 20)$, $(5, 23)$, $(6, 40)$, $(7, 28)$.

a. Which point is an outlier? [1 mark]
b. Draw a line of best fit that ignores the outlier through $(1, 12)$ and $(7, 28)$ and find its equation. [4 marks]
c. Use it to estimate y when $x = 4.5$. [2 marks]
d. Give two possible reasons for the outlier. [2 marks]

Question 18

Question 18 – Two-way stratification *(Section C · Problem solving)*

A school has 1 200 pupils: 720 boys and 480 girls. A stratified sample of 60 pupils is chosen by gender.

a. Work out the number of boys and girls in the sample. [3 marks]
b. Among the sample, 21 boys and 9 girls prefer football. Estimate the number of pupils in the school who prefer football. [4 marks]
c. Estimate the percentage of the girls in the school who prefer football. [2 marks]

Question 19

Question 19 – Evaluating a claim *(Section C · Problem solving)*

A magazine asks its readers to text in if they think the high street is getting worse. 85% of the 400 replies say it is.

a. Give two reasons why the result may not represent all adults. [2 marks]
b. The magazine claims that "85% of adults think the high street is getting worse". Is this justified? [2 marks]
c. Suggest a better way to get the information. [2 marks]
d. A random sample of 1 000 adults finds 61% agree. Which result do you trust more and why? [2 marks]

Question 20

Question 20 – Challenge: a changing population *(Section D · Challenge)*

50 birds are ringed. A second sample of 40 birds has 5 ringed. After this, 100 new unringed birds arrive.

a. Estimate the original population. [3 marks]
b. What is the new population? [1 mark]
c. A third sample of 60 birds is taken. How many ringed birds would you expect? [3 marks]
d. Explain what the capture-recapture formula would give if you used the first sample (50 ringed) and the third sample wrongly, ignoring the new birds. [3 marks]