



F04-test Skewness by Inspection: Test Pack

EDEXCEL 1ST0 · Calculator allowed · about 30 minutes

Total Marks

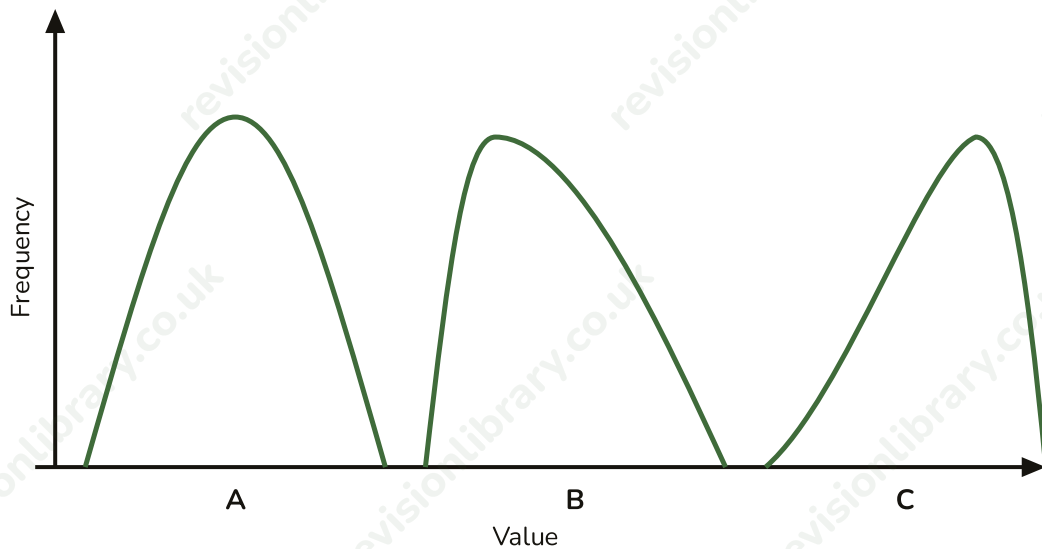
Name: _____ Date: ____ / ____ / ____

Answer ALL questions. Show all your working.

- 1** For a positively skewed distribution, write down whether the mean is usually greater than, or less than, the median.

(Total for Question 1 is 1 mark)

2 The diagram shows three distributions, A, B and C, drawn as smooth frequency curves.



- (a) State whether distribution A is symmetric, positively skewed or negatively skewed. (1)
- (b) State whether distribution B is symmetric, positively skewed or negatively skewed. (1)
- (c) State whether distribution C is symmetric, positively skewed or negatively skewed. (1)

(Total for Question 2 is 3 marks)

3 The number of pets owned by each of 9 pupils in a class is: 1, 2, 2, 3, 3, 3, 4, 4, 5.

(a) Work out the mean number of pets. (2)

.....
(b) Write down the median number of pets. (1)

.....
(c) Write down the modal number of pets. (1)

.....
(d) State, with a reason, whether this distribution is symmetric, positively skewed or negatively skewed. (1)

(Total for Question 3 is 5 marks)

4 A holiday company recorded the number of nights that 200 customers stayed at a resort. The distribution of the lengths of stay is negatively skewed. Describe what this tells you about the shape of the distribution, and about how the lengths of stay are spread either side of the median.

(Total for Question 4 is 2 marks)

5 A survey of 40 employees' annual leave days used has a lower quartile of 8, a median of 12 and an upper quartile of 15. Compare $Q_3 - Q_2$ with $Q_2 - Q_1$ to determine whether the distribution is positively skewed or negatively skewed.

(Total for Question 5 is 2 marks)

6 Priya, a Year 11 student, wants to investigate the statement: 'Most students at my school have a short journey to school, but a few travel a long way.' She plans to collect journey times, in minutes, from a sample of students.

(a) Suggest an appropriate sampling method Priya could use to choose her sample of students, and give one reason for your choice. (2)

(b) Priya collects the following journey times, in minutes, from 11 students: 5, 6, 6, 7, 8, 9, 10, 12, 15, 20, 56. Find the median journey time. (1)

(c) Calculate the mean journey time. (2)

(d) Compare the mean and the median, and use this comparison to explain whether Priya's data supports the statement that journey times are positively skewed. (2)

(Total for Question 6 is 7 marks)

7 A data set has a mean of 48, a median of 42, and a lower quartile of 38. A student claims: 'Since the mean is greater than the median, this data set must be negatively skewed.' Comment on whether the student is correct, explaining your reasoning.

(Total for Question 7 is 3 marks)

8 A data set has a mean of 52, a median of 60 and an upper quartile of 65. A student claims: 'Since the mean is less than the median, this data must be negatively skewed.'

- (a) State whether the student's claim about the type of skew is correct, based on the mean and median alone. (1)
- (b) The lower quartile of the data is 38. Show that the quartile spacing (median - lower quartile, compared with upper quartile - median) agrees with the student's conclusion. (2)
- (c) Give one reason why it is good practice to check for skewness using more than one method (for example, both the mean/median comparison and the quartile spacing). (1)

(Total for Question 8 is 4 marks)

Mark scheme · F04-test Skewness by Inspection: Test Pack

Question 1

- **B1** greater than cao
- **Answer: Greater than**

Question 2

- (a) **B1** Symmetric cao
- (a) **Answer: Symmetric**
- (b) **B1** Positively skewed cao
- (b) **Answer: Positively skewed**
- (c) **B1** Negatively skewed cao
- (c) **Answer: Negatively skewed**

Question 3

- (a) **M1** $1+2+2+3+3+3+4+4+5 (=27)$, oe divided by 9
- (a) **A1** 3 cao
- (a) **Answer: 3**
- (b) **B1** 3 cao
- (b) **Answer: 3**
- (c) **B1** 3 cao
- (c) **Answer: 3**
- (d) **B1** Symmetric, since mean = median = mode (all equal to 3), oe
- (d) **Answer: Symmetric**

Question 4

- **B1** describes the shape: most customers stayed for a long time (values bunched at the higher end), with a small tail of customers who stayed for a much shorter time
- **B1** correctly links this to spread either side of the median: the lengths of stay below the median are more spread out (a longer tail) than those above the median, which are bunched closely together
- **Answer: Most customers stayed for a long time (values bunched at the higher end), with a small tail of customers who stayed for a much shorter time; the lengths of stay below the median are more spread out than those above it, which are bunched closely together.**

Question 5

- **M1** $Q3 - Q2 = 3$ and $Q2 - Q1 = 4$ oe
- **A1** negatively skewed cao
- **Answer: Negatively skewed, since $Q2 - Q1$ (4) is greater than $Q3 - Q2$ (3).**

Question 6

- (a) **B1** valid sampling method named (e.g. simple random sample, systematic sample, stratified sample by year group/tutor group)
- (a) **B1** valid reason linked to representativeness or avoiding bias, matched to the method given
- (a) **Answer: e.g. A simple random sample, using a random number generator on the full list of students, so that every student has an equal chance of being chosen and the sample is not biased.**
- (b) **B1** 9 cao
- (b) **Answer: 9**
- (c) **M1** $5+6+6+7+8+9+10+12+15+20+56 (=154)$, oe divided by 11
- (c) **A1** 14 cao
- (c) **Answer: 14**
- (d) **B1** mean (14) is greater than the median (9)
- (d) **B1** so the data is positively skewed, which supports the statement: most journey times are short, with one or two much longer journeys stretching a tail to the right and pulling the mean up
- (d) **Answer: Mean (14) > median (9), so the data is positively skewed - this supports the statement, since most journeys are short but one long journey (56 minutes) pulls the mean up.**

Question 7

- **B1** the student is incorrect cao
- **B1** explains that a mean greater than the median is generally a sign of positive skew (a tail towards higher values), not negative oe
- **B1** states that negative skew would instead show the mean being less than the median oe
- **Answer: The student is incorrect. A mean greater than the median generally indicates positive skew, since it shows a tail stretching towards higher values; negative skew would instead show the mean being less than the median.**

Question 8

- (a) **B1** Yes - mean (52) < median (60) is consistent with negative skew cao
- (a) **Answer: Yes, negative skew**
- (b) **M1** $60-38=22$ and $65-60=5$, oe
- (b) **A1** $22 > 5$, confirming negative skew (the lower half of the data is more spread out), which agrees with the student's conclusion
- (b) **Answer: Confirmed: negative skew ($22 > 5$), which agrees with the student.**
- (c) **B1** e.g. using two methods checks that they agree, which reduces the risk of being misled by a single unusual value or measure (oe)
- (c) **Answer: Checking two methods lets you confirm they agree, reducing the risk that one unusual value or measure alone gives a misleading conclusion.**