



KS3.M-S9 Comparing Distributions

AQA KS3.M-S9 · Calculators not allowed · about 75 minutes

Name: _____ Date: ____ / ____ / ____

Answer ALL questions. Show all your working.

- 1** Write down the mean of the five test scores: 6, 9, 8, 7, 10.

(2)

- 2** State the median of the list: 12, 7, 9, 14, 11.

(1)

- 3** Give the mode of these shoe sizes: 4, 6, 5, 6, 7, 6, 5.

(1)

- 4** Calculate the range of these weights in kg: 48, 52, 45, 50, 47.

(2)

- 5** A frequency table shows times (minutes) taken by pupils to run 1 km.
Times: 4, 5, 6 with frequencies 3, 4, 3 respectively. Work out the mean time.

(2)

- 6** Class A has mean score 72 from 10 pupils. Class B has mean 68 from 10 pupils. Which class has the higher average score?

(2)

- 7** Two sets of exam scores are shown.
Set X: 55, 57, 58, 59, 60, 61, 62
Set Y: 40, 50, 60, 70, 80, 90, 100
State which set is more likely to be skewed and explain how the mean and median compare for that skew (one or two sentences).

(3)

- 8** Compare the spread of these two small data sets using range and give a one-sentence conclusion.
A: 3, 4, 5, 6, 7
B: 1, 5, 5, 6, 9

(3)

- 9** A cumulative frequency for test marks out of 20 is given as: marks 0-5: 3, 6-10: 9, 11-15: 15, 16-20: 20 (cumulative). Estimate the median mark.

(2)

- 10** A data set has mean 30 and range 12. If every value is increased by 5, what are the new mean and new range?

(3)

- 11** Two teachers compare their class test results out of 20.

Teacher P: scores 16, 15, 16, 17, 14, 16, 15, 16

Teacher Q: scores 10, 12, 14, 15, 16, 18, 19, 1

Find mean and range for both classes and state which class has the higher average and which is more consistent (smaller spread). Justify your answer with numbers.

(4)

- 12** A boxplot shows the five-number summary: min = 12, Q1 = 15, median = 18, Q3 = 22, max = 30.

State the interquartile range and give one sentence about how spread and skew are shown by this boxplot.

(3)

- 13** Which of these two datasets likely has the larger standard deviation? Give a short reason.

Set C: 49, 50, 51, 50, 49

Set D: 30, 40, 50, 60, 70

(3)

- 14** Grouped data: class widths 0-9, 10-19, 20-29 with frequencies 2, 5, 3. Use the midpoints to estimate the mean and then state which class (this or a class with mean 14 and same total frequency) has the higher average.

(4)

-
- 15** Two groups of runners have these times in minutes.

Group E: 7, 8, 7, 9, 8

Group F: 5, 9, 7, 8, 7

Calculate each group's IQR and say which group is more consistent using IQR as the measure.

(4)

-
- 16** Stretch question. A teacher compares two test score distributions out of 50.

Class G: 42, 43, 44, 45, 46, 47, 48, 49

Class H: 20, 30, 35, 40, 45, 46, 48, 50

Calculate mean and median for both classes, compute IQR for both, then recommend which class performed better overall and which was more consistent. Justify with the calculated statistics.

(5)

- 17** State the 90th percentile (the value above which 10% of data lie) for this sorted list of 10 values: 2, 4, 6, 8, 10, 12, 14, 16, 18, 20.

(2)

-
- 18** Explain in one sentence why the median is often a better measure than the mean for a distribution with outliers.

(2)

Mark scheme · KS3.M-S9 Comparing Distributions

Question 1

- **M1** sum of scores shown or stated ($6+9+8+7+10 = 40$) or method of adding
- **A1** mean = 8 cao
- **Answer: 8.0 (or 8)**

Question 2

- **B1** median = 11 cao
- **Answer: 11**

Question 3

- **B1** mode = 6 cao
- **Answer: 6**

Question 4

- **M1** correct subtraction shown: max - min ($52 - 45$)
- **A1** range = 7 kg cao
- **Answer: 7 kg**

Question 5

- **M1** correct use of ($4*3 + 5*4 + 6*3$) and total frequency 10 or equivalent
- **A1** mean = 5.0 minutes cao
- **Answer: 5.0 minutes (or 5)**

Question 6

- **M1** comparison of means stated ($72 > 68$) or statement of Class A higher
- **A1** Class A has the higher average cao
- **Answer: Class A**

Question 7

- **M1** identify Set Y as skewed (values spread with high outliers) or Set X as symmetric
- **M1** state effect on mean (pulled towards tail) or median less affected
- **A1** clear conclusion: for skewed Set Y mean is greater than median (because high values pull mean up) cao
- **Answer: Set Y is skewed; mean is greater than median because the high values pull the mean up**

Question 8

- **M1** correct range for A shown ($7 - 3 = 4$) or for B ($9 - 1 = 8$)
- **M1** both ranges stated ($A = 4$, $B = 8$)
- **A1** conclusion that B has larger spread cao
- **Answer: Range A = 4, Range B = 8; B has the larger spread**

Question 9

- **M1** recognise median position is $1/2$ of 20 = 10th value and locate it in 11-15 group (cumulative reaches 9 then 15)
- **A1** estimate median = about 10 or 11 (choosing 10 as the lowest value in 11-15 group may be accepted) cao
- **Answer: Approximately 10 or 11**

Question 10

- **M1** state new mean = old mean + 5
- **M1** state new range is unchanged (range = 12) or show reasoning
- **A1** new mean = 35 and new range = 12 cao
- **Answer: Mean = 35, Range = 12**

Question 11

- **M1** correct mean for P or correct method shown (sum = 125, mean = 15.625)
- **M1** correct mean for Q or correct method shown (sum = 105, mean = 13.125)
- **M1** range for P = $17 - 14 = 3$ and range for Q = $19 - 1 = 18$ stated
- **A1** conclusion: P has higher average ($15.625 > 13.125$) and P is more consistent (smaller range) cao
- **Answer: P mean = 15.625, range = 3; Q mean = 13.125, range = 18. So P has the higher average and is more consistent.**

Question 12

- **M1** correct IQR calculation $Q3 - Q1$ shown ($22 - 15$)
- **M1** $IQR = 7$ stated
- **A1** one-sentence comment linking box length and whisker lengths to spread or noting slight positive skew because upper whisker longer than lower one
- **Answer: IQR = 7; the upper whisker is longer than the lower so the data show a slight positive skew and more spread above the median.**

Question 13

- **M1** identify D as having larger spread (values further from mean)
- **M1** reference to values being more widely spread or giving approximate ranges (C range 2, D range 40)
- **A1** conclusion that D has the larger standard deviation with brief reason cao
- **Answer: Set D; values are spread much more widely so standard deviation will be larger**

Question 14

- **M1** use midpoints correctly (4.5, 14.5, 24.5) and multiply by frequencies or equivalent
- **M1** correct total frequency = 10 stated
- **M1** correct estimated mean computed: $(4.5 \times 2 + 14.5 \times 5 + 24.5 \times 3) / 10 = (9 + 72.5 + 73.5) / 10 = 155 / 10$
- **A1** estimated mean = 15.5 and conclusion that this class (15.5) has higher average than class with mean 14 cao
- **Answer: Estimated mean = 15.5; this class has the higher average than the class with mean 14**

Question 15

- **M1** correct ordered lists shown or used for each group
- **M1** correct $Q1$ and $Q3$ for one group (E ordered 7,7,8,8,9 $\rightarrow Q1=7, Q3=8.5$)
- **M1** correct $Q1$ and $Q3$ for the other group (F ordered 5,7,7,8,9 $\rightarrow Q1=6, Q3=8.5$)
- **A1** $IQR E = 1.5, IQR F = 2.5$ so Group E is more consistent (smaller IQR) cao
- **Answer: IQR E = 1.5, IQR F = 2.5; Group E is more consistent (smaller IQR)**

Question 16

- **M1** correct mean for G shown (sum 364, mean = 45.5) or method
- **M1** correct mean for H shown (sum 314, mean = 39.25) or method
- **M1** correct medians found (G median = 45.5, H median = 42.5)
- **M1** correct IQRs computed (G $Q1=43.5, Q3=47.5$ so $IQR=4$; H $Q1=32.5, Q3=47$ so $IQR=14.5$) or equivalent
- **A1** clear recommendation: G performed better on average (mean and median higher) and G is more consistent (smaller IQR) cao
- **Answer: G mean = 45.5, median = 45.5, IQR = 4; H mean = 39.25, median = 42.5, IQR = 14.5; recommend G: higher performance and more consistent.**

Question 17

- **M1** correct method to find position: $0.9 \times (n) = 0.9 \times 10 = 9$ th value or appropriate percentile method
- **A1** 90th percentile = 18 cao
- **Answer: 18**

Question 18

- **M1** state that outliers pull the mean towards them
- **A1** state that median is less affected by extreme values so better represents centre for skewed data
- Answer: **Because outliers pull the mean towards them, the median is less affected by extreme values and so better represents the centre for skewed distributions.**