

Representation of Data

Cambridge International AS & A Level - Mathematics 9709
Paper 5: Probability & Statistics 1 - Exam-style question bank

Original Steps to Conquer practice questions in the style of recent 9709 Paper 5 examinations - not past Cambridge questions.
Marks are shown in brackets []. Give non-exact answers to 3 significant figures (1 d.p. for angles) and show full working.

Questions

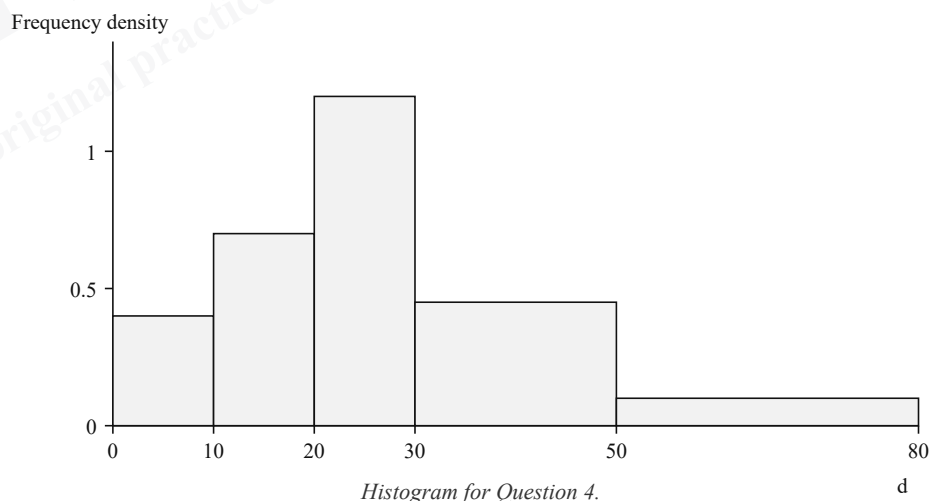
- 1 The times, in minutes, taken by 12 students to complete a task are shown below. Find the mean, median, range and interquartile range. [6]

18	22	25	26	29	31	33	35	37	40	44	50
----	----	----	----	----	----	----	----	----	----	----	----

- 2 The grouped frequency table shows the distances, d km, travelled by 35 cyclists. Using midpoints, estimate the mean and standard deviation of d . [7]

Distance d	0-10	10-20	20-30	30-50	50-80
Frequency	4	7	12	9	3

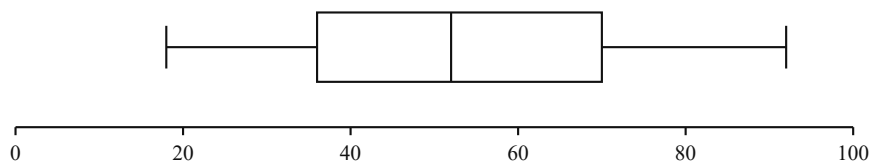
- 3 For 40 observations of x , coded values are defined by $y = \frac{x - 50}{5}$. Given that $\sum y = 24$ and $\sum y^2 = 196$, find the mean and standard deviation of the original x values. [6]
- 4 The histogram below represents the distances travelled by a group of cyclists. Use the histogram to find the frequency in the class $20 \leq d < 30$ and the total frequency represented. [5]



- 5 A cumulative frequency graph for 50 observations passes through the points shown in the table. Estimate the median, lower quartile, upper quartile and interquartile range. [7]

x	10	20	30	40	50	60
Cumulative frequency	4	11	25	38	46	50

- 6 The box-and-whisker plot below summarises the marks of a group of students. State the median and the interquartile range, and comment on the skewness of the distribution. [5]



Box-and-whisker plot for Question 6.

- 7 For data set A , $n = 20$, $\sum x = 680$ and $\sum x^2 = 23860$. For data set B , $n = 30$, $\sum x = 1110$ and $\sum x^2 = 42570$. Find the mean and standard deviation of the combined set of 50 values. [7]

- 8 The grouped data below have mean 20. Find the value of k . [5]

Class	0-10	10-20	20-30	30-40
Frequency	5	k	12	3

- 9 The stem-and-leaf diagram shows the ages of 11 competitors. Find the mode, median and range. [5]

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

Key: 1 | 2 represents 12 years.

- 10 A teacher wants to compare two classes. Class X has median 64 and interquartile range 18. Class Y has median 69 and interquartile range 10. Compare the two classes in context. [4]

Worked solutions

Full methods with final answers. Other valid methods may also earn full marks.

- 1 The total is 390, so the mean is $\frac{390}{12} = 32.5$ minutes.

There are 12 values, so the median is the mean of the 6th and 7th values.

$$\text{Median} = \frac{31 + 33}{2} = 32 \text{ minutes.}$$

$$\text{Range} = 50 - 18 = 32 \text{ minutes.}$$

$$Q_1 = \frac{25 + 26}{2} = 25.5 \quad \text{and} \quad Q_3 = \frac{37 + 40}{2} = 38.5$$

$$\text{Interquartile range} = 38.5 - 25.5 = 13 \text{ minutes.}$$

- 2 Use midpoints 5, 15, 25, 40, 65 .

$$\sum fx = 4(5) + 7(15) + 12(25) + 9(40) + 3(65) = 980$$

$$\text{Estimated mean} = \frac{980}{35} = 28.0 \text{ km.}$$

$$\sum fx^2 = 4(25) + 7(225) + 12(625) + 9(1600) + 3(4225) = 36250$$

$$\text{Variance} = \frac{36250}{35} - 28^2 = 251.714 \dots$$

$$\text{Estimated standard deviation} = 15.9 \text{ km, to 3 significant figures.}$$

- 3 $\bar{y} = \frac{24}{40} = 0.6$.

$$\text{Since } x = 5y + 50, \bar{x} = 5(0.6) + 50 = 53.0$$

$$\text{Var}(y) = \frac{196}{40} - (0.6)^2 = 4.54$$

$$\text{Var}(x) = 5^2 \text{Var}(y) = 25(4.54) = 113.5$$

$$\text{The standard deviation of } x \text{ is } \sqrt{113.5} = 10.7 \text{ to 3 significant figures.}$$

- 4 In a histogram, frequency = class width $\times$ frequency density.

$$\text{For } 20 \leq d < 30, \text{ frequency} = 10 \times 1.2 = 12$$

$$\text{The other frequencies are } 10(0.4) = 4, 10(0.7) = 7, 20(0.45) = 9 \text{ and } 30(0.1) = 3$$

$$\text{Total frequency} = 4 + 7 + 12 + 9 + 3 = 35$$

- 5 The median is the 25th value, which corresponds to $x = 30$.

The lower quartile is the 12.5th value.

$$\text{Between cumulative frequencies 11 and 25, interpolate: } Q_1 = 20 + \frac{12.5 - 11}{25 - 11} \times 10 = 21.1$$

The upper quartile is the 37.5th value.

$$\text{Between cumulative frequencies 25 and 38, interpolate: } Q_3 = 30 + \frac{37.5 - 25}{38 - 25} \times 10 = 39.6$$

$$\text{The interquartile range is } 39.6 - 21.1 = 18.5$$

- 6 The median is 52.

$$\text{The interquartile range is } 70 - 36 = 34$$

The upper whisker, from 70 to 92, is longer than the lower whisker, from 36 to 18.

The distribution is slightly positively skewed.

- 7 For the combined set, $\sum x = 680 + 1110 = 1790$.

$$\text{Combined mean} = \frac{1790}{50} = 35.8$$

$$\sum x^2 = 23860 + 42570 = 66430$$

$$\text{Variance} = \frac{66430}{50} - (35.8)^2 = 46.96$$

$$\text{Standard deviation} = \sqrt{46.96} = 6.85 \quad \text{to 3 significant figures.}$$

- 8 Use midpoints 5, 15, 25, 35 .

$$\sum fx = 5(5) + 15k + 12(25) + 3(35) = 430 + 15k$$

The total frequency is $20 + k$.

$$\frac{430 + 15k}{20 + k} = 20$$

$$430 + 15k = 400 + 20k, \text{ so } k = 6$$

- 9 The values are 12, 15, 18, 20, 20, 24, 27, 31, 33, 36, 42 .

The mode is 20.

There are 11 values, so the median is the 6th value, 24.

The range is $42 - 12 = 30$.

- 10 Class Y has the higher median, so a typical student in class Y scored more highly.

Class Y has the smaller interquartile range, so the middle half of its marks are less spread out.

Class X has lower typical performance and greater variation among the middle half of the marks.