Exercise walk-through

6.1
Sampling, averages and spread (S1–S4)

AQA · GCSE · Mathematics

Histograms, cumulative frequency and box plots (H)

- **Histograms.** In a **histogram** the area of each bar, not its height, shows the frequency.
- When the **class intervals** have different widths, the height is the **frequency density**:
- **Worked example 6 (H).** Classes $0 \leq x < 10$, $10 \leq x < 20$ and $20 \leq x < 40$ have frequencies 20, 30 and 20.
- The **class width** of the last class is 20, so the frequency densities are $20 \div 10 = 2$, $30 \div 10 = 3$ and $20 \div 20 = 1$.

$$\text{frequency density} = \frac{\text{frequency}}{\text{class width}}$$

Sources for this lesson

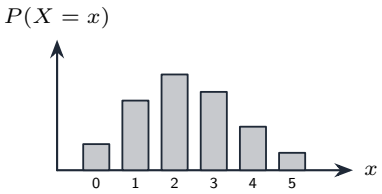
- AQA, *GCSE Mathematics (8300) specification*, subject content section 3.6 (Statistics).
- "Landon, 1,293,669; Roosevelt, 972,897", *The Literary Digest* (31 October 1936).
- P. Squire, "Why the 1936 Literary Digest poll failed", *Public Opinion Quarterly* 52 (1988), pages 125 to 133.
- AQA, *GCSE Mathematics* Higher Papers 1, 2 and 3 (June 2024), the **pie chart**, frequency table and **box plot** questions, adapted.

Common errors (2)

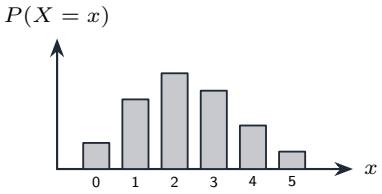
- (H) Drawing **histogram** bars with height equal to frequency when the class widths differ.
- (H) Plotting **cumulative frequency** at the midpoint instead of the upper end of the class.
- Comparing only averages; a comparison also needs a measure of spread.
- Saying that **correlation** proves cause.

Exercises 2.1·2.2

2.1 Find the mean of $2, 4, 4, 6, 9$. [1]

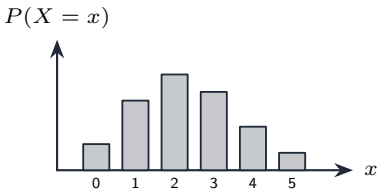


2.2 Values 1, 2 and 3 have frequencies 2, 5 and 3. Find the mean and the mode. [2]



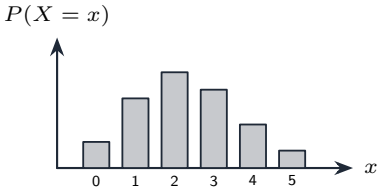
Exercises 2.1·2.2

2.1 Find the mean of $2, 4, 4, 6, 9$. [1]



2.1 $(2 + 4 + 4 + 6 + 9) \div 5 = 25 \div 5 = 5$.

2.2 Values 1, 2 and 3 have frequencies 2, 5 and 3. Find the mean and the mode. [2]



2.2 $\sum fx = 2 + 10 + 9 = 21$ over 10 values: mean $= 2.1$; mode $= 2$.

Exercise 2.3

2.3 In a random sample of 80 students, 24 walk to school. Estimate the number of walkers in a school of 600. [2]

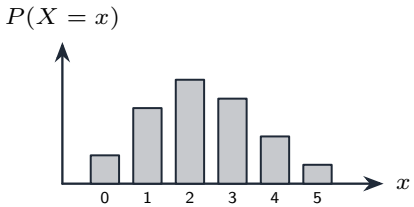
Exercise 2.3

2.3 In a random sample of 80 students, 24 walk to school. Estimate the number of walkers in a school of 600. [2]

2.3 $\frac{24}{80} = 0.3$; $600 \times 0.3 = 180$ walkers.

Setup for exercise 2.4

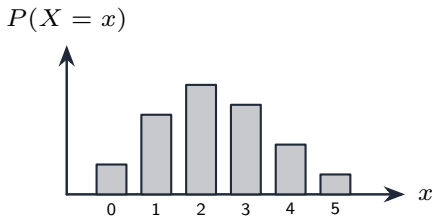
2.4 For the data $2, 4, 4, 6, 24$, find the mean and the median, and explain why they differ. [2]



2.5 Classify each as discrete or continuous: number of siblings; travel time to school. [2]

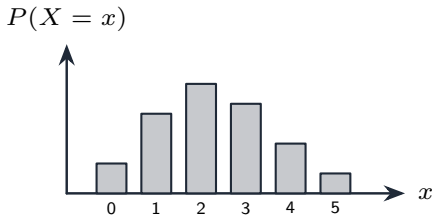
Exercise 2.4

2.6 Classes $0 \leq x < 10$, $10 \leq x < 20$, $20 \leq x < 30$ have frequencies 2, 5, 3. Estimate the mean and write down the modal class. [3]



Exercise 2.4

2.6 Classes $0 \le x < 10$, $10 \le x < 20$, $20 \le x < 30$ have frequencies 2, 5, 3. Estimate the mean and write down the modal class. [3]



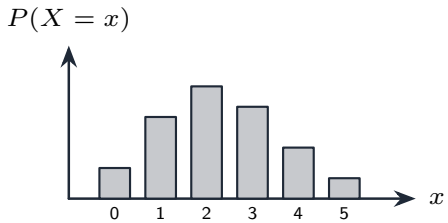
2.4 Mean $= 40 \div 5 = 8$; median $= 4$. The single large value 24 pulls the mean up while the median resists it.

2.5 Siblings: discrete (counted); travel time: continuous (measured, even if recorded to whole minutes).

2.6 Midpoints 5, 15, 25: $\sum fm = 10 + 75 + 75 = 160$, mean $\approx 160 \div 10 = 16$; modal class $10 \le x < 20$.

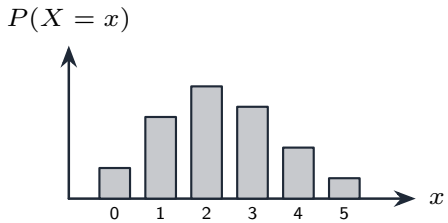
Exercise 3.1

3.1 *(June 2024 2H Q13, adapted.)* A table of lengths shows 15 values of 7.2, 18 values of 7.6 and 7 values of 8. Work out the mean. [3]



Exercise 3.1

3.1 *(June 2024 2H Q13, adapted.)* A table of lengths shows 15 values of 7.2, 18 values of 7.6 and 7 values of 8. Work out the mean. [3]



3.1 $\sum fx = 15\times7.2+18\times7.6+7\times8 = 108+136.8+56 = 300.8$; mean $= 300.8\div40 = 7.52$.

Setup for exercise 3.2

3.2 *(June 2024 2H Q16, adapted.)* Draw a box plot for: shortest value 2.4 cm, lower quartile 4 cm, median 6 cm, upper quartile 7 cm, longest value 9 cm. [3]

3.3 *(June 2024 1H Q16, adapted.)* The swimming club's median age is lower than the cycling club's, and its inter-quartile range is higher. Write two sentences comparing the clubs' ages. [2]

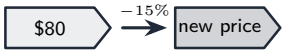


3.4 *(Original, Higher.)* A cumulative-frequency graph of 80 observations reads $Q_1 = 12$ and $Q_3 = 21$. State where these readings were taken on the cumulative axis, and work out the inter-quartile range. [3]



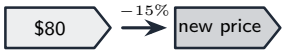
Exercise 3.2

3.5 *(Original.)* An optional online poll finds that 70% of 400 respondents cycle to school. Explain why this does not establish that 70% of the school cycles. [2]



Exercise 3.2

3.5 *(Original.)* An optional online poll finds that 70% of 400 respondents cycle to school. Explain why this does not establish that 70% of the school cycles. [2]



3.2 A rectangle from 4 to 7 with a median line at 6, whiskers to 2.4 and 9 – five plotted points.

3.3 The swimming club's average age is lower (its median is lower); its ages are less consistent (its inter-quartile range is higher).

3.4 Quartiles read at cumulative frequency $80 \div 4 = 20$ and $3 \times 80 \div 4 = 60$; $IQR = 21 - 12 = 9$.

3.5 The poll is voluntary, so keen cyclists respond more – selection bias that more responses cannot remove; the sample is not random.

Local review guidance: exact $\sum fx$ products before dividing (approximated products forfeit the mark); comparisons state an average AND a spread in context; sampling limits named as selection bias or non-response, not "not enough people". This is authored classroom guidance.