4.1 Data summaries, histograms and interpretation

Name: _____ Class: _____ Date: _____ **Total: 13 marks**

KEY WORDS frequency density

Objective

Data summaries, histograms and interpretation. Original classroom practice for IB SL; not an official IB paper.

1 Key knowledge and worked examples

■ Can one average tell the whole story?

- Two groups have the same median but different spread. One summary cannot describe both location and consistency.
- This lesson studies **frequency density** : Frequency divided by class width, used as histogram height.

■ The method

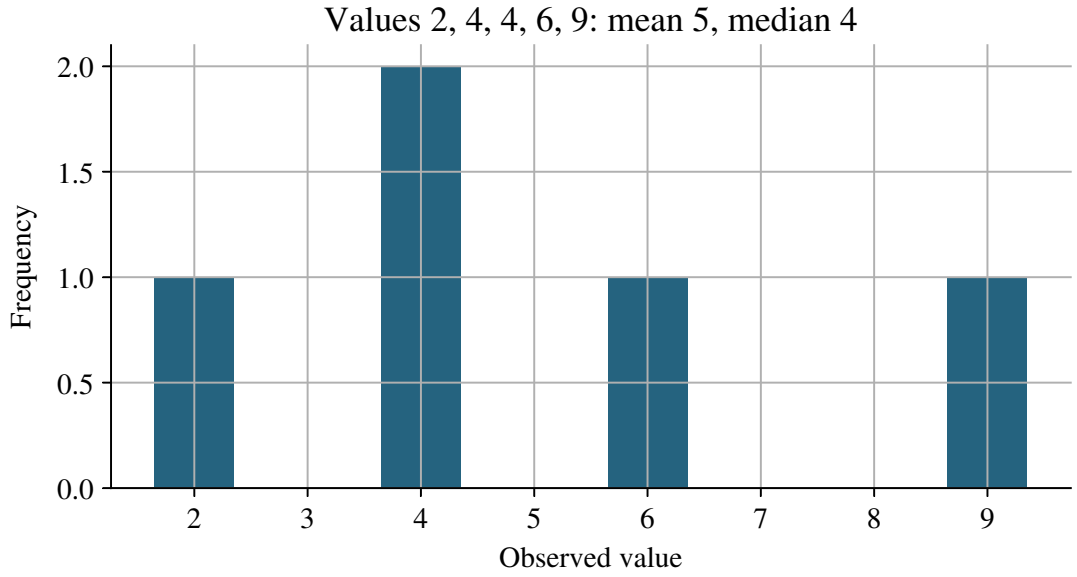
- Compare an appropriate average and spread in context.
- A histogram uses area for frequency, so height=frequency/class width.
- Grouped estimates assume representative values within intervals.

■ A worked case

$$\overline{x} = \frac{\sum x_i}{n}, \qquad \text{density} = \frac{\text{frequency}}{\text{class width}}$$

Example:

A class from 10 to 20 with frequency 30 has density 30/10=3. A class from 20 to 40 with frequency 20 has density 20/20=1. Its wider bar must not be mistaken for a larger density. For values 2,4,4,6,9, the median is 4 and mean is 5.



Where it goes wrong

- The tallest histogram bar need not contain the most observations.
- A grouped mean is an estimate.
- Correlation does not prove causation, and extrapolation extends beyond the observed range.

Not true: A histogram bar's height always equals its frequency, even with unequal class widths. With unequal class widths the height is frequency density; frequency is the area of the bar.

Applying it

- Choose a display that fits the data type.
- Give both a numerical comparison and what it means for the population; do not infer more precision than the sample supports.

In your course

- Analysis and Approaches SL, the course first assessed in 2021.
- Paper 1 is non-calculator: keep fractions, surds and π exact and show each step. On calculator papers, still write the expression before its value.

Key:

- **Frequency density**: Frequency divided by class width, used as histogram height.
- Compare an appropriate average and spread in context.
- Watch for: the tallest histogram bar need not contain the most observations.

■ **Write from a mean of 6**

- A class has frequency 24 and width 6. The data values 3, 5, 5, 8 and 9 are a separate list.
- Find the frequency density and the mean. Histogram height is density when widths differ.

2 Practice

2.1 Find the density for frequency 30 and class width 10. Show the relation, substitution and units. [2]

2.2 Find the mean of 2,4,4,6,9. Show the relation, substitution and units. [2]

2.3 Find the median of 2,4,4,6,9. Show the relation, substitution and units. [2]

2.4 Decide whether this claim is true and justify: Frequency density means Frequency divided by class width, used as histogram height. [2]

2.5 In a histogram with unequal class widths, which bar contains the most observations? Choose one answer.

- **A** The one with the largest AREA, not necessarily the tallest
- **B** Always the tallest bar, because height shows how many observations it holds
- **C** Always the widest bar, because a wider class must contain more values [1]

3 Exam-style questions

3.1 Classes 0–10 and 10–30 have frequencies 20 and 30. Find histogram densities and compare bar heights with frequencies; can the exact mean be recovered? [4]