Exercise walk-through

2.6
Statistics – Numerical Measures of Central Locations

GAC Mathematics

Statistics: collecting and displaying data

- A **population** is everyone you want to describe; a **sample** is who you actually measure.
- To generalise, the **sample** must be **representative**: chosen so that every part of the **population** has a fair chance of being included.
- **Qualitative data** are categories; **quantitative data** measure amounts, either **discrete** (counted) or **continuous** (measured).
- **Bias** enters through who you ask, who answers and how the question is worded. The *Literary Digest* poll had all three problems: its lists favoured the wealthy, and the 24% who replied were a self-selected group.

$$D_1 = \frac{20}{10} = 2 \qquad D_2 = \frac{20}{20} = 1$$

Statistics: measures of variability

- The **range** is the largest value minus the smallest, so one extreme value can change it a lot.
- The **interquartile range (IQR)** is $Q_3 - Q_1$, the spread of the middle half. A **box plot** shows the minimum, quartiles, **median** and maximum.
- The **standard deviation** measures spread using squared distances from the **mean**. For a whole **population**, $\sigma = \sqrt{\sum (x-\mu)^2/n}$; for a sample estimate, $s = \sqrt{\sum (x-\bar{x})^2/(n-1)}$. It has the same units as the data.
- Two data sets can share a mean and be completely different: a city with every day at 20 °C and one swinging between 0 °C and 40 °C have the same mean temperature.

$$\text{range} = 12 - 2 = 10 \qquad IQR = 8 - 4 = 4$$

Sources for this lesson

- GAC010 Mathematics II course outline (sequences, finance, probability and statistics).
- *Labour Law of the People's Republic of China* (1994), Article 44, on overtime pay.
- National Bureau of Statistics of China, *Statistical Communiqué on the 2023 National Economic and Social Development* (per capita disposable income, **mean** and **median**).

Exercises 2.1·2.2

2.1 Find the mean of 4, 6, 8 and 10. [2]



2.2 Find the median and mode of 2, 3, 3, 7, 9. [2]



Exercises 2.1·2.2

2.1 Find the mean of 4, 6, 8 and 10. [2]



2.1

Add the values and divide by the count.

$$\bar{x} = \frac{\sum x}{n} = \frac{4 + 6 + 8 + 10}{4} = 7$$

Mean 7.

2.2 Find the median and mode of 2, 3, 3, 7, 9. [2]



2.2

The middle of the five ordered values is 3, so median 3. The most frequent value is also 3, so mode 3.

Exercise 2.3

2.3 A frequency table has value 1 with frequency 2, value 2 with frequency 3, and value 4 with frequency 1. Find the mean. [3]



Exercise 2.3

2.3 A frequency table has value 1 with frequency 2, value 2 with frequency 3, and value 4 with frequency 1. Find the mean. [3]



2.3

Multiply each value by its frequency.

$$\bar{x} = \frac{\sum fx}{\sum f}$$

$$\bar{x} = \frac{1(2) + 2(3) + 4(1)}{2 + 3 + 1}$$

Correct weighted total and count.

$$\bar{x} = \frac{12}{6} = 2$$

Mean 2.

Exercise 2.4

2.4 The ordered data are 2, 4, 6, 8. Find the median. Then replace 8 by 80 and find the new median and new mean. [3]



Exercise 2.4

2.4 The ordered data are 2, 4, 6, 8. Find the median. Then replace 8 by 80 and find the new median and new mean. [3]



2.4

Original median $m = (4 + 6)/2 = 5$. New median remains 5 because the middle two values are unchanged.

$$\bar{x}_{new} = \frac{2 + 4 + 6 + 80}{4} = 23$$

New mean 23.

Exercise 3.1

3.1 Group A has 8 learners with mean score 65. Group B has 12 learners with mean score 75. (a) Find the combined mean. (b) A single missing score is then added, bringing the combined mean for 21 learners to 70. Find that score. (c) Explain why averaging 65 and 75 equally is not the correct first calculation. [6]



Exercise 3.1

3.1 Group A has 8 learners with mean score 65. Group B has 12 learners with mean score 75. (a) Find the combined mean. (b) A single missing score is then added, bringing the combined mean for 21 learners to 70. Find that score. (c) Explain why averaging 65 and 75 equally is not the correct first calculation. [6]



3.1

(a) Recover totals from each group.

$$T = n_A\bar{x}_A + n_B\bar{x}_B = 8(65) + 12(75) = 1420$$

Correct total.

Full solution on next slide

Full solution 3.1

3.1

(a) Recover totals from each group.

$$T = n_A\bar{x}_A + n_B\bar{x}_B = 8(65) + 12(75) = 1420$$

Correct total.

$$\bar{x} = \frac{T}{n_A + n_B} = \frac{1420}{20} = 71$$

Combined mean 71.

(b) Subtract the old total from the required new total.

$$x = 21(70) - 1420 = 50$$

Missing score 50. Equal weighting would give each group the same influence despite unequal learner counts.

Exercise 3.2

3.2 Five supplied monthly incomes are 1000, 1100, 1200, 1300 and 6400 units. (a) Find the mean and median. (b) A report says, “Most people earn the mean.” Explain what the summaries do and do not show. (c) If the highest income changes to 8400 while the others stay fixed, find the new mean and median. Which measure changes more, and why? [8]



Exercise 3.2

3.2 Five supplied monthly incomes are 1000, 1100, 1200, 1300 and 6400 units. (a) Find the mean and median. (b) A report says, “Most people earn the mean.” Explain what the summaries do and do not show. (c) If the highest income changes to 8400 while the others stay fixed, find the new mean and median. Which measure changes more, and why? [8]



3.2

(a) Mean uses all five incomes.

$$\bar{x} = \frac{1000 + 1100 + 1200 + 1300 + 6400}{5} = 2200$$

Mean 2200. Median 1200.

(b) Nobody in the supplied list earns 2200; the mean is a total-per-person summary, not

Full solution on next slide

Full solution 3.2

3.2

(a) Mean uses all five incomes.

$$\bar{x} = \frac{1000 + 1100 + 1200 + 1300 + 6400}{5} = 2200$$

Mean 2200. Median 1200.

(b) Nobody in the supplied list earns 2200; the mean is a total-per-person summary, not a statement that most incomes equal it. The median marks the ordered middle and can describe the centre without claiming all or most incomes equal it.

(c) Recalculate the total after the change.

$$\bar{x}_{new} = \frac{1000 + 1100 + 1200 + 1300 + 8400}{5} = 2600$$

New mean 2600 and median still 1200. The mean rises by 400 because it includes the changed magnitude, while the ordered middle remains unchanged. These five invented records do not represent an entire population.