

SECTION A Quick Recall (*answer in one or two lines*)

Q1. A bag has 4 red balls and 6 blue balls. You draw two balls *without* replacement. What is the probability that both are red? [2 marks]

Solution: Total balls = 10. Drawing without replacement:

$$P(\text{both red}) = \frac{4}{10} \times \frac{3}{9} = \frac{12}{90} = \frac{2}{15} \approx 0.133$$

Q2. Two events A and B satisfy $P(A) = 0.5$, $P(B) = 0.4$, and $P(A \cap B) = 0.2$. Are A and B independent? Justify your answer in *one sentence*. [2 marks]

Solution: Check: $P(A) \cdot P(B) = 0.5 \times 0.4 = 0.20 = P(A \cap B)$. Since $P(A \cap B) = P(A) \cdot P(B)$, events A and B **are independent**.

Q3. The ages (in years) of five students are: {18, 19, 19, 21, 23}. Without a calculator, find the **mean**, **median**, and **mode**. [3 marks]

Solution: Mean: $\frac{18 + 19 + 19 + 21 + 23}{5} = \frac{100}{5} = 20$

Median: Ordered data: 18, 19, **19**, 21, 23 \Rightarrow middle value = **19**

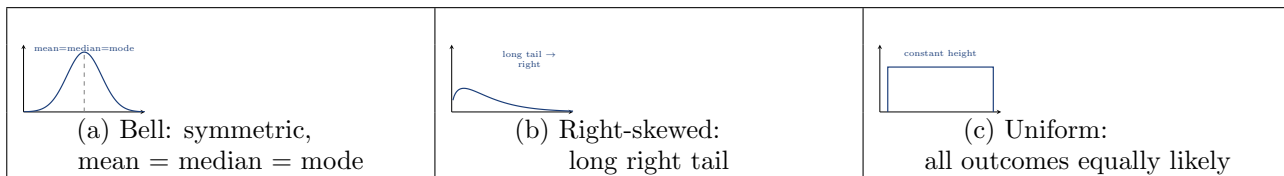
Mode: **19** (appears twice, more than any other value)

SECTION B Think & Explain

Q4. A fair coin is tossed 100 times and lands **heads 70 times**. Your friend says: “*The next 30 tosses must mostly be tails to even things out.*” Is your friend correct? Explain the concept behind your answer. [3 marks]

Solution: No, the friend is incorrect. This reasoning is known as the **Gambler’s Fallacy**. Each coin toss is an *independent* event — the coin has no memory of past outcomes. The probability of heads on any future toss remains exactly $\frac{1}{2}$, regardless of previous results. The “evening out” that the Law of Large Numbers describes happens because the *proportion* converges to 0.5 over a very large number of trials, not because future outcomes compensate for past ones.

Q5. Sketch a rough shape for each of the following distributions and label its key feature: (a) Symmetric bell-shaped (b) Right-skewed (c) Uniform. (Axes not required — freehand sketches are fine.) [3 marks]



SECTION C Problem Solving

Q6. A random variable X takes values $\{1, 2, 3\}$ with probabilities $P(X = 1) = 0.2$, $P(X = 2) = 0.5$, $P(X = 3) = 0.3$. (a) Find $E[X]$. (b) Find $E[X^2]$. (c) Hence find $\text{Var}(X)$ using $\text{Var}(X) = E[X^2] - (E[X])^2$. [4 marks]

Solution: (a) $E[X] = 1(0.2) + 2(0.5) + 3(0.3) = 0.2 + 1.0 + 0.9 = \mathbf{2.1}$

(b) $E[X^2] = 1^2(0.2) + 2^2(0.5) + 3^2(0.3) = 0.2 + 2.0 + 2.7 = \mathbf{4.9}$

(c) $\text{Var}(X) = E[X^2] - (E[X])^2 = 4.9 - (2.1)^2 = 4.9 - 4.41 = \mathbf{0.49}$

Q7. A student claims: “If the standard deviation of a dataset is zero, all values must be identical.” Prove or disprove this claim using the definition of variance. [3 marks]

Solution: The claim is **TRUE**. Proof:

By definition, for a dataset $\{x_1, x_2, \dots, x_n\}$ with mean \bar{x} :

$$\text{Var}(X) = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2$$

If $\text{SD} = 0$, then $\text{Var}(X) = 0$. Since each term $(x_i - \bar{x})^2 \geq 0$ (squares are non-negative), and their sum equals zero, every term must individually equal zero:

$$(x_i - \bar{x})^2 = 0 \quad \Rightarrow \quad x_i = \bar{x} \quad \text{for all } i.$$

Therefore all values are identical (each equals the mean \bar{x}). □

Q8. Bonus — Think deeper: You measure the heights of students in two classes. Class A has mean 170 cm and SD 5 cm. Class B has mean 170 cm and SD 15 cm. Which class would you predict has more students outside the range 160–180 cm? Why? No calculation needed — reason from the meaning of standard deviation. [5 (bonus) marks]

Solution: Class B will have more students outside [160, 180] cm.

The standard deviation measures the *spread* or dispersion of data around the mean. Both classes have the same mean (170 cm), but Class B has a much larger SD (15 cm vs. 5 cm), meaning its heights are far more spread out.

For Class A (SD = 5): the range 160–180 cm spans ± 2 SDs from the mean. By the empirical rule, roughly 95% of students fall within this range, leaving only about 5% outside it.

For Class B (SD = 15): the same range spans only ± 0.67 SDs from the mean. Far fewer students are captured in this narrower band (in SD units), so a much larger proportion — roughly 50% — fall outside the 160–180 cm range.

In essence, greater spread \Rightarrow more values in the tails \Rightarrow more students outside any fixed interval.