GAC010 Mathematics II: Probability, Statistics and Finance

GAC010-aligned practice with invented teaching conditions; the centre's brief sets the actual assessment.

What this module is, and how it is marked

- use arithmetic and geometric sequences and their sums;
- calculate pay, **deductions**, discounts, interest, **depreciation** and loan balances;
- find probabilities with tables, trees and counting;

What this module is, and how it is marked (2)

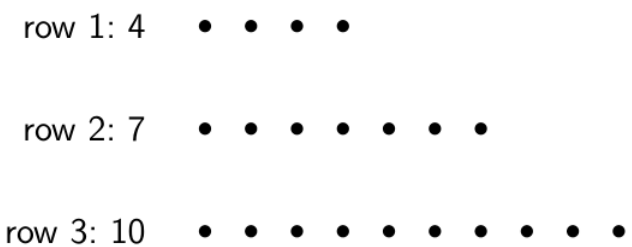
- collect, display and summarise data, and choose the right average and spread;
- use the **normal distribution** and z-scores.

Sequences and series

- The sums are $S_n = \frac{n}{2}(2a + (n-1)d)$ and $S_n = \dfrac{a(1-r^n)}{1-r}$ for $r \neq 1$.
- A geometric **series** with $|r| < 1$ has a **sum to infinity** : $S_\infty = \dfrac{a}{1-r}$.
- Keep a term separate from a total: u_n is one value, S_n is the sum of the first n.

$$u_n = a + (n-1)d \qquad \text{and} \qquad u_n = ar^{n-1}$$

Figure: Sequences and series



- Rows of 4, 7 and 10: a constant difference of 3
- Match each part of the diagram to the text.

Finance: earning money and purchasing goods

- **Gross pay** is what you earn; **net pay** is what arrives after **deductions** such as tax and social insurance.
- **Simple interest** is $I = Prt$: interest on the original amount only.
- **Compound interest** is $A = P(1 + r)^n$: interest is added to the balance and then earns interest itself, so the balances form a **geometric sequence**.

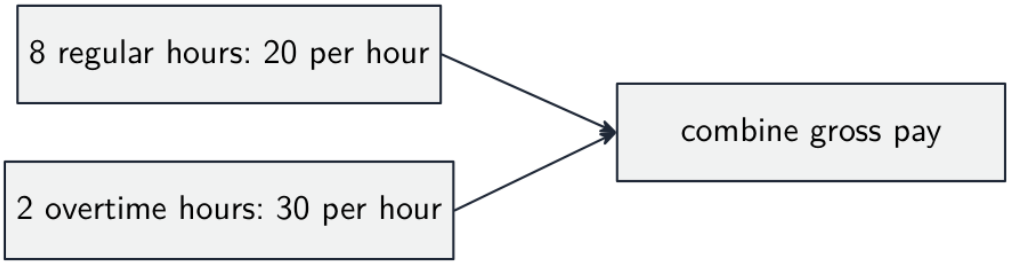
$$G = 8(20) + 2(1.5)(20) = 160 + 60 = ¥220$$

Finance: earning money and purchasing goods (2)

- **Depreciation** is a fall in value. A fixed-percentage model uses $V_n = V_0(1-r)^n$.
- A **discount** of 20% multiplies by 0.8. Two successive discounts of 20% and 10% multiply by $0.8 \times 0.9 = 0.72$, which is a 28% reduction, not 30%.

$$N = G - D = 220 - 15 = ¥205$$

Figure: Finance: earning money and purchasing goods



- A pay record: 8 regular hours and 2 overtime hours at 1.5 times the rate
- Match each part of the diagram to the text.

Finance: investing money and loans

- The **present value** of a future amount, at a constant rate r per period, is $PV = \dfrac{FV}{(1+r)^n}$: how much you would need to invest now to have FV later.
- An **annuity** is a **series** of equal payments at regular intervals.
- In a **loan repayment** schedule, each period interest is added to the balance and the payment is taken off. The order matters.
- An **APR** is an annual rate. If interest is added monthly, the rate per month is the **APR** divided by 12, and the effective yearly growth is higher: a nominal 12% with 1% a month gives $(1.01)^{12} - 1 \approx 12.68\%$.

$$B_1 = 1000(1.02) - 300 = ¥720$$

Figure: Finance: investing money and loans



- One month of a loan: ¥1,000 plus 2% interest, then a ¥300 payment
- Match each part of the diagram to the text.

Probability: key concepts

- For equally likely outcomes, $P(A) = \dfrac{\text{favourable outcomes}}{\text{total outcomes}}$.
- Events are **mutually exclusive** when they cannot both happen: $P(A \text{ or } B) = P(A) + P(B)$.
- Events are **independent** when one does not affect the other: $P(A \text{ and } B) = P(A) \times P(B)$.

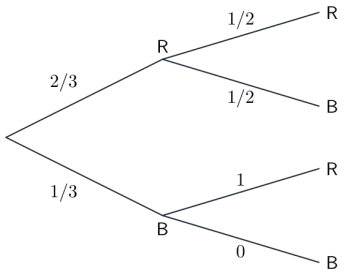
$$\frac{365}{365} \times \frac{364}{365} \times \cdots \times \frac{343}{365} \approx 0.493$$

Probability: key concepts (2)

- The **complement** rule, $P(\text{not } A) = 1 - P(A)$, is often the fastest route to an answer containing "at least one".
- **Counting techniques** give the number of outcomes: a **permutation** counts ordered selections, a **combination** counts unordered ones.
- For events that overlap, $P(A \cup B) = P(A) + P(B) - P(A \cap B)$; conditional probability is $P(A \mid B) = \dfrac{P(A \cap B)}{P(B)}$.

$$P(RR) = P(R_1) \times P(R_2 \mid R_1) = \frac{2}{3} \times \frac{1}{2} = \frac{1}{3}$$

Figure: Probability: key concepts



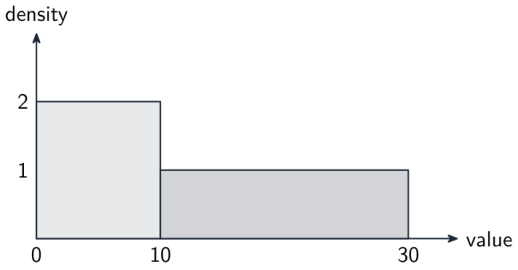
- A two-draw tree for a bag with two red and one blue counter, without replacement, labels conditional second-draw probabilities.
- Match each part of the diagram to the text.

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.

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

Figure: Statistics: collecting and displaying data



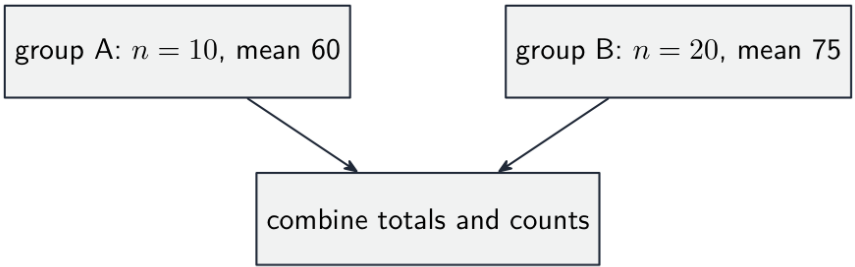
- A histogram has intervals 0 to 10 and 10 to 30 with densities 2 and 1; their areas each represent frequency 20.
- Read the axes and units first.

Statistics: measures of central location

- The **mean** is the total divided by the count. It uses every value, including extreme ones.
- The **median** is the middle value when the data are in order. Extreme values do not move it.
- The **mode** is the most frequent value; it also works for categories.

$$\bar{x} = \frac{10(60) + 20(75)}{10 + 20} = \frac{2100}{30} = 70$$

Figure: Statistics: measures of central location



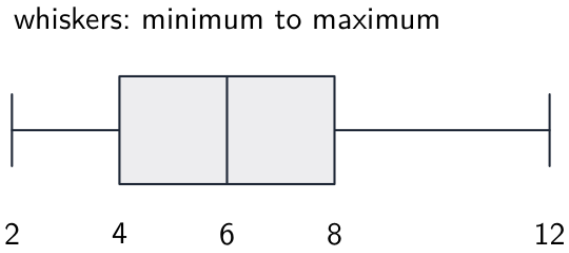
- Two groups contribute different totals: 10 learners with **mean** 60, and 20 learners with mean 75.
- Match each part of the diagram to the text.

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.

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

Figure: Statistics: measures of variability



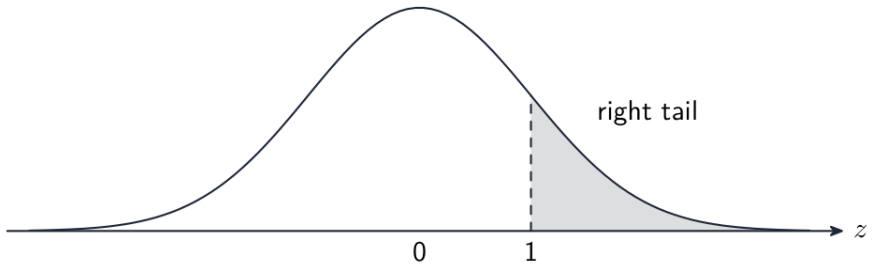
- A minimum-to-maximum **box plot** marks minimum 2, first quartile 4, **median** 6, third quartile 8 and maximum 12.
- Read the axes and units first.

Distributions: the normal distribution

- The **normal distribution** is a symmetric, bell-shaped model described by its **mean** μ and **standard deviation** σ. Many measurements, such as adult heights, are close to normal.
- About 68% of values lie within one standard deviation of the mean, 95% within two, and 99.7% within three.
- A **z-score** says how many standard deviations a value is from the mean: $z = \dfrac{x - \mu}{\sigma}$. It lets you compare values on different scales.
- Use the cumulative area $\Phi(z) = P(Z \le z)$ and symmetry: $\Phi(-z) = 1 - \Phi(z)$.

$$z = \frac{60 - 50}{10} = 1 \qquad P(X > 60) = 1 - \Phi(1) = 1 - 0.8413 = 0.1587$$

Figure: Distributions: the normal distribution



- A standard-normal curve marks z equal to 1 and shades the right-tail region beyond it.
- Read the axes and units first.

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**).

Sources for this lesson (2)

- P. Squire, "Why the 1936 Literary Digest Poll Failed", *Public Opinion Quarterly* 52 (1988).
- W. Sartorius von Waltershausen, *Gauss zum Gedächtniss* (1856), the source of the Gauss story.

Common errors

- Using n instead of $n-1$ in $a+(n-1)d$ or ar^{n-1}.
- Giving a term when the question asks for a sum, or the reverse.
- Applying **simple interest** where the question compounds, or the reverse.

Common errors (2)

- Adding successive percentage discounts: 20% then 10% off is 28% off, not 30%.
- Taking off a loan payment before adding interest when the schedule says interest first.
- Multiplying probabilities of dependent events as if they were **independent**; forgetting to change the denominator when drawing without replacement.

Common errors (3)

- Thinking a bigger **sample** is always better: a biased sample stays biased however large it is.
- Finding the **median** without first putting the data in order.
- Mixing the **population** ($\div n$) and sample ($\div (n-1)$) standard deviations.

Key terms

- **sequence**: an ordered list of numbers.
- **series**: the sum of the terms of a **sequence**.
- **common difference**: the fixed amount added each step in an **arithmetic sequence**.

Key terms (2)

- **common ratio**: the fixed number multiplied each step in a **geometric sequence**.
- **net pay**: pay after **deductions** such as tax.
- **present value**: the amount needed now to reach a future amount at a given rate.

Key terms (3)

- **mutually exclusive**: describes events that cannot happen together.
- **independent**: describes events where one does not change the probability of the other.
- **representative**: describes a **sample** that fairly reflects the **population**.

Key terms (4)

- **median**: the middle value of ordered data.
- **interquartile range**: the spread of the middle half of the data, $Q_3 - Q_1$.
- **standard deviation**: a measure of how far values typically lie from the **mean**.
- **z-score**: the number of standard deviations a value lies from the mean.

Check yourself

- Answer each question before reading the answers.
- 1.
- For the **arithmetic sequence** $7, 11, 15, \ldots$, find the 20th term and the sum of the first 20 terms. 2.
- For the **geometric sequence** $3, 6, 12, \ldots$, find the 8th term and the sum of the first 8 terms. 3.