

GRADE 11 STUDY GUIDE | 2024

Statistics and Probability

Mathematics

1. Measures of Central Tendency

Mean: $\bar{x} = (\text{sum of all values}) / n$ **Median:** middle value when data sorted in order**Mode:** value that appears most often

- For grouped data, use midpoint of each class $\times$ frequency to estimate mean
- Outliers affect the MEAN most; median is more robust
- Bimodal data has two modes

2. Measures of Dispersion (Spread)

Range = Max - Min**Variance:** $\sigma^2 = \text{Sum}[(x - \bar{x})^2] / n$ **Standard Deviation:** $\sigma = \sqrt{\text{variance}}$

Standard deviation measures how spread out data is from the mean. A small SD means data is clustered near the mean.

KEY: About 68% of data falls within 1 SD of the mean (in a normal distribution). About 95% within 2 SDs. About 99.7% within 3 SDs.

3. Five-Number Summary and Box-and-Whisker

Min | Q1 (lower quartile) | Q2 (median) | Q3 (upper quartile) | Max**IQR** = Q3 - Q1**Outlier:** value $< Q1 - 1.5 \times \text{IQR}$ OR value $> Q3 + 1.5 \times \text{IQR}$

Box-and-whisker plot: the box spans Q1 to Q3, a line marks Q2, whiskers extend to min and max (or to fences if outliers exist).

4. Histograms, Ogives and Frequency Polygons

- Histogram: bars touching, showing frequency vs class intervals
- Frequency polygon: connect midpoints of histogram tops with straight lines
- Cumulative frequency curve (ogive): plot cumulative frequency vs upper class boundary. S-shaped curve.

-

From the ogive: read off median (50%), Q1 (25%), Q3 (75%)

5. Scatter Plots and Correlation

A scatter plot shows the relationship between two variables. The line of best fit (regression line) is drawn to best represent the trend.

- Positive correlation: as x increases, y increases
- Negative correlation: as x increases, y decreases
- No correlation: no pattern visible
- Strong correlation: points clustered close to the line
- Weak correlation: points spread widely around the line

6. Probability

$$P(A) = n(A) / n(S)$$

$$P(A \text{ or } B) = P(A) + P(B) - P(A \text{ and } B)$$

$$P(\text{not } A) = 1 - P(A)$$

- Mutually exclusive events: $P(A \text{ and } B) = 0$; $P(A \text{ or } B) = P(A) + P(B)$
- Independent events: $P(A \text{ and } B) = P(A) \times P(B)$
- $P(A|B) = P(A \text{ and } B) / P(B)$ [conditional probability]
- Tree diagrams and two-way tables are useful for complex probability problems

KEY: Check if events are independent: $P(A \text{ and } B) = P(A) \times P(B)$ must hold. Check if mutually exclusive: $P(A \text{ and } B) = 0$ must hold. Events can be neither, one, or the other -- but NOT both (unless $P(A)=0$ or $P(B)=0$).

7. Counting Principles

Fundamental Counting Principle: $n_1 \times n_2 \times n_3 \times \dots \times n_k$ (for ordered choices)

Factorial: $n! = n \times (n-1) \times \dots \times 2 \times 1$

Permutations: $P(n;r) = n! / (n-r)!$ [order MATTERS]

Combinations: $C(n;r) = n! / [r!(n-r)!]$ [order does NOT matter]

Example: How many 3-letter codes can be formed from the letters A,B,C,D,E with no repetition?

Order matters, no repetition => Permutations

$$P(5;3) = 5! / (5-3)! = 120 / 2 = 60 \text{ codes}$$