

Statistics and probability

International Baccalaureate · IB Diploma · Mathematics: Analysis and Approaches · SL

Scope and prerequisites

This handout covers selected SL topics for the course first assessed in 2021; it is not the whole syllabus.

Prerequisites: read the stated quantities and units, and use arithmetic and the model conditions below.

The examples are original and fictional: they are not real measurements or real student investigations.

Data summaries, histograms and interpretation

■ 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

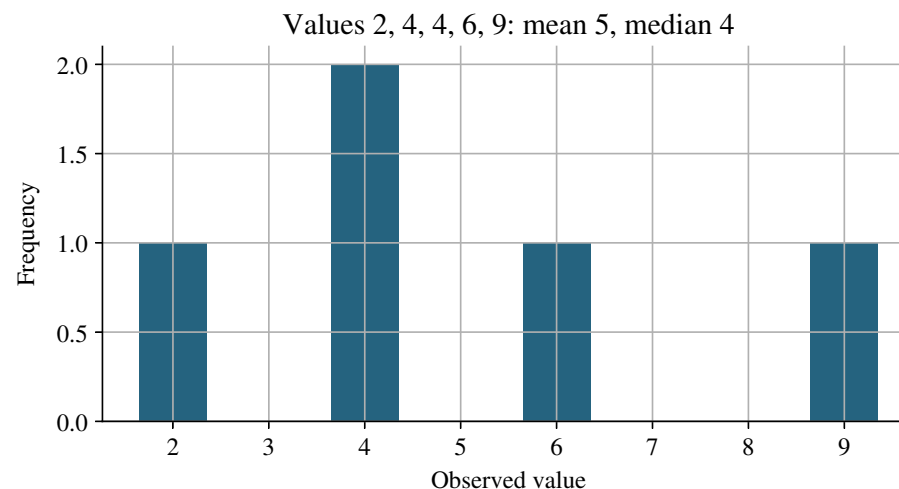
- Compare an appropriate average and spread in context.
- A histogram uses area for frequency, so height=frequency/class width.
- Grouped estimates assume representative values within intervals.

■ A worked case

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

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.



For 2, 4, 4, 6, 9: $\text{mean} = 25 \div 5 = 5$, $\text{median} = 4$

■ 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.

Probability, trees and conditional reasoning

■ What changes after the first draw?

- A bag has 3 red and 2 blue counters. Taking two without replacement changes the chance of the second colour.
- This lesson studies **conditional probability** : The probability of an event after restricting the sample space to a stated condition.

■ The method

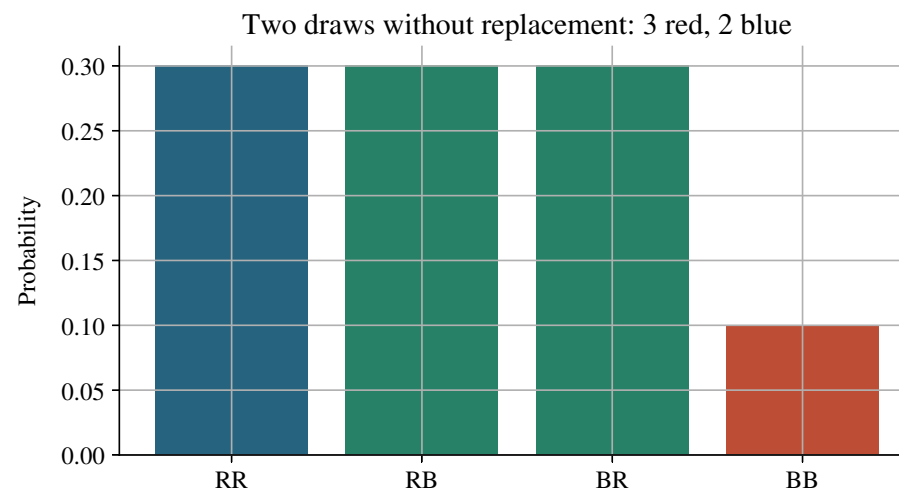
- Multiply along a tree branch and add disjoint branches.
- With replacement, the composition stays fixed.
- Conditional probability is $P(A \text{ given } B) = P(A \cap B) / P(B)$, for $P(B) > 0$.

■ A worked case

$$P(A | B) = \frac{P(A \cap B)}{P(B)}, \quad P(B) > 0$$

Example:

Without replacement, $P(\text{two red}) = 3/5 \times 2/4 = 3/10$. $P(\text{one of each}) = 3/5 \times 2/4 + 2/5 \times 3/4 = 3/5$. If $P(A \cap B) = 0.12$ and $P(B) = 0.3$, $P(A \text{ given } B) = 0.4$.



Without replacement: $P(RR) = 3/5 \times 2/4 = 0.3$ and $P(BB) = 2/5 \times 1/4 = 0.1$

■ Where it goes wrong

- Mutually exclusive means no overlap; independent means that knowing one event does not change the other's probability.
- Two disjoint events with positive probability are not independent.

Not true: Mutually exclusive events with positive probabilities must be independent. Mutually exclusive events with positive probabilities are dependent: if one happens, the other cannot, so $P(A \text{ and } B) = 0 \neq P(A)P(B)$.

■ Applying it

- A two-way table makes the restricted denominator visible.
- Before using a product $P(A)P(B)$, justify independence from the context or the supplied information.

■ 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:

- **Conditional probability:** The probability of an event after restricting the sample space to a stated condition.
- Multiply along a tree branch and add disjoint branches.
- Watch for: mutually exclusive means no overlap; independent means that knowing one event does not change the other's probability.

■ Write from a conditional probability of 0.30

- P of A and B is 0.18 and P of B is 0.60.
- Find P of A given B. Two disjoint events with positive probability are not independent.

Binomial, normal and Poisson models

■ What makes a count predictable?

- A quality inspector counts defective items. The number is random, but a model can describe its likely range.
- This lesson studies **expected value** : The probability-weighted mean of a random variable.

■ The method

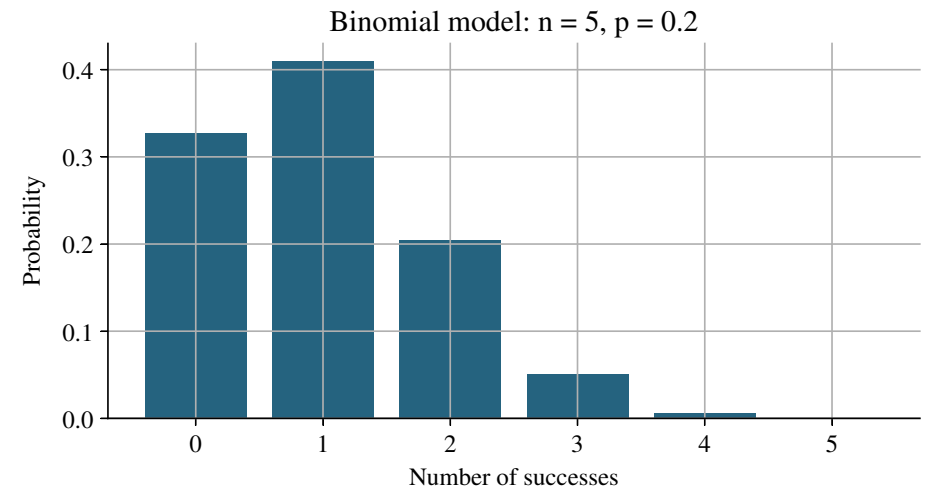
- A binomial model needs fixed n, independent trials, two outcomes and constant p.
- $E(X) = np$ and $Var(X) = np(1-p)$.
- For a normal model use $z = (x - \mu) / \sigma$.
- A Poisson model describes counts with a constant rate and appropriate independence assumptions.

■ A worked case

$$P(X = k) = \binom{n}{k} p^k (1-p)^{n-k}, \quad X \sim B(n, p)$$

Example:

For X binomial(5,0.2), $P(X=0) = 0.8^5 = 0.32768$, $E(X) = 1$ and $Var(X) = 0.8$. For a normal quantity with $\mu = 100, \sigma = 15$, the value 130 has $z = 2$. A Poisson mean of 3 per hour gives mean 6 over two hours.



$X \sim B(5, 0.2)$: $P(X = 1) \approx 0.41$ is the largest bar; $P(X = 5)$ is tiny

■ Where it goes wrong

- Not every count is binomial: changing p or dependence can invalidate it.
- For a continuous variable, the probability of one exact value is zero.
- Continuity correction matters when approximating a discrete distribution by a normal one.

Not true: Every count of successes has a binomial distribution regardless of dependence. The binomial model needs independent trials with a fixed probability; dependent trials break it.

■ Applying it

- Write the event as an inequality before using calculator distribution functions.
- Distinguish $P(X < k)$, $P(X \leq k)$ and a tail complement.
- State assumptions in context.

■ In your course

- Analysis and Approaches SL, the course first assessed in 2021.

Key:

- **Expected value:** The probability-weighted mean of a random variable.
- A binomial model needs fixed n, independent trials, two outcomes and constant p.

- Watch for: not every count is binomial: changing p or dependence can invalidate it.

■ **Write from a z-score of 1.5**

- A value 95 comes from a distribution with mean 80 and standard deviation 10.
- Find the z-score. A binomial count needs independent trials.

■ **Write from an expected count of 5**

- A binomial model has $n = 20$ trials with $p = 0.25$.
- The expected count is $np = 5$ – probability-weighted, never guaranteed.

■ **Write from a z-score of 1**

- A value 62 comes from a distribution with mean 55 and standard deviation 7.
- The z-score is $7 \div 7 = 1$ – one standard deviation above the mean.