

Statistics and probability · supported focus

International Baccalaureate · IB Diploma · Mathematics:
Applications and Interpretation · HL

Scope and prerequisites

Supported HL focus. First assessment 2021; current through 2028. First-assessment-2029 course is separate.. Remaining guide, assessment and practical requirements retain their recorded holds.

Prerequisites: read the stated quantities and units, use arithmetic and the model conditions below. Each lesson develops its own method before independent transfer.

These are original or explicitly fictional teaching examples, not actual measurements or completed assessed learner 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.

■ Choose the mathematical structure

- 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.
- State the allowed inputs and units before calculating. An equation should express the relationship, not just record a calculator entry.

■ Work through a checked case

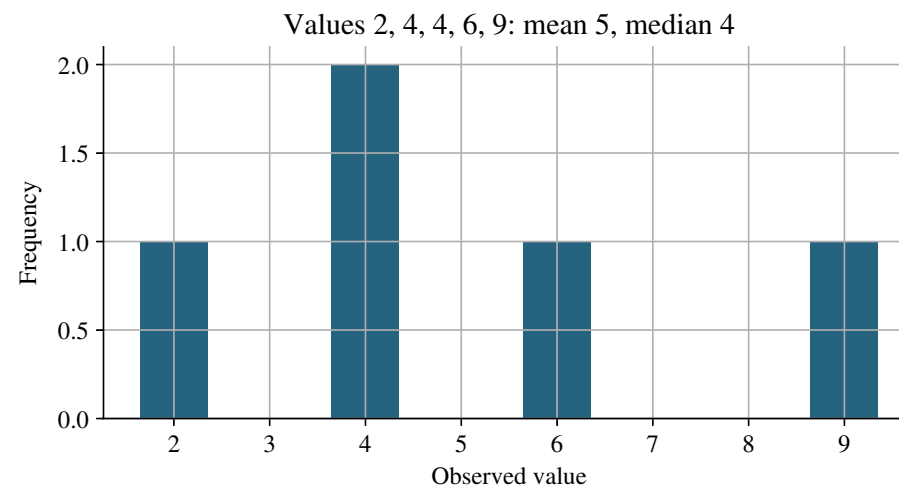
$$\bar{x} = \frac{\sum x_i}{n}, \quad \text{density} = \frac{\text{frequency}}{\text{class width}}$$

- Check the result against the starting quantities. Substitute into the original relation, or compare the graph and numerical answer where appropriate.

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.



■ Test a tempting shortcut

- 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.
- When a shortcut fails, identify the assumption it breaks. Keep an exact value until the requested final rounding.

Warn:

A histogram bar's height always equals its frequency, even with unequal class widths. This claim is false. Explain which definition or assumption it violates.

■ Interpret a new situation

- 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.
- A complete solution gives the mathematical result and explains what it means. Check that it is possible in the stated context.

■ Use this in your course

- Current first-assessment-2021 Applications and Interpretation HL. This is authored concept support; the full guide is needed to certify every objective.
- Give the method before the final answer, and use the paper’s calculator and formula rules. Review a wrong answer by locating the first invalid step.

Key:
Frequency divided by class width, used as histogram height. Choose the relationship, show the method, check its assumptions and interpret the result.

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.

■ Choose the mathematical structure

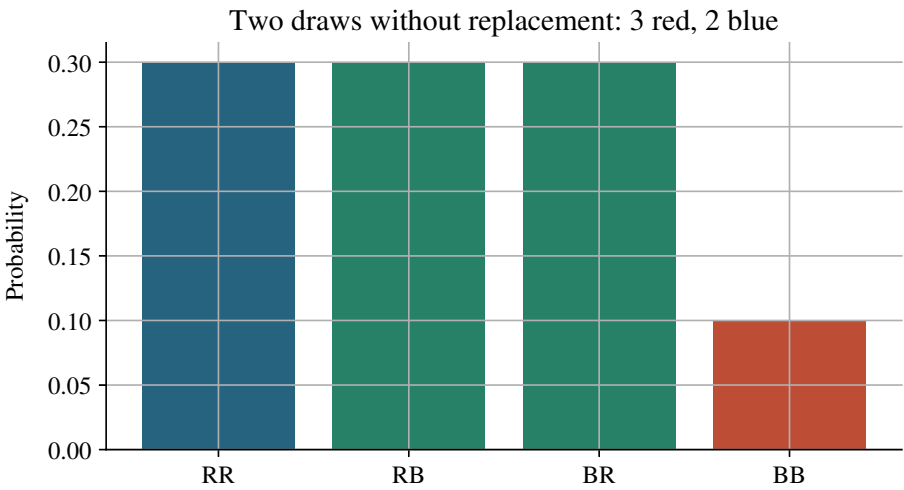
- Multiply along a tree branch and add disjoint branches. With replacement, the composition stays fixed. Conditional probability is $P(A \text{ given } B) = P(A \cap B) / P(B)$, for $P(B) > 0$.
- State the allowed inputs and units before calculating. An equation should express the relationship, not just record a calculator entry.

■ Work through a checked case

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

- Check the result against the starting quantities. Substitute into the original relation, or compare the graph and numerical answer where appropriate.

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 $P(B) = 0.3$, $P(A \text{ given } B) = 0.4$.



■ Test a tempting shortcut

- 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.
- When a shortcut fails, identify the assumption it breaks. Keep an exact value until the requested final rounding.

Warn:
Mutually exclusive events with positive probabilities must be independent. This claim is false. Explain which definition or assumption it violates.

■ Interpret a new situation

- 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.
- A complete solution gives the mathematical result and explains what it means. Check that it is possible in the stated context.

■ Use this in your course

- Current first-assessment-2021 Applications and Interpretation HL. This is authored concept support; the full guide is needed to certify every objective.
- Give the method before the final answer, and use the paper’s calculator and formula rules. Review a wrong answer by locating the first invalid step.

Key:

The probability of an event after restricting the sample space to a stated condition. Choose the relationship, show the method, check its assumptions and interpret the result.

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.

■ Choose the mathematical structure

- A binomial model needs fixed n , independent trials, two outcomes and constant p . $E(X)=np$ and $\text{Var}(X)=np(1-p)$. For a normal model use $z=(x-\mu)/\sigma$. A Poisson model describes counts with a constant rate and appropriate independence assumptions.
- State the allowed inputs and units before calculating. An equation should express the relationship, not just record a calculator entry.

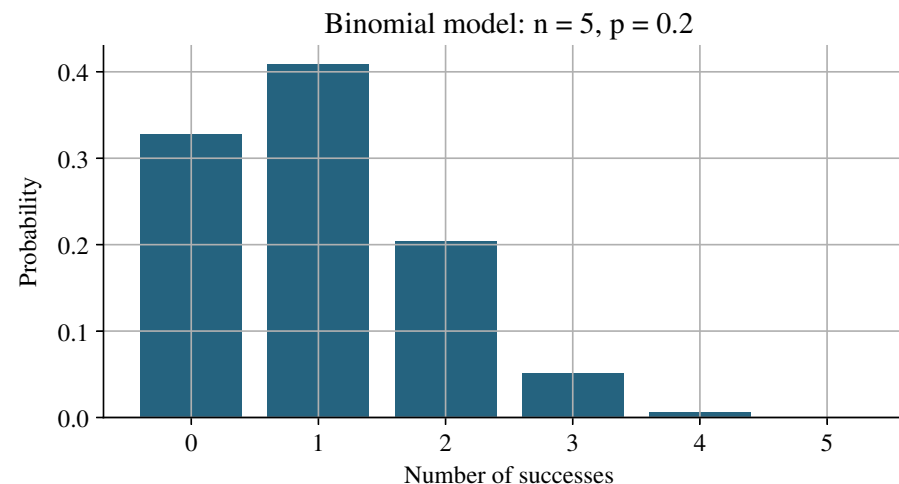
■ Work through a checked case

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

- Check the result against the starting quantities. Substitute into the original relation, or compare the graph and numerical answer where appropriate.

Example:

For X binomial(5,0.2), $P(X=0)=0.8^5=0.32768$, $E(X)=1$ and $\text{Var}(X)=0.8$. 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.



■ Test a tempting shortcut

- 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.
- When a shortcut fails, identify the assumption it breaks. Keep an exact value until the requested final rounding.

Warn:

Every count of successes has a binomial distribution regardless of dependence. This claim is false. Explain which definition or assumption it violates.

■ Interpret a new situation

- 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.
- A complete solution gives the mathematical result and explains what it means. Check that it is possible in the stated context.

■ Use this in your course

- Current first-assessment-2021 Applications and Interpretation HL. This is authored concept support; the full guide is needed to certify every objective.
- Give the method before the final answer, and use the paper's calculator and formula rules. Review a wrong answer by locating the first invalid step.

Key:

The probability-weighted mean of a random variable. Choose the relationship, show the method, check its assumptions and interpret the result.

Hypothesis testing and contextual conclusions

- **Could chance explain the result?**
 - A factory claims that only 10% of items are defective. A sample contains more defects, but chance alone may explain some difference.
 - This lesson studies **significance level** : The chosen probability threshold for rejecting a null hypothesis.

- **Choose the mathematical structure**
 - State H_0 and H_1 in population parameters before inspecting the outcome. Calculate the appropriate tail probability under H_0 . Reject H_0 when the evidence meets the specified significance rule; otherwise say there is insufficient evidence.
 - State the allowed inputs and units before calculating. An equation should express the relationship, not just record a calculator entry.

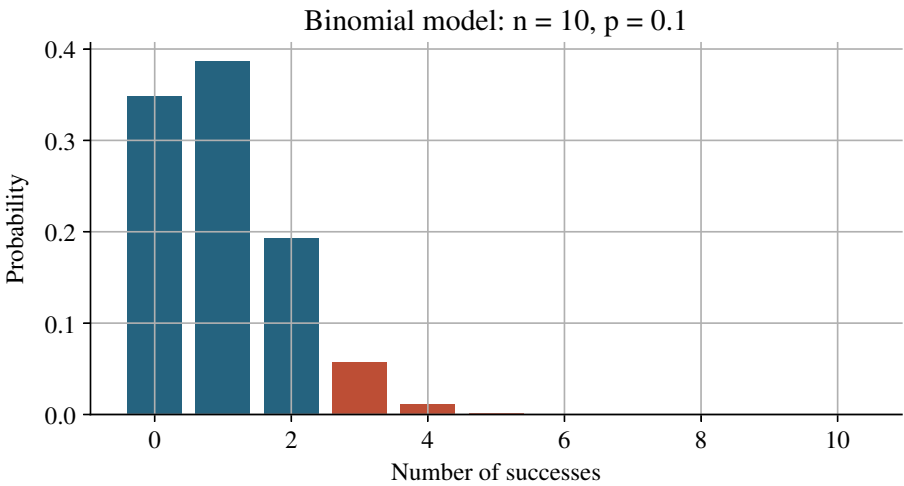
- **Work through a checked case**

$$H_0 : p = p_0, \quad H_1 : p > p_0$$

- Check the result against the starting quantities. Substitute into the original relation, or compare the graph and numerical answer where appropriate.

Example:

For $H_0:p=0.1$ and $H_1:p>0.1$ with $n=10$, observing at least 3 defects has probability $1-(0.9^{10}+10\times0.1\times0.9^9+45\times0.1^2\times0.9^8)\approx0.070191$. At 5%, this is insufficient evidence that the defect rate exceeds 10%.



- **Test a tempting shortcut**
 - Failing to reject H_0 is not proof that H_0 is true. Choose the tail from H_1 , not from whichever tail gives a small result. Statistical significance does not measure the practical size of an effect.
 - When a shortcut fails, identify the assumption it breaks. Keep an exact value until the requested final rounding.

Warn:

Failing to reject a null hypothesis proves it is true. This claim is false. Explain which definition or assumption it violates.

- **Interpret a new situation**
 - Finish with a sentence about the population and the original claim. State the model’s assumptions and consider whether the sampling procedure supports them.
 - A complete solution gives the mathematical result and explains what it means. Check that it is possible in the stated context.

- **Use this in your course**
 - Current first-assessment-2021 Applications and Interpretation HL. This is authored concept support; the full guide is needed to certify every objective.
 - Give the method before the final answer, and use the paper’s calculator and formula rules. Review a wrong answer by locating the first invalid step.

Key:
The chosen probability threshold for rejecting a null hypothesis. Choose the relationship, show the method, check its assumptions and interpret the result.

Confidence intervals, chi-square and regression

- **How uncertain is the sample mean?**
 - Two samples give different means. We need a statement about sampling uncertainty, not a promise that one sample equals the population.
 - This lesson studies **confidence interval** : An interval calculated by a procedure designed to cover the population parameter at a stated long-run rate.

- **Choose the mathematical structure**
 - For a normal mean with known σ , a 95% interval is sample mean $\pm 1.96\sigma/\sqrt{n}$. For chi-square, sum (observed-expected)²/expected with the correct degrees of freedom. Choose a test that fits the data and assumptions.
 - State the allowed inputs and units before calculating. An equation should express the relationship, not just record a calculator entry.

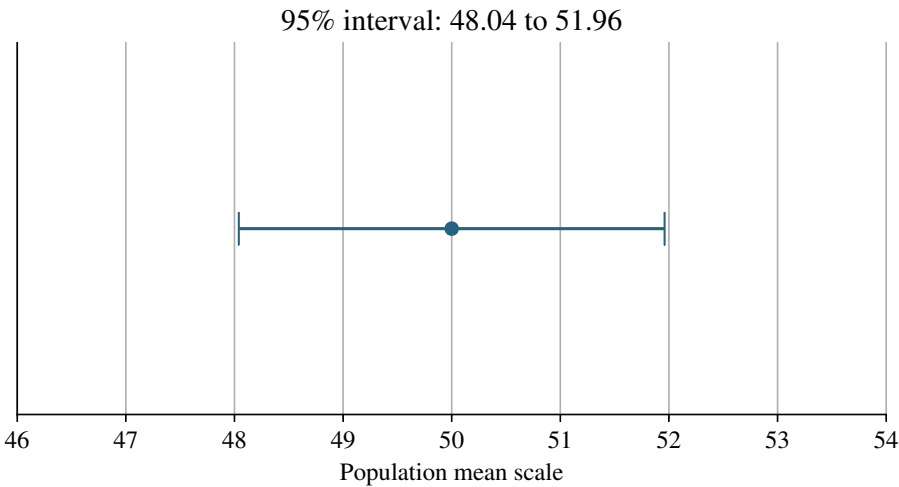
- **Work through a checked case**

$$\bar{x} \pm 1.96 \frac{\sigma}{\sqrt{n}}, \quad \chi^2 = \sum \frac{(O - E)^2}{E}$$

- Check the result against the starting quantities. Substitute into the original relation, or compare the graph and numerical answer where appropriate.

Example:

With mean=50,σ=10,n=100, standard error=1 and the 95% interval is [48.04,51.96]. For observed counts 30 and 20 against expected 25 and 25, chi-square=25/25+25/25=2, before deciding degrees of freedom and the rejection threshold.



- **Test a tempting shortcut**
 - A 95% confidence procedure does not give a 95% probability that a fixed population mean moves within this one interval. Expected counts and estimated parameters affect the validity and degrees of freedom of chi-square tests.
 - When a shortcut fails, identify the assumption it breaks. Keep an exact value until the requested final rounding.

Warn:

A strong correlation alone establishes a causal relationship. This claim is false. Explain which definition or assumption it violates.

- **Interpret a new situation**
 - Report uncertainty in the context. In regression, inspect residuals, avoid unjustified extrapolation, and distinguish a fitted relationship from a causal explanation. S3 and IB AI HL require different verified scope boundaries.
 - A complete solution gives the mathematical result and explains what it means. Check that it is possible in the stated context.

■ Use this in your course

- Current first-assessment-2021 Applications and Interpretation HL. This is authored concept support; the full guide is needed to certify every objective.
- Give the method before the final answer, and use the paper's calculator and formula rules. Review a wrong answer by locating the first invalid step.

Key:

An interval calculated by a procedure designed to cover the population parameter at a stated long-run rate. Choose the relationship, show the method, check its assumptions and interpret the result.