

Statistics and probability · supported focus

Supported focus · HL

Supported scope

- 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.
- State model conditions and distinguish fictional data from measurements.

Data · 1.1

Reasoning

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.

Data · 2.1

Reasoning

Compare an appropriate average and spread in context. A histogram uses area for frequency, so $\text{height} = \text{frequency} / \text{class width}$. 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.

Data · 3.1

Reasoning

Use the displayed relationship within its stated domain.

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

Data · 3.2

Reasoning

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

Data · 3.3

Reasoning

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.

Data · 4.1

Reasoning

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.

Data · 4.2

Reasoning

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.

Data · 5.1

Reasoning

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.

Data · 6.1

Reasoning

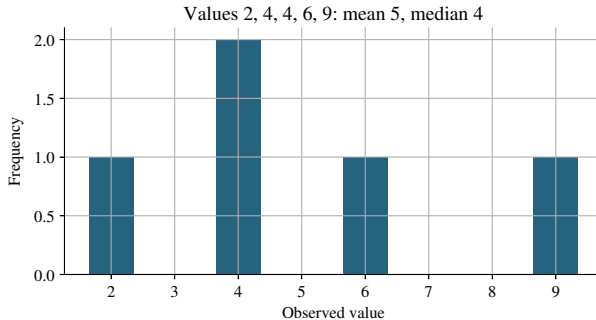
Current first-assessment-2021 Applications and Interpretation HL. This is authored concept support; the full guide is needed to certify every objective. Give the method before the final answer, and use the paper's calculator and formula rules. Review a wrong answer by locating the first invalid step.

Data · 6.2

Reasoning

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

Data · diagram



- Original teaching diagram: Data summaries, histograms and interpretation
- Use the labels and the stated model together.

Data · original transfer

Your turn

Classes 0–10 and 10–30 have frequencies 20 and 30. Find histogram densities and compare bar heights with frequencies; can the exact mean be recovered?

Data · transfer reasoning

Method and interpretation

Densities 2 and 1.5, so the higher bar has fewer observations. Areas represent frequency. Only a grouped midpoint estimate, not the exact mean, follows without raw values; state the endpoint convention.

Probability · 1.1

Reasoning

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.

Probability · 2.1

Reasoning

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.

Probability · 3.1

Reasoning

Use the displayed relationship within its stated domain.

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

Probability · 3.2

Reasoning

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

Probability · 3.3

Reasoning

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

Probability · 4.1

Reasoning

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.

Probability · 4.2

Reasoning

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

Probability · 5.1

Reasoning

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.

Probability · 6.1

Reasoning

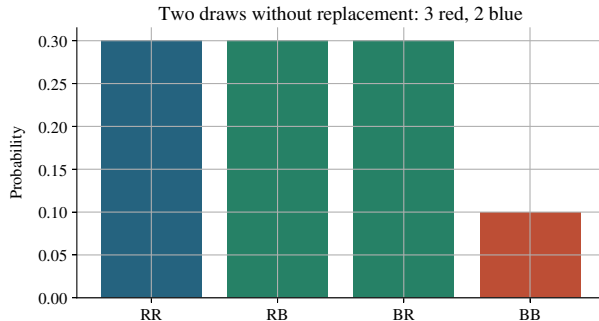
Current first-assessment-2021 Applications and Interpretation HL. This is authored concept support; the full guide is needed to certify every objective. Give the method before the final answer, and use the paper's calculator and formula rules. Review a wrong answer by locating the first invalid step.

Probability · 6.2

Reasoning

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.

Probability · diagram



- Original teaching diagram: Probability, trees and conditional reasoning
- Use the labels and the stated model together.

Probability · original transfer

Your turn

A bag has 3 red and 2 blue counters. Draw two without replacement. Given at least one red, find the probability both are red.

Probability · transfer reasoning

Method and interpretation

$P(RR) = 3/5 \times 2/4 = 0.3$. $P(\text{at least one } R) = 1 - 2/5 \times 1/4 = 0.9$. Conditional probability $= 0.3/0.9 = 1/3$. The conditioning changes the sample space.

Distributions · 1.1

Reasoning

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.

Distributions · 2.1

Reasoning

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.

Distributions · 3.1

Reasoning

Use the displayed relationship within its stated domain.

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

Distributions · 3.2

Reasoning

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

Distributions · 3.3

Reasoning

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.

Distributions · 4.1

Reasoning

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.

Distributions · 4.2

Reasoning

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

Distributions · 5.1

Reasoning

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.

Distributions · 6.1

Reasoning

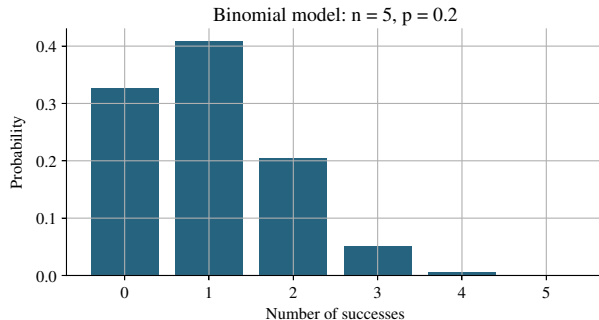
Current first-assessment-2021 Applications and Interpretation HL. This is authored concept support; the full guide is needed to certify every objective. Give the method before the final answer, and use the paper's calculator and formula rules. Review a wrong answer by locating the first invalid step.

Distributions · 6.2

Reasoning

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

Distributions · diagram



- Original teaching diagram: Binomial, normal and Poisson models
- Use the labels and the stated model together.

Distributions · original transfer

Your turn

For $X \sim \text{Binomial}(5, 0.2)$, find $P(X \geq 1)$ and the mean. Explain one situation in which the binomial model would be inappropriate.

Distributions · transfer reasoning

Method and interpretation

$1 - 0.8^5 = 0.67232$; mean $np = 1$. Trials must be independent with constant success probability and fixed count; depletion or correlated outcomes can violate this.

Testing · 1.1

Reasoning

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.

Testing · 2.1

Reasoning

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.

Testing · 3.1

Reasoning

Use the displayed relationship within its stated domain.

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

Testing · 3.2

Reasoning

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

Testing · 3.3

Reasoning

For $H_0: p=0.1$ and $H_1: p>0.1$ with $n=10$, observing at least 3 defects has probability $1-(0.9^{10}+10\times 0.1\times 0.9^9+45\times 0.1^2\times 0.9^8)\approx 0.070191$. At 5%, this is insufficient evidence that the defect rate exceeds 10%.

Testing · 4.1

Reasoning

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.

Testing · 4.2

Reasoning

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

Testing · 5.1

Reasoning

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.

Testing · 6.1

Reasoning

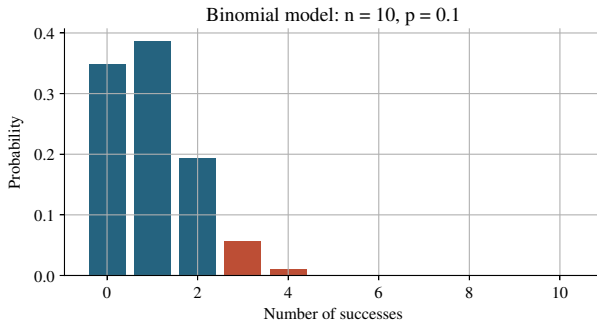
Current first-assessment-2021 Applications and Interpretation HL. This is authored concept support; the full guide is needed to certify every objective. Give the method before the final answer, and use the paper's calculator and formula rules. Review a wrong answer by locating the first invalid step.

Testing · 6.2

Reasoning

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

Testing · diagram



- Original teaching diagram: Hypothesis testing and contextual conclusions
- Use the labels and the stated model together.

Testing · original transfer

Your turn

A correctly designed test gives $p=0.04$. State decisions at 5% and 1% significance and explain why p is not the probability H_0 is true.

Testing · transfer reasoning

Method and interpretation

Reject at 5%, do not reject at 1%. p is a tail probability under H_0 and the model, not a posterior probability. Do not change the chosen threshold after seeing the data.

Advanced Statistics · 1.1

Reasoning

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.

Advanced Statistics · 2.1

Reasoning

For a normal mean with known σ , a 95% interval is $\text{sample mean} \pm 1.96\sigma/\sqrt{n}$. For chi-square, sum $(\text{observed}-\text{expected})^2/\text{expected}$ 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.

Advanced Statistics · 3.1

Reasoning

Use the displayed relationship within its stated domain.

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

Advanced Statistics · 3.2

Reasoning

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

Advanced Statistics · 3.3

Reasoning

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

Advanced Statistics · 4.1

Reasoning

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.

Advanced Statistics · 4.2

Reasoning

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

Advanced Statistics · 5.1

Reasoning

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.

Advanced Statistics · 6.1

Reasoning

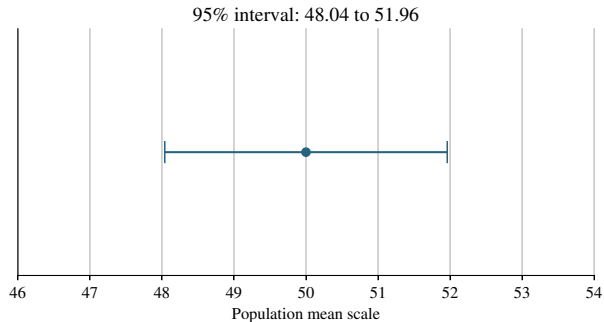
Current first-assessment-2021 Applications and Interpretation HL. This is authored concept support; the full guide is needed to certify every objective. Give the method before the final answer, and use the paper's calculator and formula rules. Review a wrong answer by locating the first invalid step.

Advanced Statistics · 6.2

Reasoning

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.

Advanced Statistics · diagram



- Original teaching diagram: Confidence intervals, chi-square and regression
- Use the labels and the stated model together.

Advanced Statistics · original transfer

Your turn

A known- σ normal model has sample mean 50, $\sigma 10$, $n 100$. Give the 95% interval, then the interval if n quadruples with mean and σ unchanged. Explain the repeated-sampling interpretation.

Advanced Statistics · transfer reasoning

Method and interpretation

SE 1:48.04 to 51.96. At $n = 400$, SE 0.5:49.02 to 50.98. The procedure covers the fixed population mean in 95% of repeated samples under assumptions; it does not assign 95% probability to the fixed mean after observing this interval.

Check yourself

- Explain frequency density using the stated model.
- Explain conditional probability using the stated model.
- Change one value in the original transfer; explain which assumption still matters.