

Inference Summary

STEP 4

How to Organize a Statistical Problem: A Four-Step Process

Confidence intervals (CIs)		Significance tests
STATE:	What <i>parameter</i> do you want to estimate, and at what <i>confidence level</i> ?	What <i>hypotheses</i> do you want to test, and at what <i>significance level</i> ? Define any <i>parameters</i> you use.
PLAN:	Choose the appropriate inference <i>method</i> . Check <i>conditions</i> .	Choose the appropriate inference <i>method</i> . Check <i>conditions</i> .
DO:	If the conditions are met, perform <i>calculations</i> .	If the conditions are met, perform <i>calculations</i> . <ul style="list-style-type: none"> • Compute the test statistic. • Find the P-value.
CONCLUDE:	Interpret your interval in the context of the problem.	Make a <i>decision</i> about the hypotheses in the context of the problem.

CI: statistic \pm (critical value) \cdot (standard deviation of statistic)

Standardized test statistic = $\frac{\text{statistic} - \text{parameter}}{\text{standard deviation of statistic}}$

Inference about	Number of samples (groups)	Interval or test	Name of procedure (TI Calculator function) Formula	Conditions
Proportions	1	Interval	One-sample z interval for p (1-PropZInt) $\hat{p} \pm z^* \sqrt{\frac{\hat{p}(1 - \hat{p})}{n}}$	Random Data from a random sample or randomized experiment <ul style="list-style-type: none"> ○ 10%: $n \leq 0.10N$ if sampling without replacement Large Counts At least 10 successes and failures; that is, $n\hat{p} \geq 10$ and $n(1 - \hat{p}) \geq 10$
		Test	One-sample z test for p (1-PropZTest) $z = \frac{\hat{p} - p_0}{\sqrt{\frac{p_0(1 - p_0)}{n}}}$	Random Data from a random sample or randomized experiment <ul style="list-style-type: none"> ○ 10%: $n \leq 0.10N$ if sampling without replacement Large Counts $np_0 \geq 10$ and $n(1 - p_0) \geq 10$
	2	Interval	Two-sample z interval for $p_1 - p_2$ (2-PropZInt) $(\hat{p}_1 - \hat{p}_2) \pm z^* \sqrt{\frac{\hat{p}_1(1 - \hat{p}_1)}{n_1} + \frac{\hat{p}_2(1 - \hat{p}_2)}{n_2}}$	Random Data from independent random samples or randomized experiment <ul style="list-style-type: none"> ○ 10%: $n_1 \leq 0.10N_1$ and $n_2 \leq 0.10N_2$ if sampling without replacement Large Counts At least 10 successes and failures in both samples/groups; that is, $n_1\hat{p}_1 \geq 10, n_1(1 - \hat{p}_1) \geq 10,$ $n_2\hat{p}_2 \geq 10, n_2(1 - \hat{p}_2) \geq 10$
		Test	Two-sample z test for $p_1 - p_2$ (2-PropZTest) $z = \frac{(\hat{p}_1 - \hat{p}_2) - 0}{\sqrt{\frac{\hat{p}_c(1 - \hat{p}_c)}{n_1} + \frac{\hat{p}_c(1 - \hat{p}_c)}{n_2}}}$ where $\hat{p}_c = \frac{\text{total successes}}{\text{total sample size}} = \frac{X_1 + X_2}{n_1 + n_2}$	Random Data from independent random samples or randomized experiment <ul style="list-style-type: none"> ○ 10%: $n_1 \leq 0.10N_1$ and $n_2 \leq 0.10N_2$ if sampling without replacement Large Counts At least 10 successes and failures in both samples/groups; that is, $n_1\hat{p}_c \geq 10, n_1(1 - \hat{p}_c) \geq 10$ $n_2\hat{p}_c \geq 10, n_2(1 - \hat{p}_c) \geq 10$

Inference about	Number of samples (groups)	Estimate or test	Name of procedure (TI Calculator function) Formula	Conditions
Means	1 (or paired data)	Interval	One-sample t interval for μ (TInterval) $\bar{x} \pm t^* \frac{s_x}{\sqrt{n}}$ with $df = n - 1$	Random Data from a random sample or randomized experiment ○ 10%: $n \leq 0.10N$ if sampling without replacement Normal/Large Sample Population distribution Normal or large sample ($n \geq 30$); no strong skewness or outliers if $n < 30$ and population distribution has unknown shape
		Test	One-sample t test for μ (T-Test) $t = \frac{\bar{x} - \mu_0}{\frac{s_x}{\sqrt{n}}}$ with $df = n - 1$	Random Data from a random sample or randomized experiment ○ 10%: $n \leq 0.10N$ if sampling without replacement Normal/Large Sample Population distribution Normal or large sample ($n \geq 30$); no strong skewness or outliers if $n < 30$ and population distribution has unknown shape
	2	Interval	Two-sample t interval for $\mu_1 - \mu_2$ (2-SampTInt) $(\bar{x}_1 - \bar{x}_2) \pm t^* \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$ df from technology or $\min(n_1 - 1, n_2 - 1)$	Random Data from independent random samples or randomized experiment ○ 10%: $n_1 \leq 0.10N_1$ and $n_2 \leq 0.10N_2$ if sampling without replacement Normal/Large Sample Population distributions Normal or large samples ($n_1 \geq 30$ and $n_2 \geq 30$); no strong skewness or outliers if sample size < 30 and population distribution has unknown shape
		Test	Two-sample t test for $\mu_1 - \mu_2$ (2-SampTTest) $t = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$ df from technology or $\min(n_1 - 1, n_2 - 1)$	Random Data from independent random samples or randomized experiment ○ 10%: $n_1 \leq 0.10N_1$ and $n_2 \leq 0.10N_2$ if sampling without replacement Normal/Large Sample Population distributions Normal or large samples ($n_1 \geq 30$ and $n_2 \geq 30$); no strong skewness or outliers if sample size < 30 and population distribution has unknown shape
Distribution of categorical variables	1	Test	Chi-square test for goodness of fit (χ^2 GOF-Test) $\chi^2 = \sum \frac{(\text{observed} - \text{expected})^2}{\text{expected}}$ with $df = \text{number of categories} - 1$	Random Data from a random sample or randomized experiment ○ 10%: $n \leq 0.10N$ if sampling without replacement Large Counts All <i>expected counts</i> at least 5
	2 or more	Test	Chi-square test for homogeneity (χ^2 -Test) $\chi^2 = \sum \frac{(\text{observed} - \text{expected})^2}{\text{expected}}$ $df = (\# \text{ of rows} - 1)(\# \text{ of columns} - 1)$	Random Data from independent random samples or randomized experiment ○ 10%: $n_1 \leq 0.10N_1$, $n_2 \leq 0.10N_2$, and so on if sampling without replacement Large Counts All <i>expected counts</i> at least 5
Relationship between 2 categorical variables	1	Test	Chi-square test for independence (χ^2 -Test) $\chi^2 = \sum \frac{(\text{observed} - \text{expected})^2}{\text{expected}}$ $df = (\# \text{ of rows} - 1)(\# \text{ of columns} - 1)$	Random Data from a random sample or randomized experiment ○ 10%: $n \leq 0.10N$ if sampling without replacement Large Counts All <i>expected counts</i> at least 5
Relationship between 2 quantitative variables (slope)	1	Interval	One-sample t interval for β (LinRegTInt) $b \pm t^*(SE_b)$ with $df = n - 2$	Linear Relationship between the variables is linear Independent observations; check the 10% condition if sampling without replacement. Normal Responses vary Normally around regression line for all x -values Equal SD around regression line for all x -values Random Data from a random sample or randomized experiment
		Test	One-sample t test for β (LinRegTTest) $t = \frac{b - \beta_0}{SE_b}$ with $df = n - 2$	