

ASSUMPTIONS/CONDITIONS FOR HYPOTHESIS TESTS and CONFIDENCE INTERVALS

One of the important tasks when applying a statistical test (or confidence interval) is to check that the assumptions of the test are not violated.

One-sample confidence interval and z-test on μ

CONFIDENCE INTERVAL: $\bar{x} \pm (\text{z critical value}) \cdot \frac{\sigma}{\sqrt{n}}$

SIGNIFICANCE TEST: $z = \frac{\bar{x} - \mu_0}{\frac{\sigma}{\sqrt{n}}}$

CONDITIONS:

- The sample must be reasonably random.
- The data must be from a normal distribution or large sample (need to check $n \geq 30$).
- σ must be known.
- The sample must be less than 10% of the population so that $\frac{\sigma}{\sqrt{n}}$ is valid for the standard deviation of the sampling distribution of \bar{x} .

One-sample confidence interval and t-test on μ

CONFIDENCE INTERVAL: $\bar{x} \pm (\text{t critical value}) \cdot \frac{s}{\sqrt{n}}$

SIGNIFICANCE TEST: $t = \frac{\bar{x} - \mu_0}{\frac{s}{\sqrt{n}}}$ where degrees of freedom $df = n - 1$

CONDITIONS:

- In theory, the data should be drawn from a normal distribution or it is a large sample (need to check that $n \geq 30$). In practice, using the t-distribution is sufficiently robust provided that there is little skewness and no outliers in the data. Look at a graph of the data.
- The data must be reasonably random.
- The sample must be less than 10% of the population.

Two-sample confidence interval and t-test on $\mu_1 - \mu_2$

CONFIDENCE INTERVAL: $(\bar{x}_1 - \bar{x}_2) \pm t^* \sqrt{\frac{(s_1)^2}{n_1} + \frac{(s_2)^2}{n_2}}$

SIGNIFICANCE TEST: $t = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{(s_1)^2}{n_1} + \frac{(s_2)^2}{n_2}}}$

CONDITIONS:

- The two samples must be reasonably random and drawn independently or, if it is an experiment, the subjects were randomly assigned to treatments.
- In theory, the data should be drawn from normal distributions or be large samples (check that $n_1 + n_2 \geq 30$). In practice, using the t-distribution is sufficiently robust provided that there is little skewness and no outliers in the data for each sample. Examine graphs of both sets of data.

NOTE: There are two ways to calculate the degrees of freedom.

Option 1. Use procedures based on the t-statistic with critical values from the t-distribution with df equal to the smaller of $n_1 - 1$ and $n_2 - 1$. This will always yield a conservative approximation for df.

Option 2. The two-sample t statistic does not have a t distribution. Moreover, the exact distribution changes as unknown population standard deviations σ_1 and σ_2 change. However, an excellent approximation is available. Most statistical software systems and the TI-83 use the two-sample t-statistic where the degrees of freedom are calculated in the formula below. This generally is NOT a whole number.

$$df \approx \frac{\left(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}\right)^2}{\frac{1}{n_1 - 1} \left(\frac{s_1^2}{n_1}\right)^2 + \frac{1}{n_2 - 1} \left(\frac{s_2^2}{n_2}\right)^2}$$

Option 1 always errs on the safe side, reporting higher p-values and lower confidence than are actually true. The gap between what is reported and what is true is quite small unless the sample sizes are both small and unequal. As the sample sizes increase, probability values based on t with degrees of freedom equal to the smaller of $n_1 - 1$ and $n_2 - 1$ will become more accurate.

Matched pairs confidence interval and t-test

- Pairing data often reduces the danger of introducing extraneous or uncontrolled factors.
- Pairing data has the theoretical effect of reducing measurement variability, which increases the accuracy of statistical conclusions.

CONFIDENCE INTERVAL : $\bar{x}_d \pm t^* \cdot \frac{s_d}{\sqrt{n}}$ where df = n-1

TEST STATISTIC : $t = \frac{\bar{x}_d - \mu_d}{\frac{s_d}{\sqrt{n}}}$ where df = n-1

CONDITIONS:

- The sample of paired differences must be reasonably random.
- The paired differences $d = x_1 - x_2$ should be approximately normally distributed or be a large sample (need to check $n \geq 30$). This procedure is robust if there are no outliers and little skewness in the paired differences. Examine a graph of the differences.

NOTES:

- In matched pairs where two measurements are taken on each experimental unit, the unit serves as its own control.
- For matched-paired data, the standard error of the statistic for the matched pairs test will be smaller than the standard error for an independent two-sample t-test because variability within the samples has been removed. It is the reason why we block in experimental design (randomized block)

Linear Regression Test on Slope and Confidence Interval

CONFIDENCE INTERVAL: $b \pm t^* SE_b$ where $df = n - 2$ and $SE_b = \frac{s}{\sqrt{\sum (x - \bar{x})^2}}$

TEST STATISTIC: $t = \frac{b - \text{hypothesized value}}{SE_b}$ where $df = n - 2$ and $SE_b = \frac{s}{\sqrt{\sum (x - \bar{x})^2}}$

CONDITIONS:

- The observations are independent.
- The true relationship is linear. Check that the scatter plot is roughly linear and that the residual plot has no pattern.
- The standard deviation of the response y about the true line is the same everywhere. Look at the residual plot and check that the residuals have roughly the same spread across all the x -values.
- For any fixed value of x , the response y varies normally about the true line. Check a histogram or stemplot of the residuals.

NOTE: Hypotheses can specify any value for slope, for example, $H_o : \beta = 1$, $H_a : \beta \neq 1$

$H_o : \beta = 0$ (there is no useful linear relationship)

$H_a : \beta \neq 0$ (there is a useful linear relationship) or $\beta > 0$ or $\beta < 0$ (if needed)

One-sample confidence interval and z-test on p

CONFIDENCE INTERVAL: $\hat{p} \pm z^* \sqrt{\frac{\hat{p}(1 - \hat{p})}{n}}$

TEST STATISTIC: $z = \frac{\hat{p} - p}{\sqrt{\frac{p(1 - p)}{n}}}$

CONDITIONS:

- The sample must be reasonably random
- The sample must be less than 10% of the population
- The sample must be large enough so that:

$n \cdot \hat{p}$ and $n(1 - \hat{p}) \geq 10$ for a confidence interval

$n \cdot p$ and $n(1 - p) \geq 10$ for the significance test

Two-sample confidence interval and z-test on $p_1 - p_2$

CONFIDENCE INTERVAL: $(\hat{p}_1 - \hat{p}_2) \pm z^* \sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2}}$

TEST STATISTIC: $z = \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{\frac{\hat{p}_c(1-\hat{p}_c)}{n_1} + \frac{\hat{p}_c(1-\hat{p}_c)}{n_2}}}$ where $\hat{p}_c = \frac{x_1 + x_2}{n_1 + n_2}$

CONDITIONS:

- The two samples must be independently drawn and reasonably random or subjects were randomly assigned to two groups.
- The sample sizes must be large enough so that: $n_1\hat{p}_1, n_1(1-\hat{p}_1), n_2\hat{p}_2, n_2(1-\hat{p}_2)$ are all five or more. (the number of successes and the number of failures must be at least 5) for the confidence interval.

The sample size must be large enough so that: $n_1\hat{p}_c, n_1(1-\hat{p}_c), n_2\hat{p}_c, n_2(1-\hat{p}_c)$ are all five or more. (the number of successes and the number of failures must be at least 5) for the significance test.

Chi-Square Test for Goodness-of-Fit

H₀ : The hypothesized distribution is correct

H_a : At least one of the categories/proportions is not correct

TEST STATISTIC: $\chi^2 = \sum \frac{(\text{observed count} - \text{expected count})^2}{\text{expected count}}$ or $\chi^2 = \sum \frac{(\mathbf{O} - \mathbf{E})^2}{\mathbf{E}}$

where $df = k - 1$ (k is # of classes)

CONDITIONS:

- The expected cell counts are all greater than or equal to 5
- The sample is reasonably random

Chi-Square Test for Homogeneity of Proportions and for Independence

H₀ : the true category proportions are the same for all populations (homogeneity).

H_a: the true category proportions are not the same for all populations.

H₀ : the two variables are independent (alternatively, there is no association)

H_a : the two variables are not independent (alternatively, there is an association)

TEST STATISTIC: $\chi^2 = \sum \frac{(\mathbf{O} - \mathbf{E})^2}{\mathbf{E}}$ where $df = (r - 1)(c - 1)$

and the expected cell count = $\frac{(\mathbf{row\ marginal\ total})(\mathbf{column\ marginal\ total})}{\mathbf{grandtotal}}$

CONDITIONS:

- The sample must be reasonably random
- The sample size must be large enough so that all expected counts are at least 1 and no more than 20% are less than 5. In particular, all expected cell counts in a 2x2 table should be 5 or more.

NOTES:

- Finding dependence between row and column variables does not imply causation, especially on data from an observational study - the typical case when using a chi-squared test for independence.
- A two-sided z-test on $p_1 - p_2$ will give the same p-value as a chi-squared test of homogeneity on a 2x2 table.