Confidence Intervals and Hypothesis Testing for Means and Proportions

	Margin of error (for confidence intervals)	Confidence Interval endpoints	Test statistic (for hypothesis tests)
Proportions	$ME = z^* \sqrt{\frac{\hat{p}\hat{q}}{n}}$	$\hat{p} \pm ME$	$z = \frac{\hat{p} - p}{\sqrt{\frac{pq}{n}}}$
Two proportions	$ME = z^* \sqrt{\frac{\hat{p}_1\hat{q}_1}{n_1} + \frac{\hat{p}_2\hat{q}_2}{n_2}}$	$\hat{p}_1 - \hat{p}_2 \pm ME$	$z = \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{\frac{\bar{p}\bar{q}}{n_1} + \frac{\bar{p}\bar{q}}{n_2}}}$ where $\bar{p} = \frac{x_1 + x_2}{n_1 + n_2}$
Single Mean	$ME = t^* \frac{s}{\sqrt{n}}$	$\bar{x} \pm ME$	$t = \frac{\bar{x} - \mu}{\frac{s}{\sqrt{n}}}$
Two Means (Independent samples)	$ME = t^* \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$	$\bar{x}_1 - \bar{x}_2 \pm ME$	$t = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$
Two Means (Matched pairs)	$ME = t^* \frac{s_d}{\sqrt{n}}$	$\bar{d} \pm ME$	$t = \frac{\bar{d}}{\frac{s_d}{\sqrt{n}}}$
Goodness of fit	N/A	N/A	$\chi^2 = \sum \frac{(\text{expected} - \text{observed})^2}{\text{expected}}$
Independence or Homogeneity	N/A	N/A	$\chi^2 = \sum \frac{(\text{expected} - \text{observed})^2}{\text{expected}}$
Correlation (confidence interval/test for the slope)	$ME = t^* SE(b_1)$ (use software)	$b_1 \pm ME$	$t = \frac{b_1}{SE(b_1)}$ (use software)

- p is the population proportion and $q = 1 - p$.
- $\hat{p}$ is the sample proportion, and $\hat{q} = 1 - \hat{p}$.
- For comparing two proportions, p_1 and q_1 are the population proportions for the two samples, and $\hat{p}_1$ and $\hat{p}_2$ are the corresponding sample proportions. $\hat{p}$ and $\hat{q}$ are the pooled sample proportions (the sample data from both groups are pooled together to form them).
- μ is the population mean
- $\bar{x}$ is the sample mean
- s is the sample standard deviation
- For independent samples: $\bar{x}_1$ and $\bar{x}_2$ are the means of two groups, s_1 and s_2 are the corresponding standard deviations, and n_1 and n_2 are the corresponding sample sizes.
- For paired data: $\bar{d}$ is the mean difference of the pairs of sample data, s_d is the standard deviation.
- z^* and t^* are critical values, and depend on the confidence level. t^* also depends on df.
- b_1 is the slope of the regression equation for the sample, β_1 is the slope of the regression line for the population.
- $SE(b_1)$ is the standard error of the slope (use computer software to get).