Type of inference	Parameter (or question for HT)	Conditions needed to use theoretical sampling distribution of test statistic	Sample Statistic	Confidence Interval formula	Theoretical distribution of Test statistic	Degrees of freedom	Test statistic SE is "standard error" Generally, use software to obtain this. See formulas to the right.	Standard Error formula for CI	Standard Error for HT	Sample size for estimating
One mean	μ	Distribution normal or CLT applies, meaning $n \geq 30$ approximately	$\overline{X}$	$\overline{X} - t^* \cdot SE \leq \mu \leq \overline{X} + t^* \cdot SE$	t distribution	$df = n - 1$	$t = \frac{\overline{X} - \mu_0}{SE}$	$SE = \frac{s}{\sqrt{n}}$	$SE = \frac{s}{\sqrt{n}}$	$n = \left(\frac{z^* \cdot \tilde{\sigma}}{\text{ME}}\right)^2$, where ME is the chosen margin of error and $\tilde{\sigma}$ is an estimate of the population standard deviation.
One proportion	p	$np \geq 10$ AND $n(1-p) \geq 10$ CI: $p = \hat{p}$ HT: $p = p_0$	$\hat{p}$	$\hat{p} - z^* \cdot SE \leq p \leq \hat{p} + z^* \cdot SE$	normal distribution	Not relevant	$z = \frac{\hat{p} - p_0}{SE}$	$SE = \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$	$SE = \sqrt{\frac{p_0(1-p_0)}{n}}$ Where p_0 is the value in the null hypothesis	$n = \left(\frac{z^*}{\text{ME}}\right)^2 \cdot \tilde{p}(1-\tilde{p})$, where ME is the chosen margin of error and we use $\tilde{p} = 0.5$ or some other value of $\tilde{p}$ if available.
Two means	$\mu_1 - \mu_2$	In EACH group, distribution normal OR CLT applies, meaning $n \geq 30$ approximately	$\overline{X}_1 - \overline{X}_2$	$(\overline{X}_1 - \overline{X}_2) - t^* \cdot SE \leq \mu_1 - \mu_2 \leq (\overline{X}_1 - \overline{X}_2) + t^* \cdot SE$	t distribution	df = smaller of $n_1 - 1$ and $n_2 - 1$ or Satterthwaite approximation	$t = \frac{(\overline{X}_1 - \overline{X}_2) - 0}{SE}$	$\text{SE} = \sqrt{\frac{s_1{}^2}{n_1} + \frac{s_2{}^2}{n_2}}$	$\text{SE} = \sqrt{\frac{s_1{}^2}{n_1} + \frac{s_2{}^2}{n_2}}$	**Standard Error (degrees of freedom) for two-sample t-procedures:** In two-sample t procedures, in order to show that the test statistic has an exact t-distribution, we must have that the two population variances are equal. In that case, $df = n_1 + n_2 - 2$ If it is not appropriate to assume that the two population variances are equal, then a "conservative" approach (does not overstate our confidence in our answers) is to use the smaller of $n_1 - 1$ and $n_2 - 1$. An adjustment can be made to the degrees of freedom to take into account how different the variances are and how different the sample sizes are. Most statistical software will use this Satterthwaite approximation as the degrees of freedom for two-sample t procedures. It is derived by a modification of the method of moments method of estimation. $df = \frac{\left(\frac{s_1{}^2}{n_1} + \frac{s_2{}^2}{n_2}\right)^2}{\frac{1}{n_1 - 1}\left(\frac{s_1{}^2}{n_1}\right)^2 + \frac{1}{n_2 - 1}\left(\frac{s_2{}^2}{n_2}\right)^2}$ Don't do this "by hand." It is included here because, as you use software, you will see degrees of freedom that do not fit the "simple" method given in this handout and in many applied statistics texts.
Two proportions	$p_1 - p_2$	$n_1\hat{p}_1 \geq 10$ AND $n_1(1-\hat{p}_1) \geq 10$ AND $n_2\hat{p}_2 \geq 10$ AND $n_2(1-\hat{p}_2) \geq 10$	$\hat{p}_1 - \hat{p}_2$	$(\hat{p}_1 - \hat{p}_2) - z^* \cdot SE \leq p_1 - p_2 \leq (\hat{p}_1 - \hat{p}_2) + z^* \cdot SE$	normal distribution	Not relevant	$z = \frac{(\hat{p}_1 - \hat{p}_2) - 0}{SE}$	$SE = \sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2}}$ where $\hat{p}_1$ and $\hat{p}_2$ are the sample proportions from the two separate samples	$\text{SE} = \sqrt{\frac{\overline{p}(1-\overline{p})}{n_1} + \frac{\overline{p}(1-\overline{p})}{n_2}}$ For testing whether the pop. prop. are equal. Here, $\overline{p}$ is the pooled proportion. $\overline{p} = \frac{\text{sum of counts from both samples}}{\text{sum of trials from both samples}}$	
Mean of differences from matched pairs data	μ_d	Distribution of differences normal OR CLT applies, meaning $n \geq 30$ approximately, where n is number of pairs.	$\overline{X}_d$	$\overline{X}_d - t^* \cdot SE \leq \mu_d \leq \overline{X}_d + t^* \cdot SE$	t distribution	$df = n - 1$ where n is the number of pairs	$t = \frac{\overline{X}_d - \mu_0}{SE}$	$SE = \frac{s_d}{\sqrt{n_d}}$ where the subscripts refer to using the differences	$SE = \frac{s_d}{\sqrt{n_d}}$ where the subscripts refer to using the differences	
Correlation coefficient	ρ	See regression model conditions. Generally linear pattern (rather than a different pattern), same variance across x-values, residuals are independent and have normal distribution.	r	$r - t^* \cdot SE \leq \rho \leq r + t^* \cdot SE$	t distribution	$df = n - 2$	$t = \frac{r - \rho_0}{SE}$	$SE = \sqrt{\frac{1 - r^2}{n - 2}}$	$SE = \sqrt{\frac{1 - r^2}{n - 2}}$	
Slope coefficient	β	See regression model conditions. Generally linear pattern (rather than a different pattern), same variance across x-values, residuals are independent and have normal distribution.	b	$b - t^* \cdot SE \leq \beta \leq b + t^* \cdot SE$	t distribution	$df = n - 2$ for simple regression	$t = \frac{b - \beta_0}{SE}$	Obtain with technology	Obtain with technology	
Test of goodness of fit	Do the data fit a particular specified distribution?	Each expected count is at least 5.	χ^2_p chi-squared, p degrees of freedom	This investigation is a test. No parameters are estimated with this procedure.	χ^2_p distribution	df = number of categories minus 1	$\chi^2 = \sum \frac{(Observed - Expected)^2}{Expected}$	Not relevant	Not relevant	
Test of association of two categorical variables	Are the two categorical variables associated?	Each expected count is at least 5.	χ^2_p chi-squared, p degrees of freedom	This investigation is a test. No parameters are estimated with this procedure.	χ^2_p distribution	r = # of rows c = # of cols $df = (r-1)(c-1)$	$\chi^2 = \sum \frac{(Observed - Expected)^2}{Expected}$	Not relevant	Not relevant	
Analysis of Variance (ANOVA) for difference of means	Is there a difference in the means of two or more groups?	In EACH group, distribution normal OR CLT applies, meaning $n \geq 30$ approximately. Variability is similar in all groups ($\frac{\sigma_{\max}}{\sigma_{\min}} \leq 2$).	$F_{p,q}$ F statistic, p and q degrees of freedom	This investigation is a test. No parameters are estimated with this procedure.	F distribution	k = # of groups n = total sample sizes $df: p = k - 1$, $q = n - k$	$F = \frac{\text{Mean Square Error Between Groups}}{\text{Mean Square Error Within Groups}}$	Not relevant	Not relevant	
Analysis of Variance (ANOVA) for regression	Is at least one variable in the model useful in predicting the response variable?	For EACH explanatory variable, same conditions as linear model for single explanatory variable. Check with residual analysis, including plots. To start, plot the residuals vs the fitted values.	$F_{p,q}$ F statistic, p and q degrees of freedom	This investigation is a test. No parameters are estimated with this procedure.	F distribution	k = # of explanatory variables n = sample size $df: p = k$, $q = n - k - 1$	$F = \frac{\text{Mean Square Error Between Groups}}{\text{Mean Square Error Within Groups}}$	Not relevant	Not relevant	

Adapted from "Theoretical Distributions. Applied Statistic Overview" by Mary Parker, October 2020