FORMULAS & Helpful Info

Pooled T-test Statistic for \( \mu_1 - \mu_2 \)  
Assumptions: Simple Random Sample, Independent Samples, Normal populations or large \( n \), Same \( \sigma \)  
\[ t = \frac{\bar{x}_1 - \bar{x}_2}{s_p} \]  
where \( s_p = \sqrt{\frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}} \)  
\[ \text{d.f.} = n_1 + n_2 - 2 \]

Pooled T-interval procedure for \( \mu_1 - \mu_2 \)  
Assumptions: Simple Random Sample, Independent Samples, Normal populations or large \( n \), Same \( \sigma \)  
\[ (\bar{x}_1 - \bar{x}_2) \pm t_{s_p} \sqrt{\frac{1}{n_1} + \frac{1}{n_2}} \]  
where \( s_p = \sqrt{\frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}} \)
\[ \text{d.f.} = n_1 + n_2 - 2 \]

Nonpooled T-test Statistic for \( \mu_1 - \mu_2 \)  
Assumptions: Simple Random Sample, Independent Samples, Normal populations or large \( n \)  
\[ t = \frac{\bar{x}_1 - \bar{x}_2}{\frac{s_1}{\sqrt{n_1}} + \frac{s_2}{\sqrt{n_2}}} \]

Nonpooled T-interval for \( \mu_1 - \mu_2 \)  
Assumptions: Simple Random Sample, Independent Samples, Normal populations or large \( n \)  
\[ \text{d.f.} = \text{largest integer less than or equal to} \left( \frac{\left(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}\right)^2}{\frac{s_1^2}{n_1 - 1} + \frac{s_2^2}{n_2 - 1}} \right) \]
\[ \text{CI:} \quad (\bar{x}_1 - \bar{x}_2) \pm t_{\left(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}\right)^2} \sqrt{\frac{1}{n_1 - 1} + \frac{1}{n_2 - 1}} \]

Paired Samples for \( \mu_1 - \mu_2 \)  
\( H_0: \mu_1 = \mu_2 \)  
\( d = \) difference variable  
Assumptions: \( d \) is normally distributed, \( n \) pairs  
\[ \text{d.f.} = n - 1 \]

\[ t = \frac{\bar{d}}{s_d / \sqrt{n}} \]

Inferences for population proportions  
Assumptions: Simple Random Sample, At least 5 successes, At least 5 failures \( \hat{p} = \frac{\text{successes}}{n} \)  
\[ \text{CI:} \quad \hat{p} \pm z \sqrt{\frac{\hat{p}(1-\hat{p})}{n}} \]

Minimum sample size for estimating \( p \) within error \( E \):  
\( n \) is the smallest whole number greater than or equal to \( n^* \)
\[ n^* = \hat{p}(1-\hat{p}) \left( \frac{z}{E} \right)^2 \]  

If no estimate for \( p \) is known  
\[ n^* = 0.25 \left( \frac{z}{E} \right)^2 \]

One Proportion Hypothesis Test  
\( H_0: p = p_0 \)  
Assumptions: Simple Random Sample, At least 5 successes, At least 5 failures  
\[ z = \frac{\hat{p} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}} \]

Inferences for 2 population proportions  
\( H_0: p_1 = p_2 \)  
Assumptions: Simple Random Sample, Independent Samples, At least 5 successes and at least 5 failures for both samples  
\[ z = \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{\frac{p_p(1-p_p)}{1} \left( \frac{1}{n_1} + \frac{1}{n_2} \right)}} \]  
where \( p_p = \frac{\text{successes}_1 + \text{successes}_2}{n_1 + n_2} \)

Two Proportion Z interval procedure for \( p_1 - p_2 \)  
Assumptions: Simple Random Sample, Independent Samples, At least 5 successes and at least 5 failures for both samples  
\[ \text{CI:} \quad (\hat{p}_1 - \hat{p}_2) \pm z \sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \hat{p}_2(1-\hat{p}_2)} \]
Chi-Square Goodness-of-fit Test  Assumptions: All expected frequencies are at least 1, At most 20% of expected frequencies are less than 5, simple random sample  
H₀: Variable is distributed as expected,  H₁: Variable is not distributed as expected  
d.f. = k – 1  where k is the number of possible values for the variable  
\[ \chi^2 = \sum \frac{(O - E)^2}{E} \]  where  O = observed value  and  E = expected value

Chi-Square Independence Test  Assumptions: All expected frequencies are at least 1, At most 20% of expected frequencies are less than 5, simple random sample  
H₀: The variables are not associated,  H₁: The variables are associated  
d.f. = (r – 1)(c – 1)  where r = number of rows and c = number of columns  
\[ \chi^2 = \sum \frac{(O - E)^2}{E} \]  Expected values = \[ \frac{RC}{n} \]  where  R = sum of row values and  C = sum of column values