Inference Formulas & Conditions

Test	Conditions	Null Hypothesis	Test Statistic	Confidence Interval
One-Sample t-test (for a population mean)	1.) Random Sample 2.) 10 % Condition: $n < \frac{1}{10}N$ 3.) Normality: Population is normal, **or** $n \geq 30$ by CLT **or** distribution not overly skewed, no outliers	$H_o: \mu = \mu_o$	$t = \frac{\bar{x} - \mu_o}{\frac{s}{\sqrt{n}}}$ df = $n - 1$	$\bar{x} \pm t^*_{n-1} \frac{s}{\sqrt{n}}$
Matched Pair t-test (for a difference in 2 **dependent** sample means)	1.) Random Sample or Assignment 2.) 10 % Condition: $n < \frac{1}{10}N$ 3.) Samples are **DEPENDENT** (Matched) 4.) Normality: $n \geq 30$ by CLT or distribution of **differences** are not overly skewed, no outliers (**n is # of pairs**)	$H_0 : \mu_d = 0$	$t = \frac{\bar{x}_d - 0}{\frac{s_d}{\sqrt{n}}}$ df = $n - 1$ (**n is # of pairs**)	$\bar{x}_d \pm t^*_{n-1} \frac{s_d}{\sqrt{n}}$ ***Subtract the lists of data to create 1-list before you start***
Two-Sample t-test (for a difference in 2 population means) **DO NOT POOL**	1.) Random Sample or Assignment 2.) 10 % Condition: BOTH $n < \frac{1}{10}N$ 3.) Samples are Independent 4.) Normality: BOTH Populations are normal, **or** BOTH $n_1, n_2 \geq 30$ by CLT **or** BOTH distributions not overly skewed, no outliers	$H_0 : \mu_1 = \mu_2$ **or** $H_0 : \mu_1 - \mu_2 = 0$	$t = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$ df = smaller of $n_1 - 1$ or $n_2 - 1$ or calculator df	$(\bar{x}_1 - \bar{x}_2) \pm t^*_{n-1}\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$
One Proportion z-test (for a population proportion)	1.) Random Sample 2.) 10 % Condition: $n < \frac{1}{10}N$ 3.) **Normality*- Large Counts** $np_0 \geq 10 \quad n(1 - p_0) \geq 10$ ***Note: Use p_0 for the test and $\hat{p}$ for the interval**	$H_0 : p = p_0$	$z = \frac{\hat{p} - p_0}{\sqrt{\frac{p_0(1 - p_0)}{n}}}$	$\hat{p} \pm z^* \sqrt{\frac{\hat{p}(1 - \hat{p})}{n}}$ **Note: The condition for CI for Normality-Large Counts is** $n\hat{p} \geq 10$ *and* $n(1 - \hat{p}) \geq 10$
Two Proportion z-test (for a difference in 2 population proportions)	1.) Random Sample or Assignment 2.) 10 % Condition: BOTH $n < \frac{1}{10}N$ 3.) Samples are Independent 4.) **Normality*- Large Counts** $n_1 p_c \geq 5 \quad n_1(1 - p_c) \geq 5$ $n_2 p_c \geq 5 \quad n_2(1 - p_c) \geq 5$ ***Note: Use p_c for the test and $\hat{p}_1, \hat{p}_2$ for the interval**	$H_0 : p_1 = p_2$ **or** $H_0 : p_1 - p_2 = 0$	$z = \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{\frac{p_c(1 - p_c)}{n_1} + \frac{p_c(1 - p_c)}{n_2}}}$ **Note:** $p_c = \frac{x_1 + x_2}{n_1 + n_2}$	$(\hat{p}_1 - \hat{p}_2) \pm z^* \sqrt{\frac{\hat{p}_1(1 - \hat{p}_1)}{n_1} + \frac{\hat{p}_2(1 - \hat{p}_2)}{n_2}}$ **Note: The condition for CI for Normality-Large Counts is** $n_1\hat{p}_1 \geq 5 \quad n_1(1 - \hat{p}_1) \geq 5$ $n_2\hat{p}_2 \geq 5 \quad n_2(1 - \hat{p}_2) \geq 5$

Note: NO 10% CONDITION WITH RANDOM ASSIGNMENT!!!

Why are the conditions important?

RANDOM – Best chance of getting a representative sample from the population to make conclusions from.

10% - Our sampling is without replacement and thus “dependent”. With less than 10% of the population in our sample. it creates an “essentially independent” environment and makes our standard error valid.

NORMAL/LARGE COUNTS – Guarantees that the sampling distribution will be approx. normal and thus our calculations will be valid.