

Two Sample Test

Comparing Two Means

- Significant difference between two samples?
One sample from one population, while the other sample from another population.
- Fail to reject H_0 : the difference between two means is just because of sampling error.
Actually drawn from one population.
- Reject H_0 : two samples do not represent one population.
- Sampling distribution of the difference between means, not sampling distribution of the mean.

- Two types: Independent or non-independent
- Independent: compare the means of two unrelated samples to determine if these two means differ significantly. (Two samples do not affect each other)
- Non-independent: matched pair (each observation in one sample is matched with each observation in another sample)
- When writing Alternative H, think about two-tailed or one-tailed.
- Null hypothesis: Two population means are the same.
- Z-test or t-test (when sample size is small...)

How to get statistic? (Indep.)

$$\textit{Statistic} = \frac{(\textit{Observed} - \textit{Null})}{\textit{S.E.}}$$

- However, observed difference, expected difference, and SE for difference, in this case.

- SE for Z test $\sqrt{\left(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}\right)}$

- SE for T test
(Equal population var.) $\sqrt{\left[\left(\frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}\right)\left(\frac{n_1 + n_2}{n_1 n_2}\right)\right]}$

- Observed= mean for the first sample – mean for the second sample.
- Null= mean for the first population – mean for the second population. Therefore, always 0.
- S is standard deviation for each sample.
- N is sample size for each sample.
- If you are doing t-test, what should you care about?
Degree of freedom! In this case, DF is (n_1+n_2-2)

Compare two proportions

- Null hypothesis: Two population proportions are the same: $(p_1 - p_2) = 0$
- Alternative: Think about one-tailed or two-tailed
- Always z-test.

$$\textit{Statistic} = \frac{(\textit{Observed} - \textit{Null})}{\textit{S.E.}}$$

- Observed: Proportion of the first sample – proportion of the second sample
- Null: proportion of the first population – proportion of the second population. Therefore, always 0.
- Pooled proportion $p = \frac{n_1 p_1 + n_2 p_2}{n_1 + n_2}$
- SE $\sqrt{\left[p(1-p) \left(\frac{1}{n_1} + \frac{1}{n_2} \right) \right]}$