MA214 Lecture 04

Two Sample Inference : CI and HT

F-statistic

Example : $F = \frac{S_1^2}{S_2^2} = \frac{(1.4)^2}{(0.8)^2} = 3.06$

The tabulated value of $F_{0.025} = 2.7559$ for (20, 15) df.

Reject the H_0 and conclude that the variances are equal.

Small sample sizes:

- Normally distributed
- Independent
- Variances from two groups are equal.

Use the F-test.

H_0 would be the variances are equal.
H_a would be the variances are unequal.

F-statistic $= \frac{\max(V_1^2, V_2^2)}{\min(V_1^2, V_2^2)}$

Reject it if $F > F_{0.025}$

Degrees of freedom: $(n_1 - 1, n_2 - 1)$
$(21-1, 16-1)$
$= (20, 15)$

If the variances are not equal.
use the Welch's t-test.

What if the samples are NOT INDEPENDENT?

The notion of "paired data".
matched-pair study.

Be aware of the mathematical assumptions behind the scene!

Paired t-test

$t = \frac{\bar{D} - D_0}{\frac{S_D}{\sqrt{n}}}$

we care about the differences between the starting and ending states.

New Method	Standard Method	Differences.
X	Y	D = X - Y

Continuous populations.

- Comparing means
- Comparing population proportions.

Example 7:

we can have a "Contingency table"
2×2

Proportions: $\hat{p}_1$ (men)

$\hat{p}_2$ (women)

Hypothesis: $H_0: p_1 - p_2 = 0$
$H_a: p_1 - p_2 > 0$

$Z = \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{\hat{p}(1-\hat{p})\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$

$\hat{p}_1 = \frac{x_1}{n_1}$, $\hat{p}_2 = \frac{x_2}{n_2}$

$\hat{p} = \frac{x_1 + x_2}{n_1 + n_2}$
pooled proportion.

$\hat{p}_1 = \frac{52}{1000} = 0.052$

$\hat{p}_2 = \frac{23}{1000} = 0.023$

$\hat{p} = \frac{52+23}{1000+1000} = 0.0375$

$Z = 3.41$

p-value $= P(Z > 3.41) = 0$ this can be seen it's less than any alpha level.

Confidence Interval for the difference between two proportions $(p_1 - p_2)$

Reject H_0

$$(\hat{p}_1 - \hat{p}_2) \pm z_{\frac{\alpha}{2}} \sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2}}$$

Example:

$\hat{p}_1 = \frac{16}{400}$

$\hat{p}_2 = \frac{24}{300}$