STA 209 – Exam 2 Formula Sheet

CI for a single population proportion (p)

Conditions

1) The data comes from a random sample.
2) $n * \widehat{p} \geq 10$ (Success Condition)
3) $n * (1 - \widehat{p}) \geq 10$ (Failure Condition)

Formula

$\widehat{p} \pm z^* \times \sqrt{\frac{\widehat{p}(1-\widehat{p})}{n}}$

z^* is the appropriate value from the normal distribution that gives us the Confidence % that we want

- 95% Confidence $\rightarrow z^* = 1.96$
- 80% Confidence $\rightarrow z^* = 1.282$
- 90% Confidence $\rightarrow z^* = 1.645$
- 99% Confidence $\rightarrow z^* = 2.576$

CI for a difference in population proportions $(p_1 - p_2)$

Conditions

1) Data for both groups is representative, both groups independent.
2) $n_1 * \widehat{p}_1 \geq 10$ (Success Condition Grp 1)
3) $n_1 * (1 - \widehat{p}_1) \geq 10$ (Failure Condition Grp 1)
4) $n_2 * \widehat{p}_2 \geq 10$ (Success Condition Grp 2)
5) $n_2 * (1 - \widehat{p}_2) \geq 10$ (Failure Condition Grp 2)

Formula

$(\widehat{p}_1 - \widehat{p}_2) \pm z^* \times \sqrt{\frac{\widehat{p}_1(1-\widehat{p}_1)}{n_1} + \frac{\widehat{p}_2(1-\widehat{p}_2)}{n_2}}$

z^* is the appropriate value from the normal distribution that gives us the Confidence % that we want

- 95% Confidence $\rightarrow z^* = 1.96$
- 80% Confidence $\rightarrow z^* = 1.282$
- 90% Confidence $\rightarrow z^* = 1.645$
- 99% Confidence $\rightarrow z^* = 2.576$

Confidence Interval for a Single Mean

Conditions

1) The *population* is Normal **OR** the sample size large enough for CLT
2) There was a representative

100(1-α)% Confidence Interval Formula (σ unknown)

$\overline{x} \pm t_{(1-\alpha/2, df=n-1)} \times \frac{s}{\sqrt{n}}$

where $t_{(1-\alpha/2, df=n-1)}$ is the $1-\alpha/2$ quantile for a t-distribution with n-1 degrees of freedom (use `qt()` to get final values).

Confidence Interval for Difference in Means

Conditions

1) The *populations* are Normal **OR** the both sample sizes $n_1 \geq 30$ and $n_2 \geq 30$
2) There was a random sample for both groups.

100(1-α)% Confidence Interval Formula (σ unknown)

$(\overline{x}_1 - \overline{x}_2) \pm t_{(1-\alpha/2, df)} \times \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$

where $t_{(1-\alpha/2, df)}$ is the $1-\alpha/2$ quantile for a t-distribution with degrees of freedom $df = min(n_1, n_2) - 1$ (use `qt()` to get final values).

Single Proportion HT

$$Z := \frac{\hat{p} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}} \sim \mathbf{N(0,1)}$$

Conditions:

- Representative sample
- $n \times p_0 \geq 10$
- $n \times (1 - p_0) \geq 10$

Difference in Proportions

$$Z :== \frac{(\hat{p}_1 - \hat{p}_2)}{\sqrt{\widehat{p}_{pool}(1 - \widehat{p}_{pool})(\frac{1}{n_1} + \frac{1}{n_2})}} \sim \mathbf{N(0,1)}$$

,

$$\text{where} \quad \widehat{p}_{pool} = \frac{x_1 + x_2}{n_1 + n_2} = \frac{n_1\hat{p}_1 + n_2\hat{p}_2}{n_1 + n_2}$$

Conditions:

- Representative, independent samples
- $n_1 \times \widehat{p}_1 \geq 10$ and $n_1 \times (1 - \widehat{p}_1) \geq 10$
- $n_2 \times \widehat{p}_2 \geq 10$ and $n_2 \times (1 - \widehat{p}_2) \geq 10$

Single Mean

$$T := \frac{\bar{x} - \mu_0}{s/\sqrt{n}} \sim \mathbf{t(df = n\text{-}1)}$$

Conditions:

- Representative sample
- Normal population **OR** n large enough for CLT to work

Difference in Means

$$T := \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} \sim \mathbf{t(df = min(}n_1\mathbf{,}\ n_2\mathbf{)\ \text{-}\ 1)}$$

Conditions:

- Representative, independent samples
- Normal populations **OR** n_1, n_2 large enough for CLT to work

Strength of Evidence Chart