STAT 512 su 2021 Lec 08 slides

Classical two-sample results comparing means, proportions, and variances

Karl B. Gregory

University of South Carolina

These slides are an instructional aid; their sole purpose is to display, during the lecture, definitions, plots, results, etc. which take too much time to write by hand on the blackboard. They are not intended to explain or expound on any material.

Setup:

$$X_1, \ldots, X_{n_1} \overset{\text{ind}}{\sim} \text{Normal}(\mu_1, \sigma_1^2) \text{ with } \bar{X} \text{ and } S_1^2$$
$$Y_1, \ldots, Y_{n_2} \overset{\text{ind}}{\sim} \text{Normal}(\mu_2, \sigma_2^2) \text{ with } \bar{Y} \text{ and } S_2^2.$$

Pivot quantity results ($\sigma_1^2 \neq \sigma_2^2$)

$$\frac{\bar{X} - \bar{Y} - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} \sim \text{Normal}(0, 1)$$

$$\frac{\bar{X} - \bar{Y} - (\mu_1 - \mu_2)}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}} \overset{\text{approx}}{\sim} t_{\hat{\nu}},$$

In the above

$$\hat{\nu} = \left(\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}\right)^2 \left[\frac{\left(\frac{S_1}{n_1}\right)^2}{n_1 - 1} + \frac{\left(\frac{S_2^2}{n_2}\right)^2}{n_2 - 1}\right]^{-1}$$

Exercise: Derive first result above and apply to CI for $\mu_1 - \mu_2$.

Pivot quantity results ($\sigma_1^2 = \sigma_2^2 = \sigma^2_{\text{common}}$)

$$\frac{\bar{X} - \bar{Y} - (\mu_1 - \mu_2)}{\sqrt{\sigma^2_{\text{common}}\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} \sim \text{Normal}(0, 1)$$

$$\frac{\bar{X} - \bar{Y} - (\mu_1 - \mu_2)}{\sqrt{S^2_{\text{pooled}}\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} \sim t_{n_1+n_2-2},$$

In the above

$$S^2_{\text{pooled}} = \frac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{n_1 + n_2 - 2}$$

Exercise: Derive above results and apply to make CIs for $\mu_1 - \mu_2$.

Pivot quantity result

$$\frac{S_1^2/\sigma_1^2}{S_2^2/\sigma_2^2} \sim F_{n_1-1,n_2-1}$$

Exercise: Derive above result and apply to make CI for σ_2^2/σ_1^2.

Large samples from non-Normal distributions:

$$X_1, \dots, X_{n_1} \overset{\text{ind}}{\sim} F_1, \text{ mean } \mu_1 \text{ and variance } \sigma_1^2 < \infty, \text{ with } \bar{X} \text{ and } S_1^2$$
$$Y_1, \dots, Y_{n_2} \overset{\text{ind}}{\sim} F_2, \text{ mean } \mu_2 \text{ and variance } \sigma_2^2 < \infty, \text{ with } \bar{Y} \text{ and } S_2^2.$$

Suppose n_1, n_2 are large: Asymptotically, $n_1/n_2 \to c \in (0, \infty)$ as $n_1, n_2 \to \infty$.

Asymptotic pivot quantity results

$$\frac{\bar{X} - \bar{Y} - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} \xrightarrow{\text{D}} Z$$

$$\frac{\bar{X} - \bar{Y} - (\mu_1 - \mu_2)}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}} \xrightarrow{\text{D}} Z$$

as $n_1, n_2 \to \infty$, where $Z \sim \text{Normal}(0, 1)$.

Exercise: Prove the above and apply to make large-sample CIs for $\mu_1 - \mu_2$.

Specific case of Bernoulli random samples:

$$X_1, \dots, X_{n_1} \overset{\text{ind}}{\sim} \text{Bernoulli}(p_1) \text{ with } \hat{p}_1$$
$$Y_1, \dots, Y_{n_2} \overset{\text{ind}}{\sim} \text{Bernoulli}(p_2) \text{ with } \hat{p}_2$$

Suppose n_1, n_2 are large: Asymptotically, $n_1/n_2 \to c \in (0, \infty)$ as $n_1, n_2 \to \infty$.

Pivot quantity result

$$\frac{\hat{p}_1 - \hat{p}_2 - (p_1 - p_2)}{\sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2}}} \xrightarrow{\text{D}} Z$$

as $n_1, n_2 \to \infty$, where $Z \sim \text{Normal}(0, 1)$.

Exercise: Apply to make large-sample CI for $p_1 - p_2$.