EPSY 5261 : Introductory Statistical Methods

Day 19

Confidence Intervals for Comparing Two Means

Learning Goals

- At the end of this lesson, you should be able to…
  - Identify when to answer a research question with a confidence interval
  - Explain the need for creating a confidence interval to do statistical inference
  - Know how to calculate a confidence interval by hand and using R Studio for a difference in means
  - Interpret a confidence interval

Confidence Intervals

- Sampling Variability = Samples vary
- We need something to quantify the uncertainty in our estimates

Confidence Intervals

Terminology

- 95% confidence interval:
  - Sample statistic +/- (2 x SE)
- Margin of error:
  - A specified number of standard errors that we add and subtract from the sample statistic to get a confidence interval.
  - Margin of error quantifies the amount of sampling error due to variation from sample to sample.

Assumptions needed to use t-distribution for two means

- Assumptions
  - Sample size is large enough for both samples ($n_1 \geq 30$ and $n_2 \geq 30$)

OR

  - Any sample that does not satisfy the sample size condition comes from a population with a normal distribution
  - For small samples, we can proceed if the distribution of data looks reasonably bell-shaped and symmetric
  - In practice, better to use bootstrap with small samples

Formula

$$CI = (\mu_1 - \mu_2) \pm t * SE$$

Table 17.1 in text

studied in EPsy 5261.

Situation	SE
Single Mean	$\frac{\text{SD}}{\sqrt{n}}$
Single Proportion	$\frac{\hat{p}(1-\hat{p})}{\sqrt{n}}$
Difference in Means	$\sqrt{\frac{\text{SD}_1^2}{n_1} + \frac{\text{SD}_2^2}{n_2}}$
Difference in Proportions	$\sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2}}$

Formula

$$CI = (\mu_1 - \mu_2) \pm t * \sqrt{\frac{SD_1}{n_1} + \frac{SD_2}{n_2}}$$

College Debt Activity

Write your final confidence interval interpretation on the white board for your group.

In what cases would we want a single confidence interval vs. a difference in means confidence interval?

Summary

- For a research question asking for an estimate, the best way to answer is with a confidence interval
- The confidence interval allows us to take into sampling account variability