EPSY 5261 : Introductory Statistical Methods

Day 17

Confidence Intervals for a Single Mean

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 mean
  - Interpret a confidence interval
  - Explain how the confidence level we choose affects our interval

Confidence Intervals

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

→ Confidence Intervals

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-interval for single mean

- Assumptions
  - Sample size is large enough (>30)

OR

  - Data comes from a population with a normal distribution
    - For small samples, we can proceed if the distribution of the sample looks reasonably normal
    - In practice, better to use a simulation method to get the standard error

Formula

$$CI = \bar{x} \pm t * SE$$

Table 17.1 in text

studied in EPsy 5261.

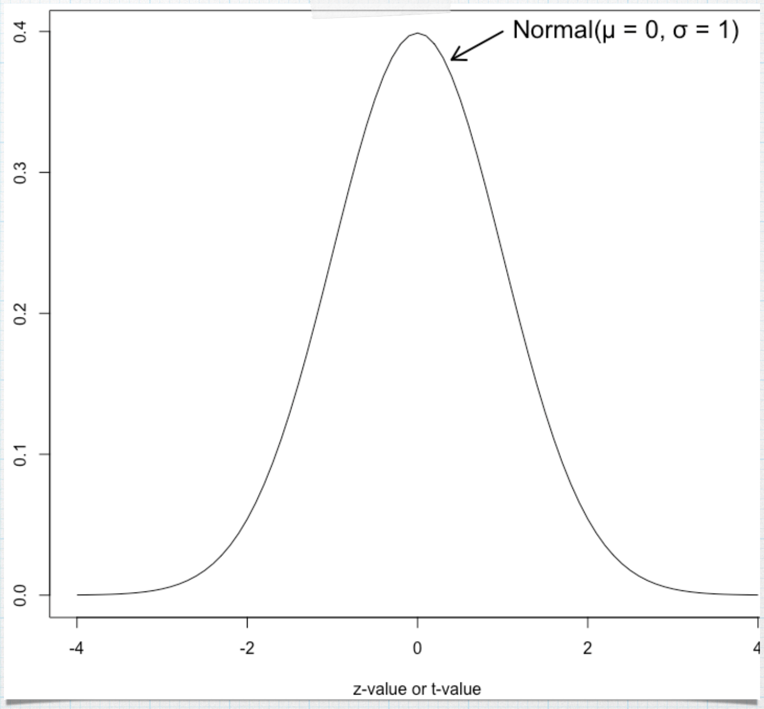
Situation	SE
Single Mean	$\dfrac{\text{SD}}{\sqrt{n}}$
Single Proportion	$\dfrac{\hat{p}(1-\hat{p})}{\sqrt{n}}$
Difference in Means	$\sqrt{\dfrac{\text{SD}_1^2}{n_1}+\dfrac{\text{SD}_2^2}{n_2}}$
Difference in Proportions	$\sqrt{\dfrac{\hat{p}_1(1-\hat{p}_1)}{n_1}+\dfrac{\hat{p}_2(1-\hat{p}_2)}{n_2}}$

Formula

$$CI = \bar{x} \pm t * \frac{SD}{\sqrt{n}}$$

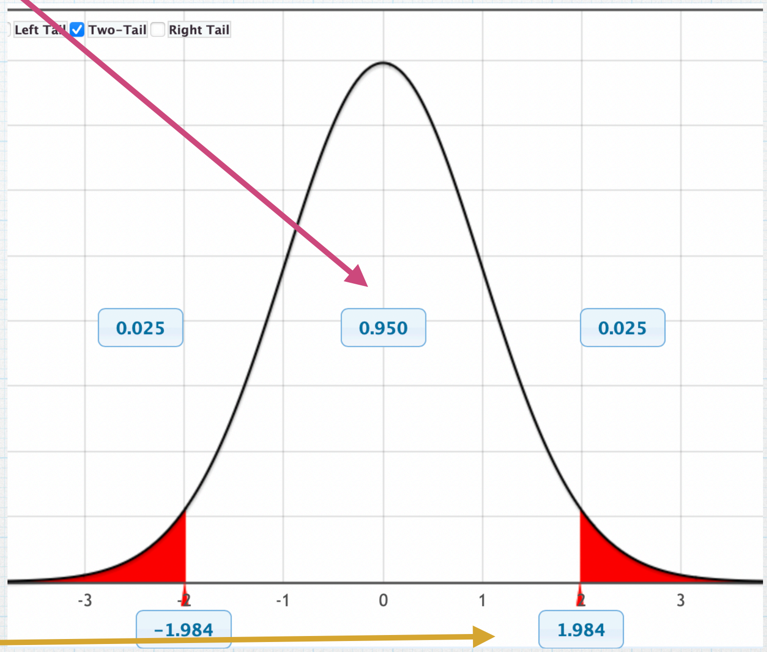
What is t*?

* Recall the t-distribution (same one as used for t-test)
* Use this to find the t* value based on the desired confidence level



For example 95% confidence

- Recall that the shape of the t-distribution is based on degrees of freedom (basically our sample size, n-1)
- To get 95% confidence for our estimate we need to look at how many standard deviations away from the mean we need to be to obtain that level of confidence
- t* = 1.984



T-distribution with sample size 100

Body Temperature Activity

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

What was the relationship between the confidence level and the 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
- With a higher confidence level we expect a larger confidence interval (more uncertainty in the estimate).