CONFIDENCE INTERVALS

Here we will calculate and plot 95% confidence intervals (CIs) for some of our helicopter data and use CIs to make some inferences. Then we evaluate the *post hoc* power of our experiment and [more usefully] *a priori* power for a subsequent experiment.

Confidence Intervals

A confidence interval is key to inference because it describes a range of expected values. A 95% confidence interval (which is standard) says there is a 95% probability (or that we are 95% confident) that the mean appears within those limits. And thus values outside those limits are improbable for that distribution.

Example 1: 0.5 $\pm$ 0.7 (mean $\pm$ 95% CI) includes zero in that interval. The mean is not significantly different from zero.

Example 2. A mean = 1.2 $\pm$ 0.3 (95% CI). A value of 0.8 lies outside that CI, and is unlikely ($p \leq$ 5%) to be obtained with new samples of the data.

Confidence intervals should be your default error bar as opposed to standard errors, standard deviations, etc. Why? Because they translate directly to inference – consider two means with 95% CIs, where CI's do not overlap. What can you infer? Could you say anything like that if SEs or SDs were graphed instead?

1. Import and attach our helicopter data from https://sciences.ucf.edu/biology/d4lab/wp-content/uploads/sites/125/2019/09/helicopter-data.txt

2. then go to this link using your web browser: http://sciences.ucf.edu/biology/d4lab/wp-content/uploads/sites/125/2017/01/CIs.R.txt then copy and paste that file into a new script file and save as an R script called CIs.R (This works around a !@#$% security block to upload an R file into the campus server)

3. Run commands in that file, following along to the comment line instructions (#...). BUT NOTICE: You need to edit the file to match OUR data. For example, this year the Design column was called ID. This is practice for re-using scripts (that you make or get online).

   What does the plot of means and 95% CIs suggest about our helicopter data?

   How would interpretations differ if you use SDs or SEs?
   Which is more honest and best matches other stats we have done?

Now try this on some other data set we have played with in class – cars, wheat, whatever.

It is really important that you present variation properly and plainly. And that you know how to interpret overlapping error bars, means, etc. And USE 95% CIs ! We will later use 95% CIs for error around regressions, too.