

WU #20

Math 58B, Spring 2023

Tuesday, April 11, 2023

Your Name: _____

Names of people you worked with: _____

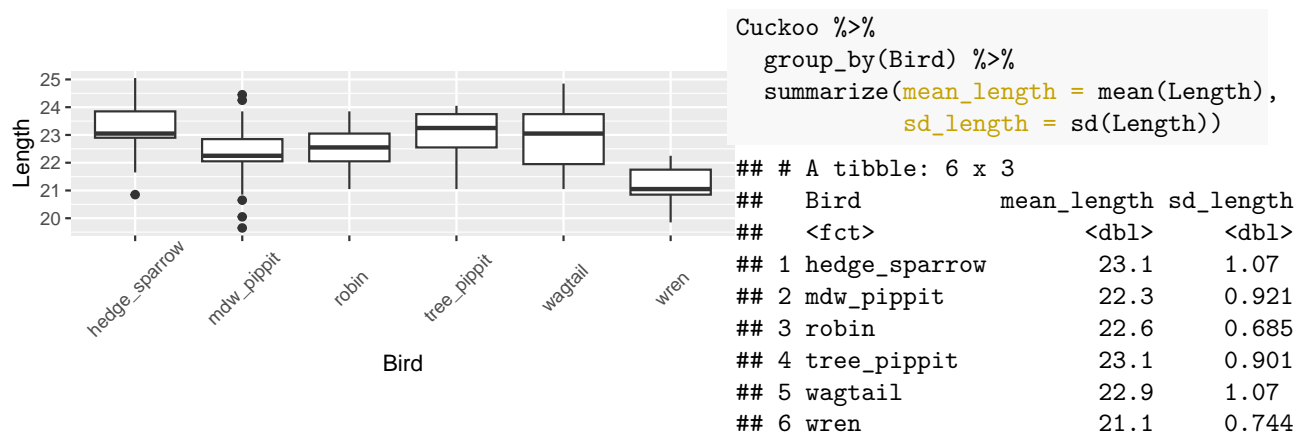
1. What are you most excited about doing this summer?
2. Where is a bootstrap distribution centered? (Your answer could be generic for any statistic, but if it is easier for you to think about, feel free to answer when considering the bootstrap distribution associated with the statistic $\bar{X}_1 - \bar{X}_2$.)
3. Consider a situation where you want to test whether the cuckoo egg lengths differ across host species of bird (on average). That is, you'd like to test the null hypothesis:

$$H_0 : \mu_{HS} = \mu_{MP} = \mu_{RB} = \mu_{TP} = \mu_{WG} = \mu_{WN}$$

The test statistic used when comparing two means was a t score.

$$t \text{ score} = \frac{(\bar{X}_1 - \bar{X}_2) - 0}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$

How might you generalize the t score so as to compare differences in sample means in the numerator and variability within a sample in the denominator. Brainstorm with your peers and try to come up with a single test statistic for the cuckoo bird problem at hand.



3. Solution:

The ANOVA test statistic is:

$$F = \frac{\frac{\sum_{i=1}^I n_i (\bar{x}_i - \bar{x})^2}{I-1}}{\frac{\sum_{i=1}^I s_i^2 (n_i - 1)}{N-I}}$$

The key to understanding the ANOVA F test statistic is recognizing that the numerator compares how different the sample means are and the denominator summarizes the extent of the variability within each group.