In-Class Activity 2: ANOVA

Hriday Baghar

2/2/2022

This exercise concerns the Morely series of experiments on the speed of light. The data are in a built-in R data set: 'morley'. The data are from 5 distinct experiments. The data from the first experiment was used in Lecture 3 to test the null hypothesis that the mean for the first experiment is equal to 792. The goal of this exercise is to use all 5 experiments to assess whether the experiments were unbiased overall, i.e., whether the overall mean is equal to 792 and to assess whether there were differences between experiments.

1. An incorrect approach is to treat the data from all 5 experiments as one sample and apply the 1-sample t-test. Why is this wrong?

2. Given that the approach in Q1 is wrong, we might try testing the null hypothesis for each experiment separately. What test would be appropriate?

3. What are the results?

4. This involves performing 5 tests so we need to be concerned about type I error inflation. How does this affect our conclusion from these test results?

5. Now suppose that we want to test the null hypothesis that the means for all experiments are equal. First create some descriptive statistics for this question.

6. What test would be appropriate to test the null hypothesis of equal means?

7. Do the assumptions for this test appear to be met?

8. Should you reject the null hypothesis with a significance level of 0.05?

9. What is the p-value for the test?

10. Now suppose that we wanted to determine which experiments have different means than others. One way to do this is to compare all possible pairs of means. What test would be used?

11. Perform this analysis.

12. What are the conclusions?

13. How do multiple testing issues affect the conclusions that can be drawn from this analysis?