ISyE 6414 Regression Analysis
Homework 4
Fall 2022

Refer to the file "homework04data01.csv", which can be found in the homework section of Canvas, to answer Questions 1–4. This data set consists of antler-length data gathered by zoologists. They are interested in seeing if we can use antler length (in inches) to classify four species (a, b, c, or d for ease of notation).

1. Produce box plots, and based on those box plots, tell whether the four species' antler lengths are different or not.

2. Use analysis of variance to test if there is a difference in mean antler length based on species. Use $\alpha = 0.05$.

3. From the Analysis of Variance results, identify and report SSE, SSTr, MSE, MSTr.

4. Construct 95% confidence intervals for each pair of species by using Tukey multiple comparisons of means. Do any of the pairings imply a difference in average antler length between species?

Refer to the file "homework04Hospital.csv", which can be found in the homework section of Canvas, to answer Questions 5–9. This data set presents data concerning the need for labor in 17 hospitals. The main objective of the regression analysis is to evaluate the performance of the hospitals in terms of how many labor hours are used relative to how many labor hours are needed. Here:

y: monthly labor hours required
x_1: monthly X-ray exposures
x_2: monthly occupied bed days
x_3: average length of patients' stays (in days).

5. Construct and solve a multiple linear regression model using this dataset. Identify and report each coefficient (β estimate).

6. Interpret each coefficient.

7. Conduct a hypothesis test using $\alpha = 0.05$ to see whether or not the predictor BedDays is statistically significant. State the null and alternative hypotheses, the test statistic, critical t-value, and your conclusion.

8. Identify the p-value for testing the predictor Xray. By using the p-value and $\alpha = 0.05$, state whether or not Xray is statistically significant. Would your answer be different if α were 0.01?

9. Identify the p-value for testing the predictor Length. By using the p-value and $\alpha = 0.05$, state whether or not Length is statistically significant. Would your answer be different if α were 0.10?