Math 4650: Final Project

Final project is due on Thursday at 1:30 p.m., April 29, 2021.

Final Exam is on Thursday at 1:30 p.m., April 29, 2021.

In November of 1969, with the Vietnam War raging, President Nixon signed an executive order instructing the Selective Service to reinstitute the draft. The order stipulated that the selection be a random process based on the birthdays of men born between January 1, 1944 and December 31, 1950. The order did not specifically state how the birthdays should be selected. The days of the year (including February 29) were written on slips of paper. The Selective Service placed these pieces of paper in separate plastic capsules that were mixed in a shoebox and then dumped into a deep glass jar. Capsules were drawn from the jar one at a time. On December 1, 1969, Rep. Alexander Pirnie, R-NY, drew the first capsule that contained the number 258 corresponding to the date September 14, so all registrants with that birthday were assigned lottery number 1. The second number drawn was 115 corresponding to April 24, and so forth. All men of draft age (born between January 1, 1944 and December 31, 1950) who shared a birth date would be called to serve at once. The first 195 birthdates drawn were later called to serve in the order they were drawn; the last of these was September 24. The fairness of the draft lottery was immediately criticized. Critics contended that the lottery process was not truly random.

The task of this project is to analyze the data set available on canvas. The file name is *1970lottery.csv*. You have been asked to provide a statistical analysis of the data and write a brief report on your findings and conclusions. **Your analysis and report should include at least the following parts:**

1. Plot draft number versus day number. Do you see any trend from this scatter plot?

2. Fit a simple linear regression of draft number on day number. To assess if the simple linear regression model is appropriate for the draft lottery data, plot the residuals and obtain the Q-Q plot of the residuals.

3. Does your analysis suggest that your fitte model is appropriate? If not, can you find a suitable transformation that improves the model in part 2?

4. Provide side-by-side boxplots of draft number for each month. How do you interpret your findings from these boxplots?

5. To test if the mean draft number for each month is the same for all months, perform the analysis of variance and related multiple comparisons (if you deem necessary) for the draft lottery data.

You may work on this project by yourself or in a group of not exceeding 3 people. Your written report should be typed not to exceed 6 pages. Please hand in one report for the group signed by all members. Each group member will receive the same grade and will be assumed to have participated fully in the project.

Your project must be submitted online in a pdf format to our class website on UNT canvas before 1:30 p.m. on April 29, 2021.