Week 13 Lab: Chi-Square and ANOVA

Please use the R markdown template to complete and submit this lab.

A. Chi-Square Analysis (10 pts)

Some students at Longwood University in Virginia were interested in people's taste preferences among various brands of bottled water. They collected data from community members, including a double-blind taste test where participants ranked water from four sources (tap water, and bottled water from Aquafina, Fiji, and Sam's Choice, presented in a random order). Some of the data from this study are stored in **WaterTaste** dataset, which contains responses from 100 participants, including their top (first) choice in the taste test, as well as their usual water source (bottled or tap/filtered).

1. Using R, load the data, check the variables, and find those that you will need to **create a contingency table for showing top choice of water versus their usual water source**. **Label your rows and columns.**
2. State valid null and alternative hypotheses for the chi-square test of independence.
3. Using R, conduct a chi-square test of independence with this data.
4. Based on your analysis, is there evidence that the top choices for taste preference are associated with whether or not people usually drink bottled water? If there is a significant association between these two variables, describe how they are related.

B. Analysis of Variance(20 pts)

The dataset ***Cereal*** shows the number of grams of fiber per serving for 30 different breakfast cereals from three different companies. The summary statistics are shown below. Conduct an analysis of variance test to determine whether there is a difference in mean number of grams of fiber per cereal between the three companies.

1. Summarize your data.

a. Create a table of summary statistics, showing the sample size, mean, and standard deviation for each company.

b.Create a plot of side-by-side boxplots. Comment on whether you are seeing any potential differencesin the mean number of grams of fiber per serving, based on the three different companies.

2. Check to see if the conditions for ANOVA are met. Interpret your results in the context of this problem, using complete sentence(s). State any limitations to using the traditional ANOVA test.

3. State the hypotheses for this ANOVA using either words or symbols.

4. Perform the ANOVA and two way comparisons.

a. Interpret the results of your ANOVA table in the context of this problem, using complete sentence(s).
b. Interpret the results of your two-way comparisons in the context of this problem, using complete sentence(s).

5. Perform an analysis of the power of this test. Comment on the power of this test. What could be done to improve the power? What other limitations, if any, do you note in this test?

6. Refer to your answers in #2 and 5. Comment on whether or not a non-parametric test would be a better option for this data. Why or why not?