RIT

STAT 614 Applied Statistics

Spring 2023

Homework #9

ANOVA

1. Consider the following partial ANOVA table:

Source of Variation	*df*	*SS*	*MS*	*F*
Treatment	*	75	25	*
Error	*	*	*	
Total	19	135		

For each sub-question show your work.

   a. Determine the df for treatment
   b. Determine the df for error
   c. Determine the SS for error
   d. Determine the MS for error
   e. Determine the F-statistic
   f. Use the rejection region method (find the critical value) to test the hypothesis of equal means

2. The data in the .csv file provided show the measurements of hemoglobin (g per 100 ml) in the blood of brown trout. The trout were placed at random in four different troughs. The fish food added to the troughs contained, respectively, 0, 5, 10, and 15 grams of sulfamerazine per 100 pounds of fish. The measurements were made on 10 fish randomly selected from each trough after 35 days.
   a. Create a summary plot of the data and comment
   b. Write down the null and alternative hypothesis for the test to see if there are any differences in mean hemoglobin among the different amounts of sulfamerazine added.
   c. Test this hypothesis using an ANOVA and $\alpha=0.05$ and draw conclusions. (Note: If using R, then treat the numeric factor as as.factor(varname) in the ANOVA)
   d. Did the outcome match what you thought would happen based on your plot, explain

3. Six observations of each of four types of cereal grain grown in a certain region were

analyzed to determine thiamin content. The data (ug/g) is provided in the .csv file.

a. Create a summary plot of the data and comment
b. Write down the null and alternative hypothesis for the test to see if there are any differences in mean thiamin content among the different types of grains.
c. Test this hypothesis using an ANOVA and $\alpha=0.05$ and draw conclusions.
d. Did the outcome match what you thought would happen based on your plot, explain.

4. An experiment was performed to investigate popcorn yield. Two factors were involved, the type of popcorn (gourmet or plain) and the size of the batch (small and large). The response was yield. The data is provided in a .csv file.
    a. Run the full two-factor ANOVA. Which effects are significant?
    b. Give an interpretation of the interaction effect.
    c. Identify the conditions that give the best outcome (highest yield) based on your analysis. Provide a graphic that help convey the results