

ADM 2304 - Assignment 3

Renad Gharz

16/05/2022

1. Midterm Scores

A Business School offers multiple sections of a second-year introduction to statistics course, which is taught by three instructors. Their names are Dr. Who, Dr. Jones (Jr.), and Prof. X. The dataset **Midterm Scores** contains the midterm scores of a random sample of students who took the course with one of these instructors over the last several years. The Business School would like to determine whether the average midterm score differs among these three instructors.

1.a

Create a side-by-side boxplot of the data and explain whether the similar variance and the nearly normality conditions for conducting an ANOVA seem to be satisfied.

1.b

In addition to a side-by-side boxplot, what other graphs can you use to check whether the conditions for using an ANOVA are satisfied? Note: You don't need to produce the graphs; only explain how you would produce them.

Answer the following questions, with exception of part g), assuming that the conditions for conducting an ANOVA are satisfied.

1.c

Use software to calculate the sample variance for each instructor and then use it to calculate the pooled variance manually. Verify that your pooled variance value is the same as the MSE value displayed in the partial ANOVA table in part d) below.

1.d

Fill in manually the missing values (1) to (6) in the ANOVA table below. Show your computations (maximum of 2 decimal places). Check your results using software.

1.e

Use the one-way ANOVA table you produced in part d) to test whether there is a significant difference in the true mean midterm score among the three instructors. Use the critical value approach and a 5% confidence level.

1.f

Use the Bonferroni method for multiple comparisons to determine which population means differ (if any) at $\alpha = 0.05$. Show your computations and clearly state your conclusion for each pairwise comparison.

1.g

Perform a Kruskal-Wallis non-parametric test to determine whether there is a difference in the midterm scores across the three instructors. Use a 5% significance level and the critical value approach. Is your conclusion consistent with your results in part e) above?

2. Utility Bills

Supposed that, when comparing utility bills, a researcher was interested in determining whether residential utility bills differed among different cities in Canada, and also whether they differed depending on the number of bedrooms in a house. As part of the study, a random sample of households in different Canadian cities was selected, and their monthly utility bills and number of bedrooms were recorded. The data is provided in the dataset **Utility Bills**.

2.a

Plot the residuals against the fitted values corresponding to the two-way ANOVA model for this analysis. What two key model assumptions can be examined with this plot and do they appear to be warranted?

2.b

Test the following hypothesis at the 1% significance level. Use software to generate the corresponding two-way ANOVA table, but show any other computations.

- Significant interaction effect between number of bedrooms and city;
- Significant main effect of the number of bedrooms factor (if warranted);
- Significant main effect of the city factor (if warranted).

2.c

2.d

2.e