**Logo

Description automatically generatedIntermediate Analytics**

**Module 5 Non-Parametric Statistical Methods and Sampling**

**ALY6015, Spring 2022**

**Week-5**

**Professor: Roy Wada**

**Submitted by: Abhinav Jain**

**Date: 05/19/2022**

Introduction

**Non-Parametric Statistical Method:** It's an approach in which data isn't assumed to come from pre-defined models with a limited number of parameters.

Non-parametric approaches have the following advantages:

One of the authentic findings is achieved when the data contains varying amounts of variability; nevertheless, non-parametric tests data need not to be normally distributed to take a test.

Non-parametric approaches have several drawbacks.

The main drawback of nonparametric machine learning algorithms is that they require more training for the data to map the functions and are relatively sluggish in training the variables if they are large.

This report used a hypothesis test which includes a Kruskal-Wallis test, Wilcoxon rank-sum test, a signed-rank test, and, a sign test. Furthermore, I calculated the Spearman rank, the test for correlation coefficient, and various techniques by using simulation techniques.

We begin with the sign test, then we have problems with Wilcoxon rank-sum test, signed-rank test, and the most intresting Kruskal-Wallis test were all utilized in this study. I have used simulation approaches to compute the spearman rank correlation coefficient and different methodologies.

**Problem1:**

**Table

Description automatically generated**

Solution 1#

Text, letter

Description automatically generated

Because the p-value is 1, the critical value is 0.3357, indicating that there is insufficient evidence to reject the null hypothesis. The solution suggests that there is insufficient evidence to reject the null hypothesis in the situation above.

**Problem 2:**

Graphical user interface, text, application

Description automatically generated

**Solution 2**

**Text

Description automatically generated**

# H0 (Owner hypothesis): 200 median

# H1 (alternative hypothesis): median!=200 The low p-value indicates sufficient evidence to reject the null hypothesis and accept the alternative hypothesis.

**#Problem 3**

**Text, table

Description automatically generated**

**Solution 3:**

**Text

Description automatically generated**

The test indicates that by implementing Wilcoxon rant sum test evidence against the null hypothesis is that the p-value is less than

0.05. As a result, need to reject the null hypothesis, whereas test the alternative

The hypothesis is that the length of each gender's prison term is equal.

**Problem 4:**

**Graphical user interface, text, application

Description automatically generated**

**Solution4:**

**Text

Description automatically generated with medium confidence**

This task solution shows that is insufficient evidence which allows to reject the null hypothesis and infer that the number of victories differs.

**Problem 5:**

**Graphical user interface, text, application

Description automatically generated**

**Solution**

**Text

Description automatically generated**

Solution5: Accept the alternative hypothesis since CV=16 is bigger than ws null hypothesis rejected.

Solution6: CV=117, null hypothesis rejected since CV is bigger than ws

Solution7: Do not reject the null hypothesis if CV=60 and CV are smaller than ws.

Solution8: If the CV is larger than ws, the null hypothesis is rejected.

**# Problem 5**

**Table

Description automatically generated**

**Solution:**

**A picture containing text

Description automatically generated**

In this solution, the rejection of the null hypothesis is the p-value is not significant not less than 0.05, and there is no change in the mean.

**Problem 6:**

**Table

Description automatically generated**

Solution:

Text

Description automatically generated

In this task, it is determined that the p-value is above the significance level, by analyzing the test it is not possible to reject the null hypothesis and cannot conclude that a link exists between daily commuter train passenger trips and service.

**# Problem7**

**Graphical user interface, text, application

Description automatically generated**

**Solution:**

Text

Description automatically generated

The average number of boxes required to win all four rewards is 12.4.

**# Problem 8Text

Description automatically generated**

**Solution:**

Text

Description automatically generated

To win the reward, an average of 10 tickets must be purchased.

**References:**

**[1] https://stat.ethz.ch/R-manual/R-devel/library/stats/html/binom.test.html**

**[2] Critical value for Wilcoxon one-sample signed-rank test in R**

DineshDinesh 11355 bronze badges & Karl Ove HufthammerKarl Ove Hufthammer 5

[**https://stats.stackexchange.com/questions/139317/critical-value-for-wilcoxon-one-sample-signed-rank-test-in-r**](https://stats.stackexchange.com/questions/139317/critical-value-for-wilcoxon-one-sample-signed-rank-test-in-r)