

Exam 1.

Student name: Javier Rentas

Date: 10/5/22

1) What is Descriptive and Inferential Statistics? **5 points**

Descriptive statistics is a summary of the data being used. Inferential statistics is when a hypothesis is tested. Mean, median mode, standard deviation, and variables fall under descriptive statistics

2) What is standard deviation? **5 points**

Standard deviation is the amount of variability in a data set. In other words, how far each score in the data set lies from the mean. The less the standard deviation, the less distance from the mean do the score lie.

3) What is median value? **5 points**

Median value is the value that lies exactly in the middle of a data set. Odd: $(n+1)/2$ Even: $(n/2)+1$

4) What is the mode and types of modes? **5 points**

The mode is the most frequently occurring number in the data set. A data set can have no mode, 1 mode, or multiple modes. Unimodal refers to 1 mode, bimodal refers to 2 modes, trimodal is 3 modes, and multimodal refers to 4 or more modes.

5) What's the difference between the range and interquartile range? **5 points**

The interquartile range is the range measured from the middle half of the data, disregarding the ends. While range takes in all the data points.

6) What is a normal distribution? **5 points**

Normal distribution is when the data is symmetrically distributed, with no skew. Most of the values in a normal distribution cluster around the center region and the measures of tendency (mean, median, mode) are the same. Even the slightest difference in the measures of tendency is considered to be non-normal distribution.

7) What is the empirical rule? **5 points**

The empirical rule serves to tell where most of the values lie in a normal distribution data set. It is a quick way to obtain an overview of the data and check for any outliers or extreme values.

8) What does correlation coefficient tell you? 5 points

The correlation coefficient describes the strength and relationship between the variables in a data set. It always ranges between -1 and 1. If positive, the variables change together in the same direction, if negative they change together in opposite directions. Taking the absolute value of the correlation coefficient gives the magnitude, the greater the value the stronger the correlation.

9) What are the assumptions of the Pearson correlation coefficient? 5 points

Both variables are on an interval or ratio level of measurement. The data from both the variables follow normal distribution. There are no outliers. The data is from a random or representative sample. Finally, a linear relationship between variables is to be expected.

10) What are the main assumptions of T-test? 5 points

The data is normally distributed. The groups compared have similar variance. The data is independent.

11) Calculate Paired T Test with R: 50 points

Using following 2 variables, calculate P value, and distribution plot.

x = c(1000,230,12333,3455,23,12,3,4,45,56,78)
y = c(1,2,3,4,5,6,7,8,9,10,11)

Email R program Exam1.R and Exam1.pdf plot to get full points.

Upload Exam1.R and Exam1.pdf plot to your Git repository.