

Exam 1.

Student name: Makaila Barnes-Powers

Date: 10/5/22

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

The properties or qualities of a dataset can be described using descriptive statistics. Descriptive statistics may be defined in terms of both the process of drawing conclusions from a collection of quantitative data and the specific quantitative observations themselves. A population can be described using descriptive statistics, either as a whole or as a particular sample. Since descriptive statistics are just explanatory, they do not become too engaged with the differences between the two types of data.

Inferential statistics' objective is to extrapolate inferences regarding a broader population from a sample that is thought to be representative of that population. Since its primary objective is to make predictions rather than present facts, inferential statistics' results often take the form of probabilities. Inferential statistics' reliability significantly depends on the correctness of the data and population representation.

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

Considering all the data, standard deviation shows the variation between a set of values and their mean. Numerous statistical tests rely heavily on this technique. Data with a low standard deviation tends to be concentrated around the mean, whereas data with a high standard deviation have a higher dispersion. In contrast to high or low standard deviations, the measured values are close to the mean whenever the standard deviation is near zero.

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

The median value is known as a central tendency measurement. When ordered numerically, it is the middlemost number or center value in a sequence of numbers. The median is also the value that is halfway into the set. For example, the median of 5, 6, and 7 is 6.

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

The value that consistently appears in each collection is referred to as the mode or modal value in statistics. The term "mode" also refers to the value or number in a data collection that occurs most frequently or with a high frequency. For example, the mode of 8, 9, 8, 7, 6, 8 would be 8 since it is the number that occurs most of the time.

The types of modes: unimodal (involving one mode within a set of data), bimodal (involving two modes within a set of data), trimodal (involving three modes within a set of data), and multimodal (involving four or more modes within a set of data).

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

There are two approaches to assess the range of values in a dataset in statistics. These techniques are known as the range and interquartile range. In a dataset, the range calculates the difference amongst the minimum and highest value. This approach provides you with the distribution of the whole data set. Whereas the interquartile range calculates the variation between a dataset's 25th and 75th percentiles. The dispersion of a data set's middle quartile is revealed by the interquartile range.

6) What is a normal distribution? 5 points

An arrangement of data known as the normal distribution describes a probability distribution that is symmetric around the mean. This method demonstrates that data close to the mean occur more frequently than data that is further away from the mean. Due to the probability density graph's bell-like appearance, the normal distribution is nonetheless commonly referred to as the bell curve. However, in statistics, it is also the most significant probability distribution because, when collated and graphed, a large amount of continuous data from psychology and nature exhibits this bell-shaped curve.

7) What is the empirical rule? 5 points

The 68-95-99.7 rule, often known as the empirical rule, is a statement regarding normal distributions. It reveals the location of the majority of the values in a normal distribution. The percentage of the data that is within one, two, or three standard deviations of the mean is indicated by this rule. Approximately 68% of values on a normal distribution are within one standard deviation of the mean, 95% are within two standard deviations of the mean, while approximately 99.7% are within three standard deviations of the mean.

8) What does correlation coefficient tell you? 5 points

How one variable changes in relation to another is indicated by the correlation coefficient. When the two travel simultaneously, their correlation is +1.0, indicating that they are moving in the same direction. They move in opposition to one another, as shown by a negative correlation coefficient.

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

The one that is most frequently employed in statistics is Pearson correlation. This evaluates how strongly and in what direction two variables are linearly related. It is simple to miss the assumptions of the Pearson correlation coefficient. The following presumptions are made: level of measurement, related pairings, lack of outliers, and linearity. To check whether the assumption is true for a Pearson correlation, each variable must be constant.

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

The validity and appropriate interpretation of the T-test relies heavily on a few assumptions. When performing a t-test, it is typical to make the following assumptions: uniformity of standard deviation variance, the range of measurement, suitability of sample size, random sampling, and the normality of data distribution.

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)
```

Mean: -1561.18

$\mu = 0$

$S^2 = SS/df = 138186403.64/(11-1) = 13818640.36$

$S^2M = S^2/N = 13818640.36/11 = 1256240.03$

$SM = \sqrt{S^2M} = \sqrt{1256240.03} = 1120.82$

T-value Calculation

$t = (M - \mu)/SM = (-1561.18 - 0)/1120.82 = -1.39$

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

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