STATISTICS WORKSHEET-1

Q1 to Q9 have only one correct answer. Choose the correct option to answer your question.

1. Bernoulli random variables take (only) the values 1 and 0.

a) True

2. Which of the following theorem states that the distribution of averages of iid variables, properly normalized, becomes that of a standard normal as the sample size increases?

a) Central Limit Theorem

3. Which of the following is incorrect with respect to use of Poisson distribution?

b) Modeling bounded count data

4. Point out the correct statement.

d) All of the mentioned

5. _____ random variables are used to model rates.

d) All of the mentioned

6. Usually replacing the standard error by its estimated value does change the CLT.

b) False

7. Which of the following testing is concerned with making decisions using data?

b) Hypothesis

8. Normalized data are centered at_____and have units equal to standard deviations of the original data.

a) 0

9. Which of the following statement is incorrect with respect to outliers?

c) Outliers cannot conform to the regression relationship

Q10and Q15 are subjective answer type questions, Answer them in your own words briefly.

10. What do you understand by the term Normal Distribution?

Normal Distribution known as Gaussian Distribution is a probability distribution that is symmetric about the mean are more frequent in occurrence than far from the mean.

11. How do you handle missing data? What imputation techniques do you recommend?

One way of handling missing values is the deletion of the rows or columns having null values. If any columns have more than half of the values as null then you can drop the entire column. In the same way, rows can also be dropped if having one or more columns values as null.

Regression imputation imputation techniques is used for missing data.

12. What is A/B testing?

A/B testing also known as split testing ,refers to a randomized experimentation process where in two or more version of a variable(web page, page element,ect.) are shown to different segments of website visitors at the same time to determine which version leaves the maximum impact and drives business metrics.

13. Is mean imputation of missing data acceptable practice?

The process of replacing null values in a data collection with the data's mean is known as mean imputation

14. What is linear regression in statistics?

Linear regression analysis is used to predict the value of a variable based on the value of another variable. The variable you want to predict is called the dependent variable. The variable you are using to predict the other variable's value is called the independent variable.

15. What are the various branches of statistics?

There are three branches of statistics . Data Collection, Descriptive Statistics and Inferential statistics.

In description statistics Data and collection data or collection, Data are described in a summarized way , where as inferential statistic . It will work large scale of work.