**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.

a) The exponent of a normally distributed random variables follows what is called the log- normal distribution

b) Sums of normally distributed random variables are again normally distributed even if the variables are dependent

c) The square of a standard normal random variable follows what is called chi-squared distribution

**d) All of the mentioned**

5. _____ random variables are used to model rates.

**c) Poisson**

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

**b) False**

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

**b) Hypothesis**

8. 4. 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?

>> The data is said to be Normally Distributed if the data is symmetrical about its mean value.

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

Missing data can be handled in many ways depending on the nature of the data. Imputation is a technique in which the missing value is replaced with some kind of value. Some of the imputation techniques are Mean and Median Imputation, Pairwise deletion, Complete Case Analysis, KNN Imputation, Regression Imputation, etc.

12. What is A/B testing?

A/B testing is used to compare two or more versions of something (for example, product feature, etc) to see which one performs better. It involves randomly assigning different groups to each version and measuring the outcome using statistical analysis. This helps organizations make data-driven decisions and optimize their strategies based on actual results.

13. Is mean imputation of missing data acceptable practice?

Yes

14. What is linear regression in statistics?

Linear Regression is a method to predict the dependent variable using a best fit line based on the independent variables. It shows significant relationships between the dependent & independent variables.

15. What are the various branches of statistics?

Inferential Statistics

Descriptive Statistics