STATISTICS WORKSHEET-1

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

Ans - 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?

Ans - a) Central Limit Theorem

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

Ans- b) Modeling bounded count data

4. Point out the correct statement.

Ans- d) All of the mentioned

5. _____ random variables are used to model rates.

Ans- c) Poisson

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

Ans- b) False

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

Ans- b) Hypothesis

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

Ans- a) 0

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

Ans - c) Outliers cannot conform to the regression relationship

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

Ans- It’s a probability distribution that is symmetric about the mean, showing that data near the mean are more frequent in occurrence than data far from the mean. In graph form, normal distribution will appear as a bell curve. It's also known as Gaussian Distribution.

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

Ans - To Handle missing date we can use deletion methods to eliminate missing data or use regression analysis to systematically eliminate data or data imputation techniques.

Best imputation technique is Multivariate Imputation by Chained Equations

12. What is A/B testing?

Ans- A/B testing is a method of comparing two versions of a webpage or app against each other to determine which one performs better. Its’s also known as split testing or bucket testing

13. Is mean imputation of missing data acceptable practice?

Ans – No it's not a good solution since mean imputation ignores feature correlation

14. What is linear regression in statistics?

Ans - In statistics, linear regression is a linear approach for modelling the relationship between a scalar response and one or more explanatory variables.

15. What are the various branches of statistics?

Ans - data collection, descriptive statistics and inferential statistics.