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

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

Ans: b) False

7. 1. 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

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

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

Ans: The normal distribution is a continuous distribution that is symmetrical on both sides of the mean. For proper normal distribution the mean, median,mode will be the same value. Mean=Median=Mode

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

Ans: Missing data reduce the statistical power of a study and can produce biased estimates, leading to invalid conclusions. Missing data is defined as the data value that is not stored for a variable. We can use Mean or Median Imputation techniques to handle missing data.

12. What is A/B testing?

Ans: A/B testing is a basic randomized control experiment. It is used to compare the two versions of a variable to find out which performs better in a controlled environment.

13. Is mean imputation of missing data acceptable practice?

Ans: Mean imputation is acceptable practice; however, this method should be avoided as far as possible because – 1. Mean imputation does not preserve the relationships among variables. 2. Mean Imputation Leads to An Underestimate of Standard Errors

14. What is linear regression in statistics?

Ans: Linear regression in statistics is a linear approach for modelling the relationship between s dependent and independent variables

15. What are the various branches of statistics?

Ans: The two major areas of statistics are known as descriptive statistics, which describes the properties of sample and population data Inferential statistics, which uses those properties to test hypotheses and draw conclusions