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.

Answer: 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?

Answer: a) Central Limit Theorem

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

Answer: b) Modeling bounded count data

4. Point out the correct statement.

Answer: d) All of the mentioned

5. _____ random variables are used to model rates.

Answer: (c) Poisson

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

Answer: b) False

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

Answer: (b) Hypothesis

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

Answer: a) 0

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

Answer: (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?

Answer: A normal distribution is a continuous probability distribution method in which most data points cluster toward the middle of a range, and the rest are positioned symmetrically towards the extremes.

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

Answer: Imputation is the best way to deal with missing data.

12. What is A/B testing?

Answer: A/B testing is a statistical method to find the most desired set of variants. Here, we compare the results of different sets of data & find the most suitable one out of them.

13. Is mean imputation of missing data acceptable practice?

Answer: Mean Imputation is not acceptable everytime while dealing with missing data, as in mean imputation, the mean of the available data of each variable in the dataset is calculated and is placed in place of missing values of label data.

14. What is linear regression in statistics?

Answer: Linear Regression is a statistical method used to predict the value of a variable, on basis of values of similar variables.

15. What are the various branches of statistics?

Answer: The branches of statistic are descriptive statistics and inferential statistics.