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. 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. 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 probability distribution that is symmetrical on both sides of the mean, so the right side of the center is a mirror image of the left side. The area under the normal distribution curve represents probability and the total area under the curve sums to one.

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

Ans: - Best techniques to handle missing data:

1) Use deletion methods to eliminate missing data.
2) Use regression analysis to systematically eliminate data.
3) Data scientists can use data imputation techniques.
4) Keeping things under control.

Imputation Techniques:

1) Complete Case Analysis (CCA).
2) Arbitrary Value Imputation.
3) Frequent Category Imputation.

12. What is A/B testing?

Ans: - A/B testing is a basic randomized control experiment. It is a way 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: - True, imputing the mean preserves the mean of the observed data. So, if the data are missing completely at random, the estimate of the mean remains unbiased.

14. What is linear regression in statistics?

Ans: - Regression analysis is a statistical method that helps us to analyze and understand the relationship between two or more variables of interest.

15. What are the various branches of statistics?

Ans: - (1) Descriptive Statistics. (2) Inferential Statistics.