STATISTICS WORKSHEET-1

Q1 to Q9 have only one correct answer. Choose the correct option to answer your question.

ANSWERS 1 TO 9

1. Bernoulli random variables take (only) the values 1 and 0.
   a) **True**
   b)

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**
   b)
   c)
   d)

3. Which of the following is incorrect with respect to use of Poisson distribution?
   a)
   b) **Modeling bounded count data**
   c)
   d)

4. Point out the correct statement
   a)
   b)
   c) **The square of a standard normal random variable follows what is called chi-squared distribution**
   d)

5. _____random variables are used 0 model rates.
   a)
   b)
   c) **Poisson**
   d)

6. Usually replacing the standard error by its estimated value does change the CLT.
   a)
   b) **False**

7. Which of the following testing is concerned with making decisions using data?
   a)
   b) **Hypothesis**
   c)
   d)

8. Normalized data are centered at_____ and have units equal to standard deviations of the original data.
   a)
   b)
   c) **1**
   d)

9. Which of the following statement is incorrect with respect to outliers?
   a)
   b)
   c) **Outliers cannot conform to the regression relationship**
   d)

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

ANSWERS 10 TO 15

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

ANS. The normal distribution is a probability distribution that is often used in statistics to model continuous random variables that have a symmetric bell-shaped distribution.
The total area under the curve of the normal distribution is equal to 1, meaning that the probabilities of all possible outcomes add up to 1.

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

ANS. The choice of imputation technique depends on the nature of the data, the amount of missing data, and the research question. However, multiple imputation Generate several plausible values for each missing value and incorporate the uncertainty in the imputation process into the analysis.

12. What is A/B testing?

ANS. A/B testing is a statistical method used to compare two versions of a product or service to determine which one performs better. It involves randomly assigning participants to one of two groups, where each group is shown a different version of the product or service, and then measuring the outcome of interest, such as the conversion rate, click-through rate, or user engagement.

13. Is mean imputation of missing data acceptable practice?

ANS. Mean imputation of missing data is a simple and commonly used method of imputation. However, it has some limitations and can introduce bias and inaccuracies in the data analysis. Therefore, its acceptability depends on the context of the analysis and the nature and extent of missing data.

14. What is linear regression in statistics?
Ans. Linear regression is a statistical method used to model the relationship between a dependent variable (also known as the response variable) and one or more independent variables (also known as explanatory or predictor variables). It is used to predict the value of the dependent variable based on the values of the independent variables.

15. What are the various branches of statistics?
Ans. Statistics is a broad field with several branches that specialize in different aspects of data analysis and modelling. Some of the main branches of statistics are descriptive statistics, inferential statistics, probability theory, bayesian statistics, time series analysis, multivariate statistics etc…