STATISTICS WORKSHEET-1

- **Please find the answers at the end as this file is an Image format and can not be edited.**

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.
   a) True
   b) False
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) Central Mean Theorem
   c) Centroid Limit Theorem
   d) All of the mentioned
3. Which of the following is incorrect with respect to use of Poisson distribution?
   a) Modeling event/time data
   b) Modeling bounded count data
   c) Modeling contingency tables
   d) All of the mentioned
4. Point out the correct statement.
   a) The exponent of a normally distributed random variables follows what is called the log- normal distribution
   b) Sums of normally distributed random variables are again normally distributed even if the variables are dependent
   c) The square of a standard normal random variable follows what is called chi-squared distribution
   d) All of the mentioned
5. _____ random variables are used to model rates.
   a) Empirical
   b) Binomial
   c) Poisson
   d) All of the mentioned
6. 10. Usually replacing the standard error by its estimated value does change the CLT.
   a) True
   b) False
7. 1. Which of the following testing is concerned with making decisions using data?
   a) Probability
   b) Hypothesis
   c) Causal
   d) None of the mentioned
8. 4. Normalized data are centered at_____and have units equal to standard deviations of the original data.
   a) 0
   b) 5
   c) 1
   d) 10
9. Which of the following statement is incorrect with respect to outliers?
   a) Outliers can have varying degrees of influence
   b) Outliers can be the result of spurious or real processes
   c) Outliers cannot conform to the regression relationship
   d) None of the mentioned

Answers:-

1) True
2) a) Central Limit Theorem
3) b) Modeling bounded count data
4) c) The square of a standard normal random variable follows what is called chi-squared
5) c) Poisson
6) b) False
7) b) Hypothesis
8) a) 0
9) 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- It is a continuous probability distribution that is symmetric about its mean. It is characterized by two parameters: the mean and the standard deviation.

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

Ans- We can handle missing data by removing column or a raw. Below are techniques for imputation

1. Mean/Median/Mode imputation, 2 KNN, 3 Multiple, Regression

12. What is A/B testing?

Ans- It is a method used to compare two versions of a product or service to determine which one performs better.

13. Is mean imputation of missing data acceptable practice?

Ans- Mean imputation of missing data is a common practice and is often used because it is simple to implement and can preserve the overall distribution of the data.

14 What is linear regression in statistics?

Ans- Linear regression is a statistical practice of calculating a straight line that specifies a mathematical relationship between two variables.

15 What are the version branches of statistics?

Ans- The two main branches of statistics are 1. descriptive statistics and 2. inferential statistics: