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.
a) True
b) False

Answer (a)

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

Answer (A)

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

Answer (B)

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

Answer (D)

5. _____ random variables are used to model rates.
a) Empirical
b) Binomial
c) Poisson
d) All of the mentioned

Answer (C)

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

Answer (B)

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

Answer (B)

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

Answer (A)

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

Answer (C)

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 It is also called the Gaussian Distribution. It is used normally in continuous distribution. In this every event is independent to one another. In this Bell-shaped curve is formed i.e., Mean=Median=Mode

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

Missing data can be handled by using the Imputation techniques, in this we fill the missing data in the dataset without reducing the actual size of the dataset by increasing the column size. Imputation technique are Numerical Variables, Categorical Variables and both.

12. What is A/B testing?
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?
Yes, it is acceptable.

14. What is linear regression in statistics?
Linear regression is used to predict the value of a variable based on the value of another variable. The variable you want to predict is called the dependent variable. The variable you are using to predict the other variable's value is called the independent variable.

15. What are the various branches of statistics?
Descriptive and Inferential are the branches of statistics.