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 for a real-valued random variable.

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

Answer :- Missing data can be handled by using imputation or data removal method. Some of the imputation techniques are:-

- Next or Previous value.
- K nearest neighbors.
- Maximum or Minimum value.
- Missing value prediction.

12. What is A/B testing?

Answer :- A/B testing is a statistical analysis framework that compares two verson of product, algorithm, or user experience design to determine which one performs batter.

13. Is mean imputation of missing data acceptable practice?

Answer :- If the data are missing completely at random, the estimate of the mean remains unbiased. And, by imputing the mean, you are able to keep your sample size up to the full sample size. So mean imputation of missing data is acceptable practice.

14. What is linear regression in statistics?

Answer :- 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?

Answer :- The two main branches of statistics are descriptive statistics and inferenital statistics.