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
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

FLIP ROBO

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

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

    Normal distribution is also called as Gaussian Distribution. It is a probability distribution that is a symmetric about the mean, showing that data near the mean are more frequent in occurrence than data far from the mean. In graph it appears as an bell curve shape.

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

    Multiple Imputation

12. What is A/B testing?

    A/B testing is also known as split testing and bucket testing, is an statistical method used in marketing, product development and other fields to compare two versions of web page, app and other elements to determine which one is perform better in terms of specific outcome or metric. A/B testing is used to make data driven decisions by comparing the impacts of different variations on user behavior or other relevant factors.

13. Is mean imputation of missing data acceptable practice?

    NO

FLIP ROBO

14. What is linear regression in statistics?

    It is a data analysis technique that predicts the value of unknown data by using other another related data and known data value.

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

FLIP ROBO