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.
   Ans: 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?
   Ans: a. Central Limit Theorem.

3. Which of the following is incorrect with respect to use of Poisson distribution?
   Ans: d. All of the mentioned
   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.
   Ans. d.All of the mentioned
   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.
   Ans: c. Poisson

FLIP ROBO

6. 10. Usually replacing the standard error by its estimated value does change the CLT.
   Ans: b. False

7. 1. Which of the following testing is concerned with making decisions using data?
   Ans: b. Hypothesis

8. 4. Normalized data are centered at_____and have units equal to standard deviations of the original data.
   Ans: a. 0

9. Which of the following statement is incorrect with respect to outliers?
   Ans: 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: A normal distribution is the proper term for a probability bell curve.
    In a normal distribution the mean is zero and the standard deviation is 1. It has zero skew and a kurtosis of 3.
    Normal distributions are symmetrical, but not all symmetrical distributions are normal.

11. How do you handle missing data? What imputation techniques do you recommend?
    Ans: a. Using deletion method to remove the missing data.
    b. replacing the missing data with mean value of the data.
    c. Using implation method

12. What is A/B testing?
    Ans:is also known as split testing, refers to a randomized experimentation process wherein two or more versions of a variable (web page, page element, etc.) are shown to different segments of website visitors at the same time to determine which version leaves the maximum impact and drive business metrics.

13. Is mean imputation of missing data acceptable practice?
    Ans:Yes, imputing the mean preserves the mean of the observed data. So if the data are missing completely at random, the estimate of the mean remains unbiased. ... Since most research studies are interested in the relationship among variables, mean imputation is not a good solution.
14. What is linear regression in statistics?
    Ans:Linear regression attempts to model the relationship between two variables by fitting a linear equation to observed data. ... A linear regression line has an equation of the form Y = a + bX, where X is the explanatory variable and Y is the dependent variable.
15. What are the various branches of statistics?
    Ans: 1. Descriptive Statistics and 2.Inferential Statistics.

FLIP ROBO