Statistics

1. Bernoulli random variables take (only) the values 1 and 0.

Ans- 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- Central Limit Theorem

3. Which of the following is incorrect with respect to use of Poisson distribution?

Ans- Modeling bounded count data

4. Point out the correct statement

Ans- All of the mentioned

5. _____ random variables are used to model rates.

Ans- Poisson

6. Usually replacing the standard error by its estimated value does change the CLT.

Ans- False

7. Which of the following testing is concerned with making decisions using data?

Ans- Hypothesis

8. Normalized data are centered at_____and have units equal to standard deviations of the original data

Ans- 0

9. Which of the following statement is incorrect with respect to outliers?

Ans- Outliers cannot conform to the regression relationship

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

Ans - Normal distribution, also known as the Gaussian distribution, is a probability distribution that is symmetric about the mean, showing that data near the mean are more frequent in occurrence than data far from the mean. In graph form, normal distribution will appear as a bell curve.

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

Ans- The best way to handle missing data is imputing the Missing Value

    1. Replacing With Arbitrary Value.
    2. Replacing With Mode.
    3. Replacing With Median.
    4. Replacing with previous value – Forward fill.
    5. Replacing with next value – Backward fill.
    6. Interpolation.

12. What is A/B testing?

Ans- An AB test is an example of statistical hypothesis testing, a process whereby a hypothesis is made about the relationship between two data sets and those data sets are then compared against each other to determine if there is a statistically significant relationship or not.

13. Is mean imputation of missing data acceptable practice?

Ans- Mean Imputation is typically considered terrible practice since it ignores feature correlation.

14. What is linear regression in statistics?

Ans- Linear regression analysis 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?

Ans- There are three real branches of statistics: Data Collection, Descriptive Statistics and Inferential Statistics.