Statistics

1. Bernoulli random variables take (only) the values 1 and 0.
    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?
    a) Central Limit Theorem

3. Which of the following is incorrect with respect to use of Poisson distribution?
    b) Modeling bounded count data

4. Point out the correct statement.
    d) All of the mentioned

5. _____ random variables are used to model rates.
    c) Poisson

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

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

8. Normalized data are centered at _____ and have units equal to standard deviations of the original data.
    a) 0

9. Which of the following statement is incorrect with respect to outliers?
    c) Outliers cannot conform to the regression relationship

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

The normal distribution, also known as the Gaussian distribution, is the most independent probability distribution in statistics for independent, random variables. Most people recognize its familiar bell-shaped curve in statistical reports. This is symmetrical around its mean, most of the observation cluster around central peak.

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

1. Deleting the record of missing values
2. creating a separate model to handle missing value
3.statistical methods like: - Mean, Median, Mode

The imputation technique I would recommend is the KNN imputer.

12. What is A/B testing?

A/B testing is basically statistical hypothesis testing, or, in other words, statistical inference. It is an analytical method for making decisions that estimates population parameters based on sample statistics.

A/B testing is one of the most popular controlled experiments used to optimize web marketing strategies. It allows decision makers to chose the best design for a website by looking at the analytics results obtained.

13. Is mean imputation of missing data acceptable practice?

Mean imputation is typically considered terrible practice since it ignores feature correlation. Second, mean imputation decreases the variance of our data while increasing bias. As a result of the reduced variance, the model is less accurate and the confidence interval is narrower.

14. What is linear regression in statistics?

Linear regression is a basic and commonly used type of predictive analysis. This 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 and the variable you are using to predict the other variables value is called the independent variable.

15. What are the various branches of statistics?

There are basically four branches of statistics:

1. Mathematical or Theoretical Statistics
2. Statistical methods or functions
3. Descriptive Statistics
4. Inferential Statistics