STATISTICS ANWERS

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-(B) Modeling Bounded Count Data

4. Point out the correct statement.

Ans-(D) All of the Mentioned

5. _____ random variables are used to model rates.

Ans- (C) All of the Mentioned

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

Ans- (B) False

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

Ans-(B) Hypothesis

8. 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

Subjective Questions and answers

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

Ans- Normal Distribution, the data occurrence is more frequent around the mean and is less far from the mean. The data is symmetrical around the means and this is also called as gaussian distribution. The standard deviations for all the values from the mean is 3.

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

Ans-In a Dataset, there is a possibility of missing data present inside it so, for handling these missing values, we use many techniques like Data Imputation Technique, Use deletion methods and use regression analysis.

Imputation techniques includes: -

1. CCA (Complete Case Analysis)
2. Arbitrary value imputation
3. Frequent category imputation

12. What is A/B testing?

Ans-A/B testing is nothing but a process of measuring the performance of two or more versions or variants of a web page, email or other marketing application. By this technique we can conclude that which variant is performing better for a given goal which results in a better product for the users.

13. Is mean imputation of missing data acceptable practice?

Ans-No, the mean imputation for missing data is not acceptable because there is a possibility of outliers in a data which will have a significant impact on mean. Using mean imputation may not create a great model.

14. What is linear regression in statistics?

Ans-Linear Regression is a statistical approach to determine the relationship between the dependent variable and independent variable. Dependent variable is a variable you want to predict and independent variable are the variable used to predict other variable's(dependent) value.

15.What are the various branches of statistics?

Ans-Statistics is a study of presentation, interpretation, collection and organization of data. There are main two branches of statistics

1.Inferetntial Statistics-It is used when it's not easy or possible to examine each member of the population.

2.Descriptive Statistics-IT is used to have a brief summary of data.it can be in numerical or graphical form.