Q1. Bernoulli random variables take (only) the values 1 and 0.

Ans: a) True

Q2. 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

Q3. Which of the following is incorrect with respect to use of Poisson distribution?

Ans: b) Modeling bounded count data

Q4. Point out the correct statement.

Ans: d) All of the mentioned

Q5. _____ random variables are used to model rates.

Ans: c) Poisson

Q6. 10. Usually replacing the standard error by its estimated value does change the CLT.

Ans: b) False

Q7. Which of the following testing is concerned with making decisions using data?

Ans: b) Hypothesis

Q8. Normalized data are centered at_____and have units equal to standard deviations of the original data.

Ans: a) 0

Q9. Which of the following statement is incorrect with respect to outliers?

Ans: c) Outliers cannot conform to the regression relationship

Q10. 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 graphical form, the normal distribution appears as a "bell curve".

Q11. How do you handle missing data? What imputation techniques do you recommend?

Ans: The missing data can be handled by:

a. Mean/Median/Mode imputation.
b. Deleting rows with missing data.
c. Using algorithm that supports missing values.

I recommend using Deleting rows with missing data as it gives an accurate value.

Q12. What is A/B testing?

Ans: A/B testing is a form of hypothesis testing and two-sample hypothesis testing to compare two versions. The control and variant, of a single variable. It is commonly used to improve and optimize user experience and marketing.

Q13. Is mean imputation of missing data acceptable practice?

Ans: Mean imputation is bad practice because it doesn't take into account feature correlation.

Q14. What is linear regression in statistics?

Ans: Linear regression is one of the statistical techniques used in predictive analysis, in this technique will identify the strength of the impact that the independent variables show on dependent variables.

Q15. What are the various branches of statistics?

Ans: Statistics have two main branches:

a. Descriptive Statistics: This usually summarizes the data from the sample by making use of an index like mean or standard deviation. The methods which are used in the descriptive statistics are displaying, organizing, and describing the data.

b. Inferential Statistics: These conclude from data which are subject to random variations like observation mistakes and other sample variation