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 c) The square of a standard normal random variable follows what is called chi-squared distribution.

5) _____ random variables are used to model rates.

Ans c) Poisson.

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) 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

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

Ans) Normal distribution, also known as (Gaussian distribution), is a probability distribution that is symmetric, bell-shaped, and continuous. It is characterized by its mean (μ) and standard deviation (σ), which determine the center and spread of the distribution, respectively. The normal distribution is a fundamental concept in statistics and probability theory, and it arises naturally in many natural and social phenomena, such as heights, weights, test scores, and errors.

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

Ans) Handling missing data is an important aspect of data analysis because missing values can affect the accuracy, validity, and reliability of statistical inferences and predictions. There are several ways to handle missing data, including:

1**.Complete case analysis**: It discards any data that has missing values in it.

2. **Pairwise deletion:** This method involves using only the available pairs of observations for each analysis, meaning that any missing data in one variable is omitted only for that particular analysis.

3**. Imputation:** his method involves replacing the missing values with estimated values based on the observed values and/or other relevant information. There are several imputation techniques.

Mean imputation: This involves replacing the missing values with the mean of the observed values of the same variable.

Median imputation: This involves replacing the missing values with the median of the observed values of the same variable.
Multiple imputation: This involves creating multiple imputed datasets based on statistical models that estimate the missing values using the observed values and other relevant predictors.

Maximum likelihood imputation: This involves estimating the missing values based on the likelihood function of the data and the missing data mechanism.

12) What is A/B testing?

Ans) A/B testing is a method of comparing two versions of a webpage, email, or any other marketing asset, to determine which version performs better in terms of a particular outcome, such as conversion rate, click-through rate, or engagement.

13) Is mean imputation of missing data acceptable practice?

Ans) Mean imputation of missing data is a common practice. it has some limitations and potential drawbacks.

Mean imputation assumes that the missing values are missing completely at random (MCAR), which means that the probability of missingness is not related to the value of the variable or any other variable in the dataset.

14) What is linear regression in statistics?

Ans) Linear regression is a statistical method used to model the linear relationship between a dependent variable and one or more independent variables. It is a commonly used technique for

predicting or estimating the value of a continuous dependent variable based on one or more independent variables.

Y = b0 + b1*X1* + *b2*X2 + ... + bn*Xn

15) 15. What are the various branches of statistics?

Ans) Types of statistics

1. Inferential statistics his branch of statistics deals with making predictions and generalizations about a population based on a sample of data.

2. Descriptive statistics: Descriptive statistics: This branch of statistics deals with the summary and analysis of data using measures such as mean, median, mode, variance, and standard deviation.