

1. a) True.
2. a) Central Limit Theorem.
3. b) Modeling bounded count data
4. c) The square of a standard normal random variable follows what is called chi-squared distribution.
5. c) Poisson
6. b) False
7. b) Hypothesis
8. a) 0
9. c) Outliers cannot conform to the regression relationship
10. Normal distribution, also known as Gaussian distribution, is a probability distribution that is symmetric, bell-shaped, and characterized by its mean ( $\mu$ ) and standard deviation ( $\sigma$ ). In a normal distribution, the mean, median, and mode are all equal and fall at the center of the distribution. Many natural phenomena, such as heights, weights, and test scores, follow a normal distribution. The normal distribution is important in statistical inference because of the central limit theorem, which states that the distribution of the sample mean approaches a normal distribution as the sample size increases.
11. Handling missing data is an important step in data analysis. There are various imputation techniques that can be used, such as mean imputation, median imputation, mode imputation, hot deck imputation, and multiple imputation. The choice of imputation technique depends on the type and amount of missing data, as well as the characteristics of the data set. It is important to carefully evaluate the impact of imputation on the results and conclusions of the analysis.
12. A/B testing is a statistical technique used in digital marketing and product development to compare two versions of a web page, advertisement, or other digital asset to determine which one performs better in terms of a predefined goal. It involves randomly assigning users to either the control group or the experimental group and then analyzing the results to determine which variation is more effective. The goal of A/B testing is to make data-driven decisions to optimize the user experience and improve business metrics such as conversion rates, click-through rates, and revenue.
13. Mean imputation is a simple technique to handle missing data by replacing missing values with the mean of the available values. However, it has limitations and may not always be the best approach, as it can introduce bias and affect the variance and correlations in the data. Other techniques, such as

regression imputation, multiple imputation, and maximum likelihood estimation, may be more appropriate depending on the data and research question.

14. Linear regression is a statistical method used to model the linear relationship between a dependent variable and one or more independent variables. It aims to find the best linear relationship between the variables, where the dependent variable can be predicted by the independent variables. The linear regression model assumes a linear relationship between the variables, normal distribution of errors, and equal variance of errors. It is widely used in various fields, including economics, social sciences, and engineering, for prediction and modeling purposes.
15. The various branches of statistics are as follows:
  - a. Mathematical or Theoretical statistics: It is the branch of statistics that deals with the development and derivation of mathematical formulas and models for statistical analysis. It involves the study of probability theory, random variables, distribution theory, and mathematical optimization techniques.
  - b. Statistical methods or functions: These are the techniques used to collect, analyze and interpret data in order to draw conclusions from it. Some of the common statistical methods include regression analysis, hypothesis testing, sampling theory, and analysis of variance.
  - c. Descriptive statistics: It is the branch of statistics that focuses on the summary and description of data. Descriptive statistics includes measures such as mean, median, mode, variance, standard deviation, and percentiles.
  - d. Inferential statistics: It is the branch of statistics that involves using data from a sample to make inferences or draw conclusions about a population. It includes techniques such as hypothesis testing, confidence intervals, and regression analysis.