1. True
2. a. Central Limit Theorem
3. b. Modeling bounded count data
4. c.
5. c. poisson
6. False
7. b.Hypothesis
8. a.0
9. c.outliers can not confirm regression relationship

10. What do you understand by the term Normal Distribution?
11. How do you handle missing data? What imputation techniques do you recommend?
12. What is A/B testing?
13. Is mean imputation of missing data acceptable practice?
14. What is linear regression in statistics?
15. What are the various branches of statistics?

10.

The term Normal distribution is also known as the Gaussian distribution, is a fundamental concept in statistics and probability theory. It is a continuous probability distribution that is characterized by a specific bell-shaped curve. The distribution is symmetrical around mean, median and mode . The distribution is described as two parameters (μ) and standard deviation (σ) where (σ) controls the spread or dispersion of the data.68% data within one standard deviation of data 95% falls two standard deviation of data and 99.7% data falls within three standard deviation of data.

11. Common approach of handling missing datas are :

a.Replace missing values with mean or median of the available data for the variable.

b.For categorical data missing value is replaced by mode of the variable.

c.K-Nearest Neighbors.

d. Percentile and Quartile impution method to fill up missing values.

12.

A/B testing is a method used to compare two versions of something, like a webpage or an ad, to see which one performs better. Users are randomly assigned to see either the original version (A) or a modified version (B), and their interactions are measured. Statistical analysis is then used to determine if the changes make a significant difference in user behaviour. A/B testing helps make data-driven decisions and improve outcomes in areas like marketing and product development.

13.

Yes, It is acceptable practice. It is simple and quick and easy imputation.

It does not require complex calculations. It does not reduce sample size .It does not introduce distortion to the mean and variance of the variable.

limitations: Mean imputation assumes a normal distribution and may not suitable for with non normal distributions.

14.

Linear regression is a statistical method used to model and analyse the relationship between two or more variables. It aims to find the straight-line equation that best fits the data points, allowing you to make predictions and understand the association between the variables.

15.

Branches of statistics include:

Descriptive: Summarizing data.

Inferential: Making predictions from samples.

Probability: Studying event likelihood.

Hypothesis Testing: Drawing conclusions.

Regression: Modeling relationships.

Time Series: Analyzing time-based data.

Multivariate: Studying multiple variables.

Nonparametric: Distribution-free methods.

Bayesian: Updating beliefs with data.

Experimental Design: Controlled experiments.

Biostatistics: Medical and biological data.

Econometrics: Economic data analysis.

Social Statistics: Societal trends.

Quality Control: Ensuring product quality.

Spatial Statistics: Analyzing spatial data.

Data Mining/ML: Extracting insights from data.