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

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?

a) Central Limit Theorem

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

b) Modeling bounded count data

4. Point out the correct statement.

Answer -- The square of a standard normal random variable follows what is called chi-squared distributio

5_____ random variables are used to model rates

c) Poisson

6. Usually replacing the standard error by its estimated value does change the CLT. a) b) False

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

b) Hypothesis

8.. Normalized data are centered at_____and have units equal to standard deviations of the original data. a) 0

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

c) Outliers cannot conform to the regression relationship

10, Normal Distribution" typically refers to the statistical concept of a normal distribution, as well as to the functionality provided by libraries like NumPy and SciPy for working with normal distributions.

11. Handling missing data is an important step in the data preprocessing phase of any analysis or modeling task. The choice of imputation technique depends on the nature of the data and the specific characteristics of the missing values. Here are some common approaches for handling missing data:

12. A/B testing, also known as split testing, is a statistical method used in marketing, product development, and other fields to compare two versions (A and B) of a variable, typically a web page, app feature, or marketing campaign, to determine which one performs better. The goal is to assess the impact of changes and make data-driven decisions about which version is more effective in achieving a desired outcome.

13, Mean imputation, where missing values are replaced with the mean of the observed values in a variable, is a simple and commonly used method for handling missing data. While it is straightforward and easy to implement, there are both advantages and limitations to mean imputation, and its acceptability depends on the context of the data and the analysis. Here are some considerations

14. Linear regression is a statistical method used for modeling the relationship between a dependent variable and one or more independent variables by fitting a linear equation to observed data. The goal of linear regression is to find the best-fitting line (or hyperplane in the case of multiple independent variables) that minimizes the sum of the squared differences between the observed and predicted values.

15. In Python, statistics is a broad field with various branches, and there are several libraries and tools specifically designed to address different aspects of statistical analysis. They are - Descriptive statistics, Inferential statistics, Bayesian statistics, Time series analysis, Multivariate statistics