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: d) All of the mentioned

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

Ans: c) Modelling contingency tables

4. Point out the correct statement.

Ans: d) All of the mentioned

5. _____ random variables are used to model rates.

Ans: c) Poisson

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

Ans: b) False

7. 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 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".

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

Ans: The first step is to form a model for the missing data. This could be the distribution of the missing data conditioned on the available data. This distribution can then be used to impute the missing data, for example via sampling from the conditional distribution.

There are Seven common Ways to Make up Data. Common Methods to Imputing Missing Data are Mean imputation, Substitution, Hot deck imputation, Cold deck imputation, Regression imputation, Stochastic regression imputation, Interpolation and extrapolation.

12. What is A/B testing?

Ans: A/B testing also known as split testing is the process of comparing two versions of a web page, email, or other marketing asset and measuring the difference in performance. You do this giving one version to one group and the other version to another group. Then you can see how each variation performs.

13. Is mean imputation of missing data acceptable practice?

Ans: Mean imputation is typically considered terrible practice since it ignores feature correlation. It reduces the variance of the imputed variables. Mean imputation shrinks standard errors, which invalidates most hypothesis tests and the calculation of confidence interval. It does not preserve relationships between variables such as correlations.

14. What is linear regression in statistics?

Ans: Linear regression analysis is used to predict the value of a variable based on the value of another variable. The variable you want to predict is called the dependent variable. The variable you are using to predict the other variable's value is called the independent variable. In statistics, linear regression is a linear approach for modelling the relationship between a scalar response and one or more explanatory variables (also known as dependent and independent.

15. What are the various branches of statistics?

Ans: There are three real branches of statistics: data collection, descriptive statistics and inferential statistics.