Statistics Worksheet - 1

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. Modelling bounded count data

4. Point out the correct statement.

Ans: d. All 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: a. True

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

Ans: b. Hypothesis

8. Normalized data are centred 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:

a. Normal distribution is a probability distribution that is symmetric about the mean, which indicates that the data near the mean is more frequent in occurrence.
b. Normal distribution will appear as a bell curve in graphical representation.
c. Mean, Median and Mode have same value for a normal distribution.
d. The tail ends of a normal distribution never touch the horizontal and extend indefinitely.

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

Ans.

a. Generally, we use mean or median of the data for the missing observations. However, it can be useful in cases where the number of missing observations is low.
b. But, when the number of missing values is more, using mean or median may lead to loss of variation in data, hence it is better to use imputations.
c. Other methods are deletion techniques to eliminate missing data
d. Regression analysis to systematically eliminate data
e. Imputation Techniques: Average Imputation

Statistics Worksheet - 1

12. What is A/B testing?

Ans.

a. A/B testing is basically a randomized control experiment in which we compare the two versions of a variable to find out which performs better in a controlled environment.

13. Is mean imputation of missing data acceptable practice?

Ans.

a. Mean Imputation is acceptable when data is missing completely at random, hence it is unbiased.
b. But when relationship among variables is important, mean imputation is not a good solution as it will lead to variation in data

14. What is linear regression in statistics?

Ans.

a. In statistics, linear regression is a linear approach for modelling the relationship between dependent and independent variables.

15. What are the various branches of statistics?

Ans.

a. Descriptive
b. Inferential
c. Decision Theory