STATISTICS WORKSHEET-1

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

Ans) 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) Central Limit Theorem

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

Ans) Modeling bounded count data

4. Point out the correct statement.

Ans) All of the mentioned

5. \_\_\_\_\_\_ random variables are used to model rates.

Ans) Poisson

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

Ans) False

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

Ans) Hypothesis

8. 4. Normalized data are centered at\_\_\_\_\_\_and have units equal to standard deviations of the original data.

Ans) 0

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

Ans) Outliers cannot conform to the regression relationship

10. What do you understand by the term Normal Distribution?

Ans: In Normal Distribution the Mean=Median=Mode=0

The distribution gives a bell like symmetric curve.

The standard deviation is equal to 01.

The skewness is Zero

The data is distributed according to the Empirical Rule, 68-95 and 99.7 percentage of data lie between 01, 02 and 03 standard deviation of the mean.

There are no outliers present.

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

Ans. Missing data can be deleted along with rows or columns.

If the number of missing data is large, then we can replace the missing data with some contact variable.

The Imputation techniques we can imply is that we can replace the Missing values with the Mean, Median and Mode of the row or column.

12. What is A/B testing?

Ans. Also called as Split Testing.

It’s a test of two Variants to understand which performs better.

Before the test we need to formulate the null hypothesis and Alternate hypothesis

Sample is selected from the data and testing is done

Then the results are compared.

13. Is mean imputation of missing data acceptable practice?

Ans Mean imputation is not a good practice as it ignores all the corelations.

Also Mean Imputation reduces the variance of the data.

14. What is linear regression in statistics?

Ans) In linear regression there is a Dependent Variable and an independent variable.

Dependent Variable values are predicted using the independent Variables.

15. What are the various branches of statistics?

Ans) Statistics is mainly divided into two categories

1. Descriptive Statistics
2. Inferential Statistics.