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) Modeling bounded count data

4**. Point out the correct statement.**

Ans) d) All of the mentioned

5. **_____ random variables are used to model rates.**

Ans) c) Poisson

6. 10**. Usually replacing the standard error by its estimated value does change the CLT.**

Ans) b) False

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

Ans) b) Hypothesis

8. 4. **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) is a probability distribution that is a symmetric in the mean show data near the mean are will occurre the data far from the mean.is called Normal Distribution or also known as the Gaussian distribution.

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

Ans) we will handle missing data-

- Pairwise Deletion
- Listwise Deletion or Dropping rows
- Dropping complete columns
  - Imputation Techniques
    - ✓ Imputation with a constant value
    - ✓ Imputation using the statistics (mean, median, mode)
    - ✓ K-Nearest Neighbor Imputation

12. What is A/B testing?

Ans) A randomized experiment process where two or more versions of a variable are shown to different segments of website at the same time to determine which version leaves the maximum impact is called A/B testing or split testing

13. Is mean imputation of missing data acceptable practice?

Ans) A mean imputation of missing data accept only terrible practice.

14. What is linear regression in statistics?

Ans) Linear regression is used to predict the value of a variable based on the value of another variable.

15. What are the various branches of statistics?

Ans) they are three various branches of statistics

- data collection,
- descriptive statistics
- inferential statistics.