STATISTICS WORKSHEET-1

Que1. Bernoulli random variables take (only) the values 1 and 0.

Ans- a) True

Que2. 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

Que3. Which of the following is incorrect with respect to use of Poisson distribution?

Ans- b) Modeling bounded count data

Que4. Point out the correct statement.

Ans- d) All of the mentioned

Que5. _____ random variables are used to model rates.

Ans- c) Poisson

Que6. 10. Usually replacing the standard error by its estimated value does change the CLT.

Ans- b) False

Que7. Which of the following testing is concerned with making decisions using data?

Ans- b) Hypothesis

Que8. 4. Normalized data are centered at_____and have units equal to standard deviations of the original data.?
Ans- a) 0

Que9.Which of the following statement is incorrect with respect to outliers
Ans- c) Outliers cannot conform to the regression relationship.

Que10. What do you understand by the term Normal Distribution?
Ans- A normal distribution is an arrangement of a data set in which most values cluster in the middle of the range and the rest taper off symmetrically toward either extreme. Height is one simple example of something that follows a normal

distribution pattern: Most people are of average height, the numbers of people that are taller and shorter than average are fairly equal and a very small number of people are either extremely tall or extremely short.
In a normal distribution, the mean, mode and median are all the same.

Que11.How do you handle missing data? What imputation techniques do you recommend?
Ans- Mostly two data imputation techniques to handle missing data: Average imputation and common-point imputation. Average imputation uses the average value of the responses from other data entries to fill out missing values. However, a word of caution when using this method – it can artificially reduce the variability of the dataset. Common-point imputation, on the other hand, is when the data scientists utilise the middle point or the most commonly chosen value.
For example, ----- on a five-point scale, the substitute value will be 3.

Que12. What is A/B testing?
Ans- A/B testing is a basic randomized control experiment. It is a way to compare the two versions of a variable to find out which performs better in a controlled environment.

Que13. Is mean imputation of missing data acceptable practice?
Ans- Yes, its True, imputing the mean preserves the mean of the observed data. So, if the data are missing completely at random, the estimate of the mean remains unbiased.

Que14. 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.

Que.15 What are the various branches of statistics?
Ans- Statistics is the branch of mathematics that deals with data. There are three real branches of statistics: data collection, descriptive statistics and inferential statistics.