Worksheet 1 Statistics

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

Ans: a) True

Q 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

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

Ans: b) Modeling bounded count data

Q 4. Point out the correct statement.

Ans: d) All of the mentioned

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

Ans: c) Poisson

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

Ans: b) False

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

Ans b) Hypothesis

Q 8. 4. Normalized data are centered at_____ and have units equal to standard deviations of the original data.

Ans: a) 0

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

Ans: c) Outliers cannot conform to the regression relationship.

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

Ans: A normal distribution is a type of continuous probability distribution in which most data points cluster toward the middle of the range, while the rest taper off symmetrically toward either extreme

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

Ans: Missing data can be dealt with in a variety of ways. Another common strategy among those who pay attention is imputation

mean imputation

substitution hot deck imputation

cold deck imputation

Regression imputation

Q 12. What is A/B testing?

Ans: A refers to 'control' or the original testing variable. Whereas B refers to 'variation' or a new version of the original testing variable.

Q 13. Is mean imputation of missing data acceptable practice?

Ans: Since most research studies are interested in the relationship among variables, mean imputation is not a good solution

Q 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.

Q 15. What are the various branches of statistics?

Ans: There are three real branches of statistics:

Data collection,

Descriptive statistics and

inferential statistics.