STATISTICS WORKSHEET-1

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

Ans= 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= Central Limit Theorem

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

Ans= Modeling bounded count data

Q-4.Point out the correct statement.

Ans= The square of a standard normal random variable follows what is called chi-squared distribution

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

Ans= Poisson

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

Ans= False

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

Ans= Hypothesis

Q-8 Normalized data are centered at_____and have units equal to standard deviations of the original data.

Ans= 0

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

Ans= Outliers cannot conform to the regression relationship

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

Ans= Normal Distribution means the distribution is symmetrical around its mean and also when we draw a graph we can see that it is bell shaped. It is easy to understand the data. The mean , mode and median all are equal and located at the center of the distribution.

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

Ans=When i get dataset there are lots of missing values in major features so i need to draw the bar chart or heatmap to understand that at which point data is missing , so for filling those missing values I can use mean method for continuous data and for categorical data I have to use mode to fill that particular missing value. Also I recommend KNN method to fill missing values. It will replace the value with the average of nearest neighbor's value.

Q-12. What is A/B testing?

Ans= A/B testing is method to compare two different things (ex: app or other user experience) to user so we can under that which is better A or B. Also help us to find which variant is more desirable and effective so we understand better.

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

Ans= Mean imputation of missing data is acceptable practice when our data is missing randomly and it is very easy to apply mean imputation. Also when our data is minimal we can use this imputation and it is quick fix for missing values at those points.

Q-14. What is linear regression in statistics?

Ans= Linear regression in statistics is include dependent variable , independent variable ,intercept ,slope and error term.here is the equation of linear regression ($y=\beta_0+ \beta_1 x+c$)

Y= Dependent variable , x= Independent variable , β_0 = Intercept , β_1 = Slope , c = error term.

Q-15. What are the various branches of statistics?

Ans= As we know that there are various branches of statistics .There are descriptive statistics , inferential statistics, probability theory,biostatistics,econometrics ,statistical computing. These are the major statistics branches.