STATISTICS WORKSHEET-1

Q1 to Q9 have only one correct answer. Choose the correct option to answer your question.

1. Bernoulli random variables take (only) the values 1 and 0.
   Answer-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?
   Answer-A(central limit theorem)
3. Which of the following is incorrect with respect to use of Poisson distribution?
   Answer-B(modeling bounded count data)
4. Point out the correct statement.
   Answer-C(The square of a standard normal random variable follows what is called chi-squared distribution)
5. _____ random variables are used to model rates.
   Answer-C(Poisson)
6. Usually replacing the standard error by its estimated value does change the CLT.
   Answer-B(True)
7. Which of the following testing is concerned with making decisions using data?
   Answer-B(Hypothesis)
8. Normalized data are centered at_____and have units equal to standard deviations of the original data.
   Answer-A(0)

9. Which of the following statement is incorrect with respect to outliers?
   Answer-C(outliers cannot conform to the regression relationship)

Q10and Q15 are subjective answer type questions, Answer them in your own words briefly.

Q10- what do you understand by the term normal distribution?

Answer- Anormal distribution of data is one in which the most greater number of data points are almost similar. They occur within a small range of value with fewer outliers on the high and low ends of the data range.

Q11- how do you handle missing data? What imputation technique do you recommend?

Answer- To handle missing data we use imputation or the removal of data. I recommend regression imputation, in this the predicted value obtained by regressing the missing variable on the variables.

Q12-what is A/B testing?

Answer- This testing is a hypothesis of statistical testing. This is the logical method for decision making.

Q13- Is mean imputation of missing data acceptable practice?

Answer- No, mean imputation underestimate the standard deviation. Mean imputation conserve the mean of the observed data.

Q14- what is linear regression in statistics?

Answer- Linear regression is a supervised machine learning model which find the best fit line between the dependent variable and independent variable.

Q15-what are the various branches of statistics?

Answer- There are 2 various branches of statistics:-

1). Descriptive statistics.

2). Inferential statistics.