STATISTICS WORKSHEET-1

1) Bernoulli random variables take (only) the values 1 and 0.
   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?
   a) Central Limit Theorem
3) Which of the following is incorrect with respect to use of Poisson distribution?
   b) Modeling bounded count data
4) Point out the correct statement.
   d) All of the mentioned
5) _____ random variables are used to model rates.
   c) Poisson
6) Usually replacing the standard error by its estimated value does change the CLT.
   b) False
7) 1. Which of the following testing is concerned with making decisions using data?
   b) Hypothesis
8) Normalized data are centered at_____and have units equal to standard deviations of the original data.
   a) 0
9) Which of the following statement is incorrect with respect to outliers?
   c) Outliers cannot conform to the regression relationship
10) What do you understand by the term Normal Distribution?
    A normal distribution is an arrangement of the data set in which most of the values cluster in the middle range and rest taper off symmetrically toward extreme. The precise shape can vary according to the distribution of the population but the peak is always in the middle and the curve is always symmetrical. Graphical representation of normal distribution is called as a bell curve.
11) How do you handle missing data? What imputation techniques do you recommend?
    Use deletion method to eliminate missing data.
    Use regression analysis to to systematically eliminate data.
    Use data impulation techniques like average impulation and common point impulation.
12) What is A/B testing?
    It is a way to compare two versions of variable to find out which one performs better in controlled environment. It is also called as a split testing or bucket testing.
13) Is mean imputation of missing data acceptable practice?
14) What is linear regression in statistics?
    Linear regression attempts to model the relationship between two variables by fitting a linear equation to observed data.
15) What are the various branches of statistics?
    There are two branches of statistics
    1 Inferential Statistics
    2 Descriptive Statistics.
    Inferential statistics used to make inference and describe about the population. These states are more useful when it's not easy or possible to examine each member of the population.
    Descriptive statistics are used to get a brief summary of data. You can have summery of data in numerical or graphical form.