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.

a) True

7. 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?

The normal distribution is a continuous probability distribution that is symmetrical around it's mean. It has a shape often referred to as a "bell curve". Many everyday data sets typically follow a normal distribution for an example the heights of adult male, the score on a test, errors in measurements.

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

I'd rather delete that row completely or give a certain value to the missing row, filling the nan with the mean or median of the column.

The imputation technique I recommend are

1. Impute missing values for continuous variable.
2. Impute missing values for categorical variable.

12. What is A/B testing?

A/B testing is a method of comparing two versions of a webpage or app against each other to determine which one performs better. A/B testing is essentially an experiment where two or more variants of a page are shown to users at random, and statistical analysis is used to determine which variation performs better for a given conversion goal.

13. Is mean imputation of missing data acceptable practice?

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.

15. What is linear regression in statistics?

Linear regression is a basic and commonly used type of predictive analysis. The dependent and independent variables show a linear relationship between the slope and the intercept.

16. What are the various branches of statistics?

- Descriptive statistics.
- Inferential statistics.