# 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.
   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?
   a) Modeling bounded count data
4. Point out the correct statement.
   a) All of the mentioned
5. _____random variables are used to model rates.
   a) Poisson
6. 10. Usually replacing the standard error by its estimated value does change the CLT.
   a) True
7. 1. Which of the following testing is concerned with making decisions using data?
   a) Hypothesis
8. 4. 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?
   a) Outliers cannot conform to the regression relationship

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

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

    The normal distribution, also known as the Gaussian or standard normal distribution, is the probability distribution that plots all of its values in a symmetrical fashion, and most of the results are situated around the probability's mean. Values are equally likely to plot either above or below the mean.

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

- Use deletion methods to eliminate missing data. The deletion methods only work for certain datasets where participants have missing fields. ...
- Use regression analysis to systematically eliminate data. ...
- Data scientists can use data imputation techniques.

12. What is A/B testing?

    A/B testing is a basic randomized control experiment. It is a way to compare the two versions of a variable to find out which performs better in a controlled environment

13. Is mean imputation of missing data acceptable practice?

    True, 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. That's a good thing.

14. What is linear regression in statistics?

    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.

15. What are the various branches of statistics?

    There are three real branches of statistics:
    (a) Data collection,
    (b) Descriptive statistics
    (c) Inferential statistics.