**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.
   1. True (yes)
   2. False
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?
   1. Central Limit Theorem (yes)
   2. Central Mean Theorem
   3. Centroid Limit Theorem
   4. All of the mentioned
3. Which of the following is incorrect with respect to use of Poisson distribution?
   1. Modeling event/time data
   2. Modeling bounded count data
   3. Modeling contingency tables (yes)
   4. All of the mentioned
4. Point out the correct statement.
   1. The exponent of a normally distributed random variables follows what is called the log- normal distribution
   2. Sums of normally distributed random variables are again normally distributed even if the variables are dependent
   3. The square of a standard normal random variable follows what is called chi-squared distribution
   4. All of the mentioned
5. random variables are used to model rates.
   1. Empirical
   2. Binomial
   3. Poisson (yes)
   4. All of the mentioned
6. Usually replacing the standard error by its estimated value does change the CLT.
   1. True
   2. False ( yes )
7. Which of the following testing is concerned with making decisions using data?
   1. Probability
   2. Hypothesis ( yes )
   3. Causal
   4. None of the mentioned
8. Normalized data are centered at and have units equal to standard deviations of the original data.
   1. 0 (yes)
   2. 5
   3. 1
   4. 10
9. Which of the following statement is incorrect with respect to outliers?
   1. Outliers can have varying degrees of influence
   2. Outliers can be the result of spurious or real processes
   3. Outliers cannot conform to the regression relationship (yes)
   4. None of the mentioned

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

1. What do you understand by the term Normal Distribution?   
   ------ **It is often called “bell curve” because it looks like bell . In normal distribution :  
    A) mean = median = mode  
    B) symmetry about centre  
    C) 50% of the value less than the mean and 50% of the value greater than the mean** .
2. How do you handle missing data? What imputation techniques do you recommend?  
   ------ **We handle missing data :  
    A) deletion technique   
    B) dropping method   
    C) Through mean , median and mode .  
    I recommend KNN .**
3. 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.**
4. Is mean imputation of missing data acceptable practice?  
   ----- yes .
5. What is linear regression in statistics?  
   -----**Linear regression is the most basic and commonly used predictive analysis. Regression estimates are used to describe data and to explain the relationship**.
6. What are the various branches of statistics?  
   ----- **Two branches, descriptive statistics and inferential statistics, comprise the field of statistics.**

