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**(✓)
   b) 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?
   a) **Central Limit Theorem**(✓)
   b) Central Mean Theorem
   c) Centroid Limit Theorem
   d) All of the mentioned

3. Which of the following is incorrect with respect to use of Poisson distribution?
   a) Modeling event/time data
   b) Modeling bounded count data
   c) Modeling contingency tables
   d) **All of the mentioned**(✓)

4. Point out the correct statement.
   a) The exponent of a normally distributed random variables follows what is called the log- normal distribution
   b) Sums of normally distributed random variables are again normally distributed even if the variables are dependent
   c) The square of a standard normal random variable follows what is called chi-squared distribution
   d) **All of the mentioned**(✓)

5. _____random variables are used to model rates.
   a) Empirical
   b) Binomial
   c) **Poisson**(✓)
   d) All of the mentioned

6. 10. Usually replacing the standard error by its estimated value does change the CLT.
   a) True
   b) False(✓)

7. 1. Which of the following testing is concerned with making decisions using data?
   a) Probability
   b) **Hypothesis**(✓)
   c) Causal
   d) None of the mentioned

8. 4. Normalized data are centered at_____and have units equal to standard deviations of the original data.
   a) **0**(✓)
   b) 5
   c) 1
   d) 10

9. Which of the following statement is incorrect with respect to outliers?
    a) Outliers can have varying degrees of influence
    b) Outliers can be the result of spurious or real processes
    c) Outliers cannot conform to the regression relationship(✓)
    d) 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?

Ans:- Normal distribution is known as the probability distribution to shown the data near about the mean. It shows the data not far the mean and represent in the curved graph.

Example:- The amount of rainfall in inches in a year for a city.

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

Ans:- The most comman way to handle the missing data i.e. delete all the rows which has missing data and proceed the file to execute.

3. What is A/B testing?

Ans:- A/B testing is the method to compare the two web pages or apps and gives the proper result and we can also check that which one is working better. It is also known as the split testing.

4. Is mean imputation of missing data acceptable practice?

Ans:- Yes, mean imputation is the acceptable practice to observed the data and tells about the missing data.

5. What is linear regression in statistics?

Ans:- Linear regression is a linear approach for modeling the relationship between dependent and independent variables. For example, we can draw a line-graph b/w **X** coordinates and **Y** coordinates and that line shows the relationship b/w both coordinates.

6. What are the various branches of statistics?

Ans:- There are three branches of statistics
1. Data collection.
2. Descriptive statistics.
3. Inferential statistic.