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**
   2. False

Ans: **True**

1. 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**
   2. Central Mean Theorem
   3. Centroid Limit Theorem
   4. All of the mentioned

Ans: **Central Limit Theorem**

1. 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
   4. All of the mentioned

**Ans: Modeling bounded count data**

1. 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

**Ans: All of the mentioned**

1. random variables are used to model rates.
   1. Empirical
   2. Binomial
   3. **Poisson**
   4. All of the mentioned

Ans : **Poisson**

1. 10. Usually replacing the standard error by its estimated value does change the CLT.
   1. True
   2. **False**

**Ans : False**

1. 1. Which of the following testing is concerned with making decisions using data?
   1. Probability
   2. **Hypothesis**
   3. Causal
   4. None of the mentioned

Ans: **Hypothesis**

1. 4. Normalized data are centered at and have units equal to standard deviations of the original data.
   1. **0**
   2. 5
   3. 1
   4. 10

Ans : **0**

1. 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**
   4. None of the mentioned

Ans: **Outliers cannot conform to the regression relationship**

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

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

Normal distribution, also known as the Gaussian distribution, is a probability distribution that is symmetric about the mean, showing that data near the mean are more frequent in occurrence than data far from the mean. In graph form, normal distribution will appear as a bell curve

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

Handle Missing Data:

1. Deletions

* Pairwise Deletion
* Listwise Deletion/ Dropping rows
* Dropping complete columns

1. Basic Imputation Techniques

* Imputation with a constant value
* Imputation using the statistics (mean, median, mode)
* K-Nearest Neighbor Imputation

1. What is A/B testing?

A/B testing (also known as bucket testing or split-run testing) is a user experience research methodology. A/B tests consist of a randomized experiment with two variants, A and B. It includes application of statistical hypothesis testing or "two-sample hypothesis testing" as used in the field of statistics. A/B testing is a way to compare two versions of a single variable, typically by testing a subject's response to variant A against variant B, and determining which of the two variants is more effective.

1. Is mean imputation of missing data acceptable practice?

Yes, Mean imputation of missing data is an acceptable practice.

1. 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

1. What are the various branches of statistics?

There are three real branches of statistics: data collection, descriptive statistics and inferential statistics.

