STATISTICS WORKSHEET-1

Q1 to Q9 have only one correct answer. Choose the correct option to answer your question.

- Bernoulli random variables take (only) the values 1 and 0.
  - True
  - False
- Which of the following theorem states that the distribution of averages of iid variables, properlynormalized, becomes that of a standard normal as the sample size increases?
  - Central Limit Theorem
  - Central Mean Theorem
  - Centroid Limit Theorem
  - All of the mentioned
- Which of the following is incorrect with respect to use of Poisson distribution?
  - Modeling event/time data
  - Modeling bounded count data
  - Modeling contingency tables
  - All of the mentioned
- Point out the correct statement.
  - The exponent of a normally distributed random variables follows what is called the log- normaldistribution
  - Sums of normally distributed random variables are again normally distributed even if the variablesare dependent
  - The square of a standard normal random variable follows what is called chi-squareddistribution
  - All of the mentioned
- _____random variables are used to model rates.
  - Empirical
  - Binomial
  - Poisson
  - All of the mentioned
- 10. Usually replacing the standard error by its estimated value does change the CLT.
  - True
  - False
- 1. Which of the following testing is concerned with making decisions using data?
  - Probability
  - Hypothesis
  - Causal
  - None of the mentioned
- 4. Normalized data are centered at_____and have units equal to standard deviations of theoriginal data.
  - 0
  - 5
  - 1
  - 10
- Which of the following statement is incorrect with respect to outliers?
  - Outliers can have varying degrees of influence
  - Outliers can be the result of spurious or real processes
  - Outliers cannot conform to the regression relationship
  - None of the mentioned

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

- What do you understand by the term Normal Distribution?
- How do you handle missing data? What imputation techniques do you recommend?
- What is A/B testing?
- Is mean imputation of missing data acceptable practice?
- What is linear regression in statistics?
- What are the various b ranches of statistics?

Normal Distribution:-

The normal distribution is a continuous probability distribution that is symmetrical on both sides of the mean, so the right side of the center is a mirror image of the left side. ... The normal distribution is often called the bell curve because the graph of its probability density looks like a bell. In Normal Distribution the mean is 0 and standard deviation is 1.

Imputation technique:-

Imputation is the process of replacing the missing data with approximate values. Instead of deleting any columns or rows that has any missing value, this approach preserves all cases by replacing the missing data with the value estimated by other available information.

A/B testing

A/B testing (also known as split testing or bucket testing) is a method of comparing two versions of a webpage or app against each other to determine which one performs better.

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. ... Since most research studies are interested in the relationship among variables, mean imputation is not a good solution.

Linear Regression in Statistics

Image result for • What is linear regression in statistics
In statistics, linear regression is a linear approach to modelling the relationship between a scalar response and one or more explanatory variables

Branches of statistics

There are mainly two type of stastics (i) descriptive &(ii) inferential.