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, properly normalized, 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- normal distribution
  - Sums of normally distributed random variables are again normally distributed even if the variables are dependent
  - The square of a standard normal random variable follows what is called chi-squared distribution
  - 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 the original 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?

Answer - Normal Distribution shows that the Mean / Median /Mode is zero and it happens in very rare scenerios. It is bell shaped curve with Standard Deviation =1. It tells thta the data is normally distributed.

- How do you handle missing data? What imputation techniques do you recommend?

Answer - Firstyly Mean or Median is calculated of the data, then at the missing places, we fill it with either Mean or Median to get the better Statistical Results.
According to me, Using Mean Imputation gives better results.

- What is A/B testing?

Answer - Two version/sets is created to use on half population to check the results whether A gives better results or B gives the better.

- Is mean imputation of missing data acceptable practice?

Answer - According to me, it is acceptable at mostly places, because mean is the average of the data, and which shows the equal distribution to each data. So, getting the missig value with Mean is acceptable practice.

- What is linear regression in statistics?

Answer - Linear Regression is used to model the relationship between two continuous variables. (Numerice Values)

- What are the various branches of statistics?

Answer - The two main branches of statistics are descriptive statistics and inferential statistics.

FLIP ROBO