WORKSHEET

STATISTICS AnswerSHEET-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

2. Which of the following theorem states that the distribution of averages of id variables, properly normalized, becomes that of a standard normal as the sample size increases?

a) Central Limit Theorem

3. Which of the following is incorrect with respect to use of Poisson distribution?

c) Modeling contingency tables

4. Point out the correct statement.

d) All of the mentioned

5. _____ random variables are used to model rates.

c) Poisson

6. Usually replacing the standard error by its estimated value does change the CLT.

b) False

7. Which of the following testing is concerned with making decisions using data?

b) Hypothesis

8. Normalized data are centered at _____ and have units equal to standard deviations of the original data.

a) 0

9. Which of the following statement is incorrect with respect to outliers?

c) Outliers cannot conform to the regression relationship

WORKSHEET

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

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

A. The normal distribution, also known as the Gaussian distribution, is a symmetric probability distribution centered on the mean, indicating that data around the mean occur more frequently than data far from it.

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

12. What is A/B testing?

A. A/B testing, like any other sort of scientific testing, is essentially statistical hypothesis testing, or statistical inference. It's an analytical decision-making approach that uses sample statistics to estimate population parameters.

13. Is mean imputation of missing data acceptable practice?

A. No, imputing the mean keeps the observed data's mean. As a result, even if all of the data is missing at random, the mean estimate remains impartial. You may keep your sample size up to the entire sample size by imputing the mean. Mean imputation will not influence your parameter estimate if all you're doing is estimating means and the data is missing fully at random. Your standard error will still be skewed. Mean imputation is not an acceptable solution since most research projects are interested in the connection between variables.

14. What is linear regression in statistics?

A. By fitting a linear equation to observed data, linear regression seeks to model the connection between two variables. One variable is regarded as an explanatory variable, while the other is regarded as a dependent variable.

15. What are the various branches of statistics?

A. There are two main branches of statistics descriptive statistics and inferential statistics respectively.

- Descriptive Statistics: The presentation and collecting of data are the focus of descriptive statistics. This is carried out at the start of a statistical analysis. It is rarely as straightforward as it appears, and the statistician must be aware of how to construct trials, choose the appropriate focus group, and prevent biases that are all too easy to introduce into the study.
- Inferential Statistics: As the name implies, inferential statistics entails deriving the appropriate inferences from descriptive statistics-based statistical analysis. The inferences are what make research significant, and inferential statistics deals with this element.