

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.

Ans: a) True

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?

Ans: a) Central Limit Theorem

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

Ans: b) Modelling bounded count data

4. Point out the correct statement.

Ans: a) The exponent of a normally distributed random variables follows what is called the log- normal distribution

5. _____ random variables are used to model rates.

Ans: c) Poisson

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

Ans: b) False

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

Ans: b) Hypothesis

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

Ans: a) 0

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

Ans: Outliers cannot conform to the regression relationship

Q10 and Q15 are subjective answer type questions, Answer them in your own words briefly.

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

Ans: A normal distribution another name is Gaussian distribution. It means that probability distribution where the values of a random variable are distributed. These values are equally distributed on the left and the right side of the central tendency. Thus, a bell-shaped curve is formed.

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

Ans: The missing values are handled through the simplest imputation method, which is replacing with the mean or median values of the dataset at large, or some similar summary statistic.

12. What is A/B testing?

Ans: A/B testing, also called split testing or bucket testing, compares the performance of two versions of content to see which one appeals more to visitors/viewers. A/B testing is a method of comparing two versions of a webpage or app against each other to determine which one performs better.

13. Is mean imputation of missing data acceptable practice?

Ans: Yes, mean imputation of missing data is acceptable practice because the imputation method substitutes reasonable guesses for missing data.

14. What is linear regression in statistics?

Ans: Linear regression is the most basic and commonly used predictive analysis. Linear regression is a simple and the simplest statistical regression technique used for predictive analysis in machine learning.

15. What are the various branches of statistics?

Ans: The branches of statistics are mainly two types. One is descriptive statistics and another is inferential statistics. Both of these are employed in scientific analysis of data and both are equally important for the statistics.