

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

Answer: 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?

Answer: a) Central Limit Theorem

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

Answer: b) Modeling bounded count data

4. Point out the correct statement.

Answer: d) All of the mentioned

5. _____ random variables are used to model rates.

Answer: c) Poisson

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

Answer: b) False

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

Answer: b) Hypothesis

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

Answer: a) 0

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

Answer: c) 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?

Answer: A normal distribution is a type of continuous probability distribution in which most data points cluster toward the middle of the range, while the rest taper off symmetrically toward either extreme.

The middle of the range is also known as the mean of the distribution.

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

Answer: Missing data can be dealt with in a variety of ways. I believe the most common reaction is to ignore it. Choosing to make no decision, on the other hand, indicates that your statistical programme will make the decision for you.

Your application will remove things in a listwise sequence most of the time. Depending on why and how much data is gone, listwise deletion may or may not be a good idea.

12. What is A/B testing?

Answer: A/B testing is a way to compare multiple versions of a single variable, for example by testing a subject's response to variant A against variant B, and determining which of the variants is more effective.

13. Is mean imputation of missing data acceptable practice?

Answer: Mean imputation is typically considered a terrible practice since it ignores feature correlation. Consider the following scenario: we have a table with age and fitness scores, and an eight-year-old has a missing fitness score. If we average the fitness scores of people between the ages of 15 and 80, the eighty-year-old will appear to have a significantly greater fitness level than he actually does.

14. What is linear regression in statistics?

Answer: 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.

15. What are the various branches of statistics?

Answer: Statistics:

Statistics is a study of presentation, analysis, collection, interpretation and organization of data

There are two main branches of statistics

- Inferential Statistic.
- Descriptive Statistic.

Inferential Statistics:

Inferential statistics used to make inference and describe about the population. These stats are more useful when it's not easy or possible to examine each member of the population.

Descriptive Statistics:

Descriptive statistics are used to get a brief summary of data. You can have the summary of data in numerical or graphical form.

I hope you find this helpful, Thank you for reading!