STATISTICS WORKSHEET-3

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

1. Which of the following is the correct formula for total variation?
   a) Total Variation = Residual Variation – Regression Variation
   b) Total Variation = Residual Variation + Regression Variation
   c) Total Variation = Residual Variation * Regression Variation
   d) All of the mentioned

Ans: d)

2. Collection of exchangeable binary outcomes for the same covariate data are called_____outcomes.
   a) random
   b) direct
   c) binomial
   d) none of the mentioned

.Ans: c)

3. How many outcomes are possible with Bernoulli trial?
   a) 2
   b) 3
   c) 4
   d) None of the mentioned

Ans: a)

4. If Ho is true and we reject it is called
   a) Type-I error
   b) Type-II error
   c) Standard error
   d) Sampling error

Ans: a)

5. Level of significance is also called:
   a) Power of the test
   b) Size of the test
   c) Level of confidence
   d) Confidence coefficient

Ans: c)

6. The chance of rejecting a true hypothesis decreases when sample size is:
   a) Decrease
   b) Increase
   c) Both of them
   d) None

Ans: a)

7. Which of the following testing is concerned with making decisions using data?
   a) Probability
   b) Hypothesis
   c) Causal
   d) None of the mentioned

Ans: b)

8. What is the purpose of multiple testing in statistical inference?
   a) Minimize errors
   b) Minimize false positives
   c) Minimize false negatives
   d) All of the mentioned

Ans: b)

9. Normalized data are centred at_____and have units equal to standard deviations of the original data
   a) 0
   b) 5
   c) 1
   d) 10

Ans: c)

10. What Is Bayes' Theorem?

Ans:

Bayes' theorem describes the probability of occurrence of an event related to any condition. It is also considered for the case of conditional probability. Bayes theorem is also known as the formula for the probability of "causes". For example: if we have to calculate the probability of taking a blue ball from the second bag out of three different bags of balls, where each bag contains three different colour balls viz. red, blue, black. In this case, the probability of occurrence of an event is calculated depending on other conditions is known as conditional probability.

11. What is z-score?

Ans:

Z-score is also known as standard score gives us an idea of how far a data point is from the mean. It indicates how many standard deviations an element is from the mean. Hence, Z-Score is measured in terms of standard deviation from the mean. For example, a standard deviation of 2 indicates the value is 2 standard deviations away from the mean. In order to use a z-score, we need to know the population mean (μ) and also the population standard deviation (σ).

12. What is t-test?

Ans:

A *t* test is a statistical test that is used to compare the means of two groups. It is often used in hypothesis testing to determine whether a process or treatment actually has an effect on the population of interest, or whether two groups are different from one another.

You want to know whether the mean petal length of iris flowers differs according to their species. You find two different species of irises growing in a garden and measure 25 petals of each species. You can test the difference between these two groups using a t test and null and alterative hypotheses.

13. What is percentile?

Ans

In statistics, percentiles are used to understand and interpret data. The *n*th percentile of a set of data is the value at which *n* percent of the data is below it. In everyday life, percentiles are used to understand values such as test scores, health indicators, and other measurements. For example, an 18-year-old male who is six and a half feet tall is in the 99th percentile for his height. This means that of all the 18-year-old males, 99 percent have a height that is equal to or less than six and a half feet. An 18-year-old male who is only five and a half feet tall, on the other hand, is in the 16th percentile for his height, meaning only 16 percent of males his age are the same height or shorter.

14. What is ANOVA?

Analysis of variance (ANOVA) is an analysis tool used in statistics that splits an observed aggregate variability found inside a data set into two parts: systematic factors and random factors. The systematic factors have a statistical influence on the given data set, while the random factors do not. Analysts use the ANOVA test to determine the influence that independent variables have on the dependent variable in a regression study.

15. How can ANOVA help?

ANOVA is helpful for testing three or more variables. It is similar to multiple two-sample t-tests. However, it results in fewer type I errors and is appropriate for a range of issues. ANOVA groups differences by comparing the means of each group and includes spreading out the variance into diverse sources. It is employed with subjects, test groups, between groups and within groups.