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

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

3. How many outcomes are possible with Bernoulli trial?

a) 2

b) 3

c) 4

d) None of the mentioned

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

5. Level of significance is also called:

a) Power of the test

b) Size of the test

c) Level of confidence

d) Confidence coefficient

6. The chance of rejecting a true hypothesis decreases when sample size is:

a) Decrease

b) Increase

c) Both of them

d) None

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

a) Probability

b) Hypothesis

c) Causal

d) None of the mentioned

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

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

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

10. What Is Bayes' Theorem?

This theorem is used to determine the conditional probability of event for continuous random variables.

11. What is z-score?

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. 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?

It is statistics tools which 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. It is also called as student test. This test is conduct when the standard deviation is not known.

13. What is percentile?

Percentiles are used to understand and interpret data. They indicate the values below which a certain percentage of the data in a data set is found. Percentiles can be calculated using the formula n = (P/100) x N, where P = percentile, N = number of values in a data set (sorted from smallest to largest), and n = ordinal rank of a given value. Percentiles are frequently used to understand test scores and biometric measurements.

14. What is ANOVA?

It is used to compare difference of means among more than 2 groups. It does this by looking a variance in the data & where that variance found, specifically, ANOVA compare the amount of variance between the group with the amount of variance within group. Null hypothesis typically is that all means are equal.

15. How can ANOVA help

Hypothesis Testing: Enables the comparison of independent and dependent variables.

Understanding Data Sets: An analyst or statistician can best determine inconsistencies in data sets.

Group Comparisons: Allows multiple groups to be compared at the same time to uncover relationships between data.