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?
11. What is z-score?
12. What is t-test?
13. What is percentile?
14. What is ANOVA?
15. How can ANOVA help?

Ans. 10:-
A theorem describing how the conditional probability of each of a set of possible causes for a given observed outcome can be computed from knowledge of the probability of each cause and the conditional probability of the outcome of each cause.

Ans. 11:-
The standard score (more commonly referred to as a z-score) is a very useful statistic because it:
(a) allows us to calculate the probability of a score occurring within our normal distribution and
(b) enables us to compare two scores that are from different normal distributions.

Ans. 12:-
The t-test is any statistical hypothesis test in which the test statistic follows a Student's t-distribution under the null hypothesis.
A t-test is the most commonly applied when the test statistic would follow a normal distribution if the value of a scaling term in the test statistic were known. When the scaling term is unknown and is replaced by an estimate based on the data, the test statistics (under certain conditions) follow a Student's t distribution. The t-test can be used, for example, to determine if the means of two sets of data are significantly different from each other.

Ans. 13:-
A percentile (or a centile) is a measure used in statistics indicating the value below which a given percentage of observations in a group of observations fall. For example, the 20th percentile is the value (or score) below which 20% of the observations may be found.

Ans. 14:-
An ANOVA test is a type of statistical test used to determine if there is a statistically significant difference between two or more categorical groups by testing for differences of means using variance.

Ans. 15:-
You might use Analysis of Variance (ANOVA) as a marketer, when you want to test a particular hypothesis. You would use ANOVA to help you understand how your different groups respond, with a null hypothesis for the test that the means of the different groups are equal. If there is a statistically significant result, then it means that the two populations are unequal (or different).