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 Bayes Theorem is the extension of Conditional probability. Conditional probability helps us to determine the probability of A given B, denoted by P(A|B). So Bayes' theorem says if we know P(A|B) then we can determine P(B|A), given that P(A) and P(B) are known to us.

Ans-11 The Z-score, also known as the standard score, is the number of standard deviations above or below the mean for a given data point. The standard deviation of a data collection represents the degree of variability within that data set.

Ans-12 The t-test is a test used for hypothesis testing in statistics and uses the t-statistic, the t-distribution values, and the degrees of freedom to determine statistical significance.

Ans-13 Percentiles are used in statistics data to provide you a number that expresses the value that a given percent of the values are lower than. Example:We have an array of the ages of all the people that working in same office.
ages = [25,31,43,48,50,41,39,60,52,32,27,46,47,55] What is 75

Ans-14 ANOVA, is a strong statistical technique that is used to show the difference between two or more means or components through significance tests. It also shows us a way to make multiple comparisons of several populations means.

Ans-15