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
   Answer-b

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
   Answer-c

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

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

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

9. Normalized data are centered at___and have units equal to standard deviations of the original data

a) 0
b) 5
c) 1
d) 10
Answer-a

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?

10.What is Bayes Theorem?

Answer- In statistics and probability theory, the Bayes' theorem (also known as the Bayes' rule) is a mathematical formula used to determine the conditional probability of events. Essentially, the Bayes' theorem describes the probability of an event based on prior knowledge of the conditions that might be relevant to the event. **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".

11.What is z-score?

Answer- A z score is simply defined as the number of standard deviations from the mean. The z-score can be calculated by subtracting mean by test value and dividing it by standard value. Where x is the test value, μ is the mean and σ is the standard value.

12.What is t-test?

Answer- A t-test is an inferential statistic used to determine if there is a significant difference between the means of two groups and how they are related. T-tests are used when the data sets follow a normal distribution and have unknown variances, like the data set recorded from flipping a coin 100 times. 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.

13.What is percentile?

Answer-Percentiles are used in statistics to give you a number that describes the value that a given percent of the value are lower than. A percentile is a measure used in statistics indicating the value below which a given percentage of observations in a group of observations fall.

14. What is ANOVA?

Answer- Analysis of Variance (ANOVA) ANOVA is a test for equality of several means. It allows us to compare the means for several groups–in one hypothesis test. It might sound intimidating, but ANOVA is simply a way to analyze several means at once.

15- How can ANOVA help?

Answer- 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.