WORKSHEET 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 residual variation + regression variation

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 binomial

3. How many outcomes are possible with Bernoulli trial?

a) 2 b) 3 c) 4 d) None of the mentioned

Ans 2

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 type 1 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

Ans level of confidence

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

a) Decrease b) Increase c) Both of them d) None

Ans increase

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 hypothesis testing

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 all of the above

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 0

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

10. What Is Bayes' Theorem?

Ans BAYES’S THEOREM – Bayes’ theorem states that the conditional probability of an event, based on the occurrence of another event is equal to the likelihood of the second event given the first event multiplied by the probability of the first event.

Bayes’ theorem provides a useful method for thinking about the relationship between a data set and a probability. In other words the theorem says that the probability of a given hypothesis being true based on specific observed data can be stated as finding the probability of observing the data given the hypothesis multiplies by the probability of the hypothesis being true regardless of the data, divided by the probability of observing the data regardless of the hypothesis

11. What is z-score?

Ans z score is the difference between the observed value and mean of the sample over standard deviation of the sample. It indicates how much a given value differs from the standard deviation. The z score is the number of standard deviations a given data point lies above or below mean.

Z score is important because it tells you not only something about the value itself, but also where the value lies int he distribution.

12. What is t-test?

Ans T-testing - a t-testing is a statistical test that compares the means of two samples. It is used in hypothesis testing with null hypothesis that the difference in group means is zero and an alternate hypothesis that the difference in group mean is different from zero.

13. What is percentile?

Ans

14. What is ANOVA?

Ans ANOVA – Analysis of variance is a statistical formula used to compare variance across the means(or average) of different group. A range of scenarios use it to determine if there is any difference between the means of different group.

15. How can ANOVA help?

Ans ANOVA is helpful for testing three or more variables. It is similar to multiple two-sample t-tests. However, it results