

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 H_0 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

Q10 and 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.A) 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.

11.A) Z-scores are important because they offer a comparison between two scores that are not in the same normal distribution. They are also used to obtain the probability of a z-score to take place within a normal distribution. If a z-score gives a negative value, it means that raw data is lesser than mean.

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

13.A) The percentile formula determines the performance of a person over others. The percentile formula is used in finding where a student stands in the test compared to other candidates. In our day-to-day life, percentile formulas are usually helpful in finding the test scores or biometric measurements.

14.A) Analysis of variance, or ANOVA, is a statistical method that separates observed variance data into different components to use for additional tests. A one-way ANOVA is used for three or more groups of data, to gain information about the relationship between the dependent and independent variables.

15.A) 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.