ASSIGNMENT – 3

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: b) Total Variation = 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: c) binomial

3. How many outcomes are possible with Bernoulli trial?

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

Ans: a) 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: a) Type-I 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: b) Size of the test

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

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

Ans: b) 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: b) Hypothesis

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 : 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

Ans : a) 0

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

10. What Is Bayes' Theorem?

We use Bayes' theorem when have to find conditional probability. When the probability of any event is depends on the happening of other events. Thus Bayes' theorem helps to find probability when we know the probability of events with which it has relationship.

11. What is z-score?

In feature engineering technique we use normalization or standard normalization we use Z score. Z score helps to convert Gaussian or Normal distribution into Standard normal distribution. If we know standard deviation away from the mean, we can find probability with the help of Z score table.

12. What is t-test?

T test is performed for hypothesis testing by comparing means of two continuous variable dataset. To perform t test, we requires difference in the means of two data sets, Standard deviation of each dataset and number of data values.

13. What is percentile?

Percentile is a statistical term which defines percentage value below which, rest of the in data set is resting. For example , I have 99.5 percentile in CAT exam, it means that 99.5% of total students appeared for CAT exams are below my score.

14. What is ANOVA?

ANOVA stands for Analysis of variance. This is a hypothesis testing tool used to determine if there is difference between means of two or more categorical data sets. Each data set must have equal number of samples.

15. How can ANOVA help?

ANOVA test helps us to understand the significant difference between the means of the independent variables, which helps in understanding the how each independent variable mean is different from each other and how independent variable is related to dependent variable.