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→ c) Both of them

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?

Ans→The Bayes theorem is a mathematical formula for calculating conditional probability in probability and statistics. In other words, it's used to figure out how likely an event is based on its proximity to another. Bayes law or Bayes rule are other names for the theorem.

11. What is z-score?

Ans→ The z-score is used to tell you how far from the mean the data point is. You calculate it using the mean and standard deviation, so it can also be said that the Z-Score is how many standard deviations below the mean the data is.

The z-score is used to standardize your normal distribution. Using the z-score, you can convert each data point into a value in terms of mean and standard deviation, effectively converting the graph into a scaled-down version. The z-score tells you how far each data point is from the mean in steps of standard deviation. So, with the mean and standard deviation, you can plot all points on our graph.

The z-score is given by :

$$\text{z-score} = \frac{\text{Data point} - \text{mean}}{\text{Standard deviation}}$$

12. What is t-test?

Ans→ The t-test compares the means (averages) of two populations to determine how different they are from each other. The test generates a

T-score and P-value, which quantify exactly how different each population is and the likelihood that this difference can be explained by chance or sampling error.

13. What is percentile?

Ans→ Percentile or centile is a value or number that represents a percentage position on a range or list of data – the person or thing at that number of values is above that number in percentage.

In statistics, percentile is used to indicate the value below which the group of percentages of data falls below

Ex: consider if your score is 75^{th} percentile, which you scored for better than 75% of people who took part in the test

It is most commonly applicable in indicating the score from the norm-referenced tests such as, SAT, GRE and LSAT.

14. What is ANOVA?

Ans→ An ANOVA test is a way to find out if survey or experiment results are significant. You're testing groups to see if there's a difference between them and analyze the difference between the means of more than two groups. A group of psychiatric patients is trying three different therapies: counseling, medication, and biofeedback.

The ANOVA, which stands for the Analysis of Variance test, is a tool in statistics that is concerned with comparing the means of two groups of data sets and to what extent they differ. In simpler and general terms, it can be stated that the ANOVA test is used to identify which process, among all the other processes, is better.

15. How can ANOVA help?

Ans→ ANOVA can also be called analysis of variance and is used in estimating or measuring variation between objects or groups. ANOVA is a collection of statistical models used in solving any variability problem or issue.

ANOVA can be applied in a lot of areas that may seem new and weird to you. It is used to get results and achieve certain objectives in an experiment. It can be applied in many areas in which would take a look at a few of these areas below.

1. Used to design an area

2. Used in identifying gender age differences

3. Used in knowing how far persons can throw a javelin

4. Used to carry out experimental designs

5. Used in analyzing variance between samples

6. Used to determine the best materials to build products for your customers

7. Used in food industries

8. Used in health care industries

9. Used in comparing the gas mileage of different vehicles

10. Used in understanding the impact of different catalyst on chemical reaction rates