

STATISTICS

WORKSHEET 3

Q1. 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

Correct Answer: A) Total Variation = Residual Variation – Regression Variation

Q2. Collection of exchangeable binary outcomes for the same covariate data are called _____ outcomes.

- a) random
- b) direct
- c) binomial
- d) none of the mentioned

Correct Answer: C) Binomial

Q3. How many outcomes are possible with Bernoulli trial?

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

Correct Answer: A) 2

Q4. 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

Correct Answer: A) Type – I Error

Q5. Level of significance is also called:

- a) Power of the test
- b) Size of the test
- c) Level of confidence
- d) Confidence coefficient

Correct Answer: B) Size of Test

Q6. The chance of rejecting a true hypothesis decreases when sample size is:

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

Correct Answer: B) Increases

Q7. Which of the following testing is concerned with making decisions using data?

- a) Probability
- b) Hypothesis
- c) Causal
- d) None of the mentioned

Correct Answer: B) Hypothesis

Q8. 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

Correct Answer: D) All of the mentioned

Q9. Normalized data are centred at _____ and have units equal to standard deviations of the original data

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

Correct Answer: A) 0

Q10. What Is Bayes' Theorem?

Bayes theorem states that the conditional probability of an event A, given the occurrence of another event B, is equal to the product of the likelihood of B, given A and the probability of A. It is given as:

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)}$$

Here, $P(A)$ = how likely A happens (Prior knowledge)- The probability of a hypothesis is true before any evidence is present.

$P(B)$ = how likely B happens(Marginalization)- The probability of observing the evidence.

$P(A/B)$ = how likely A happens given that B has happened(Posterior)-The probability of a hypothesis is true given the evidence.

$P(B/A)$ = how likely B happens given that A has happened(Likelihood)- The probability of seeing the evidence if the hypothesis is true.

Q11. What is z-score?

Z-score is also known as standard score which is used to represent the number of standard deviations by which a raw score is above or below the mean. It is usually used as a part of a z test to draw the interpretations about population data. It can positive, negative, or zero depending upon the position of the raw score with respect to the mean.

Q12. What is t-test?

It is a statistical test used to compare the means of two groups and how they are related. It is often used in hypothesis testing to determine whether a process or treatment actually has an effect on the population of interest, or whether two groups are different from one another.

Q13. What is percentile?

It is a measure in statistics indicating the value below which a given percentage of observations in a group of observations fall.

Q14. What is ANOVA?

ANOVA stands for Analysis of Variance. It is a statistical method used to test differences between two or more means. It is used to test general rather than specific differences among means. Inferences about means are made by analysing variance.

Q15. How can ANOVA help?

The ANOVA test is the initial step in analysing factors that affect a given data set. It allows a comparison of more than two groups at the same time to determine whether a relationship exists between them. Once the test is finished, an analyst performs additional testing on the methodical factors that measurably contribute to the data set's inconsistency. The analyst utilizes the ANOVA test results in an f-test to generate additional data that aligns with the proposed regression models.