

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 H_0 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. a) Power 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

Q10 and Q15 are subjective answer type questions, Answer them in your own words briefly.

10. What Is Bayes' Theorem?

ANS. Bayes' Theorem , is a mathematical formula for determining conditional probability. Conditional probability is the likelihood of an outcome occurring, based on a previous outcome occurring. Bayes' theorem is also called Bayes' Rule or Bayes' Law and is the foundation of the field of Bayesian statistics.

Formula For Bayes' Theorem : $P(A|B) = P(A \cap B) / P(B) = P(A) \cdot P(B|A) / P(B)$

where:

$P(A)$ = The probability of A occurring, $P(B)$ = The probability of B occurring, $P(A|B)$ =The probability of A given B, $P(B|A)$ = The probability of B given A, $P(A \cap B)$ = The probability of both A and B occurring

11. What is z-score?

ANS. A Z-score is a numerical measurement that describes a value's relationship to the mean of a group of values. Z-score is measured in terms of standard deviations from the mean. A Z-score of 1.0 would indicate a value that is one standard deviation from the mean. Z-scores may be positive or negative, with a positive value indicating the score is above the mean and a negative score indicating it is below the mean. In order to calculate the probability of z-score, area under the curve of distribution needs to be find out.

Formula for z-score: $z = (x - \mu) / \sigma$

12. What is t-test?

ANS. A t-test is a type of inferential statistic used to determine if there is a significant difference between the means of two groups, which may be related in certain features. The t-test is one of many tests used for the purpose of hypothesis testing in statistics. It uses standard deviation of sample itself. Thus, t-test is used when the population standard deviation is not available. It has sample size of less than .There are three types of t-tests, and they are categorized as dependent and independent t-tests. Calculating a t-test requires three key data values. They include the difference between the mean values from each data set (called the mean difference), the standard deviation of each group, and the number of data values of each group.

13. What is percentile?

ANS. Percentile can be defined as a number where a certain percentage of scores fall below a particular value. It tells where a score stands relative to other scores.

14. What is ANOVA?

ANS. Analysis of Variance is a statistical tool for comparison of groups i.e. it is used to check if the means of two or more groups are significantly different from each other. ANOVA checks the impact of one or more factors by comparing the means of different samples.. It is used to find correlation between the different groups of categorical variables. If ANOVA gives large F-Value and small P-Value, then there is a strong correlation between the variables.

15. How can ANOVA help?

ANS. ANOVA, or its non-parametric counterparts, allow you to determine if differences in mean values between three or more groups are by chance or if they are indeed significantly different. An ANOVA test is a way to find out if survey or experiment results are significant. In other words, they help you to figure out if you need to reject the null hypothesis or accept the alternate hypothesis. The proper use of ANOVA in analyzing survey data requires that a few assumptions be met including normal distribution of data; independence of cases, and equality of variance (each group's variance is equal). If these assumptions cannot be met then there are non-parametric tests available which do not require these assumptions.