

Name of the Assignment: Statistics (Worksheet – 3)

Submitted by : Shalini Joshi

Designation : Data Science Intern

Date of Submission : 31st Dec,2022

Objective Type Questions

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: 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 centered at..... and have units equal to standard deviations of the original data

a) 0 b) 5 c) 1 d)

Ans: a) 0

Subjective type Questions:

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 having occurred in similar circumstances.

Bayes' theorem relies on incorporating prior probability distributions in order to generate posterior probabilities.

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 describes the position of a raw score in terms of its distance from the mean, when measured in standard deviation units. The z-score is positive if the value lies above the mean, and negative if it lies below the mean.

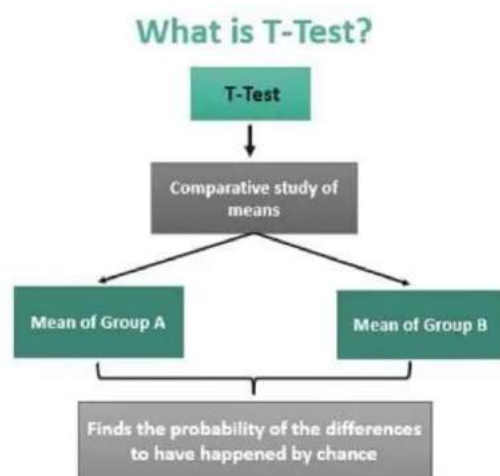
It is also known as a standard score, because it allows comparison of scores on different kinds of variables by standardizing the distribution.

The formula for calculating a z-score is $z = (x - \mu) / \sigma$, where x is the raw score, μ is the population mean, and σ is the population standard deviation.

The value of the z-score tells you how many standard deviations you are away from the mean. If a z-score is equal to 0, it is on the mean.

12. What is t-test?

Ans: A T-test is the final statistical measure for determining differences between two means that may or may not be related. The testing uses randomly selected samples from the two categories or groups. It is a statistical method in which samples are chosen randomly, and there is no perfect normal distribution.



13. What is percentile?

Ans: Percentiles are used to understand and interpret data. They indicate the values below which a certain percentage of the data in a data set is found.

Percentiles can be calculated using the formula $n = (P/100) \times N$, where P = percentile, N = number of values in a data set (sorted from smallest to largest), and n = ordinal rank of a given value.

Percentiles are frequently used to understand test scores and biometric measurements.

14. What is ANOVA?

Ans: 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.

If no true variance exists between the groups, the ANOVA's F-ratio should equal close to 1.

15. How can ANOVA help?

Ans: One-way ANOVA can help you know whether or not there are significant differences between the groups of your independent variables. You may want to test multiple independent variables (such as Location, employment status or education). When you understand how the groups within the independent variable differ (such as USA vs Canada vs Mexico, not location, employment status, or education), you can begin to understand which of them has a connection to your dependent variable.

ANOVA has uses in the following areas:

- Used to design an area; With ANOVA, you can get designs like; Randomized complete block design (RCBD) and Latin square design (LSD).
- Used in identifying gender age differences
- Used in identifying how far a person can throw javelin
- Used in analyzing variance between samples
- Used to determine the best materials to build products for your customers
- Used in healthcare and food industries

*****The End*****