

## STATISTICS WORKSHEET - 3

1. Which of the following is the correct formula for total variation?

Answer – **(b) Total Variation = Residual Variation + Regression Variation**

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

Answer – **(c) binomial**

3. How many outcomes are possible with Bernoulli trial?

Answer – **(a) 2**

4. If  $H_0$  is true and we reject it is called?

Answer – **(a) Type-I error**

5. Level of significance is also called?

Answer – **(b) Size of the test**

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

Answer – **(b) Increase**

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

Answer – **(b) Hypothesis**

8. What is the purpose of multiple testing in statistical inference?

Answer – **(d) All of the mentioned**

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

Answer – **(a) 0**

10. What Is Bayes' Theorem?

Answer – **The Bayes theorem predicts the likelihood that an event connected to any condition would occur. It is also taken into account in the situation of conditional probability. The Bayes theorem is frequently referred to as the probability of "causes" formula.**

**Bayes Theorem Statement -**

**Let  $E_1, E_2, \dots, E_n$  be a set of events associated with a sample space  $S$ , where all the events  $E_1, E_2, \dots, E_n$  have nonzero probability of occurrence and they form a partition of  $S$ . Let  $A$  be any event associated with  $S$ , then according to Bayes theorem**

$$P(E_i | A) = \frac{P(E_i)P(A|E_i)}{\sum_{k=1}^n P(E_k)P(A|E_k)}$$

for any  $k = 1, 2, 3, \dots, n$

11. What is z-score?

Answer – **The z score is a measurement of how many standard deviations a raw score is below or above the population mean. If the value is higher than the mean, it will be positive; if it is lower, it will be negative. Standard score is another name for it. It shows how far away from the mean an object is, in terms of standard deviations. The mean and population standard deviation must be known in order to use a z-score. The likelihood of a score happening inside a typical normal distribution can be calculated with the use of a z score. We may also compare two scores from other samples thanks to it. A z score table is a table for the values of, reflecting the values of the normal distribution's cumulative distribution function.**

**Formula -**

**The equation is given by  $z = (x - \mu) / \sigma$ .**

**$\mu$  = mean  $\sigma$  = standard deviation  $x$  = test value**

When we have multiple samples and want to describe the standard deviation of those sample means, we use the following formula -  $z = (x - \mu) / (\sigma / \sqrt{n})$

12. What is t-test?

Answer – To evaluate whether there is a significant difference between the means of two groups and their relationships, a t-test is an inferential statistic that is used. When data sets contain unknown variances and a normal distribution, such as the data set obtained from tossing a coin 100 times, t-tests are utilised.

In order to evaluate statistical significance, the t-test, a test used for hypothesis testing in statistics, uses the t-statistic, the values of the t-distribution, and the degrees of freedom.

13. What is percentile?

Answer – A percentile in statistics refers to a score's position in relation to other scores from the same set. Although percentile has no one fixed meaning, it is frequently described as the proportion of values in a set of data scores that are lower than a particular value.

$$P_x = \frac{x(n + 1)}{100}$$

$P_x$  = The value at which x percentage of data lie below that value

$n$  = Total number of observations

14. What is ANOVA?

Answer – By comparing the amount of variance inside each sample to the amount of variation across samples, an ANOVA is used to test for differences in population mean values. The idea that the means of two or more populations are equal is tested through variance analysis.

The ANOVA test is used in regression analysis to evaluate the influence of independent variables on the dependent variable.

15. How can ANOVA help?

Answer – Analysis of Variance is referred to as ANOVA. Ronald Fisher created a statistical test in 1918, and it has been in use ever since. Simply put, an ANOVA analysis determines if the means of three or more independent groups differ statistically. The most fundamental ANOVA is one-way ANOVA.