**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

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

3. How many outcomes are possible with Bernoulli trial?

**a) 2**

b) 3

c) 4

d) None of the mentioned

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

5. Level of significance is also called:

a) Power of the test

**b) Size of the test**

c) Level of confidence

d) Confidence coefficient

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

a) Decrease

**b) Increase**

c) Both of them

d) None

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

a) Probability

**b) Hypothesis**

c) Causal

d) None of the mentioned

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**

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

**Q10and Q15 are subjective answer type questions, Answer them in your own words briefly.**

10. What Is Bayes' Theorem?

Bayes Theorem provides a principled way for calculating a conditional probability.

Conditional probability is the likelihood of an outcome occurring, based on a previous outcome occurring.

**P(A|B)= [P(B|A) \* P(A) ] / P(B)**

A, B = events

P(A|B) = probability of A given B is true

P(B|A) = probability of B given A is true

P(A), P(B) = the independent probabilities of A and B

11. What is z-score?

**z-score gives an idea of how far from the**[mean](https://www.statisticshowto.com/probability-and-statistics/statistics-definitions/mean-median-mode/)**a data point is.** But more technically it’s a measure of how many [standard deviations](https://www.statisticshowto.com/probability-and-statistics/standard-deviation/) below or above the [population mean](https://www.statisticshowto.com/population-mean/)a [raw score](https://www.statisticshowto.com/raw-score/) is. If a Z-score is 0, it indicates that the data point's score is identical to the mean score. 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.

A z-score can be placed on a [**normal distribution**](https://www.statisticshowto.com/probability-and-statistics/normal-distributions/) curve. Z-scores range from -3 standard deviations (which would fall to the far left of the normal distribution curve) up to +3 standard deviations (which would fall to the far right of the normal distribution curve

**z = (x-µ)/σ**

Z = standard score

x = observed value

**µ** = mean of the sample

**σ** = standard deviation of the sample

12. What is t-test?

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

13. What is percentile?

In statistics, a percentile is a score below which a given percentage of scores in its frequency distribution fall or a score at or below which a given percentage fall

14. What is ANOVA?

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.

The Formula for ANOVA is:

**F= MSE/MST**

​ where:

F=ANOVA coefficient

MST=Mean sum of squares due to treatment

MSE=Mean sum of squares due to error

15. How can ANOVA help?

The ANOVA test is used to determine the influence that independent variables have on the dependent variable in a regression study.

 It help you to figure out if we need to reject the null hypothesis or accept the alternate hypothesis