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?
   1. Total Variation = Residual Variation – Regression Variation
   2. Total Variation = Residual Variation + Regression Variation
   3. Total Variation = Residual Variation \* Regression Variation
   4. All of the mentioned

Answer- B) Total Variation = Residual Variation + Regression Variation

1. Collection of exchangeable binary outcomes for the same covariate data are called outcomes.
   1. random
   2. direct
   3. binomial
   4. none of the mentioned

Answer- c) Binomial

1. How many outcomes are possible with Bernoulli trial?
   1. 2
   2. 3
   3. 4



* 1. None of the mentioned

Answer- a) 2

1. If Ho is true and we reject it is called
   1. Type-I error
   2. Type-II error
   3. Standard error
   4. Sampling error
   5. Answer- A) Type-I error
2. Level of significance is also called:
   1. Power of the test
   2. Size of the test
   3. Level of confidence
   4. Confidence coefficient
3. The chance of rejecting a true hypothesis decreases when sample size is:
   1. Decrease
   2. Increase
   3. Both of them
   4. None

Answer- B) Increased

1. Which of the following testing is concerned with making decisions using data?
   1. Probability
   2. Hypothesis
   3. Causal
   4. None of the mentioned

Answer- B) Hypothesis

1. What is the purpose of multiple testing in statistical inference?
   1. Minimize errors
   2. Minimize false positives
   3. Minimize false negatives
   4. All of the mentioned

Answer- D) All of the mentioned

1. Normalized data are centred at and have units equal to standard deviations of the original data
   1. 0
   2. 5
   3. 1
   4. 10

Answer- A) 0

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

1. What Is Bayes' Theorem?

Answer- In statistics and probability theory, the Bayes’ theorem (also known as the Bayes’ rule) is a mathematical formula used to determine the conditional probability of events.

Bayes, is a mathematical formula for determining [conditional probability](https://www.investopedia.com/terms/c/conditional_probability.asp). Conditional probability is the likelihood of an outcome occurring, based on a previous outcome having occurred in similar circumstances. Bayes' theorem provides a way to revise existing predictions or theories (update probabilities) given new or additional evidence.

​*P*(*A*∣*B*)=*P*(*B*)*P*(*A*⋂*B*)​=*P*(*B*)*P*(*A*)⋅*P*(*B*∣*A*

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*⋂*B*))= The probability of both A and B occurring​

1. What is z-score?

Answer- **Z-score is also known as standard score** gives us an idea of how far a data point is from the mean. It indicates how many standard deviations an element is from the mean. Hence, Z-Score is measured in terms of standard deviation from the mean. For example, a standard deviation of 2 indicates the value is 2 standard deviations away from the mean. In order to use a z-score, we need to know the population mean (μ) and also the population standard deviation (σ).

1. What is t-test?

Answer- T-tests are statistical [hypothesis tests](https://statisticsbyjim.com/glossary/hypothesis-tests/) that you use to analyze one or two [sample](https://statisticsbyjim.com/glossary/sample/) means. Depending on the t-test that you use, you can compare a sample mean to a hypothesized value, the means of two independent samples, or the difference between paired samples.

There are many types of t-test. Some of these are:

* The one-sample t-test, which is used to compare the mean of a population with a theoretical value.
* The unpaired two-sample t-test, which is used to compare the mean of two independent given samples.
* The paired t-test, which is used to compare the means between two groups of samples that are related.

1. What is percentile?

Answer- A percentile is **a comparison score between a particular score and the scores of the rest of a group**. It shows the percentage of scores that a particular score surpassed. For example, if you score 75 points on a test, and are ranked in the 85 th percentile, it means that the score 75 is higher than 85% of the scores.

1. What is ANOVA?

Answer- ANOVA is to test for differences among the means of the population by examining the amount of variation within each sample, relative to the amount of variation between the samples.  Analyzing variance [tests the hypothesis](https://www.simplilearn.com/tutorials/statistics-tutorial/hypothesis-testing-in-statistics) that the means of two or more populations are equal.

1. How can ANOVA help?

Answer- As an analyst, you might use Analysis of Variance (ANOVA) to test a particular hypothesis. You'd use ANOVA to figure out how your various groups react, with the null hypothesis being that the means of the various groups are equal. If the difference between the two [populations](https://www.simplilearn.com/tutorials/machine-learning-tutorial/population-vs-sample) is statistically significant, then the two populations are unequal.

