Statistics Worksheet 3

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

Ans : Total Variation = Residual Variation + Regression Variation

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

Ans : Binomial Outcome

3) How many outcomes are possible with Bernoulli trial?

Ans : 2

4) If Ho is true and we reject it is called

Ans : type 1 error

5)Level of significance is also called:

Ans : Confidence coefficient

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

Ans : increases

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

Ans ,: Hypothesis

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

Ans : All of the mentioned

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

Ans : 0

10)What Is Bayes' Theorem?

Ans : Bayes' Theorem states that the conditional probability of an event, based on the occurrence of another event, is equal to the likelihood of the second event given the first event multiplied by the probability of the first event.Bayes' Theorem calculates the conditional probability of an event, based on the values of specific related known probabilities.Bayes Theorem provides a useful method for thinking about the relationship between a data set and a probability. In other words, the theorem says that the probability of a given hypothesis being true based on specific observed data can be stated as finding the probability of observing the data given the hypothesis multiplied by the probability of the hypothesis being true regardless of the data, divided by the probability of observing the data regardless of the hypothesis.

Applications of Bayes' Theorem are widespread and not limited to the financial realm. For example, Bayes' theorem can be used to determine the accuracy of medical test results by taking into consideration how likely any given person is to have a disease and the general accuracy of the test. Bayes' theorem relies on incorporating prior probability distributions in order to generate posterior

probabilities.Bayes Theorem can be derived for events and random variables separately using the definition of conditional probability and density. Here, the joint probability P(A ∩ B) of both events A and B being true such that, P(B ∩ A) = P(A ∩ B)

11) What is z-score?

Ans : The Z-score, or standard score, is the number of standard deviations a given data point lies above or below the mean. The mean is the average of all values in a group, added together, and then divided by the total number of items in the group.To calculate the Z-score, subtract the mean from each of the individual data points and divide the result by the standard deviation. Results of zero show the point and the mean equal. A result of one indicates the point is one standard deviation above the mean and when data points are below the mean, the Z-score is negative.

In most large data sets, 99% of values have a Z-score between -3 and 3, meaning they lie within three standard deviations above or below the mean.

The basic z score formula for a sample is:

$z = (x – \mu) / \sigma$

12) What is t-test?

Ans : A t-test is an inferential statistic used to determine if there is a significant difference between the means of two groups and how they are related. T-tests are used when the data sets follow a normal distribution and have unknown variances, like the data set recorded from flipping a coin 100 times. The t-test is a test used for hypothesis testing in statistics and uses the t-statistic, the t-distribution values, and the degrees of freedom to determine statistical significance.A t-test compares the average values of two data sets and determines if they came from the same population. In the above examples, a sample of students from class A and a sample of students from class B would not likely have the same mean and standard deviation. Similarly, samples taken from the placebo-fed control group and those taken from the drug prescribed group should have a slightly different mean and standard deviation.

13)What is percentile?

Ans : 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. In statistics, a k-th percentile (percentile score or centile) is a score below which a given percentage k of scores in its frequency distribution falls (exclusive definition) or a score at or below which a given percentage falls (inclusive definition).

14)What is ANOVA?

Ans : ANOVA stands for Analysis of Variance. One-Way Analysis of Variance tells you if there are any statistical differences between the means of three or more independent groups. An ANOVA test is a type of statistical test used to determine if there is a statistically significant difference between two

or more categorical groups by testing for differences of means using variance. Another Key part of ANOVA is that it splits the independent variable into 2 or more groups. For example, one or more groups might be expected to influences the dependent variable while the other group is used as a control group, and is not expected to influence the dependent variable.

15)How can ANOVA help?

Ans : The one-way ANOVA can help you know whether or not there are significant differences between the means of your independent variables (such as the first example: age, sex, income). When you understand how each independent variable's mean is different from the others, you can begin to understand which of them has a connection to your dependent variable (landing page clicks), and begin to learn what is driving that behavior.