STATISTICS WORKSHEET-3

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

Total Variation = Residual Variation + Regression Variation

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

Binomial

How many outcomes are possible with Bernoulli trial?

Two outcomes.

If Ho is true and we reject it is called a) Type-I error b) Type-II error c) Standard error d) Sampling error

Type-I error.

Level of significance is also called:

Size of the test

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

Decrease & Increase (Both of them)

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

Hypothesis Testing.

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

All of the mentioned

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

0 b) 5 c) 1 d) 10

0

What Is Bayes' Theorem?

The Bayes' theorem (also known as the Bayes' rule) is a mathematical formula used to determine the conditional probability of events. Essentially, the Bayes' theorem describes the probability of an event based on prior knowledge of the conditions that might be relevant to the event.

Formula for Bayes' Theorem

The Bayes' theorem is expressed in the following formula:

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)}$$

Where:

- P(A|B) – the probability of event A occurring, given event B has occurred
- P(B|A) – the probability of event B occurring, given event A has occurred
- P(A) – the probability of event A
- P(B) – the probability of event B

What is z-score?

A z-score gives you an idea of how far from the mean a data point is. But more technically it tells you measure of how many std. deviation below or above the popular mean a raw score is. A z-score can be placed on a normal distribution 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

What is t-test?

A t-test is a statistical test that is used to compare the means of two groups. It is often used in hypothesis testing to determine whether a process or treatment actually has an effect on the population of interest, or whether two groups are different from one another.

What is percentile?

A percentile (or a centile) is a measure used in statistics indicating the value below which a given percentage of observations in a group of observations fall. For example, the 20th percentile is the value (or score) below which 20% of the observations may be found.

What is ANOVA?

Analysis of variance (ANOVA) is an analysis tool used in statistics that splits an observed aggregate variability found inside a data set into two parts: systematic factors and random factors. The systematic factors have a statistical influence on the given data set, while the random factors do not. Analysts use the ANOVA test to determine the influence that independent variables have on the dependent variable in a regression study.

How can ANOVA help?

ANOVA is helpful for testing three or more variables. It is similar to multiple two-sample t-tests.