Statistics Worksheet-3 (Internship-23, Konatala Mohit ID-34)

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

b) Total Variation = Residual Variation + Regression Variation

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

c) binomial

3. How many outcomes are possible with Bernoulli trial?

a) 2

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

a) Type-I error

5. Level of significance is also called:

b) Size of the test

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

b) Increase

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

b) Hypothesis

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

d) All of the mentioned

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

a) 0

10. What Is Bayes' Theorem?

Bayes' Theorem, named after 18th-century British mathematician Thomas Bayes, is a mathematical formula for determining conditional probability. 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. In finance, Bayes' Theorem can be used to rate the risk of lending money to potential borrowers. The theorem is also called Bayes' Rule or Bayes' Law and is the foundation of the field of Bayesian statistics.

11. What is z-score?

Simply put, a z-score (also called a standard score) gives you an idea of how far from the mean a data point is. But more technically it's a measure of how many standard deviations below or above the population mean a raw score is. A z-score can be placed on a normal distribution curve. Z-score range from -3 standard deviations (which would fall to the fat left of the normal distribution curve) up to +3 standard deviation (which would fall to the far right of the normal distribution curve). In order to use a z-score, you need to know the mean μ and also the population standard deviation σ.

12. What is t-test?

A t-test is a type of inferential statistics used to determine if there is a significant difference between the means of two groups, which may be related in certain features. It is mostly used when the data sets, like the data set recorded as the outcome from flipping a coin 100 times, would follow a normal distribution and may have unknown variances. A t-test is used as a hypothesis testing tool, which allows testing of an assumption applicable to a population.

13. 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.

14. 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.

15. How can ANOVA help?

ANOVA is helpful for testing three or more variables. It is similar to multiple two-sample t-tests. However, it results in fewer type I errors and is appropriate for a range of issues. ANOVA groups differences by comparing the means of each group and includes spreading out the variance into diverse sources. It is employed with subjects, test groups, between groups and within groups.