STATISTICS WORKSHEET-3

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

Ans: b) Total Variation = Residual Variation + Regression Variation

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

Ans: c) binomial

3. How many outcomes are possible with Bernoulli trial?

Ans: a) 2

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

Ans: a) Type-I error

5. Level of significance is also called:

Ans: b) Size of the test

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

Ans: d) None

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

Ans: b) Hypothesis

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

Ans: d) All of the mentioned

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

Ans: a) 0

10. What Is Bayes' Theorem?

Ans: 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 occurring. 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.

11. What is z-score?

Ans: A Z-score is a numerical measurement that describes a value's relationship to the mean of a group of values. Z-score is measured in terms of standard deviations from the mean. 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.

12. What is t-test?

Ans: 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. 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?

Ans: In statistics, a percentile (or a centile) is a score below which a given percentage of scores in its frequency distribution falls (exclusive definition) or a score at or below which a given percentage falls (inclusive definition). For example, the 50th percentile (the median) is the score below which (exclusive) or at or below which (inclusive) 50% of the scores in the distribution may be found.

14. What is ANOVA?

Ans: 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?

Ans: ANOVA, or its non-parametric counterparts, allow you to determine if differences in mean values between three or more groups are by chance or if they are indeed significantly different. ANOVA is particularly useful when analysing the multi-item scales common in market research.