STATISTICS WORKSHEETS-3

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.

a) random

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:

c) Level of confidence

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

c) 1

10. What Is Bayes' Theorem?

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. Essentially, the Bayes’ theorem describes the probability of an event based on prior knowledge of the conditions that might be relevant to the event.Bayes’ theorem is also used in various disciplines, with medicine and pharmacology as the most notable examples. In addition, the theorem is commonly employed in different fields of finance. Some of the applications include but are not limited to, modeling the risk of lending money to borrowers or forecasting the probability of the success of an investment.

The Bayes’ theorem is expressed in the following formula:



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.

11. What is z-score?

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. A Z-score of 1.0 would indicate a value that is one standard deviation from the mean. Technically, a z-score is the number of standard deviations from the mean value of the reference population.

12. 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. The t-value measures the size of the difference relative to the variation in your sample data. Put another way, T is simply the calculated difference represented in units of standard error. The greater the magnitude of T, the greater the evidence against the null hypothesis.

13. What is percentile?

A percentile (or a centile) is a score below which a given percentage of scores in its frequency distribution fall (exclusive definition) or a score at or below which a given percentage fall (inclusive definition). The term percentile and the related term percentile rank are often used in the reporting of scores from norm-referenced tests.

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. Analysis of variance (ANOVA) is a statistical technique that is used to check if the means of two or more groups are significantly different from each other. ANOVA checks the impact of one or more factors by comparing the means of different samples.

15. How can ANOVA help?

ANOVA is used to compare differences of means among more than 2 groups. It does this by looking at variation in the data and where that variation is found. Specifically, ANOVA compares the amount of variation between groups with the amount of variation within groups. 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 analyzing the multi-item scales common in market research.