WORKSHEET-3 STATISTICS

1. 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. a) random b) direct c) binomial d) none of the mentioned - 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: - a) Power 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 (Minimize errors, Minimize false positives, Minimize false negatives)

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

a) 0

10.Bayes' Theorem

Bayes Theorem provides a principled way for calculating a conditional probability. It is a deceptively simple calculation, although it can be used to easily calculate the conditional probability of events where intuition often fails. **Bayes' theorem** describes the probability of occurrence of an event related to any condition. It is also considered for the case of conditional probability. Bayes theorem is also known as the formula for the probability of "causes".

If A and B are two events Bayes theorem formula is given by,

$$P(A|B) = \frac{P(B|A)\,P(A)}{P(B)} \quad \text{where } P(B) \neq 0$$

11. Z-score

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.

It is useful to standardized the values (raw scores) of a normal distribution by converting them into z-scores because:

(a) it allows researchers to calculate the probability of a score occurring within a standard normal distribution;
(b) and enables us to compare two scores that are from different samples (which may have different means and standard deviations).

12.What is t-test?

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.

13.What is percentile?

In statistics, a percentile is a score below which a given percentage of scores in its frequency distribution falls or a score at or below which a given percentage falls.

14.What is ANOVA?

An ANOVA test is a way to find out if survey or experiment results are significant. In other words, they help you to figure out if you need to reject the null hypothesis or accept the alternate hypothesis.

15.How can ANOVA help?

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 behaviour