

STATISTICS WORKSHEET-3

10. What Is Bayes' Theorem?

Ans:- 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".

Bayes theorem is used to find the reverse probabilities if we know the conditional probability of an event.

One of the many applications of Bayes' theorem is Bayesian inference, a particular approach to statistical inference. Bayesian inference has found application in various activities, including medicine, science, philosophy, engineering, sports, law, etc. For example, we can use Bayes' theorem to define the accuracy of medical test results by considering how likely any given person is to have a disease and the test's overall accuracy. Bayes' theorem relies on consolidating prior probability distributions to generate posterior probabilities. In Bayesian statistical inference, prior probability is the probability of an event before new data is collected.

11. What is z-score?

Ans:- A z-score measures the distance between a data point and the mean using standard deviations. Z-scores can be positive or negative. The sign tells you whether the observation is above or below the mean.

For example, a z-score of +2 indicates that the data point falls two standard deviations above the mean, while a -2 signifies it is two standard deviations below the mean. A z-score of zero equals the mean. Statisticians also refer to z-scores as standard scores, and I'll use those terms interchangeably.

To calculate z-scores, take the raw measurements, subtract the mean, and divide by the standard deviation. Z-scores are important because they offer a comparison between two scores that are not in the same normal distribution. They are also used to obtain the probability of a z-score to take place within a normal distribution. If a z-score gives a negative value, it means that raw data is lesser than mean.

12. What is t-test?

Ans:- *T-test is the final statistical measure for determining differences between two means that may or may not be related. The testing uses randomly selected samples from the two categories or groups. It is a statistical method in which samples are chosen randomly, and there is no perfect normal distribution.*

A T-test studies a set of data gathered from two similar or different groups to determine the probability of the difference in the result than what is usually obtained. The accuracy of the test depends on various factors, including the distribution patterns used and the variants influencing the collected samples. Depending on the parameters,

the test is conducted, and a T-value is obtained as the statistical inference of the probability of the usual resultant being driven by chance.

The final T-test interpretation could be obtained in either of the two ways:

null hypothesis signifies that the difference between the means is zero and where both the means are shown as equal.

alternate hypothesis implies the difference between the means is different from zero. This hypothesis rejects the null hypothesis, indicating that the data set is quite accurate and not by chance.

This T-test, however, is only valid and should be done when the mean or average of only two categories or groups needs to be compared. As soon as the number of comparisons to be made is more than two, conducting this is not recommended.

13. What is percentile?

Ans:- A percentile 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 below which 20% of the observations may be found.

The term percentile and the related term *percentile rank* are often used in the reporting of scores from norm-referenced tests. For example, if a score is at the 86th percentile, where 86 is the percentile rank, it is equal to the value below which 86% of the observations may be found. In contrast, if it is in the 86th percentile, the score is at or below the value of which 86% of the observations may be found. *Every score is in the 100th percentile.*

The 25th percentile is also known as the first quartile (Q1), the 50th percentile as the median or second quartile (Q2), and the 75th percentile as the third quartile (Q3). In general, percentiles and quartiles are specific types of quantiles.

14. What is ANOVA?

Ans:- An ANOVA test is a type of statistical test used to determine if there is a statistically significant difference between two or more categorical groups by testing for differences of means using variance. A one-way ANOVA is used for three or more groups of data, to gain information about the relationship between the dependent and independent variables.

Another Key part of ANOVA is that it splits the independent variable into 2 or more groups. For example, one or more groups might be expected to influence the dependent variable while the other group is used as a control group, and is not expected to influence the dependent variable.

15. How can ANOVA help?

Ans:- You might use Analysis of Variance (ANOVA) as a marketer, when you want to test a particular hypothesis. You would use ANOVA to help you understand how your different groups respond, with a null hypothesis for the test that the means of the

different groups are equal. If there is a statistically significant result, then it means that the two populations are unequal (or different).

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 behavior.