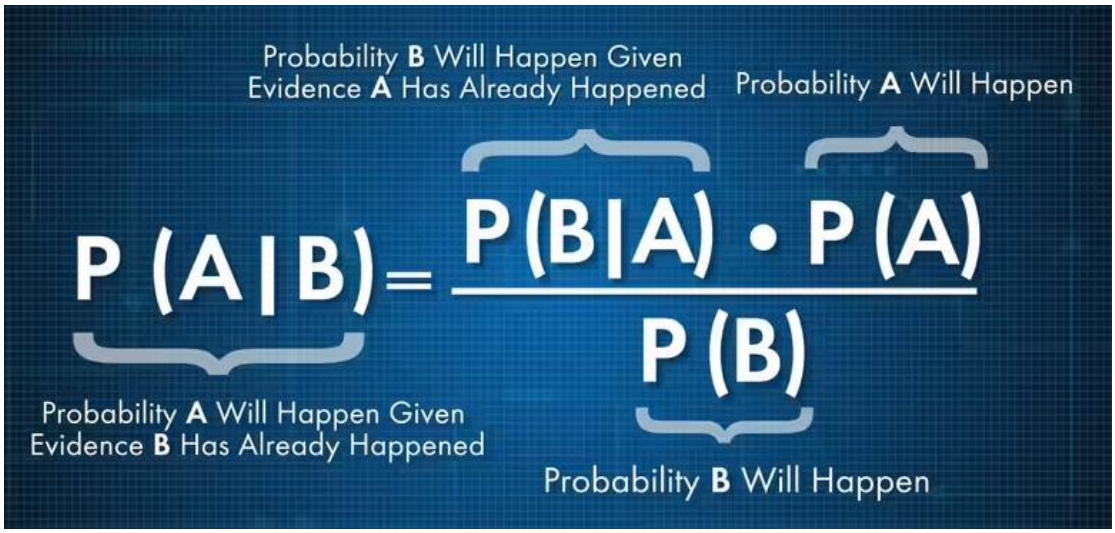
Assignment-3

Statistics(Answer sheet)

Ans.1 (b) Total Variance = Residual Variation + Regression Variation
Ans.2 (c) binomial
Ans.3 (a) 2
Ans.4 (a) Type-I error
Ans.5 (a) Power of the test
Ans.6 (b) Increase
Ans.7 (b) Hypothesis
Ans.8 (d) All of the mentioned
Ans.9 (a) 0

Ans. 10 Bayes' Theorem was named after 18th century mathematician 'Thomas Bayes'. This theorem allows us to update the predicted probabilities of an event by incorporating new information. This theorem has become a useful element in the implementation of machine learning.

Formula for Bayes' Theorem



Ans. 11 Z-score is a statistical measure that tells us how many standard deviations a specific score lies above or below the mean. For the calculation of the z-score, we need a row score, the mean, and the standard deviation. In z-score, we can find the relationship between mean and score points. It shows on bell curve.

Ans. 12 A t-test is a test used to determine if there is a statistically significant difference between the mean of two variables. It is used for hypothesis testing also. Calculating a t-test requires three fundamental data values including the difference between the mean values from each data set, the standard deviation of each group, and the number of data values. T-test can be dependent or independent.

Ans. 13 A percentile is a term that describes the comparison of the score to other scores from the same set. There is no any universal definition of percentile, it is commonly expressed as the percentage of values in a set of data scores that fall below a given value. It shows as 25%, 50%, and 75% terms.

Ans. 14 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.

Ans. 15 ANOVA is helpful for testing three or more variables. It is similar to multiple two-sample t-test. 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.