Statistical model

A Statistical Model is the use of statistics to build a representation of the data and then conduct analysis to infer any relationships between variables or discover insights

Non statistical model

it does not take a sample and use algorithms in order to validate a hypothesis. Instead, it takes input from a large volume of data and outputs the results from the data alone. This means that all the available data is taken into account.

Statistical test

A statistical test provides a mechanism for making quantitative decisions about a process or processes. The intent is to determine whether there is enough evidence to "reject" a conjecture or hypothesis about the process. ... A classic use of a statistical test occurs in process control studies.

Statistical tests are commonly classified as parametric and non-parametric tests. Parametric tests are conducted, with an assumption that the data follows a Gaussian distribution. If this assumption fails, then non-parametric tests are considered for hypothesis testing.

Most popular ones and when to use them

Chi-Squared Test

Chi-squared test is a well-known test even for those who are starting with statistical machine learning. Here, this test is used to check whether two categorical variables are related or independent. And, it is assumed that the observations used in the calculation of the contingency table are independent.

Student's t-test

Tests whether the means of two independent samples are significantly different.

Observations in each sample are independent and identically distributed (iid). Observations in each sample are normally distributed. Observations in each sample have the same variance.

Analysis of Variance Test (ANOVA)

ANOVA is another widely popular test which is used to test how independent two samples are of each other. Here the observations are assumed to follow a normal distribution without any change in the variance.

Shapiro-Wilk Test

This test is used to check whether the sample data has a Gaussian distribution.

D'Agostino's K^2 Test

Similar to Shapiro-Wilk test, this too is used to check for Gaussian distribution in data samples.

Pearson's Correlation Coefficient

A statistical test for checking correlation between two samples and whether they have a linear relationship.

Spearman's Rank Correlation

Observations in each sample are assumed that they can be ranked, for checking whether the relationship is monotonic or not.

Mann-Whitney U Test

A non-parametric statistical hypothesis test to check for independent samples and to find whether the distributions are equal or not.

Kruskal-Wallis H Test

Like previous tests, Kruskal-Wallis hypothesis test also makes the same assumptions regarding the distribution and ranking of the observations in each sample. And, the test is carried to check for the independence of the observations from each other.

Friedman Test

Friedman test checks whether the distributions of two or more paired samples are equal or not.