

Parametric Tests

Parametric Comparison Tests

1. t-Test

A t-test is a type of inferential statistic used to determine if there is a significant difference between the means of two groups, which may be related in certain features. It is mostly used when the data sets, like the data set recorded as the outcome from flipping a coin 100 times, would follow a normal distribution and may have unknown variances. A t-test is used as a hypothesis testing tool, which allows testing of an assumption applicable to a population.

a. one-sample t-Test

The mean of a single group is compared with a given mean. For example- to check the increase and decrease in sales if the average sales is given.

$$t = (x_1 - x_2) / (\sigma / \sqrt{n_1} + \sigma / \sqrt{n_2}),$$

where x_1 and x_2 are mean of sample 1 and sample 2 respectively.

b. Two-sample t-Test

i. Un-paired

The independent t-test which is also called the two sample t-test or student's t-test, is a statistical test that determines whether there is a statistically significant difference between the means in two unrelated groups. For example -comparing boys and girls in a population.

ii. Paired

Tests for the difference between two variables from the same population(pre- and post test score). For example- In a training program performance score of the trainee before and after completion of the program.

2. ANOVA

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.

a. One-way ANOVA

A one-way ANOVA evaluates the impact of a sole factor on a sole response variable. It determines whether all the samples are the same. The one-way ANOVA is used to determine whether there are any

statistically significant differences between the means of three or more independent (unrelated) groups.

b. Two-way ANOVA

A two-way ANOVA is an extension of the one-way ANOVA. With a two-way ANOVA, there are two independents. For example, a two-way ANOVA allows a company to compare worker productivity based on two independent variables, such as salary and skill set. It is utilized to observe the interaction between the two factors and tests the effect of two factors at the same time.

Parametric Correlation Tests

1. Pearson's Correlation

Pearson's correlation coefficient is the test statistics that measures the statistical relationship, or association, between two continuous variables. It is known as the best method of measuring the association between variables of interest because it is based on the method of covariance. It gives information about the magnitude of the association, or correlation, as well as the direction of the relationship.

2. Regression

Regression analysis is a reliable method of identifying which variables have impact on a topic of interest. The process of performing a regression allows you to confidently determine which factors matter most, which factors can be ignored, and how these factors influence each other.

Non-parametric Tests

Non-parametric Comparison Tests

1. One-Sample Wilcoxon Signed rank test

Wilcoxon Signed-Rank test is the equivalent non-parametric t-test and this may be used when the dependent variable is not normally distributed. The Wilcoxon signed test is designed to test hypothesis about the location (median) of a population distribution.

2. Mann Whitney's U-Test

The Mann-Whitney U test is used to compare differences between two independent groups when the dependent variable is either ordinal or continuous, but not normally distributed. For example, you could use the Mann-Whitney U test to understand whether attitudes towards pay discrimination, where attitudes are measured on an ordinal scale, differ based on gender (i.e., your dependent variable would be "attitudes towards pay discrimination" and your independent variable would be "gender", which has two groups: "male" and "female").

3. Wilcoxon Test

The Wilcoxon test, which can refer to either the rank sum test or the signed rank test version, is a nonparametric statistical test that compares two paired groups. The tests essentially calculate the difference between sets of pairs and analyze these differences to establish if they are statistically significantly different from one another.

4. Kruskal-Wallis Test

The Kruskal-Wallis H test (sometimes also called the "one-way ANOVA on ranks") is a rank-based nonparametric test that can be used to determine if there are statistically significant differences between two or more groups of an independent variable on a continuous or ordinal dependent variable. It is considered the nonparametric alternative to the one-way ANOVA, and an extension of the Mann-Whitney U test to allow the comparison of more than two independent groups.

Correlation

1. Spearman's Correlation

Spearman's correlation is equivalent to calculating the Pearson correlation coefficient on the ranked data. So ρ will always be a value between -1 and 1. The further away ρ is from zero, the stronger the relationship between the two variables. The sign of ρ corresponds to the direction of the relationship. If it is positive, then as one variable increases, the other tends to increase. If it is negative, then as one variable increases, the other tends to decrease.

2. Kendall's Tau

Kendall's Tau is a correlation suitable for quantitative and ordinal variables. It indicates how strongly 2 variables are monotonously related. Like so, Kendall's Tau serves the exact same purpose as the Spearman rank correlation. The reasoning behind the 2 measures, however, is different.

3. Regression

Nonparametric regression, like linear regression, estimates mean outcomes for a given set of covariates. Unlike linear regression, nonparametric regression is agnostic about the functional form between the outcome and the covariates and is therefore not subject to misspecification error.