Testing Assumptions

Assumption tests, such as those for normality and homoscedasticity (constant variance), are indispensable to any statistical analysis workflow. They ensure the validity, reliability, and robustness of statistical analyses. Although these tests test assumptions about parametric statistical methods, they are generally non-parametric, meaning they do not assume a specific probability distribution for the data.

Non-parametric assumption tests, in contrast to their parametric counterparts, do not require the estimation of parameters under the assumption of a specific distribution. Instead, they utilise ranking, sign tests, or other methods that do not rely on the underlying distribution's exact shape. This practical approach allows these tests to be used across a broad spectrum of data types and scenarios, without the need for stringent distributional assumptions.

The assumptions of normality and homoscedasticity are critical to linear regression and other parametric models. Normality assumes that the residuals (the differences between observed and predicted values) follow a normal distribution. In contrast, homoscedasticity assumes that the variance of the residuals is constant across all levels of the independent variable(s).

Using non-parametric tests to evaluate the validity of these assumptions, we can effectively determine whether the data meet the prerequisites for applying parametric methods, such as linear regression or analysis of variance (ANOVA). In instances where the assumptions are not met, parametric methods may not be suitable, and alternative approaches, such as transformations or non-parametric methods, must be considered.

Non-parametric assumption tests are part of the data exploration and model validation process. They allow us to examine the data without making restrictive distributional assumptions, ensure that the chosen statistical methods are appropriate and that the conclusions drawn from the analysis are reliable.

It's important to note that while assumption tests do not assume a specific distribution, they may still have underlying assumptions of their own, such as independence of observations or continuity of the distribution. Therefore, we must carefully consider the assumptions and limitations of these tests and interpret the results in conjunction with domain knowledge and the specific goals of the analysis.

Tests for Normality

Shapiro-Wilk Test

Kolmogorov-Smirnov Test

Anderson-Darling Test

Lilliefors Test

Jarque-Bera Test

Tests for Homoscedasticity

Breusch-Pagan Test

White's Test

Levene's Test

Bartlett's Test

Fligner-Killeen Test