Simfit

Tutorials and worked examples for simulation,
curve fitting, statistical analysis, and plotting.
http://www.simfit.org.uk

It is frequently useful to examine a given data set without necessarily having any particular statistical tests in mind, but rather as a preliminary to more searching investigations.

In this section the following two definitions are used.

1. A single sample supplied as a column of numbers will be referred to as a vector

2. A rectangular table of numbers will be referred to as a matrix

SIMFIT provides a selection of techniques that can be used in order to make a preliminary examination of such data sets by choosing [Statistics] from the main SIMFIT menu followed by [Data exploration].

- **Exhaustive analysis of a vector**

  Displays a table with a summary of estimated moments, ranges, coefficients of variation/skew/kurtosis. Creates plots as histograms, cumulative distributions, pie charts, bar charts, time series, centered rods, normal/half-normal scores. Performs tests on runs, and/or signs, or for a normal distribution.

- **Exhaustive analysis of a matrix**

  In addition to plotting barcharts and calculating covariance or correlation matrices it is possible to select individual rows or columns for the previous exhaustive analysis of a vector procedure.

- **Exhaustive analysis of a multivariate normal matrix**

  Numerous tests are provided to ascertain if it is reasonable to assume multivariate normality and sphericity before uncritically proceeding to employ MANOVA techniques.

- **Parametric t tests on groups across rows of a matrix**

  This creates a summary of the results from tests on rows where columns have been assigned to groups.

- **Nonparametric tests across rows of a matrix**

  Where the normality and constant variance required for t tests are not justified a similar procedure can be performed using nonparametric tests.

- **All possible pairwise tests on n vectors or a library file**

  Given sets of vectors it is possible to use a selection of parametric and nonparametric tests on all possible pairs of columns. Naturally, with such a procedure there are limits to the number of comparisons allowed, i.e. $n(n-1)/2$ for n vectors.

- **Robust analysis of 1 sample**

  Results can be obtained for parameters from winzorized samples.

- **Robust analysis of 2 samples**

  Two samples can be compared by robust methods.

Of course it is assumed that users will be aware of the limitations arising from multiple tests on the same data implied by some of these procedures, as described for the Bonferroni and related techniques in the SIMFIT reference manual. They should rather interpret the results as preliminary, as should be expected for data exploration.