Transforming Samples

This is a complex dialog box. It allows many useful mathematical and statistical operations to transform one or more columns of data, and then save the results in either the same column (replacing the original data) or the next empty column. You can use the Apply button to implement multiple choices one after another, the Do One button perform only one operation then close the dialog, or the Cancel button or top left window close box to close the dialog without doing any operations.

You must have some data in the spreadsheet before using this.

The left hand column is used to select the single source column, just click on the column you wish to use. The middle list specifies the operation to perform, and the right hand side of the dialog will display other information required for specific operations.

When you click on an operation, the options to the right appear. For the simple arithmetic operations and any with an "r" (representing a constant value or other column designator) you will see two options for r. At top, you can type in a real number to use for each value in the specified source column using the operation specified in the centre column. Alternatively, you can use the list box below the r field (by clicking on a column) to specify a column of data to use. You must select columns that have the same number of rows of data points for these operations.

For simple transformations (reciprocals, squares, square roots, absolute values, logarithms, exponents), no further options are needed, and you only require to click the Apply button to complete the transformation.

For trigonometric functions (sin, cos, tan, atan) you can check the box at right if you wish your source values to be regarded as radians not degrees.

The Series option allows you to create a matching series of numbers beginning at a starting value specified and incrementing by a specified real number. This is most useful to create a matching integer series of numbers. Matching implies you can only make the series have as many points as the source column.

The statistical operations such as z or N scores, % of sum, running sum will perform the designated operation using the whole source column, then each individual point for the specific calculation. z scores provide the value less the sample mean and divided by the sample estimate of the standard deviation. N scores are calculated after ordering the data, and then using the total count and the individual point's location in the series to allocate an equivalent value on the standard normal distribution. When you then plot these two data samples together (eg using a scatter plot), you should see a near linear relationship if the source data are nearly normally distributed.

The All Selected Columns radio button will be active when any of the first 4 operators (+ - * ÷) are selected. This will allow you to do the same operation on more than 2 columns of data (even if they are non-consecutive). This is useful for adding up more than one column quickly. Other operations can then be done on the new calculated column total.