Nonparametric Association Tests (Binary Dependent)

Author: Autumn Laughbaum, Golden Helix, Inc.

Overview

This function makes use of the scipy package, specifically the `scipy.stats.ranksums` and `scipy.stats.mannwhitneyu` functions. With one binary dependent column, the user can perform nonparametric association tests on all numeric columns.

Recommended Directory Location

Save the script to the following directory:

*..\Application Data\Golden Helix SVS\UserScripts\Spreadsheet\Analysis

Note: The Application Data folder is a hidden folder on Windows operating systems and its location varies between XP and Vista. The easiest way to locate this directory on your computer is to open SVS and select Tools > Open Folder > UserScripts Folder. If saved to the proper folder, this script will be accessible from the spreadsheet Analysis menu.

Preparing to use the Script

This script should be run from a spreadsheet containing a binary dependent column and several active numeric columns.

1. From an appropriate spreadsheet, choose Analysis > Nonparametric Association Tests (Binary Dependent). The Nonparametric Association Tests (Binary Dependent) dialog allows the user to choose the test and output options.
2. Choose the Wilcoxon Rank-sum test or Mann-Whitney test and choose to output or not the Bonferroni adjusted p-values and the –log10 p-values.
3. The resulting spreadsheet has a column containing the test statistic for each active numeric column in the original spreadsheet, a p-value column and optional –log10(P) and Bonf-P columns. If a marker map was applied to the columns of the original spreadsheet, it is reapplied to the rows or the Results spreadsheet.

For more information about the internal scipy functions see:

http://docs.scipy.org/doc/scipy/reference/generated/scipy.stats.ranksums.html#scipy.stats.ranksums

and

http://docs.scipy.org/doc/scipy/reference/generated/scipy.stats.mannwhitneyu.html#scipy.stats.mannwhitneyu