

Quantile Transformation

Author: Autumn Laughbaum, Golden Helix, Inc.

Overview

This script categorizes a numeric column into N user-specified quantiles. The cutoff points are calculated over all non-missing values and column values are compared against these cutoffs with \leq .

Recommended Directory Location

Save the script to the following directory:

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

Note: The **Application Data** folder is a hidden folder on Windows operating systems and its location varies between operating systems. The easiest way to locate this directory on your computer is to open SVS and go to **Tools > Open Folder > User Scripts Folder** and save the script in the **\Spreadsheet\Column** folder. If saved to the proper folder, this script will be accessible from a real-valued column menu.

Using the Script

1. From a spreadsheet that contains a numeric (Integer or Real) column, click on the column header and select **Quantile Transformation**.
2. Enter the **Number of quantiles:** desired and click **OK**.

The script will calculate N cutoff points and compare the values in increasing order to these cutoffs.

For example, if 4 quantiles are selected, the 25th, 50th, 75th and 100th percentiles are calculated as cutoffs (q25, q50, q75, q100). For each value in the column, if the value is \leq q25, the quantile value is 1. Else if the value is \leq q50, the quantile value is 2, and so on. Below shows an example of a real-column transformed into 4 quantiles.

Association Tests (Additive Model) - Corr/Trend P in 4 quantiles [1209]

File Edit Select Quality Assurance Analysis Plot Scripts Help

| Unsort | | R | 1 | I | 2 | R |
|--------|---------------|---|-----------------------|---|----------------------------|---|
| Map | Marker | | Corr/Trend P | | Corr/Trend P - 4 quantiles | |
| 1 | SNP_A-1909444 | | 3.18111303323955e-039 | | | 1 |
| 2 | SNP_A-2237149 | | 0.610846371495617 | | | 3 |
| 3 | SNP_A-4303947 | | 7.40390929240427e-016 | | | 1 |
| 4 | SNP_A-1886933 | | 1.12958274619396e-033 | | | 1 |
| 5 | SNP_A-2236359 | | 0.653263115294475 | | | 4 |
| 6 | SNP_A-2205441 | | 0.715777522173936 | | | 4 |
| 7 | SNP_A-2116190 | | 0.140115315945726 | | | 2 |
| 8 | SNP_A-4291020 | | 0.184087893113927 | | | 2 |
| 9 | SNP_A-1902458 | | 0.626520158003362 | | | 4 |
| 10 | SNP_A-2131660 | | 0.312328152169445 | | | 3 |

Association Tests (Additive Model) - Corr/Trend P in 4 quantiles