Paper 273-2013

An Exploratory Graphical Method for Identifying Associations in Sparse r x c Contingency Tables

Martin Lesser, Feinstein Institute for Medical Research-Biostatistics
Meredith Akerman, Feinstein Institute for Medical Research Biostatistics

ABSTRACT

We investigate a graphical method, based on scree plots, for visualizing "significant" departures between observed and expected cell frequencies in RxC contingency tables, with a large number of rows and/or columns. This method is based on Snedecor and Cochran's (1989) proposal to identify the cells with the largest values of $(O-E)2/E$, known as the contribution to chi-square. The scree plot shows the contributions plotted in descending order, so that the user can detect which cells contribute the significant departures, thus suggesting where the null hypothesis of independence may have been violated. This method may be useful in large sparse RxC tables. We used the following SAS procedures to develop a macro for producing the scree plot: PROC FREQ (chisq, cellchi2, deviation, ODS output), PROC SQL, and PROC GPLOT.

No paper was submitted for publication.

CONTACT INFORMATION

Martin Lesser
Feinstein Institute for Medical Research-Biostatistics
MLesser@nshs.edu

Meredith Akerman
Feinstein Institute for Medical Research Biostatistics
MAkerman@nshs.edu