Computer Program Exchange

TRIAGE: A SAS Program of Dissimilarity Data Diagnostics

Tony D. Thompson and Joseph Lee Rodgers
University of Oklahoma

Description

Proximity data are typically represented either by a continuous spatial model using a multidimensional scaling algorithm or by a discrete network model using a clustering algorithm (e.g., Pruzansky, Tversky, & Carroll, 1982). Recently, diagnostics have been developed to help the researcher choose an appropriate model for a given set of data. Pruzansky et al. (1982) defined statistics—the skewness of the distribution of distances and the percent violations of the ultrametric inequality—that helped to distinguish spatial and tree data in their simulation. Tversky and Hutchinson (1986) used nearest neighbor analysis to develop centrality (C) and reciprocity (R) statistics. Large values of C and R are diagnostic of an underlying tree structure. Finally, testing for violations of the triangle inequality and finding the number of necessary dimensions (from the eigenvalues of the scalar products matrix) can be used to evaluate the appropriateness of the geometric spatial model (Young & Householder, 1938).

TRIAGE is a combination of three macros which take as input a symmetric matrix of dissimilarities, and then output various diagnostic statistics. Macro NEAREST performs a nearest neighbor analysis on the data. Output includes the overall C and R statistics, each entity's nearest neighbor and centrality, and the reciprocity of each entity's nearest neighbor. Macro SKEWLONG produces skewness, elongation, and triangle inequality statistics. Macro EVAL produces eigenvalues of the scalar products matrix corresponding to the dissimilarities. Documentation of the program includes descriptions of the diagnostics and guidelines for their interpretation.

Availability

TRIAGE requires a computer system containing both the PROC MATRIX and MACRO languages. The program will be made compatible with the SAS Interactive Matrix Language (IML) upon release of IML documentation. A documented listing of the program, including a working example and sample output, is available without charge from Joseph Lee Rodgers, Department of Psychology, University of Oklahoma, Norman OK 73019, U.S.A.

References

Pruzansky, S., Tversky, A., & Carroll, J. D. (1982). Spatial versus tree representations of proximity data. *Psychometrika, 47,* 3–24.

Tversky, A., & Hutchinson, J. W. (1986). Nearest neighbor analysis of psychological spaces. *Psychological Review, 93,* 3–22.

Young, G., & Householder, A. S. (1938). Discussion of a set of points in terms of their mutual distances. *Psychometrika, 6,* 331–333.

Downloaded from apm.sagepub.com at UNIV OF OKLAHOMA LIBRARIES on January 20, 2016