NAME

crosstab − contingency table (or crosstabulation) analysis

SYNOPSIS

crosstab [<] *contingency_table* > *stderr*

DESCRIPTION

crosstab calculates measures of 1) significance of association and 2) strength of association for a contingency table. It reads either an ASCII file or an unsigned long (PFUINT) HIPS file containing the *contingency_table*. If read from *stdin* only ASCII files are allowed. The first two numbers in the ASCII file are number of rows *nr* and number of columns *nc*. The ASCII file is read row-wise.

The significance of association is measured by a chi-square test for no association between rows and columns. The corresponding significance level gives the probability of finding a larger chi-square (small values indicating a significant association). Also Cramer's V, the contingency coefficient C, and the phi coefficient are calculated. If the table is quadratic the diagonal sum and kappa (with t-test of significance) are calculated.

The strength of association (if significant) is measured by entropy measures, specifically the uncertainty of the rows given the columns, U(r|c), *vice versa*, U(c|r), and a symmetrical uncertainty measure, U(r,c).

REFERENCES

J.B. Campbell (1996). *Introduction to Remote Sensing.* Second Edition. Taylor and Francis, London.

K. Conradsen (1984). *En Introduktion til Statistik 1B.* 4. udgave. (In Danish.)

J.W. Dixon et al. (1985). *BMDP Manual.*

W.H. Press, S.A. Teukolsky, W.T. Vetterling and B.P. Flannery (1997). *Numerical Recipes in C: The Art of Scientific Computing.* Second Edition. Cambridge University Press.

AUTHOR

Allan Aasbjerg Nielsen, M.Sc., Ph.D.
IMM, Department of Mathematical Modelling
Technical University of Denmark, Building 321
E-mail aa@imm.dtu.dk, Internet http://www.imm.dtu.dk/˜aa