BULL. AUSTRAL. MATH. SOC. 62F15, 62H12, 62H17, 62K99
VOL. 29 (1984), 137-138.

A UNIFIED THEORY OF CAUSAL MODELS

HARRI TAPIO KIIVERI

Directed graphs and conditional independence ideas are used to define a class of causal models applicable to a finite set of random variables. Given the directed graph for a particular model, a factorization of the joint distribution of the random variables can be written down and rules given for reading the conditional independence relations amongst the variables. The definition does not depend on any particular distributional form and hence can be applied to models with both discrete and continuous random variables. However, the most tractable examples are those for which the joint distribution is normal, or a product of multinomials, and these are considered in detail.

In the normal case connections are demonstrated with covariance selection, structural equation models, and the analysis of covariance structures as considered by Jöreskog [4], McDonald [5] and others. Topics covered include variance components, path analysis, factor analysis and regression with errors in variables.

Discrete counterparts of the recursive models in the normal case are obvious from the general definition and these are shown to include models considered by Goodman [2] and Haberman [3]. Examples covered include latent structure analysis and contingency tables with ordered variables.

Models containing unobserved or latent variables are treated using theory for incomplete data. This results in a more compact parameterization

Received 2 November 1983. Thesis submitted to University of Western Australia, December 1982. Degree approved October 1983. Supervisor: Dr T.P. Speed.

https://doi.org/10.1017/S0004972700021353 Published online by Cambridge University Press

of many of the causal models used in the social sciences and simpler expressions for the first and second derivatives of the log likelihood function.

For solving the likelihood equations in the normal and discrete cases, algorithms such as Newton-Raphson, quasi-Newton and the $E - M$ algorithm (formalized by Dempster *et al* [*1*]) are discussed. In each case it is shown how a single program can be written to handle the class of models considered, with or without unobserved variables, using the quasi-Newton algorithm. Cases when the equations can be explicitly solved are also presented.

Identifiability of the parameters in any particular model is considered along with ways of checking the positive definiteness of the information matrix, both theoretically and numerically.

References

[*1*] Arthur P. Dempster, N.M. Laird and Donald B. Rubin, "Maximum likelihood from incomplete data via the *EM* algorithm", *J. Roy. Statist. Soc. Ser. B* 39 (1977), 1-38.

[*2*] L.A. Goodman, "Analysising qualitative/categorical data", *Log-linear models and latent structure analysis*, - (Abt Books, Cambridge, Massachusetts, 1978).

[*3*] Shelby J. Haberman, "Product models for frequency tables involving indirect observation", *Ann. Statist.* 5 (1977), 1124-1147.

[*4*] Karl G. Jöreskog, "Structural equation models in the social sciences: specification estimation and testing, *Applications of statistics*, 265-287 (Proc. Sympos. Wright State University, Dayton, Ohio, 1976. North-Holland, Amsterdam, 1977).

[*5*] R.P. McDonald, "A simple comprehensive model for the analysis of covariance structures", *British J. Math. Statist. Psych.* 31 (1978), 59-72.

CSIRO, Division of Mathematics and Statistics,
Private Bag, PO,
Wembley,
Western Australia 6104.