

not only when these are observational. We agree that the graphs suggested by Glymour and Spirtes could possibly be chosen as another description of our nondecomposable models but we do not regard them as indicating useful potential processes to generate the data, the point of our distinction.

In a recent paper, Stone (1993) elucidates requirements for particular causal interpretations. He also examines critically strongly ignorable treatment allocation. Pearl in his contribution gives an important graphical interpretation exactly of this assumption, this facilitating the judgement of the effects of interventions in a hypothesized causal process.

Several contributors mention the role of latent variables, including as a special case the occurrence of measuring errors. We agree that their use, preferably sparingly, especially in elucidating nondecomposable models, needs further study. For instance, the tetrad conditions studied by Spirtes, Glymour and Scheines (1993) for linear relations become relevant as well for binary variables having a quadratic exponential distribution. This distribution has some of the properties of the multivariate normal distribution and provides exact or approximate answers to Hill's question about graphical theory for binary distributions and to Whittaker's comments on complete independence.

Dempster favours shrinking estimates toward zero as opposed to setting parameters exactly to zero. We agree when empirical prediction is the objective, but not where essentially qualitative understanding via simple representations is involved, and the latter is our main concern.

The issue, raised by Whittaker, of labelling the edges of a graph can be solved in various ways if a single degree of freedom is attached to each edge (by partial correlation coefficients or by standardized regression coefficients, for instance). The introduction of graphs with dashed edges has, however, a different objective, because it leads to structures of independence different from those discussed by Whittaker, thus enriching the class of graphical chain models, as pointed out by Hill. Whittaker's graphs (ai) and (aii) do not represent the multivariate regression of our Figure 1c because the essential association between the two responses is omitted.

Rejoinder

David J. Spiegelhalter, A. Philip Dawid, Steffen L. Lauritzen and Robert G. Cowell

We are grateful to the discussants for their thoughtful comments: since our paper is already quite long enough we shall try to restrict our responses. We shall first deal

Whittaker points out the relation of the Schur complement to partial correlations and inverse covariance matrices. An early treatment of this in the statistical literature is by Cramér (1946, subsections 22.7, 23.4 and 23.5). The connection between partial correlation and canonical parameters in the exponential family has opened the road to defining analogous independence structures for discrete variables and for mixed discrete and continuous variables, known now as block regression (full edge) chain models.

In general distributional assumptions are necessary, in addition to the independence graph, for a full specification of a statistical model. Indeed some research hypotheses may not be possible for a particular joint distribution of specified form. For example, $X \perp\!\!\!\perp Y|A$ cannot hold without additional independences if the joint distribution is given by the linear logistic regression of the binary variable A on the bivariate normal variable (X, Y) . Similarly if (X, Y) are conditionally bivariate normal given the discrete variable A , then marginal independence of X and Y is possible only with additional independences. See Cox and Wermuth (1992b) for further details.

We were glad to see that Sobel regards our introduction of multivariate regression (dashed edge) chain graphs as a step toward more traditional analyses in the social sciences. In fact, it was one of our purposes to provide simple examples which help one to recognize similarities and distinctions between different approaches, the latter being explicitly appreciated by both Sobel and Dempster.

Because of the particular focus of our paper, we have put little emphasis on such issues as description of sample selection, checking data quality, testing model adequacy, examining the need of data transformation and comparison of the fits of different kinds of models. All of these are a normal if often difficult part of applied statistical work. From our present perspective, whether the formal aspects to the analysis are in frequentist or Bayesian terms is a secondary issue.

A special topic for further work concerns the role of graphs with both kinds of edge, for example, in representing the regression for multivariate binary data studied by Zhao and Prentice (1990) and by Fitzmaurice and Laird (1993).

with representations of causality, followed by some technical points on zero probabilities. Automatic model construction will then be considered, and whether a