Series Editor's Introduction

Latent structure models have been popular in the social and behavioral sciences for a very long time. They come in many disguises. In psychology, the major example is the Spearman model, in which intelligence is the latent variable. In sociology, Lazarsfeld contributed latent class analysis to express the idea that observed variables are jointly determined by underlying constructs that are not directly measured. In education Lawley, Rasch, Lord and others constructed the elaborate structure called "item response theory" or "latent trait theory," which tries to measure unidimensional ability in an indirect way. Taken together, I think these models and the resulting statistical techniques are the most important, and the most interesting, contribution of the social and behavioral sciences to data analysis.

Statisticians were initially very critical of these models. This was caused, to some extent, by grandiose claims from psychologists such as Cattell. Another reason was that many of the founders of latent structure analysis did not speak the lingo of mathematical statistics sufficiently well, which was reason enough to exclude them from the club of respectable scientists. It is true, there have always been statisticians interested in latent structure models, and as a rule, they came up with important contributions. We mention Lawley, Anderson, Rao, and Rubin. But only

recently have statisticians realized that latent variable models are at the core of much of modern statistics. We mention Bayes and empirical Bayes techniques, the EM algorithm, imputation methods, hidden Markov models, and state space models. In fact, the random variables that define the disturbances in the linear models are unobservable, or latent. More generally, we have come to realize that conceptualizing in terms of latent variables is a particular way of looking at the relationships between observables. Factor analysis, for example, can be formulated as a structural model for observed covariances, without even mentioning latent variables. They are useful, maybe even necessary, for the interpretation of the results.

Ton Heinen's book discusses latent variable models for discrete observed variables, concentrating on latent class and latent trait analysis. It does this mainly from the sociological perspective, and it shows that using latent class analysis as the starting point makes it possible to discuss a large and coherent class of techniques. Psychologists, no doubt, would start with continuous latent variable models such as factor analysis and latent trait analysis. As Heinen clearly shows, the sociological approach, expertly developed by Lazarsfeld, Guttman, Goodman, Clogg, and Hagenaars, is an equally natural way of organizing latent variable techniques. With considerable detail and precision it is shown that small modifications of the basic model can deal with discrete nominal, ordinal, and numerical observed variables.

A final point. It was Cliff Clogg who suggested that we should take a look at Heinen's thesis for our series. Cliff also offered to be the reviewer for this book. His endorsement is the strongest recommendation I can think of.

JAN DE LEEUW