Series Editor's Introduction

There are two essentially different approaches to analyzing association in contingency tables. Both have their origins in the early history of statistics. The first is Yule's axiomatic theory to define coefficients of association. The second is Pearson's attempt to extend the theory of correlation to contingency tables, using tetrachoric and polychoric correlation coefficients.

A great deal has been written about the fierce controversy between the two approaches. For the fascinating details we refer to Chapter 7 of the book by D. A. Mackenzie (1981). For our purposes it suffices to observe that the logical, axiomatic approach of Yule became the dominating method for contingency table analysis in sociometrics and other areas such as biometrics, while the correlational approach of Pearson continued to rule the waves in psychometrics and other areas such as educational research. The reasons are complicated, but they certainly have to do with the philosophical and methodological perspectives of the disciplines involved.

Of course both classes of methods have been developed enormously since the early contributions of Pearson and Yule. The coefficients of association, summarized in the definitive book by L. A. Goodman and W. H. Kruskal (1979), naturally lead to the more complete descriptions offered by log-linear analysis. And Pearson's tetrachoric theory was extended to test theory and factor analysis. The book by D. J. Bartholomew (1987) is a good recent summary.

In the last 20 years, however, there have been several attempts to arrive at a unifying framework, a Pearson-Yule compromise. The increasing popularity of correspondence analysis is one tendency in this direction, but more directly, the in-depth study of contingency tables with ordered categories unifies the logical and quantitative aspects of the analysis.

It makes sense to interpret the Clogg and Shihadeh book we introduce here in this light. Obviously, it comes from the sociological tradition. It incorporates ordinal and even numerical information into the classical log-linear analysis of multidimensional contingency tables. The book builds on methods introduced by Goodman, Haberman, Fienberg, and Clogg, and it presents them in a unifying framework, at a leisurely pace. Thus it shows how far one can go in the direction of correspondence analysis, the bivariate normal distribution, and latent variables from the classical Yulean point of view.

There are many forms of input-output analysis for which these models are appropriate, and applications could easily be collected from traffic research, tourism, import-export, stimulus recognition research, and citation analysis. The scope of these methods for bivariate tables is even wider than the book suggests.

Clogg and Shihadeh add a last, very useful, chapter on logistic regression, and this chapter clearly deals with data sets of a different structure, and it addresses different types of questions.

As for extensions to the multivariate situation that survey researchers and psychological or educational testers have to deal with, these are also available in various forms. The latent trait models, such as the Rasch model and its extensions, have strong connections with log-linear analysis, and the multivariate probit models with latent variables of psychometrics can be thought of as minor variations of corresponding logit models. In this sense, the Clogg and Shihadeh book can be thought of as a discussion of some of the basic building blocks that are needed for the analysis of multivariate tables. The basic idea of looking at a multivariate distribution by parametrizing the bivariate marginals is one of the key components of the Pearson approach, common to multiple correspondence analysis and to multivariate probit models. The models in this book for bivariate and multivariate tables and for regression analysis should be very useful for social research.

JAN DE LEEUW