

Fienberg S E. *The analysis of cross-classified categorical data*.
Cambridge, MA: MIT Press, (1977) 1980. 198 p.

[Department of Applied Statistics, University of Minnesota, St. Paul, MN]

This book presents an introduction to the analysis of categorical data using loglinear and logistic models, especially in the context of multivariate cross-classifications. It guides the reader from an understanding of the basic techniques for two-way tables up through research topics such as the link of the quasi-independence model with the literature on paired comparisons. [The *SSCI*® and *SCI*® indicate that this book has been cited in over 570 publications.]

Loglinear Models for Contingency Table Analysis

Stephen E. Fienberg
Departments of Statistics and
Social and Decision Sciences
Carnegie Mellon University
Pittsburgh, PA 15213-3890

November 25, 1988

In 1972 I was working on a systematic treatment of loglinear models for the analysis of categorical data with my collaborators Yvonne M.M. Bishop, Paul W. Holland, and Frederick Mosteller, subsequently published as *Discrete Multivariate Analysis*¹ (henceforth *DMA*—also a *Citation Classic*). I had begun to realize that *DMA* had too many technical details for most researchers and students who wanted to use the methodology rather than learn about the underlying theory. Thus, I prepared a set of lecture notes for our master's-level categorical data analysis course at the University of Minnesota. These notes were further developed for use in a series of short courses, and the positive response to them stimulated me to publish an expanded version as a separate book.

Several features distinguished the new book from *DMA*. Virtually all of the examples and data sets were new, and I summarized the basic theory in a four-page appendix rather than

in a detailed presentation, as in *DMA*. The new book emphasized the distinction between explanatory and response variables and how this leads to the conversion from loglinear to logit models, as well as a discussion of logistic regression. It also included several approaches and results that appeared in the literature after the work on *DMA* was complete, e.g., the use of the Darroch-Ratcliff extension to the iterative proportional fitting procedure that lies at the heart of the computational procedures used in both books.

As I was completing the manuscript on the first edition in 1976, I received a telephone call from the woman assigned to prepare the book jacket. She asked for a copy of the opening chapters and then set to work on her design. When I first saw the cover, I didn't understand the design, but then she gave me an explanation of how it linked to the basic structure of multivariate data! Not only did the book become widely used and cited, but the jacket won a special award.

As I worked on both editions, I was fortunate to have students, colleagues, and collaborators come to me with categorical data problems, the solutions for which ultimately worked their way into the book. For example, William Mason (from the University of Michigan) originally raised the issue that led to the discussion of continuation ratios for polytomous response variables in chapter 6 and some lecture notes by Holland stimulated me to include the material on retrospective epidemiological studies in chapter 7. I often included new ideas in the middle of examples in an attempt to make them more than simply numerical exercises. As a consequence, the book became an important research source for others and not simply a low-level exposition of ideas in *DMA*.

Major advances in the analysis of cross-classified categorical data since the publication of the second edition of the book include the development of association models for ordinal variables² and the use of families of goodness-of-fit statistics.³

1. Bishop Y M M, Fienberg S E & Holland P W (with the collaboration of Light R J & Mosteller F). *Discrete multivariate analysis: theory and practice*. Cambridge, MA: MIT Press, 1975. 557 p. (Cited 1,970 times.) [See also: Fienberg S E. *Citation Classic, Current Contents/Social & Behavioral Sciences* 21(14):16, 3 April 1989.]
2. Goodman L A. *The analysis of cross-classified data having ordered categories*. Cambridge, MA: Harvard University Press, 1984. 414 p. (Cited 25 times.)
3. Read T R C & Cressie N A C. *Goodness-of-fit statistics for discrete multivariate data*. New York: Springer-Verlag, 1988. 211 p.

1A-14