

how the practice of ignoring missing data led to incorrect inferences in important substantive arenas. Equally important, the work by the political methodologists promises to provide a practical solution that can be used by almost any political scientist with missing data. Political methodology is happy to import methods (and scholars) from other disciplines, but those ideas must be shown to solve important problems in political science.

**REFERENCES**

Achen, C. (1983), "Towards Theories of Data," in *Political Science: The State of the Discipline*, ed. A. Finifter, Washington, DC: American Political Science Association, pp. 69-93.  
 ——— (1985), "Editorial," *Political Methodology*, 11.  
 Bartels, L., and Brady, H. (1993), "The State of Quantitative Political Methodology," in *Political Science: The State of the Discipline II*, ed. A. Finifter, Washington, DC: American Political Science Association, pp. 121-159.  
 Beck, N., and Jackman, S. (1998), "Beyond Linearity by Default: Generalized Additive Models," *American Journal of Political Science*, 42, 596-627.  
 Beck, N., and Katz, J. N. (1995), "What to do (and not to do) With Times Series Cross-Section Data," *American Political Science Review*, 89, 634-647.  
 Beck, N., Katz, J. N., and Tucker, R. (1998), "Taking Time Seriously: Time Series-Cross-Section Analysis With a Binary Dependent Variable," *American Journal of Political Science*, 42, 1260-1288.  
 Beck, N., King, G. M., and Zeng, L. (2000), "Improving Quantitative Studies of International Conflict: A Conjecture," *American Political Science Review*, 94, 21-35.  
 Dahl, R. (1961), "The Behavioral Approach in Political Science: Epitaph for a Monument to a Successful Protest," *American Political Science Review*, 55, 763-772.  
 Geddes, B. (1990), "How the Cases You Choose Affect the Answers You Get: Selection Bias in Comparative Politics," *Political Analysis*, 2, 131-150.

Goodman, L. (1953), "Ecological Regressions and the Behavior of Individuals," *American Sociological Review*, 18, 663-666.  
 Gow, D. J. (1985), "Quantification and Statistics in the Early Years of American Political Science, 1880-1922," *Political Methodology*, 11, 1-18.  
 Hastie, T. J., and Tibshirani, R. J. (1990), *Generalized Additive Models*, New York: Chapman and Hall.  
 Herron, M. C. (2000), "Post-Estimation Uncertainty in Limited Dependent Variable Models," *Political Analysis*, 8, 83-98.  
 Jackman, S. (2000), "Estimation and Inference via Bayesian Simulation: An Introduction to Markov Chain Monte Carlo," *American Journal of Political Science*, 44, 375-404.  
 Johnston, R., Blais, A., Brady, H. E., and Crete, J. (1992), *Letting the People Decide: Dynamics of a Canadian Election*, Montreal: McGill-Queen's University Press.  
 King, G. (1998), *A Solution to the Ecological Inference Problem*, Princeton, NJ: Princeton University Press.  
 King, G., Honaker, J., Joseph, A., and Scheve, K. (1998), "Analyzing Incomplete Political Science Data: An Alternative Algorithm for Multiple Imputation," unpublished paper presented at the Annual Meetings of the American Political Science Association, Boston.  
 King, G., Keohane, R. O., and Verba, S. (1994), *Designing Social Inquiry*, Princeton, NJ: Princeton University Press.  
 King, G., Tomz, M., and Wittenberg, J. (2000), "Making the Most of Statistical Analyses: Improving Interpretation and Presentation," *American Journal of Political Science*, 44, 341-355.  
 Little, R. J. A., and Rubin, D. B. (1987), *Statistical Analysis With Missing Data*, New York: Wiley.  
 Petty, W. (1690), *Political Arithmetick*, London: Robert Clavel and Hen. Mortlock.  
 Robinson, W. S. (1950), "Ecological Correlation and the Behavior of Individuals," *American Sociological Review*, 15, 351-357.  
 Rubin, D. R. (1981), "Estimation in Parallel Randomized Experiments," *Journal of Educational Statistics*, 6, 377-400.  
 Sniderman, P. M., Brody, R. A., and Tetlock, P. E. (1991), *Reasoning and Choice: Explorations in Political Psychology*, New York: Cambridge University Press.  
 Western, B. (1998), "Causal Heterogeneity in Comparative Research: A Bayesian Hierarchical Modelling Approach," *American Journal of Political Science*, 42, 1233-1259.

**Statistics in Sociology, 1950-2000**

**Adrian E. RAFTERY**

Sociology is the scientific study of modern industrial society. Example questions include: What determines how well people succeed in life, occupationally and otherwise? What factors affect variations in crime rates between different countries, cities, and neighborhoods? What are the causes of the increasing U.S. divorce rate? What are the main factors driving fertility decline in developing countries? Why have social revolutions been successful in some countries but not in others?

Adrian E. Raftery is Professor of Statistics and Sociology, Department of Statistics, University of Washington, Seattle, WA 98195 (E-mail: [raftery@stat.washington.edu](mailto:raftery@stat.washington.edu)). The author is very grateful to Andrew Abbott, Mark Becker, Timothy Biblarz, Kenneth Bollen, Leo Goodman, Mark Handcock, Charles Hirschman, Robert Hauser, Daphne Kuo, Paul LePore, David Madigan, Peter Marsden, Stephen Matthews, Martina Morris, Michael Seltzer, Burton Singer, Michael Sobel, Elizabeth Thompson, Stanley Wasserman, Rob Warren, and Kazuo Yamaguchi for extremely helpful comments. The fact that so many colleagues helped testifies to the vibrancy of the sociological methodology community whose work is the core of this discussion. The views expressed remain resolutely personal.

The roots of sociology go back to the mid-nineteenth century and to seminal work by Auguste Comte, Karl Marx, Max Weber, and Emile Durkheim on the kind of society newly emerging from the industrial revolution. Sociology has used quantitative methods and data from the beginning, but before World War II the data tended to be fragmentary and the statistical methods simple and descriptive.

Since then, the available data have grown in complexity, and statistical methods have been developed to deal with them, with the sociologists themselves often leading the way (Clogg 1992). The trend has been toward increasingly rigorous formulation of hypotheses, larger and more detailed datasets, more complex statistical models to match the data, and a higher level of statistical analysis in the major sociological journals.

Statistical methods have had a successful half-century in sociology, contributing to a greatly improved standard of scientific rigor in the discipline. Sociology has made use of a wide variety of statistical methods and models, but I focus here on the ones developed by sociologists, motivated by sociological problems, or first published in sociological journals. I distinguish three postwar generations of statistical methods in sociology, each defined by the kind of data it addresses. The first generation of methods, starting after World War II, deals with cross-tabulations of counts from surveys and censuses by a small number of discrete variables such as sex, age group, and occupational category; social mobility tables provide a canonical example. Schuessler (1980) gave a survey that largely reflects this first-generation work.

The second generation, starting in the early 1960s, deals with unit-level data from surveys that include many variables. This generation was galvanized by Blau and Duncan's (1967) highly influential book *The American Occupational Structure*, and by the establishment of *Sociological Methodology* in 1969 and *Sociological Methods and Research* in 1972 as publication outlets. These developments marked the coming of age of research on quantitative methodology in sociology. The third generation of methods, starting in the late 1980s, deals with data that are not usually thought of as cross-tabulations or data matrices, either because the data take different forms, such as texts or narratives, or because dependence is a crucial aspect. These generations do not have clear starting points and all remain active today; like real generations, they overlap.

Today, much sociological research is based on the reanalysis of large high-quality survey sample datasets, usually collected with public funds and publicly available to researchers, with typical sample sizes in the range of 5,000–20,000. This has opened the way to easy replication of results and has helped produce standards of scientific rigor in sociology comparable to those in many of the natural sciences. Social statistics is expanding rapidly as a research area, and several major institutions have recently launched initiatives in this area.

## 1. THE FIRST GENERATION: CROSS-TABULATIONS

### 1.1 Categorical Data Analysis

Initially, much of the data that quantitative sociologists had to work with came in the form of cross-classified ta-

bles, and so it is not surprising that this is perhaps the area of statistics to which sociology has contributed the most. A canonical example has been the analysis of social mobility tables, two-way tables of father's against respondent's occupational category; typically the number of categories used is between 5 and 17.

At first, the focus was on measures of association, or mobility indices as they were called in the social mobility context (Glass 1954; Rogoff 1953), but these indices failed to do the job of separating structural mobility from exchange (or circulation) mobility. It was Birch (1963) who proposed the log-linear model for the observed counts  $\{x_{ij}\}$ , given by

$$\log(E[x_{ij}]) = u + u_{1(i)} + u_{2(j)} + u_{12(ij)}, \quad (1)$$

where  $i$  indexes rows and  $j$  indexes columns,  $u_{1(i)}$  and  $u_{2(j)}$  are the main effects for the rows and columns, and  $u_{12(ij)}$  is the interaction term, measuring departures from independence. The difficulty with (1) for social mobility and similar tables is that the number of parameters is too large for inference and interpretation; for example, in the U.S. datasets 17 categories were used, so the interaction term involves  $16^2 = 256$  parameters.

A successful general approach to modeling the interaction term parsimoniously is the association model of Duncan (1979) and Goodman (1979),

$$u_{12(ij)} = \sum_{k=1}^K \gamma_k \alpha_i^{(k)} \beta_j^{(k)} + \phi_i \delta(i, j), \quad (2)$$

where  $\delta(i, j) = 1$  if  $i = j$  and 0 otherwise. In (2),  $\alpha_i^{(k)}$  is the score for the  $i$ th row on the  $k$ th scoring dimension, and  $\beta_j^{(k)}$  is the corresponding score for the  $j$ th column; these can be either specified in advance or estimated from the data. The last term allows a different strength of association on the diagonal. [Model (2) is unidentified as written; various identifying constraints are possible.] In most applications to date,  $K = 1$ . Goodman (1979) initially derived this model as a way of describing association in terms of local odds ratios. He (Goodman 1985) later showed that this model is closely related to canonical correlations and to correspondence analysis (Benzécri 1976), and provided an inferential framework for these methodologies. Table 1 shows the actual counts for a reduced version of the most extensive U.S. social mobility study and the fitted values from an association model; the model accounts for

Table 1. Observed Counts From the Largest U.S. Social Mobility Study and Expected Values From a Goodman Association Model With 4 Degrees of Freedom, Sample Size 19,912

Father's occupation	Son's occupation									
	Upper nonmanual		Lower nonmanual		Upper manual		Lower manual		Farm	
	Observed	Expected	Observed	Expected	Observed	Expected	Observed	Expected	Observed	Expected
Upper nonmanual	1,414	1,414	521	534	302	278	643	652	40	42
Lower nonmanual	724	716	524	524	254	272	703	698	48	43
Upper manual	798	790	648	662	856	856	1,676	1,666	108	112
Lower manual	756	794	914	835	771	813	3,325	3,325	237	236
Farm	409	386	357	409	441	405	1,611	1,617	1,832	1,832

NOTE: Data from Hout (1983).

99.6% of the association in the table, and its success is evident.

Hout (1984) extended the range of application of these models by modeling the scores and diagonal terms in (2) as sums or products of covariates, such as characteristics of the occupational categories in question; this is an extension of Birch's (1965) linear-by-linear interaction model. This has led to important discoveries, including Hout's (1988) finding that social mobility is on the increase in the United States. Biblarz and Raftery (1993) adapted Hout's models to higher-dimensional tables to study social mobility in nonintact families, finding that occupational resemblance is weaker there than in intact families. From sociology, these ideas have diffused to other disciplines, such as epidemiology (Becker 1989).

An appealing alternative formulation of the basic ideas underlying (1) and (2) is in terms of marginal *distributions* rather than the main effects in (1). The resulting marginal models specify a model for the marginal distributions and a model for the odds ratios, and this implies a model for the joint distribution that is not log-linear (Becker 1994; Becker and Yang 1998; Lang and Agresti 1994).

An alternative approach that answers different questions is the latent class model (Goodman 1974; Lazarsfeld 1950). This represents the distribution of counts as a finite mixture of distributions in each of which the different variables are independent. An interesting recent application to criminology was described by Roeder, Lynch, and Nagin (1999).

## 1.2 Hypothesis Testing and Model Selection

Sociologists often have sample sizes in the thousands, and so they came up early and hard against the problem that standard  $p$  values can indicate rejection of null hypotheses in large samples, even when the null model seems reasonable theoretically and inspection of the data fails to reveal any striking discrepancies with it. The problem is compounded by the fact that there are often many models rather than just the two envisaged by significance tests, and by the need to use stepwise or other multiple-comparison methods for model selection (e.g., Goodman 1971). By the early 1980s, some sociologists were dealing with this problem by ignoring the results of  $p$  value-based tests when they seemed counterintuitive and by basing model selection instead on theoretical considerations and informal assessment of discrepancies between model and data (e.g., Fienberg and Mason 1979; Grusky and Hauser 1984; Hout 1983, 1984).

Then it was pointed out that this problem could be alleviated by instead basing model selection on Bayes factors (Raftery 1986), and that this can be simply approximated for log-linear models by preferring a model if the Bayes information criterion (BIC) = deviance - (degrees of freedom)  $\log(n)$ , is smaller (Schwarz 1978). For nested hypotheses, this can be viewed as defining a significance level for a test that decreases automatically with sample size. Since then, this approach has been used in many sociological applications of log-linear models. Kass and Wasserman (1995) showed that the approximation is

quite accurate if the Bayesian prior used for the model parameters is a unit information prior, and Raftery (1995) indicated how the methodology can be extended to a range of other models. Weakliem (1999) criticized the use of BIC on the grounds that the unit information prior to which it corresponds may be too diffuse in practice. This points toward using Bayes factors based on priors that reflect the actual information available; this is easy to do for log-linear and other generalized linear models (Raftery 1996).

## 2. THE SECOND GENERATION: UNIT-LEVEL SURVEY DATA

The second generation of statistical models responded to the availability of unit-level survey data in the form of large data matrices of independent cases. The methods that have proved successful for answering questions about such data have mainly been based on the linear regression model and its extensions to path models, structural equation models, generalized linear models, and event history models. For questions about the *distribution* of variables rather than their predicted value, however, nonparametric methods have proven useful (Handcock and Morris 1998; Morris, Bernhardt, and Handcock 1994).

### 2.1 Measuring Occupational Status

Occupational status is an important concept in sociology, and developing a useful continuous measure of it was a signal achievement of the field. Initially, the status of an occupation was equated with its perceived prestige, as measured in surveys. However, surveys could measure the prestige of only a small number of the 800 or so occupations identified in the Census. To fill in the missing prestige scores, Duncan (1961) regressed the prestige scores for the occupations for which they were available on measures of the average education and average income of incumbents of the occupation. He found that the predictions were very good ( $R^2 = .91$ ), and that the two predictors were about equally weighted. Based on this, he created a predicted prestige score for all occupations, which became known as the Duncan socioeconomic index (SEI). The SEI later turned out to be a better predictor of various social outcomes than the prestige scores themselves. Duncan's initial work has been updated several times (Hauser and Warren 1997).

In much social research, particularly in economics, current income is used as a predictor of social outcomes, but there are good reasons to prefer occupational status. It has proven to be a good predictor of many social outcomes. Jobs and occupations can be measured accurately, in contrast to income or wealth, whose measurement is plagued by problems of refusal, recall, and reliability. Also, occupational status is more stable over time than income, both within careers and between generations. This suggests that occupational status may actually be a better indicator of long-term or permanent income than current income itself. The status of occupations tends to be fairly constant both in time and across countries (Treiman 1977).

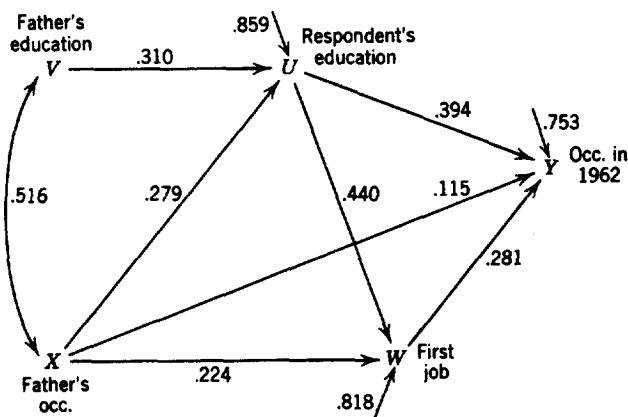


Figure 1. A Famous Path Model: The process of stratification, U.S. 1962 (Blau and Duncan 1967). The numbers on the arrows from one variable to another are regression coefficients, .516 is the correlation between V and X, and the numbers on the arrows with no sources are residual standard deviations. All the variables have been centered and scaled.

2.2 The Many Uses of Structural Equation Models

Figure 1 shows the basic path model of occupational attainment at the heart of the work of Blau and Duncan (1967) (see Duncan 1966). Wright (1921) introduced path analysis, and Blalock (1964) gave it a causal interpretation in a social science context. (See Freedman 1987 and Sobel 1998 for critique and discussion, and Abbott 1998, Pearl 1998, and Sobel 2000 for histories of causality in social science.)

Often, variables of interest in a causal model are not observed directly, but other variables are observed that can be viewed as measurements of the variables, or "constructs" of interest, such as prejudice, alienation, conservatism, self-esteem, discrimination, motivation, or ability. Jöreskog (1973) dealt with this by maximum likelihood estimation of a structural equation model with latent variables; this is sometimes called a LISREL model, from the name of Joreskog's software. Figure 2 shows a typical model of this kind; the goal of the analysis is testing and estimating the strength of the relationship between the unobserved latent variables represented by the thick arrow. Diagrams such as Figures 1 and 2 have proven useful to sociologists for specifying theories and hypotheses and for building causal models.

The LISREL framework has been extended and used ingeniously for purposes beyond those for which it was originally intended. Muthén (1983) extended it to categorical variables, and later (Muthén 1997) showed how it can be

used to represent longitudinal data, growth curve models, and multilevel data. Kuo and Hauser (1996) used data on siblings to control for unobserved family effects on socioeconomic outcomes, and cast the resulting random effects model in a LISREL framework. Warren, LePore, and Mare (2000) considered the relationship between the number of hours that high school students work and their grades; a common assumption might be that working many hours tends to depress grades. They found that although number of hours and grades do indeed tend to covary (negatively), the causal direction is the opposite: Low grades leads to many hours worked, rather than the other way round.

The advent of graphical Markov models (Spiegelhalter, Dawid, Lauritzen, and Cowell 1993), specified by conditional independencies rather than by regression-like relationships, is important for the analysis of multivariate dependencies, although they can seem less interpretable to sociologists. The relationship between these and structural equation models has begun to be understood (Spirtes et al. 1998). Also, the LISREL model seems ideally suited to Markov chain Monte Carlo (MCMC) methods (Gilks, Richardson, and Spiegelhalter 1996), and this is likely to permit useful extensions of the framework (Arminger 1998; Raftery 1991; Scheines, Hoijtink, and Boomsma 1999).

2.3 Event History Analysis

Unit-level survey data often include or allow the reconstruction of life histories. These include the times of crucial events such as marriages, divorces, births, committals to and releases from prison, job changes, and going on or off welfare.

The analysis of factors influencing the time to a single event such as death was revolutionized by the introduction of the Cox (1972) proportional hazards model. Tuma and Hannan (1984) generalized this approach to allow for repeated events, for multiple types of events, such as marriages and divorces, and for events consisting of movement between different types of states, such as different job categories.

Uses of the Cox model in medicine have tended to treat the baseline hazard nonparametrically, but in social science it has sometimes been found useful to model it parametrically. For example, Yamaguchi (1992) analyzed permanent employment in Japan where the surviving fraction (those who never change jobs) and its determinants are of key in-

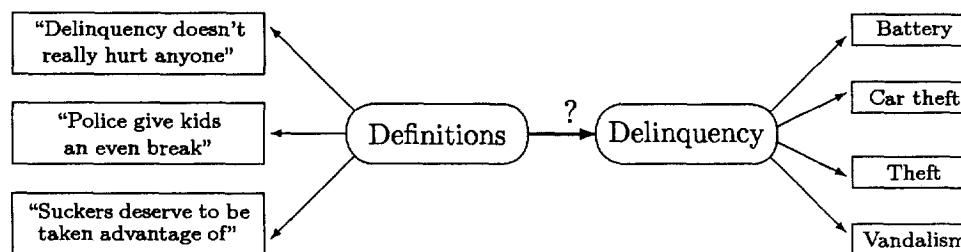


Figure 2. Part of a Structural Equation Model to Assess the Hypothesis That Learned Definitions of Delinquency Cause Delinquent Behavior (Matsueda and Heimer 1987). The key goal is testing and estimating the relationship represented by the thick arrow. The constructs of interest, "Definitions" and "Delinquency", are not measured directly. The variables inside the rectangles are measured.

terest; he found that covariates were associated both with the timing of job change and with the surviving fraction.

Social science event history data are often recorded in discrete time (e.g., by year), either because events tend to happen at particular times of year (e.g., graduating) or because of measurement constraints. As a result, discrete-time event history models have been popular (Allison 1982; Xie 1994), and in some ways these are easier to handle than their continuous-time analogs. Ways of dealing with multi-level event history data, smoothly time-varying covariates, and other complications have been introduced in this context (e.g., Fahrmeir and Knorr-Held 1997; Raftery, Lewis, and Aghajanian 1995).

One problem with social science event history data is that dropping out can be related to the event of interest. For example, people may tend to leave a study shortly before a divorce, which will play havoc with estimation of divorce rates. The problem seems almost insoluble at first sight, but Hill (1997) produced an elegant solution using the shared unmeasured risk factor (SURF) model of Hill, Axinn, and Thornton (1993). The basic trick is to observe that although one does not know which of the people who dropped out actually got divorced soon afterward, one can estimate which ones were most at risk of divorcing.

### 3. THE THIRD GENERATION: NEW DATA, NEW CHALLENGES, NEW METHODS

#### 3.1 Social Networks and Spatial Data

Social networks consist of sets of pairwise connections, such as friendships between adolescents, sexual relationships between adults, and political alliances and patterns of marriage between social groups. The analysis of data about such networks has a long history (Wasserman and Faust 1994). Frank and Strauss (1986) developed formal statistical models for such networks related to the Markov random field models used in Bayesian image analysis and derived using the Hammersley–Clifford theorem (Besag 1974). This has led to the promising “ $p^*$ ” class of models for social networks (Wasserman and Pattison 1996).

Methods for the analysis of social networks have focused mostly on small datasets with complete data. In practical applications, however, such as the effect of sexual network patterns on the spread of sexually transmitted diseases (Morris 1997), the datasets tend to be large and very incomplete, and current methods are somewhat at a loss. This is the stage that pedigree analysis in statistical genetics was at some years ago, but the use of likelihood and MCMC methods have led to major progress since then (Thompson 1998). Social networks are more complex than pedigrees in one way, because pedigrees tend to have a tree structure whereas social networks often have cycles, but progress does seem possible.

Most social data are spatial, but this fact has been largely ignored in sociological research. A major exception is Massey and Denton’s (1993) study of residential segregation by race, reviving a much older sociological tradition of spatial analysis in American society (e.g., Duncan and Duncan 1957). More recently, the field of research on fertility

and contraception in Asia (several major projects focused on China, Thailand, and Nepal) has been making fruitful use of satellite image and Geographic Information System (GIS) data (e.g., Entwisle, Rindfuss, Walsh, Evans, and Curran 1997). More extensive use of spatial statistics in sociology seems likely.

#### 3.2 Textual Data

In its rawest form, a great deal of sociological data is textual; for example, interviews, answers to open-ended questions in surveys, ethnographic accounts. How to analyze such data and draw inference from it is a largely open question. Efforts at formal analysis have focused on standard content analysis, consisting mainly of counting words in the text in different ways. It seems likely that using the context in which words and clauses appear would yield better results. Promising recent efforts to do just this include Carley’s (1993) map analysis, Franzosi’s (1994) set-theoretic approach, and Roberts’s (1997) generic semantic grammar, but the surface has only been scratched. The human mind is very good at analyzing individual texts, but computers are not (at least as yet); in this way the analysis of textual data may be like other problems such as image analysis and speech recognition. A similar challenge is faced on a massive scale by information retrieval for the Web (Jones and Willett 1997), where most search engines are based on simple content analysis methods. The more contextual methods being developed in sociology might be useful in this area as well.

Singer, Ryff, Carr, and Magee (1998) have made an intriguing use of textual data analysis, blending quantitative and qualitative approaches. They took a standard unit-level dataset with more than 250 variables per person and converted them into written “biographies.” They then examined the biographies for common features and thinned them to more generic descriptions.

#### 3.3 Narrative and Sequence Analysis

Life histories are typically analyzed by reducing them to variables and doing regression and multivariate analysis, or by event history analysis. Abbott and Hrycak (1990) argued that these standard approaches obscure vital aspects of a life history (such as a professional career) that emerge when it is considered as a whole. They proposed viewing life histories of this kind as analogous to DNA or protein sequences, using optimal alignment methods adapted from molecular biology (Sankoff and Kruskal 1983), followed by cluster analysis, to detect patterns common to groups of careers. Stovel, Savage, and Bearman (1996) used these methods to describe changes in career systems at Lloyds Banks over the past century.

Subsequently, Dijkstra and Taris (1995) extended the ideas to include independent variables, and Abbott and Bearman (1997) applied the Gibbs sampling sequence detection method of Lawrence et al. (1993), originally also developed for microbiology; this seems to work very well. The approach is interesting, and there are many open statistical questions.

### 3.4 Simulation Models

Another way to represent a social process in more detail is via a macrosimulation or microsimulation model. Such models are often deterministic and quite complicated, representing systems by different compartments that interact, and each compartment by a set of differential or difference equations. They have been used to, for example, explore the implications of different theories about how domestic politics and war interact (Hanneman, Collins, and Mordt 1995), the social dynamics of collective action (Kim and Bearman 1997), and the role of sexual networks in the spread of HIV (Morris 1997 and references therein).

A difficulty with such models is that ways of estimating the many parameters involved, of assessing the fit of the model, and of comparing competing models are not well established; all of this tends to be done by informal trial and error. Methods being developed to put inference for such models on a solid statistical footing in other disciplines may prove helpful in sociology as well (Guttrop and Walden 1987; Poole and Raftery 2000; Raftery, Givens, and Zeh 1995).

### 3.5 Macrosociology

Macrosociology deals with large entities, such as states and their interactions. As a result, the number of cases tends to be small, and the use of standard statistical methods such as regression is difficult. This was pointed out trenchantly by Ragin (1987) in an influential book. His own proposed alternative, qualitative comparative analysis, seems unsatisfactory, because it does not allow for variability of any kind and so is sensitive to small changes in the data and in the way the method is applied (Liebersohn 1994).

One solution to the problem is to obtain an at least moderately large sample size, as Bollen and Appold (1993) were able to do, for example. Often, however, this is not possible, so this is not a general solution. Another approach is to use standard regression-type models, but to do Bayesian estimation with strong prior information if available, which it often is from the practice, common in this area, of analyzing specific cases in great detail (Western and Jackman 1994). Bayes factors may also help, as they tend to be less stringent than standard significance tests in small samples and allow a calibrated assessment of evidence rather than forcing the rejection or acceptance of a hypothesis (Kass and Raftery 1995). They also provide a way of accounting for model uncertainty, which can be quite large in this context (Western 1996).

## 4. DISCUSSION

Statistical methodology has had a successful half-century in sociology, leading the way in providing models for cross-classifications and developing well-adapted methods for unit-level datasets. This has contributed to the greatly improved level of scientific rigor in sociology today.

New kinds of data and new challenges abound, and the area is ripe for statistical research. Several major institu-

tions are launching initiatives in the area. The University of Washington has just established a new Center for Statistics and the Social Sciences, UCLA's new Statistics Department grew out of social statistics, and there are other initiatives at the University of Michigan, Columbia University, UC Santa Barbara, and the universities in North Carolina's Research Triangle. Harvard's new Center for Basic Research in the Social Sciences also emphasizes social statistics. They all join the most successful effort of this kind to date, the Social Statistics Department at the University of Southampton.

## REFERENCES

- Abbott, A. (1998), "The Causal Devolution," *Sociological Methods and Research*, 27, 148-181.
- Abbott, A., and Barman, E. (1997), "Sequence Comparison via Alignment and Gibbs Sampling: A Formal Analysis of the Emergence of the Modern Sociological Article," *Sociological Methodology*, 27, 47-88.
- Abbott, A., and Hrycak, A. (1990), "Measuring Sequence Resemblance," *American Journal of Sociology*, 96, 144-185.
- Allison, P. (1982), "Discrete-Time Methods for the Analysis of Event Histories," *Sociological Methodology*, 13, 61-98.
- Arminger, G. (1998), "A Bayesian Approach to Nonlinear Latent Variable Models Using the Gibbs Sampler and the Metropolis-Hastings Algorithm," *Psychometrika*, 63, 271-300.
- Becker, M. P. (1989), "Using Association Models to Analyze Agreement Data: Two Examples," *Statistics in Medicine*, 8, 1199-1207.
- (1994), "Analysis of Cross-Classifications of Counts Using Models for Marginal Distributions: An Application to Trends in Attitudes on Legalized Abortion," *Sociological Methodology*, 24, 229-265.
- Becker, M. P., and Yang, I. (1998), "Latent Class Marginal Models for Cross-Classifications of Counts," *Sociological Methodology*, 28, 293-326.
- Besag, J. E. (1974), "Spatial Interaction and the Statistical Analysis of Lattice Systems" (with discussion), *Journal of the Royal Statistical Society, Ser. B*, 36, 192-236.
- Benzécri, J.-P. (1976), *L'Analyse des Données* (2nd ed.), Paris: Dunod.
- Biblarz, T. J., and Raftery, A. E. (1993), "The Effects of Family Disruption on Social Mobility," *American Sociological Review*, 58, 97-109.
- Birch, M. W. (1963), "Maximum Likelihood in Three-Way Tables," *Journal of the Royal Statistical Society, Ser. B*, 25, 220-233.
- (1965), "The Detection of Partial Association, II: The General Case," *Journal of the Royal Statistical Society, Ser. B*, 27, 111-124.
- Blalock, H. M. (1964), *Causal Inference in Non-Experimental Research*, New York: Harcourt, Brace.
- Blau, P. M., and Duncan, O. D. (1967), *The American Occupational Structure*, New York: Free Press.
- Bollen, K. A., and Appold, S. J. (1993), "National Industrial-Structure and the Global System," *American Sociological Review*, 58, 283-301.
- Carley, K. M. (1993), "Coding Choices for Textual Analysis: A Comparison of Content Analysis and Map Analysis," *Sociological Methodology*, 23, 75-126.
- Clogg, C. C. (1992), "The Impact of Sociological Methodology on Statistical Methodology" (with discussion), *Statistical Science*, 7, 183-207.
- Cox, D. R. (1972), "Regression Models and Life Tables" (with discussion), *Journal of the Royal Statistical Society, Ser. B*, 34, 187-220.
- Dijkstra, W., and Taris, T. (1995), "Measuring the Agreement Between Sequences," *Sociological Methods and Research*, 24, 214-231.
- Duncan, O. D. (1961), "A Socioeconomic Index for All Occupations," in *Occupations and Social Status*, ed. A. J. Reiss, New York: Free Press, pp. 109-138.
- (1966), "Path Analysis," *American Journal of Sociology*, 72, 1-16.
- (1979), "How Destination Depends on Origin in the Occupational Mobility Table," *American Journal of Sociology*, 84, 793-803.

- Duncan, O. D., and Duncan, B. (1957), *The Negro Population of Chicago*, Chicago: University of Chicago Press.
- Entwisle, B., Rindfuss, R. R., Walsh, S. J., Evans, T. P., and Curran, S. R. (1987), "Geographic Information Systems, Spatial Network Analysis, and Contraceptive Choice," *Demography*, 34, 171-187.
- Fahrmeir, L., and Knorr-Held, L. (1997), "Dynamic Discrete-Time Duration Models: Estimation via Markov Chain Monte Carlo," *Sociological Methodology*, 27, 417-452.
- Fienberg, S. E., and Mason, W. M. (1979), "Identification and Estimation of Age-Period-Cohort Effects in the Analysis of Discrete Archival Data," *Sociological Methodology*, 10, 1-67.
- Frank, O., and Stauss, D. (1986), "Markov Graphs," *Journal of the American Statistical Association*, 81, 832-842.
- Franzosi, R. (1994), "From Words to Numbers: A Set Theory Framework for the Collection, Organization and Analysis of Narrative Data," *Sociological Methodology*, 24, 105-136.
- Freedman, D. A. (1987), "As Others See Us" (with discussion), *Journal of Educational Statistics*, 12, 101-223.
- Gilks, W. R., Richardson, S., and Spiegelhalter, D. J. (1996), *Markov Chain Monte Carlo in Practice*, London: Chapman and Hall.
- Glass, D. V. (1954), *Social Mobility in Britain*, Glencoe, IL: Free Press.
- Goodman, L. A. (1971), "The Analysis of Multidimensional Contingency Tables: Stepwise Procedures and Direct Estimation Methods for Building Models for Multiple Classifications," *Technometrics*, 13, 33-61.
- (1974), "The Analysis of Systems of Qualitative Variables When Some of the Variables are Unobservable," *American Journal of Sociology*, 79, 1179-1259.
- (1979), "Simple Models for the Analysis of Association in Cross-Classifications Having Ordered Categories," *Journal of the American Statistical Association*, 74, 537-352.
- (1985), "The Analysis of Cross-Classified Data Having Ordered and/or Unordered Categories," *The Annals of Statistics*, 13, 10-69.
- Grusky, D. B., and Hauser, R. M. (1984), "Comparative Social Mobility Revisited: Models of Convergence and Divergence in Sixteen Countries," *American Sociological Review*, 49, 19-38.
- Guttorp, P., and Walden, A. T. (1987), "On the Evaluation of Geophysical Models," *Geophysical Journal of the Royal Astronomical Society*, 91, 201-210.
- Handcock, M. S., and Morris, M. (1998), "Relative Distribution Methods," *Sociological Methodology*, 28, 53-98.
- Hanneman, R. A., Collins, R., and Mordt, G. (1995), "Discovering Theory Dynamics by Computer Simulation: Experiments on State Legitimacy and Imperialist Capitalism," *Sociological Methodology*, 25, 1-46.
- Hauser, R. M., and Warren, J. R. (1997), "Socioeconomic Indexes for Occupations: A Review, Update and Critique," *Sociological Methodology*, 27, 177-298.
- Hill, D. H. (1997), "Adjusting for Attrition in Event-History Analysis," *Sociological Methodology*, 27, 393-416.
- Hill, D. H., Axinn, W. G., and Thornton, A. (1993), "Competing Hazards With Shared Unmeasured Risk Factors," *Sociological Methodology*, 23, 245-277.
- Hout, M. (1983), *Mobility Tables*, Beverly Hills, CA: Sage.
- (1984), "Status, Autonomy and Training in Occupational Mobility," *American Journal of Sociology*, 89, 1379-1409.
- (1988), "Expanding Universalism, Less Structural Mobility: The American Occupational Structure in the 1980s," *American Journal of Sociology*, 93, 1358-1400.
- Jones, K. S., and Willett, P. (1997), *Readings in Information Retrieval*, San Francisco: Morgan Kaufman.
- Jöreskog, K. G. (1973), "A General Method for Estimating a Linear Structural Equation System," in *Structural Equation Models in the Social Sciences*, ed. A. S. Goldberger and O. D. Duncan, New York: Seminar, pp. 85-112.
- Kass, R. E., and Raftery, A. E. (1995), "Bayes Factors," *Journal of the American Statistical Association*, 90, 773-795.
- Kass, R. E., and Wasserman, L. (1995), "A Reference Bayesian Test for Nested Hypotheses and Its Relationship to the Schwarz Criterion," *Journal of the American Statistical Association*, 90, 928-934.
- Kim, H., and Bearman, P. S. (1997), "The Structure and Dynamics of Movement Participation," *American Sociological Review*, 62, 70-93.
- Kuo, H. H. D., and Hauser, R. M. (1996), "Gender, Family Configuration, and the Effect of Family Background on Educational Attainment," *Social Biology*, 43, 98-131.
- Lang, J. B., and Agresti, A. (1994), "Simultaneously Modeling Joint and Marginal Distributions of Multivariate Categorical Responses," *Journal of the American Statistical Association*, 89, 625-632.
- Lawrence, C. E., Altschul, S. F., Boguski, M. S., Liu, J. S., Neuwald, A. F., and Wooton, J. C. (1993), "Detecting Subtle Sequence Signals," *Science*, 262, 208-214.
- Lazarsfeld, P. F. (1950), "The Logical and Mathematical Foundation of Latent Structure Analysis," in *Studies in Social Psychology in World War II. Vol. 4: Measurement and Prediction*, eds. E. A. Schulman, P. F. Lazarsfeld, S. A. Starr, and J. A. Clausen, Princeton, NJ: Princeton University Press, pp. 362-412.
- Lieberman, S. L. (1994), "More on the Uneasy Case for Using Mill-Type Methods in Small-N Comparative Studies," *Social Forces*, 72, 1225-1237.
- Massey, D. S., and Denton, N. A. (1993), *American Apartheid: Segregation and the Making of the Underclass*, Cambridge, MA: Harvard University Press.
- Matsueda, R. L., and Heimer, K. (1987), "Race, Family Structure, and Delinquency: A Test of Differential Association and Social Control Theories," *American Sociological Review*, 52, 826-840.
- Morris, M. (1997), "Sexual Networks and HIV," *AIDS*, 11, S209-S216.
- Morris, M., Bernhardt, A. D., and Handcock, M. S. (1994), "Economic Inequality: New Methods for New Trends," *American Sociological Review*, 59, 205-219.
- Muthén, B. (1983), "Latent Variable Structure Equation Modeling With Categorical Data," *Journal of Econometrics*, 22, 43-65.
- (1997), "Latent Variable Modeling of Longitudinal and Multilevel Data," *Sociological Methodology*, 27, 453-480.
- Pearl, J. (1998), "Graphs, Causality, and Structural Equation Models," *Sociological Methods and Research*, 27, 226-284.
- Poole, D., and Raftery, A. E. (2000), "Inference from Deterministic Simulation Models: The Bayesian Melding Approach," *Journal of the American Statistical Association*, 95, to appear.
- Raftery, A. E. (1986), "Choosing Models for Cross-Classifications," *American Sociological Review*, 51, 145-146.
- (1991), "Bayesian Model Selection and Gibbs Sampling in Covariance Structure Models," Working Paper 92-4, University of Washington, Center for Studies in Demography and Ecology.
- (1995), "Bayesian Model Selection in Social Research" (with discussion), *Sociological Methodology*, 25, 111-193.
- (1996), "Approximate Bayes Factors and Accounting for Model Uncertainty in Generalized Linear Models," *Biometrika*, 83, 251-266.
- Raftery, A. E., Givens, G. H., and Zeh, J. E. (1995), "Inference From a Deterministic Population Dynamics Model for Bowhead Whales" (with discussion), *Journal of the American Statistical Association*, 90, 402-430.
- Raftery, A. E., Lewis, S. M., and Aghajanian, A. (1995), "Demand or Ideation? Evidence From the Iranian Marital Fertility Decline," *Demography*, 32, 159-182.
- Ragin, C. (1987), *The Comparative Method: Moving Beyond Qualitative and Quantitative Strategies*, Berkeley, CA: University of California Press.
- Roberts, C. W. (1997), "A Generic Semantic Grammar for Quantitative Text Analysis: Applications to East and West Berlin News Content From 1979," *Sociological Methodology*, 27, 89-130.
- Roeder, K., Lynch, K. G., and Nagin, D. S. (1999), "Modeling Uncertainty in Latent Class Membership: A Case Study in Criminology," *Journal of the American Statistical Association*, 94, 766-776.
- Rogoff, N. (1953), *Recent Trends in Occupational Mobility*, Glencoe, IL: Free Press.
- Sankoff, D., and Kruskal, J. B. (1983), *Time Warps, String Edits, and Macromolecules*, Reading, MA: Addison-Wesley.
- Scheines, R., Hoijtink, H., and Boomsma, A. (1999), "Bayesian Estimation and Testing of Structural Equation Models," *Psychometrika*, 64, 37-52.
- Schuessler, K. F. (1980), "Quantitative Methodology in Sociology: The Last 25 Years," *American Behavioral Scientist*, 23, 835-860.
- Schwarz, G. (1978), "Estimating the Dimension of a Model," *The Annals of Statistics*, 6, 461-464.
- Singer, B., Ryff, C. D., Carr, D., and Magee, W. J. (1998), "Linking Life

- Histories and Mental Health: A Person-Centered Strategy," *Sociological Methodology*, 28, 1–52.
- Sobel, M. E. (1998), "Causal Inference in Statistical Models of the Process of Socioeconomic Achievement: A Case Study," *Sociological Methods and Research*, 27, 318–348.
- (2000), "Causal Inference in the Social Sciences," *Journal of the American Statistical Association*, 95, in this issue.
- Spiegelhalter, D., Dawid, P., Lauritzen, S., and Cowell, R. (1993), "Bayesian Analysis in Expert Systems," *Statistical Science*, 8, 219–282.
- Spirtes, P., Richardson, T. S., Meek, C., Scheires, R., and Glymour, C. (1998), "Using Path Diagrams as a Structural Equation Modeling Tool," *Sociological Methods and Research*, 27, 182–225.
- Stovel, K., Savage, M., and Bearman, P. (1996), "Ascription into Achievement: Models of Career Systems at Lloyds Bank, 1890–1970," *American Journal of Sociology*, 102, 358–399.
- Thompson, E. A. (1998), "Inferring Gene Ancestry: Estimating Gene Descent," *International Statistical Review*, 66, 29–40.
- Treiman, D. J. (1977), *Occupational Prestige in Comparative Perspective*, New York: Academic Press.
- Tuma, N. B., and Hannan, M. T. (1984), *Social Dynamics: Models and Methods*, Orlando, Fla.: Academic Press.
- Warren, J. R., LePore, P. C., and Mare, R. D. (2000), "Employment During High School: Consequences for Students' Grades in Academic Courses," *American Educational Research Journal*, to appear.
- Wasserman, S., and Faust, K. (1994), *Social Network Analysis: Methods and Applications*, Cambridge, U.K.: Cambridge University Press.
- Wasserman, S., and Pattison, P. (1996), "Logit Models and Logistic Regressions for Social Networks. I. An Introduction to Markov Graphs and  $p$ ," *Psychometrika*, 61, 401–425.
- Weakliem, D. L. (1999), "A Critique of the Bayesian Information Criterion for Model Selection" (with discussion), *Sociological Methods and Research*, 27, 359–443.
- Western, B. (1996), "Vague Theory and Model Uncertainty in Macrosociology," *Sociological Methodology*, 26, 165–192.
- Western, B., and Jackman, S. (1994), "Bayesian Inference for Comparative Research," *American Political Science Review*, 88, 412–423.
- Wright, S. (1921), "Correlation and Causation," *Journal of Agricultural Research*, 20, 557–585.
- Xie, Y. (1994), "Log-Multiplicative Models for Discrete-Time, Discrete-Covariate Event History Data," *Sociological Methodology*, 24, 301–340.
- Yamaguchi, K. (1992), "Accelerated Failure-Time Regression Models With a Regression Model of Surviving Fraction: An Application to the Analysis of 'Permanent Employment' in Japan," *Journal of the American Statistical Association*, 87, 284–292.

## Psychometrics

Michael W. BROWNE

To progress, a scientific discipline must develop methodology for obtaining measurements of relevant constructs and to extract meaning from the measurements it does have. This is not a straightforward matter in psychology. Typically, constructs of interest are not clearly defined and cannot be measured directly. In addition, the measurements that are available are subject to substantial measurement error. Consequently, the measurement process often consists of repeated attempts to measure the same construct in different ways. When the relationship between several constructs is under investigation, each of the constructs is measured repeatedly, resulting in a substantial number of measurements. Thus the statistical methodology developed for the analysis of psychological measurements is typically multivariate.

Because constructs are not clearly defined, the investigator is often not sure exactly what is being measured. This has led to the concept of a latent or hidden variable that is not measured directly. Inferences about the latent variable are deduced from interrelationships between manifest, or observed, variables.

In a broad sense, psychometrics may be regarded as the discipline concerned with the quantification and analysis of

human differences. This involves both the construction of procedures for measuring psychological constructs and the analysis of data consisting of the measurements made. In this sense, both the construction of a psychological attitude scale and the analysis of the data resulting from its application may be regarded as part of psychometrics. In a narrower sense, psychometrics is often regarded as the development of mathematical or statistical methodology for the analysis of measurement data in psychology. This methodology is primarily multivariate, and latent variables feature strongly. It is this aspect of psychometrics that I consider here.

Although many of the techniques that currently constitute psychometrics date far earlier, psychometrics emerged as a formal discipline with the formation of the Psychometric Society in 1935, with L. L. Thurstone as its first president. Thurstone had played a leadership role at the University of Chicago in providing methodology for analyzing measurements of psychological constructs, and students of his featured strongly among the founding members of the society (Horst and Stalnaker 1986). From the beginning, the Psychometric Society has had an international membership, and the proportion of members outside the United States has grown steadily over the years.

The Psychometric Society produced a journal, *Psychometrika*, whose stated aim was the "development of psychology as a quantitative rational science." Subsequently, *Psy-*

---

Michael W. Browne is Professor in the Department of Psychology and in the Department of Statistics, The Ohio State University, Columbus, OH 43210 (E-mail: [browne.4@osu.edu](mailto:browne.4@osu.edu)). The author thanks the associate editor, Mark Becker, who, in collaboration with three anonymous referees, provided constructive suggestions for improvement. He is also grateful to Michael Walker and to Robert MacCallum for helpful comments that have positively influenced this vignette.