

## Book and Software Reviews

Books and software for review are to be sent to the Book Review Editor Jaki Stanley McCarthy, USDA/NASS, Research Division, Room 305, 3251 Old Lee Highway, Fairfax, VA 22030, U.S.A.

---

SPENCER, BRUCE D., <i>Statistics and Public Policy</i> <i>Bruce D. Rogers</i> . . . . .	325
CHAPMAN, AUDREY R. (Ed.), <i>Health Care and Information Ethics. Protecting Fundamental Human Rights</i> <i>Judy Calder</i> . . . . .	329
DICKINSON GIBBONS, JEAN, <i>Nonparametric Methods for Quantitative Analysis</i> (3rd edition) <i>Ramalingam Shanmugan</i> . . . . .	331
CHOW, S.L., <i>Statistical Significance: Rationale, Validity, and Utility</i> <i>John Tarnai</i> . . . . .	332

---

**Bruce D. Spencer**, *Statistics and Public Policy*. New York, Oxford University Press, 1997. ISBN 0-19-852341-6. 304 pp., 36.95 USD.

Stanford economist Michael Boskin argues that the accelerated pace of change in the American economy is “straining the ability of the statistical system to measure economic performance accurately” (see Boskin 1997). At a time when the stock market has been volatile and performing poorly in Asia and at record levels and volatile in the United States, such a warning carries tremendous weight. Almost daily on the way to the office we hear news reports about which government, or other, statistics is going to be released that day, and in the evening we return home listening to new commentators suggesting how the stock market responded to the release of those statistics. Given record participation by individuals in the stock market through their retirement savings accounts, more people are attuned to the effect of government statistics on their lives than ever before.

In such an environment Bruce D. Spencer’s *Statistics and Public Policy* should receive a receptive audience exploring as it does the use of statistics in public administration. The book was written as a tribute to a pioneer in the field of applied statistics in public management, Richard Savage, and extends Professor Savage’s work in new directions. I have reviewed Spencer’s edited book from the perspective of a graduate professor of public administration, almost all of whose students already work full-time in the public sector. The mostly original chapters in the book demonstrate that when used as another management tool, statistics enhance the decision-making process. The book generally emphasizes ideas rather than technical issues (although there is an interesting and detailed discussion of probability analysis in DNA testing) and statistical perspectives more than public policy

perspectives. W. Allen Wallis's address to the Washington Statistical Society, "Statistics in Washington: 1935–1945," reprinted in this volume recounts the striking growth in the use of statistical analysis during that ten-year period. The chapters in the Spencer book illustrate newer applications for public policy analysis given innovative uses of statistical reasoning. The book is divided into three statistical thematic areas: the uses of statistics for description, issues in data collection, and the uses of statistics for public policy analysis.

The uses of statistics for description section begins with the article "Issues in DNA fingerprinting." Herman Chernoff's article gives a detailed accounting of the techniques of DNA fingerprinting and offers a probabilistic model which makes evident that a false identification "is much more likely to be due to an error in labeling the biological specimens, or tampering with evidence." The article demonstrates the statistical basis for the use of DNA. Interestingly, it is now approximately 100 years after the introduction of fingerprinting as a means for identifying suspects.

In Thomas Jabine's very interesting article on "The emerging field of human rights statistics," the intersection of statistics and politics is reached. In essence this chapter is concerned with the operationalization of human rights. How can we most appropriately measure human rights? In an attempt to be as little partisan as possible he has chosen to measure human rights by operationalizing principles in the Universal Declaration of Human Rights, a document that most countries of the world have subscribed to as well as subsequent human rights treaties. Jabine uses such measures as access to health care in the United States, no-name burials in Argentina, the greater economic costs of not immunizing children than immunizing them, housing discrimination and use of landmines, to illustrate how human rights might be measured. In particular there is an extremely interesting discussion of estimation procedures for determining numbers of rapes in the former Yugoslavia. David Chu and Nancy Spruill look at the defense and intelligence establishments "Making defense decisions: what role might a statistical perspective play?" and make compelling arguments for changing the focus in debates from "consensus" to a "consensus distribution" permitting an examination of alternative probabilities in intelligence, resource allocation, weapons choices and cost estimation.

On what political or social issues should researchers collect statistics? Stephen Fienberg, "Ethics, objectivity, and politics: statistics in a public policy context," argues that those who collect data need to be aware that values and political pressures play a role in choices of problems to be examined and in the statistical methods that are used. He comes down on the side of those who urge an expansion of racial and ethnic categories in federal government reporting. One important consideration that is not discussed is the possibility in this case that political "pressure" might prevent the further subdivision of a population looking for common links that an unfettered collection of racial and ethnic statistics might hinder. The issue here, of course, is what is reported in any way is dishonest or untruthful.

In a very well researched chapter Fienberg looks at four intensely debated policy areas including aids, employment discrimination and the decennial census. With much supporting information he illustrates that in those areas political pressures "imperiled the collection and dissemination of quality statistical information of material importance."

Part II of the volume focuses on the data collection for description. Given the controversy in the U.S. Congress over whether or not sampling techniques may be utilized in

the 2000 Census, Leslie Kish's "Periodic and rolling samples and censuses" is a very timely chapter. Undercounting problems and a decrease in public cooperation during a census lead Kish to argue the merits of rolling samples, which are special cumulated samples. Frequently data from a decennial census could be used when the data is well over ten years old from the time the data is released from one census until the time the next decennial statistics are released. Kish argues that using rolling samples permits an examination of a changing population over the ten-year period covered by a traditional census and provides more accurate and usable data.

While the proposal for the 2000 Census in the U.S. was not to have a rolling sample until the census in 2010, it would have utilized statistical sampling. Speaker of the House Newt Gingrich, filed suit in U.S. District Court in Washington, in February 1998, claiming that statistical sampling violates the Constitution and the Federal Census Act. The Supreme Court will ultimately decide the methods to be permitted in the next U.S. Census.

Philip Redfern extends the discussion about "Numbering the people: issues of accuracy, privacy and open government." Kish discussed address registries as used in four Nordic countries for enumerating a population. Redfern contrasts the advantages of a census with that of an address registry where individuals are assigned a personal number. For the latter, the principle disadvantage is the limited number of questions asked in the Nordic countries. Another problem is the reliance on administrative records. It is necessary to explore alternatives or supplements to traditional counts of the population to ensure greater accuracy and to overcome obstacles to a more accurate count. Address registries may not be the answer to such obstacles.

In addition to counting a population as a whole, it is important, too, to be able to analyze data for subpopulations. In particular, fairly accurate counts for members of the population with disabilities and their types could assist government planners. The problem, as Allan Sampson points out in "Surveying individuals with disabilities," is in defining disability. Examining the widely varying survey estimates of the disabled population in the U.S. Sampson examines frameworks for conceptualizing disability. He notes the differing conceptualizations between the U.S. and Europe and argues that any agreed upon definitions must also take into account how the environment affects the disability. Sampson also looks at questions about the reliability and validity of survey questions to disabled populations as well as the general question of sampling rare populations. Sampson considers the problem within the context of a census. However, once the conceptual issues are resolved an equally large question remains, how do you randomly sample disabled individuals from a general population? Are the subpopulation sampling frames adequate? That is left for another article.

In a highly technical chapter James Coleman, "Constructed social networks in the study of diffusion," is concerned with statistics making use of relations among individuals. He chooses the "diffusion of sexually transmitted diseases in a sexually interacting population" as his focus. Coleman examines how a researcher might attempt to construct a social network from a sample of the population and then study processes utilizing that network. The process appears promising to future researchers.

In Part III of the book the authors are concerned with the uses of statistics for policy analysis. A mystifying question explored by Nathan Keyfitz "Why forecasts fail and

policies are often frustrated?" is the basis for this chapter. He maintains that the way we approach short-term solutions will influence long-term effects. By looking at several models used in policy analysis and forecasting Keyfitz shows the importance of discipline-related interpretations leading to either success or failure. However, the forces of politics, economics and society cannot be treated separately, he maintains. It is the interaction effects of the three that can lead to prolonging short-term solutions.

The public sector has long belabored the question of whether to use an experimental design or not in evaluating new programs. Because the programs are 'public,' experimentation easily could become a political issue. Why should one group get a 'better deal than another' albeit for a short time? Lincoln Moses and Frederick Mosteller in "Experimentation: just do it!" make a convincing argument for the use of randomized controlled experimentation. They recognize that its utility may be limited to certain policy areas because of the nature of the problem. They also offer good curriculum advice to graduate programs in statistics and in public policy.

Stephen Dresch and Kenneth Janson in "Talents, rewards and professional choice: a general equilibrium analysis," are concerned with incomplete and inaccurate analytical models resulting in incomplete and inaccurate policy conclusions. They demonstrate these points by constructing a model of the labor market. They use simulations as a check on the policy analytical models.

In the final chapter, Sam Savage writes about "Statistical analysis for the masses." During the conduct of his nationwide seminars on management science and spreadsheets, he initially thought that seminar participants did not understand the Central Limit Theorem. He realized that he was wrong and that the participants did not understand what was meant by a probability distribution. He makes recommendations for tools that would enable users to move away from asking "what is this number?" to asking "what is this distribution?"

The Spencer volume contributes conceptually and policy-analytically to the joint fields of statistics and public policy. It will make an excellent addition to classes in applied statistics and public policy. Its use will cross disciplinary lines but will be of particular value to graduate students in public management. As a result of this edition I anticipate that we will see more papers being written that extend the Spencer book and which add new methodological ground to the study of public policy.

## Reference

Boskin, M. (1997). Some Thoughts on Improving Economic Statistics. Hoover Institution, Essays in Public Policy. <http://www-hoover.stanford.edu/publications/epp/86/86a.html>.

*Bruce D. Rogers*  
*Institute of Government*  
*Tennessee State University*  
*Box 140*  
*10th and Charlotte*  
*Nashville, TN 37203-3401, U.S.A.*  
*Phone: 615-963-7248*  
*Fax: 615-963-7245*  
*E-mail: rogers07@harpo.tnstate.edu*

**Audrey R. Chapman (ed.)**, *Health Care and Information Ethics. Protecting Fundamental Human Rights*. Kansas City, MO: Sheed and Ward Publicity, 1997. ISBN 1-55612-922-X. 480 pp., 29.95 USD.

This very important book, "Health Care and Information Ethics: Protecting Fundamental Human Rights," grew out of a project sponsored by the American Association for the Advancement of Science (AAAS), with subsequent assistance from the Robert Wood Johnson Foundation. While the title is indeed indicative of the central issues covered by the diverse group of authors who wrote the book's individual chapters, it does not even begin to hint at the broad spectrum of professionals and students who can benefit from (and even be intrigued by) the wealth of useful, well organized information and analysis the book contains. While the topics treated in this book have serious implications for both consumers and providers of health care alike, the information and analyses which the authors present also relate directly to critical issues associated with sampling and research methodology, survey research, statistics and psychometrics, computer science and information systems management, ethics, and philosophy. If the same restricted, limited view of purpose evident in the choice of title for the current volume had been applied to Michael Crichton's "Jurassic Park," for example, Crichton's work would have wound up with the enticing title "A Book About Dinosaurs," accurate, though limited as a description of its contents. In short, in my opinion it would be a shame if the somewhat 'stand-offish' title, "Health Care and Information Ethics," had the effect of discouraging individuals not engaged in either health care or information ethics from picking up this book and reading it closely from cover to cover.

The book's editor, Audrey Chapman, Ph.D., delineates the key proposition which the authors addresses from their own particular substantive vantage point. This proposition is based on the notion that access to a basic and adequate standard of health care, "consistent with a particular society's level of resources," is a basic human right. It is in this human rights context that Chapman introduces the issues which the earlier AAAS project identified as in need of redress if health care information systems are to be developed which are consistent with the four norms of fundamental human rights: (1) universality; (2) privacy; (3) nondiscrimination; and (4) consent.

Electronic health care information systems are seen by Chapman and the book's other authors as offering "new opportunities for monitoring access to health care" while "imposing additional challenges for protecting human rights." One of the project's stated goals is to identify design criteria and methodologies to allow health care information systems to be consistent with human rights. This requires improved monitoring of access to health care, which in turn requires answers to three key questions: (1) what is the minimal amount of information which needs to be collected in order to evaluate universality of access to health care; (2) which variables are most important; and (3) what are the most efficient methodological strategies to attain that goal? To answer these questions, the limitations of current data sources are discussed, and new approaches, methodologies and types of research initiatives are proposed in the first half of the book. The second half of the book is devoted to proposals for building human rights protections into the design of health care information systems.

Discussions of the kinds of data needed for monitoring access to health care are excellent. For example, Bennett's discussion of the probable effects of welfare reform

on poor women's access to health care, as well as the sources of data required to monitor changes, is illuminating. This is followed by Sugarman et al., looking at new trends towards "self-governance," the process by which individual tribes have been given tribal level control over federal resources for Indian health programs. It is clear that the results of self-governance will have to be judged using data which will have to come from the tribes themselves, instead of other federal agencies such as the U.S. Bureau of Indian Affairs. Fowler and Samuelson astutely suggest that the definitions and indicators of access found in the Institute of Medicine's 1993 study represent a good basis from which to develop a specific monitoring plan.

The ratings of existing data sources for monitoring access and universality is a tour de force of extant national and state level data collection efforts which could easily serve as a history of health-related surveillance efforts for students as well as for professionals. Descriptions of the uses of the on-going National Health Interview Survey, including periodic surveys of medical expenditures, Medicare beneficiaries, and people in nursing homes, clearly identify these surveys as valuable sources of high quality data, representing a reasonable degree of comparability over time. As such, they are recommended for use in any system designed to monitor access and universality, keeping in mind that they do not cover some of society's most vulnerable groups, including migrant workers and the homeless populations. The discussion of the strengths and weaknesses of vital records (birth and death certificates), administrative records, and Medicare data is particularly informative, as is the extensive presentation related to the problems associated with the use of ethnic and racial classifications.

The chapters dealing with issues of privacy, nondiscrimination, and disclosure should make every American's hair stand on end. The potential for discrimination in employment, based on personal health care data, is particularly problematic. The gravest threat to privacy and control over disclosure and use of an individual's health information relates to the private sector's ability to obtain and use such information. With the exception of the limited protection provided by the Americans with Disabilities Act, and through legislation that applies to specific records regarding substance abuse, laws regarding the use of health information by the private sector are simply nonexistent.

Issues related to control (or the use of individual health care information) for purposes of research and system monitoring (universality and nondiscrimination) produced a wide range of views, including the suggested need for written authorization for all disclosures and uses to Baily's position that all individuals have a responsibility to make their data available for legitimate purposes. Or, as Duncan sees it, privacy advocates and data users will have to cooperate: "In building a quality medical care system there is no tradeoff between privacy and confidentiality on the one hand and data access on the other. Both are required if high quality data is to serve health care needs."

The breadth of the issues covered in this book, as well as the depth with which some of these issues (for example, outcomes measurement) are covered, may make it unreasonable to expect some mention of the widely held notion that the quality of health care in the United States, even with all its attendant problems, is still second to no other. In addition, to leave out a discussion of the structure of our health care delivery system in the context of its integral development within a singular system of free enterprise as opposed to "European Socialism" seems a lost opportunity to better explore the difficulties associated

with developing health care information systems which can meet the test of consistency with the human rights norms of universality, privacy, nondiscrimination and consent.

Taking all of the chapters into consideration, I give this book a rating of “excellent,” and strongly recommend it to students, scholars, and practitioners interested in any aspect of health care policy or health care delivery, survey research, methodology (especially as regards sampling of rare populations), computer science and information technology, econometrics, as well as those interested in issues of data privacy, human rights, and philosophy.

*Judy Calder*  
*Center for Applied Research*  
*University of Nevada, Reno*  
*Reno, NV 89557-0017, U.S.A.*  
*Phone: (702) 784-6718*  
*Fax: (702) 784-4506*  
*E-mail: calder@scs.unr.edu*

**Jean Dickinson Gibbons**, *Nonparametric Methods for Quantitative Analysis (3rd edition)*. Columbus, OH: American Sciences Press, 1997. ISBN 0-935950-37-0. 528 pp. + index, 70.95 USD.

This book can be used as a textbook for non-statistics major students from allied disciplines such as business, education, or social science. Every chapter has data-oriented problems for students to practice. Answers to even numbered problems are given in the back. Those who wish to pursue the details of any exposition will find the list of references and of experiments very helpful. The tables in the appendix are quite extensive. In these scientific calculators, I question the importance of having a table on squares and square roots. Perhaps in the future edition, the author might consider removing it.

The book exposes clearly the basics of many Nonparametric statistical techniques. Simple knowledge of algebra and arithmetic would be sufficient to understand the book. However, a basic understanding of statistical principles would enhance a complete understanding of the material about the Nonparametric methods.

Each concept is introduced with an example just to motivate readers, and it is the best feature of the book. In each example, the author has tried to point out on what type of question the researcher might ask, what kind of data are necessary to answer those questions, and what type of Nonparametric techniques are appropriate. The concept of  $P$ -value is introduced in each example. The goodness-of-fit tests and tests on categorical data are extra features in this book which are not seen in other similar books. The additions in the third edition include computer package solutions to selected examples using MINI-TAB, SAS, or SPSS. The Lilliefors goodness-of-fit test, Johnkheere-Terpstra test for ordered alternative hypothesis in one-way ANOVA, the Page test for two-way ANOVA, and the rank von Neumann test for randomness are added in the third edition.

The author is famous for a clear presentation of the material. There are nine chapters covering the topics: introduction to Nonparametric statistics, goodness-of-fit tests, inferential procedures on location parameters based on single and multiple samples, inferences on scale parameters, distribution tests for multiple samples, association analysis, tests

for randomness, and selection procedures for binomial populations. The author could have included the selection procedures for other populations also.

There are controversial statements like “. . . the computer itself cannot think; it can only process according to the program a human mind has stored in it,” “full-fledged statisticians,” and “the debate on cookbooks” are somewhat distasteful in my opinion. Otherwise, the book contains valuable material for those who plan on using Non-parametric statistical techniques. The author has chosen a variety of data for presentation of the statistical ideas. I enjoyed reading this book, and would recommend it to users of statistical techniques from other disciplines.

*Ramalingam Shanmugan*  
*Mathematics Department*  
*University of Colorado at Denver*  
*Denver, CO 80217-3364*  
*U.S.A.*

**S.L. Chow**, *Statistical Significance: Rationale, Validity, and Utility*. London: Sage Publications, 1996. (SBN 0-7619-5205-5 (paperback). 224 pp., £13.95. ISBN 0-7619-5204-7 (cloth). 224 pp., £37.50.

Significance testing has been criticized over the years on logical and philosophical grounds, as well as blamed for the slow progress of “soft” psychology (Harlow, Mulaik, and Steiger 1997; Meehl 1978; Morrison and Henkel 1970). Chow’s primary thesis in this book is that the null-hypothesis significance test procedure (NHSTP) has been criticized for the wrong reasons. Through a comprehensive treatment of criticisms of the NHSTP, Chow outlines statistical and philosophical arguments that support his contention that these criticisms are unwarranted.

The book is organized into eight chapters, with the first chapter summarizing the basic criticisms of the NHSTP. Chapter 2 describes the features of the NHSTP and its mathematical foundations, particularly its origins in the works of Fisher and of Neyman and Pearson. Subsequent chapters address the criticisms in detail, providing a rationale for why they might be invalid. The criticisms are interwoven throughout the book, rather than addressed in separate chapters, and they revolve generally around these issues: (a) problematic aspects of both the null and research hypotheses; (b) ambiguities in the meaning of statistical significance; (c) uninformative nature of the significance test in terms of effect size, likelihood of theory corroboration, and usefulness of results; (d) what significance testing cannot do; and (e) arguments that the null hypothesis is never true.

Chapters 3 and 4 are the toughest reading in the book, because the concepts are difficult to comprehend for those untrained in the philosophy of science and metatheoretical notions. Yet, the concepts are central not only to Chow’s arguments in favor of the NHSTP, but also to the reader’s grasp of how the remaining chapters support the book’s primary thesis. Chapter 3 makes a distinction between substantive and statistical hypotheses, which Chow argues is central to an understanding of the NHSTP and its criticisms. A description of the logical relations among hypotheses in four kinds of experiments (theory corroboration, utilitarian, clinical, and generality) is presented to show that the NHSTP criticisms, while unwarranted in every case, are even more inappropriate for



some kinds of experiments than for others. The main point of Chapter 4 is to demonstrate that NHSTP is not an inductive procedure that can or should be used for theory corroboration. Instead, Chow maintains that the only role for the NHSTP is limited to assessing whether chance influences can be disregarded in hypothesis testing. At the same time, however, the NHSTP is an essential step in the process of securing evidential support for theory.

In Chapter 5 Chow argues that, because of the ambiguous and anomalous nature of the NHSTP, many of the complaints about it are based on non-statistical reasoning. Particular examples here include claims that the null hypothesis of no difference between two populations can never be true and that statistical significance does not provide information about the substantive value of the research. Criticisms of the NHSTP posit that statistical significance is largely (a) a matter of sample size, (b) related to effect size in anomalous ways, and (c) uninformative with regard to the practical value of the results. Not surprisingly, given his argument in Chapter 4, Chow also rebuts the common wisdom of abandoning statistical significance in favor of reporting effect size because of its greater value in providing evidential support.

The power of a statistical test to yield statistically significant results is related mathematically to the NHSTP, and the use of power to determine sample size has become standard operating procedure. Chapter 6 challenges the validity of power as well as the notion that power should be used instead of, or in addition to, statistical significance. Chow argues that (a) the meaning of Type II error is different for power and for the NHSTP, (b) those two meanings cannot be reconciled, and (c) graphical representation of power is inconsistent with that of the NHSTP.

Chapter 7 evaluates the Bayesian approach to statistical decision-making and its implications for the NHSTP. Chow's view is that the Bayesian approach requires a different data gathering procedure – which he refers to as the “sequential-sampling procedure” – than that typically used in a theory corroboration experiment. This requirement limits the usefulness of Bayesian methods and argues for continued use of the NHSTP.

The final chapter re-summarizes the basic criticisms of the NHSTP in terms of a set of 14 questions that could be asked of data. Chow suggests that it is unreasonable to expect that the NHSTP can provide an answer to each of these questions. Instead, the major role of significance testing is to allow a decision about whether chance can be ruled-out as an explanation of the data. Arguing that the criticisms of the NHSTP are flawed because basic metatheoretical concepts are misunderstood, Chow concludes that the NHSTP has a limited – but very important – role in empirical research. These conclusions are echoed in a more recent and broader treatment of this topic (Trout 1998), which likewise argues that most of the standard criticisms of the NHSTP are unwarranted.

Chow's work fills a void in the literature on significance testing and provides a concise, but thorough, review of the significance test controversy. Specifically, graduate students in the social and behavioral sciences generally are introduced to significance testing in their first course in statistics, but few students have the opportunities this book offers to explore the methodological rationale and the philosophical underpinnings of statistical significance testing. However, this book may require several readings to appreciate fully the complex arguments that are made.

**References**

- Harlow, L., Muliak, S., and Steiger, J. (eds.) (1997). *What if There Were No Significance Tests?* Lawrence Erlbaum: Mahwah, N.J.
- Meehl, P.E. (1978). Theoretical Risks and Tabular Asterisks: Sir Karl, Sir Ronald, and the Slow Progress of Soft Psychology. *Journal of Consulting and Clinical Psychology*, 46, 806–834.
- Morrison, D.E. and Henkel, R.E. (eds.) (1970). *The Significance Test Controversy: A Reader*. Aldine Publishing, Chicago.
- Trout, J.D. (1998). *Measuring the Intentional World*. Oxford University Press, New York.

*John Tarnai*  
*Washington State University*  
*Social and Economic Sciences Research Center*  
*133 Wilson Hall*  
*Pullman, WA 99164-4014, U.S.A.*  
*Phone: 509 335 1511*  
*Fax: 509 335 0116*  
*E-mail: tarnai@wsu.edu*