Springer Series in Statistics

Advisors:
P. Bickel, P. Diggle, S. Fienberg, U. Gather, I. Olkin, S. Zeger
Alho/Spencer: Statistical Demography and Forecasting.
Andersen/Borgan/Gill/Keiding: Statistical Models Based on Counting Processes.
Atkinson/Riani: Robust Diagnostic Regression Analysis.
Atkinson/Riani/Cerioli: Exploring Multivariate Data with the Forward Search.
Buckley: Introduction to Rare Event Simulation.
Cappé/Moulines/Ryden: Inference in Hidden Markov Models.
Chan/Tong: Chaos: A Statistical Perspective.
Chen/Shao/Ibrahim: Monte Carlo Methods in Bayesian Computation.
Coles: An Introduction to Statistical Modeling of Extreme Values.
Devroye/Lugosi: Combinatorial Methods in Density Estimation.
Efroymovich: Nonparametric Curve Estimation: Methods, Theory, and Applications.
Fienberg/Hoaglin: Selected Papers of Frederick Mosteller.
Frühwirth-Schnatter: Finite mixture and markov switching models.
Ghosh/Ramamoorthi: Bayesian Nonparametrics.
Glaz/Naus/Wallenstein: Scan Statistics.
Gu: Smoothing Spline ANOVA Models.
Györfi/Kohler/Krzyzak/Walk: A Distribution-Free Theory of Nonparametric Regression.
Hall: The Bootstrap and Edgeworth Expansion.
Härdle: Smoothing Techniques: With Implementation in S.
Hastie/Tibshirani/Friedman: The Elements of Statistical Learning: Data Mining, Inference, and Prediction.
Ibrahim/Chen/Sinha: Bayesian Survival Analysis.
Kotz/Johnson (Eds.): Breakthroughs in Statistics Volume I.
Kotz/Johnson (Eds.): Breakthroughs in Statistics Volume II.
Kotz/Johnson (Eds.): Breakthroughs in Statistics Volume III.
Küchler/Sørensen: Exponential Families of Stochastic Processes.
Kutoyants: Statistical Influence for Ergodic Diffusion Processes.
Lahiri: Resampling Methods for Dependent Data.
Le Cam: Asymptotic Methods in Statistical Decision Theory.
Li: Monte Carlo Strategies in Scientific Computing.

(continued after p. 660)
Both of us have had the good fortune to work closely with Frederick Mosteller and learn from him throughout our professional careers. We have been inspired by his creative approach to statistics and its applications and we, along with countless others, have benefited from reading his papers and books. Through this volume we hope to share the variety and depth of Mosteller’s writings with a new generation of researchers, who can build upon his insights and efforts. In many ways this volume of selected papers can be viewed as a companion to an earlier volume that we assembled with William Kruskal and Judith Tanur, *A Statistical Model: Frederick Mosteller’s Contributions to Statistics, Science, and Public Policy*, Springer-Verlag (1990), and Fred’s forthcoming autobiography, which will also be published by Springer-Verlag.

This volume contains a reasonably complete bibliography of Mosteller’s papers, books and other writings. Several of these, which Fred co-authored with his friend and long-time collaborator John Tukey, are included in the multi-volume *Collected Works of John W. Tukey* (CWJWT). Conveniently, the bibliography in CWJWT tells whether a paper appears in one of those volumes. CWJWT includes papers P10, P70, P133, and P142 from the Mosteller bibliography. We say that the bibliography is “reasonably” complete because we know that Fred continues to work on projects and to collaborate on revisions of books.

We chose the papers for this volume to give a broad perspective on Fred’s work, ranging from statistical theory through applications in a variety of domains, and reflecting his long and varied career. We iterated with Fred over a period of a couple of years until we had a collection that captured the nature of his contributions and also fit within a single volume. Each paper reflects one or more aspects of Fred’s approach to statistical research and its application, and the papers often reflect his general philosophy on how one should go about doing good science.

In preparing the papers for this volume, we scanned the original, converted the resulting image to text, and transformed the text into input for *LaTeX*. By careful proofreading we have tried to ensure that each paper is faithful to
the original. We have, however, corrected typographic errors in the original papers when we encountered them (not often) and occasionally made other straightforward changes. Also, we have restructured some tables, because of incompatibilities in page size and layout.

We are deeply indebted to a number of others who contributed to the preparation of the volume, including Cleo Youtz, Marjorie Olson, and Jessa Piaia at Harvard University, and especially to those who actually prepared parts of the \LaTeX{} document over the 15 years we have been slowly working on this volume: Valerie Baddon (at York University), Howard Fienberg, Valerie Lenhart, John Clark, Heather Wainer, and Heidi Sestrich (all at Carnegie Mellon). Heather in particular has valiantly worked with us to ensure that the format and content were indeed correct and as consistent as feasible, given the diversity of the original publications and their format styles, and she helped immeasurably in the reproduction of the figures from the original papers.

Stephen E. Fienberg December 24, 2005
David C. Hoaglin
Contents

Preface ........................................................ v

Frederick Mosteller—A Brief Biography
Stephen E. Fienberg .............................................. 1

Bibliography ................................................... 7

1. Unbiased Estimates for Certain Binomial Sampling
Problems with Applications
M.A. Girshick, Frederick Mosteller, L.J. Savage ................... 57

2. On Some Useful “Inefficient” Statistics
Frederick Mosteller ............................................. 69

3. A $k$-Sample Slippage Test for an Extreme Population
Frederick Mosteller .............................................. 101

4. The Uses and Usefulness of Binomial Probability Paper
Frederick Mosteller, John W. Tukey ................................ 111

5. The Education of a Scientific Generalist
Hendrik Bode, Frederick Mosteller, John W. Tukey, Charles Winsor .... 147

6. Remarks on the Method of Paired Comparisons: I. The
Least Squares Solution Assuming Equal Standard Deviations
and Equal Correlations
Frederick Mosteller .............................................. 157

7. Remarks on the Method of Paired Comparisons: II. The
Effect of an Aberrant Standard Deviation When Equal
Standard Deviations and Equal Correlations Are Assumed
Frederick Mosteller .............................................. 163
8. Remarks on the Method of Paired Comparisons: III. A Test of Significance for Paired Comparisons when Equal Standard Deviations and Equal Correlations Are Assumed
Frederick Mosteller ................................................. 167

9. An Experimental Measurement of Utility
Frederick Mosteller, Philip Nogee .................................. 179

10. A Mathematical Model for Simple Learning
Robert R. Bush, Frederick Mosteller ................................. 221

11. A Model for Stimulus Generalization and Discrimination
Robert R. Bush, Frederick Mosteller ................................. 235

12. The World Series Competition
Frederick Mosteller ..................................................... 251

13. Principles of Sampling
William G. Cochran, Frederick Mosteller, John W. Tukey .......... 275

14. Stochastic Models for the Learning Process
Frederick Mosteller ..................................................... 295

15. Factorial $\frac{1}{2}$: A Simple Graphical Treatment
Frederick Mosteller, D.E. Richmond ................................ 309

16. A Comparison of Eight Models
Robert R. Bush, Frederick Mosteller ................................ 317

17. Optimal Length of Play for a Binomial Game
Frederick Mosteller ..................................................... 333

18. Tables of the Freeman-Tukey Transformations for the Binomial and Poisson Distributions
Frederick Mosteller, Cleo Youtz .................................... 337

19. Understanding the Birthday Problem
Frederick Mosteller ..................................................... 349

20. Recognizing the Maximum of a Sequence
John P. Gilbert, Frederick Mosteller ................................ 355

21. The Distribution of Sums of Rounded Percentages
Frederick Mosteller, Cleo Youtz, Douglas Zahn .................... 399

22. The Expected Coverage to the Left of the $i$th Order Statistic for Arbitrary Distributions
Barry H. Margolin, Frederick Mosteller ............................ 413
23. Bias and Runs in Dice Throwing and Recording: A Few Million Throws
Gudmund R. Iversen, Willard H. Longcor, Frederick Mosteller, John P. Gilbert, Cleo Youtz ................................. 417

Frederick Mosteller .......................................................... 435

25. A Conversation About Collins
William B. Fairley, Frederick Mosteller ................................. 445

26. Statistics and Ethics in Surgery and Anesthesia
John P. Gilbert, Bucknam McPeek, Frederick Mosteller ............... 457

27. Experimentation and Innovations
Frederick Mosteller .......................................................... 473

Frederick Mosteller, Gale Mosteller ...................................... 487

29. Classroom and Platform Performance
Frederick Mosteller .......................................................... 499

30. The Clinician’s Responsibility for Helping to Improve the Treatment of Tomorrow’s Patients
Bucknam McPeek, John P. Gilbert, Frederick Mosteller ............... 513

31. Innovation and Evaluation
Frederick Mosteller .......................................................... 515

32. Combination of Results of Stated Precision: I. The Optimistic Case
Frederick Mosteller, John W. Tukey ...................................... 531

33. Combination of Results of Stated Precision: II. A More Realistic Case
Frederick Mosteller, John W. Tukey ...................................... 549

34. Allocating Loss of Precision in the Sample Mean to Wrong Weights and Redundancy in Sampling with Replacement from a Finite Population
J.L. Hodges, Jr., Frederick Mosteller, Cleo Youtz ...................... 575

35. Reporting Clinical Trials in General Surgical Journals
John D. Emerson, Bucknam McPeek, Frederick Mosteller .......... 583
36. Compensating for Radiation-Related Cancers by Probability of Causation or Assigned Shares
Frederick Mosteller ......................................................... 597

37. Methods for Studying Coincidences
Persi Diaconis, Frederick Mosteller ................................. 605

38. A Modified Random-Effect Procedure for Combining Risk Difference in Sets of $2 \times 2$ Tables from Clinical Trials
John D. Emerson, David C. Hoaglin, Frederick Mosteller .......... 623

39. The Case for Smaller Classes and for Evaluating What Works in the Schoolroom
Frederick Mosteller .......................................................... 643

40. Frederick Mosteller and John W. Tukey: A Conversation
Moderated by Francis J. Anscombe ................................. 647
Frederick Mosteller—A Brief Biography

Stephen E. Fienberg

*Carnegie Mellon University*

Frederick Mosteller celebrated his 89th birthday this year (2005) but remains professionally active and involved in several research projects. For an extensive review of many of his contributions, see the special volume *A Statistical Model* [5], prepared in his honor. John Tukey’s biography of Fred (as he is known to his friends, collaborators, colleagues, and students) in that volume is especially noteworthy.

Fred was born Charles Frederick Mosteller in Clarksburg, West Virginia, on December 24, 1916.¹ The family later moved to the Pittsburgh area, where Fred attended Schenley High School and later Carnegie Institute of Technology (now Carnegie Mellon University). In college, he was interested in mathematics and, in particular, in combinatoric problems. This inclination led him to the statistician Edwin C. Olds, who in turn steered Fred into the field of statistics. Fred completed his Sc.M. degree at Carnegie Tech in 1939 and then enrolled at Princeton University to work on a Ph.D. with Samuel Wilks. In addition to participating with Wilks and others in a wartime research group (see the discussion in [17]), Fred assisted Wilks in his role as editor of the *Annals of Mathematical Statistics*. Fred received the ASA Samuel S. Wilks Award in 1986. At Princeton Fred also began his lifelong interaction and collaboration with John Tukey, described in part in [3] and in their joint interview in *Statistical Science* [14].

Fred met his wife, Virginia (1917–2000), when he was a college freshman (she rode the same streetcar that Fred took from Wilkinsburg to the campus each day), and they were married in 1941, just as the statistical war research

¹This is a birth-date he shares with my mother, and a birthday—same day but different year—he shares with David Wallace, a collaborator of Fred’s. One of Fred’s favorite probability examples has been the well-known birthday problem, see [8]. This event also links to another of Fred’s preoccupations, the statistical analysis of coincidences [4].
effort was gearing up. Fred accepted a position in the Department of Social Relations at Harvard University in 1946, and he remained on the faculty of the university in various positions (see below) for the rest of his career. Fred and Virginia moved to their home in Belmont after the birth of their son Bill in 1947. Their daughter Gale was born in 1953.

Fred became Professor of Mathematical Statistics at Harvard in 1951 and led the effort to create the Department of Statistics, which opened in 1957. He served as the department’s first chair, from 1957 to 1969. I remember my first meeting with Fred in the fall of 1964 as I was entering graduate school. Fred took me to lunch at the Harvard Faculty Club, where he insisted that I try the horse steak. As I was busy chewing, he shared with me some handwritten notes from Tukey on assessing probability assessors, a problem that was to be my first project as one of Fred’s research assistants. Although we wrote up the results of that effort as a joint technical report, I confess it was not until a collaboration on the topic with Morrie DeGroot, some 18 years later, that I came to understand and appreciate Fred’s and John’s insights into the problem.

As I was completing my Ph.D. under Fred’s supervision in 1968, he organized a group of us to write a book built around the recent developments in categorical data analysis, especially linked to the use of log-linear models. This project ultimately produced *Discrete Multivariate Analysis: Theory and Practice* [1]. He was the guiding light behind the project and our constant editor and sometimes contributor, but in typical fashion he insisted that only Yvonne Bishop, Paul Holland, and I be listed as “authors.” Ultimately, he agreed to let us acknowledge his efforts by listing him as a “collaborator” on the title page.

During his years in the Statistics Department, Fred formally supervised 17 Ph.D. dissertations, but he served on the committees of countless others in Statistics, Social Relations, and other parts of Harvard. Over the years he was always available to offer comments on works in progress and unpublished manuscripts, and wise students took advantage of his generosity.

Fred later served as chair of two other departments, Biostatistics and Health Policy and Management, both in the School of Public Health, and he also taught courses in the Harvard Law School and the John F. Kennedy School of Government. On retirement in 1987, Fred maintained his office in the Department of Statistics, where he continued with his usual array of multidisciplinary projects, almost as if nothing had changed. At the end of 2003, he dismantled his office and relocated to the Washington, D.C. area.

Fred’s bibliography is astounding; it contains 65 books, nearly 350 papers in books and journals, 41 miscellaneous publications, and 26 reviews.
And, not surprisingly, many of these were coauthored or coedited by over 200 other individuals. My personal favorite is Fred’s classic 1964 book with David Wallace, *Inference and Disputed Authorship: The Federalist* [15], which was republished in 1984 in expanded form [16]. The intriguing analyses of *The Federalist Papers* presented by Mosteller and Wallace include one of the first major uses of Bayesian methods, and they provide an early exposition of Laplace’s method for approximating distributions.

A partial list of Fred’s varied methodological research interests includes publications on inefficient statistics, sampling, Bayesian methods, paired comparisons, the jackknife, statistics in sports, contingency table analysis, exploratory data analysis, randomized experiments, robustness and meta-analysis and research synthesis.

Depending on how one measures collaboration, John Tukey is Fred’s most extensive collaborator; but others prominent on the list include Thomas Chalmers, John Gilbert, David Hoaglin, Bucknam McPeek, and Fred’s long-time research assistant, Cleo Youtz. Other collaborators of note include John Bailar, Bill Cochran, Persi Diaconis, Milton Friedman, Bill Kruskal, Pat Moynihan, Jimmie Savage, Judy Tamur, Allen Wallis, Sam Wilks, Charlie Winsor, and Gale Mosteller (his daughter). And I am pleased to be included on the list.

Beginning in the 1950s, Fred helped lead an effort to bring probability and statistics to American high schools. He was instrumental in producing teacher’s manuals, and this involvement led to one of the early elementary statistics texts, *Probability with Statistical Applications* [13]. Fred used a version of this book as the text for his pioneering 1961 televised course on NBC’s *Continental Classroom*, which introduced him to students across the nation, young and old. Many textbooks on varied topics followed. When the American Statistical Association set up a joint committee with the National Council of Teachers of Mathematics (NCTM) in the 1960s to change the statistical content of the secondary school mathematics curriculum, Fred led the effort once again. He helped to organize, and goaded others into contributing to, the preparation of the ASA-NCTM Committee’s early products, including the 4-volume collection *Statistics by Example* [9, 10, 11, 12] and *Statistics: A Guide to the Unknown* [17], which has now appeared in multiple forms and multiple editions.

Fred has always been an organizer, and this talent was recognized by many different societies and other organizations that came to Fred for help with projects, as well as to fill leadership positions. Among the societies he has led as president are (in approximately chronological order): the Psychometric Society, the American Statistical Association, the Institute of Mathematical Statistics, the American Association for the Advancement of Science, and the
International Statistical Institute. In the 1960s, he served as Chairman of the Board of Directors of the Social Science Research Council and later as Vice-chair of the President’s Commission on Federal Statistics, which led to the creation of the Committee on National Statistics at the National Research Council (NRC). At the NRC and elsewhere he has served on so many statistical and interdisciplinary committees and task forces that one observer [6, p. 87] was led to remark:

Applied mathematicians are of course essential in most assessments, for help in designing tests, auditing calculations, and assisting in the drawing of conclusions. Much of this is routine craftwork. But one unusual, crucial role should be recognized. It is my guess that statisticians Frederick Mosteller (Harvard) and John Tukey (Princeton) have served on or assisted more technical committees than anybody else alive . . . . It is not just their ability to manipulate numbers that keeps these experts in demand, but sensibility in thinking through questions of macro-experimental design: how inquiries should be cast, what evidence and logic are applicable, how discrimination can be increased, how uncertainties and sensitivities should be probed, what inferences are allowable from evidence. Mosteller and Tukey outlined this role in an article in 1949 in which they called for education of “scientific generalists” who would master “science, not sciences” [2].

The products of many of these activities are mentioned in the bibliography that follows.

Recognition of his accomplishments has come in many forms. Fred has received honorary degrees from the University of Chicago (1973), Carnegie Mellon University (1974), Yale University (1981), Wesleyan University (1983), and Harvard University (1993). He is also an honorary fellow of the Royal Statistical Society, an honorary member of the International Statistical Institute, and an elected member/fellow of the American Academy of Arts and Sciences, the American Philosophical Society, the Institute of Medicine, and the National Academy of Sciences. He has served as the Committee of Presidents of Statistical Societies’ R.A. Fisher Lecturer, and has been honored in numerous other ways.

By the age of 89, most people have long since retired and turned to pastoral pursuits. Yet, though Fred has been officially retired for eighteen years, he remains remarkably active. Nonetheless I recall being at least partially surprised when, perusing the table of contents of a 2002 issue of Statistics in Medicine a few years ago, I came across a paper by Lincoln Moses, John Buehler, and guess who [7], on one of Fred’s longstanding research interests, meta-analysis; this was a reminder that we needed to continually update the bibliography that follows until this volume was sent to press!
Fred will remain a role model for statisticians and other scientists whom he has mentored, taught, and otherwise influenced over the years. It is our hope that reprinting a selection of his contributions to statistics in a single volume will help his work similarly to influence new generations of researchers.

References

This bibliography of Frederick Mosteller’s writings generally follows the four categories that Fred has used in his personal bibliography: books, papers, miscellaneous, and reviews. Ordinarily the publications in each category appear in chronological order. We deviate from that ordering to bring together closely related books.

We list Fred’s name as it appeared on the publication, usually “Frederick Mosteller.” For works with other authors, we list the names in the order that they appeared on the publication. For publications by a committee or panel, we list the members.

If, in addition to being one of the editors of a book, Fred is the author or coauthor of a chapter, we list that chapter under the book but not among the papers.

We give any additional information that we have, such as new editions, different printings, and translations. We realize that our information about translations is not complete, nor do we always know of reprinting of papers or chapters.

**BOOKS**


Bibliography

• Appendix IV. Frederick Mosteller. “Charts indicating confidence limits and critical differences between percentages.” pp. 297–301.


  Two parts of this book were published earlier—papers P31 and P32.


  Members of the group: Edwin C. Douglas, Frederick Mosteller, Richard S. Pieters, Donald E. Richmond, Robert E.K. Rourke, George B. Thomas, Jr., and Samuel S. Wilks. These names were inadvertently omitted from this edition.
A revised edition of this volume was published in 1959.


Members of the group: Frederick Mosteller, Richard S. Pieters, Robert E.K. Rourke, George B. Thomas, Jr., and Samuel S. Wilks. B7 and the revised edition of B6 were translated by Professor Marta C. Valincq into Spanish, and published by Comision de Educacion Estadistica del Instituto Interamericano de Estadistica Rosario (Rep. Argentina), 1961.


Extensive revision of B9.


Official textbook for Continental Classroom. Derived from B9 by abridgment and slight rewriting. Also translated into Turkish.


Second Edition of B17. A new chapter dealing with authorship work published from about 1969 to 1983 was added in the second edition. A new, lengthy Analytic Table of Contents replaced the original Table of Contents.


Also translated into Russian, 1975.

Reissue of B20.


These comprise essays by many authors, and the same essay may appear in more than one of the books.


Parts of these books were translated into Japanese, 1979.

**B35** *Statistics by Example: Exploring Data*. (with the assistance of Martha Zelinka)


**B36** *Statistics by Example: Weighing Chances*. (with the assistance of Roger Carlson and Martha Zelinka)


B37 *Statistics by Example: Detecting Patterns.* (with the assistance of Roger Carlson and Martha Zelinka)

B38 *Statistics by Example: Finding Models.* (with the assistance of Roger Carlson and Martha Zelinka)


Also translated into Russian, 2 volumes. Moscow: Statistika Publishers, 1983.


- A Report to the American People and the National Science Board.
- Source Materials.


The second edition has two new chapters—one on life tables, the other on multiple regression.


At the time of his death, William G. Cochran left an almost completed manuscript on observational studies. Lincoln E. Moses and Frederick Mosteller edited and organized the manuscript, and *Planning and Analysis of Observational Studies* is the result.


Members of the Committee: Frederick Mosteller, Chairman. Jacob I. Fabrikant, R.J. Michael Fry, Stephen W. Lagakos, Anthony B. Miller, Eugene L. Saenger, David Schottenfeld, Elizabeth L. Scott, John R. Van Ryzin, and Edward W. Webster; Stephen L. Brown, Staff Officer, Norman Grossblatt, Editor.


Members of the Committee: Frederick Mosteller, Chairman, H. David Banta, Stuart Bondurant, Morris F. Collen, Joanne E. Finley, Barbara J. McNeil, Lawrence C. Morris, Jr., Lincoln E. Moses, Seymour Perry, Dorothy P. Rice, Herman A. Tyroler, and Donald A. Young. Study Staff: Enriqueta C. Bond, Barbara Filner, Caren Carney, Clifford Goodman, Linda DePugh, Naomi Hudson, and Wallace K. Waterfall.


B59 Committee to Review the Adverse Consequences of Pertussis and Rubella Vaccines, Division of Health Promotion and Disease Prevention, Institute of Medicine, Christopher P. Howson, Cynthia J. Howe, and Harvey V. Fineberg, editors. *Adverse Effects of Pertussis and Rubella Vaccines*. Washington, D.C.: National Academy Press, 1991.


PAPERS


P11 Frederick Mosteller and John W. Tukey. “Practical applications of new theory, a review.”


