A Brief History of Statistics in Three and One-Half Chapters: A Review Essay

Stephen E. Fienberg

Statistical Science, Vol. 7, No. 2. (May, 1992), pp. 208-225.

Stable URL:
http://links.jstor.org/sici?sici=0883-4237%28199205%297%3A2%3C208%3AABHOSI%3E2.0.CO%3B2-G

Statistical Science is currently published by Institute of Mathematical Statistics.

Statistical Science
1992, Vol. 7, No. 2, 208–225

A Brief History of Statistics in Three and One-Half Chapters: A Review Essay

Stephen E. Fienberg

EDITOR'S NOTE

This article by Stephen Fienberg reviews the last three and one-half centuries of statistics and probability largely through the author's overview and synthesis of seven recent books on the topic. It is an altering and expansion of an earlier paper with the same title published in *Historical Methods*. We are partly taking this liberty because *Historical Methods* falls well outside the normal reading range of statisticians. An extension particularly worth noting is the introductory timeline on pages 210 and 211. The original version appeared in *Historical Methods* (1991 24 124–135; Heldret Publications, 1319 18th Street, N.W., Washington, D.C. 20036-1802, copyright 1991). Permission to reprint in revised form has been granted by the Helen Dwight Reid Educational Foundation.

The seven books reviewed here are as follows:

LORRAINE J. DASTON (1988). *Classical Probability in the Enlightenment.* Princeton Univ. Press, 423 pp.

GERD GIGERENZER, ZENO SWIJTINK, THEODORE PORTER, LORRAINE DASTON, JOHN BEATTY and LORENZ KRÜGER (1989). *The Empire of Chance. How Probability Changed Science and Everyday Life.* Cambridge Univ. Press, 340 pp.

ANDERS HALD (1990). *A History of Probability and Statistics and Their Applications before 1750.* Wiley, New York, 586 pp.

LORENZ KRÜGER, LORRAINE DASTON and MICHAEL HEIDELBERGER, eds. (1987). *The Probabilistic Revolution, Volume 1: Ideas in History.* MIT Press, 449 pp.

LORENZ KRÜGER, GERD GIGERENZER and MARY S. MORGAN, eds. (1987). *The Probabilistic Revolution, Volume 2: Ideas in the Sciences.* MIT Press, 459 pp.

THEODORE M. PORTER (1986). *The Rise in Statistical Thinking, 1820–1900.* Princeton Univ. Press, 333 pp.

STEPHEN M. STIGLER (1986). *The History of Statistics: The Measurement of Uncertainty before 1900.* Harvard Univ. Press, 410 pp. [Reissued in a paperback edition (1990).]

INTRODUCTION

Writing a history of some aspect of science is often a daunting task, requiring the painstaking reexamination of source materials long forgotten by contemporary scientists and the blending of a knowledge of the substance of the science in question with a broader historical sense. Furthermore, unlike those working in other areas of history, the historian of science is constantly encountering examples of lapses in the scientific etiquette and scholarship of those under study (as Stigler notes in *The History of Statistics*, "Citation was an imperfect art in the eighteenth century," p. 95), as well as instances of Stigler's Law of Eponymy, which has a less modest origin than the name might suggest and states, in its simplest form, that "No scientific discovery is named after its original inventor" (Stigler, 1980). This essay takes its title in part as a play on the titles of two recent books, a highly popular account of the history of physics and astronomy as they relate to the beginning of the universe (Hawking, 1988) and a somewhat less popular work of fictional history (Barnes, 1989), but with few pretensions of emulating the success of either of the authors. It is my intent to provide a brief but accurate overview of selective aspects of the development of the field of statistics, drawing in large part on a septet of recently published books as well as a personal assessment of the books themselves. I write, not as a historian of science, but as a statistician with a strong interest in the history of his own discipline.

A history of a scientific field such as statistics plays a special role for the field itself, helping statisticians to understand some of the origins of their work as well as a sense of what, statistically, discovery is all about. Because statistical thinking infuses so many other scientific fields today, the history of statistics plays an important backdrop to the history of science more broadly. Thus, one might well begin by asking who has actually been writing about the history of statistics?

Stephen E. Fienberg is Professor of Statistics and Law and Vice President for Academic Affairs at York University, 4700 Keele Street, North York, Ontario M3J 1P3, Canada.

In fact, only two of the many authors whose names adorn the covers of the books under review are actually statisticians, Anders Hald and Stephen Stigler. The others are economists (Mary Morgan), historians (John Beatty, Lorraine Daston and Theodore Porter), philosophers of science (Lorenz Krüger, Michael Heidelberger and Zeno Swijtink) and psychologists (Gerd Gigerenzer). The nonstatisticians each have knowledge about statistics, and they may bring to their historical inquiries a methodology that the statistician does not possess. For me, this raises two questions: What kind of training and knowledge should the historian of science have? How much substantive knowledge is necessary? The insight of statisticians into their own field may provide considerable power in the development of a history of the field, but such insights alone do not make for good historians of science. I raise these questions at the outset so that the reader may keep them in mind as I explore the historical developments chronicled in the books under review. Cowan (1987) discusses some of these issues in her review of the books by Porter and Stigler, and I attempt to explicate her views in the Epilogue as well as explain what for me constitutes the basis for a good history of statistics.

As most readers of *Statistical Science* are aware, the standard modern textbook description of statistics is as the science of collecting, organizing, interpreting and reporting data, where the data consist of observations taken in the real world. While much of statistical work is cast in the language of probability, and statistical inference uses probability to express uncertainty, probability theory, for me as a statistician, is primarily a branch of mathematics, whereas statistics is an area of science separate from mathematics. Because of the special links between probability, and statistics, any history of statistics must deal with contributions and developments in the domain of probability. In this historical overview of the field of statistics provided in the present essay, I include developments within the domain of probability, but primarily when the probabilistic developments relate to specific statistical ones.

The history that follows has four identifiable periods: 1660–1750, 1750–1820, 1820–1900 and 1900–1950. In part, the choice of periods is linked to the ideas and exposition in several of the books under review. In particular, the end of each of the first three periods marks the completion of a major probabilistic or statistical development as chronicled by one or more of the authors. Chapter 1 deals with what is best referred to as pre-history, during which many of the basic ideas in classical probability theory were developed but when there was little in the way of development of statistical ideas. Chapter 2 takes up the basic development of formal statistical methods, primarily linked to problems in the physical sciences, especially astronomy, by such eminent scientists as Carl Friedrich Gauss, Pierre Simon Laplace and Adrien Marie Legendre, as well as the Reverend Thomas Bayes. The rise of statistical theory linked to the social and biological sciences is the focus of Chapter 3, which covers the period from 1820 through 1900. Stigler argues that "the infant discipline of [statistics] may be said to have arrived" in 1900. Modern statistical methodology and theory as we know it is a twentieth century creation, even though it is rooted in these earlier developments. Thus the reader might expect a full chapter treatment of the developments over the past 90 years. During the preparation of a pair of volumes of pioneering statistical papers of the past century, entitled *Breakthroughs in Statistics* (Johnson and Kotz, 1992), the editors polled a group of prominent statisticians to determine which contributions to include. They note in their preface:

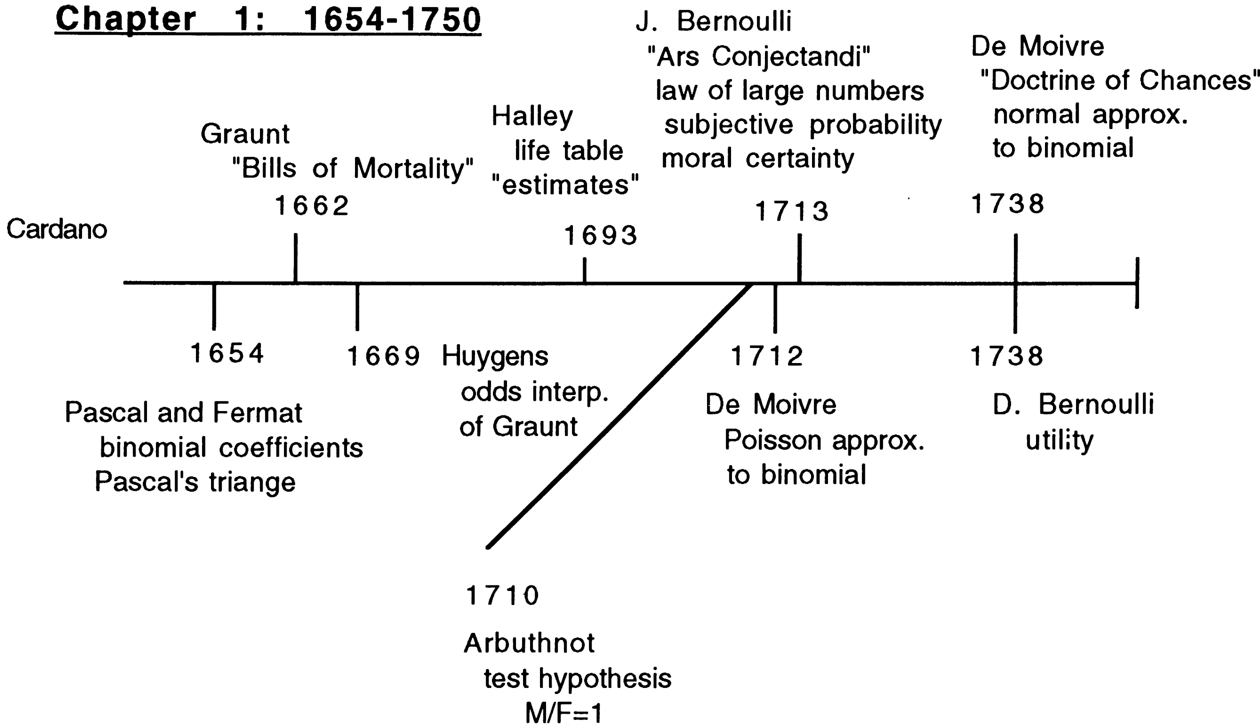
> There was remarkable near-unanimity recommending the early work of K. Pearson, "Student," R. A. Fisher, and J. Neyman and E. S. Pearson up to 1936. For the years following 1940, opinions became more diverse, although some contributions, such as Wald (1945), were cited by quite large numbers of respondents. After 1960, opinions became sharply divergent. (p. x)

My overview of the history of statistics devotes only one-half a chapter to this period ending somewhere around 1950, mainly because from this and other evidence it is clear to me that the accomplishments of the past four decades are still too fresh for us to sort out which will ultimately be viewed as lasting by those in future generations. To close my history and review, I offer an Epilogue that considers some overarching issues and returns explicitly to the books on which I draw throughout.

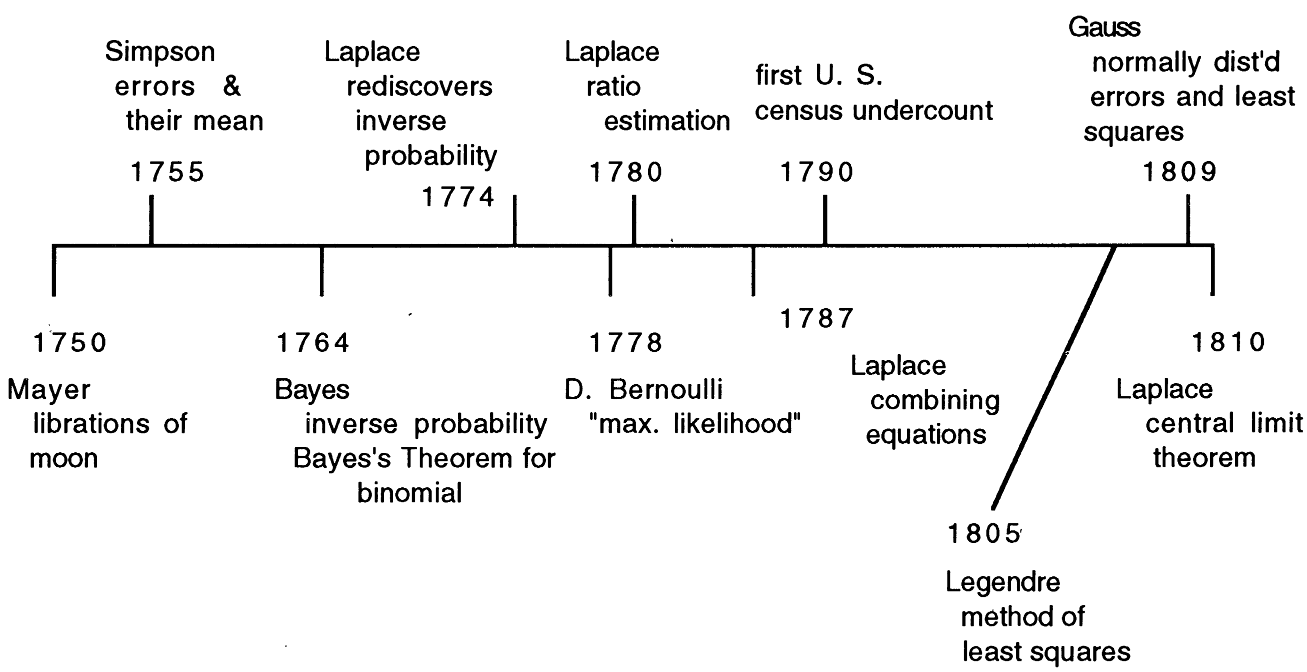
The seven books under review deal with overlapping periods and overlapping material, often leaving the reader with quite different impressions of the relative importance of some contributions to the development of the field of statistics. Because the writing of such histories of science involves considerable subjectivity, these differences in perspective make each separate volume of interest, and thus the seven books often collectively offer a better picture than a single authoritative volume. Each of the books has its virtues (and often its flaws), and in the Epilogue I will supply my assessment of those under review. Nonetheless, I note at the outset that all are valuable additions to my personal bookshelf and belong in any major library.

Before turning to my abbreviated history, let me give some indication of the coverage of each of the seven books from the perspective of my three and one-half chapters rather than simply describing the contents of individual volumes. Hald describes the prehistorical period up through 1750, as does half of the first chapter of *The Empire of Chance*. In *Classical*

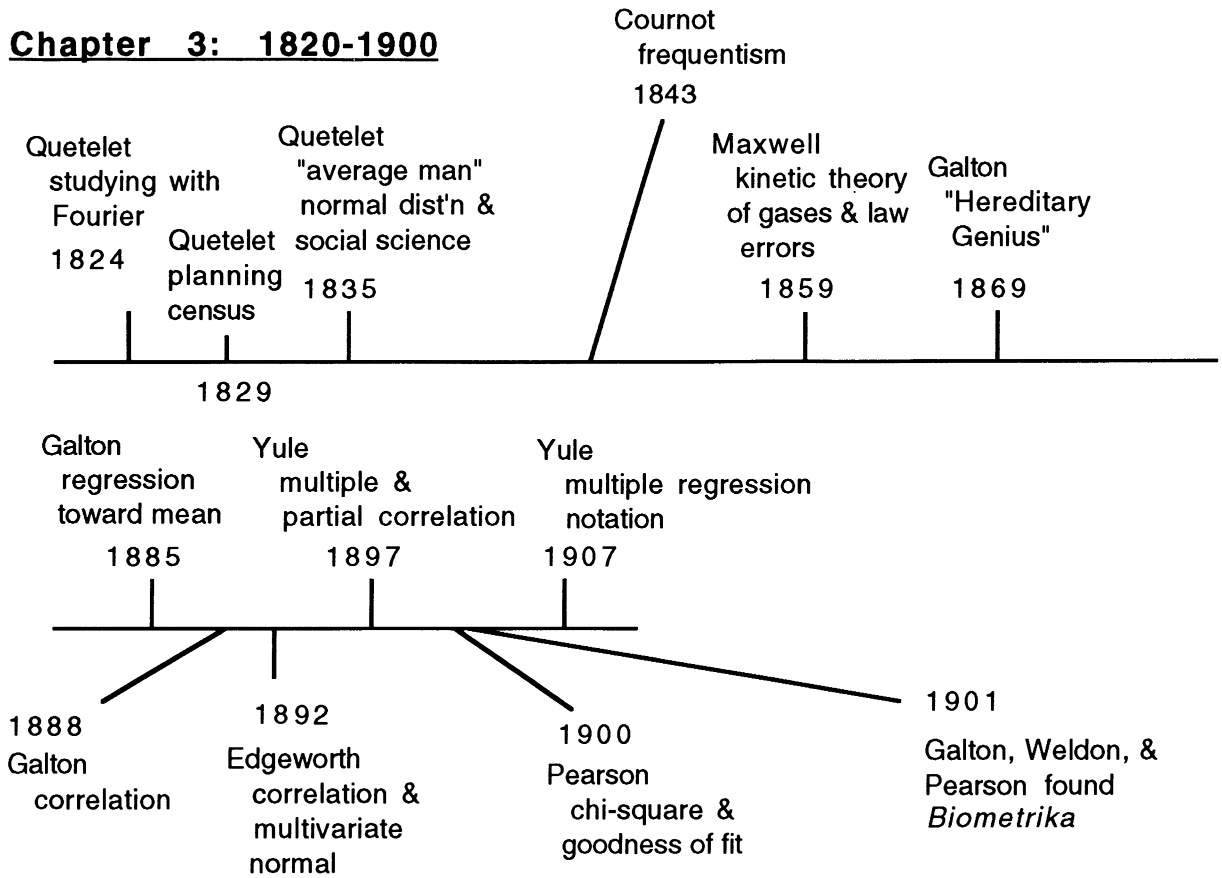
Chapter 1: 1654-1750



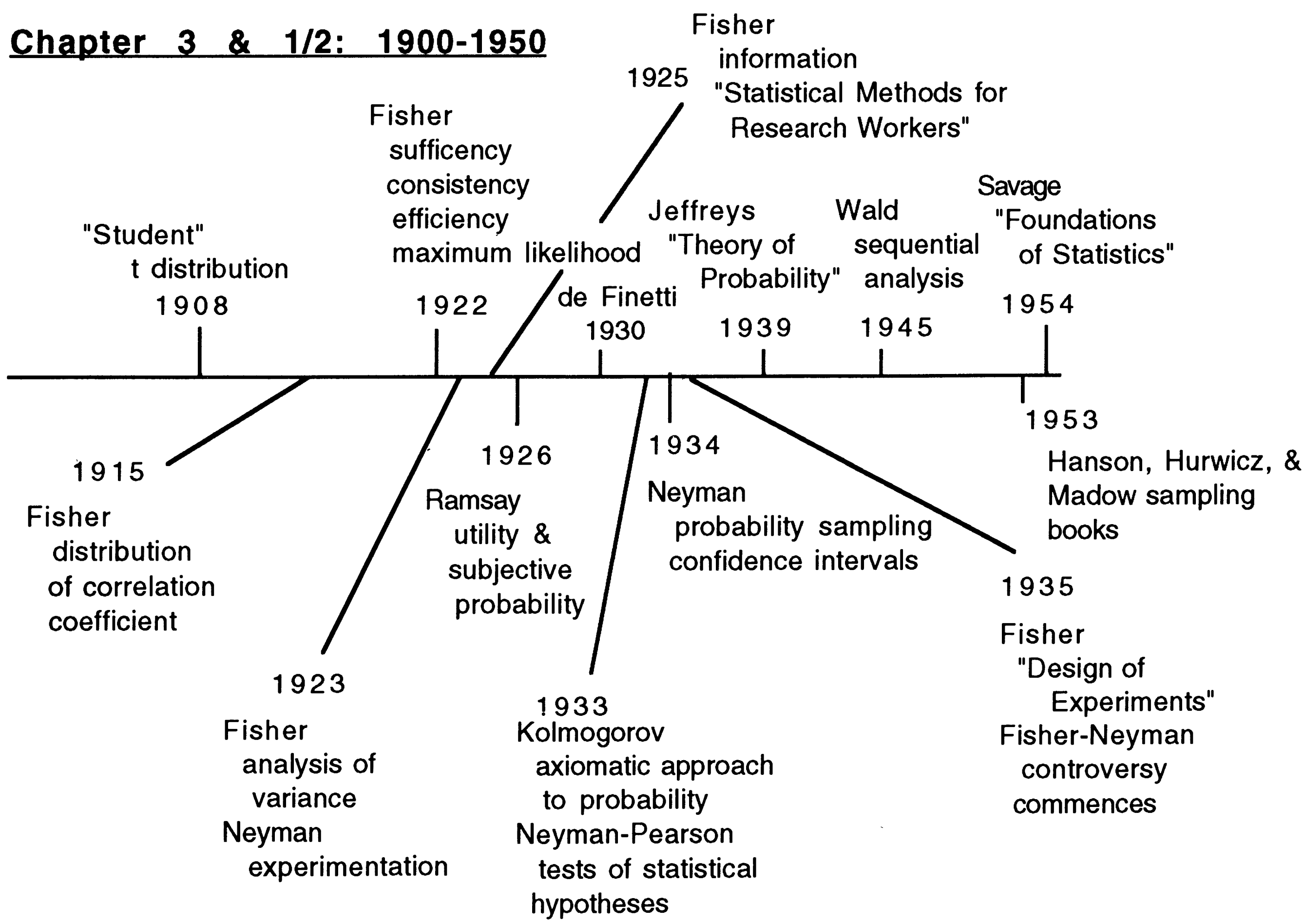
Chapter 2: 1750-1820



Chapter 3: 1820-1900



Chapter 3 & 1/2: 1900-1950



Probability in the Enlightenment, Daston covers both the prehistorical period and the next century, but mainly with a focus on the development of probability and its interpretation. She stresses the broader historical and philosophical perspectives, while Hald focuses on the more narrow technical one. Stigler also describes in detail the key contributions of Jakob Bernoulli and Abraham De Moivre, and, in some ways, the work of these two figures represents the boundary between pre-history and history. Stigler's book is my primary source for the period from 1750 to 1820, although the second half of the opening chapter of *The Empire of Chance*, parts of Daston and some individual chapters in Volume 1 of *The Probabilistic Revolution* are also useful sources. Porter and Stigler focus on the period from 1820 to 1900, with related discussions in the second chapter of *The Empire of Chance*, and selected material from Volumes 1 and 2 of *The Probabilistic Revolution*. Developments in the twentieth century are the primary focus of *The Empire of Chance*, as well as of several chapters from Volumes 1 and 2 of *The Probabilistic Revolution.*

Two somewhat older histories are worth mentioning here as well, both covering approximately the same period – the mid-nineteenth-century history by Isaac Todhunter (1865) and the early-twentieth-century lectures by Karl Pearson (1978), which were posthumously published 42 years after his death. Todhunter was long viewed as the unquestioned authority for the contributions from the periods covered by my first two chapters, but, with the current revival of interest in the history of probability and statistics, most knowledgeable readers recognize that Todhunter lacked a broad statistical perspective and that he often "meticulously reports proofs of many results which are of little interest today; conversely, he omits proofs of results of great importance" (p. 9 in Hald's book). Furthermore, Todhunter did not really set contributions into their broader scientific context. Pearson, on the other hand, reveled in the contextual material and often goes off on interesting digressions into the lives of the statisticians and methematicians whose work he describes. Unlike Todhunter who is known primarily for his historical volumes, Pearson was a distinguished statistician, and his views of the historical contributions of others are seen through the filter of his own work and ideas (which were many). His volume does not contain errors of fact, many of which are corrected in the volumes under review here.

CHAPTER 1: PRE-HISTORY THROUGH 1750

Where shall the history of statistics begin? This is the title of a brief 1960 article by Maurice Kendall (1960), who argues that "there are dangers in pursuing the roots of a subject down to its slenderest fibrils." He concludes "that statistics in any sense akin to our own cannot be traced back before about A.D. 1660" and points to the work of John Graunt as an appropriate starting point. Stigler, too, cautions us not to begin the history too early: "If all sciences require measurement – and statistics is the logic of measurement – it follows that the history of statistics can encompass the history of all science." He suggests that it is reasonable to restrict the history to the development of probability-based statistical methods. In this spirit, we label the period of the development of probability and the exposition of nonprobabilistic methods of data analysis as pre-history, and out history proper begins around 1750 (Chapter 2).

Both Daston and Hald note, as have many others before, that probability theory was originally inspired in large part by games of chance, and they recognize the formative role of the Italian mathematician, Girolamo Cardano, in developing probabilistic ideas in the sixteenth century. Hald then distinguishes three periods from 1660 up through 1750 during which crucial developments in the history of probability occurred.

From 1654 to 1665, we have the correspondence of Blaise Pascal and Pierre de Fermat and the development of results on the binomial, including binomial coefficients and Pascal's triangle. This period represents what the philosopher, Ian Hacking, has referred to as "The Emergence of Probability." Hacking (1975) asks why the time was ripe for the emergence of what is our current concept of probability at about 1660, and he reviews many of the contributions in the ensuing half century. Shafer (1990) reminds us that Pascal and Fermat did not use the word probability in their 1654 letters; rather they were thinking solely of fair prices.

Then from 1708 to 1718, after about a 50-year period of stagnation, we see a flurry of activity by Pierre Rémond de Montmort, the Bernoullis and De Moivre, during which the elementary results were molded into a coherent theory of probability. Of special importance for statistics during this period was the work of Jakob (also known as James or Jacques) Bernoulli, published posthumously in 1713, in which he not only provides a version of the law of large numbers but he also discusses the concept of probability, introducing perhaps for the first time the subjective notion that probability is personal and varies with an individual's knowledge. Stigler points out, however, that Bernoulli did more than prove the law of large numbers. He also attempted to show how to quantify the number of observations required for an observed proportion to fall within a given amount of the true proportion with "moral certainty," that is, with a chance exceeding 1000/1001 (this being the illustrative example that ends Bernoulli's book).

In the final period of probabilistic pre-history, from

FIG. 1. *Jacques Bernoulli (1654–1705).*

1718 to 1738, we see consolidation and extension. This was the period when De Moivre developed the normal approximation to the binomial, although, as Stigler notes, De Moivre thought of the normal curve more as a calculating device than as a continuous probability distribution in its own right. These results, along with the method of generating functions, were published in the second edition of this book, *The Doctrine of Chances* (De Moivre, 1718). De Moivre's derivation of the normal approximation was in many ways his attempt to improve upon Bernoulli's bound for the sample size required to achieve moral certainty, and he recognized the importance of $\sqrt{n}$, where n is the number of trials, as giving the scale on which deviations from the center of the distribution should be judged. Yet De Moivre stopped short of going beyond the binomial to a broader form of central limit theorem as we know it today, and he also stopped short of using the approximation to make inverse inferences about the binomial parameter, p.

While De Moivre's work was widely circulated and brought to a successful completion the development of what we now call classical probability theory, the period up to 1750 did not produce a theory of statistics,

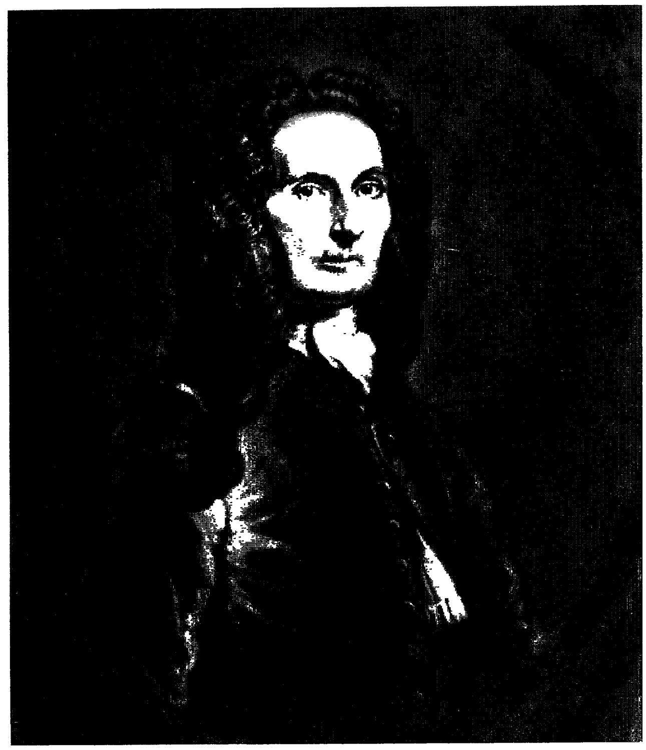


FIG. 2. *Abraham De Moivre (1667–1754), as portrayed in 1736.*

involving the application of probabilistic ideas to data. Rather the contributions to statistics prior to 1750 consisted mainly of examples of data analysis, often without the use of any explicit probabilistic ideas or the notion of uncertainty. The prototypic model of descriptive statistical analysis is that found in John Graunt's 1662 work on the *Bills of Mortality.* In it, he gave birth to a number of data analytical approaches that included (i) an examination of the trustworthiness of the data in the "bills" published over a 60-year period, (ii) a careful description of the mortality due to the plague, including a wonderful illustration of "imputation" of omitted deaths in certain years, (iii) a detailed description and analysis of the sex ratio (the ratio of the number of males to number of females) of births and deaths in London and Romsey, a country parish in Hampshire—there appears to be amazing stability in the ratios, and (iv) the development of what is at least in part an empirically based life table in order to answer questions on how many men of fighting age there were in London. Hald observes that Graunt's "statistical methods are seldom pronounced directly but are to be found in his examples" (p. 86).

Graunt's first edition was only 85 pages in length, but it stimulated work by a substantial number of others on these and other topics, and many of these individuals introduced the use of probabilistic tools into their analyses. For example, while Graunt merely speculated on the reasons for the stability of sex ratios, some 50 years later, John Arbuthnot actually attempted to test the hypothesis that the ratio is 1, using a binomial model and what is in effect a simple sign test (see p. 278 in Hald's book). This work was then followed up by Nicholas Bernoulli, who explored further the appropriateness of the binomial model for this problem. Similarly, the Dutch mathematician and physicist Christian Huygens, in correspondence with his brother, followed up on Graunt's life table in 1669, giving it a probabilistic interpretation through the use of odds, and he calculated both the expected and median lifetime, noting how the two concepts differed (see Chapter 8 in Hald's book).

Although the origin of classical probability theory was closely linked to gambling, it is important to recognize that many of the actors described in this chapter were deeply religious men (e.g., Blaise Pascal), whose work on probabilistic problems was driven in large measure by their theological concerns about proof to support the existence of God or the role of God in certain problems, such as what we now refer to as Pascal's wager concerning the existence of God. (For further details, see the discussion in Daston's book on pp. 60–63 and in Hald's book on p. 64.) There is also Jakob Bernoulli's discussion of moral certainty in *Ars Conjectandi* and Arbuthnot's 1712 statistical argument for devine providence. Virtually all of these authors believed that the world was deterministic – God leaves nothing to chance – yet they focused on probability to describe both games of chance and issues of morality and theology. Daston (Chapters 2 and 6) pursues this moral-theological theme and its link to expectations, and she includes a detailed discussion of the seminal work of Daniel Bernoulli (Jakob's nephew), who, in a 1738 paper, introduced the general idea of the modern notion of utility. She also describes how various authors pursue the theme linking statistical arguments to moral issues over the next 200 years. The ideas on utility of Daniel Bernoulli ultimately had a major impact on the work by Frank Ramsay in the 1920s and serve as a direct underpinning for modern statistical decision theory. See Savage (1954) as well as the brief discussion of this topic in Chapter 3½ below.

CHAPTER 2: THE INTRODUCTION OF INFERENCE AND THE BEGINNING OF MATHEMATICAL STATISTICS FROM 1750 TO 1820

There are two intertwined strands of statistical activity in the period beginning at about 1750 that were brought together at about 1820 in what Stigler refers to as the Gauss–Laplace synthesis and thereby set the foundations for what we have come to know as mathematical statistics. The first strand involves the development of the method of least squares for the estimation of unknown coefficients in linear equations (they were not quite the linear models we now speak about); the second strand deals with the development of probability based inferences, growing out of the work of Bernoulli and De Moivre. We pick up the story of the development of inference first.

As we left off our history with De Moivre, he had written about the normal approximation to the binomial distribution, and had stopped just short of using this approximation to make inferences about the binomial parameter, p. Chronologically, the developments over the next 50 years begin with two papers by Englishmen: a 1755 paper of Thomas Simpson, which was followed quickly by a posthumously published 1764 paper by the Reverend Thomas Bayes. In this now celebrated essay, Bayes developed and then utilized the inverse probability argument that later became associated with his name, Bayes's Theorem. He used the inverse probability approach to answer the problem that had stumped both Bernoulli and De Moivre – that is, he showed how to reason probabilistically about the binomial parameter, p, in light of the data. Then came an amazing series of results by the French mathematician and astronomer Pierre Simon Laplace, in which he first approached the binomial problem using an inverse probability argument similar to that of Bayes, and then went on to expand its use for a variety of other distributions. All of this ultimately led Laplace in 1810 to a clear formulation of the central limit theorem that could justify the use of the normal distri-

FIG. 3. *Pierre Simon Laplace (1749–1827), as portrayed in 1799.*

bution to approximate the distribution of sums (or averages) from virtually any probability distribution.

Stigler presents these statistical accomplishments out of chronological order. After describing Simpson's advances and Bayes's recognition of them, he skips ahead to Laplace, foreshadowing the story that is to come with the statement:

> Simpson had seen that the concept of error distributions permitted a back-door access to the measurement of uncertainty. Later Laplace was to slip in this same back door and come around to open the front (only to find that Bayes's key was already in the lock). (p. 95)

Stigler painstakingly analyzes Laplace's (re)discovery of Bayes's Theorem and notes the slow progress he made at the beginning. But once he understood the power of the inverse probability approach, Laplace was able to move ahead rapidly to solve the problems that had thwarted Bayes. Thus, the contribution of Bayes, amazing though it was, appears to have had little direct influence on the development of the theory of mathematical statistics in this crucial period. To supplement this discussion in Stigler's book, we refer the reader to Stigler (1986), which provides a guided tour and an English translation of Laplace's 1774 memoir.

The other key strand of developments during this period relates to the evolution of a general statistical approach for combining observations that culminated in the method of least squares. Behind virtually all of these developments was a practical series of data problems in astronomy involving observations on planetary positions, orbits and geodesic arcs. At the beginning of these developments, we have the 1750 study of Johann Tobias Mayer on the librations of the moon, in which he proposed an ingenious method for solving 27 equations in 3 unknowns through the creation of 3 equations formed by summing the 9 equations in each of 3 carefully constructed groups. Stigler contrasts Mayer's accomplishment with the 1749 failure of the distinguished mathematician Leonhard Euler to solve essentially the same problem—his explanation is Mayer's empirically based "conviction that a combination of observations increased the accuracy of the result in proportion to the number of equations combined" (p. 28). With an intervening contribution by Ruggiero Giuseppe Boscovich, who introduced the notion of a principle for combining, we come again to Laplace, who, in 1787, extended Mayer's approach by reducing a set of linear equations by combining them together in several different ways. The problem was that Laplace's approach was still ad hoc; that is, it did not involve an explicit mathematical criterion relating to the statistical aspects of the situation. Thus, different people trying to emulate Laplace's approach on a new set of equations could well produce different answers. Legendre resolved all of this with his 1805 development of the method of least squares, which invoked a statistical minimization principle and created a unique set of reduced equations, which could then provide the estimates of the unknown coefficients based on the combined observations.

But how does the method of least squares fit with the quantification of uncertainty that had been developed in the other strand of research? This crucial link was provided by the great German mathematician Carl Friedrich Gauss, who, in 1809, used a circular argument to justify the use of normally distributed error terms for systems of linear equations and then went on to show that maximizing the posterior distribution of the errors was equivalent to using the method of least squares. The paper had an immediate impact on Laplace, who, in 1810, recognized that the normal error term could be justified by his own results on the central limit theorem. This fusing of the two lines of development into a single, powerful statistical approach applicable to a broad class of physical problems is what Stigler refers to as the Gauss–Laplace synthesis. It brings to a close our second chapter and leaves us with a question asked by Stigler in his introduction: If the Gauss–Laplace synthesis essentially produced the methodology of regression analysis, why did another 75 years go by before Galton invented regression?

A quick glance back over this chapter might leave the reader with the impression that only one approach to statistical inference (i.e., the inductive process of generalizing from a sample of observations to population quantities using a probabilistic argument) appeared during this period—that is, the inverse probability method we now associate with the name of Bayes. This is not quite accurate because, in 1778, Daniel Bernoulli published a memoir on his choice of error curves in which he proposed with little justification a method of estimation that we can now recognize as what Stigler refers to as "an adumbration of maximum likelihood." Bernoulli's method was harshly criticized as arbitrary by Euler in an appended commentary, and it was quickly overshadowed by Laplace's work. Thus, Bernoulli's contributions played no known role in the later rediscovery of maximum likelihood in the twentieth century.

CHAPTER 3: THE SOCIALIZATION OF STATISTICS AND THE DEVELOPMENT OF CORRELATION AND STATISTICAL MODELS FROM 1820 TO 1900

As we open this chapter in 1824, we find the Belgian astronomer and mathematician Adolphe Quetelet in Paris learning about probability and statistics from Joseph Fourier, who had in turn learned these ideas from Laplace. Quetelet returned to Brussels and immediately attempted to apply Laplace's 1780 method of

FIG. 4. *Adolphe Quetelet (1796–1874), as portrayed in 1822.*

ratio estimation using birth and death rates as part of the analysis of past census data and in planning for an 1829 census. Although attracted by the power of this method, Quetelet finally backed away from its full-scale implementation because of the quantities that required measurement and the differences among groups in the population for which he needed to account. Henceforth, he concentrated on the macroanalysis of social data using statistical methods learned from Laplace and others, but he eschewed the probabilistic analysis of uncertainty in estimates and measurements that was inherent in the method of Laplace that he had earlier recommended.

Quetelet may well be considered the father of quantitative social science, and his two major contributions toward the statistical analysis of social data, the concept of "the average man" in his 1835 book and the later fitting of the normal distribution to several social science data sets, coupled with the interpretation of the stability of the corresponding social phenomena, dominated the field for many years. Nonetheless, Stigler argues that, despite these and other contributions, Quetelet's attempts to import the methods of statistics into the realm of social science were failures because he resisted turning to the individual level and the study of relationships among variables in order to explain the heterogeneity that he found blocking the use of the known statistical methods of his day.

Porter, on the other hand, focuses on Quetelet's view of statistical social science as social physics and sees less failure and much greater influence, extending to the development of the kinetic theory of gases. He traces a path from the pioneering 1859 paper by the Scottish physicist, James Clerk Maxwell, back to Quetelet through a review of one of his books by John Herschel and through Henry Thomas Buckle's *History of Civilization in England*, which presented an exaggerated account of Quetelet's findings of statistical regularities. Porter also finds analogies between probabilistic molecular behavior and "the statistical behavior of a free society" in later work of Maxwell and Boltzmann.

But, in the words of Porter, it was during this period that we see the rise in statistical thinking, as others venture forward along paths that Quetelet was unable to follow. Contributors to the development of statistical thinking during this period, especially with an orientation towards its application in the social and behavioral sciences, include Poisson, I. J. Bienaymé (Heyde and Seneta 1977), Wilhelm Lexis, William Farr, Auguste Comte (in a negative fashion), Antoine Augustin Cournot and Gustov Theodor Fechner (followed somewhat later by Hermann Ebbinghaus). Daston singles out Cournot's work as marking "the advent of a new interpretation of mathematical probability exclusively in terms of objective frequencies. [He] was also among the first to recognize that the classical approach to probability had been an *interpretation*, distinct from mathematical probability per se" (p. 224). Porter also suggests that it was with the rise of social statistics and the idea of statistical regularities that we see a shift from the subjective notion of probability inherent in the work of Laplace to the frequency notion which was to fully emerge only in the twentieth century. Porter also emphasizes the role of the positivist empirical philosophers, John Stuart Mill, Richard Leslie Ellis and Jakob Friedrick Fries (see also the discussion in Shafer, 1990).

Daston's chapter in Volume 1 of *The Probabilistic Revolution* provides support for this view on the rise of frequentism. On the other hand, the essay by Kamlah offers a rather different explanation of the shift to the frequency interpretation of probability, which totally ignores all of the nineteenth century contributions to statistics (as distinct from contributions to probability theory viewed as a branch of mathematics).

Yet despite all of this intellectual attention towards statistics, as the 1870s approached, there had not been a major breakthrough in the development of statistical methodology that could rank with the development of least squares and its link to the normal distribution and the central limit theorem. The period from 1880 to 1900 saw a notable change in the pace of statistical developments, especially in England as a result of the contributions and leadership of Francis Galton, Francis Ysidro Edgeworth, Karl Pearson and George Udny Yule. There was also the related work of others, such as the logician John Venn, but the lines of work

FIG. 5. *Francis Galton (1822–1911).*

carried out by this quartet of scientists seem to be the ones most critical to the development of a basic approach to statistical methodology.

Stigler describes Galton as a "romantic figure in the history of statistics, perhaps the last of the gentleman scientists" (p. 266). Before turning to statistics, he explored Africa and studied meteorological problems and heredity. In fact, it was through his study of *Hereditary Genius* (1869) that we get the first glimmerings of Galton's own ideas on regression. Over the next 16 years, he sharpened these ideas, and, in his 1885 presidential address to the anthropological section of the British Association for the Advancement of Science, Galton presented, in a table and diagram reproduced by Stigler, the first statistical description of the phenomenon of regression and its link to normal distributions. This was done in the context of an empirical example of the regression of children's heights on the average of their parents' heights. A few years later, in 1888, Galton formulated the related concept of correlation. Porter and Stigler present complementary detailed descriptions of these developments, and Stigler (1989) supplements them with excerpts of previously unpublished correspondence and a reprint of Galton's own 1890 account entitled "Kinship and Correlation."

Galton's ideas on regression and correlation were quickly picked up and extended by others. Most notable in this regard was the work of Edgeworth, who, during the 1880s, tried to take the ideas of statistical analysis developed decades earlier in astronomy and geodesy, and apply them to social and economic statistics. Because of his knowledge of least squares and inverse probability, Edgeworth was able to link Galton's concepts of regression and correlation to these earlier methods, and he did so directly in the context of multivariate normal distributions, for which he introduced the equivalent of modern notation for the correlation matrix.

Then came the contributions of Pearson and Yule, a description of which forms the final chapter of Stigler's book. Stigler attempts to provide direct evidence of the influence of Edgeworth on the evolution of Karl Pearson's ideas about statistical methods. In the early 1890's, Pearson began to write about methods for the analysis of skew curves, which led to the formulation of the Pearson family of curves. Pearson also lectured on the history of statistical methods during this period (Pearson, 1978). During some of these lectures, Pearson introduced Yule to the study of skew curves and correlation. According to Stigler, it was Yule who provided statistics with its second synthesis, one that reconciled the developments in the theory of correlation and regression with the earlier methodology of least squares and the theory of errors. In an 1897 paper, Yule presented this synthesis and introduced the concepts of multiple and partial correlation. A decade later, he introduced the modern notation for regression analysis that is still in widespread use today. Then the synthesis was truly complete.

The nineteenth century ended with other new contributions by Pearson and his collaborators in the biometrical school, for example, his chi-square test for goodness of fit in contingency tables, and his founding of the first independent methodologically oriented statistics journal, *Biometrika*, with Galton and W. F. R. Weldon (with funding from Galton). For Stigler, however, it is not the varied contributions of Pearson but the second synthesis of Yule that provides the watershed for the creation of the modern field of statistics:

> The conceptual triumphs of the nineteenth century had been the product of many minds working on problems in many fields, and one of the most striking of their accomplishments was the creation of a new discipline. Before 1900 we see many

> scientists of different fields developing and using techniques we now recognize as belonging to modern statistics. After 1900 we begin to see identifiable statisticians developing such techniques into a unified logic of empirical science that goes far beyond its component parts. There was no sharp moment of birth; but with Pearson and Yule and the growing number of students in Pearson's laboratory, the infant discipline may be said to have arrived. And that infant was to find no shortage of challenges. (p. 361)

Moreover, as Hacking (1990a) describes the situation: "By the end of the century chance had attained the respectability of a Victorian valet, ready to be the logical servant of the natural, biological and social sciences" (p. 2).

CHAPTER 3½: THE MODERN STATISTICAL ERA FROM 1900 TO 1950

And so the rest is history, or so the saying goes. But is it? The field of statistics as we know it today emerged only in the twentieth century, building on the Gauss–Laplace synthesis described in Chapter 2 and the treatment of regression by Galton, Edgeworth, Pearson and Yule described in Chapter 3.

It is only after the turn of the century that we find the rise of theories of statistical inference, using probabilistic notions in a systematic way to gather, analyze and summarize scientific data. Chapter 3 of *The Empire of Chance* captures much of this development, telling part of the story of the English statistician Ronald A. Fisher (1890–1962) and his development of the notions of a statistical model, sufficiency, likelihood, the statistical concept of randomization, the theory of experimental design and the method of the analysis of variance. To many statisticians, Fisher is the greatest statistician of the century and these contributions, most of which came in a relatively short time span in the 1920s, permanently altered the course of statistical development. For other evaluations of the contributions of Fisher, see Barnard (1990), Fienberg and Hinkley (1980), and Savage (1976). Fisher's collected works, consisting of 140 papers on genetics, 129 on statistics and 16 on other topics (in addition to various reviews) have been published in a six-volume set, and his four books continue in print. On the occasion of the 100th anniversary of his birth in 1990, Oxford University Press reissued his three statistical books in a special single volume.

Fisher's work built on that of the nineteenth century and especially on papers by Edgeworth, Pearson and Yule, as well as those by William S. Gosset, who worked at Guinness Brewery in Dublin and interacted extensively with Karl Pearson (see Pearson, 1990, and Box, 1987). Because of restrictions on publication by

FIG. 6. *Ronald A. Fisher (1890–1962) at his desk calculator, Whittingham Lodge, 1952.*

Guinness, Gosset published under the pseudonym "Student," a name still associated with the t-distribution he invented. Fisher shared Pearson's enthusiasm for Charles Darwin's theory of the origin of species by natural selection, and he pursued a highly creative career in genetics as well as his career as a statistician. Although Pearson published Fisher's 1915 paper on the distribution of the sample correlation coefficient in *Biometrika*, the two men quickly had a falling out and engaged in public debates and acrimonious printed exchanges until Pearson's death. Gosset was friendly with both Fisher and Pearson, and often played a mediating role in the clashes between these two statistical titans, although he too had his own disputes with Fisher (see the paper by Picard in Fienberg and Hinkley, 1980, and Egon Pearson's 1990 biography of Gosset).

Fisher entered Cambridge in 1909 and published his first paper while still an undergraduate in 1912. In 1919, he accepted the newly created position of statistician at Rothamsted Experimental Station. It was while at Rothamsted that Fisher's early statistical ideas were shaped and focussed. Fisher's statistical ideas had an immediate impact on the scientific analysis of data, for example, the design of experiments and the analysis of variance, but some of the concepts he introduced sparked considerable controversy. In particular, his formulation of *tests of significance* and his *fiducial* approach to probabilistically based interval estimation produced the greatest statistical controversy of the first half of the century. Fisher's fiducial method provided a method of inverting probability statements about observations given the values of parameters into probability statements about parameters given the observations without the use of Bayes's Theorem, which provides a mechanism for such an inversion. Zabell (1989) provides an account of Fisher's somewhat myopic views on the decline of inverse probability during the latter half of the nineteenth century.

Jerzy Neyman, a Polish statistician, came to England in 1925 to work in Karl Pearson's laboratory. Neyman had already written, in 1923, a remarkably original doctoral thesis on the application of statistics to agricultural experimentation, in which he explicitly used the notion of hypothetical responses corresponding to what would have been observed had the treatment allocation been different (see the 1990 *Statistical Science* translation of this part of his dissertation with accompanying commentary). In London, Neyman struck up a collaboration with Pearson's son, Egon, that led to the theory of hypothesis testing, which they claimed improved upon Fisher's significance testing approach by explicitly recognizing the role of rival hypotheses. In one of the most important papers of the century, Neyman (1934) laid the foundation for the statistical theory of sampling *and* gave the first description of the *confidence* method of interval estimation based on an infinite sequence of repeated samples. Neyman, separately and in collaboration with Pearson, then elaborated on the method of confidence intervals linking it to their theory of testing hypotheses. Hacking (1990a) and others have argued that E. B. Wilson gave the rationale for confidence interval statements prior to Neyman, and that Wilson's work had roots in an even earlier description due to C. S. Peirce in the nineteenth century. The common attribution of the ideas to Neyman may be another instance of a variant on Stigler's Law of Eponymy.

While Fisher was relatively accepting of the early papers by Neyman and Pearson, ultimately he argued vigorously against their embellishments. Fisher actually prepared a positive referee's report on their 1933 paper that appeared in the *Philosophical Transactions of the Royal Society* (Reid, 1982, pp. 103–104), and he made complimentary comments on Neyman's 1934 paper when it was presented at a meeting of the Royal Statistical Society. The dispute between Fisher and Neyman and Pearson seemed to develop out of a critical exchange over a central point in a 1935 paper by Neyman in the *Journal of the Royal Statistical Society* on statistical problems in agricultural experimentation. The rhetorical aspects of this dispute were then carried over to the basic issues of inference on which they differed. Fisher's debate with Neyman and Pearson on issues involving testing and interval estimation was replete with rhetorical flourishes. As Gigerenzer et al. note in *The Empire of Chance*:

> Fisher never perceived the emerging Neyman–Pearson theory as correcting and improving on his own work on tests of significance. Right up to his death in 1962 he rejected the key concepts of the Neyman–Pearson theory, such as "errors of the second kind," "repeated sampling from finite populations," and "inductive behavior." His recurring reproach was that Neyman and Pearson were mere mathematicians without experience in the natural sciences, and that their work reflected this insulation from all living contact with real scientific problems. (p. 98)

Over time, the flaws of Fisher's fiducial method became apparent to many statisticians, and, as is described in *The Empire of Chance*, a somewhat curious hybrid of his approach to tests of significance and the Neyman–Pearson theory quickly spread to various fields of application. In 1937, Neyman moved to the United States and helped to stimulate the adaptation and development of his ideas on sampling to large-scale national sample surveys under government auspices. (We return to the impact of his work on sampling in the Epilogue, below.) Neyman was a key figure in the rise of mathematical statistics in the United States which occurred in the late 1930s and 1940s, along with Harold Hotelling, Abraham Wald and Samuel Wilks.

A curious omission in Gigerenzer et al.'s description of "the inference experts" is one of the emergence of the Bayesian or subjectivist school of statistical inference, which occurred at approximately the same time as the development of the Neyman–Pearson theory. In the mid 1920s, Frank Ramsay, reacting to ideas of others such as John Maynard Keynes, set out an approach to probability based on personal degrees of belief and linked these to the notion of utility, drawing on the original formulation of Daniel Bernoulli back in 1738. This subjective perspective was independently justified by Bruno de Finetti (1930, 1937) in terms of coherence or consistency and ultimately synthesized with further technical work on utility by L. Jimmie Savage (1954) decades later. Coupled with Harold Jeffreys's (1939) systematic treatment of the use of Bayesian methods in statistical inference (albeit from a somewhat objective perspective that is decidedly not decisive-theoretic), these authors laid the foundation for the Bayesian revival that has occurred over the past three decades. For more on Jeffreys' contributions, see Lindley (1991) and the related articles in the Spring 1991 issue of *Chance*.

In passing, Gigerenzer et al. also duly note the key contribution to probability theory in the early 1930s by the Russian mathematician, A. N. Kolmogorov. He laid down an axiomatic approach to the theory of probability drawing on the mathematical fields of set theory and the theory of functions which launched probability as a separately identifiable subfield of mathematics. The Kolmogorov axioms and set-theoretic approach provided a mathematical foundation to the earlier classical theory of probability and provided a powerful basis for the proof of mathematical results in statistical theory. For a more personal recollection of Kolmogorov and the impact of his work, see Kendall (1991).

Gigerenzer et al. end their description of "the inference experts" with two sections on the statistics profession which describe the specialization of statistical knowledge and its institutionalization, especially in the form of statistical laboratories and departments of statistics in universities. The earliest of these was Karl Pearson's Biometric Laboratory at University College, London, which dates back to 1895. [Gigerenzer et al. give the date as 1904. That was the year when Galton established the Eugenics Record Office. Two years later it became known as the Eugenics Laboratory when Galton turned it over to Pearson to operate. The two operations were officially merged in 1911 to form the Department of Applied Statistics with Pearson as its first professor (Walker, 1978).] The Biometric Laboratory was followed, in 1932, by the Statistical Laboratory at Iowa State College, founded by George Snedecor, and the creation of a large number of departments across the United States, separate from mathematics, over the next 35 years. The special role of statistics in aid of the allied effort in World War II served in both the United States and in Great Britain as a stimulus to the development and expansion of statistics as a field and to the creation of separate departments in universities. All of this is duly noted in *The Empire of Chance*.

The post-war spurt of publications in statistical theory and methods and the creation of new statistical units in a number of American universities in the late 1940's brings to an end the developments in the history of probability and statistics as they are chronicled in the volumes under review. Although *The Empire of Chance* does describe subsequent applications in various domains, it does not do so in any systematic fashion, nor does it link these developments to the further development and specialization of statistical knowledge. Thus, I have chosen to end this review at the midpoint of the current century.

EPILOGUE

In *The Probabilistic Revolution*, the various contributing authors attempt to argue that during the nineteenth century there was a scientific revolution that produced a major paradigm shift associated with the adoption of probabilistic thinking. Indeed, Volume 1 begins with a chapter by Thomas Kuhn, who summarizes ideas from his classic 1962 book (see Kuhn, 1970), which laid out the distinction between the normal evolutionary mode of scientific progress and revolutionary change. I found the arguments in support of the occurrence of such a probability revolution to be the weakest feature of these two edited volumes of papers. Part of the problem comes from the focus that many of the authors have on probability rather than on statistics—that is, inferential issues that link probability to actual data. My interpretation of the nineteenth century history is that the spread of probabilistic ideas into several areas of science was evolutionary, rather than revolutionary in nature. The basic mathematical structure of probability was already widely accepted and the adoption of probabilistic models for phenomena was a natural extension to deterministic approaches.

In *The Empire of Chance*, Gigerenzer et al. also talk about the probabilistic revolution, specifically in the field of physics, focusing on the change in interpretation of physical phenomena that occurred in 1860 when Maxwell used the law of error to describe the velocities of gases. Their argument, which ties in to one of the foci in Porter's book, is a bit more convincing but it is restricted to the field of physics. Even in physics, the triumph of probabilistic thinking over determinism is far from complete as one can note from the current fascination of physical scientists with chaos theory and other mathematical devices that seemingly explain what others describe as stochastic behavior.

For me, if there was a scientific revolution during this period, it was really a result of the statistical ideas associated with what Stigler calls the Gauss–Laplace synthesis, which combined the normal error theory with the curve fitting method of least square into an inferential approach to the analysis of data using linear models. Yet, it took another 75 to 100 years and Galton's and Yule's formulation of regression before these ideas were used far beyond the boundaries of the astronomical problems addressed by Laplace and Gauss. Such delays are consistent with Kuhn's notions, because the adoption of a paradigm shift often requires an entirely new generation of scientists open to different ways of scientific thinking.

I actually have a second candidate to propose as the focus for a revolution in statistical thinking: the 1920s and 1930s contributions of Fisher, Neyman and Pearson for designing and making inference from randomized experiments and randomly selected samples. The truly novel component to their statistical ideas is the injection of a probabilistic component into a scientific problem through the design introduced by the scientist and the use of this component to make inferences about relevant hypothesis or population quantities. Following the application of the Fisher–Neyman–Pearson ideas to selected problems in the late 1930s and World War II, we have seen what is close to the universal adoption of their approach to virtually all areas of science. The possible exceptions to this spread are the "great observational sciences" of astronomy and physics that were so important to Laplace and Gauss.

In the traditional approach to statistics that emerged from the Gauss–Laplace synthesis and was developed by the statisticians at the turn of the century, probability enters a scientific problem as a property of the state of nature. The key new feature of the Fisher approach to experimentation and the parallel Neyman approach to sampling is the introduction of probability to the problem through a probability based randomization mechanism, such as a table of "random numbers." This externally introduced probability is not part of the state of nature, but it is used to make inferences from the experiment or sample to a population. Indeed, the methodology does not depend on whether nature is totally deterministic or at least partially stochastic. The statistical methods that have emerged over the subsequent 50 to 60 years represent an amalgam of the Fisher–Neyman randomization ideas and the ideas regarding the stochastic features of nature, with some statisticians relying solely on randomization for inference, others relying solely on statistical models with stochastic components and others on some mixture of the two.

To write a truly comprehensive history of only a selective period of an area of science often requires a heroic effort that goes unrecognized by nonhistorians. Such an effort typically involves the painstaking examination of source materials, which, in the case of Stigler, included the examination of the marginal notes of one statistician in his personal copy of an earlier work. It also attempts to unravel fact from myth, and to resolve competing claims of priority for various discoveries in circumstances where authors were not necessarily familiar with or at least did not comprehend the contribution of earlier related results of others. It requires the examination of secondary and even tertiary sources in order to reassess the influence and importance of various primary contributions. A further difficulty arises as one surveys the landscape of intellectual activities in a given period. The greater the relevant literature, the more difficult is the dual task of being comprehensive while at the same time describing the "big picture."

My overview in this review essay is garnered in large part from the books under review and involves only limited examination of primary sources. Where my familiarity with the primary sources is greatest, in the twentieth century, I take the greatest issue with the authors trying to describe this period. But, since many of the accomplishments of the twentieth century may still be too fresh for us to sort out, this disagreement might have well occurred even if I was far less familiar with the materials.

Each of the major authors in the septet of books under review succeeds in giving a reasonably good view of his or her selected topics and periods from the history of statistics. While there are often quite different perspectives and commentaries on the relative importance of specific authors and specific works, this variety is what I would expect from authors with such diverse backgrounds and interests. When the books are viewed collectively, their coverage of statistical topics is almost complete. At first, one might express surprise at the total absence of references to William Playfair, whose compelling statistical graphics have been the focus of considerable recent attention (e.g., see Costigan-Eaves and Macdonald-Ross, 1990). But Playfair worked without the use of probabilistic ideas in the school of political arithmetic that descended from Graunt, and thus his work does not play a key role in the development of statistics during the nineteenth century following the publication of his books.

The one topic I found to be "neglected" is the development of sample surveys and the related statistical methodology as a formal statistical enterprise. There are some related references in Porter's and Stigler's (especially as census-taking provides part of the foundation of survey-taking) and a somewhat obscure footnote in Daston's book (p. 360), but these discussions are brief at best. Moreover, Gigerenzer et al., who focus on twentieth century contributions, devote a scant 2–3 pages (referring primarily to secondary sources) in order to describe what I view as one of the more remarkable achievements in statistics over the past century. At least Stigler does mention in passing Laplace's development of ratio estimation using inverse probability and Quetelet's later proposal for implementation of it. But none of the books or authors explains why, after the extensive development of political arithmetic in the seventeenth and eighteenth centuries and the early nineteenth century development of census-taking and surveys as statistical activities, it took so long for random selection of sampling units to be introduced.

The following very brief account of these developments in the area of sample surveys is drawn in part from Fienberg and Tanur (1990). Although census taking goes back at least to biblical times, for most practical purposes we can skip from then to the end of the eighteenth century and the initiation of census activities in the United States. Although there is some debate as whether Canada, Sweden or the United States should be credited with originating the modern census, attention is usually focused on the United States, whose the first census was taken in 1790 (Anderson, 1988). Issues of census accuracy and undercount in the United States date back to this first census (Jefferson, 1791).

The move from censuses to surveys was slow and laborious, with much of the groundwork laid by Quetelet who helped to organize the first of a series of International Statistic Congresses in 1853. In William Farr from England and Lemuel Shattuck from the United States, Quetelet found like-minded collaborators who worked to put census taking and other forms of social data collection on a scientific basis. But it was only with the work of A. N. Kiaer, just before the turn of the century, that we began to see the development of the argument for the "representative method" for survey taking that was to be the focus of so much energy at the meetings of the International Statistical Institute meetings for the first 30 years of the twentieth century. There was a clear change in mode of inference that accompanied the breakthrough in methodology for sample surveys, which was tied to Jerzy Neyman's pathbreaking 1934 paper on the topic. When Neyman delivered a series of lectures organized by W. Edwards Deming at the U.S. Department of Agriculture Graduate School in 1937 (Neyman, 1938), he stimulated the rapid expansion of government-sponsored, probability-based sample surveys that had been begun only a few years earlier under the sponsorship of the Committee on Government Statistics led by Stuart A. Rice. The development of the theory of multi-stage, clustered probability sampling by Morris Hansen and his collaborators, which occurred primarily at the Bureau of the Census during the late 1930s and 1940s, culminated in the publication of the widely used two-volume compendium of theory and methods by Hansen, Hurwitz and Madow (1953). A good statistical history of the topic could hardly ignore the virtually independent and often complementary contributions of P. C. Mahalanobis and his students in India and by Frank Yates and William G. Cochran in England during this same period. Hansen (1987) provides a personal account of some of these developments, all of which go well beyond the 1900 boundary of the Porter and Stigler books, although not that of *The Empire of Chance*.

But an omission of this magnitude is still only a minor flaw in such an important set of books that otherwise give the reader an excellent sense of the intellectual origins of modern statistical thinking and the way in which statistical ideas were created and developed. Let me return briefly to the books under review.

Hald's book provides an interesting compendium of material on the pre-history of statistics. Some of the chapters provide real insight into the detailed discussion in source materials and to the interrelationships among various contributions. But Hald also includes some seemingly irrelevant digressions, for example, on mathematics and natural philosophy before 1650 and on the Newtonian revolution in mathematics and science. In such a major digression Hald emulates Karl Pearson, who included major chunks of material in his lectures on topics like the Newton–Leibnitz controversy, which has little direct bearing on the history of probability and statistics. Most of the material on such topics in Hald's book, however, are based on secondary or tertiary sources and thus offer few if any new insights. He also fails to include Newton's connection to the Trial of the Pyx described in Stigler (1977). Furthermore, unlike what we read in Hald's confusing account of Kepler and data analysis (which is unlinked to any other material in the book), Kepler did not show that an ellipse gave a better fit than an ovoid to the data on the planetary orbit of Mars. Rather, his choice, fortuitous though it was, was based on a somewhat arbitrary evaluation and the fact that the ellipse was more mathematically tractable than the ovoid. For a related discussion, see Fienberg (1985). In fact, recent evidence by Donahue (1988), which appeared simultaneously with the publication of Hald's book, strongly suggests that the new data presented by Kepler in his 1609 book to support his elliptical orbit theory were fabricated. But I too digress. Of far greater concern to me is the disjointed nature of Hald's presentation with material from a single source being presented in seemingly unrelated sections and chapters. A great part of the value of Hald's book for statistics is his focus on the details of the technical arguments, including the reproduction of proofs, updated somewhat so that the notation conforms somewhat more closely to the modern usage. By the time Hald gets to De Moivre's recursion formula for the "duration of play" in a two-person probabilistic game, however, we are awash in combinatoric formulae and equations. For the statistician, this suddenly reads more like a text on combinatorial probability than one on the history of probability and statistics.

Daston's book stands in sharp contrast to Hald's. It contains few technical arguments and tends to stress cross-cutting historical and philosophical themes. Indeed, her book is organized by these themes, and, instead of proceeding in a chronological fashion, it sweeps back and forth across the terrain of over 200 years of writing on probability covering themes such as expectation, the theory and practice of risk, the meaning of probability, the probability of causes and moralizing mathematics. Her prose is carefully crafted, and the text contains many insights, but I am afraid that many of them would have been lost on me had I not read Hald and Stigler first. Cowan (1987), in a review of Stigler's and Porter's books, notes the difference between two approaches to the writing of the history of science: externalism and internalism. For the prehistorical period, Daston is the externalist, always looking to the larger intellectual issues, and Hald is the internalist forever wallowing in the details of formulas and proofs.

Stigler's book is meticulous in detail, as he attempts to adjudicate among conflicting claims over interpretation of early works (e.g., of Bayes and Laplace), to speculate explicitly about what some authors meant and to try to develop a consistent interpretation of seemingly conflicting historical statements or claims of proof. He brings a modern statistical outlook to otherwise archaic technical arguments but always takes care to explain what we can infer about what was known at the time the author in question was writing. But most importantly, Stigler writes, not in a dry and boring fashion that some think appropriate for history, but in a way that draws the reader in with the fascination of solving a mystery of paramount importance. And everywhere there is a touch of wit, as in the remark about the Bernoullis:

> The Bernoullis are surely the most renowned family in the history of the mathematical sciences. Perhaps as many as twelve Bernoullis have contributed to some branch of mathematics or physics, and at least five have written on probability. So large is the set of Bernoullis that chance alone may have made it inevitable that a Bernoulli should be designated father of the quantification of uncertainty. (p. 63)

Porter's book has few technical arguments and mathematical formulas, but his prose is not light. He sets the story of Chapter 3 of our history of statistics much more firmly on its social roots and leans more upon social explanations and liberal politics than upon statistical and mathematical developments. As such it is a much easier "read" than Stigler or the related chapters of *The Probabilistic Revolution*. The contrast between Stigler and Porter is best seen in their discussions of Quetelet: Porter focuses on Quetelet's work on social statistics as social physics, while Stigler finds examples of early contributions to such topics as the analysis of variance and normal probability plots; Porter sees Quetelet as "lacking . . . the genius to formulate a usable mathematical procedure for analyzing statistical information," whereas Stigler finds ingenious methods of calculation. In her book review of Stigler and Porter, Cowan notes that they have written markedly different accounts, separated largely by the distinction between externalism and internalism. Stigler, the statistician and thus the internalist, focuses on the scientific texts, whereas Porter, the historian and externalist, searches more to learn what led key individuals to work on the problems they chose and in the manner that they did. In exploring the internalist/externalist distinction, Cowan criticizes Stigler for being so focussed on the intellectual history of statistics that he fails to note the social origins of Galton's work on correlation and his founding of the eugenics movement. Yet, without the fine-grained intellectual history, all the contextual information of the sort offered by Porter, the externalist, has at best limited value. In my readings of the two books, I found Porter more prone to generalization than Stigler, and far less knowledgeable about the importance of the statistical ideas under discussion. Nonetheless, his historical perspective provides a valuable supplement to Stigler's insights and observations on the technical details.

The two volumes of *The Probabilistic Revolution* provide some useful supplements to the material in Stigler and Porter, but their primary focus on probability rather than statistics leads them into lacunae that are diverting at best from the broad picture of the historical developments of statistics. With the exception of the brief chapter by Stigler in Volume 1, which closely resembles the Introduction to his book but with some subtle statistical differences, the chapters are not authored by statisticians, and their externalist perspectives are not in line with my own.

The septet of books described in this review form a major addition to and revamped perspective on the history of statistics. As fields, both statistics and the history of science are fortunate to have these contributions, which fill in many of the gaps in our previous understanding of the evolution of statistics as a field. The books belong in every library and on the shelves of those who take the study of the development of scientific concepts and methods seriously.

Many readers will not have the time or patience to pore over the thousands of pages represented by these volumes. For them, I recommend Stigler's book: for its coverage of the two most important chapters of the history described here and a good piece of the prehistory, as well as for its scholarship and balance between technical and contextual materials, and for the elegance of its prose. And in its newly issued paperback edition, Stigler's book is one that belongs on the shelf of every statistician and historian of probability and statistics. Hacking (1990b) argues that "[h]istorical research is best done by historians trained in archives, notwithstanding the existence of wonderful hobbyists." Stigler is not a mere hobbyist, and his book puts the lie to Hacking's suggestion [a view also expressed by Shafer (1990) in his rejoinder to Hacking].

Following Stigler, I suggest a selection of material from Porter's book, especially the intriguing sections on the influence of Quetelet on Maxwell and Boltzmann. If your interest is in the twentieth century developments and the controversies involving Fisher and Neyman, you might choose to go to their biographies (on Fisher, by his daughter, Joan: Box, 1978; on Neyman, by Constance Reid, 1982). And then, of course, you can always do quite well by piecing together a selection of historical articles and reprinted classics from *Statistical Science*.

ACKNOWLEDGMENTS

This essay was prepared in large part at the Hebrew University of Jerusalem, where I was Berman Visiting Professor during the winter of 1989–1990, and at Carnegie Mellon University. I am indebted to several colleagues for ideas, critical comments, and editorial suggestions: Myron Gutmann, John Modell, Teddy Seidenfeld, S. James Press, Stephen Stigler, Judith Tanur and Joel Tarr. Stu Hunter suggested the preparation of the timeline and provided a draft version, and Stephen Stigler generously provided several of the photographs.

REFERENCES

ANDERSON, M. J. (1988). *The American Census. A Social History.* Yale Univ. Press.

BARNARD, G. A. (1990). Fisher: A retrospective. *Chance* **3** 22–28.

BARNES, J. (1989). *A History of the World in 10½ Chapters.* Knopf, New York.

BERNOULLI, J. (1713). *Ars Conjectandi.* Thurnisiorum, Basil.

BOX, J. F. (1978). *R. A. Fisher, The Life of a Scientist.* Wiley, New York.

BOX, J. F. (1987). Guinness, Gosset, Fisher, and small samples. *Statist. Sci.* **2** 45–52.

COSTIGAN-EAVES, P. and MACDONALD-ROSS, M. (1990). William Playfair (1759–1823). *Statist. Sci.* **5** 318–326.

COWAN, R. S. (1987). Book review of Porter and Stigler. *J. Amer. Statist. Assoc.* **82** 1178–1179.

DE FINETTI, B. (1930). Fondamenti logici del ragionamento probabilistico. *Unione Matematica Italiana, Bolletino Series A* **9** 258–261.

DE FINETTI, B. (1937). La prévision: Ses lois logiques, ses sources subjectives. *Annales de l'Institut Henri Poincaré* **7** 1–68. [Translated and reprinted in *Studies in Subjective Probability* (H. E. Kyberg, Jr. and H. E. Smokler, eds.) 93–158. Wiley, New York, 1964.]

DE MOIVRE, A. (1718). *The Doctrine of Chances: or, A Method of Calculating the Probability of Events in Play.* W. Pearson, London. (2nd ed., 1738, Woodfall, London.)

DONAHUE, W. H. (1988). Kepler's fabricated figures: Covering up the mess in the *New Astronomy. J. Hist. Astronom.* **19** 217–237.

FIENBERG, S. E. (1985). Statistics and the scientific method: comments on and reactions to Freedman. In *Cohort Analysis in Social Research: Beyond the Identification Problem* (W. M. Mason and S. E. Fienberg, eds.) 371–383. Springer, New York.

FIENBERG, S. E. and HINKLEY, D. V. (1980). *R. A. Fisher: An Appreciation. Lecture Notes in Statist.* **1**. Springer, New York.

FIENBERG, S. E. and TANUR, J. M. (1990). A historical perspective on the institutional bases for survey research in the United States. *Survey Methodology* **16** 31–46.

GALTON, F. (1869). *Hereditary Genius: An Inquiry into its Laws and Consequences.* Macmillan, London. (2nd ed., 1892.)

GRAUNT, J. (1662). *Natural and Political Observations Made Upon the Bills of Mortality.* Martyn, London.

HACKING, I. (1975). *The Emergence of Probability.* Cambridge Univ. Press.

HACKING, I. (1990a). *The Taming of Chance.* Cambridge Univ. Press.

HACKING, I. (1990b). Comment: In praise of the diversity of probabilities. *Statist. Sci.* **5** 450–454.

HANSEN, M. H. (1987). Some history and reminiscences on survey sampling. *Statist. Sci.* **2** 180–190.

HANSEN, M. H., HURWITZ, W. N. and MADOW, W. G. (1953). *Sample Survey Methods and Theory* (2 vol.). Wiley, New York.

HAWKING, S. W. (1988). *A Brief History of Time.* Bantam, New York.

HEYDE, C. C. and SENETA, E. (1977). *I. J. Bienaymé: Statistical Theory Anticipated.* Springer, Berlin.

JEFFERSON, T. (1791). Letter to David Humphreys. [Reprinted in *The Papers of Thomas Jefferson. Vol. 22.* (Charles T. Cullen, ed.) 62. Princeton Univ. Press, 1986.]

JEFFREYS, H. (1939). *Theory of Probability.* Oxford Univ. Press.

JOHNSON, N. L. and KOTZ, S., eds. (1992). *Breakthroughs in Statistics* (2 vol.). Springer, New York.

KENDALL, D. G. (1991). Kolmogorov as I remember him. *Statist. Sci.* **6** 303–312.

KENDALL, M. G. (1960). Where shall the history of statistics begin? *Biometrika* **47** 447–449.

KUHN, T. S. (1970). *The Structure of Scientific Revolutions*, 2nd ed. Univ. Chicago Press.

LINDLEY, D. V. (1991). Sir Harold Jeffreys. *Chance* **4** 10–14.

NEYMAN, J. (1923). On the application of probability theory to agricultural experiments. Essay on priniciples. Section 9. *Annals of Agricultural Sciences.* **X** 1–51. [Translated and edited by D. M. Dabrowska and T. P. Speed (1990). *Statist. Sci.* **5** 463–480.]

NEYMAN, J. (1934). On the two different aspects of the representative method: The method of stratified sampling and the method of purposive selection (with discussion). *J. Roy. Statist. Soc. Ser. A* **97** 558–625.

NEYMAN, J. (1938). *Lectures and Conferences on Mathematical Statistics and Probability.* Graduate School, U.S. Dept. Agriculture, Washington, DC.

PEARSON, E. S. (1990). *"Student": A Statistical Biography of William Sealy Gosset* (R. L. Plackett and G. A. Barnard, eds.). Oxford Univ. Press.

PEARSON, K. (1978). *The History of Statistics in the Seventeenth and Eighteenth Centuries, Against the Changing Background of Intellectual, Scientific, and Religious Thought* (E. S. Pearson, ed.). Charles Griffin, London.

RAMSAY, F. P. (1926). Truth and Probability. [Reprinted in *Studies in Subjective Probability* (H. E. Kyberg Jr. and H. E. Smokler, eds.) 61–92. Wiley, New York, 1964.]

REID, C. (1982). *Neyman from Life.* Springer, New York.

SAVAGE, L. J. (1954). *The Foundations of Statistics.* Wiley, New York.

SAVAGE, L. J. (1976). On rereading Fisher. *Ann. Statist.* **4** 441–500.

SHAFER, G. (1990). The unity and diversity of probability (with discussion). *Statist. Sci.* **5** 435–462.

STIGLER, S. M. (1977). Eight centuries of sampling inspection: The trial of the Pyx. *J. Amer. Statist. Assoc.* **72** 493–500.

STIGLER, S. M. (1980). Stigler's law of eponymy. *Trans. New York Acad. Sci.* **39** 147–157.

Stigler, S. M. (1986). Laplace's 1774 memoir on inverse probability. *Statist. Sci.* 1 359–378.

Stigler, S. M. (1989). Francis Galton's account of the invention of correlation. *Statist. Sci.* 4 73–86.

Todhunter, I. (1865). *A History of the Mathematical Theory of Probability.* Macmillan, London. (Reprinted, 1949 and 1965, Chelsea, New York.)

Wald, A. (1945). Sequential tests of statistical hypotheses. *Ann. Math. Statist.* **16** 117–196.

Walker, H. M. (1978). Pearson, Karl. *International Encyclopedia of Statistics* 2 691–698. Free Press, New York.

Zabell, S. (1989). R. A. Fisher on the history of inverse probability (with discussion). *Statist. Sci.* 4 247–263.