# A History of Probability and Statistics and Their Applications before 1750

This Page Intentionally Left Blank

# A History of Probability and Statistics and Their Applications before 1750

This Page Intentionally Left Blank

# A History of Probability and Statistics and Their Applications before 1750

ANDERS HALD
Formerly Professor of Statistics
University of Copenhagen
Copenhagen, Denmark

WILEY-INTERSCIENCE

A JOHN WILEY & SONS, INC., PUBLICATION

A NOTE TO THE READER
This book has been electronically reproduced from digital information stored at John Wiley & Sons, Inc. We are pleased that the use of this new technology will enable us to keep works of enduring scholarly value in print as long as there is a reasonable demand for them. The content of this book is identical to previous printings.



Published by John Wiley & Sons, Inc., Hoboken, New Jersey.
Published simultaneously in Canada.



***Library of Congress Cataloging-in-Publication is available.***

ISBN 0-471-47129-1

Printed in the United States of America.

10 9 8 7 6 5 4 3 2 1

# Preface

Until recently a book on the history of statistics in the 19th century was badly needed. When I retired six years ago, I decided to write such a book, feeling that I had a good background in my statistical education in the 1930s, when the curriculum in statistics was influenced mainly by the writings of Laplace, Gauss, and Karl Pearson. Studying the original works of these authors I found no difficulty in understanding Gauss and Pearson, but I soon encountered difficulties with Laplace. The reason is of course that Gauss and Pearson are truly 19th century figures, whereas Laplace has his roots in the 18th century.

I then turned to the classical authors and worked my way back to Cardano through de Moivre, Montmort, Nicholas and James Bernoulli, Huygens, Fermat, and Pascal. Comparing my notes with Todhunter's *History*, I found to my surprise that his exposition of the topics in probability theory that I found most important was incomplete, and I therefore decided to write my own account.

The present book, covering the period before 1750, is an introduction to the one I had in mind. It describes the contemporaneous development and interaction of three topics: probability theory and games of chance; statistics in astronomy and demography; and life insurance mathematics.

Besides the story of the life and works of the great natural philosophers who contributed to the development of probability theory and statistics, I have told the story of important problems and methods, in this way exhibiting the gradual advance of solving these problems. I hope to have achieved a better balance than had been achieved before in evaluating the contributions of the various authors; in particular, I have stressed the importance of the works of John Graunt, Montmort, and Nicholas Bernoulli.

The contents of the book depend heavily on research carried out by many authors during the past 40 years. I have drawn freely on these sources and acknowledged my debt in the references. The manuscript was written during the years 1985–1987, so works published in 1986 and 1987 are not fully integrated in the text. Some important books and papers from 1988 are briefly mentioned.

With hesitation, I have also included some background material on the history of mathematics and the natural and social sciences because I have always felt that my students needed such knowledge. I realize of course that my qualifications for doing so are rather poor since I am no historian of science. These sections and also the biographies are based on secondary sources.

The plan of the book is described in Section 1.2.

I am grateful to Richard Gill for advice on my English in Chapters 2 and 3, to Steffen L. Lauritzen for translating some Russian papers, and to Olaf Schmidt for a discussion of Chapter 10. In particular, I want to thank Søren Johansen for discussions on the problem of the duration of play.

I am grateful to two anonymous reviewers from the publisher for valuable comments on the manuscript and for advice resulting in considerable reduction of the background material. I thank the copy editor for improving my English and transforming it into American.

I thank the Institute of Mathematical Statistics, University of Copenhagen, for placing working facilities at my disposal.

I thank the Almqvist & Wiksell Periodical Company for permission to use material in my paper published in *Scandinavian Actuarial Journal*, 1987; the International Statistical Institute for permission to use material from three papers of mine published in *International Statistical Review*, 1983, 1984, and 1986; and Springer-Verlag for permission to use material from my paper published in *Archive for History of Exact Sciences*, 1988.

I am grateful to the Department of Statistics, Harvard University, for permission to quote from Bing Sung's *Translations from James Bernoulli*, Technical Report No. 2, 1966, and to Thomas Drucker for permission to quote from his (unpublished) translation of Nicholas Bernoulli's *De Usu Artis Conjectandi in Jure*.

My first book on statistics, written fifty years ago, was dedicated to G. K., so is this one.

ANDERS HALD

*September 1988*

# Contents

This Page Intentionally Left Blank

# A History of Probability and Statistics and Their Applications before 1750

This Page Intentionally Left Blank

CHAPTER 1

# The Book and Its Relation to Other Works

## 1.1 PRINCIPLES OF EXPOSITION

This book contains an exposition of the history of probability theory and statistics and their applications before 1750 together with some background material. A history should of course give an account of the time and place of important events and their interpretations. However, opinions differ greatly on where to put the main emphasis of interpretation.

We have attempted to cover three aspects of the history: problems, methods, and persons. We describe probabilistic and statistical problems and their social and scientific background; we discuss the mathematical methods of solution and the statistical methods of analysis; and we include the background and general scientific contributions of the persons involved, not only their contributions to probability and statistics.

Since history consists of facts and their interpretation, history continually changes because new facts are found in letters, archives, and books, and new interpretations are offered in the light of deeper understanding based, in this case, on the latest developments in probability theory, statistics, and the history of science.

In the 17th and 18th centuries many problems were formulated as challenge problems, and answers were given without proofs. Some books on probability were written for the educated public and therefore contained statements without proofs. In such cases we have tried to follow the author's hints and construct a proof which we believe represents the author's intentions.

The material has been ordered more according to problems and methods than according to persons in an attempt to treat the achievements of the various authors as contributions to a general framework.

A leading principle of the exposition of probability theory and life insurance mathematics has been to rewrite the classics in uniform modern terminology and notation. It is clear that this principle may be criticized for distorting the facts. Many authors prefer to recount the old proofs with the original notation to convey the flavor of the past to the reader. There are two essential steps in modernization that we have made here. The first is to use a single letter, $p$, say, to denote a probability instead of the ratio of the number of favorable cases to the total number of cases, $a/(a+b)$, say, where $a$ and $b$ are positive integers. This change of notation conceals the fact that nearly all the probabilities discussed were constrained to rational fractions. The advantage of this notation was noted by de Moivre (1738, p. 29) who writes, "Before I make an end of this Introduction, it will not be improper to shew how some operations may often be contracted by barely introducing one single Letter, instead of two or three, to denote the Probability of the happening of one Event" and, further (on p. 30), that "innumerable cases of the same nature, belonging to any number of Events, may be solved without any manner of trouble to the imagination, by the mere force of a proper Notation." However, de Moivre did not rewrite the *Doctrine of Chances* with the new notation; he used it only in his *Annuities upon Lives* (1725 and later editions). We have followed the advice of de Moivre and rewritten the proofs in the new notation, feeling confident that the reader will keep in mind that most probabilities were defined as proper rational fractions, a fact which is nearly always obvious from the context.

The second great simplification of the proofs is obtained by the introduction of subscripts. In analyzing some complicated games of chance, for example, Waldegrave's problem, Nicholas Bernoulli and de Moivre had to use the whole alphabet divided into several sections to denote probabilities and expectations of the players corresponding to various states of the game. De Moivre achieved some simplification by using superscripts in a few cases. In many problems they gave the solution for two, three, and four players only and concluded that "the continuation of this rule is manifest," in this way avoiding a general proof which would have been rather unintelligible. Using modern notation with subscripts, it is easy to rewrite such proofs in much shorter form without invalidating the idea of the proof; in fact, we believe that our readers will get a clear idea of the proof because they are accustomed to this symbolism, just as readers in the past understood the original form of the proof because they were educated in that notational tradition.

Comparison of proofs and results in a uniform notation makes evaluating the contributions of various authors easier and minimizes the danger of attributing too much to an individual author. Furthermore, the importance of the results to the following period and to today becomes evident.

The same principle of exposition cannot be used for statistics, because statistics before 1750 was nonmathematical. We shall therefore illustrate the development of statistical methods by typical examples, giving both the original data and their analysis at the time and adding some comments from a modern point of view.

The book is written in textbook style, since our main purpose is to give an account of the most important results in the classical literature. Like most histories of mathematics and science, our exposition concentrates on results which have proved to be of lasting importance.

The persons who laid the foundation of probability theory and statistics were natural philosophers having a broader background and outlook than scientists today. The word "scientist" was coined about the middle of the 19th century, reflecting an ongoing specialization and professionalization. Nevertheless, we shall often use the words "mathematician" and "scientist" to stress certain characteristics of the persons involved.

To convey the flavor of classical works, we shall present quotations of programs from the prefaces of books, the formulation of important problems, and some heated disputes of priority.

We shall point out priorities, but the reader should be aware of the uncertainty involved by taking note of Stigler's Law of Eponymy, (Stigler, 1980), which in its simplest form states that, "No scientific discovery is named after its original inventor."

The driving force behind the development of probability theory and statistics was pressure from society to obtain solutions to important problems for practical use, as well as competition among mathematicians. When a problem is first formulated and its solution indicated, perhaps only by a numerical example, the problem begins a life of its own within the mathematical community; this leads to improved proofs and generalizations of the problem, and we shall see many examples of this phenomenon.

Finally, it should be noted that any history is necessarily subjective, since the weight and interpretation of the events selected depend on the author's interests.

For the serious student of the history of probability theory and statistics, we can only recommend that he or she follow the advice given by de Moivre (1738, p. 235), discussing the works of James and Nicholas Bernoulli on the binomial distribution: "Now the Method which they have followed has been briefly described in my *Miscellanea Analytica*, which the Reader may consult if he pleases, unless they rather chuse, which perhaps would be the best, to consult what they themselves have writ upon that Subject."

## 1.2 PLAN OF THE BOOK

A fuller title of the book would be *A history of probability theory and statistics and their applications to games of chance, astronomy, demography, and life insurance before 1750, with some comments on later developments.* The topics treated may be grouped into five categories:

Background in mathematics, natural philosophy, and social conditions
Biographies
Probability theory and games of chance
Statistics in astronomy and demography
Life insurance mathematics

Probability theory before 1750 was inspired mainly by games of chance. Dicing, card games, and lotteries, public and private, were important social and economic activities then as today. It is no wonder that intellectual curiosity and economic interests led to mathematical investigations of these activities at a time when the mathematization of science was going on. We shall distinguish three periods.

The period of the foundation of probability theory from 1654 to 1665 begins with the correspondence of Pascal and Fermat on the problem of points, continues with Huygens' treatise on *Reckoning at Games of Chance*, and ends with Pascal's treatise on the *Arithmetical Triangle* and its applications. The correspondence was not published until much later. In his treatise, Pascal solves the problem of points by recursion and finds a division rule, depending on the tail probability of the symmetric binomial. In their correspondence, he and Fermat had solved the same problem also by combinatorial methods. Huygens uses recursion to solve the problem numerically. He also considers an example with a possibly infinite number of games, which he solves by means of two linear equations between the conditional expectations of the two players. All three of them solved the problem of the Gambler's Ruin without publishing their method of solution.

After a period of stagnation of nearly 50 years, there followed a decade with astounding activity and progress from 1708 to 1718 in which the elementary and fragmentary results of Pascal, Fermat, and Huygens were developed into a coherent theory of probability. The period begins with Montmort's *Essay d'Analyse sur les Jeux de Hazard*, continues with de Moivre's *De Mensura Sortis*, Nicholas Bernoulli's letters to Montmort, James Bernoulli's *Ars Conjectandi*, Nicolaas Struyck's *Reckoning of Chances in Games*, and ends with de Moivre's *Doctrine of Chances*. Hence, by 1718 four comprehensive textbooks were available. We shall mention the most important results obtained. They discussed elementary rules of probability calculus, conditional probabilities and expectations, combinatorics, algorithms and recursion formulae, the method of inclusion and exclusion, and examples of using infinite series and limiting processes. They derived the binomial and negative binomial distributions, the hypergeometric distribution, the multivariate version of these distributions, the occupancy distribution, the distribution of the sum of any number of uniformly distributed variables, the Poisson approximation to the binomial, the law of large numbers for the binomial, and an approximation to the tail of the binomial. They solved the problem of points for a game of bowls and for the game of tennis, Waldegrave's problem, the problem of coincidences, and the problem of duration of play, and found the minimax solution for the strategic game Her.

The third period, from 1718 to 1738, was a period of consolidation and steady progress in which de Moivre derived the normal approximation to the binomial distribution, developed a theory of recurring series, improved his solution of the problem of the duration of play, and wrote the second edition of the *Doctrine of Chances*, which became the most important textbook before the publication of Laplace's *Théorie Analytique des Probabilités* in 1812.

We shall discuss these books in detail. We have, however, singled out the most important problems for separate treatment to show how they were solved by joint effort, often in competition among several authors.

Many problems were taken up by the following generation of mathematicians and given solutions that have survived until today. We shall comment on these later developments, usually ending with Laplace's solutions.

The successful development of probability theory did not immediately lead to a theory of statistics. A history of statistical methods before 1750 must therefore build on typical examples of data analysis; we have concentrated here on examples from astronomy and demography.

Astronomers had been aware of the importance of both systematic and random errors since antiquity and tried to minimize the influence of such errors in their planning of observations and data analysis. We shall discuss some data by Tycho Brahe from the end of the 16th century as an example. The mathematization of science in the beginning of the 17th century naturally led many scientists to determine not only the mathematical form of natural laws but also the values of the parameters by fitting equations to data. They inserted the best sets of observations in the equations, as many as the number of parameters, solved for the parameters, calculated the expected values, and studied the deviations between observed and calculated values. Prominent examples are Kepler's three laws on planetary motion derived from his physical theories and data collected by Copernicus and Tycho Brahe. Kepler's data were used by Newton to check his axiomatic theory. Galileo used several sets of observations on the new star of 1572 to compare two hypotheses on the position of the star. We shall also see how Newton used an interpolation polynomial to find the tangent to the orbit of a comet.

A paragon for descriptive statistical analysis of demographic data was provided by Graunt's *Natural and Political Observations made upon the Bills of Mortality* in 1662. Graunt's critical appraisal of the rather unreliable data, his study of mortality by cause of death, his estimation of the same quantity by several different methods, his demonstration of the stability of statistical ratios, and his life table set up new standards for statistical reasoning. Graunt's work led to three different types of investigations: political arithmetic; testing the stability of statistical ratios; and calculation of expectations of life and survivorship probabilities.

Petty also employed Graunt's method of analysis, although less critical, to economic data and coined the term "political arithmetic" for analyses of data of political importance. Similar methods were used by natural philosophers and theologians to analyze masses of data on human and animal populations. The many regular patterns observed were taken as proof of the existence of a supreme being and His "original design." We shall remark only slightly on this line of thought.

It is surprising that probabilists at the time recognized the importance of Graunt's work and without hesitation used their theory on games of chance to describe demographic phenomena. They wrote about the chance of a male birth and the chance of dying at a certain age.

Graunt gave a detailed description and analysis of the yearly variation of the sex ratio at birth in London and Romsey and suggested that similar investigations should be carried out in other places. Arbuthnott used some of Graunt's data extended to his own time to give a statistical proof, based on the symmetric binomial, for the existence of divine providence, a proof that was further strengthened by 'sGravesande. Nicholas Bernoulli compared the observed distribution of the yearly number of male births with a skew binomial distribution, the parameter being estimated from the data, and discussed the probability of the observed number of outliers. His investigation is the first attempt to fit a binomial to data and to test the goodness of fit. Some years later, Daniel Bernoulli used the normal approximation to the binomial in his analysis of deviations between observed and expected values of the number of male births to decide between two hypothetical values of the sex ratio.

Huygens used Graunt's life table to calculate the median and the average remaining lifetime for a person of any given age. He also showed how to calculate survivorship probabilities and joint-life expectations. His results were, however, not published, but similar results were published without proof by James Bernoulli and later proved by Nicholas Bernoulli.

The usefulness of probability theory was convincingly demonstrated by application to problems of life insurance. In the 16th and 17th centuries, states and cities sold life annuities to their citizens to raise money for public purposes. The yearly benefit of an annuity was fixed as a percentage of the capital invested, often as twice the prevailing rate of interest and independent of the nominee's age. In a report from 1671, de Witt showed how to calculate the value of an annuity by means of a piecewise linear life table combined with the age of the nominee and the rate of interest. De Witt's life table was hypothetical, although he referred to some investigations of the mortality of annuitants. In 1693 Halley constructed a life table from observations of the yearly number of deaths in Breslau, calculated the first table of values of annuities as a function of the nominee's age, and explained how to calculate joint-life annuities.

After these ingenious beginnings one would have expected rapid development of both mathematical and practical results in view of the fact that many economic contracts in everyday life depended on life contingencies, but nothing happened for about 30 years. The breakthrough came in 1725 with de Moivre's *Annuities upon Lives*, greatly simplifying both the mathematics and the calculations involved; however, as shown by Simpson, de Moivre went too far in his simplifications. Simpson therefore constructed his own life table for the population of London, and by recursion he calculated tables of values of single- and joint-life annuities for various rates of interest. In the strong competition between de Moivre and Simpson, a comprehensive theory of life annuities was created, and the necessary tables for practical applications provided.

In some chapters in this book we have supplemented the text with problems for the reader, mostly taken from the classical literature.

Although we have not included every classical paper, or every paper commenting on the classical literature, we believe that we have covered the most important ones. However, for various reasons two important results before 1750 have been omitted. The first is Cotes's rule (1722) for estimating a true value by a weighted mean, when observations are of unequal accuracy (see Stigler, 1986, p. 16); the second is Daniel Bernoulli's results (1738) on the theory of moral expectation, the utility of money, and the Petersburg problem (see Todhunter, 1865, Jorland, 1987, and Dutka, 1988).

We shall discuss our reasons for stopping our history at 1750.

By 1750 probability theory had been recognized as a mathematical discipline with a firm foundation and its own problems and methods as described by de Moivre in the *Doctrine of Chances*. A new development began with the introduction of inverse probability by Bayes (1764) and Laplace (1774b).

By 1750 statistics had still not become a mathematical discipline; a mathematical theory of errors and of estimation emerged in the 1750s, as described by Stigler (1986).

Also about 1750, the first phase of the development of a theory of life insurance had been completed. In the 1760s life insurance offices arose so that new and more accurate mortality observations became available. A theory of life assurances was developed, and new ways of calculating and tabulating the fundamental functions were invented.

The reader should note that formulae are numbered with a single number *within sections*. When referring to a formula in *another chapter* the decimal notation is used, (20.5.25) say, denoting formula (25) in §5 of Chap. 20. *Within a chapter* the chapter number is omitted so that only section and formula numbers are given.

## 1.3 A COMPARISON WITH TODHUNTER'S BOOK

The unquestioned authority on the early history of probability theory is Isaac Todhunter (1820–1884) whose masterpiece, *A History of the Mathematical Theory of Probability from the Time of Pascal to that of Laplace*, was published in 1865. Kendall (1963) has written a short biography of Todhunter in which he gives a precise characterization of his work: "The *History of the Mathematical Theory of Probability* is distinguished by three things. It is a work of scrupulous scholarship; Todhunter himself contributed nothing to the theory of probability except this account of it; and it is just about as dull as any book on probability could be."

We consider Todhunter's *History* an invaluable handbook giving a chronological review of the classical literature grouped according to authors. For the period before 1750, however, we shall argue that Todhunter's account of important topics is incomplete, that he has overlooked the significance of important contributions, and that the trend in the historical development is lost by his organization of the material.

In the many references to Todhunter's *History* in the following we shall omit the year of publication (1865) and give page references only.

As a mathematician Todhunter concentrates on the mathematical theory of probability and disregards the general background, the lives of the persons involved, and the application of their theories to statistics and life insurance.

Todhunter's book is ordered chronologically according to authors; each important author is allotted a separate chapter in which his works are reviewed page by page and commented upon. This method makes it easy for the reader to locate the contributions of each author but difficult to follow the advances made by various authors to the solution of a given problem. We have avoided this dilemma by reviewing the works of each author and referring the detailed treatment of the most important topics to separate chapters that show the historical development for each topic.

More important, however, are the different weights given to many topics by Todhunter and by us. It is not surprising that the significance of a theorem or method differs when viewed from the perspectives of 1865 and today. Todhunter meticulously reports proofs of many results which are without interest today; conversely, he omits proofs of results of great importance. We shall give some examples.

Today one of the most important and interesting topics is the development from James Bernoulli's law of large numbers for the binomial distribution through Nicholas Bernoulli's improved version of James's theorem and his approximation to the binomial tail probability to de Moivre's normal approximation. These three results are treated by Todhunter in less than two pages (pp. 72, 131, 192). He states Bernoulli's theorem without giving his proof; he has overlooked the significance of Nicholas' contribution and gives neither theorem nor proof; he states de Moivre's result for $p = 1/2$ only and indicates the proof by the remark, "Thus by the aid of Stirling's Theorem the value of Bernoulli's Theorem is largely increased." Todhunter has completely overlooked de Moivre's long struggle with this problem, the importance of de Moivre's proof as a model for Laplace's proof, and de Moivre's statement of the theorem for any value of $p$. Instead of giving the historical development of the method of proof, he gives Laplace's proof (pp. 548–552) because, as he says, previous demonstrations are now superseded by that. This is of course a very peculiar argument for a historian.

It is a common misunderstanding, perhaps due to Todhunter's incomplete account, that de Moivre gave the normal approximation only for the symmetric binomial.

The deficiency of Todhunter's method is most conspicuous in his analysis of the correspondence between Montmort and Nicholas Bernoulli, published in the second edition of Montmort's *Essay* (1713). These closely intertwined letters contain formulations of new problems, usually as a challenge to the recipient; theorems without proofs, sometimes with hints for solution; replies to previous questions; a running commentary on progress with the solution of various problems; and remarks on the contributions of other authors. A single letter often treats five to ten different topics. It is of course impossible to get the gist of these letters in a page-by-page review; rather, it is necessary to give an overview of the contents grouped by subject matter. Todhunter therefore does not realize the importance of Nicholas Bernoulli's work; perhaps he was also under the influence of de Moivre who in the later editions of the *Doctrine* tried to conceal the importance of Bernoulli's results to his own work.

The most difficult topic in probability theory before 1750 was the problem of the duration of play. It was formulated by Montmort in 1708; the first explicit solution was given by Nicholas Bernoulli in 1713. Two solutions were given by de Moivre in 1718, and these were worked out in more detail in 1730 and 1738. Todhunter gives up analyzing this important development, instead he uses Laplace's solution from 1812 to prove de Moivre's theorems. Furthermore, he does not comment on Laplace's solution from 1776 by solving a partial difference equation because this method "since [has] been superseded by that of Generating Functions" (Todhunter, p. 475).

The same procedure is used by Todhunter in his discussion of Waldegrave's problem, the probability of winning a circular tournament, which was solved incompletely by Montmort and de Moivre. A general solution was given by Nicholas Bernoulli, but Todhunter gives only Laplace's proof without noting that Bernoulli's is just as simple.

Todhunter's discussion of the strategic game Her is rather incomplete. He has overlooked the fact that Montmort gives the general form of the player's expectation under randomized strategies and that Waldegrave solves the problem numerically arriving at what today is called the minimax solution. Misled by Todhunter's account, Fisher (1934) solved the "old enigma of card play" by randomization and reached the same solution as Waldegrave did 221 years before.

Todhunter gives unsatisfactory accounts of James Bernoulli's and Montmort's probabilistic discussion of the game of tennis, of the problem of points in a game of bowls, of Montmort's discussion of the occupancy problem, of Simpson's solution of the theory of runs, and of several other problems mentioned in the following chapters.

Kendall's characterization that "it is just as dull as any book on probability could be" applies equally well to several sections of the present book. Detailed proofs of elementary theorems illustrating the historical development are necessarily dull for us, even if they were exciting for them. Pascal, Fermat, Huygens, Hudde, James Bernoulli, Montmort, Nicholas Bernoulli, de Moivre, and Struyck were all intensely interested in solving the problem of the Gambler's Ruin, which today is considered elementary. For statisticians who find examples of games of chance rather dull, it must be a consolation to know that dicing and card playing have their equivalents in sampling from infinite and finite populations, respectively.

## 1.4 WORKS OF REFERENCE

Gouraud's *Histoire* (1848, 148pp.) gives a nonmathematical and rather uncritical exposition of probability theory and insurance mathematics beginning with Pascal and Fermat and ending with Poisson and Quetelet. It contains many references and was therefore useful for Todhunter when he wrote his *History* (1865, 624pp.).

The first two chapters of Czuber's *Entwicklung der Wahrscheinlichkeitstheorie* (1899, 279pp.) covers nearly the same period as the present book but in less detail. Czuber indicates some methods of proof without giving complete proofs.

The books by Edwards (1987, 174pp.), *Pascal's Arithmetical Triangle*, and Hacking (1975, 209pp.), *The Emergence of Probability*, may be read as an introduction to the present one; they give a more detailed treatment of certain aspects of the history up to the time of Newton and Leibniz.

David (1962, 275pp.) gives a popular history of probability and statistics from antiquity through the time of de Moivre, stressing basic ideas and providing background material for the lives of the great probabilists.

Jordan's book (1972, 619pp.) contains a mathematical account of classical probability theory organized according to topics, with some references to the historical development.

The first 81 pages of Maistrov's book (1974, 281pp.) gives a sketch of the history of probability theory before 1750.

Daston's *Classical Probability in the Enlightenment* (1988, 423pp.) gives a comprehensive, nonmathematical study of the basic ideas in classical probability theory in their relation to games of chance, insurance, jurisprudence, economics, associationist psychology, religion, induction, and the moral sciences, with references to a wealth of background material. Daston's discussion of the history of probabilistic ideas is an excellent complement to our discussion of mathematical techniques and results.

Turning to books on the history of statistics, we mention first Karl Pearson's *The History of Statistics in the 17th and 18th Centuries*, Lectures given at University College, London, 1921–1933, edited by E. S. Pearson (1978, 744pp.). This is a fascinating book with an unusual freshness that conveys Pearson's enthusiasm and last-minute endeavors in preparing his lectures. It describes "the changing background of intellectual, scientific and religious thought," and gives lively biographies with digressions into the fields of mathematics and history of science. Pearson does not discuss statistics in the natural sciences but is mainly concerned with political arithmetic, demography, and the use of statistics for theological purposes. Pearson does not conceal his strong opinions on the subjects treated and the persons involved, which occasionally lead to biased evaluations.

Stigler's *The History of Statistics* (1986, 410pp.) is the first comprehensive history of statistics from 1750 to 1900; it also contains a discussion of Bernoulli's law of large numbers and de Moivre's normal approximation to the binomial.

Westergaard's *Contributions to the History of Statistics* (1932, 280pp.) gives the history of political arithmetic, population statistics, economic statistics, and official statistics before 1900, as well as a short survey of statistical theory. It is a nonmathematical, well-balanced, and scholarly work, with valuable references to the vast literature on descriptive and official statistics.

John's *Geschichte der Statistik* (1884, 376pp.) contains a description of the development of German political science, at that time called statistics, and of political arithmetic and population statistics before 1835.

Meitzen's *Geschichte, Theorie und Technik der Statistik* (1886, 240pp.) discusses the history of official statistics with the main emphasis on its development in Germany.

Following the pioneering work by M. G. Kendall and F. N. David in the 1950s and 1960s, there has been growing interest in the history of probability and statistics, and a great number of papers have been published; the most important, relating to the period before 1750, are listed in the References at the end of this book. Several important papers have been reprinted in *Studies in the History of Statistics and Probability*, Vol. 1 edited by E. S. Pearson and M. G. Kendall (1970) and Vol. 2 edited by M. G. Kendall and R. L. Plackett (1977). A *Bibliography of Statistical Literature Pre-1940* has been compiled by Kendall and Doig (1968).

A comprehensive account of the development of life insurance and its social, economic, and political background before 1914, with some remarks on mathematical results has been given by Braun in *Geschichte der Lebensversicherung und der Lebensversicherungstechnik* (1925, 433pp.).

For the biographies we have of course used the *Dictionary of Scientific Biography*, edited by C. C. Gillispie (1970–1980) and the individual biographies available.

As reference books for the history of mathematics we have used Cantor (1880–1908) and Kline (1972).

For long periods of time there existed a considerable backlog of publications of the Academies at London, Paris, Turin, etc., so that papers were read some years before they were published. Referring to such papers we have used the *date of publication*; in the list of references, however, we have usually added the date of communication to the Academy.

CHAPTER 2

# A Sketch of the Background in Mathematics and Natural Philosophy

## 2.1 INTRODUCTION

The first mathematical analyses of games of chance were undertaken by Italian mathematicians in the 16th century. The main results, which remained unpublished for nearly a century, were obtained by Cardano about 1565.

It was almost 100 years after Cardano before probability theory was taken up again, this time in France by Pascal and Fermat (1654). Their work was continued by Huygens (1657) in the Netherlands. He wrote the first published treatise on probability theory and its application to games of chance.

About the same time a statistical analysis of data on the population of London was carried out by Graunt (1662). He did not have any knowledge of probability theory.

The first contributions to life insurance mathematics were made by de Witt (1671) in the Netherlands and by Halley (1694) in England. They combined Huygens' probability theory with Graunt's life table.

Error theory and the fitting of equations to data were developed in astronomy and navigation. Outstanding contributions are due to the Danish astronomer Tycho Brahe in the late 16th century, the German astronomer and mathematician Kepler in the beginning of the 17th century, and the Italian natural philosopher Galileo.

The problems taken up were of great current interest scientifically, socially, and economically. Their solutions depended on the mathematical background and sometimes required the development of new mathematical tools.

All these activities were well under way just before the Newtonian revolution, which was of decisive importance both mathematically and philosophically to further development.

For historical background we shall sketch the principal progress in mathematics and natural philosophy of importance for our subject before 1650. However, these fields only constitute a small part of the cultural background at the time. Most natural philosophers had a very broad education, worked in many different areas, and entertained ideas which today would be called superstitious. Belief in astrology, alchemy, and magic was widespread. Cardano and Tycho Brahe are outstanding examples of the versatile men of the Renaissance. Besides being a great mathematician, physician, and scientist, Cardano believed in and practiced divination, occultism, and healing by magic. The astronomer Tycho Brahe worked also in astrology and alchemy and produced many medicaments, for example, an elixir against the then common and dangerous epidemic diseases. Both men also made important technical inventions and thus bear witness to the close relationship between science and technology.

The purpose of the present chapter is to refresh the reader's memory on some of the salient historical facts before 1650. It is, however, not possible to point to a simple causal explanation of the development of probability and statistics in terms of these facts, but the record should make it easier for the reader to review and to grasp the spirit of the time.

The exposition is necessarily very brief; it is also biased in the sense that it concentrates on the most conspicuous events in the development of mathematics and natural philosophy, and it emphasizes those events that are of particular interest for the history of probability and statistics.

## 2.2 ON MATHEMATICS BEFORE 1650

Classical Greek mathematics had been nearly forgotten in Western Europe in the early Middle Ages. The Crusades and increasing trade and travel in the Mediterranean countries from about 1100 brought the Europeans into contact with the Arabs and the Byzantines who had preserved the Greek works. During the later Middle Ages and the Renaissance, the classical works were translated, commented upon by European mathematicians, and put to good use in connection with many practical applications, such as navigation, surveying, architecture, and commercial arithmetic. A survey of the existing mathematical knowledge with a view to applications was given by Luca Pacioli (c. 1445–c. 1517) in 1494.

In the 16th century considerable progress was made in arithmetic, algebra, and trigonometry. Zero was accepted as a number, and negative and irrational numbers came gradually into use. Complex numbers occurred in the solution of quadratic equations, but they were considered "useless." The decimal system of notation was introduced for fractions, replacing the ratio of two integers.

Two prominent Italian mathematicians, Niccolò Tartaglia (c. 1499–1557) and Girolamo Cardano (1501–1576), wrote textbooks containing new results in arithmetic and algebra. For example, they gave methods for the solution of equations of the third and fourth degrees, and Cardano noted that the number of roots equaled the degree of the equation.

The French mathematician François Vieta (1540–1603) published several works on plane and spherical trigonometry in which he systematized and extended the formulae for right and oblique plane triangles and for spherical right triangles. He also found many trigonometric identities, for example, the important expression for $\sin nx$ in terms of $\sin x$. Vieta's trigonometric research was mainly inspired by problems in astronomy and surveying; however, he also showed how to use trigonometric formulae for the solution of certain algebraic equations.

Progress in algebra was hampered by the tradition that geometry was the only real mathematics, and algebraic results had therefore to be given a geometrical interpretation. For example, algebraic equations had to be written in homogenous form of at most the third degree. Vieta realized, however, that algebra could be used to prove geometrical results and to handle quantities whether or not they could be given a geometrical interpretation. Thus algebra gradually became a separate mathematical discipline independent of geometry.

The increasing use of mathematics in practice resulted in the computation and publication of many tables, particularly tables of trigonometric functions.

Texts on arithmetic and algebra in the Renaissance were written in a verbal style with abbreviations for special words: for example p for plus, m for minus and R for square root. According to Kline (1972, p. 260), the expression $(5+\sqrt{-15})(5-\sqrt{-15})=25-(-15)=40$ was written by Cardano as

5p: Rm: 15

5m: Rm: 15

25m:m: 15 qd est 40.

Gradually, symbols were introduced for the unknowns and exponents for powers. A decisive step was taken by Vieta, who used letters systematically also as coefficients in algebraic equations. The sign = for equality was proposed about the middle of the 16th century but was not universally accepted. Descartes used $\propto$ as a stylized ae (from *aequalis*), and this was still used by Bernoulli in his *Ars Conjectandi* in 1705. The symbol $\infty$ for infinity was introduced by Wallis in 1655. The letter $\pi$ was introduced in the beginning of the 18th century but de Moivre still used $c$ (derived from circumference) as late as 1756 in his *Doctrine of Chances.*

Essential steps in the free use of letters and special signs for mathematical symbols were first taken by Descartes, Newton, and Leibniz.

The most important advance in arithmetic in the 17th century was the invention of logarithms. The German mathematician Michael Stifel (1486–1567) considered in 1544 the correspondence between terms of an arithmetic and a geometric series and stated the "four laws of exponents," but he did not take the decisive step of introducing logarithms. This was done by John Napier (1550–1617), Laird of Merchiston in Scotland, a prominent politician and defender of the Protestant faith. He published two books on logarithms, the *Descriptio* (1614) and the *Constructio* (1619), the first giving the definitions and working rules of logarithms and a seven-figure table of logarithmic sines and tangents, the second containing theory and proofs. Napier considered the synchronized motion of two points, each moving on a straight line, the one with constant velocity, and the other with a decreasing velocity proportional to the distance remaining to a fixed point, the initial velocity being the same. In modern notation his model may be written as

$$dx/dt = r, \qquad x(0) = 0,$$
$$dy/dt = -y, \qquad y(0) = r,$$

with the solution $x(t) = rt$, and

$$y(t) = re^{-t} = re^{-x/r}, \qquad t \geqslant 0.$$

It follows that the Naperian logarithm, $\log_N y = x$, is a linear function of the natural (or hyperbolic) logarithm

$$\log_N y = r \log_e \frac{r}{y}.$$

Napier constructed his table of logarithms by means of a detailed tabulation of the function $g(x) = r(1-\epsilon)^x$, $x = 0, 1, \ldots,$ for $r = 10^7$ and for small positive values of $\epsilon$. He carried out these calculations personally during a period of nearly 20 years.

Since $\log_N r = 0$, and

$$\log_N(10y) = \log_N y - 23{,}025{,}842.34,$$

Napier realized that his definition of logarithms was unpractical, and in cooperation with Henry Briggs (1561–1630), professor of mathematics first in London and later in Oxford, he proposed the system of common (or Briggsian) logarithms with base 10. From 1615 Briggs devoted the main part of his time to the construction of logarithmic tables. In 1617 he published the first table of common logarithms of the natural numbers from 1 to 1000. This was followed by his *Arithmetica Logarithmetica* (1624) containing the logarithms of the natural numbers from 1 to 20,000 and from 90,000 to 100,000 to 14 decimal places, with an introduction on the construction of the table and examples of arithmetical and geometrical applications. Posthumously occurred his *Trigonometria Brittanica* (1623) containing sines, tangents, and their logarithms to 14 decimal places. Many other tables were published about the same time so that 20 years after Napier's book, a wealth of logarithmic tables was available, and for the next 300 years logarithmic tables were the most important tools for computational work. In *Napier Tercentenary Memorial Volume* (1915) Glaisher writes,

> By his invention Napier introduced a new function into mathematics, and in his manner of conceiving a logarithm he applied a new principle; but even these striking anticipations of the mathematics of the future seem almost insignificant by comparison with the invention itself, which was to influence so profoundly the whole method of calculation and confer immense benefits upon science and the world.

For more details on the history of logarithms, we refer to Naux (1966, 1971) and Goldstine (1977).

The great progress in physics and astronomy in the beginning of the 17th century by Galileo and Kepler had a profound influence on the direction of mathematics. By fitting mathematical equations to data, they obtained simple descriptions of physical phenomena, and they thus demonstrated the usefulness of mathematics in science and technology. All the great contributions to mathematics in the following centuries came from men who were as much scientists as mathematicians.

As natural tools for his work in astronomy Johannes Kepler (1571–1630) worked on interpolation; logarithms; tabulation of trigonometric functions and logarithms; the mathematics of conics, for example, the gradual change of one conic into another by a change of the parameters; the length of curves;

and the areas and volumes limited by curves and surfaces. He calculated such areas and volumes as the sum of a large number of small sections.

In the 1630s and 1640s many mathematicians worked on the area (integration) problem. Cavalieri (1598–1647) invented a method of "indivisibles," a geometrical method for finding areas and volumes by means of an infinite number of equidistant parallel linesegments and areas, respectively. Other mathematicians, such as Fermat, Roberval, Pascal, and Wallis, solved concrete problems either by Cavalieri's method or by approximating the area under a curve by the sum of the areas of suitably chosen rectangles with bases of the same length, letting the number of rectangles increase indefinitely and keeping only the main term of the sum. Using the latter method, Fermat, for example, worked out the integral of $x^n$ over a finite interval for all rational $n$ except $-1$. A general method of integration (and differentiation) had, however, to wait for the works of Newton and Leibniz in the latter part of the century.

Practical problems in optics, perspective, and cartography led Girard Desargues (1591–1661) to use projection and section as a general method in geometry, and he thus founded modern projective geometry. He studied transformation and invariance for the purpose of deriving properties of the conics from those already proved for the circle.

The most influential natural philosopher and mathematician in the first half of the 17th century was René Descartes (1596–1650). Here we shall only mention some mathematical results contained in his *La Géométrie* (1637). Descartes continued Vieta's attempts to introduce better symbolism. For example, he introduced the rule of using the first letters of the alphabet for constants and coefficients and the last for variables and unknowns. He also continued Cardano's and Vieta's algebraic works. He asserted that the number of roots in a polynomial equation $f(x) = 0$ equals the degree of $f(x)$ and, furthermore, that $f(x)$ is divisible by $x - a$ if and only if $f(a) = 0$. He considered algebra to be an extension of logic and independent of geometry and used algebra to solve geometrical construction problems. He founded what today is called analytic or coordinate geometry by introducing (oblique) coordinate axes and defining a curve as any locus given by an algebraic equation. In particular, he studied the conics and the correspondence between their algebraic and geometrical expressions. He also began the study of curves of higher degrees. The important problem of finding the tangent of a curve was solved by a combination of geometrical and algebraic reasoning.

About the same time, and independently, Fermat solved nearly the same problems in analytical geometry, but since his works were not published (they only circulated in manuscript), he did not have the same influence as Descartes.

Frans van Schooten (1615–1660), professor of mathematics in Leiden, translated *La Géométrie* into Latin, which was still the international scientific language. He also added his own commentary and taught Cartesian geometry to his students. Because his translation was much in demand, he published a much enlarged second edition, which besides his own commentaries contained essential contributions by his students Huygens, de Witt, and Hudde, whom we shall meet later in their capacities as probabilists.

The early history of combinatorics is rather obscure (see Biggs, 1979). From numerical examples given by the Indian mathematician Bhaskara about 1150, it seems that he knew the general formulae for the number of permutations of $n$ objects and the number of combinations of $r$ among $n$ objects. From the Hindus this knowledge spread to the Europeans through the Arabs.

The binomial expansion and the corresponding arithmetical triangle of coefficients originated also among Hindu and Arab mathematicians. The arithmetical triangle and its construction are explained by the Arab mathematician al-Tusi in 1265 but is not found in European works before the 16th century.

The relationship between the combinatorial formula and the binomial coefficient was recognized by Marin Mersenne (1588–1648) in 1636. Finally, a unified theory of the combinatorial numbers, the figurate numbers, and the binomial coefficients was developed by Pascal in 1654 and published in 1665.

The previous remark that the early history of combinatorics is rather obscure is no longer true after the publication of *Pascal's Arithmetical Triangle* by Edwards (1987). Here the history of the arithmetical triangle is traced back to Pythagorean arithmetic, Hindu combinatorics, Arabic algebra, and Chinese and Persian mathematics, and a meticulous study of the development in Europe is given, comprising contributions by Tartaglia, Cardano, Stifel, Mersenne, and Pascal and ending with the use of the binomial coefficients in the works of Wallis, Newton, Leibniz, and Bernoulli. We shall return to this book in §§4.3 and 5.2.

The above sketch of the history of mathematics before 1650 is essentially based on the book by Kline (1972), where details and references may be found.

## 2.3 ON NATURAL PHILOSOPHY BEFORE 1650

The 12th century saw the rise of the European university with its four faculties: theology, law, medicine, and the arts. The curriculum of the arts embraced grammar, logic, rhetoric, arithmetic, music, geometry, and astronomy. Most teachers and scholars had a clerical education, and Latin, the official language of the Church, became the universal scholarly language. University teaching and research were dominated by Christian theology and the heritage of the Greeks. The classical Greek works were read in Latin translation, often obtained by translating Arab editions. The most important works were by Aristotle in philosophy, ethics, and logic; Euclid in geometry; Ptolemy in astronomy; and Galen in medicine.

In the 12th and 13th centuries many philosophers wrote commentaries on the Scriptures and the Greek classics for the purpose of developing a general philosophy uniting these two lines of thought. The attempts to reconcile Aristotelian ideas with Christian theology resulted in a firmly established philosophy which has been called Scholasticism. Among the scholastic philosophers the most famous was Thomas Aquinas (1225–1274) whose system of thought was later authorized by the Catholic Church as the only right one. It became the dominant philosophy for about 400 years.

Aristotle's scientific results were considered authoritative, but his inductive–deductive method by which these results had been obtained was pushed into the background. Scholasticism took over Aristotelian logic with its laws of reasoning (the doctrine of syllogism) and used it to create a system of theological explanations of both natural and supernatural phenomena. It also adopted Aristotle's natural philosophy with its teleological explanations, its distinction of sublunary matter into four elements (earth, water, air, and fire), and its conception of an immutable universe of celestial bodies set into motion by God. The natural motion of terrestrial elements was supposed to be linear, whereas the motion of the heavenly bodies was circular. The immutability of the universe was in agreement with the Scriptures and with the deterministic outlook of Christian theology, which supposed that everything was created by the will of an all-powerful God.

The exegetic and speculative nature of scholasticism led to opposition, particularly among English philosophers, who advocated the importance of observation, experimentation, and induction in natural philosophy as a supplement to the revelations in the Scriptures. The most prominent spokesman of this school was William of Ockham (1285–1349), who is most known today for the maxim called "Ockham's Razor": "Entities are not to be multiplied without necessity." This philosophical principle of economy of the number of concepts used in the construction of a theory had great importance for the development of logic, mathematics, and natural philosophy.

A long period of gradual progress in wealth and knowledge was disrupted by the Black Death in 1348. In just a few years about one-third of the population of Europe died of the plague. It took about 100 years for Europe to recover from this catastrophy, which was further aggravated by the Hundred Years' War (1338–1453) between England and France.

The second half of the 15th century was a period with many inventions made by practical men, such as artisans, architects, shipbuilders, and engineers. A considerable metal-working and mining industry was developed. For example, rails were used to further transport in mines, and suction pumps were constructed for drainage. Many inventions of importance for shipping and warfare were made. Charts and instruments for navigation, such as the magnetic compass, the quadrant, and the astrolabe, were developed, and ocean-going sailing ships were built. The effectiveness of gunpowder in warfare was greatly improved by the construction of cannons and handguns. These inventions were the necessary conditions for the great voyages of exploration about 1500 and the following conquests overseas with their wide-ranging consequences for daily life in Europe.

About 1450 printing by movable type was invented, and Gutenberg set up his printing press in Mainz. The printing of illustrations from engraved metal plates also became common. The new techniques spread rapidly all over Europe. The transition from handwritten to printed books was a technological advance with revolutionary effects not only in the world of learning but in religion, politics, art, and technology as well. The next 50 years saw the printing of thousands of books both in Latin and the vernacular such that the knowledge hitherto accumulated in libraries for the few suddenly became available for the many, particularly for the laity. Printed books helped to spread the culture of the Renaissance and to further the Reformation.

The extraordinary growth of trade and industry in the Italian city-states in the 15th century created a new class of merchants and bankers who used their great wealth to support not only artists but artisans and scientists as well for the purpose of furthering the development of useful methods for the new competitive capitalist economy. The intellectual outlook gradually changed from the authoritative scholastic philosophy to a more independent way of thinking based on observations and experiments.

Freedom of thought also spread to religious matters through the Reformation in the first half of the 16th century. Protestantism in its various forms (Lutherans, Calvinists, Huguenots, Puritans, Presbyterians) spread all over Europe (except for Italy and Spain, which remained Catholic). The bible was translated from Greek into the national languages, commentaries were issued, and the laity were encouraged to read and interpret the Scriptures themselves. The struggle between Protestantism and Catholicism, mixed with strong economical and political interests, led in many countries to civil war, with increased intolerance and orthodoxy on both sides.

Among the measures taken by the Catholic Church to identify and suppress its opponents were the establishment of the Holy Office, the Inquisition in Rome in 1542, and the *Index librorum prohibitorum* (Index of prohibited books) in 1559. This list came to comprise the famous books by Copernicus, Galileo, Kepler, and Descartes. In due course, when the

Lutheran, Calvinian, and Anglican Churches acquired secular power, they also forced their dogmas on society, in particular on schools and universities. Protestantism, however, did not develop an all-embracing philosophy covering not only religious and moral but also scientific matters, and the conflict between the Reformed Churches and the ongoing scientific revolution therefore became less severe.

The works of Pythagoras and Plato were studied with renewed interest during the Renaissance, and their idea of a rational and harmonious universe describable in mathematical terms was taken over. The Church accepted the idea that an omniscient God has created the universe according to simple mathematical laws, originally unknown to man. It therefore became a praiseworthy enterprise to study and disclose these laws to obtain a better understanding of God and his "original design."

The Catholic Church, however, considered the laws of nature found by the scientists as hypotheses or practical computational devices, which were accepted as true explanations of nature only if they did not contradict the dogmas of the Church. This attitude gave rise to endless conflicts between scientists and the dogmatists of the Church, which seriously restricted the lives and modes of expression of such men as Copernicus, Cardano, Galileo, Kepler, and Descartes.

The scientific revolution in the 16th and 17th centuries, which forms the foundation of modern mathematics and science, may summarily be considered to begin with the publication of Nicholas Copernicus' *De Revolutionibus Orbium Coelestium* (On the Revolutions of the Celestial Spheres) in 1543 and reaching its culmination by Isaac Newton's *Philosophiae Naturalis Principia Mathematica* (The Mathematical Principles of Natural Philosophy) in 1687.

Among astronomers, Copernicus' book was received as a great work in mathematical astronomy comparable only to Ptolemy's *Almagest*, which had been the basis for all astronomy since about A.D. 150. However, Copernicus' heliocentric model gave no better predictions of phenomena than did Ptolemy's geocentric model, and Copernicus did not provide any empirical evidence for his hypotheses. The strength of the Copernican model was its simple and harmonious explanations of planetary motions and the natural ordering of the earth and the planets in relation to the sun. A few astronomers accepted the Copernican ideas and published books with improved versions of Copernicus' mathematics and tables, but general acceptance had to wait until the beginning of the 17th century.

Beyond the small circle of astronomers, Copernicus' main ideas of the diurnal rotation of the earth and the earth's yearly revolution around the sun just as another planet were generally rejected and ridiculed as being at variance with experience and with the Scriptures. Luther, Melanchthon, and