PROBABILITY THEORY

Teugels J. L.
University Center for Statistics, Catholic University of Leuven, Belgium

Keywords: chance mechanisms, games, gambling, problem of points, mathematical expectation, duration of play, classical ruin problem, matching problem, St. Petersburg paradox, law of large numbers, insurance, annuities, demography, theory of errors, least squares, average man, needle problem, axiomatization, St. Petersburg School, stochastic process, spatial statistics, time series, stochastic hydrology, dam theory, wind, reliability theory, wildlife sampling, forestry, fishery, design and analysis of experiments, weather modification

Contents

1. Introduction: Chance Mechanisms
1.1. Divination
1.2. Dispute Solving
1.3. Games
1.4. Gambling
1.5. Scientific Evidence
2. Early Concepts of Probability
2.1. Ancient Times
2.2. Renaissance
2.3. The Problem of Points
2.4. Other Problems from Gambling
3. The First Steps Towards a Theory of Probability
3.1. New Contributions
3.2. Still More Gambling Problems
3.3. Theoretical Contributions
4. Earliest Applications
4.1. Insurance
4.2. Demography
4.3. The Theory of Errors
4.4. Social Sciences
4.5. Geometry
5. The Axiomatization of Probability Theory
5.1. Doubts about Probability Theory
5.2. The Need for Axiomatization
5.3. The Russian School
6. Probability and Statistics in Life Support Systems
6.1. Spatial Statistics
6.2 Time Series
6.3. Image Analysis
6.4. Simulation
7. Water

7.1. Rainfall
7.2. Water Storage
7.3. Floods
7.4. Environmental Issues
7.5. Seawater
8. Energy
8.1. Stochastic Hydrology
8.2. Wind
8.3. Reliability Theory
9. Environment
9.1. Wildlife Sampling
9.2. Abundance
9.3. Forestry
9.4. System Analysis
9.5. Exposure Data
9.6. Environment and Health
10. Food
11. Agriculture
11.1. Design and Analysis of Experiments
11.2. Weather Modification
Acknowledgments
Glossary
Bibliography
Biographical Sketch

Summary

Chance mechanisms are part of the surrounding world. Life-supporting systems, in particular, have many built-in stochastic aspects. Some of these are mentioned here as illustrations and to put them into a global framework. The statistical methodology needed to deal with them is taken up in greater detail in later sections. In order to put the developments in their proper context, a short history is given of the evolution of probabilistic and statistical thinking up until the early twentieth century.

1. Introduction: Chance Mechanisms

There is a wide variety of contexts in which the concept of chance and/or probability has been used. The following examples are meant to indicate to the reader that probabilistic concepts are sometimes hidden in real life situations.

1.1. Divination

From ancient times, chance mechanisms have been used in divination. The flight of birds and the intestines of offerings are prime examples. Chance mechanisms have also been used in divination in the Bible. Leviticus ch. 16, vv. 5–10, for example, tells how Aaron cast lots over two goats to decide which one should be dealt with as an atonement. In present-day Shinto temples, the visitor is offered a box with a large number of long sticks. The visitor is expected to shake the box till one of the twigs sticks out. The stick will correspond to a specific fortune printed on a piece of paper which is retained if the fortune is a good one. If it is not, the piece of paper is fastened to a tree or a fence in the neighborhood of the temple.

1.2. Dispute Solving

Drawing straws when assigning duties is a typical instance of using chance mechanisms in daily life. The toss of a coin at the beginning of various games decides which team should start playing on which half of the field. In the Talmudic tradition the allocation of daily duties in the temple is organized by a randomization procedure.

1.3. Games

This is probably the best known and most popular application.

- The tomb of Tutankhamen contained a beautiful checker-like game—"Hounds and Jackals"—intended to help the deceased arrive safely on the other side. Successive steps on the squares of the game by the hounds and the jackals were decided by throwing astragali (made from the hard part of the foot of certain animals) or long flat two-sided sticks.
- The throwing of dice controls most pastime family games. Needless to say, all card games like whist and bridge are based on the random allocation of cards.
- Here is an example from an unexpected angle. In the eighteenth century sheet music was still rather expensive. In order to enable the common people to play music, classical composers like Haydn, Mozart, Kirnberger, and Emmanuel Bach amused themselves and their clientele by composing minuets, marches, waltzes, and so on. In these pastimes, previously written musical figures are provided by the compiler of a game: the player chooses the music for each measure by throwing dice, spinning a top, or choosing a number at random. Random mechanisms are also advocated in contemporary music in aleatoric composition.

1.4. Gambling

Most people will be unaware that the early development of the scientific theory of probability was greatly influenced by gambling. The next section illustrates this with more concrete examples. Gambling, from which present-day casinos have emerged, seems to have been one of the oldest human pastimes. Apart from the one-armed bandit or slot machine, most casino games (like roulette, baccarat, and craps) are sufficiently standardized for the probabilities of winning to be known exactly.

Lotto and lotteries also fall into this category since the drawing mechanisms are so refined that no specific outcomes are more likely than any others. In this sense, the above casino games and lotteries are perfect games of chance. For a few other casino games (like blackjack), specific strategies have been developed that are more favorable than random playing. Up to a certain extent, they can be considered as needing some form of skill.

1.5. Scientific Evidence

As another illustration of the use of chance mechanisms, one could give a vast number of applications from statistics. Here are a few less usual examples.

- The existence and measurability of telekinesis has been a touchy subject. In a number of cases, experiments have been described and analyzed that should determine once and for all whether or not telekinesis exists. Outcomes of such experiments with a medium are compared with expected outcomes under random circumstances. Substantial differences between the two sets of outcomes should then indicate some unexplainable capability of the medium.
- Since its scientific maturity, probability theory has been invoked in lawsuits. A number of examples will be given in later sections. If one combines a sufficient number of external properties of individuals and if it is taken as given that these characteristics are independent, then it is always possible to come up with somebody unique. A famous example is the *People versus Collins* case in which such a set of independence hypotheses between different events was used.
- More recently, legal verdicts have been based on DNA investigations. The forerunner of this procedure in forensic studies was the evaluation of fingerprints. From delicate calculations, scientists have discovered that specific replications in DNA patterns are so unlikely that they can hardly appear in two different creatures.

2. Early Concepts of Probability

The emergence of probabilistic concepts has been slow and unclear. One should not make the mistake of looking at historic facts with modern eyes. Most people by now have some feeling of the frequentist concept of probability. To be more specific, the probability of a certain event E in a random experiment is approximately the ratio of the frequency that E appears in independent repetitions of the experiment to the number of repetitions. This interpretation however is of rather recent origin.

2.1. Ancient Times

It is noteworthy that the Greeks valued certain throws of astragali higher than others, irrespective of their relative probabilities. Presumably, the odd shape of these astragali made estimation—a modern concept!—of the relevant probabilities impossible. As suggested by historians, the use of astragali and the drawing of lots in divination may have prevented a scientific study of the outcomes of games of chance for religious reasons.

2.2. Renaissance

It took until the middle of the seventeenth century before gamblers started to notice measurable differences between events that have nearly the same probability of occurrence. For example, Galileo wrote a paper on the observation that dice players consider 10 to be more advantageous than 9 when throwing three dice simultaneously. It can be shown by enumeration that the probability of 10 is 27/216 while that of 9 is 25/216. Modern statistics tells us that one needs very sizable experiments to prove statistically that 9 and 10 are not equally likely.

In Europe, the making of playing cards was established by the fifteenth century. However, cards were expensive and thus only the wealthy could afford to buy them. There was a tendency among scientists to formulate probabilistic problems in terms of games. In the sixteenth century and even later, only people of substantial means in terms of time and money could afford to spend their energy on scientific questions. They were also familiar with gambling terminology for the same reason.

Present-day probability texts often state problems in terms of urn models where the experimenter is asked to draw balls from urns with a given composition. This is very much like stating them in terms of gambling, which might be less familiar to many people. Urn models have the additional advantage that, depending on the specific type of application, the number of urns and their composition can often be easily interpreted. For example, urn models have been of standard usage in modeling the impact of contagious diseases in populations; also, basic and important concepts about dependent trials are often illustrated using urns with changing compositions.

2.3. The Problem of Points

From about 1500 Italian mathematicians tried to solve the following problem, known as the "problem of points" or the "division problem": two players, A and B, agree to play a series of fair games until one of them has won a specified number of games, s say. The game is interrupted when A has won a ($< s$) games while B has won b ($<s$) games. How should the stakes be divided?

The problem of points is probably the most significant instigator of probabilistic thinking. The problem itself is, of course, only one link in a long chain of successive findings. However, because of its prime importance within the development of the origins of probability theory, it deserves special treatment as a trendsetter.

The problem of points, or division problem, seems to be a very old one, and already appears in 1494 in the writings of Lucia Pacioli (ca. 1445–ca. 1514) for the case $a = 5$, $b = 2$ and $s = 6$. His suggestion is to divide the stakes in the ratio a/b. There is no probabilistic or combinatorial thinking behind his reasoning.

In 1556 Niccolo Tartaglia (ca.1499–1557) criticizes Pacioli and suggests a ratio $(s + a – b)/(s – a + b)$ if the stakes are equal. Again no scientific reasoning is provided. L. Forestani (?–1623) reasons in 1603 that Pacioli's proportion should be corrected. His view is that, allowing for chances in the games that are not yet played, fortune may be reversed. He suggests the ratio $(2s – 1 + a – b)/(2s –1 – a + b)$ where $2s –1$ is the maximum number of games that could be played anyway.

Girolamo Cardano (1501–1576) is the first to note (1539) that the division should only depend on the number of games each player is lacking in order to win, that is to say, on the quantities $m = s – a$ and $n = s – b$. Using a kind of obscure inductive argument Cardano settles for the ratio $n(n + 1)$ over $m(m + 1)$. It should be mentioned that Cardano wrote the influential *De Ludo Aleae*, a treatise on the moral, practical, and theoretical aspects of gambling with dice. He clearly states that all six sides of an honest dice are equally likely to be thrown, and he introduces chance as the ratio between the number of favorable cases and the number of equally possible cases.

The problem was ultimately solved in a fascinating correspondence between Pierre de Fermat (1601–1665) and Blaise Pascal (1623–1662) between July and October 1654. In modern terminology, the division should be done according to a proportion. This ratio should be the same as the ratio of the probabilities of winning the whole stake for each one of the players, if the game was played till the end. Through the construction of his famous arithmetic triangle, Pascal gives an explicit expression in a form that is rather too complicated to be written out here. It should be mentioned that Pascal gave a second solution, this time based on recursions. Fermat also had a solution but he employed a waiting-time argument to solve the problem in its full generality.

Christiaan Huygens (1629–1695) visited France in 1655 and heard about the problem of points. Not having access to its solution, Huygens solved the problem on his return to Holland. His 16-page manuscript *Van Rekeningh in Spellen van Geluck* (On Reckoning in Games of Chance) is considered to be the first printed treatise on probability theory. From an axiom on the value of a fair game, Huygens derives 14 propositions. The first three can be considered as founding the definition of what is now called “mathematical expectation.” Proposition 4 solves the problem of points. In his last proposition he even anticipates the use of conditional expectations. Huygens’s booklet was immediately translated into Latin in 1657 and a variety of other languages. It gained wide popularity as the only textbook prior to 1700.

2.4. Other Problems from Gambling

As has already been said, most scientists were familiar with the language of gambling, and, in this respect, the problem of points is no exception. Another such problem is that of the “duration of play.” It was suggested by Huygens and goes as follows: two players A and B have probability p and q respectively of winning a game. At the beginning of the game A has a coins while B has b coins. The loser in each round has to give one coin to his opponent. What is the probability that A will lose all his capital before winning the whole capital of B? An associated question is the determination of the average number of games played before A or B is ruined.

This problem became famous in risk theory where it is known as the “classical ruin problem.” As a limiting case, B is often taken to be a casino when the initial capital b is considered very large. De Moivre has given a full solution of the duration problem.

3. The First Steps Towards a Theory of Probability

Rapid development becomes apparent in the early years of the eighteenth century.

3.1. New Contributions

The publication of the *Essay d’Analyse sur les Jeux de Hazard* (1708) by Pierre Rémond de Montmort (1678–1719) was the first testament in the eighteenth century to an explosion of activity and competition. Referring as it does to an earlier observation about gambling, it is not surprising that the fourth part of de Montmort's book contains solutions of various problems dealing with games.

A second key publication was the *Ars Conjectandi* (1713) by James Bernoulli (1654–1705). The first part of this book is actually a comprehensive and annotated edition of Huygens's treatise. Further, Abraham de Moivre (1667–1754) published *The Doctrine of Chances* in 1718. This became the standard text for the remaining part of the century, to be superseded only in 1812 by Laplace's *Théorie Analytique des Probabilités*. The correspondence between Montmort and Nicholas Bernoulli (1687–1759), James's nephew, was also influential and appeared in the second edition of Montmort's book in 1713.

3.2. Still More Gambling Problems

Nicholas Bernoulli's name is associated with another problem from gambling, the game known as *Treize* or the "matching problem." Two players A and B each have a complete deck of n cards. Both draw cards simultaneously and one by one. If they match then A wins and if there is no match B wins. The problem is to determine the probability of winning for each of the two players.

It is easy to check that with three cards A has probability 2/3 while B only has probability 1/3. It takes a bit of effort to solve the problem for $n = 4$ where A has probability 5/8. Nicholas Bernoulli gave the general solution for finite n while de Moivre treated the case where $n = \infty$. In this case, the result involves the quantity e, the base of the natural logarithm. In current literature, the problem is known as the rencontre game and falls under the heading of "secretary problems."

Yet another problem from gambling had a very profound influence on the development of probability theory. It was also formulated by Nicholas Bernoulli but treated by Daniel Bernoulli (1700–1782) in a paper published by the St. Petersburg Imperial Academy of Sciences, which is why it entered the literature as the "St. Petersburg paradox." The game goes as follows: two players A and B toss a coin till it first comes up heads. Player B has to give two coins to player A if it comes up heads on the first toss, four coins if it is first heads on the second toss, eight if on the third, and so on. The question is to determine the amount that A should pay B at the start of the game to make it fair.

According to Huygens's rules for the mathematical expectation, the game can only be fair if A pays the same number of coins as B is expected to pay. At the nth play, this expectation is $2(1/2) + 4(1/4) + \dots 2n(1/2n) = n$. This means that, when no primary restriction is put on the number of plays, the game can only be fair if A pays an infinite amount to B. It is hard to imagine now how this kind of "solution" would have been received by the solver's contemporaries. Since its appearance in 1738 the St. Petersburg paradox has been the subject of an incredible number of scientific, philosophical, and also popular contributions.

Isaac Newton (1642–1727) could also be mentioned in connection with the early history of probability for solving a dice problem for Samuel Pepys in 1693. More intriguing is his tacit understanding of a confidence interval for the arithmetic mean when he stated that the mean length of a king's reign is "about eighteen or twenty years apiece."

-
-
-

Bibliography

Armitage P. and Colton T. (eds.) (1998). *Encyclopedia of Biostatistics*, Vol. 1–6, 4898pp. Chichester: John Wiley. [This encyclopedia focuses on medical and health sciences and should be seen as complementary to the one listed below by Kotz, Johnson and Read. It aims at a wide coverage of general statistical theory and concentrates on technical and practical aspects of biostatistics. Contributions, even the more mathematical articles, are linguistically treated so that the less mathematical reader can gain some general understanding. Most articles contain citations of the scientific literature while some include a selected bibliography for additional reading.]

Barndorff-Nielsen O.E., Gupta V.K., Perez-Abreu V. and Waymire E. (eds.) (1998). *Stochastic Methods in Hydrology. Rain, Landforms and Floods*, 207pp. (Proceedings of the CIMAT conference, Guanajuato, Mexico, 1998). Singapore: World Scientific. [Five overview lectures on stochastic and statistical methods in hydrology. Together they provide contemporary mathematical and statistical developments in river basin hydrology as they pertain to space-time rainfall, spatial landform and network structures, averages and fluctuations in water balance.]

Barnett V. and Turkman K.F. (1996, 1994). *Statistics for the Environment* (SPRUCE Proceedings) 427pp, 391pp. (Conferences Lisbon, Portugal, 1992, and Harpenden, UK, 1993. Chichester: John Wiley. [The two volumes together give a fine survey of the stochastic subjects relevant for the environment. Particular topics treated are environmental statistics, monitoring and sampling, pollution and contamination, climatology, fish populations, and forestry, plus water-related statistical subjects like rainfall and climate, sea levels and wave energy, water quality, supply, and management, and hydrological modeling.]

Christakos G. and Hristopulos D.T. (1998). *Spatiotemporal Environment Health Modelling: A Tractatus Stochasticus*, 400pp. Boston: Kluwer Academic. [This book uses stochastic concepts and methods to build links between models and techniques of environmental sciences on the one hand, and health sciences on the other. Special attention is given to associations between contaminant transport models and human exposure, between ozone distribution and disease incidence rates, between carcinogenesis and environmental conditions. Computational tools for immediate practical use are included.]

Federer W.T. (1993, 1998). *Statistical Design and Analysis for Intercropping Experiments. Vol. 1: Two Crops*, 298pp, *Vol. 2: Three or More Crops*, 262pp. New York: Springer-Verlag. [Detailed analysis of statistical issues dealing with two crops and with intercropping experiments; statistical complexities in experiments dealing with more than two crops via intercropping, based on mixing of species simultaneously or sequentially. Every chapter ends with a list of useful references, a problem set, a comprehensive summary, and a discussion. Mathematical proofs are delegated to appendices, making the book easy to read.]

Ganoulis J.G. (1994). *Engineering Risk Analysis of Water Pollution. Probabilities and Fuzzy Sets*, 306pp. Weinheim: VCH. [This book deals with a quantitative analysis of environmental issues related to water quality of natural hydrosystems. Risk and reliability in water quality are analyzed from the engineering point of view. The environmental impact from wastewater disposal on rivers, groundwater, and coastal areas is modeled. The book does not presuppose knowledge of fuzzy set theory.]

Hald A. (1990, 1998). *A History of Probability and Statistics and Their Applications before 1750*, 586pp, *A History of Probability and Statistics and Their Applications from 1750 to 1930*, 795pp. New York: John Wiley. [Currently the most comprehensive and complete history of the emergence of probabilistic and statistical (but non-mathematical) thinking up to the middle of the eighteenth century. This book also deals with aspects from demography and insurance over the same period and contains an extraordinarily complete bibliography. The second part also refers to the applications of stochastic thinking in applications.]

Klein J.L. (1997). *Statistical Visions in Time. A History of Time Series Analysis, 1662–1938*, 345pp. Cambridge: Cambridge University Press. [Historical text that emphasizes the role played by time series analysis within the development of statistics.]

Kotz S., Johnson N.L. and Read C.B. (eds.) (1982–1989). *Encyclopedia of Statistical Sciences*, Vol. 1–9 + Suppl. 6536pp. New York: John Wiley. [Systematic survey of all concepts and procedures in probability theory, stochastic processes, statistics, and all their applications. Moreover, a substantial part of the work is composed of historical material, including biographies of some of the key people in the field. All articles are covered in depth and contain a wealth of bibliographical information. Because of its completeness, the encyclopedia is considered the best basic reference on all stochastic subjects, particularly for material before 1989. A revision of this encyclopedia is currently under investigation.]

Ott W.R. (1995). *Environmental Statistics and Data Analysis*, 313pp. Boca Raton: CRC Press. [Elementary statistical concepts and tools are used in depicting and analyzing statistical data regarding pollutants. Traditional distributions are employed to model pollution data and robustness of the models is analyzed. This book provides enough illustrative material to yield immediate applicability.]

Patil G.P. and Rao C.R. (eds.) (1994). *Environmental Statistics*, 927pp. Amsterdam: North-Holland. [This volume provides an initial perception of environmental statistics with examples of major concern, opening perspectives on research, training, policy, and regulation involving statistics and environment. It can be considered as a primer in the field even though most single topics have received a fuller treatment elsewhere.]

Prabhu N.U. (1997). *Stochastic Storage Processes. Queues, Insurance Risk, Dams, and Data Communication*. 2nd edn, 206pp. New York: Springer. [Detailed treatment of storage models within the framework of continuous time. The single server queue is treated first while the second part deals with continuous time storage models. Special attention is given to insurance problems and to the theory of dams. The inclusion of problems and bibliographic information makes the book suitable as a textbook.]

Seber G.A.F. (1982). *The Estimation of Animal Abundance and Related Parameters*, 2nd edn, 654pp. London: Griffin. [Full treatment of capture-recapture procedures, and applications, as well as an extensive bibliography.]

Sheynin O. (ed.) (2000). *From Daniel Bernoulli to Urlanis. Still more Russian papers on probability and statistics*, 229pp. Egelsbach: Hänsel-Hohenhausen. [The Russian school has made important contributions to the development of stochastic concepts. This volume deals with some of the earliest contributions.]

Stigler S.M. (1999). *Statistics on the Table. The History of Statistical Concepts and Methods*, 488pp. Cambridge, MA: Harvard University Press. [A lively collection of essays examining in witty detail the history of some concepts and pitfalls when bringing statistics to the table.]

Thompson M.E. (1997). *Theory of Sample Surveys*, 305pp. London: Chapman and Hall. [Good source of the main sampling techniques and their probabilistic and statistical treatment.]

Wu C.F.J. and Hamada M. (2000). *Experiments. Planning, Analysis, and Parameter Design Optimization*, 630pp. New York: John Wiley. [A prime merit of this volume is that it extends the applications of the theory from the classical agricultural context to a platform including economic, medical, and social topics. The book includes a wealth of new methodologies besides the traditional material on design and analysis of experiments. New items are robust parameter design and designs with complex aliasing when conducting economical experiments. Generalized linear models and Bayesian methods are used for non-normal data.]

Biographical Sketch

Jef L. Teugels was born in Londerzeel, Belgium, in 1939. After obtaining the licence degree in Mathematics at the Katholieke Universiteit Leuven in 1963, he received both his M.Sc. (1966) and his Ph.D. (1967) at Purdue University, USA. He then joined the Katholieke Universiteit Leuven as Assistant Professor in 1967 and Full Professor in 1973. He taught probability and statistics courses to majors in mathematics, physics, chemistry, agriculture, and social sciences.

He has held visiting positions at Université Lovanium (Zaire), Universidade de Coimbra (Portugal), University of Cambridge (UK), University of British Columbia (Canada), Naval Postgraduate School (USA), University of North-Carolina (USA), Universitas Katolik Parahyangan (Indonesia), University of California at Santa Barbara (USA), Australian National University, University of Maryland (USA), and Keio University (Japan).

He co-authored three books and some 100 papers in the statistical literature. His main research areas started with queuing theory and analytic probability, moving gradually to stochastic and statistical issues in insurance, extreme value analysis, and environmetrics. He acts as associate editor of *Insurance: Mathematics and Economics*, the *Journal of Applied Probability*, *Advances of Applied Probability*, the *Wiley Series in Probability and Mathematical Statistics*, *Extremes*, *Journal of Statistical Planning and Inference*, *Environmetrics*, and is now editor-in-chief of *Applied Stochastic Models in Business and Industry*. He is also section editor of the *Encyclopedia of Environmetrics* (Wiley).

During the period 1975–1985 he was scientific secretary of the Bernoulli Society of which he became the president in 1995–1997. He has been on a variety of scientific and organizing committees. He is a member of the Belgian Mathematical Society, the Belgian Statistical Association, the London Mathematical Society, and the International Statistical Institute, and became a fellow of the Institute of Mathematical Statistics in 1985. He is currently vice-president of the International Statistical Institute.