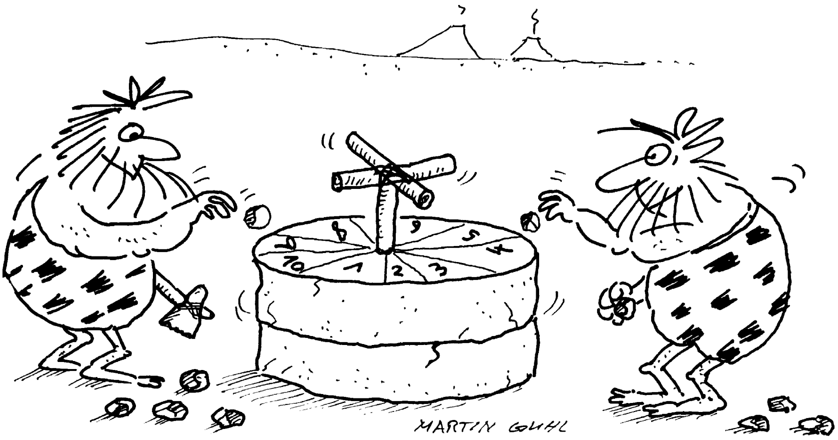
Introduction

It is difficult to say who had a greater impact on the mobility of goods in the preindustrial economy: the inventor of the wheel or the crafter of the first pair of dice. One thing, however, is certain: the genius that designed the first random-number generator, like the inventor of the wheel, will very likely remain anonymous forever. We do know that the first dice-like exemplars were made a very long time ago. Excavations in the Middle East and in India reveal that dice were already in use at least fourteen centuries before Christ. Earlier still, around 3500 B.C., a board game existed in Egypt in which players tossed four-sided sheep bones. Known as the *astragalus*, this precursor to the modern-day die remained in use right up to the Middle Ages.

In the sixteenth century, the game of dice, or craps as we might call it today, was subjected for the first time to a formal mathematical study by the Italian



The cartoons in the book are supplied by www.cartoonstock.com mathematician and physician Gerolamo Cardano (1501–1576). An ardent gambler, Cardano wrote a handbook for gamblers entitled *Liber de Ludo Aleae* (The Book of Games of Chance) about probabilities in games of chance. Cardano originated and introduced the concept of the set of outcomes of an experiment, and for cases in which all outcomes are equally probable, he defined the probability of any one event occurring as the ratio of the number of favorable outcomes and the total number of possible outcomes. This may seem obvious today, but in Cardano's day such an approach marked an enormous leap forward in the development of probability theory. This approach, along with a correct counting of the number of possible outcomes, gave the famous astronomer and physicist Galileo Galilei the tools he needed to explain to the Grand Duke of Tuscany, his benefactor, why it is that when you toss three dice, the chance of the sum being 10 is greater than the chance of the sum being 9 (the probabilities are $\frac{27}{216}$ and $\frac{25}{216}$, respectively).

By the end of the seventeenth century, the Dutch astronomer Christiaan Huygens (1629–1695) laid the foundation for current probability theory. His text *Van Rekeningh in Spelen van Geluck* (On Reasoning in Games of Chance), published in 1660, had enormous influence on later developments in probability theory (this text had already been translated into Latin under the title *De Ratiociniis de Ludo Aleae* in 1657). It was Huygens who originally introduced the concept of expected value, which plays such an important role in probability theory. His work unified various problems that had been solved earlier by the famous French mathematicians Pierre Fermat and Blaise Pascal. Among these was the interesting problem of how two players in a game of chance should divide the stakes if the game ends prematurely. Huygens' work led the field for many years until, in 1713, the Swiss mathematician Jakob Bernoulli (1654–1705) published *Ars Conjectandi* (The Art of Conjecturing) in which he presented the first general theory for calculating probabilities. Then, in 1812, the great French mathematician Pierre Simon Laplace (1749–1827) published his *Théorie Analytique des Probabilités*. This book unquestionably represents the greatest contribution in the history of probability theory.

Fermat and Pascal established the basic principles of probability in their brief correspondence during the summer of 1654, in which they considered some of the specific problems of odds calculation that had been posed to them by gambling acquaintances. One of the more well known of these problems is that of the Chevalier de Méré, who claimed to have discovered a contradiction in arithmetic. De Méré knew that it was advantageous to wager that a six would be rolled at least one time in four rolls of one die, but his experience as gambler taught him that it was not advantageous to wager on a double six being rolled at least one time in 24 rolls of a pair of dice. He argued that there were six possible outcomes for the toss of a single die and 36 possible outcomes for the toss of a pair of dice, and he claimed that this evidenced a contradiction to the arithmetic law of proportions, which says that the ratio of 4 to 6 should be the same as 24 to 36. De Méré turned to Pascal, who showed him with a few simple calculations that probability does not follow the law of proportions, as De Méré had mistakenly assumed (by De Méré's logic, the probability of at least one head in two tosses of a fair coin would be $2 \times 0.5 = 1$, which we know cannot be true). In any case, De Méré must have been an ardent player in order to have established empirically that the probability of rolling at least one double six in 24 rolls of a pair of dice lies just under one-half. The precise value of this probability is 0.4914. The probability of rolling at least one six in four rolls of a single die can be calculated as 0.5177. Incidentally, you may find it surprising that four rolls of a die are required, rather than three, in order to have about an equal chance of rolling at least one six.

Modern probability theory

Although probability theory was initially the product of questions posed by gamblers about their odds in the various games of chance, in its modern form, it has far outgrown any boundaries associated with the gaming room. These days, probability theory plays an increasingly greater roll in many fields. Countless problems in our daily lives call for a probabilistic approach. In many cases, better judicial and medical decisions result from an elementary knowledge of probability theory. It is essential to the field of insurance.† And likewise, the stock market, "the largest casino in the world," cannot do without it. The telephone network with its randomly fluctuating load could not have been economically designed without the aid of probability theory. Call-centers and airline companies apply probability theory to determine how many telephone lines and service desks will be needed based on expected demand. Probability theory is also essential in stock control to find a balance between the stock-out probability and the costs of holding inventories in an environment of uncertain demand. Engineers use probability theory when constructing dikes to calculate the probability of water levels exceeding their margins; this gives them the information they need to determine optimum dike elevation. These examples

† Actuarial scientists have been contributing to the development of probability theory since its early stages. Also, astronomers have played very important roles in the development of probability theory.

underline the extent to which the theory of probability has become an integral part of our lives. Laplace was right when he wrote almost 200 years ago in his Théorie Analytique des Probabilités:

> The theory of probabilities is at bottom nothing but common sense reduced to calculus; it enables us to appreciate with exactness that which accurate minds feel with a sort of instinct for which ofttimes they are unable to account. . . . It teaches us to avoid the illusions which often mislead us; . . . there is no science more worthy of our contemplations nor a more useful one for admission to our system of public education.

Probability theory and simulation

In terms of practical range, probability theory is comparable with geometry; both are branches of applied mathematics that are directly linked with the problems of daily life. But while pretty much anyone can call up a natural feel for geometry to some extent, many people clearly have trouble with the development of a good intuition for probability. Probability and intuition do not always agree. In no other branch of mathematics is it so easy to make mistakes as in probability theory. The development of the foundations of probability theory took a long time and went accompanied with ups and downs. The reader facing difficulties in grasping the concepts of probability theory might find comfort in the idea that even the genius Gottfried von Leibniz (1646–1716), the inventor of differential and integral calculus along with Newton, had difficulties in calculating the probability of throwing 11 with one throw of two dice. Probability theory is a difficult subject to get a good grasp of, especially in a formal framework. The computer offers excellent possibilities for acquiring a better understanding of the basic ideas of probability theory by means of simulation. With computer simulation, a concrete probability situation can be imitated on the computer. The simulated results can then be shown graphically on the screen. The graphic clarity offered by such a computer simulation makes it an especially suitable means to acquiring a better feel for probability. Not only a didactic aid, computer simulation is also a practical tool for tackling probability problems that are too complicated for scientific solution. Even for experienced probabilists, it is often difficult to say whether a probability problem is too hard to solve analytically. However, computer simulation always works when you want to get a numerical answer to such a problem. To illustrate this, consider the problem of finding the probability that any two adjacent letters are different when the eleven letters of the word Mississippi are put in random order. Seemingly, a simple probability problem. However, it turns out that the analytical solution of this problem is very difficult to obtain, whereas an accurate estimate of the answer is easily obtained by simulation, see Section 2.9.4.

An outline

Part One of the book comprises Chapters 1–6. These chapters introduce the reader to the basic concepts of probability theory by using motivating examples to illustrate the concepts. A "feel for probabilities" is first developed through examples that endeavor to bring out the essence of probability in a compelling way. Simulation is a perfect aid in this undertaking of providing insight into the hows and whys of probability theory. We will use computer simulation, when needed, to illustrate subtle issues. The two pillars of probability theory, namely, the *law of large numbers* and the *central limit theorem* receive in-depth treatment. The nature of these two laws is best illustrated through the coin-toss experiment. The law of large numbers says that the percentage of tosses to come out heads will be as close to 0.5 as you can imagine provided that the coin is tossed often enough. How often the coin must be tossed in order to reach a prespecified precision for the percentage can be identified with the central limit theorem.

In Chapter 1, readers first encounter a series of intriguing problems to test their feel for probabilities. These problems will all be solved in the ensuing chapters. In Chapter 2, the law of large numbers provides the central theme. This law makes a connection between the probability of an event in an experiment and the relative frequency with which this event will occur when the experiment is repeated a very large number of times. Formulated by the aforementioned Jakob Bernoulli, the law of large numbers forms the theoretical foundation under the experimental determination of probability by means of computer simulation. The law of large numbers is clearly illuminated by the repeated coin-toss experiment, which is discussed in detail in Chapter 2. Astonishing results hold true in this simple experiment, and these results blow holes in many a mythical assumption, such as the "hot hand" in basketball. One remarkable application of the law of large numbers can be seen in the Kelly formula, a betting formula that can provide insight for the making of horse racing and investment decisions alike. The basic principles of computer simulation will also be discussed in Chapter 2, with emphasis on the subject of how random numbers can be generated on the computer.

In Chapter 3, we will tackle a number of realistic probability problems. Each problem will undergo two treatments, the first one being based on computer simulation and the second bearing the marks of a theoretical approach. Lotteries and casino games are sources of inspiration for some of the problems in Chapter 3. Also, the theory of probability is used to put coincidences in a broader context. Nearly all coincidences can be explained using probabilistic reasoning.

The binomial distribution, the Poisson distribution, and the hypergeometric distribution are the subjects of Chapter 4. We will discuss which of these important probability distributions applies to which probability situations, and we will take a look into the practical importance of the distributions. Once again, we look to the lotteries to provide us with instructional and entertaining examples. We will see, in particular, how important the sometimes underestimated Poisson distribution, named after the French mathematician Siméon-Denis Poisson (1781–1840), really is.

In Chapter 5, two more fundamental principles of probability theory and statistics will be introduced: the central limit theorem and the normal distribution with its bell-shaped probability curve. The central limit theorem is by far the most important product of probability theory. The names of the mathematicians Abraham de Moivre and Pierre Simon Laplace are inseparably linked to this theorem and to the normal distribution. De Moivre discovered the normal distribution around 1730.† An explanation of the frequent occurrence of this distribution is provided by the central limit theorem. This theorem states that data influenced by many small and unrelated random effects are approximately normally distributed. It has been empirically observed that various natural phenomena, such as the heights of individuals, intelligence scores, the luminosity of stars, and daily returns of the S&P, follow approximately a normal distribution. The normal curve is also indispensable in quantum theory in physics. It describes the statistical behavior of huge numbers of atoms or electrons. A great many statistical methods are based on the central limit theorem. For one thing, this theorem makes it possible for us to evaluate how (im)probable certain deviations from the expected value are. For example, is the claim that heads came up 5,250 times in 10,000 tosses of a fair coin credible? What are the margins of errors in the predictions of election polls? The standard deviation concept plays a key roll in the answering of these questions. We devote considerable attention to this fundamental concept, particularly in the context of investment issues. At the same time, we also demonstrate in Chapter 5, with the help of the central limit theorem, how confidence intervals for the outcomes

† The French-born Abraham de Moivre (1667–1754) lived most of his life in England. The protestant de Moivre left France in 1688 to escape religious persecution. He was a good friend of Isaac Newton and supported himself by calculating odds for gamblers and insurers and by giving private lessons to students.

of simulation studies can be constructed. The standard deviation concept also comes into play here. The central limit theorem will also be used to link the random walk model with the Brownian motion model. These models, which are used to describe the behavior of a randomly moving object, are among the most useful probability models in science. Applications in finance will be discussed, including the Black–Scholes formula for the pricing of options. In Chapter 5 we also have a first acquaintance with Bayesian statistics. The English clergyman Thomas Bayes (1702–1761) laid the foundation for this branch of statistics. The Bayesian approach is historically the original approach to statistics, predating what is nowadays called classical statistics by a century. Astronomers have contributed much to the Bayesian approach. In Bayesian inference one typically deals with nonrepeatable chance experiments. Astronomers cannot do experiments on the universe and thus have to make probabilistic inferences from evidence left behind. This is very much the same situation as in forensic science, in which Bayesian inference plays a very important role as well. The Bayesian approach is increasingly used in modern science. In the second part of the book we delve more deeply into this approach.

The probability tree concept is discussed in Chapter 6. For situations where the possibility of an uncertain outcome exists in successive phases, a probability tree can be made to systematically show what all of the possible paths are. Various applications of the probability tree concept will be considered, including the famous Monty Hall dilemma and the test paradox.

Part Two of the book is designed for an introductory probability course and consists of Chapters 7–16. These chapters can be studied independently of Part One. Chapter 7 states the axioms of probability and derives several basic rules from the axioms. These rules include the inclusion-exclusion rule and are illustrated with many examples. Chapter 8 delves into the concept of conditional probability which lies at the heart of probability theory. The law of conditional probability and Bayes' rule for revising conditional probabilities in light of new information are discussed in depth. Chapter 9 addresses discrete random variables and introduces the concepts of expected value and variance of a random variable. Rules for the expected value and variance of a sum of random variables are treated, including the square-root rule for the standard deviation of the sum of independent random variables. Also, Chapter 9 introduces the most important discrete probability distributions such as the binomial, the Poisson and the hypergeometric distributions. Chapter 10 addresses continuous random variables and explains the concept of probability density, always a difficult concept for the beginner to absorb. Also, this chapter provides insight into the most important probability densities. In particular, the normal density and its role in the central limit theorem are discussed. Whereas Chapter 10 deals with the probability distribution of a single random variable, Chapter 11 treats joint probability distributions for two or more dependent random variables and introduces the concepts of covariance and correlation. The multivariate normal distribution is the most important joint probability distribution and is the subject of Chapter 12. The multidimensional central limit theorem is treated together with several applications, including the chi-square test. Chapter 13 defines conditional probability densities and states the law of conditional expectation. Practical applications of the theory are given. Also, Bayesian inference for continuous models is discussed. The Bayesian approach is inextricably bound up with conditional probabilities. Chapter 14 deals with the method of moment-generating functions. This useful method enables us to analyze many applied probability problems. Also, the method is used to provide proofs for the strong law of large numbers and the central limit theorem. In the final Chapters 15 and 16 we introduce a random process, known as a Markov chain, which can be used to model many-real world systems that evolve dynamically in time in a random environment. Chapter 15 deals with discrete-time Markov chains in which state changes can only occur at fixed times. The basic theory of Markov chains is presented. The theory is illustrated with many examples, including the application of absorbing Markov chains to the computation of success run probabilities in a sequence of independent trials. Also, the powerful method of Markov chain Monte Carlo simulation is treated in Chapter 15. Chapter 16 deals with continuous-time Markov chains in which the times between state changes are continuously distributed. This versatile probability model has numerous applications to queueing systems and several of these applications are discussed. Each of the chapters is accompanied by carefully designed homework problems. The text includes several hundred homework problems. The answers to the odd-numbered problems are given at the end of the book.

Cambridge University Press
978-1-107-65856-1 - Understanding Probability: Third Edition
Henk Tijms
Excerpt
More information

PART ONE

Probability in action

Cambridge University Press
978-1-107-65856-1 - Understanding Probability: Third Edition
Henk Tijms
Excerpt
More information