God . . . has afforded us only the twilight of probability; suitable, I presume, to that state of mediocrity and probationership he has been pleased to place us in here. . .

John Locke (1690)

1

Classical probabilities, 1660–1840

1.1 INTRODUCTION

In July of 1654 Blaise Pascal wrote to Pierre Fermat about a gambling problem which came to be known as the Problem of Points: Two players are interrupted in the midst of a game of chance, with the score uneven at that point. How should the stake be divided? The ensuing correspondence between the two French mathematicians counts as the founding document in mathematical probability, even though it was not the first attempt to treat games of chance mathematically (Pascal, [1654] 1970, vol. 1, pp. 33–7; Cardano, [comp. *c.* 1525] 1966). Some years later, Pascal included among his *Pensées* an imaginary wager designed to convert sporting libertines: no matter how small we make the odds of God's existence, the pay-off is infinite; infinite bliss for the saved and infinite misery for the damned. Under such conditions, Pascal argued that rational self-interest dictates that we sacrifice our certain but merely finite worldly pleasures to the uncertain but infinite prospect of salvation (Pascal, [1669] 1962, pp. 187–90).

These two famous Pascal manuscripts, the one mathematical and the other philosophical, reveal the double root of the mathematical theory of probability. It emerged at the crux of two important intellectual movements of the seventeenth century: a new pragmatic rationality that abandoned traditional ideals of certainty; and a sustained and remarkably fruitful attempt to apply mathematics to new domains of experience. Neither would have been sufficient without the other. Philosophical notions about what happens only most of the time, and about the varying degrees of certainty connected with this unreliable experience date from antiquity, as do games of chance. But before *circa* 1650, no one attempted to quantify any of these senses of probability. Nor would the spirit of mathematical enterprise have alone sufficed, for quantification requires a subject matter, an interpretation to flesh out the mathematical formalism. This was particularly true for the calculus of probabilities, which until this century had no mathematical existence independent of its applications.

1.2 THE BEGINNINGS

The prehistory of mathematical probability has attracted considerable scholarly attention, perhaps because it seems so long overdue. Chance is our constant companion, and the mathematics of the earliest formulations of probability theory was elementary. Suggestive fragments of probabilistic thinking do turn up almost everywhere in the classical and medieval learned corpus: Around 85 B.C., Cicero connected that which usually happens with what is ordinarily believed in his rhetorical writings and called both *probabile* (Cicero, 1960, pp. 85–90). In a tenth-century manuscript, a monk enumerated all 36 possibilities for the toss of two dice (Kendall, 1956), and Talmudists reasoned probabilistically about inheritances and paternity (Rabinovitch, 1973). Yet none of these flowered into a mathematics of probability.

Several plausible hypotheses about why mathematical probability came about when it did also dissolve upon inspection. Maritime insurance expanded rapidly in Italy and the Low Countries during the commercial boom of the fifteenth and sixteenth centuries, but insurers did not collect statistics on shipwrecks, much less develop a mathematical basis for pricing premiums. It was the mathematicians who later – much later – influenced the insurers, not vice versa (Maistrov, [1964] 1974; Daston, 1987). Nor did any new recognition of chance inspire the mathematicians; on the contrary, the early probabilists from Pascal through Pierre Simon Laplace were determinists of the strictest persuasion (Kendall, 1956; Hacking, 1975). One might speculate that the mathematics of combinatorics was a precondition for the earliest versions of probability theory, but the two subjects appear to have developed in tandem, with probability theory often stimulating work in combinatorics rather than the reverse (Todhunter, 1865). The Renaissance doctrine of signatures linked the evidence of things with that of words in a way that parallels the objective and subjective senses of mathematical probabilities, but Cicero and the medieval rhetorical tradition that followed him had done so long before

(Hacking, 1975; Garber and Zabell, 1979). Similarly, the passion for gambling was hardly an invention of the seventeenth century, and so could not have been the catalyst that transformed qualitative probabilities into quantitative ones. It is, in short, easier to say where mathematical probability did *not* come from.

The very earliest writings on mathematical probability do supply some clues, however. If we return to the two Pascal musings, we discover that although they are recognizably part of what came to be called the calculus of probabilities, they are not cast in terms of probabilities. The fundamental concept was instead expectation, later defined as the product of the probability of an event e and its outcome value V:

$$P(e)V = E$$

So, for example, the expectation of someone holding one out of a thousand tickets for a fair lottery with a prize of $10,000 would be $10. As the definition implies, we now derive expectation from the probability, but for the early probabilists expectation was the prior and irreducible notion.

Expectation in turn was understood in terms of a fair exchange or contract. Pascal described his solution to the Problem of Points as rendering to each player what "in justice" belonged to him. In the first published treatise on mathematical probability, *De ratiociniis in ludo aleae* (1657), the Dutch mathematician and physicist Christiaan Huygens made expectation his departure point and defined it in terms of equity: equal expectations obtained in a fair game; that is, one that "worked to no one's disadvantage" (Huygens, [1657] 1920, p. 60). Since later probabilists would *define* a fair game as one in which the players possessed equal expectations, this definition of equal expectations in terms of a fair game strikes the modern reader as circular. But for the first generation of probabilists, notions of equity were intuitively clear enough to serve as the stuff of definitions and postulates.

These intuitions drew upon a category of legal agreement that had become increasingly important in sixteenth- and seventeenth-century commercial law, the aleatory contract. Jurists defined such agreements as the exchange of a present and certain value for a future, uncertain one – staking a gamble, purchasing an annuity, taking out an insurance policy, bidding on next year's wheat crop, or buying the next cast of a fisherman's net. Pascal's wager hinged upon a similar trade of the certain enjoyment of present vices for the uncertain joy of salvation. Aleatory contracts acquired prominence and a certain notoriety as the preferred way of exonerating merchants who made loans with interest from charges of usury

(Coumet, 1970). The element of risk, argued the canon lawyers, was the moral equivalent of labor, and therefore earned the merchant his interest as honestly as the sweat of his brow would have. Thus Jesuits successfully petitioned the Sacred Congregation for Propaganda in 1645 for a special dispensation for their Chinese converts, who were charging 30% interest on loans, on the condition "there is considered the equality and probability of the danger, and provided that there is kept a proportion between danger and what is received" (Noonan, 1957, p. 289).

It was this "proportion between danger and what is received," the element of equity fundamental to all contracts, that the mathematicians attempted to quantify in almost all of the early applications of probability theory. Both the problems they addressed – gambling stakes, annuity prices, future inheritances – and the terms in which they did so – using the concept of equal expectations – bear witness to the seminal influence of the law of aleatory contracts.

Pascal's wager is an example of how reasoning by expectations had become almost synonymous with a new brand of rationality by the mid-seventeenth century. His libertine interlocutor must be led back into the Christian fold by uncertain wagers rather than theological certainties. In the sixteenth century, Reformation controversies between Protestants and Catholics on the one hand, and the revival of the sceptical philosophy of Sextus Empiricus and his school on the other, combined to undermine the ideal of certain knowledge that had guided intellectual inquiry since Aristotle. In its place gradually emerged a more modest doctrine that accepted the inevitability of less than certain knowledge, but maintained nonetheless that it was still sufficient to guide the reasonable man in precept and in practice. Aristotle's dictum from the *Nicomachean Ethics* (1094b 24–25) was much quoted: "it is the mark of an educated man to look for precision in each class of things just so far as the nature of the subject admits: it is evidently equally foolish to accept probable reasoning from a mathematician and to demand from a rhetorician demonstrative proofs."

The ultimate result of the Reformation and Counter-Reformation clashes over the fundamental principles of faith and their justification, and of the radical scepticism of Michel de Montaigne and other sixteenth-century thinkers was vastly to erode the domain of the demonstrative proof and to expand that of probable reasoning. Their immediate impact was more devastating, challenging all claims to any kind of knowledge whatsoever. Religious apologists who sought to undercut the other side's claims to legitimacy on the basis of either (ambiguous) revelation

Cambridge University Press
978-0-521-39838-1 - The Empire of Chance: How Probability Changed Science and Everyday Life
Gerd Gigerenzer, Zeno Swijtink, Theodore Porter, Lorraine Daston, John Beatty, and
Lorenz Kruger
Excerpt
More information

or (dubious) authority soon discovered that their destructive arguments were a double-edged sword. The revived pyrrhonism of the "libertins érudits" denied the reliability of even sense impressions and mathematical demonstrations; Descartes' *Meditations* began with a sceptical reverie of this extreme variety. Thus all of the traditional sources of certainty, religious and philosophical, came simultaneously under attack. Confronted with a choice between fideist dogmatism on the one hand and the most corrosive scepticism on the other, an increasing number of seventeenth-century writers attempted to carve out an intermediate position that abandoned all hope of certainty except in mathematics and perhaps metaphysics, and yet still insisted that men could attain probable knowledge. Or rather, they insisted that probable knowledge was indeed knowledge (Popkin, 1964; Shapiro, 1983).

In order to make their case for the respectability of the merely probable, these "mitigated sceptics" turned from rarified philosophical discourse to the conduct of daily life. The new criterion for rational belief was no longer a watertight demonstration, but rather that degree of conviction sufficient to impel a prudent man of affairs to action. For reasonable men that conviction in turn rested upon a combined reckoning of hazard and prospect of gain, i.e. upon expectation. Pascal's wager is about neither the bare probability of God's existence, nor the infinite bliss or misery that awaits saint or sinner, respectively. Rather, it is about the product of the two, significantly conceived in terms of a gamble, and the relationship between certain stake and uncertain pay-off, and thus a sterling example of the new rationality. Pascal's Port Royal colleagues Antoine Arnauld and Pierre Nicole made such mixed reasoning the *sine qua non* of rational judgment in their influential *Logique*, cautioning their readers that it is not enough to consider how good or bad an outcome is in itself, but also the likelihood that it will come to pass (Arnauld and Nicole, [1662] 1965, pp. 352–3). English and Dutch spokesmen for the new rationality of expectation preferred commercial analogies, but the idea was the same. John Wilkins, Anglican bishop and founding member of the Royal Society of London, argued in his *Of the Principles and Duties of Natural Religion* (1675) that just as merchants were willing to risk the perils of a long voyage in the name of profit, so "he that would act rationally, according to such Rules and Principles as all mankind do observe in the government of their Actions, must be persuaded to do the like" in matters of science and religion (Wilkins, [1675] 1699, p. 16). The emphasis upon action as the basis of belief, rather than the reverse, was key to the defense against scepticism, for as these writers were wont acidly to observe, even the most confirmed sceptic took his meals just as if the external world existed.

Expectation was thus central to the new rationality or "reasonableness," as it was sometimes called. The mitigated sceptics were less interested in equity than in rational belief, but they drew heavily upon the doctrine of aleatory contracts for examples to show that it was accepted practice and therefore reasonable to exchange a present, certain good – be it money to invest, a long-accepted scientific theory, or the indulgence of our lusts and passions – for a future, uncertain one – more money, a better theory, salvation. Mathematicians seeking to quantify the legal sense of expectations inevitably became involved in quantifying the new rationality as well. So began an alliance between mathematical probability theory and standards of rationality that stamped the classical interpretation as a "reasonable calculus"; as a mathematical codification of the intuitive principles underlying the belief and practice of reasonable men. The identification of classical probability theory with reasonableness was so strong that when the results of the one clashed with the other, it was the mathematicians who anxiously amended definitions and postulates to restore harmony, as we shall see below.

1.3 THE CLASSICAL INTERPRETATION

Thus the calculus of chance was in the first instance a calculus of expectations, and thereby an attempt to quantify the new, more modest doctrine of rationality that surfaces almost everywhere in seventeenth-century learned discourse. The first published works on the subject, from Huygens' little treatise of 1657 to Jakob Bernoulli's definitive *Ars conjectandi* of 1713, covered a range of topics that cohere only against this background. Aleatory contracts like gambling (Huygens, Pierre de Montmort, Jakob Bernoulli) and annuities (Johann De Witt, Halley, Nicholas Bernoulli), and later evidentiary problems like the evaluation of historical or courtroom testimony (John Craig, George Hooper, Nicholas and Jakob Bernoulli) constituted the domain of applications for the new theory. By the end of this period, probability had emerged as a distinct and primitive concept, although most of the applications continued to revolve around questions of expectation for some time thereafter.

Just what these probabilities measured was ambiguous from the outset, and remains a matter of controversy to this day. Originally the word "probability" had meant an opinion warranted by authority (Byrne, 1968); hence the Jesuit doctrine of probabilism, which casuists wielded to absolve almost every transgression on the grounds that one theologian or another had taken a mild view of the matter (Demain, 1935). However, the mitigated scepticism of the early seventeenth century modified even this qualitative sense of probability. The proponents of reasonableness spoke not of certainty but of certainties, ranging from the highest grade of "mathematical" certainty attained by demonstration, through the "physical" certainty of sensory evidence, down to the "moral" certainty based on testimony and conjecture. The precise descriptions of these levels varied slightly from author to author, but the notion of such an ordered scale, and the emphasis that most things admit only of moral certainty, remained a staple of the literature from Hugo Grotius' *De veritate religionis christianae* (1624) to John Locke's *Essay Concerning Human Understanding* (1690) and thereafter. When Bishop Joseph Butler claimed in 1736 that "probabilities are the very guide of life," he was by then repeating a cliché (Butler, 1736, p. iii).

In the context of these discussions, the very meaning of the word "probability" changed from its medieval sense of any opinion warranted by authority to a degree of assent proportioned to the evidence at hand, both of things and of testimony (Locke, [1690] 1959, IV. xv–xvi). These probabilities were qualitatively conceived, and owed much to the language and practice of legal evidence, as the numerous courtroom examples and analogies make clear (Daston, 1988, chapter 2). However, mathematicians like Gottfried Wilhelm Leibniz and Jakob Bernoulli seized upon the new "analysis of hazards" as a means of quantifying these degrees of certainty, and in so doing, converting the three ordered points into a full continuum, ranging from total disbelief or doubt to greatest certainty (J. Bernoulli, 1713, IV.i). Indeed, Leibniz described the fledgling calculus of probabilities as a mathematical translation of the legal reasoning that carefully proportioned degrees of assurance on the part of the judge to the kinds of evidence submitted (Leibniz, [comp. *c.* 1705] 1962, pp. 460–5). The fact that these legal probabilities were sometimes expressed in terms of fractions to create a kind of "arithmetic of proof" (for example, the testimony of a relative of the accused might count only ⅓ as much as that of an unimpeachable witness) may have made them seem mathematically tractable.

The mathematicians who set about trying to measure these probabilities in some non-arbitrary fashion came up with at least three methods: equal possibilities based on physical symmetry; observed frequencies of events; and degrees of subjective certainty or belief. (Other seventeenth-century meanings of "probability," such as the appearance of truth or the strength of analogy, were not successfully quantified.) The first was well suited to gambling devices like coins or dice but little else; the second depended on the collection of statistics and assumptions of long-term stability; and the third echoed the legal practice of proportioning degrees of certainty to evidence.

The various senses emerged from different contexts, and suggested different applications for the mathematical theory. Sets of equiprobable outcomes based on physical symmetry derived from gambling and were applied to gambling – very few other situations satisfy these conditions in an obvious way. Statistical frequencies originally came from mortality and natality data gathered by parishes and cities from the sixteenth century onwards. In 1662 the English tradesman John Graunt used the London bills of mortality to approximate a mortality table by assuming that roughly the same fraction of the population died each decade after the age of six (Graunt, [1662] 1975, pp. 29–30). (Since the bills of mortality registered only cause, not age at death, Graunt's table was based on informed guesswork about what diseases killed whom at what age, and the faith that mortality was regular.) Eighteenth-century authors collected more detailed demographic data and enlisted probability theory in order to compute the price of annuities, and later life insurance, and to argue for divine providence in human affairs.The epistemic sense of belief proportioned to evidence arose from legal theories about just how much and what kind of evidence was required to produce what degree of conviction in the mind of the judge, and inspired applications to the probabilities of testimony, both courtroom and historical, and of judgment.

Latter-day probabilists view these three answers to the question, "What do probabilities measure?" as quite distinct, and much ink has been spilt arguing their relative merits and compatibility (Nagel, 1955). In particular, a bold line is now drawn between the first two "objective" meanings of probability, which correspond to states of the world, and the third "subjective" sense, which corresponds to states of mind. Yet classical probabilists used "probability" to mean all three senses, shifting from one to another with an insouciance that bewilders their more nice-minded successors.

Why were classical probabilists able to conflate these different notions of probability so easily, and often very fruitfully? In part, because the objective and subjective senses were not then separated by the chasm that yawns between them in current philosophy. Legal theorists of the sixteenth and seventeenth centuries found it plausible to assume that conviction formed in the mind of the judge in proportion to the weight of the evidence presented, and Locke repeated the assumption in a more general context, invoking the qualitative probabilities of evidence: the rational mind assents to a claim "proportionably to the preponderancy of the greater grounds of probability on one side or the other" (Locke, [1690] 1959, vol. II, p. 366). At least two further elements were required to connect the objective and subjective senses of qualitative probabilities. First, precept had to be guaranteed in practice. It was not enough that the mind *should* apportion assent in strict relation to the evidence; it had to be shown that it actually did so. Second, the evidence had to be quantified.

The empiricist philosophy-*cum*-psychology of the late seventeenth and eighteenth centuries satisfied both desiderata. John Locke, David Hartley, and David Hume created and refined a theory of the association of ideas that made the mind a kind of counting machine that automatically tallied frequencies of past events and scaled degrees of belief in their recurrence accordingly. Hartley went so far as to provide a physiological mechanism for this mental record-keeping: each repeated sensation set up a cerebral vibration that etched an ever deeper groove in the brain, corresponding to an ever stronger belief that things would be as they had been. Hume notoriously rejected the rationality of such inferences to the future based on past experience, *pace* Locke and Hartley, but he retained the psychology that made them inevitable. Images of past experiences conjoin to heighten the vivacity of a mental impression, each repetition being "as a new stroke of the pencil, which bestows an additional vivacity on the colours." (Hume, [1739] 1975, p. 135). Since the mind irresistibly conferred belief in proportion to the vivacity of an idea, the more frequent the conjunction of events in past experience, the firmer the conviction that they would occur again. Locke and Hartley contended that this matching of belief to frequencies was rational (Hartley appealing explicitly to the calculus of probabilities; Locke, [1690] 1959, IV.xv; Hartley, 1749, vol. I, pp. 336–9). Hume replied that it was merely habitual, although his "Essay on Miracles" elevated belief based on unexceptioned past experience to at least a kind of reasonableness (Hume, [1758] 1955, chapter 10). All however concurred that the normal mind, when uncorrupted by upbringing or prejudice, irresistibly linked the subjective probabilities of belief with the objective probabilities of frequencies.

They also showed an increasing tendency to reduce all forms of evidence whatsoever to frequencies, in contrast to the legal doctrines that had originally been the prototype of degrees of belief proportioned to evidence. For the judge, the probative weight of eye-witness testimony that the accused had been seen fleeing the scene of the murder with unsheathed bloody sword derived from the quality of the evidence, not its quantity. It mattered not how many times in the past similar evidence had led to successful convictions. Locke remained very close to this legal tradition in his discussion of the kinds of evidence that create probabilities: number of witnesses, their skill and integrity, contradictory testimony, internal consistency, etc. He told the cautionary tale of the King of Siam, who dismissed the Dutch ambassador as a liar because his tales of ice-skating on frozen canals ran counter to the accumulated experience of generations of Siamese that water was always fluid. The King erred in trusting the mere quantity of his experience, without evaluating its breadth and variety. Yet Locke also made a place for "the frequency and constancy of experience" and for the number, as well as the credibility of testimonies (Locke, [1690] 1959, IV.xv). Later philosophical writings on probabilities narrowed the sense of evidence to the countable still further. Hume represents the endpoint of this evolution, in which evidence has become the sum of repeated, identical events. According to Hume, the mind not only counted; it was exquisitely sensitive to small differences in the totals: "When the chances or experiments on one side amount to ten thousand, and on the other to ten thousand and one, the judgment gives the preference to the latter, upon account of that superiority" (Hume, [1739] 1975, p. 141).

The guarantee that subjective belief was willy-nilly proportioned to objective frequencies and also, according to some authors, to physical symmetries allowed classical probabilists to slide from one sense of probability to another with little or no explicit justification. Only when associationist psychology shifted its emphasis to the illusion and distortions that prejudice and passion introduced into this mental reckoning of probabilities did the gap between subjective and objective probabilities become clear enough to demand a choice between the two. It was not so much the development and triumph of a thoroughgoing frequentist version of probability theory that marked the end of the classical interpretation, as the realization that a choice must be made between (at least) two distinct senses of probability. The range of problems to which the classical probabilists applied their theory shows that their understanding of probability embraced objective as well as subjective elements: statistical actuarial probabilities happily co-existed with epistemic probabilities of testimony in the work of Jakob Bernoulli or Laplace.