



[A] rewarding read. . . . Mr. Devlin does a remarkable job of showing just how much derived from the history-changing Pascal-Fermat correspondence."

—*Washington Times*

THE UNFINISHED GAME

PASCAL, FERMAT, and the
SEVENTEENTH-CENTURY
LETTER that MADE the
WORLD MODERN

A Tale of How Mathematics
Is Really Done



KEITH DEVLIN

Author of *The Math Gene* and
The Millennium Problems



The Unfinished Game

*Pascal, Fermat, and the
Seventeenth-Century Letter that
Made the World Modern*

KEITH DEVLIN



A Member of the Perseus Books Group
New York

Table of Contents

Title Page

NOTE TO THE READER

PREFACE

CHAPTER 1 - Monday, August 24, 1654

RISK

A LONG STRUGGLE

THE PROBLEM OF THE POINTS

CHAPTER 2 - A Problem Worthy of Great Minds

PROBABILITY

STRUGGLING TOWARD THE SOLUTION

PASCAL AND FERMAT

THE STRATEGY

CHAPTER 3 - On the Shoulders of a Giant

NUMBER WAS THE KEY

THE REMARKABLE MAN FROM MILAN

CHAPTER 4 - A Man of Slight Build

AN ABUNDANCE OF TALENTS

FIRST LETTERS

PASCAL'S UNIVERSAL METHOD

CHAPTER 5 - The Great Amateur

THE LAWYER IN TOULOUSE

CHAPTER 6 - Terrible Confusions

FERMAT STUMBLES

CHAPTER 7 - Out of the Gaming Room

ONLY THE GODS KNOW WHAT TOMORROW WILL BRING

THE MAN WHO COUNTED DEAD PEOPLE

A DUTCH TREAT WITH GREAT EXPECTATIONS

CHAPTER 8 - Into the Everyday World

THE AMAZING BERNOULLIS

THE ART OF CONJECTURE

PROBABILITY GOES TO COURT

A NEW KIND OF PROBABILITY

CHAPTER 9 - The Chance of Your Life

FEAR OF FLYING

RINGING IN THE BELL CURVE

THE PRICE ON YOUR HEAD

THE REVEREND BAYES

BAYES' FORMULA

CHAPTER 10 - The Measure of Our Ignorance

IF ONLY

MEASURING OUR IGNORANCE

PROBABILITY GOES BACK TO COURT

WHAT EXACTLY IS PROBABILITY?

THE FORESIGHT SAGA

THE KEY LETTER FROM PASCAL TO FERMAT

INDEX

Basic Ideas

Copyright Page

CHAPTER 4

A Man of Slight Build

... I replied to him [M. de Roberval] that I did not found my reasoning so much on this method of combinations, which in truth is not in place on this occasion, as on my universal method from which nothing escapes and which carries its proof with itself. This finds precisely the same division as does the method of combinations. Furthermore, I showed him the truth of the divisions between two players by combinations in this way. Is it not true that if two gamblers finding according to the conditions of the hypothesis that one lacks two points and the other three, mutually agree that they shall play four complete plays, that is to say, that they shall throw four two-faced dice all at once,—is it not true, I say, that if they are prevented from playing the four throws, the division should be as we have said according to the combinations favorable to each? He agreed with this and this is indeed proved. But he denied that the same thing follows when they are not obliged to play the four throws.

Advances in science are often delayed because the common wisdom holds that something is impossible. From the time of the ancient Greeks, it was believed that the future was in the hands of the gods—a matter of pure fate. Quantifying the way a pair of dice may fall, as Cardano had done, was one thing; predicting what future throws of the dice might bring (as in the unfinished game) was quite another.

As recently as 1756, over a hundred years after Pascal and Fermat had their famous exchange, the great French mathematician Abraham de Moivre (whom we shall meet later), the man who discovered the normal distribution that forms the basis of contemporary predictive statistics, firmly believed the future was determined by God. He wrote (in his book *Doctrine of Chances*):

Again, as it is thus demonstrable that there are, in the constitution of things, certain Laws according to which Events happen, it is no less evident from Observation, that those Laws serve to wise, useful and beneficent purposes to preserve the steadfast Order of the Universe, to propagate the several Species of Beings, and furnish to the sentient Kind such degrees of happiness as are suited to their State. But such Laws, as well as the original Design and Purpose of their Establishment, must all be from *without*; the *Inertia* of matter, and the nature of all created Beings, rendering it impossible that any thing should modify its own essence, or give to itself, or to anything else, an original determination or propensity. And hence, if we blind not ourselves with metaphysical dust, we shall be led, by a short and obvious way, to the acknowledgement of the great MAKER and GOVERNOR of all; *Himself all-wise, all-powerful and good.*¹⁵

Today, we have no difficulty in accepting that an event of pure chance may have sufficient “predictability” that we may calculate precisely the likelihoods pertaining to its outcome without needing to assume divine determination—the result really is a matter of pure chance. Indeed, today’s religious person who sees the hand of God in all things simply accepts that the laws of probability are themselves a

manifestation of God's ways. In short, we have taken the notions of undetermined random events and mathematically computed probabilities and folded them into our worldview. Quantum theory starts from the assumption that this is the nature of the very fabric of our universe.

But in the early seventeenth century, things looked very different. Pascal and Fermat not only had to figure out how to perform the calculation that resolves the problem of the unfinished game, but also had to do so within a worldview that considered what they were doing impossible. Prior to 1654, tomorrow was viewed as a matter of Fate, something over which a person had no control. In solving the problem, they were instrumental in changing that view, though as the passage from de Moivre indicates, it took a long time for the true import of their work to break through.

With the solution to the problem of the points and the subsequent acceptance of probability theory, mortals did have (a measure of) control over the future. An individual might not know with certainty what would happen, but by calculating the probabilities of various likely events, one might choose one's actions so as to minimize dangers and maximize preferred outcomes, or one might purchase insurance to facilitate recovery if things went badly. In short, Pascal and Fermat gave us the ability to manage risk.

AN ABUNDANCE OF TALENTS

Unlike Cardano, Pascal did not write an autobiography, and accordingly, we must rely on the scholarship of others to discover what kind of a man he was. In *Blaise Pascal: Mathematician, Physicist and Thinker About God*, biographer Donald Adamson writes that his subject was

a man of slight build with a loud voice and somewhat overbearing manner . . . he lived most of his adult life in great pain. He had always been in delicate health, suffering even in his youth from migraine . . . [and was] precocious, stubbornly persevering, a perfectionist, pugnacious to the point of bullying ruthlessness yet seeking to be meek and humble.¹⁶

Historian René Taton, in his biography of Pascal in the *Dictionary of Scientific Biography*, provides this assessment of Pascal's life:

At once a physicist, a mathematician, an eloquent publicist in the Provinciales . . . Pascal was embarrassed by the very abundance of his talents. It has been suggested that it was his too concrete turn of mind that prevented his discovering the infinitesimal calculus, and in some of the Provinciales the mysterious relations of human beings with God are treated as if they were a geometrical problem. But these considerations are far outweighed by the profit that he drew from the multiplicity of his gifts, [and] his religious writings are rigorous because of his scientific training.¹⁷

In December 1639, six years before Pascal had the conversation with the Chevalier de Méré that led to the famous correspondence with Fermat, the Pascal family left Paris to live in Rouen. Blaise's father, Étienne, had been appointed a tax collector for Upper Normandy. In February the following year, the young Blaise published his first genuine mathematical paper, *Essay on Conic Sections*. (The earlier "paper" he had presented to Mersenne's Academy in Paris as a teenager was a single-page affair—little

more than a student exercise, though Blaise's young age at the time made it a portent of greater things to come.)

It was in Rouen, from 1642 to 1645, that he worked on the Pascaline, his mechanical calculator. The only previous attempt to construct such a device was in 1624 by a man called Schickard, so Pascal's machine was the second mechanical calculator ever built.

The year after Blaise finished work on his device, his father slipped on the ice and broke his hip. Two young men were hired to care for Étienne. The caretakers were brothers in a nearby religious community of Jansenists. Jansenism was a proselytizing Catholic sect that preached an extreme form of asceticism, sacrifice, and strict adherence to the Scriptures as the only path to salvation. Though the two brothers seemed to have little effect on Étienne Pascal, they were more successful with the young and impressionable son, who adopted their faith. From that point forward, there was to be no pleasurable social life for young Blaise. Intellectual pursuits such as mathematics or science were also distractions to be avoided.

For a short while, Pascal renounced mathematics as he fought to save his soul from eternal damnation. The terror of such a fate eventually became too great for the young man to bear, however, and he became ill, suffering severe headaches and a partial paralysis. His doctor advised him that for the sake of his health, he should abandon the Jansenist ways and lead a life more normal for a young man. Although he would remain strongly religious for the remainder of his all-too-short life, Blaise resumed normal activities. Indeed, he did so with vigor, adding regular visits to the gaming rooms to his earlier academic pursuits. It was at the gambling table that Pascal met the Chevalier de Méré, a keen gambler with sufficient mathematical ability to figure out for himself some of the more favorable odds.

Pascal's first focus, when he resumed his researches, was in physics. He carried out a series of experiments in which he observed that the pressure of the atmosphere decreases with height and so deduced that, at some height, it must thin out to a vacuum. In September 1647, he informed Descartes of his finding when the great philosopher came on a visit. Descartes refused to accept Pascal's reasoning, and the pair argued about it for two days. Reporting on the dispute afterward in a letter to the physicist Christiaan Huygens, Descartes noted rather haughtily that Pascal "has too much vacuum in his head." In October 1647, Pascal published his findings under the title *New Experiments Concerning Vacuums*, which led to disputes with a number of scientists who, like Descartes, did not believe in a vacuum.

Étienne died in September 1651. The event prompted Blaise to write to one of his sisters describing his Christian views on death and how they applied to their deceased father. The ideas he expressed in that letter were to form the basis for *Pensées*, Pascal's most famous work of philosophy, a collection of personal thoughts on human suffering and faith in God, which he began in late 1656 and continued to work on during 1657 and 1658. *Pensées* was finally published in 1670, eight years after Pascal's death. While it is a philosophical text, *Pensées* is not totally devoid of mathematical content. Among its essays is one that has come to be known as *Pascal's Wager*, an argument that belief in God is rational. Stripped of its numerical calculations, Pascal's case is essentially that if God does not exist, one will lose little by believing in the supreme being, while if God does exist, one will lose everything by not believing. We'll look at this argument in its full mathematical glory in Chapter 7.

In 1653, Pascal wrote *Treatise on the Equilibrium of Liquids*, in which he explains what is nowadays known as Pascal's law of pressure. At the same time, he resumed his earlier investigations of conic sections and proved some important theorems in projective geometry. In particular, he completed a work titled *The Generation of Conic Sections*, most of which he had written in 1648. This was meant to be the first part of a comprehensive treatise on conics, but he never got back to it.

His *Traité du triangle arithmétique* (*Treatise on the Arithmetical Triangle*), completed and printed in 1654 but not released until 1665, contained the numerical triangle that today bears his name. Although he was not the first to study this particular mathematical structure, his insights were far more illuminating than anything written beforehand. Pascal's results concerning binomial coefficients were to lead Isaac Newton to his discovery of a hugely important result known as the general binomial theorem (the binomial theorem for fractional and negative powers).

The full title of Pascal's work, rarely given in historical accounts, was *Traité du triangle arithmétique, avec quelques autres petits traités sur la même matière*. The additional clause translates as "with some other small treatises on the same topic," and the essay covers, among other things, Pascal's fairly comprehensive treatment of everything then known about probability theory (as we now call it). He sent a copy to Fermat, who referred to it in one of his letters to Pascal.

It was around the time he started his correspondence with Fermat about the problem of the points, in the summer of 1654, that Pascal first fell ill. Although Pascal did not know it at the time, he was suffering from the first stages of the stomach cancer that would eventually kill him. In one of his letters, written in July 1654, he told Fermat, "though I am still bedridden, I must tell you that yesterday evening I was given your letter." Before his correspondence with Fermat began, Pascal had been unsure if his own initial attempt at a solution was correct, and he first discussed the matter with his colleague Pierre de Carcavi, another member of Mersenne's Academy. Carcavi in turn suggested that Pascal contact Fermat, acknowledged to be greatest mathematician alive.

A few weeks after writing the August 24 letter that is the focus of our story, Pascal almost lost his life in an accident. The horses pulling his carriage bolted, and the carriage was left hanging over a bridge above the river Seine. Although he was rescued without injury, the experience seems to have affected him psychologically, and not long afterward, he underwent another religious experience. On November 23, 1654, he pledged his life to Christianity once again and, soon after, made the first of several visits to the Jansenist monastery Port-Royal des Champs, about twenty miles southwest of Paris. He began to publish anonymous works on religious topics, including eighteen *Provincial Letters* published during 1656 and early 1657. These were written in defense of his friend Antoine Arnauld, an opponent of the Jesuits and a defender of Jansenism, who was on trial before the faculty of theology in Paris for his controversial religious works.

As the pain from the malignancy in his stomach increased, Pascal lost interest in science. In 1658, he tried to do some mathematics, much of it in correspondence with others, but that was largely as a distraction from the pain and he soon gave up. He spent his last years giving to the poor and going from church to church in Paris, attending one religious service after another.

On August 19, 1662, at age thirty-nine, Blaise Pascal died in Paris, in intense pain after the cancer

reached his brain. He left behind a legacy of ideas made all the greater by his brief correspondence with Fermat in 1654.

FIRST LETTERS

The letter from Pascal to Fermat that began their collaboration has unfortunately been lost, and our knowledge of the two men's correspondence begins with Fermat's undated reply. After a brief discussion of the part of Pascal's letter that Fermat has found problematic—concerning the ever troublesome issue of how to handle games where one player wins in fewer than the maximum number of rounds—Fermat explains the difficulty he sees in Pascal's approach:

But you proposed in the last example in your letter (I quote your very terms) that if I undertake to find the six in eight throws and if I have thrown three times without getting it, and if my opponent proposes that I should not play the fourth time, and if he wishes me to be justly treated, it is proper that I have 125/1296 of the entire sum of our wagers.

This, however, is not true by my theory. For in this case, the three first throws having gained nothing for the player who holds the die, the total sum thus remaining at stake, he who holds the die and who agrees to not play his fourth throw should take 1/6 as his reward.

And if he has played four throws without finding the desired point and if they agree that he shall not play the fifth time, he will, nevertheless, have 1/6 of the total for his share. Since the whole sum stays in play it not only follows from the theory, but it is indeed common sense that each throw should be of equal value.

I urge you therefore [to write me] that I may know whether we agree in the theory, as I believe [we do], or whether we differ only in its application.

I am, most heartily, etc.,

Fermat

Pascal acknowledges the validity of Fermat's objection in his next letter, dated Wednesday, July 29, 1654, which begins:

1. Impatience has seized me as well as it has you, and although I am still abed, I cannot refrain from telling you that I received your letter in regard to the problem of the points yesterday evening from the hands of M. Carcavi, and that I admire it more than I can tell you. I do not have the leisure to write at length, but, in a word, you have found the two divisions of the points and of the dice with perfect justice. I am thoroughly satisfied as I can no longer doubt that I was wrong, seeing the admirable accord in which I find myself with you.

Pascal was a mathematician of formidable powers. But as their respective mathematical records make plain, Fermat was far superior, and anyone who reads the entire correspondence between the two will see readily that Fermat was the dominant collaborator.

PASCAL'S UNIVERSAL METHOD

Fermat's initial reply to Pascal and the famous August 24 letter from Pascal to Fermat make it clear that the two men had adopted different approaches to solving the problem of the points. While Fermat listed explicitly all the possible combinations of outcomes that could arise in the completion of the game, Pascal adopted a different strategy. Since both methods are correct, they will of course yield the same result, as Pascal made clear in his letter:

I did not found my reasoning so much on this method of combinations, which in truth is not in place on this occasion, as on my universal method from which nothing escapes and which carries its proof with itself. This finds precisely the same division as does the method of combinations.

The "universal method" that Pascal refers to is an instance of what is generally known as a *recursive method*. Recursion is the process whereby a sequence of numbers is generated by a rule that gives the new number in terms of the last number produced (or sometimes the last two numbers). The most famous example of a recursion is the rule for generating the Fibonacci numbers. As an exercise in his book *Liber abaci*, Leonardo of Pisa (Fibonacci) posed a problem about the growth of a rabbit colony. To solve the problem, you have to generate the sequence of numbers that starts with two 1's and grows according to the rule that the new number at each stage is the sum of the previous two numbers. When you follow this rule, you get the sequence

1, 1, 2, 3, 5, 8, 13, 21, 34, 55, 89, . . .

Pascal approaches the problem of the points by looking at the quantity $e(a, b)$ that gives the share of the stake that the first player should be given if player 1 requires a winning throws to win and player 2 requires b winning throws. Clearly, if the numbers a and b are equal, then $e(a, b) = 1/2$. The idea is to see how $e(a, b)$ changes when each player wins one more throw. This leads to an algebraic expression for $e(a, b)$ in terms of $e(a - 1, b)$ and $e(a, b - 1)$, and you can solve the problem of the points by using recursion to calculate $e(2, 3)$, the desired share in the particular game they considered.

The main drawback to Pascal's approach was that it requires some complicated algebra dependent upon the theory of combinations he worked out in connection with his famous triangle. Nevertheless, when carried out correctly (not an easy matter), it does lead to the correct result, and the same approach was subsequently used by Abraham de Moivre, Joseph Louis Lagrange, Pierre-Simon Laplace, and other mathematicians to find general methods for solving a variety of other problems.

A detailed description of Pascal's recursive solution to the problem of the points quickly becomes too technical for this book. But my main reason for not including it is that Fermat's solution is simply much better.

Outside observers often assume that the more complicated a piece of mathematics is, the more mathematicians admire it. Nothing could be further from the truth. Mathematicians admire elegance and simplicity above all else, and the ultimate goal in solving a problem is to find the method that does the job in the most efficient manner. Though the major accolades are given to the individual who solves a

particular problem first, credit (and gratitude) always goes to those who subsequently find a simpler solution.

Pascal's solution to the problem of the points was far more complicated than Fermat's and required some sophisticated—and daunting!—algebra. It is hard to follow even for a professional mathematician. But for all its complicated appearance, following it (or even producing it in the first place) is, again from the perspective of the professional mathematician, routine. Fermat's approach is undoubtedly more pedestrian, requiring only that you list all the possibilities and then simply count them. But by penetrating to the heart of the problem and doing *just what is required* to get the answer, Fermat's approach shows true genius.

CHAPTER 9

The Chance of Your Life

... Consequently, as you did not have my method when you sent me the division among several gamblers, but [since you had] only that of combinations, I fear that we hold different views on the subject.

I beg you to inform me how you would proceed in your research on this problem. I shall receive your reply with respect and joy, even if your opinions should be contrary to mine.

I am etc.

And so Pascal ends his long letter.³⁷ Although the subsequent development of probability theory was rapid, it might have been even faster if, shortly after Pascal and Fermat had their exchange of letters, Isaac Newton and Gottfried Leibniz had not invented calculus, the mathematics of continuous motion and change. To live at a time when new mathematics that will change the world is being invented is a rare privilege granted to few generations. For the mathematicians in the second half of the seventeenth century, there was not one such revolution going on but two. Although the social impacts of both calculus and probability theory were huge, calculus offered by far the greater challenge to mathematicians, since the mathematics was much more exceptional. As a consequence, most of the focus of the greatest mathematical minds was on the development of calculus. Probability theory took second place. Nevertheless, progress was steady and dramatic, as mathematicians not only continued to develop the mathematics of probability itself but also found ways to apply it to everyday life.

Jakob Bernoulli's introduction of *a posteriori* probability was a major advance. The distinction between that and the older "dice game" probability was the first step toward today's notions of *subjective* and *objective* probability.³⁸ Graunt's life tables and Nikolaus Bernoulli's use of probabilities in courts of law provided a numerical measure of the reliability of an individual's knowledge of a particular event, but those probabilities were objective. What classified them as *a posteriori* (rather than *a priori*) was that the probability values assigned to various events were determined not by reasoning logically from symmetry conditions, as in a game of dice or cards, but by (the *objective* process of) counting outcomes of trials in the world. *Subjective* probability takes this a step further.³⁹ One of the first people to attempt to capture mathematically what people do when they try to make rational decisions about the future was another Bernoulli: Johann's son Daniel.

FEAR OF FLYING

As friends of Leibniz, Jakob and Johann Bernoulli were among the first mathematicians to learn about the remarkable new method of calculus. In correspondence with Leibniz, Jakob and Johann carried out

much of the early development of the subject, and many of the techniques learned by today's calculus students were developed by the two Bernoullis. (Some historians suggest that the rich subject we today call calculus was created by not two but four mathematicians: Newton, Leibniz, and the two Bernoullis.) Daniel Bernoulli, too, did pioneering work on calculus, but he took a different course from his Uncle Jakob's, concentrating on applying the methods of calculus to flowing liquids and gases. Arguably the most significant of Daniel's many discoveries was what is now known as Bernoulli's equation, which describes one of the sources of lift that keep an aircraft aloft.⁴⁰ Though the airline industry's use of this equation was two hundred years away, other work of his was also relevant to the future world of air travel. Daniel's contributions to our understanding of probability help explain the oft-cited observation that, although modern air travel is the safest way to travel, many people nevertheless feel extremely nervous boarding a plane. Some people fear flying so much that they never get on a plane at all.

Such people may well know that the probability of being involved in a major accident is far less than when they travel by car. What scares them is the often horrific nature of an airline crash and the importance they attach to such an event, however unlikely.

The fear of lightning is similar. The tiny mathematical probability of being struck by a lightning bolt is far outweighed by the significance many people attach to it.

It was this essentially human aspect of probability calculations that interested Daniel Bernoulli. Was it possible, he asked, to make specific observations about the way people actually assessed risk? In 1738, he published a seminal paper on the issue in the *Papers of the Imperial Academy of Sciences in Saint Petersburg*.⁴¹ In that paper, he introduced the key new concept of *utility*, which he based on Huygens's notion of expectation.

By taking account of both the probabilities and the payoffs, expectation was supposed to measure the value of a particular risk or wager, from an individual's perspective. The greater the expectation, the more attractive the risk. That, at least, was the theory. For many examples it seemed to work well enough. But there was a problem, most dramatically illustrated by a tantalizing puzzle, proposed by Daniel's cousin Nikolaus, that came to be known as the Saint Petersburg paradox. Here it is:

Suppose I challenge you to a game of repeated coin tosses. If you throw a head on the first toss, I pay you \$2 and the game is over. If you throw a tail on the first throw and a head on the second, I pay you \$4 and the game is over. If you throw two tails and then a head, I pay you \$8 and the game is over. We continue in this fashion until you throw a head. Each time you throw a tail and the game continues, I double the amount you will win when you eventually throw a head.

Now imagine that a friend comes along and offers to pay you \$10 to take your place in the game. Would you accept or decline? What if your friend offered you \$50? Or \$100? In other words, how much do you judge the game to be worth to you?

This is exactly what expectation is supposed to measure. So what does it work out to be in this case? Well, in principle, the game can go on indefinitely—there are infinitely many possible outcomes: H, T H, T T H, T T T H, T T T T H, and so forth. The respective probabilities of these outcomes are $1/2$, $1/4$, $1/8$, $1/16$, $1/32$, and so forth. Thus, the expectation (i.e., your expected reward) is

$$(\frac{1}{2}) \times 2 + (\frac{1}{4}) \times 4 + (\frac{1}{8}) \times 8 + (\frac{1}{16}) \times 16 + \dots$$

This infinite sum can be rewritten as

$$1 + 1 + 1 + 1 + \dots$$

Since the sum goes on forever, the expectation is therefore infinite.

According to the theory, faced with an infinite expectation, you should not give up your opportunity to play for any amount of money. But most people—even probability theorists—are tempted to take the \$10 offer and would almost certainly take \$20. Why?

Pondering this question, as well as a number of other problems he saw with expectation, led Daniel Bernoulli to replace the highly mathematical concept of expectation with the far less formal idea of utility. Utility is intended to measure the significance of a particular outcome to a specific individual. As such, it is very much a personal thing. It depends on the value someone puts on a particular event. Your utility and mine might differ.

At first glance, the move to replace the mathematically precise concept of expectation with the decidedly personal idea of utility might appear to render further scientific analysis impossible. But it doesn't. Even for a single individual, it may well be impossible to assign specific numerical values to utility. Yet Bernoulli was able to make a profound observation about utility. He wrote that the "utility resulting from any small increase in wealth will be inversely proportionate to the quantity of goods previously possessed." Bernoulli's utility law explains why, even for a moderately wealthy person, the pain of losing half of one's fortune is much greater than the pleasure or benefit of doubling it. Thus, few of us are prepared to gamble half our wealth for the chance of doubling it. Only when we are truly able to declare "What have I got to lose?" are most of us prepared to take a big gamble.

For example, suppose you and I each have a net worth of \$10,000. I offer you a single toss of a coin. Heads, I give you \$5,000; tails, you give me \$5,000. The winner comes out with \$15,000, the loser with \$5,000. Since the payoffs are equal and the probability of each of us winning is $\frac{1}{2}$, we each have an expectation of zero. In other words, according to expectation theory, it makes no difference to either of us whether we play or not. But most of us would not play, because we would view the risk of losing unacceptable. The 0.5 probability of losing \$5,000 (half our wealth) far outweighs the 0.5 probability of winning \$5,000 (an increase of 50 percent).

The utility law likewise resolves the Saint Petersburg paradox. The longer the game goes on, the greater the amount you will win when a head is finally tossed. (If the game goes on for six tosses, you are bound to win over \$100. Nine tosses, and your winnings will exceed \$1,000. If it lasts for fifty tosses, you will win over \$1 million billion.) According to Bernoulli's law, once you reach the stage where your minimum winning represents a measurable gain *in your terms*, the benefit to be gained by playing longer starts to decrease. That determines the amount for which you would be prepared to sell your place in the game.

So utility superseded expectation. A similar fate eventually befell Bernoulli's utility concept when mathematicians and economists looked more closely at human behavior. Still, it was Daniel Bernoulli who first insisted that if you wanted to apply the mathematics of probability theory to real-world problems, you had to take account of the human factor.

What made risk management a mathematical science rather than a branch of psychology—although it undoubtedly has an element of art to its application—is that the human factor is itself amenable to mathematical analysis. Leading this development was none other than a Bernoulli, though a Frenchman living in England would make the real breakthrough.

RINGING IN THE BELL CURVE

Jakob Bernoulli had shown how to determine the number of observations needed to ensure that the probability determined by the sample was within a specified amount of the true probability. This was of theoretical interest, but of little practical use. For one thing, it required that you know the true probability in advance. Second, as Bernoulli had himself shown in the case of the marbles in the jar, the number of observations required to achieve a reasonably accurate answer could be quite large. Potentially much more useful was the opposite problem: Given a specified number of observations, can you calculate the probability that they will fall within a specified bound of the true value?

The first person to try to investigate this question was Jakob's nephew Nikolaus, who worked on it while completing his late uncle's work and preparing it for publication. To illustrate the problem, Nikolaus gave an example concerning births. Assuming a ratio of 18 male births to every 17 female births, for a total of 14,000 births the expected number of male births would be 7,200. He calculated that the odds were over 43 to 1 that the actual number of male births in the population would lie between $7,200 - 163$ and $7,200 + 163$ —that is, between 7,037 and 7,363. Nikolaus did not completely answer the question, but he made sufficient progress to publish his findings in 1713, the same year his deceased uncle's book finally appeared.

Some years later, Nikolaus's ideas were taken up by Abraham de Moivre (1667-1754), a French mathematician who, as a Protestant, had fled to England in 1688 to escape persecution by the Catholics. Unable to secure a proper academic position in his adopted country, de Moivre made a living by tutoring in mathematics and—a sure sign of the growth of probability theory for risk management—consulting for insurance brokers.

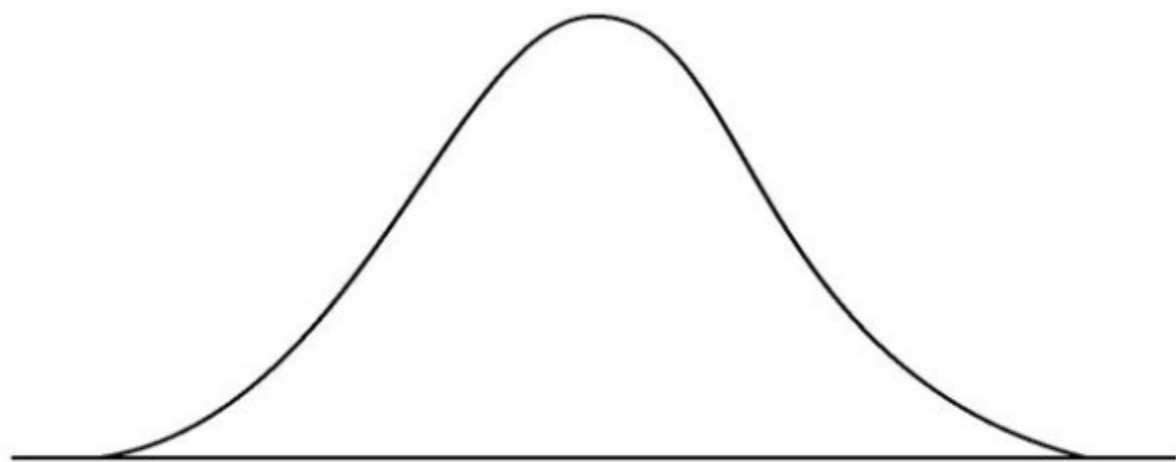


FIGURE 1. The Normal Distribution

Where Nikolaus Bernoulli had made only partial progress on his problem, de Moivre obtained a complete solution, which he published in 1733 in the second edition of his book *The Doctrine of Chances*. Using methods of calculus as well as probability theory, de Moivre showed how a collection of random observations would distribute themselves around their average value. These days, this distribution is known as the *normal distribution*. When it is represented graphically, plotting observations along the horizontal axis and the frequency (or probability) of each observation along the vertical axis, the resulting curve is shaped like a bell. For this reason it is often called the *bell curve* (see [Figure 1](#)).

The bell curve shows that most observations are clustered in the center, around the mean value. Going out from the middle, the curve slopes down symmetrically, with an equal number of observations on both sides of the mean. At first it falls slowly, then much more rapidly, finally becoming very flat toward the ends. In general, observations far from the mean are less frequent than observations close to the mean.

De Moivre discovered the bell curve by examining the behavior of random observations. Its elegant, symmetrical shape showed that beneath randomness lay an elegant geometry. From a mathematical point of view, that alone was a significant result. But that was not the end of the story. Eighty years later, the great German mathematician Karl Friedrich Gauss noticed that whenever he plotted a large number of measurements of something, the resulting curve always looked remarkably like de Moivre's bell curve. Different measurements of a particular distance on the earth's surface, for example, or of an astronomical distance, would form a bell-shaped cluster around a central value—the average of the individual measurements. The inevitable errors in taking measurements gave rise, it seemed, to a normal distribution of values.

By considering the bell curve not as a geometric feature of randomness but as an inevitable consequence of errors in measurements, Gauss realized that he could use the normal distribution to provide a framework for assessing the value of data. In particular, using the bell curve, it would be possible to assign (approximate) numerical probabilities to events, much as they could be assigned to

outcomes of a roll of dice or a spin of the roulette wheel. The closer a particular datum was to the mean, as measured on the bell curve, the greater the probability that it was correct. The techniques of probability theory developed by Pascal, Fermat, Huygens, the Bernoullis, and others could then be transferred from the gambling table to other areas of life. As a result of Gauss's contribution, the bell curve is sometimes called the *Gaussian distribution*.

The key technical concept Gauss needed to use the bell curve to assess the reliability of data was a measure of the dispersion around the mean that de Moivre himself had introduced. Known nowadays as the *standard deviation*, de Moivre's measure allows us to judge whether a set of observations is sufficiently representative of the entire population. Although based on some solid mathematics, standard deviation is a somewhat arbitrary notion that gives a numerical answer to the question "How much do the individual data points differ from the mean?" For a normal distribution, approximately 68 percent of the observations fall within one standard deviation of the mean and 95 percent fall within two standard deviations of the mean. For this reason, when newspapers and magazines publish results of surveys, they generally include a statement of the deviation. If they don't, you should view the results with suspicion. Remember, a man whose head is in an oven and whose feet are encased in ice could be said to feel just fine on average. But what a deviation!

With statistics playing such a major role in twenty-first-century life, the bell curve is an icon of our age. It allows us to assign numerical probabilities to events and thereby apply the methods of probability theory. In general, whenever there is a large population that produces data, and the data produced by each member of the population is independent of that produced by the others, then the bell curve will arise and Gauss's method provides a measure of the reliability of that data.

Following Gauss's lead, business leaders now use the bell curve to plan expansions, officials use it to determine public policy, educators use it to grade student examinations, pollsters use it to analyze their data and make their predictions, medical researchers use it to test the effectiveness of different treatments, economists use it to analyze economic performance, biologists use it to study plant growth, psychologists use it to analyze the behavior of their fellow humans, and insurance companies use it to set their premiums.

I'll look at this last application as an example.

THE PRICE ON YOUR HEAD

Few people would classify themselves as regular gamblers—in the sense, say, of frequenting casinos or participating in weekly poker games—yet a great many buy tickets in state lotteries. In so doing, they are gamblers in the technical sense of betting money in the hope of winning something, but they generally do not see themselves as gambling. They pay a small fixed fee, and their subsequent involvement is at most that of checking the list of winning numbers. Moreover, the game they are playing (in fact, it's hardly a game at all) has fixed odds and there is limited opportunity to secure an advantage. Most people who buy tickets are not aware of or even interested in the odds against winning—perhaps because the probability of winning for any single individual is essentially zero. Thus, the typical buyer of a lottery ticket is

unaffected by probability theory; his or her odds are effectively zero, and winning is a matter of pure luck.

The same is not true for the four areas where a great many people “gamble” every day: life, health, property, and savings. Apart from investing in the stock market, these are all a matter of taking on risks rather than gambling per se. But most of us—or the companies we deal with, such as the bank that lends us money—try to mitigate undesired consequences of these risks by taking out some form insurance or annuity. That is very definitely gambling, and the companies that assume our risk have to do the math right, to be able to charge us a fee that enables them to stay in business while being acceptably fair to their customers.

The oldest example of this form of gambling is the life annuity, which dates back at least to the Roman Empire. The person who takes out a typical annuity is not really gambling: you simply pay a fixed amount by a specified date in order to receive regular payments for the remainder of your life. But to the authorities or, these days, the companies that sell you the annuity, it is very much a gamble: they are betting on your life. If you die before they have paid out to you the full value of your purchase price, they have made a profit on you; if you live a long time, they take a loss. Clearly, their task is to set the purchase price so that, when averaged over all their customers, they make an acceptable profit while avoiding charges of profiteering on people’s misfortunes. There are generally laws that stipulate a minimum number of payments that must be made out to the heirs of an individual who dies before a fixed amount of time has elapsed.

Such laws existed in ancient Rome, where around 220 A.D., the jurist D. Ulpian devised the following table for converting a life annuity to one with a fixed number of payments.⁴²

AGE OF ANNUITANT	0–19	20–24	25–29	30–34	35–39	40 . . . 49	50–54	60–
DURATION OF ANNUITY (YEARS)	30	28	25	22	20	19 . . . 10	9	5

It is not known whether this table was based on observational data or was simply a legal stipulation, but the figure of thirty years’ duration for a person who dies before the age of twenty, in an era when the life expectancy was much lower, suggests the latter. Nikolaus Bernoulli discussed Ulpian’s table in his book *De usu artis conjectandi in jure*. He carried out his own computation based on Graunt’s life table and came up with a figure of twenty-one years, instead of thirty, for conversion when the annuitant is less than twenty years old at the time of death. Nikolaus examined another annuity system, offered by the magistrate of Amsterdam in 1673, and declared that the prices charged were close to those he himself had calculated.

In the Middle Ages, it was common practice for states and towns to sell annuities to raise funds. In general, the price was not dependent on the age of the annuitant, and as a result, buyers generally nominated healthy children. In his book, Bernoulli quoted an announcement made in London in 1704:

Yesterday the Queen issued an order for the sale of life annuities; in two hours £10,000 were subscribed and to date £100,000. These are the conditions for the annuities; if someone wants 10 pounds of yearly income with a possible extension to two succeeding persons, then he pays 90 pounds for one life, 100 pounds for two lives, 120 pounds for three lives following one another; and if one wishes to have 14 pounds yearly for 99 years, one pays 210 pounds for it.

A detailed study of annuities was made by a remarkable Dutchman by the name of Jan de Witt (1625-1672), who, after a brilliant early career as a mathematician, went into politics and became, at age twenty-eight, the prime minister of Holland. In need of funds to fight the English and French, he proposed the sale of annuities and wrote a paper that described how to price and structure them, based on an original and sophisticated mathematical analysis. The paper was published in 1671, one year before France's invasion of Holland, an event that led to his resignation from office and subsequent murder by a mob.

Soon after de Witt's work, the famous English scientist Edmund Halley (1656-1742) made further progress. Halley was an immensely talented and versatile scientist who did leading work in astronomy, geophysics, physics, and mathematics. He is best known today for predicting the return of a now eponymous comet, but he also wrote a paper on insurance mathematics, which he published in the Royal Society's scholarly publication *Philosophical Transactions*, of which he was the editor. Like everything else he did, it was first-rate.

Halley's study was a mathematically more sophisticated version of what John Graunt had done thirty years earlier. Halley's sources of data were the records of all births and deaths, listed according to gender and age, that had been kept in Breslau, in Silesia (in modern-day Poland), since the end of the sixteenth century. After Graunt's work, the Royal Society had been eager to locate another source of data to construct a life table, particularly one that was more complete and hence more reliable. When members heard of the Breslau records, which were indeed far better than those available to Graunt, they asked Halley to analyze them. This he duly did, presenting his results to the society in a 1693 paper with the pithy title *An Estimate of the Degrees of the Mortality of Mankind, drawn from curious Tables of the Births and Funerals at the City of Breslaw; with an Attempt to ascertain the Price of Annuities upon Lives*.

Unlike Graunt, Halley was a mathematician of great power, and his analysis was far more sophisticated than the one the major had carried out. The tables he produced provided the basis for offering annuities and life insurance policies for many years. Among many findings, Halley calculated the annual death rate and infant mortality and found that both were close to those obtained by Graunt, thereby supporting the belief that Graunt's figures were reliable.

With insurance in particular, ordinary people were able to take advantage of the ability that probability theory gives us to foresee the future (or at least quantify what it might bring). An insurance company cannot say with any certainty that you will live long or avoid major injury, but it can determine with considerable certainty what the various likelihoods are for society as a whole, or for the subgroup of "people like you," and that enables it to make a living by betting on your chances—just as a casino makes its profits not on one customer's luck but on the aggregate performance of all its customers.

With Halley's work, probability theory completed its migration from the gaming rooms to the everyday

world, where it could be used to make predictions about the future. Initially, those predictions were restricted to things that could be computed from aggregate data, like the expected age at which a person would die. But things were about to take a dramatic turn. The modern era was about to dawn.

THE REVEREND BAYES

In the case of insurance and similar applications of probability theory, prediction of the future is reliable only on average, over a large number of cases of essentially the same type. But today we use probability another way: to measure the likelihood of our being correct when we try to predict *one particular event*. Although it had to await the computer age to come to fruition, the final mathematical⁴³ step in the path from the solution of the problem of the points to today's risk-managed society was an ingenious and extremely powerful mathematical formula developed by an obscure eighteenth-century Presbyterian minister in England who studied mathematics as a hobby.

Thomas Bayes was born in London in 1702. His father, Joshua, was one of the first Nonconformist ministers to be ordained in England, which meant that his son was not allowed to study at Oxford or Cambridge. Accordingly, at seventeen, Thomas headed north to Scotland to enroll at the University of Edinburgh, where he studied logic and theology. In due course, he was ordained like his father and, in 1733, became a minister in the Presbyterian Chapel in Tun-bridge Wells, a well-to-do market town some thirty-five miles southeast of London. He served there until he retired in 1752, after which he remained in the town until his death, in 1761.

Bayes is recognized today as having a brilliant mathematical mind, but he published no original mathematics during his life. (After his death, a private notebook was found containing discussions of probability, trigonometry, geometry, solution of equations, series, differential calculus, electricity, optics, and celestial mechanics.) In 1736, he did publish *An Introduction to the Doctrine of Fluxions, and a Defence of the Mathematicians Against the Objections of the Author of The Analyst*, a predominantly philosophical article about Newton's calculus, rebutting Bishop Berkeley's attack on the logical foundations of Newton's method—but that seems to be all. Despite his lack of publications, his scientific abilities must have been known to others, for in 1742 he was elected a Fellow of the Royal Society.

When Bayes died, he bequeathed his papers to his friend Richard Price, a talented mathematician and also a Fellow of the Royal Society, who would go on to do foundational work in what later became actuarial science. Among the documents Price inherited, one item in particular caught his attention. Its title was *Essay towards solving a problem in the doctrine of chances*, and it outlined a radically new way to approach and compute probabilities. Recognizing its importance, Price submitted it to the Royal Society for publication with the following cover note:

I now send you an essay which I have found among the papers of our deceased friend Mr Bayes, and which, in my opinion, has great merit. . . . In an introduction which he has writ to this Essay, he says, that his design at first in thinking on the subject of it was, to find out a method by which we might judge concerning the probability that an event has to happen, in given circumstances,

upon supposition that we know nothing concerning it but that, under the same circumstances, it has happened a certain number of times, and failed a certain other number of times.⁴⁴

Other scholars at the Royal Society also felt Bayes' essay had merit, and it was duly published in the *Philosophical Transactions* in 1764.

The importance of Bayes' paper is not merely the formulation of the specific problem Price referred to in his cover letter, but the revolutionary and widely applicable approach to probability that Bayes adopted to solve the problem. The probabilities his method computes are probabilities attached to a person's knowledge of some event—they tell you the degree of confidence in some piece of information. In that respect, Bayes was following the work of Jakob Bernoulli, but he took things considerably further. (Unfortunately, he used the confusingly similar terms “prior probability” and “posterior probability,” but with different meaning from Bernoulli's *a priori* and *a posteriori*, so it is best to view Bayes' approach purely on its own.)

Bayes' method does not tell you how to calculate the probability of a hypothesis *de novo*. Rather, it is a method for *revising* a probability in light of new information. You start with a figure for the probability of a hypothesis, H ; that figure is called the *prior probability* for the hypothesis H . Given some new information E , you then carry out a calculation to obtain a revised probability for H ; that new figure is called the *posterior probability*. The revision is carried out by inserting the relevant figures into a mathematical formula known as Bayes' formula (or sometimes Bayes' rule).

The prior probability might be a guess or an estimate. Given sufficient new information, an application of Bayes' revision procedure can lead to a more accurate probability. By applying Bayes' method iteratively (generally using a computer) each time new information is obtained, one may turn even a fairly poor prior probability into a fairly reliable posterior probability. (On the other hand, the method is not immune to the familiar adage about excessive reliance on computer models: garbage in, garbage out.)

Because the method depends on an initial “seed” value for the (first) prior probability, for two hundred years after Bayes' method became known, it was largely ignored by the statistics and probability theory communities. Starting in the 1970s, however, it has grown in popularity, in large part because the availability of powerful computers has made it possible to run the process iteratively with large amounts of information, often overcoming the inaccuracy of a poor first prior.

BAYES' FORMULA

For the record, I'll describe Bayes' formula to calculate the probability of a certain hypothesis H , based on evidence E . (If mathematical formulas give you the willies, just skip to the next chapter, though if you do, you'll miss an opportunity to learn how to interpret a seemingly devastating medical prognosis.)

Let $P(H)$ be the probability that the hypothesis H is correct in the absence of any evidence—the *prior probability*. This is the figure you start with. Where you get it from is your business.

Let $P(H|E)$ be the probability that H is correct, given E . This is the revised estimate you want to

calculate. A quantity such as $P(H|E)$ is known as a *conditional probability*—the conditional probability of H occurring, given E .

Let $P(E|H)$ be the probability that E would be found if H were correct, and let $P(E|H_{\text{wrong}})$ be the probability that E would be found if H were false. Bayes' formula requires these figures.

To compute the new estimate, you also have to calculate $P(H_{\text{wrong}})$, the probability that H is false. This is easy: the probability that H is wrong is the same as 1 minus the probability that H is correct: $P(H_{\text{wrong}}) = 1 - P(H)$.

Bayes' formula is:

$$P(H|E) = \frac{P(H) \times P(E|H)}{P(H) \times P(E|H) + P(H_{\text{wrong}}) \times P(E|H_{\text{wrong}})}$$

One of the strengths of Bayes' approach is that it can sometimes guide us when our intuitions are wrong. For example, suppose you undergo a medical test for a relatively rare cancer. The cancer has an incidence of 1 percent among the general population. Extensive trials have shown that the reliability of the test is 79 percent. More precisely, although the test does not fail to detect the cancer when it is present, it gives a positive result in 21 percent of the cases where no cancer is present—what is known as a false positive. When you are tested, the test produces a positive diagnosis. The question is, What is the probability that you have the cancer?

If you are like most people, you will assume that if the test has a reliability rate of nearly 80 percent, and since you tested positive, the likelihood that you do indeed have the cancer is about 80 percent (i.e., the probability is approximately 0.8). Are you right?

No. Given the scenario I just described, the likelihood that you have the cancer is just 4.6 percent (i.e., the probability is 0.046). There is a less-than-5-percent chance that you have the cancer. Still a worrying possibility, of course. But hardly the scary 80 percent you thought at first.

Here is how you arrive at that figure:

$$P(H) = 0.01 \text{ (the cancer has a 1 percent incidence in the population)}$$

$$P(E|H) = 1 \text{ (the test always shows positive if the cancer is present)}$$

$$P(H_{\text{wrong}}) = 0.99 \text{ (99 percent of the population is cancer-free)}$$

$$P(E|H_{\text{wrong}}) = 0.21 \text{ (the test gives a false positive in 21 percent of cases)}$$

So by Bayes' formula

$$P(H/E) = \frac{0.01 \times 1}{(0.01 \times 1) + (0.99 \times 0.21)} = \frac{0.01}{0.2179} = 0.0459$$

Bayes' formula comes from a fairly straightforward enumeration and counting argument, much like the one Fermat used to solve the problem of the points. Rather than present the argument in an abstract way, I'll go through it using the example of the cancer test that we just solved.

To keep the arithmetic simple, I'll assume a total population of exactly 10,000 people. Since all we are ultimately concerned about is percentages, this simplification will not affect the final answer. I'll also assume that the various probabilities are reflected exactly in the actual numbers. Thus, of the total population of 10,000, precisely 100 will have the cancer and 9,900 will not.

In the absence of the test, all you can say about the likelihood of your having the cancer is that there is a 1 percent chance that you do. This is the prior probability. Then you take the test, and it shows positive. How do you revise the probability that you have the cancer?

Well, there are 100 individuals in the population who have the cancer, and for all of them, the test will correctly give a positive prediction, thereby identifying 100 individuals as having the cancer.

Turning to the 9,900 cancer-free individuals, for 21 percent of them, the test will incorrectly give a positive result, thereby identifying $9,900 \times 0.21 = 2,079$ individuals, as having the cancer.

Thus, in all, the test identifies a total of $100 + 2,079 = 2,179$ individuals as having the cancer. Having tested positive, you are among that group. (This is precisely what the test evidence tells you.) The question is, Are you in the subgroup that really does have the cancer, or is your test result a false positive?

Of the 2,179 identified by the test, 100 really do have the cancer. Thus, the probability of your being among that group is $100/2,179 = 0.0459$.

And that's it.

Despite its simplicity, Bayes' method can be tricky to apply correctly, and if you are not careful, it is easy to go astray. The issue is not so much the formula but the concept of conditional probabilities.

For instance, unscrupulous lawyers have been known to take advantage of the innumeracy of judges and juries by deliberately confusing the two conditional probabilities $P(G|E)$, the probability that the defendant is guilty, given the evidence, and $P(E|G)$, the conditional probability that the evidence would be found, assuming the defendant were guilty. Such misuse of probabilities is called the *prosecutor's fallacy*, and even ethical prosecutors sometimes succumb to it unintentionally in cases where scientific evidence such as DNA testing is involved (e.g., paternity suits and rape and murder cases). Prosecuting attorneys in such cases have been known to provide the court with a figure for $P(E|G)$, the likelihood (more precisely, the *unlikelihood*) of finding the evidence if the defendant were guilty, whereas the figure relevant to deciding guilt is $P(G|E)$, which, as Bayes' formula shows, is generally much lower than $P(E|G)$. Unless there is other evidence that puts the defendant into the group of possible suspects, this use of $P(E|G)$ is highly suspect and perhaps should be prohibited. The reason is that it ignores the initial low prior

probability that a person chosen at random is guilty of the crime in question, and, as the example of the cancer test shows, the difference can be substantial.

Bayes' formula essentially took probability inside the human mind—applying it to our knowledge of a single, one-off event. With it, humankind had another powerful tool for predicting the future. But it took mathematicians many decades to discover just how powerful it can be.

All rights reserved. No part of this book may be reproduced in any manner whatsoever without written permission except in the case of brief quotations embodied in critical articles and reviews. For information, address Basic Books, 387 Park Avenue South, New York, NY 10016-8810.

Books published by Basic Books are available at special discounts for bulk purchases in the United States by corporations, institutions, and other organizations. For more information, please contact the Special Markets Department at the Perseus Books Group, 2300 Chestnut Street, Suite 200, Philadelphia, PA 19103, or call (800) 810-4145, ext. 5000, or e-mail special.markets@perseusbooks.com.

The unfinished game : Pascal, Fermat, and the seventeenth-century letter that made the world modern / Keith Devlin. p. cm.

Includes bibliographical references and index.

eISBN : 978-0-786-72632-5

1. Probabilities. 2. Pascal, Blaise, 1623-1662—

Correspondence. 3. Fermat, Pierre de, 1601-1665—

Correspondence. 4. Mathematicians—France—Correspondence.

I. Title.

QA273D455 2008

519.2—dc22

2008012222