CHAPTER 7

CLASSICAL PROBABILITY

In a letter written in September 1749 to Frederick the Great, the King of Prussia, the great mathematician Leonhard Euler (1707–1783) wrote:

> My researches on the hydraulic machine occupying me again for some days, I take the liberty to render account of the examination of the Italian Lottery, for which it has pleased Your Majesty to charge me so graciously. First I have determined by the calculus of probabilities, how much each player ought to pay in order that the advantage were so much equal for him as for the bank.[1]

This is one of the first clear references to the *calculus of probabilities*. Before 1750, the mathematical analysis pertaining to games of chance had almost always been known as the *doctrine of chances*. After that, this subject was more frequently termed the calculus of probabilities or sometimes the *theory of probability*.

This change in terminology reflected a major shift in thinking. From this point on we can find several writers referring to the calculus of probabilities in addition to, or in place of, the doctrine of chances. Moreover, the mathematical expositions became increasingly framed in terms of probability directly, although references to chances, odds, and expectations still occurred as well. By the 1760s, the idea of mathematical

Willful Ignorance: The Mismeasure of Uncertainty, First Edition. Herbert I. Weisberg.

Landmarks of Probability

1600

1654 Pascal & Fermat
1657 Huygens

1700

1713 Bernoulli
1718 De Moivre

1764 **Bayes & Price**
1774 **Laplace**

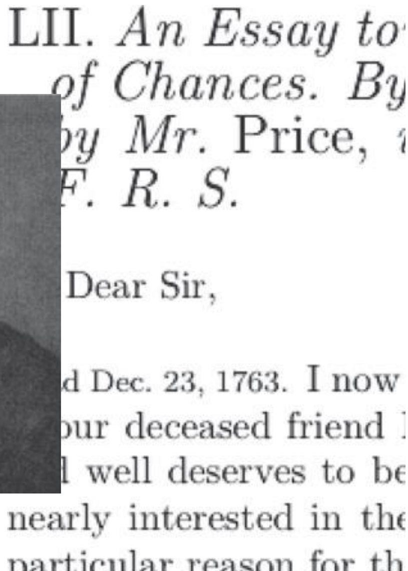
1800

1837 Poisson
1838 De Morgan
1843 Mill
1843 Cournot
1849 Ellis
1854 Boole
1866 Venn

1892 Peirce

1900

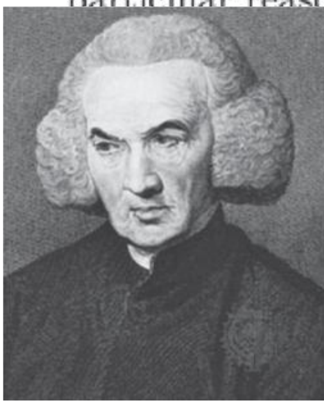
1921 Keynes
1921 Knight
1922 Fisher
1926 Ramsey
1926 Neyman & Pearson
1928 Von Mises
1937 De Finetti
1949 Reichenbach
1954 Savage
1959 Popper


LII. *An Essay to*
of Chances. By
by Mr. Price,
F. R. S.

Dear Sir,

d Dec. 23, 1763. I now
our deceased friend
l well deserves to be
nearly interested in the
particular reason for th
oper.
u know,
much es
ich he h
subject
probabil
on that v

Pierre-Simon Laplace



Richard Price

1/3 1/3 1/3

probability seems to have become so established that these earlier concepts were rarely mentioned explicitly.

During this critical transitional period, the concept of probability, defined as the *fraction of chances*, became solidified. This idea of probability as a fraction of some vaguely defined and metaphysical "chances" has been referred to by historians as *classical probability*. Concomitantly, the interpretation of this fraction of chances as the epistemic degree of certainty gained widespread currency. However, it would be a mistake to assume that probability was being applied much more broadly to civil and economic affairs. Its use was still restricted almost entirely to the narrow context of gambling in which it had originated. On the other hand, the potential to expand this new and mathematically tractable version of probability well beyond these bounds was starting to be recognized. Beginning in the early 1770s, Pierre-Simon Laplace would take the lead in this endeavor.

REVOLUTIONARY REVERENDS

Laplace was a towering figure who dominated the landscape of probability for 50 years. His monumental scientific and mathematical achievements earned Laplace the sobriquet of The French Newton. However, before we tackle the subject of Laplace's role in the evolution of probability, a slight detour is in order. A famous episode in England during the 1760s occurred that is usually considered an interesting sidebar, but may in fact have played an important role in the development of probability and statistics.

The protagonists in this little drama were two clergymen living in mid-century England. One was renowned in his time for many reasons, including his contributions related to probability and statistics, but he is almost entirely overlooked today. The other led a quiet and private life but has become quite notorious in ours, mainly for a single posthumous contribution to the theory of probability.

The Reverend Thomas Bayes

Although he is famous today for Bayes's rule, perhaps the best-known formula in the entire canon of probability, the eponymous Thomas Bayes

(1701–1761) cast almost no shadow during his rather uneventful life.[2] Very little is known about the details of his existence. His now-famous rule, or theorem, only came to light posthumously, thanks to the efforts of his friend, the Reverend Richard Price. Moreover, the essay in which the celebrated theorem was presented was ignored until it was rediscovered later, after a far more immediately influential version of essentially the same idea was independently proposed by Laplace.

The traditional rendition of the Bayesian story starts off with the demise of a little-known Protestant minister named Thomas Bayes in 1761. Among his papers were found some documents pertaining to obscure mathematical matters, in which Bayes, as an amateur mathematician, had been interested. In his will, Bayes had mentioned the Reverend Richard Price, "now I suppose preacher at Newington Green," to whom his papers should be sent. Price recognized the potential importance of Bayes's essay on probability and arranged for publication in the *Philosophical Transactions of the Royal Society*. The rest is silence.

In recent years, scholars have learned a bit more about this episode and about Bayes the man, but not very much. Thomas Bayes was born around 1701, but the exact date is uncertain. His father was a Presbyterian minister and came from a fairly well-to-do family of Sheffield, in the north of England. Their fortune had been derived from the cutlery business, for which Sheffield was renowned. At that time, those who professed religious views that were not sanctioned by the Church of England were called "nonconforming" or "dissenting" Christians. Dissenters were allowed to practice and preach but were denied certain rights, including the ability to attend Oxford or Cambridge. Consequently, Thomas was educated within the schools that were open to Dissenters at that time, and he matriculated at the University of Edinburgh in 1719.

While his course of study was designed primarily to prepare him for a career in the ministry, he must have displayed some talent for mathematics. At Edinburgh, Bayes studied under James Gregory (1666–1731), younger brother of David Gregory.[3] James had assumed the professorship of mathematics in 1692 after David left Scotland to set up shop at Oxford. From James Gregory, Bayes would undoubtedly have learned about Newton's theory of fluxions. Whether he was exposed to the doctrine of chances during his studies at Edinburgh is unknown.

probability seems to have become so established that these earlier concepts were rarely mentioned explicitly.

During this critical transitional period, the concept of probability, defined as the *fraction of chances*, became solidified. This idea of probability as a fraction of some vaguely defined and metaphysical "chances" has been referred to by historians as *classical probability*. Concomitantly, the interpretation of this fraction of chances as the epistemic degree of certainty gained widespread currency. However, it would be a mistake to assume that probability was being applied much more broadly to civil and economic affairs. Its use was still restricted almost entirely to the narrow context of gambling in which it had originated. On the other hand, the potential to expand this new and mathematically tractable version of probability well beyond these bounds was starting to be recognized. Beginning in the early 1770s, Pierre-Simon Laplace would take the lead in this endeavor.

REVOLUTIONARY REVERENDS

Laplace was a towering figure who dominated the landscape of probability for 50 years. His monumental scientific and mathematical achievements earned Laplace the sobriquet of The French Newton. However, before we tackle the subject of Laplace's role in the evolution of probability, a slight detour is in order. A famous episode in England during the 1760s occurred that is usually considered an interesting sidebar, but may in fact have played an important role in the development of probability and statistics.

The protagonists in this little drama were two clergymen living in mid-century England. One was renowned in his time for many reasons, including his contributions related to probability and statistics, but he is almost entirely overlooked today. The other led a quiet and private life but has become quite notorious in ours, mainly for a single posthumous contribution to the theory of probability.

The Reverend Thomas Bayes

Although he is famous today for Bayes's rule, perhaps the best-known formula in the entire canon of probability, the eponymous Thomas Bayes

(1701–1761) cast almost no shadow during his rather uneventful life.[2] Very little is known about the details of his existence. His now-famous rule, or theorem, only came to light posthumously, thanks to the efforts of his friend, the Reverend Richard Price. Moreover, the essay in which the celebrated theorem was presented was ignored until it was rediscovered later, after a far more immediately influential version of essentially the same idea was independently proposed by Laplace.

The traditional rendition of the Bayesian story starts off with the demise of a little-known Protestant minister named Thomas Bayes in 1761. Among his papers were found some documents pertaining to obscure mathematical matters, in which Bayes, as an amateur mathematician, had been interested. In his will, Bayes had mentioned the Reverend Richard Price, "now I suppose preacher at Newington Green," to whom his papers should be sent. Price recognized the potential importance of Bayes's essay on probability and arranged for publication in the *Philosophical Transactions of the Royal Society*. The rest is silence.

In recent years, scholars have learned a bit more about this episode and about Bayes the man, but not very much. Thomas Bayes was born around 1701, but the exact date is uncertain. His father was a Presbyterian minister and came from a fairly well-to-do family of Sheffield, in the north of England. Their fortune had been derived from the cutlery business, for which Sheffield was renowned. At that time, those who professed religious views that were not sanctioned by the Church of England were called "nonconforming" or "dissenting" Christians. Dissenters were allowed to practice and preach but were denied certain rights, including the ability to attend Oxford or Cambridge. Consequently, Thomas was educated within the schools that were open to Dissenters at that time, and he matriculated at the University of Edinburgh in 1719.

While his course of study was designed primarily to prepare him for a career in the ministry, he must have displayed some talent for mathematics. At Edinburgh, Bayes studied under James Gregory (1666–1731), younger brother of David Gregory.[3] James had assumed the professorship of mathematics in 1692 after David left Scotland to set up shop at Oxford. From James Gregory, Bayes would undoubtedly have learned about Newton's theory of fluxions. Whether he was exposed to the doctrine of chances during his studies at Edinburgh is unknown.

Almost nothing of his life between school and 1733 is known. In that year (probably) he arrived in Tunbridge Wells, located 40 miles southeast of London, to become the resident Presbyterian minister. The town was at that time a popular vacation destination, attracting many summer visitors and known originally for its reputedly health-enhancing spa. He remained there for the remainder of his life, staying on even after he retired from the pulpit in 1752. Bayes was a rather uninspiring speaker, was not very active in the social sphere, despite the rather lively atmosphere of Tunbridge Wells, and led a quiet bachelor's life.

As for mathematical pursuits, we know that in 1736 he published (anonymously) a text on Newtonian calculus meant in part as a defense against the philosophical criticisms that had been leveled by Bishop George Berkeley (1685–1753) against the infinitesimal calculus.[4] Around this time Bayes may have met Lord Philip Stanhope (1694–1773), amateur mathematician and patron of the sciences, who encouraged Bayes's mathematical interests. In 1742, Stanhope was responsible for the election of Reverend Bayes to membership in the Royal Society. It is apparent from the wording of his nomination that Bayes was regarded as an accomplished mathematician[5]:

> The Revd Thomas Bays of Tunbridge Wells, Desiring the honour of being Elected into this Society, we propose and recommend him as a Gentleman of known merit, well skilled in Geometry and all parts of Mathematical and Philosophical Learning, and every way qualified to be a valuable member of the same.

While there is scant evidence of his active participation in the Royal Society, Bayes appears to have become an advisor to Stanhope and through him to others in the Earl's wide circle. In 1755, Bayes was asked to comment on a paper by Thomas Simpson, de Moivre's erstwhile admirer and sometime nemesis.[6] Simpson had tried to address an important practical problem—how can we best estimate the true value of some quantity based on several fallible measurements? This was a critical issue for astronomers, who were trying to determine the precise position of a star or some other heavenly object for which several independent observations had been obtained.

The conventional approach to such problems had been qualitative. The "best" measurement was selected based on a detailed analysis of the conditions and instruments that had been employed to produce the observation. However, some astronomers had begun to adopt a new practice of taking a mean, or average, of the several measurements rather than trying to decide subjectively which one was the most reliable. In effect, they were applying the principle of willful ignorance. At the time it was an open question whether it was better to apply expert judgment based on all the known facts or to *ignore* the myriad of unknown possible "causes" of measurement error?

Simpson was the first to link the problem of observational errors with the doctrine of chances. His paper was in the form of a letter addressed to Lord Macclesfield that was read to the Royal Society on April 10, 1755, and appeared later that year in the *Philosophical Transactions*. Here is how he framed the issue:

> It is well known to your Lordship, that the method practiced by astronomers, in order to diminish the errors arising from the imperfections of instruments, and of the organs of sense, by taking the Mean of several observations, has not been so generally received, but that some persons, of considerable note, have been of opinion, and even publickly maintained, that one observation, taken with due care, was as much to be relied on as the Mean of a great number.
>
> As this appeared to me to be a matter of much importance, I had a strong inclination to try whether, by the application of mathematical principles, it might not receive some new light; from whence the utility and advantage of the method in practice might appear with a greater degree of evidence.

From our vantage point, it is remarkable that this problem had apparently *not* previously been regarded as a mathematical question! Simpson was breaking new ground just by considering it in this light. He went on to offer a sophisticated rationale for preferring the statistical approach of averaging several observations. Essentially, he illustrated how, under certain assumptions, the average of the measurement errors becomes smaller as the number of observations grows larger. Focusing on the *errors* rather than on the value of the quantity to be estimated was an important conceptual breakthrough that would prove very fruitful.

Simpson's paper represents one of the earliest practical applications outside games of chance of a basic concept that would later be termed a *random variable*. A random variable is a quantity that can take on different possible values with certain probabilities. For example, the outcome of rolling a single die can be considered a random variable. So can the predicted price of a particular stock at some future date. Corresponding to each possible value of the stock, a probability can, in principle, be specified. The aggregate of all these probabilities constitutes the *probability distribution* of the random variable. A particular probability distribution can be characterized in many different ways, but the mean (average) value is the most common summarization.

Simpson conceived of the observational *error* (measured in seconds of arc) as an integer value between −5 and +5. Corresponding to each of these 11 possible values (including zero) of the measurement error he posited a particular probability. In effect, the measurement error was regarded as a random variable, and the set of 11 equal probabilities was its probability distribution. Applying the doctrine of chances, he found that the probability distribution of the mean of several errors tended to become closer to zero as the number of observations increased. Specifically, the mean of the errors had a greater probability of being within any specified distance of zero than the error for any individual observation. Therefore, the mean value of the observed positions would tend to become closer to the true position as the number of observations increased.

Simpson's discussion of his analysis illustrates beautifully the gradual transition of thought taking place between the older emphasis on *chances* and the emerging focus on *probability*. He seems to vacillate, invoking probability in the context of calculations but referring to chances and odds for the ultimate expression of his results. Here is how he explains that the probability that the error is less than two arc seconds is much higher when taking the mean than when using a single observation:

> These particulars being premised, let it be now required to find what *probability*, or *chance* for an error of 1, 2, 3, 4, or 5 seconds will be, when, instead of relying on one, the mean of the observations is taken …
>
> To determine, now, the *probability* that the result comes within two seconds of the truth … the *odds*, or proportion of the *chances*, will therefore be … as 29 to 1, nearly. But the proportion, or *odds*, when a single observation is taken, is only as 2 to 1: so that the *chance* for an error exceeding two seconds, is not $\frac{1}{10}$ th part so great, from the mean of six, as from one single observation (emphasis added).

By probability he clearly intends the fraction of the chances, but referring to chances and odds was necessary to convey his meaning to the intended audience.

After agreeing to review Simpson's letter to Lord Macclesfield, Bayes must have been intrigued. His review indicated that he found the mathematical development to be quite correct. However, he raised a simple but important point about the generality claimed by Simpson for his conclusion. Simpson had argued that taking the mean was *always* likely to be the preferred approach. Bayes astutely observed that Simpson's conclusion depended on assuming that the measuring device was free from systematic bias:

> On the contrary the more observations you make with an imperfect instrument the more certain it seems to be that the error in your conclusion will be proportional to the imperfection of the instrument made use of. For were it otherwise there would be little or no advantage in making your observations with a very accurate instrument rather than with a more ordinary one, in those cases where the observation cou'd be very often repeated: & yet this I think is what no one will pretend to say.[7]

In other words, measurement accuracy depends not only upon the precision of the observational device or procedure, but also on proper calibration. Put simply, you need to be aiming at the right target. Simpson appreciated the wisdom of Bayes's remark, and in later writings amended his assumptions to exclude the case of a biased instrument.

It is possible that Simpson's paper was the stimulus for Bayes to think more deeply about matters of probability, leading to his famous essay. However, it seems more plausible that the essential ideas in the essay were being discussed by Bayes (and perhaps others) as early as the 1740s.[8] David Hartley (1705–1757) was a physician, philosopher, and founder of the "associationist" school of psychology. By around 1735, he had established a reputation in London as a respected medical practitioner and scientist and in 1736 he was elected to the Royal Society. In 1749, Hartley published a remarkable treatise combining philosophy and theology with a very thoughtful and imaginative account of psychological processes.

In his medical research activity, Hartley was one of the pioneers in appealing to statistical data, and was apparently familiar with de Moivre's *Doctrine of Chances*. Through the Royal Society, he was probably acquainted with de Moivre and with Thomas Bayes, who were also members. Hartley's writing about probability is revealing as a possible reflection of the general transition in the understanding of probability that was occurring in England at the time.

As I have mentioned, both de Moivre and Simpson had begun to think in terms of fractional probabilities, rather than chances or odds. However, neither of these mathematicians had particularly emphasized the possible link between this mathematical concept and the venerable epistemic connotation of probability. Writing in the 1740s, David Hartley saw clearly a connection between uncertainty in general and mathematical probability:

> If it be asked, upon what authority absolute certainty is represented by unity, and the several degrees of probability by fractions less than unity, in the doctrine of chances? Also, upon what authority the reasoning used in that doctrine is transferred to other subjects, and made general, as here proposed? I answer, that no person who weighs these matters carefully can avoid giving his assent; and that this precludes all objections. No skeptic would, in fact, be so absurd as to lay two to one, where the doctrine of chances determines the probability to be equal on each side; and therefore we may be sure that he gives a practical assent at least to the doctrine of chances.[9]

By the 1740s, the merging of probability as a degree of certainty and probability as a fraction of chances subject to the rules governing games of chance had apparently progressed to the extent of becoming self-evident, as it would remain for succeeding generations. Furthermore, another statement by Hartley refers to "an ingenious friend" who had communicated a solution to the "inverse problem" that, as we shall see, came to be known as Bayes's Theorem.[10] So, it is possible that Bayes's knowledge of probability and his ideas about the inverse problem may have existed for some time prior to his review of Simpson's letter to Lord Macclesfield.

The Reverend Richard Price

Richard Price (1723–1791) was virtually unknown in 1761, when he answered the call to examine the papers of his late friend, or at least acquaintance, Thomas Bayes. To statisticians who have heard the traditional story about Bayes's Theorem, Price is a rather shadowy figure, like an anonymous midwife who mediates the birth of the hero and then fades silently out of the story. Most statisticians would be surprised (as I was) to learn that Price in 1761 was destined to become one of the most famous men of the eighteenth century, and not just in England. Furthermore, a large measure of his prominence would flow indirectly from his knowledge of probability, stimulated by the encounter with Bayes's essay in 1761.

Who was Richard Price? His father, a Congregational minister in the Welsh city of Glamorgan, expected his son to follow in his ecclesiastical footsteps.[11] In 1739, after his father's unexpected death, Richard was sent to London to attend the Fund Academy, one of the schools established by the Dissenters to provide the higher education their children were barred from obtaining at the English universities. Besides preparation for the ministry, Price acquired a fairly extensive grounding in science and mathematics under the tutelage of John Eames (1686–1744), a disciple of Newton and member of the Royal Society. In fact, Eames was one of the signatories to the nomination of Thomas Bayes for membership in 1742.

After graduating and becoming ordained, Price moved to Stoke Newington, then a suburb of London. There he served as chaplain to the family of George Streatfield, and gave sermons as an assistant minister to several local congregations. In 1756, both Streatfield and a wealthy uncle of Price died, and each left him a generous bequest. This unanticipated bounty provided Price with sufficient resources to live comfortably and soon to marry his sweetheart, Sarah Blundell. The couple was happily married for the ensuing 30 years, despite Sarah's physical frailty and inability to have children. By all accounts, both Price and his wife were friendly, outgoing, and well liked. During most of their marriage they remained in Stoke Newington, where Price eventually became established as a full-time preacher.

While not particularly noted for his oratorical skills, Price was an excellent writer on both religious and secular matters. In 1758, he published *A Review of the Principal Questions and Difficulties in Morals*, the first of several publications (and numerous letters) that dealt with philosophical and theological matters.[12] Based on these writings, Price was awarded a Doctor of Divinity degree from the University of Edinburgh in 1768. These various written works reveal Price as somewhat of a contradiction, from our modern perspective.

On the one hand, he held fast to traditional ideas about Divine benevolence, literal salvation, and God's active role in shaping the world. On the other hand, his political and economic views were surprisingly modern, and even considered radical by his contemporaries. Moreover, he was in the vanguard of mathematical and scientific progress. The unifying theme in much of his thinking was the Newtonian belief in natural laws that could be understood by the employment of man's rational faculties. Price was a strong believer in individual liberty and freedom from repression, which he viewed as both morally right and confirmed by reason. One of his congregants was Mary Wollstonecraft (1759–1797), who gained fame as an early feminist. She was strongly influenced by Price's liberal philosophy.

As mentioned, in 1761 Price accepted the responsibility of editing the papers of Thomas Bayes. It has been speculated that Price's "editing" role may actually have been considerable. All we know for sure is what Price wrote in the introduction that he added, in which he was characteristically modest. "An Essay Towards Solving a Problem in the Doctrine of Chances" was finally submitted to the Royal Society by Price in 1763 and published in the *Philosophical Transactions* the following year.[13] Based on this work and perhaps other evidence of his mathematical ability, Price was elected as a member of the Royal Society in 1765. Meanwhile, Bayes's *Essay* appears to have elicited no significant interest at the time and was forgotten, although through Price it did indirectly contribute to the early development of actuarial science.

In 1764, Price was consulted by "three gentlemen" regarding a proposed plan of operation for the new Society for Equitable Assurance on Lives and Survivorships.[14] Founded 2 years earlier as the first mutual life insurance company, the Equitable was seeking methods for assessing risk, at a time when mortality data remained very sparse. Thus began a very fruitful long-term sideline for Price. He became a recognized expert on the financing of insurance, annuities, and various "friendly society" schemes to provide old-age and sickness benefits for workers and their families.

Price served as an advisor to the Equitable Society for many years, and in 1769 published *Observations on Reversionary Payments*.[15] This text eventually ran through seven editions and supplanted de Moivre's book as the classic source on actuarial methods. Price drew upon his mathematical background, including his knowledge of probability theory, to develop the most sophisticated analyses of annuities and other financial instruments to date. Furthermore, he developed new mortality data, finally rendering Halley's 1693 data from Breslau obsolete.

Price became interested in the subject of vital statistics not only for actuarial applications but also for political and philosophical reasons. He became a leading figure in the debate over the state of Britain's population. Price was among those who argued that the population was declining, and adduced statistical evidence supporting this (incorrect, as it turns out) position. He believed in the virtues of hard work and a simple agrarian lifestyle, and blamed luxurious living and increasing urbanization for the alleged population decrease.

Through his writings on financial, actuarial, and demographic matters, Price had by the 1770s become quite well known, and was regularly called upon by prominent politicians for advice. He had also developed a number of relationships in the scientific world. In particular, from 1766 on he was very close to Joseph Priestley (1733–1804), who is hailed as a pioneer of modern chemistry, and discoverer of oxygen, but was also active as a theologian, educator, and political philosopher.

Price was widely respected as kindhearted and magnanimous in his dealings, even in his philosophical disputes. He became personally acquainted with many of the leading progressive thinkers of the day, including the prison reformer John Howard (1726–1790), economist Adam Smith (1723–1790), and philosopher David Hume (1711–1776). He was generally well regarded, even by those, like Priestley and Hume, whose sometimes extreme views conflicted with his own religious beliefs. Price not only preached, but truly practiced, tolerance for free expression and participation in rational dialogue. So, it "was ironical that Price, so pacific and equable, had a genius for starting controversies."

Some of these unintended public imbroglios revolved around issues pertaining to political economy, such as whether or not Britain's population was truly in decline or how best to handle the mounting National Debt. However, in the mid-1770s international political events would catapult this mild-mannered intellectual to a much greater level of notoriety. Sometime around 1760, Price had become friendly with Benjamin Franklin (1706–1790), who was then living in London. Through Franklin, he came to know and admire many of the leading American political figures as they passed through England over the years, including John Adams (1735–1826), Thomas Jefferson (1743–1826), Thomas Paine (1737–1809), and Josiah Quincy, (1709–1784). As relations between England and the American Colonies deteriorated, Price became a strong supporter of the American cause, which he regarded as a beacon of divinely inspired light. He hoped that the example of a nation established on the ideals of liberty would inspire Britain and other countries to loosen the reins of political and religious intolerance.

On February 8, 1776, Price published a rather inflammatory pamphlet—*Observations on Civil Liberty and the War with America*.[16] He argued that American independence was not only morally justified, but would also benefit Great Britain in the long run. He even advocated the formation of a similar kind of confederacy among the states of Europe. The pamphlet sold out immediately, was reprinted several times within a few weeks, and sold over 60,000 copies in total. It set off a chain reaction of literary salvos pro and con, and may even have influenced the framers of the Declaration of Independence.

In the spring of 1777, Price wrote another pamphlet in which he made suggestions for dealing with the serious financial problems faced by the fledgling American republic.[17] A year later, he politely declined an invitation from the Continental Congress to come and assist personally in the new nation's financial administration. In the Congressional debates over the national debt and other problems relating to fiscal policy, Price was cited as a leading authority in such matters. For a man who had never set foot on American soil, Price was a most vocal and influential supporter of the American Revolution and the new nation it produced.

After the Revolutionary War, he continued an active correspondence with Franklin and many other prominent citizens of the new country who valued his advice and considered him a true friend. When the Bastille fell in France on July 14, 1789, Dr. Price believed that the cause of liberty had taken another major step forward. In November of that year, he gave an impassioned sermon that burnished his (unwanted) reputation as a flaming radical. However, his health was already beginning to decline, and he eventually died on April 26, 1791, too soon for him to witness the chaotic and tragic conclusion of the drama beginning to unfold in Paris, about which he had been so hopeful.

In a real sense, Price was a forgotten founding father of the United States. The esteem in which he was held by Americans in his time was evidenced by an extraordinary tribute received by Price in absentia in New Haven, Connecticut, in 1781. Two honorary doctorates were awarded that year at the Yale University graduation ceremony. One was given to General George Washington, Commander-in-Chief of the Continental Army and father of his country, the other to Dr. Richard Price of Newington Green, philosopher, preacher, actuary, and vicarious revolutionary.

The Famous Essay

It is a stretch to attribute much of Price's later fame to his role in editing Thomas Bayes's essay. However, this effort did lead almost immediately to membership in the Royal Society and to his early actuarial work, including the *Observations on Reversionary Payments*, which by the early 1770s had made him a respected authority on financial and statistical matters. However, as mentioned above, the direct effect of Bayes's essay on the wider world was essentially nil. Why was this work, destined to become so celebrated much later on, so completely ignored at the time?

Perhaps the best answer is that Bayes's innovations were both mathematically complex and completely original. They represented in several ways a radical shift in perspective for which his contemporaries were not yet ready. Fortunately, in Richard Price he had chosen a literary executor who could appreciate his insights and the subtlety of his reasoning. Unfortunately, today there is tremendous confusion about exactly what Bayes's Theorem really is, and what it was that Thomas Bayes can be said to have "discovered."

To understand the essence of Bayes's rule, consider the following simple problem. A single ball is to be drawn randomly from one of two urns, each of which contains 5 balls. Urn A contains 4 white balls and 1 black one; urn B contains 2 white and 3 black balls (see Figure 7.1). First one of the two urns will be chosen randomly, and then the ball will be drawn from that urn. Assume you are told that the probability is 1/3 that Urn A will be selected and is 2/3 that Urn B will be selected. You will not be allowed to see which urn is chosen.

Suppose that after the drawing you are informed that the ball is white. What is the probability that this ball came from Urn A? Note that *a priori* (beforehand) the probability of Urn A was 1/3. However, after seeing the white ball, you have additional information. How should this a posteriori

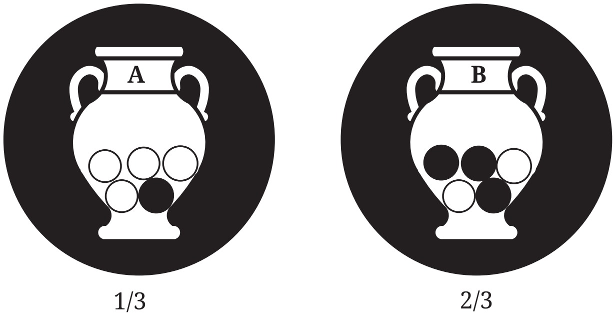


FIGURE 7.1 Example of the type of problem solved by Bayes's Theorem.

(afterward) knowledge affect the probability that Urn A was in fact the source? Since Urn A contains 4 white balls and Urn B only 2, it is obvious that the posterior probability must be greater than 1/3. But how much greater?

This is where Bayes's rule comes in. In essence, it formalizes the following intuitive reasoning. Since Urn B has an a priori probability of 2/3, we can imagine that there were actually two identical urns (Urn B1 and Urn B2) each of which contained 2 white and 3 black balls and had a selection probability of 1/3 (see Figure 7.2). So, between these three urns (all equally likely to be chosen) there would be 8 white balls in total (4 in Urn A, 2 in Urn B1 and 2 in Urn B2, all equally likely to be drawn. Because 4 of these 8 were from Urn A, the posterior probability that it was drawn from Urn A must now be 1/2 (see Figure 7.2). By the same logic, if the ball drawn had been black, the probability that the source was A would turn out to be only 1/7.

Bayes's rule is a general formula for computing unknown *conditional* probabilities. In our example, we could obtain the probability that the

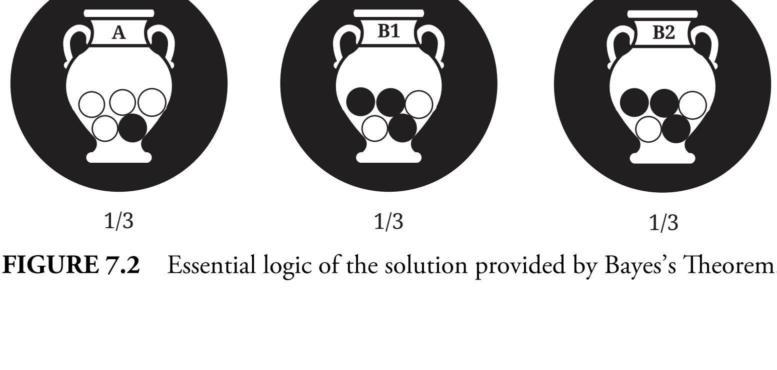


FIGURE 7.2 Essential logic of the solution provided by Bayes's Theorem.

source was Urn A (or Urn B) conditional on observing a white ball in a single drawing. The standard notation for these conditional probabilities would be $P(A|W)$ and $P(B|W)$, where the event W is the drawing of a white ball. We can think of these conditional probabilities as the "inverse" of the "direct" probabilities that are usually known, in this case $P(W|A)$ and $P(W|B)$.

In general, we will know the conditional probabilities of some *evidence* given the possible *sources* of the data.[18] We want to invert these to determine the inverse probabilities, the conditional probabilities of various sources *given the evidence*. Bayes's rule provides the link between the direct and inverse probabilities. However, there is a catch. To apply the rules, we must know the prior probabilities of the various sources. In our example these are $P(A)$ and $P(B)$.

As a simple practical example, suppose a doctor is trying to decide whether her patient suffers from a particular type of infection. Her diagnostic procedure can indicate the infection (correctly) with probability 90% if the infection is truly present. That is:

$$P(\text{test positive} \mid \text{infection}) = \frac{9}{10}$$

However, it will also (incorrectly) detect the infection with probability 10% when the patient is free of the infection:

$$P(\text{test positive} \mid \text{no infection}) = \frac{1}{10}$$

Suppose further that the doctor believes that 5% of her patients are actually infected:

$$P(\text{infection}) = \frac{1}{20}$$

If her patient tests positive, what is the probability that the infection is truly present? Applying Bayes's Theorem, we would find:

$$P(\text{infection} \mid \text{test positive}) = \frac{9}{28}$$

A simpler way to understand Bayes's Theorem is in terms of frequencies rather than probabilities. The psychologist Gerd Gigerenzer has shown that our minds have evolved to deal much more easily with information in the form of frequencies than probabilities.[19] This is not surprising since simple counting has been around for millennia, while mathematical probability is a very recent invention.

To understand how we could arrive at the correct answer via Gigerenzer's frequency approach, imagine there exists a hypothetical population of 1000 individuals, of whom 50 (5%) have the infection. The diagnostic procedure will find 45 (90%) of these. On the other hand, the procedure will also (falsely) incriminate 95 of the 950 (10%) disease-free individuals. So, the probability that a positive indication is truly an infection would be:

$$P(\text{infection} \mid \text{positive}) = \frac{45}{45 + 95} = \frac{9}{28}$$

Philosophical Significance

Most students of probability and statistics do not realize that Reverend Bayes derived this mathematical formula not for statistics, but in order to solve a particular philosophical problem of great importance. Here is how he stated this problem:

> *Given* the number of times in which an unknown event has happened and failed: *Required* the chance that the probability of its happening in a single trial lies somewhere between any two degrees of probability that can be named.

Bayes was after big game, nothing less than a general solution to the problem of *induction*—how can we generalize based on past experience? In particular, how can we draw credible inferences about the future based on events that have previously been observed?

His attempted solution was to formulate the problem in terms of conditional probabilities. Given that so many "happenings" and "failings" have been observed, what is the conditional probability that the true probability (on a single trial) is between any two specified limits? In our simple example, there are only two urns, and the proportion of white balls in each is assumed to be known. However, suppose that we know absolutely

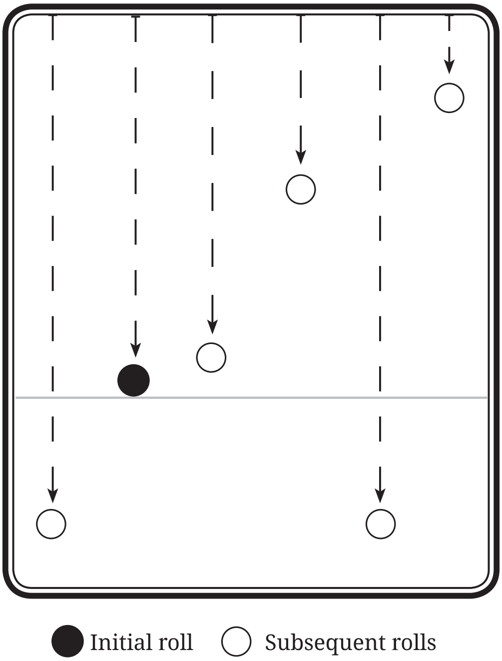


FIGURE 7.3 Bayes's famous "billiard-table" model: The position of the horizontal line marked by the black ball's landing spot is analogous to the unknown probability on each Bernoulli trial. A "success" is deemed to occur each time one of the white balls crosses the horizontal line.

nothing about the urn from which the ball is to be drawn, except that it contains 5 balls, all of which are either black or white. Then there are six possibilities for the number of white balls that the urn might contain (0, 1, 2, 3, 4, and 5). Now suppose that we do not even know how many balls are in the urn. Then, in principle, the proportion of white balls could be any fraction between 0 and 1.

To deal with this general situation, Bayes replaced the conventional urn model with a different physical analogy. He imagined a rectangular table similar to a billiard table (see Figure 7.3). In a thought experiment, Bayes considered how the table could be divided randomly into two parts by rolling a ball and marking the ball's landing spot with an invisible line. The line would determine a fraction of the total length of the table. The black ball in Figure 7.3 represents this initial roll.

This fraction determined by rolling the initial ball represented the unknown probability associated with an underlying process that has generated some observed events. This ball would then be removed after its position was secretly marked. Then another ball, represented here by a white ball, would be rolled again several times. After each such roll, a record would be kept of whether or not it landed beyond the invisible line. Each such trial on which the ball rolled past the invisible line represented an occurrence of the event.

Using this billiard-table model of how the "happenings" and "failings" were being generated, Bayes asked the following question. Suppose that you were told only the number of times that a ball landed beyond the invisible line and the number it did not. In our illustration, there would be two happenings and three failings. What could you infer about the position of the line? In principle, this problem could be solved in the same way as the simple urn problem. However, instead of only two urns, each corresponding to a different prior probability, we would have an infinite number of possible prior probabilities. Each of these probabilities would be represented by one value of the line's location expressed as a proportion of the table's length ranging from 0 to 1.

Let us denote the probability of a happening in a single trial as p. Then, Bayes made the seemingly plausible assumption that essentially all values of p are equally likely. In his physical analogy, the invisible line might have an equal chance of lying anywhere on the table. However, obviously we cannot count all the possible points on a continuum. Bayes got around this difficulty by assuming instead that the prior probability that p will lie in any specified *interval* would be equal to the length of that interval. For example, the probability that p would lie between 0.25 and 0.40 would be 0.15. Statisticians would say that p has a *uniform distribution* over the interval [0,1]. This is the continuous analog of assuming that the two urns in our simple example each have a probability 1/2 of being selected.

Conceptually, this is easy to understand. However, it turns out that the mathematical analysis of this problem is formidable and requires the application of integral calculus. Even so, only a good approximation, not an exact solution, can be derived. Consequently, a practical solution to Bayes's problem would have to await twentieth-century computing power. Bayes and Price made valiant efforts to obtain a workable approximation, using the limited tools at their disposal. The mathematical complexity of the problem may be one reason why his basic insights did not immediately catch on. However, the unconventional *conceptual* framework that Bayes adopted was also very difficult for his readers to grasp.

Unlike de Moivre and Simpson, Bayes chose not to define probability in terms of the more familiar but ambiguous concept of chances. Rather, he interpreted probability in a way that would preclude philosophical problems, as Price explained:

> He has also made an apology for the peculiar definition he has given of the word *chance* or *probability*. His design herein was to cut off all dispute about the meaning of the word, which in common language is used in different senses by persons of different opinions … Instead therefore, of the proper sense of the word *probability*, he has given that which all will allow to be its proper measure in every case where the word is used.[20]

This approach is strikingly modern. Whatever may be our interpretation of the *concept* of probability, a *measure* of this probability must adhere to certain mathematical principles, namely, the usual rules applying to games of chance.

In this way Bayes merged the epistemic sense of probability with de Moivre's mathematical definition as a fraction between 0 and 1. However, he jettisoned the underpinning of "chances" that had previously buttressed the epistemic interpretation of probability. "By *chance* I mean the same as probability," Bayes asserted. What exactly Bayes had in mind by this statement was not explained clearly by him, and is open to speculation. What seems clear is that Bayes's mathematical development required that prior and inverse probabilities be considered in some sense comparable. As a result, it became permissible to think about the probability of a probability! This mind-bending idea would have been (and was) unthinkable for Jacob Bernoulli, or even Abraham de Moivre.

As Price wrote in his introduction, neither of these giants had attempted to *invert* their laws of large numbers to derive a probability statement about the ratio of hidden chances. To apply the idea of probability, as they understood it, to these ratios would require some notion like the chances of chances. Such a recursion would have seemed to them incomprehensible. But by tossing out the baggage of chances, Bayes could begin to think of a probability as something primary. This reification of probabilities allowed him to conceive of *prior* probabilities for these probabilities of interest. Price tells us that Bayes's intention 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. He adds, that he soon perceived that it would not be very difficult to do this, provided some rule could be found, according to which we ought to estimate the chance that the probability for the happening of an event perfectly unknown, should lie between any two named degrees of probability, antecedently to any experiments made about it; and that it appeared to him that the rule must be to suppose the chance the same that it should lie between any two equidifferent degrees; which, if it were allowed, all the rest might be easily calculated in the common method of proceeding in the doctrine of chances.

Statisticians have been wrestling with the consequences of what Price called this "very ingenious solution" ever since.

Particularly controversial has been the assumption of a uniform prior probability distribution for the probability of interest, p. The idea that we can represent complete ignorance about p by assuming all possible values to be equally likely is sometimes called *Bayes's Postulate*, but more often the *principle of insufficient reason* or (following Keynes) the *principle of indifference*. Bayes himself expressed reservations about this seductively simple solution. It had the huge advantage of allowing a computation of inverse probabilities. However, it seemed too facile. How can pure ignorance do so much? This question continues to haunt statistical theory to the present day.

Bayes's essay is modern also in another way that has been little noticed by historians. It may be the first true exposition framed completely in terms of what would become the modern *theory of probability*. Bayes trafficked in probabilities, not only for mathematical convenience, as de Moivre did, but also for the final statement of his results. He did not feel compelled to convert the probability back into odds, as his predecessors did. Nor did he refer to chances in his definition of probability.

Finally, as I mentioned, Bayes and Price were motivated by the problem of induction, which was a hot intellectual topic at that time. As early as 1740, Price had been greatly influenced by the writings of Bishop Joseph Butler (1692–1752).[21] Butler, in his *Analogy of Religion,* *Natural and Revealed* (1736) had famously written: “But for us, probability is the very guide of life.” Here he was using probability in the traditional sense of **probability**, not as a mathematical concept related to games of chance. He was arguing that human rationality can be applied to discover generally reliable, albeit imperfect, knowledge in an uncertain world. This “natural religion” based on observation could supplement the “revealed religion” of scriptures.

Taking issue with this optimism about the strength of inductive inference was the famous philosopher David Hume, who championed a *skeptical* philosophy in his *Treatise on Human Nature* (1739) and *An Enquiry Concerning Human Understanding* (1748).[22] By the 1760s, Price had come to know Hume personally, and he took seriously many of Hume’s insights about the limitations of knowledge gained from experience. However, he was severely critical of Hume’s ideas on some questions. Like Butler and David Hartley, Price was much more sanguine than Hume about mankind’s justification for interpreting observed regularities in nature as causal, and he was less skeptical than Hume about the possibility of occasional miracles.

Hume had dealt with probabilistic reasoning at some length in his consideration of causality and induction. His concept of probability was very similar to the older, more qualitative probability that had motivated Jacob Bernoulli. Like Bernoulli, Hume thought of probability as a degree of justification for and against a proposition. Like Bernoulli, he assumed that both of these probabilities would be considered separately, and then compared, with the preponderance of evidence determining what was rational to accept.

When Richard Price stumbled upon his late friend’s essay (or notes), he was greatly impressed by the potential philosophical implications:

> Every judicious person will be sensible that the problem now mentioned is by no means merely a curious speculation in the doctrine of chances, but necessary to be solved in order to provide a sure foundation for all our reasonings concerning past facts, and what is likely to be hereafter … it is certain that we cannot determine, at least not to any nicety, in what degree repeated experiments confirm a conclusion, without the particular discussion of the beforementioned problem; which therefore, is necessary to be considered by any that would give a clear account of the strength of *analogical* or *inductive reasoning*; concerning which at present, we seem to know little more than that it does sometimes in fact convince us, and at other times not;

In an Appendix, Price goes on to deal with the special situation in which a particular outcome has invariably been observed in *all* prior situations of the same type. Then Bayes's rule, he says, can inform us with precision what range of probabilities it is rational to hold after any number of such observations. As this number grows very large, the range becomes very narrow. However, he sagely cautions:

> What has been said seems sufficient to shew us what conclusions to draw from *uniform* experience. It demonstrates, particularly, that instead of proving that events will *always* happen agreeably to it, there will be always reason against this conclusion. In other words, where the course of nature has been the most constant, we can have only reason to reckon upon a recurrency of events proportioned to the degree of this constancy, but we can have no reason for thinking that there are no causes in nature which will *ever* interfere with the operations of the causes from which this constancy is derived, or no circumstances of the world in which it will fail. And if this is true, supposing our only *data* derived from experience, we shall find additional reason for thinking thus if we apply other principles, or have recourse to such considerations as reason, independently of experience, can suggest.

In the context of Price's debate with Hume about miracles, this passage is quite significant. Bayes had proved, Price believed, that even a very large body of evidence, all pointing in the same direction, cannot rule out the very remote chance of a miraculous occurrence. On the other hand, also contrary to Hume's arguments, there may exist "independently of experience" other reasons to convince us of law-like natural regularities.

Price believed that man's reason could often discern the underlying laws at work. Like Jacob Bernoulli, he regarded empirical frequencies as a default basis for probability estimation when there is "total ignorance of nature." However, inferring causality based on even a relatively small number of consistent observations might sometimes be justified on the basis that "operations of nature are in general regular."

Here Price and Hume disagreed. In essence, Price was emphasizing the potential value of considering all the relevant evidence, not just the historical pattern of events. Price's religious faith fueled his optimism that the laws of nature might often be deduced, even with limited empirical evidence. He believed Bayes had demonstrated that empirical data by itself could never prove the impossibility of a thing, even of a miracle. In summary, Price was optimistic that underlying causes could often be assumed, but that deviations from causal laws could very occasionally occur.

Hume, on the other hand, believed that *all* knowledge of causes must be inferred (imperfectly) from observed regularities. A miracle for him would represent a logical contradiction, since it would be the violation of causal laws, which were nothing more than accepted generalizations that by definition did not allow exceptions. Hume held that an enormous amount of prior empirical regularity (e.g., lead weights always fall, all men eventually die) could establish practical certainty of causation. However, he was extremely skeptical that any report of the miraculous suspension of such a law would ever be credible.

FROM CHANCES TO PROBABILITY

Pierre-Simon Laplace was the pivotal figure in the transformation of the doctrine of chances into the classical theory of probability. Unlike de Moivre and Bayes, he rose meteorically to great heights and his achievements were celebrated in his own time. He became the most famous scientist and mathematician of his era, maintaining this position for decades in spite of extraordinary political and social upheavals.

Before discussing Laplace's career and impact on the evolution of probability, I will outline some relevant currents of thought in and around France that shaped his perspective. As in England, de Moivre's *Doctrine of Chances* was recognized in continental Europe as the authoritative treatment of the mathematical aspects of games of chance. Moreover, the merging of probability as a fraction of chances with epistemic probability was beginning to occur. However, the mid-eighteenth century was a time of transition, and there was still some life left in Bernoulli's idea of quantifying probability as a general logic of uncertainty.

Two well-known mathematicians who, like Bernoulli, hailed originally from Switzerland wrote about the interpretation of probability prior to 1770: Gabriel Cramer (1704–1752) and Johann Heinrich Lambert (1728–1777).[23] Their theories both took as a starting point Jacob Bernoulli's logic of combining arguments. In a series of lectures in 1744 and 1745, Cramer offered a slightly amended version of Jacob's theory. In an article published in 1764, Lambert attempted to generalize and "correct" Bernoulli's mathematical rules for combining arguments. Recall that Jacob's rules for combining arguments had led to some awkward implications, and Lambert proposed an alternative approach to get around this problem.

Neither of these works focused on empirical frequencies or on games of chance. As far as we know, these efforts represent the last serious attempts prior to the twentieth century to frame a formal logic of uncertainty in any terms *other* than those derived from the doctrine of chances. In the 1740s and 1750s, it was still possible to mean by "probability" both a degree of certainty in a general way and the more restricted concept of a fraction of chances in games and a few actuarial applications. This somewhat ambiguous status of the term left open various possible lines of future intellectual development, but also engendered some confusion. It is particularly difficult for us, looking back, to make sense of some writing about probability during this period.

This confusing state of affairs is perhaps best exemplified in a long article entitled "*Probabilité*" (Probability) written by Denis Diderot (1713–1784).[24] Diderot was a well-known member of the *philosophes*, the leading intellectuals of the Enlightenment in France. Starting in 1750, Diderot and Jean le Rond D'Alembert (1717–1783) undertook one of the most ambitious intellectual enterprises of all time. As coeditors, the pair began to publish the *Encyclopédie* (Encyclopedia), a compendium of the world's knowledge that eventually included 28 volumes and over 70,000 articles, contributed by many experts in a wide variety of disciplines. The article on "*Probabilité*" was probably written in the early 1750s. The subheadings listed below the title were Philosophy, Logic, and Mathematics.

Diderot's discussion of probability was indeed a kind of fluid amalgam that seemed to gloss over a number of distinctions that would trouble later philosophers and mathematicians. Lorraine Daston, in her classic study of *Classical Probability in the Enlightenment*, sums up Diderot's presentation:

> The project, if not the results, of Part IV of the *Ars Conjectandi* won converts among philosophers as well as mathematicians, perhaps because the legal and mathematical probabilities of evidence were seldom distinguished. The 'Probabilité' article of the *Encyclopédie* illustrates to what extent the mathematical and legal senses of probability were still intertwined … Although the author was thoroughly conversant with Jakob Bernoulli's *Ars Conjectandi*, the article is more philosophical than technical. The article borrowed heavily from Part IV of the *Ars Conjectandi*, reproducing Bernoulli's general rules for weighing evidence and conjecturing, and interpreting Bernoulli's theorem as a demonstration that 'past experience is a principle of probability for the future; we have reason to expect events similar to those we have seen happen.'[25]

Diderot may seem to us somewhat incoherent as he attempts to explain the concept of probability in a way that does justice to both its older and newer formulations.

By the 1760s, the ambiguity reflected in Diderot's article had been largely resolved by a nearly complete fusing of the epistemic and mathematical connotations of probability. The archaic **probability** that Jacob Bernoulli had tried to mathematicize was receding from consciousness. Increasingly, the "logic" of uncertain inference (including induction from observations of nature) was conceived as obeying the same mathematical principles that governed games of chance.

The narrowing and sharpening of probability represented a double-edged sword. On the one hand, philosophical analysis became more limited in scope. On the other hand, the new probability concept paved the way for great progress in collecting and analyzing statistical information. At the time, the allure of potential benefits from applying probability to various practical problems was great. The *philosophes* were generally hopeful that scientific and social dividends would flow from the rational analysis of nature and society using this new calculus. But a few skeptics were more guarded in their enthusiasm.

One of those who raised red flags was Diderot's coeditor D'Alembert, an eminent mathematician, physicist, and philosopher. As scientific editor for the *Encyclopédie*, he authored two articles on games of chance that were published in the 1750s, and he may also have contributed to Diderot's lengthy article on probability. D'Alembert was a contentious sort, well known for engaging in heated controversies with other scientists and mathematicians. With respect to probability he was something of a contrarian. While he recognized the importance of the new probability theory, he expressed various reservations about the extent to which it could directly be applied to practical problems. Some of his ideas, especially in his early writing, appear simply muddled, while others display deep philosophical insight.

Diderot and D'Alembert were dominant figures in the European intellectual landscape, and their ideas would have been well known to the young Laplace. However, a more immediate source of inspiration for his interest in probability and its potential applications was D'Alembert's younger protégé Nicolas Condorcet (1743–1794), better known as the Marquis de Condorcet. Of noble birth, Condorcet studied under D'Alembert starting in 1759 and eventually achieved great renown as a mathematician, philosopher, and political theorist.

In 1773, Condorcet was appointed acting Secretary of the French Academy of Sciences, and he became the permanent Secretary 3 years later. This venerable institution had been founded in 1666, 4 years after the Royal Society of London. However, unlike its British counterpart, the French Academy had always enjoyed state sponsorship and funding, and had hosted eminent scientists from all over Europe. (Christiaan Huygens had been one of its founding members.) Condorcet's position permitted him to exert great influence over the scientific papers selected for publication in the Academy's journals.

Condorcet's interest in probability dates from the early 1770s, although his famous treatise on the application of probability to voting behavior was not published until 1785.[26] In December 1771, Condorcet received a letter from the virtually unknown Laplace, then 22 years old. It concerned a highly original mathematical paper on esoteric methods motivated by some problems related to the doctrine of chances. Laplace was at this stage seeking to launch a career in mathematics and science, and believed that Condorcet's support would be helpful to him. Condorcet was enthusiastic about possible contributions of the probability calculus to rational decision-making in the social, political, and judicial spheres.

In addition to his philosophical and scientific pursuits, Condorcet was actively involved in political affairs, and he rose to a position of great power and influence. In August 1774, Condorcet's close friend Anne-Robert-Jacques Turgot (1727–1781) was appointed Controller-General of France, a very powerful administrative position. Turgot had Condorcet appointed Inspector General of the Mint, a post he managed to hold for 17 years, even though Turgot fell from power in 1776. So, Condorcet's enchantment with probability was not only academic; he was intimately involved with issues of public policy to which he regarded probability as highly relevant.

It seems probable that Condorcet was the muse who stimulated Laplace to complete the conceptual transition from chances to probabilities that would transform human knowledge:

> There is no question but that the technical development that made the transition was his [Laplace's]. It does seem likely, however, that it was Condorcet whose

> enthusiasm opened the prospect and pointed his talent toward the opportunity and the problems to be found in demography and statistics. Condorcet could not see very far into the problems. But he could see the subject; and he did see it; and he showed it to Laplace.

The French Newton

Pierre-Simon Laplace was born at Beaumont-en-Auge in the Normandy region of France on March 23, 1749.[27] His early life has not been well documented. His father may have been a cider merchant, who envisioned for Pierre-Simon a clerical career. In 1765, Laplace was sent off to the University of Caen to study theology. However, under the influence of two professors there, he soon turned from theology to mathematics. Armed with a letter of introduction, he ventured off to Paris in 1768 to study under D'Alembert, who at age 51 was at the height of his influence. After some initial skepticism, D'Alembert quickly became convinced of Laplace's extraordinary talent and agreed to tutor him. He also arranged for Laplace to obtain a teaching post at the *École Militaire* (Military School), the academy for training young cadets.

Before long, Laplace was cranking out a stream of original papers related to astronomy and mathematics. By 1773 this work, along with the sponsorship of D'Alembert and Condorcet, had secured him an appointment to the Academy of Sciences. As an associate member, he received financial support that allowed him to pursue full-time his various researches. Two of Laplace's earliest research reports concerned the doctrine of chances. One of these, delivered in 1774, was entitled "Memoir on the Probability of the Causes Given the Events." The essence of this work was similar to that of Thomas Bayes's posthumous essay, which was hardly known in England, let alone France. Laplace described his topic as:

> a question which has not been given due consideration before, but which deserves even more to be studied, for it is principally from this point of view that the science of chances can be useful in civil life.[28]

It is not clear whether Condorcet had in fact suggested this subject of inquiry, but he enthusiastically endorsed Laplace's contribution.

In an introduction to the published article, Condorcet wrote that the paper[29]:

> treats a branch of the analysis of chance, much more important and less known than that which forms the subject of the former Memoir; here the probability is unknown, that is to say the number of chances for or against a possible event is undetermined. It is known only that in a given number of experiments this event occurred a certain number of times, and it is required to know how from that information alone the probability of what is going to happen in the future can be stated. It is obvious that this question comprises all the applications that can be made of the doctrine of chance to the uses of ordinary life, and of that whole science, it is the only useful part, the only one worthy of the serious attention of Philosophers.

It is clear that both Condorcet and Laplace had somehow arrived at a "Bayesian" perspective about the importance of inverse probability and the method by which it could be calculated. Elaborating the mathematical and philosophical implications of this approach would become a leitmotif of Laplace's research for the rest of his long career.

Over the next 20 years, Laplace rose to a position of dominance in the French Academy of Sciences, and established his reputation in many scientific areas, especially mathematics and celestial mechanics. Laplace was very ambitious and not particularly modest about his abilities, often causing friction with some of his contemporaries. However, his intellectual superiority, coupled with political agility, stood him in good stead. In 1784, he took on an additional responsibility as an examiner at the Royal Artillery Corps, which brought him into contact with powerful political leaders in France. In this capacity in 1785, he first encountered a promising cadet by the name of Napoleon Bonaparte (1769–1821).

In 1787, Laplace began a fruitful collaboration with another of Europe's greatest mathematicians, Joseph Lagrange (1736–1813), who had recently arrived in Paris. The two geniuses shared many areas of interest and complemented each other well. While Lagrange's forte was mathematical theory, Laplace's mathematical brilliance was deployed mainly in order to solve practical problems.

The following year, at the age of 39, Laplace married Marie-Charlotte de Courty de Romanges, only 18 years old. Soon after, a son and daughter arrived. After a few quiet years, the Laplaces were caught up in the political turmoil engulfing France. They moved to a small town 30 miles outside Paris to evade the Reign of Terror in 1793, returning when it was safer in July 1794.

In 1795, the Academy of Sciences was reopened as the National Institute of Sciences and Laplace resumed his former responsibilities. Also in 1795, the Bureau of Longitudes was established, with Laplace and Lagrange as the only two mathematicians among its founders. This office was established in an attempt to revive French naval power in the wake of Britain's recent ascendance to mastery of the high seas. The Bureau was placed in charge of all institutions concerned with astronomical, geodesic, and related studies, including the Paris Observatory, of which Laplace became director. In 1796, he published the famous "nebular theory," an account of how he believed the solar system had formed from a cloud of rotating and cooling incandescent gas. His astronomical work led to an appointment as one of the 40 "immortals" of the *Académie Francaise*.

When Napoleon rose to power as First Consul of France in 1799, Laplace finagled an appointment as Minister of the Interior, but was replaced after only 6 weeks on the job. Like many master technocrats, Laplace lacked the pragmatic attitude required of a successful administrator. Bonaparte wryly observed about this episode:

> Geometrician of the first rank, Laplace was not long in showing himself a worse than average administrator; since his first actions in office we recognized our mistake. Laplace did not consider any question from the right angle: he sought subtleties everywhere, only conceived problems, and finally carried the spirit of 'infinitesimals' into the administration.[30]

Nonetheless, Napoleon appreciated Laplace's talents enough to appoint him a member, and later Chancellor, of the French Senate. Laplace received the Legion of Honor award in 1805. In that year he also helped to found the Society of Arcueil, an organization whose goal was to promote a more mathematical approach to science. As the leader of this group, he exerted great influence on the administration of scientific research and education throughout the country.

In 1812, Laplace published the first edition of his masterwork on probability, the *Theorie Analytique des Probabilités* (*Analytic Theory of*

Probabilities).[31] It contained a comprehensive overview of the mathematical theory of probability, including many of Laplace's groundbreaking original contributions. Most notably, the *TAP* included a full exposition of a result he had first presented 2 years earlier. Recall that in 1733, de Moivre had shown that the probability distribution for the total number of successes in any large number of Bernoulli trials could be closely approximated using the normal (bell-shaped) curve. I mentioned that this was the earliest version of what would much later be called the Central Limit Theorem.

Laplace went much further by generalizing this theorem. He was able to demonstrate that for virtually *any* random variables of practical importance, the probability distribution of the sum (or the average) of the values of the variables could be approximated by using the normal distribution. This mathematical *tour de force* was derived in the context of a theory of observational errors, which at that time was the main scientific application of probability theory. While this general version of the Central Limit Theorem is Laplace's most notable mathematical discovery related to probability theory, the *TAP* covered a vast terrain, well beyond anything that had gone before.

A second, and substantially expanded, edition of the *TAP* followed in 1814, with two further editions in 1820 and 1825. The Introduction to the 1814 edition was published separately as a *Philosophical Essay on Probabilities*.[32] This was a comprehensive nontechnical discussion of Laplace's conception of how the calculus of probability could be applied to a great many practical problems. Laplace regarded probability theory as not only a practical tool and mathematical discipline, but as the very essence of logical thinking about uncertain matters. In spirit he was very close to Jacob Bernoulli, whom he admired, except that by probability he meant specifically the new probability that conformed to the doctrine of chances. On the last page of his *Philosophical Essay* is his much-quoted remark that "the theory of probability is at bottom only common sense reduced to calculus."

Laplace was named a Count of the Empire 1806, and managed to weather Napoleon's inconvenient downfall in 1815. Indeed, he became the Marquis de Laplace in 1817 after the monarchy's restoration. Once again, he proved himself too useful to ignore and pliable enough to adapt to the political exigencies. He continued to be productive, although his reputation was somewhat tarnished by his blatant political opportunism over the years. His later years were also marred by the loss of his only daughter Sophie-Suzanne, who had died in childbirth in 1813. When the final edition of *TAP* was published in 1825, Laplace was 76 years old and fading. He died 2 years later.

Laplace's Philosophy of Probability

To the Marquis de Laplace many things were perfectly clear that to "lesser" minds were far from obvious. In his hands, the theory of probability became a powerful instrument. His dazzling mathematics, acute logic, and sublime self-assurance swept aside any objections to his ideas about probability. Before Laplace, probability was still somewhat ambiguous, and the mathematical doctrine of chances had not fully fused with the epistemic sense of probability. After Laplace, the merger was complete, a development that would have enormous repercussions, both practical and philosophical, over the next two centuries.

To understand Laplace's mature perspective we can turn to *A Philosophical Essay on Probabilities*. On the opening page he set out the broad scope of probability:

> I present here without the aid of analysis the principles and general results of this theory, applying them to the most important questions of life, which are indeed for the most part only problems of probability. Strictly speaking it may even be said that nearly all our knowledge is problematical; … the principal means for ascertaining truth—induction and analogy—are based on probabilities; so that the entire system of human knowledge is connected with the theory set forth in this essay.[33]

Jacob Bernoulli might have been gratified that his program of extending probability to the moral and social sciences was finally being taken seriously. However, upon reading further, he would perhaps have been chagrined at the ease with which Laplace simply *assumed* that the laws governing games of chance were applicable to "the entire system of human knowledge." Moreover, Laplace's theories reflected an extreme of Enlightenment thinking that dispensed completely with a Divine role in the administration of natural law. When asked about this absence by Napoleon, Laplace is said to have replied: “I have no need for that hypothesis.”

Laplace next stated his conception of a deterministic universe, arguing that blind chance, or fortune, is only “an illusion of the mind.” He subscribed to *the principle of sufficient reason*, according to which “a thing cannot occur without a cause which produces it.” Laplace interpreted this philosophical principle in its strongest form:

> We ought then to regard the current state of the universe as the effect of its anterior state and as the cause of the one which is to follow. Given for one instant an intelligence which could comprehend all the forces by which nature is animated and the respective situations of the beings who compose it—an intelligence sufficiently vast to submit these data to analysis—it would embrace in the same formula the movements of the greatest bodies of the universe and those of the lightest atom; for it, nothing would be uncertain, and the future, as the past, would be present to its eyes.[34]

This clockwork universe (without a clockmaker needed to maintain it) is the logical extension of Newton’s discoveries.

Jacob Bernoulli had deliberately sidestepped the problem of explaining how chance could be reconciled with God’s omniscience and omnipotence. For him it was enough to know that events unfolded *as if* they were random. De Moivre had avoided most metaphysical problems by dealing only with games of chance and actuarial problems. Laplace lived during the Enlightenment, when it was possible to aver that causal laws must be universal, that in principle every sparrow’s fall could be ascribed to potentially knowable causes.

It is only our ignorance of these laws that creates an appearance of randomness. “Probability is relative in part to this ignorance, in part to our knowledge.” Our knowledge is important because it frames for us the different chances, or *possibilities*, that might exist. “We know that of three or a greater number of events a single one ought to occur, but nothing induces us to believe that one of them will occur rather than the others.” Knowing the possibilities is useful knowledge, but not sufficient to achieve certainty. “In this state of indecision it is impossible for us to announce their occurrence with certainty.” Thus, probability is necessary to balance our knowledge and ignorance appropriately.

Laplace then launched into an exposition of the mathematical principles of probability. While he interpreted probability epistemically, as a measure of belief, he assumed (without explanation) that the usual rules of mathematical probability theory must apply:

> The theory of chance consists in reducing all the events of the same kind to a certain number of cases equally possible, that is to say, to such as *we may be equally undecided about in regard to their existence,* and in determining the number of cases favorable to the event whose probability is sought. The ratio of this number to that of all the cases possible is the measure of this probability, which is thus simply a fraction whose numerator is the number of favorable cases and whose denominator is the number of all the cases possible (emphasis added).[35]

Laplace slid back and forth between examples involving chance mechanisms (coins, urns, or dice) and practical applications of the theory (annuities, testimony, astronomy). However, he glossed over the question of exactly how the possible cases underlying the more complex natural and social phenomena resembled those in the games of chance.

This ambiguity did not seem to bother Laplace, but his failure to clarify this important issue would create many problems for others. There is a long leap between the "chances" in a lottery and those in a complex social process. I believe that Laplace's thinking may have been grounded in a conceptual framework in which metaphysical chances were meaningful, so that his readers would not be troubled by this ambiguity. Perhaps the chances for him were "obviously" metaphorical, since he believed the underlying unknown causes to be in fact strictly deterministic. So, probability assessments might vary, but mainly because different individuals would possess different *information* regarding the underlying causes: "In things which are only probable the difference of the data, which each man has in regard to them, is one of the principal causes of the diversity of opinions which prevail in regard to the same objects."

He then criticized irrational beliefs such as magic and astrology that emanate from blind superstition rather than observed data. Laplace seemed to imply that probability is based on a rational appraisal of chances that accounts for the available information pertaining to the hidden causes at work. In this way, the problem of induction is, in principle, reducible to a matter of deductive logic. However, he did not specify the precise manner in which this logic is to be implemented:

> The difference of opinions depends, however, upon the manner in which the influence of known data is determined. The theory of probabilities holds to considerations so delicate that it is not surprising that with the same data two persons arrive at different results, especially in very complicated questions.[36]

Fair enough. But what principles can guide us in our deliberations regarding probabilities implied by the available data? In particular, how can we assess whether the cases (chances) we have enumerated are really equally possible? "If they are not so, we will determine first their respective possibilities, whose exact appreciation is one of the most delicate points of the theory of chance."

He then tried to explain this point with a simple coin-tossing scenario. Suppose that we are allowed two tosses of a fair coin. If heads arises on the first toss, we win. If not, then we get to toss again, and win if a head occurs on this second throw. What is the probability of winning the game? In this situation, we might delineate three possible cases: HH, TH, and TT. Since two cases would result in success, we *might* believe the probability of winning to be 2/3. (Indeed, D'Alembert was initially confused about this before realizing the fallacy.) However, we all "know" that these cases are not equally likely. As Fermat had reasoned in 1654, we can imagine there are actually four chances that would be equally possible if we were to toss both coins, regardless of the outcome on the first toss. Three of these possible outcomes (HH, HT, and TH) are winners and one (TT) results in a loss, so the "correct" answer is 3/4.

This knowledge of the right answer derives from our understanding of the physical setup and resulting causal process in this simple case. Laplace used this simple example as a means of calling attention to the difficulty in practical problems of determining the equally likely chances. However, he did not offer much practical guidance when it came to situations in which the presumed underlying causal mechanism was not known. This "most delicate" problem had troubled D'Alembert deeply, and he remained somewhat skeptical about the ease with which mathematical laws of chance could be applied to the physical or social world. Laplace was mindful of his old mentor's philosophical reservations, but far less worried about their practical importance.

The Probability of Causes

By the 1750s the *fraction* of chances had become a natural way of measuring the probability of an event. This chance-based probability, which represented the appropriate degree of belief for events in games of chance, gradually came to be accepted as the sole way of thinking about epistemic probability more generally. Thus, mathematical probability came to assume the dual connotation of both a degree of certainty and a reflection of underlying causal structure in the real world, as perceived imperfectly through the veil of human ignorance.

Bernoulli's law of large numbers and de Moivre's improved approximation proved that the hidden ratio of chances could be discovered if the number of observations grew very large. Even a more modest number of observations would provide *some* indication of this ratio. However, neither Bernoulli nor de Moivre was able to imagine how to quantify the degree of certainty associated with an "estimate" based on a small sample. They were both wedded to the idea that practical applications of the doctrine of chances would require large enough samples to assure something like "moral certainty."

Bayes thought it might be possible to derive a precise inference about the underlying probability based on a much smaller sample. He devised a clever way of obtaining a probability statement about an underlying probability (fraction of chances). In order to do this, however, he had to adopt what Price called his "peculiar definition" of probability as meaning the same as chance. Then he could utilize the standard definition of conditional probability to obtain the desired inverse probability.

Bayes never referred explicitly to causes in his essay. However, in a "Scholium" Bayes did consider some broader philosophical implications of his analysis for the issue of causation. His work on probability was motivated by his interest in causal regularities. How much could be known about the chances for a thing to occur based only on the frequency of past occurrence? Moreover, Richard Price certainly had the philosophical implications in mind.

Laplace, on the other hand, was more confident and ambitious. He was explicitly concerned with the probability of causes very generally. He had no intention of sidestepping philosophical issues by simply treating chance and probability as identical, whatever they might really mean. So, Laplace could not find the "probability of causes" by straightforwardly applying de Moivre's definition of conditional probabilities. He needed to expand the theory of probability to encompass causal processes. This he accomplished by positing a new principle that he believed to be self-evident.

To understand Laplace's idea, suppose there are three possible causes: C_1, C_2, and C_3. The posterior probability of C_1 given the event E would be:

$$P(C_1|E) = KP(E|C_1)$$

That is, the posterior (inverse) probability of any cause such as C_1 given the event E would be proportional to the corresponding direct probability of the event given the cause. Here the proportionality constant K is necessary to make the sum of the inverse probabilities for all the possible causes add to 1. So,

$$K = \frac{1}{P(E|C_1) + P(E|C_2) + P(E|C_3)}$$

This new principle can be regarded as the special case of Bayes's Theorem when the three causes have equal prior probabilities (1/3 in this case). Laplace's rule produced the same mathematical result, but his logic was reversed. It seemed self-evident to Laplace that this fundamental principle should normally be assumed as the default, absent a reason to depart from it. That "unusual" case would happen when the various causes were known to be "unequally probable." In that special case, Laplace would weight the different possible causes in accordance with their unequal prior probabilities of happening.

Mathematically the resulting formula was identical to Bayes's Theorem, but the relevance to causation was more direct and "obvious" within Laplace's philosophical framework. Indeed, from his earliest paper on the probability of causes in 1774, the identification of "probabilities" with "causes" was evident:

> The uncertainty of human knowledge is concerned with events or with causes of events. If one is assured, for example, that an urn only contains white and black tickets in a given ratio, and one asks the probability that a ticket drawn by chance will be white, then the event is uncertain but the cause upon which the probability of its occurrence depends, the ratio of white to black tickets, is known.[37]

Bernoulli had regarded the ratio of chances as a reflection of some hidden causal process. Bayes preferred to evade the question of what exactly was meant by a probability. Laplace wrote as if the probability could be treated as somehow *equivalent* to a cause.

Insufficient Reason

Bayes explained in his Scholium the rationale for his proposal to treat all prior probabilities as equally probable. This idea later became known as the *principle of insufficient reason*. According to this principle, if we know absolutely nothing a priori about the probability p of a particular event, we should act *as if* the value of p has been determined, as in Bayes's analog, by metaphorically rolling a ball on a smooth surface and observing where it lands.

Bayes had offered a sophisticated mathematical rationale for why it *may* be reasonable to represent complete a priori ignorance by equal prior probabilities.[38] To him, it was not obvious that ignorance meant equal probabilities. He was not certain that his "billiard table" analogy could be fully justified. Laplace suffered no such compunctions. Rather than trying, like Bayes, to derive a solution using only the existing elements of the doctrine of chances, Laplace introduced a new principle, or axiom, into the theory of probability.

As I explained above, Laplace's new principle was framed in terms of causes rather than probabilities. However, he regarded causes and probabilities as effectively interchangeable. Therefore, it was quite natural for him to treat probabilities mathematically just like causes. If the probabilities associated with the unknown causes were unknown, then there would be no sufficient reason to regard these probabilities as being unequal. So, the principle of insufficient reason should, he thought, be applied to any unknown probabilities: “When the probability of an event is unknown we may suppose it equal to any value from zero to unity.” Here, however, Laplace’s failure to define precisely what the underlying cases, or chances, really represent would run into difficulties.

To avoid confusion, it is important to realize first that Bayes’s Theorem as a mathematical formula is almost trivial to derive from basic axioms of mathematical probability. Contrary to much misunderstanding, there has never been any dispute about the formula itself.[39] Moreover, there are many practical situations in which its use is entirely uncontroversial. Consider the problem of medical diagnosis discussed previously. Often the “base-rate” at which a disease is prevalent in the population is fairly well known. For example, the probability that a randomly selected individual has the disease might be 0.05, or 5%. Furthermore the respective conditional probabilities of detecting the disease given that it is or is not present may also be known.

When these prior probabilities are known, Bayes’s Theorem tells us how to put these pieces of information together to obtain the posterior (inverse) probability that a patient has the disease given that she tests positive. The difficulties arise when the prior probabilities are unknown or subject to disagreement. In particular, what is reasonable when we lack *any* information upon which to specify the prior probabilities? It is this typical situation that has generated the most heated controversies pertaining to so-called Bayesian analyses. So, the papers by Bayes in 1763 and Laplace in 1774 were more important for the philosophical Pandora’s Box they opened up than for their mathematical content.

The principle of insufficient reason seems plausible if we are dealing with a well-defined set of concrete possible events, such as the possible outcomes in a game of chance. In the absence of any evidence to the contrary, each outcome can be assumed equally possible. This is certainly a reasonable position to adopt, at least as a practical expedient. For instance, if I come to a fork in a trail, and have no idea which way to go, I may as well just flip a coin to decide. If the alternatives are specific hypotheses regarding possible causes, then we can in principle apply the same logic. We can list the possibilities and decide whether there are any relevant distinctions among them that would lead us to expect one rather than another.

However, when the assumed alternatives are mathematical probabilities, it is not clear that the same logic necessarily pertains. Probabilities *in the abstract* are simply numbers that have no context or meaning. The only way to judge the relative likelihood of occurrence for any such number would be to translate it somehow into a concrete cause, or causal processes. So, our uncertainty about probabilities regarded as naked numerical values entails much more ambiguity than our uncertainty about causes. As a result, applying the principle of insufficient reason directly to probabilities rests on a much weaker rationale than applying it to causes.

The principle of insufficient reason applied to probabilities assumes there is a one-to-one correspondence between causes and probabilities. However, since the nature of these underlying causes is completely unknown, there is no way to decide rationally whether this makes sense. This basic conundrum can be illustrated by a simple example. Suppose we know that there exist three possible causes that might have given rise to an event: C_1, C_2, C_3.

Assume first that the probabilities for the event given these three causes, respectively, are 1/8, 1/4, and 1/2. Then, by the principle of insufficient reason, our prior probabilities must be

$$P(C_1) = P(C_2) = P(C_3) = 1/3$$

Equivalently, we would have

$$P(1/8) = P(1/4) = P(1/2) = 1/3$$

However, suppose that the probability of the event given C_1 is actually 1/4 rather than 1/8. Then, we would have

$$P(1/4) = 2/3$$

$$P(1/2) = 1/3$$

Although the three possible *causes* would each have the same prior probability of 1/3, the two possible probability values would have unequal prior probabilities. So, applying the principle of insufficient reason to the two possible probability values would be arbitrary and misleading. We have no way to know whether or not it is reasonable to assume there is a one-to-one correspondence between possible probabilities and possible causes.

The same issue arises when we consider a continuum of possible probability values, as in Bayes's framework. It may seem natural to apply the principle of insufficient reason to the values of probabilities. However, we could equally well apply the principle to any other mathematical function of the probability, such as the logarithm. Indeed, we could apply the principle of insufficient reason to the *odds* rather than the probabilities. After all, probability has only recently supplanted the much older concept of odds as a standard way to represent uncertainty. We must remember that a mathematical probability is only one possible way to measure our uncertainty about a specified event. Applying the principle of insufficient reason to this particular measure can have different consequences than applying it to some other valid measure.

A Coincidence?

Condorcet referred to Bayes's essay in an introduction to one of Laplace's memoirs published in 1781. It is generally believed that Laplace must have learned about Bayes's essay around that time and was therefore unaware of it when he wrote his landmark paper on the probabilities of causes in 1774. In his *Philosophical Essay* Laplace mentioned that he admired Bayes's analysis, but found the approach a bit peculiar. Laplace quizzically remarked that Bayes found "the probability that the possibilities indicated by past experience are comprised within given limits; and he has arrived at this in a refined and very ingenious manner, although a little perplexing."[40]

As I mentioned previously, Bayes's essay was believed to have exerted no immediate influence on later developments in probability and statistics. Laplace's work on inverse probability a decade later, which *did* have great influence, was directly stimulated by Condorcet, who had a strong interest in the application of probabilistic ideas to social and political affairs. But here we have a mystery. Was Laplace truly unaware of Bayes's ideas when he conceived his own version of inverse probability? Is it possible that some hint of what Bayes had accomplished had somehow made its way across the Channel and stimulated Laplace's thinking?

Before we dismiss this speculation as entirely fanciful, let us consider that there was a substantial commerce of ideas between the leading intellectuals of England and France. It therefore seems plausible that Condorcet could have read David Hartley's *Observations on Man*. Recall that Hartley had written of his ingenious friend, probably Bayes, who had solved the problem of inverse probability. Hartley had suggested that this result might allow us to "determine the proportions, and by degrees, the whole nature, of unknown causes, by a sufficient observation of their effects."[41] Indeed, there seems to be an uncanny similarity between Laplace's formulation in 1774 and Hartley's discussion in 1749.[42] So, one possibility is that Condorcet had alluded to Hartley's "Bayesian" ideas when he was first stimulating the young Laplace's interest in the probability of causes.

Another possible agent for the transmission of Bayesian thinking across the Channel was the ubiquitous Richard Price. Price's philosophical writings were well known on the Continent. In 1767, Price had published a treatise entitled *Four Dissertations*, which included a discussion of the problem of induction in response to Hume's famous argument against belief in miracles.[43] In a very lengthy footnote, he had referred to Bayes's essay, described the general idea, and provided some examples. Furthermore, we know that Price met the Abbé Morellet (1727–1819), who had visited England in the summer of 1772, right around the time that Laplace began thinking seriously about probability.[44,45]

Morellet was one of France's most prominent *philosophes* and a close friend of Turgot, whom Condorcet had met around this same time. Morellet had written about political economy and was interested in Price's views about the population question. At a dinner in London attended by Price, along with Benjamin Franklin, Joseph Priestley, and other luminaries, Morellet found the conversation "instructive and varied." It is certainly possible that Price discussed his views on Hume and the potential importance he ascribed to the Bayes essay. If so, the general idea of inverse probability may have been conveyed to Condorcet by Morellet, who had returned to Paris in the fall of 1772. I hope that further scholarship will eventually reveal whether some link between Bayes and Laplace actually may have existed.