Chance, determinism and the classical theory of probability

Anubav Vasudevan

University of Chicago, United States

A R T I C L E   I N F O

Article history:
Received 30 June 2017
Received in revised form
13 November 2017
Available online 12 January 2018

Keywords:
Chance
Determinism
Probability
Principle of sufficient reason

A B S T R A C T

This paper situates the metaphysical antinomy between chance and determinism in the historical context of some of the earliest developments in the mathematical theory of probability. Since Hacking’s seminal work on the subject, it has been a widely held view that the classical theorists of probability were guilty of an unwitting equivocation between a subjective, or epistemic, interpretation of probability, on the one hand, and an objective, or statistical, interpretation, on the other. While there is some truth to this account, I argue that the tension at the heart of the classical theory of probability is not best understood in terms of the duality between subjective and objective interpretations of probability. Rather, the apparent paradox of chance and determinism, when viewed through the lens of the classical theory of probability, manifests itself in a much deeper ambivalence on the part of the classical probabilists as to the rational commensurability of causal and probabilistic reasoning.

1. Introduction

One of the oldest and most enduring metaphysical antinomies in the history of philosophy is that which exists between chance and determinism. How, on the one hand, can we admit the possibility that we live in a deterministic world in which every event has a prior necessitating cause and, at the same time, acknowledge the manifest fact that certain events happen by chance? The aim of this paper is to situate this metaphysical question in the historical context of some of the earliest developments in the mathematical theory of probability.

The mathematical rudiments of probability were laid in the second half of the seventeenth and eighteenth centuries, making the theory of probability, in its classical form, a product of Enlightenment thought.1 Accordingly, most of the classical theorists of probability, in keeping with the rational optimism of their day, were self-avowed determinists, accepting that the explanatory achievements of Kepler and Newton in the domain of celestial mechanics could in principle be extended so as to encompass all natural phenomena.2 There is no clearer statement of this unqualified determinism than that which appears in the following passage from Laplace:

We ought then to consider the present state of the universe as the effect of its previous state and as the cause of that which is to follow. An intelligence that, at a given instant, could comprehend all the forces by which nature is animated and the respective situation of the beings that make it up, if moreover it were vast enough to submit these data to analysis, would encompass in the same formula the movements of the greatest bodies of the universe and those of the lightest atom. For such an intelligence nothing would be uncertain, and the future, like the past, would be open to its eyes. (Laplace, Essai philosophique sur les Probabilités, L 7 vi–vii)

This now famous depiction of a causally deterministic universe is part of a more extended panegyric in which Laplace heaps praise upon the sciences for having, at last, revealed the appearance of chance in nature to be an illusion. It is the progress of science towards the ultimate aim of expelling the ‘blind chance of the

---

1 In what follows, the ‘classical theory of probability’ signifies the assortment of historical approaches to the subject which place an emphasis on combinatorial methods of reasoning, or accept, in some form, the classical definition of probability as the ratio of favorable to total possible outcomes, all such outcomes being equally possible. In accordance with this usage, classical theorists of probability include Pascal, Fermat, Huygens, Leibniz, Jacob Bernoulli, Montmort, de Moivre, and Laplace. The historical developments recounted by Todhunter (1865) and Hald (1990), with the exception of the latter’s exclusion of Laplace, can be roughly identified with the history of the classical theory of probability.

2 For representative statements of determinism in the writings of the classical theorists of probability, see B 239; Montmort (1713, xiv); and de Moivre (1756, 253). For more on the determinism of the classical theorists of probability, see Hacking (1990, ch.2); Daston (1995, 34–37); and Gigerenzer et al. (1990, 11–13).
Epicureans’ from the natural world which, for Laplace, ‘distinguishes nations and ages, and constitutes their real glory.’\(^3\)

Such rationalistic sentiments, while perhaps expressed in hyperbolic terms, are entirely in keeping with the intellectual orthodoxy of Laplace’s day.\(^4\) What gives to them an air of paradox, however, is the particular textual context in which they are made. For the above passage appears in the introduction to Laplace’s *Essai philosophique sur les Probabilités*. Thus, following his unqualified rejection of even the slightest indeterminacy in the natural causes of things, Laplace proceeds to demonstrate, by means of a protracted series of calculations, how to compute with numerical exactitude the probabilities of various chance events. No justification is offered for this stark juxtaposition in the text other than the remark, made almost in passing, that ‘probability is relative in part to [our] ignorance and in part to our knowledge.’\(^5\)

Laplace and the other classical theorists of probability thus personified, in a particularly striking way, the antinomy between chance and determinism. On the one hand, in their capacity as Enlightenment thinkers they regarded chance as a fiction, ‘a mere word,’ its apparent reality reflecting only our incapacity to see all the way through to the underlying causes of things.\(^6\) At the same time, in laying the groundwork for the mathematical theory of probability, they were busy conducting the first ever systematic investigation into the mathematical structure of chance phenomena. In light of this obvious tension it might seem inevitable that the classical theorists of probability would have made some effort to clarify the subject matter of their new ‘géométrie du hasard.’ Yet the intuitiveness of the methods they employed and the manifest practical value of the results thus obtained rendered otiose such philosophical subtleties. On the few occasions when such questions were addressed, the classical probabilists tended to prevaricate, offering a few brief qualifications to avoid any overt contradiction before returning to their calculations.\(^7\)

This lack of clarity has led to the charge of incoherence in the classical conception of probability. Since Hacking’s seminal work on the subject this charge has most often taken the form of an accusation that the classical probabilists were guilty of an unwitting equivocation between a subjective, or epistemic, interpretation of probability, on the one hand, and an objective, or statistical, interpretation, on the other.\(^8\) Thus, for example, Gigerenzer et al. write:

The classical interpretation of mathematical probability was characterized in precept by determinism and therefore by a subjective slant, and in practice by a fluid sense of probability that conflated subjective belief and objective frequencies with the help of associationist psychology. (Gigerenzer et al., 1990, 13)

Admittedly, there is some merit to this charge, for there does reside a fundamental conceptual tension at the heart of the classical theory of probability. As I hope to show, however, this tension is not best understood in terms of the contrast between subjective and objective interpretations of probability. Rather, the apparent antinomy of chance and determinism reflects a much deeper ambivalence, on the part of the classical probabilists, as to the commensurability of causal and probabilistic reasoning. This ambivalence is most clearly on display in the seemingly contradictory appeals to the principle of sufficient reason that appear in their writings. The challenge of developing a satisfactory account of the structure of rationality that can reconcile such diverse applications of this principle represents the true philosophical legacy of the theory of probability in its classical form.

The plan of the paper is as follows. In Section 2, I provide a brief introduction to the classical theory of probability, emphasizing the role that simple games of chance played in its early development. In Section 3, I examine the grounds of the judgments of ‘equi-possibility’ underlying the combinatorial methods employed by the classical probabilists in the analysis of such games, and show that, ultimately, these judgments rely on an appeal to the principle of sufficient reason. Since it was this very same principle that led the classical theorists of probability to adopt a causally deterministic metaphysics, the antimony between chance and determinism, in this context, takes the form of a paradox of sufficient reason. In Section 4, I argue that the assumptions that are needed in order to generate the paradox cannot be reconciled with the intended applications of the theory. To illustrate this point, I consider Roberval’s objection to Fermat’s combinatorial solution to the problem of points. In Section 5, I offer some brief, concluding remarks on the philosophical challenges raised by the classical theory of probability.

2. Games of chance and the classical theory of probability

On several occasions, Leibniz remarked that logic should concern itself with the study of probability, proposing that the most natural place to begin such an inquiry is with a detailed analysis of games of chance:

I have more than once said that we should have a new kind of logic which would treat of degrees of probability ... Anyone wanting to deal with this question would do well to pursue the investigation of games of chance ... carefully reasoned and with full particulars. This would be of great value ... since the human mind appears to better advantage in games than in more serious pursuits. (Leibniz, *Nouveaux Essais*, A VI.6 466)

The prudence of Leibniz’s advice is confirmed by the fact that it was through the analysis of simple games of chance, such as those which involve the tossing of coins, the casting of dice, or the drawing of cards from a shuffled deck, that the theory of probability underwent its first growth in the direction of a mature and mathematically rigorous science.\(^9\) It is worth considering briefly why this should have been the case. That is, why should so specialized an activity as dice casting have served as the backdrop for

---

3 L 7 viii.
4 As van Strien (2014, 27) reports, similar formulations of determinism in terms of an unbounded intelligence capable of predicting the future with certainty can be found in the works of several of Laplace’s contemporaries, including Maupertuis (1756, 332), Boscovich (1922, 281), Condorcet (1768, 5), and d’Holbach (1820, 51–52).
5 Sheynin (1970, 234–46) regards Laplace’s determinism as a development of the determinism of Newton and as a direct inheritance from Bernoulli and condorcet (cf. Sheynin (1976, 172–42); Hahn (1967); Hacking (1983); argues that the widespread belief in Laplacean determinism began to erode in the nineteenth century.
6 de Moivre (1756, 253).
7 Bernoulli, for example, evades the issue of the apparent conflict between determinism and contingency as follows: ‘In themselves and objectively, all things under the sun, which are, were, or will be, always have the highest certainty ... Others may dispute how this certainty of future occurrences may coexist with the contingency ... of secondary causes; we do not wish to deal with matters extraneous to our goal.’ (B 239).
9 This, of course, is not to deny that the concept of probability has a long history that predates its first mathematical treatment in the seventeenth century (see, e.g., Maistrov (2014) and Franklin (2015)).
mankind’s first systematic efforts to assess the probabilities of events.10

One answer to this question is suggested by the following passage taken from a recent textbook in econometrics:

Historically, probabilities were first proposed as a way to understand the differences noticed empirically between the likely occurrence of different betting outcomes ... Soldiers during the medieval times could attest to the differences in the empirical relative frequencies of different events related to the outcomes in [a game in which two dice are cast] ... After thousands of trials they knew intuitively that the number 7 occurs more often than any other number and that 6 occurs less often than 7 but more often than 5. This intuition was developed into something more systematic that eventually led to probability theory. (Spanos, 1999, 4–5)

As a proposed explanation of the role that simple games of chance played in the early history of probability, this account is lacking in several respects. For one thing, there is little historical evidence to suggest an emerging statistical science of dice-falls in the Middle Ages.11 Moreover, even postulating such a science may well overestimate the cognitive capacity of medieval soldiers to intuitively discern long-run statistical regularities. When two fair dice are cast repeatedly, on average, the sum of their scores will equal 7 on 6 out of every 36 casts (16.67%); it will equal 6 on 5 out of every 36 casts (13.89%). Without the aid of carefully kept records, regardless of how many times the game is repeated, it is doubtful that soldiers would ever come to ‘notice empirically’ such a minute difference in relative frequencies.12

Apart from its historical and psychological implausibility, however, there remains a still more obvious reason why the above account must be deemed unsatisfactory. For in spite of its suggestion that probabilities were first introduced in connection with games of dice, no further explanation is given as to what it is about such games that made them particularly natural candidates for probabilistic scrutiny. If, as suggested, probabilities were first proposed as a way of making sense of the observed frequencies of events in a prolonged sequence of repeated trials, then presumably probabilities could have arisen out of an attempt to explain the statistical behavior of any stable, repeatable chance process. But why then should probabilities have first been introduced in connection with so peculiar a phenomenon as the frequencies with which the various faces of cubic blocks of wood or bone show face-up when they are cast upon the floor. Why, that is, should soldiers in the Middle Ages have devoted their free time to betting on the outcomes of games of dice, as opposed to betting on some other arbitrary chance process?

As a first step towards answering this question, we should take note of the obvious, but sometimes overlooked, fact that chance processes such as the casting of dice are artifacts of human contrivance. They were engineered by human beings with specific purposes in mind and, as such, they generally possess certain features which make them particularly well-suited to serve in the tasks for which they are employed.

Take, for example, the specific case of dice casting. Throughout human history, dice have been cast for one of two reasons. On the one hand, dice were sometimes employed in religious ceremonies as instruments of ‘cleromancy,’ or divination by mechanical means.13 In the course of such a ceremony, a question would be posed by a group of devotees as to the will of the gods in regard to a certain matter, a chance process would be initiated, and the answer to the question would somehow be deduced from the outcome of the process. Thus, for example, a fortune-teller might toss a handful of marked bones onto the ground in order to assess the favorability of certain omens. Now, clearly, certain bones would be better suited to serve in this capacity than others. If, for example, one of the sides of a bone was, by far, the most likely to land face-up, and if this were owing to some manifest physical feature of the bone (e.g., if, one of its sides were significantly larger or flatter than the others), then it would strain the limits of credibility, even for the most pious ritual observer, to admit that the outcome of the cast was truly revelatory of the will of the gods. Thus, the very conditions for the credibility of an instrument of cleromancy would have naturally led to the favoring of more regular, or symmetrical, devices.

Still, if dice were only ever employed as instruments of cleromancy, there would be no real need for them to be perfectly symmetrical since it would be enough for this purpose to ensure that the outcome of the cast should not depend in any obvious way on factors wholly irrelevant to the matter to be divined.14 The real need for symmetrical dice arises in connection with the second main purpose for which dice have been employed throughout history, namely, as supplying the random element in games of chance.15 When employed in this capacity, as tools for game-playing, the ideal of a symmetrical die is inherited from that of a ‘fair’ game.16 In their attempts to devise fair conditions of play, the earliest game-designers would have naturally been led to seek out symmetrical dice, since the most direct way to ensure that a game

10 This question is not to be confused with that entertained by Hacking (2006, ch.1). Hacking is not concerned to assess why the earliest calculations of probability should have involved simple games of chance, but rather why, given the long history of gambling in human society, these calculations should have only taken place in the late Renaissance period in Europe. While Hacking’s curiosity on this point is justified, the plausibility of his claim that the modern concept of probability was a seventeenth-century contrivance is open to critique (see Garber and Zabell (1979) and Schneider, (1980)). Howson (1995), for example, argues that the sudden emergence of mathematical probability in the seventeenth century was not due to any conceptual innovation, but instead to the invention of modern mathematical notation, without which even the simplest of combinatorial calculations would have been intractable.

11 Thus, for example, Kendall (1956, 3) observes: ‘It might have been supposed that during the several thousand years of dice playing preceding, say, the year A.D. 1400, some idea of the permanence of statistical ratios and the rudiments of a frequency theory of probability would have appeared. I know of no evidence to suggest that this was so.’

12 Pace David (1962, 89, n. 1). For an equally skeptical attitude towards a statistical account of the origins of probability theory, see van Brakel (1976, 127–130).

13 For more on the use of dice as instruments of cleromancy, see David (1955, 9–10), David (1962, ch.2), and Kendall (1961).

14 This perhaps explains why astragali (hucklebones) and other irregularly shaped objects were still being used in ancient Greek divination rituals long after cubic dice had been introduced in the context of game-playing and gambling (Johnston, 2009, 99).

15 The long history of dice-playing clearly indicates an active concern to seek out symmetrical dice. At early archaeological sites throughout Europe and the Near East, researchers have unearthed numerous specimens of the astragalus, or hucklebone of animals, and it is hypothesized that one of the functions of these bones was to serve as dice for gaming (Gilmour, 1997; Minniti & Peyronel, 2005). As a first hint as to the importance of the symmetry of such devices, it may be noted that nearly all the astragali thus far unearthed have belonged to hoofed animals, such as deer, oxen, or goats. It is in these hoofed animals that the astragalus bone possesses the most nearly symmetrical shape (David, 1962, 2–3). Still more suggestively, there are many examples of astragali that have been worked and polished to give them a more cubic shape (Gilmour, 1997, 171); cf. Lewis (1988, 765). The oldest known example of a hand-carved six-sided die was recently discovered in the ‘Burnt City’ of Shahr-I Sokhta in modern day Iran, and dates from approximately 2700 BCE (Jarrige, Didier, & Quinson, 2011, 24–5).

16 Sortilege, or the ‘casting of lots’ for deliberative or juridical purposes, may be viewed as an instance of game-playing, provided we allow for high-stakes games in which the players need not be voluntary participants. As with other games of chance, considerations of fairness were paramount in the choice of the chance mechanisms employed in sortilege (Hassemer, 1967, 41).
is fair is to play the game under equal conditions so that 'no difference between the players may be noticed except that consisting in the outcome.' It is for this reason that games involving coins, cards, dice, and the like have played so prominent a role in the history of gambling. For, at least with respect to the simplest of such games, all parties involved could feel confident that no one player held an unfair advantage over any other.

Of course, a game between two players in which the winner is determined by the outcome of a single toss of a coin, while doubtless fair, is not particularly stimulating. Hence, the actual games of chance that have been played throughout history have always involved complications whereby each of the players in a game is assigned not just one but a plurality of distinct outcomes, the occurrence of any one of which would win that player the game. Thus, for example, in a game in which three dice are cast, the winner may be determined not by the scores of the dice considered separately, but rather by the sum of their scores taken together. While this variation in the conditions of play may make for a more stimulating game, it also has the effect of breaking the game's original symmetry, with the result that the modified game can no longer be judged fair. For example, if one player is declared the winner when this sum is equal to 7 and another when this sum is equal to 3, then clearly the first player has an advantage over the second.

Where then do probabilities enter the story? The earliest developments in the theory of probability were the result of attempts to restore conditions of fairness to games of chance played under asymmetrical conditions. This, for example, is the explicit subject of Huygens's 1657 treatise titled De ratiociniis in ludo aleae, one of the earliest published works on probability. Huygens's analysis was based on the principle that, in order to render an asymmetrical game fair, each player should be made to contribute an amount commensurate with that player's 'expectation' of gain, where the value of this expectation is determined in accordance with the following rule:

If I may expect either \( a \) or \( b \) ... and if the number of cases, by which \( a \) falls to me, be \( p \) and the number of cases, by which \( b \) falls, be \( q \), and supposing all the cases do happen with equal facility, then the value of my expectation is \( \frac{p\cdot a + q\cdot b}{p+q} \). (Huygens, De ratiociniis in ludo aleae, B 112)

As Bernoulli later observed, Huygens's formula for the expectation is just a certain weighted average, or 'mixture,' of the two possible returns, \( a \) and \( b \). In subsequent treatises, de Moivre would refer to the weights involved in this average, \( \frac{p}{p+q} \) and \( \frac{q}{p+q} \), as the 'probabilities' of obtaining \( a \) and \( b \), respectively. In this way, we arrive at the classical definition of probability as the ratio whose numerator is the number of possible cases that imply an event, and whose denominator is the total number of possible cases, where it assumed that all the cases 'happen with equal facility.' Historically, it was this ratio of favorable to total possible outcomes of a fair chance process—and not any long-run relative frequency—that was the intended determinant of probabilistic reasoning.

Probabilities were thus first introduced as part of a general theory of the conditions of fair play in games of chance. The general approach that was adopted by the classical theorists was to model a game of chance as an initially fair game that had been altered so that each of the players stood to gain as a result of not just one but a plurality of distinct outcomes. Combinatorial reasoning, or 'the method of combinations,' would then be applied to assess the probabilities of events, which would, in turn, be used to adjust the relative cost of participation so as to accord with a player's expected returns. Simple games of chance played a seminal role in the history of probability precisely because they admitted of the most straightforward modeling along these lines. This, in turn, was owing to the fact that the fairness of the chance processes employed in such games could be ascertained by direct appeal to their manifest symmetry.

It is important to emphasize that the classical theorists of probability did not conceive of their theory as explaining in any way the fairness of the symmetrical chance processes employed in simple games of chance. That tossing a coin or casting a die was a fair method for deciding the winner of a game, was, for the classical theorist, an intuitively obvious fact. The calculus of probabilities had no role to play in corroborating such fundamental intuitions. Indeed, unless one was willing to acknowledge that such elementary chance processes were fair, probabilistic reasoning could have no application, for the only way to determine the probability of an event was by counting the number of outcomes of a fair chance process implying that event's occurrence. In the next section, we will consider in more detail how the classical theorists of probability sought to justify such pre-probabilistic judgments of fairness. As we shall see, it is in the attempt to explicate the grounds of these judgments that the apparent tension between determinism and the classical conception of probability comes into sharpest relief.

### 3. A paradox of sufficient reason

The classical theory of probability was based on a pre-theoretical notion of a 'fair' chance process. This notion derived its intuitive significance from its paradigmatic applications in the context of simple games of chance involving the tossing of coins, the casting of dice, or the drawing of cards from a shuffled deck. By direct appeal to the fairness of such elementary chance processes, the earliest theorists of probability were able to compute the

---

17 A VI.4 92. The importance for gambling of equality of conditions is emphasized by Cardano in his sixteenth-century gambling manual, Liber de ludo aleae. He writes: "The most fundamental principle of all in gambling is simply equal conditions, e.g., of opponents, of bystanders, of money, of situation, of the dice box, and of the die itself. To the extent to which you depart from that equality, if it is in your opponent's favor you are a fool, and if in your own, you are unjust." (Cardano, 1593, 189).

18 See B 110. The first part of Bernoulli's Ars conjectandi consists of the Latin text of Huygens's De ratiociniis in ludo aleae, with annotations by Bernoulli (B 110–50). In what follows, all references to Huygens's treatise are to Bernoulli's annotated edition of the text.

19 The term 'expectation' (expectatio) was not Huygens's own term, but was rather introduced in a Latin translation of the original Dutch treatise by Huygens's correspondent, van Schooten (Freudenthal, 1980). Freudenthal (1980, 116) translates Huygens's own Dutch 'kanse' as 'chance' rather than 'expectation.'

20 B 112.

21 de Moivre (1756, 1–3).

22 Huygens, for example, does not define a fair game in probabilistic terms as one in which every party has an equal expectation of gain. Instead, he proceeds in the opposite direction, defining expectation in terms of fair games: 'One's expectation to gain any thing, is worth so much, as, if he had it, he could purchase the like expectation again in a just and equal game' (B 110). If by a 'just and equal game' Huygens simply meant one in which the expected returns are the same for all players involved, then this definition would be patently circular (cf. Daston (1980, 236–7) and Hald (1990, 68)).

23 In response to the limited scope of applicability of such 'a priori' combinatorial methods, Bernoulli proposed to estimate probabilities by the 'a posteriori' method of observing long-run relative frequencies. This led Bernoulli to establish the first limit theorem in probability theory, an integral form of the weak law of large numbers (see Hald (1990, ch.16). Despite his seminal contributions to the statistical theory of probability, it is important to emphasize that Bernoulli was not a frequentist. He viewed the 'a posteriori' method as an indirect way of assessing the ratio of favorable to total number of cases, when these cases could not be discerned by direct appeal to the symmetry of the chance process under consideration (see B 247–8; cf. Schneider (1984, 82)).
probabilities of complex events using strictly combinatorial methods, that is, by counting cases. But what exactly were the grounds on which the classical theorists of probability based their initial judgments of fairness? On this fundamental issue, the classical probabilists were embarrassingly vague. While they deftly avoided the use of probabilistic language in their depictions of fair chance processes, they would often employ phrases that were close enough in meaning to equally probable as to render it entirely unclear how these descriptions were meant to play any non-circular role in probabilistic reasoning. Thus one finds the possible outcomes of a cast of a fair die or the toss of a fair coin, described on various occasions as ‘occurring with equal ease’, ‘occurring with equal tendency,’ ‘occurring with equal facility,’ or, most notoriously, as ‘equally possible.’24 It should come as no surprise that the principle that equally possible outcomes are equally probable, would later serve as the basis for a charge against the classical probabilists of committing the fallacy of petitio principii.25

On a few occasions, however, the classical probabilists did attempt to push their analysis of fairness further. Take, for example, Bernoulli’s proposed explanation as to why it is that we can ‘exactly calculate the advantages and disadvantages of players in games of chance’26:

The originators of these games took pains to make them equitable by arranging … that all the cases happen equally easily … So, for example, … for a single die there are manifestly as many cases as the die has faces. Moreover all have equal tendencies to occur; because of the similarity of the faces and the uniform weight of the die, there is no reason why one of the faces should be more prone to land face up than another—as would be the case if the faces had dissimilar shapes or if a die were composed of a heavier material in one part than another. (Bernoulli, Ars conjectandi, B 248, italics added)

In this passage, Bernoulli argues that the reason that dice casting is considered fair is that dice have been carefully engineered so as to exhibit a certain physical symmetry: the faces of a die are all of a similar shape and its mass is uniformly distributed throughout. Moreover, the significance of this physical symmetry, for Bernoulli, rests in the fact that it precludes us from finding any reason why one of the faces of a die should be more prone to land face up than any other. One finds a similar rationale in Laplace’s formulation of the classical definition of probability:

The probability of an event is the ratio of the number of cases favorable to it, to the number of all possible cases, provided there is no reason to believe that any one of these cases as opposed to any other should occur, making them, for us, equally possible. (Laplace, Théorie Analytique des Probabilités, L 7 181, italics added)

While Laplace’s formulation differs from Bernoulli’s in subtle and important respects, he likewise appeals to a ‘rational’ symmetry, as manifest in a certain absence of reasons, to explicate the notion of a fair chance process. This, however, is as far as either Bernoulli or Laplace were willing to carry the analysis. For a more principled account of the logic underwriting their reasoning, we must turn to a thinker of a somewhat more philosophical bent, namely Leibniz.

While Leibniz contributed little to the mathematics of probability, he had an intense interest in the foundations of the subject and may even have had some influence on its early development through his correspondences with Bernoulli.27 In a brief essay, titled De incerti aestimatione, Leibniz lays down some basic axioms of probabilistic reasoning, which he then uses to carry out various probabilistic calculations (for the most part, recapitulating results already established by Huygens).28 Among the first principles he states is the following axiom:

Axiom. If players do similar things so that no difference between them may be noticed except that consisting in the outcome, then the proportion of hope to fear is the same. This can be demonstrated by metaphysical principles. Where appearances are the same, i.e., the reason for forming an opinion about the future outcome is the same, one can form about them the same judgment. (Leibniz, De incerti aestimatione, A VI.4 92)

This axiom states a sufficient condition for each of the players in a game to have the ‘same proportion of hope to fear.’ By ‘hope’ Leibniz means the probability of winning and by ‘fear’ the probability of losing.29 Now, as Leibniz later observes, if we consider a simple lottery-style game in which each player stands to win the same amount ‘in one of the outcomes and in none of the others,’ then each player has the same proportion of hope to fear, just in case each has the same hope, i.e., the same probability of winning.30 Thus, in the context of such a game, Leibniz’s axiom asserts that every player has the same probability of winning, provided the game is played under equal conditions, so that ‘no difference between [the players] may be noticed except that consisting in the outcome.’31

Leibniz’s axioms thus amounts to the claim that from equal conditions of play, it may be inferred that each of the players in a game of chance has an equal probability of winning. What is more interesting, for our present purposes, however, is what Leibniz says next. For he goes on to assert that this axiom can be justified ‘by metaphysical principles.’ While Leibniz does not specify exactly what he means by this, it is clear from the remarks that follow, as well as what he says elsewhere, that his intended metaphysical justification is based on the principle of sufficient reason. This principle, which Leibniz identifies as one of the ‘two great principles’ that govern our reasonings (the other being the principle of non-contradiction), asserts that ‘there can be found no fact that is true or existent, or any true proposition, without there being a sufficient reason for its being so and not otherwise.’32 More colloquially, the principle states:

24 Hacking [2006, ch.14] argues that the modal construction ‘equally possible’ was useful to the classical probabilists because it allowed them to take advantage of the intuitive difference between de re and de dicto senses of possibility to move back-and-forth, when convenient, between an epistemic and an aleatory interpretation of probability.
25 See Von Mises [1956, 67–8]. Similar objections to the classical definition of probability were put forward by Mill [1974, 1140–4] and Venn [1876, 26–31].
26 B 269.
27 As Hald [1990, 189] observes: ‘Leibniz did not make any direct contribution to probability theory, but in his unpublished manuscripts and his correspondence with other scientists, he had considerable influence on the philosophy of probability and the scope of its applications.’ For more on Leibniz’s possible influence on Bernoulli’s work in probability, see Schneider [1984] (cf. Franklin [2015, 365]).
28 A VI.4 92–101. For a commentary on De incerti aestimatione, see de Melo and Cussens [2004].
29 Leibniz writes: ‘Probability is the degree of possibility. Hope is the probability of winning. Fear is the probability of losing.’ (A VI.4 94) In fact, by hope and fear Leibniz means the probabilistically weighted value of winning and losing, respectively. However, at this point in the essay, he restricts his attention to games in which each player assigns the same value to winning and losing so that hope and fear are determined by the probabilities alone.
30 A VI.4 95.
31 See Cussens [2013, 5].
PRINCIPLE OF SUFFICIENT REASON: Everything that exists, exists for a reason.

How did Leibniz intend to use this principle to justify the inference from equal conditions to equal probabilities? The answer can be seen by attending to Leibniz’s subsequent remark that ‘where appearances are the same ... one can form about them the same judgment.’ As Leibniz explains elsewhere, this law is a direct corollary of the principle of sufficient reason:

It also follows [from the principle of sufficient reason] that, when in the givens everything on the one side appears the same as it is on the other side, then everything will be the same in the unknowns, that is, in the consequences. This is because no reason can be given for any difference, for this reason must certainly derive from the givens. A corollary of this, or better, an example, is Archimedes’ postulate at the beginning of the book on statics, that, given equal weights on both sides of a balance with equal arms, everything is in equilibrium. (Leibniz, *Principia logico-metaphysica*, A VI.4 1645)

Leibniz’s reasoning in the passage is clear enough. First, we are to consider the principle of sufficient reason in its contrapositive form: if there is no reason for something to exist, then it does not exist. We then apply this contrapositive form of the principle to differences between magnitudes, obtaining as a result:

PRINCIPLE OF SUFFICIENT REASON¹: If there is no reason for there to exist a difference between two magnitudes, then there is no difference in these magnitudes, i.e., these magnitudes are the same.

Throughout his writings, Leibniz makes occasional use of the principle of sufficient reason in this contrapositive, differential form. In the passage just quoted, for example, he appeals to this principle to justify the Archimedean postulate that ‘equal weights at equal distances are in equilibrium.’ As it relates specifically to probabilities, the relevant instance of the principle can be expressed as follows:

PRINCIPLE OF SUFFICIENT REASON²: If there is no reason for there to exist a difference between the probabilities of two outcomes of a chance process, then there is no difference in probability between these outcomes, i.e., their probability is the same.

For Leibniz, then, probabilistic reasoning has its justificatory basis in the principle of sufficient reason. By means of this principle, one can infer that the possible outcomes of a game of chance are all equally likely to occur, provided that there is no reason why any one such outcome should be more likely to occur than any other. In what follows, we will take it for granted that this last condition captures what the classical probabilists had in mind when they described a game of chance as fair, or its outcomes as equally possible.

A difficulty, however, looms large. For the assumption that games of chance are fair in the sense just described would appear to directly conflict with the supposition that the chance processes employed in these games are deterministic in nature. To articulate the conflict in more precise terms, consider again the specific case of dice casting. Given the principle of sufficient reason, the assumption that is needed to justify the conclusion that each of the sides of a die have an equal probability of landing face-up is the following:

DiE-1: There is no reason why any one side of the die should have a greater probability of landing face up than any other.

But, if dice casting is indeed a deterministic process, then it is difficult to see how one can accept this claim. For, if a die were to land with, say, its 3-side showing face up, then determinism would seem to imply that there must have been some reason why it did so in the form of a prior necessitating cause. Indeed, Bernoulli concedes as much when he writes:

It is most certain, given the position, velocity, and distance of a die from the gaming table at the moment when it leaves the hand of the thrower, that the die cannot but fall other than the way it actually does fall. (Bernoulli, *Ars conjectandi*, B 240)

Bernoulli’s determinism thus commits him to the following claim:

DiE-2: Whichever side of the die lands face up, there must be some reason for its having done so.

The apparent incompatibility of DiE-1 and DiE-2 is the ultimate source of the interpretive difficulties surrounding the classical conception of probability. On the one hand, by endorsing DiE-1, the classical probabilists sought to attribute to the process of dice casting a certain rational symmetry based on the regularity of the die and its uniform mass distribution. At the same time, by endorsing DiE-2, they acknowledged that, at a fundamental causal level, dice casting must be rationally asymmetrical. For whichever side of the die lands face up must possess certain properties that suffice to distinguish it from all the other sides of the die as that which was predetermined by law to have done so.

Now, given what we have said so far, there is an irony in the fact that what led the classical probabilists to accept claims such as DiE-2 and, more generally, what underlay their commitment to causal determinism, was none other than the principle of sufficient reason.³⁵ Thus, for example, Laplace writes:

The connexion between present and preceding events is based on the evident principle that a thing cannot come into existence without there being a cause to produce it. This axiom is [is] known as the principle of sufficient reason. (Laplace, *Essai philosophique sur les Probabilités*, L 7 vi)

In this context, the application of the principle is more straightforward than before, relying only on what Leibniz terms the ‘vulgar’ formulation of the principle, which asserts that ‘there is no effect without a cause.’³⁶

At the heart of the classical theory of probability there thus resides what might be termed a paradox of sufficient reason. To resolve the paradox, one must explain how this single principle can, at once, license judgments of fairness in the context of simple

³³ As van Strien (2014, 25) argues: “Laplace’s determinism is based on general principles, rather than derived from the properties of the equations of mechanics: specifically, it is based on the principle of sufficient reason and the law of continuity.”

³⁴ Leibniz (1875, 309). As van Strien (2014, 30) rightly points out, this version of the principle does not entail Laplace’s strong determinism according to which each state of the universe is the effect of the immediately preceding one. For this, Laplace needs to appeal to a law of continuity, precluding the possibility of discontinuous state transitions. For both Laplace and Leibniz, however, the principle of continuity was seen as a corollary of the principle of sufficient reason (see van Strien (2014, 29) and Mates (1980, 162–6)).
games of chance, and, at the same time, supply the metaphysical basis for the conviction that any such game has a causally predetermined outcome. It is clear that the classical probabilists felt some pressure to resolve this paradox, at least when they were led to reflect on the foundations of their theory. Moreover, they responded to this pressure in much same way, namely, by insisting that probabilities were to be understood in subjective or epistemic terms, as relative to an agent’s partial state of knowledge. Thus, for example, after observing that the outcome of a cast of a die is no less predetermined than the occurrence of an eclipse, Bernoulli writes:

Yet it is customary to count only the eclipse as necessary and to count the fall of the die … as contingent. The only reason for this is that … the study of geometry and physics has not been sufficiently perfected to enable us to calculate … these [latter] effects … Before astronomy was brought to this degree of perfection, eclipses themselves, had to be counted among future contingencies. It follows, therefore, that something can be seen as contingent by one person at one time, which may be necessary to another person at another time. (Bernoulli, Ars conjectandi, B 240)

In response to the paradox of sufficient reason, the strategy of the classical probabilists was thus to relativize probability to an agent’s subjective state of knowledge. One consequence of this approach is that, from the perspective of Laplace’s subjective interpretation of probability, nothing happens by chance—every true claim is assigned a probability of 1 and every false claim a probability of 0. When a less perfect intelligence assigns to an event an intermediate degree of probability, they are simply articulating, in numerically precise terms, the structure of their own ignorance.

Whatever else may be said of this subjectivist maneuver, it clearly succeeds in avoiding the paradox, for there is obviously no conflict between DIE-2 and the following relativized version of DIE-1:

DIE-1\*: There is no reason, as far as we know, why any one side of the die should have a greater probability of landing face up than any other.

But should the classical probabilists have felt compelled to adopt a subjective interpretation of probability in order to avoid the paradox? If one examines more closely the claims DIE-1 and DIE-2, one finds that they do not, strictly speaking, conflict with one another, as they would, for example, if DIE-1 were formulated as follows:

DIE-1\*: There is no reason why any one side of the die should land face up as opposed to any other.

While DIE-1\* is clearly in contradiction with DIE-2, it is not the assumption that is needed in order to infer from the principle of sufficient reason that the possible outcomes of a cast of a die are equally likely to occur. Instead, what is needed is DIE-1, which asserts that there is no reason why any one side of the die should have a greater probability of landing face up than any other.

Now, it may seem obvious that DIE-1 implies DIE-1\*, and, indeed, both formulations appear from time to time in the writings of the classical probabilists. Thus, for example, whereas Bernoulli says that ‘there is no reason why one of the faces should be more prone to fall than another,’ Laplace simply asserts that there is no reason to believe that ‘any one of these cases as opposed to any other should occur.’ Nevertheless, the inference from DIE-1 to DIE-1\* is not entirely straightforward. It relies on the following easily overlooked assumption:

Any reason one may have for judging that an event will occur is ipso facto a reason for judging that event to be more probable than any of its alternatives.

This assumption amounts to an assertion of the rational commensurability of causal and probabilistic reasoning. It implies that when we have a reason to judge that a certain event will occur in the form of a known deduction from prior causes, then this deduction itself provides us with a reason for judging that event to be more probable than any of its alternatives. Did the classical probabilists accept such rational commensurability? At least, in certain cases, the answer seems to be yes. Take, for example, the following remark made by Bernoulli:

Those things concerning the existence or future occurrence of which we can have no doubt—whether because of revelation, reason, sense, experience, eyewitness testimony, or other reasons—enjoy the highest, and absolute certainty. All other things receive a less perfect measure of certainty in our minds … Probability, indeed, is degree of certainty, and differs from the latter as a part differs from the whole. (Bernoulli, Ars conjectandi, B 239)

In this passage, Bernoulli treats all reasons, regardless of their source, as commensurable insofar as they confer a certain degree of certainty on our judgments. He claims, moreover, that probability simply is the measure of this certainty. Consequently, for Bernoulli, every reason, regardless of its source, is a probabilistic reason. For example, the reasons one may obtain from sense experience or eyewitness testimony for judging that a certain claim is true, count as reasons for assigning to that claim a maximal degree of probability. Thus, it seems, Bernoulli would have granted the commensurability of causal and probabilistic reasoning and so would have accepted the inference from DIE-1 to DIE-1\*.

This, however, is only part of the story. For while the classical probabilists may, in precept, have accepted the commensurability of causal and probabilistic reasoning, in practice, they did not. Indeed, as we shall see in the following section, it was precisely the denial of such commensurability that allowed the classical probabilists to apply combinatorial methods in imaginative ways to solve a variety of previously unsolved problems in probability theory. In the next section, I will illustrate this point in the context of Fermat’s solution to the problem of points.

4. The Fermat-Pascal-Roberval correspondence

It has become orthodoxy to date the beginning of the history of mathematical probability at 1654 with the brief exchange of letters between Pascal and Fermat. The Pascal-Fermat correspondence centers around a recalcitrant problem in the doctrine of chances,

[38] This convention, which dates back to the history of Todhunter (1865), is not unreasonable given the direct line of influence that can be traced from the Pascal-Fermat correspondence through the works of all the major figures in the classical tradition. Huygens was prompted to produce his treatise Ratiociniis in ludo aleae after hearing about the Pascal-Fermat correspondence during his stay in Paris in 1655 (Freudenthal, 1980, 113). An annotated edition of the Latin version of Huygens’s treatise was then reprinted in Bernoulli’s Ars conjectandi, published posthumously in 1713. News of Bernoulli’s unfinished treatise is what stimulated Montmort to publish his Essay d’analyse de hazard in 1708, which in turn led de Moivre to write his De mensura sortis in 1712 (for a detailed account of this history, see Hald (1990)). Even the half-century gap between Huygens and Bernoulli can be smoothed over by considering the works of more minor figures (see Stigler (1988) and Hald (1990, ch.12)).

37 Daston (1994, 332) claims that, for Laplace, ‘objective probability’ would have been an oxymoron. While she is correct that Laplace often falls back on subjectivist language when trying to avoid any direct conflict with his deterministic convictions, it is worth noting that, in one of his earlier memoirs, he distinguishes quite clearly between objective (absolute) and subjective (relative) notions of probability, and groups probabilistic judgments about games of chance under the former heading (L 9 384–5).

38
which has since come to be known as the problem of points, and which had been brought to Pascal’s attention by the gambler and amateur mathematician, the Chevalier de Méré.39

Suppose that two players, A and B, are engaged in a game which is played in multiple rounds. The winner of each round is determined by the outcome of some fair chance process, say, the toss of a coin. At the beginning of the game, each player stakes an equal amount of money on the agreement that the first player to win \( n \) rounds will be awarded the entirety of the stakes. Now, consider a point in the game at which player A has won \( a \) rounds to player B’s \( b \) (where \( a \) and \( b \) are both less than \( n \)). If the game were to be left unfinished at this point, what would constitute an equitable division of the stakes?

The problem of points has a long and well-documented history dating back to the late fifteenth century, and many erroneous attempts to address the problem were made prior to Fermat’s discovery of what is now acknowledged to be the correct solution.40

Fermat’s solution is based on the following argument: since the winner of the game is the first to win \( n \) rounds, supposing the game had not been interrupted, the winner would have been decided in at most \( m = (2n - 1) - (a + b) \) additional rounds. Since, however, each of the possible ways of extending the game by \( m \) rounds is equally possible, the proportion of the total stakes to which each of the players is entitled is just the proportion of the \( 2^m \) possible continuations of the game which entail victory for that player.

Suppose, for example, that player A needs 2 more rounds to win the game, and player B needs 3. Then, the game will be decided in at most 4 additional rounds. The \( 4^4 = 16 \) possible continuations of the game for four additional rounds can be enumerated as follows:

<table>
<thead>
<tr>
<th>A wins</th>
<th>B wins</th>
</tr>
</thead>
<tbody>
<tr>
<td>ABBB</td>
<td>BBBB</td>
</tr>
<tr>
<td>ABBA</td>
<td>ABB</td>
</tr>
<tr>
<td>AABA</td>
<td>BAB</td>
</tr>
<tr>
<td>AAAA</td>
<td>BBAA</td>
</tr>
<tr>
<td>BAAA</td>
<td>BBBA</td>
</tr>
<tr>
<td>BBAA</td>
<td>BBBB</td>
</tr>
</tbody>
</table>

Of these 16 possible continuations, 11 are such that player A wins 2 rounds before player B wins 3 (the remaining 5 being such that player B wins 3 rounds before player A wins 2). Hence, according to Fermat, the proportion of the total stakes to which player A is entitled is \( \frac{11}{16} \).

Despite some initial misunderstandings, Pascal was eventually brought around to see the correctness of Fermat’s method.42 In the course of their exchange, however, Pascal raised one potential objection to Fermat’s analysis, which he attributed to the mathematician Roberval:

I communicated your method to some of our gentlemen, on which M. de Roberval made me this objection: that it is wrong to base the method of division on the supposition that they are playing in four throws seeing that when one lacks two points and the other three, there is no necessity that they play four throws since it may happen that they play but two or three, or in truth perhaps four. Since he does not see why one should pretend to make a just division on the assumed condition that one player plays four throws, in view of the fact that the natural terms of the game are such that they do not throw the dice after one of the players has won ... he suspects that we have committed a paralogism. (Pascal to Fermat, F 302)

As Roberval points out, in several of the above continuations of the game, the winner is determined prior to the fourth additional round being played. Thus, for example, in the continuation AABA, player A wins after only two additional rounds have been played. But why then, Roberval asks, should we pretend, in this case, as if the game will continue for two further rounds, given that a winner has already been decided? Roberval’s objection is directed to the fact that Fermat’s analysis requires us to countenance ‘fictional’ possibilities, which, given ‘the natural terms of the game,’ will not occur. Instead, according to Roberval, the only possible continuations of the game that ought to be considered are those in which, once a winner has been decided, no further rounds are played, i.e.:

<table>
<thead>
<tr>
<th>A wins</th>
<th>B wins</th>
</tr>
</thead>
<tbody>
<tr>
<td>ABBA</td>
<td>BBBB</td>
</tr>
<tr>
<td>ABAB</td>
<td>BABB</td>
</tr>
<tr>
<td>AAAA</td>
<td>BBAA</td>
</tr>
<tr>
<td>BBBB</td>
<td>ABBA</td>
</tr>
</tbody>
</table>

Neither Pascal nor Fermat were moved by Roberval’s objection, and it left little mark on the later theoretical developments that grew out of Fermat’s combinatorial solution to the problem of points. Moreover, in subsequent historical treatments of the Pascal-Fermat correspondence, little attention is paid to this objection, and it is often introduced only to be dismissed out-of-hand.43 Todhunter, for example, devotes to Roberval’s complaints only one short paragraph in an extended and detailed summary of the correspondence, and after describing Pascal’s response to Roberval, offers a rare editorializing remark, writing: ‘Pascal puts this point very clearly.’44 In a similar vein, Leibniz, who was much enthralled with Fermat’s

---


40 For a history of the problem of points prior to Pascal and Fermat, see Hald (1990, 35–6).

41 In modern terms, Fermat’s solution yields the negative binomial distribution. If \( f(a, b) \) is the proportion of the total stakes owed to player A when A has won \( a \) games to B’s \( b \), then:

\[
f(a, b) = \frac{\sum_{i=0}^{\min(a+b, n)} \binom{n}{i} (-1)^i}{2^n}
\]

42 Initially, Pascal misinterpreted Fermat’s suggestion, imagining that the \( m \) additional rounds were to be played all at once rather than in sequence. As a result, Pascal contended that while Fermat’s method solves the problem for the two-player case, it does not solve the analogous problem for three players. For, as Pascal noted, in the three-player case, if the smallest number of additional rounds needed to decide the game are played all at once, it may be that more than one player will secure the \( n \) rounds needed to win. In a subsequent letter, Fermat corrects Pascal’s mistake, observing that if the order of the additional rounds are taken into account, then each possible continuation determines a unique winner of the game. The exchange ends with Pascal conceding that Fermat’s solution does, in fact, apply in full generality.

43 Roberval’s objection is mentioned by Todhunter (1865, 13–14), David (1962, 91), Hald (1990, 60), Franklin (2015, 308–9), and Edwards (2002, 147). Edwards’s dismissive attitude towards Roberval’s objection is fairly typical: ‘Someone with less of a feel for probability than Pascal and Fermat might object that their method of combinations includes, in its enumeration, tosses that would under no circumstance have been made. Roberval was to make just this point when he discussed the method with Roberval. To us, of course, there is no paradox, and we may instead observe that Pascal and Fermat have hit upon the notion of “embedding” one probability problem in another’.

44 Todhunter (1865, 13–14).
solution to the problem of points, describes with praise the ‘beautiful ideas on chance of Messrs. Fermat and Pascal, which Mr. Roberval either could not or did not wish to understand.’

Various objections might be offered as to why Roberval’s objection has been so unduly neglected in the literature. This may, in part, be due to the association of Roberval’s objection, in the minds of some critics, with certain obviously fallacious proposals for how to solve the problem of points. It may also be due to the fact that Fermat’s solution is so intuitively compelling. Most likely, however, it is due to the fact that Roberval’s objection is not a mathematical or technical criticism at all, but rather a philosophical one. It touches on the subtle conceptual issue of how to interpret the modal space of possibilities in which probabilistic reasoning is to take place.

That the philosophical nature of Roberval’s objection was entirely lost on Fermat is evident from the responses it elicited from him. As an initial response, Fermat pointed out that, according to his method, the game could be extended by any fixed number of rounds and the result would be the same. Thus, for example, if instead of 4 additional rounds, one assumed that the game would be played for 5, or 6, or 7 additional rounds, the proportion of the total stakes owing to player A would still be $\frac{11}{16}$. In making this observation, Fermat’s intent was presumably to show that the players should be indifferent to the prospect of extending the game by any number of additional rounds. In fact, however, all that this shows is that if the players have agreed to play some fixed number of additional rounds, then they ought not care how many additional rounds they play. It does not explain why they should not care whether they play a fixed number of additional rounds or play only until a winner has been decided.

Fermat’s second response to Roberval is somewhat more involved. It is meant to show that ‘this fiction, of extending the game to a certain number of plays,’ is only a useful mathematical shorthand which ‘serves to make the rule easy and … to reduce all the fractions to the same denomination.’ Here, Fermat has in mind the following sort of argument: consider only the possible continuations of the game which Roberval would be willing to countenance. Of these continuations, the ones in which A wins the game are:

```
ABBA
ABA BABABA
```

Thus, player A can win in either 2, 3, or 4 additional rounds. In the first case there is only one way for player A to win the game; in the second case there are two ways; and in the third case there are three. But since there are $2^2 = 4$ equally possible ways for two rounds of the game to play out, $2^3 = 8$ equally possible ways for three rounds of the game to play out, and $2^4 = 16$ equally possible ways for four rounds of the game to play out, Fermat concludes that the total proportion of the stakes owed to player A is:

$$\frac{1}{4} + \frac{2}{8} + \frac{3}{16} = \frac{4}{16} + \frac{4}{16} + \frac{3}{16} = \frac{11}{16}$$

Thus, Fermat claims, we arrive at the same solution ‘without any dissimulation’ and ‘without recurring to assumed conditions.’ This, however, is a mathematician’s feat. What Fermat has done is to replace the question:

What is the probability that A wins the game, on the assumption that 4 additional rounds are played?

with the more complicated question:

What is the probability that either (i) A wins the game in exactly 2 rounds, on the assumption that 2 additional rounds are played; or (ii) that A wins the game in exactly 3 rounds, on the assumption that 3 additional rounds are played; or (iii) that A wins the game in exactly 4 rounds, on the assumption that 4 additional rounds are played?

The advantage of recasting the question in these terms, is that with respect to any one of the events (i)–(iii), the possible outcomes which imply this event are included among the ten genuine ‘Robervalian’ possibilities enumerated above. To assess the probability of any one of these events, however, one must consider not only the outcomes favorable to it, but also the total number of possible outcomes, since the probability is the ratio of these two quantities. But, in counting the total number of possible outcomes, Fermat must again countenance fictional continuations of the game which extend beyond the point at which a winner has been decided. So, for example, in counting 8 possible ways in which the game can be continued for an additional 3 rounds, Fermat is including in his tally the ‘fictional’ continuation AAB. Thus, Fermat’s claims that the positing of such fictional continuations of the game are just a useful mathematical device that play no essential role in the analysis, is incorrect. Unless one countenances such possibilities, there is simply no combinatorial solution to the problem of points.

The disagreement between Roberval and Fermat cannot be settled by any mathematical sleight-of-hand, for it is based on a more fundamental conceptual disagreement as to the nature of the possibilities countenanced in probabilistic reasoning. Roberval first observes that, given the ‘natural terms of the game,’ the only

45 Quoted in Todhunter (1865, 14).
46 For example, Roberval’s objection is sometimes associated with D’Alembert’s erroneous claim that the probability of a fair coin’s landing heads at least once in two tosses is $\frac{2}{3}$ (see, e.g., Todhunter (1865, 14)). If Roberval had counted each of the ten genuine possibilities he identifies as equally possible, and, on these grounds, argued that player A should receive 6/10 of the total stakes, then this association would be just. But nowhere, to my knowledge, does Roberval endorse this solution to the problem of points.
47 Thus, e.g., David (1962, 91), after mentioning Roberval’s objection adds the following note: ‘It is perhaps not necessary to point out that [Fermat’s] method of solution is one which we would use today.’
48 Hald (1990, 60) even falls prey to the notion that Roberval’s objection amounts to a mere technical error. Agreeing with Pascal, he writes that Roberval’s objection fails because ‘the actual and the hypothetical play, will always give the same result for the two players, since the cases in which A wins or loses in the actual play will also be the cases in which he wins or loses in the hypothetical play.’ Such a remark seems to overlook the fact that the mapping from the hypothetical to the actual sample space is not injective.
49 F 311.
50 F 311. Franklin (2015, 309) describes Fermat’s claim that the device of fictional outcomes is simply a mathematical tool for reducing fractions to their lowest terms as ‘braeathtaking’ in its naivete, it is the remark of the purest of pure mathematicians.
51 F 311–2.
52 Indeed, Pascal seems to recognize this point. In light of Roberval’s objection, he initially concedes that ‘the method of combinations, … in truth is not in place on this occasion’ (F 302). Pascal himself proposed to determine the proportion of the total stakes owed to A by a recursive method, based on the following three conditions (where $f(a,b)$ is the proportion of the total stakes owed to A, when A has won a games to B’s $b$):

1. $f(a,b) - 1. f(a,n) = 0$
2. $f(a,a) = \frac{1}{2}$
3. $f(a,b) = \frac{f(a+1,b)+f(a,b+1)}{2}$

For more on Pascal’s recursive solution to the problem of points, see Edwards (1983, 70–5).
possible continuations of play are those in which the game concludes once a winner has been decided. Fermat then replies that in order to ‘make all the chances equal,’ so that we may apply combinatorial reasoning to the problem, we must extend the game to a certain fixed number of plays by admitting the possibility that the game might continue beyond the point at which a winner has been decided. To this, Roberval objects that doing so is illegitimate since it requires us to countenance possibilities which we know will not occur. In reply to this objection, Fermat then offers a number of unconvincing arguments that are meant to show that his method can be applied without positing such fictional possibilities. In truth, however, his only available reply to Roberval is to bite the bullet. He can only reply: ‘So be it! All that your objection goes to show, M. de Roberval, is that knowledge of ‘the natural terms of the game’ is irrelevant for assessing the probability with which player A will win.’

Consider, however, what our imagined Fermat has just been forced to admit. He has conceded that while there may be some reason, based on his knowledge of the natural terms of the game, for judging that AABA will not occur, this, by itself, does not give him any reason for excluding AABA from his probabilistic calculations. Indeed, for the purposes of these calculations, AABA may be regarded as equally possible to ABBA. Hence, in order to justify his combinatorial solution to the problem of points, Fermat must deny that any reason one may have for judging that a certain outcome will occur is ipso facto a reason for judging it to be more likely than any of its alternatives. He must, in other words, treat such reasons as incommensurable.

Recall that this incommensurability was exactly the assumption needed to block any genuine contradiction from arising in the context of the paradox of sufficient reason. Thus, the only way to accept as valid Fermat’s combinatorial solution to the problem of points is to reject the very assumption needed to generate the conflict between chance and determinism. To bring out this last point in a more vivid way, let us rehearse the above conversation between Fermat and Roberval, only now, let us endow Roberval with the unbounded intellectual capacities of Laplace’s infinite intellect. Roberval says to Fermat that given the ‘natural terms of the game’—by which he now means all of the facts describing the state of the world at the time that the game was interrupted as well as all of the laws governing its subsequent evolution—there is only one possible continuation of play that should be countenanced, namely that which is predetermined by law to occur. Fermat then replies that in order to ‘make all the chances equal,’ so that we may apply combinatorial reasoning to the problem, we must admit the possibility that the game might play out in a way that contravenes the laws of nature. Roberval then objects that to do so would be illegitimate, since it requires us to countenance possibilities which we know will not occur. To this, Fermat can only reply: ‘So be it! All that your objection goes to show, M. de Roberval, is that such perfect knowledge of the natural laws is irrelevant for assessing the probability with which player A will win the game.’

5. Conclusion

It is a generally held view that the apparent incoherence in the classical conception of probability derives from a pair of diametrically opposed theoretical pressures. On the one hand, the classical probabilists’ commitment to determinism pushed them in the direction of a subjective, or epistemic, interpretation of probability according to which judgments of probability reflect an agent’s state of partial knowledge. On the other hand, their desire to develop a mathematically rigorous science of chance phenomena led them to conceive of probability in objective terms, as ultimately grounded in the statistical regularities exhibited by such elementary chance processes as dice casting, coin tossing, and the like.

As I hope to have shown in this paper, the situation is, in fact, more nuanced than this simple account suggests. The classical probabilists’ belief in determinism did not, by itself, force them to adopt a subjective interpretation of the judgments of equiposibility appealed to in their analyses of games of chance. To the extent to which they felt compelled to adopt such an interpretation, this was owing to the additional assumption that causal and probabilistic reasoning are, in a certain sense, commensurable. Complicating matters still further, it was precisely the denial of such commensurability that enabled the classical probabilists to solve a variety of recalcitrant problems in probability theory using strictly combinatorial methods.

All told then, the fundamental philosophical challenge raised by the classical theory of probability is not that of reconciling subjective and objective interpretations of probability. It is rather the challenge of justifying the sort of rational incommensurability that afforded the classical probabilists the freedom to ignore certain causal factors that might otherwise prevent them from imaginatively constructing spaces of equipossible outcomes—the sort of freedom which, for example, allowed Fermat to ignore the consequences, cited by Roberval, of playing a game on its ‘natural terms. Unfortunately, attempts to address this philosophical challenge lies well beyond the scope of this paper. The aim of the present work has rather been to bring to light the central importance of this issue for arriving at a philosophically satisfying understanding of probability theory in its classical form. The question of the commensurability of causal deduction and probabilistic reasoning is, on the one hand, crucial for understanding the scope of applicability of the combinatorial methods employed by the classical probabilists. At the same time, it is only by answering this question that we may assess the true extent of the threat posed by determinism to an objective construal of the judgments of equiposibility on which such combinatorial methods are based.

Now, the incommensurability of probabilistic reasoning with at least some non-probabilistic modes of reasoning is hard to deny. Thus, for example, to directly observe that a coin that was just tossed is now lying on a table with its heads-face up does not, intuitively, give one any reason to doubt that the probability of its having done so is 1/2.53 But it is a further question whether the same judgment that the coin will land heads, when not the result of direct observation but rather of a deduction from prior causes, is likewise consistent with a belief that the coin is fair.54 Another way of putting this last question is as follows: how literally should we take the perceptual metaphor in Laplace’s depiction of his perfect intellect as one for whom both the future and the past are ‘open to its eyes’? Are the reasons available to such a perfect intellect for judging that a coin will land heads as probabilistically irrelevant as the reasons that you or I may have for believing this same claim on the basis of direct observation?

53 A similar point is made by Lewis (1987, 84). Lewis characterizes information obtained through direct observation as ‘inadmissible’, where inadmissible information is that which may lead us to assign a degree of credence to a claim that differs from what we believe to be the objective chance of its being true. He denies, however, that information obtained by means of a causal deduction from known laws is inadmissible (Lewis, 1987, 92–3). Accordingly, Lewis insists that ‘if our world is deterministic there are no chances in it, save chances of zero and one’ (Lewis, 1987, 120).
54 One potential ground for supposing that perceptual and causal reasons are alike in their probabilistic irrelevancy is that neither sort of reason seems to rely, in any essential way, on the conceptualization of the event judged to occur as the outcome of a ‘chance process’. Thus, for example, one can both see that a coin is showing heads-face up and deduce this fact from prior causes without treating the coin’s current position on the table as one of two possible outcomes of a ‘coin toss’.
Of course, even if this question is to be answered in the affirmative, it would still be true that Laplace's perfect intellect is not the sort of being against whom one would do well to place a bet. For once the wager has been specified, this being will be able to appeal to his comprehensive causal knowledge of the world in order to deduce whether or not the event in question will occur. If, however, such a being's deductive powers are to be understood on the model of an enlarged perceptual capacity, then such practical advantages would not be owing to his refined knowledge of the probabilities involved, any more than is the case with a gambler who cheats at cards—counting cards, after all, is still a form of cheating. Rather, from a practical point of view, the ideal nature of such a being would be reflected in the fact that he has no need to assess the probabilities of events at all. Yet if such a being were to inquire into the probability with which a fair coin will land heads, not for any practical purpose, but simply to satisfy his own curiosity, there would be nothing, in principle, to prevent him from concluding that the two possible outcomes of the toss are equally possible.

Thus, in the end, we may agree with Bernoulli that:

There is no place for conjectures in matters where one may reach complete certainty. Thus it would be pointless for an astronomer to want to conjecture about whether a particular full moon will be eclipsed or not from the fact that two or three are eclipsed every year, since he can find the truth of the matter by an infallible calculation. (Bernoulli, Ars conjectandi, B 241)

Pointless, yes, but perhaps not unintelligible. If causal deduction and probabilistic reasoning are indeed inconmensurable, then the threat to probability raised by determinism is not that such a perfect knowledge of astronomy would render one incapable of assessing the probabilities with which various astronomical events will occur, but rather that, apart from sheer curiosity, one would have no reason to care.

References

In citing the works of the classical theorists of probability, the following abbreviations have been adopted:


Todhunter, I. (1865). *A history of the mathematical theory of probability: From the time of Pascal to that of Laplace*. Macmillan.


