

Dice Games

Marie-France Bru and Bernard Bru

Abstract. Translated from the French by Glenn Shafer, the French text will appear as Chapter 1 of Volume 2 of *Les jeux de l’infini et du hasard*, by Marie-France and Bernard Bru, to be published by the Presses universitaires de Franche-Comté. The translation is published here with the permission of the publisher and the surviving author. The text has been edited to omit most references to other parts of the book. The authors extensive notes, which provide many additional references and historical details, have also been omitted.

Key words and phrases: *De vetula*, dice games, history of probability, Huygens, Jacob Bernoulli, Laplace, Montmort, normal approximation.

1. INTRODUCTION

Cubic dice appear in human societies at least as early as writing or numeration. Most museums have very good specimens, sometimes dating back to 2000 or 3000 BCE, which were excavated in Mesopotamia or in the Indus Valley (Stigler [30]). Their origin will not be discussed here, but it must at least be remarked that, unlike astragali or other objects thrown to the ground, dice, initially made of terracotta, are unique in that they tend to approach perfectly cubic dice, which are equally liable to show any one of their faces, which are marked by one to six points, thus approximating a kind of mathematical equi-probability long before the notion was imagined or named, the notion seeming thus to have emerged in full armor from a little piece of hardened clay into the sun of the Orient.

Of course, we know about mathematical cubes whose six faces are superposable by definition, studied mathematically as such ever since geometry has existed. Here, we are dealing with another type of mathematical property. Each face of the die is mathematically equivalent *when thrown*; none is given an advantage mathematically, even if it can be given an advantage physically (deliberately or not). All faces have the same “chance” to appear, and this is an equality by definition. Immediately, therefore, without having to appeal to materialistic or idealistic philosophical

theses, remarkable in all respects but foreign to our account, which is intended for children of all ages, we have a new idea, the idea of “chance,” and a mathematical object, a “mathematical die.” And “randomness,” the starting point for our thought and words about this mathematical object, instead of being that blind divinity that decides our destinies, becomes then also a mathematical randomness, a theoretical “abstract” die, as Borel [13] wrote, which falls to the ground without favoring in any way any of its faces or any player who might bet on one of them.

Another very common observation comes from dice players themselves, millions of players over thousands of years, who all know that the faces of a good die appear about equally often if thrown long enough, whereas by loading the die, you can modify this approximate equality enough to enrich yourself, if you are not first found out. The purely mathematical equality of dice then takes on a particular practical form. Certainly you cannot predict the outcome of a throw of a die, but you can bet that on a large number of throws, the six faces will appear approximately the same number of times.

We thus have multiple definitions of good dice. A mathematical definition: each face has the same chance of appearing. A statistical definition: each face appears roughly equally often in a large number of throws. And one can add juridical, psychological, epistemological, economic definitions, as many as you want. From the earliest known texts, at least those that we shall examine here, as we shall see, authors assimilate, amalgamate, with no comment or argument,

Marie-France Bru was formerly at UFR de mathématiques, Université Paris Diderot, France. She died on January 30, 2012. Bernard Bru is retired from MAP5, Université Paris Descartes, 45 rue des Saints-Pères, 75006 Paris, France (e-mail: leslogesb@gmail.com).

the mathematical equality of chances with the approximate equality of statistical frequencies. They conflate in one and the same language these two ways of defining (perhaps) the same thing. They write indifferently that one face appears more frequently than another or that it has more chances to appear. The two locutions are interchangeable. This double language, or rather this confusion of several languages, may simply be due to the fact that the authors are hard pressed to define the “chances” of which they speak, even if they understand them mathematically well enough to calculate them. Or it may be due to any other reason someone may want to propose.

Be that as it may, it is a fact that this special kind of multiple definition coalesces into a single physico-mathematical entity that we call a die and that everyone is supposed to understand clearly. The chances of the faces of this die are equal by definition, like the faces of a cube, and at the same time none of its faces is physically or morally “advantaged” if it is thrown. None can really happen much more often than another. This has always been the way, moreover, to judge of the goodness of a die, if you do not know how it was made.

At some point in the history of science, about which we know little or nothing, some scholars may perhaps have convinced themselves that as dice have their own mathematical existence, there should be a “geometry of chance,” a mathematical theory of dice, just as there is a geometric theory of cubes, and that it could be made into a calculus. And as the dice have a physical, juridical, religious existence, etc., these same scholars or others tried to apply this particular mathematics to study the “games of chance” played by their contemporaries, to determine the probability of the different possible outcomes and the *fair* value of gambles or bets, the value that maintains a kind of equality or justice among the participants, as the chances are multiplied. The calculus of dice would somehow permit us, at least ideally, to account for the frequencies of appearance of the different faces of a die, or the configurations of several dice when they are thrown together. One can thus speak, without forcing the point too far, of a (mathematical) “law” of dice, what will be called in the eighteenth century a law of probability that reason fits as closely as possible to the realities involved, just as a mathematical die is a simplified image of a good clay die. This is one example among others of the mathematical physics of Archimedes, Galileo or Newton, for whom nature is not mathematical, but can be described at least approximately in mathematical language.

Better still, and this time we have a reliable source, at the end of the seventeenth century the well-known scholar Jacob Bernoulli in Basel succeeded in putting into one and the same mathematical statement the two notions we have just introduced: the chance of one of the faces of a die and the frequency of its occurrences. This is Bernoulli’s theorem, which we examine in almost every paragraph of this work. This fundamental theorem can take very diverse forms and make use of more and more advanced mathematics, notably the mathematics of denumerable probabilities. It continues to be the object of active research.

In this chapter, we are interested only in the a priori calculus of mathematical dice, derived from their definition, six equally probable faces, what is now called the calculation of a probability law, a calculation of a particular nature that surely goes back very far in time, perhaps as far as the calculation of the dimensions of a pyramid, or a circle, or a square, though we know very little about this.

By definition, as we have said, each face of the die has the same chance of appearing. This is a hypothesis of a mathematical nature, a truth by definition, theoretical, preserved throughout a reasoning or a possible calculation. This calculation obviously depends on appropriate algebraic rules and also on the rules of the games considered. The history of dice games is as long as the history of men. It will suffice here to recall that in the Roman era (to speak simply), soldiers had the right to play dice in camp, so that Roman dice are found all over Europe. One of the most common Roman games, found in the Middle Ages and today, consists of throwing three dice and summing up the faces obtained. The first problem of the calculus of dice could therefore be this one: Do the different possible sums of points of three dice, from 3 to 18, have the same “chance” to occur, or are some favored over others, and if so, in what ratio? We shall call this the *problem of points*.¹ As for the word chance, we keep for the moment the mathematical meaning everyone knows, a kind of special weighting that allows us to begin the calculation, or as d’Alembert wrote in the preliminary discourse of the *Encyclopaedia*: a “quantity considered in the possibility of events,” which “produces the art of conjecturing, from which arises the analysis of games of chance.”

¹Translator’s note: Here, I translate the authors’ *problème des points* into English as *problem of points*. I will use *problem of division* for what is usually called *problem of points* in English: the problem of dividing the stakes when a game is ended before either player has enough points to win.

Or, if one prefers, an “abstract idea” at the foundation of the whole theory, as we know Cournot said so emphatically [15, 16].

First, a few simple observations that players must have made for a long time. There is an obvious mathematical symmetry in this problem. Points 3 and 18 (all the faces identical to 1 and to 6) necessarily have the same chance. The same seems likely true for points symmetrical with respect to the average, which is between 10 and 11, which thus must also have the same chance, although the reasoning here is less obvious (in general, it consists of changing the numbering, 1 becoming 6, 2 becoming 5 and so on).

Are these chances all equal as in the case of a single die? Certainly not. It is enough to have thrown three dice (or simply two dice as in the game of Monopoly) a few times to realize that the sums of points near the average, 10 or 11, even 9 or 12, are more advantageous, more frequent, than the extreme sums 3 or 18, or even 4 or 17. There must be a mathematical reason for this, and the problem is to find it.

2. THE PSUEDO-OVID'S *DE VETULA*

Here, we prefer the historical approach. To penetrate to the reasons of things, look at how they have gradually been revealed in the course of time, in their progression and in their ruptures, if any. In the case of the problem of points, as in all truly interesting cases, this is difficult or even impossible because the problem is lost in the mists of time and it is impossible to determine its origins and traditions in the absence of clearly identified written texts. Historians have examined this question ever since the nineteenth century, exploring the libraries at their disposal, and we have hardly progressed since then. Perhaps somewhere there is a Babylonian tablet or an Egyptian papyrus dealing with the question of points, but we do not know where they are. The first text usually cited on this question dates from the middle of the thirteenth century, a long encyclopedic poem, entitled *De vetula*, written in Latin hexameters and falsely presented as Ovid's testament. We do not know the author, but we know that he was connected to the University of Paris and that his poem probably served for a long time as a pedagogical resource in the universities and schools of the Middle Ages and the Renaissance. Be that as it may, the author by all appearances is a positive spirit, a scholar up to date on medieval science, and if his ultimate goal is indeed the conversion of souls, as was that of the builders of cathedrals, his poem *De vetula* nevertheless

possesses a certain probabilistic modernity which especially interests us here.

In the first book of the poem, the pseudo-Ovid enumerates and criticizes the pleasures of life and its profane pursuits. The author treats in particular different kinds of games, games of skill and reasoning, dice games. In the last case, the point is clearly to discourage the player from ruining himself, while awakening his intelligence. The aim seems to have been to show how chance can be reduced to calculation, even as neither this calculation nor anything else can ever dominate it. The calculation, in a certain way, puts chance in its place or, we might say, in its seat. It understands the game, but it is neutral and cannot prevent the ruin of fathers of families and the squandering of inheritances. Perhaps this is what Jean Le Fèvre, who produced a versified French adaptation of *De vetula* entitled *La Vieille ou les dernières amours d'Ovid* in the fourteenth century, meant in this trenchant title summarizing the Latin poet's point: “That he who knows how to seat the dice has no advantage in the game.”

The same moralizing intention is found in the second part of the poem, in which the author undertakes to dissuade his readers, generally clerics like himself, from indulging without measure in the pleasures of sex, detailing the disappointments of Ovid who, believing that he has finally succeeded in conquering the girl of his dreams but deceived by the night, finds himself in the bed of an “old woman,” a scabrous, rather obscene episode in the style of novels and medieval fables in which the woman is in turn idealized and trivialized. But here is not the place to detail this point any more than the third part, in which the author makes himself the defender of philosophy and theology, alone worthy of occupying a true scholar. The poem ends with Ovid's imaginary conversion to the true faith and a hymn to the Virgin Mary. In short, a medieval pedagogical poem, of which we will examine only the part that deals with the game of dice.

No doubt we should situate the calculus of dice in the Western mathematical context of the time. A relatively modest context, as it was in the time of Roman grandeur, more concerned with earthly conquests than with those of the human mind. This was deplored by Jean-Étienne Montucla in his great *Histoire des mathématiques* ([28], Volume I, page 482). The Roman colonies of the West hardly departed from this traditional disdain for mathematics, which was limited to the education of youth and the practical needs of architecture, commerce or calendars. All the more

so because mathematics had been compromised by astrology and methods of divination, whose astral fatalism and abuse were denounced and severely repressed by the Justinian code and several unanimous councils. Just the same, Montucla (ibid., page 506) observes with some relief that “the thirteenth century was almost a time of enlightenment in comparison.” It appears indeed that the author of *De vetula* knew and probably taught Boethius’s arithmetic, from which he borrows the use of tables inserted in the text, the “figures” that make it possible to visualize the “configurations of points” as well as the dice’s “ways of falling.” As J.-Y. Guillaumin [11, 22] points out in his edition of the *De institutione arithmetica*, Boethius was the first Latin author to use this pedagogical device, which he borrowed from his Greek model, Nicomachus of Gerasa’s *Introduction to Arithmetic* and which *De vetula* borrowed in turn, as certain Arab traditions did earlier (Djebbar [21]). It is also evident that the erudite thirteenth century poet used quite well the Indian positional numeration adopted by the Arabs in the ninth century. The Arabic decimal system appears in *De vetula* beginning with the first manuscript, in the form found in the twelfth–thirteenth centuries in the Latin translations of Al-Khwarizmi and in the treatises of Abraham Ibn Ezra, Fibonacci and Jordanus Nemorarius. We know that at the end of the thirteenth century, the Sorbonne Library had several manuscripts entitled “algorismus,” notably those of John of Sacrobosco, who was a professor of quadrivium² in Paris from the 1220s, and who could have had the pseudo-Ovid as a student. The positional decimal numeration, already advocated in the tenth century by Gerbert d’Aurillac (Bernelin c. 1000 [1], Beaujouan et al. [6], Allard [2], page xv, [4]), gives the calculus of dice a certain versatility and may explain in part this (perhaps) new application of the science of numbers, as it explains other applications of the time to astronomy and mechanics.

There are surely other sources for the calculus of dice that we are completely unaware of. Moreover, in reality we know almost nothing about the history of *De vetula*, and we will say no more about it. Let us rather see what people knew about “seating the dice” in the thirteenth century.

So consider three dice, which we assume to have different colors, blue, white and red, to mark their difference, their individuality. Since the first modern courses

²Translator’s note: The *quadrivium* consisted of arithmetic, geometry, astronomy and music; the *trivium* consisted of grammar, rhetoric and dialectic. These seven liberal arts prepared students for professional study in law, medicine, and the church; see Figure 1.

in the calculus of probability, Borel’s, for example, this has been the simplest way to present the matter. The author of the *De vetula* proposes to calculate in how many ways the points can appear on the three dice. He is careful to point out that there are at least two ways of doing this counting. First, there is what he calls the “punctaturae,” which Le Fèvre translates into French as “pointures,” that is, the marks that appear on the dice without taking into account their color. For example, 225 is the pointure corresponding to two dice coming up 2 and one coming up 5. And so on for the other cases. Punctatura could be translated as “what is going to be pointed or piqued,” or as “configuration of points,” or however you like. For his part, Jacob Bernoulli calls *De vetula*’s punctaturae “modes.” Here, we adopt Le Fèvre’s term, pointure, which is as good as any other and which stands out as a special mathematical term. So the poet proposes first to count the pointures.

This calculation was not new to the West. We know that Wibold of Cambrai published in the tenth century the 56 pointures of three dice by associating them with the canonical virtues, from 111, charity, to 666, humility. This counting became classic, in various forms, in the Middle Ages and the Renaissance. We also know of several counts of pointures in antiquity in the somewhat simpler case of 4-sided astragali.³ Here, there are 35 pointures, named for the most important deities, heroes or animals—Venus, the dog, etc.—which have different divinatory values. In all the known texts, there are no methods of calculation; it is a matter of simple enumeration, say in alphabetical order. In *De vetula*, the pseudo-Ovid proposes a genuine combinatorial calculation of the 56 pointures of three dice, which we will follow step by step. One starts by counting the pointures where the three dice present the same face: 111, 222, etc. There are obviously 6 of them. Then we count the pointures, such as 225, where two dice present the same face and the third a different face. In order to evaluate their number, we first choose the face that is repeated, for which there are six choices, then this choice being made, the single face from the remaining five, which makes a total of thirty, multiplying six by five. Finally, we have to calculate the number

³Translator’s footnote: The astragalus bone, from the ankles of sheep and other animals, falls on one of four sides when thrown. In Greek literature, the four ways of falling were often named with the letters A, Γ, Δ and F or the numbers 1, 3, 4 and 6, but since the four sides are visually distinct, it was not necessary to mark them. See the photographs in [23].

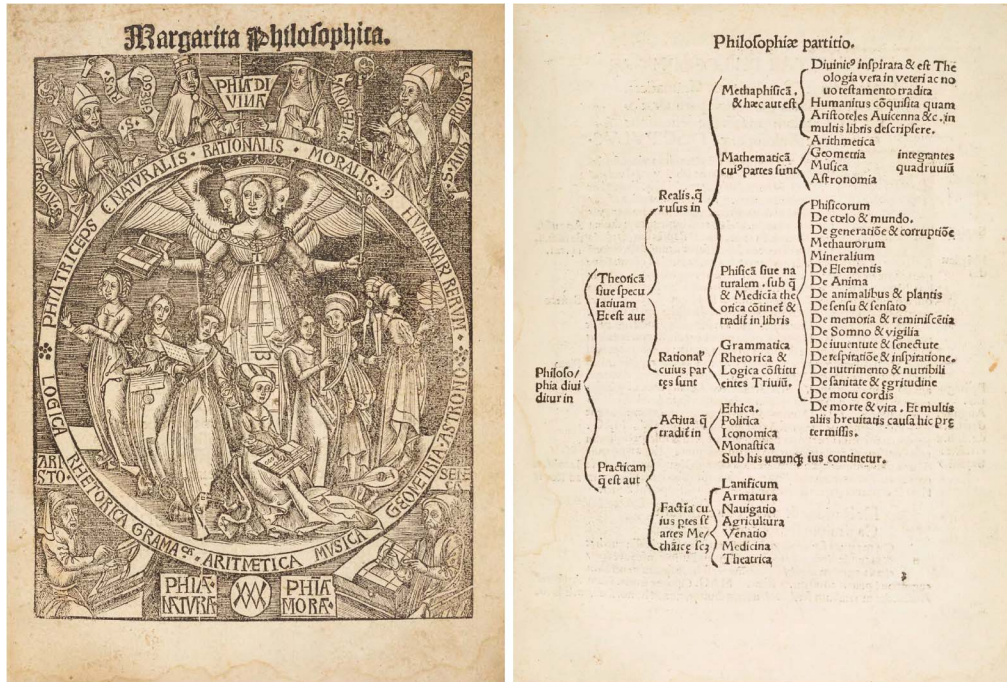


FIG. 1. Cover and third page of *Margarita Philosophica* (*Philosophical Pearl*), reproduced with permission of the Rutgers University Libraries. This sixteenth century compilation, published in 1504 by Gregor Reisch, was a comprehensive university textbook. The seven liberal arts, divided into the trivium and the quadrivium, are listed in the bottom half of the circle on the cover. The diagram on third page shows how the author fits these liberal arts into a larger map of philosophy.

of pointures corresponding to three different faces, for example, 123. We find twenty of them by enumerating the different cases according to whether the numbers marked are “continuous,” for example, 123, or “discontinuous,” for example, 124 or 135. And to be surer of ourselves, we make a table of the 56 pointures, which takes several forms in different manuscripts, culminating in the remarkable “bell-shaped” form found in the second printed edition of 1479 and its later copies (see Table 1).

The reader will have noticed that the second part of the enumeration, which leads to $6 \times 5 = 30$, uses the combinatorial rule of multiplication, which is immemorial (Biggs [10], number 1, Allard [3]) and which the author seems to know well. This rule allows in the same manner the calculation of numbers of arrangements and permutations that is found very early in Chinese, Indian and Arab mathematics (Biggs *ibid.*), in Fibonacci’s or Jordanus’s treatises, and in the collections of problems that formed the pedagogical basis of the teaching of arithmetic at the University of Paris in the thirteenth century (Beaujouan [5], Chapter XI; Allard [3]).

The bell of pointures allows the poet a first comment. The numbers go up to the middle points 10, 11 and then

back down. Thus, the points “are not of equal force.” “The largest and the smallest of them come rarely, the middle points frequently.”

But we still have not reached the crux of the matter. In fact, the middle points from 9 to 12 in this bell all have the same “force” of six pointures, and the extreme points, 3 and 4, or 17 and 18, have an equal force

TABLE 1

666							18
665							17
664	655						16
663	654	555					15
662	653	644	554				14
661	652	643	553	445			13
651	642	633	552	543	444		12
641	632	551	542	533	443		11
631	622	541	532	442	433		10
621	531	522	441	432	333		9
611	521	431	422	332			8
511	421	331	223				7
411	321	222					6
311	221						5
211							4
111							3

of one pointure. Yet the common opinion of players, from the very beginning all the way up to the Chevalier de Méré, who said it to Pascal, is that point 11, for example, comes more frequently than point 12. As for point 4, it is clearly more advantageous than point 3. We must therefore find another explanation. The “force of the pointures” is not sufficient to explain or predict the frequency of the throws.

The pseudo-Ovid now makes this remark:

Now, if we observe more carefully the pointures, there are some for which there is only one way of falling, and there are some for which there are three or six, because the manner of falling is unique when the three numbers of which we have spoken are alike. But if one of these numbers is different two are alike, three cases occur, depending on which is different. And if all are different, you will find that they can change in six ways, because if you give any position to one of the three, the other two exchange their positions, as permutation of the pointures shows. In this way, the fifty-six pointures are differentiated into two hundred sixteen ways of falling, and when they are distributed among the possible number of points for the players as they should be, you will know fully what value or loss any of them can have.

We cannot say it better. The pointure is not enough to specify the way the three dice fall. To the 56 pointures (punctaturae) correspond in reality 216 ways of falling (scemae cadendi) taking into account the color of the dice. For example, the pointure 225 corresponds to 3 ways of falling, depending on whether the 5 is on the white, the red or the blue die. The pseudo-Ovid then introduces a special technical term for the dice’s ways of falling, the *cadentiae*, which we will now translate as *chances*. The poet tells us that the number of chances makes it possible to know the players real gains and losses. In the case of a single die, each face is its own pointure and has the same chance of falling, but in the case of three dice, the chances of different pointures are different; there can be 1, 3 or 6 chances, depending on the color of the dice and, therefore, the points have different chances that one must calculate from the bell of pointures and the chances they can have. This is summarized in Table 2. Thus a total of 216 chances for the 16 possible points from 3 to 18.

TABLE 2

3	18	Punctatura	1	Cadentia	1
4	17	Punctatura	1	Cadentiae	3
5	16	Puncturae	2	Cadentiae	6
6	15	Puncturae	3	Cadentiae	10
7	14	Puncturae	4	Cadentiae	15
8	13	Puncturae	5	Cadentiae	21
9	12	Puncturae	6	Cadentiae	25
10	11	Puncturae	6	Cadentiae	27

The pseudo-Ovid stops there. The problem is completely solved. We can seat three dice. A priori, nothing stops the poet from treating the case of four dice by the same method: enumeration of pointures and then evaluation of their chances taking into account the colors. This in fact is how Jacob Bernoulli first proceeds in the first part of *Ars conjectandi*, but he concludes that this method is both tedious and long (taediosa and prolixa). He proposes another, remarkably ingenious, step-by-step method (ibid., page 24), which we examine later, but which perhaps would not have convinced the pseudo-Ovid.

Just the same, *De vetula* is a genuine “calculus of chances” in a relatively complex case. The pseudo-Ovid constructs the table of chances for the points of three dice, and in two different ways. By doing so, knowing it or not, he introduces a mathematical notion of the first magnitude, a probability law, mathematically mimicking the tables of numbers found in all commercial or agricultural accounts and already on certain Sumerian tablets 4500 years ago (Stigler [31], page 27). Was the Latin poet the first? We do not know. Was he isolated? We do not know this either, but we can doubt it and conjecture that the author of *De vetula* was taking up something he had learned elsewhere, no one knows where. He testifies at least to the existence, at the dawn of Western universities, of a calculus of perfect dice, mathematical dice and the proliferation of manuscripts and editions all the way up to the seventeenth century proves without a doubt that this thirteenth century calculus was not completely forgotten and continued to be taught here and there, down to modern times, without our being able to say more.

Things changed in the fifteenth century in Italy and in the sixteenth century throughout Europe, when mathematics acquired an autonomy and an importance it had lost long before. Luca Pacioli could even write in his dedication of *Divine Proportion* to Duke Sforza:

Among truths, as Aristotle and Averroes affirm, those of mathematics are the most

true: they are the first degree of certainty, and all the other natural sciences come after them. In this way, it is clear, Great and Powerful Duke, that all other sciences are opinions, and only these should be called certitudes. ([29], page 49 of the 1988 edition)

Kant would not say anything different. The medieval calculus of chances thus found a second life. Isolated, strange and fascinating texts inevitably attracted the attention of scholars who henceforth devoted themselves exclusively or almost exclusively to the new sciences. Chairs of mathematics were created, and we may suppose that their holders taught the calculus of chances on occasion, although we know of hardly any texts of any importance on this subject, and those usually quoted were published too late to influence the development of the calculus of dice before Huygens's treatise [25]. This treatise, the first that really counts, entitled *De ratiociniis in ludo aleae* (On calculation in games of chance), was published shortly after the exchange of correspondence between Pascal and Fermat in the summer of 1654, marking the beginning of a new era for the calculus of chances, punctuated by three eternal masterpieces, Bernoulli's *Ars conjectandi*, Moivre's *Doctrine of Chances* and Laplace's *Théorie analytique*. But we limit ourselves here just to dice games and, to stay focused, just to problem of points, posed in its general Ovidian form: given n dice with f faces, calculate the chances of obtaining a sum of points equal to s .

Pascal and Fermat dealt of course with problems of dice, but there are no written traces of their dealing with the problem of points. On the other hand, both gave satisfactory answers to the problem of division: how to divide the stake if the players stop playing. A problem whose origin, like that of the problem of points, goes back very far, but that we do not consider here, for lack of space. We also know that Pascal, Fermat, Leibniz and others developed the theory of combinations in a decisive way in the West. This theory remained the main method for calculating chances before analytic theory arrived in the eighteenth century, thanks to Moivre, Lagrange and Laplace. In large measure, it still remains the main method.

To our knowledge, the first printed trace of the problem of points in the new calculus of chances is found in Christiaan Huygens's [25] treatise, which introduces the calculus of dice as follows:

As far as the dice are concerned, we can propose the following questions: with one die,

in how many throws should we risk throwing the six or some other point? Similarly, in how many throws should we try to throw two sixes with two dice or three sixes with three dice? And many other similar questions.

To solve them, here is what must be considered. First, with one die there are six different outcomes, each of which can appear equally easily. I assume that the die has the perfect shape of a cube. Then, with two dice there are 36 different outcomes, each of which similarly can be obtained equally easily, for in conjunction with each of the outcomes for one die, each of the six outcomes of the other die can happen at the same time, and six times 6 outcomes makes 36 outcomes. Similarly, with three dice there are 216 different outcomes, because in conjunction with each of the 36 outcomes of two dice, each of the six outcomes of the third can happen and six times 36 outcomes makes 216 outcomes. It is clear, in the same way, that with four dice there are six times 216 outcomes, that is to say 1296 outcomes, and that one can thus further calculate any throw of dice, always assuming for an additional die six times as many outcomes.

Here, Huygens makes use only of the multiplication rule, already put to use in *De vetula* and by all the authors who have since addressed this question, not to mention those who used it before the thirteenth century. Huygens next deals with the problem of points for two and three dice, but does not go beyond that. His contribution is elsewhere. In fact, Huygens follows Pascal and the scholastic tradition on the hope of gain, formalizing the notion of expectation, whose central role in the theory we know, but in order to stay within the limits of the problem of points, we do not dwell on this.

3. JACOB BERNOULLI

The first important contribution to the problem of points, which really brought in a new idea that could deal in principle with any number of dice with any number of faces, is due to Jacob Bernoulli in his commentary on Huygens's treatise, which forms the first part of *Ars conjectandi*. We do not know the date of this text, which is certainly between 1685 and 1705.

TABLE 3

1	3	6	10	15	21	25	27	27	25	21	15	10	6	3	1																									
	1	3	6	10	15	21	25	27	27	25	21	15	10	6	3	1																								
		1	3	6	10	15	21	25	27	27	25	21	15	10	6	3	1																							
			1	3	6	10	15	21	25	27	27	25	21	15	10	6	3	1																						
				1	3	6	10	15	21	25	27	27	25	21	15	10	6	3	1																					
					1	3	6	10	15	21	25	27	27	25	21	15	10	6	3	1																				

On the other hand, we know that it was not read outside the Bernoulli family until after the publication of *Ars conjectandi* [8] in September 1713, that is to say after the first contributions of Montmort and Moivre, about which we will say more later.

As mentioned above, Jacob Bernoulli starts by counting the pointures and the chances in the case of four dice. But he then proposes a method that is both clear and easy, summarized in a table that allows the calculation of the chances of points in general, step by step. Bernoulli does not tell the principle of his method, the reader (especially his brother Johann) being supposed to find it for himself if he can. This principle is very simple and very general, as we will see.

To simplify the presentation, we consider n six-sided dice and we assume that we have succeeded in calculating the numbers of chances for all the points in this case (one can suppose that $n = 1$ or 2 ; this does not change the reasoning). Bernoulli proposes to show how we deduce from this the numbers of chances for the points for $n + 1$ dice.

For n dice, the points vary from n to $6n$. Let their respective numbers of chances be c_k , for $n \leq k \leq 6n$, and to make the formulas simpler to write, set $c_k = 0$ for all other integer values of k . Let us throw an extra die and try to evaluate the number of chances C_k for the sum of the faces of the $n + 1$ dice to be equal to k . There are six possibilities: the additional die falls $1, 2, \dots, 6$. If it falls i , the number of chances for the sum of the faces of the $n + 1$ dice to be equal to k is equal to c_{k-i} . Whence Bernoulli's formula:

$$(c) \quad C_k = \sum_{i=1}^6 c_{k-i}.$$

Formula (c) is what we now call a convolution formula. Suitably generalized, it allows us to calculate the probability law of a sum of independent random variables in absolute generality. But Bernoulli does not know this and he prefers to adopt an arrangement in a table starting from $n = 1$ and going step by step as far as we want.

A single die's chances are all equal to 1; we represent this with a line of 1's:

$$1\ 1\ 1\ 1\ 1\ 1.$$

Then we copy the same line five times, shifting it each time one place to the right, and we add the columns thus constituted. We obtain, by formula (c), the line of chances for the points of two dice:

$$1\ 2\ 3\ 4\ 5\ 6\ 5\ 4\ 3\ 2\ 1.$$

Then we repeat with this new line what we had done previously. That is, we write the line five times, shifting each time one unit to the right, and then we add the columns. According to the same formula (c), we obtain the chances of the points for three dice, those of *De vetula*. Let us follow the algorithm for four dice. Starting with the chances given by *De vetula* for three dice, we form Table 3.

Adding the columns gives, by formula (c), the sequence of chances for four dice:

$$1\ 4\ 10\ 20\ 35\ 56\ 80\ 104\ 125\ 140\ 146\ 140\ 125$$

$$104\ 80\ 56\ 35\ 20\ 10\ 4\ 1.$$

We can continue this way indefinitely. Bernoulli stopped with six dice to show his merit without excess of zeal. One had finally passed *De vetula*. It had taken 450 years or so.

Bernoulli's algorithm applies without change to dice with an arbitrary number of faces, but it does not give a complete formula for relatively large numbers of dice. Even if the step-by-step calculation is sure to give the answer in the end, you will be hard pressed to carry it out without error in a reasonable time when the value of n is, say, 10. And what if $n = 100$ or 1000? Nevertheless, the simplicity and elegance of the algorithm could not have failed to attract the attention of Bernoulli's nephew, Nicolas I Bernoulli, the first mathematician to have access to the manuscript of *Ars conjectandi*, when writing his famous thesis on law [9]. We may suppose that it helped attract young Nicolas to the doctrine of

chances, of which he became one of the best specialists. It is also possible that he informed Montmort quite early of his uncle's algorithm, but we have no proof of this.

4. PIERRE RÉMOND DE MONTMORT

At the beginning of the eighteenth century, all the scholars of Europe knew that the great Jacob Bernoulli had written a fundamental work on the art of conjecturing. If one of the top mathematicians of Europe, one of the masters of the new differential and integral calculus, had been interested in the theory of chance and its applications to civil, moral and economic affairs, this was a serious matter that must not be allowed to drop. So while waiting for the repeatedly delayed publication of Bernoulli's great treatise, several hastened to mark their territory in a field still virgin or nearly so. This was the case with Pierre Rémond de Montmort, who after briefly being a canon in the chapter of Notre Dame of Paris, married and was now devoting himself to the science of numbers in his castle of Montmort in Champagne. Like Father Prestet, he was from Malebranche's entourage—a man of the world and a scholar. His prefaces are remarkably written. Montmort intended to write a history of mathematics, which would have been the first in the West. He was well along in his undertaking when death overtook him.

Montmort was also one of the first Western “combinatorists,” along with Nicolas Bernoulli and Abraham de Moivre. He had that rare gift of counting configurations that seem impossible to disentangle. As he confided to Nicolas Bernoulli, you must have a special talent for it.

It seems to me that these demonstrations are not like demonstrations in geometry. Those touching on numbers and combinations are infinitely more difficult. You can have them very clearly in mind without being able to put them on paper.

In 1708, Montmort published under cover of anonymity a remarkable *Essay d'analyse sur les jeux de hazard* [20] (Essay analyzing games of chance).⁴ After Huygens's treatise, it was the first published work on

the new calculus of chances. In the Introduction, the author informs us that many of his friends “had long urged him to try to see if Algebra could not manage to determine the Banker's advantage in the game of Pharaon.” Informed about the potential work of Jacob Bernoulli, he took up the question again and had the good fortune to solve it and several others of the same kind “or even more difficult.” Montmort did not stop there. Far from the skepticism of the pseudo-Ovid or Leibniz, he made himself the advocate of a doctrine of probabilities applied to “the things of life,” the tones of which resembled those (yet to appear) of Jacob Bernoulli, or even of Condorcet a century later:

We cannot know the future, but in Games of chance, and often in other things of life, we can always know exactly how much more likely a certain thing is to happen in one way rather than in any other! And since these are the limits of our knowledge, we should at least try to reach them.

As for the problem of points, Montmort contented himself with giving a table of chances of the points in the case of 2 to 9 dice, without any indication of methods. The complexity of the calculation for more than 5 dice suggests, however, that the scholar in Champagne had a valid formula for all cases, or at least a simple algorithm, even though he said nothing about it. He did not publish his formula and his demonstration until in the second edition of the *Essay*, in 1713. Meanwhile, a mathematician of the first rank, Abraham de Moivre, a native Huguenot of Vitry-le-François exiled in London, published the formula in his memoir [18], the first version, in Latin, of his great treatise *The Doctrine of Chances* [19], which would have three English editions, continually supplemented and clarified, in 1718, 1738 and 1756. These alternating publications did not fail to arouse polemics on the true paternity of the formula and its combinatorial demonstration, which Moivre never published, but which could have differed little from Montmort's.

Without entering into too many details, we must therefore examine this demonstration, which is not without interest insofar as it makes use of two of the most famous combinatorial formulas, apparently used here for the first time in their full generality: the formula for combinations with repetition and the sieve formula. We follow as closely as possible the plan of demonstration in the second edition of Montmort's *Essay*, but not its form, which uses an arithmetic triangle without enough notation to treat the generality of

⁴Translator's note: Montmort's name did not appear in either of the book's two editions; see [24], page 287. But while being prudent with respect to political and ecclesiastical authority, Montmort was proud of his work and made sure his authorship was well known among mathematicians. Newton was among those to whom he sent copies; see [7], page 68.

the proposition in question. This is a difficulty encountered in all combinatorial demonstrations of the eighteenth century (and earlier centuries), where you treat the case of small values of the variables and pass to the general case informally, “by induction,” for lack of notation adapted to a real proof. For ease of reading, we have chosen here a middle way. We follow Montmort’s plan with modernized notation.

In the general case, we have n dice with f faces. To simplify the formulas a little, following a tradition that goes back at least to Moivre ([18], pages 220–221; [19], page 17 of the first edition, pages 35–36 of the second edition, and pages 39–40 of the third edition), we assume that the faces are numbered from 0 to $f - 1$. This does not change the chances, but the values of the points are diminished by n . We seek the number of chances that the sum of the faces of the f dice is equal to s , for s between 0 and $n(f - 1)$.

Write x_i for the result of the i th die, for $1 \leq i \leq n$. So it is a matter of finding the number of integer solutions of equation

$$(m) \quad x_1 + x_2 + \dots + x_n = s$$

such that $0 \leq x_i \leq f - 1$ for all i .

First assume that $s < f$.

In the case being considered, each of the variables x_i satisfying equation (m) is necessarily less than or equal to s and, therefore, strictly smaller than f . Hence the number of chances of obtaining the sum s is equal to the number of nonnegative integer solutions of equation (m).

We have already encountered this problem. The answer is given by the formula of combinations with repetition of s from n things, that is to say⁵

$$(0) \quad C_{s+n-1}^s$$

Suppose now that $f \leq s < 2f$.

In this case, there may be solutions of equation (m) containing variables greater than or equal to f , which cannot therefore correspond to the faces of the dice that we are considering (which are numbered from 0 to $f - 1$). So we must exclude such solutions from the count (0) made previously. Since $x_1 + x_2 + \dots + x_n = s < 2f$, no more than one of the variables x_i can be greater than or equal to f . Say that it is the first. Then $x_1 = f + t_1$, with $0 \leq t_1 < f$, and consequently

$$t_1 + x_2 + \dots + x_n = s - f,$$

all the variables being nonnegative integers and strictly less than f . The number of such solutions is obtained from the preceding formula by replacing s with $s - f$: $C_{s-f+n-1}^{s-f}$. Since there are n possibilities for the variable exceeding f , the number of chances of the point s when $f \leq s < 2f$ is

$$(1) \quad C_{s+n-1}^s - nC_{s-f+n-1}^{s-f}.$$

Next, suppose that $2f \leq s < 3f$, and consider again equation (m). Now it is obvious that at most two of the variables can exceed f . Suppose the first two exceed f . In this case, the subtraction in (1) removes twice the number of corresponding solutions, that is, the number of solutions of equation (m) satisfying $x_1 \geq f, x_2 \geq f$ and $0 \leq x_k \leq f - 1$ for k different from 1 and 2. To compensate for this excessive subtraction, we add this number back. But this number is also given by formula (0) if we set $x_1 = f + t_1$ and $x_2 = f + t_2$ with $0 \leq t_i < f$ for $i = 1$ and 2. We have

$$t_1 + t_2 + x_3 + \dots + x_n = s - 2f,$$

all variables being integers and strictly less than f . The number of these solutions is given by $C_{s-2f+n-1}^{s-2f}$. There are C_n^2 choices of two indices from n . So the number of chances for s when $2f \leq s < 3f$ is equal to

$$(2) \quad C_{s+n-1}^s - nC_{s-f+n-1}^{s-f} + C_n^2 C_{s-2f+n-1}^{s-2f}.$$

Now if $3f \leq s < 4f$ we should subtract as before the solutions containing three variables greater than f . The number of chances for s in this case is obviously

$$(3) \quad C_{s+n-1}^s - nC_{s-f+n-1}^{s-f} + C_n^2 C_{s-2f+n-1}^{s-2f} - C_n^3 C_{s-3f+n-1}^{s-3f}.$$

And so on. If $kf \leq s < (k + 1)f$, the number of chances for the sum s will be

$$(k) \quad C_{s+n-1}^s - nC_{s-f+n-1}^{s-f} + C_n^2 C_{s-2f+n-1}^{s-2f} - \dots + (-1)^k C_n^k C_{s-kf+n-1}^{s-kf}$$

for every integer k strictly less than $n - 1$.

Here is what Montmort, Moivre and the others preferred to write: the number of chances for the point s with n dice that have f faces starting with 0 is equal to

$$\frac{(s + 1)(s + 2) \dots (s + n - 1)}{(n - 1)!} - n \frac{(s - f + 1) \dots (s - f + n - 1)}{(n - 1)!} + \frac{n(n - 1)}{2} \frac{(s - 2f + 1) \dots (s - 2f + n - 1)}{(n - 1)!} - \dots,$$

⁵Translator’s note for readers accustomed to other notation:

$$C_n^k = \frac{n!}{k!(n - k)!}.$$

where the alternating sum stops as soon as one of the factors becomes zero or negative.

If we want probabilities instead of chances, we must divide this expression by the total number of chances for n dice with f faces, that is to say f^n . This gives

$$\begin{aligned} & \frac{(s+1)(s+2)\cdots(s+n-1)}{(n-1)!f^n} \\ & - n \frac{(s-f+1)\cdots(s-f+n-1)}{(n-1)!f^n} \\ & + \frac{n(n-1)}{2} \frac{(s-2f+1)\cdots(s-2f+n-1)}{(n-1)!f^n} \\ & - \dots, \end{aligned}$$

again stopping the sum as soon as one of the factors becomes zero or negative.

The general problem of points is thus completely solved. Except that as soon as the number of dice exceeds 10, the calculations, involving $n!$ and up to n alternating terms, become impractical. This problem stumped mathematicians for a long time. It was not until the beginning of the 1810s that Laplace resolved it in a masterful way.⁶

5. LAPLACE

In 1812, in the *Théorie analytique des probabilités* (Book II, Section 18) [27], Laplace gives the first known approximation to the Montmort–Moivre formula. Let us state it in the case where $f = 2b + 1$, so that the mean point is equal to nb . The probability of a sum $s = nb + k$ is approximately equal to

$$(dvl) \quad \sqrt{\frac{6}{\pi n(f^2 - 1)}} \exp\left(-\frac{6k^2}{n(f^2 - 1)}\right),$$

which traces *De vetula*’s bell of chances as closely as possible.

We will not give details in this chapter of the calculations leading to (dvl).⁷ But let us briefly consider the case where the number of faces on the dice becomes infinitely large, the number of dice remaining reasonable. This seemingly esoteric problem has a very great importance in the history of the probabilistic theory of errors, and it will not detain us for very long.

The first mathematician to consider this case was Thomas Simpson, a disciple of Moivre and soon his main competitor.

We can picture a die with f sides as a lottery wheel with f equal sectors or more simply as the segment $[0, 1]$ divided into f equal parts, the faces of the die, which are supposed to have all the same chances of being selected. If f becomes infinitely large, the faces of length $1/f$ become infinitely small and equal to dx . A face is located by the sum of infinitely many such intervals dx , taken as a unit, or by a number x between 0 and 1. We can thus see the choice of one of this infinity of faces as the random choice of a number x between 0 and 1, or as was generally said in the second half of the eighteenth century, as an “inevitable error” whose “law of facility” is constant over the interval $(0, 1)$.

The sum s of the faces of the n dice becomes infinitely large of the same order as f . So we set $1/f = dx$ and $s/f = x$, for x positive and less than n . The above formula for the probabilities becomes

$$(SL) \quad \begin{aligned} & \frac{dx}{(n-1)!} \left[x^{n-1} - n(x-1)^{n-1} \right. \\ & \left. + \frac{n(n-1)}{2}(x-2)^{n-1} \right. \\ & \left. - \frac{n(n-1)(n-2)}{2 \cdot 3}(x-3)^{n-1} + \dots \right], \end{aligned}$$

stopped as soon as one of the factors becomes zero.

The formula (SL) gives the “chance” that the sum of the n errors is between x and $x + dx$. Without the dx , it represents, in current terminology, the density of the sum of n independent random variables following the same uniform law on $(0, 1)$. It is the “facility” of the sum of n equally possible errors between 0 and 1.

We can plot the curves of these facilities for small values of n . They very quickly take the form of the Laplacian or Gaussian bells.

For $n = 2$, the curve of facility is an isosceles triangle placed on the x -axis. For $n = 3$, it is a bell formed of three pieces of parabola, and if we compare the curve corresponding to $n = 4$ with the normal law with mean 2 and variance $1/3$ [mean and variance of the sum of 4 independent variables uniformly distributed on $(0, 1)$], we observe an almost perfect superposition, which very quickly becomes an identity for larger values of n . This surprising result is a special case of the fundamental theorem proven by Laplace in 1810. The law of the sum of n independent variables with nearly any density is very close to a normal distribution with the same mean and variance, and this theorem holds

⁶Efforts by Laplace and others before 1810 are discussed by the Brus in later chapters of their book and by Anders Hald in *A History of Probability & Statistics and Their Applications before 1750* [24].

⁷Laplace’s work on this topic is discussed in great detail in other chapters of the authors’ book.

just as well for *De vetula*'s dice with a finite number of faces, as we saw above. Since the normal law has been tabulated since the beginning of the nineteenth century, we have thus the numerical solution of all possible and imaginable problems of points, and we stop there.

To conclude, let us return to the formula (dvl) above. It says that a complicated combinatorial alternating sum is nearly proportional to a simple exponential. Such an approximation was not at all obvious a priori. It took decades for it to appear as one of the basic formulas in the probability calculus. It is well known that it first emerged in the particular case of two-sided dice, coins marked 0 and 1. The sum s of the faces is then simply the number of 1's obtained (or the number of heads if 1 corresponds to heads and 0 to tails). It is no wonder that this particular case was treated first. It is one of the most remarkable results obtained by Moivre in 1733, published at the end of the second edition of his *Doctrine of Chances* in 1738 (pages 243 et seq. of the third edition). Let us recall his result in contemporary notation.

Suppose $n = 2m$ symmetrical coins marked 0 and 1 are thrown together. We denote by s the sum of the faces obtained, or the number of 1's. Since at least Pascal or Jacob Bernoulli it has been known that exactly C_n^s of the 2^n possible chances have the number of 1's equal to s . So Laplace's approximation formula follows directly in this case from Stirling's formula (e.g., Borel [12], No. 21):

$$(ms) \quad C_{2m}^s \frac{1}{2^{2m}} \approx \frac{1}{\sqrt{\pi m}} \exp\left(-\frac{(s-m)^2}{m}\right).$$

From this famous formula was born the "Gaussian curve," the analytical figure that symbolizes *De vetula*'s table of cadentiae. Laplace went from formula (ms) to formula (dvl) in his attempt to solve the problem of points for dice with any number f of faces, a problem that stumped him for almost 40 years and whose solution also gave him one of the main keys for applying his theory of probabilities to natural philosophy.

REFERENCES

- [1] BERNELIN (STUDENT OF GERBERT D'AURILLAC C. 1000) (1999). *Libre d'abaque*. Princi Néguer, Pau. Latin and French text established by B. Bakhouché with complementary notes by J. Cassinet.
- [2] ALLARD, A. (1992). *Muhammad Ibn Musa Al-Khwarizmi. Le Calcul Indien (Algorismus)*. A. Blanchard, Paris. Edition, with French translation and commentary, of the oldest Latin versions going back to the XIIIth century.
- [3] ALLARD, A. (1994). *L'enseignement du calcul arithmétique à partir des XIIe et XIIIe siècles: L'exemple de la multiplication*. Publications de l'Institut d'Etudes Médiévales, Université Catholique de Louvain, No. 16.
- [4] ALLARD, A. (1997). L'influence des mathématiques arabes dans l'occident médiéval. In *Histoire des Sciences Arabes* (R. Rashed and R. Morelon, eds.) 2 199–230. Seuil, Paris.
- [5] BEAUJOUAN, G. (1991). *Par Raison de Nombres, L'art du Calcul et des Savoirs Scientifiques Médiévaux*. Gower, Aldershot.
- [6] BEAUJOUAN, G., CARBONNE, P., CASSINET, J. et al. (1995). *Huit Siècles de Mathématiques en Occitanie*. CIHSO, Toulouse. Reprinted by Monein, Pyrémoude, 2008.
- [7] BELLHOUSE, D. R. (2011). *Abraham De Moivre: Setting the Stage for Classical Probability and Its Applications*. CRC, Boca Raton, FL.
- [8] BERNOULLI, J. (1713). *Ars conjectandi, opus posthumum. Accedit Tractatus de seriebus infinitis et epistola Gallice scripta De ludo pilae reticularis*. Impensis Thurnisiorum Fratrum, Basel. French translation by J. Peyroux, A. Blanchard, Paris, 1998. English translation with notes by E. D. Sylla, Johns Hopkins Univ. Press, Baltimore, 2006.
- [9] BERNOULLI, N. (1709). De usu artis conjectandi in jure. Doctoral thesis, Basel. Reprinted in *Die Werke von Jakob Bernoulli* 3 (1975) 287–326. French translation with notes by N. Meusnier, Paris, CAMS, 1992.
- [10] BIGGS, N. L. (1979). The roots of combinatorics. *Historia Mathematica* 6 109–136. DOI:10.1016/0315-0860(79)90074-0.
- [11] BOETHIUS, A. M. S. (1995). *De Institutione Arithmetica*. Les Belles Lettres, Paris. French translation and notes by J.-Y. Guillaumin.
- [12] BOREL, É. (1909). *Éléments de la Théorie des Probabilités*. Hermann, Paris. Further editions in 1910 and 1924, and 1950. The 1950 edition, published by Albin Michel, was translated into English by J. E. Freund and published by Prentice Hall, Englewood Cliffs, 1965.
- [13] BOREL, É. (1920). Radioactivité, probabilité et déterminisme. *Rev. Mois* 21 33–40. Reproduced as note I of the 1924 edition of [12] and in volume 4 of [14], pages 2189–2196.
- [14] BOREL, É. (1972). *Œuvres D'Émile Borel*. CNRS, Paris. 4 volumes.
- [15] COURNOT, A. (1828). De la théorie des probabilités considérée comme la matière d'un enseignement. *Le Lycée* 2 447–453. Reprinted in Volume 11 of [17].
- [16] COURNOT, A. (1843). *Exposition de la Théorie des Chances et des Probabilités*. Hachette, Paris. Reprinted with notes and index as Volume I of [17].
- [17] COURNOT, A. (1973–2010). *Œuvres Complètes*. Vrin and Presses Universitaires de Franche-Comté, Paris and Besançon. 11 titles in 13 volumes.
- [18] DE MOIVRE, A. (1710–1712) De mensura sortis, seu de probabilitate eventuum in ludis a casu fortuito pendentibus. *Philos. Trans. R. Soc. Lond.* 27 213–264.
- [19] DE MOIVRE, A. (1718). *The Doctrine of Chances: Or, A Method of Calculating the Probability of Events in Play*. W. Pearson, London. Later editions in 1738 and 1756.
- [20] DE MONTMORT, P. R. (1708). *Essay d'analyse sur les jeux de hazard*. Quillau, Paris. Published anonymously. Second edition in 1713.

- [21] DJEBBAR, A. (1985). *L'analyse combinatoire au Maghreb: L'exemple d'Ibn Mun'im (XIIe–XIIIe s.)*. Publications Mathématiques D'Orsay, number 85-01.
- [22] GUILLAUMIN, J.-Y. (2012). Boethius's *De institutione arithmetica* and its influence on posterity. In *A Companion to Boethius in the Middle Ages* (N. H. Kaylor and P. E. Phillips, eds.) 135–162. Brill, Leiden.
- [23] HAGSTROEM, K. G. (1932). *Les préludes antiques de la théorie des probabilités*. Fritze, Stockholm.
- [24] HALD, A. (1990). *A History of Probability and Statistics and Their Applications Before 1750*. Wiley, New York.
- [25] HUYGENS, C. (1657). De ratiociniis in ludo aleae. In *Exercitationum Mathematicarum, Liber V* (F. van Schooten, ed.) 511–534. Elsevier, Leiden. The Dutch original, written in 1656, was published in Amsterdam in 1660 and reprinted along with a translation into French in Huygens's *Œuvres*, volume 14, pages 1–91.
- [26] LAPLACE, P.-S. (1878–1912). *Œuvres complètes*. Gauthier-Villars, Paris. 14 volumes.
- [27] LAPLACE, P. S. (1812). *Théorie Analytique des Probabilités*. Courcier, Paris. Later editions in 1814, 1820, and 1825. The 1825 edition was reprinted in Volume 7 of [26].
- [28] MONTUCLA, J.-É. (1802). *Histoire des mathématiques*. H. Agasse, Paris. Completed and published by J. de Lalande, 4 volumes.
- [29] PACIOLI, L. (1498). *Divina Proportione*. Manuscript, Milan. First published by Paganini, Venice, 1509. Facsimile and French translation by G. Duchesne and M. Giraud, with a historical introduction by M.-T. Sarrade, Librairie du Compagnonnage, Paris, 1988.
- [30] STIGLER, S. M. (2014). Soft question, hard answer: Jacob Bernoulli's probability in historical context. *Int. Stat. Rev.* **82** 1–16.
- [31] STIGLER, S. M. (2016). *The Seven Pillars of Statistical Wisdom*. Harvard Univ. Press, Cambridge, MA.