CARMEN BATANERO, MICHEL HENRY, BERNARD PARZYSZ

THE NATURE OF CHANCE AND PROBABILITY

Chapter 1

> The theory of probabilities is at bottom nothing but common sense reduced to calculus; it enables us to appreciate with exactness that which accurate minds feel with a sort of instinct for which ofttimes they are unable to account... It teaches us to avoid the illusions which often mislead us; ... there is no science more worthy of our contemplations nor a more useful one for admission to our system of public education. (Laplace, 1986/1825, pp. 206-207)

1. INTRODUCTION

Epistemological problems play a fundamental role for mathematics educators, because analyzing the obstacles that have historically emerged in the formation of concepts can help us understand students' difficulties in learning mathematics. This is particularly important in the field of probability, where, in addition to the difficulty of understanding scientific knowledge as a theoretical interpretation of real phenomena, one has to deal with typical misconceptions and beliefs, and knowledge about future events that is often based on divinatory predictions that have arisen from a magical ancestral way of thinking. For centuries all speculation about future events was inconceivable, since the future only belonged to the omniscient and omnipotent glory of the supreme Creator as noted by Jacques Bernoulli (1713/1987) in introducing the fourth part of *Ars Conjectandi.* Mind you, this divine association was not an obstacle for players betting on games of chance; however, the quantitative control of these bets remained in the field of intuition.

Gerolamo Cardano, who connected betting to the enumeration of winning combinations, was the first to make progress in probabilistic thinking in the 16th century. However, the decisive step in probability thought was achieved by Blaise Pascal and Pierre de Fermat in their correspondence (Pascal 1654/1963a), and was exposed by Pascal in his *Traité du Triangle Arithmétique* (Pascal 1654/1963b, Edwards 1987). Ignoring metaphysics,

Graham A. Jones (ed.), Exploring probability in school: Challenges for teaching and learning, 15—37. ©2005

Pascal and de Fermat quantified the winning chances for the players in the case when a game actually stops before one of them wins the prize and where equal probabilities were not appropriate. The assumption of equiprobability of the elementary outcomes in a fair game was the first criterion to estimate the probability of a compound event made up of these outcomes.

Since then, the concept of probability has received different interpretations according to the metaphysical component of people's relationships with reality (Hacking, 1975) and thus probability is a young area where formal development has been linked to a large number of paradoxes that show the disparity between intuition and conceptual development in this field (Borovcnik, Bentz, & Kapadia, 1991). For example, many students think that the events "obtaining 5 and 6" and "6 is obtained twice" are equally likely when throwing two dice. Other examples are given in Székely (1986) and through this chapter (e.g. the first historical probability problem posed to Galileo by the Grand Duke of Tuscany).

Even today, and in spite of having a satisfactory axiomatic system, there are still controversies over the interpretation of basic concepts and about their impact on the practice of statistics. Moreover, Borovcnik and Peard (1996, p. 249) remark that probabilistic reasoning is different from logical or causal reasoning and thus counterintuitive results in probability are found even at very elementary levels. This is in contrast with other branches of mathematics where counterintuitive results are encountered only when working at a high degree of abstraction. This fact explains the existence of erroneous intuitions and learning difficulties that still persist at the high school level (Batanero, Serrano & Green, 1998; Batanero & Sanchez, this volume; Fischbein, Nello, Marino, 1991; Jones & Thornton, this volume; Langrall & Mooney; this volume Shaughnessy, 1983, 1992; Watson, this volume). A well-known example is the following: when successive players try to pick at random the shortest stick among a set of sticks it is argued that the first player has the greatest probability to get the shortest stick, because successive players might be unable to get it. The fact that the probabilities are equal for all players in this example is contrary to naive probabilistic intuition.

In this chapter we will examine different interpretations of the nature of chance, randomness, and probability and will highlight how these multiple conceptions are complementary and can influence curriculum goals. Finally we include some implications for the teaching and learning of probability in schools.

2. CHANCE AND CAUSALITY

Different Perceptions of Chance

When a human being achieves by his/her intelligence a certain perception of the world and time, he or she is confronted with the fundamental uncertainty of future events. The idea of chance is as old as civilization, but there are many different conceptions of this idea which have spanned the history of thinking up to the present day. Various explanations have tried to describe our uncertainty of future events, in particular:

- Believing in a destiny predetermined by a supernatural spirit or God;
- Assuming a personal chance factor, unequal for different individuals;
- Accepting natural necessity, ineluctably subjected to laws which still are partially unknown and which govern the world's origin and evolution;
- Arguing the inextricable complexity of the infinitesimal causes generating macroscopic phenomena, which we consider fortuitous as the only possible reasonable interpretation;
- Assuming the existence of a fundamental, chaotic and absolute natural randomness.

It is then easy to understand the difficulty of giving a definition of randomness, which includes all these interpretations. Bennett (1993) and Courtebras (2001) analyzed different historical, philosophical and psychological conceptions of chance. Some of these conceptions still appear in children and naive minds (see Langrall & Mooney; Watson; in this volume; the latter for an analysis of the role of chance language). However, continuous progress of scientific knowledge and education about rational thinking produced an evolution of dominantviews about random phenomena. A brief historical journey will be useful to appreciate this evolution.

Chance, Causality and Determinism in History

Chance mechanisms, such as cubic dice, or astragali (bones from the ankle of animals with hooves, such as the sheep) have existed since the first Sumerian, Assyrian and Egyptian civilizations, and were used to predict the future and to engage in decision-making. Games of chance were so widespread in ancient Rome as to be an object of regulation (Hacking, 1975, p. 25). However, a scientific idea of randomness was absent in the first exploratory historical phase, which extended according to Bennett (1993), from antiquity until the beginning of the Middle Ages when chance was conceived as fortune and related to causality. Within this framework, the Greek philosophers developed various points of view. For Democritus, everything on earth is the combined fruit of chance and need. Aristotle considered that chance results from the unexpected but remarkable coincidence of two or more series of events, independent of each other and due to so many different factors that the eventual result is pure chance.

Aristotle's philosophy pervaded the Middle Ages where magic thinking and superstitions were frequent in the mind of the layperson. The Renaissance progressively gave way to a deist determinism that Denis Diderot summarized in this expression: "It is written up there" (Diderot, 1796/1983). This conception was particularly well expressed by Jacques Bernoulli in introducing the fourth part of *Ars Conjectandi*: "All which benefits under the sun from past, present or future, being or becoming, enjoys itself an objective and total certainty... since if all what is future would not arrive with certainty, we cannot see how the supreme Creator could preserve the whole glory of his omniscience and omnipotence." (Bernoulli, 1713/1987, p. 14).

One hundred years later, Pierre Simon Laplace based his deterministic thinking on the "principle of sufficient reason," by virtue of which Leibniz denounced the "blind chance of epicureans" (Leibnitz, 1710/1969). After this reference, Laplace writes in his *Essai Philosophique sur les Probabilités*: "Present events are connected with preceding ones by a link based upon the evident principle that a thing cannot occur without a cause which produces it". Laplace goes on to present his point of view in a shocking formula: "We ought then to regard the present state of the universe as the effect of its anterior state and as the cause of the one which is to follow" (Laplace, 1814/1995, p. vi).

From this viewpoint, chance is only the "expression of our ignorance." Laplacian determinism was radical and dominated scientific thinking until the 19th century. It allowed no place either to natural chance, intrinsic in some situations, or to the "secondary causes" contingency, that is to give people the freedom to choose and decide. This position obviously challenged philosophers and was lengthily discussed by them; it also questioned scientists' rapport with the real world (Thom, 1990). Does determinism translate a conception of nature to its reality? Or should we understand it as a theoretical postulate about the uniqueness of evolution in an idealized world, which is represented by mathematical models whose equations integrate the formulations of the admitted laws and assumptions?

At the beginning of 20th century, Henri Poincaré remarked that ignorance of the laws governing certain natural phenomena did not necessarily involve a chance interpretation. Moreover, he noticed that, for the laws of perfect gases and Brownian motion, the regularity of macroscopic phenomena can be translated to deterministic laws, even when these phenomena are primarily random at the microscopic level. These remarks led Poincaré to declare in his *Calcul des Probabilités*: "Is it thus necessary that chance be different from the name we give to our ignorance?" (Poincaré, 1912/1987, p. 3). Poincaré gives then the following definition:

> a very small cause, which escapes us, determines a considerable effect that we cannot fail to see, and then we say that this effect is due to chance, ... it might happen that small differences in the initial conditions produce very large ones in the final phenomena... Prediction becomes impossible and we have the fortuitous phenomenon. (Poincaré, 1912/1987, p. 4-5).

Determinism remained however, impossible to circumvent for the majority of scientists in the 20th century: "God does not play dice" according to Einstein's formulation, taken again by René Thom who claims (Thom, 1986, p. 24): "in this conflict determinism-chance, Science is deterministic by reasons of principle."

The Concept of Chance in the 20th Century

In contemporary science we wonder about the existence of fundamental chance in natural phenomena, and about the possible degree of accuracy in its observations. Werner Heisenberg's uncertainty principle in quantum mechanics implies that a particle's movement can only be described by random functions and it is theoretically impossible to deterministically fix at the same time its position and speed (Kojeve, 1932/1990). The existence of intrinsic chance was accepted and developed in genetics by Jacques Monod (1970), then in thermodynamics by Ilya Prigogine and Isabelle Stengers (1979). Epistemologists such as Edgar Morin (1990) elaborated the founding concept of complexity, allowing a thorough enlightening of the chance notion, as predicted by Poincaré. Contemporary writings about chaos, determinism, chance and complexity are now very numerous. Mathematicians such as David Ruelle (1991) developed chaos theory to model complex phenomena, thus contributing to a better understanding of these phenomena.

Whatever our philosophical conceptions of chance and necessity and our epistemological conceptions of probabilities are, they are compatible with the contemporary mathematical theory of probability. In developing an axiomatic theory that was adequate to support these different interpretations, mathematics does not enter these philosophical or epistemological debates.

Random experiment is thus a primitive mathematical concept and whatever be the nature of chance in each particular random experiment, we can give probabilities for the different events, just by applying probability models, which fulfil the axioms of probability theory. But the teacher of probability needs to be aware of these interpretations, because they implicitly determine students' behaviors and answers when confronted with chance situations or when having to put their probabilistic intuitions and knowledge in practice.

3. RANDOMNESS AND PROBABILITY

First Steps

Bellhouse (2000) analyzed a 13th century manuscript, *De Vetula*, attributed to Richard de Fournival (1201-1260). In this manuscript a long epic poem is transcribed. One of its passages describes a dice game, where the players should bet on the sum of points obtained with three dice. This poem is the oldest known text establishing the link between observed frequencies and the enumeration of possible configurations: "Sixteen compound numbers are produced. They are not, however, of equal value, since the larger and the smaller of them come rarely and the middle ones frequently" (Bellhouse, 2000, p. 134).

By counting the 216 "ways of failing" (216 arrangements of the three dice), which produce 56 "observable configurations of points," the author of the poem connects each of the 16 different sums to its corresponding number of "ways of failing," achieving thus an implicit determination of their probabilities. He then advises the players to organize their bets according to their expected profit: "you will learn full well how great a gain or a loss any one of them is able to be" (Bellhouse, 2000, p. 135).

The author thus claims to be able to quantify the chances of an event to come. Let us notice that this same game, betting on the sum of three dice, motivated the Grand Duke of Tuscany to pose to Galileo (about 1620) the first known and solved probability problem in history. Although there are exactly 6 different configurations which produce either the sums 9 and 12 or the sums 10 and 11, and therefore they should be expected to have the same frequency, the observation of a long series of trials made players prefer 10 and 11 to 9 and 12. To explain this paradox, Galileo, took into account the order of number in the three dice, and gave a complete combinatorial proof of the right solution: 25 different ordered configurations for the sums 9 and 12 and 27 possible configurations for the sums 10 and 11.

During the 16th century Cardano explicitly suggested using the relative weight of favorable outcomes in a chance game to make a fair bet. In his *Liber de Ludo Aleae*, Cardano (1663/1961) advised players of the basic role of combinatorial calculation, and gave them a general rule: consider all the possibilities which represent the number of ways the favorable results can occur, and compare this number with the remainder. The mutual bet should be posed according to this proportion, so that the players can compete on equal terms (Pichard, 2001, p. 17).

Classical Interpretation

With the advent of various conceptions of probability, explanations of chance and randomness arose in terms of probability and this has continued until today. Such explanations depended upon the underlying conception of probability. The first authors interpreted their conceptions in terms of winning expectation. Pascal (1654/1963a) estimated "the value" of an interrupted game by proportionally dividing the stakes among each player's chances. In his *Traité du Triangle Arithmétique* he suggested that a fair division of stakes should be proportional to the probability of winning the whole stake by each player (Pascal, 1654/1963b).

Christiaan Huygens, inspired by Pascal, was the author of the first probability treatise: *De Ratiociniis in Aleae Ludo* (Huygens, 1657 1998). In modern terms, he showed in his third proposition that if p is the probability of a person winning a sum a, and q that of winning a sum b, then he may expect to win the sum $pa + qb$.

In the same way Gottfried Wilhelm Leibniz wrote (Leibniz, 1676/1995, p. 161):

> If a situation can lead to different advantageous results ruling out each other, the estimation of the expectation will be the sum of the possible advantages for the set of all these results, divided into the total number of results.

This classical approach, arising from Fermat's conceptions transcribed in his letter to Pascal in September 1654 (Fermat, 1853/1989, p. 154), was found in the first definitions of probability, as given by Abraham de Moivre in *The Doctrine of Chances*:

> Wherefore, if we constitute a Fraction whereof the Numerator is the number of Chances whereby an Event might happen, and the Denominator the number of all the chances whereby it may either happen or fail, that Fraction will be a proper definition of the Probability of happening (de Moivre, 1718/1967, p. 1)

Historically the first authors related the randomness of possible outcomes in a chance situation to their own uncertainty about future events. Jacques Bernoulli expressed in *Ars Conjectandi*: "Probability is in fact a degree of certainty, and differs from certainty as the part from a whole". (Bernoulli, 1713/1987, p. 16).

Pierre-Simon Laplace published his *Essai Philosophique sur les Probabilités* in 1814, already partly written in 1795 for the Meetings of Teachers' Training Schools. In this fundamental book, Laplace clearly underlined the subjective view in judging equiprobability, which is necessary for the classical definition of probability, in concrete situations. After affirming that probability is partly related to the extent of our ignorance and knowledge, he noted that:

> the theory of chance consists in reducing all the events of the same kind to a certain number of equally possible cases, that is to say, to such as we may be equally undecided about in regard to their existence,

and gave this definition as the first principle:

> probability is thus simply a fraction whose numerator is the number of favorable cases and whose denominator is the number of all cases possible (Laplace, 1814/1995, p. ix).

In this classic conception of probability we would say that an object is chosen at random out of a given class, if the conditions in this selection allow us to give the same probability for any other member of this class ("hasard du tirage au sort", according to Lahanier-Reuter, 1999). In fact it was argued that this Laplacian definition of probability was based on a subjective interpretation, associated with the need to judge the equipossibility of different outcomes. Although equiprobability is clear when throwing a die or playing a chance game, it is not the same in complex human or natural situations. Bernoulli noted this in *Ars Conjectandi*, and gave examples of epidemics and weather phenomena: equiprobability "can hardly be found in some very rare cases and does not happen apart from games of chance" (ibid. p. 40).
He then indicates how to determine the probabilities of real events: "what is not given a priori is at least possible a posteriori, that is to say, it will be possible to obtain it by observing the result of many similar examples" (ibid. p. 42). He thus suggested the possibility of an objective and frequentist estimate for the probability of a concrete event.

Frequentist Approach

Theoretical studies concerning the quantitative prediction of future events from the regularity observed in repeated trials of random phenomena only appeared three centuries ago, when Bernoulli justified a frequentist estimation of probability in giving a first proof of a main probability theorem, the Law of Large Numbers. In modern terms, this theorem can be stated as follows: when repeating the same experiment enough times, the probability that the distance between the observed frequency of one event and its probability p is smaller than a given value, can approach 1 as closely as desired.

The stabilization of frequencies for an event, after a large number of identical trials of a random experiment, had been observed over several centuries. The proof given by Bernoulli that the classical probability correctly reflects this idea of stabilized value, was interpreted as a confirmation that probability was an objective feature of random events. Given that stabilized frequencies are observable, they can be considered as approximated physical measures of this probability. As Alfred Renyi claimed: "we consider probability as a value independent of the observer, which roughly indicates with which frequency the event will happen in a long series of trials." He adds: "Mathematical theory of probability... concerns objective probabilities which can be measured like physical magnitudes" (Renyi 1966/1992, p. 26).

Moreover the frequentist approach defines probability as the hypothetical number towards which the relative frequency tends when stabilizing (von Mises, 1928/1953; Renyi, 1966/1992; Ventsel, 1973). In this conception, we assume the existence of this number for which the observed frequency is an approximated value. According to Gnedenko and Kolmogorov (1954), "mathematical probability would be a useless concept if it did not find concrete expression in the relative frequency of events resulting from long sequences of experiments, carried out under the same conditions."

However, from a practical viewpoint, the frequentist approach does not provide the probability for an event when it is physically impossible to repeat an experiment a very large number of times. It is also difficult to decide how many trials are needed to get a good estimation for the probability of an event. Moreover, we cannot give a frequentist interpretation to the probability of an event, which only occurs one time under the same conditions, such as is often found in econometrics. But the most significant criticism of the frequentist definition of probability is the difficulty of confusing an abstract mathematical object with the empirical observed frequencies, which are experimentally obtained. In von Mises'

(1928/1952) axiomatic system, probability is considered as the theoretical limit of frequencies. However such a conception raises the didactic problem of confusing model and reality, and makes the modeling process difficult to understand for students who need to use abstract knowledge about probability and random variables to solve concrete problems.

Subjective View

Even though the frequentist approach was an advance relative to the classical view, it was not free of controversy. Bayes' formula published in 1763 raised questions that belied intuition. This formula gave the probabilities of various causes when one of their consequences is observed. The probability of such a cause would thus be prone to revision as a function of new information and would lose its objective character postulated by the frequentist conception. Keynes (1921), Ramsey (1931), and de Finetti (1974) described probabilities as personal degrees of belief, based on personal judgment and information about experiences related to a given outcome (Cabriá, 1992; Hacking, 1975). De Finetti (1974) claimed that "probability does not exist." He considered that assuming an objective existence would be an erroneous and dangerous conception. Since probability is a theoretical concept, its estimated value depends on numerous factors, such as the observer's knowledge, the observation conditions or the data that he is able to collect. Therefore, we cannot say that probability exists in reality without confusing this reality with the theoretical model chosen to describe it.

In this subjective view, what is random for one person might be nonrandom for another. Randomness is no longer a physical "objective" property, but has a subjective character and probability does not measure a magnitude, such as length or weight, but only a degree of uncertainty, specific for each person (Kyburg, 1974). Émile Borel, one of the founding fathers of measure theory, suggested that "to understand some errors made in incorrect applications of probability theory, we should briefly insist on the subjective character of probability". He underlined that "the possibility of an event is always related to a certain system of knowledge and is thus not necessarily the same for all people" (Borel, 1930/1991, pp. 70-71).

In this subjectivist viewpoint, the repetition of the same situation is no longer necessary to give a sense to probability. The fact that repeated trials are no longer needed serves to expand the field of applications of probability theory, in particular to economic decisions (Saporta, 1992). Today, the neo-Bayesian school assigns probabilities to all that is dubious or unknown, even nonrandom phenomena. But, what is the scientific stature of the results which depend on judgments that vary with the observer? The solution of this dialectical debate between objectivists and subjectivists is again found in the status of the mathematical model of probability theory.

Mathematical Formalism

Throughout the 20th century, different schools contributed to the development of the mathematical formalization of probability. Borel's first view of probability as a special type of measure (Borel, 1930/1991) was used by Andrei Kolmogorov (1933/1950), by considering a set (the sample space) representing all possible outcomes in a random experience. Kolmogorov applied sets and measure theories and used Lebesgue integration to derive a satisfactory axiom system, which was generally accepted by different schools independently of their philosophical interpretation of the nature of probability. Probability is thus a mathematical object and probabilistic models can be built to describe, simplify and interpret random reality. Probability theory has proved its efficiency in many applications, but the particular derived models raise heuristic and theoretical hypotheses, which need to be evaluated empirically. Moreover, probability cannot be considered as just a special case of measure theory, since the concept of independence or the limit theorems, so relevant in probability, play a specific role. In the period from Laplace and Gauss, to Kolmogorov and Doob, many other probabilists derived these results and built an extensive framework of knowledge attracting young researchers to this interesting field (Cabriá, 1992).

Intuition of Randomness and Random Sequences

When theoretical developments about statistical inference began to reveal the importance of separating the notions of random process and random sequence, interest in finding models for processes, which provide long sequences of random digits, was born. The possibility of obtaining pseudorandom digits with deterministic algorithms also suggested the need for examining the sequence produced, regardless of the process by which it had been generated. Debate about such things led to the formalization of the concept of randomness (Fine, 1971).

Intuitively (and in particular with children), chance is perceived as being primarily unforeseeable. Thus, for example, in throwing a die six times, the sequence [1, 2, 3, 4, 5, 6] seems less likely than [2, 5, 1, 6, 4, 3]. Players hold the belief that they risk less if they choose a sequence where no regularity can be perceived a priori. According to Parzysz (2004), various concepts were created during the 20th century to take into account this unpredictability; for example, the sequences of equidistributed digits or normal numbers introduced by Borel (1909).

Fine (1971) discussed some approaches used to define a random sequence. von Mises (1928/1952) based his study of random sequence on the intuitive idea that a sequence is considered to be random if we are convinced of the impossibility of finding a method that lets us win in a game of chance where winning depends on forecasting that sequence. von Mises and Martin-Löf (1966) suggested that a random sequence does not exhibit any exceptional regularity effectively testable by any possible statistical test. Kolmogorov and Chaitin's vision (1975) of a random sequence is a highly irregular or complex sequence that cannot be reproduced from a set of instructions which is shorter that the sequence itself (Zabell, 1992, Delahaye, 1999). It is important to remark that in both the theoretical approaches of von Mises and Kolmogorov perfect randomness would only apply to sequences of infinite outcomes and therefore, randomness would only be a theoretical concept.

4. FUNDAMENTAL STOCHASTIC IDEAS

Concepts Progressively Built from School to University

A key point in teaching probability is to reflect on the main content to include at different educational levels and how this content can help prepare students for life (see Gal, this volume). We have described in the previous sections the fundamental stochastic ideas that have helped Probability theory to develop throughout history. These ideas are analyzed by Heitele (1975), who takes the view after Bruner (1960) that fundamental mathematical concepts can be studied at various degrees of formalization. These degrees of formalization are manifest in more complex cognitive and linguistic levels as one proceeds through school to university using a spiral curriculum. He also suggests that small children can build intuitive models for these fundamental ideas that later help them to establish correct analytic knowledge. This is particularly important in stochastics where the large number of paradoxes might confuse even mathematically trained people. In effect, it underscores in the case of stochastics the need to reinforce intuitive understanding before formal teaching of the topic commences. As suggested by Feller (1950) even adults are able to improve their stochastic intuition. However, wrong intuitions that are acquired early are difficult to change and can later cause difficulties in learning (Fischbein, 1975).

In developing his list of key stochastical ideas Heitele considered results from developmental psychology and the history of probability, since both sciences prove these concepts to be difficult, though powerful. We briefly analyze below some ideas mentioned by Heitele.

Random Experiment, Events and Sample Space

The idea of listing all the different possibilities in a random experiment and taking into account not only the possible outcomes, but also the different possibilities for or against to estimate a player's probability of winning, was implicit in the pioneers' work on games of chance. At the same time the first unsuccessful attempts to solve some classical probability problems were due to considering incorrect sample spaces. The notion of sample space progressively developed and was formalized by Kolmogorov, who explicitly took the set of all the different possible outcomes as a base to build a satisfactory set of axioms for probability calculus. This set of axioms quickly gave momentum to a spectacular development of this part of mathematics. Fischbein (1975) emphasized the cognitive relevance of the sample space, because small children, who are too linked to deterministic thinking, often concentrate on a single event rather than on the whole set of possible results when dealing with random situations. This same behavior has been described in Konold's (1989) "outcome approach". Jones , Langrall, Thornton, and Mogill's (1999) research involving an instructional probabilistic program with young children suggests that overcoming misconceptions related to sample space was a key factor for children who showed a growth in probabilistic thinking (see also Langrall & Mooney, this volume).

The Addition Rule

Dividing a compound event into its single constituents is a powerful way to derive complex probabilities from simpler ones (see Polaki, this volume). The second axiom of probability achieves this by allowing us to compute the probability of the compound event. As it is a general rule in mathematics, once this idea is accepted, it is progressively generalized. Starting from the union of two single disjoint events, this rule is extended to a fixed or variable number of events, and later to compute probabilities in a continuous setting, where the sum is replaced by an integral.

Independence and Conditional Probability

The complex relationship between probabilistic concepts and intuition appears in the concept of independence. As we described earlier, the concept of probability started from the study of chance games, where independence was natural: A die or a coin does not have a memory of preceding throws. It was necessary to wait until the middle of the 18^{th} century before the concept of independence was noticed and made explicit. It started as an intuitive notion: two events were considered to be independent if there was no reason to think that one of them could influence the other. The probabilistic translation of this idea is expressed by the multiplication rule:

$P(A \cap B) = P(A) \times P(B)$.

The concept of independence soon became essential in the emergence of the normal distribution, obtained by Laplace and Gauss as the limiting distribution of many "small" independent errors. With the recent foundation of probability as an axiomatic theory by Kolmogorov, an inversion between definition and concept arose because then stochastic independence was defined in terms of the multiplication rule. This new definition was criticized (von Mises, 1928/1952), because it brought an extension of the concept, which emptied it of its intuitive content. That is, some events can be stochastically independent and not be intuitively independent or vice versa.

This historical difficulty in establishing a simple link between the intuitive idea of independence and its formal definition recurs in the teaching of probability, where it can be an obstacle for students when solving conditional probability problems. Misconceptions as regards conditional probability are very commonly discussed in statistics education research and are described in other chapters of this book (e.g., Batanero & Sanchez; Tarr & Lannin; Watson).

Computing probabilities in compound experiments requires one to analyse whether the experiments are dependent or not. Here we compose the experiments themselves and not just the events in the same experiment. Therefore Heitele (1975) suggests that the study of compound experiments can lead students to perceive mathematics' facility to build complex models based upon simpler ones.

Equidistribution of Probability

The ideas described above, though very powerful, do not help us in finding an objective criterion to start assigning probability to simple events. A possible strategy is accepting Laplace's equiprobability rule in situations where there is physical symmetry or where it is possible to apply the indifference principle. Even when this strategy seems natural in games of chance and other experiments with finite sample spaces, it is not free from subjective judgment, as described above. Besides young children sometimes do not easily accept equiprobability for cases which are obvious to adults, because they personalize random generators or believe in "lucky numbers" (Truran, 1996).

Combinatorics

Combinatorics is not simply a calculus tool for probability, but there is a close relationship between the two topics. At a cognitive level, according to Piaget and Inhelder (1951/1975), if the subject does not possess combinatorial capacity, he is not able to use the idea of probability, except in cases of very elementary random experiments. On the one hand, from a mathematical point of view the connection between probability and combinatorics is particularly noticeable in compound experiments. This is the case because the task of generating the sample space of a compound event requires the application of a combinatorial constructive process on the events that comprise the compound event. On the other hand, arrangements and combinations may be defined by means of compound experiments (ordered sampling with/without replacement, non-ordered sampling with/without replacement). It is not surprising then that we use tree diagrams to facilitate both the understanding of combinatorial configurations and compound random experiments (analyses of elementary combinatorics and of students' combinatorial reasoning, are given in this volume: Batanero & Sánchez; English; Polaki).

Random Variable and Distributions Models

One of the most powerful ideas in probability was born in the 20th century and served to expand its applications beyond games of chance, as well as to solve many paradoxes and difficulties. Random variables appear in many different contexts in everyday life and the number of distribution models for random variables as well as their applications is enormous. Again, possible generalizations or extensions of this idea appear in bivariate and multivariate random variables, as well as in stochastic processes. Associated with random variables is the idea of expectation. Expectation is a very natural aspect of games of chance, where it appeared very early in the historical development of probability. An intuitive introduction to the notion of random variable and expectation at any early age might provide the background for later formal understanding of probability models, such as the binomial, geometric, uniform, exponential or normal distributions.

Laws of Large Numbers

The progressive stabilization of the relative frequency of a given outcome in a large number of trials, that has been observed for centuries and was translated by Bernoulli to a mathematical theorem, served as a justification for the frequentist definition of probability, as we have seen. Modern generalizations of this theorem are known as Laws of Large Numbers. These laws lead to connections between probability and statistics and they give validity to statistics as a methodological tool in experimental sciences. Regularity in the distribution of independent unpredictable outcomes implies the possibility of discovering mathematical models in randomness and then getting some control over it (separating random and nonrandom components in natural phenomena). This idea again is not free of difficulties, because the specific nature of random convergence is difficult to grasp and long runs, coincidences, or unexpected patterns are counterintuitive (see Watson; Batanero & Sanchez; this volume).

Sampling

Given that we are rarely able to study complete populations our knowledge is based on samples, which have two different features: representativeness and variability. Because samples are (or should be) representative of the population, we expect them to be similar to the population but, at the same time, variability implies that one sample is different from another. Psychologists such as Kahneman, Slovic and Tversky (1982) suggest that we put too much emphasis on representativeness and are not sufficiently cognizant of random sampling fluctuation and the effect of sample size on sampling variability.

Modeling and Simulation

During the 20^{th} century, probabilistic knowledge was organized into a true mathematical theory, like other branches of mathematics such as geometry. Starting from social practices and the interpretation of tangible reality, this scientific or mathematical approach served to overcome the debates about the nature of the objects concerned and to accommodate the various philosophical conceptions about the nature of chance. The abstract character of probability's axiomatic foundation allows the possibility of utilizing models that are formalized in its symbolic system and developed to represent problems arising from reality. The modeling of concrete situations is today a compulsory step in the operation of scientific knowledge and, moreover, probability is a field where simple models can be composed in a powerful way. Therefore, the teaching of statistics and probabilities should incorporate the learning of modeling.

Heitele (1975) also included the idea of simulation among his list of fundamental stochastic ideas. Simulation might be used as a pseudoconcrete model for many different real situations and it offers the possibility of working without mathematical formalism when analyzing random situations. Simulation then can act as an intermediary between reality and the mathematical model. As a didactic tool it can serve to improve students' probabilistic intuition, to teach them the different steps in the work of modeling, and to help them discriminate between model and reality.

5. IMPLICATIONS FOR TEACHING

The above discussion shows the multifaceted nature of probability (Cabriá, 1992), and in particular its duality (Hacking, 1975); it also suggests that teaching cannot be limited to one of these different perspectives because they are dialectically and experientially intertwined. Probability can be viewed as an a priori degree of uncertainty and, at the same time, as a personal degree of belief (De Finetti, 1974).

The controversies with respect to the development of the theory and philosophy of probability have also influenced teaching (Henry, 1997b; see also Greer & Mukhopadhyay, this volume, for a detailed analysis of factors affecting the place and contents of probability in the mathematics curriculum). Before 1970, the classical view of probability based on combinatorial calculus dominated the secondary school curriculum. Later it was complemented with an axiomatic approach in the so-called "modern mathematics" era. On the one hand combinatorial reasoning is difficult and students often found this approach to be very hard. On the other hand the multiple applications of probability to different sciences were hidden and probability was considered by many secondary school teachers as a subsidiary part of mathematics, since it only dealt with chance games.

With the increasing recent interest in statistics at school level and with continuing computer development, there is a growing interest in an experimental introduction of the notion of probability as a limit of the stabilized frequency. Probability has now been turned into a theoretical tool that is used to approach problems that have arisen from statistical experiences. Probabilistic modeling of statistical questions is moreover central in the educational process because it enables students to decide the best solution to some paradoxes that appear even in apparently simple problems, and are predicated on the basis of confusion between model and reality (Girard, 1997). Let us consider, for example, the probability of getting at least one tail when flipping a coin twice. Some famous mathematicians gave different solutions to this simple problem because they applied different models for the sample space in this experiment. Thus, D'Alembert in 1776 argued in the *French Encyclopedia* that this probability was 2/3, since he considered three different equiprobable cases: getting a tail in the first flip (in this case the game is over), getting a head in the first flip and a tail in the second, or getting two successive heads. In his *Essai Philosophique*, Laplace found that the solution of the problem was 3/4, assuming equiprobability for the four events (tail, tail), (tail, head), (head, tail), (head, head). This is a very simple but paradoxical problem that we can propose to our students. They can simulate the game to decide experimentally which of the two previous models better fits the situation and later try to explain mathematically why one solution is preferable to the other.

Interpreting random situations in terms of probabilistic models will serve to overcome the controversy between classic, subjective and frequentist approaches. This modern conception will give probability the status of a mathematical object that quantifies what Popper (1959) described as the propensity for a given outcome to appear more easily or frequently than others.

From a didactic viewpoint, it is desirable to distinguish three different stages in the modeling process (Henry, 1997a, 2001). Special attention should be paid to the first stage, where students work at the concrete level observing a real situation and describing it in their own words. This description already involves some abstraction and simplification of reality, insofar as choices need to be made vis-à-vis what is relevant in the situation with respect to the problem studied (working hypotheses). This description is controlled by a theoretical look, that is, scientific knowledge based on general prebuilt models. For example, we can describe the "yes" or "no" response from a person taken at random in an electoral survey, by the sampling of a marble from a Bernoulli's urn which contains marbles of two colors in proportion p and $1-p$. Moreover, we start from the working hypotheses to represent such descriptions by a system of simple and structured relationships among idealized objects: it is the pseudoconcrete model level. In the voting example, we can assume that the probability of answering "yes" is the same for any person in the population (independent of his gender, age or social position) and that it does not change in a short interval of time.

The second modeling stage is the formalization of the model, which presupposes being able to represent the pseudoconcrete model in a symbolic system suitable for probabilistic calculus. To do this it is necessary to introduce a mathematical reference sample space Ω, characterize the events as parts of this unit, translate the working hypotheses into model assumptions and finally define the probability distribution as an abstract measure on Ω. Probabilistic theory then allows a solution to the problem posed. In the voting example, we introduce the idea of random binomial distribution to estimate the probability of having x out of n positive votes in a group of n people and use the normal approximation for large values of n.

At the third stage, it is advisable to go back to the initial questions and translate the mathematical results in terms of the pseudoconcrete model. This will make these results meaningful in providing some answers, which should be related to the working hypotheses to appreciate the relevance of these answers in the real situation (model validation). In our example, we might use the theoretical model to check the hypothesis that a change in the value of p took place in the general population as a consequence of some political actions or that the vote of young people is different from the vote in the general population.

The development of computers has added an important resource to simulation in statistics and probability with random digit generators. However, using computers as simulation tools requires characterizing a model of the simulated situation and makes it still more necessary to explicitly state the working hypotheses. To be theoretically acceptable the simulation should correspond to the same theoretical model as the random experience we are trying to reproduce. In order for it to be didactically effective, that is, for students who have no theoretical models available to accept it as a simulation of the given experience, the simulation should be as close as possible to the experience itself. It is in fact only by working with different simulations, and recognizing their analogy with the same experience that students can abstract the idea of a model and make it a powerful tool in problem solving.

Finally, we remark that a pure experimental approach is not sufficient in the teaching of probability. Even when a simulation can help to find a solution to a probability problem arising in a real world situation, the simulation cannot prove that this is the most relevant solution, because it depends on the hypotheses and the theoretical setting on which the model is built. A genuine knowledge of probability can only be achieved through the study of some formal probability theory. However, the acquisition of such formal probability theory by the students should be gradual and supported by their stochastic experience (see Pfannkuch, this volume).

REFERENCES

Batanero, C., Serrano, L., & Green, D. R. (1998). Randomness, its meanings and educational implications. *International Journal for Mathematics Education in Science and Technology, 29*(1), 11-123.

Bellhouse, D. R. (2000). De Vetula: a medieval manuscript containing probability calculations. *International Statistical Review, 68*(2), 123-136.

Bennett, D. J. (1993). *The development of the mathematical concept of randomness; educational implications* (Doctoral dissertation, New York University, 1993). Dissertation Abstracts International, 54, 449A.

Bernoulli, Jacques (1987). *Ars Conjectandi- 4ème partie (N. Meunier, Trans.)* Rouen: IREM. (Original work published in 1713).

Borel, E. (1909). Presque tous les nombres réels sont normaux [Almost all the real numbers are normal]. *Rendiconti del Circulo Matemático di Palermo, 27*, 247-271.

Borel, E. (1991). *Valeur pratique et philosophie des probabilités* [Practical and philosophical value of probability]. Paris: Jacques Gabay (Original work published in 1930).

Borovcnik, M., Bentz, H. J., & Kapadia, R. (1991). A probabilistic perspective. In R. Kapadia & M. Borovcnik (Eds.), *Chance Encounters: Probability in Education* (pp. 27-73). Dordrecht: Kluwer.

Borovcnik, M., & Peard, R. (1996). Probability. In A. J. Bishop, K. Clements, C. Keitel, J. Kilpatrick, & C. Laborde (Eds.), *International handbook in mathematics education* (Part 1, pp. 239-288). Dordrecht: Kluwer.

Bruner, J. S. (1960). *The process of education.* Cambridge, MA: Harvard University Press.

Cabriá, S. (1992). *Filosofía de la probabilidad* [Philosophy of probability]. Valencia, Spain: Tirant lo Blanc.

Cardano, G. (1961). *The book on games of chances.* New York: Holt, Rinehart & Winston. (Original work Liber de Ludo Aleae; included in Opera Omnia, Vol.1, published in 1663).

Chaitin, G. J. (1975). A theory of program size formally identical to information theory. *Journal of the Association for Computing Machinery, 22*, 329-340.

Courtebras, B. (2001). Sur quelques conceptions du hasard [Regarding some conceptions of chance]. In M. Henry (Coord.), *Autour de la modélisation en probabilités* (pp. 95-132). Commission Inter-IREM Statistique et Probabilités, Besançon: Presses Universitaires Franc-Comtoises.

De Finetti, B. (1974). *Theory of probability.* London: John Wiley.

Delahaye, J. P. (1999). *Information complexité et hasard* [Information, complexity and chance] (2nd ed.). Paris: Hermes Science Publications.

Diderot, D. (1983). *Jacques le Fataliste* [Jacques the Fatalist]. Paris: Le Livre de Poche (Original work published 1796).

Edwards, A. (1987). *Pascal's arithmetical triangle*. London: Oxford University & Charles Griffin.

Feller, W. (1950). *An introduction to probability theory and its applications* (Vol. 1). New York: John Wiley.

Fermat, P. (1989). *Précis des œuvres mathématiques*. Paris: Jacques Gabay (Original work published 1853).

Fine, T. L. (1971). *Theories of probability. An examination of foundations*. London: Academic Press.

Fischbein, E. (1975). *The intuitive sources of probabilistic thinking in children*. Dordrecht, The Netherlands: Reidel.

Fischbein, E., Nello, M. S., & Marino, M. S. (1991). Factors affecting probabilistic judgments in children in adolescence. *Educational Studies in Mathematics, 22*, 523-549.

Girard, J. C. (1997). Modélisation, simulation et expérience aléatoire [Modeling, simulation and random experience]. In *Enseigner les probabilités au lycée* (pp. 73-76). Reims: Commission Inter-IREM Statistique et Probabilités.

Gnedenko, B. V., & Kolmogorov, A. N. (1954). *Limit distributions for sums of independent random variables* (K. L. Chung, Trans.). Cambridge, MA: Addison-Wesley.

Hacking, I. (1975). *The emergence of probability*. Cambridge, MA: Cambridge University Press.

Heitele, D. (1975). An epistemological view on fundamental stochastic ideas. *Educational Studies in Mathematics, 6*, 187-205.

Henry, M. (1997a). Notion de modèle et modélisation dans l'enseignement [Notion of model and modeling in teaching]. In *Enseigner les probabilités au lycée* (pp. 77-84). Reims: Commission Inter-IREM Statistique et Probabilités.

Henry, M. (1997b). L'enseignement des statistiques et des probabilités [Teaching of statistics and probability]. In P. Legrand (Coord.), *Profession enseignant: Les maths en collège et en lycée* (pp. 254-273). Paris: Hachette-Éducation.

Henry, M. (2001). (Coord.), *Autour de la modélisation en probabilités* [Around modeling in probability]. Besançon: Presses Universitaires Franc-Comtoises.

Huygens, C. (1998). *Du calcul dans les jeux de hazard* [The calculus in games of chance]. In Œuvres complètes, t. 14, La Haye, 1888-1950. (Original work Ratiociniis in Aleae Ludo, published 1657).

Jones, G. A., Langrall, C. W., Thornton, C. A., & Mogill, A. T. (1999). Student's probabilistic thinking in instruction. *Journal for Research in Mathematics Education, 30*(5), 487-519.

Kahneman, D., Slovic, P., & Tversky, A. (Eds.), *Judgement under uncertainty: Heuristics and biases*. New York: Cambridge University Press.

Keynes, J. M. (1921). *A treatise on probability*. London: Macmillan.

Kojeve, A. (1990). *L'idée du déterminisme dans la physique classique et dans la physique moderne* [The idea of determinism in classical and modern physique]. Paris: Librairie Générale Française, le Livre de Poche-essais (Original work published 1932).

Kolmogorov, A. (1950). *Foundations of probability's calculation.* New York: Chelsea Publishing Company (Original work, Grundbegriffe der wahrscheinlichkeitsrechnung, published 1933).

Konold, C. (1989). Informal conceptions of probability. *Cognition and Instruction, 6*, 59-98.

Kyburg, H. E. (1974). *The logical foundations of statistical inference.* Boston: Reidel.

Lahanier- Reuter, D. (1999). *Conceptions du hasard et enseignement des probabilités statistiques* [Conceptions of chance and teaching of statistics and probability]. París: Presses Universitaires de France.

Laplace, P. S. (1986). *Essai Philosophique sur les Probabilités* [Philosophical essay on Probabilities]. Paris: Christian Bourgois (Fifth edition of the original work, published 1825).

Laplace P. S. (1995). *Théorie analytique des probabilités* [Analytical theory of probabilities]. Paris: Jacques Gabay. (Original work published 1814).

Leibniz, G. W. (1969). *Théodicée.* Paris: Garnier-Flammarion (original work published 1710).

Leibniz, G. W. (1995). L'estime des apparences [The estimation of appearance]. In M. Parmentier (Ed.), *21 manuscrits sur les probabilités, la théorie des jeux, l'espérance de vie.* Paris: Vrin (Original work published 1676).

Martin-Löf, P. (1966). The definition of random sequences. *Information and Control, 9*, 602-619.

Mises, R. von (1952). *Probability, statistics and truth* (J. Neyman, O. Scholl, & E. Rabinovitch, Trans.). London: William Hodge and company. (Original work published 1928).

Moivre, A. de (1967). *The Doctrine of Chances* (3rd ed.). New York: Chelsea Publishing (Original work published 1718).

Monod, J. (1970). *Le hasard et la nécessité, essai sur la philosophie naturelle de la biologie moderne* [Chance and necessity, an essay on natural philosophy in modern biology]. Paris: Editions du Seuil.

Morin, E. (1990): *Science avec conscience* [Science with conscience]. Paris: Editions du Seuil, Points Sciences.

Parzysz B. (2004). Quelques questions à propos des tables et des générateurs pseudo-aléatoires [Some questions about pseudo random numbers, tables, and generators]. In B. Chaput (Coord.), *Statistique au lycée.* Paris: Commission Inter-IREM & APMEP.

Pascal, B. (1963a). Correspondance avec Fermat [Correspondence with Fermat]. In *Oeuvres Complètes* (pp. 43-49). París: ed. Seuil. (Original letter 1654).

Pascal, B. (1963b). Traité du triangle arithmétique [A treatise on the arithmetical triangle]. In *Oeuvres Complètes* (pp. 57-62). París: ed. Seuil. (Original work published 1654).

Piaget, J. & Inhelder, B. (1975). *The origin of the idea of chance in children.* (L. Leake Jr., P; Burrell, & H. D. Fischbein, Trans.). New York: Norton. (Original work published 1951)

Pichard, J. F. (2001). Les probabilités au tournant du XVIIIe siècle [Probability at the beginning of XVIII century]. In M. Henry (Coord.), *Autour de la* *modélisation en probabilités* (pp. 13-46). Besançon: Presses Universitaires Franc-Comtoises.

Poincaré, H. (1987). *Calcul des probabilités* [Probability calculus]. París: Jacques Gabay (Original work published 1912).

Popper, K. (1959). The propensity interpretation of probability. *British Journal of Philosophy of Science, 10,* 25-42.

Prigogine, I., & Steengers, I. (1979). *La nouvelle alliance* [The new alliance]. Paris: Gallimard, Folio Essais.

Ramsey, F. (1931). Truth and probability. In R. B. Braithwaite (Ed.), *The foundations of mathematics and other logical essays* (pp 156-198). London: Kegan Paul.

Renyi, A. (1992). *Calcul des probabilités* [Probability calculus] (L. Félix, Trans.) Paris: Jacques Gabay (Original work published 1966) Ruelle, D., (1991). *Hasard et chaos* [Randomness and chaos]. Paris: Odile Jacob.

Saporta, G. (1992). *Probabilités, analyse des données et statistiques* [Probability, data analysis and statistics]. Paris: Technip.

Shaughnessy, J. M. (1983). Misconceptions of probability, systematic and otherwise: Teaching probability and statistics so as to overcome some misconceptions. In D. R. Grey, P. Holmes, V. Barnett, & G. M. Constable (Eds.), *Proceedings of the First International Conference on Teaching Statistics* (pp. 784-801). Sheffield, UK: Teaching Statistics Trust.

Shaughnessy, J. M. (1992). Research in probability and statistics. In D. A. Grouws (Ed.), *Handbook of research on mathematics teaching and learning* (pp. 465-494). New York: Macmillan.

Székely, G. J. (1986). *Paradoxes in probability theory and mathematical statistics.* Dordrecht, The Netherlands: Reidel.

Thom, R. (1986). Preface. In Édition of P. S. Laplace, *Essai philosophique sur les probabilités,* (pp. 5-27). Paris: Christian Bourgois, Épistémé.

Thom, R. (1990). Postface au débat sur le déterminisme [Postface to the determinism debate]. In P. Krzysztof (Ed.), *La querelle du déterminisme, Philosophie de la science d'aujourd'hui* (pp. 266-279), Paris: Gallimard, Le débat.

Truran, J. M. (1996). Children's misconceptions about the independence of random generators. In L. Puig & A. Gutiérrez (Eds.), *Proceedings of the 20th conference of the International Group for the Psychology of Mathematics Education* (pp. 331-338). University of Valencia, Spain.

Ventsel, H. (1973). *Théorie des probabilités* (Probability theory). Moscow: Mir.

Zabell, S. L. (1992). The quest for randomness and its statistical applications. In F. Gordon & S. Gordon (Eds.), *Statistics for the XXI Century* (pp. 139-166). Washington, DC: The Mathematical Association of America.