# How can we assign a probability to an event?

Hugo LAVENANT[1]

February 2019

This a translation done by Hassan AÏT-KACI — hak@acm.org — of the original technical note in French authored by Hugo LAVENANT entitled "*Comment attribuer une probabilité à un évènement ?*"

*Il faut réfléchir pour mesurer et non pas mesurer pour réfléchir.*[2]

Gaston Bachelard, *La formation de l'esprit scientifique*.

**Abstract**

This document aims at a discussion of the different interpretations that can be made of the calculus of probabilities with the question of how to assign a probability to an event: upon which elements, whether objective or subjective, does one base oneself in order to assign probabilities to events in areas as diverse as games of chance, social phenomena, or athletic competitions? The law of large numbers, which appears to reduce all probability evaluation to a limit of frequency, is discussed.

## Introduction

As surprising as this may appear, the calculus of probabilities, that is the mathematical theory that concerns probabilities, does not say what the probability of an event actually *is*. In no opus of mathematics could you find a proof that the probability of getting "tails" when flipping a coin is 1/2; worse still, even if that was the case, the opus would not teach you what the probability of getting tails actually *means*. The objective of the calculus of probabilities is, given the probability of certain events, to calculate the probability of other events that depend on the given ones. Now, if you ensure a mathematician that the probability of obtaining tails when flipping a coin is 1/2, s/he would be able to tell you, or even *prove* to you, that the probability of obtaining tails exactly twice in three independent successive coin tosses is 3/8. However, this still does not teach you what this 3/8 means. Let there be no misunderstanding though: it is precisely because of this limitation that the

[1] Laboratoire de Mathématiques d'Orsay — hugo.lavenant@u-psud.fr.

[2] One must think in order to measure but not measure in order to think.

calculus of probabilities took off and realized intellectual constructions of great beauty which honor the human mind. But when it is question to use the results of the calculus of probabilities in concrete situations, to make decisions or as arguments in a debate, the meaning of the probability arises inevitably, whether implicitly or explicitly.

This document will try to bring to light some elements of answer to the following question: how can we assign a probability to an event? That is, upon which criteria can we base ourselves to aver that the probability of obtaining tails is 1/2? It is necessary to define some terms, at least to draw a rough picture that will be later made more precise. An *event* is said to be random, risky, or also fortuitious, if there is some uncertainty concerning its happening. Of course, this definition is insufficient and begs being discussed in more details, but such is not the object of the present document; we shall assume that everyone has a sufficiently clear idea of what a fortuitious event is. To give some examples, such is the case of the coin that can fall on tails (or heads), but also of the fact that an individual dies in the coming year, or that a sport team wins its next game. A *random experiment* describes the conditions in which a random event can happen or not. The possible results of the experiment are the *issues*; therefore, a random event is one of the possible issues of a random experiment. In the three examples mentioned above, a coin toss, an individual's life, or a game match are random experiments respectively associated to the corresponding random events. A central question is to know whether an experiment is reproducible; that is, whether the conditions in which an event possibly happens can be reproduced *all other things being equal*. Lastly, the *probability* of an event is a measure of its degree of certainty. The only constraint, is that probability of an event is a number that must be comprised between 0 and 1; the value 0 being assigned to an impossible event, and the value 1 to a certain event.

In an attempt to clarify our project with a metaphor, let us compare the calculus of probabilities to geometry: one of the objectives of geometry is to express relationships between lengths, such as Pythagoras's theorem; but geometry does not mention what a length is, and especially it does not say how a length is measured. It does not specify whether to use a ruler if the length to be measured is a straight line, or a string if what is to be measured is the perimeter of a curved shape like a circle. It does not specify what precautions to take in order for the measurement to be as precise as possible. On the other hand, geometry can prove theorems that explain how to measure a length indirectly; for example, in order to measure the distance between Earth and the Moon, there is no need to build a gigantic ruler as it suffices to measure the various angles under which the Moon is seen from Earth and some lengths on Earth in order to derive the length we seek. It is exactly the same for the calculus of probabilities: it suffices to replace "length" by "probability." In particular, the calculus of probabilities can teach us how to measure some probabilities indirectly; namely, by reducing the measure of the probability of an event to that of another that is easier to obtain. The difference is that, while there is a consensus regarding how to measure lengths, such is not the case for probabilities.

We next review threes manners of assigning probabilities, without pretense of exhaustivity nor historical contextualization; then, we shall discuss the result that seems to attempt to unify these conceptions: the law of large numbers.

# Three ways of assigning probabilities

## The "classical" conception

If we were to give only one name to symbolize the classical conception of probabilities that would be that of Pierre Simon de Laplace (1749–1827). In order to evaluate the probability of an event $E$ in a random experiment, one must begin by determining the cases (the issues, taking the terminology of the introduction) that are equally probable, then define:

$$\text{Probability}(E) = \frac{\text{number of cases } E \text{ happens}}{\text{number of possible cases}}.$$

A paradigmatic example is that of rolling a (six-faced) die: if the die is balanced, then each of the six faces constitutes an "equally probable case," such that in order to compute the probability of obtaining any of the faces, for example "five," there is only one favorable case among the six that are possible, and therefore:

$$\text{Probability}(\text{"five"}) = \frac{1}{6}.$$

If the event in question is composed of several elementary cases, it suffices to enumerate: for example, there are three ways of obtaining an even number (namely, "two," "four," or "six") so that:

$$\text{Probability}(\text{"obtain an even number"}) = \frac{3}{6} = \frac{1}{2}.$$

Another example that is abundantly used is that of picking a ball in an urn: if an urn contains $p$ white balls and $q$ black balls which are all otherwise identical except for color, then there are $p + q$ balls and so $p + q$ equally probable cases, and $p$ cases are favorable for picking a white ball, so that:

$$\text{Probability}(\text{"picking a white ball"}) = \frac{\text{number of white balls}}{\text{total number of balls}} = \frac{p}{p+q}.$$

That is to say that the probability of picking a white ball is equal to the proportion of white balls in the urn. In general, the problem of assigning probabilities is reduced to an *enumeration* problem, since it suffices to count, to enumerate, the favorable cases and the possible cases. This type of problem is not always simple but it has the advantage of possessing an answer that is well determined, and more importantly that is determinable thanks to purely mathematical tools in which the notion of randomness or chance is eliminated.

In reality, however, this definition begs clarification: how can we determine the "equally probable" cases? Since if this term is not made precise the definition is circular in that it presupposes a concept of equiprobability in order to obtain that of probability. In the foregoing examples, and in games of chance in general, it is considerations of symmetry that enable concluding: if the die is perfectly symmetrical, it is hard to see why a face would be more probable that the others; if the balls are identical, there is no reason why a ball would have more chance to be picked over another, *etc.* This is perfectly natural, almost instinctive, and this is why games of chance have been used to explain the meaning of the probability of an event. Thus, to explain the meaning of, for example, a death rate of 3 for a thousand, we can use an urn containing 3 black balls and 997 white balls, and the death of an individual is assimilated to picking a black ball, and the survival to picking a white ball. One must keep in mind that this is only an analogy; it does not define exhaustively the meaning of death rate.

But how can we proceed beyond games of chance? We could say that the equally probable events are those about which we lack the same amount of information, those for which we are equally indecisive. This permits recognizing explicitly that probabilities are a means to measure our uncertainty, that they are a tool that allows us to express our lack of knowledge: probabilities are *epistemic*; that is, relative to our knowledge, or more precisely its imperfection. Laplace's demon,[3] having before his eyes all that happened and that will happen to the world, does not need probabilities; we humans do. But in reality the question has now shifted to quantifying our lack of knowledge: how can we judge whether we are "equally indecisive"? Either no information is available to us, but in this case one wonders on what authority we can pretend quantifying that about which we know absolutely nothing; or, we have pieces of information that are of equal weight, but then again we must explain how to weigh information. It is undoubtedly for this reason that it is hard to assign classical probabilities when conditions of symmetry do not apply: a trick die is already out of scope,[4] but what about probabilities of death?

Another severe criticism can be expressed against classical probabilities which becomes relevant when the number of equally probable cases is infinite as this gives rise to Bertrand's Paradox. Indeed, if the number of possible cases and that of favorable cases are both infinite, the quotient defining the probability does not have a univocal sense. The paradox may be formulated as follows: let us assume that you have before you a glass containing a a mixture of 20 cl of water and an unknown quantity of alcohol, which you know to be comprised between 0 and 20 cl. What is the probability that the quantity of alcohol is comprised between 10 and 20 cl? Intuitively, you have a tendency to answer 1/2 since, in the absence of information, the cases "between 0 and 20 cl" and "between 10 and 20 cl" are equally probable. However, the problem can also be stated as follows: the proportion (the degree) of alcohol in this glass is comprised between 0° (if there is 0 cl) and 50° (if there is 20 cl). And a quantity of alcohol between 10 and 20 cl corresponds to a

[3] To Laplace, the world is perfectly deterministic; that is, the state of the future of the universe is entirely determined by the present. The "demon of Laplace," possessing an absolute intelligence and a perfect knowledge of only the present state of the universe, is capable by computation of knowing the future and the past as clearly as the present.

[4] We must be careful with such a statement: historically, probabilities were introduced in the 17th century to think about fair contracts in the presence of uncertainty. They were strongly connected to the notion of *fairness* and enabled finding the correct price. The idea of a trick die, or more generally of an unfair game of chance, was unthinkable from the perspective of probabilities at that time.

proportion of 33° (more precisely: 1/3 of alcohol if there is 10 cl) and 50°. However, the probability of the proportion of alcohol to be comprised between 33° and 50° knowing that it varies between 0° and 50° is, according to the classical method, 1/3 (the computation is $(50-33)/50 \simeq 1/3$). To the same question, namely "what is the probability that the quantity of alcohol is comprised between 10 cl and 20 cl?" we find two different answers: 1/2 (if we reason in terms of quantity of alcohol) and 1/3 (if we reason in terms of proportion of alcohol). Two answers to the same question: this shows indeed that this question is ill-stated; that is, that the classical method is insufficient.

Is there really nothing that can be salvaged in the classical conception, must it be all discarded in the dungeons of history? Not quite. First of all because, as pointed out above, it is not a matter of chance if rolling a die or picking a ball from an urn remain the most used examples to introduce the notion of probability, our mind intuitively perceiving the concept of equiprobability in this situations since it is reduced to that of symmetry. But also because modern methods enable us to push the classical method further by incorporating the cases where the available pieces of information do not put us in a situation of equal indecision: the incorporation is achieved by taking the distribution of probability that encodes the available pieces of information without any additional information, that is which maximizes the "disorder" among all situations that contain the available information,[5] so that the probability reflects our knowledge and our lack of knowledge at a given time.

## The frequentist conception

This conception is often presented as the only one worthwhile. If the underlying principle is rather simple, a rigorous definition and a real exploration of the consequences thereof can be attributed to Richard von Mises (1883–1953). The basic idea is that an event is seen as the issue of a random experiment, experiment which defines the conditions for the possible happening of the event and which must be reproducible. When it is reproduced a certain number of times, the frequency of an event $E$ is defined as the ratio of the number of times $E$ was realized over the total number of experiments, that is by:

---

[5] Reference is made here to the concept of *entropy* as defined in Information Theory. If a random experiment has $n$ pairwise disjoint issues $E_1, E_2, \dots, E_n$, in order to assign a distribution of probability $\mathrm{p} = (p_1, p_2, \dots, p_n)$, we choose the distribution **p** that maximizes the entropy:

$$S(\mathrm{p}) = -\sum_{k=1}^{n} p_k \ln(p_k)$$

among all the distributions of probability satisfying certain constraints corresponding to the available information on the random experiment. For example, we can impose that the expectation of a random variable, computed from the distribution **p**, be a fixed value. Maximizing the entropy guarantees that the distribution of probability does not encode more information than what is contained in the constraints on **p**.

$$\text{Frequency}(E) = \frac{\text{number of experiments that realize } E}{\text{total number of experiments}}.$$

This definition is quite different from that of the classical conception: in the latter enumeration was with respect to the possible issues of the experiment, while here it is with respect to the results of the experiments actually carried out. This definition depends on how many experiments have been carried out and it is not possible to identify frequency and probability without falling into paradoxical situations. For example, if a die is rolled and settles showing face "five," taking into account only this experiment, the conclusion is that:

$$\text{Frequency}(\text{“five”}) = \frac{1}{1} = 1.$$

Yet, a probability of 1 corresponds to an event that is certain and it is not by making a single experiment that we can conclude that the event is certain always to occur: it is not possible to identify probability with frequency. To palliate this problem, we must carry out a sufficient large number of experiments. If for example a die is rolled 100 times and settles on the number "five" 18 times, then:

$$\text{Frequency}(\text{“five”}) = \frac{18}{100} \simeq \frac{1}{6}.$$

The larger the number of experiments, the more the frequency of realization of an event $E$ approaches a limit value, and it is this value that is defined as the probability of $E$: the latter is measured more and more precisely by computing the frequencies of realization of $E$ for a larger and larger number of experiments.

This constraint of "large" number of experiments demands some clarification. First of all, the fact the frequency of realization of an event approaches more and more a limit value by making a large number of experiments is an *empirical* property which is not postulated *a priori* but is verified experimentally: in this sense, the probabilities are *objective*; they describe the regularities of the world around us. It is only by actually rolling a die a large number of times that we can ensure that the frequency of realization of the face "five" is closer and closer to 1/6, and in this sense the casinos provide us the best proof of this fact, since they are indeed places where a large number of dice are rolled. Another example of empirically observed regularity is that of social phenomena such as weddings. Here, the event is identified to an individual, and the result of the experimentally is positive if this individual gets married within the coming year, and negative otherwise. It was a surprise, in the 19th century, to discover that the frequency of weddings was of unfailing regularity, and that the more the number of individuals (that is of experiments) increased, the more the frequency of individuals getting married within the year approached a limit value. It was a surprise because a wedding was seen as the expression of free will, so finding a regularity in such a phenomenon was unexpected. Once this regularity was observed, it is possible to give meaning to the probability that an individual get married within the year; it is the limit of the proportion of persons that get married within the year in a larger and larger population.

A possible objection is that it is not possible actually to realize an infinite number of experiments, and therefore the "true" probability of an event will never be known. In reality, thanks to the calculus of probabilities, it is possible to estimate the gap between the observed frequency and the "real" probability, and so for a large but finite number of experiments, the frequency gives the probability up to some *known* uncertainty.[6] And thinking about it, this is the same thing for all physical quantity: a quantity such as the length of a table cannot be defined with infinite precision; any measure of the length of a table is flawed due to uncertainty depending on the measuring device (with a double-decimeter ruler the uncertainty is of the order of one millimeter), and the "true" value of the length could be defined as the limit of the measure as more and more precise measure devices are used. But, one may argue, in some cases such as social phenomena the number of experiments can never exceed the number of people on Earth; hence, the probability can never be known up to a precision as small as wanted. Here again, a comparison with the concept of lengths is relevant: the length of a table is never defined up to an arbitrary precision; at a microscopic scale the edge of the table is not straight but irregular, temperature can fluctuate and modify slightly the length, *etc.*, so that an irreducible imprecision plagues any measure of length. Thus, the concept of probability in this respect is as defined as that of length[7] since we know, thanks to the calculus of probability, the conditions in which the frequency gives a good approximation of the probability.

Another objection concerns the reproducibility of the experiments: rigorously speaking no experiment is similar to another, and this is all the more obvious when probabilities are defined for social phenomena. No wedding is identical to another; the statistical study of social facts goes through considering phenomena to be equivalent that are not so. This criticism is not unjustified but can be refined. The probability of an event depends on the random experiment of which it is seen as an issue; this is a property of the whole set of experiments and not of a particular case: as for rolling a die, the probability of obtaining the number "five" is not a property of the next die roll, but of a very large set of rolls indeed; as for weddings, the probability of getting married within the year does not concern one individual in particular but a whole social group, and so can vary depending on the group to which an individual belongs. Thus, the probability of getting married for a 25-year-old male individual who is still a student is not the same if he is considered to belong to the group of 25-year old males or that of students.[8] If we specify too many conditions characterizing the social group that any single individual must belong to (25-year old male,

[6] This is the concept of confidence interval: such an interval, centered around the observed frequency, indicates where the true value of the probability lies very probably.

[7] This does not necessarily mean that these concepts are well defined: thinking that there exists a "true" value that is real, objective, of the length of a table or of the probability of getting a "five" when rolling a die, that our measurements attempt to reveal it, or at least get close to it, is not a conception that is accepted by all, but it is argued here that the problem is not any different for the probability or the length.

[8] Even if this begs verification, it is surely higher in the first case than in the second.

student, living in Paris, diabetic, practicing a sport, amateur of broccoli, *etc.*), the social group could end up being restricted to a single individual and the probability does not make any sense any longer since the number of experiments is too small. Choosing a random experiment, that is of an equivalence class of distinct phenomena, is constrained by two opposite requirements between which a compromise must be found: the equivalence class must contain sufficiently many elements so that the frequency approaches the probability sufficiently well, while at the same time the distinct elements of the equivalence class must be sufficiently similar for the notion of "reproducibility" to make sense.

A problem subsists: it must be possible to distinguish, among the results of repeated experiments whose frequencies of event realization converge, those which possess additional regularities from those corresponding to actual random experiments. Let us think of days of the week: although the frequency of "Thursday" is $1/7$, to say that the probability of a given day of being "Thursday" is $1/7$ appears devoid of meaning: the sequence of week days is not a random sequence; on the contrary, it exhibits a foolproof regularity. In order to palliate this flaw, it is necessary to have a criterion defining what a random sequence is. If we consider an experiment repeated a large number of times and we are interested in whether an event $E$, supposed to happen with probability $p$, is realized, we say that the sequence of results of the experiments is *random* if there is no way to earn money by betting on the happening of $E$. That is to say that if before each experiment, a gambler can bet any amount of money $S$ of his or her choice on $E$ happening (in which case, s/he wins $S/p$ if $E$ actually happens) or on $E$ not happening (in which case, s/he wins $S/(1-p)$ if $E$ does not happen), and such that this gambler has only access to the results of the previous experiments, then there is no algorithm, no procedure, that will guarantee him/her (*via* the choice of the amounts $S$ of the bets) making a profit: over the long term, s/he will neither win nor lose money. Going back to the day-of-the-week example, a gambler would quickly understand that it is best to bet on "Thursday" 7 days after the last "Thursday" came about and then win money for sure. On the contrary, for games of chance, there does not exist any miraculous procedure that will guarantee winning money for sure, and there again this judgment does not follow from *a priori* considerations but from an empirical observation indeed; such observation having been made in the casinos.[9] Here also the hypothesis that a sequence of experiments is random is a *global* property, which depends on the whole set of experiments, not on a particular case. Hence, it makes no sense to say that a single event is random; it always is so relatively to a series of experiments of which it is only one of the instances.

Is the frequentist conception satisfactory? Not quite. First of all because the calculus of probabilities as it is today, although strongly inspired by this conception, presupposes in

---

[9] Richard non Mises compares this observation with that of the impossibility of perpetual motion: just as physics does not prove that perpetual motion does not exist but on the contrary observes it and elevates this impossibility to a postulate (the second principle of thermodynamics), the calculus of probabilities does not prove that it is impossible to win money for sure in a casino but on the contrary elevates this impossibility to the rank of an axiom, of a criterion allowing to distinguish a random sequence from another.

reality stronger hypotheses.[10] But also because if the experiments are not sufficiently reproducible or do not happen in sufficiently large number, by the frequentist conception one must renounce using the calculus of probabilities altogether. However, the concept of probability is sometimes used in such cases, notably as we shall see in what follows when it concerns the unique law of chance.

## The subjectivist conception

This is perhaps the most unusual of the three conceptions, that we shall present below in the form given to it by the work of the philosopher and mathematician Bruno de Finetti (1906–1985), who is able, in an elegant manner, to rest his philosophical pretentions on mathematical results. For him, the probability of an event is a *measure of the degree of belief* regarding this event's happening. The probability of an event is the result of the judgment of a human subject; therefore, it can vary from a subject to another for the same event: it is in this sense that it is *subjective*. It differs from the epistemic probability because the latter is relative to knowledge and so is identical for different rational subjects knowing the same things, while the subjective probability could still differ for such subjects. The classical example, besides the usual games of chance, are sporting events. Before a soccer game takes place, its issue is uncertain and everyone has a certain belief as to which team will win, the probability effectively allows quantifying this belief. Depending on the available pieces of information, on the sensitivity and experience of each, the beliefs and therefore the probabilities differ. Another example is prospecting for oil: a geologist is mandated by an oil company to indicate whether or not there is oil in a given region. The geologist investigates, assembles pieces of information and acquires a certain degree of belief, resting on the collected pieces of information and his/her own knowledge, regarding the presence of oil, or lack thereof, in the region in question. This degree of belief must be integrated into a complex decision process of the company: the decision to drill at this location depends on the potential presence of oil, or lack thereof, but also on the cost of drilling, or perhaps several potential drillings in other regions, on the economic objectives of the enterprise, *etc.* However, the geologist is not aware of all these details, and this is why s/he expresses her/his degree of belief using a probability: the latter is an operational means for expressing in a language comprehensible by all her/his level of certainty. Thanks to this language, the company can integrate this expert's advice in its decision process without having to acquire the competence of the geologist, and without requiring that the latter know all the other stakes involved in the decision.

It is necessary to have a way to operationalize the concept of probability, that is assign a number value to the degree of belief. The key tool is the *bet*. A subject assigns a probability $p$ to an event $E$ if s/he is ready to bet a sum of money $pS$ in exchange for winning the reward sum $S$ if $E$ actually happens. In other words, we measure the probability that is assigned by a subject to an event by looking at the odds at which s/he is willing to bet.[11] To

[10] For the connoisseurs, enumerable additivity ($\sigma$-additivity) of probability can be put in question by the frequentist conception even though it is one of Kolmogorov axioms.

[11] The probability is then defined as the inverse of the odds.

dissipate any misunderstanding, several remarks are due. First of all, the bet is supposed to be made for any reward sum $S$ as long as it is not too large (for the sums that are too large people start being subject to risk aversion). In particular the reward sum can be negative, which is the same as betting on the opposite event with probability $1 - p$. Thus, the subject had better bet at the odds that appear fair to her/him, at which s/he is neither losing nor winning. Second, and indeed in order to counter risk aversion, the subject is obliged to make the bet and propose a probability $p$. Last, this faculty of expressing one's degree of belief using the betting method can only be acquired by experience. Bruno de Finetti trained his students by making them bet regularly on the games of the teams of the Italian soccer league, and as time went by they acquired a finer intuition of what probability measures. The probability is not a number given in some arbitrary way; it permits *translating* a conviction as to the happening or not of an event, and translation by the betting method requires practice, experience, to work correctly.

So the probability of an event has a meaning only before the event takes place, since it is here to quantify the uncertainty. Once the event has occurred, only certainty remains and probability no longer has any reason to be. A probability is therefore associated to a singular event, contrary to the preceding conception where it had a sense only for the repetition of a random experiment and so depended on the complete set of results of the different experiments. While the probability has no meaning after the potential realization of the event, it can however be updated before in case new pieces of information become available. Take the soccer game example, if one of the two teams scores, then the subjective probability of the scoring team winning will surely increase: this does not mean that the new probability is more precise or "fairer": the probability simply changes because the pieces of information that are accessible to the subject (here, the fact that a goal was scored) have changed.

Part of the beauty of this method is that it possible to fall back on the usual rules of the calculus of probabilities. To take an example, let us assume that a subject (let's call her Alice) has been led to place a bet on the issue of a soccer game between Bourges and Caen. She must choose among three options: (A) Bourges wins, (B) Caen wins, and, (C) tied game. She assigns to these three options the respective probabilities $p_A$, $p_B$, and $p_C$. Now, let us suppose that $p_A + p_B + p_C \neq 1$: this violates the rules of the calculus of probabilities, since one and only one of the three options will happen. Then, there is a betting system, that is to say a choice of three sums of money $S_A$, $S_B$, and $S_C$, such that if Alice bets on (A), (B), or (C), the respective sums $S_A$, $S_B$, and $S_C$, at the odds indicated by $p_A$, $p_B$, and $p_C$, then she loses money for sure. To be sure to understand this result well, a numerical application may help; so let us suppose that $p_A = 0.5$, $p_B = 0.4$, and $p_C = 0.3$, so that $p_A + p_B + p_C = 1.2 > 1$. Then, if Alice bets 5€ on (A) (and so wins 10€ if (A) is realized), 4€ on (B) (and so wins 10€ if (B) is realized), 3€ on (C) (and so wins 3€ if (C) is realized), then she bets in total $5 + 4 + 3 = 12$€ and wins 10€ whatever the outcome is: she loses 2€ for sure. In a more general manner, with the only constraint being that a subject is coherent, that is to say that s/he does not accept a betting system in which s/he loses for sure, then it is possible to deduce the rules of the calculus of probabilities.

The subjectivist conception possesses the advantage of enjoying a great coherence (at least more than the two others) and of being of great conceptual beauty. It is sometimes accused of authorizing probability assignments that are "crazy" since arbitrary, but in reality one must see the betting method as a means to measure the judgment of a subject regarding an uncertain event, the degree of belief pre-exists to the measure and the latter requires training, experience, to take on a real meaning. Another criticism, or at least another point to clarify, is to know whether the objective probabilities can be made to constrain the subject probabilities: in the case where the frequency of the results of a random experiment that is repeated a large number of times approaches a limit value, must the evaluation of subjective probability of the result of the random experiment agree with the limit of the frequency? This will be discussed in the next section. Still, a major criticism persists: this concept of probability is fundamentally *antirealistic*,[12] which poses a problem when it is used in theoretical physics. In statistical physics, which makes abundant use of the concept of probability, it is wished that probabilities have an objective sense, or at least an intersubjective sense (that is to say the assignment of a probability is the same for all the subjects), because we want the probabilities to reflect a reality that is in one sense or the other independent of us. And in fact the subjective probabilities miss this aspect.

## The law of large numbers

It would be dishonest to discuss these different conceptions of probability without mentioning the result that is supposed to unify them by championing the frequentist interpretation; namely, the law of large numbers. Schematically, it states that *if a random experiment is repeated a large number of times so that the different experiments are independent from one another, then the frequency of the realizations of an event will be close to the probability of this event*. But in this formulation it is not specified what conception of probability, and in fact the status of law of large numbers remains to be clarified: is it a mathematically proven theorem or an empirical law? To answer these questions, it is necessary to explicate the meaning of this law in each of the three conceptions presented above.

### The classical conception

To simplify the situation, we shall consider an experiment possessing only two equally probable issues, "`success`" and "`failure`," so that the probability of "`success`" is equal to $1/2$. If $n$ independent experiments are carried out, the number of equally probable cases for the whole set of $n$ experiments is $2^n$: for example, if 3 experiments are carried out, there are 8 equally probable possible issues:[13]

```
"success, success, success"
"failure, success, success"
"success, failure, success"
```

[12] One of treatises on probability by de Finetti starts with this voluntarily provocative sentence: "Probability does not exist."

[13] The independence hypothesis is indeed expressed by saying that these issues are equally probable.

```
“failure, failure, success”
“success, success, failure”
“failure, success, failure”
“success, failure, failure”
“failure, failure, failure.”
```

The law of large numbers states that if $n$ is large, among the $2^n$ equally possible cases for the series of $n$ experiments a majority will correspond to situations where the frequency of `success` in the $n$ experiments is close to $1/2$. In other words, if we fix a gap to be $\epsilon > 0$ as small as we wish, then for sufficiently large $n$ the probability (as the proportion of `success` cases in over the $2^n$ equiprobable possible cases) that the frequency of `success` in the $n$ experiments is comprised between $1/2 - \epsilon$ and $1/2 + \epsilon$ is close to 1. For example, if we choose $\epsilon = 0.1$, then by carrying out $n = 500$ experiments, the probability for the frequency of `success` in these 500 experiments to be between 0.4 and 0.6 is 0.95.

Let us clarify this result: the two probabilities appearing in this law of large numbers are probabilities in the classical sense; that is, as the ratios of `success` cases over the total number of cases. The law of large numbers is therefore a *theorem*; it is a result that can be proven mathematically and whose proof relies on enumeration: it does not involve any chance. Is the question settled then? No, because this law of large numbers does not say that in a very large number of experiments, the frequency of an event is close to the probability of the event; rather, it says that *it is very probable* that the frequency of the event be close to the probability of the event.[14] To complete this line of reasoning, it should be possible to say that *if an event possesses a probability that is very close to* 1*, then in practice it will occur* (or equivalently, if an event possesses a probability that is very close to 0, then then in practice it will not occur), statement which can be called *unique law of chance*.[15] However, the “classical” definition of the probability does not say why, if the

[14] Connoisseurs will have noticed that only the so-called “weak” law of large numbers, that is of convergence *in probability* and not *almost-sure* convergence, is mentioned. The first reason is that convergence in probability is the one used in practice because it gives useful information for a large finite number of experiments; the second is that it is simpler to present. But more importantly that the strong law of large numbers does not resolve the problem of the unique law of chance (see below), because it is always necessary to explain why an almost sure event (which is not an event that is certain) always occurs in practice.

[15] This appellation in due to Émile Borel. He even gave an estimate of the probability above which an event never occurs in practice: for events that concern a unique individual, it is set to $10^{-6}$. That is to say that if an event possesses a probability that is less that a millionth, then it is safe to behave in practice as if it will not occur. This number is not arbitrary, it comes from the observation that if an individual had to take into account the complete set of events whose probability is less that $10^{-6}$, s/he would quickly be paralyzed and could no longer do anything. If on the other hand an event concerned only a single individual among the whole of humanity, its probability must be less that $10^{-15}$ in order for anyone to act as if it does not occur. For example, at an individual’s scale, a subject must behave as if s/he was going to lose at the game of lotto, but a the scale of a country’s population, it is not longer possible to behave as if everybody would lose at the game of lotto.

probability of an event is extremely close to 1, one has to behave in practice as if it occurs certainly. The classical conception does not explain in a logical manner why we practically never observe events whose probability is almost 0 and always those whose probability is almost 1. This is a hypothesis that falls out of the framework of classical conception, so that within the latter it is not possible to *prove* the convergence of the frequencies.

### The frequentist conception

The frequentist conception gives the impression on the contrary to start from the law of large numbers since it defines probabilities as limits of frequency. In particular, and as it has already been said, convergence of frequencies is an *empirical* observation, it is not provable. However, there does indeed exist a law of large numbers that does not correspond to this hypothesis made of convergence of frequencies, but that specifies the "speed" of convergence. It consists of the same statement as that of classical conception, namely that for any $\varepsilon > 0$ that we pick, it is possible to derive a number $n$ of experiments sufficiently large so that with a probability close to 1, the observed frequency of an event's realization of probability $p$ will be comprised between $p - \varepsilon$ and $p + \varepsilon$, and it is also possible to quantify how large $n$ must be.[16] This quantification is useful because it enables estimating the gap between the "real" probability and the observed frequency, as well as determining the degree to which the measure of the probability made by observing the frequency is precise.

Here too let us clarify this result: first of all, it is *derived* from two premisses: the hypothesis of convergence of frequencies, and the hypothesis of randomness of the sequence of experiments (that is, there does not exist a betting system that can guarantee winning for sure when taking into account past issues of the experiments). In this law of large numbers, the probabilities are to be understood in their frequentist interpretation: namely, that the probability of an event is equal to the limit of the frequency of its realizations if the number of experiments is large, and that the probability that in $n$ experiments the frequency of the realization of an event is comprised between $p - \varepsilon$ and $p + \varepsilon$ (where $p$ is the probability of the event) is close to 1 means that by carrying out a large number of times sequences of $n$ experiments, in the major number of these sequences the observed frequency of the event's occurrence is comprised between $p - \varepsilon$ and $p + \varepsilon$. As well as in the classical case, the frequentist law of large numbers says that *it is very probable* that the frequency of realization of the event be close to its probability, where however the expression "it is very probable" takes another meaning. In the frequentist conception, an event is very probable if, when the random experiment is repeated a large number of times, it is realized in the major number of cases. Taking again the numerical example given above (the one in which the probability of the event "success" has probability 1/2), if we repeat a large number of times sequences of 500 experiments, then in the major number of these sequences, the frequency of success will be comprised between 0.4 and 0.6. But this does not explain why the unique law of chance must apply: in most cases we are confronted with a unique

[16] In its mathematical form, it states that if $n$ experiments are carried out then with probability greater than $1 - \frac{1}{4n\varepsilon^2}$ the gap between the frequency of an event and its probability will not exceed $\varepsilon$.

occurrence of an extremely probable event, not several; yet, the probability of an isolated event makes no sense in the frequentist conception. In practice, we never carry out several sequences of 500 experiments in order to verify that the frequency of success varies only so slightly from one sequence to another. Just like the classical conception, albeit for different reasons, the frequentist conception is not able to justify only with internal arguments the unique law of chance: since probability is always defined relatively to a sequence of experiments, it is not possible to explain why it must impose a specific behavior facing a singular event even if the probability of this event is very close to 1. However, the unique law of chance recommends in fact, when facing a singular event whose probability is close to 1, to behave in practice as if it was to happen. If the frequentist conception can "prove" the law of large numbers, it cannot formally justify the pragmatic use made of it.

### The subjective conception

The subjectivist conception being indeed anchored in decision making, the unique law of chance is natural: almost by definition, if I assign to an event a probability close to 1 then I behave in practice as if it was certain. The necessity that prompts me to act like this is not located in the world, it is within me, in the sense that the probability that I assign corresponds to a judgment, not the state of the external world. This is the reason why the problem of justifying the unique law of chance disappears in a subjectivist conception, whereas in the other conceptions it has to do with how a state of the external world dictates a rational individual a way to behave.

But another problem poses itself then for the subjectivist conception: how can the regularities of the world have an influence on the probabilities that I assign? If for example I observe 100 tosses of heads-or-tails and "tails" comes out 52 times, why should I assign the probability of coming out tails a value close to 1/2? It would even be possible to try to combine both the existence of objective and subjective probabilities, the latter being equal to the former whenever these are available; thus infinitely repeatable events would be described as limits of the frequencies of their realization and the subjective probabilities could apply to singular events. However this is not what de Finetti accepts as valid; he proposes another vision relying upon an elegant mathematical theorem. Let us suppose that a sequence of random experiments with two issues ("`success`" and "`failure`") occurs in front of me, for example as a chance game. Let us suppose that I make the hypothesis that the number of `success` among the $n$ tosses does not depend on the order of tosses, that is to say for example that I assign an equal probability to 1 "`success`" followed by 5 "`failures`" as to 3 "`failures`" followed by 1 "`success`" then 2 "`failures`." This supposition is a subjective judgment that I am led to make if the experiments seem to be identically and independently reproduced. Then, simply with this hypothesis of order indifference and the rules of the calculus of probabilities (which, let us be reminded, are derived from a requirement of coherence), I must modify my probability estimates with the results of the next issues of the game, in the same formal manner as if the frequency of `success` were to converge whenever the number of experiments increases more and more.[17] The "law of large numbers" shows how, by adding this subjective judgment of indifference in the order of successes, probabilistic reasoning *constrains* one to assign to the probability a value that is closer and closer to the observed frequency whenever the number of experiments increases.

What to retain from this discussion? First of all that beyond a vague idea of convergence of frequencies toward the probability, the formulations of the law of large numbers turn out to be more and more subtle. In particular they allow to estimate the number of experiments that must be realized so that it becomes very probable that the observed frequency of realization of an event be close to its probability, which information is extremely useful to know up to what precision the frequencies approach the probabilities. However, sooner or later, this argument becomes circular, since we must explain why in practice, we must behave as if an event that is very probable occurs for sure, which is explained by neither the classical nor the subjective conceptions from their internal structure alone. The subjectivist conception is in a sense built to address this very question, and it can also afford the luxury of explaining why, and in which conditions, it is rational to behave as if the frequencies were to converge as the number of realizations of similar experiments augments. But the problem that remains is that if we do that, this rejects all notions of statistical regularities of phenomena that could be *objectively* described since the probabilities exist only in our mind.

## Conclusion

How could be reconcile these different visions? Rather than counting points to designate a "winner" way that is the best manner to assign probabilities, it is without a doubt more useful to see the study of these irreconcilable viewpoints as an exploration of what is possible.[18] Each of these conceptions thoroughly explores a specific viewpoint on probabilities, exhibits the questions to which the viewpoint can respond, the main advantages as well as the problems that it gives rise to, aspects that are ungraspable. For

---

[17] For the connoisseurs, this formulation is not precise enough. The result of de Finetti (for a modern presentation one can consult J. F. C. Kingman's article, "Uses of exchangeability" in the *The annals of Probability*, 1978, Vol. 6, No. 2, 183–197) can be formulated as follows: if $X_1, X_2,$ … are random variables taking values 0 ("failure") or 1 ("success") so that for any permutation $\sigma$ of $\{1,2,\dots\}$ the law for $(X_1, X_2, \dots, X_n)$ is the same as the one for $\left(X_{\sigma(1)}, X_{\sigma(2)}, \dots, X_{\sigma(n)}\right)$ (order indifference), then the law of the random process $(X_n)_{n\geq 1}$ is a convex combination of laws of i.i.d. Bernoulli processes. This means that there exists a measure $\mu$ of probability over $[0,1]$, such that the probability of having $r$ successes and $n-r$ failures among $(X_1, X_2, \dots, X_n)$ be equal to $\int_0^1 \binom{n}{r} p^r \, (1-p)^{n-r} d\mu(p)$. Thus, the hypothesis of order indifference does not imply that the process is independent; rather, it is a "mixture" of independent processes. The empirical frequency $\frac{1}{n}\Sigma_{k=1}^{n} X_k$ converges then toward a random variable observing the $\mu$ law.

[18] This idea was heard in a conference given by Soazig le Bihan.

example, if the objective probabilities are well-defined attributes of phenomena, just as lengths, they come amiss when it is needed to capture singular events and in particular cannot directly catch the unique law of chance which helps determine our decisions. Only a rigourous analysis of the concepts being used, which we hope to have sketched in this document, will allow reaching such conclusions and in particular to differentiate the real insufficiency from that which only has the appearance of it.

It is neither a static analysis, a viewpoint settled into fixed positions: all the interest of having several interpretations is to be able to navigate from one to the other. Indeed, one must not forget that probabilities can serve an argumentative scheme and feed a rhetorical network. It is precisely their hermeneutic character that has made them be used by such diverse actors as casino gamblers, physicists, sociologists, betting addicts, *etc.*,… To illustrate this diversity of uses, let us look at climate sciences. Climatologists must manage three kinds of uncertainty: the one linked to the natural variability of climate, which can be said to describe itself objectively with the analysis of the frequencies of the phenomena; the one linked to the imperfection of our knowledge of the climatic system; and, the one linked to human behavior, since it is difficult to predict the future state of our civilization and therefore of our emissions of greenhouse gases. The magic provided by probabilities allows, in the reports intended for the general public or decision-makers, to combine these fundamentally different uncertainties into a unique number that can be remembered and used outside the scientific sphere.

A better knowledge of the interpretations of probabilities enables better clarifying their role in an argumentation, notably by pointing out possible slips of meaning: how many times an objective probability, based on statistical observations, is identified to the individual propensity of an individual. How many times games of chance are used to give a meaning to a probability when the context is not appropriate? Such a powerful and polysemic tool as the calculus of probabilities deserves better than just being used without questioning the meaning of the answers it provides.

## Bibliography


- Émile Borel, *Valeur pratique et philosophie des probabilités*, 1939, éditions Jacques Gabay, Paris.
- Bruno de Finetti, "*La prévision : ses lois logiques, ses sources subjectives*," in *Annales de l'I.H.P.*, tome 7, no. 1, 1937, pp. 1–68.
- Bruno de Finetti, *Probability, induction and statistics—The art of guessing*, 1972, John Wiley & Sons.
- Alan Hájek, "Interpretations of Probability," *The Stanford Encyclopedia of Philosophy* (Winter 2012 Edition), Edward N. Zalta (ed.).
- Richard von Mises, *Probability, statistics and truth*, 1957 (*2e édition anglaise*), Dover Publications, New York.
- Jan von Plato, *Creating modern probability*, 1994, Cambridge University press, Cambridge.


## Additional links to related topics

The following bibliography is not part of the original in French; it was added by the translator for self-documentation.


- Henri Poincaré, "La Logique de l'infini," in *Dernières pensées*, chap. 4 (1913).
- Émile Borel, *Le hasard*, Libraire Félix Alcan, Paris (1920).
- Thierry Martin, "Les probabilités négligeables selon Émile Borel," in *Histoire des mathématiques, Images des mathématiques, La recherche mathématique en mots et en images*, Centre National de la Recherche Scientifique (février 2018).
- Philippe Sentis, "La notion de hasard : différentes définitions et leurs utilisations," in *Laval théologique et philosophique*, 61(3), 463–496 (octobre 2005).