What Probabilities Measure

Mikaël Cozic[1]

and Bernard Walliser[2,3]

Prisme N°24

September 2012

[1] Mikaël Cozic is an Associate Professor of the philosophy of science and knowledge at the University Paris-Est Créteil (UPEC) and Director of the "Decision, Rationality and Interaction" team at the French Institute for History, Philosophy of Science and Technology (IHPST), a joint research project of the University of Paris I, École Normale Supérieure (Ulm) and the French National Centre for Scientific Research (CNRS). His research focuses on the philosophy of economics and on formal theories of rationality.

[2] Bernard Walliser is a Professor at the École Nationale des Ponts et Chaussées, a Director of Research at the School for Advanced Studies in the Social Sciences (EHESS) and a researcher at the Paris School of Economics. Author of several books, his research focuses on the connections between economic theory and the cognitive sciences, as well as on the epistemology of economics.

[3] We would like to thank Anouk Barberousse, Philippe Mongin, and Denis and Hervé Zwirn for helpful discussions on this article, as well as Josselin Garnier and Laurent Mazliak for their careful reading of later versions. We are particularly grateful to Isabelle Drouet for essential comments and suggestions.

© Cournot Centre, September 2012

Summary

Probability is one of the fundamental tools that modellers use to describe and explain. It can represent both the properties of all kinds of events (social, psychological or natural) and agents' degrees of belief. Probability raises formidable conceptual challenges, which are the object of the philosophy of probability. The definition of probability is based on an often implicit ontology, and its evaluation raises specific epistemological problems. The purpose of this article is to outline a conceptual framework within which the fundamental categories of philosophers of probability and probabilists can communicate.

© Cournot Centre, September 2012

Contents

© Cournot Centre, September 2012

Introduction

There are a variety of formal models intended to represent agents' beliefs about their environment. Limiting ourselves to models of epistemic logic, beliefs are expressed in two distinct contexts. In a syntactic context, they are defined on propositions[4] by means of individual belief operators, which express the fact that an agent believes one or another proposition to be true. For example, $\mathrm{B}_i p$ means that individual i believes the proposition p to be true. In a semantic context, beliefs are defined on possible worlds by means of individual accessibility relations, indicating the worlds that an agent considers to be possible from a given world. In both cases, agents' beliefs are based on their shared environment, but also on the beliefs of others about these primary beliefs. Under relatively weak conditions, the two contexts are equivalent, thanks to the definition of principles of correspondence.

Beliefs can be seen either as an all-or-nothing matter or as a matter of degree. In a set-theoretic approach, the belief is all or nothing, in the sense that an agent either believes or does not believe that a given proposition is true. In a probabilistic approach, the belief is graduated, in the sense that the agent expresses a *degree of belief* about a given proposition. Probabilistic beliefs can be related to set-theoretic beliefs in syntax or in semantics. In syntax, one can consider that the belief operator applies to a proposition p if the individual's degree of belief in p weakly exceeds a certain threshold, for instance equal to one (Locke's thesis). In semantics, one can consider that the relation of accessibility represents the support of a probability distribution over possible worlds from a given world. Using this correspondence, the properties defined in a set-theoretic context are transferred (in a way that is not always immediate) to a probabilistic approach.

In what follows, we limit ourselves to the study of a single agent forming probabilistic beliefs about his physical or social environment. This means that we do not consider crossed beliefs, that is, one agent's beliefs about what another agent believes. Direct beliefs are naturally expressed in the semantics of possible worlds,

[4] For simplicity's sake, we do not make the customary philosophical distinction between statements and propositions.

© Cournot Centre, September 2012

each world representing one possible state of the environment. As the real world is fixed (by the modeller), an agent's beliefs are simply represented by a probability distribution, or more generally, by a probability on events considered as sets of worlds. A probability is said to be "extreme" if it takes the value 0 or 1; otherwise it is "non-extreme".

This article starts by specifying the conceptual framework in which probability is defined. It then situates the different schools that have proposed particular interpretations of probability within this framework. Finally, it discusses the operations commonly defined on probabilities according to these different interpretations. The aim is thus to build a bridge between the technical and philosophical analysis of probability. For probabilists, it seeks to bring to light the numerous implicit presuppositions that influence their work. For philosophers, it seeks to highlight the formal conditions that are assumed to be satisfied prior to calculation.

The first section introduces the general characteristics of probabilities, while the second section describes the characteristics of the random situation being analysed. The third section presents the main typologies of probabilities and the fourth section crosses them, providing examples of each category thus obtained. The fifth section describes the possible protocols for the empirical measurement of probabilities. The sixth section situates the four main schools that have conceptualized probability over the course of history. The seventh section looks at the combination of probabilities, and the eighth section examines their revision. The ninth section discusses the methodological problems raised by the comparison of categories of probability. The tenth section presents some generalizations of the concept of probability. The eleventh, twelfth and thirteenth sections examine the application of probability to statistics, the empirical sciences and epistemology.

© Cournot Centre, September 2012

1. Characteristics of a probability

Every probability is first characterized by its *attributor*, in this case the individual who expresses the probability. Traditionally, the attributor has one of two different roles, depending on his motives. He may be the *modeller* himself, seeking to build a probabilistic model of the system being studied. For example, the modeller assigns a certain probability to the emission of particles by a radioactive source. Or he may be an ordinary *agent* seeking to represent the context in which he is operating, for the purposes of knowledge or action. For example, a bettor assigns a probability to the likelihood of a certain horse winning a race. Ultimately, however, even when the probability is expressed by an agent, it is the modeller who considers its evaluation by that agent.

Every probability is also characterized by its *object*, in other words a system that produces a certain phenomenon. It is assumed that this phenomenon can be separated from its context and that it occurs in various different modes or instances. In practice, it can be realized in two contexts: static or dynamic, individual or collective. It is called a *serial* phenomenon if it can occur repeatedly in the same system. For example, the modeller assigns a given probability to the outcome of a well-defined throw of the dice. It is called a *populational* phenomenon if it occurs in distinct individuals of a same population. For example, the modeller will assign a given probability to the likelihood that an individual in a given population suffers from a particular disease.

Lastly, every probability is characterized by the date at which it is applied to the phenomenon, in relation to the date at which the phenomenon occurred. *Ex ante* probabilities are defined before the occurrence of the phenomenon. Thus, a voter expresses a probability that a given candidate will be elected before the election has taken place. *Ex post* probabilities are defined after the occurrence of the phenomenon. Thus, a jury member pronounces a probability that the defendant is guilty after the crime has been committed. Furthermore, the date at which the phenomenon is evaluated may differ from the date at which this evaluation is pronounced. Today, a researcher may consider that the probability of discovering an HIV vaccine was a given value yesterday and will be another value tomorrow. These

© Cournot Centre, September 2012

two dates usually coincide, however, as the probability assigned is applied instantaneously.

Returning to the object of the probability, it is always possible to consider successive occurrences of a serial phenomenon as a population, in which case one can treat it as a populational case. Conversely, one can consider that the individuals in a population are chosen sequentially (in a random order to be specified), thus defining successive occurrences of the phenomenon. It is worth noting, however, that a phenomenon can be both serial and populational at the same time. This is the case, for example, when the same horses compete in several successive races. If the attributor of the probability is the bettor, the object of the probability is the result of one of the races, and more precisely the position in which a given horse finishes the race.

Whatever the phenomenon considered, the attributor can attribute a probability from one of two different perspectives, which depends on the way we consider the sequential or parallel experiments (or trials) in which the phenomenon occurs. The phenomenon is *specific* if the attributor is interested in a single instance of its occurrence (a *token*). For example, the modeller is interested in the next throw of the dice, or the doctor in the illness of the patient sitting in front of him. The phenomenon is *generic* if the attributor is interested in a set of occurrences (a *type*). For example, the modeller is interested in an undefined throw of the dice, or the doctor in the illness of any patient.

Furthermore, the phenomenon can be defined in relation to a more or less finely differentiated system, but the modes of the phenomenon are assumed to be the same in each set that is identified. The phenomenon is *univocal* if the system involved is considered to be identical in all the instances of its occurrence. Successive throws of well-defined dice or the illness of patients belonging to a homogeneous group can be considered as such. The phenomenon is *multivocal* if there is a precise typology of the systems involved that produce the same phenomenon, with a probability being defined for each of the categories considered. This is the case for the probability relating to dice of different shapes and colours, or populations of patients of different ages and genders.

© Cournot Centre, September 2012

2. Analytical framework

As presented above, the modeller's description of the object of the probability involves three types of criteria. *Phenomenal criteria* make it possible to differentiate the possible modes of the phenomenon under examination. They have to be identical in each achievable experiment. *Factual criteria* make it possible to characterize the conditions of each specific experiment of the phenomenon (if it is generic). They can vary from one experiment to another. *Structural criteria* make it possible to define and classify the systems at the root of the phenomenon (if it is multivocal). They are invariant across different experiments.

In practice, the modeller is guided by the causal relations that she knows to exist between these criteria. She makes the hypothesis that the factors explaining the modes of the phenomenon comprise both the factual and the structural criteria. Conversely, the phenomenal criteria express the effects generated by these factors. The probabilities announced depend solely on the structural criteria, however, which exert a deterministic influence. They are not affected by the factual criteria, the influence of which – whether or not it is deterministic – is incorporated in the evaluation of the probability. Thus, the modeller is not seeking to build a model to describe completely the influence between criteria, but only to measure the dependence of the probabilities (and therefore of the phenomenal criteria) on the structural criteria.

In the case of a horse race, the phenomenal criterion is simply the position in which a given horse finishes the race. The factual criteria include the weather, the heaviness of the track or the jockey's skill, but also the nature of the saddle or the time of the race. The structural criteria include the age and sex of the horse, its racing stable and bloodline and the identity of its owner. The attributor will then, for example, define the probability of a given horse winning a race as a function of its age and sex. If he possesses additional information about the racing stable, he will refine the category to which the horse belongs and adjust its probability accordingly. The state of the track, on the other hand, varies from one race to another and is one of the causes that is incorporated in the probability.

© Cournot Centre, September 2012

In this way, every probability is defined within a *reference frame*, which precisely sets the characteristics of the system and the modes of the phenomenon. Conventions of equivalence establish that it is indeed the same phenomenon that is reproduced in parallel or sequential instances (fixed structural characteristics, variable factual characteristics). The first condition is that the situations compared should be similar. This condition must be satisfied in every set generated by the differentiation of the system. The second condition requires that the experiments considered be independent of each other. In a temporal context, this means that the experiments should not be repeated at intervals that are either too close together (to avoid correlation) or too far apart (to avoid excessive change). In a spatial context, it means that the individuals in the population should interact in a stable manner with regard to the phenomenon in question.

In the case of the dice, they should not be modified in any way over the course of successive throws, and every throw should be governed by the same protocol. The dice are also assumed to be exempt from wear and tear and to have no memory of past occurrences. In the case of the disease, the patients are considered to be interchangeable and subject to the same environmental conditions. They must also exert reciprocal influences that are fixed over the short term (contagion). In the case of a horse race, it is assumed that the horses do not age and that the racing conditions (weather, state of the track) do not vary or are insignificant. In particular, the horses' past results should not influence the results of future races, but this condition is often unfulfilled (handicaps).

The reference frame is, of course, a deliberate construct of the modeller. Thus, although the attribution of a probability is an epistemic operation, its definition is based on ontological presuppositions. The categories that characterize the phenomenon, like the similarities between phenomena, are defined by the modeller. When the attributor of the probability is an ordinary agent, it is also assumed that he adopts the same typologies as the modeller. Furthermore, the reference frame pre-exists the evaluation of the probability and cannot be directly affected by the results of the trials. Thus, if the modeller has pre-defined the structural criteria that differentiate the systems, these criteria cannot be brought into question by

© Cournot Centre, September 2012

observation of the results of experiments (the problem of resilience). If a coin is judged *a priori* to be unbiased (see §5), then an observed imbalance between "heads" and "tails" cannot lead to the conclusion that the coin is biased.

3. Typologies of probabilities

3.1 Basic typology

In the above explanatory framework, we can pursue our analysis of the factual criteria (hereafter referred to as factors) from the attributor's point of view. A factor is *explicit* if the attributor knows that it exerts some influence over the phenomenon. It is *implicit* if the attributor is unaware of the effect it may have on the phenomenon, in one way or another. In some cases, the attributor cannot identify any explicit factor making it possible to distinguish between two experiments, and the phenomenon is said to be *primitive.* A typical example is a radioactive source randomly emitting various types of radiation. Such a phenomenon cannot be related to any identified prior causes and must be studied independently. In all other cases, at least one explicit factor is identified and the phenomenon is said to be *emergent.* Through iteration, such a phenomenon can be related to prior factors, which may themselves be generated by even earlier factors.

Furthermore, an explicit factor is *known* if the attributor can observe the value it takes in each instance of the phenomenon. An explicit factor is *unknown* if the attributor cannot observe it with precision, for either physical or ethical reasons. Knowledge of a phenomenon varies, however, depending on whether the phenomenon is considered to be specific or generic. In a specific throw of the dice, the initial impetus is deterministic, but is in practice impossible to measure. In a generic throw, the impetus varies, and to an outside observer, the phenomenon appears to be completely random. Generic dice throwing corresponds to a random phenomenon deliberately constructed by the attributor. In a way, it transforms explicit factors into implicit ones.

Finally, an implicit factor is *regular* if the attributor can consider that its influence is governed by a random law, whatever that law might be. An implicit factor

© Cournot Centre, September 2012

is *irregular* if the attributor considers its influence to be erratic, in other words it cannot be represented by a probability law. This classification, however, depends on the sophistication of the analysis of the phenomenon. It is very difficult to distinguish between a stochastic phenomenon and a phenomenon produced on some other basis, especially if it is produced artificially without using random draws. A scientific or technological innovation to which one assigns a probability despite its erratic nature is a typical example of an irregular phenomenon.

By crossing these last two criteria, we can produce a first typology of probabilities, in four large categories (in this table, by convention, an absent factor is considered to be regular and known).

Implicit factors **Explicit factors**	**All regular**	**Some irregular**
All known	irreducible probability: *radioactive source*	approximate probability: *technological innovation*
Some unknown	controlled probability: *dice throwing*	radical probability: *horse racing*

An *irreducible probability* relates to a phenomenon considered to be primitive (or with factors that are all regular and known). The probability then expresses an intrinsic property of the phenomenon, independent of all prior factors. Thus, the radioactive source emits particles according to classic random laws, which are not derived from probabilities related to deeper factors. The probability is irreducible, because the phenomenon is assumed to be intrinsically random; it cannot be reduced to a kind of determinism with hidden variables. Quantum mechanics is often considered to be the only expression of a non-deterministic phenomenon (at least in the natural world). At a more macroscopic level, the random character disappears due to the aggregation of microscopic laws.

A *controlled probability* relates to a phenomenon that is explained by factors that are all explicit, but some of which are unknown. The probability then sums up the attributor's ignorance of the value actually taken by these factors. Dice throwing,

© Cournot Centre, September 2012

realized by human hand or machine, and stated in its specific form, is the simplest illustration of this. The probability is controlled if we hold that the basic process is perfectly deterministic and the attributor cannot observe the initial conditions of the throw, which prevents her from predicting the result (from among several possible values) by means of the laws of mechanics (see §4.1). In this case, we are dealing with a "random event in the sense of Cournot", due to the crossing of two independent causal chains. In the throw of the dice, the moving dice meet the surface on which they roll in a complex way.

An *approximate probability* relates to a phenomenon of which the explicit factors are known and the possible implicit factors prove to be irregular. The probability then reflects the best estimate that one can make of the phenomenon's occurrence. Technological innovation appears to match this description if we set aside the socio-economic conditions of its occurrence. The probability is approximate if we hold that it sums up the irregular implicit factors by assigning a certain probability to them. Here again, in the social sciences, innovations are often considered to be the only phenomena of a random nature, apart from possible elements of free will associated with behaviour.

A *radical probability* relates to a phenomenon that is explained by a combination of irregular implicit factors and unknown explicit factors. The probability then expresses the joint influence of these two types of factors (which can be confused). This is the case for horse races or any competition between several entities (elections). The factors involved are numerous and their influences tangled. The probability is radical in the sense that it simply introduces a certain regularity where there was no evident regularity *a priori*.

These distinctions are made by the modeller on the basis of his current knowledge, and they can vary over time. So, for example, the modeller may discover explicit factors that transform an irreducible probability into a controlled probability. Radioactivity may be explained by deeper factors behaving as "hidden variables". Conversely, implicit factors may be discovered that transform a radical probability into a controlled probability. Thus, for an election in a limited setting, all the relevant factors can be identified if not known. Lastly, unknown factors may become more

© Cournot Centre, September 2012

easily observable, transforming a random phenomenon into a deterministic one. This would be the case if the impetus of the throw of the dice could be understood, making it possible to predict its result perfectly.

3.2 Ontological and epistemic probabilities

According to traditional philosophy, a second typology of probabilities is based on the distinction between a property attributed to the phenomenon itself and a property located in the mind of the modeller:

- an *ontological probability* (*chance*) reflects the degree of occurrence of a phenomenon, independently of the mental states of its attributor; it takes the value α for an event E if this event has an intrinsic probability $Ch(E) = \alpha$ of occurring;
- an *epistemic probability* (*credence*) reflects the degree of belief that the attributor has in the occurrence of the phenomenon, taking into account her uncertainty; it takes the value α for an event E if the attributor i believes its occurrence to the degree $Cr_i(E) = \alpha$.

The ontological probability of an event is the *object* of the attributor's judgement, just as any physical property of an entity can be taken as an object by the attributor. In the same way that she may evaluate well or badly the weight or length of an object, so she can evaluate well or badly the probability assigned to an event, measured according to what is assumed to be a cardinal scale (see §10). By contrast, the epistemic probability assigned to an event is a *mode* of the judgement that the attributor makes on that event. The spectrum of values of probabilities can then be more or less fine-grained, refining the basic three-way division (believing the event will occur, that it will not occur, and neither believing that it will or will not occur).

Of course, the ontological (postulated) probability $Ch(E)$ that is the object of the attributor's judgement remains to be discovered. The attributor can announce her estimate of this probability, which is denoted $Ch^*(E)$. The latter has an epistemic nature and can be interpreted as a sure belief of the attributor about $Ch(E)$. A pluralist (see §3.4) could even use a hierarchy of probabilities (see §10) and analyse

© Cournot Centre, September 2012

the estimate Ch*(E) as an epistemic probability defined on an ontological probability: Cr_i (Ch(E) = α) = 1. This situation is analogous to that concerning the speed of light, which is not actually known, but the estimated value is nevertheless almost universally accepted.

3.3 Objective and subjective probabilities

In the same vein, an alternative typology of probabilities is based on the fact that the phenomenon may or may not be perceived in the same way by all the attributors:

- an *objective probability* characterizes a situation in which every attributor has sufficient information about the phenomenon and is capable of perfect reasoning on that information; such a probability is *universal* in that it is the same for all the attributors; for an event E, it is denoted Po(E);
- a *subjective probability* characterizes a situation in which an attributor's information is incomplete and/or her rationality limited; such a probability is *personal* in the sense that it differs from one attributor to another; for an event E and an attributor *i*, it is denoted Ps_i(E).

Unlike the previous distinction, this one is more a matter of degree than of binary opposition, since a probability can be viewed as more or less objective. The distinction remains relatively blurred, because it depends on an idea of perfect information and perfect rationality (epistemic objectivity). In theory, information is perfect when all the relevant factors have been made explicit and observed – which remains wishful thinking. In practice, information is held to be perfect from the moment that additional information no longer modifies the chosen probability, which introduces a sort of self-reference. As for perfect rationality, it depends on respect of probability axioms (see §7).

In the case of epistemic probabilities, the question of whether the link between information and probabilities is univocal is highlighted by the two following extreme attitudes:

© Cournot Centre, September 2012

- for the radical objectivist, each body of information univocally determines the degrees of belief adopted by a rational attributor; a corollary affirms that the probability can only be modified if the attributor receives additional information;[5]
- for the radical subjectivist, for any given body of information, all degrees of belief are possible for each attributor, provided they obey the probability axioms.

4. Comparison of typologies

4.1 Crossing the typologies

If we cross the first two typologies, we can consider that the irreducible and approximate probabilities (the first line in the first table) are of an ontological nature. They correspond to factors that are all known, and they only differ in their modes of influence. In contrast, controlled and radical probabilities (the second line in the table) are epistemic. They involve factors that are poorly known by the attributor. This amounts to saying that the only element of uncertainty for the attributor is that he cannot know some of the explicit factors. This is obvious in the cases of both dice throwing and horse racing. Another element of uncertainty, however, concerns the list of factors that influence the phenomenon, whatever the possibility of measuring them.

If we cross the first and third typologies, we can consider that irreducible and controlled probabilities (the first column in the table) are objective. On the other hand, approximate and radical probabilities (the second column in the table) are subjective. This amounts to saying that the main reason for objectivity is the fact of not having implicit criteria or considering them to be perfectly regular. The probability of a radioactive source is objective, because the decay law is perfectly known. The probability of a throw of the dice is objective, because it is the result of a

[5] For a recent discussion of the idea that a body of information univocally determines probabilities, see Levi (2010).

© Cournot Centre, September 2012

calculation based on objective data. As the laws governing the movement and landing of the dice are given by mechanics, it is possible to define, for each initial condition, the result of the throw. In the space of initial conditions (direction and speed of the throw), this gives rise to basins of attraction for each of the possible outcomes. Moreover, if we observe that these basins are distributed homogeneously over the space (because they are governed by the principle of indiscernibility), we can deduce that the probabilities of the different outcomes are uniform.[6]

Finally, if we cross the last two typologies, we obtain our initial table, but with different headings.

	Objective probability	**Subjective probability**
Ontological probability	irreducible probability *radioactivity*	approximate probability *technical innovation*
Epistemic probability	controlled probability *dice throwing*	radical probability *horse racing*

At first glance, the four families of probabilities obtained by crossing the two typologies appear to be cogent. It could also be argued, however, that the last two typologies introduced are highly congruent. First, an ontological probability appears to be more objective insofar as the attributor really can access the necessary information. Second, an epistemic probability appears to be more subjective insofar as a degree of belief is, *a priori*, only based on incomplete information.

[6] The idea that objective probabilities can emerge from the properties of a deterministic physical system, and in particular from the way that the initial conditions are associated with the types of final conditions, can be compared with the "method of arbitrary functions" that is often ascribed to Poincaré (1896/1912). It has given rise to a series of results of which von Plato (1983) gives an overview. These results show that for certain physical systems, a very large set of distributions over initial conditions lead (approximately) to the same distribution over final conditions. These results have played an important role in recent discussions about the possibility of objective probabilities within deterministic systems (see Strevens, 2011).

© Cournot Centre, September 2012

Nevertheless, the mixed categories do have an intuitive validity. First, it is possible to make a judgement of ontological and subjective probability if a random property is attributed to a phenomenon without a consensus on its evaluation. So, for example, although radioactivity is an objective physical phenomenon, its measurement may be the subject of debate between scientists. Second, it is possible to make a judgement of epistemic and objective probability if the degree of belief expressed is based on sufficient information to oblige every attributor to endorse it. So, for example, all the bettors might agree on the same estimation for the result of a race. The perimeter of these two situations of attribution depends on the norms that the attributor adopts for the formation of degrees of belief.

The main problem is to know whether the two mixed forms of probability (approximate, controlled) can be reduced to the two pure forms (irreducible, radical). First, a controlled probability can be reduced to an irreducible probability through the artifice of moving from a specific situation to a generic one. The conditions governing the throw of the dice are obscured and the attributor only considers the results of successive throws, which are no longer different from the rays emitted by a radioactive source. Second, an approximate probability can be reduced either to an irreducible probability – if a regularity is discovered behind the erratic appearance – or to a radical probability if some hitherto unknown criteria are revealed.

Consequently, it is common philosophical practice to confuse the two pairs of concepts, although the first pair (epistemic vs. ontological) is better defined conceptually (Carnap, 1950/1962; Gillies, 2000). Probabilists generally do not distinguish between the two pairs of concepts either, mainly using the second vocabulary (objective vs. subjective), which appears to be more operational. In what follows, we will introduce whichever of the two distinctions most enlightens the particular problem studied.

© Cournot Centre, September 2012

4.2 Attitudes towards multiple interpretations

The different interpretations of probabilities are reflected in distinct schools of thought (see §6). Attitudes towards this plurality of interpretations are themselves eminently variable. The *monists* believe that only one of these interpretations is legitimate, whereas the *pluralists* accept a plurality of interpretations without further ado. The *non-contextualists* believe that each recommended interpretation should be independent of the situation analysed, while the *contextualists* consider that the situation can modify the interpretation.

By crossing these two typologies, we obtain four possible positions, where the example (a, ab) corresponds to two interpretations a and (a and b) successively applied in a first and second situation.

	Non-contextualist	**Contextualist**
Monistic	one sole interpretation valid in every situation: (a,a)	one sole interpretation that varies according to the situation: (a,b)
Pluralistic	several interpretations, each valid in every situation: (ab,ab)	diverse interpretations in every situation: (a,ab)

This difference in attitude is particularly relevant for the distinction between ontological and epistemic probabilities. Monists consider that only one of these conceptions is legitimate or necessary in all relevant situations of probability attribution. Ontological monists uphold that $Cr_i(E)$ can be reduced to $Ch^*(E)$, as each degree of belief is in fact only a belief in a chance of occurrence. Epistemic monists argue that $Ch^*(E)$ can be reduced to $Cr_i(E)$, that is, that a belief in chance is no more than an ontological projection of a personal degree of belief. Conversely, pluralists accept several conceptions at the same time, as $Ch^*(E)$ and $Cr_i(E)$ simultaneously make sense, although their meanings are different. They may then consider that in some contexts of probability attribution, only one of the conceptions is necessary, whereas both conceptions are required in other contexts.

© Cournot Centre, September 2012

5. Measuring probabilities

We can define four modes of measuring probabilities, based successively on their past occurrences, their immediate characteristics and their future consequences. More precisely, there are four natural modes of measurement by the modeller, each of which corresponds to one of the categories of probability defined above.

	Objective probability	**Subjective probability**
Ontological probability	irreducible probability: *frequency*	approximate probability: *declaration submitted to subjective constraints*
Epistemic probability	controlled probability: *estimate subject to objective constraints*	radical probability: *revelation*

Irreducible probabilities are typically inferred from a *frequency* in a sequence S of experiments. Of course, this evaluation only applies if the phenomenon is repeated in analogous conditions. The frequency of a given modality of the phenomenon is simply the number of experiments that resulted in this modality divided by the total number of experiments. The actual frequency is relative to a finite sequence S that only allows one to approximate the probability. The hypothetical frequency (Reichenbach, 1949/1971; von Mises, 1928/1957) is relative to an infinite virtual sequence, which converges towards the probability (if it exists) by virtue of the law of large numbers.[7] Thus, one can estimate the probabilities of decay of radioactive particles by observing the corresponding frequencies.

[7] The law of large numbers can be written: $\mathrm{Pr}\left(\left| \mathrm{Ch}(E) - \mathrm{Fr}_n(E) \right| \leq \varepsilon \right) \to 1$ if $n \to \infty$, involving three probabilities: $\mathrm{Ch}(E)$ is the assumed true ontological probability of the phenomenon, $\mathrm{Fr}_n(E)$ the frequency of the phenomenon after n trials, and Pr the calculated probability of a hypothesis for the modeller.

© Cournot Centre, September 2012

Controlled probabilities are estimated on the basis of *objective constraints* that bound and may even determine the values of probabilities. These constraints relate to factual characteristics, explicit and unknown, which determine the phenomenon in question. They are generated by the structural characteristics of the system under study. This is true for the conditions of symmetry or invariance that are imposed on some phenomena through the laws that govern them. Thus, the throw of a die of regular form with any impetus leads to the prediction of equiprobability in the occurrence of the die landing on one of its faces.

Approximate probabilities are measured by simple *verbal declaration* of the agents, bounded by subjective constraints. They result from pure and simple questioning of the agent, provided she is willing to answer. They are often considered unreliable, however, compared with the deeper probabilities the agent is supposed to possess. They may be deliberately biased for strategic reasons involving the supply of information to others. They are also unintentionally biased by the agent's distorted awareness of her own beliefs. Thus, each individual can declare the probability that she assigns to a technological innovation according to her subjective knowledge of the processes of innovation.

Radical probabilities are obtained by *revelation* on the basis of the choices made (under uncertainty) by the agents that carry them. They result from a process of abduction held by the modeller, who uses the actions observed to reveal the underlying probabilities. For example, he can hypothesize that an individual believes to the degree α that a certain horse is going to win, if α is the maximum sum the individual is willing to place on a bet that pays her one euro if the proposition is true and nothing otherwise. The agent is not necessarily aware of the probabilities revealed in this way. These probabilities are often multivocal, inasmuch as they result from non-deductive reasoning. They are based on a postulated model of decision-making, usually the maximization of expected utility, and they require knowledge of the agent's preferences concerning the consequences of her choices. Indeed, these preferences, in conjunction with the agent's beliefs, determine her choices. So, for example, the bets that an agent places on a horse race reveal her underlying beliefs.

© Cournot Centre, September 2012

A same method of measurement can perfectly well apply to different categories of probabilities. In practice, the methods differ essentially according to whether the probabilities are objective or subjective. For the former, frequencies apply from the moment that the phenomenon is truly repetitive (over time or within a population). They can be applied equally well to radioactivity as to dice throwing. Estimates based on objective constraints apply to any phenomenon for which one has sufficient indications about its causes, which is more restrictive. They are valid for dice throwing but not radioactivity. For subjective probabilities, bets on the possible outcomes of the phenomenon can be used, provided agents are prepared to submit to this process. They concern elections as well as innovations. Declarations are often more difficult to obtain, at least in a quantitative form. They are valid as long as the modeller has a sufficiently well-structured model of the phenomenon. They apply to innovations, but can also apply to elections.

Conversely, probabilities measured by different methods involving different agents can be assigned to one and the same phenomenon. Thus, the probability that a list of symptoms is associated with one or another disease can be evaluated by all four methods. The frequency with which the symptoms lead to the disease can be measured over a sample of patients. Estimates based on constraints can also produce orders of magnitude of the ratios of probabilities according to the closeness of the underlying models. Moreover, a doctor could make an estimate of this probability on the basis of his past qualitative experience. Lastly, the probability can be inferred from a bet that the doctor makes about the state of the patient through the action that he decides to take. Of course, these measurements may diverge from each other, insofar as they are governed by different logics and use different information.

© Cournot Centre, September 2012

6. Classical interpretations

In the continuation of the general typologies presented in §3, let us turn to more specific interpretations of probabilities[8], which have been proposed over the course of time. These interpretations are intended to provide a conceptual and sometimes methodological analysis of a concept of probability used in scientific practice or everyday life. They seek to answer specific questions that arise in the use of probabilities.

The frequentist approach, pioneered by Jacques Bernoulli (1713) and developed by Fisher (1925) and von Mises (1928/1957), is based on the (asymptotic) *identification* of probabilities with the frequencies of occurrence of generic events in successive experiments (or with proportions, in the case of parallel experiments). It therefore rejects the attribution of probabilities to specific events: according to von Mises, "this is one of the most important consequences of our definition of probability".[9] Frequentism is essentially justified by the law of large numbers, which affirms that the frequency of an event converges towards its probability. Convergence of this random variable happens either in probability (weak law) or almost surely (strong law). The frequentist interpretation is intended to provide an ontological concept of probability: probabilities are supposed to express empirical properties of the agent or modeller's environment, which they can then discover.

The propensity approach, introduced by Popper (1959), also relates to the ontological concept of probability. Propensity describes the causal power exerted by certain material conditions in the production of the phenomenon studied. It was developed to explain the probability of specific events, but actually exists in two versions. According to the first (*long run*, Popper, Gillies), the propensity of a set of repeatable conditions is their tendency to produce frequencies when they are instantiated repeatedly. According to the second (*single case*, Popper, Miller, Fetzer), propensity expresses the power exerted by a complete set of relevant conditions to produce a specific event. Ontologically, the propensity interpretation is less

[8] See Hajek (2010) for a general discussion.

[9] Quoted by Gillies (2000), p.115.

© Cournot Centre, September 2012

parsimonious than frequentism, because it postulates that the system has capacities to act (causally) on the phenomenon in question. Methodologically, it is less operational in that it does not provide a specific protocol for measuring probabilities.

The logical approach was introduced by Keynes (1921) and Carnap (1950/1962).[10] "Logical" (or "inductive") probabilities are sometimes considered qualitative (see §9), but they can also be quantitative. Above all, they apply more to propositions than to material phenomena. More precisely, if p and q are two statements, the logical probability $Ind(p,q)$ is a conditional probability expressing the degree to which p entails q (maximal if p entails q and minimal if p entails non-q). In this way, they make it possible to extend the relations of logical consequence to probabilized relations. The status of these probabilities is supposed to be the same as the one for logical notions: they are neither psychological properties of agents nor empirical properties of our universe. Nevertheless, they are close to the epistemic interpretation of probabilities, under the following hypothesis: if all the information possessed by an agent can be expressed in the proposition e, then the degree to which he should believe a hypothesis h is given by $Ind(e,h)$. Logical probabilities, however, are also presented as being objective, and even as the best illustration of radical objectivism (see §3.3). This approach, however, is not directly operational, because it does not propose any precise protocol for measuring probabilities.

The subjectivist approach was developed by Ramsey (1931), de Finetti (1937) and Savage (1954/1972). It proposes an epistemic and subjective interpretation in terms of degrees of belief, but which is intended to be operational. It is based on the idea that beliefs are not directly observable, but are revealed in the virtual or real actions that are based on them, particularly bets. It can then be applied to a specific experiment, provided the agent is willing to express a choice relating to the phenomenon concerned. Beliefs can therefore be measured thanks to their "causal" properties in the agent's action. This approach is sometimes backed up by theories of convergence, which show that the subjectivist probability converges towards an asymptotic value (independent of the prior probability) if it is revised using Bayes' rule (see §7) with sufficiently rich information.

[10] For a recent defence, see Maher (2006). For a succinct presentation, see Hajek (2010, §3.2).

© Cournot Centre, September 2012

Lastly and more recently, an objective Bayesian approach (Jaynes, 2003; Williamson, 2009) has been proposed. It is Bayesian in that it seeks to provide an epistemic interpretation of probabilities. It is objective in that it argues, in contradiction to the radical subjectivists (de Finetti), that the constraints weighing on the beliefs of a rational agent are much stronger than the simple respect of probability axioms. Two additional constraints are added: alignment with partial empirical frequencies (calibration) and the respect of theoretical constraints (usually the maximization of entropy).

There is no univocal correspondence between the four classical interpretations and the categories outlined in §4. There are, however, some evident relations that must be pointed out. The frequentist and propensity approaches are clearly ontological, the former also claiming to be objective. The subjectivist approach applies above all to situations of radical, epistemic and subjective probability. The logical approach is more ambiguous, because it has both an epistemic dimension (in the evaluation of probabilities) and an objective dimension (in the calculative use of probabilities). Furthermore, some authors are non-contextual *monists* (Keynes, 1921; von Mises, 1928/1957; Popper, 1959; de Finetti, 1974) while others are *pluralists* (Ramsey, 1931; Carnap, 1950/1962; Mellor, 1971; Lewis, 1980/1986; Gillies, 2000).

7. Combining probabilities

For a long time, the most influential definition of probability, attributed to Laplace (1812), was the following: the probability of an event is the ratio of the number of cases where the event occurs to the total number of possible cases. The modern mathematical definition of probability, which takes into account the possibility of an infinite number of cases, can be traced back to Kolmogorov's (1933/1950) axiomatization of probability. Probability is considered as a numerical function whose domain is an algebra $\mathcal{A}$ over a non-empty set W of worlds. The function P from $\mathcal{A}$ to $\mathbb{R}$ is a probability function if and only if the three following axioms are satisfied:

© Cournot Centre, September 2012

(P1) non-negativity: $P(E) \geq 0$, for all $E \in \mathcal{A}$

(P2) normalization: $P(W) = 1$

(P3) finite additivity: if $E \cap E' = \emptyset$, $P(E \cup E') = P(E) + P(E')$, for all $E,E' \in \mathcal{A}$

Both the set W and the set $\mathcal{A}$ result from conventions made by the attributor. The first is never exhaustive in terms of the characteristics of the phenomenon to be taken into account. The second does not consider all the events that differ from each other materially, but only a subset. The conditions P1 and P2 are essentially technical. Only the condition of finite additivity P3 is a real constraint imposed on probabilities. It can be extended to a condition of countable additivity P'3 by considering a countable number of elements of $\mathcal{A}$. It has the following property as a consequence:

(P'3) monotonicity: if $E \subseteq E'$, then $P(E) \leq P(E')$

Ontological probabilities satisfy Kolmogorov's axioms (at least with countable additivity). First, we can verify that frequencies in finite sequences of experiments satisfy these axioms. This result also holds in the framework proposed by von Mises.[11] It is less evident for other ontological interpretations. Following Lewis (1980/1986), some authors use a strategy of circumvention. They postulate that epistemic probabilities align themselves (under certain circumstances) with ontological probabilities (see §9) and they deduce that the former obey Kolmogorov's axioms to the extent that the latter do.

Epistemic probabilities also satisfy Kolmogorov's axioms (including countable additivity). Thus, Savage (1954/1972) subjects an agent's preferences over acts to a list of axioms, some of which are specific to probabilities (in practice, ordinal

[11] See Howson and Urbach (1993), pp. 208–209.

© Cournot Centre, September 2012

probabilities, see §9). Consequently, the agent's degrees of belief obey the basic axioms of probability theory, as long as we accept certain hypotheses about the way these degrees of belief are expressed in the agent's choices. This result generalizes the *Dutch Book argument* (de Finetti, 1937), which applies to the bets made by an agent if his utility is known to the modeller. It affirms that if the agent does not respect the usual probability axioms, the modeller can build sequences of lotteries that lead to the sure ruin of the agent.

Logical probabilities raise a particular problem in that they are often applied to propositions, and notably hypotheses. These hypotheses no longer correspond necessarily to observable phenomena, but to underlying structures. In particular, they express relations between phenomena and more precisely between the phenomenal and structural criteria of a phenomenon. They may be deterministic or random. The probability of a hypothesis is usually interpreted epistemically and expresses a degree of belief. The hypothesis is assumed to be true or false, but the attributor is uncertain about its truth value. Nevertheless, the probability of a hypothesis can also be interpreted ontologically by considering that it is written into nature. Nature is then assumed to be random, not only in its direct manifestations, but also in its deeper structures.

In many respects, however, these hypotheses obey the general framework described previously. Thus, a hypothesis allows for distinct modalities even if one at most is assumed to be true.[12] It is therefore possible to build events that are sets of hypotheses and that are themselves probabilized. It gives rise to successive experiments, which constitute so many empirical tests of the hypothesis. It is not the modalities of the hypothesis that are observed, however, but the data relating them to the phenomena studied. In practice, the hypothesis (obtained by abduction or made *a priori*) is tested in its modalities with regard to these data. Its probability results from a backward inference from effects to cause, seeking to assign a probability to the causes when the consequences are known (see §8). For example,

[12] For a well-formed hypothesis, there are at least two corresponding modalities: one where it is false and one where it is true. When the hypothesis sets the value θ of a parameter in a space Θ, then we can consider by extension that Θ is the set of modalities of the hypothesis.

© Cournot Centre, September 2012

one might hypothesize that a coin is unbiased and then carry out a series of tosses to test the validity of the hypothesis.

8. Probability change

Probability change consists in transforming a prior probability P- into a posterior probability $P+$ following the reception of a message $M \in \mathcal{A}$. Most change rules consider the algebra $\mathcal{A}$ as invariant over the set of all possible worlds. This means excluding new possibilities. The most widely used change rule is the Bayesian rule of conditionalization (which only applies if the prior probability of the message is non-zero):

$$P+(E) = P(E / M) = P-(E \cap M) / P-(M) \text{ if } P-(M) > 0$$

In this rule, the weight of the worlds that are excluded by the message are allocated homothetically to the worlds that are still possible. Other rules have also been proposed. In Lewis' (1976) "imaging" rule, the weight of a world that is excluded by the message is allocated to the closest world that remains possible (within an appropriate distance).

In the literature, three contexts of change are traditionally considered according to the content of the messages received. The last ones involve both phenomenal and structural criteria and are presented in the form of events:

- in the context of *revising*, the message excludes certain modalities of the phenomenon. It may or may not contradict the initial belief;
- in the context of *updating*, the message indicates that certain modalities of the phenomenon have changed in a given direction. It does not contradict the initial belief inasmuch as the system concerned has changed;
- in the context of *focusing*, the message restricts the reference class of the phenomenon (to a sub-population of the whole population, for example). It thus changes the attributor's point of view without really supplying any new information.

© Cournot Centre, September 2012

In the medical example, patients suffer from variants A, B and C of a given disease in proportions that are influenced by some of the patients' characteristics (sex, age). A revising message indicates that, given the prevailing local conditions, the population considered is not affected by the variant A (or that women are never affected). An updating message indicates that, given the appearance of a new drug, all the patients suffering from variant A (or all the women) have recovered. A focusing message indicates that the variants of the disease (and in particular its absence) are now only considered for the patients suffering from variants B and C (or only for women).

Bayes' rule can be justified in various contexts and for various types of probabilities. In practice, if it is justified for an ontological probability, then it is also justified for an epistemic probability by virtue of the alignment principle (see §9). The reverse, however, is evidently not true. There are two types of justifications, examined in turn:

- cognitive justifications are only based on probabilities considered as the attributor's beliefs, apart from any use that may be made of those beliefs;
- pragmatic justifications are based on the attributor's decisions, themselves assumed to be based on her beliefs.

In the *focusing* context, Bayes' rule applies naturally to every form of probability. This is because the initial distribution of probabilities is homothetically transformed into a new distribution over a more limited class of worlds. Within each subclass, the ratios of probabilities express the proportions of possible worlds of each type and they are therefore conserved. For example, the new distribution of probabilities (assumed to be ontological) of the variants of the disease now only apply to women, when applying Bayes' rule.

In the *revising* context, applying Bayes' rule to ontological probabilities is more complicated. It is not the true probability that changes, but only its estimated value. Moreover, the message gives an indication about the distribution of probabilities, but it could have given other indications that would be equally relevant but lead to different final distributions. For example, if variant A is the only one

© Cournot Centre, September 2012

present in a population (and this is known to the modeller), the message may indicate equally well that variant C is absent or that variant B is absent. For epistemic probabilities, on the other hand, there is a direct cognitive justification for Bayes' rule in terms of general axioms for belief change, transposed to probabilities from AGM axioms (Walliser and Zwirn, 2002).

In the *updating* context, Bayes' rule can very well apply to ontological probabilities. Since the universe has changed, the true ontological probabilities Ch may also have changed. The estimated probabilities Ch^* can then be adapted accordingly. As a necessary condition, it must be assumed that "nothing else has occurred" in the system other than what is indicated by the message. In the example of the disease, Bayes' rule can be used to calculate the new proportion of each variant, provided that the relative occurrence of variants B and C has not changed. Moreover, for epistemic probabilities, it is the Lewis rule (imaging) that results from the general axioms of belief change, but it coincides with Bayes' rule for exhaustive messages.

As for the pragmatic justifications of Bayes' rule, they apply essentially to epistemic probabilities. In decision theory, the rule is justified by Savage's (1954/1972) axioms when these latter are interpreted in a dynamic setting. A similar justification exists in terms of the *Dutch Book,* again extended to dynamic choices (Lewis, 1999).

In what precedes, messages have always taken a set-theoretic form. But they can also take a probabilistic form. In other words, they can display a distribution of probabilities over a set of possible events. Thus, seen through an imperfect lens, a die lands on number one with some probability and on number two with the complementary probability. Bayes' rule has been generalized to this case (Jeffrey, 1992): the posterior probability of a world is simply a weighted sum (with the relative probabilities of the recorded events) of the conditional probabilities of the world obtained for each event.

Lastly, the conditional probabilities allow us to consider probability-based rules of inference. This probabilistic reasoning can be compared to classical deductive reasoning (for the modeller) or to the reasoning described by epistemic logic (for an

© Cournot Centre, September 2012

agent). The proposition E is replaced by the proposition $Pr(E) \geq \alpha$ or $Pr(E) = \alpha$, and the aim is no longer to preserve the *truth* of E, but its *probability* (Suppes, 1966; Adams, 1975). It is easy to see that several of the fundamental inference rules of deductive logic have no direct probabilistic counterpart. For example, with the conditional operator $\supset$, monotony allows us to infer $E \wedge G \supset F$[13] from $E \supset F$. With conditional probabilities, however, it is perfectly possible that $Pr(E/F) > Pr(E \wedge G/F)$. Similar remarks can be made for the *modus ponens* that allows us to infer F from E and $E \supset F$. If we conserve the conditional operator, we obtain the following result: if $Pr(E \supset F) = \alpha$ and $Pr(E) = \beta$, then $\alpha + \beta - 1 \leq Pr(F) \leq \alpha$.[14] If one prefers to use the conditional probability, then the result is: if $Pr(E/F) = \alpha$ and $Pr(E) = \beta$, then $\alpha\beta \leq Pr(F) \leq 1$[15] (Wagner, 2004).

Using this type of reasoning, we can build "Bayesian networks", which take the form of graphs linking variables according to conditional probabilities. The probabilities then propagate through the network from node to node. These networks can take the form of "Bayesian hierarchies" when the agent has to deal with nested random effects. In all these cases, conditional probabilities appear as the primary concept and traditional probabilities are simply particular cases. If $Pr(E/F)$ is a basic conditional probability, then $Pr(E) = Pr(E/W)$.

9. Ontological and methodological problems

In their ontological interpretation, probabilities are properties of events. As a large part of the philosophical literature on probability confirms, these properties are difficult to identify. Three discussions seek to analyse them in more detail.

The first discussion turns on whether non-extreme ontological probabilities exist in a deterministic context. It brings into opposition *compatibilists* (Levi, 2010; Maher, 2009; Hoefer, 2007) and *incompatibilists* (Lewis, 1980/1986; Loewer, 2001).

[13] $\wedge$ is the symbol of conjunction.

[14] In particular, if $\alpha = \beta = 1 - \varepsilon$ for $\varepsilon \leq ½$, then $Pr(F) \geq 1 - 2\varepsilon$.

[15] In particular, if $\alpha = \beta = 1 - \varepsilon$ for $0 \leq \varepsilon < 1$, then $Pr(F) \geq (1 - \varepsilon)^2$.

© Cournot Centre, September 2012

In the terms of §3, the former recognize irreducible and controlled probabilities as authentic ontological probabilities, while the latter only recognize irreducible probabilities. In fact, most philosophers accept the existence of non-extreme probabilities in the case of dice throwing; rejecting it would limit the use of probabilities to crude situations like radioactivity. The incompatibilists must also explain why we are so inclined to consider the probabilistic properties of dice as objective properties with which they are endowed.

The second discussion concerns the existence or not of non-extreme probabilities of specific events, which can be expressed on the occasion of a particular experiment. It brings into opposition *particularists* and *non-particularists*, with frequentists being the most overt representatives of the second group. Among other things, the non-particularists must explain why and by what principles do our attributions of probabilities continually shift from generic events to specific events.

The third discussion concerns the existence or not of non-extreme ontological probabilities for phenomena that have already occurred. It therefore concerns the evolution of probabilities over time – these probabilities being *temporal* for some authors (Lewis, 1980/1986) and *non-temporal* for others (Hoefer, 2007). For the former, contrary to *ex ante* probabilities, *ex post* probabilities are necessarily extreme (although possibly unknown). If the proposition *p* affirms that an event will occur at time t, then after t, the ontological probability of x is 1 if the event actually occurred and 0 otherwise. As Lewis says, "what's past is no longer chancy". As for epistemic probabilities, they vary naturally according to the information received.

The problem also arises of the transition between ontological probabilities and epistemic probabilities. This transition is only possible when both types of probability have been carefully defined on the same object. This requirement applies more to ontological probabilities, since epistemic probabilities are always defined if we accept that every uncertainty can be the subject of a bet. When it is possible, the transition from ontological probabilities to epistemic probabilities follows what we might call an *alignment principle*. This principle affirms that the epistemic probabilities Cr_i of an agent can be aligned purely and simply on the ontological

© Cournot Centre, September 2012

probabilities Ch if the agent should happen to know them. In its simplest form proposed by Miller (1966), it can be expressed as follows:

Miller's principle: $Cr_i(E / Ch(E) = \alpha) = \alpha$

A more sophisticated formulation is proposed by Lewis (1980/1986) in the following form (where the ontological probability is dated):

Principal principle: $Cr_i(E / [Ch_t(E) = \alpha] \wedge d) = \alpha$

In this expression, d is a proposition (or an event) which is both (i) compatible with the proposition that $Cr_t(E) = \alpha$ and (ii) admissible at time t. In other words, it only influences the beliefs of i about E by influencing her beliefs about the chances of E. Thus, a proposition stating that E has occurred is not admissible. On the other hand, a historical proposition, reporting on particular events that occurred before t, is admissible. The same holds true for a proposition concerning the dependency of ontological probabilities with respect to history, and more precisely a conditional whose antecedent involves events of this kind and whose consequent is a proposition on the ontological probability. For Lewis, the principal principle "seems to capture all we know about chance".

A variant of the alignment principle comes into play in the context of focusing. Let Eg be the result of a generic experiment and Es the result of a specific experiment. In the example of the disease, Eg concerns the illness of an individual of a certain type and Es the illness of a particular patient. According to the modified alignment principle, the ontological probability corresponding to the generic experiment is simply projected into an epistemic probability on the specific experiment:

Focusing principle: $Cr_i\ (Es / Ch(Eg) = \alpha) = \alpha$

© Cournot Centre, September 2012

The focusing principle performs two operations. First, like the alignment principle, it makes it possible to move from an ontological probability to an epistemic probability. Then it makes it possible to move from a generic event to a specific event. In this respect, it helps to account for the fact that we can assign a probability to a specific event, as the particularists maintain. This principle, however, raises the problem of the formal distinction between generic and specific events.

There is some disagreement over the status of the alignment principle. Lewis appears to stipulate that the principal principle is a principle of rationality, which applies to the degrees of belief of every rational agent, but does not justify them. Van Fraassen (1989) argues that it is impossible to find a justification for the alignment principle that does not suffer from circularity. By contrast, Howson and Urbach (1993) defend a version of the alignment principle adapted to the interpretation of probabilities that they support (von Mises' hypothetical frequentism), and they put forward a pragmatic justification of this version.

The opposite transition from epistemic probabilities to ontological probabilities is less conceptualized. It is generally considered in a dynamic setting, with an agent successively updating his epistemic probabilities. The question assumes a different meaning depending on whether one is a monist or pluralist subjectivist. For monists (de Finetti, 1937), ontological probabilities do not exist. It is then a matter of explaining why some probabilistic attributions appear to be objective, especially the vast inter-subjective agreement that exists between them: as a result of learning, it is a "psychological fact" (de Finetti) that inter-individual differences disappear. For pluralists, it is a matter of showing under what conditions epistemic probabilities converge towards ontological probabilities, which are assumed to exist independently of the former. These ideas are illustrated in different results of convergence,[16] typically obtained when information is sufficiently rich and numerous and when the initial epistemic probabilities do not diverge excessively from either the real world or each other.

[16] See Earman (1992), chap. 6 for a recent overview and an examination of the philosophical significance of these results, and Gillies (2000, pp. 69-83) for a presentation and critique of the founding result of de Finetti (1937).

© Cournot Centre, September 2012

10. Extensions of probabilities

The variant that comes closest to Kolmogorov's axioms concerns *ordinal probabilities*, which correspond to attributions of the kind: E is strictly more probable than E'. Ordinal probabilities (also called qualitative or comparative probabilities) are defined by a binary relation on an algebra $\mathcal{A}$ of a non-empty set W (Kreps, 1988; Fishburn, 1994; Fine, 1973). The binary relation denoted $>$, which is asymmetric and negatively transitive, is an ordinal probability if the three following conditions are satisfied for all $E, E', E'' \in \mathcal{A}$:

(PO1) not $(\varnothing > E)$
(PO2) $W > \varnothing$
(PO3) if $(E \cup E') \cap E'' = \varnothing$, then $E > E'$ if $(E \cup E'') > (E' \cup E'')$

A probability distribution P on $\mathcal{A}$ represents a relation $>$ on the same algebra if and only if, for any pair of events E and E',

$$[P(E) > P(E')] \text{ if } [E > E']$$

If a binary relation can be represented by a probability distribution, then it satisfies (PO1)–(PO3). On the other hand, a relation that satisfies these conditions is not necessarily representable by a probability distribution. A number of properties are sufficient, however, to guarantee the representability (sometimes unique) of an ordinal probability. A very particular case of ordinal (or cardinal) probabilities involves "Boolean probabilities", where the probabilities of events can only take the values 0 or 1.

A priori, ordinal probabilities are open to the same interpretations as the standard quantitative probabilities. In fact, they have above all attracted the attention of supporters of an epistemic interpretation of probabilities. For those attached to introspective methods of revealing degrees of belief, comparative

© Cournot Centre, September 2012

judgements appear to be more realistic. In certain behavioural methods of revelation, like that of Savage (1954/1972), ordinal probabilities are the starting point for the measurement of quantitative probabilities, since the judgements revealed by the choices of an agent are of the type: "agent i judges E to be more probable than E'".

A different type of extension of probabilities concerns *hierarchical probabilities,* which study probabilistic judgements of an attributor on pre-defined probabilities (Walliser and Zwirn, 2011). To this end, we define successive strata k formed respectively of worlds, meta-worlds and so on (k-worlds). Within each level between two layers, the agent makes a judgement in each $(\mathrm{k}+1)$-world on the lower k-worlds. This judgement can be of two kinds. A set-theoretical judgement associates a subset of k-worlds with each $(\mathrm{k}+1)$-world. A probabilistic judgement associates a distribution of probabilities over the k-worlds with each $(\mathrm{k}+1)$-world. On the basis of simple rules, it is possible to calculate an upper and a lower value for each event defined on the basic worlds.

If we simply take two levels (three strata), three interesting structures have been studied in the literature. Two-level probabilities (Skyrms, Baron, Kyburg) result from probabilistic judgements at each level (probabilities on probabilities). Belief functions (Dempster, Shafer) are the result of set-theoretical judgements at the lower level and probabilistic judgements at the higher level. The families of probabilities are the result of probabilistic judgements at the lower level and set-theoretical judgements at the higher level. The last two structures prove to be interchangeable under certain conditions. They are expressed as restrictions on the "Choquet capacities", that is to say measurements on the events that satisfy the axioms $\mathrm{P1}$ and $\mathrm{P2}$ and the monotonicity axiom. In particular, they make it possible to take into account situations of total ignorance, which the standard probabilities do not allow.

The worlds of each level can themselves be interpreted in two different ways. Physical worlds correspond to entities endowed with material existence. From a physical world, one defines a distribution of ontological probabilities on the worlds of lower levels. Mental worlds correspond to mental states of an agent. From a psychological world, one defines a distribution of epistemic probabilities on the worlds of lower levels. A "subordination principle" entails that the higher levels are

© Cournot Centre, September 2012

necessarily psychological and the lower levels physical. This means that an agent can only have epistemic probabilities on ontological probabilities. Moreover, a "non-schizophrenia principle" stipulates that an agent cannot form a (non-extreme) epistemic probability on his own epistemic beliefs of a lower level (he cannot hesitate between two beliefs).

If we limit ourselves to one or two levels, only two structures are possible (setting aside the set-theoretic or probabilistic nature of the evaluations). An elementary belief simply carries a judgement on a phenomenon. So, for example, an agent can make a judgement about the colour of a ball. A compound belief carries a judgement on a meta-phenomenon composed of a collection of phenomena. Thus, an agent can make a judgement about the composition of an urn containing balls. A set-theoretical judgement expresses the contents of the urn that the agent considers possible. A probabilistic judgement expresses the "ambiguity" (second-order uncertainty) of the agent about the contents, in other words the degree of confidence he has in each of the different possible contents. On the other hand, an agent cannot believe that he has different beliefs about the colour of a given ball.

General principles similar to the alignment principle make it possible to transform physical worlds into psychological worlds (projection principle) and *vice versa* (anti-projection principle). Likewise, ontological probabilities are transformed into epistemic probabilities and *vice versa*. The agent takes as his own the ontological probabilities that he is supplied with or thinks up a device that gives reality to his epistemic probabilities.

These transformations make it possible to define equivalences between the three standard contexts of belief change (revising, updating and focusing) through adaptation of the structures concerned. The change rules are easy to determine following simple principles for structures that are essentially ontological. They are then transposed to change structures of a more epistemic nature. We thus find the numerous change rules for hierarchical structures that had already been proposed on an intuitive basis (Walliser and Zwirn, 2011). Most of these rules make it possible to change probabilities even when the prior probability of the message is zero.

© Cournot Centre, September 2012

11. Application to statistical inference

Statistical methods consist in comparing a random model with the data from two perspectives. First, *statistical inference* can be used to determine unknown parameters of the relations of the model in the light of the observations. Second, *statistical tests* allow one to judge the validity of the relations of the model in the light of the observations. In both cases, the parameter plays the role of a hypothesis that the statistician seeks to probabilize.

In statistical inference, a random variable X is assumed to have a probability distribution $f(\Theta, X)$ that depends on a random parameter Θ. This distribution is given *a priori* as a constraint and cannot be modified by observations. We possess a sample of observations of the variable X, namely $x = (x_1, x_2, \ldots x_i, \ldots x_I)$. We seek to estimate the value of the parameter so as to minimize the "distance" between the theoretical distribution and the observations. To do so, we start with a prior distribution of the parameter Θ, namely $\pi(\theta)$. We can then calculate the posterior probability distribution of Θ, namely $\pi(\theta/x)$. Furthermore, we can calculate the likelihood of the hypothesis Θ, namely Π_i $(f(\Theta, x_i)$. In each case, the statistician looks for a confidence interval at α % of the parameter $[\underline{\theta}, \theta]$, the interpretation of which may vary.[17]

As an example, the statistician's basic operation is linear regression. He starts with a cloud of points corresponding to pairs of observations (x_i, y_i). He looks for the parameters a and b of the straight line of the equation $y = \mathrm{a}x + \mathrm{b}$ that comes closest to the points. The method of least squares measures this distance in terms of the sum of the squares of the deviations (measured vertically) between the straight line and the points. The straight line of the regression is supposed to represent the law governing the relation between the explanatory variable x and the explained variable y. The random errors with respect to the law, expressed by the deviations between the points and the straight line, have two interpretations. The ontological interpretation focuses on the deviations that arise with respect to the law. The

[17] For a precise analysis of the relations between different approaches to inference, see Courgeau (2012).

© Cournot Centre, September 2012

epistemic interpretation focuses on the modeller's incomplete understanding of the form of the law.

The *classical approach* consists in searching for the parameter by the method of maximum likelihood (or the method of moments). Under classical hypotheses, this method coincides with the method of least squares. In fact, the probability $f(\Theta, X)$ is *a priori* objective and the statistician's part of convention lies in choosing the maximum likelihood rule (or another rule) to calculate or situate the parameter. The confidence interval for the parameter means that there is a probability $1\text{-}\alpha$ of obtaining the sample x that has been drawn, if the hypothesis is true: $\Pr\ (x\ /\ \theta \in [\underline{\theta}, \theta]\ ) \geq 1\text{-}\ \alpha$. In other words, $(1\text{-}\alpha)$ % of the confidence intervals associated with the samples x contain the true value θ, which is hardly an intuitive interpretation.

The *Bayesian approach* involves the calculation of a posterior probability of the parameter on the basis of a prior probability. In fact, if the probability $f(\Theta, X)$ is *a priori* objective, the probability $\pi(\theta)$ is subjective (attached to the statistician), as is the posterior probability $\pi(\theta/x)$. The confidence interval means that there is a probability α that the value of the parameter is in the interval: $\Pr\ (\theta \in [\underline{\theta}, \theta]\ /\ x) \geq 1\ \text{-}\ \alpha$. This time, the parameter is indeed in an interval with probability α. The key problem is the origin of the prior probability, which we can at the very least subject to sensibility calculations.

The aim of statistical tests is to compare a reference hypothesis (called the null hypothesis) h_0 with the hypothesis tested (called the alternative hypothesis) h. For any hypothesis h, two types of error are considered. A first type error consists in rejecting the hypothesis even though it is actually true. A second type error consists in accepting the hypothesis although it is false. These errors can be probabilized as a function of the observations made. One possible test then consists in accepting the hypothesis if errors of both types are below conventional thresholds.

© Cournot Centre, September 2012

12. Application to empirical sciences

12.1 General considerations

Many models in the empirical sciences are stochastic in nature. Two interpretations are directly associated with the random errors introduced into the modelling. The ontological interpretation supposes that the phenomenon really is of an indeterminist nature. The epistemic interpretation supposes that the modeller does not know the real structure of the model (omitted explanatory variables, erroneous analytic form of a relation) and introduces random errors to make up for his ignorance. A third interpretation focuses on the comparison of the model with the data. More precisely, the methodological interpretation introduces probabilities to take into account errors in measuring the data.

So probabilities are objective in many circumstances, including the social sciences where the conditions are independent and difficult to repeat. In physics, quantum mechanics introduces (irreducible) probabilities, assumed to be intrinsic, concerning the instantaneous state of elementary particles. Likewise, statistical mechanics introduces probabilities linked to the (excessively complex) distribution of molecules in a gas. In biology, probability is involved in the random mutations that affect the genes of animal species or in gene transcription errors. In the social sciences, Durkheim used a populational approach to define the probability of individuals committing suicide as a function of their religion.

The use of subjective probability is less frequent, except in forecasting to appraise the uncertainty of the proposed scenarios. In physics, weather forecasts are assigned probabilities that reflect our incomplete knowledge of meteorological laws. In economics, macroeconomic forecasts are subject to uncertainties of multiple origins and are also probabilized. Subjective probabilities are used to build “cones of uncertainty” from the present, reflecting the fact that uncertainty increases with the horizon of the forecast.

© Cournot Centre, September 2012

12.2 Game theory

In the social sciences, one of the domains that makes the most intensive use of probability is game theory, the general framework for analysing strategic relations between agents. It identifies three main sources of uncertainty that affect not only the modeller, but also the agents represented in the model. These uncertainties have been systematically represented in probabilistic form, whatever their interpretation. It is only recently that non-probabilistic measurements have been introduced to express the third form of uncertainty (belief functions).

The objective uncertainty of the modeller reflects an intrinsic indeterminism of the system. It operates in the behaviour of agents, due to free will or a "trembling hand" (slight discrepancy between the intention and execution of an action). It intervenes in collective phenomena resulting from numerous causal chains, as in the case of innovations or external macroeconomic shocks.

The subjective uncertainty of the modeller derives from her profound ignorance of complex phenomena. It intervenes already at the level of the numerous factors introduced into production and cost functions, which combine technical and human factors. Subjective uncertainty also intervenes at the level of the functions of agents' behaviour, which only take into account a limited number of explanatory factors. In particular, it plays a role in the mental states that influence behaviour (beliefs, preferences).

The subjective uncertainty of agents concerns states of nature and the "types" of other agents (that is, the characteristics that determine their choices). It affects their factual information (past observations), their structural information (permanent structures) or their future information (predictions). It affects the passive context (states of nature), the characteristics of their opponents (actions, types) and even their own characteristics (actions, types). It is generally assumed that the agents adopt a prior probability of an objective nature (common to all agents) about the uncertain variables. This is known as a *common prior assumption*. They receive messages of a subjective nature about these variables; subjective in the sense that they differ from one individual to another (sets of information). Finally, they form a posterior probability of a subjective nature based on the previous elements.

© Cournot Centre, September 2012

As an example, we can examine the interpretation given to a mixed strategy implemented by an agent, namely a probability distribution over his pure strategies. The first interpretation considers this strategy as the expression of a deliberately random behaviour of the agent, which can be explained in several ways. The second interpretation relates to the average behaviour of a population whose members are pursuing different pure strategies. The third interpretation relates to the uncertain view that an agent has about the behaviour of his opponent.

A similar analysis can be made of models of preferences including a stochastic element, developed in the theory and psychology of decision making and sometimes applied in game theory. The first interpretation relates to the alternative "states of mind" that the agent can have (possibly according to external circumstances). The second interpretation relates to the dispersion of preferences through a population. The third interpretation expresses the uncertainty to which the agent is subject as regards the "type" of other agents (encapsulating the determinants of their choices).

13. Application to epistemology

In philosophy, probability is mainly but not exclusively[18] used in confirmation theory and inductive reasoning. Since the pioneering work of Hempel (1945), this theory has sought to clarify and codify the concepts of confirmation and disconfirmation as they are used in scientific and everyday reasoning. The question is usually formulated as follows: if H is a hypothesis and E is a proposition that summarizes empirical observations, under what conditions can we consider that E confirms (resp. disconfirms) H?[19]

Confirmation theories were first developed in a logical context, without involving probability. Today, the main works adopt a probabilist framework and are

[18] It also plays an important role in the theories of causality and knowledge.

[19] For a general presentation, see Zwirn and Zwirn (1996).

© Cournot Centre, September 2012

based on an epistemic interpretation of probability. This dominant trend is known as *Bayesian confirmation theory* (BCT).[20]

BCT uses an *incremental* (and not absolute) concept of confirmation: H is confirmed by E, not if E makes H very probable, but if E increases the probability of H:

$$E \text{ confirms } H \text{ if } \mathrm{P}(H / E) > \mathrm{P}(H)$$

This qualitative concept of confirmation is often associated with a corresponding quantitative concept that varies according to the measures chosen. The popularity of BCT stems from its capacity to account for a large number of intuitions about confirmation. It provides elements of response – more or less convincing – to classic problems of confirmation, such as the Raven paradox or the Duhem–Quine problem. It also leads to the following results:

(1) all else being equal, the more probable a proposition E is, given a hypothesis H, the more H will be confirmed by E;
(2) all else being equal, the *less* probable the proposition E is *a priori*, the more H will be confirmed by E (the "surprise principle");
(3) E confirms H if and only if $\mathrm{P}(E \mid H) > \mathrm{P}(E \mid \neg H)$.

One of the major sources of difficulty for BCT derives from its subjectivity. From the standpoint of this theory, there is nothing to prevent the proposition E from confirming H for one individual i but not for another individual j. According to some scholars, this excessive liberalism prevents BCT from accounting for the objectivity at work in scientific reasoning. One possible response consists in restricting the applications of BCT to certain favourable cases where individual probabilities are similar or even identical. This is non-trivial since it has to hold for the conditional probability $\mathrm{P}(E/H)$, but also for the plain probability $\mathrm{P}(\mathrm{E})$. It is the case if the subjective probabilities result from objective probabilities through the alignment

[20] For detailed discussions, see Earman (1992) and Howson and Urbach (1993). For a more succinct presentation, see Cozic (2011).

© Cournot Centre, September 2012

principle (or focusing principle) or if the subjective probabilities converge when the information is sufficiently rich (see §8). It is possible, however, to question BCT and characterize the concept of confirmation only according to likelihood ("likelihoodism").

14. Conclusion

Two polar-opposite ways of measuring probabilities are operationally distinguished in this essay. Frequential probabilities are inherent to phenomena and obtained simply by measuring the frequency of their occurrence in repeated experiments. Revealed probabilities are agents' degrees of belief about phenomena and are obtained by revelation from their actions. The other types of probabilities are situated between these two extremes and are based on less rigorous methodologies. The more objective ones no longer result from repeated experiments, but from simple sequential or populational data. As for the more subjective ones, they are no longer associated with well-identified agents, but are very free estimations of probabilities that enable calculations.

The epistemic and subjective approaches to probability are often described as "Bayesian". In fact, this term is open to several different meanings that can be distinguished by the use of probability in the behaviour of any agent. At the first level, an agent is Bayesian if he expresses the uncertainty about his environment in the form of epistemic probabilities. At the second level, an agent is Bayesian if he performs his reasoning in keeping with the calculation of probabilities, in particular by using Bayes' rule in a dynamic context. At the third level, an agent is Bayesian if he makes his decisions according to the criterion of expected utility maximization. At the three levels, Bayesian approaches apply to any decision maker who perceives and acts on his environment. They also apply to the modeller, especially statisticians and epistemologists who are Bayesians at the second and third levels.

Numerous bridges have been laid down between the ontological/objective and epistemic/subjective approaches. Above all, these bridges have led to the objective approach being treated as a particular or asymptotic case of the subjective approach. This is true in the calculation of probabilities, where the objective approach

© Cournot Centre, September 2012

appears as a specific form of the subjective approach with complete information. It is also true in statistics, where classical statistics appears as a particular case of Bayesian statistics.

© Cournot Centre, September 2012

References

Adams, E. (1975), *The Logic of Conditionals*, Dordrecht, Hollande: Reidel.

Carnap, R. (1950/1962), *Logical Foundations of Probability*, Chicago: University of Chicago Press.

Courgeau, D. (2012), *Probability and Social Science: Methodological Relationships between the two Approaches*, Dordrecht, Holland, and New York: Springer.

Cozic, M. (2011), "Confirmation et induction", in A. Barberousse, D. Bonnay and M. Cozic (eds), *Précis de philosophie des sciences*, Paris: Vuibert.

Earman, J. (1992), *Bayes or Bust? A Critical Examination of Bayesian Confirmation Theory*, Cambridge, MA: MIT Press.

Fine, T. (1973),*Theories of Probability: An Examination of Foundations*, New York: Academic Press.

de Finetti, B. (1937), "La prévision: ses lois logiques, ses sources subjectives", *Annales de l'I.H.P.*, 7(1), pp. 1–68.

de Finetti, B. (1974), *Theory of Probability*, 1, New York: Wiley, 1990.

Fisher, R.A. (1925), *Statistical Methods for Research Workers*, Biological Monographs and Manuals Series, Edinburgh: Oliver and Boyd.

Gillies, D. (2000), *Philosophical Theories of Probability,* London: Routledge.

Hajek, A. (2010), "Interpretations of Probability", *The Stanford Encyclopedia of Philosophy (Spring 2010 Edition)*, in Edward N. Zalta (ed.),

URL = http://plato.stanford.edu/archives/spr2010/entries/probability-interpret/

Howson, C. and P. Urbach, (1993), *Scientific Reasoning: The Bayesian Approach*, La Salle, IL: Open Court.

Hoefer, C. (2007), "The Third Way on Objective Probability: A Sceptic's Guide to Objective Chance", *Mind*, 116(463), pp. 549–596.

Jaynes, E. (2003), *Probability Theory: The Logic of Science*, Cambridge: Cambridge University Press.

Jeffrey, R. (1992), *Probability and the Art of Judgment*, Cambridge: Cambridge University Press.

Keynes, J.M. (1921), *A Treatise on Probability*, London: Macmillan.

Kolmogorov, A.N. (1933/1950), *Foundations of Probability*, New York: Chelsea Publishing Company.

Laplace, P.S. (1812), *Théorie analytique des probabilités*, Paris: Coursier.

Levi, I. (2010), "Probability Logic, Logical Probability and Inductive Support", *Synthese*, 172, pp. 97–118.

Lewis, D. (1976), "Probabilities of Conditionals and Conditional Probabilities", *The Philosophical Review*, LXXXV(3), pp. 297–315.

Lewis, D. (1980/1986), "A Subjectivist's Guide to Objective Science", *Philosophical Papers*, 2(19), pp. 83–132.

© Cournot Centre, September 2012

Lewis, D. (1999), "Why Conditionalize?", *Papers in Metaphysics and Epistemology*, pp. 403–407.

Loewer, B. (2001), "Determinism and Chance", *Studies in the History and Philosophy of Modern Physics*, 32, pp. 309–320.

Maher, P. (2006), "The Concept of Inductive Probability", *Erkenntnis*, 65, pp. 185–206.

Maher, P. (2009), "Physical Probability", in C. Glymour, W. Wei and D. Westerstahl (eds), *Logic, Methodology and Philosophy of Science*, London: College Publications.

Miller, D. (1966), *Critical Rationalism. A Restatement and Defence*, Chicago and Lasalle, IL: Open Court.

von Mises, (1928/1957), *Probability, Statistics and Truth*, New York: Macmillan.

von Plato, J. (1983), "The Method of Arbitrary Functions", *British Journal for the Philosophy of Science*, 34, pp. 37–47.

Poincaré, R. (1896/1912), *Calcul des probabilités*, Paris: Gauthier-Villars.

Popper, K. (1959), "The Propensity Interpretation of Probability", *British Journal for the Philosophy of Science*, 10, pp. 25–42.

Ramsey, F.P. (1931), "Truth and Probability", in R. Braithwaite (ed.), *The Foundations of Mathematics and Other Logical Essays*, London: Routledge and Keegan Paul.

Reichenbach, H. (1949/1971), *The Theory of Probability*, Berkeley: University of California Press.

Savage, L. (1954/1972), *The Foundations of Statistics,* New York: John Wiley and Sons.

Strevens, M. (2011), "Probability out of Determinism", in C. Beisbart and S. Hartmann (eds), *Probabilities in Physics*, Oxford: Oxford University Press.

Suppes, P. (1966), "Probabilistic Inference and the Concept of Total Evidence", in J. Hintikka and P. Suppes (eds), *Aspects of Inductive Inference*, Amsterdam: North Holland, pp. 49–55.

Wagner, C. (2004), "*Modus Tollens* Probabilized", *British Journal of Philosophy of Science*, 50(4), pp. 747–753.

Walliser, B. and D. Zwirn (2002), "Can Bayes' Rule Be Justified by Cognitive Rationality?", *Theory and Decision*, 53, pp. 95–135.

Walliser, B. and D. Zwirn (2011), "Change Rules for Hierarchical Beliefs", *International Journal of Approximate Reasoning*, 52(2), pp. 166–183.

Williamson, J. (2009), "Philosophies of Probability", in A.Irvine, (ed.), *Handbook of the Philosophy of Science*, 4, pp. 493–533.

Zwirn, D. and H. Zwirn (1996), "Metaconfirmation", *Theory and Decision*, 41, pp. 195–228.

Zwirn, H. (2002), "Principe de Miller et révision des croyances probabilistes", in P. Livet (ed.), *Révision des croyances*, Paris: Lavoisier.

© Cournot Centre, September 2012