Four concepts of probability

Mario Bunge
Foundations and Philosophy of Science Unit, McGill University, Montreal, Canada
(Received May 1981)

Four concepts of probability are examined: the mathematical concept and its personalist, frequentist, and propensity interpretations. The first two interpretations are shown to be at variance with the standard calculus of probability. The personalist concept is invalid because the probability function makes no room for any persons; and the frequency interpretation in mathematically incorrect because the axioms that define the probability measure do not contain the (semiempirical) notion of frequency. On the other hand the propensity interpretation of probability is found to be mathematically unobjectionable and the one actually employed in science and technology and compatible with both a possibilist ontology and a realist epistemology.

Introduction

The concept of probability has fascinated numerous philosophers since its inception three and a half centuries ago. Scientists too, whether basic or applied, have often taken part in philosophical discussions on probability. Notwithstanding such discussions, which have been numerous and often spirited, there is still considerable divergence of opinion concerning the interpretation of probability. This is probably due to the fact that the choice of interpretation is largely a matter of philosophy. Not that philosophy is necessarily inconclusive, but it does colour all thinking on fundamental questions.

Up until one century ago the philosophy of probability was dominated by subjectivism: probability was regarded as a measure of the credibility (or uncertainty, or weakness) of our beliefs. This interpretation had an ontological basis: since the universe was deemed to be strictly deterministic, probability had to be resorted to because of our ignorance of details. (God had no use for probability.) The paradigm case was the kinetic theory of gases: here the basic laws were deterministic but probability was called for because of our ignorance of the initial positions and velocities of the individual molecules.

About a century ago an alternative view emerged, namely the frequency interpretation.[1] According to this view probabilities are long run values of relative frequencies of observed events. While this was a step in the direction of objectivity, it remained half way, because it was concerned with observations rather than with objective facts. Probability was regarded as a feature of human experience rather than as a measure of something objective. Like the subjectivistic interpretation, the frequency interpretation is still very much alive – if not *de jure* at least *de facto*.

A third interpretation of probability began to emerge at the time of World War I with reference to statistical mechanics and other stochastic theories, namely the so-called propensity interpretation. According to this view probability values measure the strength of a tendency or disposition of some event to happen. This objectivist interpretation, which can be found in Smoluchowski,[2] Fréchet[3] and a few others, has been gaining ground among philosophers, particularly since it was adopted by Popper.[4]

There are then three main views on the nature of (applied) probability: the subjectivist, the frequency, and the propensity interpretations. Until the birth of quantum theory in 1926 the first interpretation was just as compatible with objectivism as with subjectivism, for one could argue that the basic laws are deterministic, probability being required only because of our empirical ignorance of details. But quantum mechanics and quantum electrodynamics, with their basic stochastic laws, changed the relation of probability to philosophy: from then on the subjectivistic philosophy of probability is compatible only with a subjectivistic philosophy willing to hold that the stochastic laws of quantum mechanics and other scientific theories would cease to hold the day people stopped thinking about atoms, molecules, photons, and other objects with stochastic behaviour.

The frequency interpretation has had a similar fate. While originally it could be espoused by realists as well as by empiricists, ever since the quantum theory was born realists cannot accept it because to them the laws of atoms and the like are not supposed to depend upon our observation acts. Thus an atom in an excited state has a definite objective probability of decaying to a lower energy state within the next second, whether or not somebody is counting the actual events of this kind in a large assembly of atoms of the same kind. In other words, the propensity interpretation of probability accords well with a realistic interpretation of the quantum theory. But this argument will not persuade someone who is not a realist or who, being a realist, doubts that the quantum theory is here to stay. He will demand more general reasons, i.e. reasons that can be used with reference to all scientific theories.

0307-904X/81/050306-07/$02.00

The purpose of the present paper is to supply such reasons: to show that the subjectivistic and the frequency interpretations are untenable, whereas the propensity interpretation accords well with both the mathematical theory of probability and the stochastic theories of contemporary science. To this end it will prove convenient to start by giving a brief characterization of the theory whose interpretations are at stake, namely the probability calculus.

The abstract concept

Up until four decades ago there was some confusion in the foundations of probability. The confusion consisted in a lack of distinction between the mathematical theory of probability and its various interpretations and applications. So much so that the theory was often presented as if it dealt with physical events. That stage was overcome by Kolmogoroff's work.[5] This work made it clear that the probability calculus is a branch of pure mathematics – this being why it can be applied in so many different fields of research. Let us give a quick review of the gist of Kolmogoroff's axiom system in its elementary version. (There are of course alternative formulations, in particular Renyi's, but one is enough for our purposes.)

The calculus of probability presupposes ordinary logic (the predicate calculus with identity), elementary set theory, ring theory (a branch of abstract algebra), real analysis, and measure theory. But the foundations of probability theory can be understood without the help of any sophisticated mathematics. Indeed the theory has just two basic (or primitive or defining) concepts with a simple mathematical structure. These are the notions of an event (understood in a technical sense) and of probability measure, which occur in statements of the form 'The probability of event x equals y'. In principle any set qualifies as an 'event', and the probability of such an 'event' is a real number assigned to it by the probability function.

More precisely, the probability function Pr is defined on a family F of sets such that the union and the intersection of any two members of F be in F, and also that F be closed under complementation. In sum, F must be a σ algebra, in the sense that its members obey the laws of the algebra of sets extended to countably infinite unions. This algebraic structure is not arbitrary but is demanded by the applications of the calculus. Thus given the probabilities of the events x and y, we must be able to compute the probabilities of the complex events 'x and y', 'x or y', and 'not x', and even the probability of an infinite disjunction of events. (Note that in the applications we have to do with events proper, not just with abstract sets. But note also that, since real events cannot be negative or disjunctive, the calculus of probability applies to possibilities not actualities.) As soon as any of the events referred to by the expression 'x or y' is actualized, the expression 'the probability of x or y becomes pointless. See Bunge.[6]

We are now ready for a formal definition of the probability concept, namely thus. Let F be a σ algebra on a nonempty set S, and $Pr: F \to [0, 1]$ a real-valued bounded function on F. Then Pr is a probability measure on F if and only if it satisfies the following conditions:

(i) for any countably infinite collection of pairwise disjoint sets in F, the probability of their union equals the sum of their individual probabilities. (In particular, if x and y are in F, and $x \cap y = \emptyset$, then $Pr(x \cup y) = Pr(x) + Pr(y)$.);
(ii) $Pr(S) = 1$.

Note that the theory based on these sole assumptions is semiabstract insofar as it does not specify the nature of the elements of the basic set S nor, *a fortiori*, those of the probability space F. On the other hand the range of Pr is fully interpreted: it is not an abstract set but the unit interval of the real line. Hence the semi. Were it not for the semantic indeterminacy of the domain F of the probability function, the calculus could not be applied everywhere, from physics and chemistry to biology and sociology. As long as the probability space F is not specified, i.e. as long as no model is constructed, probability has nothing to do with possibility, propensity, randomness, or uncertainty.

An application of any abstract or semiabstract theory to some domain of reality consists in enriching the theory with two different items: (*a*) a model or sketch of the object or domain of facts to which the theory is to be applied, and (*b*) an interpretation of the basic concepts of the theory in terms of the objects to which it is to be applied. Shorter: a factual scientific concept f is a mathematical concept m together with an interpretation I that assigns m a set of facts; i.e. $f = \langle m, I \rangle$. (For details see Bunge.[7])

In particular, an application of probability theory consists in joining the above definition of probability measure (or some of its consequences) with (*a*) a stochastic model – e.g. a coin flipping model or an urn model or what have you, and (*b*) a set of interpretation (or correspondence or semantic) assumptions sketching the specific meanings to be attached to a point x in the probability space F, as well as to its measure $Pr(x)$. As long as these additional assumptions are not introduced, the probability theory is indistinguishable from measure theory, which is a chapter of pure mathematics: only those specifics turn the semiabstract theory into an application of probability theory or part of it.

In other words, the general and semiabstract concept Pr of probability measure is defined (via a set of axions) in pure mathematics. Each factual interpretation I_i of the domain F of Pr, as well as of the values $Pr(x)$ of the probability measure (for x in F), yields a factual probability concept $f_i = \langle F, Pr, I_i \rangle$, where i is a natural number. These various factual probability concepts belong to factual science, not to pure mathematics: they are the probabilities of atomic collisions, of nuclear fissions, of genic mutations, of survival up to a certain age, of learning a certain item on first presentation, of moving from one social group to another, and so on and so forth.

What the various specific (or interpreted) probability concepts have in common is clear, namely the mathematical concept of probability Pr. This shows that the attempts of the subjectivists and of the empiricists to define the general concept of probability either in psychological terms (degrees of belief) or in empirical terms (frequencies of observations) were bound to fail: maximal generality requires deinterpretation, i.e. abstraction or semiabstraction. (For the notions of interpretation and of numerical degree of abstraction see Bunge.[7])

We can now approach the problem of weighing the claims of the three main doctrines on the nature of (applied) probability.

Probability as credibility

The subjectivistic (personalist, Bayesian) interpretation of probability construes every probability value $Pr(x)$ as a measure of the strength of someone's belief in x, or as the accuracy of his information about x (de Finetti,[8] Jeffreys,[9] Savage[10]). There are a number of objections to this view.

The first objection, of a logical nature, was raised towards the end of the last section, namely that one does not succeed in constructing a general concept by restricting oneself to a specific interpretation. However, a personalist might concede this point, grant that the general concept of probability belongs in pure mathematics, and claim just that the subjectivist interpretation is the only applicable, or useful, or clear one. However, this strategy will not save him, for he still has to face the following objections.

The second objection, of a mathematical nature, is that the expression '$Pr(x) = y$' makes no room for a subject u and the circumstances v under which u estimates his degree of belief in x, under v, as y. In other words, the elementary statements of probability theory are of the form '$Pr(x) = y$', not '$Pr(x, u, v) = y$'. And such additional variables are of course necessary to account for the fact that different subjects assign different credibilities to one and the same item, as well as for the fact that one and the same subject changes his beliefs not just in the light of fresh information but also as a result of sheer changes of mood. In sum, the subjectivist or personalist interpretation of probability is adventitious, i.e. compatible with the mathematical structure of the probability concept.

Even if the former objection is waived aside as a mere technicality – which it is not – a third objection is in order, namely this. It has never been proved in the psychological laboratory that our beliefs are so rational that in fact they satisfy all of the axioms and theorems of probability theory. On the contrary, there is experimental evidence pointing against this thesis. For example, most of us experience no difficulty in holding pairs of beliefs that, on closer inspection, turn out to be mutually incompatible. Of course the subjectivist could circumvent this objection by claiming that the 'calculus of beliefs' is a normative theory not a descriptive one. He may indeed hold that the theory defines 'rational belief', so that anyone whose behaviour does not conform to the theory departs from rationality instead of refuting the theory. In short he may wish to claim that the theory of probability is the theory of rationality – a philosophical theory rather than a psychological one. This move will save the theory from refutation but it will also deprive it of confirmation.

A fourth objection is as follows. A belief may be construed either as a state of mind (or a brain state) or as a proposition (or statement). If the former then the probability $Pr(x)$ of belief x can be interpreted as a measure of the objective strength of the propensity or tendency for x to occur in the given person's mind (or brain). But this would of course be just an instance of the objectivist interpretation and would be totally alien to the problem of the likelihood of x or even the strength of a subject's belief in the truth of x. On the alternative construal of beliefs as statements – which is the usual strategy of the Bayesians – we are faced with the problem of formulating rules for assigning them probabilities. So far as I know there are no such (nonconventional) rules for allotting probabilities to propositions. In particular, nobody seems to have been able to assign probabilities to scientific hypotheses – except of course arbitrarily. Surely the subjectivist is not worried by this objection: his whole point is that prior probabilities must be guesstimated by the subject, there being no objective tests, whether conceptual or empirical, to estimate the accuracy of his estimates. But this is just a roundabout way of saying that personalist probability is just a flight of fancy that must not be judged by the objective standards of science.

Our fifth objection is but an answer to the claim that probability values must always be assigned on purely subjective 'grounds', i.e. on no grounds whatever. If probability assignments were necessarily arbitrary then it would be impossible to account for the scientific practices of (a), setting up stochastic models of systems and processes and (b), checking the corresponding probability assignments with the help of observation, measurement, or theory. For example, genetic theory assigns definite objective probabilities to certain genic mutations and recombinations, and experimental biology is in a position to test those theoretical values by contrasting them with observed frequencies. (On the other hand nobody knows how to estimate the probability of either data or hypotheses. We do not even know what it means to say that such and such a statement has been assigned this or that probability.) In sum, the subjectivist interpretation of probability is at odds with the method of science: in science (a), states of things and changes of state, not propositions, are assigned probabilities, and (b), these assignments, far from being subjective, are controlled by observation, measurement or experiment, rather than being arbitrary.

Our sixth and last objection is also perhaps the most obvious of all: if probabilities are credibilities, how come that all the probabilities we meet in science, whether pure or applied, are probabilities of states of concrete things – atoms, molecules, fields, organisms, populations, societies, or what not – or probabilities of events occurring in things of that kind, no matter what credence the personalist probabilist may assign either the facts or the theories about such facts? Moreover, many of the events in question, such as atomic collisions and radiative transitions, are improbable or rare, yet we cannot afford to dismiss them as being hardly credible.

The personalist might wish to rejoin that, as a matter of fact, we often do use probability as a measure of certainty or credibility, for example when we have precious little information and when we apply the Bayes–Laplace theorem to the hypothesis/data relation. However, both cases are easily accounted for within the objectivist interpretation, as will be shown presently.

Case 1: incomplete information concerning equiprobable events

Suppose you have two keys, A and B, the first for your house and the second for your office. The probability that A will open the house door is 1, and the probability that it will open the office door is 0; similarly for key B. These are objective probabilities: they are physical properties of the four keylock couples in question. Suppose now that you are fumbling in the dark with the keys and that you have no tactual cues as to which is which. In this case the two keys are (empirically) equivalent before trying them. Whichever key you try, the probability of your choosing the right key for opening either door is 1/2. This is again an objective property, but not one of the four key-lock pairs: it is an objective property of the four you-key-lock triples. Of course these probabilities are not the same as the previous ones: we have now taken a new domain of definition of the probability function. And surely the new probability values might be different for a different person, e.g. one capable of distinguishing the keys (always or with some probability) by some tactual cues. This relativity to the key user does not render probability subjective, any more than the relativity of motion to a reference frame renders motion subjective. Moreover, even when we assign equal probabilities to all the events of a class, for want of precise information about them, we are supposed to check this hypothesis and change it if it proves empirically false. In short, incomplete information is no excuse for subjectivism.

Case 2: inference with the help of the Bayes–Laplace theorem

This is of course the stronghold of the personalist school. However, it is easily stormed. Firstly, recall that the Bayes–Laplace theorem is derivable from the mere definition of conditional probability without assuming any interpretation, whether personalist or objectivist. (Indeed, the definition is: $Pr(x|y) = Pr(x \cap y)/Pr(y)$. Exchanging x and y, dividing the two formulae, and rearranging, we obtain the theorem: $Pr(y|x) = Pr(x \cap y) \cdot Pr(y)/Pr(x)$.) Secondly, since there are no rules for assigning probabilities to propositions (recall our fourth objection), it is wrong to set x = evidence statement (e), and y = hypothesis (h) in the above formula, and consequently to use it as a principle of (probabilistic or statistical) inference. However, if we insist on setting $x = e$ (evidence) and $y = h$ (hypothesis), when we must adopt an indirect not a literal interpretation: $Pr(h)$ is not the credibility of hypothesis h but the probability that the facts referred to by h occur just as predicted by h. $Pr(e)$ is the probability of the observable events described by e; $Pr(h|e)$ is the probability of the facts described by h, given – i.e. it being actually the case – that the events referred to by e occur; and $Pr(e|h)$ is the probability of the event described by e, given that the facts referred to by h happen. This is the only legitimate interpretation of the Bayes–Laplace theorem because, as emphasized before, scientific theory and scientific experiment allow us to determine only the probabilities of (certain) facts, never the probabilities of propositions concerning facts. A byproduct of this analysis is that all the systems of inductive logic that use the Bayes–Laplace theorem interpreted in terms of hypotheses and data are wrong-headed.

In view of the objections raised against subjectivism, some of its proponents say now that it should not be construed as a theory of probability but as a normative theory of rational behaviour under risk, i.e. as decision theory (cf. Wald[11]). (The assumption is that a rational agent chooses the course of action that is likely to maximize the product of the utility of an outcome by its subjective probability.) But this position is untenable. First decision theory uses probabilities and thus presupposes an independent theory of probability but is not such a theory. Secondly, the probabilities occurring in decision theory are subjective, whereas a rational person is supposed to attempt to act always on objective probabilities – i.e. on probabilistic laws not on sheer guesses. (In real life, people who maximize their expected utilities using subjective probabilities are said to indulge in wishful thinking, not to engage in rational behaviour.)

The upshot of our analysis is that the personalist interpretation of probabilities is mistaken and irrelevant to science and technology.

Probability as frequency

If we cannot use the subjectivist interpretation then we must adopt an objectivist one. Now, many objectivists believe that the only viable alternative to the personalist interpretation is the frequency interpretation. The latter reduces to asserting that '$Pr(x) = y$' means that the relative long run frequency of event x equals number y or, rather, some rational number close to y.[1,11,12]

A first objection that can be raised against the frequency interpretation of probability – and *a fortiori* against the identification of the two – is that they are different functions altogether. Indeed, whereas Pr is defined on a probability space F (as we saw in the section on the abstract concept), a frequency function f is defined, for every sampling procedure π, on the power set $\mathscr{P}(F^*)$ of a finite subset F^* of F, namely the set of actually observed events. i.e.:

$$Pr: F \to [0, 1] \quad \text{but} \quad f: \mathscr{P}(F^*) \times \Pi \to Q$$

where Π is the set of sampling procedure (each characterized by a sample size and other statistical parameters) and Q is the set of proper fractions in $[0, 1]$.

Our second objection follows from the former: a probability statement does not refer to the same things as the corresponding frequency statement. Indeed, whereas a probability statement concerns usually a single (though possible complex) fact, the corresponding frequency statement is about a set of facts and moreover as chosen in agreement with certain sampling procedures. (Indeed, it follows from our previous analysis of the frequency function that its values are $f(x, \pi)$, where x is a member of the family of sets $\mathscr{P}(F^*)$ and π a member of Π.) For example, one speaks of the frequency with which one's telephone is observed (e.g. heard) to ring per unit interval, thus referring to an entire set of events rather than to a single event, which is on the other hand the typical case of probability statements. Of course probabilities can only be computed or measured for event types (or categories of events), never for unique events such as my writing this article. But this does not prove that, when writing '$Pr(x) = y$', we are actually referring to a set x of events: though not unique, x is supposed to be a single event. In other words, where probability statements speak about single events, frequency statements speak about sets of observed (or at least observable) events. And, since they do not say the same, they cannot be regarded as identical.

To put the same objection in a slightly different way: The frequency interpretation of probability consists in mistaking percentages for probabilities. Indeed, from the fact that probabilities can sometimes be estimated by observing relative frequencies, the empiricist probabilist concludes that probabilities are identical with relative frequencies, which is like mistaking sneezes for colds. Worse: frequencies alone do not warrant inferences to probabilities: by itself a percentage is not an unambiguous indicator of randomness. A selection mechanism, whether natural or artificial, if random, authorizes the interpretation of a frequency as a measure of a probability. For example, if you are given the percentage of events of a kind, and are asked to choose blindfolded any of them, then you can assign a probability to your correctly choosing the item of interest out of a certain reference class. In short, the inference goes like this: percentage and random choice – – → probability. (The line is broken to suggest that this is not a rigorous, i.e. deductive inference, but just a plausible one.)

Surely not all frequencies are observed: sometimes they can be calculated, namely on the basis of definite stochastic models, such as the coin flipping model (or Bernoulli sequence). But in this case too the expected frequency differs from the corresponding probability. So much so that the difference is precisely the concern of the laws of large numbers of the probability theory. One such theorem states that, in a sequence of Bernoulli trials, such as coin flippings, the frequency f_n of successes or hits in the first n trials approaches the corresponding probability (which is constant, i.e. independent of the size n of the sample). Another theorem states that the probability that f_n deviates from the corresponding probability p by more than a preassigned number tends to zero as n goes to infinity. (Note that there are two probabilities and one frequency at stake in this theorem.) Obliterate the difference between probability and frequency, and the heart of the probability calculus vanishes. This then is our third objection to the frequency interpretation of probability, namely that it cannot cope with the laws of large numbers. For further technical objections see Ville.[12]

Our fourth argument is of an ontological nature, namely this. While a frequency is the frequency of the actual occurrence of facts of a certain kind, a probability may (though it need not) measure the possibility of a fact, or rather the strength of such a possibility. Consequently identifying probabilities with frequencies (either by definition or by interpretation) implies (a), rejecting a real or physical possibility, thus forsaking an understanding of all the scientific theories which, like quantum mechanics and population genetics, take real possibility seriously, and (b), confusing a theoretical (mathematical) concept with an empirical one.

The correct procedure with regard to the probability-frequency pair is not to identify them either by way of definition or by way of interpretation, but to clarify their mutual relation as well as their relations to the categories of possibility and actuality. We submit that frequency estimates probability, which in turn measures or quantitates possibility of a kind, namely chance propensity (Bunge[6]). And, while probability concerns possibles, frequency concerns actuals and moreover, in the applications, it always concerns observed actuals.

In other words: there is no valid frequency interpretation of probability; what we do have are statistical estimates of theoretical probability values. Moreover frequencies are not the sole estimators or indicators of probability. For instance, in atomic and molecular physics transition probabilities are often checked by measuring spectral line intensities or else scattering cross sections. And in statistical mechanics probabilities are estimated by calculating entropy values on the basis of either theoretical considerations (with the help of formulae such as Boltzmann's) or measurements of temperature and other thermodynamic properties. In short, probabilities are not frequencies and they are not interpretable as frequencies although they can often (by no means always) be estimated with the help of frequencies.

To be sure frequencies, when joined to plausible random mechanisms, supply a rough indication of probability values and serve to check probability calculations. Hence probabilities and frequencies, far from being unrelated, are in some sort of correspondence. Yet this correspondence is complex and is far from complete. In fact (a), an event may be possible and may even have been assigned a nonvanishing probability without ever having been observed to happen, hence without being assigned a frequency; (b), conversely, certain events can be observed to occur with a certain frequency without, however, being assigned a nonvanishing probability.

In sum, the frequency interpretation of probability is inadmissible for a number of technical and philosphical reasons. Let us therefore look for an interpretation of probability free from the fatal flaws of the frequency interpretation.

Probability as propensity

Recall from the first section of the paper that the probability calculus has two specific undefined notions: those of probability space F and probability measure Pr. And remember that a full interpretation of a mathematical formalism involves interpreting all of its primitives. Since F is not interpreted in the pure calculus of probability, and only the range of the probability measure Pr is interpreted (in mathematical terms, namely as the unit interval of the real line), that calculus is semi-interpreted or, equivalently, it is semiabstract.

A mathematical interpretation of the probability calculus, i.e. one remaining within the context of mathematics, consists in specifying the mathematical nature of the members of the domain F or Pr, e.g. as sets of points on a plane, or as sets of integers, or in any other way compatible with the algebraic structure of F. Such an interpretation of the probability space F would yield a full mathematical interpretation of the probability theory. (Likewise, interpreting the elements of a group as translations, or as rotations, yields a full mathematical interpretation of the abstract theory of groups.) Obviously, such a mathematical interpretation is insufficient for the applications of probability theory to science or technology. Here we need a factual interpretation of the calculus.

A factual interpretation of probability theory is obtained by assigning both F and every value $Pr(x)$ of Pr, for x in F, factual meanings. One such possible interpretation consists in taking the basic set S, out of which F is manufactured, to be the state (or phase) space of a thing. In this way every element of the probability space F is a bunch of states, and $Pr(x)$ becomes the strength of the propensity or tendency the thing has to dwell in the state or states x. Similarly, if x and y are states (or sets of states) of a thing, the conditional probability of y given x, i.e. $Pr(y|x)$, is interpreted as the strength of the propensity or tendency for the thing to go from state(s) x to state(s) y. This then is the propensity interpretation of probability.

This is not an arbitrary interpretation of the calculus of probability. Given the structure of the probability function and the interpretation of its domain F as a set of facts (or events or states of affairs), the propensity interpretation is the only possible interpretation in factual terms. Indeed, if F is a set of facts, then $Pr(x)$, where x is in F, cannot but be a property of the individual fact x. This is, contrary to the frequency view (see previous sections), probability is not a collective or ensemble property, i.e. a property of the entire set F, but a property of every individual fact, namely its propensity to happen. What are ensemble properties are, of course, the normalization condition $Pr(S) = 1$ (recall the first section) and derived functions such as the moments of a probability distribution, its standard deviation if it has one, and so on. (This consideration suffices to ruin the frequency school, according to which probability is a collective or ensemble property.)

This point is of both philosophical and scientific interest. Thus some biologists hold that, because the probability of survival can be measured only on entire populations, it must be a global property of a population rather than a property of each and every member of the population. (Curiously enough they do not extend this interpretation to the mutation probability.) The truth is of course that, while each probability function Pr is a property of the ensemble F, its values $Pr(x)$ are properties of the members of F.

It is instructive to contrast the propensity to the frequency interpretations of probability values, assuming that the two agree on the nature of the probability space F. (This assumption is a pretence: not only frequentists like von Mises but also Popper, the philosophical champion of the propensity interpretation, have stated that facts have no probabilities unless they occur in experimentally controlled situations. In fact they emphasize that probabilities are mutual properties of a thing and a measurement set-up, which of course makes it impossible to apply stochastic theories to astrophysics. For a detailed examination of various versions of the propensity interpretation, in particular Popper's and my own, see Settle[13].) The contrast between the propensity and the frequency interpretations is displayed in the following *Table*.

Table Potentialist versus actualist interpretation of probability

$p = Pr(x)$	Propensity	Frequency
0	x has (almost) nil propensity	x is (almost) never the case
$0 < p \ll 1$	x has a weak propensity	x is rare
$0 \ll p < 1$	x has a fair propensity	x is fairly common
$p \cong 1$	x has a strong propensity	x is very common
$p = 1$	x has an overpowering propensity	x is (almost) always the case

Note the following points. Firstly, although a probability value is meaningful, i.e. it makes sense to speak of the single fact propensity, it is so only in relation to a definite probability space (e.g. with reference to a precise category of trials). Likewise a frequency value makes sense only in relation to a definite sample-population-sampling method triple. For example, the formula 'x is rare' presupposes a certain set of occurrences, to which x belongs, among which x happens to be infrequent.

Second, in the case of continuous distributions, zero probability is consistent with very rare (isolated) happenings. That is, even if $Pr(x) = 0$, x may happen, though rarely as compared with other events represented in the probability space. (All breakthroughs or revolutions, in any field, have low probability, perhaps vanishing probability, yet they happen and are the most important events.) Consequently a fact with probability 1 can fail to happen. (Recall that any set of rational numbers has zero Lebesgue measure. Entire sets of states and events are assigned zero probability in statistical mechanics for this very reason even though the system of interest is bound to pass through them. This is what 'almost never' is taken to mean in that context, namely that the states or events in question are attained only denumerably many times.)

Third, the frequency column should be retained alongside the propensity interpretation though in a capacity other than interpretation or definition. Indeed, although the frequency column fails to tell us what '$Pr(x) = y$' means it does tell us under what conditions such a formula is true. Long run frequency is in short a truth condition for probability statements. Besides, frequency statements have a heuristic value. For example, if p means a transition probability, then the greater p, the more frequent or common the transition.

Fourth, note again that the present propensity interpretation differs from Popper's in that the latter requires the system of interest to be coupled to an experimental device. No such hang-up from the frequency (or empiricist) interpretation remains in our own version of the propensity interpretation. Nor do we require that only events proper (i.e. changes of states) be assigned probabilities, as an empiricist must, since states may be unobservable. States too may be assigned probabilities, and in fact they are assigned in many a stochastic theory, such as statistical mechanics and quantum theories. (The statistical mechanical measure of entropy is a function of the thermodynamic probability of a state; or, as Planck put it, measures the preference (*Vorliebe*) for certain states over others.) In other words not only transition probabilities (which are conditional) but also absolute probabilities can be factually meaningful.

Fifth, note that the propensity (or any other) interpretation of probability is to be distinguished from the probability elucidation (or exactification) of the intuitive or presystematic notion of propensity, tendency, or ability. In the former case one attaches factual items to a concept, whereas in the latter one endows a factual concept with a precise mathematical structure. In science (and also in ontology) we need both factual interpretation and mathematical elucidation.

Sixth, the propensity interpretation presupposes that possibilities can be real or physical rather than being just synonymous with our ignorance of actuality. On the other hand according to the frequency interpretation there is no such thing as a chance propensity for a single thing: there would be only limiting frequencies defined for entire ensembles of things or for whole sets of events in a single thing, such as a sequence of throws of a coin, or a family of radiative transitions of a kind. Indeed, the phrase '$Pr(x) = y$' is, according to the frequency school, short for something like 'The relative frequency of x in a large ensemble (or a long sequence) of similar trials is observed to approach y'. This view is refuted by the existence of microphysical theories concerning a single thing, such as a single atom, to be sharply distinguished from a theory about an aggregate of coexisting atoms of the same kind. Another example: in principle, genetics is in a position to calculate the probability of any gene combination – which, given the staggering number of possibilities, is likely to be a one-time event. A relative frequency is a frequency of actuals, hence it cannot be identical with a possibility (measured by a probability). Unlike frequencies, probabilities do measure real (physical) possibilities. Therefore if we take real possibility seriously, i.e. if we are possibilists rather than actualists, we must favour the propensity over the frequency interpretation.

In short, there are a number of reasons for favouring the propensity interpretation of probability over its rivals.

Concluding remarks

We have examined four concepts of probability, each of them in three respects: mathematical validity, scientific viability, and philosophical plausibility. The concepts in question are the following: the semiabstract concept defined implicitly by the theory of probability, the notion of personalist (or subjective or Bayesian) probability, the frequency conception, and the propensity interpretation.

We have taken the majority view that the mathematical concept of probability is adequately characterized by the standard theory of probability developed along the lines of Kolmogoroff's work. Both the personalist and the frequency notions turn out to be mathematically untenable because they are at variance with the standard theory of probability. The personalist concept is mathematically invalid because the probability measure defined in the calculus of probability makes no room for any persons. And the frequency view is mathematically incorrect becase (a), the axioms of the probability calculus do not contain the (semiemepirical) notion of frequency, and (b), some of the key theorems of the theory of probability, such as the laws of large numbers, concern the differences between probability values and frequencies. Only the propensity interpretation of probability was found mathematically unobjectionable.

As for the use of the various probability concepts in science and technology, the semiabstract concept cannot be applied without further ado: it must be turned into a concept with a factual meaning. This is done by interpreting the basic space S (out of which the probability space F is constructed) in terms of factual items such as the states or the changes of state of a concrete thing. Consequently an arbitrary value $Pr(x)$ of the probability function, for x in F, means the weight or strength of the state(s) x, or else the tendency or propensity for event(s) x to happen. This is the meaning to be assigned to statements of the form '$Pr(x) = y$' occurring in the stochastic theories of pure and applied science. Such statements are objective and in general also testable.

On the other hand science and technology have no use for the personalist view precisely because it is subjective. Who but a dogmatist or a biographer could be interested in pronouncements such as 'Bayesian X attaches credence Y to theory Z'? As for the frequency view, it is not viable in science either because it conflates calculated probabilities with observed frequencies, thus preventing the latter from discharging the function of testing the former. However, the frequency conception has at least some heuristic power, which the personalist does not. For example, if p is the viability (survival probability) of organisms of a certain kind, present in number N, then pN may be interpreted as the fraction of surviving organisms – although strictly speaking pN is only the average number of survivors. We may be permitted to reason that way provided it helps and no traces of such heuristic props remain in the end. And they must not remain if only because actuality (such as the actual fraction of survivors) should not be confused with possibility (as measured by the most probable number of survivors). Besides, the propensity concept is at least as heuristically fertile as the frequency misconception.

The propensity interpretation of probability is then the only one that fits the mathematical theory of probability and is also adaptable to science. Moreover, it is the only one that fits in with a realistic theory of knowledge, whereas the Bayesian view is consistent with a subjectivist epistemology, and the frequency view invites an empiricist theory of knowledge. Finally, both the Bayesian and the frequency views presuppose classical determinism of the Laplacean style: they equate possibility with conceptual possibility and they deny the reality of randomness. On the other hand the propensity view takes real possibility seriously and admits the reality of randomness. If a concrete thing, be it atom, organism, or community, has propensity $Pr(x)$ to be in state(s) x, or to experience change(s) x, then this is a property the thing possesses independently of our beliefs – a property that can sometimes be checked by observing actual frequencies but not be confused with the latter.

In short, the propensity interpretation of probability is consistent with the standard theory of probability and with scientific practice, as well as with a realist epistemology and a possibilist ontology. None of its rivals has these virtues.

References

1 Venn, J. 'The logic of chance' (3rd edn), Macmillan, London, 1888
2 Smoluchowski, M. von, 'Uber den Begriff des Zufalls und den Ursprung der Wahrscheinlichkeitsgesetze in der Physik. *Naturwissenschaften* 1918, 6, 253
3 Fréchet, M. 'Les définitions courantes de la probabilité' Reprinted in 'Les mathématiques et le concret', pp. 157–204. Presses Universitaires de France, Paris, 1946
4 Popper, K. R. 'The propensity interpretation of the calculus of probability and the quantum theory'. In S. Körner (ed.), 'Observation and interpretation'. Butterworths, London, 1957
5 Kolmogoroff, A. 'Grundbegriffe der Wahrscheinlichkeitsrechnung', Springer-Verlag, Berlin, 1933
6 Bunge, M. 'The furniture of the world', Vol. 3 of 'Treatise on basic philosophy', D. Reidel, Dordrecht, 1977
7 Bunge, M. 'Interpretation and truth', D. Reidel, Dordrecht, 1974
8 Finetti, B. de, 'Probability, induction and statistics', John Wiley, New York, 1972
9 Jeffreys, H. 'Scientific inference' (3rd edn), Cambridge University Press, Cambridge, 1975
10 Savage, L. J. 'The foundations of statistics', John Wiley, New York, 1954
11 Wald, A. 'Statistical decision functions', John Wiley, New York, 1960
12 Ville, J. 'Etude critique de la notion de collectif', Gauthier-Villars, Paris, 1939
13 Settle, T. 'Induction and probability unfused'. In 'The philosophy of Karl Popper' (P. A. Schilpp, ed.), Book II, pp. 697–749. La Salle, Ill.: Open Court, 1979