Under consideration for publication in Math. Struct. in Comp. Science

Debate On The Concept Of Probability

Received 28 october 2011, 30 Novembre 2011

1. Opening of the session

Giuseppe Longo

The main promotor of this special issue, Mioara Mugür-Schachter, organized for us a final debate for the one day conference dedicated to the theme of the issue. Here are her questions, as an informal and open guidelines for the speakers, who were asked to provide a written version of their answers and reflections.

1.1. Opening questions

In order to induce a common structure into the various contributions, we have called upon answers to the following four definite questions.

- 1 Do you associate some own and definite *significance* to the assertion that in a given factual situation, there 'exists' a probability law?
- 2 Do you feel satisfied with the law of large numbers as a definition of the individual numerical probabilities to be assigned to the events produced by a random phenomenon?
- 3 (Question addressed to the contributors that are interested in modern biology). What do you think of the current use and role of randomness and of probability in Biology? Can a suitable concept of biological randomness be directly borrowed from Physics and if so from which physical theory?
- 4 (Question addressed to the contributors that work in the field of quantum physics): Do you think that the quantum mechanical 'postulate' of probability can somehow be derived? Do you think that this postulate can be applied to any *factually* realizable microstate?

1.2. The debate

The answers given by each participant are reproduced in alphabetic order and in the form conceived by that participant. They are followed by comments expressed by other participants.

C. S. Calude: Probability, Randomness, Entropy: Convenience or Necessity? Answer to question 1. No.

Answer to question 2. Randomness is not absolute, it is a matter of degree. The degree of randomness relevant for the studyof a phenomenon *depends on* the phenomenon and the model adopted (which implies the choiceof some probability space). The law of large numbers corresponds to one of the possible degrees of randomness. Contextually, it can be adequate or inadequate. For example, for quantum random generators he law of large numbers is too weak to be adequate.

Answer to question 3. Randomness in biology seems to be used in a metaphoric sense rather than a mathematical one, so I don't feel qualified to comment on it.

Answer to the first question 4. . Yes: working on this issue we have obtained partial results.

Answer to the second question 4. . No

General remarks. Probability, randomness and entropy are presented from two different perspectives: Kolmogorov's axiomatic approach, and algorithmic information approach.

The first, the classical theory, develops a calculus with unspecified "random events".

The second theory defines and studies various "(algorithmic) random objects", like random finite strings or random infinite sequences. The adjective "algorithmic" comes from the essential role played by the theory of computation in the second approach.

One can show that "true/pure randomness" does not exist from a mathematical point of view: only degrees of randomness, stronger or weaker, can be defined and studied. Randomness of various degrees should not be feared only, but also revered: indeed, randomness is extremely useful in different contexts, in particular, in enhancing computation. Quantum randomness can andwill be better understood, hence it will play an increasingly important role.

Comments

MMS. For me your answer to the question 2 is highly interesting. It suggests that you regard "randomness" as just another face of "probability". Such a view, when connected with my own contribution to this volume, might quite unexpectedly entail a mathematically expressed but semantic definition of degrees of randomness produced by any given factual random phenomenon (a global degree, as well as "local" degrees of randomness per event).

Maria Luisa Dalla Chiara

The four questions proposed to our debate represent some very crucial and fundamental problems of probability theory. Since I am not close to biological research, I will only try and comment on the first, the second and the fourth question.

Let me start with a general (and perhaps trivial) remark: It seems to me that many

Debate On The Concept Of Probability

results of contemporary investigations show that a pluralistic view in probability cannot be avoided both from the mathematical and from the empirical point of view.

Apparently, Kolmogorov's axiomatic theory represents a kind of perfect general mathematical framework, which all particular applications of the notion of probability should be compared with. To what extent can this theory be regarded as a purely syntactical approach? In all axiomatic versions of theories (either in mathematics or in physics) syntax plays an essential role. However, syntax without semantics is, in a sense, 'blind'. Generally, the proposal of a set of axioms leads, in a natural way, to investigate a class of *semantic models* (or *interpretations*) for the axioms in question; and very often such models are connected with some intuitive ideas that had originally stimulated the axiomatization-attempt. Paradigmatic examples that might be recalled in this connection are, of course, innumerable, both in the case of *concrete* mathematical theories (like arithmetic, set theory,...) and of *abstract* algebraic theories (like group theory, lattice theory, ...). Apparently, semantics represents an essential part of any axiomatic enterprise.

As is well known, in the particular case of Kolmogorov's axiomatization, different kinds of models (which may be connected with different possible applications of the general notion of probability) have been proposed. We need only think of the relevant role played by the *subjective* interpretations of probability. In the case of empirical sciences (and of physics, in particular), *frequency-models* have often been considered the most attractive and adequate interpretations with respect to the concrete scientific practice. But what is the precise notion of frequency that is involved in such models and what is the role of the *law of large numbers* in this connection?

Let us refer to the case of physics. For all physical theories (from classical to quantum mechanics) the interaction between the unsharp world of the experimental situations and the *sharp* world of the mathematical representatives of experimental concepts plays an important role, although the formal reconstruction of such crucial interactions is not always clear. Apparently, the concept of probability occurs in both worlds, as a sharp and as an unsharp concept at the same time. In most cases, what physicists can in fact measure or test are relative frequencies that are to be correlated with the sharp probability-functions dealt with in the theoretical framework. And the comparison between the unsharp and the corresponding sharp concepts can be investigated by taking into serious account the measurement-precisions that are involved in the experimental situations. As expected, in the case of probability, such precisions are naturally connected with the number n of cases that have been experimented. From a formal point of view, the relative frequencies that can be physically measured can be reconstructed in the context of approximate models, where approximation represents an intrinsic feature that cannot be simply regarded as a kind of *epistemic* limitation. In this framework, there is no need to refer to the critical *limit* of relative frequency-sequences (which represents the well known crucial difficulty of the frequency-definitions of probability, proposed by von Mises and Reichenbach). I wonder whether the *factual probabilities* studied in the article 'On the concept of probability' by Mioara Mugur Schächter might be semantically reconstructed by means of convenient approximate models. Accordingly, it seems to me that the first question (Do you associate some own and definite significance to the assertion that, in a given factual situation, there 'exists' a probability law?) might be usefully discussed in the framework of an approximate semantics, where factual (or empirical) probabilityfunctions can only be determined up to a given **precision**, which essentially depends on the particular concrete situation under investigation. Following this perspective, it seems natural to give a negative answer to the second question (*Do you feel satisfied with* the law of large numbers as a definition of the individual numerical probabilities to be assigned to the events produced by a random phenomenon?).

Let me now turn to the case of quantum mechanics and to the fourth question (Do you think that the quantum mechanical 'postulate' of probability can somehow be derived? Do you think that this postulate can be applied to any factually realizable microstate?). I wonder: from what should the quantum mechanical 'postulate' of probability be derived? The interplay between the unsharp experimental world and the sharp world of the mathematical representatives of physical concepts is perhaps more intriguing in quantum theory than in classical mechanics, also because, in the quantum case, uncertainty and approximation have become an essential part of the theoretical framework itself. What is particularly critical is the operation that assigns a *pure* (or a *mixed*) state to an individual quantum object or to a statistical ensemble of equally prepared objects. Anyway, such operation is currently done in quantum mechanics, where the so-called Born-probability plays an essential role in the context of the Hilbert-space mathematical environment that is associated to quantum phenomena. Since the Born-rule represents one of the primitive and constitutive ingredients of axiomatic quantum theory, I cannot see how this rule could be derived from something else. Interestingly enough, such rule has shown a deep *stability* in the framework of generalizations or applications of quantum theory (as an example, we might recall the case of unsharp quantum theory or the case of quantum computation). One of the most important interactions between quantum mechanics and probability theory is due to the fact that the quantum world gives rise to some deeply non-Kolmogorovian behaviours. Counterexamples to Kolmogorovaxiomatization have been brought into light according to different modalities, in different quantum situations. A standard example arises in the framework of the quantum logical approach, where (following Birkhoff and von Neumann's proposal) quantum events are mathematically represented by closed subspaces of a Hilbert space. In such a case the structure of all events turns out to be an orthomodular lattice, which is weaker than a Boolean algebra (structure that is requested by Kolmogorov's axioms). More recently, a different counterexample has been investigated in the framework of quantum computation, where quantum computational events are mathematically represented by pieces of quantum information (identified with qu-registers and mixtures of qu-registers). Also in this case, the structure of all events turns out to be deeply non-Boolean. For instance, the conjunction of two events is generally non-idempotent (the event A and A may be different from the event A), while the non-contradiction principle may be violated (contradictions are not necessarily impossible).

Let me conclude with a general intuitive remark. The construction of different kinds of models seems to represent a characteristic feature of what we call 'creative activities' both in the case of formal and of empirical sciences. In this sense, scientific imagination seems to be very close to artistic imagination. By using a theatrical metaphor we might say that, as happens to artists, scientists also are used to invent scenes and stories in some possible worlds, which can be intuitively regarded as a kind of ideal stages.

Comments

MMS. To begin with, I specify that throughout this debate I consider exclusively the frequential concept of probability.

Once this stated I fully agree with your remark that a syntax that would not incorporate any feature of semantic source is not conceivable. For a formal system *is* no more than just a global meta-description of the whole class of scientific descriptions of entities from a given domain of material reality: these constitute the very matter out of which the considered syntax is constructed by a particular sort of abstraction that filters out the deductive relations among semantic features that act inside that class.

But this does not exclude essential differences between the basic descriptions from that class and any corresponding syntax drawn from these: Inside the semantic matter offered by the basic descriptions concerned by a syntax there emerge quasi continuously, by a sort of effervescence maintained by a multitude of more or less anonymous researchers, new or modified basic description of various and moving contents and with mutually distinct statuses (experimental, theoretical, technical); whereas the emergence of a corresponding syntax is a — relatively — very rare event which, once realized, is endowed with a relatively long-lasting stability. Furthermore the basic descriptions are constrained by the aim to create some consensual new knowledge concerning entities from the considered domain of reality, permitting new predictions or uses. While the corresponding syntactic meta-description is constrained by the fundamentally different aim to just abstractly transform — rapidly, securely and without altering the truth value — sets of basic descriptions that are explicitly known, into new basic descriptions that are as yet not known in an explicit form; this meta-description works like a conceptual machine of which the general structure is strictly and stably codified. So even though the genesis of a syntactic structure is intimately determined by the corresponding semantic matter, once constructed this syntactic structure is quite distinctly individualized (similarly, in this respect, to what happens for living beings).

Now, in order to put to work an achieved syntactic machine, one has to *introduce* in it basic descriptions brought into a state on which this syntactic machine be capable to perform the desired transformations. A syntactic system, by its very essence of a meta-description of exclusively the *general* features of the basic descriptions concerning entities from a whole semantic domain, does not offer any possibility to *generate* inside it the semantic prime matter, the *data* that specify this or that particular semantic basic description so as to be workable via the considered syntax. In many cases of application of mathematical tools, just trained common intuition suffices for introducing the semantic problems into a syntactic tool corresponding to them. But in the case of probabilities this is *not* so: the semantic to be fed into Kolmogorov's syntax *must explicitly contain the definition of the numerical distribution of the probabilities of the involved individual events*, and this is a very sophisticated datum that has to be deliberately constructed inside the semantic matter itself. The probabilistic semantic has to be worked out like a

key to be introduced into a corresponding keyhole pre-organized inside the syntax. This seems to be little realized. The fact that the numerical distribution of the probabilities of the involved individual events cannot be obtained inside the probabilistic syntax via formal transformations, that it is a factual datum that has to be introduced from outside the syntax, that Kolmogorov's syntax can only offer *location* to such a numerical distribution, seems to be rather nebulously grasped. So much so that R.J. Solomonoff felt it necessary to stress this fact[†]:

 \ll Probability theory tells how to derive a new probability distribution from old probability distributions...It does *not* tell how to get a probability distribution from data in the real world. \gg

(The mentioned sort of formal transformation of a given probability law into another one is performable by use of the general features introduced by the definition of a probability *measure* and of the relations posited in a probability space between "elementary events" and "events" from an algebra put on the set of elementary events).

This illustrates strikingly the general fact that, if one wants a formal system to be useful with respect to the semantic domain incorporated in its structure, so as to constitute with it a "physical theory", then certain *matching relations* must be imposed upon, both, the basic descriptions from this semantic domain and the corresponding syntax.

Why should this task be left to a spontaneous evolution of the probabilistic thinking that might accomplish it at any remote time, or never?

This brings us to your important remarks concerning the interactions between the unsharp world of factual situations and the sharp world of mathematical representations. These remarks seem to recommend tolerance of certain non-effective definitions. I quite fully agree with these remarks, in particular with the way in which they involve the notions of mathematical limit versus precision, epistemic limitation, and reconstruction by approximate model (by the way, this is precisely what I have tried). Indeed nothing would justify a draconian interdiction of any one among these methods. But I stress that in the case of the concept of a factual probability law, the definition offered by the law of large numbers is also vitiated by a sort of circularity, not only by a noneffectiveness that might be regarded as only an "unsharp" feature common to all the natural "laws". And this, in its essence, is connected with the use of the principle of Laplace, so with the indefinite oscillations that this principle permits between a priori assumptions and a posteriori "verifications", for which the specification of an effective modality of realization might introduce another indefinite recession: all this, considered globally, indicates strongly that the in the case of the concept of probability the "matching relations" between the semantic concept of probability, and Kolmogorov's syntax, simply are not coherently worked out.

As for a deductive introduction of the Born-rule — apart from the obvious logical obstacle — I would like to point out that the interest of the rule itself appears as strongly diminished as soon that it is realized that among the *physically realizable* microstates,

[†] R.J. Solomonoff [1957] ("An inductive inference machine", *IRE National Record*, 5, (quoted in Segal, J., [2003], *Le zéro et le un*, Syllepse).)

those for which the state vector can be formally specified, are relatively very rare in consequence of the requirements of: the existence, in principle, of a hamiltonian (like for any bound state); the possibility to express mathematically this hamiltonian; the solvability of the corresponding Schrödinger equation; and in fine, the possibility to express mathematically the limiting conditions, so as to specify the solution out of an infinite family; etc.). The Born-rule is just a postulate of an *in principle limited applicability*, that includes the word "probability". But this word points toward a semantic concept that even in the macroscopic classical thinking is not yet fully worked out; and in order to fill this gap inside the *classical* thinking, the unique procedure proposed so far, I think, is contained in the present volume, and it requires "complexifications" that *cannot* be associated with fundamental quantum mechanics, of the type of that proposed by de Broglie and then by Bohm, or of some other type.

When one realizes this, it appears how far we are in fact from the miraculously precise power of "probabilistic" prediction of quantum mechanics such as it is asserted in a non restricted way. This power does *not* exist for *any* realizable microstate, it exists only for a relatively very small class of microstates, either for natural but *simple* bound eigenstates of a total energy with discrete spectrum of eigenvalues, or free states that are deliberately generated so as to permit probabilistic predictions that are known *by construction* (like in the case of the macroscopic games of chance). Such a predictability by construction is realized in the case of the free superposition microstates considered in the domain of quantum computation; but in this case the probabilistic predictability, *as such*, is moreover only approximate, because of the continuous character of the spectrum of the momentum.

Stefano Galatolo

Answer to question 1. My first answer is "probably yes".

More seriously the question is deep, and if considered generally, it may have some ambiguity.

The choice of a suitable probabilistic model is related to the choice of a conceptual (mathematical) model for the factual situation we mean to describe. If we restrict our attention to factual situations modelled by some precise class of mathematical models, the question become more concrete. For example, in dynamical systems the question becomes related to the choice of a "natural" measure on the system's phase space.

If we are interested in the long time behaviour of the system and in statistics, then many important information on the behaviour of the system are encoded in invariant measures, and if you want to select (among the many measures that are invariant, the most physically meaningful ones, we arrive to the concept of natural (physical) invariant measure which is much studied in the theory of dynamical systems (see our contribution to the special issue).

Now, the practical question becomes "can we have an estimation for the **probability** of a given event?" This corresponds to compute the natural invariant measure, given the

description of the system. Sometimes this is possible in principle, sometimes not (see our contribution). Yet the search for efficient algorithms is open.

Answer to question 2. The law of large numbers (ergodic theorem and similar) is a necessary relation between the statistical behaviour of a system and the probabilities of the underlying chosen model. In practice it can be used to extract information about a suitable model to be considered by the statistics of the observed phenomena. In this sense it could be a "definition".

By observing the system's statistical behaviour we can decide the probabilities of a good model to represent it. Moreover, as remarked in the discussion, the statement of the law of large number involves two times the use of the concept of probability. The ergodic theorem (a general version of the law of large numbers) says that if we consider an ergodic system with invariant measure μ , time averages computed along μ -typical orbits coincides with space average with respect to μ . More precisely, for any suitable observable f it holds

$$\lim_{n \to \infty} \frac{S_n^f(x)}{n} = \int f \, d\mu,$$

r

for μ almost each x, and where

$$S_n^f = f + f \circ T + \ldots + f \circ T^n$$

Here μ is used as a measure to calculate $\int f d\mu$ (the space average) but also is used to define a suitable concept of almost each (the law is respected everywhere but a set of zero μ -probability). Here, depending on what is meant to be described we could choose another measure do define another suitable *almost each* concept, as for example a measure having some "physical" meaning on the underlying space where the dynamics act (the Lebesgue measure for example). This also can help to make a good choice on the probabilistic model for the phenomena. In dynamical systems for example, this can be made by using the Lebesgue measure for the a. e. concept and look to which μ makes the above equation work. This is related to the concept of physical measure mentioned above.

Answer to question 3. I think that even this problem is related to the choice of a suitable conceptual model. In biological situations this is very hard, partly because of the complexity of the involved systems (too complex models are useless). I think this is one of the main open problems of the whole science.

Differential equations and related concepts turned out to be a wonderful tool to modelize and understand Mechanics. What about biological systems?

Comments

MMS. It seems to me, but I am not quite certain, that you think that a "mathematical model" (which you identify with "conceptual model" and, may be, even with "formal

Debate On The Concept Of Probability

system"?) is able to generate the factual particular data that are necessary for treating a definite concrete problem (a sort of Platonic conception on the mathematical representations and their unrestricted power?).

If I am right, you certainly are not alone to hold this view. In this case my own reaction consists of a generalization of certain remarks contained in my preceding comment on Marisa Dalla Chiara's answers: Not only formal systems, but also mathematical models, are always placed on a meta-level of conceptualization with respect to the modelled basic scientific descriptions of this or that definite particular physical case or phenomenon or situation from a given domain of physical reality. They always are extracted from a class of such basic factual descriptions, in consequence of a process of *abstraction* of data that singularize a given basic description of a particular case. Therefore — quite essentially, by construction — the mathematical models, like the formal systems, once constructed, are unable to generate back, inside themselves, a specification of factual data that singularize a given particular basic description and have been abstracted away. Such a specification can only be found again in the initial basic descriptions. In the case, now, of the concept of probability, Kolmogorov's syntax with its general concept of a probability measure cannot generate the particular factual numerical probability distribution that characterizes the description of a concrete given probabilistic situation. While on the other hand, in general, the basic descriptions of particular probabilistic situations do not offer either an effective and non-circular specification of these numerical distributions of probabilities. Precisely this is the whole problem raised by the nowadays concept of probability: its basic semantic matter is not achieved so as to match satisfactorily the syntactic machine constructed by Kolmogorov so as to draw some profit from its existence.

And I cannot see how *other* domains from the mathematical theory of abstract measures (Lebesgue measures, ergodic theorem, etc.) that *less* intimately connected with probabilistic problems than a Kolmogorov probability measure, could compensate for this gap that is quite specific of the concept of probability: this gap is of *semantic* nature, **not** of mathematical nature.

Annick LESNE: How physicists perceive the syntax of Kolmogorov (dialogue with Mioara Mugur-Schächter)

AL. My position is that adopting a probabilistic approach is only an epistemic choice, justified by the scale of the description, the observations that are to be explained, or the predictions that are to be done.

MMS^{\ddagger} This is a deep remark that I do not in the least contest. I quite agree with you that the fact of considering, or not, an empirical situation as being "probabilistic", depends on the scale of observation, and so, that in *this* sense, one is facing an epistemic choice. But the problem that I raise — and that K himself has strongly raised — is quite *independent* of the scale of observation. It consists of the fact that when

[‡] I ask the reader to excuse the repetitions that mark my comments of the answers of the different participants: these repetitions are recursively induced by a view that appears to be more or less uniformly shared by the other participants, and with which I quite decidedly do not agree.

you do choose to consider that this or that empirical situation is probabilistic with respect to the involved scale of observation, then — in general — you cannot make use of Kolmogorov's syntax in order to treat mathematically the questions tied with this empirical probabilistic situation: the very **possibility** of principle to make use of the modern probabilistic syntax of Kolmogorov (K) requires conditions that are not insured by the nowadays "theory" of probabilities: Kolmogorov's syntax — like any syntax introduces a certain structure of abstract *void places* ("squares", cases). These, in the case of K's syntax, are those denoted $(\Pi, U), [U, \tau, p]$ (where the significance of the symbols is well known). The general mathematical structure of the triad $[U, \tau, p]$, is specified, but the particular numerical content of the general mathematical concept "p" (a probability measure) is not specified. This particular numerical content is a semantic datum of which it is supposed that it *can* be introduced when it is desired to make use of Kolmogorov's syntax for treating a problem tied with a particular empirical probabilistic situation. But in fact, in general this assumption is **not** fulfilled. There does not exist a general consensual way for specifying the numerical distribution of probabilities that corresponds to a given empirical situation unanimouslt considered to be "probabilistic". K has called this situation "the problem of applicability" (cf. the manuscript of C. Porter), while I myself, not being informed of Kolmogorov's vocabulary, have called it "the aporia of Kolmogorov's syntax". (This aporia is an analogue of the situation that would realize if, when one tried to express in terms of an equation of the abstract form $ax^2 + bx + c = 0$, a problem considered to belong to the domain of applicability of the syntax of the equations of the second degree, such an expression in fact would not be realizable, because of some systematic obstacles in the way of the process of identification of the factual element to be represented by "x" and of the factual elements that specify the numerical values of the coefficients a, b and c (cf. also the quotation from Solomonoff in my previous comment)).

In short, it seems to me that the essential difference between our views is that you think that, when this is desired, the syntax of K *can* be made use of by the physicists. Whereas I think to have shown that in the present phase of development of the semantic concept of probability, this is not possible even if it is desired, because the semantic concept of probability does not provide a way for defining the involved factual probability law, nor does it prescribe a useful way to connect the semantic probabilistic problem to K's syntax.

AL. This perfectly summarizes our respective viewpoints and it underlines the gap between probability theory and statistical estimation, which addresses precisely the issue of determining empirical probability distribution. I agree that K's theory is only a syntax. However, I defend the practical position that it is possible to devise the model directly within a probabilistic framework that is consistent with K's syntax, by exploiting tools of probability theory (e.g. limit theorems) to compute derived properties (e.g. predict a collective behaviour). The probabilistic model is so to say constructed *ab initio*. The bridge with experimental data, or other models is done at a higher level, e.g. of averages and moments.

For me, the origin of randomness in classical physics is well-understood: it comes from the sensitivity of the dynamics to its initial conditions and surroundings, such that the outcome is influenced in a qualitative way by minute factors, e.g. the way one initially grasps the dice, the hygrometry or a car passing in the street. As we are not able to identify all the potential initial or external influences, we straightawayturn to a probabilistic description of the outcome.

This has been formalized for chaotic dynamics, where one investigates the stationary probability distribution in the phase space (what is termed the "invariant measure") rather than trying to follow exactly one specific trajectory (see for instance the contribution of Stefano Galatolo and coworkers in this special issue).

There is usually no ambition to *reconstruct* the stochastic process that actually led to this stationary distribution. This, however, *could be done in an effective way* by using, for instance, Langevin's approach, i.e. by adding an effective term of noise to a deterministic equation to better account of the reality). Again the parameters of the added noise are tuned by matching some observable features. The probabilistic model obtained in this way is then suitable for making further predictions.

My favourite example in this respect is that of diffusion [Castiglione et al. 2008]. Several coexisting descriptions can be considered, either deterministic ones(molecular dynamics, Fick's law and the diffusion equation) or stochastic (master equation, Fokker-Planck equation, Langevin equation). All do properly define the same diffusion coefficient. The choice and relevance of a given formalism is here determined by the scale of the description, i.e. the observable quantities onewants to account for: either the detailed positions and velocities of all the molecules, or their distribution, or the probability of presence of one molecule, or the local concentration of molecules. No one is "more true" than another, they simply account for different perceptions of the reality, hence they are more or less useful depending on the experimental data or goals.

Another example is the theoretical approach of disordered media. In order to understand the generic properties of disordered media (e.g. alloys or porous rocks), one abandons a detailed description (the position and size of the holes in the porous rock) and turns to a probabilistic description. Investigating a random medium captures the properties common to all disordered media, whereas the results obtained in investigating a given disordered medium mix up generic properties and specific behaviour of the considered sample. This recovers the core principle of statistical mechanics, consisting in working with statistical ensembles instead of a specific realization of the gas or liquid sample.

The first mention of an effective probabilistic approach can be traced back to the daemon of Laplace, who would be able to predict exactly the future, given a complete knowledge of the present state of the Universe. Since getting such a complete knowledge is obviously impossible for anyone else, one relies on a "principle of indifference", which can be seen as a probabilistic approach. The principle of maximum entropy, later formalized by the approach of Laplace, provides a constructive way to choose the shape of the probability distribution, e.g. a parametrized family. The parameters are then estimated by fitting the predicted moments to the observed moments.

Much debate arose about the different interpretations of the very notion of probability [Cox 1946]. Basically, the frequentist interpretation (implicitly or explicitly referring to a set of independently repeated events, e.g. coin tossing, or the probability of developing a cancer during the year for a typical individual in a given population) is opposed to a notion of a "single-event" (e.g. the chance that there exists another inhabited planet in our galaxy, or the risk that Mr. Smith will develop a cancer this year), sometimes called propensity. Although this debate possesses much importance on philosophical grounds, I think that it is of no concern for a physicist, insofar as all the mathematical framework of probability theory is equally valid in both views. Adopting one or the other is a matter of personal feeling; this will not change the equations and the quantitative predictions that follow.

The former probabilities will be estimated as frequencies over a (large) sample of observations, the latterpropensities will be estimated within a Bayesian approachin which any additional piece of information or observationis used to improve a prior estimate.

MMS. If I understand correctly, you argue that K's syntax acts as just a structure that inspires and guides the empirical ways of dealing with this or that concrete probabilistic situation, in order to construct a numerical distribution of real numbers that satisfies the general conditions imposed by K upon the abstract concept of a probability measure?

If this is indeed your view, I quite agree that human intuition is marvellous and that in any given empirical probabilistic situation it finally succeeds to determine a probability distribution, by appropriately combining various sorts of techniques, either pre-existing techniques or newly invented ones.

This however does not lead to what can be properly called a "theory" of probabilities. In order for such a theory to exist it is necessary to be able to specify general conditions and rules that define a quite general uniform way of associating K's probabilistic syntax, with any empirical probabilistic situation (like in the case of the formalisms of classical mechanics, of electromagnetism, of thermodynamics, of quantum mechanics). Indeed a mathematical "theory" of some definite domain of reality consists, by definition of the concept, of a mathematical syntax that is connectable in a *generally* specified and *generally* performable way, with that domain of reality.

This being stated, it is true that the absence of a theory of probabilities is not a catastrophe. The physicists always succeed to throw together a set of real numbers that express a numerical distribution of probabilities compatible with, both, the considered empirical situation and K's abstract concept of a probability measure. But is that a reason for not trying to construct a theory of probabilities in the full sense of this locution? Nothing hinders to set such a construction as an aim, as K did. With respect to such an aim, the present situation concerning the concept of probability can be regarded as a problem to be solved.

AL. As for quantum mechanics, in my viewpoint, that of a classical physicist, chance in our surrounding world, for instance coin tossing or throwing of a dice, differs from the probabilistic nature of quantum mechanics. Whether the classical chance should be called "subjective" and the quantum chance called "objective" can be an interesting matter of debate. At least, I underline that the corresponding mathematical notions differ: the central notion for classical chance is that of stationary distribution, whereas the probabilities in quantum mechanics correspond to square amplitudes of underlying wave functions. There is no analogue of a wave function in the classical physics.

Mioara MUGUR-SCHÄCHTER

Answer to question 1. Yes, I do (in the context of this discussion the term 'exists' points either toward a materially constructed and known 'pre-existence' — as for instance in the case of games of chance — or, more generally, toward only a possibility of conceptual construction). I hold that the "significance" of the existence of a probability law] (not its existence alone) consists of a global that in general is unknown, and of which the directly perceived events are observable fragments. The factual probability law on the involved set of events, considered globally, is a parcelled cryptic expression of this unknown global form.

Answer to question 2. No, I am not, for the following reasons.

Consider a random phenomenon (Π, U) . The current mathematical expression of the theorem ("law") of large numbers, is:

$$\forall j, \forall (\epsilon, \delta), \qquad \exists N_0 : \forall (N \ge N_0) \Rightarrow \mathcal{P}[(|n(e_i)/N - p(e_i)|) \le \epsilon] \ge (1 - \delta)$$

This reads: Given a universe of events $\{e_j\}, j = 1, 2, ...q$. *IF* a *factual* probability law $\{p(e_j)\}, j = 1, 2, ...q$ on the universe $\{e_j\}$ *does* exist, *THEN* — for every e_j and every pair (ϵ, δ) of two arbitrarily small real numbers ϵ and δ — there exists an integer N_0 such that, when the number N of 'identical' reproductions of the experiment Π becomes equal to N_0 or bigger than N_0 , the meta-*probability* \mathcal{P} of the meta-event denoted $[(|n(e_j)/N - p(e_j)|) \leq \epsilon]$ that [the absolute value of the difference $(|n(e_j)/N - p(e_j)|$ between the relative frequency $n(e_j)/N$ of the event e_j and the individual numerical value of the factual probability $p(e_j)$ of that event e_j be smaller than or equal to ϵ], becomes bigger or equal to $(1 - \delta)$.

This can also be expressed by saying that: when N tends toward infinity the relative frequency $n(e_j)/N$ of the event e_j "tends in probability" — namely via the meta**probability** \mathcal{P} — toward the *a priori unknown* value of the probability $p(e_j)$ of the event e_j . This formulation stresses "the frequential definition the concept of probability" contained in the law of large numbers.

So the "frequential concept of probability" specifies the factual probability distribution $\{p(e_j)\}, j = 1, 2, \ldots q$ of the individual numerical values of the events e_j , as a set of *probable and limit-values* of the relative frequencies $n(e_j)/N$ when N increases toward infinity[§].

This definition is essentially distinct from Kolmogorov's purely formal definition of a probability "measure" where the individual numerical values $\{p(e_j)\}, j = 1, 2, \ldots q$ are not taken into consideration, so, a fortiori, they are not defined.

Initially Kolmogorov has *founded* his syntax on this frequential definition because he

[§] In Christopher Porter's contribution to this volume it is very interestingly shown that this definition — asserted by von Mises in full generality — has been subjected by this author to also *other* conditions concerning the choice of "relevant" sequences of N events e_j .

thought that it was a sound factual, semantic definition[¶]. But finally he rejected this semantic definition because of its limit-character that entails non-effectiveness:

« I have already expressed the view...that the basis for the applicability of the results of the mathematical theory of probability to real random phenomena must depend in some form on the frequency concept of probability, the unavoidable nature of which has been established by von Mises in a spirited manner... (But) The frequency concept (of probability) which has been based on the notion of limiting frequency as the number of trials increases to infinity, does not contribute anything to substantiate the **applicability** of the results of probability theory to real practical problems where we have always to deal with a finite number of trials. (My brackets and italics). »

In fact however the law of large numbers is liable to even much more extended and strong criticisms than non-effectiveness alone. This leaps to one's eyes as soon as one distinguishes between the assertion of bare existence of a factual probability law, and the definition of the numerical contents of this law:

A. In the theorem of large numbers the existence, on the considered universe of events e_j produced by a random experiment Π , of a factual probability law $\{p(e_j)\}, j = 1, 2, \ldots, q$, is not proved, it is just posited. And the theorem of large numbers asserts strictly nothing concerning the significance to be assigned to this posited existence. This circumstance is not only an abstract conceptual gap. It entails a quite crucial pragmatic consequence: According to the law of large numbers, the initially unknown but researched definition of the individual numerical value $p(e_j)$ emerges via convergence toward the mathematical limit of the value of the countable finite corresponding relative frequency $n(e_j)/N$. But this convergence itself is based upon the postulation of the existence of the two probability laws $\{p(e_j)\}$ and $\mathcal{P}[(|n(e_j)/N - p(e_j)|) \leq \epsilon]$, without in any way specifying in what sort of physical features or circumstances these two existences consist. Such an approach obviously favours the absence of a deliberate organization of the conditions of convergence.

Now, nothing interdicts to admit that the factual relative frequencies $n(e_j)/N$ can play a role of ideal and probable specification-by-progressive-materialization of the mathematical limits $p(e_j)$. But this can be conceived *only* if the process of evolution of the relative frequencies $n(e_j)/N$ is materially constrained by *that* what the unknown limit $p(e_j)$ designates inside the material world. For otherwise **why** should a convergence realize?

And the same can be asked concerning the meta-probability $\mathcal{P}[(|n(e_j)/N - p(e_j)|) \leq \epsilon]$. In short: emergence of the knowledge of a numerical qualification ' $p(e_j)$ ' via some material effect of a non-specified material circumstance, upon the evolution of the relative frequencies $n(e_j)/N$, is conceivable inside the nowadays scientific representation of reality. But for a genuine understanding of such an effect and for its pragmatic domination, an *independent* factual definition of the assumption of existence of a probability law

In my own vocabulary the term "semantic" points toward significance with respect to some acting but unspecified net of references of factual nature — physical as well as psychical (aims) — and assigned spontaneously, in an implicit way, outside any formal logical or logical-mathematical system. However a semantic character of an assertion does not hinder a mathematical expression of this assertion, worked out by the help of mathematical concepts: I distinguish radically between the assertion itself and its various possible expressions.

and of the way in which this assumption translates into physical constraints that entail mathematical convergence of the relative frequencies $n(e_i)/N$, seems to be a necessity.

B. This focuses even more critical attention upon the structure of the frequential definition of the initially unknown factual probability $law\{p(e_j)\}$: This definition mixes factual elements in the proper sense — namely finite countable relative frequencies — with the mathematical concept of limit that is quite essentially non-factual: von Mises has ill-constructed the mathematical expression of his 'semantic' concept of probability. He has operated a deficient translation from factuality into mathematical terms, a forced, an overdone translation. The non-effectiveness refused by Kolmogorov is the consequence of this deficiency of translation. And let us add in this context that nothing whatsoever imposes the necessity of a mathematical method for constructing the semantic, factual numerical distribution of the individual events involved in a factual probabilistic situation: any sort of operational algorithm can do.

C. Finally the law of large numbers is flawed by also a peculiar sort of logical circularity. Via the dependence between the probabilities $p(e_j)$ and the meta-probabilities $\mathcal{P}[(|n(e_j)/N - p(e_j)|) \leq \epsilon]$ it introduces in the definition of a factual probability law $p(e_j)$ a logical regression that develops along a "vertical" hierarchy of levels of conceptualization: a sort of circularity spread out on a spiral.

Together, the reasons listed above lead to reject the law of large numbers as an acceptable semantic definition of the frequential concept of probability.

Before closing my answer to the question 2, let us consider a sub-question of 2 raised by this answer:

Sub-question 2. How can one explain the flaws pointed out above in the law of large numbers?

The theorem of large numbers has been first worked out by Jacob Bernoulli and later by Richard von Mises, and these were not idiots. Furthermore, during a long time it has been *accepted* by Kolmogorov, as well as, up to this very day, by practically all the mathematicians specialized in the theory of probability and by practically all the physicists who work with probabilities, even if the physicists who work in the domain of quantum mechanics have perceived the necessity of an extended concept of probability. So, what happened?

Answer to sub-question 2. The semantic concept of frequential probability began to acquire some inner structure only starting from the 17^{th} century, first via a work of Blaise Pascal (1654), but mainly by Jacob Bernoulli's well-known first version of the concept of a "law of large numbers" (1690, published in 1715). This was a mathematical expression deliberately worked out such as to apply quite specifically to games of chance, that is, to *artefacts* that function on two superposed levels and are simultaneously subject to two mutually compatible aims: On a basic level they must provide, for individual players, the possibility to gain or to lose possessions, so amusement charged with emotions; while on a global meta-level they must insure systematic gains for the owner of the game. For such artefacts the factual, numerical probability law was **known** by construction. So the question of existence of such a law and the question of its numerical content simply did not arise.

More than 200 years later (1931) Richard von Mises published Wahrscheinlichkeitsrechnung und ihre Anwendungen in der Statistik und theoretischen Physik.

During this long time the domain of applicability of the term 'probability' underwent a surreptitious extension. Namely, from the realm of fabricated games of chance, to the realm of, also, natural events. Von Mise's work, quite explicitly, referred to this whole enlarged domain of applicability inside which — in general now — the numerical form of the factual probability law is *unknown* and has to be identified. It seems that some contemporaries were clearly aware of this extension and of its implicit consequences. For instance, Alexandre Ostrowski has written:

 \ll His dynamical personality incited to forget his major deviations. He even obtained for giveness for his theory of probabilities. \gg

This suggests that Ostrovski still conceived that applicability of the concept of probability was restricted to *exclusively* games of chance. Indeed, since for a game of chance the probability law $\{p(e_j)\}$ is known by construction, in this case:

- * The necessity to postulate the 'existence' of this law disappears.
- * The problem of its significance vanishes, since obviously this significance consists of the deliberate realization of an artefact that generates precisely this law.
- * So a frequential definition of the law is no more necessary, and the argument of ineffectiveness vanishes.
- * And since the law $\{p(e_j)\}$ possesses by construction an independent materialized definition, any logical circularity vanishes also.

Whereas for a random phenomenon that introduces an a priori *unknown* probability $law\{p(e_i)\}$, all the exposed flaws do emerge.

This opposition seems to specify the problem with which one is faced. The answers to the questions 1 and 2 leave open the following challenge:

Identify:

- * The general specific criteria for asserting:
 - The "existence" of a factual probability law.
 - The **meaning** to be assigned to the assertion of "existence".
 - A definition of this factual law that shall be general, independent and effective.

The very interesting contribution of C. Porter in this volume conveys data that are still unknown to most, and in particular were unknown to myself. These show how Von Mises, and then Kolmogorov, each one in his own way, have tried to answer these questions via purely mathematical criteria of "place-invariance". But no one achieved his answer. And so long that this challenge is not met there is no general "theory of probabilities". There is only the probabilistic syntax of Kolmogorov, applicable to games of chance (so a "theory of games of chance"), and on the other hand, a vast factual domain of random phenomena in connection with all of which the term "probability" is made use of in the *absence* of a definite common semantic specified sufficiently for being connectable with Kolmogorov's syntax. In my contribution to this volume I have tried to meet this challenge. In so far that it will be judged that I succeeded that will have to be explained by:

- The systematic descriptional relativizations introduced as required by the "general method of relativized conceptualization".
- A constant sharp distinction between elements of semantic nature and elements of syntactic nature.
- The modification of the rule on connection between the semantic data (Π, U) and the corresponding Kolmogorov probability space $[U, \tau, p(\tau)]$: the universe of observable outcomes U is inserted in the algebra τ considered to be the total algebra defined on a universe U^c constructed out of the elements from U: instead of making use of Kolmogorov's connecting rule $[(\Pi, U) \rightsquigarrow [U, \tau, p(\tau)]$ (where " \rightsquigarrow " is just a sign to be read "connected with"), we have established reasons why one has to make use of the connecting rule $(\Pi, U) \rightsquigarrow [U^c, \tau \supset U, p(\tau)]$. Such a modification could never have been conceived in the absence of a clear distinction between semantic elements and syntactic elements, which in its turn would not have been achievable in the absence of systematic descriptional relativizations.

Hovever — quite remarkably — the result is essentially tied with the classical thinking alone: it is not *directly* applicable to the quantum mechanical descriptions of microstates.

Answer to question 3. I am not a biologist and I did not form a structured opinion concerning the questions 3. However, given the general present state of development of the concept of probability, I suppose that any application to biology is necessarily trapped inside at least the same difficulties as those that arise in physics.

Answer to the two questions 4. My answer is negative: Quantum mechanics represents phenomena tied with what we call states of microsystems (microstates) (Dirac). The representations have been realized via several different formal systems (syntaxes). The most expressive and current representation is realized by use of the mathematical syntax of Hilbert vectors. I restrict my argument to this syntax, but it can be transposed to the other ones also.

The elements of the syntax of Hilbert vectors are connectable with elements of the semantic concept of a microstate, via the current languages and a system of postulates that is equivalent to a set of coding rules. The quantum mechanical "probability postulate" of Born asserts that the probability of an observable quantum mechanical event — always an observable outcome of an act of 'measurement' performed on the studied microstate — can be coded inside the Hilbert syntax by an expression that is calculable from the projection of the Hilbert vector that represents the considered microstate, upon a syntactic element that codes for the result of the considered act of measurement. Now,

* A relation of which the nature is representational, "linguistic", just *a coding-rule*: "this formal element points (via usual language) toward that element from the semantic concept of microstate" cannot be proved. Quite essentially it is *a conventional choice*. The coding of the semantic features assigned to the concept of microstate, and the representations of these features inside the syntax of Hilbert vectors, connects two

universes of radically distinct natures, it is a **bridge**. This bridge cannot be entirely absorbed in only one of these two universes, namely in the syntax only. But proofs can be realized only with elements that are entirely contained inside a formalized system.

** The semantic definition that I propose for the concept of a factual probability law seems to be performable only inside the *classical* conceptualization. This, I think, entails that — rigorously speaking — a semantic concept of "quantum mechanical probability" *simply is not constructible*.

These are the reasons for which my answer to the first question 4 is negative.

Concerning the second question 4 my answer is equally negative: It is far from being possible to associate a definite Hilbert vector to any microstate that can be generated via definite physical operations (without mentioning all the that can be just conceived). And when no definite Hilbert vector representation is available, the observable relative frequency of this or that observable result of an act of measurement is simply is cut from the quantum mechanical formalism with its postulates and theorems (namely Gleason's theorem, that presupposes an independently achieved definition of a semantic concept of probability). In such cases — and these are the most current ones — the probability postulate cannot be connected in any strict sense with the observable data, not even in principle. One is confined to statistical investigations and to conclusions in statistical terms. Then the miraculous precision of the "essentially probabilistic" predictions of quantum mechanics recedes far away.

Karl SVOZIL

To me, the issue of "the unreasonable effectiveness of mathematics in the natural sciences", including probability theory, as exposed by Eugene P. Wigner in the 1959 Richard Courant Lecture delivered at New York University, remains as mind-boggling as ever.

A priori, I cannot see any reason why formal methods should be applicable to the phenomena in the Universe; less so to phenomena which are postulated and believed to occur at random.

Maybe the strongest position in this regard is the Pythagorean assertion that "the Universe is mathematics," or "God computes." Inside a system of beliefs of this sort, the mathematics we create and apply are an intrinsic, subjective, and necessarily reflexive and self-referential auto-view of this mathematical Universe.

Such a scenario might even allow for a dualistic setup in which we, as transcendental beings (with respect to the perceived Universe), are moving "freely" inside an algorithmically created Universe.

Also, as far as I am aware of, all claims of a "lawful" clockwork Universe, and also of a Universe (partly) governed by chance, are metaphysical conjectures. Because in the former case it is not possible to prove in any strict sense that this claim can be substantiated (i.e., tested or falsified) not only for the small class of physical behaviours we have access to, but in general, for all times and everywhere. And in the latter case, how could we operationally or formally obtain certainty and knowledge referring to an

Debate On The Concept Of Probability

"absence of all laws of behaviour", even if one restricts these laws to effectively computable ones? A plethora of recursion theoretic theorems state the contrary.

So, one is relegated to just *beliefs* that the Universe is fully or only partly lawful. I personally tend to think that Sigmund Freud's advice to his fellow psychoanalysts is the best advice one could give also to researchers in physics and biology: to adopt a contemplative strategy of (an evenly-suspended attention), always being aware of the dangers caused by "temptations to project into science as generally valid theories, what [the analyst] in dull self-perception recognizes as the peculiarities of his own personality."

Comments

MMS. YES. All the assertions you mention — to refuse them — are manifestations of what I, with my own vocabulary, call naïve realism, according to which man can know "reality" such as it "truly is", in some unconceivable absolute way. When in fact only human descriptions of fragments somehow separated from inside the "physical reality" and then qualified accordingly to the human grids of qualification, can be "known": knowledge involves in a non removable way qualifications, qualifications involve in a non removable way grids for qualification (of biological nature or artificial-and-biological nature) and these grids mark the resulting qualifications, the description, by non removable relativities to themselves. These hide the physical reality as much as they reveal it. So "knowledge of physical reality such as it is in itself" is just a self-contradicting concept; while the knowable human descriptions are artefacts, not reality. In this sense I feel in deep agreement with your denunciations of fallacy and with their daring formulations.

Nevertheless for any living being the descriptions of reality achieved via its sensorial apparatuses are essential, because without them it would not be able to attain its vital aims. And the descriptions achieved by men via their biological instruments and the prolongations of these by apparatuses and reasoning, have produced the evolution of humanity. This fact, that descriptions, knowledge produced by living beings, permit to attain the *aims* of the living beings, can indeed be felt to be a miracle, in this definite sense that, like you, I cannot see an explanation of another nature than just metaphysical postulates. But even these are difficult to formulate. Indeed if one posits some sort of correspondence or coherence between human knowledge and reality, this presupposes possibility of a comparison and, in the absence of any qualification assignable to one of the two compared entities, one hits again one's mind against a self-contradicting concept. I cannot find an explanation expressible in defendable rational terms. So one feels constrained to keep to just asking oneself and marvelling. As Wittgenstein wrote, « that which cannot be said must be left in silence » (I imagine the English version starting from the French one).

But this does not suppress the fact that it might be *very* useful to explicitly understand and *control* the processes of organization of human knowledge and the ways in which it emerges, acts, and changes. This, I believe — if subjected to an appropriate explicit method of construction of knowledge — does even offer a *methodological* answer to your striking question « how could we operationally or formally obtain certainty and knowledge referring to an "absence of all laws of behaviour", even if one restricts these laws to effectively computable ones? ».

2. GENERAL conclusion on this special issue of MSCS

On a basic level, this volume offers a view on various contexts and ways in which the concepts of probability and of randomness are made use of nowadays. And on a metalevel — via questions concerning more explicitly the particular case of the concept of probability — it also investigates how the specialists who make current use of this concept conceive the general relations between a domain of semantic data and a corresponding syntactic structure.

The first idea of this issue has emerged in the mind of the editor of MSCS while he was becoming aware that nobody seemed to be troubled, nor at least to notice, that there exists no logically acceptable and effective *general* method for specifying explicitly the factual numerical distribution of the probabilities of the individual events involved in a given particular situation that on the other hand is consensually considered to be probabilistic. That is why the questions proposed for the final debate are focused exclusively upon this definite point.

And when now the debate that has been realized is considered globally, it confirms the initial perception that originated it: From the majority of the reactions there emanates a serene and kind unison of unwillingness to assign particular importance to the questions proposed for discussion.

The fact is clear that indeed, after nearly 30 years since the author of the unique modern probabilistic syntax has publicly claimed that this syntax cannot be applied to factuality, nobody feels that this conceptual situation involves a problem that should be considered and solved.

We have thought that this fact deserves being exposed to contemplation and reflection. Not only for its specific relation with the concept of probability, but also as an illustration of the somewhat obscure and trembling, diverse and sometimes outraged intimate views by which the supra-individual abstract whole called scientific thinking draws substance and a distilled stable form from the individual minds, wherefrom gush the favoured directions of research while other directions vanish into absence of focus, like a jet of water absorbed in sand. By this illustration the present special issue of MSCS can also be regarded as an investigating incursion into the evolution of scientific thinking.

Mioara Mugur-Schächter Guest editor of this special issue of MSCS

Giuseppe Longo