Eötvös-HPS 01-03

A Physicalist Interpretation of Probability[*†]

László E. Szabó
Theoretical Physics Research Group of the Hungarian Academy of Sciences
Department of History and Philosophy of Science
Eötvös University, Budapest

Abstract

There is no such property of an event as its "probability." Rather, I argue that probability is a reducible concept, supervening on physical quantities characterizing the state of affairs corresponding to the event in question. The term "probability" can be used only collectively: it means different dimensionless $[0,1]$-valued physical quantities (measures) in the different particular situations. I also argue that "probability" is not the limiting value of relative frequency, and not even necessarily related to the notion of frequency. In some cases, the conditions of the sequential repetitions of a particular situation are such, however, that the "probability" (the corresponding physical quantity) is approximately equal to the relative frequency of the event in question. Sometimes we do not know the value of the physical quantity X, identified with the "probability" of an event A. In this case, if we are convinced about the relationship between X and the relative frequency of A, we can measure X by counting the relative frequency of A. Furthermore, I will argue that "probability" has nothing to do with indeterminism and, on the other hand, has nothing to do with "lack of knowledge."

Introduction

This paper develops a new interpretation of probability, which sidesteps the usual difficulties intrinsic to the various standard interpretations and, on the other hand, in some sense incorporates much of the intuition behind them. I call it *Physicalist Interpretation*. The term "physicalist" is borrowed from the philosophy of mind. According to the physicalist account of mind, the mental

[*]Talk presented on the International Interdisciplinary Workshop on Determinism, Ringberg Castle, Germany, June 4 - 8, 2001.
[†]The research was partly supported by the OTKA Foundation, No. T025841, T032771.

is completely describable in physical terms. In other words, the mental supervenes[1] on the physical. Similarly, the main idea of the physicalist interpretation of probability is that there is no such property of an event as its "probability." Rather, I argue that probability is a reducible concept, *supervening on physical quantities* characterizing the state of affairs corresponding to the event in question. More radically, I claim that probability is a concept which can be completely eliminated from the scientific discourse. The term "probability" can be used only collectively: it means different dimensionless $[0, 1]$-valued physical quantities, more precisely, different dimensionless normalized measures composed by different physical quantities in the different particular situations.

According to this approach, the physical quantity identified with "probability" is not the limiting value of relative frequency, and not even necessarily related to the notion of frequency. Although, in some cases, the conditions of the sequential repetitions of a particular situation are such, that the "probability" (the corresponding physical quantity) is approximately equal to the relative frequency of the event in question. Neither is "probability" a *new* objective property of a system, expressing its propensity to behave in a certain way, although, the physical quantities characterizing the system are definitely capable to describe such a propensity. In this approach, "probability" is not the measure of the degree of belief of an agent in one proposition or another. However, according to the physicalist account of mind, one can imagine a collection of physical quantities characterizing the agent's brain, which compose a dimensionless measure playing the same role in a typical betting scenario as the "subjective probability."

Difficulties of the standard interpretations of probability

We need to make a distinction between *mathematical* and *realistic* interpretations of probability. Mathematical interpretation means that the formal mathematical structure PROBABILITY THEORY is represented in ANOTHER FORMAL MATHEMATICAL STRUCTURE. Discussing the Bertrand-paradox,[2] for example, we consider a representation of the probability-theoretic notions in geometrical terms. (The alleged "paradox" consists in the simple fact that we have a kind of freedom in constructing such a representation.) It is a mathematical representation when probabilities are represented by the limiting values of convergent relative-frequency-like infinite sequences, or when the ("subjective") probabilities are represented in game-theoretic terms.

The mathematical interpretations do not raise difficulties at all. Our concern is, however, not a mathematical but a realistic interpretation. A realistic interpretation is nothing but the way in which we apply probability theory to

[1] I mean supervenience in the sense of *local* supervenience in Chalmer's terminology (D. J. Chalmers, *The Conscious Mind*, Oxford University Press, 1996).

[2] M. Kac and S. M. Ulam, *Mathematics and Logic*, Dover Publications, NY, 1968.

the real world. The similarity between the mathematical and realistic interpretations is that in both cases we construct a representation of PROBABILITY THEORY in ANOTHER LANGUAGE. But, in case of a realistic interpretation this other language *must be, in final analysis, translatable into empirical terms*.

In this section I would like to briefly review the standard interpretations,[3] and to illustrate that none of them is tenable, because none of them provides a sound definition of what probability is.

Classical Interpretation

The classical interpretation goes back to Laplace. According to his definition the probability of an outcome is the ratio of favorable cases to the number of equally possible cases. Two outcomes are meant "equally possible" if we have no reason to prefer one to the other (principle of indifference). Consider the following often quoted example: A symmetric die has six faces numbered 1-6. When it is tossed in the standard way there are six possible outcomes. The probability of getting an <even number> is $\frac{3}{6}$, for three of the possible outcomes (2, 4, 6) are favorable.

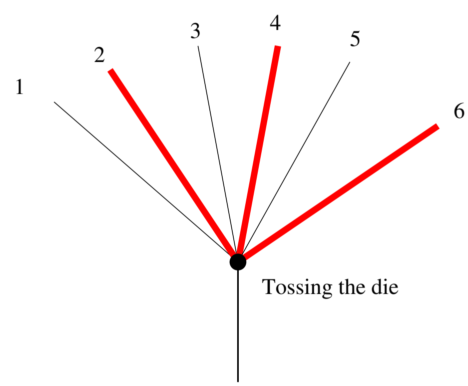

Figure 1: *In the moment of tossing the die the history of the universe is branching into six branches. Three of them are favorable*

Since we are talking about a realistic interpretation, we must translate the probability-theoretic notions into an observational language. But how can we perform such a translation into empirical terms? In no way, it seems. For either we obtain a concept of probability which is brutally indifferent to the facts of the world, or we must tacitly refer to another, for instance the frequency interpretation of probability. In the case of the above example of the symmetric die, the essential fact is that in the moment of tossing the die the history of the universe is (objectively or epistemically, all the same) branching into six branches (Fig. 1). More precisely, the possible histories can be sorted into six

[3] J. Earman and W. Salmon, The Confirmation of Scientific Hypotheses, In M. H. Salmon, *et al.*, *Introduction to Philosophy of Science*, Prentice Hall, Englewood Cliffs, New Jersey, 1992.

classes corresponding to the outcomes 1-6. There are three branches that correspond to the event <even number>. Therefore, $p\left(< \text{event number} >\right) = \frac{3}{6}$. All other facts of the world are negligible. For instance, if the die were biased, the probability of getting an even number would be the same. 'But a biased die is not symmetric anymore, – we could argue – so one cannot apply the principle of indifference in the same way!' This argument is, however, problematic, because it appeals to two different conceptions of probability. Of course, the biased die is not symmetric in respect of some properties. The mass-density, for example, is not symmetric. But even a standard die is not completely symmetric in all properties. Otherwise we were not able to differentiate the six outcomes. For instance, different numbers are written on the different faces of the die. So, it seems, we must conclude, that there are relevant and irrelevant asymmetries: only those asymmetries are relevant, which can influence the probabilities of the six outcomes. 'But what kind of "probability" do we mean here?', we should ask ourselves. And the only possible answer could be something like this: It is an observable fact that the biased die produces one outcome *more often* than the other. That is, we should refer to the frequency interpretation of probability.

Relative Frequency Interpretation

The frequency interpretation is based on the following idea: Probability of an outcome is not a concept which could be assigned to an individual experiment, but rather it is assigned to a long sequence of repeated experiments. Denote $\mathcal{A}$ the Boolean algebra of the outcome events, and let $p : \mathcal{A} \to [0,1]$ be the probability to be interpreted in empirical terms. Performing an experiment, each of the possible outcome events does or does not occur, and this fact can be described by a suitable "outcome function" (classical two-valued truth function, if you want) $u : \mathcal{A} \to \{0,1\}$ satisfying the following conditions:

$$\begin{aligned} u\left(\emptyset\right) &= 0 \\ u\left(\neg A\right) &= 1 - u\left(A\right) \\ u\left(A \wedge B\right) &= u\left(A\right) u\left(B\right) \end{aligned}$$

where the outcome function has value 0 if the corresponding event does not occur, and has value 1 if does.

Consider now the sequential repetitions of an experiment. Let

$$u_1, u_2, \ldots, u_N, \ldots \tag{1}$$

be the sequence of the outcome functions we obtain. For each N we define the relative frequency function as follows:

$$\nu_N : A \in \mathcal{A} \mapsto \nu_N\left(A\right) = \frac{1}{N} \sum_{i=1}^{N} u_i\left(A\right) \in [0,1] \tag{2}$$

One can easily prove that if the sequence

$$\nu_1, \nu_2, \ldots, \nu_N, \ldots \tag{3}$$

is pointwise convergent, then the function $p = \lim_{N\to\infty} \nu_N$ is a probability function on $\mathcal{A}$, satisfying the Kolmogorov axioms.[4]

Here we encounter the first difficulty of the Frequency Interpretation: the limit of an infinite sequence is independent of the first N elements, where N can be an arbitrarily large number. In other words, the observed relative frequency in any finite sample is *irrelevant* to the probability. How, then, we are supposed to find out what these probabilities are? Or if we have a hypotheses about the value of a probability, how can we empirically confirm this hypotheses, if there is no logical relationship between what we observe in a finite sample and the value we would like to confirm?

The second difficulty is that (1) is the sequence of the *real* outcomes of the consecutive experiments, therefore, there is no guarantee that the sequence (3) is convergent. Note, that the randomness of the outcomes does not guarantee the convergence. To illustrate this, consider the following simple example: We flip two coins. If the result is <Heads> & <Heads> then the outcome of the experiment is <1>, otherwise the outcome is <0>. Repeat the experiment until the relative frequency of <1> is less than 0.4. Then we change the roles of <0> and <1>, and repeat the experiment until the relative frequency of <1> becomes larger than 0.6. Then we change again, and so on. The sequence of outcomes obtained through this method is completely random, but there is no limiting value of the relative frequencies. So, if we insist that probability is nothing but limiting relative frequency, we must conclude that probabilities $p(<0>)$ and $p(<1>)$ *do not exist*.

This conclusion is, however, counter-intuitive, because we "know" that in each run of the experiment probabilities $p(<0>)$ and $p(<1>)$ do exist, for instance $p(<1>) = 0.25$ for a while, then it changes for 0.75, then changes for 0.25 again, and so on.

The above example also throws light on the third problem of the Frequency Interpretation. Namely, that it does not account for the probability of the outcome of an individual experiment, which probability, on the other hand, is a meaningful concept in our intuition.

Propensity Interpretation

Karl Popper's Propensity Interpretation aimed to solve the problem how to assign probability to the outcome of an individual experiment. While the Classical and the Frequency Interpretations try to reduce the notion of probability to other, already known concepts, the Propensity Interpretation identifies probability with a new quantity, called "propensity", expressing the measure of the "probabilistic causal tendency" of the system to behave in a certain way.

[4]This is actually a trivial consequence of the Pitowsky theorem (I. Pitowsky, *Quantum Probability – Quantum Logic*, Lecture Notes in Physics **321** , Springer, Berlin 1998).

The standard objection against propensity says that it does not provide an admissible interpretation of probability. Consider two events A and B which are causally related, therefore $p(B|A) > p(B)$. $p(B|A)$ can be interpreted as propensity, for we can speak meaningfully of the tendency of a cause to produce the effect. Consider now $p(A|B)$ which is also a meaningful concept, in the sense that it can be easily expressed by the Bayes rule: $p(A|B) = \frac{p(A \wedge B)}{p(B)}$. However, $p(A|B)$ cannot be interpreted as propensity, the argument goes, because it does not make sense to talk about the causal tendency of the effect to have been produced by one cause or another.

This usual objection is, however, not acceptable, in my view, because it is based on a misinterpretation of conditional probability, in general. Conditional $p(B|A)$ is nothing but the ratio $\frac{p(A \wedge B)}{p(A)}$ and, in general, it has nothing to do with "the tendency of a cause A to produce the effect B." As in any other interpretation of probability, the correlation $p(A|B) > p(A)$ is only a necessary but not a sufficient condition for a causal relationship. The origin of the misunderstandings is that conditional probability is often saddled with completely unjustified meaning: it does not mean, for example, the value for which the probability of event B *changes* when event A happens, or it is not equal to the probability of event B if the system is *prepared such* that event A occurs with probability 1, etc. In case of dicing, for example, the conditional probability $p\left(<4> \mid <\text{even}>\right)$ is nothing but the rate $\frac{p(<4>)}{p(<\text{even}>)} = \frac{1}{3}$. But, it does not mean that $p\left(<4>\right) = \frac{1}{3}$ if the die is prepared such that $p\left(<\text{even}>\right) = 1$. In this case, the conditional probability $p\left(<4> \mid <\text{even}>\right)$ would not be a well-defined notion, because the "$p\left(<\text{even}>\right) = 1$" preparation does not correspond to a unique condition: if the die is biased in such a way that $p\left(<2>\right) = 1$, then $p\left(<\text{even}>\right) = 1$, and $p\left(<4>\right) = 0$. While in another case, if it is biased such that $p\left(<4>\right) = 1$, then, again, $p\left(<\text{even}>\right) = 1$, but $p\left(<4>\right) = 1$.

There is, on the contrary, a more difficult problem with propensity. In Propensity Interpretation, probability – propensity – is a separate quantity, which is not expressed in terms of other, empirically defined quantities. How, then, is the numerical value of propensity determined? We have no starting point for the empirical test of the value of propensity. Consequently, there is no empirical basis for such a proposition as "the probability of getting <Heads> is $\frac{1}{2}$", and the whole talk about probabilities loses empirical control. We do not even know whether propensities satisfy Kolmogorov axioms, or not.

Subjective Interpretation

Finally we must briefly mention the Subjective Interpretation. It identifies probability with a person's degree of conviction or belief in one proposition or another.[5] Surprisingly, Subjective Interpretation is, in my view, a realistic

[5] I use the term "subjective probability" in the usual text-book sense, as a degree of belief. And this concept is different from Olimpia Lombardi's "degree of knowledge" – "probability due to

one, since it aims to apply probability theory to the real world. For, when we say that "Mr. Smith has subjective probability 0.9 for that the lovely bay mare Willow will come in first", we are talking about a *real* person's degree of belief in a *real* event, a real winning of a real mare, shown by the photo-finish.

Beyond the obvious problem that subjective probability could be applied only for a very restricted part of reality, Subjective Interpretation suffers from the same difficulty as the propensity approach: It claims the existence of a separate quantity, the degree of belief of a person, which is not yet empirically defined. How, then, is the numerical value of the degree of belief determined?

Physicalist Interpretation of Probability

As we have seen, although each of the standard interpretations can grasp something from our intuition about probability, *none of them can provide an ultimate explanation, in empirical terms, of what probability is.* How is it possible, on the other hand, that physics and other empirical sciences can apply the formal theory of probability, without perceiving anything from this unanswered fundamental question? In the second part of the paper I shall make an attempt to develop a new interpretation of probability, which perhaps can resolve this contradiction.

The key idea of my proposal is that probability is a concept which can be completely eliminated from the scientific discourse. This fact explains why the standard interpretations are unable to give a sound definition of probability, and also explains why empirical sciences can manage without such a definition.

Thesis 1 *There is no such property of an event as its "probability." What we call probability is always a physical quantity characterizing the state of affairs corresponding to the event in question.*

Consider the following example: A gun is hinged in such a way that it can shoot uniformly into a square of size $a \times a$ on the wall (Fig. 2). Inside, there is a round target of radius R and an air-balloon of radius r, in front of the target. What is the probability that the balloon bursts out (event A)? What is the probability that the shot hits the target (event B)? And what is the conditional probability of that the balloon bursts out, given that the bullet hits the target?

The physicist's standard answer to these questions is the following:

$$\begin{aligned} p(A) &= \frac{\pi r^2}{a^2} \\ p(B) &= \frac{\pi R^2}{a^2} \\ p(A|B) &= \frac{r^2}{R^2} \end{aligned}$$

ignorance," which is rather identical with Professor Primas' "epistemic probability."

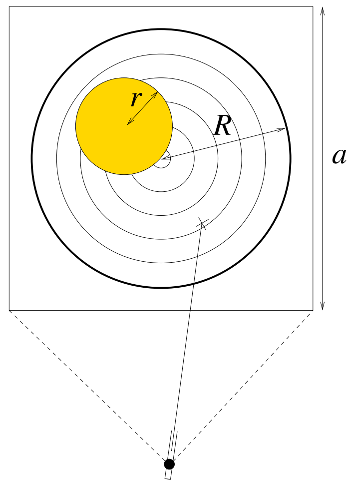

Figure 2: *A gun is hinged in such a way that it can shoot uniformly into a square of size $a \times a$ on the wall. Inside, there is a round target of radius R and an air-balloon of radius r, in front of the target. What is the probability that the balloon will burst out? What is the probability that the bullet hits the target? And what is the conditional probability of that the balloon bursts out, given that the bullet hits the target?*

Let it remain in obscurity how the physicist arrives at these results. What is important is the fact that "probability" is expressed in terms of known, well-defined physical quantities, composing a dimensionless normalized measure on a space, the measurable subsets of which represent the outcome events in question. My suggestion is to give up the independent concept of "probability" with an overall context-independent meaning. In my view, this is what we must learn from the failure of the standard interpretations. The reason why none of these standard approaches can provide a sound meaning for the term "probability" is that there is no such a property of an event as its "probability." That is, when we say that $p(A) = \frac{\pi r^2}{a^2}$, we do not mean that there is a known, well-defined quantity, $p(A)$, on the left hand side, which is, contingently, equal to $\frac{\pi r^2}{a^2}$. We just mean, that the measure $\mu(...) = \frac{\text{area of ...}}{a^2}$ satisfies the Kolmogorov axioms and shows many other features we usually assign, intuitively, to probability.

In case of a completely *different scenario*, "probability" is identified with a dimensionless normalized measure composed by completely *different physical quantities*. So, the best what we can say about probability is the following:

Thesis 2 *The term "probability" can be used only collectively: it means different dimensionless $[0,1]$-valued physical quantities, more precisely, different dimensionless normalized measures composed by different physical quantities in the different particular situations.*

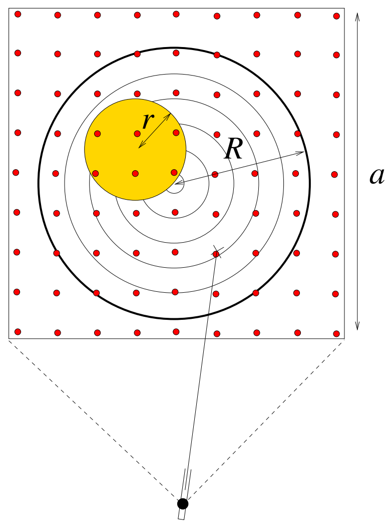

Figure 3: *If the size of the balloon is constant and the uniform distribution of the shots on the square is provided then the relative frequency of event A is approximately equal to $\frac{\pi r^2}{a^2}$.*

From the point of view of the everyday practice of sciences, the most important question is how probability is related to relative frequency. According to the above two Theses, it cannot be claimed, in general, that probability is equal to the limiting relative frequency, first of all because we do not know what probability is, in general. In the above example, we used the term "probability" for the quantity $\frac{\pi r^2}{a^2}$. In general, it has nothing to do with the relative frequency of event A. The value of $\frac{\pi r^2}{a^2}$ – although it is a well-defined number in each individual experiment, so, in this sense, "probability" is a meaningful notion for an individual event – can change during the sequential repetitions of the experiment (we can change the size of the balloon, for example), therefore there is no guarantee that the sequence of relative frequencies will converge to a limiting value. But, in particular cases, if $\frac{\pi r^2}{a^2}$ is constant and the uniform distribution of the shots on the square is provided (Fig. 3), the relative frequency of event A is approximately, $o\left(\frac{1}{N}\right)$, equal to "probability" $\frac{\pi r^2}{a^2}$. (And this is not a probability-theoretic result but it is an elementary fact of kinematics.) In general,

Thesis 3 *The physical quantity identified with "probability" is not the limiting value of relative frequency, and not even necessarily related to the notion of frequency. In some cases, the conditions of the sequential repetitions of a particular situation are such, however, that the probability (the corresponding physical quantity) is approximately equal to the relative frequency of the event in question.*

Assume now, that we do not know the size of the balloon, therefore, we do not know the value of $\frac{\pi r^2}{a^2}$, i.e., the value of $p(A)$, but we know that it is constant, and we can also guarantee the uniform distribution of the shots. In this case, we can measure, $o\left(\frac{1}{N}\right)$, the "probability" $p(A)$, that is, $\frac{\pi r^2}{a^2}$, by measuring the relative frequency of A. That is,

Thesis 4 *Sometimes we do not know the value of the physical quantity X, corresponding to the "probability" of an event A. In this case, if we are convinced about the relationship between X and the relative frequency of A, we can measure X by counting the relative frequency of A.*

The physical quantity $\frac{\pi r^2}{a^2}$ exists and has a well-defined value, independently whether the laws of nature governing the shooting and the motion of the bullets are deterministic or not. Moreover, the relationship between $\frac{\pi r^2}{a^2}$ and the relative frequency of A (if there is such a relationship at all) is not influenced by the deterministic or indeterministic character of the physical process in question. The relative frequency can be equal to $\frac{\pi r^2}{a^2}$ even if the uniform distribution of the shots are provided through a deterministic ergodic process, by the random number generator of a computer, for instance.

Similarly, nothing can influence the value of $\frac{\pi r^2}{a^2}$, which would be related to our knowledge about the details of the process. Similarly, if the condition of the uniform distribution of shots is satisfied, this value will be approximately equal to the relative frequency of A, independently of whether we know the direction of the subsequent shot, or not.

Finally, we have to emphasize that it is a matter of fact, whether the distribution of the shots is uniform or not. A priori we must not suppose that it is uniform, only because we have no information about how the directions of the consecutive shots are determined, and, on this basis, we have no reason to prefer one direction to the other.

So, our last three Theses are the following:

Thesis 5 *The value of the physical quantity identified with "probability" is not influenced by the fact whether the process in question is indeterministic or not. And a priori there is no reason to suppose that this value can be only* 0 *or* 1*, only because the process is deterministic.*

Thesis 6 *The value of the physical quantity identified with "probability" is not influenced by the extent of our knowledge about the details of the process.*

Thesis 7 *Neither the value of the physical quantity identified with "probability," nor the existence of the conditions under which this value and the relative frequency of the corresponding event are approximately equal can be knowable a priori.*

Although the standard interpretations do not provide a coherent definition in empirical terms, they grasp many important aspects of our intuition of probability. The physical quantity like $\frac{\pi r^2}{a^2}$, in our example, seems to fit very well to these intuitive descriptions of probability: 1) In some sense it reflects the ratio of favorable cases to the number of equally possible cases. 2) Under suitable circumstances it is approximately equal to the relative frequency measured during the sequential repetitions of the experiment. 3) It is meaningful and has a definite value in each individual experiment. 4) In the example we investigated, the rate $\frac{\pi r^2}{a^2}$ expresses indeed the measure of the "tendency" of the whole system to behave in such a way that the balloon will burst out.

Of course, in the above context we could not deal with subjective probability. According to the physicalist account of mind, however, one can imagine a collection of physical quantities characterizing an agent's brain, which compose a dimensionless measure playing the same role in a typical betting scenario as the "degree of belief."

So, our physicalist account of probability grasps a big part of the intuition behind the standard approaches.