File opened October 5, 2026
Version: November 2, 2026

Objective Probability as an Implicitly Defined Theoretical Term

John D. Norton[1]
Department of History and Philosophy of Science
University of Pittsburgh

Attempts to formulate an explicit definition of objective probability necessarily fail through circularities, since there are no more basic terms to use in the definition. Objective probabilities can be defined implicitly, as we do for other basic terms in our sciences, by means of the axioms of probability theory, where the truth of the axioms for the system in question is to be affirmed empirically. Much of the familiar understanding of objective probabilities can be relocated into an interpretive overlay whose content is not required to conform strictly with the axioms. This two-part characterization of probability is found in Kolmogorov's celebrated 1933 treatise on the foundations of probability.

1. Introduction: the Problem of Defining Objective Probability

Chances or objective probabilities arise commonly in physical theories. They are objective in the sense of expressing objective facts about some physical system, independently of our thoughts and beliefs. The most straightforward case is the Born rule in quantum mechanics. It assigns probabilities to outcomes of measurements on quantum system. In classical deterministic theories, probabilities still arise as effective theories. The outcomes of trials with physical randomizers such as die casts and coin tosses are governed by a robust and long-standing effective theory. In it, the symmetries of the chance set up lead us to assign equal probabilities to each outcome, even when we may imagine that the fullest but inaccessible theoretical treatment is classically deterministic. Probabilities arise as basic terms in a similar

[1] I thank both Caterina Sisti, whose work on the history of probability led to this paper, and Gal Ben Porath for helpful discussion.

way in the statistical physics that governs thermal systems, even when allow that a deterministic classical phase space ultimately governs the thermal processes.

The challenge is to provide an account of the chances or objective probabilities in such cases. Efforts to provide an explicit definition for them in non-probabilistic terms have failed. The best known and, in my view, the most promising of such attempts was advanced by von Mises (1957) and Reichenbach (1949). They sought to define objective probabilities in terms of the relative frequency of successes in repeated, independent trivial. This general approach has been subjected to sustained critical assault, such as in Hájek (1996, 2009). The most promising strategy among these attempts is to use some version of a law of large numbers in the probability calculus. This law *almost* enables us to connect probabilities with non-probabilistic relative frequencies of outcomes. But only *almost,* since it only assures us that a specified connection with frequencies obtains *probably*, even if with arbitrary high probability. The result is that the attempt to provide an explicit definition collapses into a circularity.[2]

The temptation is to define chances through subjective probabilities or credences in a Bayesian analysis. The difficulty is that the chances sought are not subjective. No amount of conditionalization can eliminate the subjectivity. For any body of evidence, there is always a sufficiently perverse prior probability that outweighs the trace of objective chance introduced by conditionalizing on the evidence. The approach can enjoy some success in the sense that the credences of many agents may converge, if their priors are not too disparate and they all conditionalize repeatedly on the same rich body of evidence. There is an essential condition that still must be provided by analysis external to the calculation: the body of evidence must be the right one that captures the objective character of the chances. Without it, the fact of convergence is misleading. It is not a recovery of an objective chance.

The solution, it will be argued here, is to abandon attempts to find an explicit definition for objective chances. As a general matter, the demand for explicit definitions can be met for

---

[2] Popper (1959) sought to escape the problems of the relative frequency approach by defining objective probabilities in terms of the propensity of a chance set up to produce sequences of outcomes. To succeed, the approach must give an account of these sequences such that the relative frequencies match the probabilities. The problems of the original relative frequency approach are return.

derivative terms in a theory, whose substance can be specified by reference to other more basic terms. However, these basic terms are basic because they cannot be defined through other terms without circularity. They are definitional termini. There is nothing more basic to which we can resort. Yet we do have an adequate grasp of the terms in so far as we can employ them reliably in our scientific discourses. Such terms are defined implicitly, that is, by their roles in the basic suppositions of the theory.

We should treat objective chances similarly and define them implicitly. The propositions forming the implicit definitions are facts about the system, independent of our thoughts and beliefs, and must be affirmed empirically as are all such facts about the system. The need for empirical affirmation of the axioms is distinctive of the case of objective probabilities. The same approach cannot be used to define subjective probabilities, where they are understood to be *rational* credences, for the axioms must then be affirmed as somehow expressing norms of rationality.[3]

An axiom system for probability already provides a serviceable implicit definition for objective probabilities. In so far as these implicit definitions conform with the strict mathematical properties of an additive measure, they lack essential elements of our practical understanding of probabilities. These additional elements can be supplied in an interpretive overlay that is not constrained by the formal conditions of the axiom system. In it, we may include our expectation that an outcome of very high probability will happen, even though the formal analysis allows for a small probability of its failure.[4] Efforts to include these sorts of interpretive aspects of objective probability into an explicit definition is partly responsible for the continuing failure of these efforts.

---

[3] The exception arises if we treat subjective probabilities within the confines of the material theory of induction, as described in Norton (2021, Ch.10). For then the rational norms are based on background facts. One result of the material theory, however, is that such rational norms vary from context to context according to the background facts and will only sometimes be expressible by an additive probability measure.

[4] The looser character of this interpretive overlay conforms with Ben Porath's (2024, Ch. 2) insistence that objective probabilities have an essentially subjective aspect.

The proposal is for a two-part characterization of objective probabilities. An axiom system provides a precise, implicit definition of the central theoretical term, probability; and an interpretive overlay enables us to connect this theoretical term with our practice. This two-part characterization is not new. As will be recounted in Section 2, it emerged in the early part of the twentieth century as a favored approach and was given a most compact expression in Kolmogorov's (1933) treatise on probability theory. The approach faded from the literature and is now all but lost.

This note presents an elementary development of the approach and seeks to articulate more clearly the rationale for it. My hope is that it will help revive the visibility of the approach in the more recent literature. To this end, Section 3 below will recount how implicit definitions work in several example. Each, we shall see, is accompanied by an interpretive overlay that is not part of the implicit definition. The overlays serve as an aid for us is forming serviceable mental pictures of the term and its use. They can do so quite effectively since they are not part of the formal theory and thus not constrained to conform with its strictures.

Section 4 will display an implicit definition for a simple case of the objective probability of coin tosses. It is sufficient to establish how an implicit definition for objective probability can proceed. More complicated examples would be more elaborate but not introduce any further foundational ideas.

The objectivity of the objective probabilities, it is recalled in Section 5, arises because the propositions comprising the axioms are facts about the physical system at issue. They must be affirmed empirically as must all other such facts. The section reviews how such affirmation arise in several familiar cases from quantum mechanics and statistical physics.

Section 6 will review how a precise link is made between the objective probability of success in tossing heads and the frequency of heads in many independent coin tosses. The result is a version of the weak law of large numbers, as it applies to this case.

Section 7 will review two versions of an interpretive overlay that lies outside the formal analysis but contributes to our understanding of objective probabilities. In one, employed by Kolmogorov, Cournot's principle directs us to interpret outcomes of very low probability as ones that will not occur and, consequently, that those of very high probability as ones that will occur. The principle does not conform with the formal theory since it accords a practical certainty to

high probability outcomes, whereas the formal theory still allows for unlikely failures. The principle is admissible in so far as it is located in the interpretive overlay.

In the second, Lewis' "principal principle" or the material theory of induction authorize us to set subjective probabilistic credences equal numerically to the objective probabilities of the formal analysis. The authorization is limited to the likelihoods employed in Bayes' theorem, not for the probabilities of the full Bayesian analysis. The latter requires in addition a probability space that assigns probabilities to the different possible values of the objective probability of outcomes of a single coin toss.

2. Kolmogorov's *Grundbegriffe*

Andrey Kolmogorov's (1933) *Grundbegriffe der Wahrscheinlichkeitsrechnung*, translated as *Foundations of the Theory of Probability* (1956), is now routinely cited as the definitive axiomatic development of probability theory that employs the methods of mathematical measure theory. There is, however, much more in the work of foundational and philosophical importance.

The definitive study of these further aspects and of the historical background to Kolmogorov's *Grundbegriffe* is Glenn Shafer and Vladimir Vovk (2006) expansive study. They document how Kolmogorov did not himself conceive of his work as novel scholarship. Rather his work was a compact synthesis of new work in measure theory and probability theory from a few preceding decades. Thus, Kolmogorov acknowledged in his Preface (1933, p.III) a work on measure theory by Fréchet still in preparation. Shafer and Vovk also recount in some detail how Kolmogorov conceived his treatment as giving the two-part characterization of probability to be recalled in this note.[5] The first part is his celebrated axiom system. For the second part, Kolmogorov employed what is also known as "Cournot's principle." It asserts that practical import accrues to probability, as implicitly defined by his axioms, if we treat events of very low probability as ones that we are certain in practice will not occur.

[5] Shafer and Vovk (2006, p.2) quote Kolmogorov's explicit recognition of the need for a second part in his remark to Fréchet in a letter of 1939 "You are also right in attributing to me the opinion that the formal axiomatization should be accompanied by an analysis of its real meaning."

A casual reader of Kolmogorov's rich and dense treatise may easily overlook these foundational elements and the clarity of Kolomogorov's understanding of them. His remarks are brief, as are the presentations of the main results in his treatise. A casual reader may jump from the first sentence of Kolmogorov's Preface "The purpose of this monograph is to give an axiomatic foundation for the theory of probability." to the statement of the axioms. Such a reader would then miss Kolomogorov's explanation of how he conceived his project.

To begin, he was quite aware that his account did not offer an explicit definition of probability. In the introductory paragraphs of Chapter 1, Kolmogorov (1956) described how his approach conforms with Hilbert's celebrated conception of axiomatic systems, such as implemented in Hilbert's (1903) *Foundations of Geometry*: (p. 1, Kolomogorov's emphasis)

> In accordance with the above, in §1 the concept of a *field of probability* is defined as a system of sets which satisfies certain conditions. What the elements of this set represent is of no importance in the purely mathematical development of the theory of probability (cf. the introduction of basic geometric concepts in the *Foundations of Geometry* by Hilbert, or the definitions of groups, rings and fields in abstract algebra).
>
> Every axiomatic(abstract) theory admits, as is well known, of an unlimited number of concrete interpretations besides those from which it was derived. …

Kolmogorov continued to note that other foundations for probability are possible. He contrasted his approach specifically with the *explicit* definitions found in earlier work, such as the relative frequency approach of von Mises: (1956, p. 2)

> There are other postulational systems of the theory of probability, particularly those in which the concept of probability is not treated as one of the basic concepts, but is itself expressed by means of other concepts. [Footnote[6]] However, in that case, the aim is different, namely, to tie up as closely as possible the mathematical theory with the empirical development of the theory of probability.

---

[6] Here Kolmogorov includes a brief footnote reference to two works of Richard von Mises and one of S. Bernstein. They include a 1931 von Mises work with the title *Wahrscheinlichkeitsrechnung*, which is likely von Mises (1931).

In allowing that implicit definitions admit "an unlimited number of concrete interpretations," Kolmogorov was conforming with Hilbert's conception of axiomatic systems. Hilbert's attitude to the core terms in his axiom systems was, famously, captured by a report of a remark made Hilbert in discussion. Blumenthal, (1935, pp. 402-403) in his biographical chapter on Hilbert, attributes to him the remark on his axiomatization of geometry:[7]

> "At any time, one must be able to say 'tables, chairs, beer mugs' in place of 'points, lines, planes.'

It follows that something more is needed, beyond the axiom system, if the terms in Kolmogorov's axioms are to have practical import. To this end, the section §2 following the statement of Kolmogorov's axioms §1 is titled: (1933, pp.2-3)[8]

The Relation to the World of Experience

Das Verhältnis zur Ehrfahrungswelt

Paired footnotes to the axioms in §1 and this new section §2 make clear that the new section is not part the mathematics but pertains to questions of meaning. The footnote to the heading "§1. Axioms" reads: (1956, p. 2)

> The reader who wishes from the outset to give a concrete meaning to the following axioms, is referred to §2.

The corresponding footnote to the heading of §2 begins: (1956, p.3)

> The reader who is interested in the purely mathematical development of the theory only, need not read this section, since the work following it is based only upon the axioms in §1 and makes no use of the present discussion.

Kolmogorov's treatise is still cited today as the standard source for a measure theoretic axiom system for probability. With the notable exception of the work of Shafer and Vovk, Kolmogorov's two-part characterization of probability seems to have all but disappeared from

---

[7] "Man muß jederzeit an Stelle von 'Punkte, Geraden, Ebenen' 'Tische, Stühle, Bierseidel' sagen können."

[8] This heading is *mistranslated* in the (1956, p.3) English edition as "The Relation to Experimental Data."

the literature.[9] Even just using an implicit definition for objective probability seems so natural that I fully expect others to have explored it in more recent literature. However, such explorations are elusive enough that I can say with confidence that they would have a very small presence in the literature. A welcome if rare instance is provided by Elliott Sober's (2010, 2024) reflections on objective probability within his longer discussions of evolutionary theory in biology.[10] He wrote: (2024, p. 243)

> An interpretation of probability must find some antecedently well understood concept that helps one define what probability means in some or all contexts of use, but why is this type of clarification needed when we say that the probabilities used in evolutionary theory are objective? We know that probabilities must conform to the axioms of probability, and we have a grasp of their epistemology. Why do we need anything more?

Here I answer this last question. The implicit definition of objective probability provided by an axiom system must be supplemented by an interpretive overlay that provides the practical aspects of our understanding of objective probabilities.

3. Implicit Definitions for Basic Terms and Interpretive Overlays

Implicit definitions assert enough truths about a quantity or concept that its application in cases of interest become univocal. For basic quantities and concepts, the defining assertions cannot be combined to form an explicit definition of the quantity or concept, that is, one whose formulation makes no mention of the quantity or concept being defined. Rather, implicit definitions of basic quantities are inevitably circular. The quantity or concept being defined is either included explicitly in the definition or it is included in the definitions of the terms used in the initial definition.

Implicit definitions are commonly accompanied by an interpretive overlay that assists us in informal understanding of the quantity or concept. These interpretive overlays can figure prominently in the discussions of the quantity or concept. However, they are merely aids to

---

[9] For some discussion, see Shafer and Vovk (2006, §5.2.3) "Why was Kolmogorov's philosophy not more influential?"

[10] I thank Shimin Zhao for alerting me to Sober's remarks.

human comprehension and play no essential role in the strict definition. The assertions of the implicit definition suffice to fix the quantity or concept. The remainder of this section has several illustrations.

3.1 Time

The most famous instance of the problem of demanding explicit definitions is Augustine's (1900, p. 244) 4th century ACE lament over time. "What, then, is time? If no one ask of me, I know; if I wish to explain to him who asks, I know not." The difficulty is immediately apparent then as now, once we try to formulate an explicit definition for time. For example, after following Einstein's operational analyses of relativistic time, we may seek to define time as that which is read by clocks.[11] The circularity lies in deciding which devices are clocks. Are they devices that cycle through their physical states, with periods of equal *time*?

The lament serves to highlight how the demand for an explicit definition confuses us. Relativity theory provides an elaborate physical theory of time as part of spacetime. In practical terms, we all know what time is, as well we know anything; and we do so in spite of not being able to give an explicit definition for it. For we know how to use all manner of temporal notions within our discourse. Relativity theory and this discourse defines time implicitly.

An interpretive overlay assists us in forming appropriate mental pictures. Time, we sometimes imagine, flows like a river; or the passage of time is like a spotlight moving over the various moments, laid out like the frames of a motion picture strip. This interpretive overlay, however, plays no role in the formal theories of time, such as in the precise formulations of modern relativity theory.

3.2 Temperature

A more recent illustration is provided by the introduction of temperature T into thermodynamics. It enters the laws of thermodynamics through an implicit definition in the "zeroth law." Here is an early formulation of it: (Fowler and Guggenheim, 1956, p. 56, their emphasis)

---

[11] Einstein's (1905, p. 39) simplest example is: "If, for instance, I say, 'That train arrives here at 7 o'clock,' I mean something like this: 'The pointing of the small hand of my watch to 7 and the arrival of the train are simultaneous events.'"

> As a natural generalization of experience we introduce the postulate: *If two assemblies are each in thermal equilibrium with a third assembly, they are in thermal equilibrium with each other.* From this it may be shown to follow that the condition for thermal equilibrium between several assemblies is the equality of a certain single-valued function of the thermodynamic states of the assemblies, which may be called the temperature *t*…

Further, similar assertions enhance the implicit definition. Most important is that heat only passes spontaneously from one body to another if the first is at a higher temperature.[12]

We also may add an interpretive overlay to assist us in forming mental pictures of temperature. For example, we may think of temperature as some measure of the density of heat in a hot body and imagine that the greater the density the more the heat seeks to escape the body. A lower temperature body lacks that density and is happy to accept heat. Thus, we develop an informal understanding of why heat moves from hotter to colder bodies. The mental picture, however, is not part of the formal thermodynamic theory.

3.3 Causation

A third example, stated briefly here, is provided in the philosophy of causation. Woodward's interventionist account of causation defines relations over variables implicitly as causal by this condition: an intervention on the cause is associated with a change in the effect, but not *vice versa*. In later work, Woodward provided a simple formulation adapted to variables governed by deterministic relations: (2021, §2)

> (M) Where *X* and *Y* are variables, *X* causes *Y* iff there are some possible interventions that would change the value of *X* and if were such intervention to occur, a regular change in the value of *Y* would occur.

The definition is implicit since an intervention is itself a causal relation, albeit one with specific additional characteristics suited to the specifics of its role in the definition.

---

[12] Much more can be said. The two results given so far only fix temperature up to an arbitrary monotonic rescaling. Subsequent analysis by Thomson of idealized, reversible heat engines reduces the rescaling admissible to linear functions and shows that this absolute temperature scale is the same as read by an ideal gas thermometer.

The interpretive overlay involves familiar images of the cause somehow reaching out to bring about the effect. We can use causes defined in accord with this definition as a means to manipulate effects, where we conceive of causal processes to be much like a puppeteer pulling strings to activate a marionette.

4. Objective Probability Defined Implicitly

Objective probability *p* arises as an aspect of a physical process or as a term in a physical theory, which is identified here as the formal theory governing the objective probability *p*. The objectivity resides in the independence of the probability from human thoughts and beliefs. The proposal is that we define the *p* of objective probabilities in the same way as the examples in Section 3, that is, implicitly. The formal properties of *p* are to be fixed implicitly by the relations in which it appears.

We saw above in Section 2 that Kolmogorov introduced his axioms for probability to provide an implicit definition for probability. For completeness, I recall his axioms here for a finitely additive probability measure: (1956, p. 2, Kolmogorov's emphasis)

> Let *E* be a collection of elements $\xi, \eta, \zeta, \ldots$, which we shall call *elementary events*, and *F* a set of subsets of *E*; the elements of the set *F* will be called *random, events*.
>
> I. *F is a field of sets.*
>
> II. *F contains the set E.*
>
> III. *To each, set A in F is assigned a non- negative real number* $\mathsf{P}(A)$*. This number* $\mathsf{P}(A)$ *is called the probability of the event A.*
>
> IV. $\mathsf{P}(E)$ *equals* 1.
>
> V. If A and B have no element in common, then,
>
> $$\mathsf{P}(A+B) = \mathsf{P}(A) + \mathsf{P}(B)$$
>
> A system of sets, *F*, together with a definite assignment of numbers $\mathsf{P}(A)$, satisfying Axioms I-V, is called a *field of probability*.

This axiom systems introduces some mathematical complications that are not relevant to the simple example to be pursued below. Most significantly, results in measure theory had shown that it will not always be possible to assign a [σ-]additive probability measure to every subset of

the total outcome set E. Some may be non-measurable. The field introduced in Axiom II gives us the flexibility to assign a probability measure only to some subsets of E.

This section will treat a simple case. It will suffice to illustrate how objective probabilities can be defined implicitly. The same idea could be illustrated in more complicated or more general scenarios. But they would not substantially add to what is displayed in this simple case. The case is the familiar one of many identical, independent systems each of which can adopt one of two states. Such a system is the tossing of many coins, where each lands with the outcome of H (heads) or T (tails). Another is the measurement of the spin of many independent spin-1/2 particles. The measurement outcomes are "up" or "down." Again, we might track the motion of a single thermal molecule in a chamber filled with a dilute gas. At any moment, the molecule may be in "left" L or "right" R part of the chamber.

This simplest of examples already displays the sufficiency of implicit definitions of objective probability. Because of the great familiarity of the example, I will give only an informal development that shows the key parts. For concreteness, take the case of n independent coin tosses for coins 1, 2, …, n. For each coin toss, the objective probability p for each outcome is $p(H) = r$, $p(T) = (1–r)$. The following further conditions define its applications implicitly:

Real value. p is a real-value function with range [0,1] defined over the subsets of the outcome spaces of each individual coin toss and the product outcome spaces of some sequence of coin tosses.

Normalization. For each outcome space, p(outcome space) = 1.

Additivity. If A and B are disjoint subsets of the same outcome space, $p(A \cup B) = p(A) + p(B)$.

Independence. If S and T are disjoint, independent outcome spaces and $A \subseteq S$ and $B \subseteq T$ so that $AB \subseteq S \times T$, then $p(AB) = p(A) \bullet p(B)$.

These propositions describe the properties of the function p that assigns real numbers to in the range of 0 to 1 to subsets of the outcome spaces. The application to specific cases requires further conditions for each case. If the coin tosses are fair, the physical symmetries of the tosses lead us to set $p(H) = 1/2$. In other cases, r can take on different values.

5. The Essential Factual Foundation for the Axioms

For objective probabilities to be objective, they must express something factual, independent of our thoughts and beliefs, about the systems of interest. This condition is met by assuring that the axioms used are such facts and that they are empirically supported. The details of how their factual character and empirical support is secured are not simple. They involve deep foundational investigations in significant sciences. A few examples just touch on what is involved in them.

5.1 Quantum Systems

The most straightforward case arises in quantum mechanics. The spatial degrees of freedom of a particle is described by its complex-valued wave function $\psi(x)$. The norm-squared of the wave function $|\psi(x)|^2$ is formally a probability density over the measured positions x possible for the particle. The quantity $|\psi(x)|^2 dx$, for some small interval dx, formally has the properties of a probability and verifies the axioms. The empirical foundation for this result is spread through the two decades and more of empirical work in quantum theory, leading up to the solidification of the theory in the late 1920s.

The case of spin measurements on a spin half particle is more abstruse, but it makes this connection to an additive measure more transparent. If we have such a particle whose spin is aligned at angle θ to the "spin up" $z+$ direction, its state is, in the simplest case:

$$|\psi(\theta)> = \cos(\theta/2)\ |z+> + \sin(\theta/2)\ |z->$$

The amplitudes for the mutually exclusive spin up ($z+$) and spin down ($z-$) measurement outcomes are $\cos(\theta/2)$ and $\sin(\theta/2)$ respectively. Their norms-squared are $\cos^2(\theta/2)$ and $\sin^2(\theta/2)$. They provide the objective probability $p(z+)$ of a measurement outcomes of spin up and $p(z-)$ of spin down, since each has real value in 0 to 1 and they obey additivity by the familiar trigonometric identity

$$\cos^2(\theta/2) + \sin^2(\theta/2) = 1.$$

That is:

$$p(z+) + p(z-) = 1.$$

5.2 Deterministic Classical Systems

The statistical physics of thermal systems depends on a probability measure defined over the possible states of the system. The awkwardness for the case of classical physical systems is

that the time development in the phase space representing the physical system is deterministic. The most natural way to introduce an additive measure is to consider the evolution of the system over a long period of time and to associate what becomes the objective probability of a state with the fraction of the time the system spends in the region of phase space associated with that state.

Such was Einstein's assumption in his 1903 "A Theory of the Foundations of Thermodynamics." He offered a "precisification" ("*präzisieren*") of the postulate ("*Vorassetzung*") underlying his account, where the system occupies a region Γ of the phase space for a portion τ of the total time T. It is that[13] (p. 172) "… a system that adopts a stationary state, has, for each region Γ, a magnitude τ/T with a definite limiting value for $T=\infty$." The fractional occupation time, τ/T, satisfies both the normalization and additivity properties and thus serves formally at an objective probability.

The complication is that Einstein's postulate, although widely adopted tacitly, still lacks a fully secure demonstration for thermal systems in general. However, I am confident that something like it does provide the factual foundation of the objective probabilities of statistical physics.

The familiar case of coin tosses, if assumed within the context of classical theories, faces a similar problem, in so far as its underlying dynamics is deterministic. In long-standing practice that has spanned centuries, we treat coin tosses as if they were governed by an effective theory in which coin toss outcomes are governed by objective probabilities. Extensive investigations have sought to provide some form of justification in the factual properties of coin tossing systems. Poincaré's (1912, pp. 148-150) method of arbitrary functions seems to me the most promising.

6. The Connection to Relative Frequencies

Objective probabilities are closely connected to frequencies. The connection does not arise from an identity of objective probabilities and frequencies. Rather the connection is that the assertions used above to define implicitly objective probabilities correspond closely[14] with those

[13] "…eines Systems, welches einen stationären Zustand annimt, so besitzt die Grösse τ/T für jedes Gebiet Γ einen bestimmten Grenzwert."

[14] Closely, but not exactly. Objective probabilities are real-valued. Frequencies are rational valued. In more complicated cases of infinite outcomes spaces, frequencies are only finitely

characterizing frequencies. This closeness is the basis for Kolmogorov's "empirical deduction of the axioms," as he put it, in the interpretive §2 of his treatise.

The connection with frequencies is made precise with the various laws of large numbers in the probability theory literature. They assert, loosely, that as the number of independent trials increases, the objective probability rises towards one of an ever-closer match between the objective probability of success on a single trial and the frequency of successes. Such a law is not an explicit definition. The closeness of the match of objective probability and frequency is asserted only as probable.

The simplest version of the weak law of large numbers is supplied by the binomial or Bernoulli distribution for the number of heads *H* in many, independent coin tosses. For *n* tosses, the objective probability *p* for *m* heads *H* is given as:

$$p(m, n|r) = \binom{n}{m} r^m (1-r)^{n-m}$$

where the binomial coefficient is $\binom{n}{m} = \frac{n!}{m!(n-m)!}$. The quantity of interest is the frequency m/n which we expect to be drawn closer in many trials to the objective probability r of heads in an individual trial. This distribution of m has a mean of nr and variance of $\sigma^2(m) = nr(1-r)$. It follows that the mean of the corresponding objective probability distribution for the frequency m/n is r and that the frequency is distributed around this mean with a variance of $\sigma^2(m/n) = r(1-r)/n$.

We read from these last results that the variance $\sigma^2(m/n)$ in the frequencies can be made arbitrarily small by choosing a sufficiently large number n of trials. Thus, the objective probability that the actual frequency m/n is in some arbitrarily small interval around r can be brought as close as we like to unity by choosing a sufficiently large n.[15] This is a statement

---

additive and not countably additive, which latter is commonly supposed for objective probabilities.

[15] The more precise version of this result uses Chebyshev's inequality. It entails that the objective probability that the frequency m/n differs from r by more than $k\sigma^2(m/n)$ is less than or equal to $1/k^2$. Since the deviation $k\sigma^2(m/n)$ can be made arbitrarily small by choosing a large enough n, we can be assured that the frequency m/n lies within that arbitrarily small interval with

within the formal analysis of the connection between objective probabilities and frequencies for this case of a coin toss.

We interpret this as the close connection of relative frequencies with the objective probabilities. But it is not an explicit definition of the objective probability in terms of frequencies. The usual interpretation that we *expect* a relative frequency close to p is part of a non-definitional interpretive overlay, to which we now turn.

7. Interpretive Overlays for Objective Probability

The results of the last section are the sorts of results that can be recovered for a notion of objective probability that is defined implicitly within the formal theory. They do not exhaust what we feel we learn through these results. These further ideas are not part of the formally derived theory. They are located in the interpretive overlay and thus need not conform exactly with the strictures of the formal theory. There are two versions of the overlay sketched below.

7.1 What We Expect: Cournot's Principle

We interpret the objective probability p of some outcome as an informal measure of how likely the outcome is to occur. A value of p of one half, at the midpoint of the range, is interpreted as indicating an outcome that is as likely as not. Deviations to higher and lower values correspond to outcomes that are increasingly more likely and less likely. These judgments are vague. As stated so far, they give us no means to distinguish clearly the likelihood of an outcome with p=0.6 from those p=0.65, beyond saying that the second is a little bit more likely than the first.

An elaboration of this approach brings some improvement to these looser considerations. It is known as Cournot's principle, which belongs in the interpretive overlay. The principle allows us to treat events of very low probability as ones that will not happen; and those of very high probability as ones that will. Adopting Cournot's principle allows us to sharpen the interpretive judgments above of likelihood. To apply it, we must first find a way to associate outcomes of intermediate objective probability with ones of very high and very low probability.

---

objective probability $0.99 = 1–1/10^2$, say, by choosing $k = 10$; or with objective probability $0.9999 = 1–1/100^2$ by choosing k=100; and so on.

These last outcomes are ones to which Cournot's principle can be applied. We can then use the result to illuminate outcomes of intermediate probability.

For example, we might ask just how likely is an outcome with objective probability $p = 0.60$. We can use the law of large numbers above to derive the formal result that the frequency of success in very many independent trials will be very close to 0.6 *with very high objective probability*. Cournot's principle uses that result to give us the assurance that the frequency *will be* very close to 0.6. We thereby have the means to distinguish more carefully the import of outcomes with objective probability of 0.6 from those with objective probability of 0.65. The first is associated in the interpretive overlay with assured frequencies of success of 0.6 and the latter with assured frequencies of success of 0.65.

These last reflections show how the connection between objective probabilities and frequencies can be made. Cournot's principle strips away the probabilistic qualifications on the connection in the formulation of the law of large numbers. It allows us to assert a direct connection between objective probabilities and frequencies. However, it only allows us to do so in the interpretive overlay. The simple connection is not part of the formal theory.

Cournot's principle enjoyed some prominence in earlier efforts to understand probability. Since the principle is employed, but not by name, in Kolmogorov's treatise, Shafer and Vovk (2006, §2.2) provide a rich and informative history of it leading up to Kolmogorov's treatise. Shafer (2023) provides a still richer compilation of writing on the Cournot principle.

The content of the principle appears as the essential element in the interpretive second section §2 of Kolmogorov's treatise. It is formulated for an event A with probability $\mathsf{P}(A)$ that can be realized under conditions S: (1956, p. 4)

> If $\mathsf{P}(A)$ is very small, one can be practically certain that when conditions S are realized only once, the event A would not occur at all.

The original German (Kolmogorov, 1933, p.4) is "… kann man praktisch sicher sein…" which I would read as "… one can be certain in practice…"

This version of Kolmogorov's proposal in §2 is preceded immediately by another version in Kolmogorov's proposal:

> One can be practically certain that if the complex of conditions S is repeated a large number of times, n, then if m be the number of occurrences of event A, the ratio m/n, will differ very slightly from $\mathsf{P}(A)$.

This version follows from the shorter version above, if that version is coupled with a law of large numbers.

Kolmogorov's interpretive rule continued to enjoy some prominence in the literature. It is given detailed examination under the name "Cournot's principle" in Hirsch and Nolfi (1954, p. 137). There, one version of the principle is:[16]

> If the probability p of the occurrence of an event E is sufficiently small, then one should proceed as if E will definitely not occur.

To recapitulate, this way of assigning meaning to probability depends on separating the implicit definition of objective probability p and its formal treatment in the probability theory from the understanding of p developed in the interpretive overlay. For Cournot's principle can only be maintained as a matter of practical application. It cannot be a part of the formal analysis since even the highest probability is compatible with the event failing to occur.

7.2 Credences Expressed as Subjective Probabilities

We can express these expectations in the language of subjective probabilities, if we find it appropriate to represent our credences by subjective probabilities. We do so by equating numerically our credence in some outcome, expressed as a subjective probability, with the objective probability of that outcome. Thus, if an outcome has an objective probability of 0.6, then our credence in its occurrence is expressed by a subjective probability of 0.6.

The connecting of objective and subjective probabilities is not automatic since they have different meanings. It requires some form of authorization. One frequently cited form is David Lewis' (1980) "principal principle," which asserts that we should match our credences as subjective probabilities to the objective probabilities, if the outcomes admit objective probabilities.

The material theory of induction also supplies an authorization. According to the analysis of Norton (2021, Ch.10), if we are to employ a probabilistic representation of any form of uncertainty, we have a positive obligation to display a warrant for the probabilistic representation in background facts. Such a warrant is present if we are forming credences over outcomes in a

---

[16] "Wenn für ein Ereignis E die Wahrscheinlichkeit p des Eintretens genügend klein ist, so soll man so tun ob E sicher nicht eintreten wird."

physical system governed by objective probabilities. The factual conditions asserted by the objective probabilities provide the warranting facts.

This matching gives us less than a full Bayesian analysis in which we might seek a posterior probability for r given m success in n trials. We secure only the conditional probability $p(m,n|r)$ above, now as a subjective probability, for the likelihood term in Bayes' theorem. What remains unauthorized is that other credences in the analysis are to be represented by probabilities. We must ask what warrants representing our uncertainty over the possible values of r by something as specific as a probability measure?

8. Conclusion

We have, I have argued here, been too ready to accept the need for an explicit definition of a basic quantity like objective probability. The impossibility of such a definition is a distinctive feature of truly basic quantities. They are so foundational that we cannot fully express them in terms of other quantities without entering into some form of circularity. It follows that efforts to find an explicit definition for a basic quantity like objective probability cannot succeed. They will forever be surrounded by an aura of futility and failure. The resolution offered here is that we have needlessly posed for ourselves an intractable problem by seeking an explicit definition at the outset. Recalling the analysis given by Kolmogorov in his celebrated treatises, objective probabilities can be given quite serviceable implicit definitions. We can then look to an interpretive overlay for many of the results we normally associate with objective probabilities but cannot be expressed directly in the strictly formal analysis. This two-part approach of implicit definitions and interpretive overlays allows us to keep the precision of a formal analysis and still retain interpretations that cannot be stated directly in the formal analysis.

References

Ben Porath, Gal (2024) *Conceptual and Ontological Foundations of Stochastic Dynamics*. PhD Dissertation, University of Pittsburgh.

Blumenthal, Otto (1935) "Lebensgeschichte," pp. 401-29 in *David Hilbert Gesammelte Abhandlungen*. Dritter Band. Berlin: Julius Springer.

Einstein, Albert (1903) "Eine Theorie der Grundlagen der Thermodynamik," *Annalen der Physik*, **11**, pp. 170-87.

Einstein, Albert (1905) “On the Electrodynamics of Moving Bodies,” p. 39 in H. A. Lorentz, et al, *The Principle of Relativity.* London: Methuen, 1923. Dover, reprint, n.d.

Fowler, Ralph H. and Guggenheim, Edward A. (1965) *Statistical Thermodynamics*. Cambridge: Cambridge University Press.

Hájek, Alan (1996) “Mises Redux" -- Redux: Fifteen Arguments against Finite Frequentism,” *Erkenntnis* **45**, No. 2/3, *Probability, Dynamics and Causality*, pp. 209-227.

Hájek, Alan (2009) “Fifteen Arguments Against Hypothetical Frequentism” *Erkenntnis*. **70**, pp. 211-35.

Hilbert, David (1902) *Grundlagen der Geometrie*. 2nd ed. Leipzig: B. G. Teubner.

Hirsch, M. Guy and Nolfi, Padrot (1954) “Sur un Aspect Paradoxal de la Théorie des Probabilités,” *Dialectica*, **8** No. 2, pp. 125-144.

Kolmogorov, Andrey N. (1933) *Grundbegriffe der Wahrscheinlichkeitsrechnung.* Berlin: Julius Springer.

Kolmogorov, Andrey N. (1956) *Foundations of the Theory of Probability*. Trans., ed. N. Morrison. New York: Chelsea Publishing Company.

Lewis, David (1980) “A Subjectivist's Guide to Objective Chance,” pp. 263–293 in *Studies in Inductive Logic and Probability*, Vol. II, ed. Richard C. Jeffrey. Berkeley: University of California Press,

von Mises, Richard (1931) *Wahrscheinlichkeitsrechnung und ihre Anwendung in der Statistik und theoretischen Physik*. Leipzig und Wien: Deuticke.

von Mises, Richard (1957) *Probability, Statistics and Truth*. Revised English edition, New York: Macmillan.

Norton, John D. (2021) *The Material Theory of Induction*. BSPSOpen/University of Calgary Press.

Poincaré, Henri (1912) *Calcul des Probabilités*. Paris: Gauthier-Villars.

Popper, Karl R. (1959) “The Propensity Interpretation of Probability,” *The British Journal for the Philosophy of Science*. **10**, pp. 25-42.

Reichenbach, Hans (1949) *The Theory of Probability*. 2nd ed. Trans E. H. Hutten and M. Reichenbach. Berkeley: University of California Press.

Shafer, Glenn and Vovk, Vladimir (2006) “The Sources of Kolmogorov’s *Grundbegriffe*,” *Statistical Science*, **21**, No. 1, pp. 70–98.

Shafer, Glenn (2023) "'That's what all the old guys said.' The many faces of Cournot's principle," http://www.probabilityandfinance.com/articles/60.pdf

Sober, Elliott (2010) "Evolutionary Theory and the Reality of Macro-Probabilities," pp. 133-161 in *The Place of Probability in Science: In Honor of Ellery Eells (1953–2006).*Dordrecht: Springer.

Sober, Elliott (2024) "Big Picture Questions," Ch. 9 in *The Philosophy of Evolutionary Theory : Concepts, Inferences, and Probabilities.* Cambridge University Press.

Woodward, James F. (2021) "Downward Causation Defended," pp. 217-51 in J. Voosholz and M. Gabriel, eds., Top-Down Causation and Emergence. New York: Springer.