Typicality First

Isaac Wilhelm

(please cite version in *The Philosophical Quarterly*)

Abstract

Instances of the law of large numbers are used to model many different physical systems. In this paper, I argue for a particular interpretation, of those instances of that law, which appeals to typicality. As I argue, the content of that law, when used to model physical systems, is that the probability of an event typically—rather than probably—approximates the frequency with which that event occurs.

1 Introduction

This paper is about perhaps the most basic, and most important, relationship between frequencies and probabilities: the relationship described by the law of large numbers. That law articulates a formal, mathematical relationship which connects frequencies to probabilities. Particular instances of that law are invoked throughout the sciences: many scientific theories use instances of that law to relate (i) the frequencies with which certain physical events occur, and (ii) the probabilities of those physical events. But there is disagreement over how, exactly, the relationship between (i) and (ii) should be understood.

According to a view which I call 'Probability First', the law of large numbers establishes a probabilistic relationship between frequencies and probabilities. Roughly put, Probability First says that with high probability, the probability of an event occurring is very close to the frequency with which that event occurs. In slogan form: probabilities *probably* equal frequencies.

According to a view which I call 'Typicality First', the law of large numbers establishes a different, typicalistic relationship between frequencies and probabilities. That relationship, according to Typicality First, should be understood in terms of typicality rather than probability. Roughly put, Typicality First says that typically, the probability of an event occurring is very close to the frequency with which that event occurs. In slogan form: probabilities *typically* equal frequencies.

Both views are implicitly referenced, and implicitly endorsed, throughout the literature.[1] But they are rarely discussed explicitly or at length. My primary goal, in what follows, is to remedy that. My secondary goal is to argue for Typicality First.

As will become clear, Typicality First deepens our understanding of the law of large numbers. For according to Typicality First, that law does more than just describe how frequencies connect to probabilities. That law also describes how typicality and probability connect to each other: the connection is mediated by the frequencies which, in typical cases, the probabilities approximate. So ultimately, the law of large numbers establishes a bridge between the notion of typicality and the notion of probability. That law describes how those physical notions interrelate.

Because of this, Typicality First deepens our understanding of the relationship between empirical frequencies and physical probabilities too. For according to Typicality First, that relationship should be understood in terms of typicality. Probabilities typically approximate frequencies, in the sense that in the vast majority of cases—in typical cases, that is—the probabilities and the frequencies are very close to each other. So typicality is the key to understanding why, and how, frequency and probability converge.

[1]For descriptions of the law of large numbers which endorse—or at least resemble—Probability First, see (Berkovitz, 2015, p. 633; Bricmont, 2020, p. 39; de Finetti, 1970/2017, pp. 277-278; Lewis, 1980, p. 273). For descriptions of the law of large numbers which endorse—or at least resemble—Typicality First, see (Dürr et al., 1992, p. 870; Dürr et al., 2017, pp. 135-142; Dürr & Lazarovici, 2009, pp. 58-61; Dürr & Teufel, 2009, pp. 102-106; Goldstein, 2012, p. 66; Hubert, 2021, p. 5265; Lazarovici & Reichert, 2015, p. 701; Maudlin, 2011, p. 315; 2020, p. 242).

Strikingly, as I explain later, the ideas underlying Typicality First support an elegant view of what was achieved by Kolmogorov's measure-theoretic axiomatization of probability. Very roughly put, according to that view, the mathematical theory of size—namely, measure theory—supports a sized-based representation of the probability axioms. Size facts, in particular, can act as mathematical 'stand-ins' for probability facts. The discussion of Typicality First will ultimately help illuminate, and motivate, this view of Kolmogorov's impressive achievement.

In Section 2, I discuss what typicality is, and I explain the differences between typicality and probability. In Section 3, I set the stage for the discussion to come: I introduce the main case on which I focus here, and I explain a few formal notions which the law of large numbers invokes. In Section 4, I present the law of large numbers, and I formulate precise versions of Typicality First and Probability First. In Section 5, I present an argument for Typicality First, and I defend that argument against some objections. Finally, in Section 6, I explain how the key idea in that argument—namely, that measures are used to express size facts—is compatible with probabilities featuring in the law of large numbers: in so doing, I propose and defend the view that size facts can act as stand-ins for probability facts, which I then connect to my preferred interpretation of Kolmogorov's axiomatization of probability.

In addition to formulating precise versions of Typicality First and Probability First, providing arguments in favor of the former, presenting an interesting interpretation of Kolmogorov's achievement, and exploring the law of large numbers' physical applications, this paper has another goal as well. There is often lots of confusion, and unclarity, in discussions of typicality and probability. Philosophical explorations of those two notions, both in the literature and in conversation, often get derailed by seemingly relevant, but actually irrelevant, concerns. One of my main goals, in this paper, is to flag—and dispel—many such troublesome confusions. The parts of the paper in which I do that, in what follows, are generally labeled '*Confusion dispelled*'.

2 Typicality and Probability

Typicality and probability are, in several important ways, different from one another. In this section, I discuss four differences between them. To start, I explain what typicality is. Then I explain how typicality differs from probability.

For present purposes, it is convenient to use an account of typicality which invokes mathematical measures (Wilhelm, 2022; forthcoming; Frigg, 2009; Frigg & Werndl, 2012; Goldstein, 2012; Maudlin, 2020). In particular, let Γ be a set, let P be a property, and let Γ_P be the set of all elements in Γ which have P. Let M be a measure over an algebra of sets which contains Γ_P, such that $M(\Gamma)$ is finite. Then P is *typical* in Γ if and only if $M(\Gamma_P)$ is very close to $M(\Gamma)$. In other words, P is typical in Γ just in case the size of the set Γ is roughly equal to the size of the set of elements in Γ which have P.

Note that this account of typicality appeals to mathematical measures. Formally, a mathematical measure is a function which satisfies certain conditions. But described using intuitive terms, mathematical measures assign sizes to sets. Mathematical measures can be used to represent facts about how big various sets are.

For example, let Γ be the set of all possible microstates of a gas confined to a box. Let P be the property of evolving to the equilibrium macrostate relatively quickly; so Γ_P is the set of all microstates that, pretty quickly, reach equilibrium. Then as can be shown, Γ_P is roughly as big as Γ is. There is a natural measure, of the sizes of these sets, which implies as much: the 'modified Lebesgue measure'.[2] So P is typical in Γ: the property of evolving to equilibrium relatively quickly, in other words, is typical in the set of all possible microstates for the gas in question.

Confusion dispelled: do not make the mistake of rejecting the views I develop, in this paper, by pointing out that other accounts of typicality do not appeal to mathematical measures. There are many accounts of typicality, of course: some appeal to cardinality

[2]For detailed discussions of this measure, and the corresponding famous result about the sizes of Γ and Γ_P, see (Boltzmann, 1877/2015; Goldstein, 2001; Lazarovici & Reichert, 2015).

(Wilhelm, 2022); some appeal to purely topological notions (Frigg, 2009); and some appeal to other mathematical devices. Each account formalizes a distinct, but still perfectly respectable, notion of what it is for something to be typical. So there is no conflict between the measure-theoretic account and these other accounts, since the notions of typicality that all of these accounts formalize are different. And in this paper, I focus on the measure-theoretic account for the simple reason that, obviously, mathematical measures feature in the law of large numbers: so a measure-theoretic account of typicality is appropriate here.

It is worth introducing some terminology. Say that a 'typicality fact' is a fact about what is, or is not, typical. In the example of the gas, for instance, the fact that P is typical in Γ counts as a typicality fact. And say that a 'probability fact' is a fact about the probability of some proposition. The fact that a radium atom decays with thus-and-so probability, for example, is a probability fact.

One might wonder whether typicality is just probability in disguise. For a property to be typical in a set, one might claim, is just for the probability of having that property—for elements in that set—to be extremely high. Facts about typicality, in other words, are just facts about probability.

But that is not so. In what follows, I present four differences between typicality and probability. For lack of space, I do not argue for these differences here. Instead, I describe them, and point to those places in the literature where the arguments are developed in detail.

Regarding the first difference: typicality and probability have different formal, logical, and grammatical features. Probability facts are expressed using upward continuous mathematical measures; typicality facts need not be expressed by mathematical measures which are upward continuous (Goldstein, 2012, p. 71). The semantic contents of sentences which invoke probability are different from the semantic contents of sentences which invoke typicality (Crane & Wilhelm, 2020, pp. 178-179). And typicality is often represented using different formal frameworks from those used to represent probability (Maudlin, 2020, p. 234).

Regarding the second difference: what it is to be typical is, importantly, distinct from what it is to be probable. Typicality facts concern the sizes of sets; probability facts concern the likelihoods of events (Goldstein et al., 2010, p. 3211). Probability is intimately related to randomness; typicality is not (Lazarovici & Reichert, 2015, pp. 698-699). As I prefer to put it, typicality—though not probability—is ultimately about size, that is, about 'how much' of something is a certain way.

Confusion dispelled: do not object by observing that size is often formalized using cardinality, or topology, rather than the measures of the measure-theoretic account of typicality. That observation is correct, of course: size can be formalized using cardinality, topology, and many other mathematical notions. But that observation is irrelevant here. All that matters, for my purposes, is that size can also be formalized using mathematical measures, in the way that the measure-theoretic account of typicality describes. In other words, what matters is simply that mathematical measures represent one reasonable way—not the only reasonable way, but one reasonable way—to formalize a notion of size. So those who understand typicality in this way are not committed to rejecting the claim that cardinality, or topology, or various other mathematical notions, can be used to formalize notions of size as well. I focus on measure-theoretic formalizations of size here because, again, measures play an important role in the law of large numbers.

Regarding the third difference: typicality facts and probability facts support different sorts of explanations. Some probabilistic explanations cannot be interpreted as typicality explanations, and some typicality explanations cannot be interpreted as probabilistic explanations (Wilhelm, 2022, pp. 574-577). It is sometimes argued that typicality facts, for instance, support the best explanations of quantum phenomena in Bohmian mechanics (Dürr et al., 1992, p. 856).[3]

Regarding the fourth difference: typicality and probability feature in different rationality principles (Wilhelm, forthcoming, p. 5). In particular, typicality guides rational belief

[3]This is somewhat controversial. Valentini (2020) argues that these explanations of quantum phenomena, based on typicality facts, are not successful; so other sorts of explanations are needed. Norsen (2018) argues that the best explanations, of these quantum phenomena, combine facts about typicality with Valentini's results.

in accord with the Typical Principle. Probability, in contrast, guides rational credence in accord with the Principal Principle.

So typicality and probability differ in several important ways. Typicality and probability have different formal, logical, and grammatical properties. What it is to be typical differs from what it is to be probable. Typicality and probability support different kinds of explanations. And typicality and probability figure in different principles of rationality. So typicality and probability are indeed distinct.

Confusion dispelled: for reasons related to the above, throughout this paper, I avoid using the potentially misleading phrase 'probability measure'. In mathematics, a 'probability measure' is defined as any mathematical measure whatsoever which has certain formal properties. Plenty of typicality facts are expressed using mathematical measures which have those properties, and so plenty of typicality facts are expressed using 'probability measures' in the mathematical sense of that phrase. But those typicality facts are still distinct from probability facts, for all the reasons just given: they may be expressible using different formal or logical resources; they make different claims about the world; they support different kinds of explanations; and they figure in different rationality principles. So a measure in a typicality fact can be a 'probability measure'—in the mathematical sense of 'probability measure'—and yet that typicality fact may not be a probability fact. And so to avoid confusion over all this, I do not use the phrase 'probability measure' here.

3 Stage-Setting

In this section, I introduce the main physical system on which I will focus: an infinite sequence of coin flips. I also explicate some aspects of the formalism which the law of large numbers—discussed in detail later—invokes. Ultimately, I explain how that formalism can be used to represent facts about the frequency with which a coin lands heads.

Before beginning, it is worth making a quick disclaimer. Throughout this paper, I discuss a specific instance of the law of large numbers: its application to a case involving coin flips. I focus on that instance of the law merely in order to keep things simple. But all of my points generalize to other instances of the law of large numbers, such as instances which feature in contemporary scientific theories of statistical mechanics and quantum mechanics.

Similarly, to keep things simple, I focus on the strong version of the law of large numbers. For that version is the easiest to explain informally. All my points generalize, however, to other versions of that law, like the weak version.

Confusion dispelled: do not object to what follows by claiming that finite sequences, and not infinite sequences, are what scientists observe. It is true, obviously, that actual scientific experiments never perform infinite numbers of trials. But that is irrelevant for my purposes here. For the same points could be made about finite sequences, as described by the weak law of large numbers (Gut, 2005). In particular, everything which I say about (i) infinite sequences approximating probabilities in a set of cases whose measure is 1, could be said about (ii) finite sequences approximating probabilities in a set of cases whose measure is extremely close to 1.

Now for a terminological point. The law of large numbers is a mathematical theorem, a bit of pure mathematics. Nevertheless, many physical systems are well-modelled by particular instances of the law of large numbers: namely, physical systems which satisfy the assumptions upon which the law of large numbers is based. Those particular instances of the law, which serve to model physical systems, are my focus here; not the purely mathematical law of large numbers. So when I refer to an 'interpretation' of that law—or when for brevity, I simply refer to that law itself, dropping all mention of interpretations—I mean an interpretation of a particular instance of that law: an interpretation of an instance of the law which models some physical system or other. I do not mean an interpretation of the mathematical theorem: this is not a paper about platonism, nominalism, or anything like that.

Confusion dispelled: relatedly, do not claim that, since the axioms which imply the law of large numbers are probability axioms, that law must be interpreted in accord with

Probability First rather than Typicality First. It is true, of course, that the mathematical law of large numbers follows from the axioms of mathematical probability. But Probability First, and Typicality First, are not accounts of the interpretation of the mathematical law of large numbers. They are accounts of the interpretation of particular applications of the law of large numbers to physical systems. Because of this, terms like 'probability axiom' can be misleading in the present context, for much the same reasons that—as I explained in Section 2—the term 'probability measure' can be misleading. Just as a measure can be a 'probability measure' in the mathematical sense, and yet be used to express a typicality fact rather than a probability fact, some axioms may be 'probability axioms' in the mathematical sense, and yet imply a theorem whose descriptions of physical systems are best interpreted using typicality rather than probability. And that is what this paper is about: the extent to which Probability First, or Typicality First, provides the best interpretation of those descriptions—based on the law of large numbers—that feature in science.

Now to discuss coin flips, and the instance of the law of large numbers which describes them. Suppose that a fair coin is going to be flipped infinitely many times. In each flip, exactly one of two possible outcomes obtains: the coin lands heads, or the coin lands tails. Let h represent the former outcome, and let t represent the latter outcome.

Let Ω be the set of all countably infinite sequences whose members are h and t. Each sequence in Ω represents a possible way for the coin flips to unfold. The sequence $\langle h, h, h, \ldots \rangle$, for instance, represents the possibility in which the coin always lands heads. Similarly, in the sequence $\langle h, t, t, h, \ldots \rangle$, the coin first lands heads, then lands tails, then lands tails, then lands heads, and so on.

The law of large numbers is formulated, in part, using random variables. Roughly put, random variables are functions which can be used to record the outcomes of probabilistic processes. For each natural number i, let X_i be a random variable which says, of each possible infinite sequence of coin flips, whether or not the ith flip in that sequence lands heads. A little more precisely, take any natural number i, and take any possible infinite sequence of coin flips ω in Ω. Then X_i has the following properties.

(1) If the ith member of ω is h, then $X_i(\omega) = 1$.

(2) If the ith member of ω is t, then $X_i(\omega) = 0$.

In other words, X_i is a function which uses 0 and 1 to represent the outcome of the ith flip.

Random variables can be used to express facts about the relative frequency with which the coin lands heads. To see how, take any possible infinite sequence of coin flips ω. And take any natural number n. Then consider the expression below.

$$\frac{1}{n}\sum_{i=1}^{n} X_i(\omega)$$

For each i, if the coin lands heads in the ith flip in ω, then $X_i(\omega)$ is 1; otherwise, $X_i(\omega)$ is 0. So the total number of 1s, in the sum above, equals the number of times—in the first n flips of the sequence ω—that the coin lands heads. The result of summing up all those 1s, and dividing that number by n, is the relative frequency with which the coin lands heads in the first n flips of the ω sequence.

Facts about relative frequencies, of the coin landing heads, can be used to express facts about the frequency with which the coin lands heads across infinite—rather than just finite—numbers of flips. To see how, once again, take any possible infinite sequence of coin flips ω. Then consider the expression below.

$$\lim_{n\to\infty} \frac{1}{n}\sum_{i=1}^{n} X_i(\omega) \tag{1}$$

As explained above, for each n, the expression $\frac{1}{n}\sum_{i=1}^{n} X_i(\omega)$ represents the relative frequency with which the coin lands heads in the first n flips of the ω sequence. Given that, (1) says the following. Consider what happens to that relative frequency as n gets bigger and bigger without bound; this corresponds to the '$\lim_{n\to\infty}$' part of the expression above. The resulting quantity is called the 'limiting relative frequency', in sequence ω, of the coin landing heads. It represents how frequently, throughout the entire ω sequence, heads occurs.

The law of large numbers expresses a fact about the relationship between probabilities and limiting relative frequencies. There is disagreement, however, over exactly what that fact is: Typicality First provides one account of that fact, and Probability First provides another. Let us see how.

4 Typicality First and Probability First

In this section, I present precise versions of both Typicality First and Probability First. By way of preparation, I introduce a mathematical measure which the law of large numbers invokes. Then I present a version of the law of large numbers. Finally, I formulate two different interpretations of that law: Typicality First, and Probability First.

There is a natural, intuitive—and most importantly, physically justified, when used in the model of the physical system consisting of fair coin flips—measure of the subsets of Ω. This measure, call it 'M', says how big various subsets of Ω are: for each such subset A, $M(A)$ is the size of A.[4] For example, let H_1 be the set of all sequences in Ω whose first member is h: so H_1 represents all the possible infinite sequences of coin flips in which the first flip lands heads. Then $M(H_1) = \frac{1}{2}$: that is, the size of H_1 is $\frac{1}{2}$.

There is another, related way to think about the sizes which M assigns to sets in Ω. It turns out that $M(\Omega) = 1$: the size of Ω itself, in other words, is 1. Because of that, for each set A to which M assigns a size, $M(A)$ represents the proportion of Ω which A contains. The equation $M(H_1) = \frac{1}{2}$, for example, can be understood as saying this: H_1 contains half of the sequences in Ω. Or to put it another way: in 50% of the possible infinite sequences of coin flips, the first flip lands heads.

An extremely important observation: it is no accident that, intuitively, mathematical

[4]Strictly speaking, M—which is basically the Lebesgue measure—is defined over the smallest σ-algebra Σ containing all cylinder sets in Ω, where $C \subseteq \Omega$ is a cylinder set just in case for some natural number n and for some $a_1, \ldots, a_n \in \{h, t\}$, $C = \{(\omega_1, \omega_2, \ldots) \in \Omega \mid \omega_1 = a_1, \omega_2 = a_2, \ldots, \omega_n = a_n\}$. Then M is defined as the unique measure from Σ to $[0, 1]$ such that (i) $M(\Omega) = 1$, and (ii) for all cylinder sets $C \subseteq \Omega$ such that $C = \{(\omega_1, \omega_2, \ldots) \in \Omega \mid \omega_1 = a_1, \omega_2 = a_2, \ldots, \omega_n = a_n\}$, $M(C) = \frac{1}{2^n}$. For more, see (Tao, 2011).

measures like M quantify the sizes of sets. For the mathematical theory of measures was designed to do just that. Mathematical measures were initially used to extend the Riemannian theory of integration. And integration is, intuitively, about the sizes of sets, and the sizes of the shapes which those sets compose (Hawkins, 2002, p. 63; Paunić, 2002, p. 3; Stillwell, 2010, p. 531).

The measure M, along with the random variables from Section 3, can be used to define the probability of any given flip landing heads. The definition, in rough outline, is as follows. For each natural number i, let H_i be the set of all sequences in Ω whose ith member is h. The probability of flip i landing heads is identified with the expected value of the random variable X_i, and the expected value of X_i is defined in terms of M. The definition implies that the probability of flip i landing heads equals $M(H_i)$, which—a straightforward computation shows—equals $\frac{1}{2}$.[5] In other words, a fact about measures can be wielded to articulate a fact about probabilities.[6]

With all that as background, here is the law of large numbers. Basically, it expresses a fact about the relationship between (i) the limiting relative frequency, in any given sequence, of the coin landing heads, and (ii) the probability, in any given flip, of the coin landing heads.

$$M\left(\left\{\omega \in \Omega \;\middle|\; \lim_{n\to\infty} \frac{1}{n}\sum_{i=1}^{n} X_i(\omega) = \frac{1}{2}\right\}\right) = 1 \qquad \text{(LLN)}$$

Intuitively, (LLN) says the following. Consider the set of all possible infinite sequences of coin flips such that the limiting relative frequency of heads, in those sequences, equals the

[5]To derive this result, let $E(X_i)$ represent the expected value of X_i, and let T_i be the set of all sequences in Ω whose ith member is t. Then $E(X_i) = \int_\Omega X_i(\omega)dM(\omega) = 0 \cdot \int_{T_i} dM(\omega) + 1 \cdot \int_{H_i} dM(\omega) = \int_{H_i} dM(\omega) = M(H_i) = \frac{1}{2}$, where the first equality follows from the definition of expected value and the other equalities follow from measure theory. A nice consequence of this: for all i and j, $E(X_i) = E(X_j)$; that is, all the random variables have the same expected value. This captures the fact that each flip of the coin has the same probability of landing heads.

[6]In Section 6, I discuss the fact that $M(H_i)$ equals the probability of flip i landing heads. As I argue, in that equality, a fact about size—namely, the fact that the set H_i has size $\frac{1}{2}$—serves as a stand-in for a probability fact.

probability of the coin landing heads. Then that set, according to (LLN), has measure 1. In other words, that set takes up 100% of Ω.[7]

With all that as background, we are ready to see the two interpretations of the law of large numbers: Typicality First, and Probability First. These interpretations disagree over how to interpret the measure M. As a result, they disagree over the physical fact that (LLN) expresses.

To start, here is a precise formulation of Typicality First. It says that according to the law of large numbers, typically, the coin lands heads 50% of the time.

<u>Typicality First</u>

(LLN) says that the property of having a limiting relative frequency of $\frac{1}{2}$, for the coin landing heads, is typical in Ω. The instance of measure M, in (LLN), serves to express a typicality fact: it helps express the fact that a certain set is roughly as big as Ω is.

In other words, (LLN) says that typically, the frequency with which the coin lands heads equals the probability of the coin landing heads. In slogan form: probabilities typically equal frequencies.

Now for a precise formulation of Probability First. It says that according to the law of large numbers, with probability 1, the coin lands heads 50% of the time.

<u>Probability First</u>

(LLN) says that for any possible infinite sequence of coin flips ω, the probability of that sequence having a limiting relative frequency of $\frac{1}{2}$ is equal to 1. The instance of measure M, in (LLN), serves to express a probability fact: it helps express the fact that a certain event obtains with unit probability.

[7] Of course, that set is also a proper subset of Ω.

In other words, with probability 1, the frequency with which the coin lands heads equals the probability of the coin landing heads. In slogan form: probabilities probably equal frequencies.

Typicality First and Probability First are distinct accounts of the law of large numbers. According to Typicality First, that law expresses a typicality fact: it says something about typical limiting relative frequencies. According to Probability First, that law expresses a probability fact: it says something about how probable certain possible infinite sequences are. And for all the reasons explained in Section 2, typicality and probability are distinct. So these are distinct interpretations of what the law of large numbers says.

Before concluding this section, three potential confusions should be addressed. The first and second concern how Typicality First, and Probability First, relate to actual scientific practice. The third concerns the uniformity of the measure M.

First confusion dispelled: note that Typicality First and Probability First are not competing accounts of what particular scientists have in mind, when those scientists use the law of large numbers. As a survey of the literature shows—see the citations in Section 1, for instance—some scientists prefer Typicality First while other scientists prefer Probability First. This paper is compatible with, but not primarily about, theories of the psychological and sociological underpinnings of those preferences. This paper is informed by, but not merely a reproduction of, what scientists say about the law of large numbers. Ultimately, Typicality First and Probability First are competing accounts of the proper interpretation of the applications of that law to physical systems. So the project here is interpretive, not descriptive.

Second confusion dispelled: it would be a mistake to read Typicality First as claiming that all instances of the law of large numbers, which practicing scientists actually use, are false when M is understood to express a probability fact. Similarly, it would be a mistake to read Probability First as claiming that all instances of the law of large numbers, which practicing scientists actually use, are false when M is understood to express a typicality fact. Both understandings of M may well render some instances of the law of large numbers, which practicing scientists actually use, as true. For the difference between Typicality First and Probability First does not concern the truth values of various interpretations of various instances of the law of large numbers; both views are compatible with both of the sentences "Probabilities typically equal frequencies" and "Probabilities probably equal frequencies" being true, in many cases. The difference between Typicality First and Probability First concerns, again, what the best interpretation of those instances of the law of large numbers is. According to Typicality First, the best interpretation invokes typicality; according to Probability First, the best interpretation invokes probability. The difference between Typicality First and Probability First, in other words, is about which physical interpretation of the measure M is best, not about whether different physical interpretations of the measure M have implications for the truth values of certain sentences.

Third confusion dispelled: do not ascribe any special significance to the fact that the measure M is uniform. It is tempting to assume, mistakenly, that typicality facts can only be expressed by uniform mathematical measures. If that were right, one might worry that if the example in this paper were changed, so that the fair coin were replaced by an unfair coin, then the corresponding instance of the law of large numbers could not be interpreted using typicality – and so the corresponding version of Typicality First would be false. But this is wrong. It is true, of course, that if the original example were changed so that it featured an unfair coin, then M would no longer be the right measure to use, when formulating the relevant instance of the law of large numbers. But it does not follow that the corresponding version of Typicality First would be false. Instead, the corresponding version of Typicality First would simply imply that when it comes to using certain sets to model the physical system consisting of the unfair coin, the sizes of those sets are—for the purposes of the modelling—best quantified using a non-uniform measure.[8] And there is nothing problematic about that: non-uniform measures are often used to quantify the sizes of sets in models.

[8] I discuss this in more detail in Section 5, when dispelling another, related confusion about M.

An example will help illustrate the issue. Consider a coin with a bias of $\frac{2}{3}$ towards heads. The right measure to use, when defining the probability of any given flip of this coin landing heads, is $M_{\frac{2}{3}}$: this is the measure which, for each natural number i, assigns a size of $\frac{2}{3}$ to H_i. The corresponding instance of the law of large numbers is this.

$$M_{\frac{2}{3}}\left(\left\{\omega \in \Omega \;\middle|\; \lim_{n\to\infty} \frac{1}{n}\sum_{i=1}^{n} X_i(\omega) = \frac{2}{3}\right\}\right) = 1 \qquad \text{(LLN 2)}$$

According to the corresponding version of Typicality First, the property of having a limiting relative frequency of $\frac{2}{3}$, for the unfair coin landing heads, is typical in Ω: the instance of measure $M_{\frac{2}{3}}$, in (LLN 2), serves to express the typicality fact that a certain set is roughly as big as Ω is. And that is perfectly fine. So the non-uniformity of $M_{\frac{2}{3}}$ does not create any problems for typicality-based views of the law of large numbers. Those views are perfectly compatible with non-uniform measures featuring in that law.

In short, the uniformity of M—and correspondingly, the fairness of the coin—is irrelevant for my purposes here. This is worth belaboring because sometimes, it is assumed that non-uniform measures pose problems for typicality. But that is not so. Plenty of typicality measures are non-uniform: take the measure used in Bohmian mechanics (Dürr et al., 1992), or the measure which Everett used in his relative-state formulation of quantum theory (Barrett, 2017), or the modified Lebesgue measure mentioned in Section 2. Non-uniform measures can do just as good a job of quantifying the sizes of various sets as uniform measures can.

5 An Argument for Typicality First

In this section, I give an argument for Typicality First. Basically, the argument is this: Typicality First is worth endorsing because it respects the intuitive interpretation of mathematical measures. After presenting the argument, I respond to an objection.

The argument is based on the intuitive interpretation—given in Section 4—of what mathematical measures are used to express. Recall that according to this intuitive interpretation, mathematical measures quantify the sizes of sets. The measure M, for instance, says how big various subsets of Ω are.

This interpretation of M generates the following interpretation of the law of large numbers: that law is about the size of a particular set of possible infinite sequences of coin flips. In particular, let S be the property of being a sequence in Ω whose limiting relative frequency of heads equals the probability of heads. So roughly put, S is the property—had by some sequences and not others—of the coin landing heads with a limiting relative frequency of $\frac{1}{2}$. Let Ω_S be the subset of Ω consisting of all and only those sequences which have S. So Ω_S is the set which features in (LLN). Therefore, (LLN) says that the size of Ω_S, according to M, is 1. In other words, according to M, Ω_S is as big as Ω itself.

Typicality First does an extremely good job of respecting all this. The reason is simple: as explained in Section 2, typicality facts are about the sizes of sets. So given Typicality First, the instance of M in (LLN) expresses a fact about a particular set's size. And that interpretation of M, as assigning a size to a set, is just the intuitive interpretation of M given before.

In fact, Typicality First actually follows directly from the intuitive interpretation of mathematical measures. To see how, recall the measure-theoretic account of typicality from Section 2: property P is typical in set Γ just in case the set Γ_P—the set containing all and only the elements of Γ which have P—is roughly as big as Γ itself. Then recall the property S: the property, that is, of having a limiting relative frequency of heads which equals the probability of heads. And consider the set Ω_S of all and only the sequences in Ω which have this property. Then $M(\Omega_S) = 1$; this is just what (LLN) says. Also, as a matter of pure mathematics, $M(\Omega) = 1$. Now suppose that the measure M is interpreted in the intuitive way described above, as quantifying the sizes of sets. Then (LLN) and the equation $M(\Omega) = 1$ both express facts about the sizes of Ω_S and Ω, respectively: and jointly, those two equations say that Ω_S is roughly as big as Ω is. Therefore, by the measure-theoretic account of typicality, S is typical in Ω. So putting all this together: given the intuitive interpretation of mathematical measures, (LLN) says that the property S—of having a limiting relative frequency of heads which equals the probability of heads—is typical in the set of all possible infinite sequences of coin flips. In other words, given the intuitive interpretation of mathematical measures, Typicality First holds.

This argument for Typicality First—call it the 'measure argument'—takes the intuitive interpretation of measures seriously. According to that intuitive interpretation, measures quantify the sizes of sets. Typicality First says likewise for the instance of the measure M which appears in the law of large numbers. That instance of M helps express a fact about the size of one set in particular: the set of sequences in which the limiting relative frequency of heads approximates the probability of heads. And since that set is so big, the property corresponding to that set is typical in the set of all possible infinite sequences of coin flips.

One might object to Typicality First as follows. There are many different ways of measuring the sizes of infinite sets. In the case of Ω, for instance, there are many different measures which assign sizes to the sets which Ω contains: M is one such measure, but there are others. So why think that M is the right measure to use, for describing the sizes of the subsets of Ω? Why does M do a better job of describing those sets' sizes than other measures do? Absent any answers to these questions, one might claim that M should not be interpreted as a way of quantifying size at all.

Let me offer two different responses to this objection. The first response appeals to the method of arbitrary functions, and in particular, the physical dynamics which underwrite the coin toss.[9] Basically, we should think that M is the right measure to use, for describing the sizes of the subsets of Ω when modelling the particular coin in question, because M is generated by basically any reasonable mathematical probability function over the degrees of freedom which are most relevant for the tossing of the coin. Specifically, take an arbitrary mathematical probability function which satisfies some fairly minimal constraints. Given

[9] Many thanks to a reviewer for drawing my attention to this response.

some strong but—for present purposes—entirely natural idealizations about what counts as a fair coin toss (Keller, 1986), this mathematical probability function evolves, under the dynamics, to a distribution which agrees with M. That is why, when modelling the tossing of this particular coin, M does a better job of describing the sizes of various subsets of Ω than other measures do.[10]

The second response appeals to posits about the sizes of various sets of physical possibilities. Basically, we should think that M is the right measure to use, for describing the sizes of the subsets of Ω when modelling the particular coin in question, because there are objective physical facts about the physical sizes of various sets of physically possible sequences of flipping that coin, and M is the mathematical measure which correctly captures those facts. Specifically, for each coin, there exists (i) a corresponding set of all physically possible sequences of the outcomes of flipping that coin infinitely many times, and (ii) a corresponding physical measure which quantifies the sizes of various subsets of that set. For every coin, one and the same mathematical set represents that coin's corresponding set of physically possible sequences: namely, Ω. But for different coins, different mathematical measures represent those coins' different corresponding physical measures: some coins' corresponding physical measures are best represented by M, while other coins' corresponding physical measures—like the unfair coin discussed at the end of Section 4—are best represented by $M_{\frac{2}{3}}$.[11] That is why, when modelling the tossing of this particular coin, M does a better job of describing the sizes of various subsets of Ω than other measures do: there is a physical measure—which figures in the physical facts about the sizes of various subsets of the set of all physically

[10]For accessible and thorough discussions of the method of arbitrary functions, see (Strevens, 2013, pp. 53-59; Suárez, 2020, pp. 51-61). For arguments against the claim that the method of arbitrary functions helps show that dynamics make various probabilities objective, see (de Canson, 2022).

[11]This does not presuppose that the coin in question, whose physical measure is best represented by M, is fair. Rather, this only presupposes that M is the right mathematical measure to use, when quantifying the physical measure associated with the coin in question; fair coins can then be defined, subsequently, as coins whose physical measures are all best represented by the particular mathematical measure M. Note that in all this, no appeal was made to the law of large numbers; it was not assumed that a fair coin is one for which the law of large numbers implies that the coin lands heads with a probability of $\frac{1}{2}$. Rather, in all this, appeals were made only to physical possibilities, physical measures over those possibilities, and certain mathematical measures best representing those physical measures.

possible sequences of the outcomes of flipping this particular coin—which best corresponds to the mathematical measure M.

Each response has its own attractive features and unattractive features. The first response is attractively non-committal, since it makes very few posits about what sorts of facts exist: it addresses the objection by relying solely, or at least primarily, on facts about the underlying dynamics. But the first response is unattractively limited, since it only applies to chancy systems which satisfy the assumptions of the method of arbitrary functions. The second response is attractively unlimited, since it applies to basically all chancy systems whatsoever, including those which fail to satisfy the assumptions just mentioned. But the second response is unattractively committal, since it makes several additional posits about what sorts of facts exist: it addresses the objection by relying, in large part, on posited facts about the physical sizes of various sets of physical possibilities. Ultimately, the reader is welcome to adopt whichever response they prefer – just note that adopting the first response amounts to restricting Typicality First to systems which satisfy the assumptions of the method of arbitrary functions.

Confusion dispelled: related to all this, do not claim that the above two responses contradict the point—from the end of Section 4—that in examples featuring unfair coins, other measures should be used to quantify the sizes of sets. When subsets of Ω are being used to model physical systems, there is no one, objective, system-independent fact about what the 'right' size of this-or-that subset of Ω is. The right size depends on the details of the physical system being modelled, and those details can vary from one physical system to the next. If the physical system features the particular coin being discussed, for instance, then the measure which correctly captures the sizes of the relevant sets is M. If the physical system features the unfair coin discussed at the end of Section 4, in contrast, then the measure which correctly captures the sizes of the relevant sets is $M_{\frac{2}{3}}$. The physical facts—about the underlying dynamics, or about the physical sizes of various sets of physical possibilities, or whatever—determine which measure well-models the physical system at issue.

To summarize: Typicality First holds because it respects the standard interpretation of mathematical measures as quantifying the sizes of sets. In fact, when supplemented with the account of typicality from Section 2, the standard interpretation of mathematical measures implies Typicality First. So Typicality First is quite plausible.[12]

6 The Stand-In View and Kolmogorov's Axiomatization of Probability

In this section, I present a particular view—call it the 'stand-in view'—of the relationship between size and probability. The stand-in view helps address an objection to, and provide general support for, the intuitive interpretation of mathematical measures. In addition, the stand-in view follows from an elegant account of exactly what Kolmogorov achieved, in his axiomatization of probability theory in terms of the theory of measures.

By way of preparation, it is worth making an observation about the measure M in (LLN). The observation is that in (LLN), M plays two different roles, one explicit and one implicit. The explicit role is straightforward: M assigns a measure to a certain set of possible infinite sequences of coin flips.

The implicit role is somewhat more complicated: the probability in (LLN), namely $\frac{1}{2}$, is defined in terms of M. To see how, start by noting that the probability of flip i landing heads—which, because of various assumptions, is the same probability for all i whatsoever—is what the quantity $\frac{1}{2}$, in (LLN), represents. Then recall that as mentioned earlier, the probability of flip i landing heads equals the expected value of H_i, and as a routine calculation shows,[13] that expected value equals $M(H_i)$. Therefore, $M(H_i)$ equals the probability of flip i landing heads: in other words, $M(H_i) = \frac{1}{2}$. So the probability $\frac{1}{2}$, which (LLN) features, is

[12]There might be a way to amend this argument, so that it supports Probability First rather than Typicality First. Perhaps certain probability facts can be interpreted as stand-ins for size facts – in much the way that, as discussed in Section 6, certain size facts can be interpreted as stand-ins for probability facts. Or perhaps the stand-in view discussed in Section 6 can be wielded, directly, to argue for Probability First. That is definitely worth exploring in future work; for lack of space, however, I do not explore it here.

[13]See footnote 5.

defined in terms of M.

One might use all these facts, concerning the implicit role which M plays, to formulate an objection to the intuitive interpretation of mathematical measures. For according to the intuitive interpretation, mathematical measures quantify the sizes of sets. So the intuitive interpretation implies that M assigns sizes to various sets used to compute the probabilistic quantity—namely, $\frac{1}{2}$—that (LLN) features. But as was just observed, in its implicit role, M is used to express a probability, not a size. And so, one might claim, the intuitive interpretation of mathematical measures is false in this case. In (LLN), M is not always being used to quantify the sizes of various sets: instead, in the derivation of the quantity $\frac{1}{2}$ which (LLN) features, M is being used to express facts about the probabilities of events.

To defend the intuitive interpretation of mathematical measures against this objection, I propose the 'stand-in' view of size and probability. The stand-in view is this: measure-theoretic facts about the sizes of sets often function as stand-ins for facts about the probabilities of events. So in accord with the intuitive interpretation, mathematical measures always express facts about size. But sometimes, those size facts function as stand-ins for probability facts. And that is how mathematical measures—which, according to the intuitive interpretation, express size facts—can also be used to express probabilistic quantities: by expressing facts about size which act as surrogates for facts about probabilities.

Here is an illustration of the stand-in view. In its implicit role in (LLN), M helps express a bunch of different size facts. For instance, in the equation $M(H_i) = \frac{1}{2}$ which is used to derive the probabilistic quantity that (LLN) features, M helps express the fact that $\frac{1}{2}$ is the size of H_i. According to the stand-in view, that fact about size serves as a stand-in for a fact about probability. In particular, the fact that H_i has size $\frac{1}{2}$ serves as a stand-in for the fact that $\frac{1}{2}$ is the probability of flip i landing heads. So given the stand-in view, there is no problem, here, for the intuitive interpretation of mathematical measures. In its implicit role, M helps express a size fact; and that size fact serves as a stand-in for the fact about the probabilistic quantity of $\frac{1}{2}$ which (LLN) invokes.

The stand-in view follows from a particularly elegant way of thinking about exactly what Kolmogorov achieved, in his measure-theoretic axiomatization of probability. On this way of understanding the achievement, Kolmogorov showed that the resources of the theory of size—namely, the resources of measure theory—can be used to formalize the basic axioms to which probabilities conform.[14] This achievement of Kolmogorov's was facilitated by the simple and beautiful fact that sizes can be stand-ins for probabilities.

An analogy will illustrate the idea here.[15] Consider two different interpretations of how set theory can be used to describe the natural numbers. According to the first interpretation, natural numbers and sets are numerically identical, and so facts about natural numbers are numerically identical to facts about sets. Perhaps 0 is numerically identical to $\varnothing$, 1 is numerically identical to $\{\varnothing\}$, and so on; and correlatively, perhaps facts about addition are numerically identical to facts about certain set-theoretic operations. According to the second interpretation, certain sets are stand-ins for natural numbers, and so certain facts about sets serve as stand-ins for various corresponding number-theoretic facts. Perhaps $\varnothing$ is a stand-in for 0, $\{\varnothing\}$ is a stand-in for 1, and so on; and correlatively, perhaps facts about certain set-theoretic operations are stand-ins for facts about addition.

The second interpretation is analogous to the stand-in view of size. Both use a certain mathematical formalism to express facts which function as stand-ins for various other facts. In the case of the second interpretation, the mathematical formalism is set theory, and the facts expressed in that formalism function as stand-ins for facts about natural numbers. In the case of the stand-in view, the mathematical formalism is measure theory, and the facts expressed in that formalism—which are facts about size—function as stand-ins for facts about probabilities.

Kolmogorov's achievement was, of course, impressive. Its striking simplicity and ele-

[14]Kolmogorov formalizes probability in the following way. Let E be a set of elementary events. Let $\mathfrak{F}$ be a field of subsets of E; and suppose E is in $\mathfrak{F}$. Then a probability function P is a function from $\mathfrak{F}$ to the non-negative reals such that (i) $P(E) = 1$, and (ii) for all disjoint sets A and B, $P(A \cup B) = P(A) + P(B)$ (Kolmogorov, 1933/1950, p. 2).

[15]Thanks to an anonymous reviewer for pointing out this analogy.

gance, however, often leads people to make an interpretive mistake. Instead of taking facts about size to be *stand-ins* for facts about probabilities, one might be tempted to mistakenly *identify* the size facts with the probability facts. This is, basically, what proponents of Probability First do. Since the probabilistic quantity $\frac{1}{2}$ in (LLN) is defined in terms of M, they insist on interpreting M as conveying probabilistic information. They insist on interpreting (LLN), and the explicit instance of M in that equation, as being about the extremely high probability of any given sequence having a limiting relative frequency which equals a probability. But the M in that equation, like all measures, expresses a fact about size: it says that the size of a certain set is extremely large. And facts about large sizes are typicality facts, not probability facts.

It is worth pointing out that Kolmogorov himself may have endorsed the stand-in view. For Kolmogorov begins his famous monograph on probability with the following passage.

> "[The task of giving an axiomatic foundation for the theory of probability] would have been a rather hopeless one before the introduction of Lebesgue's theories of measure and integration. However, after Lebesgue's publication of his investigations, the *analogies* between measure of a set and probability of an event, and between integral of a function and mathematical expectation of a random variable, became apparent. These *analogies* allowed of further extensions; thus, for example, various properties of independent random variables were seen to be in *complete analogy* with the corresponding properties of orthogonal functions" (1933/1950, p. v, emphasis added).

Kolmogorov's use of the word 'analogy', when describing the relationship between measure theory and probability, suggests that he would endorse the stand-in view. For Kolmogorov does not say that measure-theoretic notions are numerically identical to probabilistic notions. Rather, Kolmogorov says that measure-theoretic notions and probabilistic notions bear lots of striking analogies to each other. And that is, basically, another way of stating the stand-in view: size facts can be stand-ins for probability facts, insofar as size-related notions—as formalized in measure theory—are analogous to probability-related notions.[16]

But none of my claims hang on this. It seems plausible that Kolmogorov would have been sympathetic to the stand-in view. But perhaps not. Regardless, the stand-in view still provides an attractive account of how size and probability relate. And the stand-in view helps show that mathematical measures, like M, can be used to express probability facts in equations like (LLN): those measures simply express those probability facts indirectly, by expressing size facts which function as stand-ins for those probability facts.

7 Conclusion

The law of large numbers expresses one of the most basic, and important, relationships between frequencies and probabilities. As I have argued, that relationship should be understood in terms of typicality. The law of large numbers says that typically, the probabilities approximate empirical frequencies.

Ultimately, the difference between Typicality First and Probability First concerns different physical interpretations of a mathematical measure. Typicality First interprets that measure in terms of typicality. Probability First interprets that measure in terms of probability.

In closing, let me emphasize that physical interpretations of mathematical measures are worth exploring in detail. Everyone agrees, of course, that mathematical measures are extremely useful when it comes to formulating physical theories. Unsurprisingly, there is less agreement about how mathematical measures in physical theories should be interpreted. But there is much to gain from discussing, and exploring, physical interpretations of mathematical measures. Such discussions can uncover subtle distinctions, such as the distinction between

[16] This is compatible with the view that measure theory is extremely powerful, and has other uses too: for instance, that measure theory can be used to ground physical probabilities in physical facts (Rosenthal, 2010; Strevens, 2008).

Typicality First and Probability First. And such discussions can lead to attractive views about semantics – like the stand-in view – or about various mathematical innovations – like Kolmogorov's measure-theoretic approach to probability. So it is worth studying, more closely, the physical interpretation of mathematical measures.

Acknowledgements

Thanks to Heather Demarest, Sheldon Goldstein, Dustin Lazarovici, Barry Loewer, the audience at CHPS, an editor at *The Philosophical Quarterly*, and two anonymous reviewers, for much helpful feedback and discussion.

References

Allori, V. (2022). The paradox of deterministic probabilities. *Inquiry.* doi:10.1080/0020174X.2022.2065530

Barrett, J. A. (2017). Typical worlds. *Studies in History and Philosophy of Modern Physics, 58*, 31–40.

Boltzmann, L. (2015). On the Relationship between the Second Fundamental Theorem of the Mechanical Theory of Heat and Probability Calculations Regarding the Conditions for Thermal Equilibrium (K. Sharp & F. Matschinsky, Trans.). *Entropy, 17*, 1971–2009. (Original work published 1877)

Bricmont, J. (2020). Probabilistic Explanations and the Derivation of Macroscopic Laws. In V. Allori (Ed.), *Statistical Mechanics and Scientific Explanation* (pp. 31–64). Singapore: World Scientific.

Berkovitz, J. (2015). The Propensity Interpretation of Probability. *Erkenntnis, 80*, 629–711.

Crane, H., & Wilhelm, I. (2020). The Logic of Typicality. In V. Allori (Ed.), *Statistical Mechanics and Scientific Explanation* (pp. 173–229). Singapore: World Scientific.

de Canson, C. (2022). Objectivity and the Method of Arbitrary Functions. *British Journal for the Philosophy of Science, 73*(3), 663–684.

de Finetti, B. (2017). *Theory of Probability.* West Sussex: Wiley. (Original work published 1970)

Dürr, D., Froemel, A., & Kolb, M. (2017). *Einführung in die Wahrscheinlichkeitstheorie als Theorie der Typizität.* New York, NY: Springer.

Dürr, D., Goldstein, S., & Zanghì, N. (1992). Quantum Equilibrium and the Origin of Absolute Uncertainty. *Journal of Statistical Physics, 67*(5/6), 843–907.

Dürr, D., & Lazarovici, D. (2020). *Understanding Quantum Mechanics.* New York, NY: Springer.

Dürr, D., & Teufel, S. (2009). *Bohmian Mechanics.* New York, NY: Springer.

Frigg, R. (2009). Typicality and the Approach to Equilibrium in Boltzmannian Statistical

Mechanics. *Philosophy of Science, 76*, 997–1008.

Frigg, R. & Werndl, C. (2012). Demystifying Typicality. *Philosophy of Science, 79*, 917–929.

Goldstein, S. (2001). Boltzmann's Approach to Statistical Mechanics. In J. Bricmont, D. Dürr, M. C. Galavotti, G. Ghirardi, F. Petruccione, & N. Zanghì (Eds.), *Chance in Physics* (pp. 39–54). Heidelberg: Springer.

Goldstein, S. (2012). Typicality and Notions of Probability in Physics. In Y. Ben-Menahem & M. Hemmo (Eds.), *Probability in Physics* (pp. 59–71). New York, NY: Springer-Verlag.

Goldstein, S., Lebowitz, J., Mastrodonato, C., Tumulka, R., & Zanghì, N. (2010). Normal typicality and von Neumann's quantum ergodic theorem. *Proceedings: Mathematical, Physical and Engineering Sciences, 466* (2123), 3203–3224.

Gut, A. (2005). *Probability.* New York, NY: Springer.

Hawkins, T. (2002). *Lebesgue's Theory of Integration* (2nd ed.). Providence, RI: American Mathematical Society.

Hubert, M. (2021). Reviving Frequentism. *Synthese, 199*, 5255–5284.

Keller, J. (1986). The probability of heads. *American Mathematical Monthly, 93*, 191–197.

Kolmogorov, A. (1950). *Foundations of the Theory of Probability* (N. Morrison, Trans.). New York, NY: Chelsea Publishing Company. (Original work published 1933)

Lazarovici, D., & Reichert, P. (2015). Typicality, Irreversibility and the Status of Macroscopic Laws. *Erkenntnis, 80*, 689–716.

Lewis, D. (1980). A Subjectivist's Guide to Objective Chance. In W. L. Harper, R. Stalnaker, & G. Pearce (Eds.), *Ifs* (pp. 267–297). Boston, MA: D. Reidel.

Maudlin, T. (2011). Three Roads to Objective Probability. In C. Beisbart & S. Hartmann (Eds.), *Probabilities in Physics* (pp. 293–319). New York, NY: Oxford University Press.

Maudlin, T. (2020). The Grammar of Typicality. In V. Allori (Ed.), *Statistical Mechanics and Scientific Explanation* (pp. 231–251). Singapore: World Scientific.

Norsen, T. (2018). On the Explanation of Born-Rule Statistics in the de Broglie-Bohm

Pilot-Wave Theory. *Entropy, 20*(422), 1–26.

Paunić, D. (2002). History of Measure Theory. In E. Pap (Ed.), *Handbook of Measure Theory* (pp. 1–26). Amsterdam: Elsevier.

Pitowsky, I. (2012). Typicality and the Role of the Lebesgue Measure in Statistical Mechanics. In Y. Ben-Menahem & M. Hemmo (Eds.), *Probability in Physics* (pp. 41–58). New York, NY: Springer.

Rosenthal, J. (2010). The natural-range conception of probability. In G. Ernst & A. Hüttemann (Eds.), *Time, Chance and Reduction* (pp. 71–91). New York, NY: Cambridge University Press.

Stillwell, J. (2010). *Mathematics and Its History.* New York, NY: Springer.

Strevens, M. (2008). *Depth.* Cambridge, MA: Harvard University Press.

Strevens, M. (2013). *Tychomancy.* Cambridge, MA: Harvard University Press.

Suárez, M. (2020). *Philosophy of Probability and Statistical Modelling.* Cambridge: Cambridge University Press.

Tao, T. (2011). *An Introduction to Measure Theory.* Providence, RI: American Mathematical Society.

Valentini, A. (2020). Foundations of Statistical Mechanics and the Status of the Born Rule in de Broglie-Bohm Pilot-Wave Theory. In V. Allori (Ed.), *Statistical Mechanics and Scientific Explanation* (pp. 423–477). Singapore: World Scientific.

Wilhelm, I. (2022). Typical. *The British Journal for the Philosophy of Science, 73*, 561–581.

Wilhelm, I. (forthcoming). The Typical Principle. *The British Journal for the Philosophy of Science.*