

Zero Probability

(Unpublished)

Dan David November

Abstract

In probability textbooks, it is widely claimed that zero probability does not mean impossibility. But what stands behind this claim? In this paper I offer an explanation to this claim based on Kolmogorov's formalism. As such, this explanation is relevant to all interpretations of Kolmogorov's probability theory. I start by clarifying that this claim refers only to nonempty events, since empty events are always considered as impossible. Then, I offer the following three reasons for the claim that nonempty events with zero probability are considered as possible: The main reason is simply because they are nonempty. Hence, they are considered as possible despite their zero probability. The second reason is that sometimes the zero probability is taken to be an approximation of some infinitesimal probability value. Such a value is strictly positive and as such does not imply impossibility in a strict sense. Finally, the third reason is that there are interpretations according to which the same event can have different probabilities. Specifically, it is assumed that an event with exactly zero probability (that does not approximate an infinitesimal value) can have strictly positive probabilities. This means that such an event can be possible which implies that its zero probability does not mean impossibility.

1. The main claims

It is commonly held that zero probability does not mean impossibility. For example, when Kolmogorov discusses the relation of his probability theory to experimental data, he writes that: " $P(A)=0$ does not imply the impossibility of A ." (Kolmogorov, 1933, p. 5). And Jeffreys in his classical textbook "Theory of Probability" asserts that: "[...] a proposition can have probability 0 and yet be possible [...]" (Jeffreys, 1998, p. 65). Von Plato claims that this idea, that zero probability does not mean impossibility, was around even before Kolmogorov formalized his theory. To support his claim, he quotes Poincaré from his 1896 book on the calculus of probability: "[...] with an infinity of possible results, probability 0 does not always mean impossibility, and probability 1 not certainty." (Von Plato, 1994, p. 7). But what is the reason behind the claim that zero probability does not mean impossibility? Interestingly, according to Hájek this question not been addressed explicitly in the literature: "Indeed, given how many probability textbooks go out of their way to caution the reader that 'probability 0 does not imply impossible', it is perhaps surprising that more is not made of it, at least in philosophical circles." (Hájek, 2003, pp. 285–6). In this paper I try to take on Hájek's challenge and explain the reasons for this claim. The explanation I offer is based solely on Kolmogorov's formalism and hence relevant to all interpretations of Kolmogorov's probability theory.

Explaining the claim that zero probability does not mean impossibility is important mainly because it fills a lacuna in the field of philosophy of probability. But it is also important because it reveals new insights on the connection between Kolmogorov's theory and its interpretations. Specifically, my explanation reveals that Kolmogorov's formalism can express different senses of 'possibility'. These senses are relevant to all its interpretations but are generally not explicitly addressed by them. My explanation also reveals that the claim that the same event can have different probabilities, is in fact interpretation-dependent. Thus, it does not stem from Kolmogorov's theory. This means that it cannot be mathematically justified or refuted based solely on Kolmogorov's theory. This conclusion has important implications on the Bayesian framework which I do not address in this paper.

In this paper I offer an explanation to the claim that zero probability does not mean impossibility, based on analyzing the structure of Kolmogorov's probability space. I begin (in section 2) with a brief description of the connection between Kolmogorov's theory and its interpretations. I continue (in section 3) with a reminder of Kolmogorov's definition of the probability space. Then (in section 4) I discuss the first sense of 'possibility' expressed by nonempty events in Kolmogorov's formalism. The second sense (which I discuss in section 5) is expressed by events having strictly positive probabilities. I then distinguish between two cases that are expressed by zero probability in Kolmogorov's formalism: infinitesimal probabilities (discussed in section 6) and exactly zero probability (section 7). The latter case raises questions regarding equivalence relations between events in Kolmogorov's theory (which I address in section 8). Finally, I conclude.

2. Interpretations of probability and Kolmogorov's probability theory

"Interpretations of probability"¹ is a term commonly used to describe all the different theories which deal mainly with the notions: 'event' and 'probability'. These theories express the large disagreement that exists among philosophers regarding these notions. The interpretations give different answers to the question: "what are 'event' and 'probability'?". A very important aspect of the required answer is the mathematical part of the definitions of 'event' and 'probability'. This part is standardly (though not universally) given by Kolmogorov's probability theory. According to Gillies: "The theory of probability has a mathematical aspect and a foundational or philosophical aspect. There is a remarkable contrast between the two. While an almost complete consensus and agreement exists about the mathematics, there is a wide divergence of opinions about the philosophy." (Gillies, 2000, p. 1).

Mathematical probability theories (such as Kolmogorov's) are treated as distinct from their interpretations. According to Lyon, the two types of theories aim to answer different questions regarding the notion of probability: "In philosophy of probability, there are two main questions that we are concerned with. The first question is: what is the correct mathematical theory of

¹ For good surveys of the interpretations of probability see: (Gillies, 2000; Hájek, 2012) and (Von Plato, 1994) for a more historical perspective.

probability? [...] These formal theories of probability tell us how probabilities behave, how to calculate probabilities from other probabilities, but they do not tell us what probabilities *are*. This leads us to the second central question in philosophy of probability: just what are probabilities? [...] philosophers have tried to answer this question. Such answers are typically called *interpretations of probability*, or philosophical theories of probability." (Lyon, 2010, p. 93). In other words, mathematical probability theories provide the mathematical parts of the definitions of 'event' and 'probability' while their interpretations describe the non-mathematical (interpretive) parts. This dichotomy is not entirely accurate, because the mathematical and the interpretive parts are connected and affect each other. The mathematical parts have implications for the interpretive parts (such as imposing restrictions on them) and vice versa. This means that Kolmogorov's theory *explicitly* deals with the mathematical parts, but also *implicitly* deals (at least partially) with the interpretive parts. The interpretive aspects which are implicitly described by Kolmogorov's theory are common to all interpretations of his theory, simply because these aspects stem from the restrictions imposed by the mathematical parts. Realizing this is important for my explanation since its focus is on the interpretive features which stem from Kolmogorov's theory. This focus is what makes my explanation relevant to all interpretations of Kolmogorov's theory.

A "complete" probability theory (a theory that is composed out of a mathematical probability theory and an interpretation of it) describes what I call: "probabilistic states". In other words, anything describable by a probability theory is a probabilistic state. According to Gyenis and Rédei: "Interpretations of probability are typical classes of applications of probability theory, classes consisting of applications that possess some common features, which the interpretation isolates and analyses." (Gyenis & Rédei, 2014, p. 19). Thus, it can be said that each of the interpretations of Kolmogorov's theory tries to provide in non-mathematical terms a general description of a collection of probabilistic states describable by his theory.

The main characteristic of a probabilistic state is that it involves at least two different possibilities (or alternatives, or options). For example, a probabilistic state of an ideal coin toss involves the possibility of the coin landing on 'heads' and the possibility of it landing on 'tails'. Similarly, a probabilistic state of an agent picking an integer from the set $\{0,5,2,7,8,11\}$ includes each of the six options of picking one of these integers and perhaps also the option of not picking any of them.

A state which includes only one possibility is normally not considered a probabilistic one². The possibilities of a probabilistic state are mutually exclusive and exhaustive, which means that there are no other alternatives which are relevant to it. Moreover, it is assumed that one of these possibilities necessarily occurs (or manifests, or happens). For example, an ideal coin toss *does not* include the possibility of the coin landing on its side (and hence landing not on 'heads' or 'tails') - this option is simply not part of the ideal state. It is important to clarify that possibilities are not events! Thus, a possibility does not have a probability. (Contrary to an event which is a set of possibilities that does have a certain probability.) This clarification is important for a claim which I will make later in this paper, that a nonempty event is considered as possible simply because it contains possibilities. In the meantime, this crude characterization of probabilistic states seems to capture the common feature of all the things describable by Kolmogorov's theory.

Here I would like to mention that Kolmogorov's theory has reached a status of orthodoxy despite having known objections to it³. It is largely preferred over other mathematical probability theories which have different and not necessarily equivalent formal definitions of 'event' and 'probability'⁴. Hence Kolmogorov's theory is almost universally treated to as *the* (mathematical) probability theory. As a result, the different interpretations are commonly thought of as interpretations of Kolmogorov's theory⁵. The fact that Kolmogorov's formalization of 'event' and 'probability' is so widely accepted is the reason I focus on it.

In the next section I present a definition of Kolmogorov's probability space. I highlight the points that are relevant to my explanation of the claim that zero probability does not mean impossibility.

² Interestingly, Kolmogorov's probability space can in fact describe probabilistic states which include only one possibility. However, such a state, by definition, has exactly one event with probability 1 and hence does not seem to mathematically describe anything that is normally considered probabilistic.

³ See Lyon (2010) for a general discussion on the problems Kolmogorov's theory has in relation to the different interpretations. Also see Lyon (2016) for a more detailed discussion on some of the possible objections to Kolmogorov's axioms. And see Hájek (2003) for an objection to the definition of conditional probability which is a specific part of Kolmogorov's theory.

⁴ See for example: Goossens (1979); Popper (1938, 1955, 1959, Chapter 8); Rényi (1955).

⁵ The main interpretations of probability are in fact *not* interpretations of Kolmogorov's theory in a strict sense. See Hájek (2012) and especially Lyon (2016) on this issue.

3. The definition of the probability space

My explanation of the claim that zero probability does not mean impossibility is based on Kolmogorov's probability space's structure. In this section I give its definition⁶ and emphasize some aspects which are important for my explanation.

A probability space is defined as a triple $\langle \Omega, \Sigma, P \rangle$ consisting of the following three components: a sample space (Ω), a σ -algebra (Σ) and a probability measure (P).

The probability space's components are defined as follows:

1. A sample space (Ω) - a nonempty set

The members of the sample space are sometimes called "elementary events" but nevertheless, they are *not* a mathematical formalization of 'events'. 'Events' are defined by the σ -algebra component (the definition of which will be given shortly). However, these members are indeed elementary in the sense that 'events' and 'probabilities' depend on them. Moreover, these members formally describe the different possibilities (or alternatives, or options) which are the characterizing feature of every probabilistic state⁷. Recall that these possibilities are mutually exclusive and exhaustive so as the members of the sample space.

'Events' are formally defined by the σ -algebra component⁸:

2. A σ -algebra (Σ) (defined over the sample space) - a subset of the power set of the sample space (i.e. a set of subsets of Ω) which satisfies the following three conditions:
 - 2.1. Σ is not empty (or, equivalently, Ω is in Σ)
 - 2.2. Σ is closed under complementation (i.e. if A is in Σ then so is $\Omega \setminus A$)

⁶ The definition of Kolmogorov's probability space presented in this paper, is one of several standard ways to define it. See Billingsley (1995, p. 23) for a similar yet more rigorous definition.

⁷ See November (2018, Chapter 1) for a detailed analysis of the interpretive meaning of the sample space component of Kolmogorov's probability space.

⁸ More accurately, since the sample space is necessary for the σ -algebra's definition (the σ -algebra is defined over a given sample space), 'events' are formally defined by both components together.

2.3. Σ is closed under countable unions (i.e. if $A_1, A_2, A_3 \dots$ are in Σ , then so is $A = A_1 \cup A_2 \cup A_3 \cup \dots$)

The members of Σ are the mathematical definition of 'events'. Which means that mathematically 'events' are sets that together form a σ -algebra. Since the σ -algebra is defined over a specific sample space, mathematically 'events' are sets of members of a sample space. Hence, loosely speaking, an event (A) can be thought of as a collection of some of the possibilities which constitute a given probabilistic state. The happening (or occurrence, or manifestation) of any one of A 's possibilities is the happening (or occurrence, etc.) of A . Similarly, the complementary event of event A (A^c) is a collection of all the possibilities which are not in A . Hence, A^c can be thought of as the case that A does not happen. Thus, when a possibility happens (or manifests, etc.) it is either in A , which means that A happens, or in A^c (and hence not in A) which means that A does not happen.

Notice that the σ -algebra always contains the sample space event and the empty event (i.e. $\emptyset, \Omega \in \Sigma$). Hence these events can be called the "mandatory events". Another important point to notice is that the σ -algebra is also connected to a specific probability measure. Realizing this is important for understanding the relation between events and probabilities. As I will show in sections 7 and 8, it is especially important for defining an identity relation between events.

3. A probability measure (P) - a real valued function defined over Σ which satisfies the following conditions:

3.1. P is non-negative

3.2. $P(\emptyset) = 0$

3.3. P is countably additive (which means that for all countable collections $\{E_i\}$ of pairwise disjoint sets, $P(\cup_i E_i) = \sum_i P(E_i)$)

3.4. P returns values in the unit interval $[0,1]$ and $P(\Omega) = 1$

The values assigned to the members of the σ -algebra by the probability-measure function are the mathematical definition of 'probabilities' in Kolmogorov's theory. In other words, mathematically 'probabilities' are the values of a function from a σ -algebra of a given probability space to the unit interval which satisfies certain conditions (described by the definition of the probability measure).

Two points regarding the definition of the probability measure are important for my explanation of the claim that zero probability does not mean impossibility. The first concerns the probability values of the two mandatory events. By definition, the probability of the empty event is 0 and the probability of the sample space event is 1. Each of these events cannot have any other probability value. This fact about these events distinguishes them from other events. Any event, which is not a mandatory event, can have any probability value assigned to it by the probability measure function⁹. This is because the same σ -algebra can have infinitely many probability measures defined over it¹⁰. Thus, all events necessarily have probabilities, but only the mandatory events necessarily have *particular* probability values (the empty event necessarily has 0 probability and the sample space event necessarily has 1). All other events can have any probability value in the range $[0,1]$. The second point is that probabilities are *real* numbers between 0 and 1 inclusively. This means that according to this definition, probability values are not, and cannot be, surreal numbers. And more specifically they cannot be infinitesimals¹¹. This fact is important for cases of supposedly infinitesimal probability. Loosely speaking, in such cases, when an event with infinitesimal probability is described by Kolmogorov theory, it is assigned zero probability by the probability measure function. The zero probability approximates its infinitesimal probability value¹². I elaborate this point in section 6.

In the literature of philosophy of probability, it is quite common to find other mathematical definitions of 'event' and 'probability' which seem to be different from Kolmogorov's definition of the probability space. Specifically, 'events' are commonly mathematically defined using an algebra without the explicit mention of a sample space. However, since an algebra is a type of an algebraic structure, it is in fact a set with operations defined on it. This means that an algebra *always* has an underlying set which plays the same role as the sample space plays in Kolmogorov's definition

⁹ Assuming that the probabilities of all the events together satisfy the definition of the probability measure.

¹⁰ Except for when the σ -algebra is a trivial one (a σ -algebra that contains only the two mandatory events). The only probability measure that can be defined over a trivial σ -algebra is its corresponding trivial probability measure which assigns 1 to the sample space event and 0 to the empty event.

¹¹ Probabilities also cannot be numbers which involve infinitesimals in their definitions, such as probabilities of the form $p = (r \pm \varepsilon)$, $0 \leq p \leq 1$, when r is a real number in the interval $[0,1]$ and ε is an infinitesimal.

¹² In the domain of the reals, zero can approximate an infinitesimal similarly to how a rational number in the domain of the rational numbers can approximate a real number. For example, $3\frac{14}{100}$ is an approximation of π .

(see Gilbert & Nicholson (2004, p. 4)). Thus, the common way of mathematically defining 'events' and 'probabilities' using only an algebra and a probability measure, is just a partial description of the relevant Kolmogorovian probability space because it lacks an explicit mention of the sample space. Realizing this is important since it means that the explanation I present in this paper is relevant even to those who mathematically define 'events' without the explicit mention of a sample space.

A word about notation, in the rest of this paper I use "event" and "probability" (without inverted commas) to denote the two notions 'event' and 'probability' that have both a mathematical part and an interpretive part¹³. I use "k-event" and "k-probability" (without inverted commas) to denote the mathematical parts of these notions given by Kolmogorov's probability theory. In other words, given a probability space (S), a k-event is a member of the σ -algebra of S and a k-probability is a value of the probability measure of S .

In the following sections I discuss two different senses of 'possibility' that are mathematically expressed by Kolmogorov's probability space. I would like to clarify that these senses of 'possibility' are not the senses (or kinds) of possibility that appear in discussions about modality. In such discussions it is common to distinguish for example between logical possibility, physical possibility, metaphysical possibility and other kinds of possibility¹⁴. It is possible (no pun intended) to try and identify such senses of 'possibility' with those expressed by Kolmogorov's probability space, but since it is not relevant for my explanation, it is not part of this paper.

4. The first sense of possibility – being a nonempty event

¹³ An 'event' can be a state of the world (the actual world or a counterfactual one), a proposition or a sentence in a formal language etc. and a 'probability' can be either chance, propensity, credence or degree of belief etc. depending on one's choice of interpretation of probability.

¹⁴ See Kment (2017) on this topic.

The first sense of possibility expressed by Kolmogorov's probability space stems from the definition of k-events as sets. Events that are mathematically described by nonempty k-events (hereinafter, nonempty events) are commonly considered as possible. While events that are mathematically described by the empty k-event (hereinafter, empty events) are *always* considered as impossible¹⁵. Roughly, being a nonempty event means that the event is a nonempty collection of the different possibilities (or alternatives, or options) which constitute a particular probabilistic state. And in that sense, the event is possible. In contrast to empty events which are empty collections of those possibilities and hence are impossible.

It is easy to understand this sense of possibility by analyzing the empty k-event. The empty k-event is an empty subset of a given sample space. Hence, any event that is correctly described by the empty k-event, loosely speaking, is an empty collection of the different possibilities (or options, etc.) which constitute a particular probabilistic state. In other words, given a probabilistic state, the empty event is the case that none of its possibilities occurs (or manifests, or happens). But since these possibilities are exhaustive, one of them *necessarily* occurs. Recall that this characteristic of a probabilistic state is an essential one. It is true for every probabilistic state. Thus, the case that none of the possibilities occurs, is not possible. To illustrate this point, recall the example of an ideal coin toss. It includes only two possibilities: the coin landing on 'heads' or 'tails'. The option that the coin does not land on either 'heads' or 'tails' simply does not exist in this probabilistic state. The empty event in this example does not include any of the possibilities which can occur in this particular probabilistic state and hence it is impossible. This is true for any empty event in any probabilistic state – they are simply impossible.

This point can be made in a different way. Recall that the empty k-event is the complementary k-event of the sample space k-event. The sample space k-event is composed out of all the members of a given sample space. It mathematically describes an event which is a collection of all the different possibilities (or alternatives, or options, etc.) which constitute a particular probabilistic state. In other words, given a probabilistic state S , the sample space k-event describes the event that any one of the possibilities which constitutes S occurs (or manifests, etc.). And since one of these possibilities necessarily occurs (there are no other possibilities), the event described by the

¹⁵ The empty k-event is even referred to as "an impossible event" by Kolmogorov himself, see Kolmogorov (1933, p. 5,6)

sample space k -event is certain to happen. As a result, its complementary event, the empty event, is certain *not* to happen. Recall that a complementary event of event A can be thought of as the case that A does not happen. Thus, the empty event can be thought of as the case that the sample space does not happen. But since the sample space event is certain to happen, the case that it does not happen is impossible. In other words, it is impossible for the empty event to happen. Which again brings us to the conclusion that any empty event is impossible. In short, the empty event is impossible in the sense that it does not correspond to any of the possibilities which constitute a particular probabilistic state. This means that any event that is not the empty event, i.e. any nonempty event, is possible in the sense that it does correspond to at least one of these possibilities.

The claim that zero probability does not mean impossibility does not refer to empty events. Empty events are always considered as impossible. Notice that an empty event is both empty *and* has probability zero. That is because it is described by the empty k -event which has zero probability by definition. Thus, empty events are impossible because they are empty *and* because they have zero probability.

In the next sections I analyze the senses of possibility expressed by the probabilities of the events. This will enable me to clarify why nonempty events with zero probability are commonly considered as possible. Very roughly, I will claim that they are considered as possible because they are nonempty and despite having zero probability.

5. The second sense of possibility – having a strictly positive probability

This short section concerns with the sense of possibility expressed by strictly positive probability values of events. Very roughly, events that have strictly positive probability values are considered as possible. The exact sense of possibility expressed by having a strictly positive probability depends on one's choice of interpretation of probability. For example, according to some objective

interpretations, having a strictly positive probability means that the relative frequency of that event in a series of trials is strictly positive. And according to some subjective interpretations, having a strictly positive probability means that the degree of belief of an agent in that event is strictly positive, etc. In general, according to each of the main interpretations of probability, an event that has a strictly positive probability is considered as possible in a sense determined by the interpretation.

It is sometimes claimed that events that have *low*, yet strictly positive probability values are not "practically" possible. For example: "[...] a sufficiently high probability can be considered 'practically certain,' and a sufficiently low correspondingly 'practically impossible.'" (Von Plato, 1994, p. 44). This claim is important; however, it relies on making a distinction between "practical" possibility and "theoretical" possibility or something of that sort. Such distinctions are based on some metaphysical assumptions which are not part of (and are not mathematically described by) Kolmogorov's formalism. These distinctions are closely related to the types of possibility which appear in the literature about modality. As such they are not discussed in this paper. The purpose of this paper is to explain why zero probability does not mean impossibility based on Kolmogorov's formalism without relying on additional assumptions which are not necessarily common to all its interpretations.

Events that have zero probability are also commonly considered as possible for several reasons which I discuss in the next sections.

6. Probability zero – the case of infinitesimal probability

Arguably¹⁶ there are probabilistic states in which there are events with infinitesimal probabilities. Such cases cannot be accurately described by Kolmogorov's probability space simply because k -probabilities values cannot be infinitesimals. Recall that k -probability values are real numbers in

¹⁶ I use "arguably" here since there are different objections to the idea of infinitesimal probabilities. See for example: Elga (2004); Williamson (2007).

the range $[0,1]$ and as such they are not infinitesimals. Hence, infinitesimal probabilities are either not described in Kolmogorov's theory or are described *inaccurately* (or approximately) by zero k-probability.

For example, a fair lottery over the natural numbers is often described as a case in which there are (or there should be) events with infinitesimal probabilities. For the lottery to be fair, all events corresponding to one result of the lottery (i.e. that the result is a natural number n) should have the same probability. The question is, what is this probability? It turns out that in Kolmogorov's framework answering this question is not simple. In fact, strictly speaking, this lottery cannot be described by Kolmogorov's theory. The reason is that the k-probability of all k-events describing these events must be a real number that is either zero or strictly positive. In the former case, the k-probability of their union is 0 and, in the latter, it is ∞ . Notice that their union is the sample space k-event (Ω). Hence, in both cases the k-probability of the sample space k-event is not 1 (it is either 0 or ∞) in contrast to the definition of the probability measure which asserts that $P(\Omega) = 1$. This seems to imply that according to Kolmogorov's theory there cannot be a fair lottery over the natural numbers. More generally, it is impossible to describe a uniform distribution over any *infinite* set of nonempty events using Kolmogorov's theory¹⁷. Thus, according to Kolmogorov's theory, infinitely many nonempty events cannot be uniformly distributed (they can only be nonuniformly distributed). To avoid this conclusion, some scholars have suggested that the range of Kolmogorov's probability measure should be changed to include infinitesimals. Their suggestion is based on the idea that in there are cases in which there are events with infinitesimal probabilities.

Event with infinitesimal probability that are mathematically described by k-events with zero k-probability are commonly considered as possible despite their mathematical description. For these events, the zero k-probability is not thought of as an *accurate* mathematical description of the described infinitesimal probability. The key point is that the infinitesimal probability of the described event is indeed small (infinitesimally so) but it is not zero! - it is strictly positive. And

¹⁷ The inability to describe a uniform distribution over any *infinite* set of nonempty events, is considered as a major drawback to Kolmogorov's formalism (see Lyon (2016)). Mainly because there does not seem to be an a priori reason for rejecting a uniform distribution of probabilities over an infinite set of events while accepting any nonuniform distribution (that satisfies the probability measure's definition). There are different attempts to amend this drawback. For example, see Gwiazda (2010) for a solution using "Asymptotic Density" (phrased as a response to this problem as it is presented in Armstrong & McCall (1989)), and see Wenmackers & Horsten (2012) for a solution using non-standard analysis.

since commonly an event with a strictly positive probability is not considered as impossible (but rather as possible) an event with infinitesimal probability is also not considered as impossible¹⁸.

In other words, when an event with infinitesimal probability is described by a k-event with zero k-probability, this description is taken to be inaccurate (or an approximation). Hence the zero k-probability is not taken to imply (or to indicate) that the described event is impossible. Notice that a nonempty event with an infinitesimal probability is considered as possible in at least two senses of possibility: it is nonempty, and it has a strictly positive probability. This is despite being formally described by a k-event with zero k-probability. Thus, the zero k-probability is not taken to mean impossibility because it is just a mathematical approximation of the event's infinitesimal probability. In other words, the zero k-probability is a mathematical description which approximates a probability value that is *not* zero. In the next section I discuss cases where the probability value is exactly zero.

7. Probability zero – the case of exactly zero probability

This section deals with nonempty events with zero probability. Pay attention that the probability of these events is *exactly* zero and not an infinitesimal number. Such events are mathematically described by nonempty k-events with zero k-probability. Since the probability of the described events is exactly zero, the zero k-probability that mathematically describes it, is an *accurate* description. In contrast to cases of infinitesimal probabilities when the zero k-probability only *approximates* these probabilities.

Nonempty events with zero probability are commonly considered as possible despite having zero probability. The reason is twofold. First, these events are nonempty and hence are considered as possible in the sense discussed in section 4 (briefly, a nonempty event is possible in the sense that it is a nonempty collection of possibilities which constitute a particular probabilistic state). The second reason is roughly that such events are thought of as *contingently* having zero probability

¹⁸ One plausible claim is that an event with an infinitesimal probability is *practically* impossible. But as I already mentioned above, such a claim relies on some distinction between practical and theoretical possibility which I do not address in this paper.

(as opposed to *necessarily* having it). Loosely speaking, even if a nonempty event with exactly zero probability is considered as impossible, it is not considered as *necessarily* impossible. Thus, it is possible (in a sense which will be discussed in the next sections) for such an event to be possible (in the sense discussed in section 5). This claim, that it is possible for such events to be possible, is based on the idea that the same event can have different probabilities. More accurately, the idea is that the same nonempty and non-sample-space event can have different probabilities. Recall that an empty event necessarily has probability 0 and a sample space event necessarily has probability 1, thus they cannot have other probabilities. This idea specifically implies that nonempty events with zero probability can have *strictly positive* probability values. This is important since events with strictly positive probabilities are commonly considered as possible according to each of the main interpretations of probability (when the exact sense of possibility is determined by the interpretation). Hence the claim that a nonempty event has zero probability only contingently, means that it can have strictly positive probability values. This implies that such an event *is* possible, loosely speaking.

The idea that the same event can have different probabilities is important for my explanation of the second reason why nonempty events with zero probability are possible. Simply because it is a necessary condition for the claim that these events have zero probability only contingently. Clearly, if the same event *cannot* have different probabilities, then a nonempty event with zero probability cannot have strictly positive probabilities. In this case, a nonempty event with zero probability can be considered as possible *only* in the sense discussed in section 4. In other words, if a nonempty event with zero probability cannot have strictly positive probabilities, it can be considered as possible only because it is nonempty and *despite* having zero probability.

Obviously, the idea that the same event can have different probabilities depends on the exact meaning of "same event". This idea assumes that the identity relation between events does *not* depend on their probabilities. Such a relation implies that the probability of an event is not one of its essential properties. More accurately, any event necessarily has *a* probability value, but not necessarily that *particular* probability value (unless it is an empty event or a sample space event that necessarily have 0 and 1 probability values respectively).

On the face of it, the claim that the probability of an event is not one of its essential properties seems quite reasonable. For example, it seems plausible to claim that the probability that a given

coin lands 'tails', can be any value in the range $[0,1]$. The probability can be $1/2$ when the coin is fair, and all relevant conditions are "normal" so to speak. Or it can be any value in the range $[0,1]$ when the coin is biased or if the conditions are "abnormal", etc. Hence, this example seems to show that the same event - a coin lands 'tails' - can have different probability values, including zero probability.

The key point behind the idea that the same event can have different probabilities is that the underlying identity relation between events does not depend on their probabilities. Notice that objecting to such identity relation amounts to rejecting the assumption that the probability of an event is not one of its essential properties. In other words, objecting to the idea that the same event can have different probabilities, means that each event necessarily has a *specific* probability value. Thus, any change in the event's probability would mean that the event itself has changed as well. This conclusion may pose a problem for some interpretations of probability. In other words, holding that the probability of an event is one of its essential properties might be problematic for some interpretations. Specifically, it does not seem to sit well with the Bayesian framework in which it is commonly said that the probability of a given event is updated in light of new evidence.

Identity relations of both types, those which depend on the events' probabilities and those which do not, can be justified in different ways. However, and this is the crucial point, none of these relations stems from Kolmogorov's mathematical definition of k-events! In other words, the definition of k-events is compatible with both the claim that the probability of an event is one of its essential properties and with its negation. The reason is roughly that events with different probabilities (whether they are considered as the same event with different probabilities or different events altogether) are mathematically described by different probability spaces. But in Kolmogorov's theory there is no explicit definition according to which two k-events are said to be two mathematical descriptions of the same event. In other words, there is no equivalence relation between k-events in the sense that they describe the same events. Thus, Kolmogorov's theory cannot support nor refute the claim that the probability of an event is one of its essential properties. I elaborate this point in the next section.

To sum up, nonempty events with exactly zero probability are commonly considered as possible for two reasons: The first is that they are nonempty and hence are possible in the sense discussed in section 4. The second reason is that the fact that they have zero probability is commonly thought of as contingent and not as a necessary fact. This means that such events can have strictly positive probability values. Loosely speaking, it is possible for such events to be possible in the sense discussed in section 5 and hence they are possible. This second reason is based on the idea that the same event can have different probabilities. In other words, the identity relation between events does not depend on their probabilities. However, this idea does not stem from Kolmogorov's definition of k-events! Moreover, it cannot be justified or refuted by their definition. This means that the second reason for claiming that nonempty events with zero probability are possible, is not based on Kolmogorov's formalism! It is possible to try and add to Kolmogorov's theory different equivalence relations between k-events which will express either the idea that the same event can have different probabilities or its negation. But in both these cases, such additions will not serve as a mathematical justification or refutation of this idea. Simply because they are additions to the theory and not an integral part of it. In the next section I discuss some important aspects concerning these putative equivalence relations.

8. Equivalence relations between k-events

In the previous section I mentioned two reasons for the claim that nonempty events with zero probability are possible. The first is that they are nonempty and the second is that they can have strictly positive probability values. The second reason means that such events are possible because, loosely speaking, they can be possible. This explanation rests on the idea that the same event can have different probabilities. However, this idea does not stem from Kolmogorov's definition of k-events. Moreover, it is neither justified nor refuted by Kolmogorov's theory. Kolmogorov's theory does not include an explicit definition of when two k-events are said to be "equivalent" in the sense that they mathematically describe the same event(s). In other words, there is no explicit equivalence relation between k-events in Kolmogorov's theory. Such a relation is necessary for

one to mathematically justify or refute the idea that the same event can have different probability values.

In this section I demonstrate that Kolmogorov's theory can accommodate both types of equivalence relations between k-events: those which support the claim that the same event can have different probabilities, and those which undermine it. I explain that deciding between these types cannot be done based solely on Kolmogorov's theory. Thus, the answer to the question whether the same event can have different probabilities, turns out to be interpretation-dependent. This means that the claim that nonempty events have zero probability only contingently, is interpretation-dependent.

In Kolmogorov's theory there is no explicit equivalence relation between k-events in the sense that they mathematically describe the same event. This means that there is no way of determining whether two different k-events mathematically describe the same events or not. Part of the problem is that there is no explicit (and non-trivial) identity relation between k-events¹⁹. Thus, there is no explicit way of determining whether two k-events are the same or not. Specifically, the set-theoretic identity relation, is not an adequate equivalence relation between k-events (and is also not an identity relation between them).

Recall that according to their definition, k-events are sets which together compose a σ -algebra. Thus, they can be compared as sets. According to the set-theoretic identity relation, two k-events are identical if and only if they have the same members. However, this criterion is insufficient for determining that two k-events are equivalent in the sense that they describe the same event.

For example, according to this relation, the k-event $k_a = \{1\}$ which is part of the σ -algebra $\Sigma_{2,1} = \{\emptyset, \Omega_2, \{1\}, \{2\}\}$ defined over the sample space $\Omega_{2,1} = \{1,2\}$ is identical to the k-event $k_b = \{1\}$ which is part of the σ -algebra $\Sigma_{\mathbb{P}\mathbb{N}} = P(\mathbb{N})$ (i.e. $\Sigma_{\mathbb{P}\mathbb{N}}$ is the power set of the natural numbers) defined over the sample space $\Omega_{\mathbb{N}} = \mathbb{N}$ (where \mathbb{N} is the set of natural numbers). However, despite being identical in set-theoretical terms, the k-events k_a and k_b are commonly not considered as equivalent mathematical descriptions of the same events. On the other hand, k_a and the k-event

¹⁹ According to the trivial identity relation between k-events, given a probability space, each of its k-events is identical only to itself. This means specifically that k-events from different probability spaces are never the same k-event.

$k_c = \{3\}$ which is part of the σ -algebra $\Sigma_{2,2} = \{\emptyset, \Omega_{2,2}, \{3\}, \{4\}\}$ defined over the sample space $\Omega_{2,2} = \{3,4\}$ are commonly considered as two mathematical descriptions of the same events despite being set-theoretically different.

These examples show that the set-theoretic identity relation is not suitable to be an equivalence relation between k-events. Two k-events can be set-theoretically identical and still be considered as non-equivalent and vice versa: two k-events can be considered as equivalent and yet be set-theoretically different. The reason is roughly that k-events are defined as part of given probability spaces, but the set-theoretical identity relation is oblivious to that. More specifically, given two k-events, the structures of the σ -algebras to which they belong, seem to be necessary for determining whether they are equivalent or not.

For example, the following two probability spaces contain set-theoretically different k-events. However, these probability spaces seem to be two different mathematical descriptions of the same probabilistic states. They seem to have the same structure despite having different components and thus different k-events. The major difference between them is the members of their sample spaces. In the first probability space, the members of the sample space are $\{1,2\}$ and in the second they are $\{3,4\}$. All other differences between these probability spaces are derived from this difference between their sample spaces. Let $PS_{2,1} = \langle \Omega_{2,1}, \Sigma_{2,1}, P_{2,1} \rangle$ where $\Omega_{2,1} = \{1,2\}$, $\Sigma_{2,1} = \{\emptyset, \Omega_{2,1}, \{1\}, \{2\}\}$ and $P_{2,1}$ assigns the following values to the k-events in $\Sigma_{2,1}$: ($P_{2,1}(\emptyset) = 0$, $P_{2,1}(\Omega_{2,1}) = 1$, $P_{2,1}(\{1\}) = \frac{1}{2}$, $P_{2,1}(\{2\}) = \frac{1}{2}$). And $PS_{2,2} = \langle \Omega_{2,2}, \Sigma_{2,2}, P_{2,2} \rangle$ where $\Omega_{2,2} = \{3,4\}$, $\Sigma_{2,2} = \{\emptyset, \Omega_{2,2}, \{3\}, \{4\}\}$ and $P_{2,2}$ assigns the following k-probabilities to the k-events in $\Sigma_{2,2}$: ($P_{2,2}(\emptyset) = 0$, $P_{2,2}(\Omega_{2,2}) = 1$, $P_{2,2}(\{3\}) = \frac{1}{2}$, $P_{2,2}(\{4\}) = \frac{1}{2}$).

The probability spaces $PS_{2,1}$ and $PS_{2,2}$ are commonly considered as equivalent. The difference between their corresponding sample spaces seem to be irrelevant for determining which probabilistic states can be mathematically described by them. In other words, any probabilistic state mathematically describable by $PS_{2,1}$ is also describable by $PS_{2,2}$ and vice versa. Given a probabilistic state mathematically describable by $PS_{2,1}$, it is possible to describe it using $PS_{2,2}$ simply by mapping the members '1' and '2' from $PS_{2,1}$ to the members '3' and '4' from $PS_{2,2}$.

respectively. Thus, any events describable by the k-events $\{1\}$ and $\{2\}$ which belong to $PS_{2,1}$ can be described by the k-events $\{3\}$ and $\{4\}$ which belong to $PS_{2,2}$ respectively, and vice versa. Loosely speaking, the "names" of these equivalent k-events seem to be irrelevant for determining the set of events mathematically describable by them²⁰.

Hence, it seems that an equivalence relation between k-events must rely on some sort of an equivalence relation between either σ -algebras (and sample spaces) or whole probability spaces. In other words, the question is, given different probability spaces, does an equivalence relation between their k-events involves their probability measure components or only the σ -algebras? The first option means that an equivalence relation between k-events does not depend on their k-probabilities and the latter means that it does. The problem is that in Kolmogorov's theory there is no explicit equivalence relation between probability spaces or their components²¹. The question whether an equivalence relation between k-events, relies on an equivalence relation between σ -algebras or between whole probability spaces, is directly related to the idea that the same event can have different probabilities. An equivalence relation between σ -algebras ignores the k-probabilities of their k-events, and thus can support this idea. While an equivalence relation between probability spaces, does not ignore the k-probabilities of the k-events and as such, undermines this idea.

This point can be nicely illustrated by the following example. In this example there are two probability spaces that have the same sample space and σ -algebra components but different probability measures. This means that the σ -algebra components are equivalent because they are in fact the same σ -algebra. The answer to the question are the k-events in these two probability spaces equivalent or not, reflects one's position on the idea whether the same event can have

²⁰ Notice that there are infinitely many more probability spaces which are equivalent to $PS_{2,1}$ in the sense that they mathematically describe the same set of probabilistic states. For example, any probability space of the following form is commonly considered as equivalent to $PS_{2,1}$: Let $PS_{2,n,m} = \langle \Omega_{2,n,m}, \Sigma_{2,n,m}, P_{2,n,m} \rangle$ be a probability space where n and m are natural numbers and $n \neq m$, $\Omega_{2,n,m} = \{n, m\}$, $\Sigma_{2,n,m} = \{\emptyset, \Omega_{2,n,m}, \{n\}, \{m\}\}$ and $P_{2,n,m}$ assigns the following k-probabilities to the k-events in $\Sigma_{2,n,m}$: ($P_{2,n,m}(\emptyset) = 0$, $P_{2,n,m}(\Omega_{2,n,m}) = 1$, $P_{2,n,m}(\{n\}) = \frac{1}{2}$, $P_{2,n,m}(\{m\}) = \frac{1}{2}$). Since there are infinitely many natural numbers, there are infinitely many probability spaces of this form.

²¹ Kolmogorov's theory has implicit trivial equivalence between probability spaces and between σ -algebras. These relations are simply the identity relations which are also equivalence relations by definition. According to these relations, a σ -algebra or a probability space is equivalent only to itself.

different probabilities or not. Let $PS_{2,1}$ be the probability space defined above (i.e. $PS_{2,1} = \langle \Omega_{2,1}, \Sigma_{2,1}, P_{2,1} \rangle$), and let $PS_{2,3} = \langle \Omega_{2,1}, \Sigma_{2,1}, P_{2,3} \rangle$ be a probability space that has the same sample space and σ -algebra components as $PS_{2,1}$ but a different probability measure. The probability measure $P_{2,3}$ assigns the following k-probabilities to the k-events in $\Sigma_{2,1}$: ($P_{2,3}(\emptyset) = 0$, $P_{2,3}(\Omega_{2,1}) = 1$, $P_{2,3}(\{1\}) = \frac{1}{3}$, $P_{2,3}(\{2\}) = \frac{2}{3}$). In other words, the k-events $\{1\}$ and $\{2\}$ which belong to $PS_{2,1}$ have the k-probabilities 1/2 and 1/2 respectively, while the k-events $\{1\}$ and $\{2\}$ which belong to $PS_{2,3}$ have the k-probabilities 1/3 and 2/3 respectively.

Claiming that the k-events which belong to $PS_{2,1}$ are equivalent to those which belong to $PS_{2,3}$ seems to support the idea that the same event can have different probabilities. Since, for example, any event described by the k-event $\{1\}$ can have probability 1/2 or probability 1/3²². On the other hand, claiming that the k-events which belong to $PS_{2,1}$ are not equivalent to those which belong to $PS_{2,3}$, undermines this idea. Since the only difference between the k-event $\{1\}$ of $PS_{2,1}$ and the k-event $\{1\}$ of $PS_{2,3}$, is their corresponding k-probabilities. This means that the only difference between the events mathematically described by the k-event $\{1\}$ of $PS_{2,1}$ and those described by $\{1\}$ of $PS_{2,3}$, is their probabilities. Thus, if the events describable by the k-event $\{1\}$ of $PS_{2,1}$ are said to be different than those describable by $\{1\}$ of $PS_{2,3}$, then having different probability values *is* what makes them different. This implies that the same event cannot have different probabilities because having different probabilities is what makes otherwise identical events, different.

Lastly, I would like to clarify why the fact that the same k-event can have different k-probabilities, does not imply that the same event can have different probabilities. The reason is roughly because the relation between probabilistic states and probability spaces, is many to many. This means that it is possible that every time a k-event has a different k-probability, it mathematically describes a different event.

²² More accurately, the event described by the k-event $\{1\}$ can have a probability value mathematically described by the k-probability 1/2 or a value described by the k-probability 1/3. Assuming that these k-probabilities are accurate descriptions of the event's probability (and not an approximation of it) implies that the same event can have different probabilities.

Recall that it is commonly held that the same probabilistic state can be mathematically described by different probability spaces (as was shown in the above examples). It is also commonly held that the same probability space can mathematically describe different probabilistic states. For example, the probability space $PS_{2,1}$ which was defined above ($PS_{2,1} = \langle \Omega_{2,1}, \Sigma_{2,1}, P_{2,1} \rangle$, $\Omega_{2,1} = \{1,2\}$, $\Sigma_{2,1} = \{\emptyset, \Omega_{2,1}, \{1\}, \{2\}\}$ and $P_{2,1}: (P_{2,1}(\emptyset) = 0, P_{2,1}(\Omega_{2,1}) = 1, P_{2,1}(\{1\}) = \frac{1}{2}, P_{2,1}(\{2\}) = \frac{1}{2})$) can be a mathematical description of different probabilistic states. Such as an ideal coin toss of a fair coin, or the parity of the number of leaves on a given tree (having no information regarding this tree, trees in general, leaves, or any other relevant information.). And seemingly, there can be infinitely many other such states. Hence this example illustrates the claim that the same probability space can describe different probabilistic states. This claim together with the claim that the same probabilistic state can be described by different probability spaces, means that the relation between probabilistic states and probability spaces is many to many.

The fact that the relation between probabilistic states and probability spaces is many to many, is important. Roughly because it clarifies why the fact that the same k-event can have different k-probabilities does not imply that the same event can have different probabilities. As was mentioned in section 3, according to the definition of the probability space, there can be infinitely many probability measures defined over the same σ -algebra. This means that the same k-event (according to the identity relation between σ -algebras) can have different k-probabilities (unless it is the empty k-event or the sample space k-event). However, since the relation between probabilistic states and probability spaces is many to many, this fact about k-events does not imply that the same event can have different probabilities. It is possible that for each k-probability a k-event has, that k-event mathematically describes a *different* event. In other words, it is possible that all probability spaces that have the same sample space and σ -algebra components but different probability measures, mathematically describe different probabilistic states and specifically different events.

Hence, the fact that there can be infinitely many probability measures defined over the same σ -algebra, does not help to settle the question whether an equivalence relation between k-events, should depend on their k-probabilities or not. This means that this mathematical fact does not answer the question whether the same event can have different probabilities or not. Thus, the

answer to this question depends on the choice of interpretation of probability. This means that the claim that nonempty events with zero probability have it only contingently, is also interpretation-dependent. In other words, interpretations according to which the same event can have different probabilities, can accommodate the claim that nonempty events with zero probability are possible because they have it only contingently. While other interpretations, cannot.

In this section I showed that Kolmogorov's theory can accommodate two different types of equivalence relations between k -events: those which support the claim that the same event can have different probabilities, and those which undermine it. I explained that deciding between these types cannot be done within Kolmogorov's theory. As a result, the question whether the same event can have different probabilities turns out to be interpretation-dependent. This means that the claim that nonempty events with zero probability have it only contingently, is also interpretation-dependent. Thus, nonempty events with zero probability are commonly claimed to be possible mainly because they are nonempty. This claim can also be made on the grounds that the zero probability is assumed to be a contingent fact and not a necessary one. But this latter reason depends on the choice of interpretation and not on Kolmogorov's formalism.

9. Conclusion

In this paper I took on Hájek's challenge and provided an explanation for the widely accepted claim that zero probability does not mean impossibility. The explanation I offered is relevant to all interpretations of Kolmogorov's probability theory since it is based solely on his formalism. More accurately, I have explained why events described by nonempty k -events with zero k -probability are commonly claimed to be possible, despite having zero probability. I have claimed that there are two different senses of possibility expressed by Kolmogorov's probability space. One sense is expressed by k -events being nonempty and the other by k -probabilities being strictly positive.

According to the first sense, events described by nonempty k -events are commonly considered as possible because they are nonempty. Being a nonempty event means roughly that the event is a

nonempty collection of the different possibilities (or alternatives, or options) which constitute a particular probabilistic state. Hence a nonempty event is possible in the sense that it corresponds to such possibilities. Likewise, an empty event is always considered as impossible since it does not correspond to any of the possibilities which constitute a particular probabilistic state. This means that the claim that zero probability does not mean impossibility, refers only to nonempty events.

According to the second sense of possibility, events that have probabilities described by strictly positive k -probabilities, are commonly considered as possible. The exact sense of possibility is determined by the choice of interpretation of probability. This sense of possibility is relevant for explaining the claim that zero probability does not mean impossibility, for two reasons: The first concerns the case where the zero k -probability is said to describe infinitesimal probabilities. In other words, the zero k -probability does not describe a probability value which is exactly zero. Thus, the described events have probabilities which are strictly positive and hence are possible. The second reason concern the case where the zero k -probability describes a probability value which is exactly zero. In this case, the nonempty events with exactly zero probability are considered as possible because they are nonempty but also because it is claimed that they can have strictly positive probability values and thus be possible.

More accurately, I distinguished between cases when the zero k -probability mathematically describes infinitesimal probabilities and cases where it describes exactly zero probability. In the first case, the zero k -probability is in fact a mathematical description which approximates a strictly positive probability value that is not zero. Thus, it does not indicate that the described event is impossible in the sense expressed by zero probability. In this case, the nonempty k -events with zero k -probability describe nonempty events that have strictly positive (though infinitesimal) probabilities. Such events are considered as possible in the two aforementioned senses of possibility.

In the second case, the zero k -probability describe a probability value which is *exactly* zero (and not an infinitesimal). Thus, the described events are nonempty with exactly zero probability. In this case, these events are commonly considered as possible mainly because they are nonempty. However, there is another reason for claiming that such events are possible. This reason is based on the claim that the same event can have different probabilities. If this claim is true, then it might

be possible for nonempty events with zero probability to have strictly positive probabilities and thus be possible. Loosely speaking, it is possible for such events to be possible and that is why they *are* considered as possible.

I have showed that the claim that the same event can have different probabilities does not stem from Kolmogorov's theory. In fact, Kolmogorov's theory is consistent with both this claim and its negation. More specifically, in order to mathematically justify this claim, one has to assume an equivalence relation between k -events (in the sense that they describe the same event) that does not depend on their k -probabilities. The crucial point is that such an assumption is not part of Kolmogorov's theory. Kolmogorov's theory can accommodate two different types of equivalence relations between k -events: those which support this claim and those which undermine it. The first kind is based on some equivalence relation between σ -algebras and the second, on an equivalence relation between probability spaces. Thus, it turns out that the question whether the same event can have different probabilities is interpretation-dependent. Specifically, it implies that the claim that nonempty events with zero probability can have strictly positive probabilities, is interpretation-dependent. This means that claiming that nonempty events with exactly zero probability are possible because they can have other probability values, depends in fact, on the choice of interpretation and not on Kolmogorov's formalism.

In summary, it is commonly claimed that zero probability does not mean impossibility because of the following reasons: The main reason is that nonempty events are considered as possible regardless of their probabilities, simply because they are nonempty. Another reason is that sometimes the "zero probability" is meant to refer to zero *k-probability* which approximates some infinitesimal probability value. Such a value is strictly positive and thus does not imply impossibility. And the last reason is that there are interpretations in which it is assumed that the same event can have different probabilities. Specifically, an event with exactly zero probability can have strictly positive probabilities. This implies that such an event can be possible which means that its zero probability does not mean impossibility.

Bibliography

- Armstrong, D. M., & McCall, S. (1989). God's lottery. *Analysis*, 49(4), 223–224.
<https://doi.org/10.1093/analys/49.4.223>
- Billingsley, P. (1995). *Probability and Measure*. Wiley-Interscience (3rd ed.). New York.
- Elga, A. (2004). Infinitesimal chances and the laws of nature. *Australasian Journal of Philosophy*, 82(1), 67–76. <https://doi.org/10.1080/713659804>
- Gilbert, W. J., & Nicholson, W. K. (2004). *Modern Algebra with Applications* (2nd ed.). Hoboken, N.J.: Wiley-Interscience.
- Gillies, D. (2000). *Philosophical Theories of Probability*. Routledge (3rd ed.). London, New York.
- Goosens, William K. (1979). Alternative axiomatizations of elementary probability theory. *Notre Dame Journal of Formal Logic*, 20(1), 227–239.
- Gwiazda, J. (2010). Probability, Hyperreals, Asymptotic Density, and God's Lottery.
- Gyenis, Z., & Rédei, M. (2014). Defusing Bertrand's paradox. *British Journal for the Philosophy of Science*, 0, 1–25. <https://doi.org/10.1093/bjps/axt036>
- Hájek, A. (2003). What conditional probability could not be. *Synthese*, 137(3), 273–323.
<https://doi.org/10.1023/B:SYNT.0000004904.91112.16>
- Hájek, A. (2012). Interpretations of Probability. In *The Stanford Encyclopedia of Philosophy* (Winter 2012 ed.).
- Jeffreys, H. (1998). *Theory of Probability* (3rd ed.). New York: Oxford University Press.
- Kment, B. (2017). Varieties of Modality. In *The Stanford Encyclopedia of Philosophy* (spring 2017 ed.).
- Kolmogorov, A. N. (1933). *Foundations of the Theory of Probability*. Chelsea.
- Lyon, A. (2010). Philosophy of Probability. In *Philosophies of the Sciences: a guide* (pp. 92–125).
<https://doi.org/10.1002/9781444315578.ch5>

- Lyon, A. (2016). Kolmogorov's Axiomatisation and its Discontents. In *The Oxford Handbook of Probability and Philosophy* (pp. 155–166).
- November, D. D. (2018). *Ontological Implications of the Probability Space*. The Hebrew University of Jerusalem.
- Popper, K. R. (1938). A set of independent axioms for probability. *Mind*, 47(186), 275–277.
- Popper, K. R. (1955). Two autonomous axiom systems for the calculus of probabilities. *The British Journal for the Philosophy of Science*, 6(21), 51–57.
- Popper, K. R. (1959). *The Logic of Scientific Discovery*. Routledge.
- Rényi, A. (1955). On a new axiomatic theory of probability. *Acta Mathematica Academiae Scientiarum Hungaricae*, 6(3–4), 285–335. <https://doi.org/10.1007/BF02024393>
- Von Plato, J. (1994). *Creating Modern Probability: Its Mathematics, Physics, and Philosophy in Historical Perspective*. Cambridge University Press.
- Wenmackers, S., & Horsten, L. (2012). Fair infinite lotteries. *Synthese*, 190(1), 37–61. <https://doi.org/10.1007/s11229-010-9836-x>
- Williamson, T. (2007). How probable is an infinite sequence of heads? *Analysis*, 67(3), 173–180. <https://doi.org/10.1093/analys/68.3.247>