The Indifference Principle, Its Paradoxes and Kolmogorov's Probability Space

Dan David November

Abstract

In this paper I show that the validity of the Indifference Principle (IP), in light of its related paradoxes, is still an open question. I do so by offering an analysis of the Indifference Principle and its related paradoxes in the way they occur in Kolmogorov's probability theory. As such this analysis is relevant to all interpretations of his theory. I describe the conditions that any mathematical formalization of the Indifference Principle must satisfy. Consequently, I show that a mathematical formalization of the Indifference Principle has to be a set of constraints on probability spaces which mathematically describe the same events. I claim that the question whether IP-related paradoxes undermine the validity of the Indifference Principle depends on whether such a set of constraints exist. Since currently there is no mathematical proof of the existence of such a set of constraints, nor of the impossibility of there being one, the validity of the Indifference Principle remains undecided.

1. The Indifference Principle

Bertrand's Paradox and similar ones (such as the Box Factory Paradox and the Wine-Water Paradox) are commonly presented as examples which undermine or even refute the Indifference Principle (IP). There is a vast literature about IP and its related paradoxes. Many scholars have offered various solutions to these paradoxes with the aim of salvaging IP (Jaynes 1973; Strevens 1998; Bangu 2010; Mikkelson 2004; Tissier 1984; Di Porto et al. 2011, 2010; Burock 2005), while others have objected to them. Some scholars have claimed that the paradoxes simply cannot be solved and hence IP faces a real problem (Shackel 2007; Deakin 2006; Howson and Urbach 2006; Rowbottom 2013; Milne 1983). Other scholars have claimed that the various paradoxes are not "real" paradoxes to begin with and hence that IP is valid (Gyenis and Rédei 2014; Marinoff 1994; Aerts and de Bianchi 2014).

The literature about IP and its related paradoxes keeps growing (for example, Drory 2015; Gyenis and Rédei 2014; Aerts and de Bianchi 2014), which shows the ongoing interest in this topic. More importantly, it shows that there is no consensus about which of the proposed explanations for IP's paradoxes is correct. This lack of consensus can be attributed in part to the lack of consensus regarding which of the interpretations of probability is the "right" one[1]. This is because almost all the proposed explanations for IP's paradoxes are offered on the basis of a particular interpretation of probability (sometimes implicitly). As such, these explanations are commonly based on assumptions that are part of these interpretations and not necessarily shared by other interpretations of probability. Usually these assumptions are important for the different explanations, since they concern the nature of probability (and hence IP and its paradoxes). To the best of my knowledge there have been hardly any attempts to explain IP and its related paradoxes without relying on assumptions belonging to particular interpretations of probability[2]. My aim in this paper is to fill this lacuna.

In this paper I explain IP and its paradoxes based on how they occur in Kolmogorov's probability theory. This theory is widely considered *the* standard mathematical probability theory and the various interpretations of probability are commonly considered interpretations of Kolmogorov's

[1] See Gillies (2000), Hájek (2012) for good surveys of the various interpretations of probability. See Von Plato (1994) for a more historical perspective and Lyon (2010) for a short survey which highlights the major problems each of the main interpretations faces.

[2] Deakin's (2006) discussion of the Wine-Water Paradox is a step in this direction.

theory[3]. The distinction between mathematical probability theories such as Kolmogorov's and philosophical probability theories known as "interpretations of probability" is widely accepted in the philosophy of probability. Roughly, the difference between these two types of theories is that they address different questions about the basic notions of 'probability' and 'event'[4]. According to Lyon, mathematical probability theories "[...] tell us how probabilities behave, how to calculate probabilities from other probabilities, but they do not tell us what probabilities *are*." (Lyon 2010, 93). The latter question is answered by interpretations of probability. Hence it can be said that mathematical probability theories provide the mathematical parts of the definitions of 'probability' and 'event', while the interpretations provide the non-mathematical parts.

In this paper I analyze IP and its related paradoxes in a way they could (or should) be mathematically formalized within Kolmogorov's theory. My analysis relies on the common assumption that Kolmogorov's probability space is a correct mathematical description of the notions of 'probability' and 'event'. This assumption seems to be common to all the main interpretations of probability. As such my analysis is relevant to all interpretations of Kolmogorov's theory. In particular, I describe the conditions that a mathematical formalization of IP within Kolmogorov's theory must satisfy and show that such a formalization has to be a set of constraints (C) on probability spaces that have the equivalent σ-algebra components. These σ-algebras are equivalent in the sense that they mathematically describe the same events. As a result, the question of whether IP-related paradoxes undermine the validity of IP turns out to depend on whether such a C exist. Since currently there is no mathematical proof of the existence of such a C (or a proof of the impossibility of there being one), the question remains open.

The term "Principle of Indifference" was first coined by Keynes. As he put it, "[t]he Principle of Indifference asserts that if there is no known reason for predicating of our subject one rather than

[3] However, see Lyon (2016) and Hájek (2012) for objections to this common view.
[4] 'Probability' here is used as a general term for notions such as credence, degree of belief, propensity, and chance. Similarly, 'event' is used as a general term for anything that has a 'probability' (such as a proposition, a state of the world, or a metaphysical event).

another of several alternatives, then relatively to such knowledge the assertions of each of these alternatives have an equal probability. Thus equal probabilities must be assigned to each of several arguments, if there is an absence of positive ground for assigning unequal ones." (Keynes 1921, 45). Interestingly, in the literature there are other phrasings of IP which are slightly different than Keynes'. For example, Bartha and Johns (2001, 109) describe IP as follows: "In the absence of any known reason to assign two outcomes different probabilities, they ought to be assigned the same probability". Howson and Urbach (2006, 266) describe IP as asserting that "[...] equal parts of the possibility space should receive equal probabilities relative to a null state of background information"(Howson and Urbach 2006, 266)(Howson and Urbach 2006, 266)(Howson and Urbach 2006, 266)(Howson and Urbach 2006, 266)(Howson and Urbach 2006, 266)(Howson and Urbach 2006, 266) and Mellor (2005, 28) states, "[IP] says that evidence which gives us no reason to think that any one of a number of mutually exclusive possibilities [...] is more probable than any other will give those possibilities equal epistemic probabilities", and the list goes on. These different phrasings show that there is no consensus regarding the exact phrasing of IP. However, almost all of them seem to have the same three components that I describe in the next section. Roughly, in all of IP's phrasings it is assumed that events are comparable and that equivalent events should have equal probabilities if there is no information indicating otherwise. The differences between the various phrasings are mainly in the events' equivalence criteria. These differences depend on the exact definition of 'events', which in turn depends on the choice of interpretation of probability.

IP is commonly treated as an epistemic principle in the sense that it is a principle that an agent can use when she has to assign subjective probabilities (such as degrees of belief or credences) to propositions (i.e. "subjective events"). As such, IP is commonly thought of as relevant *only* to subjective interpretations of probability. However, IP also appears in discussions relying on the classical interpretation and others in which probabilities and events are considered objective. In such discussions, IP is treated as a tool that agents use to infer the events' *objective* probabilities (commonly referred to as "chances" or "propensities"). Thus, although IP is commonly thought of as a way of assigning subjective probabilities, it can also be seen as a way of inferring objective probabilities. This means that using IP for assigning probabilities to events does not necessarily imply that these probabilities are subjective. Hence IP is relevant to both subjective and objective interpretations of probability.

Since my purpose in this paper is to explain IP and its related paradoxes within Kolmogorov's theory, it is sufficient to think of IP as a description of a connection between *mathematical* events and *mathematical* probabilities. The question of whether or not IP is taken to be an epistemic principle seems to be irrelevant to IP's mathematical formalization itself. Treating IP as a description of a mathematical connection between events and probabilities in Kolmogorov's probability theory is one of the keys which make my analysis relevant to all interpretations of this theory.

In the next section I give a general characterization of IP as a connection between events and probabilities in the aforementioned mathematical sense.

2. A characterization of the Indifference Principle

In this section I characterize the connection described by IP between events and probabilities. This characterization seems to be common to almost all the different phrasings of IP regardless of whether they rely on a subjective or an objective interpretation.

The connection described by IP has the following three components:

1. Presumption – Events[5] are comparable and thus can be considered equivalent (or equal) in some sense.
2. Assertion – Equivalent events have (or should have) equal probabilities.
3. Conditionalization – There is no information indicating otherwise. Thus, equivalent events have equal probabilities *if* (sometimes *iff*) there is no information indicating otherwise.

The connection between events and probabilities described by IP's assertion component is that *equivalent events have equal probabilities*[6]. This connection relies on IP's presumption that

[5] Here 'events' is used as a general term which refers to alternatives, outcomes, possibilities, parts of the possibility space, etc.

events can be considered equivalent in some sense, such as "equipossible cases" (Fraassen 1989, 298), "equal parts of the possibility space" (Howson and Urbach 2006), "possibilities of which we have equal ignorance" (Shackel 2007, 151), and the like. Equivalent events are not necessarily identical (but identical events are trivially equivalent because they are the same event). The important point is that different events can be considered equivalent according to some criterion. The specifics of the criterion commonly depend on the choice of interpretation of probability. Obviously, the criterion cannot be that the events have equal probabilities because then IP would be the trivial assertion that events with equal probabilities have equal probabilities. The key point is that IP relies on there being a way of comparing events and deciding that they are equivalent which does not depend on their probabilities. This presumption is a necessary condition for IP's assertion. If events cannot be considered equal, IP's assertion does not hold.

According to IP's conditionalization component, IP's assertion is also conditional on there being no information indicating that the equivalent events do not have equal probabilities[7]. This is expressed as "[...] in the absence of reasons to the contrary" (Fraassen 1989, 299), "[...] relative to a null state of background information" (Howson and Urbach 2006, 266), "[...] iff we have insufficient reason to consider any one of these outcomes more or less likely than any other" (Mikkelson 2004, 137), "[...] if we have no grounds for preferring one [outcome] over any other [...]" (Norton 2008, 47), "[...] in the absence of any known reason to assign two events differing probabilities [...]" (Strevens 1998, 231), and the like.

IP also implicitly relies on the trivial assumption that the probabilities of the events are not given. It is obvious that if they are, then the connection between them and the events is already given *explicitly* and there is no need for IP to connect them. In other words, when the events' probabilities are given there are two alternatives: equivalent events either have equal probabilities or they do not. If they do, then IP is redundant. If they do not, then IP cannot be used because the condition that there is no information indicating that the equivalent events *do not* have equal probabilities is violated. In Section 5 I show that given a σ-algebra which

---

[6] IP also implicitly asserts that if one event (e_1) is greater than another event (e_2) (in the sense in which the events are comparable), then e_1 has (or should have) greater probability than e_2.

[7] Some have argued that this condition is insufficient, and that IP must also depend on information which *positively* indicates that equivalent events have equal probabilities. This is an important point concerning IP, but it does not affect the argument presented in this paper.

mathematically describes the events, the connection drawn by IP is described by a particular probability space whose σ-algebra component is the given one. Thus, this apparently trivial assumption is important because it means that the probability space which describes the events is *not* already given.

Another important implicit assumption concerning IP is that an application of it results in a unique assignment of probabilities. More precisely, given a set of events, an application of IP results in a unique assignment of probabilities to these events. This is a crucial assumption, since, as I show below, it is at the heart of all IP-related paradoxes. Roughly, in all such paradoxes, an event is assigned *different* probabilities by different applications of IP, and this result is considered paradoxical. Hence, IP's assertion that equivalent events have equal probabilities (as long as there is no information indicating otherwise) also implicitly implies that the given equivalent events have *specific* (equal) probability values. In other words, IP asserts that equivalent events have specific equal probabilities and no other equal probability values.

In the next section I discuss how applying IP can lead to paradoxes.

3. The Indifference Principle-related paradoxes

Applying IP arguably is the most common way of assigning probabilities to events. In many cases it is used implicitly. The following paradigmatic examples are commonly presented as successful applications of IP:

1. The assignment of probability 1/2 to the event of a coin toss landing on heads (or tails). The two events ('heads' and 'tails') are considered equivalent, since there is no information indicating that the coin is biased, and hence are assigned equal probabilities.
2. The assignment of probability 1/6 to each of the events of a die toss landing on one of its faces, assuming that the die is unbiased.

3. The assignment of probability 1/2 to the event of a dart hitting the left (or right) half of a target board. Assuming that the dart hits the board, the two events ('left' and 'right') are considered equivalent (if there is no other information about the board, the marksman or any other relevant factor) and hence have equal probabilities.

Very roughly, in all these examples some events are considered equivalent in some sense and hence, according to IP, have equal probabilities.

IP-related paradoxes are often presented as examples of unsuccessful applications of IP in the sense that they do not result in a unique assignment of probabilities to events, in contrast to the implicit assumption mentioned in the previous section. Roughly, in each of these paradoxes there is a question about the probability of a given event. Each of these questions is answered by applying IP. However, this is done in more than one way and as a result the event in question is assigned different probabilities by the different applications of IP. These different assignments of probabilities are commonly treated as paradoxical because of the aforementioned assumption that an application of IP to given events results in a *unique* assignment of probabilities to the given events.

A simple example of an IP-related paradox is the Box Factory Paradox: "A factory produces cubes with side-length between 0 and 1 foot; what is the probability that a randomly chosen cube has side-length between 0 and 1/2 a foot? The tempting answer is 1/2, as we imagine a process of production that is uniformly distributed over side-length. But the question could have been given an equivalent restatement: A factory produces cubes with face-area between 0 and 1 square-feet; what is the probability that a randomly chosen cube has face-area between 0 and 1/4 square-feet? Now the tempting answer is 1/4, as we imagine a process of production that is uniformly distributed over face-area. This is already disastrous, as we cannot allow the same event to have two different probabilities [...]. But there is worse to come, for the problem could have been restated equivalently again: A factory produces cubes with volume between 0 and 1 cubic feet; what is the probability that a randomly chosen cube has volume between 0 and 1/8 cubic-feet? Now the tempting answer is 1/8, as we imagine a process of production that is uniformly distributed over volume. And so on for all of the infinitely many equivalent reformulations of the problem (in terms of the fourth, fifth, … power of the length, and indeed in terms of every non-zero real-valued exponent of the length). What, then, is *the* probability of the event in question?" (Hájek 2012, sec. 3.1)

In the Box Factory Paradox, IP is applied several times. Each application of IP gives a different answer to the question about the probability that a randomly chosen cube would have a certain size (and hence a certain side-length, face-area and volume). The paradox is the fact that the same event is attributed different probabilities by different applications of IP. In this sense, the Box Factory Paradox is an example of an unsuccessful use of IP.

Bertrand's Paradox is perhaps the most famous example of an unsuccessful application of IP. It is commonly described as the having three different answers to the following question (taken from Clark (2007, 22)): "What is the chance that a random chord of a circle is longer than the side of an inscribed equilateral triangle?

(1) The chords from a vertex of the triangle to the circumference are longer if they lie within the angle at the vertex. Since that is true of one-third of the chords, the probability is one-third.

(2) The chords parallel to one side of such a triangle are longer if they intersect the inner half of the radius perpendicular to them, so that their midpoint falls within the triangle. So the probability is one-half.

(3) A chord is also longer if its midpoint falls within a circle inscribed within the triangle. The inner circle will have a radius one-half and therefore an area one-quarter that of the outer one. So the probability is one-quarter."

The above three answers to the question presented in Bertrand's Paradox are said to be the results of different applications of IP. Hence the paradox is that an "[a]pplication of the principle of indifference is supposed to suffice for solving probability problems. Probability problems have, of their nature, unique solutions, because a solution is a *single* function from the events of interest into [0, 1]. The solution being a function entails that each event has a unique probability. Yet different ways of applying the principle here result in different probabilities for the same event." (Shackel 2007, 152)

Another important IP-related paradox is the Wine-Water paradox: "We are given a glass containing a mixture of water and wine. All that is known about the proportions of the liquids is that the mixture contains at least as much water as wine, and at most, twice as much water as wine. The range for our assumptions concerning the ratio of water to wine is thus the interval 1 to 2. Assuming that nothing more is known about the mixture, the indifference or symmetry principle or any other similar form of the classical theory tells us to assume that equal parts of this interval have equal probabilities. The probability of the ratio lying between 1 and 1.5 is thus 50%, and the other 50% corresponds to the probability of the range 1.5 to 2.

But there is an a1ternative method of treating the same problem. Instead of the ratio water/wine, we consider the inverse ratio, wine/water; this we know lies between 1/2 and 1. We are again told to assume that the two halves of the total interval, i.e., the intervals 1/2 to 3/4 and 3/4 to 1, have equal probabilities (50% each); yet, the wine/water ratio 3/4 is equal to the water/wine ratio 4/3. Thus, according to our second calculation, 50% probability corresponds to the water/wine range 1 to 4/3 and the remaining 50% to the range 4/3 to 2. According to the first calculation, the corresponding intervals were 1 to 3/2 and 3/2 to 2. The two results are obviously incompatible." (Von Mises 1981, 77)

Like the Box Factory Paradox and Bertrand's Paradox, the Wine-Water Paradox presents us with a question about the probability of an event and we are given different answers to this question. Each of these answers is the result of an application of IP. As such they are said to be "incompatible" and thus the Wine-Water Paradox is considered a paradox.

To sum up, IP-related paradoxes can be described as follows:

1. There is a question about the probability of a given random event.
2. The question is answered by applying IP.
3. The problem is that the question is given *different* answers by different applications of IP.

4. The different answers are considered paradoxical because of the aforementioned assumption that applications of IP to given events result in a unique assignment of probabilities to the events.

Before I move on to describe how IP and its related paradoxes are manifested in Kolmogorov's theory, I wish to address the claim that these paradoxes are not well-defined. Some writers have claimed that the questions posed (or the terms used) in the different IP-related paradoxes are not well-defined (they are vague, ambiguous, underdetermined, etc.). For example: "[...] Bertrand's original problem is vaguely posed [...]" (Marinoff 1994, 1), "Bertrand's problem cannot undermine Laplace's principle [i.e. IP] provided that the former is posed in non ambiguous terms [...]" (Aerts and de Bianchi 2014, 1), "[...] all so-called paradoxes to PI [i.e. IP] are simply disagreements and ambiguity in relation to sample space identification [...]" (Burock 2005, 2). In particular, it has been claimed that the term 'random' is the cause of the problem: "[randomness] is a notoriously difficult concept. In fact, it may not even be a single concept at all, but a cluster of concepts [...]" (Bangu 2010, 33), "Bertrand's paradox is of course not a logical paradox. The different results arise from assigning three different meanings to the phrase 'at random' [...]" (Tissier 1984, 19).

As mentioned above, there is no consensus among philosophers regarding the exact meaning of the notion 'probability', and the same goes for 'randomness'. This lack of consensus is manifested in the different interpretation of probability theory. So, philosophically speaking, IP-related paradoxes and, in particular, the term 'random' might not be well-defined. However, arguably anything that is considered random according to some interpretation of probability can be mathematically described by Kolmogorov's probability theory. This suggests that IP-related paradoxes can be correctly described by his theory. The key point is that when such paradoxes are described by Kolmogorov's theory, they do not seem to be any more vague or ambiguous than any other mathematical question concerning the probability of a given event! Such questions are exactly the sort that appear in IP-related paradoxes. In Kolmogorov's theory, they cannot be answered unless the relevant probability spaces are given. And in all IP-related paradoxes, these spaces are assumed to be given by applying IP. However, applying IP in Kolmogorov's theory relies on having a mathematical formalization of IP and, as I explain in the rest of this paper, this is an open problem. Thus, it would be more accurate to say that the problem with IP-related paradoxes is that the mathematical formalization of IP is still an open question, rather than claim that they are not well-defined.

Since my aim in this paper is to explain IP and its related paradoxes based on the way they are manifested in Kolmogorov's probability theory, in the next section I give a brief description of Kolmogorov's probability space definition and show how IP relates to it.

4. Kolmogorov's probability space

In this section I describe Kolmogorov's definition of probability space[8]. I highlight several points which are important for understanding IP and its related paradoxes and, in particular, how IP can be mathematically formalized in Kolmogorov's theory.

Kolmogorov's probability space is defined as a triple $<\Omega,\Sigma,P>$ consisting of the following 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 referred to as "elementary events". However, this name is misleading since these members are *not* a mathematical formalization of 'events'. 'Events' are mathematically defined as the members of the σ-algebra component (which is defined just below). Nevertheless, there is a sense in which the sample space and its members are

[8] The definition I present in this paper is one of several standard ways of defining Kolmogorov's probability space. See Billingsley (1995, 23) for a slightly different definition.

indeed elementary, and it is the fact that Kolmogorov's mathematical formalizations of 'events' and 'probabilities' depend on them[9].

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

The members of Σ are Kolmogorov's mathematical formalization of 'events'. In other words, mathematically 'events' are sets of members of a sample space that together form a σ-algebra. Moreover, since the σ-algebra component is part of a given probability space, it is also connected to a specific probability measure (the probability measure component is defined below). This means that mathematically 'events' are sets of members of a sample space that have *specific* probability values (described by the probability measure). This fact is very important for the formalization of IP in Kolmogorov's theory. It is especially important to the formalization of IP's assertion that equivalent events have *specific* equal probabilities.

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. $\mathrm{P}(\emptyset) = 0$
   3.3. P is countably additive (which means that for all countable collections $\{E_i\}$ of pairwise disjoint sets $\mathrm{P}(\bigcup_i E_i) = \sum_i \mathrm{P}(E_i)$)
   3.4. P returns results in the unit interval [0,1] and $\mathrm{P}(\Omega) = 1$

The value assigned to a member of the σ-algebra ($e \in \Sigma$) by the probability measure function (i.e. P(e)) is the probability of *e*. This means that mathematically 'probabilities' are defined as the values of a real function from a σ-algebra to the unit interval which satisfies certain conditions (described by the definition of the probability measure).

[9] For a detailed analysis of the interpretive meaning of the sample space component, see November (2018a).

A key point regarding the possible mathematical formalizations of IP in Kolmogorov's theory is that according to the definition of the probability space the same mathematical events can have different probability values. This point relies on the assumption that the identity relation between mathematical events is the identity relation between σ-algebras[10]. In other words, two mathematical events are the same iff they belong to the same σ-algebra and are set-theoretically identical (i.e. contain the same members). According to the definition of the probability space, there can be infinitely many probability measures defined over the same σ-algebra (except for trivial σ-algebras[11]). This means that the same mathematical events can have infinitely many different probability values. Each time a σ-algebra has a different probability measure defined over it, however, it belongs to a different probability space. In other words, each time an event has a different probability, it necessarily belongs to a different probability space.

This point is very important for understanding both IP's mathematical formalization and its related paradoxes. Recall that IP's assertion that equivalent events have equal probabilities implicitly implies that the equivalent events have *specific* equal probability values. Thus, they have these values and not any other equal probabilities. Loosely speaking, it is assumed that an application of IP results in a *unique* assignment of probabilities to events. Naively, the mathematical formalization of this assumptions seems to be the claim that given a set (S) of all probability spaces which have the same σ-algebra component (i.e. contain the same events), an application of IP is a way to select exactly one probability space out of S. However, this formalization is inaccurate because *different* σ-algebras can be considered equivalent in the sense that they mathematically describe the same events. Similarly, different probability spaces (due to their different σ-algebra components) can be considered equivalent in the sense that they mathematically describe the same events when having the same specific probabilities. Thus, a more precise formalization of the above assumption that an application of IP ends with a unique assignment of probabilities to events is that it is a way to select all *equivalent* probability spaces out of all probability spaces which have *equivalent* σ-algebra components. In other words, an

[10] Kolmogorov's theory does not actually include an explicit definition of an identity relation between events. See November (2018b) for a thorough discussion of this issue.

[11] A trivial σ-algebra is a σ-algebra that contains only the sample space event and the empty event. The only probability measure that can be defined over a given trivial σ-algebra is its corresponding trivial probability measure, which assigns 1 to the sample-space event and 0 to the empty event.

application of IP is a way to select all probability spaces which mathematically describe the same events when having specific probabilities out of all the probability spaces which mathematically describe the same events.

The idea that different probability spaces are equivalent in the sense that they can mathematically describe the same events having specific probabilities seems to be widely accepted. For example, the probability spaces $PS_{2,1}$ and $PS_{2,2}$ are commonly considered equivalent: $PS_{2,1} = \langle \Omega_{2,1}, \Sigma_{2,1}, \mathrm{P}_{2,1} \rangle$ where $\Omega_{2,1} = \{1,2\}$, $\Sigma_{2,1} = \left\{\emptyset, \Omega_{2,1}, \{1\}, \{2\}\right\}$ and $\mathrm{P}_{2,1}$ assigns the following probabilities to the events in $\Sigma_{2,1}$: ($\mathrm{P}_{2,1}(\emptyset) = 0$, $\mathrm{P}_{2,1}\left(\Omega_{2,1}\right) = 1, \mathrm{P}_{2,1}(\{1\}) = \frac{1}{3}, \mathrm{P}_{2,1}(\{2\}) = \frac{2}{3}$). And $PS_{2,2} = \langle \Omega_{2,2}, \Sigma_{2,2}, \mathrm{P}_{2,2} \rangle$ where $\Omega_{2,2} = \{3,4\}$, $\Sigma_{2,2} = \left\{\emptyset, \Omega_{2,2}, \{3\}, \{4\}\right\}$ and $\mathrm{P}_{2,2}$ assigns the following values to the events in $\Sigma_{2,2}$: ($\mathrm{P}_{2,2}(\emptyset) = 0$, $\mathrm{P}_{2,2}\left(\Omega_{2,2}\right) = 1, \mathrm{P}_{2,2}(\{3\}) = \frac{1}{3}, \mathrm{P}_{2,2}(\{4\}) = \frac{2}{3}$). The differences between $PS_{2,1}$ and $PS_{2,2}$ seem to be irrelevant to their expressive power. Anything describable by $PS_{2,1}$ is also describable by $PS_{2,2}$ and vice versa, simply by mapping the sample-space members '1' and '2' from $PS_{2,1}$ to the members '3' and '4' from $PS_{2,2}$ respectively. Arguably (see November (2018b)), the events mathematically describable by $PS_{2,1}$ and $PS_{2,2}$ are also describable by the probability space ($PS_{2,3}$) *but* with different probabilities: $PS_{2,3} = \langle \Omega_{2,3}, \Sigma_{2,3}, \mathrm{P}_{2,3} \rangle$ where $\Omega_{2,3} = \{5,6\}$, $\Sigma_{2,3} = \left\{\emptyset, \Omega_{2,3}, \{5\}, \{6\}\right\}$ and $\mathrm{P}_{2,3}$ assigns the following probabilities to the events in $\Sigma_{2,3}$: ($\mathrm{P}_{2,3}(\emptyset) = 0$, $\mathrm{P}_{2,3}\left(\Omega_{2,3}\right) = 1, \mathrm{P}_{2,3}(\{5\}) = \frac{1}{4}, \mathrm{P}_{2,3}(\{6\}) = \frac{3}{4}$). The σ-algebra components of the above three probability spaces, $PS_{2,1}$, $PS_{2,2}$ and $PS_{2,3}$, seem to be equivalent in the sense that any set of events mathematically describable by one of them is also describable by the other two. (This can be done by mapping the members of $\Omega_{2,3}$ to those of $\Omega_{2,1}$ (or $\Omega_{2,2}$)). Thus, loosely speaking, all three probability spaces mathematically describe the same events because their corresponding σ-algebra components are equivalent. But only $PS_{2,1}$ and $PS_{2,2}$ describe them as having the same specific probabilities (1/3 and 2/3), while according to $PS_{2,3}$ their probabilities are different (1/4 and 3/4). In other words, $PS_{2,1}$ and $PS_{2,2}$ are equivalent probability spaces while $PS_{2,3}$ is not equivalent to them in the sense that it describes the events as having other probabilities.

In this paper, I put aside the issues of equivalence relations between probability spaces and between σ-algebras. Instead, I focus on the much simpler case of probability spaces that have the same σ-algebra component. In other words, I take the set (S_i) of all probability spaces which have the *same* σ-algebra component as a representative case of the set of all probability spaces which have *equivalent* σ-algebra components (S_e). S_i is contained in S_e, because identical σ-algebra components are trivially equivalent. I show that even in this much simpler case, there is no known mathematical formalization of IP which manages to select exactly one probability space out of S_i, and there is no known proof that there cannot be such a formalization. This implies that there is no known mathematical formalization of IP which manages to select the set of all equivalent probability spaces out of S_e, and that there is no known proof that there cannot be such a formalization.

The fact that, according to Kolmogorov's probability space definition, the same mathematical events can have different probabilities is important in understanding IP-related paradoxes. As mentioned above, in all IP-related paradoxes the same events are assigned different probabilities. These assignments are said to be the result of different applications of IP to these events. The assignments of different probabilities to these events by applications of IP are considered paradoxical because of the aforementioned implicit assumption that such applications should result in a unique assignment of probabilities to these events. It is important to understand that, according to Kolmogorov's definition of the probability space, assignments of different probabilities to the same events in general are *not* considered paradoxical. In other words, assignments of different probabilities by means which are not applications of IP are not paradoxical. Such assignments are simply different probability spaces that have the same σ-algebra components (and hence the same sample space components[12]) but different probability measure components (i.e. $<\Omega,\Sigma,P_1>$ and $<\Omega,\Sigma,P_2>$ where $P_1 \neq P_2$). (In the rest of this paper I refer to such probability spaces as: "same-events spaces"). Assignments of different probabilities to the same events by different applications of IP are considered paradoxical *only* because it is

[12] Probability spaces that have the same σ-algebra components necessarily have the same sample-space components. This is because a σ-algebra always contains as a member the sample space which the σ-algebra is defined over . This means that σ-algebras which are defined over different sample spaces have different members and thus *are* different.

implicitly assumed that an application of IP selects exactly one probability space from a given set of all same-events spaces.

In the next section I elaborate on several important issues regarding IP's mathematical formalization in Kolmogorov's theory.

5. IP's formalization in Kolmogorov's probability space

In this section I discuss important issues regarding IP's mathematical formalization in Kolmogorov's probability theory. Roughly, I claim that this formalization has to be a set of constraints on probability spaces.

Any mathematical formalization of IP has to address the three components of IP (its Presumption, Assertion and Conditionalization) and its underlying assumptions. The key issue is IP's assertion that equivalent events have *specific* equal probability values. Thus, a formalization of IP has to mathematically describe this connection between events and *specific* probabilities. Such a connection is mathematically described by a probability space. More precisely, given a σ-algebra which mathematically describes the events, the connection described by IP is mathematically described by a specific probability space whose σ-algebra component is the given one. Since there can be infinitely many probability spaces whose σ-algebra component is the same, a mathematical formalization of IP has to be some mathematical way of uniquely selecting exactly one probability space from a given set of same-events spaces. In other words, IP's mathematical formalization has to include constraints on sets of same-events spaces. Ideally, given a set of constrains (C) which is a mathematical formalization of IP, for each set (S) of same-events spaces there is only one probability space (s) in S which satisfies the constraints in C, and C is such that equivalent events in s have equal probabilities. Any such set of constraints would be a good mathematical formalization of IP in the sense that it successfully formalizes IP's assertion.

In addition to IP's assertion, it seems that a mathematical formalization of IP also has to address IP's presumption and conditionalization. Unfortunately, unlike IP's assertion, it is not clear how these components should be mathematically formalized within Kolmogorov's theory. In the rest of this section, I show that IP's presumption that events are comparable should be formalized as an order relation on the σ-algebra component. However, it is not clear what are the specifics of such a relation or even whether such a relation exists in Kolmogorov's theory. I also show that IP's conditionalization should be mathematically formalized as a set of constraints on sets of same-events spaces. However, due to the formalization of IP's assertion, the formalization of IP's conditionalization turns out to be redundant. Moreover, it is not clear that IP's conditionalization can be formalized in a non-redundant way in Kolmogorov's theory.

Recall that IP's presumption is that events are comparable. Generally, when things are claimed to be comparable (non-mathematically), this claim can be mathematically formalized as an order relation on the set of those things[13]. However, because there seem to be different non-mathematical senses in which *events* are compared and considered equivalent, it is not clear which mathematical order relation on the σ-algebra should be used to mathematically describe their comparability.

There are two obvious candidates for an order relation on the σ-algebra component. Both stem from the fact that by Kolmogorov's definition events are sets. Unfortunately, both are unsuitable to be the mathematical formalization of IP's presumption. The first order-relation is based on the inclusion relation between sets. By this relation, a set A is considered strictly bigger than a set B iff A includes B and B does not include A (i.e. $A > B \Leftrightarrow (B \subseteq A \ and \ A \not\subseteq B)$). This means that A and B are considered equivalent iff A includes B and B includes A (i.e. $A \equiv B \Leftrightarrow (B \subseteq A \ and \ A \subseteq B)$)[14]. However, this relation fails to be a good mathematical description of the presumption that events are comparable, because it is a *partial* order relation. This means that by

[13] "Two elements a and b of A [where $(A, \leq)$ is a partially ordered set] are said to be *comparable* if either $a \leq b$ or $b \leq a$." (Potter 2004, 104)

[14] More precisely, in set theory such sets are in fact the same set. I.e. they are considered *identical* (*A*=*B*) and not just equivalent ($A \equiv B$).

this relation it is possible that not all events in a given σ-algebra component are comparable! And that is a problem for IP's formalization because if there are events that cannot be compared, one cannot use IP to assign them probabilities.

The second candidate for an order relation on the σ-algebra component is based on the events' cardinality. By this relation, a set A is considered strictly bigger than a set B iff A's cardinality is strictly bigger than B's cardinality (i.e. $A > B \Leftrightarrow C(B) > C(A)$). This also means that A and B are considered equivalent iff A's cardinality is equal to that of B (i.e. $A \equiv B \Leftrightarrow C(B) = C(A)$). This order relation is a total order relation which means that all events are comparable. However, it cannot serve as a mathematical description of IP's presumption that events are comparable because of IP's assertion that equivalent events have equal probabilities. The reason is that there are cases for which comparing events by their cardinality and assigning them probabilities accordingly, using IP, contradicts Kolmogorov's definition. In particular, the problem arises in cases when the sample space has cardinality of at least $\aleph_0$ and there are two mutually exclusive events whose union is the sample space event (i.e. $e_1, \mathrm{e}_2 \mid e_1 \cup e_2 = \Omega$) and they all have the same cardinality (i.e. $C(e_1) = C(e_2) = C(\Omega)$). In such cases, by IP all three events should have the same probability since they all have the same cardinality (i.e. $p(e_1) = p(e_2) = p(\Omega)$) and by Kolmogorov's definition $p(\Omega) = p(e_1) + p(e_2)$ and $p(\Omega) = 1$. Unfortunately, there is no way to satisfy these three conditions together in Kolmogorov's theory. In such cases, either $p(\Omega) = 0$ or $p(\Omega) \neq p(e_1) + p(e_2)$, in contrast to Kolmogorov's definition, or the probabilities of events with equal cardinality are not equal. The latter alternative means that when IP's presumption is formalized with an order relation based on cardinality, IP's assertion can turn out to be false. This means that the mathematical formalization of IP's presumption cannot be the cardinality-based order relation.

The above two failed attempts to formalize IP's presumption are based solely on the fact that events are sets. This seems to suggest that any other attempt to formalize IP's presumption will require an order relation on the σ-algebra component which would be based on some additional information. As such, any other attempt to formalize IP's presumption would require an addition to Kolmogorov's definition of events. (For example, Shackel's assumption, which I discuss below, is that the σ-algebra component has a measure defined over it which is not a probability measure, and thus the order relation between events is based on their measure values.)

The third component of IP is its conditionalization that equivalent events have equal probabilities *only if there is no information indicating otherwise*. Here the main question is how to mathematically describe the (non-mathematical) information which indicates that equivalent events do not have equal probabilities. It seems that there can be very different pieces of information which indicate the above, and thus it is not clear that all of them can be described by the same mathematical description. However, they all have one thing in common, which is that they indicate that equivalent events *do not* have equal probabilities. In other words, they are different pieces of information which refute the possibility that equivalent events have equal probabilities. Thus, a formalization of any such piece of information would be a mathematical way of selecting one or more probability spaces from a given set of same-events spaces (S) such that equivalent events *do not* have equal probabilities in the selected spaces. In other words, a formalization of such information would be a set of constraints (C_{cond}) on a given S which restricts it to one or more spaces in which equivalent events do not have equal probabilities.

However, the mathematical formalization of IP's assertion as a set of constraints (C) on a set of same-events spaces, already precludes C_{cond} and thus makes it redundant! Recall that IP's assertion is conditioned on IP's conditionalization, which means that C holds iff there is no C_{cond} given. But this is trivially true because any s in S which satisfies C is a space in which equivalent events have equal probabilities, and thus s does not satisfy C_{cond}. This means that when IP's assertion is formalized as C, IP's conditionalization is redundant and there is no need to explicitly formalize it. As a result, it turns out that the mathematical formalization of IP in Kolmogorov's theory must be a set of constraints on same-events spaces that is accompanied by a suitable order relation on the σ-algebra component.

In the next section I discuss the important implication of IP's formalization for its validity. But first I would like to address Shackel's discussion of Bertrand's Paradox. This discussion is important mainly because it includes one of the few attempts to explicitly mathematically formalize IP. Shackel mathematically defines IP as follows: "**Principle of Indifference for Continuum Sized Sets.** For a continuum sized set X, given a σ-algebra, Σ, on X and a measure, μ, on Σ, and given that we have no reason to discriminate between members of Σ with equal measures, then we assign equiprobability to members of Σ with equal measures: For all x, y in Σ, if $\mu(x) = \mu(y)$ then $P(x) = P(y)$. (This can easily be achieved by setting $P(x) = \mu(x)/\mu(X)$ for all x in Σ.)" (Shackel 2007, 159). This mathematical formalization of IP includes a formalization of IP's presumption and assertion, but not of IP's conditionalization. According to Shackel, the formalization of IP's presumption is that events are comparable by their *measure values*. Hence equivalent events are events that have equal measure values. In other words, Shackel implicitly defines an order relation on σ-algebras that is based on the measure values of the events. Shackel even provides a method for assigning probabilities to events which guarantees that events with equal measure values (i.e. equivalent events) have specific equal probability values[15].

The problem with Shackel's formalization of IP lies in his formalization of IP's presumption. The key point is that this formalization relies on a strong assumption that is *not* part of Kolmogorov's theory. Shackel assumes that each σ-algebra is given with a "regular" measure defined over it (i.e. a measure which is *not* a probability measure, at least not necessarily). This "regular" measure can be seen as an additional component of Kolmogorov's probability space. (By definition, the probability space includes only a probability measure component defined over the σ-algebra component). Due to this assumption, Shackel's formalization faces a serious problem: mathematically there can be infinitely many measures defined over any given σ-algebra, so it is unclear which one of these measures should be taken as the "regular" measure in Shackel's mathematical formalization of IP[16]. Without some way of choosing one of these "regular" measures, Shackel's mathematical formalization of IP's presumption is incomplete. This means that Shackel's formalization of IP remains ill-defined and hence, despite its appeal, cannot serve as a mathematical formalization of IP.

---

[15] Shackel relies on the fact that a probability measure is a special kind of measure. Conditions 3.1-3.3 of the aforementioned probability measure's definition are just the formal definition of the mathematical notion of measure. Any measure which satisfies condition 3.4 is a probability measure.

[16] This problem was already noticed by Gyenis and Rédei, who state: "[...] there are infinitely many measures μ on *S* [a Boolean σ-algebra of certain subsets of *X*] that could in principle be taken as ones that define a probability *p*. Which one should be singled out that yields a *p* that could in principle be interpreted as expressing epistemic indifference about elements in *X*?" (Gyenis and Rédei 2014, 7)

The mathematical formalization of IP I have described above clarifies another issue concerning IP-related paradoxes. Roughly, it shows that there is a certain type of solutions which is simply misguided. In particular, some writers (such as Di Porto et al. (2010, 2011)) believe that probabilities can *only* be obtained by measuring actual frequencies (since they adhere, perhaps implicitly, to an objective interpretation of probability). This means that in their view IP-related paradoxes can be solved by measuring the relevant frequencies. I call this type of solutions the "physical approach". The above mathematical formalization of IP shows that this approach cannot succeed.

Recall that every IP-related paradox includes a question about the probability of a given event and that this question is given different answers by different applications of IP. Roughly, a physical-approach solution to a given IP-related paradox "solves" the paradox by running the "right" experiment. The right experiment is supposed to express the correct way that IP should be applied. More precisely, a physical-approach solution to an IP-related paradox is the claim that the probability of the event in question can be obtained by measuring the relative frequency of the event when running the experiment where IP is applied correctly.

However, the physical-approach to IP-related paradoxes simply does not address the real problem they raise. As I have shown above, the mathematical formalization of IP is a set of constraints (*C*) on same-events spaces (*S*). The problem is finding a *C* such that for every *S* there is only one probability space in *S* which satisfies the constraints in *C*. This is a mathematical problem. As such, it does not depend on a particular interpretation of probability. In particular, it does not depend on the interpretation being an objective one or on relative frequencies.

It is indeed commonly held that there is a connection between relative frequencies and probabilities (expressed by the laws of large numbers). This means that relative frequencies can confirm or cast doubts on apriori assignments of probabilities. However, and this is the crucial point, relative frequencies can confirm or undermine assignments of probabilities only after these assignments have been made! This means that they can confirm or undermine probabilities determined by applications of IP only *after* IP has been used. And IP can be used only after it has been mathematically formalized. Hence measuring relative frequencies cannot serve as a mean for mathematically formalizing IP and as a result it cannot solve IP-related paradoxes.

6. The implication of IP's formalization for its validity

The major question concerning IP's mathematical formalization as a set of constraints on same-events spaces is whether such a set exists. Is there a set of constraints (*C*) which manages to constrain *every* set of same-events spaces in such a way that only one space in each of these sets satisfies *C*? This question is highly important, mainly because a positive answer would settle the debate concerning whether or not IP-related paradoxes are genuine paradoxes. A positive answer would mean that there is *at least one* set of constraints that is always satisfied by only one probability space in every set of same-events spaces. In other words, there are no cases when IP's set of constraints (*C*) fails to constrain a given set of same-events spaces (*S*) so that only one probability space in *S* satisfies the constraints in *C*. Hence if the answer to the existence question is positive, then there are no cases of IP-related paradoxes. This means that all IP-related paradoxes are not genuine paradoxes and the aforementioned debate is settled.[17]

On the other hand, a negative answer to the existence question means that there is no mathematical formalization of IP as a set of constraints (*C*) that is always satisfied by only one probability space in *every* set of same-events spaces (*S*). In other words, there are cases where either zero or more than one probability space in the given *S* satisfy *C*. This means that there is at least one genuine IP-related paradox. Thus, it is possible that one or more of Bertrand's Paradox, the Box Factory Paradox and the Wine-Water Paradox are indeed genuine IP-related paradoxes.

A negative answer to the existence question can also be seen as a *solution* to IP-related paradoxes in the sense that it refutes one of their premises. Recall that IP-related paradoxes rely on the assumption that IP is always sufficient for a unique assignment of probabilities to events. However, a negative answer to the existence question would be a mathematical proof that this assumption is plain wrong. This means that a negative answer would refute one of the premises of IP-related paradoxes and thus solve them.

---

[17] A positive answer to the existence question raises an interesting question regarding the uniqueness of *C*: is this *C* unique or is there more than one such set of constraints? I do not address this question in this paper.

A negative answer to the existence question would also imply that IP is not a valid principle in the sense that applying it does not *always* result with a unique assignment of probabilities to events. As a result, the discussion surrounding IP should change: Instead of focusing on whether or not IP is a valid principle (since a negative answer to the existence question shows that it is not), the discussion should be focused on describing the cases where it is safe to use IP, if there are any.

The above discussion regarding the existence question sheds new light on Jaynes' famous paper about Bertrand's Paradox, "The Well-Posed Problem". In this paper Jaynes presents one of the most famous solutions offered for Bertrand's Paradox. Roughly, Jaynes believes that the question in Bertrand's Paradox, as in any other probability problem, must have a unique solution: "The essential point is this: If we start with the assumption that Bertrand's problem has a definite solution *in spite of the many things left unspecified*, then the statement of the problem automatically implies certain invariance properties [...]" (Jaynes 1973, 480). Jaynes adds that "[t]he transformation group, which expresses these invariances mathematically, imposes definite restrictions on the form of the solution, and in many cases fully determines it." (Jaynes 1973, 488). Recall that a solution to Bertrand's Paradox is the probability of the event that a random chord of a circle is longer than the side of an inscribed equilateral triangle. This solution is mathematically described by a probability space. Hence Jaynes' invariance requirements for a solution to Bertrand's Paradox can be thought of as (and even translated into) a set of constraints on the set of same-events spaces which mathematically describe the question in Bertrand's Paradox.

Now the questions are: what exactly is Jaynes' set of constraints (C_j) and, more importantly, does it provide a positive answer to the existence question? In other words, does C_j manage to constrain *every* set of same-events spaces in such a way that only one space in each of these sets satisfies C_j? Unfortunately, it seems that the answer is 'no', even according to Jaynes: "There remains the interesting, and still unanswered, question of how to define precisely the class of problems which can be solved by the method illustrated here. There are many problems in which we do not see how to apply it unambiguously; von Mises' water-and-wine problem is a good example." (Jaynes 1973, 490). In other words, Jaynes believes that his set of constraints does not manage to constrain every set of same-events spaces to just one. In particular, Jaynes' set of constraints fails in the case of the Wine-Water Paradox[18]. This means that Jaynes does not provide a positive answer for the existence question and the status of IP as a valid principle remains unknown.

I hope that the above discussion clarifies the importance of the existence question of IP's mathematical formalization. However, despite its importance, to the best of my knowledge there has not been any explicit attempt to answer it. This is mainly because IP's mathematical formalization has not been widely covered in the literature (notable exceptions are Shackel (2007) (discussed above) and Gyenis and Rédei (2014)). This does not mean that IP does not have a mathematical formalization; On the contrary, since almost all writers about IP claim that they apply IP in different circumstances, it seems that each of them has in mind some implicit mathematical formalization of it. These formalizations are used implicitly in each of the writers' calculations when they apply IP. However, since commonly IP is not *explicitly* formalized, it is not clear whether it has one specific (implicit) mathematical formalization which can be considered *the* mathematical formalization of IP. More importantly, this means that currently the existence question is still unanswered and thus the question regarding IP's validity remains open.

[18] Moreover, Drory (2015, 458) states that Jaynes' set of constrains is not even sufficient to solve Bertrand's Paradox! He claims, "[E]ach of these solutions [the three answers that appear in Bertrand's Paradox] can be supported by invariance requirements and even by the very same requirements, in the sense that they will be all called rotation, scaling and translation invariance". In other words, Drory believes that each of the three probability spaces which mathematically describe the three answers satisfies C_j (Jaynes' invariance requirements). Hence C_j fails to be a good mathematical formalization of IP. See also Nathan (1984) for a more general objection to several of Jaynes' assumptions and Friedman (1975) for a criticism of the acceptability of Jaynes' invariance criteria.

Bibliography

Aerts, Diederik, and Massimiliano Sassoli de Bianchi. 2014. "Solving the Hard Problem of Bertrand's Paradox." *Journal of Mathematical Physics* 55 (8). https://doi.org/10.1063/1.4890291.

Bangu, Sorin. 2010. "On Bertrand's Paradox." *Analysis* 70 (1):30–35. https://doi.org/10.1093/analys/anp124.

Bartha, Paul, and R Johns. 2001. "Probability and Symmetry." *Philosophy of Science* 68 (3):S109–22. https://doi.org/10.1086/392902.

Billingsley, Patrick. 1995. *Probability and Measure*. *Wiley-Interscience*. 3rd ed. New York.

Burock, Marc. 2005. "Indifference, Sample Space, and the Wine/Water Paradox." PhilSci Archive. http://philsci-archive.pitt.edu/id/eprint/2487.

Clark, Michael. 2007. *Paradoxes from A to Z: Second Edition*. *Paradoxes from A to Z: Second Edition*. Second ed. https://doi.org/10.4324/9780203962367.

Deakin, Michael. 2006. "The Wine/Water Paradox: Background, Provenance and Proposed Resolutions." *The Australian Mathematical Society Gazette* 33 (3):200–205.

Drory, Alon. 2015. "Failure and Uses of Jaynes' Principle of Transformation Groups." *Foundations of Physics* 45 (4):439–60.

Fraassen, Bas C. van. 1989. *Laws and Symmetry*. New York: Oxford University Press. https://doi.org/10.1093/0198248601.001.0001.

Friedman, Kenneth S. 1975. "A Problem Posed." *Foundations of Physics* 5 (1):89–91. https://doi.org/10.1007/BF01100318.

Gillies, Donald. 2000. *Philosophical Theories of Probability*. *Routledge*. 3rd ed. London, NewYork.

Gyenis, Z., and M. Rédei. 2014. "Defusing Bertrand's Paradox." *British Journal for the Philosophy of Science* 0:1–25. https://doi.org/10.1093/bjps/axt036.

Hájek, Alan. 2012. "Interpretations of Probability." *The Stanford Encyclopedia of Philosophy*. http://plato.stanford.edu/archives/win2012/entries/probability-interpret/.

Howson, Colin, and Peter Urbach. 2006. *Scientific Reasoning: The Bayesian Approach*. *Open Court Company*. 3rd ed. Chicago.

Jaynes, E. T. 1973. "The Well-Posed Problem." *Foundations of Physics* 3 (4):477–92. https://doi.org/10.1007/BF00709116.

Keynes, John Maynard. 1921. *A Treatise on Probability*. London: Macmillan and Co.

Lyon, Aidan. 2010. "Philosophy of Probability." In *Philosophies of the Sciences: A Guide*, 92–125. https://doi.org/10.1002/9781444315578.ch5.

———. 2016. "Kolmogorov's Axiomatisation and Its Discontents." In *The Oxford Handbook of Probability and Philosophy*, 155–66.

Marinoff, L. 1994. "A Resolution of Bertrand's Paradox." *Philosophy of Science* 61 (1):1–24. https://doi.org/10.2307/188286.

Mellor, D H. 2005. *Probability: A Philosophical Introduction*. London, NewYork: Routledge. https://doi.org/10.4324/9780203504864.

Mikkelson, Jeffrey M. 2004. "Dissolving the Wine/Water Paradox." *British Journal for the Philosophy of Science* 55 (1):137–45. https://doi.org/10.1093/bjps/55.1.137.

Milne, Peter. 1983. "A Note on Scale Invariance." *British Journal for the Philosophy of Science* 34 (1):49–55. https://doi.org/10.1093/bjps/34.1.49.

Mises, Richard Von. 1981. *Probability, Statistics and Truth*. *Dover Publications*. 2d rev. en. New York: Dover Publications.

Nathan, A. 1984. "The Fallacy of Intrinsic Distributions." *Philosophy of Science* 51 (4):677–84. http://www.jstor.org/stable/187982.

Norton, John D. 2008. "Ignorance and Indifference." *Philosophy of Science* 75 (1):45–68. https://doi.org/10.1086/587822.

November, Dan D. 2018a. "Interpretive Implications of the Sample Space." http://philsci-archive.pitt.edu/15588/.

———. 2018b. "Zero Probability." http://philsci-archive.pitt.edu/15590/.

Plato, Jan Von. 1994. *Creating Modern Probability: Its Mathematics, Physics, and Philosophy in Historical Perspective*. Cambridge University Press.

Porto, P Di, B Crosignani, A Ciattoni, and H C Liu. 2010. "Bertrand's Paradox: A Physical Solution." *Solutions*, August, 3. http://arxiv.org/abs/1008.1878.

———. 2011. "Bertrand's Paradox: A Physical Way out along the Lines of Buffon's Needle Throwing Experiment." *European Journal of Physics* 32 (3):819–25.

Potter, Michael D. 2004. *Set Theory and Its Philosophy: A Critical Introduction*. New York: Oxford University Press.

Rowbottom, Darrell P. 2013. "Bertrand's Paradox Revisited: Why Bertrand's 'solutions' Are All Inapplicable." *Philosophia Mathematica* 21 (1):110–14. https://doi.org/10.1093/philmat/nks028.

Shackel, Nicholas. 2007. "Bertrand's Paradox and the Principle of Indifference." *Philosophy of Science* 74 (April):150–75. http://www.jstor.org/stable/10.1086/519028.

Strevens, M. 1998. "Inferring Probabilities from Symmetries." *Noûs* 32 (2):231–46. http://www.jstor.org/stable/2671966.

Tissier, P. E. 1984. "Bertrand's Paradox." *The Mathematical Gazette* 68 (443):15–19. https://doi.org/10.2307/3615385.