Intuitionistic probablism in epistemology

Kino Zhao
yutingz3@uci.edu

Unpublished Draft as of November 12, 2018

Abstract

This paper examines the plausibility of a thesis of probabilism that is based on intuitionistic logic and exposits the difficulties faced by such a program. The paper starts by motivating intuitionistic logic as the logic of investigation along a similar line as Bayesian epistemology. It then considers two existing axiom systems for intuitionistic probability functions – that of Weatherson (2003) and of Roeper and Leblanc (1999) – and discusses the relationship between the two. It will be shown that a natural adaptation of an accuracy argument in the style of Joyce (1998) and de Finetti (1974) to these systems fails. The paper concludes with some philosophical reflections on the results.

Intuitionism in mathematics refers to a class of programmes, more or less following the tradition of Brouwer, that challenges the classical way of viewing what mathematics does and what is allowed to be taken for granted in the practice of mathematics. Famously, intuitionism rejects the classical law of excluded middle (LEM), which states that $P \vee \neg P$ can be asserted truthfully at any time, regardless of what P is. Intuitionistically, P can only be asserted if there is some constructive procedure that establishes P, whereas $\neg P$ can only be asserted if there is some constructive procedure that leads one to conclude, from P, a contradiction[1]. If one cannot accomplish either of these for some given P, then one cannot assert $P \vee \neg P$.

A major motivation behind this view in the philosophy of mathematics is that, if one finds the idea of mathematical objects existing in an acausal Platonic heaven problematic, then one is likely to also find the claim that we can make true or false statements

[1] The gloss here follows (roughly) the interpretation of intuitionistic logic found in Heyting (1956). Different schools of intuitionism and constructive mathematics at large differ on what counts as a legitimate constructive procedure, but this is beyond the point of the current paper.

about mathematical objects we have never thought about before problematic. Instead, intuitionists hold that statements about some generic, unknown mathematical object do not have truth values until we can have a reasonably direct access to it through construction.

One might have similar worries for sentences in non-mathematical domains. For example, one might think that future contingents are, at the present time, neither true nor false. This is raised by Weatherson (2003) as one of the considerations that might cause one to question the law of excluded middle as it applies to future contingents. It is not true, now, that there will be a sea battle tomorrow. Neither is it true, now, that there won't be. Hence, one cannot truthfully claim, now, that "either there will be, or there won't be, a sea battle tomorrow". The view is that one should be allowed to refrain from making claims about matthers about which one has no evidence one way or another.

Intuitionistic logic thus has a natural, albeit underexplored, application to the epistemology of evidence evaluation. "Subjective probabilistic semantics should yield intuitionism", declares Harman (1983), "since what settles a bet is not the mere truth or falsity of a proposition but rather the *discovery* that the proposition is true or the *discovery* that it is false". Similarly, van Benthem (2009) observes that intuitionistic logic registers "*procedural information* about our current investigative process" (p.255, emphasis original). The thought is that the kind of "knowledge gap" that exists before we have complete information on a subject matter is better reflected using a logic that allows "truth value gaps" in a similar way.

In formal epistemology, the well established Bayesian tradition studies evidence evaluation by developing formal procedures that guide one's treatment of each piece of relevant information. Central to Bayesian epistemology is the thesis of probabilism, which states that a rational agent should have a credence function that obeys the laws of probability theory (e.g. Jeffrey, 1992). Little attention has been paid to the logic underlying such functions, but it has traditionally been assumed that the logic is classical. LEM is either listed as an axiom (such as in Joyce, 1998) or derivable as a consequence of defining probability functions from a Boolean algebra. Let us call *classical probabilism* the thesis that requires a rational agent's credence function to be a classical probability function.

The current paper investigates the plausibility of an intuitionistic probabilism – one that requires a rational agent's credence functions to be intuitionistic probability functions. It does so by considering two existing axiom systems for intuitionistic probability functions: that of Weatherson (2003) and of Roeper and Leblanc (1999). The paper examines the relationship between these two systems and the difficulties of adapting an accuracy argument in the style of Joyce (1998) and de Finetti (1974) to these systems. It will be shown that neither of these two systems allows a straightforward adaptation of the accuracy argument. This is partly due to the fact that there exist infinitely many valuations even just for finitely many proposition letters (a result by Gödel, 1986, p.223 ff.). Having infinitely many valuation functions blocks the geometric intuition that is necessary for the accuracy argument, because the argument relies on the existence of some Euclidean-like distance that defines the inaccuracy measure. This feature is lost in the transition from a space of finite dimensions to one of infinite dimensions.

The paper is organized as follows. Section 1 motivates the attempt at adapting intuitionistic logic to Bayesian epistemology. Section 2 presents a brief formal introduction to intuitionistic logic and its semantics. Section 3 introduces the two axiom systems mentioned above. Section 4 discusses a natural way of generalizing the accuracy argument to the system of intuitionistic probability functions provided in Roeper and Leblanc (1999) and argues that the natural proof strategy fails. It will be clear how the failure of this adaptation also precludes the success of an adaptation of the Weatherson axioms. Section 5 concludes with philosophical reflections on the results.

1 Intuitionism in epistemology

As briefly mentioned before, intuitionism in mathematics is (partly) motivated by a skeptism over mathematical Platonism. The thought is that we would like to ground the truth of mathematical claims not in some facts in a mind-independent world of abstracta, but in some mind-dependent processes like proofs and constructions. Relatedly, if one would like to reject the mind-independent existence of mathematical entities, then one would not want the meaning of mathematical claims to be dependent on these (non-existent) entities in the same way that the sense of empirical claims seem to depend on observable objects or properties. To address this worry, the intuitionis holds that "the meaning of a sentence is to be given, not by the conditions under which it is true, where truth is conceived as a relationship with some external reality, but by the conditions under which it is proved, its proof conditions – where a proof is a (mental) construction of a certain kind" (Priest, 2008, p.138).

We may take the same considerations outside of mathematics and talk about evidence instead of constructions and verification instead of proofs. That is, we may require the assertion of a claim to be based on evidence in a strong sense, such that the absence of evidence does not count as evidence of absence. Similarly, we might want to have a theory of meaning where the meaning of a sentence is given not by its truth conditions, but by its verification conditions.

This last consideration is also discussed in Weatherson (2003), who observes that "[a] standard objection to classical Bayesianism is that it has no way of representing complete uncertainty" (p. 114). Even if I have no evidence for or against p, the classical Bayesian still requires me to assign credence 1 to the classical tautology $p \vee \neg p$. By the Axiom of additivity, this means I must have non-zero credence assignment to p and to $\neg p$. If one does not want to give up additivity, one needs to seriously consider giving up $p \vee \neg p$ as a tautology.

Here is another way of understanding the intuitionistic worldview. It is often said that classical logic is the logic of truth. One way to think about classical possible worlds is to see each world as a full description of a possible state of affairs or a maximally consistent set of true sentences. That is, for every proposition, if it is not true at a world, then its negation is true at that world (hence LEM). A classical world is a possible version of reality, where every question has a definitive answer. Conversely, an intuitionistic world is a state of knowledge during an investigation, where one has a consistent set of beliefs but also open questions that need to be answered.

It is worth noting that intuitionistic logic is not the only, or even the primary, logic that is used to handle truth gaps in the literature. There are a number of three-valued logics that were designed to manage neither-true-nor-false sentences like fictional contingents or future contingents as well as both-true-and-false sentences like the Liar sentence (see, e.g., Priest, 2008). As we will discuss later in the paper, probabilism behaves much better with finitely valued logics than it does with intuitionistic logic.

Nevertheless, intuitionistic logic has a natural place as the logic of investigation. A truth gap in a fiction, such as the colour of Sherlock Holmes's shirt on a particular day, has a certain sense of inevitability to it – that is, it seems to be the kind of thing that we cannot know *in principle*, because there is nothing to be known. Similarly, people who reject classical logic for future contingents tend to do so because they believe that future contingents do not yet have a truth value at all and so, again, are in principle unknowable. Both of these ways of seeing truth gaps may motivate the introduction of a third truth value – an *in principle unknowable* value – which we can then develop a logic for. However, an open question in an investigation is not especially different from an answered question. They are formulated in similar ways and, presumeably, call for the same kind of answers or answer strategies. The only difference is that one is currently known and the other is currently unknown. Representing an open question as a lack of truth values in the intuitionistic sense rather than as something completely different in three-valued logics seems much more natural.

The above reasoning is not meant to be a knockdown argument for the superiority of intuitionistic logic over three-valued logics in epistemology. Rather, it is meant to motivate intuitionistic logic as a natural candidate to consider as a replacement of classical logic.

To conclude this section, let me lay out the motivation once again: the central feature of Bayesian epistemology is its ability to guide investigation and the resulting changing of beliefs. The thesis of probabilism states that an agent's system of credences needs to satisfy particular constraints in order to be successful in the Bayesian framework. While classical logic is often said to be the logic of truth, intuitionistic logic can be seen as the logic of investigation. Consequently, it is natural to examine the plausibility of probabilism with intuitionistic logic as its underlying logic. The next section lays out both the formal framework and the philosophical conceptualization of what an intuitionistic probabilism might mean.

2 Kripke semantics for intuitionism

The use of Kripke semantics for intuitionistic logic is controversial. For example, Dummett (1977, p.166) writes[2]:

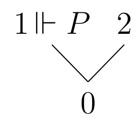
> Since intuitionistic logic is founded on a rejection of the whole notion of objectively determined truth-values independent of our capacity for recognizing them, it does not appear as if valuation systems are going to be of any help towards formulating notions of completeness for intuitionistic logic.

Later in this section, however, I will attempt to motivate the adoption of Kripke semantics along the line of intuitionistic logic as the logic of investigation sketched out

[2]Troelstra and van Dalen (1988, p.75) express a similar worry about Kripke semantics in particular: that it does not connect to the BHK interpretation of intuitionistic appoaches to mathematics.

above. For a more extensive introduction to either Kripke semantics or other semantics for intuitionism, see Troelstra and van Dalen (1988) and Chagrov and Zakharyaschev (1997, Ch.2).

Formally, a Kripke model is given by a triple $\mathcal{K} = \langle K, \leq, \Vdash \rangle$, where K is a set of worlds or *nodes*, $\leq$ is an asymmetric *accessibility relation* among those worlds, and $\Vdash$ is a *forcing* relation between worlds and propositions. A typical Kripke model is often drawn as a tree structure[3], where the accessibility relations "grow" only upward. An example tree is given below[4].



The truth conditions for all nodes $k \in K$ obey the following rules (see also Kripke, 1963; Troelstra and van Dalen, 1988, p.77).

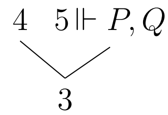
- $k \Vdash P$ and $k \leq k'$ implies $k' \Vdash P$.
- $k \Vdash A \wedge B$ if and only if $k \Vdash A$ and $k \Vdash B$.
- $k \Vdash A \vee B$ if and only if $k \Vdash A$ or $k \Vdash B$.
- $k \Vdash A \rightarrow B$ if and only if for all k' with $k \leq k'$, if $k' \Vdash A$ then $k' \Vdash B$.
- $k \Vdash \neg A$ if and only if for all k' with $k \leq k'$, $k' \nVdash A$.

In the above toy model, we see that P and hence $P \vee Q$ (where Q is arbitrary) are both true at world 1. However, $\neg P$ is true at world 2 but not at world 0. This means that $P \vee \neg P$ is not true at world 0.

Connectives like $\vee$ and $\wedge$ are treated straightforwardly in ways similar to classical logic. The material conditional $\rightarrow$, however, requires special care. Intuitionistically, $P \rightarrow Q$ is true at a node just in case, for every node above it, if P holds then Q holds. One side effect of this definition is that the equivalence between $P \rightarrow Q$ and $\neg P \vee Q$ is not a valid in intuitionistic logic. Consider the following counter model:

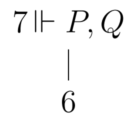
[3] Strictly speaking, Kripke models are not proper trees, since it is permissible that the branches merge. Since nothing hinges on this technical point, we will occasionally refer to these models as trees for ease of presentation.

[4] The worlds need not be, and often are not, labelled. I have labelled them here for clarity of explanation.



In this model, $P \to Q$ holds on all three nodes, and yet neither $\neg P$ nor Q holds on node 3, which means $\neg P \vee Q$ does not hold on node 3.

Similarly, $P \wedge Q$ is not equivalent to $\neg(\neg P \vee \neg Q)$ like the classically valid De Morgan's law. To see this, consider the following counter model:



Here, $\neg P \vee \neg Q$ does not hold at 6, or 7, which means $\neg(\neg P \vee \neg Q)$ holds at 6. But $P \wedge Q$ does not hold at world 6.

The feature that $\wedge$ and $\to$ are not reducible to $\neg$ and $\vee$ will turn out to be significant for our discussion of Dutch book arguments later. For now, let us focus on how a probability theory can be built on intuitionistic logic.

To see how this semantics fits in with our narrative about investigations, notice first that the nodes at the end of the branches are classical. That is, every proposition is either forced on that node, or its negation is. Further, if node j follows node i, or $i \leq j$ in our notation, then the set of propositions true in i is a subset of the propositions true in j. A branch, then, can be seen as an investigative path, where more and more information are acquired as we move from node to node, and the path leads to a possible state of affairs. Finally, we can interpret a tree as a theory (in the nontechnical sense) about the different ways in which an investigation may unfold.

In the classical setting, each world assigns a valuation 0 or 1 to each proposition p. Suppose a number n of worlds are indistinguishable to me (i.e., my information state is insufficient to pick out the real world), and they differ on their valuations of p, then I assign a probability of p between 0 and 1. Formally, suppose we have $v_1, \ldots, v_n$ as valuations of n worlds. Suppose further we have constants $a_1, \ldots, a_n$, with $0 \leq a_i \leq 1$ and $\sum_{i=1}^{n} a_i = 1$. Then, classically, $Pr(P) = \sum_{i=1}^{n} a_i v_i(P)$ is a probability function. This type of weighted averaging of valuations is often called convex combinations, or probabilistic averages, of valuations. These functions form the *convex hull* of the valuations, which is the smallest convex set containing the valuations (Aliprantis and Border, 2006, p.182-183). It is a key insight of de Finetti (1992) that the "coherent" credence functions are exactly those that form the convex hull of classical truth values.

This grounds the important geometric intuitions that support the later stages of the accuracy argument (e.g., Joyce, 1998, p.582-3 and Pettigrew, 2016, p.82, 107 ff.).

Moving to the intuitionistic setting, we can make sense of probability in a similar way. We can treat each tree as a world in the sense that we can see a number of different investigative paths being compatible with our current state, and they may lead to a number of different conclusions. However, I may be wrong about what the current state is (i.e. which tree structure I'm working with), as well as which investigative step I'm currently at (i.e., which node I'm on).

In the current paper, I will treat convex combinations of intuitionistic valuation functions in what I consider as a natural way, following the above narrative. Let w be valuation functions on nodes. For any proposition P on any node i of a Kripke tree, $w_i(P) = 1$ if that node forces P, $w_i(P) = 0$ if it does not. For a Kripke tree with m many nodes, let there be $b_1, \ldots, b_m$ with $0 \leq b_i \leq 1$ and $\sum_{i=1}^{m} b_i = 1$, let us set $v(P) = \sum_{i=1}^{m} b_i w_i(P)$. That is, v of P of a model is the convex combination of the pure world valuation of the nodes. We then define probability of P as convex combination of v's of different models in the same way: $Pr(P) = \sum_{i=1}^{n} a_i v_i(P)$ for $0 \leq a_i \leq 1$ and $\sum_{i=1}^{n} a_i = 1$. Philosophically, one can interpret this as an agent's not being able to discern which structure is the true structure and which world in that structure is the true world. In what follows, I call any functions that can be derived in the above way *convex combinations of intuitionistic valuation functions.*

3 Two systems of intuitionistic probability

This section introduces two existing sets of axioms for intuitionistic probability functions, and discusses their relationship.

Weatherson (2003) introduces a 4-axiom system that he adapts from the classical 3-axiom system of Kolmogorov (1960) they are as follows:

- (P0) $\mathbb{P}(\top) = 1$
- (P1) $\mathbb{P}(\bot) = 0$
- (P2) If $A \vdash B$ then $\mathbb{P}(A) \leq \mathbb{P}(B)$
- (P3) $\mathbb{P}(A \vee B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \wedge B)$

He then develops a Dutch book argument and a converse Dutch book argument for these axioms. I will briefly mention the implications of these arguments later in this section. It is worth noting that, in Weatherson's set up, if I buy into a bet on p and $\neg(p \vee \neg p)$ happens, then I lose the bet on p. This way of setting up the bet blurs the distinction between intuitionistic logic and three-valued logics. As we will see, this difference is significant for accuracy argument.

Another feature to note about these axioms is that they are fairly opaque. With classical probability functions, one only needs three axioms describing the behaviour of $\neg$ and $\vee$. Classical logic is semantically complete with just those two connectives. However, as we saw in section 2, connectives like $\rightarrow$ and $\wedge$ are not interdefinable with $\neg$ and $\vee$. We should expect an adequate axiom system of intuitionistic probability function to also guide us on our treatment of those connectives.

The axiom system provided in Roeper and Leblanc (1999) is significantly more expansive for just this reason. Their system contains 9 axioms (listed in Appendix A). In addition to giving explicit rules governing connectives like $\wedge$ and $\rightarrow$, these axioms also have the following nice properties (Roeper and Leblanc, 1999, p.229, 231):

Theorem 3.1 (Weak Soundness)**.** *If P is any absolute probability function for intuitionistic propositional logic and A any theorem of that logic, then $P(A) = 1$.*

Theorem 3.2 (Weak Completeness)**.** *If A is a statement of L that is not a theorem of intuitionistic propositional logic, then there exists an absolute probability function P for intuitionistic propositional logic with $P(A) \neq 1$.*

More relevant to our narrative, these axioms – I shall call them the RL axioms – are developed using a Kripke semantics in the way described in section 2 above. In the process of proving the weak completeness theorem, Roeper and Leblanc showed that the average of all the nodes on a Kripke tree satisfies the RL axioms. With the following proposition (proved in Appendix A)

Proposition 3.3. *If $P_1, \dots P_i, \dots$ satisfy the RL axioms, and $a_1, \dots a_i, \dots$ are such that $0 \leq a_i \leq 1$, and $\sum_{i=1}^{\infty} a_i = 1$, then $\mathbb{P} = \sum_{i=1}^{\infty} a_i P_i$ also satisfy the RL axioms.*

one can conclude that any probability function generated in the way described in section 2 satisfy the RL axioms.

We now have two axiom systems that are both well motivated as candidate axiomatizations of intuitionistic probability functions: the Weatherson axioms for their natural adaptation from classical probability axioms, whereas the RL axioms for their nice properties relating to Kripke semantics for intuitionistic logic. Unfortunately, these two axiom systems do not coincide (proved in Appendix B):

Proposition 3.4. *Any function that satisfies the RL axioms also satisfies the Weatherson four axioms, but not vice versa.*

As briefly mentioned before, the Dutch book arguments presented in Weatherson (2003) do not distinguish between intuitionistic logic and a three-valued logic where the third value behaves like falsity. But if one does find his arguments convincing, the above proposition would lead one to conclude that the RL axioms are necessary but not sufficient to define intuitionistic probability functions. As we shall see in the next section, this question becomes a lot more complicated once we dig deeper into exactly what we take an intuitionistic probability function to be.

4 Accuracy argument

This section examines the prospect of an accuracy argument for intuitionistic probabilism. I first identify key elements of the classical accuracy argument, before providing a detailed explanation on how one of them fails for one system of intuitionistic probability. I conclude this section with some remarks on additional difficulties faced by intuitionistic probabilism.

The accuracy argument for classical probabilism can be roughly separated into three steps (cf. de Finetti, 1992, Joyce, 1998, Pettigrew, 2016). Firstly, one shows that classical probability functions are exactly those that are convex combinations of classical valuation functions. Secondly, one shows that one can understand inaccuracy as a "distance" measure between a credence function and a valuation function. Finally, one observes that, due to convexity of the probability functions, any credence function that is not probabilistic has a probabilistic projection that is closer (by the inaccuracy distance measure) to all the valuation functions. The below argument concerns the first step of this argument for the RL axioms.

Classically, the axioms of Kolmogorov coincide with convex combinations (also often called "probabilistic combinations") of classical valuations because the worlds are taken to form a Boolean algebra. Intuitionistically, however, this has to be argued for. This is because intuitionistic valuations are interdependent among worlds. In fact, one can show that the functions defined by the RL axioms do not coincide with convex combinations of intuitionistic worlds, at least in the way defined in section 2.

To see this, we need to look at another standard result from intuitionistic logic. Given that both modal logic and intuitionistic logic can be based on Kripke frames, it should not be surprising that there are similarities between the metatheories of these two branches of nonclassical logic. In particular, just as there are metatheorems in modal logic which relate validities to frame conditions, so there are metatheorems in intuitionistic logic which relate validities to frame conditions. One of these states that one can define a formula (cf. Chagrov and Zakharyaschev, 1997, p.44, Proposition 2.40)

$$\varphi_n = p_0 \vee (p_0 \to p_1) \vee \cdots \vee ((p_0 \wedge \cdots \wedge p_{n-1}) \to p_n) \tag{1}$$

for each $n \geq 1$, such that:

Proposition 4.1. *A frame $K = \langle W, R \rangle$ validates φ_n iff each rooted subframe of K contains $\leq n$ points.*

Using this formula, we can find a probability function $\mathbb{P}$ that satisfies the RL axioms, but cannot be written as a convex combination of intuitionistic valuation functions. The construction of this probability function proceeds as follows. Firstly, find a series of rooted frames, $K_1, K_2, \ldots$, each K_i containing exactly i many nodes. For each frame K_i, define a probability function P_{K_i} such that it satisfies RL axioms, and $P_{K_i}(\varphi_{i-1}) < 1$, where φ_0 is chosen to be some classical contradiction. This is possible for all i, since the rooted frame K_i has more than $i-1$ many nodes, and hence does not validate φ_{i-1} by Proposition 3.4.

Define $\mathbb{P} = \sum_{i=1}^{\infty} 2^{-i} P_{K_i}$. Since $0 < 2^{-i} < 1$ and $\sum_{i=1}^{\infty} 2^{-i} = 1$, and all P_{K_i} satisfy RL axioms, $\mathbb{P}$ satisfies RL axioms by Proposition 3.3. Suppose, for the sake of reductio, that $\mathbb{P}$ can be written as convex combination of valuation functions v_j. That is, there are finite models $L_1, \ldots, L_n$ such that $v_{L_1}, \ldots, v_{L_n}$ are valuation functions of $L_1, \ldots, L_n$ in the way described in section 2.2. By reductio hypothesis, $\mathbb{P} = \sum_{i=1}^{n} a_i v_i$ for some $0 < a_i < 1$ and $\sum_{i=1}^{n} a_i = 1$. Each L_i has some number of nodes that is less than some large natural number N. By Proposition 3.4, $L_1 \ldots L_n$ all validate φ_N. This means that each valuation function has to assign 1 to φ_N, and so $\mathbb{P}(\varphi_N) = 1$. However, we have defined $\mathbb{P} = \sum_{i=1}^{\infty} 2^{-i} P_{K_i}$, such that $P_{K_{N+1}}(\varphi_N) < 1$. We arrive at a contradiction.

The above serves as a proof for the following theorem:

Theorem 4.2. *There exists a function $\mathbb{P}$ that satisfies the RL axioms for intuitionistic probability functions, but is not in the convex hull of the intuitionsitic valuation functions.*

Together with Proposition 3.4, we arrive at the following conclusion: the convext combinations of intuitionistic valuations form a proper subset of functions satisfying the RL axioms, which in turn form a proper subset of functions satisfying the Weatherson axioms.

How should we make sense of the above result? A modest conclusion is that the RL or the Weatherson axioms for intuitionistic probability and the valuation functions as defined in section 2 do not fit together in the way classical accuracy argument demands. This might mean that we need to add additional axioms to the RL system, or that there is a better way of defining intuitionstic valuation functions, or that the intuitionistic accuracy argument, if there is one, is radically different from its classical counterparts.

Moreover, there is something intrinsically unfitting between the ideas of intuitionistic epistemology and accuracy dominance. Suppose we have somehow bypassed the above problem, and established that intuitionistic probability functions are exactly those that are convex combinations of intuitionistic valuation functions, however defined. The next step should be to define some kind of "distance from the truth" measure that we then argue to be shorter for elements of the convex set than elements from outside. In classical probabilism, the preferred score is the "Brier score", which is defined as $(v(P) - c(P))^2$ for any credence function c, valuation function v, and proposition P. For example, if I give P a probability of 0.6 and P is in fact true, then the inaccuracy measure for my credence is calculated as $(1 - 0.6)^2 = 0.16$.

The problem with generalizing something like the Brier score to the intuitionistic setting is that there are infinitely many intuitionistic valuations, even when there are only finitely many atomic propositions (Gödel, 1986, p.223 ff.). Hence, the natural geometric intuitions which motivate the Brier score are not obviously available in the intuitionsitic setting. To see why, consider how we visualize the "inaccuracy distance" between an estimation of a proposition and the possible valuations of that proposition: we usually see the two possible valuations T and F as the end points of the two axis on a two-dimensional plain, where the points on the plain represent the agent's crendence on the truth and falsity of that proposition. The "distance" between a credence and the true answer is then understood as the geometric distance between the two points[5]. This is the most natural way of understanding what is meant for a credence judgment to be "closer to" or "further away from" a valuation. As our world includes more and more sentences, the dimension of this geometry increases. However, as long as we don't have infinitely many valuations (which can be guaranteed, classically, by only having finitely many propositions), the same geometric intuition holds. The transition from this straightforward intuition on Euclidean distance to the case of infinite dimensions is a much more radical one, but it is one that an adequate adaption of accuracy argument to intuitionistic probabilism needs to face.

5 Concluding remarks

Those who would like to investigate non-classical probabilism are likely to be motivated by skepticisms over future contingents or completely unknown propositions in a way discussed in section 1. One key difference between someone who is epistemically concerned in this way and someone following the traditional mathematical intuitionism line is that, for the epistemist, the world is still fundamentally classical. Consequently, a successful thesis of probabilism has to assess the performance of probabilistic credence functions against a classical world. Intuitionistic logic is much more faithful to the dialectic of an incomplete investigation in a classial world than the kind of three-valued logics investigated in Williams (2012), where the world is treated as three-valued as well. This unique philosophical advantage is, I believe, largely responsible for the formal difficulties witnessed in this paper. It is much easier to say "the world is n-valued and so our credence functions should base on an n-valued logic" than it is to match an infinite-valued credence function with a bivalent world. Of course, all the formal difficulties are not meant to discount the philosophical considerations associated with this narrative. The lesson might just be that we should radically restructure what we think arguments for probabilism should be like if we want to develop an intuitionistic probabilism.

On the other hand, one might worry that there exists a kind of "double-counting" with this approach. Suppose I receive a piece of information, p, that is supposedly relevant to some other proposition, q, how should I adjust my assessment of q? Bayesian epistemology suggests that I raise or lower my existing probability of p, whereas intu-

[5] For a more detailed exposition on scoring rules, see Pettigrew (2016).

itionistic logic following the narrative of the current paper suggest that I, as it were, add p to my q-information collection. Once I finish my collection, I can declare q true (or false). These two procedures seem to be parallel ways of achieving the same end. A forciful combination of them may create limitations that neither procedure possesses on their own.

In this paper, I started by drawing attention to a natural connection between intuitionistic logic and epistemology, and used it to motivate a thesis of probabilism where the underlying logic is intuitionistic. I then described intuitionistic semantics and pointed out some difficulties in defining intuitionistic probability functions. After that, I took a closer look at the axioms provided in Weatherson (2003) and Roeper and Leblanc (1999), and discussed the relationship between the two. Finally, I showed how one natural adaptation of the accuracy argument to intuitionistic probabilism fails for both of these systems. I ended the paper with some philosophical advantages and reservations on the project of intuitionistic probabilism.

Appendix A

Let L be a set of propositions in a language. Roeper and Leblanc (1999) list the following 9 axioms for what they call absolute intuitionistic probability functions.

AI1. $0 \leq P(A)$

AI2. $\max_{A \in L}\{P(A)\} = 1$

AI3. $\max_{B \in L}\{P(A \wedge B)\} = P(A)$

AI4. $P(A \wedge B) \leq P((A \wedge B) \wedge B)$

AI5. $P(A \wedge (B \wedge C)) = P(A \wedge (C \wedge B))$

AI6. $P(A \wedge (B \wedge (C \wedge D))) = P(A \wedge ((B \wedge C) \wedge D))$

AI7. $P(A \wedge (B \vee C)) = P(A \wedge B) + P(A \wedge C) - P(A \wedge (B \wedge C))$

AI8. $P((A \supset B) \wedge C) = \max_{D \in L}\{P(C \wedge D) | P((D \wedge A) \wedge B) = P(D \wedge A)\}$

AI9. $P(A \wedge \neg B) = \max_{C \in L}\{P(A \wedge C) | P(C \wedge B) = 0\}$

The goal of this appendix is to show the following

Proposition (3.3)**.** *If $P_1, \dots P_i, \dots$ satisfy the RL axioms, and $a_1, \dots a_i, \dots$ are such that $0 \leq a_i \leq 1$, and $\sum_{i=1}^{\infty} a_i = 1$, then $\mathbb{P} = \sum_{i=1}^{\infty} a_i P_i$ also satisfy the RL axioms.*

Proof. **AI1** and **AI2** are obviously satisfied by $\mathbb{P}$.

For **AI3**, observe that

$$\max_{B \in L}\{\sum_{i=1}^{\infty} a_i P_i(A \wedge B)\} \leq \sum_{i=1}^{\infty} \max_{B \in L}\{a_i P_i(A \wedge B)\}$$

Then we have

$$\sum_{i=1}^{\infty} \max_{B \in L}\{a_i P_i(A \wedge B)\} = \sum_{i=1}^{\infty} a_i \max_{B \in L}\{P_i(A \wedge B)\} = \sum_{i=1}^{\infty} a_i P_i(A)$$

This means

$$\max_{B \in L}\{\mathbb{P}(A \wedge B)\} = \max_{B \in L}\{\sum_{i=1}^{\infty} a_i P_i(A \wedge B)\} \leq \sum_{i=1}^{\infty} a_i P_i(A) = \mathbb{P}(A)$$

For the other direction, we have

$$\mathbb{P}(A) = \sum_{i=1}^{\infty} a_i P_i(A) = \sum_{i=1}^{\infty} a_i P_i(A \wedge A) \leq \max_{B \in L}\{\sum_{i=1}^{\infty} a_i P_i(A \wedge B)\} = \max_{B \in L}\{\mathbb{P}(A \wedge B)\}$$

The proofs of **AI4-7** are similar. I will give the proof of **AI7** in detail as illustration.

$$\begin{aligned}
&\mathbb{P}(A \wedge B) + \mathbb{P}(A \wedge C) - \mathbb{P}(A \wedge (B \wedge C)) \\
=&\sum_{i=1}^{\infty} a_i P_i(A \wedge B) + \sum_{i=1}^{\infty} a_i P_i(A \wedge C) - \sum_{i=1}^{\infty} a_i P_i(A \wedge (B \wedge C)) \\
=&\sum_{i=1}^{\infty} a_i \big(P_i(A \wedge B) + P_i(A \wedge C) - P_i(A \wedge (B \wedge C))\big) \\
=&\sum_{i=1}^{\infty} a_i P_i(A \wedge (B \vee C)) \\
=&\mathbb{P}(A \wedge (B \vee C))
\end{aligned}$$

For **AI8**, consider Lemma 9 of Roeper and Leblanc (1999, p.226), we have $P(((A \supset$

$B) \wedge A) \wedge B) = P((A \supset B) \wedge A)$.

Setting $D = A \supset B$, we have $\mathbb{P}((D \wedge A) \wedge B) = \sum_{i=1}^{\infty} a_i P_i((D \wedge A) \wedge B) = \sum_{i=1}^{\infty} a_i P_i(D \wedge A) = \mathbb{P}(D \wedge A)$ by Lemma 9. This means

$$\mathbb{P}((A \supset B) \wedge C) = \mathbb{P}(D \wedge C) \leq \max_{D \in L}\{\mathbb{P}(C \wedge D) | \mathbb{P}((D \wedge A) \wedge B) = \mathbb{P}(D \wedge A)\}$$

For the other direction, consider a D that satisfies $\mathbb{P}((D \wedge A) \wedge B) = \mathbb{P}(D \wedge A)$. This means $\sum_{i=1}^{\infty} a_i P_i((D \wedge A) \wedge B) = \sum_{i=1}^{\infty} a_i P_i(D \wedge A)$. First, we need to show that for each i, $P_i((D \wedge A) \wedge B) = P_i(D \wedge A)$.

To see this, consider Lemma 1 of Roeper and Leblanc (1999, p.225), which states $P(A \wedge B) \leq P(A)$. This means $P_i((D \wedge A) \wedge B) \leq P_i(D \wedge A)$ for each P_i satisfying the RL axioms.

Let $I = \{i : P_i((D \wedge A) \wedge B) = P_i(D \wedge A)\}$ and let $J = \{j : P_j((D \wedge A) \wedge B) < P_j(D \wedge A)\}$. Since the numbers are all non-negative, we have $\sum_{i=1}^{\infty} P_i((D \wedge A) \wedge B) = \sum_{i \in I} P_i((D \wedge A) \wedge B) + \sum_{j \in J} P_j((D \wedge A) \wedge B)$ and likewise $\sum_{i=1}^{\infty} P_i(D \wedge A) = \sum_{i \in I} P_i(D \wedge A) + \sum_{j \in J} P_j(D \wedge A)$. By hypothesis, the two sums indexed by $i \in I$ are equal. But then we are left with $\sum_{j \in J} P_j((D \wedge A) \wedge B) = \sum_{j \in J} P_j(D \wedge A)$. This means that the set J must be empty. Namely, $P_i((D \wedge A) \wedge B) = P_i(D \wedge A)$ for all i.

We now have that, for all D with $\mathbb{P}((D \wedge A) \wedge B) = \mathbb{P}(D \wedge A)$, we must have $P_i((D \wedge A) \wedge B) = P_i(D \wedge A)$ for all i for that D as well. Since all P_i satisfies **AI8**, we have $P_i((A \supset B) \wedge C) = \max_{D \in L}\{P_i(C \wedge D) | P_i((D \wedge A) \wedge B) = P_i(D \wedge A)\}$. This means

$$\begin{aligned}\sum_{i=1}^{\infty} a_i P_i((A \supset B) \wedge C) &= \sum_{i=1}^{\infty} a_i \max_{D \in L}\{P_i(C \wedge D) | P_i((D \wedge A) \wedge B) = P_i(D \wedge A)\} \\ &\geq \max_{D \in L}\{\sum_{i=1}^{\infty} a_i P_i(C \wedge D) | P_i((D \wedge A) \wedge B) = P_i(D \wedge A)\}\end{aligned}$$

$$\mathbb{P}((A \supset B) \wedge C) \geq \max_{D \in L}\{\mathbb{P}(C \wedge D) | \mathbb{P}((D \wedge A) \wedge B) = \mathbb{P}(D \wedge A)\}$$

For **AI9**, since we have $\neg B \equiv B \supset \bot$. By **AI8**, we have $\mathbb{P}(A \wedge (B \supset \bot)) = \max_{C \in L}\{\mathbb{P}(A \wedge C) | P(C \wedge B) = \mathbb{P}((C \wedge B) \wedge \bot)\}$. Since $P_i((C \wedge B) \wedge \bot) = 0$ for any P_i satisfying RL axioms, **AI9** follows.

□

Appendix B

Proposition .1 (3.4)**.** *Any function that satisfies the RL axioms also satisfies the Weatherson four axioms, but not vice versa.*

Proof. For $\mathbb{P} \models RL \Rightarrow \mathbb{P} \models W$:

Observe that (P0) is equivalent to the weak soundness theory of Roeper and Leblanc (1999), while (P3) corresponds to their Lemma 6 (p.225, immediately provable from RL axiom AI7 and AI9, both listed below in the appendix).

For (P1), we start with Lemma 10 (p.226), which states $\mathbb{P}(\top \wedge \neg\top) = 0$. Together with the observation that $\bot \equiv \neg\top$, we get $\mathbb{P}(\top \wedge \bot) = 0$. According to lemma 11 (p.226) and (P0), we have $\mathbb{P}(\top \wedge \bot) = \mathbb{P}(\bot) = 0$.

For (P2), suppose $A \vdash B$, then $\vdash A \to B$. By weak soundness, we have $\mathbb{P}(A \to B) = 1$. By Lemma 12 (p.226), we have $\mathbb{P}(A) = \mathbb{P}((A \to B) \wedge A)$. Since $\mathbb{P}(A \to B) = 1$, we have $\mathbb{P}((A \to B) \wedge A) \leq \mathbb{P}(A \wedge B)$ by Lemma 9 (p.226). By Lemma 1 (p.225), we have $\mathbb{P}(A \wedge B) \leq \mathbb{P}(B)$. We have thus established $\mathbb{P}(A) \leq \mathbb{P}(B)$.

For $\mathbb{P} \models W \not\Rightarrow \mathbb{P} \models RL$:

consider such $\mathbb{P}$, defined as follows

$$\mathbb{P}(\varphi) = \begin{cases} 1 & \text{if } IPC \vdash \varphi \\ 0 & \text{if } IPC \nvdash \varphi \end{cases}$$

where IPC is intuitionistic propositional calculus.

It is easy to check that $\mathbb{P}$ satisfies Weatherson's P0 - P2.

Recall, P3 says: $\mathbb{P}(A \vee B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \wedge B)$

Case 1: $IPC \vdash A$, and $IPC \vdash B$, then $\mathbb{P}(A \vee B) = \mathbb{P}(A \wedge B) = \mathbb{P}(A) = \mathbb{P}(B) = 1$. P3 holds.

Case 2: $IPC \vdash A$, but $IPC \nvdash B$, then $\mathbb{P}(A \vee B) = \mathbb{P}(A) = 1$ and $\mathbb{P}(A \wedge B) = \mathbb{P}(B) = 0$. P3 holds.

Case 3: $IPC \nvdash A$, and $IPC \nvdash B$, then $IPC \nvdash A \vee B$, $\mathbb{P}(A \vee B) = \mathbb{P}(A) = \mathbb{P}(B) = 0$ and $\mathbb{P}(A \wedge B) = 0$. P3 holds.

To see how $\mathbb{P} \not\models RL$, consider Axiom 8 of RL which states $\mathbb{P}((A \supset B) \wedge C) = \max_{D \in L}\{\mathbb{P}(C \wedge D) | \mathbb{P}((D \wedge A) \wedge B) = \mathbb{P}(D \wedge A)\}$

Let C and D be tautologies, and $A = \neg\neg B$, and B be atomic. Then:

Since $IPC \nvdash A \supset B$, $IPC \nvdash (A \supset B) \wedge C$, $\mathbb{P}((A \supset B) \wedge C) = 0$.

Since B is atomic and $A = \neg\neg B$, $IPC \nvdash (D \wedge A) \wedge B$ and $IPC \nvdash D \wedge A$. This means $\mathbb{P}((D \wedge A) \wedge B) = 0 = \mathbb{P}(D \wedge A)$. We have $\max_{D \in L}\{\mathbb{P}(C \wedge D) | \mathbb{P}((D \wedge A) \wedge B) = \mathbb{P}(D \wedge A)\} = 1 \neq \mathbb{P}((A \supset B) \wedge C)$. This probability assignment does not satisfy RL's Axiom 8. $\square$

Bibliography

Aliprantis, C. D. and Border, K. C. (2006). *Infinite dimensional analysis.* Springer, Berlin, third edition. A hitchhikers guide.

Chagrov, A. and Zakharyaschev, M. (1997). *Modal logic*, volume 35 of *Oxford Logic Guides.* The Clarendon Press, Oxford University Press, New York.

de Finetti, B. (1974). *Theory of Probability: A Critical Introductory Treatment. Transl. by Antonio Machi and Adrian Smith.* J. Wiley.

de Finetti, B. (1992). Foresight: Its Logical Laws, Its Subjective Sources. In Kotz, S. and Johnson, N. L., editors, *Breakthroughs in Statistics. Vol. I*, Springer Series in Statistics: Perspectives in Statistics, pages 134–174. Springer, New York.

Dummett, M. (1977). *Elements of Intuitionism.* Oxford Logic Guides. Clarendon, Oxford.

Gödel, K. (1986). *Collected Works. Vol. I. Publications 1929-1936.* Clarendon, New York.

Harman, G. (1983). Problems with probabilistic semantics. In Orenstein, A. and Stern, R., editors, *Developments in Semantics*, pages 243–237. Haven.

Heyting, A. (1956). *Intuitionism. An Introduction.* North-Holland, Amsterdam.

Jeffrey, R. (1992). *Probability and the Art of Judgment.* Cambridge University Press.

Joyce, J. M. (1998). A nonpragmatic vindication of probabilism. *Philosophy of science*, 65(4):575–603.

Kolmogorov, A. N. (1960). *Foundations of the Theory of Probability.* Chelsea Pub Co, 2 edition.

Kripke, S. A. (1963). Semantical analysis of modal logic I. Normal propositional calculi. *Zeitschrift fur mathematische Logik und Grundlagen der Mathematik*, 9(56):67–96.

Pettigrew, R. (2016). *Accuracy and the Laws of Credence.* Oxford University Press.

Priest, G. (2008). *An introduction to non-classical logic: From if to is.* Cambridge University Press.

Roeper, P. and Leblanc, H. (1999). Absolute probability functions for intuitionistic propositional logic. *Journal of philosophical logic*, 28(3):223–234.

Troelstra, A. S. and van Dalen, D. (1988). *Constructivism in Mathematics. An Introduction. Volume 1.* Elsevier, Amsterdam.

van Benthem, J. (2009). The information in intuitionistic logic. *Synthese*, 167(2, Knowledge, Rationality & Action):251–270.

Weatherson, B. (2003). From classical to intuitionistic probability. *Notre Dame Journal of Formal Logic*, 44(2):111–123.

Williams, J. R. G. (2012). Generalized probabilism: Dutch books and accuracy domination. *Journal of philosophical logic*, 41(5):811–840.