Dogmatism, Probability and Logical Uncertainty

David Jehle   Brian Weatherson

2012

Many epistemologists hold that an agent can come to justifiably believe that p is true by seeing that it appears that p is true, without having any antecedent reason to believe that visual impressions are generally reliable. Certain reliabilists think this, at least if the agent's vision is generally reliable. And it is a central tenet of dogmatism (as described by James Pryor) that this is possible. Against these positions it has been argued (e.g. by Stewart Cohen and Roger White) that this violates some principles from probabilistic learning theory. To see the problem, let's note what the dogmatist thinks we can learn by paying attention to how things appear. (The reliabilist says the same things, but we'll focus on the dogmatist.)

Many epistemologists hold that an agent can come to justifiably believe that p is true by seeing that it appears that p is true, without having any antecedent reason to believe that visual impressions are generally reliable. Certain reliabilists think this, at least if the agent's vision is generally reliable. And it is a central tenet of dogmatism (as described by Pryor (2000) and Pryor (2004)) that this is possible. Against these positions it has been argued (e.g. by Cohen (2005) and White (2006)) that this violates some principles from probabilistic learning theory. To see the problem, let's note what the dogmatist thinks we can learn by paying attention to how things appear. (The reliabilist says the same things, but we'll focus on the dogmatist.)

Suppose an agent receives an appearance that p, and comes to believe that p. Letting Ap be the proposition that it appears to the agent that p, and $\rightarrow$ be the material conditional, we can say that the agent learns that p, and hence is in a position to infer $Ap \rightarrow p$, once they receive the evidence Ap.[1] This is surprising, because we can prove the following.[2]

Theorem 1
If Pr is a classical probability function, then
$\Pr(Ap \rightarrow p \mid Ap) \leqslant \Pr(Ap \rightarrow p)$.

---

[1] We're assuming here that the agent's evidence really is Ap, not p. That's a controversial assumption, but it isn't at issue in this debate.

[2] Popper and Miller (1987) prove a stronger result than Theorem One, and note its significance for probabilistic models of learning.

(All the theorems are proved in the appendix.) We can restate Theorem 1 in the following way, using classically equvalent formulations of the material conditional.

Theorem 2

If Pr is a classical probability function, then

- $\Pr(\neg(Ap \wedge \neg p) \mid Ap) \leqslant \Pr(\neg(Ap \wedge \neg p))$; and
- $\Pr(\neg Ap \vee p \mid Ap) \leqslant \Pr(\neg Ap \vee \mathrm{p})$.

And that's a problem for the dogmatist if we make the standard Bayesian assumption that some evidence E is only evidence for hypothesis H if $\Pr(H \mid E) > \Pr(H)$. For here we have cases where the evidence the agent receives does not raise the probability of $Ap \rightarrow p$, $\neg(Ap \wedge \neg p)$ or $\neg Ap \vee p$, so the agent has not received any evidence for them, but getting this evidence takes them from not having a reason to believe these propositions to having a reason to get them.

In this paper, we offer a novel response for the dogmatist. The proof of Theorem 1 makes crucial use of the logical equivalence between $Ap \rightarrow p$ and $((Ap \rightarrow p) \wedge Ap) \vee ((Ap \rightarrow p) \wedge \neg Ap)$. These propositions are equivalent in classical logic, but they are not equivalent in intuitionistic logic. Exploiting this non-equivalence, we derive two claims. In Section 1 we show that Theorems 1 and 2 fail in intuitionistic probability theory. In Section 2 we consider how an agent who is unsure whether classical or intuitionistic logic is correct should apportion their credences. We conclude that for such an agent, theorems analogous to Theorems 1 and 2 fail even if the agent thinks it extremely unlikely that intuitionistic logic is the correct logic. The upshot is that if it is rationally permissible to be even a little unsure whether classical or intuitionistic logic is correct, it is possible that getting evidence that Ap raises the rational credibility of $Ap \rightarrow p$, $\neg(Ap \wedge \neg p)$ and $\neg Ap \vee p$.

1 Intuitionistic Probability

In Weatherson (2003), the notion of a $\vdash$-probability function, where $\vdash$ is an entailment relation, is introduced. For any $\vdash$, a $\vdash$-probability function is a function Pr from sentences in the language of $\vdash$ to [0, 1] satisfying the following four constraints.[3]

(P0) $\Pr(p) = 0\$$ if p is a $\vdash$-antithesis, i.e. iff for any X, $p \vdash X$.
(P1) $\Pr(p) = 1$ if p is a $\vdash$-thesis, i.e. iff for any X, $X \vdash p$.
(P2) If $p \vdash q$ then $\Pr(p) \leqslant \Pr(q)$.
(P3) $\Pr(p) + \Pr(q) = \Pr(p \vee q) + \Pr(p \wedge q)$.

---

[3]We'll usually assume that the language of $\vdash$ is a familiar kind of propositional calculus, with a countable infinity of sentence letters, and satisfying the usual recursive constraints. That is, if A and B are sentences of the language, then so are $\neg A$, $A \rightarrow B$, $A \wedge B$ and $A \vee B$. It isn't entirely trivial to extend some of our results to a language that contains quantifiers. This is because once we add quantifiers, intuitionistic and classical logic no longer have the same anti-theorems. But that complication is outside the scope of this paper. Note that for Theorem 6, we assume a restricted language with just two sentence letters. This merely simplifies the proof. A version of the construction we use there with those two letters being simply the first two sentence letters would be similar, but somewhat more complicated.

We'll use $\vdash_{CL}$ to denote the classical entailment relation, and $\vdash_{IL}$ to denote the intu-itionist entailment relation. Then what we usually take to be probability functions are $\vdash_{CL}$-probability functions. And intuitionist probability functions are $\vdash_{IL}$-probability functions.

In what follows we'll make frequent appeal to three obvious consequences of these axioms, consequences which are useful enough to deserve their own names. Hopefully these are obvious enough to pass without proof.[4]

Weatherson (2003) discusses what happens if we make P2* or P3* an axiom in place of either P2 and P3. It is argued there that this gives us too many functions to be useful in epistemology. The arguments in Williams (2012) provide much stronger reasons for believing this conclusion is correct.

(P1*) $0 \leqslant \Pr(p) \leqslant 1$.
(P2*) If p $\dashv\vdash$ q then $\Pr(p) = \Pr(q)$.
(P3*) If $p \wedge q$ is a $\vdash$-antithesis, then $\Pr(p) + \Pr(q) = \Pr(p \vee q)$.

$\vdash$-probability functions obviously concern unconditional probability, but we can easily extend them into conditional $\vdash$-probability functions by adding the following axioms.[5]

(P4) If r is not a $\vdash$-antithesis, then $\Pr(\cdot \mid r)$ is a $\vdash$-probability function; i.e., it satisfies P0-P3.
(P5) If $r \vdash p$ then $\Pr(p \mid r) = 1$.
(P6) If r is not a $\vdash$-antithesis, then $\Pr(p \wedge q \mid r) = \Pr(p \mid q \wedge r)\Pr(q \mid r)$.

There is a simple way to generate $\vdash_{CL}$ probability functions. Let $\langle W, V \rangle$ be a model where W is a finite set of worlds, and V a valuation function defined on them with respect to a (finite) set K of atomic sentences, i.e., a function from K to subsets of W. Let L be the smallest set including all members of K such that whenever A and B are in L, so are $A \wedge B$, $A \vee B$, $A \rightarrow B$ and $\neg A$. Extend V to V^*, a function from L to subsets of W using the usual recursive definitions of the sentential connectives. (So $w \in V^*(A \wedge B)$ iff $w \in V^*(A)$ and $w \in V^*(B)$, and so on for the other connectives.) Let m be a measure function defined over subsets of W. Then for any sentence S in L, $\Pr(S)$ is $m(\{w: w \in V^*(S)\})$. It isn't too hard to show that Pr is a $\vdash_{CL}$ probability function.

There is a similar way to generate $\vdash_{IL}$ probability functions. This method uses a sim-plified version of the semantics for intuitionistic logic in Kripke (1965). Let $\langle W, R, V \rangle$ be a model where W is a finite set of worlds, R is a reflexive, transitive relation defined on W, and V is a valuation function defined on them with respect to a (finite) set K of atomic sentences. We require that V be closed with respect to R, i.e. that if $x \in V(p)$ and xRy, then $y \in V(p)$. We define L the same way as above, and extend V to V^* (a function from L to subsets of W) using the following definitions.

---

[4]In the original, the next three paragraphs were footnoted, but I no longer like having numbered things in footnotes.

[5]For the reasons given in Hájek (2003), it is probably better in general to take conditional probability as primitive. But for our purposes taking unconditional probability to be basic won't lead to any problems, so we'll stay neutral on whether conditional or unconditional probability is really primitive.

$w \in V^*(A \wedge B)$ iff $w \in V^*(A)$ and $w \in V^*(B)$.
$w \in V^*(A \vee B)$ iff $w \in V^*(A)$ or $w \in V^*(B)$.
$w \in V^*(A \rightarrow B)$ iff for all w' such that wRw' and $w' \in V^*(A)$, $w' \in V^*(B)$.
$w \in V^*(\neg A)$ iff for all w' such that wRw', it is not the case that $w' \in V^*(A)$.

Finally, we let m be a measure function defined over subsets of W. And for any sentence S in L, $\Pr(S)$ is $m(\{w: w \in V^*(S)\})$. Weatherson (2003) shows that any such $\Pr$ is a $\vdash_{IL}$ probability function.

To show that Theorem 1 may fail when $\Pr$ is $\vdash_{IL}$ a probability function, we need a model we'll call M. The valuation function in M is defined with respect to a language where the only atomic propositions are p and Ap.
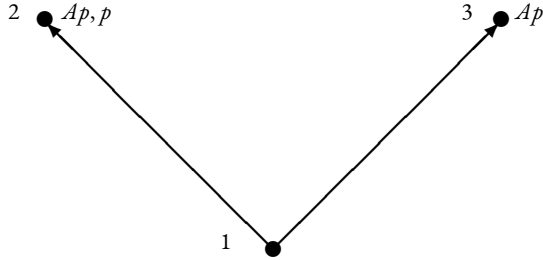
$$W = \{1, 2, 3\}$$
$$R = \{\langle 1, 1 \rangle, \langle 2, 2 \rangle, \langle 3, 3 \rangle, \langle 1, 2 \rangle, \langle 1, 3 \rangle\}$$
$$V(p) = \{2\}$$
$$V(Ap) = \{2, 3\}$$

Graphically, M looks like this.

We'll now consider a family of measures over m. For any $x \in (0, 1)$, let m_x be the measure function such that $m_x(\{1\}) = 1 - x$, $m_x(\{2\}) = x$, and $m_x(\{3\}) = 0$. Corresponding to each function m_x is a $\vdash_{IL}$ probability function we'll call $\Pr_x$. Inspection of the model shows that Theorem 3 is true.

Theorem 3.
In M, for any $x \in (0, 1)$,

1. $\Pr_x(Ap \rightarrow p) = \Pr_x((Ap \rightarrow p) \wedge Ap) = x$
2. $\Pr_x(\neg Ap \vee p) = \Pr_x((\neg Ap \vee p) \wedge Ap) = x$
3. $\Pr_x(\neg(Ap \wedge \neg p)) = \Pr_x(\neg(Ap \wedge \neg p) \wedge Ap) = x$

An obvious corollary of Theorem 3 is

Theorem 4.

For any $x \in (0, 1)$,

1. $1 = \Pr_x(Ap \to p \mid Ap) > \Pr_x(Ap \to p) = x$
2. $1 = \Pr_x(\neg Ap \lor p \mid Ap) > \Pr_x(\neg Ap \lor p) = x$
3. $1 = \Pr_x(\neg(Ap \land \neg p) \mid Ap) > \Pr_x(\neg(Ap \land \neg p)) = x$

So for any x, conditionalising on Ap actually raises the probability of $Ap \to p$, $\neg(Ap \land \neg p)$ and $\neg Ap \lor p$ with respect to $\Pr_x$. Indeed, since x could be arbitrarily low, it can raise the probability of each of these three propositions from any arbitrarily low value to 1. So it seems that if we think learning goes by conditionalisation, then receiving evidence Ap could be sufficient grounds to justify belief in these three propositions. Of course, this relies on our being prepared to use the intuitionist probability calculus. For many, this will be considered too steep a price to pay to preserve dogmatism. But in section 2 we'll show that the dogmatist does not need to insist that intuitionistic logic is the correct logic for modelling uncertainty. All they need to show is that it *might* be correct, and then they'll have a response to this argument.

2 Logical Uncertainty

We're going to build up to a picture of how to model agents who are rationally uncertain about whether the correct logic is classical or intuitionistic. But let's start by thinking how an agent who is unsure which of two empirical theories T_1 or T_2 is correct. We'll assume that the agent is using the classical probability calculus, and the agent knows which propositions are entailed by each of the two theories. And we'll also assume that the agent is sure that it's not the case that each of these theories is false, and the theories are inconsistent, so they can't both be true.

The natural thing then is for the agent to have some credence x in T_1, and credence $1-x$ in T_2. She will naturally have a picture of what the world is like assuming T_1 is correct, and on that picture every proposition entailed by T_1 will get probability 1. And she'll have a picture of what the world is like assuming T_2 is correct. Her overall credal state will be a mixture of those two pictures, weighted according to the credibility of T_1 and T_2.

If we're working with unconditional credences as primitive, then it is easy to mix two probability functions to produce a credal function which is also a probability function. Let $\Pr_1$ be the probability function that reflects the agent's views about how things probably are conditional on T_1 being true, and $\Pr_2$ the probability function that reflects her views about how things probably are conditional on T_2 being true. Then for any p, let $Cr(p) = x\Pr_1(p) + (1-x)\Pr_2(p)$, where Cr is the agent's credence function.

It is easy to see that Cr will be a probability function. Indeed, inspecting the axioms P0-P3 makes it obvious that for any $\vdash$, mixing two $\vdash$-probability functions as we've just done will always produce a $\vdash$-probability function. The axioms just require that

probabilities stand in certain equalities and inequalities that are obviously preserved under mixing.

It is a little trickier to mix conditional probability functions in an intuitive way, for the reasons set out in Jehle and Fitelson (2009). But in a special case, these difficulties are not overly pressing. Say that a ⊢-probability function is **regular** iff for any p, q in its domain, $\Pr(p \mid q) = 0$ iff $p \wedge q$ is a ⊢-antitheorem. Then, for any two regular conditional probability functions $\Pr_1$ and $\Pr_2$ we can create a weighted mixture of the two of them by taking the new unconditional probabilities, i.e. the probabilities of p given T, where T is a theorem, to be weighted sums of the unconditional probabilities in $\Pr_1$ and $\Pr_2$. That is, our new function $\Pr_3$ is given by:

$$\Pr_3(p$$

In the general case, this does not determine exactly which function $\Pr_3$ is, since it doesn't determine the value of $\Pr_3(p \mid q)$ when $\Pr_1(q \mid T) = \Pr_2(q \mid T) = 0$. But since we're paying attention just to regular functions this doesn't matter. If the function is regular, then we can just let the familiar ratio account of conditional probability be a genuine definition. So in general we have,

$$\Pr_3(p$$

And since the numerator is 0 iff q is an anti-theorem, whenever $\Pr(p \mid q)$ is supposed to be defined, i.e. when q is not an anti-theorem, the right hand side will be well defined. As we noted, things get a lot messier when the functions are not regular, but those complications are unnecessary for the story we want to tell.

Now in the cases we've been considering so far, we've been assuming that T_1 and T_2 are empirical theories, and that we could assume classical logic in the background. Given all that, most of what we've said in this section has been a fairly orthodox treatment of how to account for a kind of uncertainty. But there's no reason, we say, why we should restrict T_1 and T_2 in this way. We could apply just the same techniques when T_1 and T_2 are theories of entailment.

When T_1 is the theory that classical logic is the right logic of entailment, and T_2 the theory that intuitionistic logic is the right logic of entailment, then $\Pr_1$ and $\Pr_2$ should be different kinds of probability functions. In particular, $\Pr_1$ should be a $\vdash_{CL}$-probability function, and $\Pr_2$ should be a $\vdash_{IL}$-probability function. That's because $\Pr_1$ represents how things probably are given T_1, and given T_1, how things probably are is constrained by classical logic. And $\Pr_2$ represents how things probably are given T_2, and given T_2, how things probably are is constrained by intuitionistic logic.

If we do all that, we're pushed towards the thought that the if someone is uncertain whether the right logic is intuitionistic or classical logic, then the right theory of probability for them is intuitionistic probability theory. That's because of Theorem 5.

Theorem 5
Let $\Pr_1$ be a regular conditional $\vdash_{CL}$-probability function, and $\Pr_2$ be a

regular conditional $\vdash_{IL}$-probability function that is not a $\vdash_{CL}$-probability function. And let Pr_3 be defined as in the text. (That is, $\mathrm{Pr}_3(A) = x\mathrm{Pr}_1(A) + (1\text{-}x)\mathrm{Pr}_2(A)$, and $\mathrm{Pr}_3(A \mid B) = (\mathrm{Pr}_3(A \wedge B))/(\mathrm{Pr}_3(B))$.) Then Pr_3 is a regular conditional $\vdash_{IL}$-probability function.

That's to say, if the agent is at all unsure whether classical logic or intuitionistic logic is the correct logic, then their credence function should be an intuitionistic probability function.

Of course, if the agent is very confident that classical logic is the correct logic, then they couldn't rationally have their credences distributed by any old intuitionistic probability function. After all, there are intuitionistic probability functions such that $\mathrm{Pr}(p \vee \neg p) = 0$, but an agent whose credence that classical logic is correct is, say, 0.95, could not reasonably have credence 0 in $p \vee \neg p$. For our purposes, this matters because we want to show that an agent who is confident, but not certain, that classical logic is correct can nevertheless be a dogmatist. To fill in the argument we need,

Theorem 6
Let x be any real in $(0, 1)$. Then there is a probability function Cr that (a) is a coherent credence function for someone whose credence that classical logic is correct is x, and (b) satisfies each of the following inequalities:

$$\mathrm{Pr}(Ap \rightarrow p \mid Ap) > \mathrm{Pr}(Ap \rightarrow p)$$
$$\mathrm{Pr}(\neg Ap \vee p \mid Ap) > \mathrm{Pr}(\neg Ap \vee p)$$
$$\mathrm{Pr}(\neg(Ap \wedge \neg p) \mid Ap) > \mathrm{Pr}(\neg(Ap \wedge \neg p))$$

The main idea driving the proof of Theorem 6 which is set out in the appendix, is that if intuitionistic logic is correct, it's possible that conditionalising on Ap raises the probability of each of these three propositions from arbitrarily low values to 1. So as long as the prior probability of each of the three propositions, conditional on intuitionistic logic being correct, is low enough, it can still be raised by conditionalising on Ap.

More centrally, we think Theorem 6 shows that the probabilistic argument against dogmatism is not compelling. The original argument noted that the dogmatist says that we can learn the three propositions in Theorem 6, most importantly $Ap \rightarrow p$, by getting evidence Ap. And it says this is implausible because conditionalising on Ap lowers the probability of $Ap \rightarrow p$. But it turns out this is something of an artifact of the very strong classical assumptions that are being made. The argument not only requires the correctness of classical logic, it requires that the appropriate credence the agent should have in classical logic's being correct is one. And that assumption is, we think, wildly implausible. Even if the agent should be *very confident* that classical logic is the correct logic, it shouldn't be a requirement of rationality that she be absolutely certain that it is correct.

So we conclude that this argument fails. A dogmatist about perception who is at least minimally open-minded about logic can marry perceptual dogmatism to a probabilistically coherent theory of confirmation.

This paper is one more attempt on our behalf to defend dogmatism from a probabilistic challenge. Weatherson (2007) defends dogmatism from the so-called "Bayesian objection". And Jehle (2009) not only shows that dogmatism can be situated nicely into a probabilistically coherent theory of confirmation, but also that within such a theory, many of the traditional objections to dogmatism are easily rebutted. We look forward to future research on the connections between dogmatism and probability, but we remain skeptical that dogmatism will be undermined solely by probabilistic considerations.

Appendix: Proofs

Theorem 1
If Pr$ is a classical probability function, then
$\Pr(Ap \to p \mid Ap) \leqslant \Pr(Ap \to p)$.

Proof: Assume Pr is a classical probability function, and $\vdash$ the classical consequence relation.

1. $Ap \to p \dashv \vdash ((Ap \to p) \wedge Ap) \vee ((Ap \to p) \wedge \neg Ap)$
2. $\Pr(Ap \to p) = \Pr(((Ap \to p) \wedge Ap) \vee ((Ap \to p) \wedge \neg Ap))$ (from 1, P2^*^)
3. $\Pr((Ap \to p) \wedge Ap) \vee ((Ap \to p) \wedge \neg Ap)) = \Pr((Ap \to p) \wedge Ap) + \Pr((Ap \to p) \wedge \neg Ap)$ (from P3*)
4. $\Pr((Ap \to p) \wedge Ap) = \Pr(Ap)\Pr(Ap \to p \mid Ap)$ (from P6)
5. $\Pr((Ap \to p) \wedge \neg Ap) = \Pr(\neg Ap)\Pr(Ap \to p \mid \neg Ap)$ (from P6)
6. $\Pr(Ap \to p) = \Pr(Ap)\Pr(Ap \to p \mid Ap) + \Pr(\neg Ap)\Pr(Ap \to p \mid \neg Ap)$ (from 2, 4, 5)
7. $(Ap \to p) \wedge Ap \dashv \vdash \neg Ap$
8. $\Pr((Ap \to p) \wedge Ap) = \Pr(\neg Ap)$ (from 7, P2^*^)
9. $\Pr(Ap \to p \mid \neg Ap) = 1$ or $\Pr(\neg Ap) = 0$ (from 8, P6)
10. $\Pr(Ap \to p \mid Ap) \leqslant 1$ (from P4, P5)
11. $\Pr(Ap \to p) \geqslant \Pr(Ap)\Pr(Ap \to p \mid Ap) + \Pr(\neg Ap)\Pr(Ap \to p \mid Ap)$ (from 6, 9, 10)
12. $\vdash Ap \vee \neg Ap$
13. $\Pr(Ap \vee \neg Ap) = 1$ (from 12, P1)
14. $\Pr(Ap) + \Pr(\neg Ap) = 1$ (from 13, P3*)
15. $\Pr(Ap \to p) \geqslant \Pr(Ap \to p \mid Ap)$ (from 11, 14)

Note (11) is an equality iff (8) is. The only step there that may not be obvious is step 10. The reason it holds is that either Ap is a $\vdash$-antitheorem or it isn't. If it is, then it entails $Ap \to p$, so by P5, $\Pr(Ap \to p \mid Ap) \leqslant 1$. If it is not, then by P1*, $\Pr(x \mid Ap) \leqslant 1$ for any x, so $\Pr(Ap \to p \mid Ap) \leqslant 1$.

Theorem 2
If Pr is a classical probability function, then

- $\Pr(\neg(Ap \wedge \neg p) \mid Ap) \leqslant \Pr(\neg(Ap \wedge \neg p))$; and

- $\Pr(\neg Ap \vee p \mid Ap) \leqslant \Pr(\neg Ap \vee p)$.

Proof: Assume Pr is a classical probability function, and $\vdash$ the classical consequence relation.

1. $Ap \to p \dashv\vdash \neg(Ap \wedge \neg p)$
2. $\Pr(Ap \to p) = \Pr(\neg(Ap \wedge \neg p))$ (1, P2^*^)
3. $\Pr(Ap \to p \mid Ap) = \Pr(\neg(Ap \wedge \neg p) \mid Ap)$ (1, P4, P5)
4. $\Pr(Ap \to p) \geqslant \Pr(Ap \to p \mid Ap)$ (Theorem 1)
5. $\Pr(\neg(Ap \wedge \neg p) \mid Ap) \geqslant \Pr(\neg(Ap \wedge \neg p))$ (2, 3, 4)
6. $Ap \to p \dashv\vdash \neg Ap \vee p$
7. $\Pr(Ap \to p) = \Pr(\neg Ap \vee p)$ (6, P2^*^)
8. $\Pr(Ap \to p \mid Ap) = \Pr(\neg Ap \vee p \mid Ap)$ (6, P4, P5)
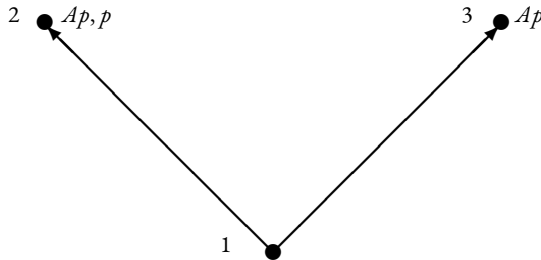9. $\Pr(\neg Ap \vee p \mid Ap) \geqslant \Pr(\neg Ap \vee p)$ (4, 7, 8)

The only minor complication is with step 3. There are two cases to consider, either Ap is a $\vdash$-antitheorem or it isn't. If it is a $\vdash$-antitheorem, then both the LHS and RHS of (3) equal 1, so they are equal. If it is not a $\vdash$-antitheorem, then by P4, $\Pr(\cdot \mid Ap)$ is a probability function. So by P2^^, and the fact that $Ap \to p \dashv\vdash \neg(Ap \,\square\, \neg p^*)$, we have that the LHS and RHS are equal.

Theorem 3.
In M, for any $x \in (0, 1)$,

1. $\Pr_x(Ap \to p) = \Pr_x((Ap \to p) \wedge Ap) = x$
2. $\Pr_x(\neg Ap \vee p) = \Pr_x((\neg Ap \vee p) \wedge Ap) = x$
3. $\Pr_x(\neg(Ap \wedge \neg p)) = \Pr_x(\neg(Ap \wedge \neg p) \wedge Ap) = x$

Recall what M looks like.

The only point where $Ap \to p$ is true is at 2. Indeed, $\neg(Ap \to p)$ is true at 3, and neither $Ap \to p$ nor $\neg(Ap \to p)$ are true at 1. So $\Pr_x(Ap \to p) = m_x(\{2\}) = x$. Since Ap is also true at 2, that's the only point where $(Ap \to p) \wedge Ap$ is true. So it follows that $\Pr_x((Ap \to p) \wedge Ap) = m_x(\{2\}) = x$.

Similar inspection of the model shows that 2 is the only point where $\neg(Ap \wedge \neg p)$ is true, and the only point where $\neg Ap \vee p$ is true. And so (*b*) and (*c*) follow in just the same way.

In slight contrast, Ap is true at two points in the model, 2 and 3. But since $m_x(\{3\}) = 0$, it follows that $m_x(\{2, 3\}) = m_x(\{2\}) = x$. So $\Pr_x(Ap) = x$.

Theorem 4.

For any $x \in (0, 1)$,

    a. $1 = \Pr_x(Ap \to p \mid Ap) > \Pr_x(Ap \to p) = x$
    b. $1 = \Pr_x(\neg Ap \vee p \mid Ap) > \Pr_x(\neg Ap \vee p) = x$
    c. $1 = \Pr_x(\neg(Ap \wedge \neg p) \mid Ap) > \Pr_x(\neg(Ap \wedge \neg p)) = x$

We'll just go through the argument for (*a*); the other cases are similar. By P6, we know that $\Pr_x(\neg(Ap \wedge \neg p) \mid Ap) \Pr_x(Ap) = \Pr_x((Ap \to p) \wedge Ap)$. By Theorem 3, we know that $\Pr_x(Ap) = \Pr_x((Ap \to p) \wedge Ap)$, and that both sides are greater than 0. (Note that the theorem is only said to hold for $x > 0$.) The only way both these equations can hold is if $\Pr_x(\neg(Ap \wedge \neg p) \mid Ap) = 1$. Note also that by hypothesis, $x < 1$, and from this claim (*a*) follows. The other two cases are completely similar.

Theorem 5

Let $\Pr_1$ be a regular conditional $\vdash_{CL}$-probability function, and $\Pr_2$ be a regular conditional $\vdash_{IL}$-probability function that is not a $\vdash_{CL}$-probability function. And let $\Pr_3$ be defined as in the text. (That is, $\Pr_3(A) = x\Pr_1(A) + (1\text{-}x)\Pr_2(A)$, and $\Pr_3(A \mid B) = (\Pr_3(A \wedge B))/\{\Pr_3(B)\}$.) Then $\Pr_3$ is a regular conditional $\vdash_{IL}$-probability function.

We first prove that $\Pr_3$ satisfies the requirements of an unconditional $\vdash_{IL}$-probability function, and then show that it satisfies the requirements of a conditional $\vdash_{IL}$-probability function.

If p is an $\vdash_{IL}$-antithesis, then it is also a $\vdash_{CL}$-antithesis. So $\Pr_1(p) = \Pr_2(p) = 0\$$. So $\Pr_3(A) = 0x + 0(1\text{-}x) = 0$, as required for **(P0)**.

If p is an $\vdash_{IL}$-thesis, then it is also a $\vdash_{CL}$-thesis. So $\Pr_1(p) = \Pr_2(p) = 1$. So $\Pr_3(p) = x + (1\text{-}x) = 1$, as required for **(P1)**.

If $p \vdash_{IL} q$ then $p \vdash_{CL} q$. So we have both $\Pr_1(p) \leqslant \Pr(q)$ and $\Pr_2(p) \leqslant \Pr_2(q)$. Since $x \geqslant 0$ and $(1\text{-}x) \geqslant 0$, these inequalities imply that $x\Pr_1(p) \leqslant x\Pr(q)$ and $(1\text{-}x)\Pr_2(p) \leqslant (1\text{-}x)\Pr_2(q)$. Summing these, we get $x\Pr_1(p) + (1\text{-}x)\Pr_2(p) \leqslant x\Pr_1(q) + (1\text{-}x)\Pr_2(q)$. And by the definition of $\Pr_3$, that means that $\Pr_3(p) \leqslant \Pr_3(q)$, as required for **(P2)**.

Finally, we just need to show that $\Pr_3(p) + \Pr_3(q) = \Pr_3(p \vee q) + \Pr_3(p \wedge q)$, as follows:

$$
\begin{aligned}
\Pr_3(p) + \Pr_3(q) \quad &= x\Pr_1(p) + (1\text{-}x)\Pr_2(p) + x\Pr_1(q) + \\
&\quad (1\text{-}x)\Pr_2(q) \\
&= x(\Pr_1(p) + \Pr_1(q)) + (1\text{-}x)(\Pr_2(p) + \\
&\quad \Pr_2(q))
\end{aligned}
$$

$$= x(\Pr_1(p \vee q) + \Pr_1(p \wedge q)) +$$
$$(1\text{-}x)(\Pr_2(p \vee q) + \Pr_2(p \wedge q))$$
$$= x\Pr_1(p \vee q) + (1\text{-}x)\Pr_2(p \vee q) +$$
$$x\Pr_1(p \wedge q) + (1\text{-}x)\Pr_2(p \wedge q)$$
$$= \Pr_3(p \vee q) + \Pr_3(p \wedge q), \text{ as required}$$

Now that we have shown $\Pr_3$ is an unconditional $\vdash_{IL}$-probability function, we need to show it is a conditional $\vdash_{IL}$-probability function, where $\Pr_3(p \mid r) =_{\mathrm{df}}$ $(\Pr_3(p \wedge r))/(\Pr_3(r))$. Remember we are assuming that both $\Pr_1$ and $\Pr_2$ are regular, from which it clearly follows that $\Pr_3$ is regular, so this definition is always in order. (That is, we're never dividing by zero.) The longest part of showing $\Pr_3$ is a conditional $\vdash_{IL}$-probability function is showing that it satisfies (P4), which has four parts. We need to show that $\Pr(\cdot \mid r)$ satisfies (P0)-(P3). Fortunately these are fairly straightforward.

If p is an $\vdash_{IL}$-antithesis, then so is $p \wedge r$. So $\Pr_3(p \wedge r) = 0$, so $\Pr_3(p \mid r) = 0$, as required for (P0).

If p is an $\vdash_{IL}$-thesis, then $p \wedge r \dashv \vdash r$, so $\Pr_3(p \wedge r) = \Pr_3(r)$, so $\Pr_3(p \mid r) = 1$, as required for (P1).

If $p \vdash_{IL} q$ then $p \wedge r \vdash_{IL} q \wedge r$. So $\Pr_3(p \wedge r) \leqslant \Pr_3(q \wedge r)$. So $(\Pr_3(p \wedge r))/(\Pr_3(r) \leqslant (\Pr_3(q \wedge r))/(\Pr_3(r))$. That is, $\Pr_3(p \mid r) \leqslant \Pr_3(q \mid r)$, as required for (P2).

Finally, we need to show that $\Pr_3(p \mid r) + \Pr_3(q \mid r) = \Pr_3(p \vee q \mid r) + \Pr_3(p \wedge q \mid r)$, as follows, making repeated use of the fact that $\Pr_3$ is an unconditional $\vdash_{IL}$-probability function, so we can assume it satisfies (P3), and that we can substitute intuitionistic equivalences inside $\Pr_3$.

$$\Pr_3(p|r) + \Pr_3(q|r) = \frac{\Pr_3(p \wedge r)}{\Pr_3(r)} + \frac{\Pr_3(q \wedge r)}{\Pr_3(r)}$$
$$= \frac{\Pr_3(p \wedge r) + Pr(q \wedge r)}{\Pr_3(r)}$$
$$= \frac{\Pr_3((p \wedge r) \vee (q \wedge r)) + \Pr_3((p \wedge r) \wedge (q \wedge r))}{\Pr_3(r)}$$
$$= \frac{\Pr_3(p \vee q) \wedge r) + \Pr_3((p \wedge q) \wedge r)}{\Pr_3(r)}$$
$$= \frac{\Pr_3(p \vee q) \wedge r)}{\Pr_3(r)} + \frac{\Pr_3((p \wedge q) \wedge r)}{\Pr_3(r)}$$
$$= \Pr_3(p \vee q|r) + \Pr_3(p \wedge q|r) \text{ as required}$$

Now if $r \vdash_{IL} p$, then $r \wedge p \dashv\vdash_{IL} p$, so $\Pr_3(r \wedge p) = \Pr_3(p)$, so $\Pr_3(p \mid r) = 1$, as required for (P5).

Finally, we show that $\Pr_3$ satisfies (P6).

$$\Pr{}_3(p \wedge q | r) = \frac{\Pr_3(p \wedge q \wedge r)}{\Pr_3(r)}$$

$$= \frac{\Pr_3(p \wedge q \wedge r)}{\Pr_3(q \wedge r)} \frac{\Pr_3(q \wedge r)}{\Pr_3(r)}$$

$$= \Pr{}_3(p | q \wedge r) Pr\, 3\, (q | r) \text{ as required}$$

Theorem 6 Let x be any real in $(0, 1)$. Then there is a probability function Cr that (a) is a coherent credence function for someone whose credence that classical logic is correct is x, and (b) satisfies each of the following inequalities:

$\Pr(Ap \to p \,|\, Ap) > \Pr(Ap \to p)$
$\Pr(\neg Ap \vee p \,|\, Ap) > \Pr(\neg Ap \vee p)$
$\Pr(\neg(Ap \wedge \neg p) \,|\, Ap) > \Pr(\neg(Ap \wedge \neg p))$

We'll prove this by constructing the function Pr. For the sake of this proof, we'll assume a very restricted formal language with just two atomic sentences: Ap and p. This restriction makes it easier to ensure that the functions are all regular, which as we noted in the main text lets us avoid various complications. The proofs will rely on three probability functions defined using this Kripke tree M.
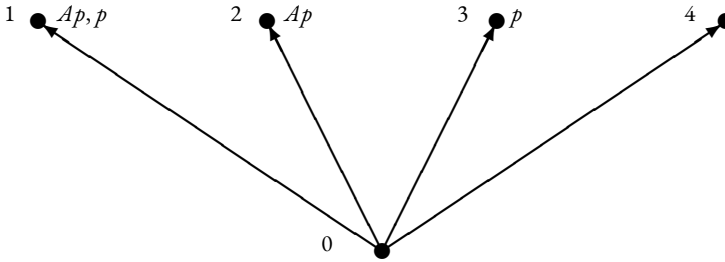
We've shown on the graph where the atomic sentences true: Ap is true at 1 and 2, and p is true at 1 and 3. So the four terminal nodes represent the four classical possibilities that are definable using just these two atomic sentences. We define two measure functions m_1 and m_2 over the points in this model as follows:

	$m(0)$	$m(1)$	$m(2)$	$m(3)$	$m(4)$
m_1	0	$x/2$	$(1\text{-}x)/2$	¼	¼
m_2	$x/2$	$(1\text{-}x)/4$	$(1\text{-}x)/4$	¼	¼

We've just specified the measure of each singleton, but since we're just dealing with a finite model, that uniquely specifies the measure of any set. We then turn each of these into probability functions in the way described in section 1. That is, for any proposition X, and $i \in \{1, 2\}$, $\Pr_i(X) = m_i(M_X)$, where M_X is the set of points in M where X is true.

Note that the terminal nodes in M, like the terminal nodes in any Kripke tree, are just classical possibilities. That is, for any sentence, either it or its negation is true at a terminal node. Moreover, any measure over classical possibilities generates a classical probability function. (And vice versa, any classical probability function is generated by a measure over classical possibilities.) That is, for any measure over classical possibilities, the function from propositions to the measure of the set of possibilities at which they are true is a classical probability function. Now m_1 isn't quite a measure over classical possibilities, since strictly speaking $m_1(\{0\})$ is defined. But since $m_1(\{0\}) = 0$ it is equivalent to a measure only defined over the terminal nodes. So the probability function it generates, i.e., Pr_1, is a classical probability function. Of course, with only two atomic sentences, we can also verify by brute force that Pr_1 is classical, but it's a little more helpful to see why this is so. In contrast, Pr_2 is not a classical probability function, since $Pr_2(p \vee \neg p) = 1 - x/2$, but it is an intuitionistic probability function.

So there could be an agent who satisfies the following four conditions:

- Her credence that classical logic is correct is x;
- Her credence that intuitionistic logic is correct is $1-x$;
- Conditional on classical logic being correct, she thinks that Pr_1 is the right representation of how things probably are; and
- Conditional on intuitionistic logic being correct, she thinks that Pr_2 is the right representation of how things are.

Such an agent's credences will be given by a $\vdash_{IL}$-probability function Pr generated by 'mixing' Pr_1 and Pr_2. For any sentence Y in the domain, her credence in Y will be $xPr_1(Y) + (1-x)Pr_2(Y)$. Rather than working through each proposition, it's easiest to represent this function by mixing the measures m_1 and m_1 to get a new measure m on the above Kripke tree. Here's the measure that m assigns to each node.

	$m(0)$	$m(1)$	$m(2)$	$m(3)$	$m(4)$
m	$x(1-x)/2$	$(3x^2 - 2x + 1)/4$	$(1-x^2)/4$	¼	¼

As usual, this measure m generates a probability function Pr. We've already argued that Pr is a reasonable function for someone whose credence that classical logic is x. We'll now argue that $Pr(Ap \rightarrow p \mid Ap) > Pr(Ap \rightarrow p)$.

It's easy to see what $Pr(Ap \rightarrow p)$ is. $Ap \rightarrow p$ is true at 1, 3 and 4, so

$Pr(Ap \rightarrow p) = m(\{1\}) + m(\{3\}) + m(4)$
$= (3x^2 - 2x + 1)/4 + ¼ + ¼$
$= (3x^2 - 2x + 3)/4$

Since Pr is regular, we can use the ratio definition of conditional probability to work out $Pr(Ap \rightarrow p \mid Ap)$.

$Pr(Ap \rightarrow p \mid Ap) = (Pr((Ap \rightarrow p) \wedge Ap))/(Pr(Ap))$
$= m(1)/(m(1) + m(2)) \backslash$
$= ((3x^2 - 2x + 1)/4) / ((3x^2 - 2x + 1)/4 + (1-x^2)/4)$

$$= (3x^2 - 2x + 1) / ((3x^2 - 2x + 1) + (1-x^2))$$
$$= (3x^2 - 2x + 1) / 2(x^2 - x + 1)$$

Putting all that together, we have

$$Pr(Ap \rightarrow p | Ap) > Pr(Ap \rightarrow p)$$

$$\Leftrightarrow \qquad \frac{3x^2 - 2x + 3}{4} > \frac{3x^2 - 2x + 1}{2(x^2 - x + 1)}$$

$$\Leftrightarrow \qquad 3x^2 - 2x + 3 > \frac{6x^2 - 4x + 2}{x^2 - x + 1}$$

$$\Leftrightarrow \quad (3x^2 - 2x + 3)(x^2 + x + 1) > 6x^2 - 4x + 2$$

$$\Leftrightarrow \quad 3x^4 - 5x^3 + 8x^2 - 5x + 3 > 6x^2 - 4x + 2$$

$$\Leftrightarrow \qquad 3x^4 - 5x^3 + 2x^2 - x + 1 > 0$$

$$\Leftrightarrow \quad (3x^2 + x + 1)(x^2 - 2x + 1) > 0$$

$$\Leftrightarrow \qquad (3x^2 + x + 1)(x - 1)^2 > 0$$

But it is clear that for any $x \in (0,1)$, both of the terms of the LHS of the final line are positive, so their product is positive. And that means $Pr(Ap \rightarrow p \mid Ap) > Pr(Ap \rightarrow p)$. So no matter how close x gets to 1, that is, no matter how certain the agent gets that classical logic is correct, as long as x does not reach 1, conditionalising on Ap will raise the probability of $Ap \rightarrow p$. As we've been arguing, as long as there is any doubt about classical logic, even a vanishingly small doubt, there is no probabilistic objection to dogmatism.

To finish up, we show that $Pr(\neg Ap \vee p | Ap) > Pr(\neg Ap \vee p)$ and $Pr(\neg(Ap \wedge \neg p) | Ap) > Pr(\neg(Ap \wedge \neg p))$. To do this, we just need to note that $Ap \rightarrow p$, $\neg Ap \vee p$ and $\neg(Ap \wedge \neg p)$ are true at the same points in the model, so their probabilities, both unconditionally and conditional on Ap, will be identical. So from $Pr(Ap \rightarrow p \mid Ap) > Pr(Ap \rightarrow p)$ the other two inequalities follow immediately.

References

Cohen, Stewart. 2005. "Why Basic Knowledge Is Easy Knowledge." *Philosophy and Phenomenological Research* 70 (2): 417–30. doi: 10.1111/j.1933-1592.2005.tb00536.x.

Hájek, Alan. 2003. "What Conditional Probability Could Not Be." *Synthese* 137 (3): 273–323. doi: 10.1023/B:SYNT.0000004904.91112.16.

Jehle, David. 2009. "Some Results in Bayesian Confirmation Theory with Applications." PhD thesis, Cornell University.

Jehle, David, and Branden Fitelson. 2009. "What Is the 'Equal Weight View'?" *Episteme* 6 (3): 280–93. doi: 10.3366/E1742360009000719.

Kripke, Saul. 1965. "Semantical Analysis of Intuitionistic Logic." In *Formal Systems and Recursive Functions*, edited by Michael Dummett and John Crossley. Amsterdam: North-Holland.

Popper, Karl, and David Miller. 1987. "Why Probabilistic Support Is Not Inductive."

Philosophical Transactions of the Royal Society of London. Series A, Mathematical and Physical Sciences 321 (1562): 569–91. doi: 10.1098/rsta.1987.0033.

Pryor, James. 2000. "The Sceptic and the Dogmatist." *Noûs* 34 (4): 517–49. doi: 10.1111/0029-4624.00277.

———. 2004. "What's Wrong with Moore's Argument?" *Philosophical Issues* 14 (1): 349–78. doi: 10.1111/j.1533-6077.2004.00034.x.

Weatherson, Brian. 2003. "From Classical to Intuitionistic Probability." *Notre Dame Journal of Formal Logic* 44 (2): 111–23. doi: 10.1305/ndjfl/1082637807.

———. 2007. "The Bayesian and the Dogmatist." *Proceedings of the Aristotelian Society* 107: 169–85. doi: 10.1111/j.1467-9264.2007.00217.x.

White, Roger. 2006. "Problems for Dogmatism." *Philosophical Studies* 131 (3): 525–57. doi: 10.1007/s11098-004-7487-9.

Williams, J. R. G. 2012. "Gradational Accuracy and Non-Classical Semantics." *Review of Symbolic Logic* 5 (4): 513–37. doi: 10.1017/S1755020312000214.