Arguments For—Or Against—Probabilism?

Alan Hájek

1 Introduction

On Mondays, Wednesdays, and Fridays, I call myself a probabilist. In broad outline I agree with probabilism's key tenets: that

1. an agent's beliefs come in degrees, which we may call credences; and that
2. these credences are rationally required to conform to the probability calculus.

Here, 'the probability calculus' refers to at least the finite fragment of Kolmogorov's theory, according to which probabilities are non-negative, normalized (with a top value of 1), and finitely additive. Probabilism is a simple, secund theory. Indeed, it achieves such an elegant balance of simplicity and strength that, in the spirit of Ramsey's and Lewis's accounts of 'law of nature', I am inclined to say that probabilism codifies the synchronic laws of epistemology. Or so I am inclined on those days of the week.

But on the remaining days of the week I am more critical of probabilism. A number of well-known arguments are offered in its support, but each of them is inadequate. I do not have the space here to spell out all of the arguments, and all of their inadequacies. Instead, I will confine myself to four of the most important arguments—the Dutch Book, representation theorem, calibration, and gradational accuracy arguments—and I will concentrate on a particular inadequacy in each of them, in its most familiar form.

I think it is underappreciated how structurally similar these four arguments for probabilism are. Each begins with a mathematical theorem that advertises credences

---

2 The Ramsey/Lewis account has it that a law of nature is a theorem of the best theory of the universe—the true theory that best balances simplicity and strength. I say 'synchronic' laws of epistemology to allow for there being further 'diachronic' laws about how credences should update in the face of evidence.
Call an agent who violates probability theory incoherent.\textsuperscript{5} Call a set of bets, each of which you consider fair, and which collectively guarantee your loss, a Dutch Book against you. The Dutch Book theorem tells us that if you are incoherent, there exists a Dutch Book against you. Note the logical form: a conditional with an existentially quantified consequent. The antecedent speaks of a violation of probability theory; the consequent states the existence of something that follows from such a violation. We will see this form again and again.

So much for the theorem. What about the argument for probabilism? It is so simple that it can be presented entirely in words of one syllable:\textsuperscript{6}

You give some chance to $p$: it is the price that you would pay for a bet that pays a buck if $p$ is true and nothing otherwise. You give some chance to $\neg p$: it is the price that you would pay for a bet that pays a buck if $p$ is false and nothing otherwise. And so on. Now, if you failed to live up to the laws of chance, then you could face a dire end. A guy—let’s make him Dutch—could make a set of bets with you, each fair by your lights, yet at the end of the day you would lose, come what may. What a fool you would be! You should not tempt this fate. So you should bet in line with the laws of chance.

This argument is invalid. For all the Dutch Book theorem tells us, you may be just as susceptible to Dutch Books if you obey probability theory. Maybe the world is an unknown place, and we’re all suckers! (Compare: it’s certain that if you pursue a career in philosophy, you will eventually die; but that’s hardly a reason to avoid a career in philosophy.) This possibility is ruled out by the surprisingly neglected, yet equally important Converse Dutch Book theorem: if you obey probability theory, then there does not exist a Dutch Book against you. So far, so good for probabilism.

But nothing can rule out the following mirror-image theorem, since it is clearly true. With an eye to the financial gains that are in the offing, let’s call it the Czech Book theorem.

**Czech Book Theorem**

If you violate probability theory, there exists a set of bets, each of which you consider fair, which collectively guarantee your gain.

The proof of the theorem is easy: just rewrite the proof of the original Dutch Book theorem, replacing ‘buying’ by ‘selling’ of bets, and vice versa, throughout. You thereby turn the original ‘Dutch Bookie’ who milks you into a Czech Bookie’ whom you milk. Call a set of bets, each of which you consider fair, and which collectively guarantee your gain, a Czech Book for you. The Czech Book theorem tells us that if you are incoherent, there exists a Czech Book for you. It is a simple piece of mathematics, and there is no disputing it.

So much for the theorem. I now offer the following argument against probabilism, again in words of one syllable. It starts as before, then ends with a diabolical twist:

\textsuperscript{3} I say ‘invalid’ to convey that the fault with each argument is that the conclusion does not follow from the theorem, rather than that the theorem is false. There’s a sense in which any argument for a necessary truth $p$ is valid—even ‘not $p$’. After all, it is not possible for the premises of the argument to be true and the conclusion false. So if probabilism is a necessary truth, then the argument ‘Snow is white: probabilism’ is valid in this sense. But philosophers often use ‘invalid’ in a different sense, according to which an argument is invalid if it is missing key steps needed to show us that its conclusion follows from its premises. This is the sense that I intend in this paper.

\textsuperscript{4} This section streamlines an argument given in my (2005), which concentrated solely on the Dutch Book argument.

\textsuperscript{5} de Finetti used the word ‘incoherent’ to mean ‘Dutch bookable’, while some other authors use it as I do. It will be handy for me to have this word at my disposal even when I am not discussing Dutch books.

\textsuperscript{6} The homage to George Boolos will be obvious to those who know his (1994).
2.1 Saving the Dutch Book Argument

If you survey the vast literature on Dutch Book arguments, you will find that most presentations of it focus solely on sets of bets, making use of your 

If you violate probability theory, there does not exist a set of bets, each of which you consider fair, and collectively guarantees your loss.

The proof of this theorem is straightforward. Let us assume, for the sake of contradiction, that you have found a set of bets, each of which you consider fair, and collectively guarantees your loss. Let the bets be denoted by $B_1, B_2, \ldots, B_n$. Then, by the property of fair bets, for each $i$, you assign a price $p_i$ to $B_i$ such that

$$p_i = \frac{1}{2} \text{ for all } i.$$ 

But this contradicts the fact that the bets are fair, as the expected value of each bet is not zero. Therefore, your assumption must be false, and there does not exist a set of bets, each of which you consider fair, and collectively guarantees your loss.

The Dutch Book argument is a powerful tool in probability theory, but it has been criticized for being too restrictive. Some philosophers have argued that it is not always possible to find a set of bets that collectively guarantees your loss. However, the Dutch Book argument is still widely used in decision theory and economics.

References:

- Good (1963) argues that the Dutch Book argument is not always applicable, as it assumes that all agents have the same beliefs.
- Jeffrey (1992) argues that the Dutch Book argument is not always applicable, as it assumes that all agents have the same beliefs.
- Bovens (1997) argues that the Dutch Book argument is not always applicable, as it assumes that all agents have the same beliefs.
- Williamson (1995) argues that the Dutch Book argument is not always applicable, as it assumes that all agents have the same beliefs.
2.2 The Dutch Book Argument: Merely Dramatic?

Rosenquist (1986) argued that the Dutch Book argument for probability is not compelling because it is based on hypothetical scenarios that are not relevant to real-world decisions. He claimed that the argument is merely dramatic and does not provide a strong reason to believe in the probability of an event.

However, this argument has been criticized by many philosophers of probability, including Skyrms (1975) and Ramsey (1926). They argue that the Dutch Book argument is not merely dramatic and that it provides a strong reason to believe in the probability of an event. According to Skyrms, the Dutch Book argument is based on a clear and principled decision rule that is necessary and sufficient for coherence.

The argument goes as follows: if you believe in a proposition with probability p, then you should be willing to accept a bet on that proposition at odds of 1:p. If you are not willing to accept the bet, then you are incoherent. If you are incoherent, then you are vulnerable to a Dutch Book, which is a set of bets that guarantees a profit no matter what happens.

For example, suppose you believe that it will rain with probability 0.5. If you are not willing to accept a bet on rain at odds of 1:1, then you are incoherent. If you are incoherent, then you are vulnerable to a Dutch Book, which consists of betting on rain at odds of 1:1, betting on no rain at odds of 1:1, and then collecting the winnings if it rains and the losses if it does not.

This argument is not merely dramatic because it provides a clear and principled decision rule that is necessary and sufficient for coherence. It is not a matter of opinion, but a matter of logical necessity. If you are not willing to accept the bet, then you are incoherent.

So, is the Dutch Book argument merely dramatic, or is it a compelling reason to believe in the probability of an event? The answer depends on your philosophical stance on probability. If you believe in subjective probability, then the argument is merely dramatic. If you believe in objective probability, then the argument is compelling.

Sources:

Note: This section is a brief introduction to the Dutch Book argument and its philosophical implications. For a more detailed discussion, please refer to the original sources.
attitudes. Either the Dutch Book bets or the Czech Book bets could be used to establish the existence claim. This again is a conditional with an existentially quantified consequent. Now I don’t have a mirror-image theorem to place alongside it, in order to undercut it.

However, nor have I seen the converse of this more fundamental putative theorem; still less am I aware of anyone claiming to have proved it. It seems to be a live possibility that if you obey probability theory, then there also exists a set of propositions to which you have inconsistent attitudes—not inconsistent in the sense of being Dutch-bookable (the converse Dutch Book theorem assures us of this), but inconsistent nonetheless. That is, I have not seen any argument that in virtue of avoiding the inconsistency of Dutch-bookability, at least some coherent agents are guaranteed to avoid all inconsistency. Without a proof of this further claim, it seems an open question whether probabilistically coherent agents might also have inconsistent attitudes (somewhere or other). Maybe non-extremal credences, probabilistic or not, necessarily manifest a kind of inconsistency. I don’t believe that, but I don’t see how the Dutch Book argument rules it out. The argument needs to rule it out in order to preempt the possibility of a partners-in-crime defense of non-probabilism: the possibility that we are all epistemically damned whatever we do. Indeed, if all intermediate credences were ‘inconsistent’ in this sense, then this sense of inconsistency would not seem so bad after all. I said earlier that the original Dutch Book argument, understood in terms of monetary losses, is invalid; the converse Dutch Book theorem came to its rescue (even though this theorem is surprisingly neglected). Now I am saying that the Ramsey-style Dutch Book argument, understood as dramatizing an inconsistency in attitudes, is similarly invalid; it remains to be seen if the converse of the putative theorem (italized in the previous paragraph) will come to its rescue.

I say ‘putative theorem’ because its status as a theorem is less clear than before—this status is disputed by various authors. Schick (1986) and Maher (1993) question the inconsistency of the attitudes at issue regarding the additivity axiom. They reject the ‘package principle’, which requires one to value a set of bets at the sum of the values of the bets taken individually, or less specifically, to regard a set of bets as fair if one regards each bet individually as fair. The package principle seems especially problematic when there are interference effects between the bets in a package—e.g., the placement of one bet is correlated with the outcome of another. For example, you may be very confident that your partner is happy; you will pay 90 cents for a bet that pays a dollar if so. You may be fairly confident that the Democrats will win the next election; you will pay 60 cents for a bet that pays a dollar if they win. So by the package principle, you should be prepared to pay $1.50 for both bets. But you also know that your partner hates you betting on political matters and inevitably finds

3 Representation Theorem-Based Arguments

The centerpiece of the argument for probabilism from representation theorems is some version of the following theorem, which I will not dispute.

Representation Theorem

If all your preferences satisfy certain ‘rationality’ conditions, then there exists a representation of you as an expected utility maximizer, relative to some probability and utility function.

(The ‘rationality’ constraints on preferences are transitivity, connectedness, independence, and so on.) The contrapositive gets us closer to the template that I detect in all the arguments for probabilism:

If there does not exist a representation of you as an expected utility maximizer, relative to some probability and utility function, then there exist preferences of yours that fail to satisfy certain ‘rationality’ conditions.

14 I suppose you are safe from such inconsistency if you obey probability theory trivially with a consistent assignment of 1’s and 0’s, corresponding to a consistent truth-value assignment. Let us confine our attention, then, to non-trivial probability assignments, which after all are the lifeblood of probabilism.

15 Thanks here to Kenny Easwaran.
Focusing on the probabilistic aspect of the antecedent, we have a corollary that fits the conditional-with-an-existentially-quantified-consequent form:

If your credences cannot be represented with a probability function, then there exist preferences of yours that fail to satisfy certain 'rationality' conditions.

The antecedent involves a violation of the probability calculus; the consequent states the existence of a putatively undesirable thing that follows: some violation of the 'rationality' conditions on preferences. In short, if your credences cannot be represented with a probability function, then you are irrational.

I will dispute that probabilism follows from the original theorem, and a fortiori that it follows from the corollary. For note that probabilism is, in part, the stronger thesis that if your credences violate probability theory, then you are irrational (a restatement of what I called tenet (2) at the outset). It is clearly a stronger thesis than the corollary, because its antecedent is weaker: while your credences cannot be represented with a probability function entails 'your credences violate probability theory', the converse entailment does not hold. For it is possible that your credences violate probability theory, and that nonetheless they can be represented with a probability function. Merely being representable some way or other is cheap, as we will see; it's more demanding actually to be that way. Said another way: it's one thing to act as if you have credences that obey probability theory, another thing to actually have credences that obey probability theory. Indeed, probabilism does not even follow from the theorem coupled with the premises that Maher adds in his meticulous presentation of his argument for probabilism, as we will also see.

The concern is that for all we know, the mere possibility of representing you one way or another might have less force than we want; your acting as if the representation is true of you does not make it true of you. To make this concern vivid, suppose that I represent your preferences with Voodooism. My voodoo theory says that there are warring voodoo spirits inside you. When you prefer A to B, then there are more A-favouring spirits inside you than B-favouring spirits. I interpret all of the usual rationality axioms in voodoo terms. Transitivity: if you have more A-favouring spirits than B-favouring spirits, and more B-favouring spirits than C-favouring spirits, then you have more A-favouring spirits than C-favouring spirits. Connectedness: any two options can be compared in the number of their favouring spirits. And so on. I then 'prove' Voodooism: if your preferences obey the usual rationality axioms, then there exists a Voodoo representation of you. That is, you act as if there are warring voodoo spirits inside you in conformity with Voodooism. Conclusion: rationality requires you to have warring Voodoo spirits in you. Not a happy result.

Hence there is a need to bridge the gap between the possibility of representing a rational agent a particular way, and this representation somehow being correct. Maher, among others, attempts to bridge this gap. I will focus on his presentation, because he gives one of the most careful formulations of the argument. But I suspect my objections will carry over to any version of the argument that infers the rational obligation of having credences that are probabilities from the mere representability of an agent with preferences obeying certain axioms.

Maher claims that the expected utility representation is privileged, superior to rival representations. First, he assumes what I will call interpretivism:

an attribution of probabilities and utilities is correct just in case it is part of an overall interpretation of the person's preferences that makes sufficiently good sense of them and better sense than any competing interpretation does (1993, 9).

Then he maintains that, when available, an expected utility interpretation is a perfect interpretation:

if a person's preferences all maximize expected utility relative to some p and u, then it provides a perfect interpretation of the person's preferences to say that p and u are the person's probability and utility functions.

He goes on to give the argument from the representation theorems:

... we can show that rational persons have probability and utility functions if we can show that rational persons have preferences that maximize expected utility relative to some such functions. An argument to this effect is provided by representation theorems for Bayesian decision theory.

He then states the core of these theorems:

These theorems show that if a person's preferences satisfy certain putatively reasonable qualitative conditions, then those preferences are indeed representable as maximizing expected utility relative to some probability and utility functions (1993, 9).

We may summarize this argument as follows:

**Representation Theorem Argument**

1. (Interpretivism) You have a particular probability and utility function iff attributing them to you provides an interpretation that makes:

- (i) sufficiently good sense of your preferences and
- (ii) better sense than any competing interpretation.

2. (Perfect interpretation) Any maximizing-expected-utility interpretation is a perfect interpretation (when it fits your preferences).

3. (Representation theorem) If you satisfy certain constraints on preferences (transitivity, connectedness, etc.), then you can be interpreted as maximizing expected utility.

4. The constraints on preferences assumed in the representation theorem of 3 are rationality constraints.

Therefore (generalizing what has been established about 'you' to 'all rational persons').

**Conclusion:** [All] rational persons have probability and utility functions (1993, 9)

The conclusion is probabilism, and a bit more, what we might call utilitism.

According to Premise 1, a necessary condition for you to have a particular probability and utility function is their providing an interpretation of you that is better than
If all your preferences satisfy the same ‘rationality’ conditions, then you can be interpreted as maximizing non-expected utility, some rival to expected utility, and in particular as having credences that violate probability theory.

How can this be? The idea is that the rival representation compensates for your credences’ violation of probability theory with some non-standard rule for combining your credences with your utilities. Zynda (2000) proves this mirror-image theorem. As he shows, if you obey the usual preference axioms, you can be represented with a sub-additive belief function, and a corresponding combination rule. For all that

Maher’s argument shows, this rival interpretation may also be ‘perfect’.

According to probabilism, rationality requires an agent’s credences to obey the probability calculus. We have rival ways of representing an agent whose preferences obey the preference axioms; which of these representations correspond to her credences? In particular, why should we privilege the probabilistic representation? Well, there may be reasons. Perhaps it is favoured by considerations of simplicity, fertility, consilience, or some other theoretical virtue or combination thereof—although good luck trying to clinch the case for probabilism by invoking these rather vague and ill-understood notions. And it is not clear that these considerations settle the issue of what rational credences are, as opposed to how they can be fruitfully modelled. (See Eriksson and Hájek 2007 for further discussion.) It seems to be a further step, and a dubious one at that, to reify the theoretical entities in our favourite model of credences.

It might be objected that the ‘rival’ representations are not really rival. Rather, the objection goes, they form a family of isomorphic representations, and choosing among them is merely a matter of convention; whenever there is a probabilistic representation, all of these other representations impose exactly the same laws on rational opinion, just differently expressed. First, a perhaps flat-footed reply: I understand ‘probabilism’ to be defined via Kolmogorov’s axiomatization of probability. So, for example, a non-additive measure is not a probability function, so understood. That said, one might want to have a broader understanding of ‘probabilism’, encompassing any transformation of a probability function and a correspondingly transformed combination rule for utility that yields the same ordinal representation

preferenced. If that is what is intended, then probabilism should be stated in those terms, and not in the flat-footed way that is entirely standard. We would then want to reexamine the arguments for probabilism in that light—presumably with a revised account of ‘credence’ in terms of betting, a revised statement of what ‘calibration’ consists in, and revised axioms on ‘gradational accuracy’. But I am getting ahead of myself—calibration, and gradational accuracy are just around the corner!

In any case, I believe that my main point stands, even with a more liberal understanding of ‘probabilism’: the representation theorem argument is invalid. We have the theorem:

if you obey the preference axioms, then you are representable by a credence function that is a suitable transformation of a probability function.

But to be able validly to infer probabilism in a broad sense, we need the further theorem:

if you obey the preference axioms, then you are not also representable by a credence function that is not a suitable transformation of a probability function.

It seems that the status of this is at best open at the moment. The representation theorem argument for probabilism remains invalid until the case is closed in favour of the further theorem.

4 The Calibration Argument

The centerpiece of the argument is the following theorem—another conditional with an existentially quantified consequent—which I will not dispute.

**Calibration Theorem**

If \( c \) violates the laws of probability then there is a probability function \( c^+ \) that is better calibrated than \( c \) under every logically consistent assignment of truth-values to propositions.

Calibration is a measure of how well credences match corresponding relative frequencies. Suppose that you assign probabilities to some sequence of propositions—for example, each night you assign a probability to it raining the following day, over a period of a year. Your assignments are (perfectly) calibrated if proportion \( p \) of the propositions to which you assigned probability \( p \) are true, for all \( p \). In the example, you are perfectly calibrated if it rained on 10% of the days to which you assigned probability 0.1, on 75% of the days to which you assigned probability 0.75, and so on. More generally, we can measure how well calibrated your assignments are, even if they fall short of perfection.

The clincher for probabilism is supposed to be the calibration theorem. If you are incoherent, then you can figure out a priori that you could be better calibrated by being coherent instead. Perfect calibration, moreover, is supposed to be A Good Thing, and a credence function that is better calibrated than another one is thereby supposed to be superior in at least one important respect. Thus, the argument con-

---

16 Only arguably. In fact, I think that it does not follow, because the preference axioms are not all rationality constraints.

17 I thank Harry Field and Jim Joyce for independently offering versions of this objection to me.
I argue elsewhere (MS) that perfect calibration may be *A Rather Bad Thing*, as does Seidenfeld (1985) and Joyce (1998). More tellingly, the argument, so presented, is invalid.

I will not quarrel with the calibration theorem. Nor should the probabilist quarrel with the following 'mirror-image' theorem:

If \( c \) violates the laws of probability then there is a *non-probability* function \( c^+ \) that is better calibrated than \( c \) under every logically consistent assignment of truth-values to propositions.

Think of \( c^+ \) as being more coherent than \( c \), but not *entirely* coherent. If \( c \) assigns, say, 0.2 to rain and 0.7 to not-rain, then an example of such a \( c^+ \) is a function that assigns 0.2 to rain and 0.75 to not-rain. If you are incoherent, then you can find a priori that you could be better calibrated by *staying in coherent*, but in some other way.  

So as it stands, the calibration argument is invalid. Given that you can improve your calibration situation *either* by moving to some probability function *or* by moving to some other non-probability function, *why* do you have an incentive to move to a probability function? The answer, I suppose, is this. If you moved to a non-probability function, you would only recreate your original predicament; you would know a priori that you could do better by moving to a probability function. Now again, you could *also* do better by moving to yet another non-probability function. But the idea is that moving to a non-probability function will give you no rest; it can never be a stable stopping point. Still, the argument for probabilism is invalid as it stands. To shore it up, we had better be convinced that at least some probability functions are stable stopping points.

The following converse theorem would do the job:

If \( c \) obeys the laws of probability then there is no another function \( c^+ \) that is better calibrated than \( c \) under every logically consistent assignment of truth-values to propositions.

I offer the following near-trivial proof: Let \( P \) be a probability function. \( P \) can be perfectly calibrated—just consider a world where the relative frequencies are exactly as \( P \) predicts. As required by calibration. (If \( P \) assigns some irrational probabilities, then the world will have to provide infinite sequences of the relevant trials, and calibration will involve agreement with limiting relative frequencies.) At that world, no other function can be better calibrated than \( P \). Thus, \( P \) cannot be beaten by any other function, come what may, in its calibration index—*for short, \( P \) is not calibration-dominated.* Putting this result together with the calibration theorem, we have the result that probability functions are exactly the functions that are not calibration-dominated.

The original calibration argument for probabilism, as stated above, was invalid, but I think it can be made valid by the addition of this theorem. However, this is not yet a happy ending for calibrationists. If you are a fan of calibration, surely what matters is being well calibrated in the *actual* world, and being coherent does not guarantee that. A coherent weather forecaster who is wildly off step with the actual relative frequencies can hardly plead that at least he is perfectly in step with the relative frequencies in some other possible world! (Compare: someone who has consistent but wildly false beliefs can hardly plead that at least his beliefs are true in some other possible world!)

5 The Gradational Accuracy Argument

Joyce (1998) rightly laments the fact that ‘probabilists have tended to pay little heed to the one aspect of partial beliefs that would be of most interest to epistemologists: namely, their role in representing the world’s state’ (576). And he goes on to say: ‘I mean to alter this situation by first giving an account of what it means for a system of partial beliefs to accurately represent the world, and then explaining why having beliefs that obey the laws of probability contributes to the basic epistemic goal of accuracy.’

The centerpiece of his ingenious (1998) argument is the following theorem—yet another conditional with an existentially quantified consequent—which I will not dispute.

**Gradational Accuracy Theorem**

If \( c \) violates the laws of probability then there is a probability function \( c^+ \) that is strictly more accurate than \( c \) under every logically consistent assignment of truth-values to propositions (Joyce 2004, 143).

---

18 To be sure, the mirror-image theorem gives you no advice as to which non-probability function you should move to. But nor did the calibration theorem give you advice as to which probability function you should move to. Moreover, for all the theorem tells us, you can worsen your calibration index, come what may, by moving from a non-probability function to a ‘wrong’ probability function. Here’s an analogy (adapted from Aaron Bronfman and Jim Joyce, personal communication). Suppose that you want to live in the best city that you can, and you currently live in an American city. I tell you that for each American city, there is a better American city. (I happen to believe this.) It does not follow that you should move to Australia. If you do not know which Australian city or cities are better than yours, moving to Australia might be a backward step. You might choose Coober Pedy.

That said, the calibration argument may still be probative, still diagnostic of a defect in an incoherent agent’s credences. To be sure, she is left only with the general admonition to become coherent, without any advice on how specifically to do so. Nevertheless, the admonition is non-trivial. Compare: when an agent has inconsistent beliefs, logic may still be probative, still diagnostic of a defect in them. To be sure, she is left only with the general admonition to become coherent, without any advice on how specifically to do so. Nevertheless, the admonition is non-trivial.

19 Seidenfeld (1985) has a valuable discussion of a theorem due to Pratt and rediscovered by Dawid that may seem to yield the desired result. Its upshot is that if an agent is probabilistically coherent, and updates by conditionalization after each trial or feedback information about the result of that trial, then in the limit calibration is achieved almost surely (according to her own credences). This is an important result, but it does not speak to the case of an agent who is coherent but who has not updated on such an infinite sequence of feedback information, and indeed who may never do so (e.g., because she never gets such information).
Saying that incoherent measures are ‘always less accurate than they need to be’ suggests that they are always unnecessarily inaccurate—that they always could be more accurate. But this would not distinguish incoherent measures from coherent measures that assume non-extremal values—that is, coherent measures that are not entirely opiniated. After all, such a measure could be more accurate: an opiniated measure that assigns 1 to the truth and 0 to all false alternatives to it, respectively, is more accurate. Indeed, if a coherent measure \( P \) assigns non-extremal values, then necessarily there exists another measure that is more accurate than \( P \). In each possible world there exists such a measure, namely an opiniated measure that gets all the truth values right in that world. More disturbingly, if a coherent measure \( P \) assigns non-extremal values, then necessarily there exists an incoherent measure that is more accurate than \( P \); for example, one that raises \( P \)'s non-extremal probability for the truth to 1, while leaving its probabilities for falsehoods where they are. (The coherent assignment \( 1/2, 1/2 \) for the outcomes of a coin toss, \(<\text{Heads}, \text{Tails}>\), is less accurate than the incoherent assignment \( 1, 1/2 \) in a world where the coin lands Heads, and it is less accurate than the incoherent assignment \( 1/2, 1 \) in a world where the coin lands Tails.) But this had better not be an argument against the rationality of having a coherent intermediate-valued credence function—that would hardly be good news for probabilism!

The reversal of quantifiers in Joyce’s theorem appears to save the day for probabilism. It isn’t merely that:

if your credences violate probability theory, in each possible world there exists a probability function that is more accurate than your credences.

More than that, by his theorem we have that:

if your credences violate probability theory, there exists a probability function such that in each possible world, it is more accurate than your credences.

The key is that the same probability function outperforms your credences in each possible world, if they are incoherent. Thus, by the lights of gradational accuracy you would have nothing to lose and everything to gain by shifting to that probability function. So far, so good. But we had better be convinced, then, that at least some coherent intermediate-valued credences do not face the same predicament—that they cannot be outperformed in each possible world by a single function (probability, or otherwise). Well, let’s see.

With the constraints on reasonable gauges of accuracy in place, Joyce (1998) proves the gradational accuracy theorem. He concludes: ‘To the extent that one accepts the axioms, this shows that the demand for probabilistic consistency follows from the purely epistemic requirement to hold beliefs that accurately represent the world’ (2004, 143).
of the constraint that being weakly dominated is a necessary condition for being weakly dominated according to $S$.

Being weakly dominated according to a reasonable scoring rule is an undesirable property of a function: holding that function is apparently precluded, since there is another function that is guaranteed to do no worse, and that could do better, by $S$’s lights. The constraint of coherent admissibility on a reasonable scoring rule $S$ is that $S$ will never attribute the undesirable property of being weakly dominated to a coherent admissibility function.

**Coherent Admissibility.** No coherent admissibility function is weakly dominated according to $S$.

The constraint of truth-directedness on $S$ is that $S$ should favour a credence function over another at a world if the former’s assignments are uniformly closer than the latter’s to the truth values in that world:

**Truth Directedness.** If $b$’s assignments are uniformly closer than $c$’s to the truth values according to $v$, then $S(b, v) < S(c, v)$.

Now we can state the final theorem of Joyce’s paper (this volume):

**Theorem.** Let $S$ be a scoring rule defined on a partition $X = \{X_1\}$. If $S$ satisfies TRUTH DIRECTEDNESS and COHERENT ADMISSIBILITY, and if $S(b, v)$ is finite and continuous for all $b$ in $B_X$ and $v$ in $V_X$, then

(i). every incoherent credence function is strongly dominated according to $S$ and, moreover, is strongly dominated by some coherent credence function, and

(ii). no coherent credence function is weakly dominated according to $S$.

(i). is the counterpart to the original gradational accuracy theorem, and it is striking that the domination of incoherent functions follows from the non-domination of coherent functions, and seemingly weak further assumptions. (ii), entails the converse theorem that I have contended was missing from Joyce’s (1998) argument.

So does this lay the matter to rest, and give us a compelling argument for probabilism? Perhaps not. For (ii), just is the constraint of Coherent Admissibility on scoring rules—the rules have been pre-selected to ensure that they favour coherent credence functions. In short, Coherent Admissibility is question-begging with respect to the converse theorem. (By contrast, the other constraints of Truth Directness, continuity, and finiteness are not.) Another way to see this is to introduce into the debate Mr. Incoherent, who insists that credences are rationally required to violate the probability calculus. Imagine him imposing the mirror-image constraint on scoring rules:

**Incoherent Admissibility.** No incoherent credence function is weakly dominated according to $S$.

Then (i). would be rendered false, and of course the mirror-image of (ii). would be trivially true, since it just is the constraint of Incoherent Admissibility. From a neutral standpoint, which precludes the issue in favour of neither coherence nor incoherence, offhand it would appear that Incoherent Admissibility is on all fours with Coherent Admissibility.

How would Joyce convince Mr. Incoherent that Coherent Admissibility is the correct constraint to impose, and not Incoherent Admissibility, using premises that they ought to share? Joyce offers the following argument that any coherent credence function can be rationally held (under suitable conditions), and that this in turn limits which scoring rules are acceptable. He maintains that it is ‘plausible’ that there are conditions under which any coherent credence function can be rationally held. After all, for any assignment of probabilities ($p_X$) to $(X_1)$ it seems that a believer could, in principle, have evidence that justifies her in thinking that each $X_i$ has $p_i$ as its objective chance. Moreover, this could exhaust her information about $X_1$’s truth-value. According to the ‘Principled Principle’ of Lewis (1980), someone who knows that the objective chance of $X_1$ is $p_X$, and who does not possess any additional information that is relevant to questions about $X_1$’s truth-value, should have $p_X$ as her credence for $X_1$. Thus, ($p_X$) is the rational credence function for the person to hold under these conditions. In light of this, one might argue, the following restriction on scoring rules should hold:

**Minimal Coherence:** An epistemic scoring rule should never preclude, a priori, the holding of any coherent set of credences.

(263–297, this volume).

So we have here a putative reason to impose Coherent Admissibility on a reasonable scoring rule $S$. It obviates the putatively unacceptable situation in which a coherent credence function is precluded, insofar as it is weakly dominated according to $S$—unacceptable, since for any coherent assignment of credences, one could have evidence that it corresponds to the objective chances. To complete the argument, we apparently have no parallel reason for imposing Incoherent Admissibility on $S$, for one could not have evidence that an incoherent assignment of credences corresponds to the objective chances.

However, it is not clear to me that any assignment of probabilities could correspond to the objective chances, still less that one could have evidence for any particular assignment that this is the case. There may necessarily be chance gaps to which some other probability functions could nevertheless assign values. For example, propositions about the chance function (at a time) might be ‘blind spots’ to the function itself but could be assigned probabilities by some other function. Perhaps
there are no higher-order chances, such as the chance of: the chance of Heads is 1/2, even though the chance of Heads is 1/2, is a proposition, and thus fit to be assigned a value by some probability functions. Or perhaps such higher-order chances are defined, but they are necessarily 0 or 1; and yet some other probability function could easily assign an intermediate value to ‘the chance of Heads is 1/2’.

Moreover, it is clear to me that not any coherent credence function can be rationally held. For starters, any coherent credence function that violates the Principle Principle cannot be—and presumably Joyce agrees, given his appeal to it in his very argument. Indeed, if there are any constraints on rational credences that go beyond the basic probability calculus, then coherent violations thereof are counterexamples to Joyce’s opening claim in the quoted passage. Choose your favourite such constraint—the Reflection Principle, or regularity, or the principle of indifference, or what have you. My own favourite is a prohibition on Moore paradoxical credences, such as my assigning high credence to ‘p & my credence in p is low’ or to ‘p & I don’t assign high credence to p’.

Epistemically rational credence is more demanding than coherent credence, then there will be coherent credences that are rationally precluded. More power to an epistemic scoring rule, I say, if it precludes the holding of them!

So I am not persuaded by this defence of the Coherent Admissibility constraint, as stated. And to the extent that one is moved by this defence, it would seem to provide a more direct argument for coherence—from the coherence of chances and the Principle Principle—without any appeal to scoring rules.

Now, perhaps a slight weakening of the constraint can be justified along the lines of Joyce’s argument. After all, some of the problematic cases that I raised involved higher-order probability assignments of one kind or another (higher order chances, the Principal Principle, the Reflection Principle, and the prohibition on Moore paradoxical credences), and the others (regularity and the principle of indifference) are quite controversial. So perhaps Joyce’s argument goes through if we restrict our attention to partitions (X_n) of probability-free propositions, and to purely first-order probability assignments to them. Then it seems more plausible that any coherent assignment of credences across such a partition should be admissible, and that a scoring rule that ever judges such an assignment to be weakly dominated is unreasonable.

The trouble is that then Joyce would seem to lose his argument for probabilism tout court, as opposed to a watered-down version of it. Probabilism says that all credences are rationally required to conform to the probability calculus—not merely that credences in probability-free propositions are so required. Consider, then, a credence function that is coherent over probability-free propositions, but that is wildly incoherent over higher-order propositions. It is obviously defective by probabilist lights, but the concern is that its defect will go undetected by a scoring rule that is confined to probability-free propositions. And how probative are scoring rules that are so confined, when an agent’s total epistemic state is not so confined, and should be judged in its entirety?

6 Conclusion

I began by confessing my schizophrenic attitude to probabilism. I have argued that the canonical statements of the major arguments for it have needed some repairing. Why, then, am I sympathetic to it at the end of the day, or at least at the end of some days? Partly because I think that to some extent the arguments can be repaired, and I have canvassed some ways in which this can be done, although to be sure, I think that some other problems remain. To the extent that they can be repaired, they provide a kind of triangulation to probabilism. And once we get to probabilism, it provides us with many fruits. Above all, it forms the basis of a unified theory of decision and confirmation—it combines seamlessly with utility theory to provide a fully general theory of rational action, and it illuminates or even resolves various hoary paradoxes in confirmation theory.

I consider that to be the best argument for probabilism. Sometimes, though, I wonder whether it is good enough.

So on Mondays, Wednesdays and Fridays, I call myself a probabilist. But as I write these words, today is Saturday.

References


23 Thanks here to Kenny Easwaran and (independently) Michael Titelbaum.

24 Thanks here to Jim Joyce and (independently) Kenny Easwaran.

25 See Earman (1992), Jeffrey (1992), and Howson and Urbach (1993), among others.

26 For very helpful comments I am grateful to: Selim Berker, David Chalmers, James Chase, John Cusbert, Lina Eriksson, James Ladyman, Aidan Lyon, Ralph Miles, Katie Steele, Michael Titelbaum, and especially Kenny Easwaran (who also suggested the name ‘Czech Book’), Justin Fisher, Franz Huber, Carrie Jenkins, Jim Joyce, and Andrew McConigal.