THE DIALECTICS OF ACCURACY ARGUMENTS FOR PROBABILISM

ALEXANDER R. PRUSS

Abstract. Scoring rules measure the deviation between a credence assignment and reality. Probabilism holds that only those credence assignments that satisfy the axioms of probability are rationally admissible. Accuracy-based arguments for probabilism observe that given certain conditions on a scoring rule, the score of any non-probability is dominated by the score of a probability. The conditions in the arguments always include propriety: the claim that the expected accuracy of p is not beaten by the expected accuracy of any other credence c by the lights of p if p is a probability. I argue that if we think through how a non-probabilist is likely to respond to pragmatic arguments for probabilism, then the non-probabilist will accept a condition *stronger* than propriety for the same reasons that the probabilist gives for propriety, but this stronger condition is incompatible with the other conditions that the probabilist needs to run the argument. This makes it unlikely for the probabilist's argument to be compelling.

1. Introduction

Scoring rules measure the inaccuracy or deviation from reality of credence function. Strictly proper scoring rules have the property that for any credence function that satisfies the axioms of probability, the mathematical expectation of the score of a credence function p by the lights of p is strictly better than the mathematical expectation of any other credence function c by the lights of p. Credence functions need not satisfy the axioms of probability, but assuming strict propriety and modest auxiliary assumptions, it has been shown that the score of a credence function that does not satisfy the axioms of probability is strictly dominated by the score of a one that does satisfy these axioms. These results have been interpreted by epistemologists as supporting probabilism, the thesis that reasonable credences will always be probabilistically consistent (e.g., [5], [4], [8]).

However, there has been little sympathetic engagement by defenders of probabilism with the kind of views a non-probabilist is likely to hold. Accuracy arguments are not the only arguments for probabilism: historically, pragmatic arguments have had pride of place in the probabilist's arsenal.??refs It is thus worth thinking what a non-probabilist who already has a well-developed answer to pragmatic arguments is likely to say about the accuracy arguments. I will take as a paradigm of this a non-probabilist who accepts Pruss's $\mathrm{LSI}^{\uparrow}$ level set integral [11] as yielding a prevision for basing practical decisions on inconsistent credences satisfying certain axioms, and thereby escapes all Dutch Book and some dominance arguments. I will argue that such a non-probabilist has three plausible replies to the accuracy arguments. The upshot will be that accuracy arguments provide some evidence for probabilism, but not very much.

First, a quick review of scoring rules. Let Ω be a finite sample space, encoding the possible situations that the crdences concern. Let the credence functions $\mathcal{C}$ be the functions from the power set of Ω to the interval $[0,1]$.[1] Let $\mathcal{P}$ be the subset of $\mathcal{C}$ which consists of the functions satisfying the axioms of probability. An *inaccuracy scoring rule* is a function s from a set $\mathcal{F} \supseteq \mathcal{P}$ of credence function to $[M,\infty]^{\Omega}$ for some finite M, where A^B is the set of functions from B to A. Then $s(c)(\omega)$ for $c \in \mathcal{F}$ measures the accuracy of the credence function c when we are in fact at $\omega \in \Omega$, with higher values being worse, less accurate.

Given a probability $p \in \mathcal{P}$ and an extended real function f on Ω, let $E_p f$ be the expected value with respect to p defined in the following way to avoid multiplying infinity by zero:

$$E_p f = \sum_{\omega\in\Omega, p(\{\omega\})\neq 0} p(\{\omega\}) f(\omega).$$

We then say that a scoring rule s is *proper* on $\mathcal{F} \supseteq \mathcal{P}$ provided that for every $p \in \mathcal{P}$ and every $c \in \mathcal{F}$, we have $E_p s(p) \le E_p s(c)$, that it is *strictly proper* there provided the inequality is always strict, and that it is *quasi-strictly proper* there provided that it is proper and the inequality is strict when $p \in \mathcal{P}$ and $c \in \mathcal{F} - \mathcal{P}$.

Propriety captures the idea that if an agent adopts a probability function p as their view, then by the agent's lights there can be no improvement in the expected score from switching to a different credence function. Strict propriety captures the idea that such an will expect other credence functions to be inferior. Proper and strictly proper scoring rules have been widely studied: for instance, see [2], [3], [8], [9], [14].

A scoring rule is said to be *additive* provided that $\mathcal{F} = \mathcal{C}$ and there is a collection of functions $(s_A)_{A\subseteq\Omega} : \mathbb{R} \times \{0,1\} \to [M,\infty]$ for a finite M such that for all $c \in \mathcal{F}$ and $\omega \in \Omega$:

$$s(c)(\omega) = \sum_{A\subseteq\Omega} s_A(c(A), 1_A(\omega)),$$

where $1_A(\omega)$ is 1 if $\omega \in A$ and 0 otherwise.

The set of probabilities $\mathcal{P}$ can be equipped with the topology resulting from its natural embedding ψ in the $|\Omega|$-dimensional cube $[0,1]^{\Omega}$, where

[1] Some of our discussion will then be simplified by not considering negative credences and credences greater than one. The kind of non-probabilist that we will be considering will be one that will place some reasonable constraints on credences, and making credences range from 0 to 1 certainly seems reasonable.

$\psi(p)(\omega) = p(\{\omega\})$. Thus, a sequence of probabilities (p_n) converges to a probability p just in case $p_n(\{\omega\}) \to p(\{\omega\})$ for all $\omega \in \Omega$.

A scoring rule is *probability-continuous* provided that the restriction of s to $\mathcal{P}$ is a continuous function to $[M, \infty]^\Omega$ equipped with the Euclidean topology.

Say that a credence assignment c_1 is *(strictly) s-dominated* by c_2 provided that $s(c_2)(\omega) < s(c_1)(\omega)$ for all $\omega \in \Omega$.

Predd, et al. [9] showed that if s is a probability-continuous additive strictly proper scoring rule, then for any non-probability c, there is a probability p such that p strictly s-dominates c. In other words, any forecaster whose forecast fails to be a probability can find a forecast that is a probability and that is strictly better no matter what. Recently, Pettigrew [7] announced that this result holds without the assumption of additivity, merely assuming probability-continuity. While Pettigrew's proof was flawed, correct proofs can be found in Nielsen [6] and Pruss [10, 13]. Nielsen's proof also extended the result to the quasi-proper case. Finally Pruss [13] showed that certain non-trivial conditions weaker than continuity suffice for the domination result.

2. A way to be a non-probabilist

Let's call credence assignments that could be rationally permitted "rationally admissible". The probabilist says that only credences that are probabilities are rationally admissible. The non-probabilist denies this, and says that some non-probabilities are rationally admissible. But the non-probabilist is unlikely to take a credence function that assigns 0 to all tautologies and 1 to all other propositions to be rationally admissible. All that is needed to be a non-probabilist is to say that rational admissibility allows at least one credence assignment that isn't a probability function. Our non-probabilist is likely to have some constraints on which credences are rationally admissible. Reasonable minimal constraints would be $c(\varnothing) = 0$ (*Zero*) and $c(\Omega) = 1$ (*Normalization*), but we can expect others.

Now, historically, accuracy arguments are not the only arguments for probabilism. There are also pragmatic arguments, especially ones based on Dutch Books and utility domination.??refs A non-probabilist should have something to say about these. Now whether a given credence assignment gives rise to unfortunate pragmatic consequences depends on how we link credences with decisions. For instance, the link can be given by a preference comparison where $f \precsim_c g$ if and only if the portfolio with the utility function g is at least as desired as the portfolio with utility function f in the light of credence c.[2] One will, then, want $\precsim_c$ to satisfy some formal axioms when

[2]This way of doing things already assumes that we treat two portfolios with the same utility function interchangeably. This assumption is not satisfied for every method of linking decisions to credences. It is not satisfied, for instance, by the method presupposed by De Finetti's pragmatic arguments for probabilism. In the context of preference comparisons derived from previsions, the interchangeability of of portfolios with the same utility

c is rationally admissible. For instance, it is very reasonable to want $\precsim_c$ to satisfy the *strict dominance* axiom that if $f < g$ everywhere, then $f \prec_c g$, where $f \prec_c g$ if and only if $f \precsim_c g$ but not $g \precsim_c f$. Otherwise, $\precsim_c$ would allow one to accept a portfolio whose yield is guaranteed to be worse than that of another portfolio.

In the case of classical decision theory, the link between credence and decisions uses mathematical expectation. If one is choosing between portfolio with different utility functions from Ω to $\mathbb{R}$, and one has a probabilistic credence assignment p, then one should opt for a portfolio for which the mathematical expectation $E_p f$ of the utility function f is biggest. Now, $E_p f$ is normally defined only for a probability p. To adapt this to decisions made in the light of non-probabilistic credences, one can adapt $E_p f$ to get some sort of "prevision" $V_c f$ of the utility f given a non-probabilistic credence c. Given such a prevision, one can then define a corresponding preference comparison by saying that $f \precsim_c g$ if and only if $V_c f \le V_c g$. Say that V_c is *strongly monotonic* providing that if $f < g$ everywhere, then $V_c f < V_c g$.

Recently, Pruss [11] offered level set integrals as a way of calculating a prevision of a utility given a non-probabilistic credence. One of his two ways of doing so is:

$$\mathrm{LSI}_c\, f = \int_0^\infty c(\{\omega : f(\omega) > y\})\, dy,$$

for non-negative real-valued f. If f is allowed to be negative, then we define $\mathrm{LSI}_c\, f = -\alpha + \mathrm{LSI}_c(\alpha + f)$ where α is a sufficiently large real number that $\alpha + f$ is everywhere non-negative.[3] (It turns out that the definition does not depend on the choice of α.) This agrees with mathematical expectation $E_c f$ when c is a probability function. Given the reasonable Zero and Normalization constraints on c, Pruss [11], working with utilities that are everywhere finite, proves that LSI_c is strongly monotonic if and only if c satisfies the Monotonicity Axiom that $c(A) \le c(B)$ if $A \subseteq B$. Let $\mathcal{M}$ be the set of credences that satisfy Zero, Normalization and Monotonicity. Pruss also shows that decision-making procedures with credences in $\mathcal{M}$ using level set integrals can avoid many pragmatic arguments for probabilism (though one of his proofs is flawed; see our Appendix for a fix).

Note that while Pruss only defined level set integrals for real-valued f, we can extend the definition to cases where f takes values in $[-\infty, \infty)$ or in $(-\infty, \infty]$ (but not both, so we avoid $\infty - \infty$), which we will need in the

function is equivalent to saying that the prevision is "integral-like" [11]. Because our task in this paper is to consider how the accuracy arguments for probabilism fare against the most plausible versions of non-probabilism, and because a preference comparison that allows one to distinguish two portfolios that have exactly the same utility function—say, because they arrange the wagers in the portfolio in a different but equivalent way—is *eo ipso* problematic, this is a reasonable assumption in our context. We should not expect a smart non-probabilist to make such a problematic pragmatic distinction.

[3]Our LSI_c corresponds to $\mathrm{LSI}_c^{\uparrow}$ in [11]. We ignore his $\mathrm{LSI}_c^{\pm}$, because it does not commute with positive affine transformations, as we would expect in a utility prevision, since utilities are normally thought to be defined only up to positive affine transformations.

application to scoring rules. Specifically, for $M \geq 0$, let $f_M(\omega) = f(\omega)$ if $|f(\omega)| \leq M$ and $f_M(\omega) = M \operatorname{sgn} f(\omega)$ (where $\operatorname{sgn} x$ is 1 if $x > 0$ and -1 if $x < 0$), and let

$$\mathrm{LSI}_c\, f = \lim_{M \to \infty} \mathrm{LSI}_c\, f_M.$$

This agrees with the previous definition when f is finite-valued, and Zero, Normalization and Monotonicity continue to imply strong monotonicity (see the Appendix).

While Pruss does not himself advocate non-probabilism, it is reasonable to think that something like his approach is a plausible model of a non-probabilist response to pragmatic arguments for probabilism: impose some formal constraints, such as Zero, Normalization and Monotonicity, on credences and find a decision procedure that allows escape from many pragmatic arguments for probabilism.

3. A reaction to the accuracy arguments

Now, let us consider how our non-probabilist is likely to react to the accuracy arguments for probabilism. The accuracy theorist says that given an inaccuracy scoring rule s that satisfies certain conditions, for any non-probabilistic credence c, there is a probability p such that $s(c) > s(p)$ everywhere on Ω, i.e., no matter what, c is more inaccurate according to s than p is. Thus, it is concluded, it is irrational to adopt c as one's credence, since one would be sure to be less inaccurate to adopt p.

How convincing this line of thought is depends on whether the non-probabilist should be expected to agree that s correctly measures the inaccuracy of a credence function. If we simply adopted the nearly trivial scoring rule where $t(c)(\omega) = 0$ if c is a probability and $t(c)(\omega) = 1$ if it's not, then we would have the radical domination result that $t(c) > t(p)$ whenever c isn't a probability and p is, but of course the non-probabilist is not going to agree that t is a good measure of inaccuracy, and will rightly insist that t is *ad hoc*.

The usual proceeding in accuracy arguments for probabilism is not so *ad hoc*. Rather, one imposes constraints on the scoring rule s that appear plausible, and proves that these imply that for every non-probability c there is a probability p such that $s(c) > s(p)$ everywhere.

Common to all the versions of these arguments is *propriety*: we assume that $E_p s(p) \leq E_p s(c)$ whenever p is a probability and c is a credence different from p. The thought is that given a probability p there should't be another credence, c, which by the lights of p would be expected to be less inaccurate. If we had such p and c, then an agent who had credence assignment p would be rationally required to switch to c on accuracy grounds without any evidence, and this is implausible.

Now, our non-probabilist may well find propriety compelling. However, a non-probabilist thinks that some non-probabilities—say, those in $\mathcal{M}$—are rationally admissible. And they will thus think that the above argument about evidenceless switching should apply to the rationally admissible non-probabilities as well, as long as one replaces comparisons of the mathematical expectation E with the appropriate preference comparison.

Let $\mathcal{A} \subseteq \mathcal{C}$ be the set of credences satisfying the conditions our non-probabilist thinks yield rational admissibility and suppose $\precsim_c$ is the associated preference comparison. Then the non-probabilist impressed by the reasoning behind propriety is going to insist that $-s(c) \precsim_r -s(r)$ whenever $r \in \mathcal{A}$ and c is any credence other than r, a condition I will call $(\mathcal{A}, \precsim)$-propriety; if $\precsim_c$ is derived from a prevision V_c, I will also call it $(\mathcal{A}, V)$-propriety, with the condition then being equivalent to $V_r(-s(c)) \leq V_r(-s(r))$. The reason for the negative signs is that $\precsim_r$ is meant for utilities, and our inaccuracy scores are *dis*utilities.[4]

So far we have no disagreement between the person offering the accuracy argument for probabilism and the non-probabilist. The non-probabilist is likely to think that all the probabilities are rationally admissible so $\mathcal{P} \subset \mathcal{A}$, and that $f \precsim_p g$ if and only if $E_p f \leq E_p g$ for a probability p. In that case, $(\mathcal{A}, \precsim)$-propriety will be a stronger condition than ordinary propriety, i.e., $(\mathcal{P}, E)$-propriety.

Propriety does not, of course, yield the strict domination results that are supposed to trouble non-probabilists. After all, the completely trivial scoring rule T such that $T(c)(\omega) = 0$ for all c and ω is proper, but gives no reason to prefer probabilities to non-probabilities. But the accuracy-arguer adds some additional conditions on s on top of propriety. For instance, they may add strict or quasi-strict propriety and continuity on the probabilities. Such conditions guarantee that for any non-probability c there is a probability p that strictly s-dominates c.

At this point, however, the probabilist offering an accuracy argument runs into a serious problem. For while our non-probabilist was liable to find propriety compelling, they only found it compelling as a special case of a stronger requirement, $(\mathcal{A}, \precsim)$-propriety. Now in order to be pragmatically plausible, the preference comparison $\precsim_r$ should satisfy strict dominance. But the following four statements are logically incompatible:

(1) $\precsim_r$ satisfies strict dominance for r in $\mathcal{A}$
(2) $\mathcal{A}$ is not a subset of $\mathcal{P}$
(3) s is $(\mathcal{A}, \precsim)$-proper
(4) for any $c \notin \mathcal{P}$ there is a $p \in \mathcal{P}$ such that $s(c) > s(p)$ everywhere.

We will call (4) "the domination thesis" from now on.

To see that the four conditions above are incompatible, note that if we choose r in $\mathcal{A} - \mathcal{P}$ (which we can by (2)), then the domination thesis (4) implies there is a p distinct from c such that $s(r) > s(p)$ everywhere, which

[4]In the case of mathematical expectation for a probability r, we have $E_r(-f) = -E_r(f)$ and so our condition is equivalent to the more familiar $E_r s(r) \leq E_c s(r)$. However, it is in general not true that $\mathrm{LSI}_r(-f) = -\mathrm{LSI}_r(f)$: see the Appendix for what is actually true.

by strict dominance (1) implies that $-s(r) \prec_r -s(p)$, and that contradicts the $(\mathcal{A}, \precsim)$-propriety condition (3).

In particular, whatever conditions the accuracy-arguer would place on s that imply the domination thesis are incompatible with our non-probabilist's requirement that s be $(\mathcal{A}, \precsim)$-proper, given that our non-probabilist takes some non-probability to be admissible and can be expected to be working with a preference comparison that satisfies strict dominance, perhaps one derived from a prevision that satisfies strong monotonicity.

In other words, the accuracy-arguer offers some set of conditions C on a scoring rule (e.g., strict propriety and continuity in the case of Pettigrew [7]) and proves that C plus propriety implies the domination thesis. If the accuracy-arguer has done their job well, C and propriety will be plausible to our non-probabilist. But if propriety will be plausible to the non-probabilist, likewise $(\mathcal{A}, \precsim)$-propriety will be plausible. Thus our non-probabilist will see C and $(\mathcal{A}, \precsim)$-propriety as reasonable constraints to put on a scoring rule. But if $\mathcal{P} \subset \mathcal{A}$ and $\precsim$ agrees with E on the probabilities, then the accuracy theorist's argument for the domination thesis from propriety and C shows that no scoring rule satisfies C and $(\mathcal{A}, \precsim)$-propriety.

At this point, one might wonder if there are any non-trivial scoring rules that satisfy $(\mathcal{A}, \precsim)$-propriety for any plausible examples of $\mathcal{A}$ and $\precsim$. The answer is positive in the case of $(\mathcal{M}, \mathrm{LSI})$. In fact any bounded scoring rule s defined only on the probabilities and proper there (where s is bounded provided there is a finite K such that $|s(p)(\omega)| < K$ for all p and ω) can be extended to a $(\mathcal{M}, \mathrm{LSI})$-proper scoring rule on all the credences. To see that, note that s's values $s(p)$ are functions from Ω to $\mathbb{R}$ and $\mathbb{R}^\Omega$ can be thought of as n-dimensional Euclidean space, where n is the cardinality $|\Omega|$ of Ω. Let V be the topological closure of the set $\{-s(p) : p \in \mathcal{P}\}$. For any fixed $u \in \mathcal{M} - \mathcal{P}$, the prevision LSI_u is known to be a continuous function from $\mathbb{R}^\Omega$ to $\mathbb{R}$??ref, and since V is compact, it attains a maximum at one or more points of V. Choose any one of these points, and call it α_u.[5] Then let $s(u) = -\alpha_u$. For any $u \in \mathcal{M} - \mathcal{P}$, the point $-s(u)$ maximizes LSI_u over V. The same can be seen to be true for $u \in \mathcal{P}$ by propriety of our original score s on $\mathcal{P}$, the fact that any point of V is a limit of a sequence of values of $-s$, and the fact that LSI_u agrees with E_u for u a probability. Then for any $u, v \in \mathcal{M}$ we have $\mathrm{LSI}_u(-s(u)) \geq \mathrm{LSI}_u(-s(v))$ because LSI_u is maximized over V at $-s(u)$ and $-s(v) \in V$. Finally, we need to define $s(c)$ where $c \in \mathcal{C} - \mathcal{M}$. The simplest solution is just to let $s(c)(\omega) = \infty$ for all ω (any point that is dominated by some point in V will also work).

The accuracy-arguer now needs to convince the non-probabilist to hold on to C and weaken $(\mathcal{A}, \precsim)$-propriety to mere propriety. The non-probabilist

[5] This can be done as a direct application of the Axiom of Choice, but we can also do it constructively. Identifying $\mathbb{R}^\Omega$ with $\mathbb{R}^n$, order it lexicographically. The set of points of where LSI_u attains its maximum is closed (since it's the pre-image of the closed set $\{\max_V \mathrm{LSI}_u\}$ under the continuous function LSI_u) and hence compact, and so it will have a lexicographically first element. Let α_u be that element.

has several options. First, they can reject C. Second, they can accept both C and $(\mathcal{A}, \precsim)$-propriety as plausible constraints on scoring rules but say that it is an unfortunate fact that no scoring rule satisfies both, and hence a scoring rules are not a good way to evaluate the accuracy of a credence assignment. Third, propriety itself could be rejected. In all cases, the argument from accuracy for probabilism will carry little weight.

I will next survey the variety of versions of C that have been offered and the reasons that can be given for them, and discuss the three options for a non-probabilist response.

4. The condition C

The condition C is added to propriety to yield the domination thesis. Several versions of C are known. The earliest known fairly general version of C was

(CASP) probability-continuity, additivity and strict propriety. [9]

It was later seen that additivity can be dropped, and two other versions were offered:

(CSP) probability-continuity and strict propriety [7, 6, 10]

(CQSP) probability-continuity and quasi-strict propriety [6].

In the same line of development, one can get a weakening of CQSP where probability-continuity is replaced by a continuity condition at the probabilities with an infinite score combined with a complicated geometric condition [13], but as no one has formulated a philosophical reason to accept the geometric condition beyond the fact that it is entailed by CQSP, we do not need to consider this version.

There is also one other fairly recent line of development. Campbell-Moore and Levinstein [1] prove that

(CASTD) probability-continuity, additivity and strict truth-directedness

together with propriety implies strict propriety, and so we can take C to be CASTD. Here, strict truth-directedness says that if c' is truer than c at ω, then $s(c')(\omega) < s(c)(\omega)$. A credence assignment c' is truer than another credence assignment c at ω provided that for every event A, if $\omega \in A$, then $c'(A) \geq c(A)$, and if $\omega \notin A$, then $c'(A) \leq c(A)$, and in at least one case the inequality is strict.

5. Rejecting C

In order for the accuracy argument to succeed against our non-probabilist, the relevant version of C must be sufficiently plausible to overcome the plausibility that $(\mathcal{A}, V)$-propriety has to our non-probabilist. But the candidates for C, while all plausible, are not *that* plausible, especially to the non-probabilist. To see this, let us consider their ingredients.

Additivity is primarily a simplifying assumption rather than a philosophically plausible constraint on what counts as accuracy. Global features of a credence function could turn out to be relevant to the function's accuracy. One might, for instance, think that there is an additional epistemic utility in having gotten *everything* right that goes beyond the value of each individual thing one got right. Moreover, one way to see the implausibility of additivity as anything beyond a simplifying assumption is to reflect on the likely reaction of an accuracy theorist who has a strong commitment to strict or at least quasi-strict propriety to the news that there is no additive quasi-strictly proper scoring rule for probabilities (whether countably or finitely additive) defined on every subset of an infinite sample space [12, Proposition 3]. The reaction is likely to be to search for non-additive quasi-strictly proper scoring rules rather than rejection of the accuracy framework on the grounds that without additivity, scoring rules are not plausible measures of accuracy. The fact that additivity does not have that much independent plausibility casts a shadow over arguments based on CASP and CASTD.

Probability-continuity has a significant degree of initial plausibility—it does seem that a slight change in credences should result only in a slight change in score. Though even this is not completely clear. For there might turn out to be probability thresholds that have significance in the case of beliefs about important matters. For instance, for propositions that are particularly epistemically central to one's view of the world, such as that life has (or lacks) meaning, that moral realism is (or is not) true, that God does (or does not) exist, or that we live (or do not live) in a simulation, if the proposition is true, there may be a discontinuous jump in epistemic utility as one goes from assigning a credence lower than 1/2 to assigning a credence greater than 1/2. Or it may be the case that there is a threshold such that one does not count as knowing when one's credence lies below that threshold, and if knowledge has a special value, then the epistemic utility of a credence in a truth may discontinuously jump as we cross that threshold. Furthermore, one might think that the epistemic utility of a credence in a falsehood may discontinuously go down when the credence hits one—it seems extra bad to be sure of a falsehood—or with Descartes one might think that there is a special value in being certain of a truth, so the epistemic utility jumps discontinuously as the credence hits one.

While *quasi-strict propriety* is logically weaker than strict propriety, there does not appear to be any reason to accept it beyond the reasons for strict propriety, unless one has a prior objection to non-probabilistic credences that the non-probabilist will take to be question-begging. For quasi-strict propriety, in the absence of strict propriety, expressly disadvantages non-probabilistic credences with respect to the scoring rule by requiring that any non-probabilistic credence have a poorer expectation than p by the lights of p for any probabilistic credence p, without requiring that a probabilistic credence other than p have such a poorer expectation. Thus, despite the logical weakening in the premises, the argument based on CQSP thus has little if any weight beyond the one based on CSP.

Strict truth-directedness has some plausibility, though it only appears in CASTD conjoined with additivity, which is not particularly plausible. Without additivity, Theorem 3 in the Appendix shows that truth-directedness and continuity do not yield the domination thesis.

And there is some reason for the proponent of probabilism to be suspicious of strict truth-directedness. Suppose that in the case of a fair coin toss, I have a probabilistic credence p that the probability of heads is 0.50 and that of tails is 0.50 as well. Next, suppose that I sustain a head injury that causes my credence for tails to shift to 0.49, everything else remaining the same, so I still assign 0.50 to heads, inconsistency notwithstanding. If it turns out that in fact the coin does land heads, it is not clear that I am better off epistemically for having shifted my credence slightly in the truth-ward direction, when I have done so at the cost of inconsistency. But since my new credences are truer, a strictly truth-directed scoring rule will give me a higher score.

Furthermore, Theorem 2 in the Appendix shows that strict truth-directedness of a scoring rule is incompatible with $(\mathcal{M}, \mathrm{LSI})$-propriety if Ω has at least two points. Thus the non-probabilist who likes $\mathcal{M}$ and LSI will have good reason to be suspicious of strict truth-directedness.

We are finally left with *strict propriety* as such. Now, this has some initial plausibility. Just as it seemed likely that one would not be required by inaccuracy minimization to change one's probabilistic credences evidencelessly, it is fairly plausible that one would not be permitted to do so, and hence if one's probability is p, then the p-expected score of a different credence should be strictly worse. ??ref

But there are several problems in this line of thought.

First, it could well be that there are many permissible ultimate priors for rational credences. On subjective Bayesianism, any coherent (and maybe regular) credence assignment can function as the ultimate priors, but one need not be a subjective Bayesian to think that there is some freedom. But if there is any freedom in the ultimate priors, then it is unclear why it would be irrational for someone to reverse-engineer their current credences and the evidence they have received back to their original priors, and then switch those original priors to some other set of permissible ultimate priors, and then re-impose the evidence on top of this, thereby changing one's credences evidencelessly. Moreover, on any view on which there are non-formal constraints on the ultimate priors, it should be not only permissible but required that if one should discover that one's ultimate priors did not satisfy these constraints, then one should backtrack and fix one's priors and readjust one's current credences.

Second, even if one grants that one would not be permitted to change credences evidencelessly, it is not clear that this prohibition would have to come from expected accuracy optimization. One can have two levels of commitment to the accuracy-theoretic framework. More weakly one could hold that it provides *a* constraint on one's rational credences, or more strongly one could hold that it accounts for *all* the constraints on one's rational credences. Only the stronger commitment to the accuracy-theoretic framework yields the argument for strict propriety. For the weaker commitment is compatible with there being a rule of rationality separate from the accuracy-theoretic framework that forbids evidenceless switches of credence away from probabilistic credences.

Third, the argument for strict propriety, i.e., strict $(\mathcal{P}, E)$-propriety, is plausible precisely because we think probabilistic credences in $\mathcal{P}$ are rationally admissible. (If we did not think some credence assignment to be rationally admissible, we should have no problem with a scoring rule permitting—or even requiring—an evidenceless change from that credence assignment.) Thus the likely principle behind the argument for strict propriety is really that it is impermissible to change *rationally admissible* credences evidencelessly. What this supports—assuming we can defend it from the two previous objections—is the thesis that

(5) $-s(c) \prec_r -s(r)$ if r is rationally admissible, c is different from r and $\precsim$ is the correct preference ordering.

Now, our non-probabilist thinks that the set of rationally admissible credences contains some non-probabilities. Given this, the argument pushes her to accept that what one might call strict $(\mathcal{A}, \precsim)$-propriety—that $-s(c) \prec_r -s(r)$ for all $r \in \mathcal{A}$ and $c \in \mathcal{C} - \{r\}$.

But remember that our non-probabilist, in order to get out of pragmatic problems, is assumed to have a preference structure $\precsim$ that satisfies strong monotonicity. If, further, $\mathcal{A} \supset \mathcal{P}$ and $\precsim$ agrees with E on the probabilities, then strict $(\mathcal{A}, \precsim)$-propriety will be impossible given probability-continuity. For strict $(\mathcal{A}, \precsim)$-propriety will imply strict $(\mathcal{P}, E)$-propriety, which together with probability-continuity will imply the domination condition which is incompatible $(\mathcal{A}, \precsim)$-propriety (given strong monotonicity of $\precsim$ for credences in $\mathcal{A}$). Thus, our non-probabilist is likely to want to reject strict $(\mathcal{A}, \precsim)$-propriety, and likewise strict propriety as it's based on the same reasoning.

Interestingly, Theorem 2 in the Appendix shows that in the special case of strict $(\mathcal{M}, \mathrm{LSI})$-propriety, the assumption of probability-continuity in this argument can be dropped. Indeed, if Ω has at least two points, no $(\mathcal{M}, \mathrm{LSI})$-proper scoring rule is quasi-strictly proper. In fact, more strongly, for any $(\mathcal{M}, \mathrm{LSI})$-proper scoring rule, there is a probability and a non-probability that get the exact same score (same everywhere on Ω), which rules out quasi-strict propriety as well as pretty much excluding the possibility for an accuracy-based argument for probabilism with a $(\mathcal{M}, \mathrm{LSI})$-proper scoring rule. Thus if our non-probabilist takes the members of $\mathcal{M}$ to be rationally admissible with LSI as the associated prevision, and hence finds $(\mathcal{M}, \mathrm{LSI})$-propriety very plausible, then they will be very suspicious of the premises for any accuracy-based argument for probabilism. (It is worth noting that in the Appendix, these observations are proved with Monotonicity replaced by the weaker Subadditivity condition.)

Thus, in fact, given the conflict between C and $(\mathcal{A}, \precsim)$-propriety, and given the weakness of the arguments for any of the versions of C, it does not appear irrational for our non-probabilist to reject all the versions of C.

Our arguments in this section may seem akin to a trivial *modus tollens* response: the conjunction of the conditions in C implies probabilism is true, so the non-probabilist concludes that this conjunction is false. However, the point is a little subtler. One of the conditions in C, namely propriety, is more plausible than the others. By our non-probabilist's lights, the argument for propriety is an argument for a stronger thesis, $(\mathcal{A}, \precsim)$-propriety, which thesis ends up being incompatible with the conjunction of the other conditions.

6. Pessimism about scoring rules

But the non-probabilist also has a different response available, which is to accept that both $(\mathcal{A}, \precsim)$-propriety and some version of C are correct conditions to impose on a reasonable scoring rule. Of course, it immediately follows from this that there is no reasonable scoring rule. Is this an unacceptable conclusion?

Recently, Pruss [12] has proved (assuming the Axiom of Choice) various negative results about scoring rules in infinite contexts. For instance, in the case of finitely additive probabilities defined on all subsets of an infinite space, there is no strictly proper scoring rule, and in the case of continuum-many coin tosses and countably additive probabilities defined on the usual product σ-algebra, there is no strictly proper scoring rule either.

Now, one plausible reaction to those negative results would be to conclude that strict propriety is an unreasonable condition on a scoring rule. That would lead to the non-probabilist's responding to the arguments for probabilism by rejecting condition C, as in the previous section.

But there is another possible reaction if we are impressed by the idea that strict propriety is a correct requirement for a scoring rule to capture the concept of accuracy. We could conclude that in the infinite contexts there is are no good scoring rules: the scoring rule approach should simply be put aside. And this has a parallel for the non-probabilist, who can say that in contexts where non-probabilistic credences are appropriate, there are no good scoring rules. Scoring rules *should* satisfy C and be $(\mathcal{A}, \precsim)$-proper, but because no scoring rule does that, scoring rules are not a good tool for analyzing credences in contexts where non-probabilistic credences are an option.[6]

[6] It is worth noting that in the infinite case, there is a third solution. Pruss [12] shows that it is possible to construct strictly proper scoring rules in the infinite contexts if instead of requiring the values of the scores to be extended real numbers, we allow scores to take values in some larger set such as nets of real numbers (at the same time, he notes that it is not known at present whether an argument for probabilism can be run in that context). This way out of the negative results does not appear to have a parallel in our finite-space non-probabilist context.

7. Further constraints to the non-probabilist's rescue

So far, we have assumed that the set $\mathcal{A}$ of rationally admissible credences contains all the probabilities. But this might be incorrect. First, some formal epistemologists accept regularity as a further constraint on rationality: all non-empty events must have non-zero probability. If so, then non-regular probabilities, ones that assign zero probability to some non-empty event, will not be rationally admissible and hence will not be members of $\mathcal{A}$. In that case, $(\mathcal{A}, \precsim)$-propriety will not imply propriety, even if $\precsim$ agrees with mathematical expectation E on the probabilities. And the usual arguments for propriety are implicitly or explicitly based on the assumption that all probabilities are rationally admissible. If some credence is not rationally admissible, it is not particularly surprising if by its own lights it is recommended that one change to a different credence. Granted, there is a price to be paid for requiring regularity even in the finite case: updates will have to be something like Jeffrey conditionalization to maintain regularity.??ref

Moreover, regularity need not be the only additional constraint on rationally admissibility. It could turn out that there are also non-formal constraints based on the subject matter. Objective Bayesians think there are non-formal constraints on the priors. Now consider this story. Original priors are rationally required to be regular and friendly to induction (e.g., elegant regularities should not have astronomically low probabilities), but at the same time rationality requires classical Bayesian conditionalization rather than Jeffrey conditionalization. In that case, a probability assignment that is regular *and* not friendly to induction is one that a perfectly rational agent could *never* have. For our perfectly rational agent's original priors would be regular but friendly to induction. And our perfectly rational agent's posteriors would never be regular, since regularity is lost in classical Bayesian conditionalization—the complement of one's evidence will come to have credence zero. So a regular probability assignment not friendly to probabilism would never be rationally inadmissible.

As soon as any probability function is omitted from the set of rationally admissible credences, the philosophical arguments for strict or even non-strict $(\mathcal{P}, E)$-propriety fail, since the arguments are only plausible for probability functions that are rationally admissible. All we will have will be arguments for strict or non-strict $(\mathcal{P}_0, E)$-propriety, where $\mathcal{P}_0$ is the subset of probabilities that are rationally admissible. And that won't be good enough for the domination thesis.

Indeed, even $(\mathcal{P}, E)$-propriety is not enough for the domination thesis if all we have is strict $(\mathcal{P}_0, E)$-propriety, with $\mathcal{P}_0$ a proper subset of $\mathcal{P}$. For let s be any probability-continuous strictly proper score. Fix a probability $q \in \mathcal{P} - \mathcal{P}_0$. Define $s'(p) = s(p)$ for any probability p and $s'(c) = s(q)$ for any non-probability c. Then s' is a probability-continuous score which fails in strict propriety only at q, which is not rationally admissible. But s' does not have the domination property. For $s'(c)$ for a non-probability

c is not dominated by the score of any probability, since $s'(c) = s(q)$, and no s-score of a probability by a proper scoring rule s is dominated by the s-score of any probability (if $s(p_1)$ were dominated by $s(p_2)$, we would have $E_{p_1} s(p_1) > E_{p_1} s(p_2)$, contrary to propriety), so $s'(c) = s(q)$ is not dominated by the s- or, equivalently, s'-score of any probability. Hence s' does not satisfy the domination thesis, despite having probability-continuity, $(\mathcal{P}, E)$-propriety and strict $(\mathcal{P}_0, E)$-propriety.

It is thus crucial to the accuracy arguments for probabilism that all probabilities be rationally admissible, and while previously we had our non-probabilist grant that, it need not in fact be granted.

8. Conclusions

We have imagined a non-probabilist who thinks that rationally admissible credences include some class of non-probabilities and who has a decision procedure based on these credences that is helpful in practical cases. For example, the non-probabilities held to be rationally admissible could be ones satisfying some formal axioms weaker than the Kolmogorov axioms, say Zero, Normalization and Monotonicity, in which case the decision procedure could be based on level set integrals [11]. Propriety is defined with respect to the probabilities: no score of a *probability* p is beaten by the lights of p by the score of any other credence. And accuracy-based arguments for probabilism all presuppose propriety. But the reason why propriety appears plausible is because a typically more general thesis appears plausible: no score of a *rationally admissible* credence r is beaten by the lights of r by the score of any other credence. Given that as our non-probabilist thinks that some non-probabilities are rationally admissible, if they find the considerations behind propriety compelling, they will accept the extension of propriety to their preferred class of credences. But now if our non-probabilist's decision procedure is strongly monotonic, as it needs to be to escape pragmatic arguments for probabilism (and as it will be in our example case of an advocate of $\mathcal{M}$ and LSI), and assuming that our non-probabilist accepts all probabilities are rationally admissible, then the extended propriety thesis ends up being logically incompatible with the rest of the premises of the accuracy theorist's argument for probability (e.g., strict propriety and probability-continuity).

At this point, three ways were seen for the non-probabilist to continue the discussion: deny one or more of the premises incompatible with the non-probabilistically extended propriety thesis, most likely strict propriety; grant that all the conditions the probabilist wants to put on a scoring rule are correct, but conclude that there is no such thing as a good scoring rule; or insist that not all probabilities are rationally admissible.

In all of the above, it was assumed that the non-probabilist positively thinks that some non-probabilities are rationally admissible. However, it is worth noting that the accuracy-theoretic arguments may still have significant force against someone who is merely agnostic about whether any non-probabilities are rationally admissible. Such a theorist might find it plausible to think that a good scoring rule will satisfy strict propriety and continuity with respect to the rationally admissible credences, and may be confident that all probabilities are rationally admissible, while being agnostic on whether any non-probabilities are. In that case, learning that strict propriety and continuity cannot hold with regard to a strict superset of the probabilities (given strong monotonicity of the decision procedure) will give the theorist reason to think that *only* the probabilities are rationally admissible. In fact, it is interesting to note that if the class of potentially rationally admissible credences includes $\mathcal{M}$ and the decision procedure is based on LSI, then one can even drop the assumption of continuity from the argument in light of the fact proved in the Appendix that there is no strictly $(\mathcal{M}, \mathrm{LSI})$-proper scoring rule.

Similarly, a non-probabilist who is inclined to think that some non-probabilities are rationally admissible but is significantly more strongly committed to the rational admissibility of the probabilities may find the accuracy-theoretic arguments to carry some weight.

Thus, the accuracy-based arguments for probabilism carry some weight, but are very far indeed from significantly disturbing a committed non-probabilist who already knows how to respond to pragmatic arguments.

Appendix: Some technical results

Level set integrals and monotonic credences. Given a credence c, let $c^*(A) = 1 - c(\Omega - A)$ (for a probability p we have $p^* = p$). Then given Zero and Normalization, we have $\mathrm{LSI}_c(-f) = -\mathrm{LSI}_{c^*} f$ when f is a function that takes values either in $(-\infty, \infty]$ or in $[-\infty, \infty)$. We only need to check this for f having finite values. Moreover, because $\mathrm{LSI}_c(\alpha + f) = \mathrm{LSI}_c f$ (by the well-definition of $\mathrm{LSI}_c f$), we may suppose f takes values in $[0, L]$ for finite L and then:

$$
\begin{aligned}
\mathrm{LSI}_c f &= \int_0^L c(\{\omega : f(\omega) > y\})\, dy \\
&= \int_0^L (1 - c^*(\{\omega : -f(\omega) \geq -y\}))\, dy \\
&= L - \int_0^L c^*(\{\omega : -f(\omega) \geq -y\})\, dy \\
&= L - \int_0^L c^*(\{\omega : -f(\omega) > -y\})\, dy \\
&= L - \int_0^L c^*(\{\omega : -f(\omega) > t - L\})\, dt \\
&= L - \int_0^\infty c^*(\{\omega : L - f(\omega) > t\})\, dt \\
&= -\mathrm{LSI}_{c^*}(-f),
\end{aligned}
$$

where the first and last equalities used Zero and Normalization (applied to c) respectively, and the move from considering the level set $\{\omega : -f(\omega) \geq -y\}$ to considering the level set $\{\omega : -f(\omega) > -y\}$ depended on the fact that the two level sets are equal except perhaps when y is one of the finitely many values of f.

In [11, Lemma 1g], it is incorrectly stated (with some trivial translation to our setting) that if f is negative and finite, then $\mathrm{LSI}_c f = -\mathrm{LSI}_c(-f)$. The right hand side should instead be $-\mathrm{LSI}_{c^*}(-f)$. The only place where [11] uses the incorrect/ claim appears to be in the proof of his Theorem 1 in the case of what is called there $\mathrm{LSI}_P^{\uparrow}$, where it is shown that decisions using level set integrals avoid Dutch Books. To fix the problem, in the statement of the theorem in the case of $\mathrm{LSI}_P^{\uparrow}$ one needs to replace Non-Negativity with the axiom that credences have value at most 1, and instead of the argument given in the proof, use the formula $\mathrm{LSI}_P^{\uparrow} f = -\mathrm{LSI}_{P^*}^{\uparrow}(-f)$ to establish that $\mathrm{LSI}_P^{\uparrow} f < 0$ if $f < 0$ everywhere, noting that P^* satisfies Non-Negativity if $P \leq 1$ everywhere.

Theorem 1. *If $c \in \mathcal{M}$ and f and g are functions on Ω with values in $[-\infty, \infty]$ such that $f < g$ everywhere, then $\mathrm{LSI} f < \mathrm{LSI} g$, with both level set integrals well-defined.*

Proof. Since $f < g$ everywhere, g cannot take the value $-\infty$ anywhere and f cannot take the value $+\infty$ anywhere. Let $M_0 = 1 + \max(\max f, \max(-g))$. This is finite, and if $M \geq M_0$, then

$$f_M \leq f_{M_0} < g_{M_0} \leq g_M$$

everywhere. By [11, Theorem 2], we then have

$$\mathrm{LSI}_c f_M \leq \mathrm{LSI}_c f_{M_0} < \mathrm{LSI}_c g_{M_0} \leq \mathrm{LSI}_c g_M.$$

Taking the limit as $M \to \infty$, we conclude that $\mathrm{LSI}_c f < \mathrm{LSI}_c g$. □

Say that a credence c satisfies Subadditivity provided that $c(A) + c(B) \leq c(A \cup B)$ whenever A and B are disjoint. Given that our credences take values in $[0, 1]$, Subadditivity implies Zero and Monotonicity. Let $\mathcal{S}$ be the credences that satisfy Normalization and Subadditivity.

Say that a scoring rule s is probability-distinguishing provided that if p and c are credences in its domain with p a probability and c not a probability, then $s(p)$ cannot equal $s(c)$ everywhere on Ω. Being probability-distinguishing is a necessary condition for quasi-strict propriety as well as for the domination thesis (4), and indeed is plausibly a necessary condition for any accuracy argument for probabilism based on s.

Theorem 2. *Let s be a $(\mathcal{S}, \mathrm{LSI})$-proper scoring rule defined for all credences in $\mathcal{S}$. Then s is neither probability-distinguishing, nor quasi-strictly proper, nor strictly truth-directed.*

Write $v \cdot w$ for the dot product of two vectors.

Lemma 1. *Let $V \subseteq (-\infty, M]^2$ be a non-empty closed convex set. Then there is a point z of V and a vector $v = (v_1, v_2)$ in the positive quadrant $(0,\infty)^2$ such that for all $w \in V - \{z\}$ we have $v \cdot w < v \cdot z$.*

Proof. Let $v(\theta) = (\cos\theta, \sin\theta)$ be the unit vector at angle θ. For $\theta \in (0, \pi/2)$, the set of points z of V that maximize $v(\theta)\cdot z$ is a finite (perhaps degenerate) line segment which is a subset of a line normal to $v(\theta)$ (finitude follows from the fact that any line segment contained in $(-\infty, M]^2$ and orthogonal to a vector in the positive quadrant is finite). Let $a(\theta) = (a_1(\theta), a_2(\theta))$ and $b(\theta) = (b_1(\theta), b_2(\theta))$ be the endpoints of that line segment, chosen so that $a_2(\theta) \le b_2(\theta)$.

I claim that if $0 < \theta < \phi < \pi/2$, then $b_2(\theta) \le a_2(\phi)$.

For let $z = (z_1, z_2) = b(\theta)$ and $w = w(w_1, w_2) = a(\phi)$. Then:

$$w_1 \cos\theta + w_2 \sin\theta = v(\theta) \cdot w \le v(\theta) \cdot z = z_1 \cos\theta + z_2 \sin\theta$$

and so

$$w_1 \cos\theta \cos\phi \le z_1 \cos\theta \cos\phi + z_2 \sin\theta \cos\phi - w_2 \sin\theta \cos\phi$$

Also:

$$z_1 \cos\phi + z_2 \sin\phi = v(\phi) \cdot z \le v(\phi) \cdot w = w_1 \cos\phi + w_2 \sin\phi$$

and so

$$z_1 \cos\theta \cos\phi + z_2 \cos\theta \sin\phi - w_2 \cos\theta \sin\phi \le w_1 \cos\theta \cos\phi.$$

Putting our two inequalities together:

$$\begin{aligned} z_1 \cos\theta \cos\phi + z_2 \cos\theta \sin\phi - w_2 \cos\theta \sin\phi \\ \le z_1 \cos\theta \cos\phi + z_2 \sin\theta \cos\phi - w_2 \sin\theta \cos\phi \end{aligned}$$

and so:

$$z_2(\cos\theta \sin\phi - \sin\theta \cos\phi) \le w_2(\cos\theta \sin\phi - \sin\theta \cos\phi).$$

Since $\cos\theta \sin\phi - \sin\theta \cos\phi = \sin(\phi - \theta) > 0$, we have $z_2 \le w_2$, which is what we wanted.

So if $0 < \theta < \phi < \pi/2$, then $b_2(\theta) \le a_2(\phi)$ as desired. Thus for any sequence of angles $\pi/4 = \theta_0 < \cdots < \theta_n = \pi/2$, we have:

$$\begin{aligned} b_2(\theta_n) &= b_2(\theta_0) + \sum_{i=1}^{n} (b_2(\theta_i) - a_2(\theta_i) + a_2(\theta_i) - b_2(\theta_{i-1})) \\ &\ge b_2(\theta_0) + \sum_{i=1}^{n} (b_2(\theta_i) - a_2(\theta_i)). \end{aligned}$$

Hence:

$$b_2(\theta_n) \ge \sum_{\theta \in (\pi/4, \pi/2)} (b_2(\theta) - a_2(\theta)).$$

Since $b_2(\theta) - a_2(\theta) \ge 0$, and the only way an uncountable sum of non-negative values can converge if all but countably many summands are zero, it follows that for uncountably many $\theta \in (\pi/4, \pi/2)$ we have $b_2(\theta) = a_2(\theta)$.

Choose one such θ. Then $b(\theta)$ and $a(\theta)$ have the same second coordinate but both lie on a line at angle $\theta + \pi/2$. Since that line is not horizontal, the only way they can have the same second coordinate is if $b(\theta) = a(\theta)$. Letting $v = v(\theta)$, it follows that only one point $z = a(\theta) = b(\theta)$ maximizes $v \cdot z$ over V, as desired. □

If $\Omega = \{1, 2\}$ and $\alpha, \beta \in [0, 1]$, let $r_{\alpha,\beta}$ be the credence in $\mathcal{M}$ such that $r_{\alpha,\beta}(\{1\}) = \alpha$ and $r_{\alpha,\beta}(\{1\}) = \beta$. This is a member of $\mathcal{S}$ if and only if $\alpha + \beta \leq 1$ and a member of $\mathcal{P}$ if and only if $\alpha + \beta = 1$. It will be useful to note that

$$\mathrm{LSI}_{r_{\alpha,\beta}} f = \begin{cases} f(1) + \beta(f(2) - f(1)) & \text{if } f(1) < f(2) \\ f(2) + \alpha(f(1) - f(2)) & \text{otherwise.} \end{cases}$$

Lemma 2. *Let $\Omega = \{1, 2\}$. Suppose that $s : \mathcal{S} \to [-\infty, M]$ is such that $\mathrm{LSI}_r\, s(r) \geq \mathrm{LSI}_r\, s(u)$ for any $r, u \in \mathcal{S}$. Then there are $\alpha, \beta, \gamma \in (0, 1)$ such that $\gamma \leq \alpha$ and $\beta \leq 1 - \alpha$ with exactly one of the two inequalities strict and $s(r_{\alpha,1-\alpha}) = s(r_{\gamma,\beta})$ everywhere.*

Proof. By abuse of notation, we will identify $\mathbb{R}^\Omega$ with $\mathbb{R}^2$ in the natural way (a function $f \in \mathbb{R}^\Omega$ corresponds to the point $(f(1), f(2))$ in $\mathbb{R}^2$).

Let U be the set of finite values of s and let V be the closed convex hull of U.

By Lemma 1, there is a point $z = (z_1, z_2)$ of V and a vector $v = (v_1, v_2)$ in the positive quadrant $(0, \infty)^2$ such that for all $w \in V - \{z\}$ we have $v \cdot w < v \cdot z$.

Without loss of generality we can suppose $z_1 \geq z_2$ for the rest of our proof. For if we can prove the result in that case, then if for any z chosen as above we had $z_1 < z_2$, we could have swapped the labels on the two elements of Ω and applied the case we are about to prove.

Given the choice of z, let p be the probability such that $p(\{i\}) = v_i/(v_1 + v_2)$ for $i = 1, 2$. Then $z = s(p)$. For every $w \in U$ is such that

$$E_p w = \mathrm{LSI}_p\, w \leq \mathrm{LSI}_p\, s(p) = E_p s(p)$$

and hence $v \cdot w \leq v \cdot s(p)$. It follows this is true for every $w \in V$, and hence $v \cdot z \leq v \cdot s(p)$ which by choice of z can only be true if $s(p) = z$.

Let $\alpha = p(\{1\})$ and choose any $\beta \in (0, 1 - \alpha)$. Then $s(p)(2) = z_2 \geq z_1 = s(p)(1)$, and so:

$$\mathrm{LSI}_{r_{\alpha,\beta}}\, s(p) = s(p)(2) + \alpha(s(p)(1) - s(p)(2)) = \mathrm{LSI}_p\, s(p).$$

Now, if $s(r_{\alpha,\beta})(1) \geq s(r_{\alpha,\beta})(2)$, then:

$$\mathrm{LSI}_{r_{\alpha,\beta}}\, s(r_{\alpha,\beta}) = s(r_{\alpha,\beta})(2) + \alpha(s(r_{\alpha,\beta})(1) - s(r_{\alpha,\beta})(2)) = \mathrm{LSI}_p\, s(r_{\alpha,\beta}).$$

On the other hand if $s(r_{\alpha,\beta})(1) < s(r_{\alpha,\beta})(2)$, then

$$\begin{aligned} \mathrm{LSI}_{r_{\alpha,\beta}}\, s(r_{\alpha,\beta}) &= s(r_{\alpha,\beta})(1) + \beta(s(r_{\alpha,\beta})(2) - s(r_{\alpha,\beta})(1)) \\ &< s(r_{\alpha,\beta})(1) + (1 - \alpha)(s(r_{\alpha,\beta})(2) - s(r_{\alpha,\beta})(1)) = \mathrm{LSI}_p\, s(r_{\alpha,\beta}). \end{aligned}$$

So in either case:

$$\mathrm{LSI}_p\, s(r_{\alpha,\beta}) \geq \mathrm{LSI}_{r_{\alpha,\beta}}\, s(r_{\alpha,\beta}) \geq LSI_{r_{\alpha,\beta}} s(p) = \mathrm{LSI}_p\, s(p) \geq \mathrm{LSI}_p\, s(r_{\alpha,\beta}).$$

It follows that $\mathrm{LSI}_p\, s(r_{\alpha,\beta}) = \mathrm{LSI}_p\, s(p)$. But LSI_p is the same as E_p for p a probability. Considering $s(r_{\alpha,\beta})$ and $s(p)$ as vectors in $\mathbb{R}^2$, it follows that $v \cdot s(r_{\alpha,\beta}) = v \cdot s(p)$. Thus by the choice of z, we have $s(p) = z = s(r_{\alpha,\beta})$ with $\alpha + \beta < 1$. And $p = r_{\alpha,1-\alpha}$. Letting $\gamma = \alpha$ we are done. □

Proof of Theorem 2. That s is not quasi-strictly proper follows from its not being probability-distinguishing, so all we need to show is that s is neither probability-distinguishing nor strictly truth-directed.

Let $\Omega_2 = \{\omega_1, \omega_2\}$ and let $\mathcal{S}_2$ be the credences on Ω_2 that satisfy Normalization and Subadditivity. For a credence r on Ω_2 and $A \subseteq \Omega$, let $\bar{r}(A) = r(A \cap \Omega_2)$. It is easy to check that if $r \in \mathcal{S}_2$, then $\bar{r} \in \mathcal{S}$. Define the scoring rule $s_2(r)$ for $r \in \mathcal{S}_2$ by $s_2(r)(\omega) = s(\bar{r})(\omega)$ for $\omega \in \Omega_2$. Then $\mathrm{LSI}_r(-s_2(v)) = \mathrm{LSI}_{\bar{r}}(-s(\bar{v}))$ for any r and v in $\mathcal{S}$, so s_2 is $(\mathcal{S}_2, \mathrm{LSI})$-proper if s is $(\mathcal{S}, \mathrm{LSI})$-proper. It is also easy to see that if s_2 fails to be probability-distinguishing or strictly truth-directed, so does s.

So we can assume without loss of generality that $\Omega = \{1, 2\}$. Let $s'(u) = -s(u)$. Then for any $r, u \in \mathcal{S}$, we have

$$\mathrm{LSI}_r\, s'(r) = \mathrm{LSI}_r(-s(r)) \geq \mathrm{LSI}_r(-s(u)) = \mathrm{LSI}_r\, s'(u)$$

by $(\mathcal{S}, \mathrm{LSI})$-propriety. Choose α, β, γ as in Lemma 2 applied to s'. Then $s(r_{\alpha,1-\alpha}) = s(r_{\gamma,\beta})$ everywhere and so s is not probability-distinguishing. If $\gamma < \alpha$ and $\beta = 1 - \alpha$, then $r_{\alpha,1-\alpha}$ is truer than $r_{\gamma,\beta}$ at 1, and if $\gamma = \alpha$ and $\beta < 1 - \alpha$, then $r_{\alpha,1-\alpha}$ is truer than $r_{\gamma,\beta}$ at 2. Thus s cannot be strictly truth-directed. □

The credences $\mathcal{C}$ are the space of functions from the powerset of Ω to $[0, 1]$ and can be equipped in the natural way with $2^{|\Omega|}$-dimensional Euclidean topology. This agrees with the topology on $\mathcal{P} \subset \mathcal{C}$ that was used to define probability-continuity.

Say that a scoring rule s is probability-distinguishing provided that if $p \in \mathcal{P}$ and $c \in \mathcal{C} - \mathcal{P}$, then $s(p)(\omega) \neq s(c)(\omega)$ for some ω. If a scoring rule is proper but not probability distinguishing, then it cannot be quasi-strictly proper and also it cannot satisfy the domination thesis (4). To see the latter point, observe that no score of a probability can be s-dominated by the score of a probability given propriety, since if p were s-dominated by q, then $E_p s(p) > E_p s(q)$, contrary to propriety. So if the score of a non-probability c equaled that of a probability, we wouldn't have the domination thesis for c.

Theorem 3. *Let s be any proper truth-directed scoring rule defined on the probabilities $\mathcal{P}$ on Ω where $|\Omega| = 2$. Then s can be extended to a truth-directed, proper but not probability-distinguishing scoring rule defined on all of $\mathcal{C}$. Furthermore, the extension can be taken to be a continuous function from $\mathcal{C}$ to $[M, \infty]$ if s is probability-continuous.*

Proof. Without loss of generality $\Omega = \{1,2\}$. Let p_α be the probability such that $p_\alpha(\{1\}) = \alpha$. Note that p_α is truer than p_β at 1 if and only if $\alpha > \beta$ and at 2 if and only if $\alpha < \beta$.

Let $\alpha(c) = 1/2 + (c(\{1\}) - c(\{2\}))/2$ for any credence c. Now define

$$s'(c) = s(p_{\alpha(c)}) + c(\varnothing) + 1 - c(\Omega)$$

for $c \in \mathcal{C}$. Note that this agrees with the original definition on $\mathcal{P}$, since if c is a probability, $\alpha(c) = c(\{1\})$. For simplicity, write s in place of s'.

We now need to show that s thus extended is truth-directed, proper but not quasi-strictly proper.

Propriety is easy. Let p be any probability and c any credence. If c is a probability, we have $E_p s(p) \le E_p s(c)$ by propriety restricted to the probabilities. If c is not a probability, we have $E_p s(p) \le E_p s(p_{\alpha(c)}) \le E_p s(c)$, since $s(c) \ge s(p_{\alpha(c)})$ everywhere.

Lack of probability distinguishing follows from the fact that if c satisfies Zero and Normalization but is not in $\mathcal{P}$, then $s(c) = s(p_{\alpha(c)})$ everywhere.

We now prove truth-directedness. All we need to prove is that if c is truer than d at 1, then $s(c)(1) < s(d)(1)$; the case where c is truer than d at 2 is essentially the same. Furthermore, by forming a chain of credences between c and d that differ on only one set, we need to prove that $s(c)(1) < s(d)(1)$ in each of the following cases:

(i) c and d agree on all events except $\varnothing$, where $c(\varnothing) < d(\varnothing)$
(ii) c and d agree on all events except Ω, where $c(\Omega) > d(\Omega)$
(iii) c and d agree on all events except $\{1\}$, where $c(\{1\}) > d(\{1\})$
(iv) c and d agree on all events except $\{2\}$, where $c(\{2\}) < d(\{2\})$

The inequality $s(c)(1) < s(d)(1)$ is obvious in cases (i) and (ii).

Now suppose we have case (iii) or (iv). In both cases we have $\alpha(c) > \alpha(d)$. Then $p_{\alpha(c)}$ is truer at 1 than $p_{\alpha(d)}$, and so by truth-directedness of s on $\mathcal{P}$ we have $s(c)(1) = s(p_{\alpha(c)}) < s(p_{\alpha(d)}) = s(d)(1)$.

Finally, the continuity claim is clear from our definition of the extension s. □

References

[1] Catrina Campbell-Moore and Benjamin A. Levinstein. 2021. "Strict Propriety is Weak", *Analysis* 81:8–13.
[2] A. P. Dawid, and M. Musio. 2014. "Theory and applications of proper scoring rules", *Metron* 72:169–183.
[3] T. Gneiting, and A. E. Raftery. 2007. "Strictly proper scoring rules, prediction, and estimation", *Journal of the American Statistical Association* 102:359–378.
[4] H. Greaves and D. Wallace. 2006. "Justifying Conditionalization: Conditionalization Maximizes Expected Epistemic Utility", *Mind* 115:607–632.
[5] J. M. Joyce. 1998. "A Nonpragmatic Vindication of Probabilism", *Philosophy of Science* 65:575–603.
[6] Michael Nielsen. Forthcoming. "On the Best Accuracy Arguments for Probabilism", *Philosophy of Science.*

[7] Richard Pettigrew. 2021. "Accuracy-First Epistemology Without the Additivity Axiom", *Philosophy of Science* 89:128–151.

[8] Richard Pettigrew. 2016. *Accuracy and the Laws of Credence*. Oxford: Oxford University Press.

[9] Joel B. Predd, Robert Seiringer, Elliott H. Lieb, Daniel N. Osherson, H. Vincent Poor, and Sanjeev R. Kulkarni. 2009. "Probabilistic Coherence and Proper Scoring Rules", *IEEE Transactions on Information Theory* 55:4786–4792.

[10] Alexander R. Pruss. 2021. "Proper Scoring Rules and Domination". `https://arxiv.org/abs/2102.02260`

[11] Alexander R. Pruss. 2021. "Avoiding Dutch Books Despite Inconsistent Credences". *Synthese* 198:11265–11289.

[12] Alexander R. Pruss. 2022. "Accuracy, Probabilism and Bayesian Update in Infinite Domains". `http://philsci-archive.pitt.edu/21169/`

[13] Alexander R. Pruss. 2021. "Necessary and Sufficient Conditions for Domination Results for Proper Scoring Rules". `https://arxiv.org/abs/2103.00085`

[14] R. L. Winkler, J. Munoz, J. L. Cervera, J. M. Bernardo, G. Blattenberger, J. B. Kadane, D. V. Lindley, A. H. Murphy, R. M. Oliver, and D. Ríos-Insua. 1996. "Scoring rules and the evaluation of probabilities", *Test* 5:1—60.