STRICT PROPRIETY IS WEAK

CATRIN CAMPBELL-MOORE AND BENJAMIN A. LEVINSTEIN

Abstract. Considerations of accuracy—the epistemic good of having credences close to truth-values—have led to the justification of a host of epistemic norms. These arguments rely on particular ways of measuring accuracy. In particular, the accuracy measure should be strictly proper. However, the main argument for strict propriety only supports weak propriety. But, strict propriety follows from weak propriety given strict truth-directedness (which is non-negotiable) and additivity (which is both very common and plausible). So no further argument is necessary.

1. Our Main Point

Considerations of accuracy—the epistemic good of having credences close to truth-values—have led to the justification of a host of epistemic norms, including probabilism (Joyce, 1998, 2009), conditionalization (Greaves and Wallace, 2006; Briggs and Pettigrew, 2018), the Principal Principle (Pettigrew, 2016), and the Principle of Indifference (Pettigrew, 2014).

For these arguments to succeed, accuracy-theorists need to tell us how to measure the accuracy of credences at a world, where accuracy is a function just of credences and truth-values. I.e.:[1]

$$\mathsf{Acc} : \mathsf{Credences} \times \mathsf{Worlds} \to \mathbb{R}$$

Instead of directly arguing for a particular measure, accuracy theorists typically argue for constraints on what counts as a legitimate measure of accuracy.

One fundamental constraint simply encodes the idea that the higher your credences in truths, and the lower your credences in falsehoods, the more accurate your credences are:

Strict Truth Directedness.
If $c(\phi) \leq b(\phi)$ for all ϕ true at w, and $b(\phi) \geq c(\phi)$ for all ϕ false at w, with at least one strict inequality, then $\mathsf{Acc}(c, w) < \mathsf{Acc}(b, w)$.

This says that when a credence function c is modified by increasing the credence assigned to some propositions that are true in w, or reducing the credence to some propositions false in w, $\mathsf{Acc}(c, w)$ is strictly increased.

A further constraint, which is still very plausible, says that the accuracy of your entire credence function at a world is just the sum of the accuracy of each of your credences in individual propositions at that world.[2]

We are very grateful to Richard Pettigrew and Jason Konek for helpful feedback on a draft. Also thanks to Julia Staffel, Kenny Easwaran, Pavel Janda and two anonymous referees. Catrin Campbell-Moore's research was supported by a Leverhulme Trust Research Fellowship.

[1] We assume a fixed, finite collection of propositions $\mathcal{F}$ containing at least one contingent proposition. The credences are functions $b, c : \mathcal{F} \to [0, 1]$. One could also allow accuracy values of $-\infty$, see Remark 3.

[2] For a defence, see Pettigrew (2016, §4.1).

Additivity.
There is $\mathsf{a} : [0,1] \times \{\mathsf{t}, \mathsf{f}\} \longrightarrow \mathbb{R}$, with $\mathsf{Acc}(c, w) = \sum_{\phi} \mathsf{a}(c(\phi), w(\phi))$.[3]

The constraint of present focus, which is vital for the validity of the accuracy-theorists' results, is strict propriety, which says that a probabilistic credence function expects itself to be more accurate than any other credence function.[4]

Strict propriety.
If b is probabilistic then $\mathrm{Exp}_b \mathsf{Acc}(c)$ is *uniquely* maximised at $c = b$. ($\mathrm{Exp}_b \mathsf{Acc}(c)$ denotes the expected accuracy of credence function c according to b.[5])

This needs much more justification than our previous constraints. The main style of argument for strict propriety runs something like the following (Joyce, 2009: 277-9; Gibbard, 2007; Oddie 1997: 537-8; Greaves and Wallace, 2006: 621): Firstly, rational agents should be immodest, i.e. should expect their own credences to be doing best at the pursuit of accuracy. After all, if your credence function expects something else to be doing better, then your own credence function advises you (even in the absence of further evidence) to adopt an alternative credence function instead; so, you can't rationally rely on it. Secondly, all probability functions could be rational given the right circumstances.[6] This then, it is claimed, justifies the principle of strict propriety.

However, as has been noticed by Pettigrew (2011, §4.1) and Mayo-Wilson and Wheeler (2016: 65),[7] this argument only supports *weak* propriety:

Weak propriety.
If b is probabilistic then $\mathrm{Exp}_b \mathsf{Acc}(c)$ is maximised (possibly non-uniquely) at $c = b$.

So long as your credences are amongst the best, they are not self-undermining. If a measure of accuracy mandated that you change your credence to a different one regardless of what evidence you possessed, then it seemingly causes needless and capricious doxastic changes and forbids perfectly rational views. But merely thinking that some other credences are equally as good does not lead to this irrationality. Moreover, these authors claim that this argument cannot naturally be extended to obtain strict propriety.[8]

But, it turns out that strict propriety is a mathematical consequence of weak propriety given our two additional principles of strict truth directedness and additivity. Merely weakly

---

[3] $w(\phi) = \begin{cases} \mathsf{t} & \phi \text{ is true in } w \\ \mathsf{f} & \phi \text{ is false in } w \end{cases}$. The sum is taken over $\phi \in \mathcal{F}$.

[4] Joyce's (2009) dominance argument for probabilism uses a weaker principle: Coherent Admissibility. But there are similar difficulties motivating the strong version of the principle that is needed for his result. Others (e.g. Pettigrew, 2016; Joyce, 1998) give an alternative collection of axioms which entail strict propriety, but the argument for them is much less direct.

[5] $\mathrm{Exp}_b \mathsf{Acc}(c) := \sum_w b(w) \times \mathsf{Acc}(c, w)$. If $\{w\} \notin \mathcal{F}$ and $b(w)$ is ill-defined, then we define strict propriety as $\sum_w b^*(w) \times \mathsf{Acc}(c, w)$ is maximised at $c = b$ for any b^* probabilistic on $\mathcal{F}^*$, an extension of $\mathcal{F}$ which includes such $\{w\}$ (see Pettigrew, 2016, §2.2). For additive Acc, the choice of b^* doesn't matter, so we can still legitimately refer to $\mathrm{Exp}_b \mathsf{Acc}(c)$.

[6] More carefully, for every probability space $\langle \Omega, \mathcal{F}, \mathrm{Pr} \rangle$, there is an isomorphic space $\langle \Omega', \mathcal{F}', \mathrm{Pr}' \rangle$ such that Pr' is rationally permissible (in the right circumstances). This formulation is then compatible with further restrictions to probabilism. See Joyce (2009, Footnote 17).

[7] Note that Mayo-Wilson and Wheeler (2016) is interested in the imprecise setting where our argument doesn't apply. Fallis (2007: 221) appeals to a similar argument to endorse weak but not strict propriety.

[8] Joyce (2009) in fact argues that, firstly, credences that could be rationally *required* expect themselves to be strictly optimal, and secondly, any probability function could be rationally required. This stronger 'secondly' premiss is less plausible, and arguments that the 'firstly' premiss is only plausible with a *weakly* optimal consequent might similarly hold for Joyce's version. Furthermore, by our argument here, it is sufficient to argue for weak propriety.

proper additive accuracy measures will thus fail to be strictly truth directed; in fact they will be constant on some region.[9]

So, no further argument for strict propriety is necessary. One just needs to argue for weak propriety, and strict propriety comes for free.

2. The result

The first part of our argument works just with a. We define:

$$\mathrm{Exp}_x\, \mathsf{a}(y) := x \times \mathsf{a}(y, \mathsf{t}) + (1 - x) \times \mathsf{a}(y, \mathsf{f})$$

which is the expected accuracy of adopting credence y in ϕ, as evaluated according to a probability function b with $b(\phi) = x$. We define truth directedness and propriety as expected for a.

Proposition 1. *Suppose* a *is weakly proper and strictly truth directed. Then it is strictly proper.*

This result is an immediate consequence of Schervish's theorem on representations of scoring rules (1989, Theorem 4.2) (as Schervish et al. (2009, Lemma 5) note), but we can also show it directly.

Proof. We need to show that $\mathrm{Exp}_x\mathsf{a}(x) > \mathrm{Exp}_x\mathsf{a}(y)$ for $y \neq x$. Consider some z which lies between y and x. We will show that $\mathrm{Exp}_x\mathsf{a}(z) > \mathrm{Exp}_x\mathsf{a}(y)$; that is, that $\mathrm{Exp}_x\mathsf{a}(y)$ strictly improves as we move y towards x. To do this, we compare $\mathrm{Exp}_x\mathsf{a}(z) - \mathrm{Exp}_x\mathsf{a}(y)$ to $\mathrm{Exp}_z\mathsf{a}(z) - \mathrm{Exp}_z\mathsf{a}(y)$, which we know is ≥ 0 by weak propriety at z.

Consider first the case where $x > z > y$.

$$\mathrm{Exp}_x\mathsf{a}(z) - \mathrm{Exp}_x\mathsf{a}(y) = \underset{\vee}{x} \times \overbrace{(\mathsf{a}(z, \mathsf{t}) - \mathsf{a}(y, \mathsf{t}))}^{>0} + \underset{\wedge}{(1 - x)} \times \overbrace{(\mathsf{a}(z, \mathsf{f}) - \mathsf{a}(y, \mathsf{f}))}^{<0} \tag{1}$$

$$\mathrm{Exp}_z\mathsf{a}(z) - \mathrm{Exp}_z\mathsf{a}(y) = z \times (\mathsf{a}(z, \mathsf{t}) - \mathsf{a}(y, \mathsf{t})) + (1 - z) \times (\mathsf{a}(z, \mathsf{f}) - \mathsf{a}(y, \mathsf{f})) \tag{2}$$

Since a is strictly truth directed and $z > y$, $\mathsf{a}(z, \mathsf{t}) - \mathsf{a}(y, \mathsf{t}) > 0$ and $\mathsf{a}(z, \mathsf{f}) - \mathsf{a}(y, \mathsf{f}) < 0$. Since $x > z$, and thus $(1 - x) < (1 - z)$, in equation 1 there is more weight than in equation 2 on something positive, and less on something negative. Therefore

$$\mathrm{Exp}_x\mathsf{a}(z) - \mathrm{Exp}_x\mathsf{a}(y) > \mathrm{Exp}_z\mathsf{a}(z) - \mathrm{Exp}_z\mathsf{a}(y).$$

And since $\mathrm{Exp}_z\mathsf{a}(z) \geq \mathrm{Exp}_z\mathsf{a}(y)$, $\mathrm{Exp}_x\mathsf{a}(z) > \mathrm{Exp}_x\mathsf{a}(y)$.

For $x < z < y$, an analogous argument shows that $\mathrm{Exp}_x\mathsf{a}(z) > \mathrm{Exp}_x\mathsf{a}(y)$. (The inequalities annotating equations 1 and 2 are reversed.)

We have thus shown that $\mathrm{Exp}_x\mathsf{a}(z) > \mathrm{Exp}_x\mathsf{a}(y)$. By weak propriety, $\mathrm{Exp}_x\mathsf{a}(x) \geq \mathrm{Exp}_x\mathsf{a}(z)$ and thus $\mathrm{Exp}_x\mathsf{a}(x) > \mathrm{Exp}_x\mathsf{a}(y)$, as required. □

Additivity then allows us to turn this into a result about accuracy of an entire credal state.

Corollary 2. *Suppose* Acc *is weakly proper, additive and strictly truth directed. Then it is strictly proper.*

[9] This follows from proposition 1 using the fact that weakly proper a are weakly truth directed (Schervish, 1989, Lemma 1). Schervish, Seidenfeld, and Kadane (2009, Lemma 5) also noted this as a consequence of Schervish's representation result.

Proof. We first note that for additive Acc with associated a, if b is probabilistic then $\mathrm{Exp}_b\mathsf{Acc}(c) = \sum_\phi \mathrm{Exp}_{b(\phi)}\mathsf{a}(c(\phi))$. This allows us to observe that if a is strictly proper, so is Acc (as each $\mathrm{Exp}_{b(\phi)}\mathsf{a}(b(\phi)) \geq \mathrm{Exp}_{b(\phi)}\mathsf{a}(c(\phi))$, with some $>$).

One then needs to show that a is weakly proper and strictly truth directed. Consider some b probabilistic, with $b(\psi) = x$, and let c differ from it just by having $c(\psi) = y$. By using the aforementioned fact, the only difference between $\mathrm{Exp}_b\mathsf{Acc}(b)$ and $\mathrm{Exp}_b\mathsf{Acc}(c)$ is $\mathrm{Exp}_x\mathsf{a}(x)$ and $\mathrm{Exp}_x\mathsf{a}(y)$. Since $\mathrm{Exp}_b\mathsf{Acc}(b) \geq \mathrm{Exp}_b\mathsf{Acc}(c)$, also $\mathrm{Exp}_x\mathsf{a}(x) \geq \mathrm{Exp}_x\mathsf{a}(y)$; i.e. a is weakly proper. For truth directedness, observe that the only difference between $\mathsf{Acc}(c, w)$ and $\mathsf{Acc}(b, w)$ is in $\mathsf{a}(x, w(\psi))$ and $\mathsf{a}(y, w(\psi))$. A truth directed move of y to x is similarly a truth directed move of c to b, so $\mathsf{Acc}(b, w) > \mathsf{Acc}(c, w)$ and thus $\mathsf{a}(x, w(\psi)) > \mathsf{a}(y, w(\psi))$. By noting that we can consider either $w(\psi) = \mathsf{t}$ or $w(\psi) = \mathsf{f}$, we see a is strictly truth directed. □

Remark 3. One can also allow $-\infty$ in the range of the accuracy measure as long as we make the assumption that if $\mathsf{Acc}(c, w) = -\infty$, then there is some ϕ such that $w(\phi) = \mathsf{t}$ and $c(\phi) = 0$ or $w(\phi) = \mathsf{f}$ and $c(\phi) = 1$. The equivalent assumption for a says that $-\infty$ can only appear at $\mathsf{a}(0, \mathsf{t})$ or $\mathsf{a}(1, \mathsf{f})$. The definitions of strict truth directedness need to be weakened to allow that $\mathsf{Acc}(b, w) = \mathsf{Acc}(c, w) = -\infty$ even when b is a truth directed improvement on c.[10] And we should take $0 \times -\infty := 0$ in the expected accuracy formulas.

To check that Proposition 1 still holds, with the infinity assumption on a, one should observe that the only infinite value we could have in equations 1 and 2 is $\mathsf{a}(y, \mathsf{t}) = -\infty$, in which case $\mathrm{Exp}_x\mathsf{a}(y) = -\infty < \mathrm{Exp}_x\mathsf{a}(z)$.

For Corollary 2, one should also check that the infinity assumption on Acc entails that on a. To do this, consider c with $c(\psi) = x$ and $c(\phi) = 0.5$ for $\phi \neq \psi$. If $\mathsf{a}(x, \mathsf{t}) = -\infty$, then for w with $w(\psi) = \mathsf{t}$, $\mathsf{Acc}(c, w) = -\infty$, so $x = 0$; and similarly for f. To show that $\mathrm{Exp}_b\mathsf{Acc}(c) = \sum_\phi \mathrm{Exp}_{b(\phi)}\mathsf{a}(c(\phi))$ still holds, note that for probabilistic b, $b(\phi) = 0$ iff $b(w) = 0$ for all $w(\phi) = \mathsf{t}$ and $b(\phi) = 1$ iff $b(w) = 0$ for all $w(\phi) = \mathsf{f}$. Then for the weak and strict propriety relationships, note that $\mathrm{Exp}_x\mathsf{a}(x)$ is always finite. For strict truth directedness, pick b and c, which just differ on ψ, to have $b(\phi) = c(\phi) = 0.5$ for all $\phi \neq \psi$ (they do not need to be probabilistic).

References

Briggs, R. and R. Pettigrew (2018). An accuracy-dominance argument for conditionalization. *Noûs*, Online First.

Fallis, D. (2007). Attitudes toward epistemic risk and the value of experiments. *Studia Logica 2*, 215–246.

Gibbard, A. (2007). Rational credence and the value of truth. In T. Gendler and J. Hawthorne (Eds.), *Oxford Studies in Epistemology*, Volume 2, pp. 143–164. Oxford University Press.

Greaves, H. and D. Wallace (2006). Justifying conditionalization: Conditionalization maximizes expected epistemic utility. *Mind 115*(459), 607–632.

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

Joyce, J. M. (2009). Accuracy and coherence: Prospects for an alethic epistemology of partial belief. In F. Huber and C. Schmidt-Petri (Eds.), *Degrees of Belief*, Volume 342, pp. 263–297. Springer.

[10]The unweakened version of strict truth directedness for a is equivalent to the weakened one plus the infinity assumption; but not so for Acc.

Mayo-Wilson, C. and G. Wheeler (2016). Scoring imprecise credences: A mildly immodest proposal. *Philosophy and Phenomenological Research 93*(1), 55–78.

Oddie, G. (1997). Conditionalization, cogency, and cognitive value. *British Journal for the Philosophy of Science 48*(4), 533–541.

Pettigrew, R. (2011). An improper introduction to epistemic utility theory. In R. de Henk, S. Hartmann, and S. Okasha (Eds.), *EPSA Philosophy of Science: Amsterdam 2009*, pp. 287–301. Springer.

Pettigrew, R. (2014). Accuracy, risk, and the principle of indifference. *Philosophy and Phenomenological Research 91*(1), 35–59.

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

Schervish, M. (1989). A general method for comparing probability assessors. *The Annals of Statistics 17*, 1856–1879.

Schervish, M. J., T. Seidenfeld, and J. B. Kadane (2009). Proper scoring rules, dominated forecasts, and coherence. *Decision Analysis 6*(4), 199–278.