Joyce's Argument for Probabilism*

Patrick Maher†
Department of Philosophy
University of Illinois at Urbana-Champaign

James Joyce's 'Nonpragmatic Vindication of Probabilism' gives a new argument for the conclusion that a person's credences ought to satisfy the laws of probability. The premises of Joyce's argument include six axioms about what counts as an adequate measure of the distance of a credence function from the truth. This paper shows that (a) Joyce's argument for one of these axioms is invalid, (b) his argument for another axiom has a false premise, (c) neither axiom is plausible, and (d) without these implausible axioms Joyce's vindication of probabilism fails.

1. Introduction. James Joyce (1998) offered a new argument for the conclusion that a person's credences ought to satisfy the laws of probability. The argument assumes that it makes sense to talk about the distance that a person's credence function is from the truth in any possible world. Joyce also states six axioms that, he says, any adequate measure of this distance must satisfy. Joyce's "Main Theorem" is that these axioms imply that if a person's credence function b does not satisfy the laws of probability then there exists another credence function b^* which does satisfy the laws of probability and which is closer to the truth in every possible world. Joyce assumes that an epistemically rational person strives to have credences that are as close to the truth as possible and so concludes that an epistemically rational person's credences will satisfy the laws of probability.

Joyce's Main Theorem assumes that the set of propositions is countable. However, the case in which the set of propositions is countably infinite raises mathematical issues that I judge not worth pursuing here, so in this article I will consider only the special case in which the set of propositions

*Received February 2001; revised October 2001.

†Send requests for reprints to the author, Department of Philosophy, 105 Gregory Hall, 810 South Wright Street, Urbana, IL 61801; e-mail: p-maher@uiuc.edu.

Philosophy of Science, 69 (March 2002) pp. 73–81. 0031-8248/2002/6901-0003$10.00
Copyright 2002 by the Philosophy of Science Association. All rights reserved.

https://doi.org/10.1086/338941 Published online by Cambridge University Press

is finite. In that special case, at least, Joyce's Main Theorem is correct.[1] But there remains the question of whether his axioms are true.

In Section 3 I show that Joyce's argument for one of his axioms is invalid. In Section 4 I show that his argument for another axiom rests on a false premise. In Section 5 I argue that both these axioms are implausible. In Section 6 I show that without the implausible axioms Joyce's vindication of probabilism fails.

2. Concepts and Notation. I will use the following concepts and notation. Except for 'n', '$\mathcal{X} \to \mathcal{Y}$', and '$\mathcal{I}$', this notation is Joyce's.

Ω: A finite algebra of propositions. (Joyce allows Ω to be countable but I am here considering only the special case in which it is finite.)
n: The number of elements in Ω.
$\Re$: The set of real numbers.
$\mathcal{X} \to \mathcal{Y}$: The set of all functions from $\mathcal{X}$ to $\mathcal{Y}$.
B: The set $\Omega \to \Re$. Joyce refers to elements of B as "credence functions."
V: The set of functions in $\Omega \to \{0,1\}$ that correspond to consistent truth value assignments to the elements of Ω (with 1 representing truth and 0 falsity). Elements of V can be thought of as possible worlds. Note that $V \subset B$.
$\mathcal{I}$: The set of all $I \in B \times V \to \Re$ which are such that, for any $b \in B$ and $\omega \in V$, $I(b, \omega)$ is a rationally permissible measure of the inaccuracy of b in the possible world ω.

In the preceding definitions I said B and V are subsets of $\Omega \to \Re$. Here I was following what Joyce seems to be saying in the following passage:

> B is the family of all credence functions defined on a countable Boolean algebra of propositions Ω and V is the subset of B containing all consistent truth-value assignments to members of Ω. (590f.)

However, there are other places (583 for example) where Joyce says or implies that B and V are subsets of $\Re^n$.

Let $X_1, \ldots, X_n$ be some enumeration of the elements of Ω. Then we can identify any $b \in \Omega \to \Re$ with the point

$$(b(X_1), \ldots, b(X_n)) \in \Re^n.$$

Although Joyce does not mention it, I take it that he is assuming such an identification of $\Omega \to \Re$ with $\Re^n$. For simplicity, and to make my treatment

1. There are a few small slips in Joyce's proof but they can be corrected. For example, Joyce's proposition II (598) does not follow from the axiom Structure that he cites but it does follow from Normality. Also, his Lemma-3 (599) is false if $c = m$, as is possible, but it is easy to show that the Main Theorem does hold if $c = m$.

https://doi.org/10.1086/338941 Published online by Cambridge University Press

correspond to Joyce's, I will likewise assume this identification in what follows.

3. Weak Convexity. The axiom that Joyce (596) calls "Weak Convexity" may be stated as follows:

> **Weak Convexity:** If $b, b^* \in B$, $\omega \in V$, $I \in \mathcal{I}$, $I(b,\omega) = I(b^*, \omega)$, and $m = \frac{1}{2}b + \frac{1}{2}b^*$ then $I(b, \omega) \geq I(m, \omega)$ with identity only if $b = b^*$.

Here is Joyce's argument for Weak Convexity quoted in full (with two typographic errors corrected):

> To see why Weak Convexity is a reasonable constraint on gradational inaccuracy notice that in moving from b to m an agent would alter each degree of belief $b(X)$ by adding an increment of $k(X) = \frac{1}{2}[b^*(X) - b(X)]$. She would add the same increment of $k(X)$ to each $m(X)$ in moving from m to b^*. To put it in geometric terms, the "vector" k that she must add to b to get m is the same as the vector she must add to m to get b^*. Furthermore, since $b^* = b + 2k$ the change in belief involved in going from b to b^* has the *same direction* but a *doubly greater magnitude* than the change involved in going from b to m. This means that the former change is more *extreme* than the latter in the sense that, for *every* proposition X, both changes alter the agent's degree of belief for X in the same direction, either by moving it closer to one or closer to zero, but the b to b^* change will always move $b(X)$ twice as far as the b to m change moves it. Weak Convexity is motivated by the intuition that extremism in the pursuit of accuracy is no virtue. It says that if a certain change in a person's degrees of belief does not improve accuracy then a more radical change in the same direction and of the same magnitude should not improve accuracy either. Indeed, this is just what the principle says. If it did not hold, one could have absurdities like this: "I raised my confidence levels in X and Y and my beliefs became less accurate overall, so I raised my confidence levels in X and Y again, by exactly the same amounts, and the initial accuracy was restored." (596f.)

Joyce is here claiming that Weak Convexity follows from a premise that could be stated formally thus:

> **Premise 1:** If $b,k \in B$, $\omega \in V$, $I \in \mathcal{I}$, and $I(b + k, \omega) \geq I(b, \omega)$ then $I(b + 2k, \omega) \geq I(b, \omega)$.

Define $\rho \in B \times \Omega \rightarrow \Re$ by the condition

https://doi.org/10.1086/338941 Published online by Cambridge University Press

$$\rho(b,\omega) = \sum_{X\in\Omega} |b(X) - \omega(X)|.$$

A proof of the following theorem is given in Section 8.1.

Theorem 1: The assumption that $\rho \in \mathcal{I}$ is consistent with Premise 1 but not with Weak Convexity.

This theorem shows that Premise 1 does not entail Weak Convexity and so Joyce's argument for Weak Convexity is invalid.

4. Symmetry. The axiom that Joyce (596) calls "Symmetry" may be stated as follows:

Symmetry: If $I \in \mathcal{I}$ and $I(b, \omega) = I(b^*, \omega)$ then, for any $\lambda \in [0, 1]$,
$I(\lambda b + (1 - \lambda)b^*,\omega) = I((1 - \lambda)b + \lambda b^*, \omega)$.

Joyce offers a "rationale" or argument for Symmetry. As he presents it, this argument depends on Weak Convexity and we have just seen that Joyce has failed to justify that. However, Joyce's essential point here seems not to depend on Weak Convexity, so I shall try to extract this essential point from his presentation.

Speaking of a particular kind of violation of Symmetry, Joyce writes: "Given the initial symmetry of the situation this would amount to an unmotivated bias in favor of one set of beliefs or the other." (597)
This suggests that Joyce believes:

Premise 2: If $I(b, \omega) = I(b^*, \omega)$, $\lambda \in [0, 1]$, and
$I(\lambda b + (1 - \lambda)b^*, \omega) > I((1 - \lambda)b + \lambda b^*, \omega)$
then I has an unmotivated bias in favor
of $\lambda b + (1 - \lambda)b^*$ as against $(1 - \lambda)b + \lambda b^*$.

I suppose that Joyce also believes:

Premise 3: If I has an unmotivated bias then $I \notin \mathcal{I}$.

Premises 2 and 3 entail Symmetry, so I will take this to be Joyce's argument for Symmetry. The argument is valid but I will now argue that Premise 2 is false.

Consider the following example. (Here each $b \in B$ will be regarded as a point in $\Re^n$ with its ith coordinate denoted b_i.)

$$\begin{aligned}
\omega &= (1, 0, \omega_3, \ldots, \omega_n)\\
b &= (0, 0, \omega_3, \ldots, \omega_n)\\
b^* &= (3/2, 1/2, \omega_3, \ldots, \omega_n)\\
\lambda &= 2/3\\
c &= \lambda b + (1 - \lambda)b^* = (1/2, 1/6, \omega_3, \ldots, \omega_n)\\
c^* &= (1 - \lambda)b + \lambda b^* = (1, 1/3, \omega_3, \ldots, \omega_n)
\end{aligned}$$

https://doi.org/10.1086/338941 Published online by Cambridge University Press

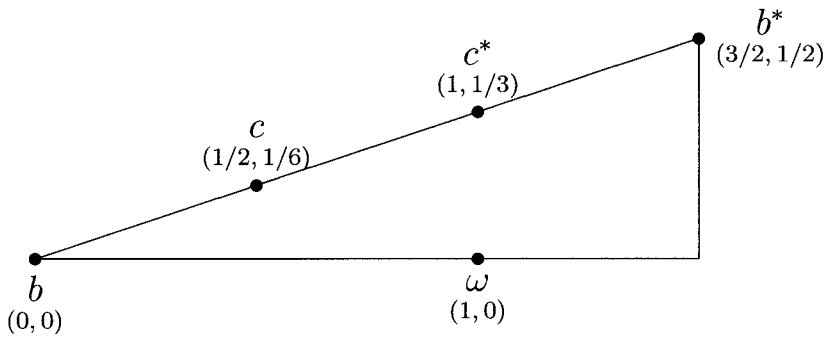


Figure 1. Counterexample to Premise 2.

Figure 1 shows the situation in the plane of the first two coordinates. Since

$$\rho(b, \omega) = \sum_{X \in \Omega} |b(X) - \omega(X)| = \sum_{i=1}^{n} |b_i - \omega_i|$$

we have

$$\rho(b, \omega) = \rho(b^*, \omega) = 1$$

but

$$\rho(c, \omega) = 2/3 > 1/3 = \rho(c^*, \omega).$$

According to Premise 2, it follows that ρ has an unmotivated bias in favor of c as against c^*. But the definition of ρ shows no bias; ρ treats all belief functions in the same way. Furthermore, the inequality $\rho(c, \omega) > \rho(c^*, \omega)$ can be "motivated" as follows: Moving from b to c improves accuracy with respect to X_1 by 1/2 and moving from b^* to c^* does the same. On the other hand, moving from b to c *reduces* accuracy with respect to X_2 by 1/6, while moving from b^* to c^* *increases* accuracy with respect to X_2 by the same amount.

Hence Premise 2 is false. Although $\rho(b, \omega) = \rho(b^*, \omega)$ and $\rho(c, \omega) > \rho(c^*, \omega)$, ρ does not have an unmotivated bias in favor of c as against c^*.

5. These Axioms Not Plausible. We have seen that Joyce's arguments for Weak Convexity and Symmetry are both unsound. It also seems to me that neither of these axioms is self evident. I will now give a reason for thinking that in fact these axioms are false.

It is natural to measure the inaccuracy of b with respect to the proposition X in possible world ω by $|b(X) - \omega(X)|$. It is also natural to take the total inaccuracy of b to be the sum of its inaccuracies with respect to each proposition. But these two assumptions together imply that the overall

https://doi.org/10.1086/338941 Published online by Cambridge University Press

inaccuracy of b in ω is $\rho(b, \omega)$. Hence, it is natural to use $\rho(b, \omega)$ as a measure of the inaccuracy of credence function b in the possible world ω.

Furthermore, there are other contexts in which measures analogous to ρ seem the most natural ones to use. For example, imagine that a student takes three examinations, with the maximum possible score for each examination being 100. Suppose the student scores 92 on the first examination, 98 on the second, and 97 on the third. I now ask: How far short has the student fallen from the goal of a perfect score on all three examinations? I think the most natural answer to this question is $8 + 2 + 3 = 13$ points. But that answer assumes a measure with the same form as ρ.

These considerations make it plausible that $\rho \in \mathcal{I}$. But if $\rho \in \mathcal{I}$ then, by Theorem 1, it follows that Weak Convexity is false. The example used in Section 4 shows that it also follows that Symmetry is false. I conclude that, in the absence of any cogent argument for them, Weak Convexity and Symmetry are both implausible.

6. Other Axioms Not Sufficient. I have argued against two of Joyce's axioms but have not criticized his other four axioms. In this section I will consider what can be inferred from these other axioms alone.

Joyce's first four axioms may be stated as follows. (I have altered Joyce's formulations of these axioms to make them clearer and more precise.)

> **Structure:** If $I \in \mathcal{I}$ and $\omega \in V$ then
> a. $I(b, \omega)$ is a non-negative continuous function of b.
> b. For all $X \in \Omega$ and $\varepsilon > 0$ there exists $\delta > 0$ such that, for all $b \in B$, if $|b(X)| > \delta$ then $I(b, \omega) > \varepsilon$.
>
> **Extensionality:** $\mathcal{I}$ is a subset of $B \times V \to \Re$.
>
> **Dominance:** If $I \in \mathcal{I}$ and $b(X) = b^*(X)$ for every $X \in \Omega$ other than Y then $I(b, \omega) > I(b^*, \omega)$ iff $|b(Y) - \omega(Y)| > |b^*(Y) - \omega(Y)|$.
>
> **Normality:** If $I \in \mathcal{I}$ and $|b(X) - \omega(X)| = |b^*(X) - \omega^*(X)|$ for all $X \in \Omega$ then $I(b, \omega) = I(b^*, \omega^*)$.

It is easy to prove:

> **Theorem 2:** The assumption that $\rho \in \mathcal{I}$ is consistent with Structure, Extensionality, Dominance, and Normality.

Hence these four axioms cannot be used to argue against the assumption that $\rho \in \mathcal{I}$.

https://doi.org/10.1086/338941 Published online by Cambridge University Press

As I mentioned in the Introduction, Joyce's strategy for justifying probabilism is to show that if a credence function b does not satisfy the axioms of probability then there exists another credence function b^* that does satisfy the axioms of probability and is closer to the truth in every possible world. On the other hand, we have the following theorem (proved in Section 8.2):

> **Theorem 3:** If Ω contains at least one contingent proposition then there exists $b \in B$ which does not satisfy the axioms of probability and is such that, for all $b^* \in B$ that do satisfy the axioms of probability, there exists $\omega \in V$ for which $\rho(b, \omega) \leq \rho(b^*, \omega)$.

So if $\rho \in \mathcal{I}$ then Joyce's vindication of probabilism fails. But Theorem 2 says that the assumption that $\rho \in \mathcal{I}$ is consistent with Joyce's first four axioms. Hence, without the implausible axioms of Weak Convexity and Symmetry, Joyce's vindication of probabilism fails.

7. The Norm of Gradational Accuracy. Besides the six axioms, Joyce's argument for probabilism assumes that an epistemically rational person strives to have credences that are as close to the truth as possible. Joyce (579) calls this "The Norm of Gradational Accuracy (NGA)." I do not think that NGA is obviously true. Our concern with truth might more plausibly be understood as a concern to accept true propositions and not false ones, rather than to have credences that are close to the truth according to some favored measure of distance. Credences help determine what propositions it is rational to accept and this gives a connection between credences and truth. (See Maher 1993, ch. 6 for an elaboration of this approach.) I see no need to attach an additional epistemic value to having credences close to the truth; an advantage of not doing so is that we then do not need to make unjustifiable decisions about how to measure this distance. So NGA seems to me more dubious than the laws of probability that Joyce is attempting to justify.

However, the main point of this paper has been to show that, even if we grant Joyce NGA, his vindication of probabilism is still unsuccessful because two of his axioms are unjustified and implausible.

8. Proofs.

8.1. Proof of Theorem 1. Thinking of b, k, and ω as points in $\Re^n$, let their ith coordinates be denoted b_i, k_i, and ω_i respectively. Then we have:

https://doi.org/10.1086/338941 Published online by Cambridge University Press

$$\begin{aligned}\rho(b + k, \omega) &= \sum_{i=1}^{n} |b_i + k_i - \omega_i| \\ &= \sum_{i=1}^{n} |\tfrac{1}{2}(b_i - \omega_i) + \tfrac{1}{2}(b_i + 2k_i - \omega_i)| \\ &\leqslant \sum_{i=1}^{n} (\tfrac{1}{2}|b_i - \omega_i| + \tfrac{1}{2}|b_i + 2k_i - \omega_i|) \\ &= \tfrac{1}{2}\rho(b, \omega) + \tfrac{1}{2}\rho(b + 2k, \omega).\end{aligned}$$

Hence

$$\begin{aligned}\rho(b + 2k, \omega) &\geqslant 2\rho(b + k, \omega) - \rho(b, \omega) \\ &= \rho(b + k, \omega) + [\rho(b + k, \omega) - \rho(b, \omega)] \\ &\geqslant \rho(b, \omega), \text{ if } \rho(b + k, \omega) \geqslant \rho(b, \omega).\end{aligned}$$

Thus the assumption that $\rho \in \mathcal{I}$ is consistent with Premise 1.

Now consider the case in which, for some ω,

$$\begin{aligned}b &= (\omega_1 + 1, \omega_2, \omega_3, \ldots, \omega_n) \\ b^* &= (\omega_1, \omega_2 + 1, \omega_3, \ldots, \omega_n).\end{aligned}$$

Then $\rho(b,\omega) = 1 = \rho(b^*,\omega)$. Also, $m = (\omega_1 + \frac{1}{2}, \omega_2 + \frac{1}{2}, \omega_3, \ldots, \omega_n)$ and so $\rho(m,\omega) = 1 = \rho(b,\omega)$. Since $b \neq b^*$, this shows that it is not consistent with Weak Convexity to have $\rho \in \mathcal{I}$.

8.2. Proof of Theorem 3. For each $X \in \Omega$ let n_X denote the number of $\omega \in V$ for which $\omega(X) = 1$. Define $b \in B$ by the condition that, for all $X \in \Omega$,

$$b(X) = \begin{cases}1 \text{ if } n_x > n/2 \\ 0 \text{ if } n_x \leqslant n/2\end{cases}.$$

For each $\omega \in V$ let X_ω be the conjunction of all $Y \in \Omega$ for which $\omega(Y) = 1$. Then $\omega(X_\omega) = 1$ and $\omega'(X_\omega) = 0$ if $\omega' \neq \omega$; hence $n_{X\omega} = 1$ for all $\omega \in V$. Since there is at least one contingent proposition in Ω, $n \geq 2$ and so $n_{X\omega} \leq n/2$. Hence $b(X_\omega) = 0$ for all $\omega \in V$. But $\{X_\omega: \omega \in V\}$ is a partition of Ω, so b does not satisfy the laws of probability.

For any $X \in \Omega$,

$$\begin{aligned}\sum_{\omega \in V} |b(X) - \omega(X)| &= n_x|b(X) - 1| + (n - n_x)|b(X)| \\ &= n_x (1 - b(X)) + (n - n_x)b(X), \\ &\quad \text{since } 0 \leqslant b(X) \leqslant 1 \\ &= n_x + (n - 2n_x)b(X) \\ &= \begin{cases}n - n_x \text{ if } n_x > n/2 \\ n_x \quad\quad \text{if } n_x \leqslant n/2\end{cases}. \qquad (1)\end{aligned}$$

https://doi.org/10.1086/338941 Published online by Cambridge University Press

Let b^* be any element of B that satisfies the laws of probability. Then we likewise have

$$\sum_{\omega \in V} |b^*(X) - \omega(X)| = n_x + (n - 2n_x)b^*(X).$$

Since $0 \leq b^*(X) \leq 1$, it follows that

$$\sum_{\omega \in V} |b^*(X) - \omega(X)| \geq \begin{Bmatrix} n - n_X & \text{if } n_X > n/2 \\ n_X & \text{if } n_X \leq n/2 \end{Bmatrix}. \tag{2}$$

Comparing (1) and (2) we see that, regardless of the value of n_X,

$$\sum_{\omega \in V} |b(X) - \omega(X)| \leq \sum_{\omega \in V} |b^*(X) - \omega(X)|.$$

Since this holds for all $X \in \Omega$, we have

$$\sum_{X \in \Omega} \sum_{\omega \in V} |b(X) - \omega(X)| \leq \sum_{X \in \Omega} \sum_{\omega \in V} |b^*(X) - \omega(X)|.$$

Reversing the order of the summations and applying the definition of ρ then gives:

$$\sum_{\omega \in V} \rho(b,\omega) \leq \sum_{\omega \in V} \rho(b^*, \omega).$$

Hence $\rho(b, \omega) \leq \rho(b^*,\omega)$ for at least one $\omega \in V$.

REFERENCES

Joyce, James M. (1998), "A Nonpragmatic Vindication of Probabilism", *Philosophy of Science* 65: 575–603.

Maher, Patrick (1993), *Betting on Theories.* Cambridge: Cambridge University Press.

https://doi.org/10.1086/338941 Published online by Cambridge University Press