Collectivist Foundations for Bayesian Statistics

Conor Mayo-Wilson and Aditya Saraf

Abstract

What (if anything) justifies the use of Bayesian statistics in science? The traditional answer is that Bayesian statistics is simply an instance of orthodox expected utility theory. Thus, Bayesian statistical methods, like principles of utility theory, are justified by norms of individual rationality. In particular, most Bayesians argue that a scientist's credences must satisfy the probability axioms if she adheres to norms of practical and epistemic (individual) rationality. We argue that, to justify Bayesian statistics as a tool for science, it is necessary that a scientist's *public credences* (i.e., her degrees of belief *qua* scientist) obey the probability axioms. We claim that norms of collective science help justify this restricted view, termed *public probabilism*.

What (if anything) justifies the use of Bayesian statistics in science? The traditional answer is that Bayesian statistics is simply an instance of orthodox expected utility theory, and so Bayesian statistical methods, like principles of utility theory, are justified by norms of *individual rationality*. For example, one who has credences that violate the probability axioms is Dutch-Bookable, accuracy dominated, and so on. And if one is Dutch-Bookable or accuracy dominated, so the standard arguments go, then one violates some norm of practical or epistemic (individual) rationality.

In this paper, we explore an alternative defense of scientists' use of Bayesian statistical methods: norms of *collective science* may work in tandem with norms of individual rationality to justify the mathematical machinery of Bayesian statistics. We focus on *probabilism*, the thesis that experimenters' credences must obey the probability axioms.

The paper contains two major sections. In §1, we criticize (versions of) three of the most common arguments for probabilism: Dutch Book arguments, accuracy dominance arguments, and representation-theorem based arguments.[1] We argue that none of the three justifies the claim that ra-

[1] For a summary of Dutch book arguments, see [Vineberg, 2011] and references therein. For accuracy dominance arguments, see [Joyce, 1998] and [Pettigrew, 2016]. For use of representation theorems, see [Savage, 1972] and [Krantz et al., 2006].

tional credences obey an Archimedean axiom. The first two merely assume credences are real numbers and so must have that property. Proponents of the third argument do try to justify the Archimedean axiom but they often appeal to implausible behaviorist assumptions to do so.

In §2, we argue that, to justify the use of Bayesian statistics in science, a scientist's prior distribution cannot be interpreted as an her personal credences. Instead, a prior distribution must represent what we call the scientist's *public* credences, which are the beliefs that she adopts in virtue of her professional and social obligations as a scientist. We then argue that a norm of collective science, which we call *epistemic cooperativeness*, can justify the claim that a scientist's public credence should satisfy the Archimedean axiom. Our argument provides some evidence for the claim that, together, norms of individual rationality and collective science may justify *public probabilism*, the thesis that scientists' public credences should obey the probability axioms.

Should one conclude that Robinson Crusoe may rationally violate probabilism but that scientists must heed Bayesian norms? Our arguments leave open that possibility. But our arguments are better understood as opening a different avenue for defending the claim that scientists' behavior should abide by decision-theoretic norms. To be clear, our arguments do *not* establish that the traditional, individualistic strategy for defending probabilism is hopeless; we argue only that some common individualistic arguments share a weakness. We also do not definitively establish that scientists' public credences should abide by the probability axioms. After all, we focus on only one axiom that is necessary for a probabilistic representation of credence.

The paper, therefore, should be understood as an attempted proof of concept: we aim to show that collective norms for scientific inquiry might better justify what are often taken as norms for rational, individual behavior.

1 Individualistic Foundations of Bayesianism

The fundamental thesis of Bayesianism is *probabilism*, which asserts that, given an algebra of events $\mathcal{A}$, one's credences over $\mathcal{A}$ ought to be representable by a probability measure $P : \mathcal{A} \to [0, 1]$.[2] What does it mean for

[2] In this paper, we define probability measures as functions of *sets* in an algebra. Some computer scientists and philosophers prefer to define probability measures as functions from sets of *sentences* in some formal language to real numbers. The difference between the two axiomatizations has philosophical importance but only in ways tangential to our goals in this paper. Additionally, we assume only that the probability axioms require finite (rather than σ/countable) additivity, though we do not rule out the possibility

credences to be representable by a probability measure? We first distinguish theories of credence along two dimensions: (1) absolute vs. comparative and (2) mental vs. behavioral. These distinctions will be critical for our argument in the next section that rational comparative credence might violate the Archimedean axiom.

1.1 What is credence?

Some philosophers and social scientists argue an agent's credence $c(A)$ in event A is what we call *absolute*: $c(A)$ is a property of the agent and a single event. In contrast, some decision theorists assume that humans and perhaps other agents possess only *comparative* credences among sets of events. An agent's comparative credences are typically represented by a binary relation $\succeq$ on a space of events, where $A \succeq B$ represents the claim that the agent believes A to be at least as likely as B.[3] Under the absolute interpretation, one's credal function c is said to be representable by a probability measure if c is itself a probability measure. Under the comparative interpretation, one's credences are representable by a probability measure just in case there is a probability function P such that $A \succeq B$ if and only if $P(A) \geq P(B)$.[4]

Absolute and comparative accounts of credence can, in turn, be either behaviorist or mentalist.[5] For example, a behaviorist might *identify* an agent's absolute credence $c(A)$ with the agent's fair price $Pr(A)$ on A, i.e., the price she would be willing both to buy and sell a gamble that pays precisely \$1 if A occurs and is worthless otherwise. Alternatively, a behaviorist might claim $A \prec B$ precisely if, when forced to choose between (i) receiving a prize if A occurs or (ii) receiving the same prize when B occurs, the agent would prefer the latter option. This comparative, behaviorist account of credence, therefore, identifies credence with the outcomes of *forced choices*.

that collective norms might justify countable additivity in ways that axioms of individual rationality might not.

[3] Historically, binary representations are the most common way of representing comparative credence. However, it is well-known that some aspects of human perception are best represented by higher arity relations. For instance, see [Krantz et al., 2006, p. 139]'s discussion of difference structures. Further, choice behavior is often represented by *choice functions* [Seidenfeld et al., 2009, Sen, 1971], which take sets of arbitrary size and output the options that are "choice-worthy." So it is not beyond the realm of possibility that credence should similarly be modeled by relations of higher arity or some mathematical apparatus like choice functions. We focus on binary relations for simplicity.

[4] What we call representability is often called "agreement" in decision theory. Some decision theorists aim only for what is called "partial" or "almost" agreement between $\preceq$ and a probability measure. See [Fishburn, 1986] for survey.

[5] For an enlightening discussion of behaviorist vs. mentalist views, see [Okasha, 2016].

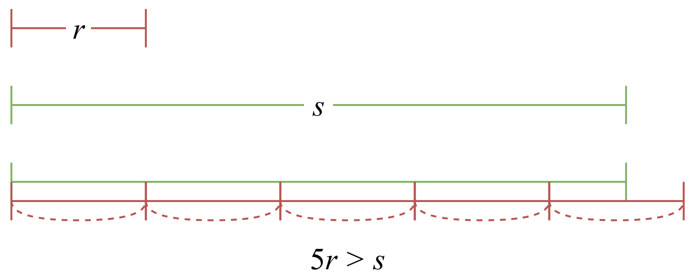

Figure 1: An example of the Archimedean property for real numbers.

Most philosophers and social scientists reject behaviorist accounts of credences.[6] According to those scholars, an agent's fair prices might be a crude measurement of her credence, but fair prices and credence are not the same. Instead, credence is said to be a mental state (or mental disposition) that is causally related to behavior, but not the sole determinant.[7] We do not defend mentalism, but we identify problems with several common, naive behaviorist interpretations in §1.3. Our reasons for discussing behaviorism will be clearer after we discuss the Archimedean axiom.

1.2 The Archimedean Axiom

Real numbers satisfy an Archimedean axiom: for any two real numbers $r, s > 0$, there is some natural number n such that $n \cdot r > s$. See Figure 1 for a visual example. Thus, if credences are numerically representable, then they must obey an Archimedean-like condition. In this section, we formulate a (simple) comparative Archimedean condition – one that can be expressed solely in terms of $\preceq$ and the set-theoretic structure among events in an algebra.[8] Then, we provide an example of comparative credences that seem rational but violate the condition. Thus, our Archimedean condition seems not to be justified by norms of individual rationality.

Before diving into the mathematical details, we sketch our argument informally. Many philosophers and scientists (e.g., Nicolas of Cusa) have believed that space is infinite and that the number of planets is likewise

[6]See [Eriksson and Hajek, 2007] for an especially trenchant criticism of behaviorist views.

[7]For example, see [Walley, 1991, §1.3]'s discussion of "theoretical" interpretations of probability.

[8]For similar Archimedean conditions for comparative probability relations, see [Krantz et al., 2006, p. 204].

infinite. Suppose, for the sake of argument, that Nicolas of Cusa considered two hypotheses about the number of planets that contained water: θ_1, that exactly one planet, Earth, contains water, and θ_2, that at least two planets contain water. It is rationally permissible, we claim, for Cusa to believe that, for any positive whole number n, the proposition that "There are at least two planets with water even though exactly n many planets will be observed by humans and, of those observed, none but Earth contains water" is at least as probable as θ_1. Space is infinite after all, so no matter how many planets Cusa observes, he will only have observed an infinitesimal fraction of all planets. Yet if he holds such beliefs, we claim that his comparative credences are not real-valued. In greater detail, consider the following condition.

Archimedean axiom for credence: Every bounded and disjoint sequence of events is finite, where a sequence of events $A_1, A_2, \ldots$ is *bounded and disjoint* if it has three properties:

1. $A_i \cap A_j = \emptyset$ if $i \neq j$,
2. $A_i \succeq A_1$ for all i, and
3. $A_1 \succ \emptyset$.

Why is the condition above analogous to the Archimedean axiom for real numbers? Think of the events $A_1, A_2, \ldots$ as line segments, just as real numbers were depicted by line segments above. Condition 1 says the line segments do not overlap; condition 2 says all of the segments are at least as long the first one, and condition 3 says the first segment has positive length. So if there were an infinite sequence of bounded and disjoint events $A_1, A_2, \ldots$, then its union $B = \bigcup_{n \in \mathbb{N}} A_n$ would be representable by an infinite line and would be "infinitely more probable" than A_1.

If an an agent's comparative credence relation $\preceq$ is representable by a finitely additive probability measure P (i.e. $A \preceq B \Leftrightarrow P(A) \leq P(B)$), then $\preceq$ satisfies the above Archimedean axiom for credence.[9] Thus, if it

[9] Proof: Otherwise, there is an infinite bounded and disjoint sequence $A_1, A_2, \ldots$, and thus for all n:

$$\begin{aligned} P(\bigcup_{k \leq n} A_k) &= \sum_{k \leq n} P(A_k) \text{ as } A_i \cap A_j = \emptyset \text{ if } i \neq j \\ &\geq n \cdot P(A_1) \text{ as } A_1 \preceq A_i \text{ for all } i \end{aligned}$$

Since $\emptyset \prec A_1$, we know $P(A_1) = r > 0$. Hence, by the Archimedean property of real is rationally permissible for one's credences $\preceq$ to violate the Archimedean axiom above, then it is rationally permissible to have credences that violate the probability axioms. We now defend the antecedent of that conditional.

Again, let θ_1 represent the proposition "Exactly one planet, Earth, contains water" and θ_2 represent "At least two planets contain water." Let $A_1 = \theta_1$, and for any natural number $n \geq 2$, let B_n represent the proposition that "exactly n many planets will be observed by humans but, of those observed, none but Earth contain water", and let $A_n = B_n \cap \theta_2$ be the conjunction of B_n and θ_2. Because it is impossible for humans to observe both exactly five and exactly six planets, the events A_n are disjoint and satisfy condition 1 above. Because the Earth has water, one should regard A_1 as strictly more likely than the impossible event $\emptyset$, as condition 3 requires. Finally, it is permissible to regard A_n as at least as plausible as A_1 – if one believes there are infinitely many planets – because one might believe there are many other planets that contain water that have yet to be observed. So it is rationally permissible for one's credences to satisfy condition 2.

Importantly, the example shows that, even when only *two* hypotheses are under investigation, one's credences might rationally violate the Archimedean condition if infinite amounts of data are available.

1.3 The Archimedean Axiom and Behaviorism

Our criticism of the Archimedean axiom requires rejecting two behaviorist theories of credence that entail the condition. So in this section, we first provide reasons for rejecting those behaviorist theories and then explain how non-behaviorist assumptions are employed in the argument above.

First, note that the two crude behaviorist views discussed in §1.1 entail that credence obeys an Archimedean condition. According to the first – where one *identifies* an agent's credence $c(A)$ in A with her fair price – credences are Archimedean because (i) prices are real numbers and (ii) real numbers, by definition, satisfy an Archimedean condition.

This strategy for defending the Archimedean axiom, however, faces a serious problem. If credence were nothing more than a disposition to bet in a particular way in highly artificial situations, then it would be of little philosophical interest. When philosophers and social scientists discuss credence, however, their working hypothesis is that there is some underlying disposition or mental state that is causally related to a wide variety of other mental states (e.g., desires and regrets) and behaviors. Fair prices

numbers, there is some sufficiently large n such that $P(\bigcup_{k \leq n} A_k) \geq n \cdot r > 1$, which is a contradiction.

might be one crude way of measuring that underlying disposition or mental state, but they aren't the only way. Further, sometimes fair prices might be wildly misleading: it is not at all clear that, when forced to offer prices in a laboratory, one's behavior will identify an underlying attitude that can consistently explain or predict one's behavior in other scenarios.

The second behaviorist strategy – in which one identifies comparative credences with outcomes of particular forced choices – is often used to justify the Archimedean axiom for $\preceq$ in a different way: one can only make *finitely* many choices, and hence, there are no infinite sequences of bounded disjoint events.[10] But this behaviorist strategy is even cruder than the last. It identifies credence not just with forced choices, but with forced choices that *have already been made*. Credences so understood are, we conjecture, theoretically useless for predicting future behavior and for explaining past behavior (as credence just *is* the outcome of forced choices).

There are, of course, other ways that one might try to identify credence with behavior. We cannot prove that no such attempt will be successful. Rather, we claim only that if one identifies credence with behavior in a way that make it easy to justify the Archimedean axiom, then one typically faces the difficult task of defending the claim that credence, so defined, plays any important predictive or explanatory purpose in describing *other* behaviors.[11]

To see why our argument requires rejecting the above two behaviorist protocols, consider our earlier example involving Nicolas of Cusa. A naive behaviorist might object that (i) Nicolas of Cusa's credences are Archimdean because they equal his fair prices or (ii) Nicolas of Cusa could only finitely many choices, and thus, his credence relation cannot be defined over an infinite space. To the first objection, we reiterate what we have said above: Cusa's fair prices at any give time are likely little use in explaining or predicting his behavior in other contexts. To the latter objection, we respond that it is perfectly plausible to compare the likelihood that n or m planets have water, for any numbers n and m. Such comparisons require one to adopt either (i) a mentalist framework, in which agents are attributed infinitely many non-occurrent credences or (ii) a behaviorist framework in which credence is understood as a disposition that is partially manifested in infinitely many possible scenarios. We see no reason to rule out such theories of credence.

[10] See [Krantz et al., 2006, pp. 25-26] for discussion of this defense.

[11] For what it's worth, we think philosophers and social scientists should aim to develop a plausible, *operationalist* account of credence, but there are many possible ways of measuring "mental" states other than through behavior (e.g., through fMRI scans, eye-tracking experiments, etc.).

2 Collectivist Foundations for Bayesianism

2.1 Public Probabilism

In the previous section, we argued that three standard arguments for probabilism – which appeal exclusively to norms of individual rationality – share a common weakness: they do not justify the Archimdean axiom. We now argue that, luckily, to justify the use of Bayesian statistics in science, one need not assume that a scientist's *private* credences satisfy the probability axioms.

Our view is not entirely novel. Many statisticians already reject the view that the prior probability distributions appearing in scientific journals should be interpreted as the author's credences. For example, [Gelman and Shalizi, 2013, p. 19], two prominent statisticians, write, "… the prior distribution $p(\theta)$ is one of the assumptions of the model and does not need to represent the statistician's personal degree of belief in alternative parameter values." They continue:

> We do not have to worry about making our prior distributions match our subjective beliefs still less about our model containing all possible truths. Instead we make some assumptions, state them clearly, see what they imply, and check the implications. This applies just [as] much to the prior distribution as it does to the parts of the model showing up in the likelihood function [Gelman and Shalizi, 2013, p. 20].

We will not reconstruct Gelman and Shalizi's view in detail, but the quotation raises an interesting possibility: Bayesian statistical methods might be justified without discussing rational belief at all!

Unlike Gelman and Shalizi, we do think of prior probabilities as representing belief *in some sense*, but not as representing the experimenter's *actual* credences, nor even as idealizations of those credences. Instead, prior probabilities are, we claim, best interpreted as a model of a scientist's *public* (or "professional") credences.

We think of modeling scientists' beliefs in much the same way one might model the beliefs of a juror in a criminal trial. Famously, jurors are required to ignore some types of evidence (e.g., hearsay), and they are obliged to consider other types (e.g., exhibits introduced during trial). A model of a juror's decision-making – whether descriptive or normative – would be wildly inaccurate if it considered the private beliefs of the juror qua citizen; what matters are the beliefs of the juror qua juror. We refer to these as the juror's *public* beliefs. The juror's public beliefs are, we think, rightfully called "beliefs" because they explain the juror's courtroom behavior in much the way her private beliefs explain her behavior in her private life.[12]

Analogously, participation in a scientific community requires one to adopt certain beliefs and to modify those beliefs in a particular way, regardless of one's private opinions. A model of scientific decision-making – whether normative or descriptive – might likewise be inaccurate if it confuses the beliefs of a scientist qua scientist with her beliefs qua private citizen.

For example, consider the paleontologist Marcus Ross, who identifies as a creationist.[13] Ross professes to believe that the universe is only 10,000 years old despite having written a dissertation on a marine animal that is widely accepted to be 65 million year old. When asked whether the reasoning in his dissertation was sound, Ross responded, "Within the context of old age and evolutionary theory, yes. But if the parameter is different, portions of it could be completely in error." Here, Ross himself distinguishes between his private beliefs and what he advocates on the pages of scientific journals. Any attempt to explain or predict Ross' professional behavior on the basis of personal religious convictions, therefore, is doomed to failure.

Similarly, any attempt to *evaluate* Ross' scientific work should not confuse his scientific reasoning with his religious thinking. Ross' publications, grant proposals, and teaching – his behavior qua scientist – might be scientifically rigorous and satisfy all the demands of normative decision theory. Of course, Ross' behavior *considered as a whole* may be deeply irrational: his professional choices seem incoherent given his private beliefs. Yet that irrationality is irrelevant to the coherence of his professional behavior.

In general, when we participate in different institutions (e.g., as jurors, journal referees, etc.), we may be required to act as if our beliefs differ from our private beliefs for two reasons. First, the evidence that we are required to consider (or not) in virtue of the institution's goals may differ from the evidence that is available to us as individuals. Second, procedural constraints require us to *evaluate* evidence in particular ways (e.g., rubrics in hiring committees).

[12] Nothing in our argument hinges on us calling the juror's attitudes "beliefs." Some philosophers might prefer to say the juror "accepts" certain propositions [Cohen, 1992, Levi, 1967, Maher, 1993, Van Fraassen, 1980]. For those who prefer the language of acceptance, our thesis is that scientists have some professional, graded acceptance-like attitude that must obey the probability axioms. [Fleisher, 2018] defends a similar thesis, namely, that the hypotheses a scientists *endorses* should be selected to maximize expected (epistemic) utility. Comparing our view to Fleisher's is beyond the scope of this paper.

[13] Our description of Ross's beliefs is based on [Rosin, 2007]'s report.

Because a scientist's private credences may differ from her public ones (i.e., her credences qua scientist), probabilism – if understood as a thesis about a scientist's private credences – is neither necessary nor sufficient to justify the use of Bayesian statistics in science. Instead, the use of Bayesian statistics requires the truth of what we call *public* probabilism, which asserts that a scientist's public credences should obey the probability axioms, at least on a restricted set of propositions in her domain of expertise.[14]

2.2 Epistemic Cooperativeness

In the next section, we argue that a norm of scientific inquiry, which we call *epistemic cooperativeness* (or just "cooperativeness" for short), requires that scientists' credences obey an Archimedean condition. In this section, we define "cooperativeness" and argue it is a norm of contemporary science.[15]

Epistemic cooperativeness, in our sense, is a relation between a researcher, a scientific question, and a community, which may include both experts and laypersons. Roughly, we say a researcher is epistemically cooperative (or just "cooperative", for short) if she is *publicly* open-minded towards all "live" hypotheses about the question at hand. For example, Galileo was cooperative towards his peers and the general public with respect to the question of whether a heliocentric or geocentric model best explained everyday phenomena and astronomical data at his time. Whether Galileo privately believed that geocentrism was unfathomably stupid is irrelevant to whether he was cooperative in our sense. Galileo's published writings and his correspondences with Church officials contain a serious engagement with geocentrism.

More precisely, cooperativeness constrains a scientist's (1) evaluation of evidence and (2) experimental design. When analyzing data, a cooperative researcher cannot employ methods that foreclose the possibility of finding evidence for rival, live hypotheses and against her own pet hypothesis. When designing experiments and gathering data, a cooperative researcher must ensure that her experimental protocol has some ability to distinguish among rival live hypotheses (if there are any) and to find faults with any hypotheses she seeks to confirm.[16]

[14] For instance, a computer scientist need not have an informed stance on climate change.

[15] Although sociologists of science have studied norms of science extensively (e.g., see "The Normative Structure of Science" in [Merton, 1973]), our notion of "epistemically cooperative" cuts across several different purported scientific norms (e.g., Merton's norms of "disinterestedness", "universalism", and "communism"), and so we give it a new name.

[16] In [Mayo, 1996, 2018]'s terminology, experiments and data-analysis methods must be "severe tests" of the experimenter's hypothesis. Unlike Mayo, we do not think the philo-

We emphasize three features of our definition of cooperativeness. First, what counts as "live" can change; astronomers no longer need to test geocentric models. Second, a scientist may be cooperative but extremely critical of her peers' work. Third, cooperativeness does not require "impartiality" on "non-partisanship." A cooperative scientist may seek to discredit others and to confirm her own pet hypotheses. In fact, an epistemically cooperative scientist may be motivated entirely by fame, monetary prizes, and personal grudges against her rivals.[17] To be cooperative, in our sense, requires only that a scientist's public behavior meets the two conditions above.

What is required to meet the first condition of epistemic cooperativeness, i.e., to avoid foreclosing "the possibility of finding evidence for rival hypotheses and against [one's] pet hypothesis"? Here is a weak necessary condition that will play a central role in our argument.

Weak Cooperativeness: If H_1 and H_2 are live hypotheses and a scientist receives an indefinite amount of evidence in favor of H_1 and against H_2, then she must eventually *not* regard H_2 as infinitely more likely than H_1.[18]

The reader might wonder why what we call "cooperativeness" is a norm that arises only for scientists engaged in collective inquiry. Why, for example, would an inquisitive person deserted on a remote island not be bound by the same norms? Why not call the virtues we have described "open-mindedness" or "vigilance against error", which are phrases that seem to describe virtues that a researcher might exhibit in isolation?

We do not deny the existence of purely individualistic epistemic norms.

sophical motivation for severe-testing prohibits the use of Bayesian statistical methods, but cooperativeness may constrain the prior distribution representing the scientist's public credences. Of course, to consider whether cooperativeness constrains the "choice" of a prior *probability* distribution, one must first argue that a scientist's public credences are representable by a probability distribution at all, which is the entire point of this paper.

[17] We take no stance on whether such behavior or motivations violate other scientific, moral, or social norms. However, there are compelling arguments that "epistemically impure" motives might improve science [Kitcher, 1990, 1995, Strevens, 2003].

[18] The requirement that the acquired evidence is both *for* H_1 and *against* H_2 is essential here. Suppose H_1 is the hypothesis that the value of some unknown physical constant θ is *exactly* π, and let H_2 be the hypothesis that θ is close to but not equal to π, e.g., H_2 asserts $\theta \in [\pi - \epsilon, \pi + \epsilon] \setminus \{\pi\}$ for some small ϵ. Although measurements might confirm over time that the value of θ matches π to further and further decimal places, such measurements don't distinguish H_1 from H_2. Although H_1 and H_2 are rival hypotheses, every piece of data that is compatible with H_1 is likewise compatible with H_2. So one is not obliged to regard H_1 as more probable than H_2, even if one acquires an indefinite amount of evidence for H_1. Thanks to Blinded for review for asking us to make this clear.

We claim only that those norms, if they exist, are weaker than what many philosophers have imagined. To see why, consider the question, "*Which* hypotheses should a scientist take seriously when designing experiments and evaluating data?" A scientist is not obliged to consider *all* hypotheses that explain her data for there are innumerable such hypotheses that have yet to be articulated.

We claim that the set of hypotheses that a scientist is obliged to consider depends upon, among other things, the history of her field of study and the current work of her peers. Why? A scientist would be considered negligent if, each time she acquired novel data or designed a new experiment, she ignored all past research and considered only several new hypotheses that she personally devised. Practically speaking, the norm to consider rival hypotheses is often enforced through peer review, where referees for grant-giving institutions and journals evaluate whether the proposer or author has discussed and cited "relevant" or "appropriate" literature."[19] Similar practices show many scientists' obligations depend upon their institutional roles (e.g., as employees of public universities, members of academic societies, journal referees, etc.) and upon the interests of society at large.

One might object that, although cooperativeness is a collective norm, Weak Cooperativeness is not. The objector might grant that the set of hypotheses that a scientist must consider is determined (in part) by professional and social obligations, but that once that set is determined, Weak Cooperativeness amounts to the duty to be responsive to evidence. Such a demand is a paradigmatic norm of individual (epistemic) rationality.

We disagree. As we show in the next section, Weak Cooperativeness entails that a scientist's credences should obey the Archimedean condition. Thus, if a rational agent's credences may violate the Archimedean condition, then Weak Cooperativeness is not a requirement of individual rationality. Our earlier example involving Nicholas of Cusa demonstrates this, as the example suggests that he is individually rational but not weakly cooperative. In particular, he does not seriously engage with the hypothesis that Earth is the only planet with water. The objector, we think, mistakes a vague, individualistic epistemic norm (to consider evidence) with a rather precise, formal consequence of epistemic cooperativeness.

At the very least, we think the normative force of Weak Cooperativeness is *strengthened* by a scientist's obligations towards others. Imagine an uncooperative scientist S who *publicly* believes a hypothesis H and might

[19] For example, referees for *Nature* are asked, "does this manuscript reference previous literature appropriately? If not, what references should be included or excluded?"[noa].

continue to believe H (publicly) in the face of indefinite evidence to the contrary. In such a case, S's colleagues would be unlikely to seek out her paper defending H, for her colleagues would know that S might conclude that H in spite of large amounts of disconfirming evidence; after all, S's dogmatic stance towards H is *public*.

If S's data were publicly accessible, then expert readers could evaluate the data themselves. But data are not always publicly accessible, and reanalyzing another researcher's work is often costly and time-consuming. Equally importantly, some scientists may need to rely on S's results but lack the technical knowledge to re-analyze S's data. Finally, even if S honestly discloses her data, her colleagues might worry that she ignored or failed to report relevant evidence because, by supposition, some pieces of evidence will not sway S at all. In short, if a scientist violates Weak Cooperativness, then her research is likely to be ignored by her colleagues, and if it is not ignored, it might not be trusted. Thus, even if Weak Cooperativeness is partially supported by norms of individualistic rationality, its force is strengthened by considering a scientist's obligations to her peers.

Is cooperativeness so described actually a norm of science? As we have discussed above, peer review (of articles and grant proposals) suggests that scientists are required to evaluate how well their evidence bears on a variety of live hypotheses. Contemporary calls for pre-registered trials - with the requirement to include a detailed data-analysis plans – provide further evidence that scientists are discouraged from using methods that preclude the possibility of finding support for their rivals' hypotheses [Nosek et al., 2018].

2.3 Non-Archimedean credences violate epistemic cooperativeness

We now argue that if a scientist initially regards one hypothesis H_1 as "infinitely more probable" than another H_2 – in other words, her prior credences over hypotheses violate the Archimedean condition – then she might acquire an indefinite amount of data favoring H_2 and nonetheless continue to believe that H_1 is infinitely more plausible than H_2.[20] Thus, her beliefs violate Weak Cooperativeness.

Let Θ be a set of competing, live hypotheses. Following standard statistical terminology, we call elements of Θ *simple* hypotheses. We call a

[20]Some philosophers have often argued that being appropriately "open-minded" *requires* that one's credences be representable by hyperreals, which violates an Archimedean condition. See [Easwaran, 2014] for a summary and criticism of those arguments.

subset $H \subseteq \Theta$ a *composite* hypothesis, representing a disjunction of simple hypotheses.

Imagine a scientist designs an experiment with outcomes in Ω, and so her credence relation $\preceq$ orders events in the Cartesian product $\Theta \times \Omega$. Given $H \subseteq \Theta$ and $E \subseteq \Omega$, the set $H \times E$ thus represents the event that (i) one of the hypotheses in H is true, and (ii) one of the experimental outcomes in E occurs. Below, we will often write H instead of $H \times \Omega$ to represent the event that one of the simple hypotheses in H is true and similarly we often write E instead of $\Theta \times E$ to represent the event that some outcome in E is observed. Finally, we will combine these conventions and write $H \cap E$ to denote $(H \times \Omega) \cap (\Theta \times E) = H \times E$.

To represent how the scientist's credences change over time, we make the following assumption about her *conditional credence* $H|E$ that H is true given E is observed:

Assumption †: If (1) $H_1 \cap E \preceq H_2 \cap E$ and (2) $H_1|E$ and $H_2|E$ are well-defined, then $H_1|E \preceq H_2|E$.

The assumption is true if (i) one's credences are representable by a numerical probability function P and (ii) one's conditional credence $P(H|E)$ is defined by the ratio formula $P(H|E) = \frac{P(H \cap E)}{P(E)}$.[21] However, we think the assumption has plausibility outside those circumstances as well. In particular, it might have non-trivial consequences even if one's theory of probability permits conditioning on events with zero probability.

Now suppose that the scientist's credences violate the Archimedean condition, i.e., she believes one simple hypothesis θ_2 is "infinitely more probable" than another θ_1. Formally, assume there is an infinite, bounded and disjoint sequence of observable events $E_1, E_2, \ldots \subseteq \Omega$ such that the experimenter regards $\theta_2 \cap E_n$ as at least as probable as θ_1 for all n.[22]

Now consider what happens if the scientist first learns $E = \bigcup_{n \in \mathbb{N}} E_n$ (i.e., that at least one E_n is true) and then, at a discrete set of times $1, 2, \ldots$, the scientist learns $\neg E_1, \neg E_2$, and so on for all n. By stage n, therefore, the scientist has learned $E \cap \bigcap_{k \leq n} \neg E_k$, which is equivalent to learning that $F_n = \bigcup_{k > n} E_k$. Thus, the scientist's conditional credence at stage n in θ is equal to $\theta|F_n$. Using our earlier example, Nicolas of Cusa learns at stage 1 that one planet other than Earth has been scoured for water unsuccessfully;

[21]Proof: If $P(E) > 0$, then $P(H_1|E) \geq P(H_2|E)$ if and only if $P(H_1 \cap E) \geq P(H_2 \cap E)$.

[22]For brevity, we often write $\theta \cap A$, $\theta|A$, and $\theta \times A$ instead of $\{\theta\} \cap A$, $\{\theta\}|A$, and $\{\theta\} \times A$ respectively.

at stage 2, he two planets have been scoured, and so on.

We claim that, at every stage of inquiry, the experimenter still regards θ_2 as infinitely more probable than θ_1. More precisely, we claim that, for any stage n, there is an infinite, bounded and disjoint sequence of observable events $G_1, G_2, \ldots \subseteq \Omega$ such that *conditional on* F_n, the experimenter regards $\theta_2 \cap G_k$ as at least as probable as θ_1 for all k. In fact, we show the claim holds by letting $G_k = E_{k+n}$, assuming the relation $\preceq$ is transitive and has the property that $A \succeq B$ whenever $B \subseteq A$ (i.e., the experimenter regards A is at least as likely as B whenever B entails A).

Claim: $\theta_2 \cap E_{k+n} | F_n \succeq \theta_1 | F_n$ for all n and all $k \geq 1$.
Proof: By assumption, $\theta_2 \cap E_m \succeq \theta_1$ for all m. Thus, $\theta_2 \cap E_{k+n} \succeq \theta_1$ for all n and all $k \geq 1$. Since $\theta_1 \cap F_n \subseteq \theta_1$, it follows that $\theta_2 \cap E_{k+n} \succeq \theta_1 \succeq \theta_1 \cap F_n$ for all n and all $k \geq 1$. Now $(\theta_2 \cap E_{k+n}) \cap F_n = \theta_2 \cap E_{k+n}$ since $E_{k+n} \subseteq F_n = \bigcup_{m>n} E_m$. Thus, by transitivity, $(\theta_2 \cap E_{k+n}) \cap F_n \succeq \theta_1 \cap F_n$. By Assumption †, it follows that $\theta_2 \cap E_{k+n} | F_n \succeq \theta_1 | F_n$ for all n and all $k \geq 1$, which is what we desired to show. $\Box$

Importantly, nothing in our argument prevents us from also assuming that the evidence $\neg E_n$ favors θ_1 over θ_2 *for every natural number* n![23] In short, as we stated above, the scientist dogmatically continues to believe θ_2 is true, no matter how much data she acquires in favor of θ_1.

To return to the example of the previous section, our fictitious version of Nicolas of Cusa might not be individually irrational to continue believing θ_2, that "At least two planets contain water", no many how many waterless planets he observes. But if his peers regard hypothesis θ_1 "Only Earth contains water" as a legitimate competitor to θ_2, then Nicolas of Cusa violates the norm of epistemic cooperativeness by collecting data that he knows might never lead him to take his peers' view seriously.

There are at least three limitations of our argument. First, we have not argued that a scientist with non-Archimedean credences is *guaranteed* to remain dogmatic; we have argued that the scientist believes there is a positive probability she will remain convinced in a hypothesis, even in the face of an indefinite amount of contrary evidence. Second, our argument does not show that the experimenter cannot regard certain *experimental outcomes* as infinitely more probably than others. Finally, our argument

[23] Here, one can be agnostic about what "favoring" means, but one way of specifying that is via a qualitative law of likelihood that asserts the some evidence E favors H_1 over H_2 precisely if $E|H_1 \succ E|H_2$.

only works in finite hypothesis spaces: nothing we have said prevents a scientist from regarding some infinite (composite) hypothesis $\{\theta_n\}_{n\in\mathbb{N}}$ as infinitely more probable than another hypothesis.

Despite the limitations, we think the above argument shows how norms of collective science might be used to justify, at least partially, axioms that are necessary for an experimenter's public credences to obey the probability axioms. Importantly, we do not claim that norms of collective science are sufficient, by themselves, to justify public probabilism; norms of individual rationality might still have some role to play.

3 Objections and Replies

Norms often conflict, and so norms of individual rationality likely conflict with norms of collective science [Mayo-Wilson et al., 2011]. Thus, one might object that our appeal to both types of norms to justify public probabilism is illegitimate, as contradictory norms could justify any thesis whatsoever.

Such an objection proves too much, as it applies to any normative thesis about a juror's duties, a doctor's obligations, a sporting referee's responsibilities, and so on. For instance, Bayesian decision theory entails one should never turn down free information [Good, 1967], but jurors are required to turn down free evidence if it is inadmissible.[24] However, despite the apparent conflict between norms for jurors and norms of individual rationality, no one thinks that "any thesis whatsoever" about jurors' obligations can be justified by the two types of norms.

Where does the objection go wrong? To answer that question, note there are at least two ways the objection can be disambiguated. First, our critics might claim that a scientist's duties might conflict with her duties as a private citizen, in the same way a criminal lawyer's obligation to defend a client she knows to be guilty might conflict with her duty, qua private citizen, to report known criminal activity. Of course, we agree that private and professional obligations often conflict. But notice the conflicts in the lawyer's private and professional duties do nothing to challenge the thesis that, qua lawyer, she is obliged to defend her client. Similarly, the fact that a scientist has conflicting private and professional duties does little to refute our thesis that, qua scientist, she should have probabilistic credences.

Second, the objector might continue that our defense of public probabilism is importantly disanalogous from the examples involving lawyers and jurors. Lawyers' professional duties qua lawyers *override* their obligations

[24] Thanks to Blinded for review for this example.

qua private citizens; similar remarks apply to jurors. Conflict is avoided by specifying a priority of duties. In contrast, a defense of public probabilism might, we have claimed, require appeals to both norms of collective science and norms of individual rationality; if one of those types of norms outranks the other in cases of conflict, one may be unable to use both types in defense of public probabilism.

We think this objection gains plausibility by trading on ambiguity. It may be plausible that *some* norms of collective science conflict with *some* norms of individual rationality. But thusfar, we have little reason to suspect the conflicting norms are the ones required to justify public probabilism.

4 Conclusions and Future Work

We have argued that, although rational private credences might violate an Archimedean axiom, a scientist's public credences over hypotheses must obey an Archimedean axiom. But one might imagine that other foundations of probabilism are suspect. So our argument is only a small step towards justifying public probabilism, and it is a very small step towards justifying the use of Bayesian statistics in science. Future work, therefore, ought to investigate to what extent collective norms of science can justify (i) conditionalization and (ii) other axioms required for probabilism that may be poorly justified by individual rationality.

Equally importantly, we think our framework opens up potential avenues for justifying other parts of Bayesian statistical practice that might seem incompatible with orthodox decision theory. For example, some Bayesian statisticians argue that one ought to "choose" supposedly non-informative priors when analyzing data; others argue that scientists ought to choose priors that have good frequentist properties.[25] From the standpoint of traditional decision theory, analyzing data using a probability distribution that deviates from one's personal credences is potentially irrational. Just as it would be irrational for Alison to choose an action just because it maximizes Bill's subjective expected utility, a Bayesian statistician who chooses a prior in one of the two ways just specified may advocate statistical decisions that would fail to minimize expected loss by her own lights. Our argumentative strategy in this paper suggests a novel way to rationalize these two methods for "choosing" a prior: norms of science might require a statistician to adopt public credences that differ from her private ones.

[25] See Excursion/Chapter 6 of [Mayo, 2018] for a discussion of various ways in which contemporary Bayesians choose priors.

References

Peer-review policy | Nature Research. URL https://www.nature.com/nature-research/editorial-policies/peer-review#the-review-process.

L. Jonathan Cohen. *An essay on acceptance and belief*. Oxford: Clarendon Press, 1992.

Kenny Easwaran. Regularity and hyperreal credences. *Philosophical Review*, 123(1):1–41, 2014.

Lina Eriksson and Alan Hajek. What are degrees of belief? *Studia Logica*, 86(2):183–213, 2007. URL http://link.springer.com/article/10.1007/s11225-007-9059-4.

Peter C. Fishburn. The axioms of subjective probability. *Statistical Science*, 1(3):335–358, 1986. URL http://projecteuclid.org/euclid.ss/1177013611.

Will Fleisher. Rational endorsement. *Philosophical Studies*, 175(10):2649–2675, 2018. Publisher: Springer.

Andrew Gelman and Cosma Rohilla Shalizi. Philosophy and the practice of Bayesian statistics. *British Journal of Mathematical and Statistical Psychology*, 66(1):8–38, 2013.

Irving John Good. On the principle of total evidence. *The British Journal for the Philosophy of Science*, 17(4):319–321, 1967.

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

Philip Kitcher. The Division of Cognitive Labor. *The Journal of Philosophy*, 87(1):5–22, 1990.

Philip Kitcher. *The Advancement of Science: Science Without Legend, Objectivity Without Illusions*. Oxford University Press, USA, 1995.

David H. Krantz, R. Duncan Luce, Patrick Suppes, and Amos Tversky. *Foundations of Measurement Volume I: Additive and Polynomial Representations*. Dover Publications, Mineola, NY, December 2006. ISBN 978-0-486-45314-9.

Isaac Levi. *Gambling with Truth*. MIT Press, Cambridge, 1967.

Patrick Maher. *Betting on theories*. Cambridge University Press, 1993.

Deborah G. Mayo. *Error and the growth of experimental knowledge*. University of Chicago Press, 1996.

Deborah G. Mayo. *Statistical inference as severe testing*. Cambridge: Cambridge University Press, 2018.

Conor Mayo-Wilson, Kevin J. Zollman, and David Danks. The Independence Thesis: When Individual and Social Epistemology Diverge. *Philosophy of Science*, 78(4):653–677, 2011.

Robert K. Merton. *The sociology of science: Theoretical and empirical investigations*. University of Chicago press, 1973.

Brian A. Nosek, Charles R. Ebersole, Alexander C. DeHaven, and David T. Mellor. The preregistration revolution. *Proceedings of the National Academy of Sciences*, 115(11):2600–2606, 2018. Publisher: National Acad Sciences.

Samir Okasha. On the interpretation of decision theory. *Economics & Philosophy*, 32(3):409–433, 2016.

Richard Pettigrew. *Accuracy and the Laws of Credence*. Oxford University Press, 2016. ISBN 978-0-19-873271-6.

Hanna Rosin. Rock of Ages, Ages of Rock (Published 2007). *The New York Times*, November 2007. ISSN 0362-4331. URL https://www.nytimes.com/2007/11/25/magazine/25wwln-geologists-t.html.

L. J. Savage. *The foundation of statistics*. Dover publications, 1972.

Teddy Seidenfeld, Mark J. Schervish, and Joseph B. Kadane. Coherent choice functions under uncertainty. *Synthese*, 172(1):157, March 2009. ISSN 1573-0964. doi: 10.1007/s11229-009-9470-7. URL https://doi.org/10.1007/s11229-009-9470-7.

Amartya K. Sen. Choice functions and revealed preference. *The Review of Economic Studies*, 38(3):307–317, 1971. Publisher: JSTOR.

Michael Strevens. The Role of the Priority Rule in Science. *The Journal of philosophy*, 100(2):55–79, 2003.

B.C. Van Fraassen. *The scientific image*. Oxford University Press, USA, 1980.

Susan Vineberg. Dutch book arguments. 2011. URL http://stanford.library.usyd.edu.au/archives/spr2012/entries/dutch-book/.

Peter Walley. *Statistical reasoning with imprecise probabilities*. Number 42 in Monographs on Statistics and Applied Probability. Chapman & Hall, New York, 1991.