

Bayesian Humility

Adam Elga*

September 8, 2015

Abstract

Say that an agent is *epistemically humble* if she is less than fully confident that her opinions will converge to the truth, given appropriate evidence. Is such humility rationally permissible? According to *the orgulity argument* (Belot 2013): the answer is “yes” but long-run convergence-to-the-truth theorems force Bayesians to answer “no.” That argument has no force against Bayesians who reject countable additivity as a requirement of rationality. Such Bayesians are free to count even extreme humility as rationally permissible. Furthermore, dropping countable additivity does not render Bayesianism more vulnerable to the charge that it is excessively subjective.

1 Introduction

Presented with Bayesian confirmation theory, it is easy to feel cheated. One might have hoped for a substantive, detailed account of what sorts of evidence support what sorts of scientific hypotheses. Instead one is told how one’s evidence determines reasonable attitudes toward such hypotheses *given a prior* (an initial probability function). And one is told that different priors deliver different outputs, even for the same batch of total evidence.

One might worry that given this dependence, Bayesianism is ill-placed to explain the significant agreement observed among reasonable scientists, or to

*Penultimate draft. Please cite published version: *Philosophy of Science* 83 (July 2016) pp. 305–323. For helpful comments and discussion, thanks to Andrew Bacon, Gordon Belot, Cian Dorr, Kenny Easwaran, Kevin Kelly, Bas van Fraassen, Jim Hawthorne, Teddy Seidenfeld, Brian Weatherson, Jonathan Wright, and anonymous reviewers. I thank John Burgess for introducing me to the dualities between measure and category when he was my junior paper advisor in the spring of 1995. For financial support I am grateful to the David A. Gardner ’69 Magic project (through Princeton University’s Humanities council), the PIIRs Research Community on Systemic Risk, and a 2014-15 Deutsche Bank Membership at the Princeton Institute for Advanced Study. For an extremely congenial work environment in July 2014, thanks to Mindy and Gene Stein.

deliver an objective account of confirmation in science.^{1,2}

In the face of this worry it is natural to seek comfort from some remarkable long-run “convergence-to-the-truth” and “washing out” theorems. These theorems show that unless priors differ radically, differences between them become negligible in the long run, under the impact of a stream of common evidence. This is sometimes thought to take the sting out of the above worry, by showing that many differences between priors don’t end up mattering.

But Belot (2013) argues that rather than helping Bayesianism, such convergence theorems are a liability to it. The argument is that the theorems preclude Bayesians from counting as rational “a reasonable modesty” about whether one’s opinions will approach the truth in the long run.

I will argue:

1. Long-run convergence theorems are no liability to *finitely additive Bayesianism*, a version of Bayesianism that rejects countable additivity as a requirement of rationality.³ Defenders of finitely additive Bayesianism are free to count any amount of humility about convergence to the truth—even extreme pessimism—as rationally permissible.⁴
2. Long-run convergence theorems are of no help to Bayesians responding to concerns about scientific agreement. In contrast, *short-run* convergence theorems (Howson and Urbach 2006, 238; Hawthorne 2014, §5) are of some help. Those theorems do not require countable additivity.

Let me take these points in turn, starting with a brief explanation of long-run convergence theorems, and an assessment of the charge that such theorems count against Bayesianism.

¹For expressions (but not always endorsements) of this worry, see Chalmers (1999, 133) as cited in Vallinder (2012, 8), Easwaran (2011, §2.6), Earman (1992, 137), Howson and Urbach (2006, 237), Hawthorne (2014, §3.5).

²Note that for present purposes, a confirmation theory may count as Bayesian even if it imposes constraints on priors more restrictive than mere coherence. Thanks here to Cian Dorr.

³Here I apply observations from Juhl and Kelly (1994, 186) and Howson and Urbach (2006, 28–29).

⁴Weatherson (2014) convincingly argues that *unsharp* Bayesians (who hold that states of graded opinion should be represented not by probability functions, but rather by sets of probability functions) are also free to count humility about convergence to the truth as rationally permissible.

2 Long-run Convergence Theorems

To introduce the long-run convergence theorems that are wielded in Belot (2013), consider an immortal investigator whose evidence consists of successive digits from a countably infinite binary sequence (a sequence consisting of 0s and 1s). The investigator receives one digit of the sequence per day, and is interested in H , a fixed hypothesis about the whole sequence. For example, H might be the proposition that after a certain point, the sequence consist of all 1s. (For convenience I treat interchangeably a proposition about the sequence, and the corresponding set of sequences for which that proposition holds.) From here on, I will assume that this setup is in place unless otherwise noted.

Now apply Bayesian confirmation theory to this setup. In particular, suppose that the investigator starts with a prior probability function P defined over an appropriate domain that includes H , updates by conditionalization each time she receives a digit, and is certain of all of the above.

Formal construction Let \mathcal{C} denote $\{0, 1\}^{\mathbb{N}}$, the collection of all functions from \mathbb{N} to $\{0, 1\}$. Thus each element of \mathcal{C} is a (countably) infinite binary sequence. For any finite or infinite binary sequence x and any positive integer i , we write $x(i)$ to denote the i^{th} term of x —that is, the binary digit that x assigns to i —and we write x^i to denote the set $\{y \in \mathcal{C} : x(i) = y(i) \text{ for } 1 \leq i \leq n\}$ of sequences from \mathcal{C} whose first i digits match x . By “string” we shall mean “finite binary sequence”, and for convenience we count the null sequence—the sequence of length 0—as a string. Given string s , let $[s]$ denote the set of sequences from \mathcal{C} that start with s .

We assume that \mathcal{C} is a Cantor space. That is, we assume that \mathcal{C} is endowed with the product topology. In this context, the family of all sets $[s]$ (for strings s) is a denumerable topological basis for the Cantor Space \mathcal{C} . Now let \mathcal{B} denote the Borel σ -algebra generated by this topology. The pair $\langle \mathcal{C}, \mathcal{B} \rangle$ accordingly forms a measurable space. In this paper we restrict attention to probability functions on the Borel σ -algebra \mathcal{B} over \mathcal{C} .⁵

We represent the prior beliefs of an ideal Bayesian-rational investigator by a probability function P on \mathcal{B} : a nonnegative, finitely additive set function on \mathcal{B} satisfying $P(\mathcal{C}) = 1$. We assume throughout that P is finitely additive: $P(A_1 \cup \dots \cup A_n) = P(A_1) + \dots + P(A_n)$ whenever $\{A_1, \dots, A_n\}$

⁵Thanks here to an anonymous reviewer, whose suggested exposition I adopt almost verbatim.

is a finite set of pairwise disjoint members of \mathcal{B} . We will sometimes but not always make the stronger assumption that P is countably additive: $P(A_1 \cup A_2 \cup \dots) = P(A_1) + P(A_2) + \dots$ whenever $\{A_1, A_2, \dots\}$ is a finite or countably infinite set of pairwise disjoint members of \mathcal{B} .

For every string s such that $P([s]) > 0$, we assume that when the investigator sees s as the first digits of the observed sequence, her new probability function is $P(\cdot \mid [s])$, the result of conditionalizing P on $[s]$. For the purposes of this paper, we needn't specify what happens when an investigator observes a string to which she had previously assigned probability 0.

Now return to our investigator. Before seeing any digits, she might wonder: in the limit of seeing more and more digits, how likely is it that I will arrive at the truth about H ? In other words, how likely is it that my probability for H will converge to 1 if H is true and to 0 otherwise?

A pessimistic answer to that question is: I am unlikely to converge to the truth (about H). A more confident answer is: I will probably converge to the truth. A maximally confident answer is: my probability that I will converge to the truth equals 1.

A celebrated long-run convergence theorem entails that *if the investigator's probability function is countably additive*, then she is committed to the maximally confident answer:

Theorem For any countably additive probability function P on \mathcal{B} and any hypothesis $H \in \mathcal{B}$,

$$P(\{x : \mathbf{1}_H(x) = \lim_{i \rightarrow \infty} P(H \mid x^i)\}) = 1,$$

where $\mathbf{1}_H$ is the indicator function for H (taking value 1 or 0 according to whether its argument is or is not a member of H).

Proof This is an immediate consequence of the Lévy zero-one law (Lévy 1937).⁶

So: countably additive Bayesian confirmation theory entails that in the above situation, rationality requires investigators to have *full* confidence that their opinions will converge to the truth.

⁶For a fairly accessible proof, see Halmos (1974, Theorem 49B, p. 213). For additional discussion, and a related application of Doob's martingale convergence theorem, see Schervish and Seidenfeld (1990, §4). For an explanation emphasizing the role that countable additivity plays in a similar proof, see Kelly (1996, 325–327).

But Belot (2013) gives an ingenious argument that rationality requires no such thing.

3 *The Orgulity Argument*

Here is a stripped-down exposition of what I shall call “the orgulity argument”—the argument from Belot (2013) that poses the strongest challenge to Bayesianism.⁷

Explaining the argument will require a bit of setup. As before, let H be a hypothesis about the infinite binary sequence that is in the domain of the investigator’s probability function. Examples of such hypotheses include: that the sequence eventually becomes periodic, that it ends with the pattern “01010101 . . .”, that it is computed by a Turing Machine (that the function $x : \mathbb{N} \rightarrow \{0, 1\}$ giving successive digits of the sequence is a computable function) or that it contains infinitely many 0s. (Note that H is not required to be countable. Cf. Belot (2013, n. 32).)

Say that an investigator is **open-minded** with respect to H if for every finite batch of evidence, there is a finite extension of it that would lead her to assign probability greater than $1/2$ to H , and also a finite extension of it that would lead her to assign probability less than $1/2$ to H .⁸ An open-minded investigator commits to never irrevocably making up her mind about whether H or not- H is

⁷Besides putting forward the orgulity argument, Belot (2013) has another main goal: to show that in certain contexts, convergence to the truth by Bayesian investigators is much harder to achieve than is commonly supposed. In particular, suppose that an investigator observes independent draws from a fixed chance process. It is well known that if the chance process has only a finite set of possible outcomes, then provided that the investigator’s prior is suitably spread out over the space of such processes, there is chance 1 that her prior will weakly converge to the probability function that assigns all probability to the true process. (For an explanation of the relevant sort of convergence, see Belot (2013, 489, n. 14, n. 17).) In this sense, there is chance 1 that the investigator will converge to the truth. In contrast, Belot (2013) notes that there is no such assurance of convergence to the truth if the chance process has a countably infinite set of positive-probability outcomes, and argues that in that setting, convergence to the truth is in a certain sense atypical (Belot 2013, 492).

These observations are not—and Belot (2013) does not mean them to be—criticisms of Bayesianism. (Bayesians are in no way committed to thinking that convergence to the truth is typical when sampling from a chance process with infinitely many outcomes.) Rather, the observations are presented “by way of stage setting and because I suspect that the results in question are not as widely known among philosophical Bayesians as they might be” (Belot 2013, 484).

⁸Here I adopt the suggestion from Weatherson (2014) to modify the definition of “open-minded” given in Belot (2013, 496) to introduce a pleasing symmetry. Nothing of substance hinges on this.

more likely.

Formally, we say that a probability function P is *open-minded* with respect to a hypothesis $H \in \mathcal{B}$ if for every string s , there are strings t and t' , each of which is an extension of s , such that $P(H \mid [t]) > 1/2$ and $P(H \mid [t']) < 1/2$. (A string t is said to be an extension of a length- n string s when t is at least as long as s and $s(i) = t(i)$ for $i \in \{1, \dots, n\}$.)

Remark If a hypothesis fails to be dense in \mathcal{C} , or if its complement $\mathcal{C} \setminus H$ fails to be dense, then some initial segment of the observed sequence decisively settles whether H is true. As a result, no probability function is open-minded with respect to such a hypothesis—and hence, no probability function is open-minded with respect to every $H \in \mathcal{B}$. For example, no probability function is open-minded with respect to the universal hypothesis \mathcal{C} , since for any probability function P and any string s , $P(\mathcal{C} \mid [s]) = 1$ whenever that conditional probability is well defined. Similarly, no probability function is open-minded with respect to the hypothesis H_{111} that the sequence contains at least three 1s, since for any probability function P and any extension t of the string 111, $P(H_{111} \mid t) = 1$ whenever that conditional probability is well defined.

In contrast, there are many countably additive probability functions that are open-minded with respect to hypotheses H that are both dense in \mathcal{C} and have dense complements. This is true even if H is uncountable.

For an example of a countably additive prior that is open-minded with respect to a countably infinite hypothesis, see Belot (2013, n. 34, n. 37). For an example of a countably additive prior P that is open-minded with respect to an uncountable hypothesis H , let H be the set $L_{2/3}$ of sequences in which 1 occurs with limiting relative frequency $2/3$. (The limiting relative frequency of 1s in a sequence is defined to be $\lim_{n \rightarrow \infty} n^{-1} \sum_{i=1}^n x(i)$.) Now let P be $(B_{1/3} + B_{2/3})/2$, where B_p is the *Bernoulli measure* with bias p (the countably additive probability function on \mathcal{B} that treats the digits of the sequence as independent random quantities, each with probability p of taking value 1).

To show that P is open-minded with respect to H , let s be any string of length n . We must show that there exist finite extensions t and t' of s so that $P(H \mid [t]) < 1/2$ and $P(H \mid [t']) > 1/2$. We can do that by letting t be the

result of appending many 0s to s , and letting t' be the result of appending many 1s to s .⁹

With the above setup in place, here is the first premise of the orgulity argument:

Premise 1 It is rationally permissible to be open-minded with respect to some hypothesis H .

Now take any Bayesian investigator with a countably additive prior who is open-minded about a hypothesis H , and consider the set T of sequences that get her to converge to the truth about H . Formally, consider a countably additive probability function P that is open-minded with respect to H , and let $T = \{x \in \mathcal{C} : \mathbf{1}_H(x) = \lim_{i \rightarrow \infty} P(H \mid x^i)\}$.

We noted above that convergence theorems entail that this investigator must assign probability 1 to T . We will now see that T is in one sense a “tiny” set.

Start by defining *the Banach-Mazur game*. In this game two players generate an infinite binary sequence together, starting with the empty sequence. The players alternate moves; at each move a player extends the sequence by appending whatever finite block of digits she wishes. The goal of the player who moves second is to have the resulting infinite sequence fall outside of some fixed set G .¹⁰

G is said to be **meager** if there exists a winning strategy for the second player in this game—in other words, if the second player can force the generated

⁹The details: Let t be the result of appending $2n$ 0s to s , and let t' be the result of appending $2n$ 1s to s . We now have that:

$$P(H \mid [t]) < 1/2 \iff P(H \cap [t])/P([t]) < 1/2 \tag{1}$$

$$\iff P(H \cap [t]) < (1/2)P([t])$$

$$\iff (1/2)(B_{1/3}(H \cap [t]) + B_{2/3}(H \cap [t])) < (1/2)(1/2)(B_{1/3}([t]) + B_{2/3}([t]))$$

$$\iff B_{2/3}([t]) < (1/2)(B_{1/3}([t]) + B_{2/3}([t])) \tag{2}$$

$$\iff B_{2/3}([t]) < B_{1/3}([t]), \tag{3}$$

where (1) follows from the definition of conditional probability because $P([t]) > 0$, and (2) holds because by the strong law of large numbers, $B_{1/3}(H) = 0$ (hence $B_{1/3}(H \cap [t]) = 0$) and $B_{2/3}(H) = 1$ (hence $B_{2/3}(H \cap [t]) = B_{2/3}([t])$). The right-hand side of (3) is true because the proportion of 1s in t is closer to $1/3$ than it is to $2/3$. Dual reasoning shows that $P(H \mid [t']) > 1/2 \iff B_{2/3}([t']) > B_{1/3}([t'])$, the right hand side of which is true because the proportion of 1s in t' is closer to $2/3$ than it is to $1/3$.

¹⁰This is actually the special case of the Banach-Mazur game appropriate to the present context. For a general discussion, see Oxtoby (1980).

sequence to *avoid* G . When a set of sequences is meager, it is “tiny” in one sense—it is *easy to avoid*. (This is just one of several equivalent characterizations of the meager sets.) It is sometimes said that sequences “typically” have a property if the set of sequences that fail to have the property is meager.

The following fact, which is a consequence of Belot’s mathematical observations, drives the orgulity argument:

Fact (Belot 2013, 498–499) Consider a Bayesian investigator who is open-minded with respect to a hypothesis. The set of sequences that get the investigator to converge to the truth about the hypothesis is meager. In other words: “typical” sequences *prevent* the investigator from converging to the truth about the hypothesis. Formally: for every countably additive probability function P on \mathcal{B} and hypothesis $H \in \mathcal{B}$, if P is open-minded with respect to H , then T is meager. (Recall that T is the set of sequences for which P converges to the truth about H .)

Proof In the Banach-Mazur game, player 2 can force the generated sequence to be one that prevents the investigator from converging to the truth about H by at each of her turns appending “a string of bits that causes P to [assign probability greater than $1/2$ to H] followed by a string of bits that causes P to [assign probability less than $1/2$ to H].” (Belot 2013, n. 41) Player 2 can implement this strategy because P is open-minded.

Since we can see the truth of this fact, so can a reasonable open-minded investigator. She can see that typical sequences prevent her from converging to the truth about H . Given this, it seems permissible for her to have less than full confidence that she will converge to the truth about H . That is the next premise of the argument:

Premise 2 If it is rationally permissible to be open-minded about a hypothesis, then it is rationally permissible to have less than full confidence that one will converge to the truth about that hypothesis.

From Premises 1 and 2 we get the conclusion of the orgulity argument:

Conclusion It is rationally permissible to have less than full confidence that one will converge to the truth about some hypothesis.

We saw in §2 that countably additive Bayesianism entails the negation of this conclusion. So if the argument is sound, then that theory stands refuted.

4 *The Premises of the Orgulity Argument*

Before advocating a reaction to the orgulity argument, let me address the plausibility of its premises.

A defender of countably additive Bayesianism might try to reject Premise 1 by claiming that open-mindedness is always irrational. That proposal is unappealing because (in the presence of countable additivity), it entails:

MaxCon For every hypothesis and every rational cautious investigator, there is some finite sequence of evidence that would get the investigator to become *maximally confident* about that hypothesis.

In the above statement, to say that an investigator is *cautious* is to say that no finite sequence of digits would get her to be maximally confident about whether the next digit is 0 or 1. And to say that an investigator is *maximally confident* about a claim is to say that her probability for that claim is either 0 or 1. For a proof that (assuming countable additivity) denying Premise 1 entails MaxCon, see Appendix A.

To see why MaxCon is false, consider an example. Suppose that H is the claim that the sequence contains infinitely many 0s. MaxCon entails that if a Bayesian investigator is cautious and rational, some finite sequence of digits would get her to assign probability 0 or 1 to H . But that is absurd. It is absurd that rationality in every case requires cautious investigators to count some finite string of digits as decisively settling whether the whole string contains infinitely many 0s.

So Premise 1 cannot reasonably be resisted.

What about Premise 2? Belot (2013, 500) considers an opponent who rejects Premise 2 for the following reason: sequences of evidence digits that prevent an open-minded investigator from converging to the truth are skeptical scenarios, and the investigator may therefore reasonably assign them total probability zero. Belot responds that such sequences are not skeptical scenarios. Whatever one thinks of that response, however, an additional response is available: Even granting that the scenarios in question are skeptical scenarios, it does not immediately follow that they deserve zero probability.

As an example, consider a regularity that is very well confirmed: that gravity is an attractive. Here is a skeptical scenario: one year from now, gravity will suddenly turn repulsive. Given our evidence, that scenario deserves only a

miniscule amount of probability. But that the scenario is a skeptical one does not immediately show that it deserves absolutely *no* probability. So simply calling failure-to-converge scenarios “skeptical scenarios” does not on its own make it reasonable to reject Premise 2.

Furthermore, the Fact gives at least some initial support to Premise 2. It is unsettling to think that to be rational one must have full confidence that one will converge to the truth, given that “typical” sequences prevent one from doing so.

That said, there is at least one objection to Premise 2 that is worth taking seriously. One might follow de Finetti (1970, 34–35) and argue that there is no reason to treat as unlikely or epistemologically negligible those hypotheses that are “small” from a topological or set-theoretic point of view. In the present case, one might affirm that open-minded agents should be fully confident that they will converge to the truth. When it is pointed out that this means assigning probability zero to a property of sequences that is “typical” in the topological sense, one might simply reply: “So what? That notion of typicality has no relevance to this case.”

What would advance the orgulity argument against this objection is an independent reason for thinking that topological notions of size are relevant in the present case.¹¹

Bottom line: the orgulity argument has some force as an objection to countably additive Bayesianism, though there is at least one defensible line of resistance to its second premise.

5 *Finitely Additive Bayesianism Permits Humility*

Happily, the orgulity argument has no force at all against finitely additive Bayesianism, a version of Bayesianism that rejects countable additivity as an across-the-board requirement of rationality.¹² That is because *finitely additive Bayesians*—those Bayesians who reject countable additivity as a requirement of rationality—can comfortably accept the conclusion of the argument. They can accept that it is rationally permissible for an open-minded investigator in the sequence situation to be less than fully confident that she will converge to the

¹¹Thanks here to an anonymous reviewer, whose comments I draw on.

¹²Juhl and Kelly (1994, 185–188) and Howson and Urbach (2006, 28–29) make similar points in response to the concern that the long-run convergence theorems yield implausibly strong constraints on rationality.

truth.¹³

Indeed, they can (if they wish) accept something much stronger. Let us say that an investigator in the sequence situation is *completely pessimistic* if she is fully confident that she will “converge to the false” — that her probability for H will converge to 0 if H is true and to 1 otherwise. It turns out that some open-minded investigators with finitely additive priors are completely pessimistic. (For a proof, see Appendix B.)

It follows that finitely additive Bayesians are free to count even complete pessimism as being rationally permissible. That is as much humility as anyone can demand. Furthermore, finitely additive Bayesianism has significant independent appeal.^{14,15}

Moral: the orgulity argument has no force against finitely additive Bayesianism, a viable alternative to countably additive Bayesianism.

6 *Long-Term Convergence Theorems do Not Address the Charge of Excessive Subjectivity*

Recall from §1 the motive given for appealing to long-run convergence theorems: The deliverances of Bayesianism depend on a choice of prior probability function. As a result, Bayesianism faces the charge of being excessively subjective, and of not sufficiently explaining agreement among reasonable scientists.

In response to those charges, it is tempting to appeal to long-run convergence

¹³Note that for an important class of finitely additive probability functions—“strategic” functions (Dubins and Savage 1965, Chapter 2; Purves and Sudderth 1976, §2)—generalizations exist for the convergence result described in §2 (Purves and Sudderth 1983, Theorem 1; Chen 1977; Zabell 2002, §2; Seidenfeld 1985, Appendix). So to in general avoid the convergence conclusions, it is not enough to merely give up countable additivity as a requirement of rationality. One must hold that it is rationally permissible to have a probability function that fails to be strategic (and thereby fails to be coherent according to the notion of coherence proposed in Lane and Sudderth (1985)—for further discussion see Zabell 2002, §3). Thanks here to an anonymous reviewer.

¹⁴See, for example, de Finetti (1970), Savage (1954) and Levi (1980). Works that at least take very seriously the hypothesis that countable additivity should be rejected include Seidenfeld and Schervish (1983), Dubins and Savage (1965), Kelly (1996), and Juhl and Kelly (1994).

¹⁵Of course there are objections to rejecting countable additivity as well; an assessment of the costs and benefits of doing so is beyond the scope of this discussion. Such an assessment would need to address concerns about susceptibility to infinite Dutch Books (Bartha 2004; Seidenfeld and Schervish 1983), the possibility of paradoxical-seeming failures of conglomerability (Kadane et al. 1986; Schervish et al. 1984), the possibility of uniform distributions over countably infinite spaces (de Finetti 1970, 122), violations of intuitive comparative dominance principles (Easwaran 2013), as well as considerations of general mathematical utility (Dubins and Savage 1965).

theorems in order to show that differences between rational priors eventually disappear.¹⁶ Bayesians who reject countable additivity in response to the orgulity argument cannot appeal to those theorems for that purpose, since those theorems assume countable additivity.¹⁷ So it might seem that such Bayesians give up a valuable defense against charges of excessive subjectivity.

But in fact, they do not. For the long-run convergence theorems are red herrings. The theorems provide Bayesians with no defense against charges of excess subjectivity. To see why, let me distinguish two versions of the charge of excess subjectivity, and explain why long-run convergence theorems do not address either of them.

First let us pin down a target. Let *Core Bayesianism* be the conjunction of the following claims:

1. Ideally rational agents have personal probability functions which represent their degrees of belief.
2. Conditionalization on new evidence is a rational way of updating one's probability function.

Now for the first version of the charge, posed in characteristically trenchant fashion in Earman (1992, 137):

if in the face of currently available evidence you assign a high degree of belief to the propositions that Velikovsky's *Worlds in Collision* scenario is basically correct, that there are canals on Mars, that the earth is flat, etc., you will rightly be labeled as having an irrational belief system. And if you arrived at your present beliefs within the framework of Bayesian personalism, then the temptation is to say that at worst there is something rotten at the core of Bayesian personalism and at best there is an essential incompleteness in its account of procedural rationality.

¹⁶§2 describes only one particularly simple long-run convergence theorem. Other notable long-run convergence theorems are discussed in Blackwell and Dubins (1962) and Schervish and Seidenfeld (1990, §3).

¹⁷As described in n. 13, versions of the convergence theorems exist that replace the assumption of countable additivity with the strictly weaker assumption of being a strategic measure. However, such theorems are of no use to Bayesians who wish to avoid the orgulity argument by rejecting countable additivity. For the measures that are modest about whether they will converge to the truth must of course fail to be strategic.

The argument is this: (1) starting with a bizarre prior and conditionalizing on available evidence will result in irrational opinions, but (2) Core Bayesianism does not entail that those opinions are irrational, and (3) a true and complete theory of scientific inquiry would entail this, so (4) Core Bayesianism is false or incomplete.

This argument—call it *the problem of deviant priors*—is a fair challenge. Various replies are possible. One might, for example, embrace pure subjectivism and admit that it can be rational to believe that the Earth is flat, based on current evidence. Or one might adopt an objective Bayesianism of one kind or another, and endorse or seek constraints on priors that rule out the deviant ones as irrational.

However, in no case does it help the Bayesian to appeal to long-term convergence theorems. For the most such theorems might show is that in the limit of infinite inquiry, the opinions of a scientist starting with a deviant prior would likely to approach those of an ordinary scientist. And that conclusion would do nothing to answer the problem of deviant priors, which concerns a single time (now) at which an agent has opinions alleged to be irrational (Earman 1992, 148).

A less extreme challenge to Bayesianism is to explain the significant amount of agreement among actual scientists. One might pointedly ask the Bayesian: assuming the truth of Bayesianism, what explains scientific consensus about the basic structure of matter, the rough characteristics of the solar system, the boiling points of various liquids, the mechanism of photosynthesis, and so on? Is it merely a coincidence that the scientific method often gets scientists to rapidly agree? Call this *the problem of scientific agreement*.

In answering this problem, long-term convergence theorems are again of no help. For again, the most they could hope to show is that in the limit of infinite inquiry, scientists would likely approach agreement. That would not touch the question of why scientist have reached so much agreement already (Howson and Urbach 2006, 238).

7 *Short-Term Convergence Theorems Help Address the Problem of Scientific Agreement*

The previous section argued that long-term convergence theorems do not help the Bayesian answer charges of excessive subjectivity. One might wonder: do *any* convergence theorems help the Bayesian address those charges?

When it comes to the problem of deviant priors, the answer is “no”. All

convergence theorems place some constraints on priors, and opponents will always be able to ask about priors that violate the constraints.

But when it comes to the problem of scientific agreement, the answer is “at least a bit”. The remainder of this section describes some *short-run* convergence theorems, and explains how they partly address the problem of scientific agreement.

As a toy example of such a theorem, consider two Bayesian agents who are about to observe what they both regard to be independent random draws from an urn containing 100 red and green balls in an unknown proportion. Straightforward calculations show that unless the agents start out with radically different opinions about what the proportion is, they will be confident that their opinions about the urn’s composition will become extremely similar—not just in the limit of infinite draws, but *soon* (after a small number of draws). This result holds because the assumption of independent sampling is so powerful: it is easy for a small number of samples to vastly confirm one hypothesis about the composition of the urn over another. As a result, it is easy for initial differences of opinion about that composition to get swamped.

This toy theorem addresses a toy instance of the problem of scientific agreement. It explains why in the urn setup, unless scientists have extreme differences in priors, observing even a small number of samples from the urn will bring them into significant agreement. That gives the Bayesian a satisfying explanation of why agreement in such cases is so common.

The above story can be generalized to many situations in which scientists start with similar statistical models, and use repeated experiments to estimate the value of certain model parameter. For example, suppose that a particular team all agrees on the likely behavior of an apparatus designed to measure the speed of light, conditional on hypotheses about that speed. Then the above story might explain why those team members rapidly come to agree about the speed of light, as they observe the behavior of the apparatus. (For short-run convergence results for cases of roughly this kind, see Savage (1954, §3.6), Edwards et al. (1963, 541–545), Earman (1992, 142–143), Howson and Urbach (2006, 239), and Hawthorne (2014, §4.1).)

That shows that short-run convergence theorems are of at least some help in addressing the problem of scientific agreement. But there is more work to be done. For in most cases, scientists do not start with similar statistical models. What then? Here the Bayesian may appeal to short-run convergence results that

rely on weaker assumptions.

I won't address this strategy in detail, but will briefly say why it is promising.

Start by adapting the setup from Hawthorne (1993, Theorem 6): consider ideal Bayesian agents who initially disagree about a finite set of hypotheses. Suppose that the agents are exposed to a common stream of observations relevant to those hypotheses. Crucially, do *not* assume that the agents regard successive observations as independent random draws. Instead impose a much weaker condition: that the agents expect successive observations to be, on average, at least slightly informative about the hypotheses in question.¹⁸

Given these conditions, we may pick one hypothesis H and ask the following question. Supposing that H is true, how many observations would it take to make it at least 99% likely that the observations confirm H over its rivals to some given degree? Remarkably, the proof of Hawthorne (1993, Theorem 6) supplies answers to such questions.¹⁹ And the answers to such questions can be parlayed into explicit bounds on how fast we may expect the agents to converge on the truth, given the truth of each competing hypothesis. If convergence is sufficiently fast in some particular domain, that would provide a satisfying Bayesian explanation for scientific agreement in that domain. It remains for future work to fill out the details of this approach.

The bottom line is that short-run convergence theorems provide a worked-out and convincing answer to the problem of scientific agreement in certain special cases, and there is reason to think that this answer can be extended to include many additional cases.

Happily, none of the short-run convergence theorems mentioned above rely on countable additivity. So finitely additive Bayesians can freely appeal to them.

¹⁸I give here only the barest sketch of the wonderful Hawthorne (1993, Thm. 6), which is explained in a simplified form in Hawthorne (2014, §5).

¹⁹The answer depends on such quantities as the degree to which the agents expect the observations to be informative about the hypotheses, for a particular technical notion of informativeness defined in Hawthorne (1993, §3.2).

*Appendix**A Proof That Countably Additive, Cautious Investigators Are Either Decisive or Extremely Open-Minded*

In §4 it is claimed that under the assumption of countable additivity, the falsity of Premise 1 entails MaxCon. Here we prove a slightly stronger claim from which the above claim easily follows.

Say that an investigator with prior P is **decisive** with respect to $H \in \mathcal{B}$ if for some string s , $P(H \mid [s])$ equals 0 or 1.

Say that an investigator with prior P is **extremely open-minded** with respect to H if for every $\epsilon > 0$ and for every evidence string, there is a finite extension of that evidence that would lead her to assign probability less than ϵ to H , and also a finite extension of it that would lead her to assign probability greater than $1 - \epsilon$ to H .

Recall that an investigator with prior P is said to be **cautious** if for no string s , $P([s0] \mid [s]) \in \{0, 1\}$, where $s0$ denotes the concatenation of s and the string 0.

Note that a cautious investigator assigns strictly positive probability to every string. For suppose that a probability function P assigns probability 0 to some string. Then there must be strings s and sd such that sd is the result of appending a single digit d to s , and $P([s]) > 0$ while $P([sd]) = 0$. (Recall that the length-0 sequence z counts as a string, and so $P([z]) = 1$.) So $P([sd] \mid [s]) = 0$, hence $P([s0] \mid [s]) \in \{0, 1\}$, and an investigator with prior P is not cautious.

Now for the main claim:

Claim: Suppose that an investigator is cautious and has a countably additive prior P , and let H be any member of \mathcal{B} . Then the investigator is decisive with respect to H , or extremely open-minded with respect to H , or both.

Proof: Consider a cautious investigator with countably additive prior P who is not decisive with respect to H . We will show that she is extremely open-minded with respect to H .

Take any string s . Note that $P([s]) > 0$ (since the investigator is cautious) and $0 < P(H \mid [s]) < 1$ (since she is not decisive about H). Let $P'(\cdot) = P(\cdot \mid [s])$ be the result of conditionalizing P on $[s]$. For any $p \in \mathbb{R}$, let M_p be the set of

sequences x for which $\lim_{i \rightarrow \infty} P'(H \mid x^i) = p$. By the long-run convergence theorem described in §2,

$$P'(P' \text{ converges to the truth about } H) = 1.$$

So $P'((H \cap M_1) \cup (\bar{H} \cap M_0)) = P'(H \cap M_1) + P'(\bar{H} \cap M_0) = 1$. But $1 > P(H \mid [s]) = P'(H) \geq P'(H \cap M_1)$, so $0 < P'(\bar{H} \cap M_0) \leq P'(M_0)$. So M_0 is nonempty and hence there exists a sequence x such that $\lim_{i \rightarrow \infty} P'(H \mid x^i) = 0$. So for any $\epsilon > 0$ there exists an n such that $P'(H \mid x^n) < \epsilon$. It follows that for any $\epsilon > 0$, there is a finite extension s' of s such that $P(H \mid [s']) < \epsilon$. A similar argument shows that for any $\epsilon > 0$ there is a finite extension s' of s such that $P(H \mid [s']) > 1 - \epsilon$. So the investigator is extremely open-minded. ■

B Proof of The Existence of an Open-Minded, Completely Pessimistic Finitely Additive Probability Function:

In the following definitions, p ranges over reals in the open unit interval $(0, 1)$ and i ranges over \mathbb{N} .

Let L_p denote the set of infinite binary sequences whose limiting relative frequency of 1 equals p :

$$L_p = \left\{ x \in \mathcal{C} : p = \lim_{n \rightarrow \infty} n^{-1} \sum_{i=1}^n x(i) \right\}.$$

Given the formal setup of this paper, the *Bernoulli measure with bias p* is the unique countably additive probability function B_p on \mathcal{B} such that for each $x \in \mathcal{C}$ and each $n \in \mathbb{N}$:

$$B_p(x^n) = p^{\sum_{k=1}^n x(k)} \cdot (1 - p)^{n - \sum_{k=1}^n x(k)}.$$

Thus B_p treats the unknown digits of the successive terms of the true but unknown infinite binary sequence as independent and identically distributed binary random quantities, each with probability p of taking the value 1 (and probability $1 - p$ of taking the value 0). Observe that by the Strong Law of Large Numbers it follows that $B_p(L_p) = 1$.

In the following, the *Bernoulli flip-flopper with bias p up to trial i* is the unique countably additive probability function B_p^i on \mathcal{B} such that for each $x \in \mathcal{C}$ and

$n \in \mathbb{N}$:

$$B_p^i(x^n) = B_p(x^{\min(i,n)}) \cdot B_{1-p}(x^n \mid x^i).$$

Thus B_p treats the first i digits of x as independent and identically distributed binary random quantities, each with probability p of taking the value 1, and any remaining digits as independent and identically distributed binary random quantities, each with probability $1 - p$ of taking the value 1. Observe that since B_p^i applies B_p only up to trial i and thereafter applies B_{1-p} , it follows from the Strong Law of Large Numbers that $B_p^i(L_{1-p}) = 1$.²⁰

The proof to follow depends on the notion of a Banach limit, which is introduced here in a way that closely follows Rao and Rao (1983, 39-40). Let l_∞ be the space of bounded sequences of reals. A Banach limit $T : l_\infty \rightarrow \mathbb{R}$ is a non-negative linear functional on l_∞ such that $T(\langle 1, 1, \dots \rangle) = 1$ and for any $\langle y_1, y_2, \dots \rangle \in l_\infty$, $T(\langle y_1, y_2, y_3, \dots \rangle) = T(\langle y_2, y_3, y_4, \dots \rangle)$.

It can be shown that for any Banach limit T and any $y \in l_\infty$, if $\lim_{i \rightarrow \infty} y(i)$ exists, then $T(y) = \lim_{i \rightarrow \infty} y(i)$. For that reason, one may think of a Banach limit as generalizing the notion of the limit of a sequence of reals. When emphasizing this connection, we sometimes write $\text{blim}_{i=1}^\infty f(i)$ for a Banach limit of the sequence $\langle f(1), f(2), \dots \rangle$.

Banach limits are not unique, and we are guaranteed of their existence only nonconstructively, by way of the Axiom of Choice or a weaker axiom such as the Ultrafilter Lemma. Dependence on Banach limits is what makes the proof below nonconstructive.²¹

Claim: There exists an open-minded, completely pessimistic finitely additive probability function. That is, there exists an open-minded finitely additive probability function P on \mathcal{B} and a hypothesis $H \in \mathcal{B}$ such that

$$P(\{x : \mathbf{1}_H(x) = \lim_{i \rightarrow \infty} P(H \mid x^i)\}) = 0.$$

Proof: Let blim be a Banach limit. Now define P_0 and P_1 as follows. For any $H \in \mathcal{B}$, let $P_0(H) = \text{blim}_{i=1}^\infty B_p^i(H)$, and let $P_1(H) = \text{blim}_{i=1}^\infty B_{1-p}^i(H)$.

It is easy to check that P_0 and P_1 are finitely additive probability measures. For example, whenever H and H' are disjoint sets of sequences, $P_0(H \cup$

²⁰Thanks here to an anonymous reviewer, whose suggested exposition I adopt almost verbatim.

²¹For proofs that merely finitely additive probability functions can sometimes expect with probability 1 to “converge to the false” about a proposition, and results concerning when this phenomenon can arise, see Dubins (1975) and Schervish et al. (1984, Thms 3.1, 3.3). Thanks here to Teddy Seidenfeld.

$$H') = \text{blim}_{i=1}^{\infty} B_9^i(H \cup H') = \text{blim}_{i=1}^{\infty} (B_9^i(H) + B_9^i(H')) = \text{blim}_{i=1}^{\infty} B_9^i(H) + \text{blim}_{i=1}^{\infty} B_9^i(H') = P_0(H) + P_0(H').$$

(Informally, we can think of P_0 and P_1 in the following way: P_0 treats the sequence as if a large initial segment of it is generated by tosses of a coin biased toward “1”, and the rest by a coin biased toward “0”. P_1 treats the sequence in exactly the opposite way. In each case, the initial segment is expected to be extremely long, in the following sense: for every n , however large, P_0 and P_1 treat the first n digits as if they are part of the initial segment. That is what forces P_0 and P_1 to be merely finitely additive.)

Let $P = (P_0 + P_1)/2$. P is clearly a finitely additive probability function. We will now complete the proof by showing that P is open-minded and completely pessimistic with respect to the hypothesis L_9 . Note that for any $x \in L_9$ and for any i ,

$$P(L_9 \mid x^i) = \frac{P(L_9 \cap x^i)}{P(x^i)} = \frac{(1/2)(P_0(L_9 \cap x^i) + P_1(L_9 \cap x^i))}{(1/2)(P_0(x^i) + P_1(x^i))} \quad (4)$$

$$= \frac{(1/2)(0 + P_1(x^i))}{(1/2)(P_0(x^i) + P_1(x^i))} \quad (5)$$

$$= \frac{P_1(x^i)}{P_0(x^i) + P_1(x^i)} = \frac{1}{1 + P_0(x^i)/P_1(x^i)} \quad (6)$$

$$= \frac{1}{1 + B_9(x^i)/B_1(x^i)}. \quad (7)$$

(4) holds by definition. (5) holds because $P_1(L_9) = 1$ and $P_0(L_9) = 0$, since for each i , $B_1^i(L_9) = 1$ and $B_9^i(L_9) = 0$ by the strong law of large numbers. (6) is simple algebra. (7) holds because for any binary sequence x and any natural number i , $P_0(x^i) = B_9(x^i)$ and $P_1(x^i) = B_1(x^i)$. To see why, note that $P_0(x^i) = \text{blim}_{j=1}^{\infty} B_9^j(x^i) = \text{blim}_{j=i}^{\infty} B_9^j(x^i) = \text{blim}_{j=i}^{\infty} B_9(x^i) = B_9(x^i)$.²²

Now consider what happens to (7) as i approaches infinity: The proportion of 1s in the first i digits of x approaches .9 (since $x \in L_9$). As a result, $B_9(x^i)/B_1(x^i)$ grows without bound, and hence (7) approaches 0. So when $x \in L_9$, $\lim_{i \rightarrow \infty} P(L_9 \mid x^i) = 0$. A similar argument shows that when $x \in L_1$, $\lim_{i \rightarrow \infty} P(L_1 \mid x^i) = 0$.

It follows that P is open-minded, since for any initial segment of digits,

²²Note that even though P was defined by way of Banach limits, it is indeed a probability function. As a result, we are free to calculate using conditional probabilities as usual, even though the product of two Banach limits need not equal the Banach limit of the corresponding product.

appending a large enough finite block consisting of 90% 1s will force P to assign a probability to L_9 that is arbitrarily close to 1, and appending a large enough finite block consisting of 90% 0s will force P to assign a probability to L_9 that is arbitrarily close to 0.

It also follows that P is completely pessimistic, since $P(L_9 \cup L_{.1}) = 1$, and the above argument shows that P converges to the wrong verdict about L_9 for any sequence in $L_9 \cup L_{.1}$. ■

References

- Paul Bartha. Countable additivity and the de Finetti lottery. *British Journal for the Philosophy of Science*, 55(2):301–321, 2004.
- Gordon Belot. Bayesian orgulity. *Philosophy of Science*, 80(4):483–503, October 2013.
- David Blackwell and Lester Dubins. Merging of opinions with increasing information. *The Annals of Mathematical Statistics*, 33(3):882–886, 1962. ISSN 00034851. URL <http://www.jstor.org/stable/2237864>.
- Alan F. Chalmers. *What Is This Thing Called Science?* Hackett Publishing Company, Indianapolis, 3rd edition edition, 1999.
- Robert Chen. On almost sure convergence in a finitely additive setting. *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete*, 37(4):341–356, 1977. ISSN 0044-3719. doi: 10.1007/BF00533425. URL <http://dx.doi.org/10.1007/BF00533425>.
- Bruno de Finetti. *Theory of probability: a critical introductory treatment*, volume 1. Wiley, New York, 1970. Translated by A. Machi and A. Smith.
- L.E. Dubins and L.J. Savage. *How to gamble if you must: inequalities for stochastic processes*. McGraw-Hill series in probability and statistics. McGraw-Hill, 1965.
- Lester E. Dubins. Finitely additive conditional probabilities, conglomerability and disintegrations. *The Annals of Probability*, 3(1):89–99, 1975.
- John Earman. *Bayes or Bust?* MIT Press, Cambridge, Massachusetts, 1992.
- Kenny Easwaran. Bayesianism II: Applications and criticisms. *Philosophy Compass*, 6(5):321–332, 2011. ISSN 1747-9991. doi: 10.1111/j.1747-9991.2011.00398.x. URL <http://dx.doi.org/10.1111/j.1747-9991.2011.00398.x>.
- Kenny Easwaran. Why countable additivity? *Thought*, 2:53–61, 2013. doi: 10.1002/tht.60.
- W. Edwards, H. Lindman, and L. J. Savage. Bayesian statistical inference for psychological research. *Psychological Review*, 70(3):193–242, 1963.

- P.R. Halmos. *Measure Theory*. Number 9 in Graduate Texts in Mathematics. Springer, 1974.
- James Hawthorne. Bayesian induction is eliminative induction. *Philosophical Topics*, 21(1):99–138, 1993.
- James Hawthorne. Inductive logic. In Edward N. Zalta, editor, *The Stanford Encyclopedia of Philosophy*. The Stanford Encyclopedia of Philosophy, summer 2014 edition, 2014.
- Colin Howson and Peter Urbach. *Scientific reasoning: The Bayesian approach*. Open Court, third edition edition, 2006.
- Cory Juhl and Kevin T. Kelly. Realism, convergence, and additivity. *PSA: Proceedings of the Biennial Meeting of the Philosophy of Science Association*, 1994: 181–189, 1994.
- Joseph B. Kadane, Mark J. Schervish, and Teddy Seidenfeld. Statistical implications of finitely additive probability. In *Bayesian inference and decision techniques*, chapter 5, pages 59–76. Elsevier Science Publishers, 1986.
- Kevin T. Kelly. *The Logic of Reliable Inquiry*. Logic and computation in philosophy. Oxford University Press, 1996.
- David A. Lane and William D. Sudderth. Coherent predictions are strategic. *The Annals of Statistics*, 13(3):pp. 1244–1248, 1985. ISSN 00905364. URL <http://www.jstor.org/stable/2241138>.
- Isaac Levi. *The enterprise of knowledge*. MIT Press, Cambridge, 1980.
- Paul Lévy. *Théorie de l'Addition des Variables Aléatoires*. Gauthiers-Villars, Paris, 1937.
- J.C. Oxtoby. *Measure and Category: A Survey of the Analogies between Topological and Measure Spaces*. Graduate Texts in Mathematics. Springer, 1980.
- Roger A. Purves and William D. Sudderth. Some finitely additive probability. *The Annals of Probability*, 4(2):259–276, 04 1976. doi: 10.1214/aop/1176996133. URL <http://dx.doi.org/10.1214/aop/1176996133>.

- Roger A. Purves and William D. Sudderth. Finitely additive zero-one laws. *Sankhyā: The Indian Journal of Statistics, Series A (1961-2002)*, 45(1):pp. 32–37, 1983. ISSN 0581572X. URL <http://www.jstor.org/stable/25050411>.
- K.P.S.B. Rao and B. Rao. *Theory of Charges: A Study of Finitely Additive Measures*. Pure and Applied Mathematics. Elsevier Science, 1983.
- Leonard Savage. *The foundations of statistics*. John Wiley and Sons, 1954.
- M. Schervish, T. Seidenfeld, and J. Kadane. The extent of non-conglomerability of finitely additive probabilities. *Z. f. Wahrscheinlichkeitstheorie*, 66:205–226, 1984.
- Mark Schervish and Teddy Seidenfeld. An approach to consensus and certainty with increasing evidence. *Journal of Statistical Planning and Inference*, 25(3): 401–414, July 1990.
- Teddy Seidenfeld. Calibration, coherence, and scoring rules. *Philosophy of Science*, 52(2):pp. 274–294, 1985. ISSN 00318248. URL <http://www.jstor.org/stable/187511>.
- Teddy Seidenfeld and Mark J. Schervish. A conflict between finite additivity and avoiding Dutch Book. *Philosophy of Science*, 50:398–412, 1983.
- Aron Vallinder. Solomonoff induction: A solution to the problem of the priors? Master’s thesis, Lund University, September 2012.
- Brian Weatherson. Belot on Bayesian orgulity. Manuscript, April 2014. URL <http://brian.weatherson.org/papers.html>.
- S. L. Zabell. It all adds up: The dynamic coherence of radical probabilism. *Proceedings of the Philosophy of Science Association*, 2002(3):98–103, 2002.