

ARTICLE

# Beyond finite additivity

Colin Howson<sup>†</sup>

Professor of Philosophy at the University of Toronto

**Corresponding author:** Peter Urbach; Email: [pmurbach@aol.com](mailto:pmurbach@aol.com)

(Received 13 August 2019; revised 27 October 2019; accepted 31 March 2020; first published online 02 August 2024)

*This article was written by Colin Howson and submitted to the journal before his death. Peter Urbach graciously agreed to review and assist in the final publication of this article, and to serve as corresponding author during these final stages.*

## Abstract

There is a Dutch Book argument for the axiom of countable additivity for subjective probability functions, but de Finetti famously rejected the axiom, arguing that it wrongly renders a uniform distribution impermissible over a countably infinite lottery. Dubins however showed that rejecting countable additivity has a strongly paradoxical consequence that a much weaker rule than countable additivity blocks. I argue that this rule, which also prohibits the de Finetti lottery, has powerful independent support in a desirable closure principle. I leave it as an open question whether countable additivity should be adopted.

## 1. Introduction

Kolmogorov's continuity axiom (his Axiom V), equivalent with the other axioms to the rule of countable additivity, lies at the heart of modern mathematical probability theory, indispensable for the proofs of the celebrated "with probability one" theorems (the most familiar of which is probably the strong law of large numbers). These theorems not only figure among the great achievements of twentieth-century pure mathematics but also have been central to many ground-breaking applications in physics, particularly ergodic theory, as well to the well-known strong "convergence of opinion" theorems of subjective Bayesianism.<sup>1</sup> But one of the great architects of the subjective Bayesian theory, Bruno de Finetti,<sup>2</sup> claimed, on the basis of a very powerful result that he proved,<sup>3</sup> that finitely additive probability grounds sufficiently strong versions of the "strong" theorems based as they are on repeated independent trials, and he held in addition that because countable additivity forbids a countably infinite

---

<sup>†</sup> Deceased

<sup>1</sup> For an extended discussion of these see Earman 1992.

<sup>2</sup> I have added Ramsey and Savage to my personal personalist pantheon.

<sup>3</sup> The de Finetti Exchangeability Theorem.

lottery from being modeled by a uniform distribution (such a lottery is now called a de Finetti lottery), just as a finite lottery is, countable additivity should be rejected as a general rule.<sup>4</sup>

But we shall see that admitting the de Finetti lottery has a paradoxical consequence, a fact admitted by de Finetti though he attempted to dismiss concern about it. I will argue that the concern is fully justified because the consequence does represent a genuine inconsistency. Though it is blocked by countable additivity, this by itself is not sufficient to justify the latter's adoption because the consequence in question is blocked by a much weaker rule having, I will claim, strong independent support. All this in due course; first it will be useful to review some facts about coherence and countable additivity in the context of the de Finetti lottery.

## 2. The de Finetti lottery

Countable additivity prohibits the de Finetti lottery because each ticket would have to have probability zero and a countable sum of zeros is zero. But de Finetti argued that if a uniform distribution over a finite partition, and a uniform probability density over a bounded interval in  $\mathbb{R}^n$  are permissible, then this should be the case also for a countably infinite partition; otherwise, he pointed out, we are forced by what poses as a "purely formal" axiom to demand a heavily biased probability distribution in which a finite number of outcomes must receive nearly all the probability: "What is strange is simply that a formal axiom, instead of being *neutral* with respect to the evaluations ... imposes constraints of the above kind" (1974, 122).

De Finetti's case for the permissibility of the de Finetti lottery has some plausibility, but it seems to be in clear conflict with his famous no-Dutch Book criterion of consistency (coherence) for any assignment of subjective probabilities considered as normalized fair betting odds.<sup>5</sup> As he pointed out, his infinite lottery is Dutch Bookable. Anyone offering odds of zero on each ticket would be forced to lose one dollar by an opponent staking a dollar on each ticket winning; That person would win a dollar on the ticket drawn and lose nothing on the others.<sup>6</sup>

However, as de Finetti pointed out, Dutch Book arguments depend on the assumption that a sum of bets fair to the agent is fair to the agent. He called this "the hypothesis of rigidity with respect to risk" (*ibid.*, 82), and noted that because of the phenomenon of risk-aversion it is strictly *false* (*ibid.*, 74). Nevertheless, according to him it is nevertheless an acceptable approximation to a rigorous utility-based argument for the finitely additive axioms and the multiplication rule because for these a sum of at most three bets is needed and the betting is assumed to be for money-sums small enough to be roughly linear in utility. For countable additivity,

<sup>4</sup> I should emphasise that de Finetti rejected countable additivity only as rule to be applied in all cases; he certainly did not reject it where he thought it useful to employ it.

<sup>5</sup> He proved that this criterion of consistency for fair betting quotients is equivalent to your truth-value estimates not being able to be dominated (penalty uniformly reduced) according to the so-called Brier scoring rule (1974, 88–89). He came to regard this second criterion as preferable because the former is vulnerable to contamination by strategic considerations.

<sup>6</sup> It is straightforward to show that in conjunction with the other axioms, and subject to the condition that only a finite sum changes hands, countable additivity is a necessary and sufficient condition for invulnerability to a Dutch Book.

however, as we have seen an infinite number of bets is required, and de Finetti saw the countable Dutch Book argument therefore as begging the question that rigidity extends to the infinite case (1972, 91). Indeed, utility-based theories like Savage's do not extend to providing a justification for countable additivity; for that, a separate continuity assumption is required (e.g., as in Villegas 1964). Nor do the well-known accuracy arguments for the finitely additive axioms appear to extend to countable additivity without such additional assumptions (Pettigrew 2016, 222).

Accordingly, de Finetti restricted the definition of coherence to finite sums of bets: A set of subjective probability assignments is coherent just in case no *finite* subset is vulnerable to a Dutch Book. Because the opponent in the Dutch Book against the de Finetti lottery must make bets on each of the infinite number of tickets to ensure a win, the uniform-0 distribution is, according that definition, coherent despite being Dutch Bookable. De Finetti's resolution is certainly formally adequate: The restricted definition of coherence permits the uniform-0 distribution and is sufficient for Dutch-Book-argument proofs of all the finitely additive probability axioms and the multiplication theorem, but not the continuity axiom. As we shall see in the following section, however, merely finitely additive probability functions seem to come at a considerable price.

### 3. More paradox

In a personal communication to de Finetti, the probabilist and mathematician Lester Dubins showed that when regarded as a hypothesis about a data source the de Finetti lottery generates a bizarre posterior probability distribution (de Finetti 1972, 105). In Dubins's example<sup>7</sup> a number  $n \in \mathbb{N}$  is to be announced (on a screen, or by some other method). There are two propositions A and B, each having prior probability  $\frac{1}{2}$ , such that your probability of getting the number  $n$  conditional on A is uniformly 0, while on B it is  $2^{-(n+1)}$ . So A states that the data source generates a de Finetti lottery. Let  $X_n$  state that the number  $n$  is the number displayed (we can regard  $X_i$  as abbreviating the statement that a random variable X, defined on the sample space N with  $X(m) \equiv m$ , takes the value  $i$ ). A straightforward Bayes's Theorem calculation yields  $P(A|X_n) = 0$  and  $P(B|X_n) = 1$  for every  $n$  (the conditional probabilities are all well-defined since by the theorem of total probability  $P(X_i)$  is positive and equal to  $P(X_i|A)P(A) + P(X_i|B)P(B) = 0 + 2^{-1} \cdot 2^{-(i+1)} = 2^{-(i+2)}$ ). It follows that before observing the value of X you will know for certain in advance that the probability of A will be 0 and that of B will be 1 after you conditionalize.<sup>8</sup>

This certainly sounds paradoxical; Kadane et al. call it "reasoning to a foregone conclusion" (1996). De Finetti conceded its paradoxical appearance, but commented that nevertheless:

<sup>7</sup> I am taking N here to be  $\{1, 2, 3, \dots\}$ .

<sup>8</sup> This is a striking example of the phenomenon of *nonconglomerability*. A probability function is conglomerable in a partition  $\{C_i: i \in \mathbb{N}\}$  just in case for every D in the domain of P, if for every  $i \in \mathbb{N}$ ,  $P(D|C_i)$  lies in any given interval, then so does  $P(D)$ . It turns out that, subject to certain quite general conditions, every merely finitely additive probability function whose range is infinite is nonconglomerable in some countably infinite partition (Schervish et al. 1984). By contrast, it is straightforward to show that countably additive probabilities are conglomerable in every countable partition.

[a]ll these surprises are but the inevitable unforeseen complications met in every field when we pass from the finite to the infinite, and they are called paradoxical only until we become accustomed to them. (1972, 106)

This is surely disingenuous. The Dubins example is more than just a “complication”: Acknowledging *ex ante* that after conditionalizing on whatever result you observe you will be certain that A is false and that B is true, being at the same time maximally uncertain of A (prior =  $\frac{1}{2}$ ), may not be an outright contradiction but it is nevertheless—at any rate intuitively—extremely close to one.<sup>9</sup> De Finetti’s own gloss of  $P(E|H)$  as “the probability that You attribute to E if You think that in addition to Your present information . . . it will become known to You that H is true (and nothing else)” (1974, 134; emphasis in the original) seems to commit him to the claim that you should now regard 0 and 1 as the probabilities of A and B, respectively, though he follows with the caveat that the gloss is only “a preliminary guide to the meaning of . . .  $P(E|H)$  [and we] ought to warn the reader . . . against an overhasty acceptance of these initial explanations” (ibid., 135). Nevertheless, he attempted to bypass the objection by advancing a rule for betting on A or B given any  $X_i$ , in fact an infinite family of rules parametrized by  $i$ , each recommending betting on B if the observed value is no greater than  $i$ , and on A if greater. He pointed out that the expected gain from this strategy increases with  $i$  and has no upper bound (1972, 105–6). While that betting strategy is sound enough, it simply brushes under the carpet the apparent—I believe real—inconsistency in the prior assignments  $P(A) = P(B) = \frac{1}{2}$  (or indeed any nonextreme values) and the conditional probabilities  $P(A|X_i) \equiv 0$  and  $P(B|X_i) \equiv 1$ .

In what follows, I will show that the Dubins assignments are inconsistent with a principle, of ancient pedigree, that certainty is carried through deduction from premises to conclusions. Granted it, we shall see that not only are the Dubins conditional assignments ruled inadmissible but also the de Finetti lottery.

#### 4. Closing off certainty

The finitely additive probability axioms harmonise nicely with first-order logic (FOL). It is a consequence of first-order completeness that if A is a logical consequence of a set G of sentences in FOL, then A is a consequence of some finite conjunction  $G_m$  of members of G. It is also easily proved by induction that for every  $n \in \mathbb{N}$ , if  $n$  propositions in the domain of a finitely additive probability function each has probability one so does their conjunction. Hence, if each member of G has probability one so does  $G_m$ . It also follows from the finitely additive probability axioms that if B entails C, then  $P(B) \leq P(C)$ . Hence, we conclude that if each member of G has probability one then so does A. Thus, we seem to have a formal corroboration of the traditional belief that deducibility carries certainty—at any rate probability-1 certainty—with it (in this case the strongest form of certainty short of deductive) from even infinitely many premises to conclusions.

But probably the majority of calculations in mathematical probability involve infinitary propositions and their consequences. The de Finetti lottery is a particularly simple example defined in an infinitary probability system. In the

<sup>9</sup> Compare Uffink: “If it is certain beforehand that a probability value will be revised downward, this value must have been too high to start with, and could not have been a faithful representation of our opinion” (1996, 68).

usual  $\sigma$ -algebra formalism, the outcome-space is  $N$ , and the algebra, or field, of events is some appropriate subfield of the power set of  $N$ , including possibly the power set.<sup>10</sup> The disjunction over the possible exclusive outcomes of the lottery is represented by the event  $\cup_{i \in N} X_i \cap \hat{E}p_{i \in N}(\sim X_i \cup \sim \cup_{j \neq i \in N} X_j)$  (I am taking  $\sim$  as complement) where the  $X_i$  are defined as in section 2. Probability theory came of age in dealing with infinitary events, of which the celebrated “with probability one” theorems are the outstanding exemplars. The simplest and best-known of these is the Strong Law of Large Numbers, attributing probability one to a formula of the form  $\hat{E}p_{i \in N} \cup_{j \in N} \hat{E}p_{k \in N} A(i, j, k)$ , an infinite conjunction followed by an infinite disjunction followed by a formula free in  $i, j$ , and  $k$ . In the 1930s Kolmogorov’s famous monograph laying the measure-theoretic foundations of modern probability theory made  $\sigma$ -fields, the loci of such infinitary formulas, the typical context of investigation.

But within the finitely additive probability calculus there is no corresponding theorem that probability one is inherited by the logical consequences of such infinitary propositions; in fact, it is not true. At this point we face the possible objection that the relation of logical consequence is not defined for any logic incorporating such propositions. The objection is unfounded. In fact, formal logics with that property have been extensively studied since the 1950s, and one merely extends the formalism of FOL to accommodate countably infinite conjunctions and disjunctions, in much the same way that a field of sets is extended to a  $\sigma$ -field. This is the so-called  $L_{\omega_1, \omega_0}$  family, where the ordinal  $\omega_1$  signifies countable conjunctions and disjunctions, and  $\omega_0$  finite quantifier strings;  $L_{\omega_0, \omega_0}$  is of course just FOL.<sup>11</sup> Now the de Finetti lottery can be represented by (among equivalents) the sentence  $\forall x(x = a) \wedge \bigvee_{i \in \omega} B_i(a) \wedge \bigwedge_{i \in \omega} (B_i(a) \rightarrow \bigwedge_{j \neq i \in \omega} \sim B_j(a))$  (this says that the individual domain has just one outcome—the draw—and the  $B_i$  are its denumerably many possible ticket numbers).

The only distinctive rule of proof for  $L_{\omega_1, \omega_0}$  is an  $\omega$  rule permitting  $\bigwedge_{i \in \omega} A_i$  to be inferred from  $A_1, A_2, \dots$ , so its proofs can be infinite (though countable) and may have a complex ordinal structure. Though there is a completeness theorem, for countable sets of assumptions, compactness fails, and the proof-predicate is hardly effective in the usual recursively enumerable sense like that of FOL. Thus there is a powerful lobby including Boolos and Shapiro who because of its strong categoricity properties believe that full second order logic<sup>12</sup> (SOL) is fundamental, yet SOL-validity is as far from being effective as it is possible to be: Its set of codes of valid sentences is not even in the analytic hierarchy, let alone the arithmetical.<sup>13</sup> Although the

<sup>10</sup> Kadane and O’Hagan (1995) show that there is a variety of ways for the uniform distribution to be achieved; for example, there is a “random” measure on the entire power set of  $N$  giving “even” the probability 1/2. The events in the subjective theory are better described as propositions: They are ascribed truth-values and are the objects of an individual’s partial beliefs. The truth-value of  $A$  for any  $s$  in the outcome-space is given by the indicator function (random variable)  $I_A(s)$ , taking the value 1 if  $s$  makes  $A$  true and 0 if  $s$  makes  $A$  false (though 1 and 0 are purely conventional values).

<sup>11</sup> For more details of this and other infinitary languages/logics see Bell (2000). Note that the standard model of arithmetic can be characterized up to isomorphism in  $L_{\omega_1, \omega_0}$ .

<sup>12</sup> “Full” means that the domain of the  $n$ -place relation variables is the power set of  $D^n$ , where  $D$  is the domain of the individual variables.

<sup>13</sup> This does not of course mean that one can’t prove an extensive class of SOL sentences true: We can for example prove the second-order quantifier rules corresponding to the first-order ones

semantics of SOL may at first blush seem perspicuous, it is often criticized for its intimate relation to axioms of set theory that are much less transparently “logical” than contentiously set-theoretical: For example, there is a sentence of SOL that is valid just in case the continuum hypothesis is true. By contrast  $L_{\omega^1, \omega^0}$  is much less vulnerable, though admittedly vulnerability is a question of degree. Even FOL is not wholly innocent of association with powerful set-theoretical principles: Its metatheory appeals to *all* set-theoretic structures (what are they?), truth in a model is third-order (McGee, 2000, 73), while a theorem of Trakhtenbrot (1950) implies that FOL-completeness requires the Axiom of Infinity (note that  $\omega$  is a strongly inaccessible cardinal). Dana Scott, who established several of the major results for infinitary logic wrote that he

feels that he has justified the contention that  $L_{\omega^1, \omega^0}$  is the proper generalization of  $L_{\omega^0, \omega^0}$  to denumerable formulas. In fact we have seen several reasons for claiming that  $L_{\omega^1, \omega^0}$  plays the same role for  $L_{\omega^0, \omega^0}$  that the theory of Borel sets and  $\sigma$ -fields plays for the ordinary fields of sets. (1965, 341)

Scott’s observation does however tend to hide an important point, which is that the theory of Borel sets and  $\sigma$ -fields functions, within the ambient set theory, as its own logical structure. We do not need the extra elaboration of a full logical language because everything we want to say in the infinitary case, being essentially propositional, can be said in that formalism, and it contains the relation of logical consequence between its own propositions in the relation of set-theoretical inclusion: A is a consequence of B just in case  $B \subseteq A$ .

That granted, we now return to the problem of extending the finitely additive axioms so that probability one is closed under consequence. It is easy to see that countable additivity suffices because if each of the members of a countably infinite set Q has probability one then a simple consequence of the Kolmogorov Continuity axiom is that so does  $\hat{E}p_{i \in \mathbb{N}}Q$ , and if A is a consequence of Q then it is a consequence of  $\hat{E}p_{i \in \mathbb{N}}Q$ , so  $1 = P(\hat{E}p_{i \in \mathbb{N}}Q) \leq P(A)$ . But countable additivity is an unnecessarily strong axiom for this purpose. Clearly, all we need is a strict consequence of the equivalent continuity axiom: *If each of the members of a countably infinite set Q has probability one then so does  $\hat{E}p_{i \in \mathbb{N}}Q$ .*<sup>14</sup> I will call this the C-minus rule (“Continuity minus”). It is also easy to see that it is the *weakest* rule having the property that probability one is closed under logical consequence from a countable set of premises.

The word “countable” is important. With FOL we need no extra rule about the cardinality of the premisses because the completeness theorem says that independently of cardinality, if C is a consequence and if the premisses have probability one then so does C. But what about infinitary rules in non-FOL languages? Indeed, even in the  $\sigma$ -algebras of ordinary probability theory we seem to face a problem. If for uncountably

---

$(\forall X(A \rightarrow B) \rightarrow (\forall X A \rightarrow \forall X B))$ , etc., where X is a second order variable). One simply can’t prove *all* the valid statements from some recursively enumerable set of logical axioms.

<sup>14</sup> It is strictly weaker than the Continuity axiom because there is a strictly positive probability measure on any infinite separable Boolean algebra that is not countably additive (Horn and Tarski, 1948: Theorem 3.10; this uses the Axiom of Choice. A separable Boolean algebra is one in which every element joins a dense set; the power set algebra  $\mathbb{N}$  is separable). Trivially the limit of the probabilities of the sequence  $\langle 1, 1, 1, \dots \rangle$  is the limit of the sequence.

many  $A_\xi$ ,  $P(A_\xi) = 1$  implies that  $P(\hat{E}_p; A_\xi)$  is equal to 1, that would spell the end of continuous distributions, and indeed of mathematical probability. Many if not most of the important spaces that one meets in measure theory are isomorphic to Lebesgue measure in  $R^n$ , or else on  $[0,1]$  (in ergodic theory, for example, they are often the completion of the closure of the cylinder sets of  $R^N$ , or what comes to the same thing  $2^N$ <sup>15</sup>), and these contain all the singletons  $\{x\}$  over whichever uncountable set of reals we are considering. Any continuous distribution over these each has measure 0 while the whole space has probability 1. But the continuous density functions are merely mathematical devices for computing the probabilities of intervals in  $R^n$ , and of the measurable sets they generate, down to the degenerate intervals  $\{x\}$ . The rationals are of course countable. The irrationals are certainly uncountable, but nobody measures them, even in physics, except up to a finite point in any sequence in their decimal (or other) expansion, in other words up to some strictly nonempty interval. Any such interval can be refined without limit, but it is always nonempty, and no interval can be partitioned into more than countably many nondegenerate subintervals.

So we have now identified a window, between uncountable conjunctions of probability one statements, for which there should be no closure of consequence, and FOL, which fails to identify the fault at the heart of the Dubins paradox, which is that consequence should be closed under probability one. *That window is filled by C-minus.* I will consider later whether there are grounds for adopting countable additivity but we will now see how the C-minus rule prohibits the anomalous distribution in the Dubins example.

### 5. Dubins’s paradox—lost

The notation will be the same as in the earlier discussion, but to speed things up I will change  $P(A|X_i) \equiv 0$  into the equivalent  $P(\sim A|X_i) \equiv 1$ . A straightforward calculation shows that  $P(\sim A|X_i) \leq P(X_i \rightarrow \sim A)$ , where in terms of the basic Boolean operations  $X_i \rightarrow \sim A$  is just  $\sim X_i \cup \sim A$ . Thus we infer that each of the countably infinitely many propositions  $X_i \rightarrow \sim A$  has probability 1. By the C-minus rule we now infer that  $\hat{E}_{p; i \in N} (X_i \rightarrow \sim A)$  has probability 1. However,  $\hat{E}_{p; i \in N} (X_i \rightarrow \sim A) \leftrightarrow (\cup_{j \in N} X_j) \rightarrow \sim A$ , so the latter has probability 1. But  $\cup_{j \in N} X_j$  is necessarily true, so  $P(\sim A) = 1$  and  $P(A) = 0$ . Hence, given C-minus,  $P(A) = \frac{1}{2}$  is probabilistically inconsistent with  $P(A|X_i) \equiv 0$  – which, intuitively, is as it should be.

That is not all that is inconsistent with the C-minus rule: *Clearly, so is the de Finetti lottery.* The statements “1 does not win,” “2 does not win,” . . . , “n does not win,” . . . all have probability one in that lottery and so, given the C-minus rule, does their conjunction. But one of those numbers must win, and that statement too has probability one. But I think there is no reason to mourn the de Finetti lottery: It has the property that *whichever number you pick, however large, the probability is one that the number of the winning ticket will be greater*, or to put it another way, every initial segment of  $N$  has probability zero except the supremum, which has probability one. One might conceivably try to defend this with the reasoning of the so-called Wang’s

<sup>15</sup> That  $R^N = 2^N$  often surprises people unused to transfinite cardinal arithmetic because  $R$  is not merely infinite but uncountable.

Paradox: *Every positive integer is small*, where the proof is by induction (because if  $n$  is small, so presumably is  $n+1$ ). I leave the reader to judge how convincing that is.

An alternative policy mooted to refute the Dubins paradox is to reject the rule conditionalization that determines the new probabilities  $Q(A)$  and  $Q(B)$ , a strategy that even de Finetti might be thought to be questioning in his own remarks, quoted earlier, about optimal betting strategies depending on the value of  $i$  in  $X_i$ <sup>16</sup>: In this way some posterior distribution depending on  $i$  could be chosen and the  $Q(A)$ ,  $Q(B)$  simply disregarded. The argument for jettisoning conditionalization depends on an allegedly further coherence condition called a Reflection Principle, stating that your current probability of  $A$ , conditional on your future probability  $Q(A)$  of  $A$  being  $p$ , should be  $p$ . That is the form in which the principle was introduced by van Fraassen (1984), but it is reasonably clear that to acquire normative force it should be supplemented by the additional condition that  $Q$  be obtained by an “epistemically sound” method (it is not necessary for the argument to define this any more precisely, if indeed it can be so defined). So amended, the principle seems sound. But (so the argument proceeds) the assumption that conditionalization is an epistemically sound means of reaching  $Q(A)$  from the prior  $P(A) = 1/2$  given any observation report, results in an assignment incompatible with that prior. For let us assume you have done the conditionalization calculation and know that  $Q(A) = 1$ . So you know with certainty what your conditionalization-updated probability of  $A$  will be. But a probability conditioned on a proposition with probability 1 equals the unconditional probability, that is the prior probability. By reflection, therefore, that prior probability must be 1. But it is surely absurd to demand that, independently of all the information you used to generate the priors of  $A$  and  $B$ , they *must* be 1 and 0, respectively. That granted the argument seems nothing less than a *reductio* of the assumption that, at any rate in the context of this example, conditionalization is an epistemically sound method of updating priors (Howson, 2014).

This solution seems a strategy of despair: Conditionalization is so deeply rooted in Bayesian—methodology—it is the method by which we learn from experience according to that philosophy, turning Bayesianism from an academic exercise in probability into what its advocates claim is the foundation of all inductive inference—that to abandon it is tantamount to abandoning Bayesianism. But fortunately, there is no need to impugn conditionalization. There is another reason for placing the Dubins example under suspicion, and that is the de Finetti lottery: Remove it, with the assistance of  $C$ -minus, and the problem vanishes. The question then is whether we should add countable additivity to the list of desirable items.

## 6. Conclusion

The most unwavering and trenchant opposition to continuity was of course that of de Finetti, citing the alleged bias inherent in the de Finetti lottery in its front-loading of the probabilities. It was also a point seized on by Wenmackers and Horsten, who like de Finetti reject countable additivity precisely because it violates the “intuition [that] fairness is absolutely central to our concept of a lottery,” including the lottery over

<sup>16</sup> I pointed out that de Finetti in effect presented an infinite family of posterior distributions yielding superior expected gains by comparison with those generated by conditionalizing, when judged from the prior standpoint.



N (2013, 41).<sup>17</sup> If, as I think we should, we deny assumed fairness a central, or indeed any, role in deciding the issue of countable additivity, what else is there? Countable additivity does of course prohibit nonconglomerability not only in the case of Dubin's problem but in general. But prohibiting nonconglomerability simply because it might seem counterintuitive is not itself a justification: Other things have seemed highly counterintuitive that have been admitted into the canon because they have been accompanied by principles and methods almost universally agreed to have advanced if not revolutionised an entire discipline (like Lebesgue measure). Moreover, Schervish et al. (1984) have shown that every probability function admitting uncountable partitions is nonconglomerable in at least one partition (de Finetti had earlier cited the Borel paradox as a case in point [1972, 204]).

But there are of course other considerations. There is, as we have seen, a Dutch Book argument for countable additivity; admittedly its dependence on extending the assumption of rigidity to infinite combinations of bets might render it questionable to some, but possibly little if any more so than the Dutch Book arguments that remain for many people the staple of the finitely additive axioms. And, of course, there is the rich trove of "with probability one" theorems that continuity/countable additivity generates. While the distinction between "objective" and "subjective" is certainly overcrude (there is a multiplicity of subdivisions within each), it is probably appropriate to place the scientific applications within the former. Here, it seems that countable additivity has been adopted practically without question because of what is seen as an indispensable role within highly successful explanatory models forming the foundations of modern physical science. On the purely subjective side there are, of course, the probability-one convergence-of-opinion results, but to cite these as *evidence* for countable additivity would clearly be to beg the question.

But that C-minus should be added I am much more confident about. The failure to close off probability one under countable consequence seems to me entirely arbitrary, particularly given that the infinitary formulas of standard  $\sigma$ -algebras have been for well over a century the probabilist's stock in trade. If, however, we do go further and adopt the Axiom of Continuity, then perhaps we should do so in the same cautious spirit as Kolmogorov, the proposer of the axiom, when he said only that it "has been found expedient in researches of the most diverse sort" (1956, 15).

## References

- Bell, John L. 2000. "Infinitary Logic." In the *Stanford Encyclopedia of Philosophy*, edited by Edward N. Zalta. <https://plato.stanford.edu/entries/logic-infinitary/>
- Bernstein, Allen R., and Frank Wattenberg. 1969. "Nonstandard Measure Theory." In *Applications of Model Theory to Algebra, Analysis and Probability*, edited by Wilhelmus A. J. Luxemburg, 171-86. Holt, Rinehart & Winston.
- de Finetti, Bruno. 1972. *Probability, Induction and Statistics*. Wiley.

<sup>17</sup> Wenmakers and Horsten, however, also reject de Finetti's uniform-0 distribution because it violates the condition they call SUM: The probability of a combination of tickets is the sum of the individual probabilities (2013, 40). Instead, they assign a uniform positive infinitesimal distribution to the singleton members of N that adds hyperfinitely to 1 (Bernstein and Wattenberg [1969] showed that a uniform infinitesimal probability distribution can be defined over the unit interval, considered as an *uncountable* lottery, also summing to 1). But as Pruss (2012) points out, assigning infinitesimals as values rather than reals merely reproduces "up to an infinitesimal" the anomalous Dubins assignment.

- de Finetti, Bruno. 1974. *Theory of Probability*, vol. 1. Wiley.
- de Finetti, Bruno. 1980 [1937]. "Foresight: Its Logical Laws, Its Subjective Sources"; translated from the French and reprinted in *Studies in Subjective Probability*, edited by Kyburg and Smokler, 53–11.
- Earman, John. 1992. *Bayes or Bust? A Critical Examination of Bayesian Confirmation Theory*. Cambridge, MA: MIT Press.
- Horn, Alfred, and Alfred Tarski. 1948. "Measures in Boolean Algebras." *Transactions of the American Mathematical Society* 64:467–97.
- Howson, Colin. 2014. "Finite Additivity, Another Lottery Paradox and Conditionalisation." *Synthese* 191: 989–1012.
- Kadane, Joseph B., and Anthony O'Hagan. 1995. "Using Finitely Additive Probabilities: Uniform Distributions on the Natural Numbers." *Journal of the American Statistical Association* 90:626–31.
- Kadane, Joseph B., Mark J. Schervish, and Teddy Seidenfeld. 1996. "Reasoning to a Foregone Conclusion." *Journal of the American Statistical Association* 91:1228–35.
- Kolmogorov, A. N. 1956 [1933]. *Foundations of the Theory of Probability*. New York: Chelsea Publishing Company. (English translation of *Grundbegriffe der Wahrscheinlichkeitsrechnung*, 1933).
- Kyburg, Henry E., and Howard E. Smokler, eds. 1980. *Studies in Subjective Probability*, 2nd ed. Wiley.
- McGee, Vann. 2000. "Everything." In *Between Logic and Intuition: Essays in Honor of Charles Parsons*, edited by Gila Sher and Richard Tieszen, 54–79. Cambridge University Press.
- Pettigrew, Richard. 2016. *Accuracy and the Laws of Credence*. Oxford University Press.
- Pruss, Alexander R. 2012. "Infinite Lotteries, Perfectly Thin Darts and Infinitesimals." *Thought* 1:81–89.
- Schervish, Mark J., Teddy Seidenfeld, and Joseph B. Kadane. 1984. "The Extent of Non-Conglomerability of Finitely Additive Probabilities." *Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete* 66:205–26.
- Scott, Dana. 1965. "Logic with Denumerably Long Formulas and Finite Strings of Quantifiers." In *Symposium on the Theory of Models*, edited by J.W. Addison, L. Henkin, and A. Tarski, 329–41. North Holland.
- Trakhtenbrot, Boris A. 1950. "The Impossibility of an Algorithm for the Decidability Problem for Finite Classes." *Proceedings of the USSR Academy of Sciences* 70:569–72. (in Russian).
- Uffink, Jos. 1996. "The Constraint Rule of the Maximum Entropy Principle." *Studies in History and Philosophy of Modern Physics* 27B:47–81.
- van Fraassen, Bas C. 1984. "Belief and the Will." *Journal of Philosophy* 81:235–56.
- Villegas, C. 1964. "On Qualitative Probability  $\sigma$ -Algebras." *Annals of Mathematical Statistics* 35:1787–96.
- Wenmackers, Sylvia, and Leon Horsten. 2013. "Fair Infinite Lotteries." *Synthese* 190:37–61.