Beyond finite additivity

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 which a weaker rule than countable additivity blocks. I argue that this rule, which also prohibits the de Finetti lottery itself, has powerful independent support in a desirable closure principle. I leave it as an open question whether countable additivity itself should be adopted.

Keywords. Countable additivity; de Finetti lottery; probability one; logical consequence.

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 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.[[1]](#footnote-1) But one of the great architects of the subjective Bayesian theory, Bruno de Finetti, claimed, on the basis of a very powerful result that he himself proved , 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 since countable additivity forbids a countably infinite lottery from being modelled 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.

But we shall see that admitting the de Finetti lottery has a paradoxical consequence, a fact admitted by de Finetti himself though he attempted to dismiss concern about it. I will argue that the concern is fully justified since 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 since 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 since 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 Rn 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 normalised fair betting odds.[[2]](#footnote-2) As he himself pointed out, his infinite lottery is Dutch Bookable. Anyone offering odds of 0 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.[[3]](#footnote-3)

However, as de Finetti himself pointed out, Dutch Book arguments depend on the assumption that a sum of bets fair to the agent is itself fair to the agent. He called this ‘the hypothesis of rigidity with respect to risk’ (1974, 82), and noted that because of the phenomenon of risk-aversion it is strictly *false* (1974, 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, since 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, on the other hand, 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 (for example 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. Since the opponent in the Dutch Book against the de Finetti lottery has to 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 a number n∈N[[4]](#footnote-4) is to be announced (on a screen, or by some other method). There are two propositions A and B, each having prior probability ½, 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 Xn state that the number n is the number displayed (we can regard Xi as abbreviating the statement that a random variable X, defined on the sample space N with X(m) ≡ m, takes the value i). A straightforward Bayes’s Theorem calculation yields P(A|Xn) = 0 and P(B|Xn) = 1 for every n (the conditional probabilities are all well-defined since by the theorem of total probability P(Xi) is positive and equal to P(Xi|A)P(A) + P(Xi|B)P(B) = 0 + 2-1.2-(i+1) = 2-(i+2)). It follows that before observing the value of X you will know for certain that the posterior probability of A will be 0 and that of B will be 1.[[5]](#footnote-5)

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:

[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 conditionalising on whatever result you observe you will be certain that A is false and that B is true, while being *at the same time* maximally uncertain of A (prior = ½), may not be an outright contradiction but it is nevertheless – at any rate intuitively – extremely close to one.[[6]](#footnote-6) 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’ (1974, 135). Nevertheless he attempted to bypass the objection by advancing a rule for betting on A or B given any Xi, in fact an infinite family of rules parametrised 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, 105-106). 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) = ½ (or indeed any non-extreme values) and the conditional probabilities P(A|Xi) ≡ 0 and P(B|Xi) ≡ 1.

In what follows, I will show that the Dubins assignments are actually 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 itself.

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 Gm of members of G. It is also easily proved by induction that for every n∈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 Gm. It also follows from the finitely additive probability axioms that if B entails C then P(B) ≤ 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 usual σ-algebra formalism, the outcome-space is N, and the algebra, or field, of events of the power set of N. [[7]](#footnote-7) The disjunction over the possible exclusive outcomes of the lottery is represented by the event ⋃i∈N Xi ∩ ⋂i∈N(~Xi ∪ ~⋃j≠i∈N Xj) (I am taking ~ as complement) where the Xi are defined as in section 2 above. 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 ⋂i∈N⋃j∈N⋂k∈N A(i,j,k), an infinite conjunction followed by an infinite disjunction followed by an infinite conjunction 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 σ-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*. Indeed, it is not true. At this point, however, 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 nineteen-fifties, and one in particular 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 σ-field. This is the so-called Lω1,ω0 family, where the ordinal ω1 signifies countable conjunctions and disjunctions, and ω0 finite quantifier strings; Lω0,ω0 is of course just FOL.[[8]](#footnote-8) Now the de Finetti lottery can be represented by (among equivalents) the sentence ∀x(x = a) ∧ \/i∈ω Bi(a) ∧ /\i∈ω (Bi(a) → /\j≠i∈ω ~Bj(a)) (this says that the individual domain has just one outcome – the draw - and the Bi are its denumerably many possible ticket numbers).

The only distinctive rule of proof for Lω1,ω0 is an ω rule permitting /\Ai∈ω to be inferred from A1, A2, … , 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 (r.e.) sense like that of FOL. Indeed, validity for Lω1,ω0 is in the analytic hierarchy (it is a Σ12 subset of the reals; Barwise 1970, 321). Of course, not everyone believes that effectiveness in the r.e. sense is necessary: thus there is a powerful lobby including Boolos and Shapiro who, because of its strong categoricity properties, believe that full second order logic[[9]](#footnote-9) (SOL) is fundamental, yet SOL-validity is as far from being effective as it is possible to be.[[10]](#footnote-10) Although the semantics of SOL may at first blush seem perspicuous, it is often criticized for its intimate relation to axioms of set theory which are much less transparently ‘logical’ than contentiously set-theoretical: for example, there is a sentence of SOL which is valid just in case the continuum hypothesis is true. By contrast Lω1,ω0 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 ω is a strongly inaccessible cardinal). Dana Scott, who established several of the major results for infinitary logic (including the important Scott Undefinabilty Theorem) wrote that he

feels that he has justified the contention that Lω1,ω0 is the proper generalization of Lω0,ω0 to denumerable formulas. In fact we have seen several reasons for claiming that Lω1,ω0 plays the same role for Lω0,ω0 that the theory of Borel sets and σ-fields plays for the ordinary fields of sets. (1965, 341).

Scott’s observation does tend to hide an important point, which is that the theory of Borel sets and σ-fields functions serve, within the ambient set theory, as a logical structure in its own right. We do not need the extra elaboration of a full logical language, since 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 ⊆ 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, since 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 ⋂i∈NQ, and if A is a consequence of Q then it is a consequence of ⋂i∈NQ, so 1 = P(⋂i∈NQ) ≤ P(A).[[11]](#footnote-11) 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 ⋂i∈NQ*. 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 premises have probability one then so does C. But what about infinitary rules in non-FOL languages? Indeed, even in the σ-algebras of ordinary probability theory we seem to face a problem. If for *uncountably* many Aξ, P(Aξ) = 1 implies that P(⋂Aξ) 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 Rn, and these contain all the singletons {x} over whichever uncountable set of real 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 Rn, and of the measurable sets they generate, down to the degenerate intervals {x}. The rationals are of course countable, and while the irrationals are certainly uncountable, nobody actually 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 non-empty interval. Any such interval can be refined without limit, but it is always non-empty, 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 itself, 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|Xi) ≡ 0 into the equivalent P(~A|Xi) ≡ 1. A straightforward calculation shows that P(~A|Xi) ≤ P(Xi→~A), where in terms of the basic Boolean operations Xi→~A is just ~Xi∪~A. Thus we infer that each of the countably infinitely many propositions Xi→~A has probability 1. By the C-minus rule we now infer that ⋂i∈N (Xi →~A) has probability 1. However, ⋂i∈N (Xi →~A) ⬄ (⋃j∈NXi) →~A, so the latter has probability 1. But ⋃j∈N Xi is necessarily true, so P(~A) = 1 and P(A) = 0. Hence, given C-minus, P(A) = ½ is probabilistically inconsistent with P(A|Xi) ≡ 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 itself*. 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 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.

6. Conclusion

The most unwavering and trenchant opposition to continuity was of course that of de Finetti himself, citing the alleged bias inherent in the de Finetti lottery in its front-loading of the probabilities. It was also a point also 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 N itself (2013, 41)). If, as I think we should, we deny 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 particular 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 which 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, Kadane and Seidenfeld (1984) have shown that every probability function admitting uncountable partitions is nonconglomerable in at least one partition (de Finetti himself had earlier cited the Borel paradox as a case in point (1972, 204)).

But there are 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 more so than the Dutch Book arguments which remain for many people the staple of the finitely additive axioms. And of course there is the rich trove of ‘with probability one’ theorems which continuity/countable additivity generates. 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.

Whether countable additivity should or should not be adopted I leave as an open question. But that C-minus itself 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 σ-algebras have been for well over a century the probabilist’s stock-in-trade. If we do go further and adopt the Axiom of Continuity, then perhaps we should do so in the same cautious spirit as the proposer of the axiom himself, when he said only that it ‘has been found expedient in researches of the most diverse sort’ (1956, 15).

References

Barwise, J. 1970. *Admissible sets and structures*, Springer.

Bell, J.L. 2000. *Infinitary logic*, Stanford Encyclopedia of Philosophy.

Bernstein, A. R., and Wattenberg F. 1969. ‘Nonstandard Measure Theory’, ed. W. A. J. Luxemburg, *Applications of Model Theory to Algebra, Analysis and Probability*, Holt, Rinehart & Winston, 171-186.

de Finetti, B. 1937. ‘Foresight: its logical laws, its subjective sources’; translated from the French and reprinted in Kyburg and Smokler 1980, 53–11.

de Finetti, B. 1972. *Probability, Induction and Statistics*, Wiley.

de Finetti, B. 1974. *Theory of probability,* vol. 1, Wiley.

Earman, J. 1992. *Bayes or Bust? A Critical Examination of Bayesian Confirmation Theory*, MIT Press. \

Kadane, J.B. and O’Hagan, A. 1995. ‘Using finitely additive probabilities: uniform distributions on the natural numbers’, *Journal of the American Statistical Association*, 90, 626-631.

Kadane, J. B., Schervish, M. J. and Seidenfeld, T. 1996. ‘Reasoning to a foregone conclusion’, *Journal of the American Statistical Association*, 91, 1228–1235.

Kolmogorov, A.N. 1956. *Foundations of the Theory of Probability*, Chelsea (English translation of *Grundbegriffe der Wahrscheinlichkeitsrechnung*, 1933).

Kyburg H. and Smokler H. eds. 1980. *Studies in Subjective Probability*, second edition, Wiley.

McGee, V. 2000. ‘Everything’, eds. G. Sher and R. Tieszen, *Between Logic and Intuition: Essays in Honor of Charles Parsons*, Cambridge University Press, 54-79.

Pettigrew, R. 2016. *Accuracy and the laws of credence*, Oxford University Press.

Pruss, A. R. 2012. ‘Infinite lotteries, perfectly thin darts and infinitesimals’, *Thought*, *1*, 81–89.

Schervish, M.J., Seidenfeld, T. and Kadane, J.B. 1984. ‘The extent of non-conglomerability of finitely additive probabilities’, *Zeitschrift für Wahrscheinlichskeitstheorie und verwandte Gebiete*, 66, 205-226.

Scott, D. 1965. ‘Logic with denumerably long formulas and finite strings of quantifiers’, *Symposium on the Theory of Models*, eds. J.W. Addison, L. Henkin and A. Tarski, North Holland, 329-341.

Scott, D. and Krauss, P. 1966. ‘Assigning Probabilities to Logical Formulas’, *Aspects of Inductive Logic*,

eds. J. Hintikka and P. Suppes, North Holland, 219-264.

Trakhtenbrot, B. 1950. ‘The impossibility of an algorithm for the decidability problem for finite classes’ (in Russian), *Proceedings of the USSR Academy of Sciences*, 70, 569-72.

Uffink, J. 1996. ‘The constraint rule of the maximum entropy principle’, *Studies in History and Philosophy*

*of Modern Physics*, 27B, 47–81.

Villegas, C. 1964: 'On Qualitative Probability σ-Algebras', *Annals of Mathematical Statistics*, 35, 1787-96.

Wenmackers, S. and Horsten, L. 2013. ‘Fair infinite lotteries’, *Synthese*, 190, 37-61.

1. For an extended discussion of these see Earman 1992. [↑](#footnote-ref-1)
2. 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 since the former is vulnerable to contamination by strategic considerations. [↑](#footnote-ref-2)
3. 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. [↑](#footnote-ref-3)
4. I am taking N here to be {1, 2, 3, …} [↑](#footnote-ref-4)
5. This is a striking example of the phenomenon of *nonconglomerability*. A probability function is conglomerable in a partition {Ci: i∈N} just in case for every D in the domain of P, if for every i∈N, P(D|Ci) 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. [↑](#footnote-ref-5)
6. 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). [↑](#footnote-ref-6)
7. 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) IA(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). [↑](#footnote-ref-7)
8. For more details of this and other infinitary languages/logics see Bell (2000). The standard model of arithmetic can be characterized up to isomorphism in Lω1,ω0. [↑](#footnote-ref-8)
9. ‘Full’ means that the domain of the n-place relation variables is the power set of Dn, where D is the domain of the individual variables. [↑](#footnote-ref-9)
10. 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 (∀X(A→B) →(∀XA→∀XB)), 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. [↑](#footnote-ref-10)
11. Scott and Krauss (1966) show that a countably additive probability can be defined on the sentences of Lω1,ω0. [↑](#footnote-ref-11)