THE IMPOSSIBILITY OF GENERATING COMPARATIVE PROBABILITIES FROM PRIMITIVE CONDITIONAL PROBABILITIES

ALEXANDER R. PRUSS

Abstract. There are four well-known models of fundamental objective probabilistic reality: classical probability, comparative probability, non-Archimedean probability, and primitive conditional probability. I offer two desiderata for an account of fundamental objective probability, comprehensiveness and non-superfluity. It is plausible that classical probabilities lack comprehensiveness by not capturing some intuitively correct probability comparisons, such as that it is less likely that $0 = 1$ than that a dart randomly thrown at a target will hit the exact center, even though both classically have probability zero. We thus want a comparison between probabilities with a higher resolution than we get from classical probabilities. Comparative and non-Archimedean probabilities have a hope of providing such a comparison, but for known reasons do not appear to satisfy our desiderata. The last approach to this problem is to employ primitive conditional probabilities, such as Popper functions, and then argue that $P(0 = 1 \mid 0 = 1 \text{ or hit center}) = 0 < 1 = P(\text{hit center} \mid 0 = 1 \text{ or hit center})$. But now we have a technical question: How can we reconstruct a probability comparison, ideally satisfying the standard axioms of comparative probability, from a primitive conditional probability? I will prove that, given some plausible assumptions, it is impossible to perform this task: conditional probabilities just do not carry enough information to define a satisfactory comparative probability. The result is that of the models, no one satisfies our two desiderata. We end by briefly considering three paths forward.

1. Introduction

There is a quest to capture fundamental probabilistic phenomena—whether of the epistemic or the physically chancy sort—in a mathematical way. There are four particularly common formal tools for such modeling: classical probability, comparative (also known as qualitative) probability, non-Archimedean (especially hyperreal) probability, and primitive conditional probability. We have two desiderata for a model of probability. First, *comprehensiveness* calls on the model to capture all of the fundamental probabilistic features of reality, and, second, *non-superfluity* asks that the model not include additional information beyond that reality. The *probabilistic reality* of this paper is constituted by the phenomena of chance, stochasticity and objective probability, rather than the credences of limited agents like us, which we know are not modeled by the formal tools except in an idealized way.

I will begin by expanding on the requirement of the desiderata, sketching some easy and/or already known observations on how the four best-known models fare with regard to them, and then move on to discuss the main results of this paper, whose main upshot is that the primitive conditional probability model fails to be comprehensive. Thus we have good reason to think that the primitive conditional probability model is not *the* right model of fundamental probabilistic reality.

At the same time, the easy/known observations give us good reason to think that the classical probability and comparative probability fail comprehensiveness, while the non-Archimedean model fails non-superfluity. Thus, it seems that *none* of the four models is fully satisfactory. This leaves three ways out of the difficulties: a model beyond the four models, a model combining two or more of the four models, or biting the bullet on one of the critiques. I will briefly sketch a promising example of each strategy.

2. The four models and two desiderata

The two desiderata on a model of probabilistic reality aimed at in this investigation are comprehensiveness and non-superfluity. Now, useful models in general do come in a variety of degrees of realism, as the case of physics nicely illustrates. A physical model of a situation that neglects friction will fail to be comprehensive, but can nonetheless be useful for both applications and understanding. And for many purposes a finite and discrete reality can be usefully approximated by an infinitary continuous model: thus, in fluid dynamics, the distribution of pressure is taken to be a function, typically continuous, over an uncountably infinite space, even though in fact the pressure is constituted by the interactions of a finite number of particles. Similarly, a model of probabilistic reality can include more or less structure than the reality (e.g., a discrete random walk can be usefully approximated by a continuous Brownian motion or vice versa) in which case one respectively fails in non-superfluity or comprehensiveness, and yet one can still gain understanding from the model.

But if we really want to *capture* the reality, we want comprehensiveness and non-superfluity. The quest under discussion is not for useful approximations, but to find out something about what the fundamental probabilistic features of reality are or are not. Models of probabilistic reality that lack comprehensiveness or non-superfluity are useful: they give us understanding of phenomena have applications. None of the arguments should be taken to deny this. A reader who is not interested in the more ambitious pursuit of a model that is comprehensive and non-superfluous can stop here—though perhaps some of the technical results can still give some insight into the logical interconnections between the models and desiderata, and the negative results of the paper may move one to abandon the more ambitious quest.

The best-developed model of probabilistic reality is classical probability. An event A is assigned a real number $P(A)$ between 0 and 1, satisfying plausible axioms. This model captures genuine information about probabilistic phenomena, and all the information contained in the model appears to really reflect features of the reality at least in physically chancy situations and in epistemic cases with idealized Bayesian agents (real agents have imprecise probabilities of some sort).

The classical probability model would appear to be superfluous with regard to subjective probabilities for limited agents in the case of continuous distributions. Plausibly only a countable, and likely only finite, number of probability values can serve as a possible credence for an agent like us, since there are only countably many expressions in our language, and presumably only finitely many of them can fit into our minds, while a continuous distribution has uncountably many real-number values. However, with regard to objective probabilities, classical probability appears to satisfy non-superfluity as long as we have continuous distributions in the world, such as spinners, darts thrown at targets, or radioactive decay[1], or in some other situations such as a countably infinite sequence of fair and independent coin tosses.

However, the classical model appears to lack comprehensiveness. If I throw a dart uniformly randomly at a circular target, it seems much more likely that the dart lands somewhere on the horizontal line through the center than that it lands at the very center itself, and if I spin a uniform spinner, it appears more likely, and by a factor of 1.6, that the selected angle is a multiple of five degrees than that it is a multiple of eight degrees. But on a classical probability model, all these events have the same probability, namely zero, as does the seemingly even more improbable contradiction that $0 = 1$. It appears we need something finer-grained to capture all of probabilistic reality.

The second model is comparative or qualitative probabilities. A *partial comparative probability* (cf. [8]) on an algebra $\mathcal{F}$ of events (a set of subsets of a probability space Ω with $\mathcal{F}$ closed under unions and complements) is a reflexive and transitive relation $\precsim$ satisfying:

- Pos: $\varnothing \precsim A$, and
- Add: if $A \cup B$ is disjoint from C, then $A \precsim B$ if and only if $(A \cup C) \precsim (B \cup C)$.

When $A \precsim B$, we say that B is at least as likely as A. To get a *total* comparative probability, we add the totality condition that $A \precsim B$ or $B \precsim A$ for all A and B in $\mathcal{F}$. We write $A \prec B$ for the claim that $A \precsim B$ but not $B \precsim A$, and we write $A \sim B$ for the claim that $A \precsim B$ and $B \precsim A$.[2]

[1] I am grateful to an anonymous reader for this argument.

[2] Note that we are departing from de Finetti's notation [1], who uses $A \sim B$ to mean that there is no comparison between A and B, i.e., (in our notation) neither $A \precsim B$ nor $B \precsim A$.

It is very plausible that all the information conveyed by an appropriate comparative probability model is genuine: we have non-superfluity. Furthermore, unless one specifically assumes the qualitative probabilities to be Archimedean, the "zero-probability" problem for the classical model disappears on the comparative probability model: we can have primitive comparisons between classically zero-probability events.

But the comparative model still seems to omit important information. If I have an unfair coin that favors heads, the comparative probability can capture the fact that heads is more likely than tails, but omits quantitative information as to how much more likely it is, which information was nicely captured by classical probabilities. Indeed, we have good reason to think that if the classical model is non-superfluous, then the comparative model is non-comprehensive, as it leaves out some of the information from the classical model.

Now, granted, given a richer space of events, one can recover some of this information. For instance, if additionally our system contains a hundred-sided die with all sides equally likely, we might say that our unfair coin yielding heads is more likely than getting a number in the range from 1 to 57 on the die but less likely than getting a number in the range 1 to 58. Indeed, given a rich enough space of events, we can express the precise numerical probability using qualitative probabilities [9, 10]. However, although non-hypothetical frequentists will deny this, there surely could be a world with a single unfair coin flip and no other physically chancy events, and in that world, comparative probabilities do not capture all the information about how much more likely heads is than tails. Thus, we have reason to think comparative probabilities are also non-comprehensive in general.

The third model is non-Archimedean probabilities, which are like classical probabilities, except with additivity restricted to finite additivity, and the probabilities ranging over some non-Archimedean field—most commonly a field of hyperreals—that extends the real numbers by including infinitesimals other than zero.[3]

This tool can nicely capture the distinctions between various events that on a classical model have zero probability by assigning them different infinitesimal probabilities. While this model is representatively very rich, and hence likely comprehensive, it has the opposite fault of superfluity: it has been argued that it carries a vast amount of bogus information [5, 2, 14]. The choice of a specific non-Archimedean field does not appear to reflect genuine facts about the probabilistic phenomena, and there are infinitely many possible ways of assigning infinitesimal probabilities within a specific field that do not appear to reflect any differences between probabilities phenomena.

[3]An infinitesimal is an α such that $|\alpha| < 1/n$ for all natural numbers n. A non-Archimedean field will have all the standard relations on the reals—addition, subtraction, multiplication, division, and comparison—but will include "infinite" elements I such that $n < |I|$ for all natural numbers n. The reciprocal of an infinite element is a non-zero infinitesimal, and conversely.

The fourth model, primitive conditional probabilities, carries strictly more information than classical probabilities and strictly less than non-Archimedean probabilities.[4] We can take conditional probabilities $P(A \mid B)$ as primitive assignments of real numbers to pairs of events A and B, satisfying appropriate axioms (typically those of Popper functions [21]), rather than defining them by the classical ratio $P(A \mid B) = P(A \cap B)/P(B)$ which is undefined when $P(B) = 0$.

More precisely, if $\mathcal{F}$ is an algebra of subsets—the events—of a non-empty space Ω, a *conditional probability* or *Popper function* on $\mathcal{F}$ is an assignment of a real number $P(A \mid B)$ to any pair of events such that:

(P1) $P(\cdot \mid B)$ is a finitely additive probability function when $B \neq \varnothing$
(P2) $P(A \cap B \mid C) = P(A \mid C)P(B \mid A \cap C)$
(P3) if $P(A \mid B) = P(B \mid A) = 1$, then $P(C \mid A) = P(C \mid B)$
(P4) $P(A \mid \varnothing) = 1$.

Condition (P4) can be dropped if we restrict $P(A \mid B)$ to being defined when $B \neq \varnothing$, but makes it simpler to state some claims. For convenience, we will write $P(A) = P(A \mid \Omega)$ for the unconditional finitely-additive probability defined by P.

A *full* conditional probability is one where $\mathcal{F}$ is the powerset of Ω.

It is fairly plausible that all the information conveyed by a good primitive conditional probability model is genuine: differences in conditional probabilities really do reflect differences in probabilistic phenomena. Thus, the non-superfluity of the primitive conditional probability model is quite defensible.

Moreover, a primitive probability model is capable of using conditional probabilities to zoom in on zero-probability events and compare them. For instance, we can say that it is more likely that the dart uniformly aimed at the circular target will hit the horizontal line rather than the center because the conditional probability of its hitting the horizontal line on the disjunction of its hitting the horizontal line and the center is higher than the conditional probability of its hitting the center on the same condition, and that the conditional probability of the spinner yielding a multiple of five degrees is 1.6 times the conditional probability of the spinner yielding a multiple of eight degrees, both conditionally on its yielding a multiple of five or eight degrees. Hence, the zero-probability problem for the comprehensiveness of

[4] They carry strictly less information than non-Archimedean probabilities as the underdetermination theorem in [14] shows that there are distinct non-Archimedean probabilities that define the same conditional probability. On the other hand, there are primitive conditional probabilities for a fair spinner that are rotationally invariant [11, 12, 13]. This is easily checked to imply that the conditional probability of the spinner landing on a multiple of five degrees given that it lands on a multiple of five or eight degrees is greater than that of the spinner landing on a multiple of eight degrees on the same condition, and hence the model distinguishes outcomes that are classically indistinguishable, and hence carries strictly more information than the classical model.

the classical model is solved. The primitive conditional probability model is, thus, quite promising.

The point of this paper, however, is to argue that conditional probability models, like classical and comparative probability models, miss out on some of the real probabilistic information, and hence also fail to be comprehensive. The form of the argument is:

(A) For every probabilistic situation, there is an appropriate comparative probability model all of the information in which is genuine and non-superfluous.
(B) There is in general insufficient information in conditional probability models to reconstruct the comparative probability model for the situation being modeled.
(C) Therefore, there is genuine information about probabilistic situations that is missed out by conditional probability models.

I take (A) to be very plausible: probabilistic reality includes probabilistic comparisons between events, for an appropriate set of axioms for probabilistic comparisons. I will take these axioms to be the standard axioms for total or at least partial comparative probability. Thus the focus of the paper will be on arguing for (B).

The literature contains two attempts at generating a total probabilistic comparison out of a conditional probability. The first is apparently mistakenly attributed by Pruss [15] to de Finetti [1]:[5]

TDiff: $A \precsim B$ if and only if $P(A - B \mid A \,\Delta\, B) \le P(B - A \mid A \,\Delta\, B)$,

where $A - B$ is the difference of A and B, i.e., the intersection of A with the complement of B, while $A \,\Delta\, B$ is the symmetric difference between A and B, i.e., $(A - B) \cup (B - A)$. The second was given by Pruss [15] and perhaps Easwaran [2]:[6]

Union: $A \precsim B$ if and only if $P(A \mid A \cup B) \le P(B \mid A \cup B)$.

[5]De Finetti [1, p. 567] writes that in the case where neither $A \subset B$ nor $B \subset A$ holds, "[i]t is a question of comparing the conditional probabilities $[P(A - B \mid A \,\Delta\, B)]$ and $[P(B - A \mid A \,\Delta\, B)]$ (whose sum = 1), saying that A is more or less probable than B according to whether the first expression is greater than the second or is less." In [15, p. 3534n10], this is apparently misunderstood to imply that when the two conditional probabilities are equal, then A and B are equiprobable, perhaps because it was assumed that de Finetti was working with a total ordering, which his discussion of cases of non-comparability implies he is not.

[6]Easwaran [2, p. 17] offers the suggestion (it may not be a definition) that $A \prec B$ if and only if $P(A \mid A \cup B) < P(B \mid A \cup B)$. Had it been one's intention to define a total (pre)ordering, then in the case where neither $A \prec B$ nor $B \prec A$ holds, one would need to take A and B to be equiprobable, i.e., $A \precsim B$ and $B \precsim A$, and so we would have Union. But given that Easwaran, like de Finetti [1] before him, appears only interested in a partial ordering (see [2, pp. 24–25]), he has not given us an answer to the question of when exactly A and B are equiprobable (though again it was apparently not his intention to do so). If we equip Easwaran's account with the thesis that A and B are equiprobable if and only if $A = B$, then the resulting partial order will be subject to the same criticisms as the PDiff below is.

Neither is satisfactory. In fact, it is shown in the Appendix that TDiff does not in general make $\precsim$ a transitive relation.[7] On the other hand, it is also shown in the Appendix that Union sometimes violates a very plausible constraint on $\precsim$, namely the Complementation Axiom:

Comp: $A \precsim B$ if and only if $B^c \precsim A^c$,

where A^c is the complement of the event A. Violating Comp is probably less bad than violating transitivity, but both are highly plausible constraints on a probability ordering. (Note that Comp is easily seen to be a special of the Additivity Axiom for comparative probabilities.)

De Finetti [1, pp. 566–567] may have suggested this partial probabilistic comparison (see also [2, p. 17]):

PDiff: $A \precsim B$ if and only if $A \subseteq B$ or $P(A - B \mid A \Delta B) < P(B - A \mid A \Delta B)$.

This comparison does satisfy all the axioms of a partial comparative probability, as will be shown in the Appendix.

All three relations defined above satisfy the following "zooming" intuition:

Zoom: If $P(A \mid C) < P(B \mid C)$ for some C containing $A \cup B$, then $A \prec B$,

where we say that $A \prec B$ if and only if $A \precsim B$ but not $B \precsim A$. This zooming condition is central to our intuition that primitive conditional probabilities rightly capture the comparisons between the zero probability sets which suggested the insufficiency of classical probability.

The Zoom condition together with (P1) implies Regularity:

Reg: If A is non-empty, then $\varnothing \prec A$.

This can be interpreted as saying that a possible event is more likely than an impossible one, and encodes much of the intuition we started the paper with. To see the implication from Zoom to Reg, note that $P(\varnothing \mid A) = 0 < 1 = P(A \mid A)$ for any non-empty A by (P1).

However, I will argue for (B) by proving that there is no total and regular comparative probability satisfying the standard axioms for comparative probabilities and which is "canonically" defined in terms of conditional probabilities and boolean relations on sets (a precise necessary condition for canonicity will be given). The same holds for partial comparative probabilities if we add a weak and plausible constraint about how the comparative probabilities must handle the case of a uniform spinner. Furthermore, the result remains true even if we restrict ourselves to the case of comparative probabilities for a uniform spinner and events that are countable unions of intervals of angles.

Our result, with the method of proof behind it, is in the same spirit as Thong's recent argument [18, 19] that a number of proposals about relating regular comparative probabilities to conditional probabilities and subset relations fail. But our result provides stronger support for a general pessimistic conclusion about grounding comparative probabilities in conditional probabilities by not relying on refuting particular proposals for such a grounding.

[7] Contrary to the implicature of [15, p. 3534n10].

We can conclude that of the four most common tools for modeling probabilistic situations, only non-Archimedean probabilities capture all of the information in a probabilistic situation. However, if [5, 2, 14] are correct that non-Archimedean models is superfluous, then no one of the four tools is satisfactory. In the concluding section, we discuss three possible ways forward.

Finally, note that by the task of capturing "probabilistic reality" I mean more than capturing all the *particular* probabilistic facts of our world. Rather, I mean the modeling of the *kinds* of stochastic features that our world contains. For instance, a number of our arguments will involve uniform spinners. Now, it may be that our universe does not actually have any uniform spinners. But plausibly the kinds of stochastic features that the non-uniform spinners of our "imperfect" universe exhibit—say, numerical chances—would also be found in universes which have uniform spinners, and so an account of what probabilistic phenomena fundamentally are should extend to those universes. The point here is analogous to how a philosopher of science in searching for an account of nomic reality is looking for an account that can account for laws of nature in universes where matter is differently arranged from how it is arranged in our universe.

3. Main results and arguments

3.1. **Canonicity.** We now need to rigorously state our result that a conditional probability function together with boolean set operations cannot canonically define a probability comparison satisfying the requisite assumptions. What would it mean to define a probability comparison in terms of a conditional probability? We would presumably have to have some sort of a biconditional with $A \precsim B$ on the left-hand-side and expressions involving A, B and conditional probabilities on the right-hand-side, with TDiff, Union and PDiff providing simple examples, and more complicated examples perhaps involving quantification over other events.

Here, we need some way to ensure the right-hand-side of the biconditional avoids cheating. Given a conditional probability P on Ω, there is a finitely-additive hyperreal probability Q that is regular in the Bayesian sense that $Q(A) > 0$ for all non-empty A and that generates P using the formula $P(A \mid B) = \operatorname{Std} Q(A \cap B)/Q(B)$ for $B \neq \varnothing$ and $P(A \mid \varnothing) = 1$ [9, 10], where $\operatorname{Std} x$ is the real number closest to the finite hyperreal x.[8] But then Q generates a probability comparison by stipulating $A \precsim B$ if and only if $Q(A) \leq Q(B)$, and this $\precsim$ satisfies all the axioms of a total probability comparison as well as Zoom. Now, there are set theories with a global choice function[9] F such that $F(U) \in U$ for every non-empty set U, and one could

[8]A hyperreal x is finite provided that there is a (standard) natural number N such that $|x| < N$, and $\operatorname{Std} x = \sup\{y \in \mathbb{R} : y < x\}$.

[9]Such a theory is consistent if and only if standard Zermelo-Fraenkel set theory is consistent, since it's a conservative extension of Zermelo-Fraenkel set theory with choice [4].

use such a global choice function F to define the ordering $\precsim$ given P: just let C_P be the set of all total comparative probabilities $\precsim$ on the same algebra as P is defined on that satisfy ZOOM, use the fact just proved that C_P is non-empty, and define $\precsim_P$ as the relation $F(C_P)$. To put the point vividly, suppose God exists, and that for any non-empty set, God has a favorite member. Then we can stipulate that for any conditional probability P on a space Ω, the probability comparison defined by P is God's favorite member of the nonempty set of total probability comparisons on the same algebra as P is defined on that satisfy ZOOM.

But this is cheating as we haven't defined $\precsim$ just by relying on the conditional probability P and boolean relations between events, but by also relying additional information contained in God's preferences, or in whatever set-theoretic machinery is involved in defining the global choice function. That's not what's going on in our paradigm examples TDIFF, UNION, and PDIFF.

The intuition behind our canonicity condition will be that $\precsim$ is defined in terms of the boolean relations and operations on events (e.g., $\subseteq$, $\cup$, $\cap$, $(\cdot)^c$) and in terms of some mathematical vocabulary of real numbers, in addition to logical terminology, but not in terms of other details. Suppose we have one full conditional probability P_1 defined on a standard 451 mm diameter dart board at which a dart is uniformly randomly thrown and another full conditional probability P_2 defined on a one kilometer radius circular region of space in another galaxy at which a perfectly spherical asteroid is thrown in a uniformly distributed way, and there is a bijection ρ between the points of the dart board and points of the region such that $P_2(\rho A \mid \rho B) = P_1(A \mid B)$ for all subsets A and B of the dart board[10] where $\rho C = \{\rho(x) : x \in C\}$ for a subset C of the dart board. Then we can say that the two situations are isomorphic in terms of conditional probabilities, and canonicity will require that the comparative probability orderings in the two situations generated by P_1 and P_2 respectively are also order-isomorphic under ψ, since all of the boolean relations and operations are preserved by the isomorphism. The differences in size of targets and physical nature of projectiles are irrelevant for a canonical definition of conditional probabilities.

More generally, if we have a conditional probability P on an algebra $\mathcal{F}$ of subsets of Ω, and a bijection ρ of Ω to another space Ω', we will have the algebra $\mathcal{F}^\rho = \{\rho A : A \in \mathcal{F}\}$ of subsets of Ω', and the isomorphic conditional probability P^ρ on $\mathcal{F}^\rho$ defined by $P^\rho(A \mid B) = P(\rho^{-1}A \mid \rho^{-1}B)$. If $\precsim_Q$ is the comparative probability defined by the conditional probability Q, then we will then have this canonicity requirement:

CAN: $A \precsim_P B$ if and only if $\rho A \precsim_{P^\rho} \rho B$, for all events A and B, conditional probabilities P, and bijections ρ from the space that P is defined on.

---

[10] This condition need not be automatically satisfied for all uniform P_1 and P_2, as we should be open to the possibility that there could be more than one uniform conditional probability. (Cf. note 19, below.)

This condition is satisfied by any definition of $\precsim_P$ that in an intuitive sense "depends only on the values of P and boolean relations between events", rather than on set theoretic facts about the specific identities of the sets that P is defined on, divine preferences, or any other things that are not preserved by the isomorphism between $\mathcal{F}$ and $\mathcal{F}'$ induced by ρ. In particular, it is clearly satisfied by TDiff, Union and PDiff.

I do not know if Can fully captures all the intuitions behind the notion of a canonical definition[11], but it provides a very plausible necessary condition for canonicity.[12]

3.2. **Total comparative probabilities.** Let us begin with the special case of total comparative probabilities.[13]

Let $\mathbb{T}$ be the unit circle, and let $\mathcal{F}$ be the algebra of countable unions of circular intervals or arcs, where a circular interval is the set of points at angles θ where θ is constrained to lie in some interval of real numbers (including the trivial interval $[a,a] = \{a\}$). Our main result for total comparative probabilities is:

Theorem 1. *There is no assignment of regular total comparative probabilities $\precsim_P$ to conditional probabilities P on $\mathcal{F}$ where the assignment satisfies* Can.

This result will be an immediate consequence of our Theorem 2 below together with the existence of a strongly rotationally invariant full conditional probability on the circle [11, 12, 13], where we say that P is strongly

[11]We might also consider a slightly stronger condition, namely that for any boolean algebra isomorphism ρ from an algebra $\mathcal{F}$ to an algebra $\mathcal{F}'$, if we define $P^\rho(A \mid B) = P(\rho^{-1}A \mid \rho^{-1}B)$, then $\precsim_P$ and $\precsim_{P^\rho}$ are order-isomorphic under ρ. In Can this is only required in the special case of boolean algebra isomorphisms generated by bijections of the underlying space (the two conditions are equivalent in the case where $\mathcal{F}$ includes all singletons). But since our technical results only require the weaker condition, we might as well as stick to it.

[12]We can think of the above characterization of canonicity in terms of isomorphism as a semantic characterization (in the sense of the semantic/syntactic distinction in logic). We might instead try to characterize canonicity syntactically, by requiring there to be a formula $\Phi(P, A, B)$ using some specified language L such that for any conditional probability P, the relation defined by $A \precsim_P B$ if and only if $\Phi(P, A, B)$ is a comparative probability satisfying additional conditions like Zoom. Specifying the language L so that it be sufficiently rich to allow any reasonable definition of $\precsim$ while at the same time not letting in cheating definitions is not that easy. We could always wonder if some further ingredients shouldn't be allowed into the language, and hence whether one is being fair to someone who thinks there is a good definition of comparative probability in terms of conditional probability but doesn't know what it is. But in any case, it is very plausible that whatever plausible formula $\Phi(P, A, B)$ we might give, the formula should preserve boolean-algebra isomorphisms generated by bijections, so we would have $\Phi(P, A, B)$ if and only if $\Phi(P^\rho, \rho A, \rho B)$, and hence Can would be satisfied.

[13]Although there may be cases where it's plausible to deny totality, there presumably are some who find totality plausible—for instance, anyone attracted to the fine-grainedness of hyperreal probabilities is likely to find totality plausible.

invariant under some collection G of bijections of Ω onto Ω (in our case, the rotations) provided that $P(A \mid B) = P(gA \mid B)$ whenever $A \cup gA \subseteq B$.

Thus, conditional probabilities do not contain enough information to define a total comparative probability.

There is a way to see why this should be expected to be true. The results in [12, 13] show that are a number of situations, such as uniform spinners, where one can find full conditional probabilities that satisfy strong symmetry conditions, but there do not exist total comparative probabilities that satisfy even weak symmetry conditions. Because of this, assigning a total comparative probability requires one to break the symmetry in the conditional probability, and there is no canonical way of doing so.

3.3. **Partial comparative probabilities.**

3.3.1. *Some technical results.* What if we drop totality and consider partial comparative probabilities?

The de Finetti-inspired ordering given by PDiff satisfies the axioms of partial comparative probability as well as Zoom (see proof in the Appendix), and clearly satisfies Can. It follows from Theorem 1 that it is not total, but perhaps we can live with that. There might be events that are just incomparable.[14]

The PDiff proposal has a very counterintuitive consequence, since it implies:

Unfair There are no distinct equiprobable events.

Here, A and B are equiprobable provided that $A \sim B$, i.e., both $A \precsim B$ and $B \precsim A$. But by PDiff this can only happen if $A = B$.[15] Given Unfair, there can be no fair coins, dice, lotteries or spinners. This is highly implausible.

Indeed, this is so implausible that it provides a plausible argument that de Finetti did not intend PDiff. While it is fairly clear from de Finetti [1, pp. 566-567] that he intended to say that $A \prec B$ if and only if A is a proper subset of B or $P(A - B \mid A \,\Delta\, B) < P(B - A \mid A \,\Delta\, B)$, it is not completely clear under what circumstances he would want to say that A and B are equiprobable. In PDiff, I took the interpretation that he would want to say this only if $A = B$, and hence in defining $A \precsim B$, I simply replaced the proper subset relation in de Finetti's definition of $\prec$ with a subset relation when defining $\precsim$.

[14]Fishburn [3] suggests that highly disparate events, such as predictions about a card in a game given incomplete information and predictions about future population growth, will sometimes be incomparable. A more technical possibility is that there is likely to be incomparability between events such as a uniform spinner landing on a Lebesgue non-measurable set A and its landing on a measurable set B, when the measure of B lies between the lower and upper measures of A.

[15]The same point is true for the version of Easwaran's [2] ordering when we interpret it as partial in the way discussed above in note 6.

There is some reason to think this is not an uncharitable interpretation. For the alternative is that we try to save the possibility of a fair coin (say) by having some equiprobability relation $\sim$ on the events that is weaker than identity, and in particular that holds between heads (H) and tails (T), and then let $\precsim$ be defined as the disjunction $A \prec B$ or $A \sim B$. But this doesn't work.

To see that it doesn't work, consider our fair coin, with its equiprobable disjoint and mutually exhaustive events H and T, and suppose that we also have some zero-probability non-empty event E that is not a subset of H, say the event of an independent random dart hitting the center of a target, or the event that the coins in some countably infinite collection of independent fair coins that is also independent of our first coin all land heads. Let $H' = H \cup E$. Note that $P(H - H' \mid H \Delta H') = P(\varnothing \mid E - H) = 0$ but $P(H' - H \mid H \Delta H') = P(E - H \mid E - H) = 1$, so by the clear part of de Finetti's definition we have $H \prec H'$. If we have $T \sim H$, then we will have to have $T \prec H'$ (since if H' is more probable than something equiprobable with T, it is more probable than T). But H, $H - T$, $T - H$ and $T \Delta H$ differ from H', $H' - T$, $T - H'$ and $T \Delta H'$, respectively, by subsets of the zero-probability set E, and hence respectively have the same unconditional probabilities. Since $H - T = H$, $T - H = T$, and $T \Delta H = \Omega$, we have $P(H - T \mid H \Delta T) = 1/2 = P(T - H \mid H \Delta T)$, and hence similarly $P(T - H' \mid T \Delta H') = 1/2 = P(H' - T \mid T \Delta H')$, since $P(T \Delta H') = P(\Omega) = 1$. We also do not have $T \subset H'$. Thus, we do not have $T \prec H'$ on the account in question, and hence we cannot have $T \sim H$. (This argument is similar to one given by Thong [19, 18].)

At this point, it is reasonable to ask whether there might be some other way of canonically defining a partial comparative probability in terms of a conditional probability that avoids the problems with PDiff and as well as with variants where $\sim$ is weaker than identity. I will argue that the answer is negative.

In the spirit of Thong's argument [19, 18], suppose we have a uniform spinner choosing a random point on the unit circle $\mathbb{T}$. By the work of Parikh and Parnes [11] or Theorem 1 in [12, 13] (together with the fact that group of rotations on the circle is commutative), there is a strongly rotationally invariant full conditional probability P on $\mathbb{T}$, and presumably some such strongly rotationally invariant full conditional probability then models our uniform spinner. But now consider the event T_0 of the point being at an angle between $0°$, inclusive, and $180°$ exclusive, and the event T_1 of its being at an angle between $180°$, inclusive, and $360°$, exclusive. Intuitively we expect these events to be probabilistically equivalent, i.e., $T_0 \sim T_1$, even if we do not insist that *all* events that are rotationally equivalent are probabilistically equivalent (there exist events A and B that are rotationally equivalent but where nonetheless A is a proper subset of B, and there we do not expect probabilistic equivalence [16]).

But it will turn out that there is no way to canonically generate a regular partial probability comparison out of a conditional probability that satisfies ZOOM and allows one to have $T_0 \sim T_1$ when $\precsim$ is generated out of the appropriate strongly rotationally invariant conditional probability.

First, however, note that the same intuitions that pull one to accepting regularity also pull towards Strong Regularity:

SREG: If A is a proper subset of B, then $A \prec B$.

Strong regularity follows from regularity and additivity, since by additivity if A is a proper subset of B, then $\varnothing \prec B - A$ if and only if $A \prec B$ (since both $\varnothing$ and $B - A$ are disjoint from A and $\varnothing \cup A = A$ while $(B - A) \cup A = B$). However, assuming strong regularity is weaker than assuming additivity and regularity[16], and we saw that regularity follows from ZOOM, so if we can formulate a negative result in terms of strong regularity instead of additivity and regularity, it is better to do so.

Identify our spinner probability space with $\mathbb{T}$. Let T_0 be the set of points on $\mathbb{T}$ with angles in $[0°, 180°)$ and let T_1 be the set of points on $\mathbb{T}$ with angles in $[180°, 360°)$. As before, let $\mathcal{F}$ be all countable unions of circular intervals on $\mathbb{T}$.

Theorem 2. *Suppose that for every conditional probability P there is a transitive relation $\precsim_P$ satisfying both* SREG *and* CAN. *Then for every strongly rotationally invariant conditional probability Q on $\mathcal{F}$, we have neither $T_0 \precsim_Q T_1$ nor $T_1 \precsim_Q T_0$.*

Recall that there *is* a strongly rotationally invariant conditional probability Q on all subsets of $\mathcal{T}$ [11, 12, 13], and hence on $\mathcal{F}$, and it is plausible that some such probability appropriately models our spinner. Thus, we have a powerful argument that there is no way to define partial comparative probabilities in terms of conditional probabilities.

As noted before, Theorem 1 follows immediately from Theorem 2 and the existence of Q.

The above argument was run with strong rotational invariance. Weak invariance under a symmetry ρ is the even more plausible condition that $Q(\rho A \mid \rho B) = Q(A \mid B)$, which may appeal to some readers. I do not know if strong rotational invariance in Theorem 2 can be replaced with weak rotational invariance. However, if we replace strong rotational invariance with weak invariance under reflections about the center of the circle, the result remains true. The reason is that weak invariance under reflections implies strong invariance under reflections [17], and strong invariance under reflections implies strong invariance under rotations, since any non-trivial rotation can be generated by two reflections about different lines. Moreover,

[16] E.g., let $\Omega = \{1, 2, 3\}$, and define $A \lesssim B$ if and only if either (a) $A = B$, or (b) $|A| < |B|$, or (c) $A = \{1, 3\}$ and $B = \{2, 3\}$. It is easy to see that this is a strongly regular partial ordering, but it is not additive since $\{1, 3\} \lesssim \{2, 3\}$ but not $\{1\} \lesssim \{2\}$.

not only is there a strongly rotationally invariant conditional probability on $\mathbb{T}$, there is a strongly reflection-invariant one.[17]

3.3.2. *The equiprobability of T_0 and T_1.* Now, not everyone will accept it as a *constraint* on a uniform spinner that T_0 and T_1 are equiprobable. One might, after all, assign non-classical probabilities to a countably infinite fair lottery in such a way that all singletons are equiprobable and that getting an even number is equiprobable with getting an odd number, but one might also do so in such a way that the evens are not equiprobable with the odds, and one might think that *both* count as fair.[18]

For a reader who is not convinced that uniformity requires the equiprobability of T_0 and T_1, I offer a somewhat more complex argument. Given regularity, there indeed is plausibly no such thing as *the* uniform spinner.[19] But then it is intuitively plausible among the uniform spinners, there is at least one where T_0 and T_1 are equiprobable, even if there are other uniform spinners where they are not comparable. Next note that as observed earlier there actually *is* a strongly rotationally invariant conditional probability Q on all subsets of the circle. Thus we would expect there to be possible uniform spinners modeled by such invariant probabilities. The hypothesis I am arguing against is that comparative probabilities are canonically derived from conditional probabilities. On that hypothesis, the comparative probabilities for a uniform spinner where T_0 and T_1 are equiprobable will have to be derived from conditional probabilities. According to Theorem 2, they must then be derived from a uniform spinner with conditional probabilities that are *not* strongly rotationally invariant. But this is quite counterintuitive. If there are uniform spinners modeled by a strongly rotationally invariant conditional probability, and there are uniform spinners where T_0 and T_1 are equiprobable, we would expect this equiprobability to be generable from a strongly rotationally invariant conditional probability, assuming that comparative probabilities are generated from conditional probabilities.

---

[17] This follows from Theorem 1 in [13] (see corrected proof in [12]) and the fact that there is no reflection-paradoxical subset of the circle. For if A is a reflection-paradoxical subset of the circle, then $\{\theta : e^{i\theta} \in A\}$ is a reflection-paradoxical subset of the real line, but that is impossible since the isometries of the real line are a supramenable group. [20, Th. 14.21 and Cor. 14.25].

[18] I am grateful to an anonymous reader for this objection.

[19] Consider two different ways of running a spinner, both of which are intuitively uniform. On the first one, we just "uniformly" spin the spinner, and report the final angle. On the second one, take the final angle θ from the first spinner, then double it to get 2θ, and then report the angle in $[0, 360^\circ)$ degrees that is equivalent to 2θ. Both appear to be uniform, but if all singletons are equally likely with the first spinner, they are still equally likely with the second—but each singleton is twice as likely as with the first spinner, since the second spinner reports the same angle ϕ in $[0, 360^\circ)$ when the first spinner yields $\phi/2$ as when the first spinner yields $\phi/2 + 180^\circ$.

3.3.3. *One-way additivity.* Finally, observe that the total ordering given by UNION, while it must fail to have additivity by Theorem 1 (since it has all the other properties mentioned), does satisfy the one-way additivity axiom:

ADD1: If C is disjoint from $A \cup B$ and $A \precsim B$, then $A \cup C \precsim B \cup C$,

which is perhaps the even more intuitive half of additivity.

As show in the Appendix, UNION is unsatisfactory because it fails the very plausible axiom COMP. One might wonder if one could get a good probability ordering out of a conditional if we replaced additivity with ADD1 and COMP. However, it is easy to show that ADD1 conjoined with COMP implies full additivity, so this is not tenable.[20]

4. CONCLUSIONS

Two attempts to define of total comparative probabilities in terms of full conditional probabilities fail to be satisfactory, one by violating the very plausible Complementarity Axiom and the other, even worse, by violating transitivity. This failure is no accident: there is no canonical definition of total comparative probabilities in terms of full conditional probabilities. Thus, there is an important sense in which a comparative probability assignment contains information that goes beyond the information in a full conditional probability. One might have hoped that recovering partial comparative probabilities will work better than recovering total ones. However, on the plausible assumption that a fair spinner is equally likely to fall in the range from $0°$, inclusive, to $180°$, exclusive, as it is to fall in the range from $180°$, inclusive, to $360°$, exclusive, this hope is also undercut by our results.

We thus have good reason to think that primitive conditional probabilities, like classical and comparative probabilities, fail the comprehensiveness desideratum on a model of probabilities. On the other hand, the non-Archimedean model appears to fail non-superfluity by containing too much information [5, 2, 14].

We end by sketching examples of the three possible ways out of the difficulty.

First, we can combine models that are individually non-comprehensive but also non-superfluous to get a model that we hope is both comprehensive and non-superfluous. For instance, the classical and comparative models had different difficulties with comprehensiveness, while neither appeared to suffer from superfluity of information. The classical approach had trouble with zero-probability events while the comparative model had trouble with distinctions between probabilities in worlds with small probability spaces.

[20] Suppose we have ADD1 and COMP, and suppose $A \cup C \precsim B \cup C$ with C disjoint from $A \cup C$. Then $(B \cup C)^c \precsim (A \cup C)^c$ by COMP. Thus, $(B \cup C)^c \cup C \precsim (A \cup C)^c \cup C$ by ADD1. By COMP and De Morgan, $(A \cup C) \cap C^c \precsim (B \cup C) \cap C^c$, so $A \precsim B$ since C is disjoint from A and B.

Perhaps we can then say that probabilistic reality has two aspects: a numerical and a comparative one, and both numerical and comparative probabilities are fundamental. Such a combination may have been what de Finetti [1] actually had in mind. Combining primitive conditional probabilities with comparative probabilities is another option.[21] That said, such combinations appear inelegant.

Second, we can consider or develop a model beyond the four most common ones. For example, Koopman [7] introduced conditional comparative probabilities and James Hawthorne [6] has developed this suggestion in detail. Such conditional comparative probabilities might be able to surmount the small-worlds objection to taking comparative probabilities as the correct model of probabilistic phenomena on the grounds. Recall the difficulty posed by a world containing a single toss of an unfair coin, where unconditional comparative probabilities were not able to specify how unfair the coin is. But very speculatively (and going beyond what Koopman and Hawthorne likely intend for their accounts) one could try to enrich the probability space by including counterfactual options that are physically unavailable in a given world. Thus, in a small world with one unfair coin, one might include in the probability space physically unavailable (and hence classically zero probability) throws of many-sided dice, and then compare the unconditional chance of our unfair coin being heads to chances of throws of many-sided dice *conditionally* on such throws being made. That said, there is something uncomfortable about understanding the chances of coin tosses in our world in terms of physically unavailable dice throws. And inclusion of such tosses may seem to violate the non-superfluity desideratum[22], though whether this is so is not completely clear, since reality might include facts about counterfactual situations.

Finally, we might dispute the critiques of one or more of the four models. For example, of the four, the classical probability model has the advantage that it is the one that most mathematicians and statisticians use and has shown itself especially rich in practical and theoretical applications. Perhaps we can follow standard mathematical practice and bite the bullet on the zero-probability problem, denying that the event of $0 = 1$ is less likely than the spinner's landing at a specific point. On the other hand, if the intuitions about comparing zero-probability events are taken as fixed, one might simply take comparative probabilities as fundamental and bite the bullet on the problem of small-worlds, denying that there is a fact about whether the coin in the single-toss world is fair or not.[23]

[21] On the other hand, combining primitive conditional probabilities with classical probabilities gains nothing, since primitive conditional probabilities already contain all of the information of the classical model. And a combination including non-Archimedean probabilities is going to suffer from the superfluity in non-Archimedean probabilities.

[22] I am grateful to an anonymous reader for this point

[23] I am grateful to two anonymous readers for a number of comments that have greatly improved the philosophical argumentation and clarity of this paper.

Appendix: Some proofs

We begin by giving the promised counterexample to the transitivity of TDiff. Let $\Omega = \{a, b, c\}$, and define a non-Archimedean finitely-additive probability Q on Ω by $Q(\{a\}) = Q(\{b\}) = 1/2 - \alpha$ and $Q(\{c\}) = 2\alpha$ for an infinitesimal $\alpha > 0$. Define $P(A \mid B) = \operatorname{Std} Q(A \cap B)/Q(B)$ for $B \neq \varnothing$, where $\operatorname{Std} x$ is the standard part of x, i.e., the real number closest to x, assuming x is finite.[24] Stipulate $P(A \mid \varnothing) = 1$. Then P satisfies all the axioms of a full conditional probability.

Now, let $A = \{a\}$ and $B = \{b, c\}$. Then $A - B = A$ and $B - A = B$, and $A \Delta B = \Omega$. By TDiff, we have $A \sim B$ (i.e., $A \precsim B$ and $B \precsim A$) since

$$P(A \mid \Omega) = \operatorname{Std}(1/2 - \alpha) = 1/2 = P(B \mid \Omega).$$

Next let $A' = \{a\}$ and $B' = \{b\}$. Then $A' - B' = A'$ and $B' - A' = B'$ while $A' \Delta B' = \{a, b\}$. Thus:

$$P(A' \mid A' \Delta B') = P(\{a\} \mid \{a, b\}) = 1/2 = P(\{b\} \mid \{a, b\}) = P(B' \mid A' \Delta B'),$$

and so $A' \sim B'$. We have thus shown that $\{b, c\} \sim \{a\}$ and $\{a\} \sim \{b\}$. If $\precsim$ were transitive, it would follow that $\{b, c\} \sim \{b\}$, so $\{b, c\} \precsim \{b\}$. But:

$$\begin{aligned} P(\{b, c\} - \{b\} \mid \{b, c\} \Delta \{b\}) &= P(\{c\} \mid \{c\}) = 1 \\ &> 0 = P(\{b\} - \{b, c\} \mid \{b, c\} \Delta \{b\}), \end{aligned}$$

since $\{b\} - \{b, c\} = \varnothing$. So by TDiff we would have $\{b\} \prec \{b, c\}$, contradicting $\{b\} \sim \{b, c\}$.

The same P and Ω also show that the Pruss definition Union violates Comp. Let $A = \{a, b\}$ and $B = \Omega$. Then $P(A \mid A \cup B) = 1 = P(B \mid A \cup B)$, so $A \sim B$ by Union. But $A^c = \{c\}$ and $B^c = \varnothing$ while $P(A^c \mid A^c \cup B^c) = 1$ and $P(B^c \mid A^c \cup B^c) = 0$, so we do not have $A^c \precsim B^c$.

Proposition 1. *Say that $A \precsim B$ provided that $A \subseteq B$ or $P(A - B \mid A \Delta B) \prec P(B - A \mid A \Delta B)$. This is a partial comparative probability satisfying* Zoom.

Proof. Reflexivity and Pos are obvious. Moreover, $\precsim$ is the union of two orderings, one defined by $A \subseteq B$ and the other by $P(A - B \mid A \Delta B) \prec P(B - A \mid A \Delta B)$. The former obviously satisfies additivity. To see that the latter satisfies additivity, note that if C is disjoint from $A \cup B$, and $A' = A \cup C$ and $B' = B \cup C$, then $A - B = A' - B'$, $B - A = B' - A'$ and $A' \Delta B' = A \Delta B$, so we have additivity.

That leaves transitivity and Zoom. For transitivity, suppose $A \precsim B$ and $B \precsim C$, and we must prove $A \precsim C$. If $A \subseteq B$ and $B \subseteq C$, we are done.

Next suppose $A \subseteq B$, $P(B - C \mid B \Delta C) \prec P(C - B \mid B \Delta C)$ and $C \subseteq D$. Then

$$P(B - C \mid (B \Delta C) \cup (A \Delta D)) \leq P(C - B \mid (B \Delta C) \cup (A \Delta D)),$$

[24] A value x in a non-Archimedean field is finite provided that $-N < x < N$ for some integer N. We can define $\operatorname{Std} x = \sup\{y \in \mathbb{R} : y < x\}$.

with equality only if

$$P(B \Delta C \mid A \Delta D) = 0,$$

since if $P(U \mid W) < P(V \mid W)$ and $U \cup V \cup W \subseteq W'$, then $P(U \mid W') \leq P(V \mid W')$ with equality only if $P(W \mid W') = 0$, since $P(U \mid W') = P(U \mid W)P(W \mid W')$.

Suppose we don't have equality. Then

$$P(A - D \mid (B \Delta C) \cup (A \Delta D)) < P(D - A \mid B \Delta C \cup (A \Delta D)),$$

since $A - D \subseteq B - C$ and $C - B \subseteq D - A$. Since $A - D \subseteq A \Delta D$ and $D - A \subseteq A \Delta D$, it follows that:

$$P(A - D \mid A \Delta D) < P(D - A \mid A \Delta D),$$

and so $A \precsim D$. Suppose now that we have $P(B \Delta C \mid A \Delta D) = 0$. Then $P(A - D \mid A \Delta D) = 0$, since $A - D \subseteq B - C \subseteq B \Delta C$. But unless $A \Delta D = \varnothing$,

$$1 = P(A - D \mid A \Delta D) + P(D - A \mid A \Delta D),$$

so

$$P(A - D \mid A \Delta D) = 0 < 1 = P(D - A \mid A \Delta D),$$

and once again $A \precsim D$. Suppose now that $A \Delta D = \varnothing$, in other words that $A = D$. Then $C \subseteq D = D \subseteq B$, which is incompatible with $P(B - C \mid B \Delta C) < P(C - B \mid B \Delta C)$.

The remaining case of transitivity we need to handle is where $P(A - B \mid A \Delta B) < P(B - A \mid A \Delta B)$ and $P(B - C \mid B \Delta C) < P(C - B \mid B \Delta C)$. Letting $E = (A \Delta B) \cup (B \Delta C)$, we then have:

$$P(A - B \mid E) \leq P(B - A \mid E) \tag{1}$$

and

$$P(B - C \mid E) \leq P(C - B \mid E), \tag{2}$$

and the only we way we can have equality in both (1) and (2) is if

$$P(A \Delta B \mid E) = 0 = P(B \Delta C \mid E),$$

which is impossible by additivity and the definition of E, unless $A \Delta B = B \Delta C = \varnothing$, which itself is incompatible with our assumptions. So, by additivity from (1) and (2), together with strictness in at least one of them, we must have:

$$P((A - B) \cup (B - C) \mid E) < P((B - A) \cup (C - B) \mid E). \tag{3}$$

Now, in general, $A - C = ((A - B) \cup (B - C)) - D$ where $D = (B - (A \cup C)) \cup (A \cap C)$. Note that $D \subseteq (A - B) \cup (B - C)$. Swapping A and C, we see that $C - A = ((C - B) \cup (B - A)) - D$ and $D \subseteq (C - B) \cup (B - C)$. From (3) it follows by additivity that:

$$P(A - C \mid E) < P(C - D \mid E).$$

Since $A \Delta C \subseteq E$, it follows that $P(A - C \mid A \Delta C) < P(C - A \mid A \Delta C)$, and so $A \precsim C$.

It remains to show Zoom. Suppose $P(A \mid C) < P(B \mid C)$. Then $P(A - (A \cap B) \mid C) < P(B - (B \cap A) \mid C)$ by additivity, and so $P(A - B \mid C) < P(B - A \mid C)$. Since $(A - B) \cup (B - A) \subseteq A \,\Delta\, B \subseteq A \cup B \subseteq C$, it follows that $P(A - B \mid A \,\Delta\, B) < P(B - A \mid A \,\Delta\, B)$, and the proof is complete. □

Proof of Theorem 2. Let Q be strongly rotationally invariant. We will show that $T_1 \precsim_Q T_0$ implies a contradiction. By symmetry (rotate the whole set-up by π), it will follow that $T_0 \precsim_Q T_1$ also implies a contradiction.

Thus to obtain a contradiction, assume $T_1 \precsim_Q T_0$.

Write $\mathbb{T} = \{e^{i\theta} : \theta \in [0, 2\pi)\}$. Let $\phi \in (0, \pi)$ be an irrational multiple of π. Let $w_k : \mathbb{R} \to [k\pi, (k+1)\pi)$ for $k = 0, 1$ be functions such that $w_k(x) - x$ is an integral multiple of π (i.e., $w_k(x) = x - \pi\lfloor x/\pi \rfloor + k\pi$). Let $Z_k = \{w_k(n\phi) : n \in \mathbb{N}\}$ where $\mathbb{N}$ is the non-negative integers.

Now define a bijection ρ of $\mathbb{T}$ onto $\mathbb{T}$ as follows, where $\theta \in [0, 2\pi)$:

$$\rho(e^{i\theta}) = \begin{cases} e^{i(\theta+\phi)} & \text{if } \theta \in Z_0 \cap [0, \pi - \phi) \\ e^{i(\theta+\phi-\pi)} & \text{if } \theta \in Z_0 \cap [\pi - \phi, \pi) \\ e^{i(\theta-\phi)} & \text{if } \theta \in Z_1 \cap [\pi + \phi, 2\pi) \\ e^{i(\theta-\phi+\pi)} & \text{if } \theta \in Z_1 \cap (\pi, \pi + \phi) \\ 1 & \text{if } \theta = \pi \\ e^{i\theta} & \text{otherwise.} \end{cases}$$

Write $\mathbb{T}U = \{e^{i\theta} : \theta \in U\}$ for a set U of real numbers. Let $A_1 = \mathbb{T}(Z_0 \cap [0, \pi))$, $A_1' = \mathbb{T}(Z_0 \cap (0, \pi))$, $A_2 = \mathbb{T}(Z_1 \cap (\pi, 2\pi))$, $A_2' = \mathbb{T}(Z_1 \cap [\pi, 2\pi))$, $A_3 = \mathbb{T}\{\pi\} = \{-1\}$, $A_3' = \mathbb{T}\{0\} = \{1\}$, and $A_4 = A_4' = \mathbb{T} - \mathbb{T}(Z_0 \cup Z_1)$. Then A_1, A_2, A_3, A_4 are a partition of $\mathbb{T}$ and so are A_1', A_2', A_3', A_4', while ρ is a bijection of A_i onto A_i' for each i. Hence ρ is a bijection of $\mathbb{T}$ onto $\mathbb{T}$. Note too that all the A_i and A_i' are countable unions of circular intervals.

Furthermore, ρ restricted to each individual set A_i is a rotation. It follows by the finite additivity of Q and its strong rotational invariance that Q is strongly ρ-invariant, i.e., if $E \cup \rho[E] \subseteq F$, then:

$$\begin{aligned} Q(E \mid F) &= \sum_{i=1}^{4} Q(E \cap A_i \mid F) \\ &= \sum_{i=1}^{4} Q(\rho[E \cap A_i] \mid F) \\ &= Q\left(\bigcup_{i=1}^{4} \rho[E \cap A_i] \mid F\right) \\ &= Q(\rho[E] \mid F). \end{aligned}$$

Now let τ be rotation by angle π, and let $\psi(z) = \tau(\rho(z))$. Since Q is strongly rotationally invariant and strongly ρ-invariant, it is strongly ψ-invariant. Hence by [17, p. 279], Q is weakly ψ-invariant, i.e., $Q^\psi = Q$.

Then by CAN, since $T_1 \precsim_Q T_0$, we have $\psi[T_1] \precsim_Q \psi[T_0]$. But $\rho[T_0] = \mathbb{T}(0,\pi)$ and $\rho[T_1] = \mathbb{T}[\pi, 2\pi]$, so $\psi[T_0] = \mathbb{T}(\pi, 2\pi)$ and $\psi[T_1] = \mathbb{T}[0,\pi]$. Thus

$$\mathbb{T}[0,\pi] \precsim_Q \mathbb{T}(\pi, 2\pi). \tag{4}$$

Then by Strong Regularity and our assumption that $T_1 \precsim_Q T_0$, we have:

$$\mathbb{T}(\pi, 2\pi) \prec_Q T_1 \precsim_Q T_0 \prec \mathbb{T}[0,\pi],$$

which by transitivity contradicts (4). □

References

[1] de Finetti, Bruno. 2017. *Theory of Probability.* Transl. A. Machí and A. F. M. Smith. New York: Wiley.

[2] Easwaran, Kenny. 2014. "Regularity and hyperreal credences." *Philosophical Review* 123:1–41.

[3] Fishburn, 1986. "The axioms of subjective probability." *Statistical Science* 1:335–345.

[4] Gaifman, Haim. 1975. "Global and local choice functions." *Israel Journal of Mathematics* 22:257–265.

[5] Hájek, A. 2003. "What conditional probability could not be." *Synthese* 137:273–323.

[6] Hawthorne, James. 2017. "A logic of comparative support: Qualitative conditional probability relations representable by Popper functions." In: A. Hájek and C. Hitchcock (eds.), *Oxford Handbook of Probability and Philsoophy.* Oxford: Oxford University Press, 277–295.

[7] Koopman, B. O. 1940. "The axioms and algebra of intuitive probability." *Annals of Mathematics* 41:269–92.

[8] Krantz, D. H., Luce, R. D., Suppes, P., and Tversky, B. 1971. *Foundations of Measurement: Volume I: Additive and Polynomial Representations*, San Diego: Academic Press.

[9] Krauss, P. H. 1968. "Representation of conditional probability measures on Boolean algebras." *Acta Mathematica Academiae Scientiarum Hungarica* 19:229-241.

[10] McGee, Vann. 1994. "Learning the impossible." In: Ellery Eells and Brian Skyrms (eds.), *Probability and Conditionals.* Cambridge: Cambridge University Press, 1994.

[11] Parikh, Rohit, and Parnes, Milton. 1974). "Conditional probabilities and uniform sets." In: A. Hurd and P. Loeb (eds.), *Victoria Symposium on Nonstandard Analysis.* Berlin: Springer, 180–194.

[12] Pruss, Alexander R. 2022. "Correction: Non-classical probabilities invariant under symmetries", *Synthese* 200:385.

[13] Pruss, Alexander R. 2021. "Non-classical probabilities invariant under symmetries", *Synthese* 199:8507–8532.

[14] Pruss, Alexander R. 2021. "Underdetermination of infinitesimal probabilities." *Synthese* 198:777–799.

[15] Pruss, Alexander R. 2014. "Regular probability comparisons imply the Banach-Tarski Paradox." *Synthese* 191:3525–3540.

[16] Pruss, Alexander R. 2013. "Null probability, dominance and rotation." *Analysis* 73:682–685.

[17] Pruss, Alexander R. 2013. "Two kinds of invariance of full conditional probabilities." *Bulletin of the Polish Academy of Sciences – Mathematics* 61:277–283.

[18] Thong, Joshua. 2024. "Correction to: Primitive conditional probabilities, subset relations and comparative regularity." *Analysis* 84:593–594.

[19] Thong, Joshua. 2024. "Primitive conditional probabilities, subset relations and comparative regularity." *Analysis* 84:547–555.

[20] Tomkowicz, Grzegorz, and Wagon, Stan. *The Banach-Tarski Paradox*, 2nd ed., Cambridge: Cambridge University Press, 2016.

[21] van Fraassen, Bas C. 1976. "Representation of conditional probabilities." *Journal of Philosophical Logic* 5:417–430.