Clarifications on Countable Additivity

Kenny Easwaran

1 Introduction

Let Ω be the set of *states* that an agent deems possible, and let $\mathcal{F}$ be an algebra (or σ-algebra) of subsets of Ω, called *events*, and let $P\colon \mathcal{F} \to \mathbb{R}$ be the agent's degrees of belief. In this context, the theory of probability is the theory of what functions are eligible to play the role of P. To classify these theories, we use the concept of a *partition*, which is a set of events such that every element of Ω is in exactly one event in the set. I have elsewhere discussed questions about functions sending $\mathcal{F}$ to some other field than $\mathbb{R}$, but in this document I won't consider those. However, I do consider some attempts to replace the real-valued function by a partial ordering.

As I noted in my "Conditional Probability" (2019), theories of probability lie on a spectrum from weak theories with complex models, to strong theories with only simple models. Merely finitely additive probability is the theory that says that a function is eligible iff the values it assigns are all non-negative, and the values it assigns to any finite partition add up to 1. Countably additive probability strengthens the axioms by requiring that the values assigned to any countable partition must add up to 1. Fully additive probability strengthens this further, by requiring that the values assigned to any partition of any cardinality must add up to 1.

The theory of fully additive probability is equivalent to the theory of discrete probability. According to this theory, there is a countable sequence of numbers $\{p_i\colon i \in \mathbb{N}\}$, with $\sum_{i\in\mathbb{N}} p_i = 1$, and a countable sequence of states $\{s_i\colon i \in \mathbb{N}\}$, such that for any $A \in \mathbb{F}$, $P(A) = \sum_{i\in\mathbb{N}} p_i \cdot \chi_A(s_i)$, where $\chi_A(\omega)$ is the characteristic function of A, yielding 1 if $\omega \in A$ and 0 otherwise. An even stronger theory is the theory of finite probability, in which all but finitely many of the p_i are 0. In both the theory of discrete and finite probability, since the set consisting of all states whose singleton has probability 0 is also a set with probability 0, these theories are in many ways just as if we restricted Ω to be countable or finite.

The theory of finitely additive probability is the weakest theory, whose acceptable functions include some that are exceedingly complex. The theory of finite probability is the strongest theory, whose acceptable functions are the simplest. As I said in that earlier work, I suspect that some of the philosophical appeal of finitely additive probability has involved confusion between the weakest theory (with the most complex functions) and the theory with the simplest

functions (that is so strong that it can't be axiomatized). In my paper, "Why Countable Additivity?" (2013b), I aimed to defend the mathematical orthodoxy of the theory of countably additive probability, which is middling in both axiomatic strength and complexity of functions. I did this by proposing some philosophical principles that give rise to this theory, without explicitly building in any assumptions about cardinality, like the additivity principles I used above to characterize the theories.

Seidenfeld, Kadane, Schervish, and Stern (henceforth SKSS), in "Finite Additivity, Countable Additivity, and the Comparative Principle" (2024) raise some important mathematical issues around the principles I developed in that paper (and one developed by Stewart and Nielsen (2021), another response to my paper). Some of SKSS's results are phrased in dense mathematical language. In this note, I try to explain the important aspects of their results in ways that are more understandable, to help clarify their dialectical force. I show why I am dialectically unmoved by many of the points that they make. I still accept merely countably-additive probability as the best theory of probability for epistemology. However, I take their objection to my attempted extension of this argument to a comparative (rather than numerical) framework as a very significant one.

Their paper is in five sections. Section 1 is an overview. Section 2 discusses the infinitary betting arguments from my paper. Section 3 discusses issues in numerical applications of my comparative principle. Section 4 discusses issues in qualitative applications of my comparative principle. Section 5 discusses Stewart and Nielsen's use of a principle of disintegrability to analyze the strength of my comparative principle. My note follows the structure of their section 2, 3, 4, and the reader might find it helpful to have their paper and this note both on hand to help understand both.

2 Infinitary betting arguments

2.1 Fair vs favorable bets

Define a *bet* to be a function $f : \Omega \to \mathbb{R}$ that is $\mathcal{F}$-measurable and takes on only two values. Equivalently, a bet f can be defined by a triple (E, a, b), where $E \in \mathcal{F}$, and $a, b \in \mathbb{R}$, where, for $\omega \in \Omega$, $f(\omega) = a$ if $\omega \in E$ and $f(\omega) = b$ otherwise. We say that a, b are the *payoffs* of the bet. We can say that a bet is *favorable* iff $a \cdot P(E) + b \cdot P(\neg E) > 0$, *unfavorable* iff $a \cdot P(E) + b \cdot P(\neg E) < 0$, and *fair* iff $a \cdot P(E) + b \cdot P(\neg E) = 0$.

We can say that one bet is *sweetened* with respect to one or more of its payoffs compared to another, iff that payoff is strictly higher, and all other payoffs are at least as high. When $0 < P(E) < 1$, for every favorable bet on E, there are fair bets on E with respect to which it is sweetened for one payoff, for the other payoff, or both. But when $P(E) = 0$, a favorable bet must have the $\neg E$ payoff sweetened compared to a fair bet, and when $P(E) = 1$, a favorable bet must have the E payoff sweetened. (Sweetening only the other payoff of

either type of fair bet simply yields another fair bet.) So I suggest that it is better to think of a favorable bet generally as one where *all* payoffs have been sweetened.

As I show in "Why Countable Additivity?", P is countably additive iff there is no set of favorable bets whose sum is everywhere negative. As I and SKSS note, P is fully additive iff there is no set of fair bets whose sum is everywhere negative. In their first challenge in Section 2, the say, "We do not see why the bookie is obliged to accept an infinite set of strictly favorable gambles but not an infinite set of fair gambles." Their idea is that since I need to oblige bookies to accept infinite sets of favorable bets in order to rule out merely finitely-additive probability functions, they think I should also oblige bookies to accept infinite sets of fair bets, and this would rule out the merely countably-additive probability functions as well. If I thought bookies were obliged to accept both favorable bets and fair bets individually, then there would be no clear reason why I should think they are obliged to accept sets of one without being obliged to accept sets of the other.

But I don't think bookies are obliged to accept both favorable bets and fair bets, even individually. I claim that bookies are generally obliged to accept favorable bets and to reject unfavorable bets, but deny any substantive general claim about what bookies are obliged or permitted to do with fair bets, individually or in combination.

For motivation, note that if there are non-empty events E with $P(E) = 0$, then there are fair bets where the agent deems it possible that they lose but impossible that they gain, namely the bet represented by the triple $(E, -1, 0)$. Similarly, if F is any probability 0 event disjoint from E, then $(E \cup F, -1, 0)$ and $(F, +1, 0)$ are two bets that are individually fair, whose sum is the previous bet. I don't believe that the payoffs of the bets, and the probabilities of the relevant events, are sufficient to say which, if any of these "fair" bets are ones that the agent is obliged to accept, obliged to reject, or permitted to do either. I only claim that *favorable* bets are ones that agents are obliged by their probabilities to accept.

2.2 Favorable called-off bets

The second challenge they raise in Section 2 involves called-off bets based on a conditional probability function. They assume that there is a conditional probability function such that $P(A|B) \in \mathbb{R}$ whenever $A, B \in \mathcal{F}$ and $B \neq \emptyset$. This is the theory of "Coherent Conditional Probabilities", derived from the work of Dubins and de Finetti, but as they note in footnote 5, I instead accept the theory of "Regular Conditional Distributions", derived from the work of Kolmogorov. I give lengthier discussions of both theories in Sections 2 and 3 of my (2019).

We can define a general *three-way bet* as a measurable function $f\colon \Omega \to \mathbb{R}$ that takes on only three values. We can say that a three-way bet is a *fair*

called-off bet at stakes x on A given B iff it satisfies

$$f(\omega) = \begin{cases} x \cdot (1 - P(A|B)) & \text{if } \omega \in A \cap B, \\ -x \cdot P(A|B) & \text{if } \omega \in B \setminus A, \\ 0 & \text{otherwise.} \end{cases}$$

They then define a *favorable called-off bet on A given B* to be a three-way bet where the payoffs in both the $A \cap B$ and the $B \setminus A$ case are sweetened compared to a fair called-off bet on A given B, while the payoff outside of B is still 0. Again, when $0 < P(A|B) < 1$, by comparing to a fair called-off bet with different stakes, we could re-describe it by just sweetening the first payoff or just sweetening the second, but in order to deal with all cases uniformly, it is better to describe their concept of favorability by sweetening both the first and second payoffs. (As I mention later, I am worried about a concept of favorability where only a proper subset of the payoffs are sweetened, but I grant it to explain their argument.)

For $E \in \mathcal{F}$, define a *partition of E* to be a set of events such that every element of E is in exactly one event in the set. Citing Schervish et al. (2017), SKSS note that, in the theory of Coherent Conditional Probabilities, if a probability function fails to be fully additive, then there is some event E, and some partition $\Pi = \{h_i \colon i \in I\}$ such that $P(E) < \inf_{i \in I} P(E|h_i)$. Using this, they construct a set of called-off bets on E given the members of the partition, that are all favorable in their sense, that, together with a favorable bet on E, have a negative sum.

I have two responses to this challenge. The first is to object to their account of a favorable called-off bet. As suggested above, I think that sweetening some subset of the payoffs of a bet is not generally sufficient to convert a fair bet into a favorable one. In particular, in the partitions needed for their argument, each $P(h_i)$ is 0, and I am unconvinced that their sweetening (which leaves the 0 payoff in $\neg h_i$ situations unchanged) is sufficient to make the agent consider these called-off bets to be truly favorable. But I am not fully convinced that this is a sufficient response, and instead put more weight on my second.

The second response is to call into question the theory of Coherent Conditional Probabilities that they assumed in this argument, which is essential to create the example. In their footnote 5, they note that I prefer the theory of Regular Conditional Distributions. But they note, "rcd's are not defined unless both unconditional and conditional probability is countably additive. Hence, in the debate whether probability might be merely finitely additive, the question here is begged by appeal to rcd's." It's true that it would be question-begging for me to appeal to the full theory of Regular Conditional Distributions to respond to this challenge. But this is not my response.

The theory of Coherent Conditional Probabilities assumes that for every event A, and non-empty event B, there is one value $P(A|B)$ that is the conditional probability of A given B. The theory of Regular Conditional Distributions can be stated with a weaker structural assumption, that for every event A, and every non-empty event B, and every epistemologically relevant partition

Π containing B, there is a value $P_\Pi(A|B)$, but there is no general assumption about how $P_\Pi(A|B)$ and $P_{\Pi'}(A|B)$ must relate. The theory of Coherent Conditional Probabilities can be interpreted as starting with this same framework, but imposing as an additional axiom that $P_\Pi(A|B) = P_{\Pi'}(A|B)$.

Without this additional requirement, the argument of Schervish et al. (2017) doesn't go through. (Their construction involves changing partitions several times while assuming that conditional probabilities remain the same.) If countable additivity can be established for unconditional probability (as I think the argument involving favorable two-way bets mentioned in the previous subsection does), then further arguments may be able to establish the full theory of Regular Conditional Distributions. (Various such arguments are put forward by Easwaran (2008, 2013a, 2019), Nielsen (2022).) But the full theory of Regular Conditional Distributions is not required to resist the argument SKSS make — just the rejection of the structural assumption that conditional probability depends only on the two events, and not on the conditioning partition.

3 The Comparative Principle for numerical probabilities

The core novel argument of my (2013b) involved what I called "the Comparative Principle", which SKSS sometimes refer to as $\mathcal{C}$. When P_1 and P_2 are probabilities on the same set Ω of states and $\mathcal{F}$ of events, the numerical version of the Comparative Principle for P_1 can be stated as follows:

> There is no partition Π and no probability function P_2 such that for every $E \in \Pi$, $P_2(E) > P_1(E)$.

In their Section 3, they claim to provide two challenges to the numerical version of the Comparative Principle. However, neither one is actually a challenge to the Comparative Principle itself. Rather, the first is a setting in which countably additive probability functions don't exist, while the second is a setting in which countably additive probability functions are automatically fully additive. I point out that the first setting is not one that I aimed to talk about, and I claim the second setting is no more a challenge to the countably additive theory of probability than the existence of finite probability spaces is.

3.1 Amer's Theorem about free Boolean algebras

The first challenge SKSS raise cite the theorem that there is no countably-additive probability function on a free Boolean algebra with infinitely many generators. The important thing to note about this theorem is that the concept of "probability function" and of "partition" that it appeals to are distinct from the ones that I have been working with so far, and thus the concept of "countable additivity" that is impossible in this setting is not the one I care about.

I have defined a probability function as taking inputs from an algebra $\mathcal{F}$ of events that are subsets of a set Ω of states. A partition is defined to be a set

of events such that every element of Ω is in exactly one of them. The setting of Amer's theorem does not involve Ω. Instead, $\mathcal{F}$ is defined purely algebraically. As SKSS note, we can think of the free Boolean algebra on a set of generators as the Lindenbaum-Tarski algebra — that is, we can represent each generator as a distinct atomic sentence letter A_i, and each element of the algebra can be represented as a finite combination of these letters with the Boolean operations $\wedge, \vee, \neg$, with two such combinations representing the same element of the algebra iff the formulas they encode are classically logically equivalent. For $\phi, \psi \in \mathcal{F}$, we say $\phi \preceq \psi$ iff $\phi \wedge \psi = \phi$, or equivalently iff $\phi \vee \psi = \psi$. This corresponds to ϕ classically logically entailing ψ.

It is natural to identify this algebra with one presented set-theoretically, but as I will show shortly, Amer's result is importantly in tension with this identification. Let Ω be the set of all truth-value assignments to the sentence letters that generate $\mathcal{F}$. Then any Boolean combination of sentence letters can be identified with the set of truth-value assignments that make it true. If ϕ and ψ are Boolean combinations associated with sets S_1 and S_2 of truth-value assignments, then it is straightforward to check that $\phi \wedge \psi$ is associated with $S_1 \cap S_2$, $\phi \vee \psi$ is associated with $S_1 \cup S_2$, and $\neg\phi$ is associated with $\Omega \setminus S_1$. Similarly, if two formulas are classically logically equivalent, then they are associated with the same set.

However, as SKSS note at a few points in the appendix, the concept of partition that is used in Amer's result is not the same as what one would expect from this identification. The concept of a "partition" used in this result (I will use scare quotes to distinguish it from the set-theoretic concept defined above) is that $\Pi = \{\psi_i \colon i \in I\}$ is a "partition" iff for $i \neq j$, $\psi_i \wedge \psi_j = \bot$, while $\bigvee_{i \in I} \psi_i = \top$, where $\bot$ and $\top$ are any contradiction or tautology respectively. $\wedge$ is the ordinary Boolean operation mentioned above, and it corresponds to $\cap$ in the set-theoretic identification, while $\bigvee$ is new, and does *not* correspond to $\bigcup$ in the set-theoretic identification.

In particular $\bigvee$ is defined as the infinitary join operation. That is, $\bigvee_{i \in I} \psi_i$ is defined as the unique ϕ, if there is one, such that each $\psi_i \preceq \phi$, and such that for any $\theta \in \mathcal{F}$ with each $\psi_i \preceq \theta$, $\phi \preceq \theta$. That is, it is the least upper bound (if a unique one exists) within the algebra of the set of elements, or equivalently, the logically strongest formula (if a unique one exists) within the language that is entailed by each formula in the set.

To see how this differs from the set-theoretic union, let $\{A_i : i \in I\}$ be the set of sentence letters that generate $\mathcal{F}$, and consider $\bigvee_{i \in I} A_i$. This infinite join is the logically strongest formula in the language that is entailed by each sentence letter. If I is infinite, then within the language, the only formula that is entailed by each sentence letter is the tautology, and thus there is a unique strongest formula entailed by each sentence letter. (If some formula is not a tautology, then it is certainly not entailed by any sentence letter it doesn't contain.) However, set-theoretically, the union of the sets of truth-value assignments making each sentence letter true is *not* the set of all truth-value assignments. In particular, this union will not include the single truth-value assignment making all sentence letters false. This gets missed within

the Lindenbaum-Tarski algebra, because there is no formula true only on this assignment, because formulas can only be built with finitary Boolean operations.

Now we can see how the argument they give goes wrong if we try to apply the proof of Amer's algebraic result to some clearly countably-additive probabilities on the set-theoretic algebra.

Let Ω be the set of all truth-value assignments on a countable sequence of sentence letters, let $\mathcal{F}$ be the sets corresponding to finite Boolean formulas on these sentence letters, and let P be the standard countably-additive probability function on this space that treats each sentence letter as probabilistically independent, with probability 1/2. That is, for any ϕ, $P(\phi)$ will just be the fraction of rows in the relevant truth-table for ϕ on which ϕ is true. The proof SKSS give in their Appendix shows how to construct a "partition" such that the sum of the probabilities of its elements is less than ϵ. I will walk through the steps in their proof for this particular probability function, and $\epsilon = 1/2$.

First, we choose k with $1/(2^k - 1) < \epsilon$, so in this case $k = 2$. Then we enumerate successive blocks of sentence letters of length $k, 2k, 3k, \ldots$. That is $b_1 = \{A_1, A_2\}$, $b_2 = \{A_3, A_4, A_5, A_6\}$, and so on. Now we use each block b_j to generate a sentence β^j_m whose probability is at most $1/2^{kj}$. This is always possible, because there are 2^{kj} distinct conjunctions of sentence letters and negations of sentence letters in b_j. In this case, because the probability distribution is so simple, we can just let β^j_m be the conjunction of all the sentence letters in b_j. Then we generate sentences ψ_j by letting

$$\psi_j = \neg\beta^1_m \wedge \neg\beta^2_m \wedge \ldots \wedge \neg\beta^{j-1}_m \wedge \beta^j_m.$$

By construction, each β^j_m has probability at most $1/2^{kj}$, and the ψ_j conjoin each one with the negations of the previous ones to make them pairwise disjoint. That is, ψ_1 says the first two sentence letters are true, ψ_2 says that at least one of the first two is false but the next four are all true, ψ_3 says that at least one of the first six is false but the next six are all true, ψ_4 says that at least one of the first 12 are false but the next eight are all true, and so on. It's straightforward to see that these propositions have probabilities less than $\frac{1}{4}, \frac{1}{4^2}, \frac{1}{4^3}, \frac{1}{4^4}, \ldots$, and so the sum of their probabilities is less than $\frac{1}{3}$. And for the same reason that the infinite join of the sentence letters themselves had to be the tautology, the infinite join of these ψ_i must also be the tautology, and so the ψ_i form a "partition".

Thus, the probability function that gives each sentence letter independent probability 1/2 of being true, which is countably-additive as a set-theoretic probability function, is not "countably-additive" on the Lindenbaum-Tarski algebra. In fact, the construction used above can generate a "partition" on which the sum of the probabilities is arbitrarily small, using this probability function that is clearly countably additive on the full set Ω of truth-value assignments.

As Snow Zhang suggested to me, this can be seen even more clearly with the *fully* additive probability function that assigns all probability to the singleton of the truth-value assignment that makes all sentence letters false. Then we can consider the "partition" A_1, $\neg A_1 \wedge A_2$, $\neg A_1 \wedge \neg A_2 \wedge A_3$, ... Because each element of this "partition" requires one of the sentence letters to be true, it has

probability 0. So this one "partition" witnesses maximal failure of "countable additivity". Which is not surprising, because all probability mass is concentrated on the one element of Ω that fails to be within any element of this "partition".

Thus, while Amer's theorem is an interesting result about "probability functions" on free Boolean algebras, it does not seem to pose any challenges to probability functions defined set theoretically.

3.2 Ulam matrices

The second challenge they propose involves the concept of an Ulam matrix. For our purposes, the actual concept of an Ulam matrix is not essential. What is essential is that if the cardinality of Ω is not extremely special ("real-valued measurable"; it is in fact independent of ZFC whether any real-valued measurable cardinalities exist!) then there is an Ulam matrix on Ω, which is some collection of subsets of Ω, such that if all of them are members of $\mathcal{F}$, then any countably additive probability on $\mathcal{F}$ must in fact be fully additive (i.e., discrete, giving all its probability to some countable set of points).

I claim that it is no challenge to the merely countably additive theory of probability to point out that there are some probability spaces in which countably additive probabilities are fully additive. We already know of some such probability spaces, since (for instance) any countable probability space, or finite probability space, is such that every countably additive probability is automatically fully additive. The point of the merely countably additive theory of probability isn't to rule out the simple discrete probability spaces or finite probability spaces or the fully additive probability functions on those spaces — it is to rule out the very complex merely finitely additive probability functions, while allowing for the moderately complex merely countably additive probability functions that exist on many well-known well-behaved probability spaces.

Since Borel measure and Lebesgue measure are countably additive without being fully additive, it is clear that the σ-algebras that they are defined on must not contain an Ulam matrix. Similarly for any other familiar algebra on which anyone has defined a countably additive probability function that is not fully additive. In fact, it is hard to demonstrate that any σ-algebra other than a (nearly) complete power set *does* include an Ulam matrix. It is not hard to show that if the cardinality of Ω is $\aleph_1$, then its full power set must include an Ulam matrix — but it is rare for theorists of countably additive probability to work with $\mathcal{F}$ as the full power set of Ω. (By contrast, this is common with discrete or finite probability, where the power set of Ω is relatively simple, and also common with finitely additive probability, where theorists already embrace greater complexity.)

4 The Qualitative Comparative Principle

The challenge they raise in Section 4 of their paper is the strongest challenge to my views. It shows that I must deny an initially plausible principle in order for my Qualitative Comparative Principle to entail countable additivity without entailing full additivity. My current inclination is towards thinking that the purely comparative formulation of probability is weaker than I had thought.

I start by discussing the motivation for the qualitative version of the Comparative Principle in my (2013b), even though I now think it does less work than I had hoped. In some contexts it is useful to consider a concept of comparative probability that is, in a sense, finer-grained than numerical probability. To motivate this finer-grained comparative probability, consider a situation in which every real number in the interval $[0, 1]$ is a possibility, but each is equally likely, and yet we might want to be able to note that each real number in this interval is strictly more likely than any real number outside the interval, because the latter are impossible. This set of comparisons can't be represented with only a numerical probability, because the only way to represent each real number in the interval as equally likely with a numerical probability would set $P(\{x\}) = 0$ for any x, and thus would fail to directly represent the fact that $\{x\}$ inside the interval is more likely than $\{x\}$ outside the interval.

We can do this by supplementing P with a *qualitative probability* represented by relations $\rhd, \trianglerighteq, \equiv$ on the elements of $\mathcal{F}$. We assume that all three relations are transitive, that $\rhd$ is irreflexive while the others are reflexive, that $A \trianglerighteq B$ iff either $A \rhd B$ or $A \equiv B$, and that $A \equiv B$ iff $A \trianglerighteq B$ and $B \trianglerighteq A$. We further require that these comparisons *almost agree* with P; that is, if $P(A) > P(B)$ then $A \rhd B$, but we make no further assumptions about the comparison of A and B if $P(A) = P(B)$. We say that $\rhd$ is *regular* iff for every nonempty $A \in \mathcal{F}$, $A \rhd \emptyset$. If, additionally, we have $\{x\} \equiv \{y\}$ for every $x, y \in [0, 1]$, then a regular qualitative probability would satisfy the motivations of the previous paragraph.

In my earlier paper, I hypothesized that we could extend this concept of qualitative probability to a set of comparisons $\succ, \succeq, \approx$ across two probabilities P_1 and P_2, where $(P_1, A) \succ (P_2, B)$ is interpreted as "P_1 makes A more probable than P_2 makes B". I wanted to impose a Qualitative Comparative Principle for P_1:

> There is no partition Π and no probability P_2 such that for every $E \in \Pi$, $(P_2, E) \succ (P_1, E)$.

I then hypothesized that this Qualitative Comparative Principle would entail countable additivity without entailing full additivity.

SKSS show that, given some specific assumptions about $\succ, \succeq, \approx$ and the relations $\rhd_1, \trianglerighteq_1, \equiv_1$ and $\rhd_2, \trianglerighteq_2, \equiv_2$ that almost agree with P_1 and P_2, if $\rhd_1$ is regular, then P_1 satisfies the Qualitative Comparative Principle iff it is fully additive. While many of the assumptions they make seem appropriate to me, there are a few that I question. I admit that the specific principles I must reject seem initially plausible, but I currently think that it is ok to reject them. If these assumptions can be shown to be unnecessary for their argument, or if

they can be properly philosophically motivated in a way that I accept, then I would agree that this would refute the hypothesis I made about the Qualitative Comparative Principle. But I am not yet fully convinced. Instead, I just think that the theory of purely comparative probability is much weaker than I had initially thought.

The further assumptions SKSS make that do seem perfectly well-motivated to me are the following. Their Axiom 2: $\succ, \succeq, \approx$ extend each of $\rhd_1, \trianglerighteq_1, \equiv_1$ and $\rhd_2, \trianglerighteq_2, \equiv_2$, respectively. That is, if $A \rhd_1 B$ then $(P_1, A) \succ (P_1, B)$, and similarly. Their Axiom Qual_2: for each $A \in \mathcal{F}$, $\Omega \trianglerighteq A \trianglerighteq \emptyset$. That is, nothing is qualitatively more probable than certainty or less probable than impossibility. Their Axiom Qual_4: Whenever $A \cap C = B \cap C = \emptyset$, then $A \cup C \trianglerighteq B \cup C$ iff $A \trianglerighteq B$, and $A \cup C \rhd B \cup C$ iff $A \rhd B$. This is the analog of finite additivity for qualitative probability. (Their Axiom Qual_3 is a consequence of each $\rhd_i$ almost agreeing with P_i.)

The assumptions they make that I am suspicious of are the following. One part of their Axiom Qual_1 is that all pairs of events are comparable. That is, for any $A, B \in \mathcal{F}$, either $A \trianglerighteq B$ or $B \trianglerighteq A$ or both, or equivalently, either $A \rhd B$ or $B \rhd A$ or $A \equiv B$. (The rest of their Axiom Qual_1 is just the assumptions about transitivity and reflexivity that I made above.) Their Axiom 3 (with a slight change of notation):

> Suppose that, restricted to a subalgebra $\mathcal{G} \subseteq \mathcal{F}$, $\trianglerighteq_1 = \trianglerighteq_2$. Then, for $A, B \in G$, $(P_1, A) \succeq (P_2, B)$ iff $A \trianglerighteq_1 B$, and $(P_2, A) \succeq (P_1, B)$ iff $A \trianglerighteq_1 B$.

That is, if the two qualitative probabilities agree on some subalgebra, then qualitative comparisons *across* the two probability functions *also* agree.

They give their argument first in a "heuristic example", with particular probability functions P_1 and P_2, and then more generally. In the heuristic example, they show that the existence of a regular uniform qualitative probability on a space like the half-open unit interval is sufficient to lead to a violation of the Qualitative Comparative Principle, and their more general argument shows how to construct an example like this for any merely countably additive probability function supplemented with a regular qualitative probability. I will just discuss the heuristic example, to show how I must reject the assumptions Axiom Qual_1 and Axiom 3 mentioned above. Rejecting the general argument just involves rejecting these same assumptions, but the instances that I must reject in the heuristic example are some of the most plausible ones, so if it is reasonable to reject these instances, then I think it is reasonable to reject the full argument.

Their heuristic example involves a space Ω whose elements are of the form (H, x) or (T, x), where $x \in (0, 1]$. We think of P_1 as the regular uniform probability on the subset consisting of the states (H, x), where each state of the form (T, x) is treated as impossible, and P_2 as the regular uniform probability on the subset consisting of the states (T, x), where each state of the form (H, x) is treated as impossible.

Consider the four families of events $A_x = \{(H, x), (T, x)\}$, $H_x = \{(H, x)\}$, $T_x = \{(T, x)\}$, and $C_x = \{(H, x), (T, x/2), (T, .5 + x/2)\}$. Note that $\rhd_1$ and $\rhd_2$

agree on the algebra generated by the events A_x, and so Axiom 3 requires that $(P_1, A_x) \approx (P_2, A_x)$ — though this instance is plausible in any case because P_1 and P_2 were both motivated to be uniform on their respective parts of Ω, and A_x just contains the single corresponding state of each.

Now since P_1 treats each state (T, x) as impossible, we must have $A_x \equiv_1 H_x \equiv_1 C_x$, since A_x and C_x differ from H_x only in including some impossible states, and similarly $A_x \equiv_2 T_x$. But since P_2 is said to be uniform, it seems plausible that $T_x \equiv_2 T_{x/2}$ (this is the step where they make use of the comparability claim from Axiom Qual_1 in their full argument). By regularity of P_2 (together with Axiom Qual_4) we have $T_{x/2} \lhd_2 C_x$, since $C_x = H_x \cup T_{x/2} \cup T_{.5+x/2}$, and $\emptyset \lhd_2 T_{.5+x/2}$. Putting these comparisons all together using Axiom 2, we get that $(P_1, C_x) \prec (P_2, C_x)$. But the C_x form a partition of Ω, and thus this is a violation of the Qualitative Comparative Principle.

In order to resist this consequence, I must in general either: deny that $(P_1, A_x) \approx (P_2, A_x)$, even though the P_i are identical distributions over two corresponding domains, and A_x just consists of one corresponding state from each; or deny that $T_x \equiv_2 T_{x/2}$. The first is a denial of their Axiom 3, and the second is a denial of the comparability claim from their Axiom Qual_1. In the heuristic example, this amounts to denying either that a regular qualitative probability can in fact be uniform over the interval $(0, 1]$, or that corresponding elements in the domains of two uniform regular qualitative probabilities on structurally identical domains must be equiprobable.

The Qualitative Comparative Principle being incompatible with uniform regular qualitative probabilities on the interval $(0, 1]$ would be an interesting extension of the well-known fact that countable additivity is incompatible with uniform numerical probability on a countably infinite domain. But the situation is not quite as extreme as that for countably additive probabilities on a countable domain. In that case, it is necessary that many singletons of states in the domain have distinct numerical probabilities, and thus one of them is strictly more likely than the other. But in the case I describe here, the natural thought is that singletons of the domain are *incomparable* to each other, rather than one being strictly more likely than the other.

To motivate this incomparability, note that if x is chosen uniformly at random from the interval $(0, 1]$, then x^2 is chosen *non*-uniformly from the interval $(0, 1]$, concentrated towards the lower numbers. Just as we might have naively thought that we should have $\{x\} \equiv \{y\}$ when the probability is uniform, we might have naively thought that we should have $\{x^2\} \rhd \{y^2\}$ when $x < y$, since the distribution is concentrated towards the lower numbers. These naive thoughts are incompatible, and I think the best response is just to say that many singletons are incomparable in probability in both distributions, rather than saying that they are equal, or have one specific comparison or the other.

To motivate the denial of $(P_1, A_x) \approx (P_2, A_x)$, note that we can describe an isomorphic version of the case somewhat differently. Consider a case where Ω' consists of the states (H, x) where $x \in (0, 1]$ and (T, x) where $x \in (0, 2]$ that we get just by re-labeling the state (T, x) from Ω as the state $(T, 2x)$ in Ω'. Then P_1 is regular and uniform on the states (H, x) with $0 < x \leq 1$, while P_2'

is regular and uniform on the states (T, x) with $0 < x \leq 2$. Under the original labeling, it was plausible that $(P_1, H_x) \approx (P_2, T_x)$, because they were uniform over domains that seemed intuitively to be "just as long". But under the new labeling, this is much less plausible, because P_2 is uniform over a domain that is "twice as long" as the domain over which P_1 is uniform. I think that the most plausible response is to deny the comparability of singletons from the two domains, despite the intuitive temptation to assume it.

Perhaps there are some stronger philosophical arguments that can be given in favor of comparability of the relevant probabilities, or perhaps a version of SKSS's result can be demonstrated that doesn't require the two principles that I deny. In that case, I would agree that the Qualitative Comparative Principle entails more than just countable additivity. But for now, I conjecture that the Qualitative Comparative Principle, together with the slightly weaker structural assumptions I make on qualitative probability, entail only countable additivity, together with the familiar claim that probability over a countable domain must give some singleton a strictly higher probability than another, and the novel claim that probability over an uncountable domain must make many singletons incomparable rather than allowing for all to be equiprobable.

This ends up meaning that I am committed to a lot of incomparability in apparently "uniform" probability distributions. This is a severe limitation of what I had hoped to achieve with a theory based on the Qualitative Comparative Principle.

References

Easwaran, K. (2008). *The Foundations of Conditional Probability*. PhD thesis, University of California, Berkeley.

Easwaran, K. (2013a). Expected accuracy supports conditionalization — and conglomerability and reflection. *Philosophy of Science*, 80(1):119–142.

Easwaran, K. (2013b). Why countable additivity? *Thought*, 1(4):53–61.

Easwaran, K. (2019). Conditional probability. In Pettigrew, R. and Weisberg, J., editors, *Open Handbook of Formal Epistemology*, chapter 4, pages 131–198. PhilPapers.

Nielsen, M. (2022). Regular conditional probabilities and strictly proper loss functions. *Statistics & Probability Letters*, 185.

Schervish, M. J., Seidenfeld, T., and Kadane, J. B. (2017). Nonconglomerability for countably additive measures that are not κ-additive. *Review of Symbolic Logic*, 10:284–300.

Seidenfeld, T., Kadane, J. B., Schervish, M. J., and Stern, R. B. (2024). Finite additivity, complete additivity, and the comparative principle. *Erkenntnis*.

Stewart, R. and Nielsen, M. (2021). Conglomerability, disintegrations and the comparative principle. *Analysis*, 81:479–488.