Qualitative Probability and Infinitesimal Probability

Nicholas DiBella

(Work in progress. September 7, 2018)

Abstract

Infinitesimal probability has long occupied a prominent niche in the philosophy of probability. It has been employed for such purposes as defending the principle of regularity, making sense of rational belief update upon learning evidence of classical probability 0, modeling fair infinite lotteries, and applying decision theory in infinitary contexts. In this paper, I show that many of the philosophical purposes infinitesimal probability has been enlisted to serve can be served more simply and perspicuously by appealing instead to *qualitative probability*—that is, the binary relation of one event's being at least as probable as another event. I also that show that qualitative probability has comparable (if not greater) representational power than infinitesimal probability. These considerations suggest that qualitative probability provides a superior framework to infinitesimal probability for theorizing about a variety of philosophical contexts.

1 Introduction

A (positive) infinitesimal is a number that is 'infinitely smaller' than any positive real number yet is still positive. More precisely, an infinitesimal x is a positive number such that, for any positive real number y, there is no natural number n such that $y < nx$. Probability functions that can take on infinitesimal values have been employed for a number of philosophical purposes—to defend the principle of regularity (Lewis [1980], Skyrms [1980]), to make sense of rational belief update upon learning evidence of classical probability 0 (Lewis [1980], Skyrms [1980]), to model fair infinite lotteries (Wenmackers and Horsten [2013]), to apply decision theory in infinitary contexts (Hájek [2003a], Pivato [2014]), to measure objective chance (Lewis [1986], Hofweber [2014]), among others.

Although the use of infinitesimal probability to serve these purposes has been met with a fair amount of criticism,[1] I will not add to the criticisms or defenses in this paper. My main aim is simply to show that we don't *need* infinitesimal probability to serve many of these purposes. We can serve them more simply and perspicuously by appealing instead to *qualitative probability*—that is, the binary relation of one event's being at least as probable as another event.[2] As a result, critics of infinitesimal probability may find qualitative probability to be an attractive alternative framework for theorizing about infinitary probabilistic scenarios.

Another aim of this paper is to demonstrate the extraordinary representational power of qualitative probability. Although it is commonly thought—as, for example, Meacham and Weisberg [2011] claim—that qualitative probability is representationally impoverished compared to numerical probability, the case against qualitative probability has been greatly overstated. I will show that, not only does qualitative probability have comparable—if not greater—representational power than real-valued probability, it even has comparable—if not greater—representational power than infinitesimal probability.[3]

Several non-classical theories of probability have been developed that allow probability functions to take on infinitesimal values. In this paper, I focus on the theory recently developed by Benci *et al.* [2013, 2016]—Non-Archimedean Probability Theory (NAP). However, many of my points will apply to other theories of infinitesimal probability as well.[4]

The structure of this paper is as follows. In Section 2, I review NAP as well as the motivations for adopting it. In particular, I discuss four intuitively plausible desiderata for a theory of probability that Benci *et al.* argue

[1]See Benci *et al.* [2016, Section 4] for notable objections.

[2]Qualitative probability is also known as 'comparative probability' or, in epistemic contexts, 'comparative confidence'. See Fine [1973] and Fishburn [1986] for overviews.

[3]Thus, the present paper substantially extends the work of Stefánsson [forthcoming], who argues against Meacham and Weisberg's claim.

[4]See Benci *et al.* [2016, Section 2.3] for discussion of other such theories.

NAP satisfies. In Section 3, I critically assess these desiderata and provide weakened—but no less plausible—versions of them that are neutral between numerical probability and qualitative probability. In Section 4, I provide a theory of qualitative probability—which I call Qualitative Probability Theory (QP)—that I argue satisfies these weakened desiderata at least as well as NAP. In Section 5, I argue that QP gives rise to additional philosophical dividends that Benci *et al.* argue NAP gives rise to. In Section 6, I describe two types of 'conceptually possible' probabilistic scenarios that NAP cannot represent but which QP readily can. Technical results are proven in the Appendix.

2 Non-Archimedean Probability Theory (NAP)

Benci *et al.* [2016] argue that any theory of probability should satisfy four intuitively plausible desiderata:

1. **Regularity.** The probability of any possible event is strictly larger than the probability of any impossible event.

2. **Totality.** Every event has a probability. That is, every subset of the sample space is assigned a probability value.

3. **Perfect Additivity.** The probability of an arbitrary union of mutually disjoint events is equal to the sum of the probabilities of the separate events (where 'sum' has to be defined in an appropriate way in the infinite case).

4. **Weak Laplacianism.** The probability theory allows for a mathematical representation of any 'conceptually possible' probabilistic scenario. In particular, the theory allows for a mathematical representation of a uniform probability distribution on sample spaces of any cardinality—finite, countable, or uncountable—as well as many other probability ratios between the atomic events.

As Benci *et al.* [2016, Section 2.2] observe, classical probability theory—characterized by the axioms of Kolmogorov [1950]—does not satisfy all of these desiderata for arbitrary sample spaces. To overcome these limitations, Benci *et al.* have developed an alternative theory of probability—Non-Archimedean Probability Theory (NAP). On NAP, probability functions can take on real values as well as other values in 'non-Archimedean' fields—that is, fields that extend the reals and contain infinitesimals. Benci *et al.* argue that NAP, unlike classical probability theory, does satisfy all of the above desiderata for arbitrary sample spaces. They take this to be a substantial point in favour of adopting NAP as a theory of probability.

In this section, I describe the structure of NAP and clarify the above motivations for adopting it. Since the axioms of NAP are analogous to the axioms of classical probability theory, I begin by reviewing classical probability theory in Section 2.1. In Section 2.2, I lay out the axioms of NAP. In Section 2.3, I describe Benci *et al.*'s argument that NAP satisfies the above desiderata.

2.1 Classical probability theory

The central notion of classical probability theory is that of a probability triple, $\langle \Omega, \mathcal{F}, P_K \rangle$. Ω is known as the 'sample space'. It is the set of all possible outcomes (or 'atomic events'). $\mathcal{F}$ is known as the 'event space'. It is a subset of $\mathcal{P}(\Omega)$—the power set of Ω—and is a σ-algebra on Ω. Finally, P_K is a function that satisfies the following axioms:

K0. Domain and Range. The events are the elements of $\mathcal{F}$, and the probability function is a function

$$P_K : \mathcal{F} \to \mathbb{R},$$

where $\mathbb{R}$ is the set of real numbers.

K1. Non-negativity. For all A in $\mathcal{F}$,

$$P_K(A) \geq 0.$$

K2. Normalization. $P_K(\Omega) = 1$.

K3. Finite Additivity. If A and B are events and $A \cap B = \emptyset$, then

$$P_K(A \cup B) = P_K(A) + P_K(B).$$

K4. Continuity. Let

$$A = \bigcup_{n \in \mathbb{N}} A_n,$$

where $A_n \subseteq A_{n+1}$ are elements of $\mathcal{F}$ and $\mathbb{N}$ is the set of natural numbers. Then,

$$P_K(A) = \lim_{n\to\infty} P_K(A_n).$$

Note that the conjunction of K3 and K4 is equivalent to the requirement that P_K satisfies countable additivity. Countable additivity is the countable extension of finite additivity:

> Countable Additivity. Let $A = \{A_1, A_2, \ldots\}$ be a countable set of mutually disjoint events. Then
>
> $$P_K(A_1 \cup A_2 \cup \ldots) = P_K(A_1) + P_K(A_2) + \ldots$$

As I describe in the next section, NAP generalizes finite and countable additivity by generalizing Continuity.

2.2 The axioms of NAP

In this section, I lay out the axioms of NAP. I omit some of the technical details, as they will not matter for the purposes of this paper. For complete details, see Benci *et al.* [2013, 2016].

The central notion in NAP is that of a NAP space, $\langle \Omega, P, \mathcal{U}\rangle$. As in classical probability theory, Ω is the sample space, i.e., the set of possible outcomes. Here, however, P is a generalized probability function—defined on *all* of the subsets of Ω. Moreover, P takes on values in a field that may contain both real numbers and infinitesimal numbers. Finally, $\mathcal{U}$ has no analogue in classical probability theory. It is a free ultrafilter on the set of all non-empty finite subsets of Ω. As I will describe shortly, $\mathcal{U}$ is necessary for formulating a generalized infinitary additivity principle.

Here are the first four axioms of NAP:

NAP0. Domain and Range. The events are all of the subsets of Ω. The NAP function P is total:

$$P : \mathcal{P}(\Omega) \to \mathcal{R},$$

where $\mathcal{R}$ is a 'superreal' field—that is, an ordered field that contains the real numbers as a subfield. Unlike in classical probability theory, the range $\mathcal{R}$ is not fixed for every sample space. Rather, $\mathcal{R}$ is partially determined by the cardinality of Ω. In general, the larger Ω is, the larger $\mathcal{R}$ is.[5]

NAP1. Regularity. $P(\emptyset) = 0$ and, for every non-empty set A in $\mathcal{P}(\Omega)$,

$$P(A) > 0.$$

NAP2. Normalization:

$$P(\Omega) = 1.$$

NAP3. Finite Additivity. If A and B are events and $A \cap B = \emptyset$, then

$$P(A \cup B) = P(A) + P(B).$$

The final axiom of NAP is a generalization of Continuity. To spell it out, define conditional probability via the ratio formula:

$$P(A|B) = \frac{P(A \cap B)}{P(B)},$$

where A and B are subsets of Ω and B is non-empty.[6]

NAP4. Non-Archimedean Continuity. For any event A and any non-empty finite subset λ of Ω,

$$P(A|\lambda) \in \mathbb{R}. \tag{1}$$

Additionally,

$$P(A) = \lim_{\lambda \uparrow \Omega} P(A|\lambda), \tag{2}$$

where $\lim_{\lambda \uparrow \Omega} P(A|\lambda)$ is a number in $\mathcal{R}$ that is partially determined by the values of $P(A|\lambda)$ for every non-empty finite subset λ of Ω. Benci *et al.* [2016] call this number the 'Ω-limit'. The Ω-limit is not unique. It depends on the specification of a free ultrafilter $\mathcal{U}$ on the set of all non-empty finite subsets of Ω. I omit the details here.

[5] In particular, if Ω is uncountable, then $\mathcal{R}$ contains infinitesimals.

[6] Note that, because $\mathcal{R}$ is a field, this ratio is well-defined for any such A and B.

Note that NAP0–NAP3 are straightforwardly analogues of K0–K3. Further, Benci *et al.* [2016, Section 3.5] show that NAP4 entails an infinitary additivity principle that is a generalization of finite and countable additivity.

In general, it is not possible to define the ultrafilter $\mathcal{U}$ explicitly. Indeed, to show that there are models of the above axioms, Benci *et al.* appeal to the axiom of choice. So, there is an essentially non-constructive aspect to NAP functions. I will return to this point later.

2.3 NAP and the four desiderata

It readily follows that NAP (more precisely, NAP functions) satisfy the first three desiderata above. **Regularity** is satisfied via NAP1, **Totality** is satisfied via NAP0, and **Perfect Additivity** is satisfied via NAP4.

Benci *et al.* also argue that NAP satisfies **Weak Laplacianism**—that is, that any 'conceptually possible' probabilistic scenario can be represented by a NAP function. In what follows, I will understand a probabilistic scenario to be specified by three items:

1. A sample space Ω.
2. A collection S of subsets of Ω.
3. A collection C of probabilistic constraints that the elements of S satisfy.

A probabilistic scenario $\langle \Omega, S, C \rangle$ is conceptually possible, then, just in case it is conceptually possible to have a set Ω of outcomes such that the elements of S satisfy the constraints in C.

For example, Benci *et al.* take a fair lottery on $\mathbb{N}$ in which the set $\mathbb{E}$ of even integers is exactly as probable as the set $\mathbb{O}$ of odd integers to be a conceptually possible probabilistic scenario. This scenario is specified as follows:

1. $\Omega = \mathbb{N}$.
2. $S = \{\{1\}, \{2\}, \{3\}, \ldots, \mathbb{E}, \mathbb{O}\}$.
3. Probabilistic constraints.
   (a) For every $i, j \in \mathbb{N}$: $\{i\} \approx \{j\}$.
   (b) $\mathbb{E} \approx \mathbb{O}$,

where '$A \approx B$' means that A is exactly as probable as B. (This notation will prove useful later.)

Although Benci *et al.* do not explicitly argue that NAP can represent *every* conceptually possible probabilistic scenario, they do show that NAP can represent many more such scenarios than classical probability theory can. (See Benci *et al.* [2013, Section 5] and Benci *et al.* [2016, Section 3.5].) I discuss these scenarios in Section 4.3.3.

3 Reconsidering the Desiderata

As stated, the desiderata of Section 2 presuppose that any acceptable theory of probability must be one of *numerical* probability. For example, **Perfect Additivity** talks about 'summing' probabilities, which only makes sense if probability is understood as numerical. As such, these desiderata immediately disqualify any theory of *qualitative* probability from being acceptable. This seems an intuitively undesirable consequence, given the large and influential literature on qualitative probability.[7] It seems more plausible that our desiderata for a theory of probability should be formulated in a manner that is neutral between qualitative and numerical probability.

Hence, I do two things in this section. In Section 3.1, I reformulate these desiderata in a manner that is neutral between qualitative and numerical probability. In Section 3.2, I critically assess these desiderata.

3.1 Reformulating the desiderata

To reformulate the desiderata of Section 2 in a manner that is neutral between qualitative and numerical probability, I will generalize **Regularity** and **Totality** and provide a weakened—but no less plausible—version of **Perfect Additivity**.

First, consider **Regularity**. The following desideratum generalizes it:

[7] Axiomatic theories of qualitative probability have been developed by de Finetti [1937], Koopman [1940a], Scott [1964], Suppes and Zanotti [1982], and others. Although it may no longer be so widespread, the view that qualitative probability is somehow more fundamental than numerical probability was held by a number of notable authors in the history of probability, including Keynes [1921], de Finetti [1937], Koopman [1940a], and Savage [1954]. See Stefánsson [2017] for a contemporary defense of the view.

1. **Generalized Regularity**. Any possible event is strictly more probable than any impossible event.

Because **Generalized Regularity** is a comparative statement, it is readily a desideratum for any theory of numerical probability as well as any theory of qualitative probability.

Second, consider **Totality**. The following desideratum generalizes it:

2. **Complete Comparability**. For any events A and B, A is at least as probable as B, or B is at least as probable as A.

Let A and B be arbitrary events. Note that **Complete Comparability**, in conjunction with the claim that probability is numerical, implies that the numerical probability of A is at least as great as that of B or the numerical probability of B is at least as great as that of A. Moreover, this claim implies that both A and B have numerical probabilities, which implies that every event has a probability. Thus, **Complete Comparability** generalizes **Totality**. Because **Complete Comparability** is a comparative statement, it is readily a desideratum for any theory of qualitative probability as well.

Third, consider **Perfect Additivity**. Here is a weakening of it:

3. **Weakened Perfect Additivity**. A bigger disjunction of mutually disjoint events is at least as probable as a smaller disjunction of mutually disjoint events, provided that each disjunct of the former is at least as probable as each disjunct of the latter.

As before, **Weakened Perfect Additivity** is readily a constraint on any theory of qualitative probability. Additionally, it seems no less plausible a desideratum for a theory of probability than **Perfect Additivity** is for a theory of numerical probability. I now show that any theory of numerical probability that satisfies **Perfect Additivity** also satisfies **Weakened Perfect Additivity**.

Let A and B be unions of mutually disjoint events A_i and B_j, indexed by $i \in I_A$ and $j \in I_B$, respectively. Then, **Perfect Additivity** entails:

$$\begin{aligned} P(A) &= \sum_{i \in I_A} P(A_i) \\ P(B) &= \sum_{j \in I_B} P(B_j), \end{aligned}$$

where the sums are defined in an appropriate way when the index sets I_A, I_B are infinite. Now, on any reasonable definition of 'sum', $P(A) \geq P(B)$ if the following conditions hold:

(1) $|I_A| > |I_B|$, and

(2) For every A_i, B_j, $P(A_i) \geq P(B_j)$.

Thus, if we think of unions as (possibly infinite) disjunctions, then **Perfect Additivity** implies **Weakened Perfect Additivity** when probability is understood as numerical. Hence, any theory of numerical probability that satisfies **Perfect Additivity** also satisfies **Weakened Perfect Additivity**.

Finally, although **Weak Laplacianism** talks about 'probability ratios' between atomic events, I will later show—in Section 4.2—that we can understand probability ratios in terms of qualitative probability. Because the rest of **Weak Laplacianism** is already neutral between numerical and qualitative probability—in particular, a uniform probability distribution on a sample space is simply one in which every atomic event is exactly as probable as every other atomic event—I will not modify this desideratum in what follows.

3.2 Questioning the desiderata

Benci *et al.* [2016] claim that **Regularity**, **Totality**, **Perfect Additivity**, and **Weak Laplacianism** are desiderata for a theory of probability. However, it is worth asking whether these items—or their more neutral counterparts—really are such desiderata. Arguably the most questionable items are **Totality** and **Complete Comparability**.

Although **Totality** is presupposed in much of the probability literature, it is difficult to find explicit arguments for **Totality**. That said, Benci *et al.* do provide one argument in favour of **Totality**—namely, that it allows one's probability theory to avoid non-measurable sets, which Benci *et al.* state are 'widely regarded as "pathologies" by probability theorists' (*ibid.*, p. 27). This argument is a bit quick, however. While it is true that non-measurable sets are *mathematically* pathological in the sense that they complicate technical matters in classical probability theory, it is not obvious that such sets should be banned as a precondition to theorizing.

More generally, any theory of numerical probability that satisfies **Totality** rules out the possibility of what Hájek [2003b] calls 'probability gaps'—

that is, events that are assigned no probability value whatsoever. Nonmeasurable sets are a notable alleged example of probability gaps, but there are others. Hájek [2003b, pp. 278–280] cites additional such examples from decision theory, statistics, and theories of subjective probability. Here is an adaptation of a more pedestrian example from Fishburn [1986, p. 339]: is the event (A) that the population of Mexico will exceed 175 million by 2030 at least as probable as the event (B) that the first card drawn from my old and probably incomplete bridge deck will be a heart? As A and B concern rather disparate subject matters, it is not implausible that no probability comparison can be made between them. So, it is not implausible that **Totality**—and, by extension, **Complete Comparability**—fail in this case.[8]

Complete Comparability also seems to be in tension with **Weak Laplacianism**. Even if there are no actual probability gaps—in decision theory, statistics, or elsewhere—it certainly seems that we can coherently *imagine* probabilistic scenarios that involve probability gaps. By **Weak Laplacianism**, any theory of probability should be able to mathematically represent such scenarios. Nonetheless, no theory that satisfies **Totality** can represent them. In particular, then, NAP fails to satisfy **Weak Laplacianism**.

For these reasons, **Complete Comparability** does not appear to be a genuine desideratum for a theory of probability. So, in what follows, I will only aim to describe a theory of qualitative probability that satisfies **Generalized Regularity**, **Weakened Perfect Additivity**, and **Weak Laplacianism**. Henceforth, I will refer to these items as the 'weakened desiderata'.

4 Qualitative Probability

Qualitative probability is the binary relation of one event's being *at least as probable* as another event.[9] Historically, the main interest in studying qualitative probability has been to prove various 'representation theorems' connecting qualitative probability to numerical probability.[10] Nonetheless, qualitative probability may be studied as a subject in its own right, and

[8] See Keynes [1921, Chapter 3] for additional such cases.

[9] Unless otherwise specified, I leave the interpretation of 'at least as probable' open here and in what follows.

[10] See Fishburn [1986] for an overview.

that is precisely what I will do in this paper. As I will show, qualitative probability provides a simple and powerful tool for representing probabilistic scenarios that involve infinite sample spaces.

To fully appreciate the representational power of qualitative probability, it is fruitful to study qualitative *conditional* probability. Qualitative conditional probability is the *quaternary* relation of its being the case that event A, given event B, is at least as probable as event C, given event D. When appealing to qualitative conditional probability, we need not employ any special technical innovations (as we must do in NAP) to make sense of such comparisons when B and D have classical probability 0.

The first axiomatization of qualitative conditional probability was provided by Koopman [1940a]. In this section, I provide a theory of qualitative conditional probability—which I call Qualitative Probability Theory (QP)—that is an extension of Koopman's axiomatization. Although a number of alternative axiomatizations have been proposed since Koopman's, nearly all of Koopman's axioms have figured as axioms or theorems in subsequent axiomatizations.[11] As such, Koopman's axioms constitute a common core of intuitively plausible constraints on qualitative conditional probability.

The plan for this section is as follows. In Section 4.1, I lay out the axioms of QP. In Section 4.2, I show how we can understand probability ratios in QP. In Section 4.3, I argue that QP satisfies the weakened desiderata of Section 3.1 at least as well as NAP. In Section 4.4, I argue that QP satisfies these desiderata more simply and perspicuously than NAP.

4.1 The axioms of Qualitative Probability Theory (QP)

Notation:

- $A|B \succeq C|D$: A, given B, is at least as probable as C, given D.
- $A|B \approx C|D$: both $A|B \succeq C|D$ and $C|D \succeq A|B$. That is: A, given B, is exactly as probable as C, given D.
- $A|B \succ C|D$: $A|B \succeq C|D$ and it is not the case that $C|D \succeq A|B$. That is: A, given B, is strictly more probable than C, given D.
- $A \succeq B$: $A|\Omega \succeq B|\Omega$. Similarly for $A \succ B$ and $A \approx B$.

[11] See Krantz *et al.* [1971, pp. 221–222] for discussion.

The central notion of QP is that of a QP-space, $\langle \Omega, \succeq \rangle$. As in classical probability theory and NAP, Ω is the collection of possible outcomes. Let $\mathcal{P}_0(\Omega)$ be the set of non-empty elements of $\mathcal{P}(\Omega)$, i.e., the power set of Ω. Then, a QP-relation $\succeq$ is a (possibly partial) binary relation on $\mathcal{P}(\Omega) \times \mathcal{P}_0(\Omega)$ that satisfies the following axioms. For simplicity, I leave universal quantification over events implicit in the axioms that follow.

The first nine axioms are due to Koopman:

QP1. **Verified Contingency.** $k|k \succeq a|h$.

QP2. **Implication.** If $a|h \succeq k|k$, then $h \subseteq a$.

QP3. **Reflexivity.** $a|h \succeq a|h$.

QP4. **Transitivity.** If $c|l \succeq b|k$ and $b|k \succeq a|h$, then $c|l \succeq a|h$.

QP5. **Antisymmetry.** If $b|k \succeq a|h$, then $\neg a|h \succeq \neg b|k$.

QP6. **Composition.**
Suppose:

(a) $\emptyset \neq a_1 \subseteq b_1 \subseteq c_1$ and $\emptyset \neq a_2 \subseteq b_2 \subseteq c_2$.

(b) $a_2|b_2 \succeq a_1|b_1$.

(c) $b_2|c_2 \succeq b_1|c_1$.

Then: $a_2|c_2 \succeq a_1|c_1$.[12]

QP7. **Decomposition.**
Suppose:

(a) $\emptyset \neq a_1 \subseteq b_1 \subseteq c_1$ and $\emptyset \neq a_2 \subseteq b_2 \subseteq c_2$.

(b) $a_1|b_1 \succeq a_2|b_2$.

(c) $a_2|c_2 \succeq a_1|c_1$.

Then: $b_2|c_2 \succeq b_1|c_1$.[13]

QP8. **Alternative Presumption.** If $r|s \succeq a|(b \cap h)$ and $r|s \succeq a|(\neg b \cap h)$, then $r|s \succeq a|h$.

QP9. **Subdivision.**
Suppose $a_1, \ldots, a_n, b_1, \ldots, b_n$ are such that:

[12] Koopman actually states two axioms of composition; the other is analogous. I follow Koopman [1940b]'s formulations here, and in QP7, which are simpler than the logically equivalent formulations of Koopman [1940a].

[13] Koopman actually states four axioms of decomposition; the others are analogous.

(a) $(a_i \cap a_j) = (b_i \cap b_j) = \emptyset$ if $i \neq j$.
(b) $a = (a_1 \cup \ldots \cup a_n) \neq \emptyset$ and $b = (b_1 \cup \ldots \cup b_n) \neq \emptyset$.
(c) $a_n|a \succeq \ldots \succeq a_1|a$.
(d) $b_n|b \succeq \ldots \succeq b_1|b$.

Then: $b_n|b \succeq a_1|a$.

Here is the final axiom of QP:

QP10. **Qualitative Perfect Additivity.**
Let $a = \bigcup_{i \in I_a} a_i$ and $b = \bigcup_{j \in I_b} b_j$, for index sets I_a, I_b. Suppose:

(a) $|I_a| > |I_b|$.
(b) $(a_i \cap a_j) = (b_i \cap b_j) = \emptyset$ if $i \neq j$.
(c) For every $i \in I_a, j \in I_b$: $a_i|c \succeq b_j|c$.

Then: $a|c \succeq b|c$.

Remark. If one insists upon **Complete Comparability**, one may treat $\succeq$ as a total binary relation on $\mathcal{P}(\Omega) \times \mathcal{P}_0(\Omega)$. That is, one may supplement the above axioms with the following one:

QP0. **Qualitative Totality.** $a|b \succeq c|d$ or $c|d \succeq a|b$.

For the reasons I stated in Section 3.2, however, I won't assume that **Qualitative Totality** is an axiom of QP in what follows.

4.2 QP and probability ratios

It is natural to speak of one event as being 'twice' or 'one-third' as probable as another event. In this section, I show how we can make sense of probability ratios between events in terms of qualitative conditional probability. The basic idea is to understand probability ratios between arbitrary events as being relative to a set of 'reference' events such that the probability ratios between the latter events are intuitively immediate. I discuss three types of probability ratios: rational, real-valued, and infinitesimal.[14]

[14] See Stefánsson [forthcoming, Section 3.1] for an alternative (albeit less general) approach to understanding probability ratios in terms of qualitative unconditional probability.

4.2.1 Rational ratios

To make sense of rational probability ratios in QP, it is useful to employ Koopman [1940a]'s notion of an 'n-scale'.

Definition. Let an n**-scale** be a set of n events $\{s_1, \ldots, s_n\}$ such that:

1. $S_n = (s_1 \cup \ldots \cup s_n) \neq \emptyset$.
2. $s_i \cap s_j = \emptyset$ for $i, j = 1, \ldots, n$ such that $i \neq j$.
3. $s_i | S_n \approx s_j | S_n$ for $i, j = 1, \ldots, n$.[15]

An n-scale $\{s_1, \ldots, s_n\}$ is simply a finite fair lottery: given that some member of $\{s_1, \ldots, s_n\}$ obtains, s_1 is exactly as probable as s_2, which is exactly as probable as s_3, and so on.

As the following theorem proven by Koopman [1940a] shows, n-scales behave much like rational numbers.

N**-Scale Equivalence.** *Suppose:*

1. *For every positive integer n, there is some n-scale.*
2. *$\{u_1, \ldots, u_n\}$ is an n-scale and $\{v_1, \ldots, v_m\}$ is an m-scale.*
3. *k and l are integers such that $0 \leq k \leq n$ and $0 \leq l \leq m$.*
4. *$\frac{k}{n} \geq \frac{l}{m}$.*

Then:

$$(u_{i_1} \cup \ldots \cup u_{i_k}) | (u_1 \cup \ldots \cup u_n) \succeq (v_{j_1} \cup \ldots \cup v_{j_l}) | (v_1 \cup \ldots v_m). \tag{3}$$

If $\frac{k}{n} > \frac{l}{m}$, then replace '$\succeq$' with '$\succ$' above.[16]

To illustrate this theorem, suppose $\{u_1, u_2, u_3\}$ and $\{v_1, v_2, v_3, v_4, v_5\}$ are a 3-scale and 5-scale, respectively. Intuitively, u_1 is $\frac{1}{3}$ as probable as $(u_1 \cup u_2 \cup u_3)$, and $(v_1 \cup v_2)$ is $\frac{2}{5}$ as probable as $(v_1 \cup v_2 \cup v_3 \cup v_4 \cup v_5)$. Since $\frac{2}{5} > \frac{1}{3}$, we would expect that $(v_1 \cup v_2)$, given $(v_1 \cup v_2 \cup v_3 \cup v_4 \cup v_5)$, would be more

[15] Cf. Koopman [1940a], Definition 1.
[16] Cf. Koopman [1940a], Theorem 14.

probable than u_1, given $(u_1 \cup u_2 \cup u_3)$. N**-Scale Equivalence** vindicates this intuition, provided there is an n-scale for every positive integer n.

We can now see how to make sense of rational probability ratios in QP. Let a and b be events such that $a \subseteq b$, and suppose there is an n-scale V_n consisting of atomic events for every positive integer n.[17] Further, suppose that $a|b \approx U_m|V_n$, where U_m is some m-member subset of V_n. Intuitively, then, a is $\frac{m}{n}$ times as probable as b. Moreover, N**-Scale Equivalence** (along with **Transitivity**) ensures that, for any m'-member subset $U_{m'}$ of $V_{n'}$, $a|b \succeq U_{m'}|V_{n'}$ if $\frac{m}{n} \geq \frac{m'}{n'}$ and that $U_{m'}|V_{n'} \succeq a|b$ if $\frac{m'}{n'} \geq \frac{m}{n}$. Thus, intuitively, a is at least $\frac{m}{n}$ times as probable as b and no more than $\frac{m}{n}$ times as probable as b. That is, a is *exactly* $\frac{m}{n}$ times as probable as b. We can generalize beyond the case when $a \subseteq b$ as follows.

Let m and n be positive integers. Say that a is $\frac{m}{n}$ times as probable as b—equivalently, b is $\frac{n}{m}$ times as probable as a—just in case:

- For every positive integer k, there is a k-scale V_k consisting of atomic events.
- For any positive integers m', n' such that $\frac{m'}{n'} = \frac{m}{n}$, there is some positive integer k such that:

  (i) $a|(a \cup b) \approx U_{m'}|V_k$ for any m'-member subset $U_{m'}$ of V_k.

  (ii) $b|(a \cup b) \approx U_{n'}|V_k$ for any n'-member subset $U_{n'}$ of V_k.

Intuitively, a is $\frac{m'}{k}$ times as probable as $(a \cup b)$, and b is $\frac{n'}{k}$ times as probable as $(a \cup b)$. So, intuitively, a is $(\frac{m'}{k})/(\frac{n'}{k}) = \frac{m'}{n'} = \frac{m}{n}$ times as probable as b. We can also understand a being *at least* $\frac{m}{n}$ times as probable as b—equivalently, b as being *at most* $\frac{n}{m}$ times as probable a—by replacing '$\approx$' with '$\succeq$' in (ii) and with '$\preceq$' in (iii).

Although the nature of k in (i)–(iii) will generally depend on specific features of a given probabilistic scenario being represented by a QP-space, it is worth noting that k is simply given by $(m' + n')$ when a and b are disjoint. That this is the case is ensured by the following qualitative analogue of finite additivity proven by Koopman [1940a]:

Qualitative Finite Additivity. *Suppose:*

[17] The appeal to atomic events is not strictly necessary, but it will simplify the ensuing discussion.

1. $(a_1 \cap b_1 \cap h_1) = (a_2 \cap b_2 \cap h_2) = \emptyset$.
2. $a_1|h_1 \succeq a_2|h_2$.
3. $b_1|h_1 \succeq b_2|h_2$.

Then: $(a_1 \cup b_1)|h_1 \succeq (a_2 \cup b_2)|h_2$.[18]

Suppose a is $\frac{m}{n}$ times as probable as b and that $a \cap b = \emptyset$. Then, for any positive integer k, there is a k-scale $V_k = (v_1 \cup \ldots \cup v_k)$, for atomic events $v_1, \ldots, v_k$. Moreover, for any m', n' such that $\frac{m'}{n'} = \frac{m}{n}$, there is some positive integer k such that (i) $a|(a \cup b) \approx U_{m'}|V_k$ for any m'-member subset $U_{m'}$ of V_k and (ii) $b|(a \cup b) \approx U_{n'}|V_k$ for any n'-member subset $U_{n'}$ of V_k. Next, observe that $|U_{m'} \cup U_{n'}| = |U_{m'}| + |U_{n'}| - |U_{m'} \cap U_{n'}| = m' + n' - |U_{m'} \cap U_{n'}| \leq |V_k| = k$. So, $m' + n' \leq k$. Thus, we may let $U_{m'} = (v_1 \cup \ldots \cup v_{m'})$ and $U_{n'} = (v_{m'+1} \cup \ldots \cup v_{m'+n'})$, so that $U_{m'} \cap U_{n'} = \emptyset$. Then, by **Qualitative Finite Additivity**, $(U_{m'} \cup U_{n'})|V_k \succeq (a \cup b)|(a \cup b)$. Additionally, by **Verified Contingency**, $(a \cup b)|(a \cup b) \succeq V_k|V_k$. So, by **Transitivity**, $(U_{m'} \cup U_{n'})|V_k \succeq V_k|V_k$, i.e., $(v_1 \cup \ldots \cup v_{m'+n'})|V_k \succeq V_k|V_k$. Now suppose $m' + n' < k$. Then, by N**-Scale Equivalence**, $V_k|V_k \succ (v_1 \cup \ldots \cup v_{m'+n'})|V_k$, contrary to what was just found. So, it must be that $m' + n' = k$.

Let us take stock. The above analysis ensures that we can make sense of rational probability ratios in QP by making reference to n-scales—that is, finite fair lotteries. Note that such lotteries needn't be *actual* lotteries; they may merely be hypothetical. So, statements about rational probability ratios between events a and b can simply be understood as shorthand for statements involving qualitative probability comparisons among a, b, $(a \cup b)$, and various events—call them 'reference' events—involving hypothetical fair lotteries. For example, the statement that a is $\frac{2}{3}$ as probable as $(a \cup b)$ can be understood as meaning that a, given $(a \cup b)$, is just as probable as the 'reference' event that one picks one of 2 given tickets in a hypothetical fair lottery of 3 tickets, given that one picks some ticket in that lottery.[19] Since N**-Scale Equivalence** ensures that all n-scales behave like rational numbers, the nature of the n-scales with which we make sense of such statements is immaterial.

[18] Cf. Koopman [1940a], Theorem 5.

[19] More generally, the statement means that a, given $(a \cup b)$, is just as probable as the event that one picks one of m tickets in hypothetical fair lottery of n tickets, given that one picks some ticket in that lottery, for any m and n such that $\frac{m}{n} = \frac{2}{3}$.

Note that, because statements about rational probability ratios involve reference to n-scales, we can only make sense of such statements between events of a given QP-space $\langle \Omega, \succeq \rangle$ if that QP-space has a particular sort of structure. In particular, Ω must include some set Ω_{ref} of 'reference' outcomes out of which the relevant n-scales are built, and $\succeq$ must satisfy various 'reference' constraints constitutive of rational probability ratios holding among events in $\mathcal{P}(\Omega)$. More precisely, to make sense of the statement that event a is $\frac{m}{n}$ times as probable as event b, the QP-space $\langle \Omega, \succeq \rangle$ must be equipped with the following structure:

1. 'Reference' outcomes:
   a. For some countably infinite set Ω_{ref}, $\Omega_{ref} \subseteq \Omega$.
2. 'Reference' constraints:
   a. For every positive integer k, there is a k-member subset V_k of Ω_{ref} such that $\{v\}|V_k \approx \{v'\}|V_k$, for every $v, v' \in V_k$.
   b. For any positive integers m', n' such that $\frac{m'}{n'} = \frac{m}{n}$, there is some positive integer k such that:
      (i) $a|(a \cup b) \approx U_{m'}|V_k$ for any m'-member subset $U_{m'}$ of V_k.
      (ii) $b|(a \cup b) \approx U_{n'}|V_k$ for any n'-member subset $U_{n'}$ of V_k.

Remark. Although there is, for every positive integer n, an n-member subset of Ω_{ref} that is a fair sub-lottery of Ω, Ω_{ref} need not itself be a fair sub-lottery of Ω. The assumption that it is leads to a qualitative version of de Finetti [1972]'s paradox of non-conglomerability. See DiBella [2018] for discussion.

All of that said, I have not yet shown that there are QP-spaces with the aforementioned sort of structure because I have not yet shown that there *are* QP-spaces, i.e., models of the axioms of QP. My discussion so far only establishes that, if the axioms of QP are consistent with the above sorts of constraints, then we can make sense of rational probability ratios between events of a given QP-space in the above manner. I demonstrate that the axioms of QP are indeed consistent with the above sorts of constraints—as well as the constraints I will soon describe for making sense of real-valued and infinitesimal probability ratios—in the Appendix.

4.2.2 Real-valued ratios

We can understand real-valued probability ratios more generally in terms of rational probability ratios. The method I will describe is analogous to the method of constructing the reals from the rationals via Dedekind cuts.

Let r be an arbitrary positive real number. Say that a is at least r times as probable as b just in case:

(i) For every positive rational number $r' \leq r$, a is at least r' times as probable as b.

Similarly, say that a is at most r times as probable as b just in case:

(ii) For every positive rational number $r'' \geq r$, a is at most r'' times as probable as b.

Then, we can say that a is exactly r times as probable as b just in case (i)–(ii) hold. From the analysis of the previous section, (i)–(ii) entail that there is a k-scale V_k of atomic events for every positive integer k. Additionally, (i) entails the following:

(i′) For any positive integers m, n such that $r \geq \frac{m}{n}$, there is some positive integer k such that:

  (a) $a|(a \cup b) \succeq U_m|V_k$ for any m-member subset U_m of V_k.

  (b) $b|(a \cup b) \preceq U_n|V_k$ for any n-member subset U_n of V_k.

Similarly, (ii) entails the following:

(ii′) For any positive integers m, n such that $\frac{m}{n} \geq r$, there is some positive integer k such that:

  (c) $a|(a \cup b) \preceq U_m|V_k$ for any m-member subset U_m of V_k.

  (d) $b|(a \cup b) \succeq U_n|V_k$ for any n-member subset U_n of V_k.

As before, when a and b are disjoint, k is simply given by $(m + n)$. Additionally, note that the above analysis readily reduces to the analysis of the previous section when r is rational.

4.2.3 Infinitesimal ratios

We can understand infinitesimal probability ratios by appealing to real-valued probability ratios. In particular, we may say that a is infinitesimally probable relative to b just in case, for every positive real number r, b is at least r times as probable as a.

More generally, for any non-negative real number r, we can say that a is at most *infinitesimally more or less than* r times as probable as b just in case:

(a) For every positive real number $r' < r$, a is at least r' times as probable as b.

(b) For every positive real number $r'' > r$, a is at most r'' times as probable as b.

Note that the above constraints are identical to (i)–(ii) from the previous section except that they now involve strict inequalities. The case $r = 0$ corresponds to a's being infinitesimally probable relative to b.

In what follows, I will employ the following notation:

- $a \approx_r b$: a is r times as probable as b.
- $a \sim_r b$: a is at most infinitesimally more or less than r times as probable as b.

Note that, if $a \approx_r b$, then $a \sim_r b$, but the converse is not the case.

4.3 QP and the weakened desiderata

I now argue that QP satisfies the weakened desiderata of Section 3.1 at least as well as NAP.

4.3.1 Generalized Regularity

In the context of QP, **Generalized Regularity** amounts to the following claim:

Qualitative Regularity. For any non-empty event A, $A \succ \emptyset$.

It is easy to show any QP-relation satisfies **Qualitative Regularity**, as it is a straightforward consequence of Theorem 3 proven in Koopman [1940a]. According to this theorem, if $(a \cap h) \subset (b \cap h)$, then $b|h \succ a|h$. **Qualitative Regularity** follows when $a = \emptyset, b = A$, and $h = \Omega$.

Any QP-relation also satisfies a qualitative version what Benci *et al.* call the 'Euclidean' principle. According to the Euclidean principle—of which **Regularity** is a special case—any set has a strictly larger probability than any of its proper subsets. As Benci *et al.* show, any NAP function satisfies this principle. In the context of QP, this principle corresponds to the following claim:

> **Qualitative Euclidean Principle**. For any events A and B such that $A \subset B$, $B \succ A$.

This principle is also a straightforward consequence of Koopman's Theorem 3.

4.3.2 Weakened Perfect Additivity

In the context of QP, **Weakened Perfect Additivity** merely amounts to the claim that QP10, **Qualitative Perfect Additivity**, is true when $C = \Omega$. Hence, any QP-relation satisfies **Weakened Perfect Additivity**.

4.3.3 Weak Laplacianism

As I said in Section 2.3, Benci *et al.* do not show that *every* conceptually possible probabilistic scenario can be represented by a NAP function. Rather, they only show that NAP functions can represent five notable types of infinitary probabilistic scenarios that classical probability functions cannot represent:

1. A uniform distribution on a sample space Ω of any cardinality—finite, countable, or uncountable—as well as distributions on Ω involving many other real-valued probability ratios between the atomic events.

2. A fair lottery on the set $\mathbb{N}$ of natural numbers in which the probability of each subset A of $\mathbb{N}$ is the asymptotic density of A.[20]

[20] More precisely, a fair lottery on $\mathbb{N}$ in which the probability $P(A)$ differs no more than an infinitesimal from the asymptotic density $d(A)$ of A (if it exists). The reason for this

3. A fair lottery on the set $\mathbb{Q}$ of rational numbers in which the probability of an interval is proportional to the length of that interval.

4. A fair lottery on the set $\mathbb{R}$ of real numbers in which the probability of an interval is proportional to the length of that interval.

5. A countably infinite sequence of fair coin tosses.

In the Appendix, I use the machinery of Section 4.2 to show that each of the above scenarios can be represented by a QP-space (in the sense I make precise there). I also show that total versions of the above scenarios—that is, versions of these scenarios that include **Complete Comparability** as a constraint—can be represented by QP-spaces. Although further work is needed, these results suggest that QP satisfies **Weak Laplacianism** at least as well as NAP.

As the discussion in the Appendix shows, these results also establish that QP is consistent and admit of models involving sample spaces of any cardinality. To my knowledge, this is (surprisingly) the first demonstration of consistency for a theory of qualitative conditional probability.

4.4 The simplicity and perspicuity of QP

I have argued that QP satisfies the weakened desiderata of Section 3.1 at least as well as NAP. I now argue that QP satisfies these desiderata more simply and perspicuously than NAP. More precisely, I argue: (1) we can construct QP-spaces that satisfy these desiderata more simply than we can construct NAP functions that satisfy them, and (2) QP-spaces provide more perspicuous representations of probabilistic scenarios than NAP functions.

First, in order to construct NAP functions that satisfy these desiderata, we must appeal to such sophisticated mathematical machinery as the axiom of choice, free ultrafilters, and non-standard limit processes. (Although the axiom of choice entails the existence of the relevant ultrafilters and non-standard limits, one still needs to employ the latter machinery to construct NAP functions.) By contrast, to construct QP-spaces that satisfy the desiderata, the only substantive mathematical machinery we need to appeal to is

qualification is that there are subsets A, B of $\mathbb{N}$ such that $A \subset B$ yet $d(A) = d(B)$. So, in order to satisfy the Euclidean principle, it must be that $P(B) > P(A)$. See Benci *et al.* [2013, p. 142] for further discussion. A similar qualification holds for the fourth scenario.

that of partially ordered sets (of which QP-spaces are a special class). Moreover, as I show in the Appendix for a variety of (non-total) probabilistic scenarios, we can construct a QP-space that represents a given such scenario merely by appealing to the axioms of QP and the constraints of the scenario in question.[21] So, QP satisfies the desiderata more simply than NAP.

Next, a typical NAP function that can represent a given probabilistic scenario has a good deal of artifactual structure. For example, Benci *et al.* show that some NAP functions that can represent a fair lottery on $\mathbb{N}$ assign the set $\mathbb{O}$ of odd numbers greater probability than the set $\mathbb{E}$ of even numbers, while others assign these sets the same probability. As Benci *et al.* note, this difference is due to the arbitrary choice of free ultrafilter used to construct these functions. So, it is merely an artifact of some NAP functions that $\mathbb{O}$ is more probable than $\mathbb{E}$. Moreover, the fact that this probability difference is a representational artifact only becomes clear when we consider the whole family of NAP functions that can represent this scenario and then determine which probability orderings vary among the family and which do not. In general, one cannot "read off" the probabilistic facts of a probabilistic scenario straight from any particular NAP function that can represent it.

By contrast, consider a probabilistic scenario $\langle \Omega, S, C \rangle$ that can be represented by a QP-space $\langle \Omega', \succeq \rangle$ in the manner I describe in the Appendix. As I show there, $\langle \Omega', \succeq \rangle$ can be constructed such that $a|b \succeq c|d$ if and only if $a|b \succeq c|d$ is a logical consequence of C_{QP} and the axioms of QP, where C_{QP} is a characterization of C in terms of qualitative conditional probability. Intuitively, $\succeq$ faithfully represents the probabilistic constraints in C: $\succeq$ has all of the structure of these constraints, just enough additional structure to make it a QP-relation, and no further structure. Although the specific 'reference' outcomes that the QP-space employs to make sense of any probability ratios in C is arbitrary, any set of reference outcomes that satisfies the constraints of Section 4.2 suffices to define a QP-relation with the aforementioned features. In this respect, the choice of reference items that determines a QP-space is importantly dissimilar to the choice of ultrafilter that determines a NAP function. To wit: a QP-space's reference outcomes do not lead to artifactual orderings in its QP-relation, but a NAP function's ultrafilter can lead to artifactual orderings in its probability assignments. So, it is far

[21] As I show in the Appendix, we may employ the axiom of choice to construct a QP-space that represents a scenario in which **Complete Comparability** is a constraint. However, in such a case, we still need not appeal to free ultrafilters or non-standard limit processes.

easier to discern what is artifactual and what is not for a QP-space than it is for a NAP function. Unlike the case with NAP functions, one can "read off" the probabilistic facts of a probabilistic scenario straight from a single QP-space that can represent it. Thus, QP-spaces provide more perspicuous representations of probabilistic scenarios than NAP functions.[22]

5 Additional Dividends

Benci *et al.* [2016, Section 5] argue that NAP, in addition to satisfying the desiderata of Section 2, has further philosophical advantages over classical probability theory. In this section, I argue that QP has three notable such advantages as well. These advantages involve: (1) decision theory, (2) learning from evidence, and (3) chance and credence.[23]

5.1 Decision theory

Benci *et al.* argue that a decision theory that employs NAP functions can lead to subtler—and more rational—choices than a decision theory that employs only classical probability functions (*ibid.*, pp. 27, 36–37).

For example, suppose you are deciding whether to bet on the occurrence of $\{1\}$ or $\{1, 2\}$ in a fair lottery on $\mathbb{N}$. (Suppose further that the utility of winning is the same for each bet and that the utility of losing is the same for each bet.) Because any classical probability function assigns both events probability 0, a decision theory that employs only classical probability functions will have the implausible consequence that you should be indifferent between the two bets. By contrast, because NAP functions satisfy the Euclidean principle (cf. Section 4.3.1), any NAP function that can represent

[22]The issue of perspicuity in representation is closely related to the issue of non-uniqueness in representation that is discussed by Benci *et al.* [2016, Section 6.1]. While it is plausible that (as Benci *et al.* argue) non-uniqueness is not a problem in itself, non-uniqueness is pragmatically problematic insofar as it leads to misleading representational artifacts of the sort I discussed in the previous paragraph.

[23]Because of space limitations, I omit the other philosophical advantages discussed by Benci *et al.* Like the advantages I discuss below, these other advantages arise from the facts that NAP functions are regular and that one can conditionalize on events of classical probability 0 in NAP. Since QP-relations satisfy **Qualitative Regularity** and readily make sense of comparisons of the form $A|B \succeq C|D$ when B and D have classical probability 0, QP plausibly has these other advantages as well.

a fair lottery on $\mathbb{N}$ assigns greater probability to $\{1, 2\}$ than to $\{1\}$. As a result, a decision theory that employs NAP functions will have the plausible consequence that you should prefer betting on the larger set to betting on the smaller set.

Because QP-relations satisfy a qualitative version of the Euclidean principle, a decision theory based on QP will plausibly also imply that you should prefer betting on the larger set to betting on the smaller set. So, it is plausible that a decision theory based on QP will lead to subtler—and more rational—choices in cases like the above example as well.

Benci *et al.* also argue that a decision theory based on NAP can lead to subtler decisions by virtue of the plurality of NAP functions that can represent a given probabilistic scenario. For example, suppose you are now deciding whether to bet on the occurrence of an even number or the occurrence of an odd number in a fair lottery on $\mathbb{N}$. (As before, suppose also that the utility of winning is the same for each bet and that the utility of losing is the same for each bet.) As I said in Section 4.4, some NAP functions that can represent this scenario assign the set $\mathbb{O}$ of odd numbers greater probability than the set $\mathbb{E}$ of even numbers, while others assign these sets the same probability. Thus, Benci *et al.* conclude that you may reasonably favour betting on $\mathbb{O}$ over $\mathbb{E}$. By contrast, if you apply a decision theory that employs only classical probability functions, then you should be indifferent between these two bets.

Although the recommendation to prefer betting on $\mathbb{O}$ to betting on $\mathbb{E}$ is indeed subtler than the recommendation a decision theory based on classical probability would make here, the former recommendation appears to be simply wrong upon reflection. As I said in Section 4.4, it is merely a representational artifact of some NAP functions that $\mathbb{O}$ is more probable than $\mathbb{E}$. Intuitively, then, the spread of NAP values assigned to these sets should have no relevance to decision-making. Yet, by claiming that this spread can be relevant to decision-making, Benci *et al.* imply that merely representational artifacts can be relevant to decision-making. This consequence seems implausible.

As I said in Section 4.4, a wide variety of conceptually possible probabilistic scenarios can be faithfully represented by QP-spaces. So, there appear to be prospects for developing a QP-based decision theory in a manner that is faithful to a wide variety of decision problems. When applied to decision problems involving infinite sample spaces, such a decision theory would be subtler than a decision theory based on classical probability theory but not as implausibly subtle as the NAP-based decision theory suggested by Benci *et al.*

Although I do not have a detailed decision theory based on QP to offer, it is not implausible that such a theory could be developed. Such a theory would presumably have a different character than a decision theory based on NAP (or classical probability theory). Because NAP is a theory of numerical probability, plausibly it can be incorporated in a familiar sort of framework that takes an act to be rational just in case it maximizes expected utility.[24] By contrast, because QP is a theory of qualitative probability, it cannot be readily incorporated in such a framework. That said, Fine [1973, Chapter 2, Section G] develops a decision theory based on axioms of qualitative (unconditional) probability. Moreover, on Fine's theory, dominance reasoning—not expected-utility maximization—is essential in determining the rationality of an act. Easwaran [2014b] develops a broadly similar decision theory in which qualitative probability plays a central role as well. Crucially, neither of these authors takes numerical probabilities and utilities to be integrated from the start (though they can be integrated in various circumstances). While neither of these authors appeals to QP specifically, their work suggests that one could develop a QP-based decision theory along similar lines. I leave such a project open for future work.

5.2 Learning from evidence

It is widely held that, when one learns new evidence, one should update one's degrees of belief by conditionalizing on the evidence. That is, if one's initial subjective probability function is P_i, then one's new subjective probability function P should be given by $P(\cdot) = P_i(\cdot|E)$ upon learning E (and nothing stronger). Because classical probability theory (with the ratio formula for conditional probability) is silent on the value of $P_i(\cdot|E)$ when $P_i(E) = 0$, it is also silent on how one should update one's degrees of belief upon learning evidence to which one has antecedently assigned classical probability 0. Intuitively, this consequence seems problematic since it seems that one can learn such evidence. (See McGee [1994].)

NAP solves this problem by assigning all possible events positive probability, so $P(A|B)$ is defined by the ratio formula whenever B is possible.

[24] Although Benci *et al.* do not offer a detailed decision theory based on NAP, they note that non-Archimedean expected-utility theories have been developed by Pivato [2014] and Pedersen [unpublished].

Hence, one may consistently adopt NAP and accept conditionalization as the appropriate belief update rule in all contexts.

To my knowledge, the question of how to update one's comparative conditional confidence relation—that is, one's *subjective* qualitative conditional probability relation—upon learning new evidence has not been discussed in the literature. Nonetheless, a straightforward qualitative version of conditionalization suggests itself:

> **Qualitative Conditionalization**
> Let the initial comparative conditional confidence relation of agent S be $\succeq_i$. Then, when S learns evidence E (and nothing stronger), S's new comparative conditional confidence relation $\succeq$ should be given as follows:
>
> $$A|B \succeq C|D \text{ if and only if } A|(B \cap E) \succeq_i C|(D \cap E).$$

Because QP readily makes sense of comparisons of the form $A|B \succeq C|D$ when B and D have classical probability 0, it follows that QP—supplemented with **Qualitative Conditionalization**—can make sense of how one should update one's comparative conditional confidence relation upon learning evidence that has classical probability 0.

5.3 Credence and chance

Lewis [1980]'s Principal Principle can be roughly stated as follows:

$$\text{Prob}(A|Ch(A) = x) = x,$$

where 'Prob' is an agent's subjective probability function, Ch is a chance measure, and x is a real number between 0 and 1. Benci *et al.* point to the following problem with the principle when 'Prob' is taken to satisfy classical probability theory:

> In classical probability theory, it would seem that if A represents the value of a continuous observable (say, a position measurement for an electron in a superposition state), $Ch(A)$ will be zero in a non-determinist context for every value A. Hence, according to Lewis's principal principle, $\text{Prob}(A|Ch(A) = 0) = 0$. This will render any probability conditional on the posterior $\text{Prob}(A)$ undefined. (p. 28)

This is a problem because, intuitively, it makes sense to conditionalize on events that have classical probability 0. For example, intuitively, it seems that $\mathrm{Prob}(\{1\}|\{1,2,3\}) = \frac{1}{3}$ for a fair lottery on $\mathbb{N}$. Clearly, NAP has the resources to address this problem. As I said in the previous section, in NAP, we may conditionalize on any non-empty event via the ratio formula because every non-empty event is assigned a positive probability value. Indeed, in NAP, $\mathrm{Prob}(\{1\}|\{1,2,3\}) = \frac{1}{3}$ since $\mathrm{Prob}(\{1\}) = \mathrm{Prob}(\{2\}) = \mathrm{Prob}(\{3\}) > 0$.

Of course, QP also has the resources to make sense of comparisons of the form $A|B \succeq C|D$ when B and D have classical probability 0. Indeed, it is easy to show that, in QP, $\{1\}|\{1,2,3\} \approx \{2\}|\{1,2,3\} \approx \{3\}|\{1,2,3\}$ since $\{1\}|\Omega \approx \{2\}|\Omega \approx \{3\}|\Omega$.

It is worth noting that QP also has the resources to formulate a purely qualitative version of the Principal Principle. To do so, let $\succeq_{Ch}$ be the objective *qualitative* chance relation—that is, that relation of one event A's being at least as objectively probable as event B—and let $\succeq$ be the comparative conditional confidence relation of a rational agent S (cf. Section 5.2).[25] Then, a straightforward qualitative version of the Principal Principle suggests itself:

Qualitative Principal Principle
For any events A, B, and p:
$A|\Omega \succeq p|[A \succeq_{Ch} p \succeq_{Ch} B] \succeq B|\Omega$.

Although this principle is a bit unwieldy to state in plain English, it is simply a version of the original principle in which assignments of numerical probability have been replaced with relations of qualitative probability.

6 Representational Advantages of QP over NAP

In this section, I describe two 'conceptually possible' probabilistic scenarios that NAP cannot represent but which QP readily can. The existence of these scenarios suggests that QP has even greater representational power than NAP.

[25] For simplicity, I treat the objective qualitative chance relation here as a qualitative *un*conditional probability relation.

6.1 Relatively infinitesimally probable atomic events

Consider the following probabilistic scenario:

> *Toss a fair coin. If it lands Heads, pick 1. If it lands Tails, pick a number at random from* $\{2, 3, 4, \ldots\}$.

Formally, this is a scenario $\langle \Omega, S, C\rangle$ that is characterized as follows:

1. $\Omega = \{1, 2, 3, \ldots\}$.
2. $S = \{\{1\}, \{2\}, \{3\}, \ldots\} \cup \{\{2, 3, 4, \ldots\}\} \cup \{\Omega\}$.
3. Probabilistic constraints C.
   (a) $\{1\}|\Omega \approx \{2, 3, 4, \ldots\}|\Omega$.
   (b) $\{i\}|\{2, 3, 4, \ldots\} \approx \{j\}|\{2, 3, 4, \ldots\}$, for any $i, j \geq 2$.

Constraint (a) follows from the stipulation that the coin is fair: since Heads is exactly as probable as Tails, picking 1 is exactly as probable as picking some number greater than 1. Constraint (b) follows from the stipulation that you randomly pick an integer greater than 1 given that the coin lands Tails.

As it turns out, no NAP function can represent this scenario, even though it seems to be conceptually possible.

Proof. Suppose for reductio that some NAP function P could represent the above scenario. By Eq. (1) of the Non-Archimedean Continuity axiom, $P(\{2\}|\{1, 2\})$ is some real number. Moreover, the above constraints imply that $P(\{1\}) = \frac{1}{2}$ and that $P(\{i\}) = P(\{j\})$ for any $i, j \geq 2$.[26] Now suppose that $P(\{2\})$ is some real number x. By Regularity, $x > 0$. However, then Finite Additivity entails that there is some integer $n > 2$ such that $P(\{2\} \cup \{3\} \cup \ldots \cup \{n\}) = (n-1)x > 1$. This violates the Euclidean principle since $(\{2\} \cup \{3\} \cup \ldots \cup \{n\}) \subsetneq \Omega$, for any integer $n > 2$, yet $P(\Omega) = 1$. Hence, it must be that $P(\{2\})$ is infinitesimal.

Finally, by the ratio formula,

$$\begin{aligned} P(\{2\}|\{1, 2\}) &= \frac{P(\{2\} \cap \{1, 2\})}{P(\{1, 2\})} \\ &= \frac{P(\{2\})}{P(\{1\}) + P(\{2\})}, \end{aligned}$$

[26] Constraint (a), together with Normalization and Finite Additivity, implies that $P(\{1\}) = P(\{2, 3, 4, \ldots\}) = \frac{1}{2}$. Additionally, by the ratio formula for conditional probability, constraint (b) implies that $P(\{i\}) = P(\{j\})$ for any $i, j \geq 2$.

which must be infinitesimal since $P(\{1\}) = \frac{1}{2}$ and $P(\{2\})$ is infinitesimal. But we saw above that $P(\{2\}|\{1,2\})$ is a real number. Contradiction. $\square$

By contrast, as I show in the Appendix, there exists a QP-space that can represent the above scenario. Although the Non-Archimedean Continuity axiom of NAP (ironically) rules out scenarios in which one atomic event is infinitesimally probable relative to another atomic event, there are no analogous axioms in QP that rule out such scenarios.

6.2 Proper-class-sized sample spaces

Consider the following probabilistic scenario:

> *God picks a set at random.*[27]

Formally, this is a scenario $\langle \Omega, S, C \rangle$ that is characterized as follows:

1. $\Omega = \{x | x \text{ is a set}\}$.
2. $S = \{\{x\} | x \text{ is a set}\} \cup \{\Omega\}$.
3. Probabilistic constraints C.
   (a) For any sets x, y: $\{x\}|\Omega \approx \{y\}|\Omega$.

Note that the collection Ω of all sets is a proper class—that is, a collection that is 'too large' to be a set. Because NAP is formulated in set-theoretic terms, it cannot represent any probabilistic scenario whose sample space is a proper class. Moreover, it is unclear whether NAP can be easily modified to represent such a scenario, as NAP is formulated using essentially set-theoretic machinery like free ultrafilters and the axiom of choice. Thus, NAP cannot represent the above scenario, even though it seems to be conceptually possible.[28]

Although I formulated QP in set-theoretic terms in Section 4.1, it is easy to reformulate QP in class-theoretic terms. For example, the union of two

[27] For definiteness, suppose that God picks a ZF set at random. This scenario is perhaps the maximal generalization of the so-called 'de Finetti lottery', which may be thought of as a scenario in which God picks an integer at random. See Bartha [2004] for discussion of the latter.

[28] At least, insofar as we assume that it is coherent to quantify over all sets. See Rayo and Uzquiano [2006] for potential worries about this assumption.

classes A and B is simply the class that contains each element of A and each element of B, and the complement of A with respect to Ω is the class that contains the elements of Ω that are not in A. We may then define a class-theoretic notion of a Boolean algebra and understand the axioms of QP accordingly. The only axiom that requires additional care to reformulate is **Qualitative Perfect Additivity**, which involves the assumption that one collection has a greater cardinality than another collection. Although 'cardinality' is an essentially set-theoretic notion, we can reformulate this axiom with the stipulation that collection A is 'larger' than collection B just in case: (i) A is a proper class and B is not or (ii) A has a greater cardinality than B. In the Appendix, I show that QP can be reformulated in class-theoretic terms so as to ensure that there is a QP-space that can represent the above scenario.

The above probabilistic scenario is, admittedly, a rather artificial one. Nonetheless, proper-class-sized sample spaces are clearly relevant to epistemology. In particular, for any cardinality κ, it seems epistemically possible that there exist exactly κ-many things.[29] Since the collection of all cardinalities is a proper class, it is plausible that epistemic space—that is, the collection of all epistemic possibilities—is also a proper class. Thus, if we wish to do probabilistic epistemology in a manner that takes into account the totality of epistemic space, it is crucial that our epistemology employ a theory of probability that allows for proper-class-sized sample spaces. QP appears singularly suited to this task.

7 Conclusion

Infinitesimal probability has long enjoyed a prominent niche in the philosophy of probability. In this paper, I have shown that many of the philosophical purposes infinitesimal probability has been enlisted to serve can be served by appealing instead to qualitative probability—in particular, qualitative *conditional* probability. Moreover, I showed:

- QP satisfies various theoretical desiderata more simply and perspicuously than NAP.
- QP has comparable—if not greater—representational power than NAP.

[29] See Chalmers [2011, p. 90] and Pruss [2013, pp. 236–237] for arguments to this effect.

I close with some open questions about the relation between qualitative probability and infinitesimal probability.

First, as I discussed in Section 4.3.3, several notable infinitary probabilistic scenarios that can be represented by NAP functions can also be represented by QP-spaces. Is it the case that *every* probabilistic scenario that can be represented by a NAP function can also be represented by a QP-space? If so, this fact—in conjunction with the fact (cf. Section 6) that some probabilistic scenarios can be represented by QP-spaces but not by NAP functions—would establish that QP has strictly greater representational power than NAP.

Second, it remains to be seen whether any representation theorems connecting NAP functions to QP-relations can be established. For example, for a given NAP function P, is there a QP-relation $\succeq$ such that $P(A|B) \geq P(C|D)$ if and only if $A|B \succeq C|D$? Prior investigations suggest that some such representation theorems may be had. In particular, Hawthorne [2016] establishes representation theorems connecting qualitative conditional probability relations that satisfy Koopman's axioms—that is, QP1–QP9—to Popper functions. Additionally, Brickhill and Horsten [unpublished] prove a representation theorem connecting Popper functions to NAP functions. Given these connections, it seems reasonable to expect systematic connections among NAP functions and QP-relations as well.

Third, as I said in Section 6, some conceptually possible probabilistic scenarios can be represented by QP-spaces but not by NAP functions. It remains to be seen how other theories of numerical probability that allow for infinitesimals fare in this respect. For example, it may be that a theory of probability that employs Conway [1976]'s surreal numbers is better suited than NAP to represent the scenario of Section 6.2, as there are proper-class-many surreals.

Finally, infinitesimal probability has long been subject to criticism.[30] Do analogues of standard objections to infinitesimal probability apply to QP? *Prima facie*, some such objections do not carry over to QP. For example, objections—such as those of Hájek [2003b, Section 5] and Easwaran [2014a, Section 5.4]—that stem from the essential use of the axiom of choice to show the existence of infinitesimals do not straightforwardly carry over to QP. This is because the axiom of choice is generally not needed to define QP-spaces. (In the Appendix, I only employ the axiom of choice to define *total* QP-

[30] See Benci *et al.* [2016, Section 4] for notable objections.

spaces.) However, it remains to be seen whether other objections do carry over to QP and, if so, whether QP can be adequately defended in light of them.

8 Appendix

I do four things in this Appendix. First, I prove a theorem, **QP Representation**, that states general conditions under which a given probabilistic scenario is representable by a QP-space. Second, I show that **QP Representation** entails that the five probabilistic scenarios of Section 4.3.3 are all representable by QP-spaces. Third, I show that total versions of these five scenarios—that is, versions of these scenarios that include **Complete Comparability** as a constraint—are all representable by QP-spaces. Fourth, I show that the two scenarios of Section 6 are representable by QP-spaces as well.

8.1 QP Representation

In this section, I state and prove the key representation theorem of the paper. Here is the theorem.

QP Representation. *Suppose $\langle \Omega, S, C \rangle$ is a probabilistic scenario that satisfies the following conditions:*

1. *Each constraint in C has one of the following forms:*

   (i) *For every $\omega \in \Omega$, $\{\omega\} \approx_{r_\omega} \{\omega_0\}$, where ω_0 is a particular member of Ω, r is a non-negative real number, and $r_{\omega_0} = 1$.*

   (ii) *For particular $a, b \in S$ such that $a \subseteq b$, $a \approx_r b$, where $0 \leq r \leq 1$.*

   (iii) *For particular $a, b \in S$ such that $a \subseteq b$, $a \sim_r b$, where $0 \leq r \leq 1$.*

   Additionally, C does not contain constraints of both forms (ii) and (iii).

2. *For any $b \in S$, let $S_b = \{a \in S$ such that $a \approx_r b$ or $a \sim_r b$ for some $r\}$. Then, S_b is a semi-algebra over Ω.*

3. *There is some non-negative, (possibly partial) real-valued function f on S such that, if $a \subseteq b$ and ($a \approx_r b$ or $a \sim_r b$), then $r = \frac{f(a)}{f(b)}$. Moreover, f is only defined on finite subsets of Ω or infinite subsets of Ω of cardinality $|\Omega|$.*

4. *Let $a \in S$. Suppose $a = (a_1 \cup \ldots \cup a_n)$, for some positive integer n and disjoint $a_i \in S$. Further, suppose, for each a_i, that there is some real number r_i such that $a_i \approx_{r_i} a$ or $a_i \sim_{r_i} a$. Then, $r_1 + \ldots + r_n = 1$.*

Then, C can be characterized by a collection C_{QP} of constraints of the form $a|b \succeq c|d$. Moreover, $\langle \Omega, S, C \rangle$ can be represented by a QP-space $\langle \Omega', \succeq' \rangle$ in the following sense:

1. $\Omega \subseteq \Omega'$.
2. *$\succeq'$ satisfies C_{QP}.*

NOTE. The proof of this theorem, and proofs of the other results mentioned above, are still under construction.

References

Bartha, P. [2004]: 'Countable Additivity and the de Finetti Lottery', *British Journal for the Philosophy of Science*, **55**, pp. 301–321.

Benci, V., Horsten, L., and Wenmackers, S. [2013]: 'Non-Archimedean Probability', *Milan Journal of Mathematics*, **81**, pp. 121–51.

Benci, V., Horsten, L., and Wenmackers, S. [2016]: 'Infinitesimal Probabilities', *British Journal for the Philosophy of Science*, **0**, pp. 1–44.

Brickhill, H. and Horsten, L. [unpublished]: 'Popper Functions, Lexicographical Probability, and Non-Archimedean Probability'.

Chalmers, D. J. [2011]: 'The Nature of Epistemic Space', in A. Egan and B. Weatherson (*eds*), *Epistemic Modality*. Oxford: Oxford University Press, pp. 60–107.

Conway, J. [1976]: *On Numbers and Games*. London: Academic Press.

de Finetti, B. [1937]: 'La Prévision: Ses Lois Logiques, Ses Sources Subjectives', *Annales de l'Institut Henri Poincaré*, **17**, pp. 1–68.

de Finetti, B. [1972]: *Probability, Induction and Statistics*. New York: Wiley.

DiBella, N. [2018]: 'The Qualitative Paradox of Non-Conglomerability', *Synthese*, **195**, pp. 1181–1210.

Easwaran, K. [2014a]: 'Regularity and Hyperreal Credences', *Philosophical Review*, **123**, pp. 1–41.

Easwaran, K. [2014b]: 'Decision Theory without Representation Theorems', *Philosophers' Imprint*, **14**, pp. 1–30.

Fine, T. [1973]: *Theories of Probability.* New York: Academic Press.

Fishburn, P. [1986]: 'The Axioms of Subjective Probability'. *Statistical Science*, **1**, pp. 335–58.

Hájek, A. [2003a]: 'Waging War on Pascal's Wager', *Philosophical Review*, **113**, pp. 27–56.

Hájek, A. [2003b]: 'What Conditional Probability Could Not Be', *Synthese*, **137**, pp. 273–323.

Hawthorne, J. [2016]: 'A Logic of Comparative Support: Qualitative Conditional Probability Relations Representable by Popper Functions'. In A. Hájek and C. Hitchcock (*eds*), *The Oxford Handbook of Probability and Philosophy.* Oxford: Oxford University Press, pp. 277–95.

Hofweber, T. [2014]: 'Infinitesimal Chances', *Philosophers' Imprint*, **14**, pp. 1–34.

Keynes, J. M. [1921]: *A Treatise on Probability.* London: Macmillan.

Kolmogorov, A. N. [1950]: *Foundations of the Theory of Probability.* New York: Chelsea.

Koopman, B. O. [1940a]: 'The Axioms and Algebra of Intuitive Probability', *Annals of Mathematics*, **41**, pp. 269–92.

Koopman, B. O. [1940b]: 'The Bases of Probability', *Bulletin of the American Mathematical Society*, **46**, 763–774.

Krantz, D. H., Luce, R. D., Suppes, P., and Tversky, A. [1971]: *Foundations of Measurement* (Vol. 1). New York: Academic Press.

Lewis, D. [1980]: 'A Subjectivist's Guide to Objective Chance', in R. C. Jeffrey (*ed.*), *Studies in Inductive Logic and Probability*, Volume 2, Berkeley, CA: University of California Press, pp. 113–22.

Lewis, D. [1986]: 'Causation' in *Philosophical Papers*, Vol. 2. Oxford: Oxford University Press.

McGee, V. [1994]: 'Learning the Impossible', in E. Eells and B. Skyrms (*eds*), *Probability and Conditionals: Belief Revision and Rational Decision*, Cambridge: Cambridge University Press, pp. 179–99.

Meacham, C. and Weisberg, J. [2011]: 'Representation Theorems and the Foundations of Decision Theory', *Australasian Journal of Philosophy*, **89**, pp. 641–63.

Pedersen, A. P. [unpublished]: 'Strictly Coherent Preferences, No Holds Barred'.

Pivato, M. [2014]: 'Additive Representation of Separable Preferences over Infinite Products', *Theory and Decision*, **77**, pp. 31–83.

Pruss, A. [2013]: 'Probability, Regularity, and Cardinality', *Philosophy of Science*, **80**, pp. 231–40.

Rayo, A. and Uzquiano, G. [2006]: 'Introduction to *Absolute Generality*', in A. Rayo and G. Uzquiano (*eds*), *Absolute Generality*, Oxford: Oxford University Press.

Savage, L. J. [1954]: *The Foundations of Statistics*. New York: Wiley.

Scott, D. [1964]: 'Measurement Structures and Linear Inequalities', *Journal of Mathematical Psychology*, **1**, pp. 233–47.

Skyrms, B. [1980]: *Causal Necessity*. New Haven: Yale University Press.

Stefánsson, H. O. [2017]: 'What Is "Real" in Probabilism?', *Australasian Journal of Philosophy*, **95**, pp. 573–87.

Stefánsson, H. O. [forthcoming]: 'On the Ratio Challenge for Comparativism', *Australasian Journal of Philosophy*.

Suppes, P. and Zanotti, M. [1982]: 'Necessary and Sufficient Qualitative Axioms for Conditional Probability'. *Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete*, **60**, pp. 163–9.

Wenmackers, S. and Horsten, L. [2013]: 'Fair Infinite Lotteries', *Synthese*, **190**, pp. 37–61.