Stat Methods Appl (2014) 23:501–518
DOI 10.1007/s10260-014-0282-7

On the equivalence of conglomerability and disintegrability for unbounded random variables

Mark J. Schervish · Teddy Seidenfeld · Joseph B. Kadane

Accepted: 10 August 2014 / Published online: 29 August 2014

Abstract We extend a result of Dubins (Ann Probab 3:89–99, 1975) from bounded to unbounded random variables. Dubins showed that a finitely additive expectation over the collection of bounded random variables can be written as an integral of conditional expectations (disintegrability) if and only if the marginal expectation is always within the smallest closed interval containing the conditional expectations (conglomerability). We give a sufficient condition to extend this result to collections of random variables that have finite expected value and whose conditional expectations are finite and have finite expected value.

Keywords Finite additivity · Law of total probability · Daniell integral · Coherence

Mathematics Subject Classification Primary 60A05 · Secondary 28C05

1 Introduction

In discussions of the foundations of probability, a longstanding topic of debate is whether to require, beyond being finitely additive, that probabilities are countably

M. J. Schervish (✉) · T. Seidenfeld · J. B. Kadane
Department of Statistics, Carnegie Mellon University, Pittsburgh, PA 15213, USA
e-mail: mark@cmu.edu

J. B. Kadane
e-mail: kadane@stat.cmu.edu

T. Seidenfeld
Department of Philosophy, Carnegie Mellon University, Pittsburgh, PA 15213, USA
e-mail: teddy@stat.cmu.edu

additive. Specifically, we take the following three axioms to constitute the theory of finitely additive probability. Let $\{\Omega, \mathcal{B}\}$ be a measurable space. For all $A, B \in \mathcal{B}$,

Axiom 1: $0 \leq P(A) \leq 1$.
Axiom 2: $P(\Omega) = 1$.
Axiom 3: If $A \cap B = \emptyset$, then $P(A) + P(B) = P(A \cup B)$.

Countable additivity, which is taken by Kolmogorov (1956, p. 15) as an "expedient", requires the following. Let $\{A_i\}_{i=1}^{\infty}$ be elements of $\mathcal{B}$.

Axiom 4: If $A_i \cap A_j = \emptyset$ for all $i \neq j$, then $P\left(\cup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} P(A_i)$.

Call a probability *merely finitely additive* if it satisfies the first three axioms but fails the fourth one. In this paper, we assume only the first three axioms.

One major distinction between merely finitely additive probabilities and countably additive probabilities involves the theory of conditional probability. We take it as non-controversial that conditional probability satisfies the following product rule:

$$P(A \cap B) = P(B|A)P(A). \tag{1}$$

Furthermore, when the conditioning event has positive probability, $P(A) > 0$, we can use (1) to fix conditional probability by unconditional probability: $P(B|A) = P(A \cap B)/P(A)$.

However, when the conditioning event is null, $P(A) = 0$, the countably additive theory denies that conditional probability is defined given such an event A . Rather, the countably additive theory defines conditional probability through the Radon-Nikodym theorem as a solution to an integral equation with respect to a sub-σ-field.

Definition 1 Let $\mathcal{A}$ be a sub-σ-field of $\mathcal{B}$. Then $P(\cdot|\mathcal{A})(\cdot)$ is a *regular conditional probability* on $\mathcal{B}$, given $\mathcal{A}$ if

1. For each $\omega \in \Omega$, $P(\cdot|\mathcal{A})(\omega)$ is a probability on $\mathcal{B}$,
2. for each $B \in \mathcal{B}$, $P(B|\mathcal{A})(\cdot)$ is an $\mathcal{A}$-measurable function, and
3. for each $A \in \mathcal{A}$, $P(A \cap B) = \int_A P(B|\mathcal{A})(\omega) dP(\omega)$.

That is, $P(B|\mathcal{A})(\cdot)$ is a version of the Radon-Nikodym derivative of $P(\cdot \cap B)$ with respect to P defined on the sub-σ-field $\mathcal{A}$.

Kolmogorov (1956, [Section 5.2]) points out that regular conditional probabilities admit the so-called "Borel Paradox". To summarize the paradox, let $\mathcal{A}$ and $\mathcal{A}'$ be the σ-fields generated by two different random variables X and X' respectively. The paradox finds an event $C \in \mathcal{B}$, a value x of X, and a value x' of X' such that the two events $\{X = x\}$ and $\{X' = x'\}$ are identical, but $P(C|X = x) \neq P(C|X' = x')$. That is, with $B = \{X = x\} = \{X' = x'\}$, we find that $P(C|B)$ depends on which σ-field we choose for conditioning. It is well-known that, for a specific pair X and X', the sets of all x and x' values that lead to this paradox form sets of probability 0. However, Kadane et al. (1986) illustrates how one can make the paradox occur with positive probability by considering more than countably many random variables at a time.

In contrast to the countably additive theory, (see Krauss 1968; Dubins 1975) finitely additive conditional probabilities can be fully defined given each non-empty subset of Ω, while satisfying the following generalization of the product rule:

For all A, B, and C such that $B \cap C \neq \emptyset$, $P(A \cap C|B) = P(C|B)P(A|B \cap C)$.

However, the cost for these finitely additive conditional probabilities includes the penalty that they may fail to satisfy the integral property of regular conditional probabilities (clause 3 in Definition 1). In particular, there can exist a denumerable partition $\pi = \{A_i\}_{i=1}^{\infty}$ and an event B such that

$$P(B) \neq \sum_{i=1}^{\infty} P(B|A_i)P(A_i). \tag{2}$$

If (2) holds, we say that P *fails to be disintegrable* in the partition π. (See Definition 9 below for a precise definition.) Here is an elementary illustration, due to Dubins (1975) and discussed further by Kadane et al. (1996).

Example 1 Let $\Omega = \{0, 1\} \times \{1, 2, \ldots\}$. Let P be a finitely additive probability that satisfies the following:

- $P((1, i)) = 2^{-i-1}$ for $i = 1, 2, \ldots$,
- $P(B) = 1/2$, where $B = \{(0, 1), (0, 2), \ldots\}$,
- $P((0, i)) = 0$ for $i = 1, 2, \ldots$.

Define the partition $\pi = \{A_i\}_{i=1}^{\infty}$, where $A_i = \{(0, i), (1, i)\}$. Since $P(A_i) = 2^{-i-1} > 0$ for all i, we have

$$P(B|A_i) = \frac{P(B \cap A_i)}{P(A_i)} = 0,$$

for all i. Hence $\sum_{i=1}^{\infty} P(B|A_i)P(A_i) = 0 \neq P(B)$.

The concept of disintegrability is relevant to understanding some otherwise anomalous features of Bayesian statistical inference that arise when using so-called *improper* priors. These are instances of the so-called *marginalization paradoxes* of Dawid et al. (1973). As Kadane et al. (1996, [Section 5]) explains, an improper prior, e.g. Lebesgue measure over the whole real line, corresponds to a merely finitely additive prior probability on the real line. Each unit interval has equal probability, i.e. probability 0. Even when the formal posterior computed from the improper prior turns out to be countably additive, the joint (finitely additive) probability may fail to be disintegrable in the partition determined by the data.

In Example 1, we also see that the conditional probabilities $P(B|A_i)$ have the property that

$$\text{there exists } \epsilon > 0 \text{ such that } P(B) > P(B|A_i) + \epsilon \text{ for every } i. \tag{3}$$

De Finetti (1930) says that such a probability *fails conglomerability* in the partition. (See Definition 9 for a more precise definition.) Schervish et al. (1984) show that each merely finitely additive probability fails conglomerability in some denumerable partition, which is not possible for countably additive probabilities. As we saw above, for nondenumerable partitions, the countably additive theory imposes disintegrability on the definition of conditional probability. Consequently the theory is not able to define $P(A|B)$ for arbitrary sets A and B, as we saw in the Borel Paradox.

Dubins (1975) established an important equivalence between conglomerability and disintegrability of finitely additive expectations, which de Finetti (1974) calls (coherent) previsions (see Definition 2 below). Dubins showed that, with respect to the collection of bounded random variables, by replacing countably additive probability and conditional probability with the more general concepts of (finitely additive) expectations and conditional expectations, then a finitely additive expectation function is disintegrable in a partition if and only if its conditional expectations are conglomerable in that partition. It follows easily from (3) that, when a finitely additive expectation fails conglomerability in a partition π, then it fails to be disintegrable in π. The converse inference is the heart of Dubins' result.

In this paper we extend Dubins' result to particular collections of unbounded random variables. In Sect. 2 we review de Finetti's concept of coherent previsions and conditional previsions, and describe a theory of finitely additive integrals/expectations for unbounded random variables. In Sect. 3 we review Dubins' result and discuss how to extend conglomerability to unbounded random variables. In Sect. 4 we give conditions under which the equivalence between disintegrability and conglomerability extends to unbounded random variables. In particular, we restrict attention to random variables for which previsions and conditional previsions are finite. Example 5 illustrates how the equivalence can fail if our conditions are not met. We offer a concluding discussion in Sect. 6.

2 Background

Let Ω be a fixed non-empty set, and define a *random variable* to be a real-valued function X defined on Ω. We require that $X(\omega)$ be finite for all $\omega \in \Omega$, but not necessarily bounded. All of the collections of random variables discussed in the definitions and results of this paper are allowed to include unbounded random variables unless explicitly stated otherwise.

The concept of coherent prevision on a collection of random variables was introduced by de Finetti (1974).

Definition 2 Let $\mathcal{U}$ be a collection of random variables defined on Ω. A function $P : \mathcal{U} \to \mathbb{R}$ is called a *prevision*. We say that P is *incoherent* if there exists a finite subset $\{X_1, \ldots, X_n\}$ of $\mathcal{U}$ and scalars $\alpha_1, \ldots, \alpha_n$ and $\epsilon > 0$ such that, for all $\omega \in \Omega$,

$$\sum_{i=1}^{n} \alpha_i [X_i(\omega) - P(X_i)] < -\epsilon. \tag{4}$$

If P is not incoherent, we say that P is *coherent*.

An equivalent, and sometimes more convenient, way to define coherent prevision is to say that P is coherent if, for every finite subset $\{X_1, \ldots, X_n\}$ of $\mathcal{U}$ and all scalars $\alpha_1, \ldots, \alpha_n$,

$$\sup_{\omega \in \Omega} \sum_{i=1}^{n} \alpha_i [X_i(\omega) - P(X_i)] \geq 0. \tag{5}$$

It is not difficult to see that this is equivalent to Definition 2. As defined above, a coherent prevision P must assume only finite values, otherwise (5) would be impossible and/or undefined. There are ways to generalize the concept of coherent prevision to allow infinite values. (See Berti et al. 2001; Crisma and Gigante 2001; Crisma et al. 1997; Schervish et al. 2014b for some of these generalizations). Such generalizations play no role in the results of this paper. Our theorems apply only to sets of random variables for which all previsions (and conditional previsions) are finite.

There are many coherent previsions on the set $\mathcal{X}$ of bounded random variables, and each of them is a finitely additive probability when restricted to the collection of indicator functions of subsets of Ω. That is, using the standard notation of letting the name of an event stand for its indicator function, $P(\Omega) = 1$, $P(A \cup B) = P(A) + P(B)$ when $A \cap B = \emptyset$, and $P(A) \geq 0$ for all $A \subseteq \Omega$. By finite additivity and linearity of coherent prevision, if $X = \sum_{i=1}^{n} \alpha_i A_i$ is a simple function (one that assumes only finitely many distinct values) the prevision of X equals $\sum_{i=1}^{n} \alpha_i P(A_i)$. This resembles the formula for the integral of a simple function in the usual measure theoretic derivation. To carry the resemblance further, the value of P for every bounded X is uniquely determined from the finitely additive probability by means of the fact that

$$P(X) = \sup_{\text{simple } Y \leq X} P(Y) = \inf_{\text{simple } Y \geq X} P(Y). \tag{6}$$

The first equation in (6) is the same way that the Lebesgue integral of a nonnegative function X is defined in terms of the integrals of simple functions. Indeed P can be expressed as a finitely additive integral. Generalizing from the definition of Daniell integral in Royden (1968, [Chapter 13]), we can call a coherent prevision a finitely additive Daniell integral. Definition 3 applies to both bounded and unbounded functions. It also applies to functions defined on arbitrary spaces and to finitely additive set functions that are not probabilities.

Definition 3 Let $\mathcal{L}$ be a linear space of functions defined on a common space Γ such that $\mathcal{L}$ contains all constants. Let L be a linear functional defined on $\mathcal{L}$ that satisfies $L(X) \geq 0$ whenever $X(\omega) \geq 0$ for all ω. Then L is called a *nonnegative linear functional* or a *finitely additive Daniell integral* over $\mathcal{L}$. We can write $L(X) = \int_\Omega X(\omega) L(d\omega)$. If L is a nonnegative linear functional such that $L(1) = 1$, we call L a *finitely additive expectation.*

The finitely additive Daniell integral is equivalent to the integral as developed by Dunford and Schwartz (1958, [Chapter III]) for bounded random variables while remaining equivalent to the notion of coherent prevision (finitely additive expectation) for unbounded random variables. (See Proposition 1 below.) For unbounded random

variables, not all real-valued coherent previsions admit an integral representation in the sense of Dunford and Schwartz (1958). On the other hand, all real-valued coherent previsions admit a representation as a finitely additive Daniell integral. For discussion of the problem, see Berti et al. (2001), Berti and Rigo (2000), Berti and Rigo (2002), Schervish et al. (2008), Seidenfeld et al. (2009).

Here is an example of a finitely additive expectation that is not countably additive. This example will be used later to illustrate our main result.

Example 2 Let P be a countably additive probability on a measurable space $(\Omega, \mathcal{B})$ such that Ω has infinitely many elements, and let $\mathcal{X}$ be the collection of all bounded measurable random variables. It is not difficult to show that the only coherent prevision for each bounded random variable is $P(X)$ equal to its countably additive expected value. Also, let Y be a random variable that is unbounded above and bounded below. Suppose that Y has finite expected value y by the usual countably additive definition. We will show that $P(Y) = p$ is coherent with the previsions of all of the bounded random variables if and only if $p \geq y$. For the "only if" direction, note that $y = \sup_{\text{bounded } X \leq Y} P(X)$, and coherence requires that $P(X) \leq P(Y)$ for every bounded $X \leq Y$. It follows that $y \leq P(Y)$ is necessary for coherence. For the "if" direction, suppose that $p \geq y$. We prove that P extends to a nonnegative linear functional L on the span $\mathcal{L}$ of the bounded random variables and Y with $L(1) = 1$. Every element of $\mathcal{L}$ has a unique representation as $\alpha Y + X$ for some real α and some bounded random variable X. Define $L(\alpha Y + X) = \alpha p + P(X)$, which satisfies $L(1) = 1$ and is clearly linear and well-defined. To see that L is nonnegative, note that $\alpha Y + X \geq 0$ only if $\alpha \geq 0$, in which case, $L(\alpha Y + X) \geq \alpha y + L(X)$, which equals the countably additive integral of $\alpha Y + X$, which in turn is nonnegative.

To see that P is not countably additive when $p > y$, let $\{X_n\}_{n=1}^{\infty}$ be a countable sequence of bounded random variables that increase monotonically to Y, such as $X_n = \min\{Y, n\}$. Then $Y - X_1 = \sum_{n=2}^{\infty}(X_n - X_{n-1})$, with each term $X_n - X_{n-1} \geq 0$. Countable additivity would require that

$$P(Y - X_1) = \sum_{n=2}^{\infty} P(X_n - X_{n-1}).$$

But the left-hand side is $p - P(X_1)$ while the right-hand side is $y - P(X_1)$, which are not equal.

Although the space $\mathcal{L}$ in Definition 3 may contain unbounded functions, L must assume only finite values since it is a linear functional. (de Finetti, 1974, Section 3.9) proves the following result which makes clear the connection between finitely additive expectation and coherent prevision.

Proposition 1 *P is a finite coherent prevision on a set $\mathcal{U}$ of random variables if and only if there exists a finitely additive expectation L on a linear space $\mathcal{L}$ that contains $\mathcal{U}$ and all constants such that $L(X) = P(X)$ for every $X \in \mathcal{U}$.*

Coherent conditional prevision can be defined in a manner similar to coherent prevision.

Definition 4 Let $\mathcal{Q}$ be a collection of pairs (X, h) where X is a random variable and h is a nonempty subset of Ω. A function $P(\cdot|\cdot) : \mathcal{Q} \to \mathbb{R}$ is called a *conditional prevision*. We say that P is *coherent* if, for every finite subset $\{(X_1, h_1), \ldots, (X_n, h_n)\}$ of $\mathcal{Q}$ and scalars $\alpha_1, \ldots, \alpha_n$ $\omega \in \Omega$,

$$\sup_{\omega \in \Omega} \sum_{i=1}^{n} \alpha_i h_i [X_i(\omega) - P(X_i|h_i)] \geq 0.$$

If P is not coherent, we say that P is *incoherent*.

By comparing Definitions 2 and 4, it is easy to see that $P(X)$ must be the same as $P(X|\Omega)$ for every random variable X. This makes coherent prevision the special case of coherent conditional prevision in which, for every pair $(X, h) \in \mathcal{Q}$, we have $h = \Omega$. Hence, whenever we refer to a prevision $P(X)$ we mean that $(X, \Omega) \in \mathcal{Q}$ and $P(X) = P(X|\Omega)$.

When $P(h) = 0$ coherence (as defined in Definition 4) is insufficient to ensure that $P(X|h)$ has even the most basic intuitive properties. For example, if $P(h) = 0$ then $P(h|h) \neq 1$ is coherent. To insure that conditional probabilities behave as much like probabilities as possible, De Finetti introduces an additional assumption in (1974, [Appendix 16]) as do Regazzini (1987), and Crisma and Gigante (2001). For example, Crisma and Gigante (2001), de Finetti (1974), Regazzini (1987) choose a stronger definition of coherent conditional prevision, which we call DRCG-coherence in Definition 5:

Definition 5 Let $\mathcal{Q}$ be a collection of pairs (X, h) where X is a random variable and h is a nonempty subset of Ω. A conditional prevision P defined on $\mathcal{Q}$ is called *DRCG-coherent* if, for every finite subset $\{(X_1, h_1), \ldots, (X_n, h_n)\}$ of $\mathcal{Q}$ and scalars $\alpha_1, \ldots, \alpha_n$ $\omega \in \Omega$,

$$\sup_{\omega \in h_0} \sum_{i=1}^{n} \alpha_i h_i [X_i(\omega) - P(X_i|h_i)] \geq 0, \tag{7}$$

where $h_0 = \bigcup_{i=1}^{n} h_i$.

On the other hand, Dubins (1975) makes a weaker assumption that we generalize here for use with unbounded random variables.

Definition 6 Let π be a partition of Ω, and let P be a finite coherent conditional prevision defined on a set $\mathcal{Q}$. We say that P *contains a π-strategy* if, for each $h \in \pi$:

1. $\mathcal{L}_h = \{X : (X, h) \in \mathcal{Q}\}$ is a nonempty linear space that contains h and all constants,
2. $P(\cdot|h)$ is a finitely additive expectation on $\mathcal{L}_h$, and
3. $P(X|h) = c$ for each random variable X that equals the constant c on the whole set h.

Instead of the third condition in Definition 6, Dubins (1975) assumes $P(h|h) = 1$, which is equivalent to the third condition when all random variables are bounded. Theorem 1, our main result, assumes that P contains a π-strategy.

It is straightforward to show that, for every partition π, every finite DRCG-coherent prevision either contains a π-strategy or can be extended so as to contain a π-strategy. Rather than strengthen the definition of coherence, we prefer to add assumptions to theorems as needed. In Schervish (2014b, [Example 2]) we illustrate our reason for this preference.

3 Conglomerability and disintegrability

We turn now to precise definitions of conglomerability and disintegrability. Let π be a partition of Ω. That is, π is a collection $\{h : h \in \pi\}$ of mutually disjoint subsets of Ω such that their union is Ω. The conditional prevision of each random variable X given each element h of π is denoted $P(X|h)$. In order to make sense out of the loose phrase "the integral of conditional expectations," we need to be precise about what it means to integrate a conditional prevision. In this section, all integrals are intended in the sense of Definition 3.

3.1 The Integral of a conditional expectation

Nonnegative linear functionals behave in many ways like the countably additive Lebesgue integral. One property that they share is the following transformation property that we use.

Lemma 1 *Let $\mathcal{L}$ be a linear space of real-valued functions on Ω that includes all constants, let Θ be a set, and let $\Pi : \Omega \to \Theta$ be a function. Let L be a nonnegative linear functional defined on $\mathcal{L}$. Let $\mathcal{V}$ be a linear space of real-valued functions on Θ that includes all constants and such that, $Z(\Pi) \in \mathcal{L}$ for every $Z \in \mathcal{V}$. Define $L_\Pi(Z) = L(X)$, where $X = Z(\Pi)$. Then L_Π is a nonnegative linear functional on $\mathcal{V}$.*

Proof If $Z_1, Z_2 \in \mathcal{V}$ and $\alpha, \beta \in \mathbb{R}$, then

$$\begin{aligned} L_\Pi(\alpha Z_1 + \beta Z_2) &= L[\alpha Z_1(\Pi) + \beta Z_2(\Pi)] = \alpha L[Z_1(\Pi)] + \beta L[Z_2(\Pi)] \\ &= \alpha L_\Pi(Z_1) + \beta L_\Pi(Z_2). \end{aligned}$$

If $Z(\theta) \geq 0$ for all $\theta \in \Theta$, then $X(\omega) = Z(\Pi(\omega)) \geq 0$ for all $\omega \in \Omega$. Hence $L_\Pi(Z) = L(X) \geq 0$. □

Definition 7 We call L_Π in Lemma 1 the *integral induced from L by* Π.

Our primary use of Lemma 1 is to give meaning to the concept of integrating over a partition. In particular, we will give a precise meaning to "integral of conditional expectations." To that end, let π be a partition and let X be a random variable such that $P(X|h)$ is defined and finite for all $h \in \pi$. Define

$$X_\pi(h) = P(X|h), \text{ for all } h \in \pi, \tag{8}$$

$$\Pi(\omega) = \text{ that unique } h \in \pi \text{ such that } \omega \in h, \tag{9}$$

$$\begin{aligned} P(X|\pi)(\omega) &= X_\pi(\Pi(\omega)) \\ &= P(X|h), \text{ for that unique } h \in \pi \text{ such that } \omega \in h, \end{aligned} \tag{10}$$

so that X_π is a real-valued function on π, $P(X|\pi)$ is a real-valued function on Ω, and $P(X|\pi) = X_\pi(\Pi)$. Let $\mathcal{L}$ be a linear space of random variables that contains all constants, the domain of P, and $P(X|\pi)$. Assuming that it is possible to extend P to $\mathcal{L}$ in such a way that all previsions are finite, then we can apply Lemma 1 with $\Theta = \pi$, $\mathcal{V}$ equal to the linear span of the constants and the random variable X_π, L equal to the extension of P, and Π as defined in (9). Let P_π denote the integral induced from P by Π, denoted L_Π in Lemma 1. With the notation just introduced, Lemma 1 implies that

$$P_\pi(X_\pi) = \int_\pi P(X|h)P_\pi(dh) = \int_\Omega P[X|\Pi(\omega)]P(d\omega) = P[P(X|\pi)], \tag{11}$$

if any of the four terms in (11) is finite. In summary, we have the following:

Definition 8 Let π be a partition of Ω, and let X be a random variable such that $P(X|h)$ is defined and finite for all $h \in \pi$ and $P[P(X|\pi)]$ is defined and finite. Then (11) is called *the integral of the conditional expectations given* π.

3.2 Precise definitions

Definition 9 Let $\mathcal{W}$ be a collection of random variables. Let π be a partition, and let P be a finite coherent conditional prevision on a set $\mathcal{Q}$ that contains both $\mathcal{W} \times \pi$ and $\mathcal{W} \times \{\Omega\}$. We say that P *is conglomerable in* π *with respect to* $\mathcal{W}$ if, for each $X \in \mathcal{W}$,

$$\inf_{h \in \pi} P(X|h) \le P(X) \le \sup_{h \in \pi} P(X|h). \tag{12}$$

We say that P *is disintegrable in* π *with respect to* $\mathcal{W}$ if, for each $X \in \mathcal{W}$,

$$P(X) = \int_\pi P(X|h)P_\pi(dh). \tag{13}$$

In view of (11), there is an alternative way to express that P is disintegrable in a partition.

Proposition 2 *P is disintegrable in π with respect to $\mathcal{W}$ if and only if, for each $X \in \mathcal{W}$, $P(X) = P[P(X|\pi)] = P_\pi(X_\pi)$.*

When the conclusion of Proposition 2 holds, Schervish et al. (2014a) says that P satisfies the Law of Total Previsions in π.

Readers of Dubins (1975) will note that the definition of conglomerable in Definition 9 looks different from the corresponding definition that Dubins gave in Dubins

(1975). Specifically, the definition in Dubins (1975) is that P is conglomerable in π with respect to the collection $\mathcal{W}$ if

$$\text{for all } X \in \mathcal{W},\ P(X|h) \geq 0 \text{ for all } h \in \pi \text{ implies } P(X) \geq 0. \tag{14}$$

Definition 9 is a straightforward generalization of the definition that de Finetti (de Finetti 1974, p. 143) gives for indicators of events. Definition 9 and (14) are equivalent when $\mathcal{W} = \mathcal{X}$, the collection of all bounded random variables. The proof relies on the fact that $\mathcal{X}$ is a linear space and contains all constants. The two definitions are not necessarily equivalent for every collection that is not a linear space and/or does not contain all constants.

Example 3 Consider the same situation as Example 1. Let $\mathcal{W}$ be the collection of all nonnegative bounded random variables. Hence, $\mathcal{W}$ is not a linear space. Because each $X \in \mathcal{W}$ is nonnegative, it follows that $P(X|A_i) \geq 0$ for all i and $P(X) \geq 0$. Hence, (14) holds. On the other hand, let $X((j,i)) = j$ for all $j = 0, 1$ and $i = 1, 2, \ldots$. Then $P(X|A_i) = 1$ for all i while $P(X) = 1/2$ and P is not conglomerable by Definition 9.

In order to maintain the spirit of Dubins' definition when $\mathcal{W}$ is not a linear space or does not contain all constants, we need to strengthen (14).

Definition 10 Let $\mathcal{W}$ be a collection of random variables. Let π be a partition, and let P be a finite coherent conditional prevision on a set $\mathcal{Q}$ that contains both $\mathcal{W} \times \pi$ and $\mathcal{W} \times \{\Omega\}$. We say that P is *D-conglomerable in π with respect to $\mathcal{W}$* if the following is true. For all $X \in \mathcal{W}$ and all real c,

- $P(X|h) \leq c$ for all $h \in \pi$ implies $P(X) \leq c$, and
- $P(X|h) \geq c$ for all $h \in \pi$ implies $P(X) \geq c$.

We now show that Definition 10 is equivalent to conglomerability from Definition 9 for finite previsions.

Lemma 2 *Let $\mathcal{W}$ be a collection of random variables and let π be a partition. Let P be a finite coherent conditional prevision on a set $\mathcal{Q}$ that contains both $\mathcal{W} \times \{\Omega\}$ and $\mathcal{W} \times \pi$. Then P is conglomerable in π with respect to $\mathcal{W}$ if and only if P is D-conglomerable in π with respect to $\mathcal{W}$.*

Proof For the "if" direction, suppose that P is D-conglomerable in π with respect to $\mathcal{W}$. Let $X \in \mathcal{W}$, and let $c_1 = \inf_{h \in \pi} P(X|h)$ and $c_2 = \sup_{h \in \pi} P(X|h)$. Finiteness of P implies that $c_1 < \infty$ and $c_2 > -\infty$. If $c_1 = -\infty$, the second bullet in Definition 10 is vacuous. If c_1 is finite, then $P(X|h) \geq c_1$ for all $h \in \pi$ and Definition 10 says that $P(X) \geq c_1$. Similarly if $c_2 = \infty$, the first bullet in Definition 10 is vacuous. If c_2 is finite, then $P(X|h) \leq c_2$ for all $h \in \pi$ so that $P(X) \leq c_2$. Hence (12) holds.

For the "only if" direction, suppose that P is conglomerable in π with respect to $\mathcal{W}$. Let $X \in \mathcal{W}$. Then (12) holds. Let c be a real number. If $P(X|h) \geq c$ for all $h \in \pi$, then $c \leq \inf_{h \in \pi} P(X|h) \leq P(X)$ by (12). Similarly, if $P(X|h) \leq c$ for all $h \in \pi$, then $c \geq \sup_{h \in \pi} P(X|h) \geq P(X)$. □

4 Extending the equivalence of conglomerability and disintegrability to unbounded variables

Lemma 3 shows that disintegrability implies conglomerability for arbitrary collections.

Lemma 3 *Let $\mathcal{W}$ be a collection of random variables. Let π be a partition, and let P be a finite coherent conditional prevision on a set $\mathcal{Q}$ that contains both $\mathcal{W} \times \pi$ and $\mathcal{W} \times \{\Omega\}$. Suppose that P is disintegrable in π with respect to $\mathcal{W}$. Then P is conglomerable in π with respect to $\mathcal{W}$.*

Proof Let $X \in \mathcal{W}$, and define X_π by (8). Because P_π is coherent,

$$\inf_{h \in \pi} X_\pi(h) \leq P_\pi(X_\pi) \leq \sup_{h \in \pi} X_\pi(h). \tag{15}$$

By disintegrability, $P(X) = P_\pi(X_\pi)$, hence (15) implies (12). Since the above argument applies to all $X \in \mathcal{W}$, P is conglomerable in π with respect to $\mathcal{W}$. □

In light of Lemma 3, every set $\mathcal{W}$ of random variables, falls into one of three classes relative to P and a partition π.

Definition 11 Let $\mathcal{W}$ be a collection of random variables. Let π be a partition, and let P be a finite coherent conditional prevision on a set $\mathcal{Q}$ that contains both $\mathcal{W} \times \pi$ and $\mathcal{W} \times \{\Omega\}$. We say that

- *$\mathcal{W}$ is of Class 0 relative to P and π* if P is neither conglomerable nor disintegrable in π with respect to $\mathcal{W}$.
- *$\mathcal{W}$ is of Class 1 relative to P and π* if P is conglomerable in π with respect to $\mathcal{W}$ but P is not disintegrable in π with respect to $\mathcal{W}$.
- *$\mathcal{W}$ is of Class 2 relative to P and π* if P is both conglomerable and disintegrable in π with respect to $\mathcal{W}$.

Theorem 1 of Dubins (1975) can be reexpressed as saying that, for each partition π and each coherent conditional prevision P, the collection $\mathcal{X}$ of bounded random variables is either of Class 0 or of Class 2 but never of Class 1 relative to P and π. In Example 5, we give an example of a coherent prevision P, a partition π, and a collection $\mathcal{Y}$ of random variables such that $\mathcal{X} \subset \mathcal{Y}$ and $\mathcal{Y}$ is of Class 1 relative to P and π. The following result is a straightforward consequence of the class definitions.

Proposition 3 *Let $\mathcal{W}$ be a collection of random variables. Let π be a partition, and let P be a finite coherent conditional prevision on a set $\mathcal{Q}$ that contains both $\mathcal{W} \times \pi$ and $\mathcal{W} \times \{\Omega\}$. If $\mathcal{W}$ is of Class 0 relative to P and π, then every superset of $\mathcal{W}$ is also of Class 0. If $\mathcal{W}$ is of Class 2 relative to P and π, then every subset of $\mathcal{W}$ is also of Class 2.*

Our extension of Dubins' theorem gives a sufficient condition for a collection $\mathcal{W}$ of random variables to be not of Class 1. Berti and Rigo (1992, [Theorem 3.1]) prove a similar theorem for bounded random variables.

Theorem 1 *Let $\mathcal{W}$ be a collection of random variables. Let π be a partition, and let P be a finite conditional prevision on a set $\mathcal{Q}$. Assume that P contains a π-strategy and that $\mathcal{Q}$ contains both $\mathcal{W} \times \pi$ and $\mathcal{W} \times \{\Omega\}$. Assume that $(P(X|\pi), \Omega) \in \mathcal{Q}$, and that*

$$\text{for every } X \in \mathcal{W},\ X - P(X|\pi) \in \mathcal{W}. \tag{16}$$

Then, with respect to the collection $\mathcal{W}$, P is conglomerable in π if and only if P is disintegrable in π.

Proof Because $P(X|\pi)(\omega) = P(X|h)$ for all $\omega \in h$, $P[P(X|\pi)|h] = P(X|h)$ for all $h \in \pi$ and all $X \in \mathcal{W}$. By linearity of $P(\cdot|h)$, we get $P[X - P(X|\pi)|h] = 0$, hence

$$\inf_{h\in\pi} P[X - P(X|\pi)|h] = 0 = \sup_{h\in\pi} P[X - P(X|\pi)|h].$$

We have assumed that $X - P(X|\pi) \in \mathcal{W}$. If P is conglomerable in π with respect to $\mathcal{W}$, then $P[X - P(X|\pi)] = 0$, from which it follows that $P(X) = P[P(X|\pi)]$, so that P is disintegrable in π with respect to $\mathcal{W}$. If P is disintegrable in π then P is conglomerable in π by Lemma 3. □

The key assumption in Theorem 1 is (16). For an arbitrary collection $\mathcal{W}$, define

$$\mathcal{W}_- = \{X - P(X|\pi) : X \in \mathcal{W}\},$$
$$\mathcal{W}_+ = \mathcal{W} \cup \mathcal{W}_-.$$

The following results (the second of which is trivial) help to distinguish some collections of random variables by their classes.

Lemma 4 *Let $\mathcal{W}$ be a collection of random variables. Let π be a partition, and let P be a finite conditional prevision on a set $\mathcal{Q}$. Assume that P contains a π-strategy and that $\mathcal{Q}$ contains both $\mathcal{W} \times \pi$ and $\mathcal{W} \times \{\Omega\}$. Then*

1. *$\mathcal{W}_+$ satisfies (16),*
2. *$\mathcal{W}$ is of Class 2 relative to P and π if and only if $\mathcal{W}_+$ is also of Class 2, and*
3. *If $\mathcal{W}$ is not of Class 2 relative to P and π, then $\mathcal{W}_+$ is of Class 0.*

Proof For part 1, let $X \in \mathcal{W}$ so that $X - P(X|\pi) \in \mathcal{W}_+$. Also $X_\pi = [P(X|\pi)]_\pi$, hence $[X - P(X|\pi)]_\pi$ is identically 0, and

$$X - P(X|\pi) - [X - P(X|\pi)]_\pi(\Pi) \in \mathcal{W}_+.$$

For part 2, the "if" direction is immediate from Proposition 3. For the "only if" direction, note that for every $Y \in \mathcal{W}_-$, Y_π is identically 0 and $P(Y) = 0$ if $\mathcal{W}$ is of Class 2. For part 3, Theorem 1 says that $\mathcal{W}_+$ is either of Class 0 or Class 2. If $\mathcal{W}$ is not of Class 2, then no superset of it, such as $\mathcal{W}_+$, can be of Class 2. Hence $\mathcal{W}_+$ must be of Class 0. □

Proposition 4 *Let $\mathcal{W}$ be a collection of random variables. Let π be a partition, and let P be a countably additive prevision on a set $\mathcal{Q}$ that contains both $\mathcal{W} \times \pi$ and $\mathcal{W} \times \{\Omega\}$. If every element of π has positive probability, then $\mathcal{W}$ is of Class 2 relative to P and π.*

One subtle point concerning Proposition 4 is that P can be a countably additive prevision on the collection of all bounded random variables but fail to be countably additive on a collection that includes unbounded random variables. Examples 2, 4, and 5 illustrate this circumstance. As such, these examples illustrate how a probability can be both conglomerable and disintegrable in every partition relative to the class of bounded random variables, but not so with respect to larger classes that include unbounded random variables.

5 Examples

In this section, we extend Example 2 to illustrate both situations in which the conditions of Theorem 1 are satisfied and situations in which the conditions are not satisfied. Example 4 contains examples of Theorem 1 in which a collection $\mathcal{W}$ of random variables is of Class 0 relative to P and π_1 for a partition π_1, while $\mathcal{W}$ is of Class 2 relative to P and π_2 for a different partition π_2. Example 5 contains three different collections of random variables (one of which is the $\mathcal{W}$ from Example 4) that are of Classes 0, 1, and 2 relative to the same partition. In particular, the conditions of Theorem 1 fail for the collection that is of Class 1.

Example 4 We will make use of the construction described in Example 2. Let Ω be the set of ordered pairs of integers from 1 on up. Define $P(\{(x, y)\}) = 2^{-x-y}$ for all $x, y \geq 1$. Then P is countably additive as a probability on Ω. Let Y be the unbounded random variable $Y(x, y) = y$ for all x, y, whose countably additive expected value is 2. We saw in Example 2 that a necessary and sufficient condition for $P(Y) = p$ to be coherent with the previsions for the bounded random variables is $p \geq 2$. In this example, we choose $P(Y) = 4$, and extend P to the linear span of Y and bounded random variables as in Example 2. Consider the following two partitions:

$$\pi_1 = \{h_x : x \geq 1\},$$
$$\pi_2 = \{g_y : y \geq 1\},$$

where $h_x = \{(x, y) : y \geq 1\}$ for each $x \geq 1$, and $g_y = \{(x, y) : x \geq 1\}$ for each $y \geq 1$. It is straightforward to show that $P(h_x) = 2^{-x}$ for all $x \geq 1$, and $P(g_y) = 2^{-y}$ for all $y \geq 1$. Hence, for all x and y, $P(g_y|h_x) = 2^{-y}$ and $P(h_x|g_y) = 2^{-x}$. For each bounded random variable X, $P(X)$ must equal its countably additive expected value.

In order to satisfy the conditions of Theorem 1, we will assign finite coherent previsions and finite coherent conditional previsions given h_x, g_y for all x and y and given π_1 and π_2 to all random variables in the linear span $\mathcal{W}$ of the set $\mathcal{Y} = \mathcal{X} \bigcup \{Y, P(Y|\pi_1), P(Y|\pi_2)\}$, where $\mathcal{X}$ is the set of all bounded random variables. Because each nonempty subset of Ω has positive probability, P will contain both a π_1-strategy and a π_2-strategy so long as P is coherent.

First, note that $P(Y|\pi_2) = Y$. Second,

$$P(Y|\pi_1)(x, y) = P(Y|h_x) = P(Yh_x)2^x,$$

so that $P(Y|\pi_1)$ will be defined as soon as we choose coherent values for $P(Yh_x)$ for all $x \geq 1$. Note that $P(Yh_x)2^x$ might be a bounded or an unbounded sequence depending on what values we choose for $P(Yh_x)$. We don't need $Yh_x \in \mathcal{W}$ for each x, but we do need $P(Y|\pi_1) \in \mathcal{W}$. So we need to choose coherent values for $P(Yh_x)$ for all x.

An argument similar to the one given in Example 2 shows that, for each x, $P(Yh_x)$ must be assigned a value greater than or equal to 2^{-x+1}. Let $r_x = P(Yh_x) - 2^{-x+1}$ for all x, so that for all x,

$$\begin{aligned} P(Yh_x) &= r_x + 2^{-x+1}, \\ P(Y|\pi_1)(x, y) &= 2 + r_x 2^x. \end{aligned} \tag{17}$$

Since $Y = \sum_{x=1}^{\infty} Yh_x$, and $Yh_x \geq 0$ for all x, it follows that the following are necessary conditions for (17) to be coherent assignments:

$$r_x \geq 0, \text{ and } \sum_{x=1}^{\infty} r_x \leq 2. \tag{18}$$

Because $2 + r_x 2^x \geq 0$ for all x, we have

$$P[P(Y|\pi_1)] \geq \sum_{x=1}^{\infty}\sum_{y=1}^{\infty}(2 + r_x 2^x)(2^{-x-y}) = 2 + \sum_{x=1}^{\infty} r_x, \tag{19}$$

with equality in (19) required if $r_x 2^x$ is bounded. To verify that $r_x 2^x$ can be either bounded or unbounded while satisfying all of the other conditions above, consider the following two examples: for all x, either $r_x = 2^{-x}$ or $r_x = 0.6^x$.

We show next that (18) and (19) are also sufficient for coherence. Assume that (18) and (19) hold. Let

$$\begin{aligned} u &= 2 + \sum_{x=1}^{\infty} r_x, \text{ and,} \\ v &= P[P(Y|\pi_1)] - u. \end{aligned}$$

Note that $v \geq 0$ with $v = 0$ if $r_x 2^x$ is bounded. The most general gamble allowed by (17) and the previsions from Example 2 is

$$Z = \beta(Y-4) + \sum_{x=1}^{\infty}\sum_{y=1}^{\infty}\alpha_{x,y}[I_{\{(x,y)\}} - 2^{-x-y}] + \delta[P(Y|\pi_1) - u - v]$$
$$+ \sum_{i=1}^{n}\gamma_i\left(\left[\sum_{y=1}^{\infty} y I_{\{x_i,y\}}\right] - r_{x_i} - 2^{-x_i+1}\right), \tag{20}$$

where $x_i, \ldots, x_n \geq 1$ are distinct integers, and $\{\alpha_{x,y} : x, y \geq 1\}$ is a bounded collection of numbers. When $\omega = (x, y)$, the value of (20) is

$$Z(x, y) = \beta(y-4) + \alpha_{x,y} - q + \delta[2 + r_x 2^x - u - v] \tag{21}$$
$$+ \sum_{i=1}^{n}\gamma_i[y I_{\{x_i\}}(x) - 2^{-x_i+1}] - t,$$

where $t = \sum_{i=1}^{n}\gamma_i r_{x_i}$, and $q = \sum_{x=1}^{\infty}\sum_{y=1}^{\infty} 2^{-x-y}\alpha_{x,y}$, both finite numbers.

If, for every choice of $n, \beta, \delta, \gamma_1, \ldots, \gamma_n$, and $\{\alpha_{x,y} : x, y \geq 1\}$, there exists a pair (x, y) such that (21) is nonnegative, then the previsions are coherent. If $\beta > 0$, then (21) is unbounded above along sequences with fixed x and unbounded y. Similarly, if $r_x 2^x$ is unbounded, then $\delta > 0$ makes (21) unbounded above along sequences with fixed y and unbounded x. Hence, for the rest of the proof we will assume that $\beta \leq 0$ and either $r_x 2^x$ is bounded (in which case $v = 0$) or $r_x 2^x$ is unbounded and $\delta \leq 0$. The weighted average of (21) with weights 2^{-x-y} is

$$\sum_{x=1}^{\infty}\sum_{y=1}^{\infty} 2^{-x-y} Z(x, y) = -2\beta - t - v\delta. \tag{22}$$

Since either $\delta \leq 0$ or $v = 0$ implies that $-v\delta \geq 0$, the right-hand side of (22) is at least $-2\beta - t$. If $t \leq -2\beta$, then at least one value of $Z(x, y)$ must be nonnegative, because a weighted average of the $Z(x, y)$ values is nonnegative. For the remainder of the proof, assume that $t > -2\beta$. Since all $r_{x_i} \geq 0$, $\sum_{i=1}^{n} r_{x_i} \leq 2$, and $t = \sum_{i=1}^{n}\gamma_i r_{x_i}$, there must be at least one γ_i that is greater than $-\beta$. Let x_0 be an x_i such that $\gamma_0 \equiv \gamma_i > -\beta$. When $\omega = (x_0, y)$, (21) equals

$$Z(x_0, y) = \alpha_{x_0,y} - q + (\gamma_0 + \beta)y - \sum_{i=1}^{n}\gamma_i 2^{-x_i+1} - t\delta[2 + r_{x_0}2^{x_0} - u - v]. \tag{23}$$

Now, $(\gamma_0 + \beta)y$ is unbounded above as $y \to \infty$ because $\gamma_0 + \beta > 0$, and everything else in (23) is bounded. Hence (21) takes nonnegative values and the previsions that satisfy (18) and (19) are coherent. In summary, $P(Yh_x)$, $P(Y|\pi_1)$, and $P[P(Y|\pi_1)]$ are coherent if and only if there is a nonnegative sequence $\{r_x\}_{x=1}^{\infty}$ such that

$$\sum_{x=1}^{\infty} r_x \leq 2,$$
$$P(Y|\pi_1)(x, y) = 2 + r_x 2^x,$$
$$P[P(Y|\pi_1)] \geq 2 + \sum_{x=1}^{\infty} r_x.$$

We will choose examples of r_x that satisfy the above conditions while keeping $r^x 2^x$ bounded. This makes $P(Y|\pi_1)$ bounded and $\mathcal{W}$ the linear span of $\mathcal{Y} = \mathcal{X} \bigcup \{Y\}$.

For $i = 1, 2$, it follows from Theorem 1 that P is conglomerable in π_i if and only if P is disintegrable in π_i. To see whether P is disintegrable in π_i, we need to check whether (13) holds for every random variable in $\mathcal{W}$. Because P is countably additive on the indicator functions, (13) holds for each bounded X in every partition. If (13) holds for X_1 and X_2, it holds for $\alpha_1 X_1 + \alpha_2 X_2$ for all real α_1, α_2 because $P(\cdot|h)$ is linear. So, P is disintegrable in a partition if and only if (13) holds for Y.

Since $Y = P(Y|\pi_2)$, Y satisfies (13) in π_2 so that P is disintegrable in π_2 with respect to $\mathcal{W}$ and $\mathcal{W}$ is of Class 2 relative to π_2. In order for Y to satisfy (13) in π_1, we need

$$4 = P(Y) = P[P(Y|\pi_1)] = u,$$

i.e., $\sum_{x=1}^{\infty} r_x = 2$. For example, if $r_x = 3.5 \times 3^{-x}$, then $u = 3.75$ and P is not disintegrable in π_1. If $r_x = 4 \times 3^{-x}$, then P is disintegrable in both partitions.

We conclude this section with an example that involves three collections of random variables $\mathcal{X} \subset \mathcal{Y} \subset \mathcal{W}$ such that $\mathcal{X}$ and $\mathcal{W}$ satisfy the conditions of Theorem 1 while $\mathcal{Y}$ does not. Also, $\mathcal{X}$ is of Class 2, $\mathcal{Y}$ is of Class 1, and $\mathcal{W}$ is of Class 0.

Example 5 Let $\mathcal{X}, Y, \pi_1, \mathcal{Y}$, and $\mathcal{W}$ be as in Example 4. Define $r_x = P(Yh_x) - 2^{-x+1}$ as in Example 4 so that $P(Y|h_x) = 2 + r_x 2^x$. We saw in Example 4 that $r_x \geq 0$ and $\sum_{x=1}^{\infty} r_x \leq 2$ were necessary and sufficient for the stated previsions to be coherent. Let $r_x = 3.5 \times 3^{-x}$ as in one of the cases at the end of Example 4. Then $P(Y|\pi_1)$ is bounded, $\mathcal{W}$ satisfies the conditions of Theorem 1, and $P[P(Y|\pi_1)] = 3.75$. Here, $\mathcal{W}$ is of Class 0 relative to P and π_1. Because P is countably additive on $\mathcal{X}$, we know that $\mathcal{X}$ is of Class 2 relative to P and π_1. Because we have not included $Y - P(Y|\pi_1)$ in $\mathcal{Y}$, we see that $\mathcal{Y}$ does not satisfy (16). To see that $\mathcal{Y}$ is of Class 1, note that $P[P(Y|\pi_1)] = 3.75 \neq 4 = P(Y)$ while

$$\inf_{h \in \pi} P(Y|h) = 2 \leq 4 = P(Y) \leq \sup_{h \in \pi} P(Y|h) = \frac{13}{3}.$$

Hence, $\mathcal{Y}$ is of Class 1 relative to P and π_1.

6 Discussion

Conglomerability and disintegrability are familiar concepts in the countably additive theory of probability, although the names may not be as familiar as the concepts. The law of total probability or "tower property" of conditional expectations is essentially disintegrability, namely that the mean of a conditional mean is the marginal mean. With disintegrability taken for granted, conglomerability is simply an instance of the property of countably additive expectations that the mean of a random variable lies in the closed convex hull of its range. Of course, the countably additive theory guarantees disintegrability by allowing the conditional probabilities of events to change with the partition on which one conditions. The well-known Borel paradox is a classic example of how this happens. In the countably additive theory Kadane et al. (1996) illustrates how pervasive the Borel paradox is. If one insists on $P(X|h)$ having a meaning for every random variable X and every nonempty event h, then not even the countably additive theory can guarantee disintegrability in every partition.

The finitely additive theory of probability avoids the Borel paradox, but at the price of having its conditional probabilities fail conglomerability. Dubins (1975) shows that for coherent previsions over the linear space of all bounded random variables, conglomerability in a partition is equivalent to disintegrability in that same partition. In this paper we extend the concept of a coherent prevision to a collection of unbounded random variables with finite previsions and finite conditional previsions. We show that a finitely additive version of the Daniell integral gives an extension of disintegration to this collection of unbounded variables. When coherent previsions are defined for a linear space of such random variables and they contain a π-strategy, conglomerability and disintegrability of previsions in π are equivalent conditions.

As a final note, it is important to keep in mind that the concepts of conglomerability and disintegrability are defined with respect to a collection of random variables. The larger the collection of random variables, the more conditions of the form (12) and (13) that each concept requires. That is, in order for P to be conglomerable in π with respect to a collection $\mathcal{W}$, (12) must hold for every $X \in \mathcal{W}$. Similarly, for P to be disintegrable in π with respect to $\mathcal{W}$, (13) must hold for every $X \in \mathcal{W}$. Consider the collection $\mathcal{X}$ of all bounded random variables and a larger collection $\mathcal{W}$ that satisfies the conditions of Theorem 1 (such as the $\mathcal{W}$ in Example 5). Let $\mathcal{Y}$ be an intermediate collection (such as the $\mathcal{Y}$ in Example 5) so that $\mathcal{X} \subset \mathcal{Y} \subset \mathcal{W}$. If P is conglomerable in π with respect to $\mathcal{W}$, then $\mathcal{W}$ is of Class 2 relative to P and π and so are $\mathcal{Y}$ and $\mathcal{X}$. Similarly, if P is disintegrable in $\mathcal{W}$ with respect to π, then all three collections are of Class 2. However, the equivalence of conglomerability and disintegrability does not carry over from larger collections to smaller collections. The reason is that $\mathcal{W}$ might be of Class 0 while $\mathcal{Y}$ is of Class 1 and $\mathcal{X}$ is of Class 2. Indeed, this is precisely what occurs in Example 5.

References

Berti P, Regazzini E, Rigo P (2001) Strong previsions of random elements. Stat Methods Appl 10:11–28

Berti P, Rigo P (1992) Weak disintegrability as a form of preservation of coherence. J Italian Stat Soc 1:161–181

Berti P, Rigo P (2000) Integral representation of linear functionals on spaces of unbounded functions. Proc Am Math Soc 128:3251–3258

Berti P, Rigo P (2002) On coherent conditional probabilities and disintegrations. Ann Math Artif Intell 35:71–82

Crisma L, Gigante P (2001) A notion of coherent conditional prevision for arbitrary random quantities. Stat Methods Appl 10:29–40

Crisma L, Gigante P, Millossovich P (1997) A notion of coherent prevision for arbitrary random quantities. J Italian Stat Soc 6:233–243

Dawid A, Stone M, Zidek J (1973) Marginalization paradoxes in bayesian and structural inference. J Royal Stat Soc B 35:189–233

Dubins L (1975) Finitely additive conditional probabilites, conglomerability and disintegrations. Ann Probab 3:89–99

Dunford N, Schwartz J (1958) Linear operators (Part 1). Wiley, New York

de Finetti B (1930) Sulla preprieta conglomerativa della probabilita subordinate. Rend. R. Inst. Lombardo (Milano) 63:414–418

de Finetti B (1974) Theory of probability (2 vols.). Wiley, New York

Kadane J, Schervish M, Seidenfeld T (1986) Statistical implications of finitely additive probability. In: Goel P, Zellner A (eds) Bayesian inference and decision techniques. North-Holland, Amsterdam, pp 59–76

Kadane JB, Schervish MJ, Seidenfeld T (1996) Reasoning to a foregone conclusion. J Am Stat Assoc 91(435):1228–1235

Kolmogorov A (1956) Foundations of the theory of probability. Chelsea, New York

Krauss PH (1968) Representation of conditional probability measures on boolean algebras. Acta Math Sci Hungar 19:229–241

Regazzini E (1987) De finetti's coherence and statistical inference. Ann Stat 15:845–864

Royden HL (1968) Real analysis. Macmillan, London

Schervish MJ, Seidenfeld T, Kadane JB (1984) The extent of non-conglomerability of finitely additive probabilities. Z. f. Warscheinlichkeitstheorie 66:205–226

Schervish MJ, Kadane JB, Seidenfeld T (2008) Fundamental theorems of prevision and asset pricing. Int J Approx Reason 49:148–158

Schervish MJ, Seidenfeld T, Kadane JB (2014a) Dominating countably many forecasts. Ann Stat 42:728–756

Schervish MJ, Seidenfeld T, Kadane, JB (2014b) Supplement to "Dominating countably many forecasts". doi:10.1214/14-AOS1203SUPP

Seidenfeld T, Schervish MJ, Kadane JB (2009) Preference for equivalent random variables: a price for unbounded utilities. J Math Econ 45:329–340