Extension of coherent lower previsions to unbounded random variables

Matthias C. M. Troffaes, Gert de Cooman
Universiteit Gent, Onderzoeksgroep SYSTeMS,
Technologiepark – Zwijnaarde 914, 9052 Zwijnaarde, Belgium
{matthias.troffaes,gert.decooman}@rug.ac.be

Abstract

We consider the extension of coherent lower previsions from the set of bounded random variables to a larger set. An *ad hoc* method in the literature consists in approximating an unbounded random variable by a sequence of bounded ones. Its 'extended' lower prevision is then defined as the limit of the sequence of lower previsions of its approximations. We identify the random variables for which this limit does not depend on the details of the approximation, and call them previsible. We thus extend a lower prevision to previsible random variables, and we study the properties of this extension.

Keywords: imprecise probabilities, lower prevision , coherence, unbounded random variable, Dunford integral

1 Introduction

When modelling a system, it often occurs that we do not know all of its aspects, or that we wish to discard certain details in order to simplify the system analysis. These situations give rise to uncertainty, and consequently, we are challenged to find an appropriate system description which takes the uncertainty into account. One particularly successful way of doing so, consists in using a probabilistic description of the uncertainty. Despite its many successes, however, there are quite a number of situations in which this method does not lead to sensible results, simply because there may not be sufficient information available to allow us to select a *single* probability distribution as an appropriate model. There are in the literature a number of uncertainty models that do not assume the uncertainty to be described by a single probability distribution. Among these, Walley's behavioural theory of imprecise probabilities [12] stands out as a very satisfactory choice, certainly from the foundational point of view. It has a clear behavioural interpretation, which leads naturally to a theory of decision making under uncertainty. Moreover, it unifies a large number of other uncertainty models, such as 2-monotone capacities [2], possibility measures [3, 13], comparative and modal probabilities [11], and convex sets of probability measures [8].

One important shortcoming of the existing theory of coherent lower previsions is that it only deals with random variables that are bounded, whereas in engineering, for instance, applications involving unbounded random variables abound. To give only a few examples, the following classes of problems would certainly benefit from an extension of imprecise probability theory able to deal with unbounded random variables: (i) the estimation of unbounded quantities, such as the time to failure of a component in a system [9]; and (ii) optimisation involving an unbounded cost [4]. An intuitive, *ad hoc* way of dealing with an unbounded random variable is to approximate it by a sequence of bounded ones, and to use limit arguments in order to extend notions defined in the context of the bounded random variables to their unbounded counterparts, in the hope that the eventual result will not depend on the exact form of the approximation. Similar types of construction exist in the theory of integration—we shall use them as a source of inspiration.

Our main objectives in this paper are twofold: (i) to construct an extension of coherent lower previsions from bounded random variables to a larger set; and (ii) to study of the properties of this extension in order to motivate that its result can be seen as a coherent lower prevision in its own right. The paper is organised as follows. We give a brief introduction to the theory of imprecise probabilities in Section 2. In Section 3 we introduce the basic concepts of our extension. Important properties are listed in Section 4. Finally, in Section 5 we show that for linear previsions, there is a Dunford integral representation for their extension to unbounded random variables.

2 Imprecise probabilities

We start with a brief introduction to the most important aspects of the existing behavioural theory of imprecise probabilities that are relevant to the problem at hand. More details can be found in [12]. Consider an agent who is uncertain about something, say, the outcome of some experiment. If the set of possible outcomes is Ω, then a *random variable* is a mapping from Ω to $\mathbb{R}$, and it is interpreted as an uncertain reward: if $\omega \in \Omega$ turns out to be the actual outcome of the experiment then the agent receives the amount $X(\omega)$, expressed in units of some linear utility. Bounded random variables are also called *gambles*. They play a very important part in the existing theory. Denote the set of all gambles by $\mathscr{L}(\Omega)$.

The information the agent has about the outcome of the experiment will lead him to accept or reject transactions whose reward depends on this outcome. We can formulate a model for his uncertainty by looking at a specific type of transaction: buying gambles. The agent's *lower prevision* (or supremum acceptable buying price) $\underline{P}(X)$ for a gamble X is the highest price s such that he is disposed to buy the gamble X for any price strictly lower than s. If the agent assesses a supremum acceptable buying price for every gamble X in a subset $\mathcal{K}$ of $\mathscr{L}(\Omega)$, the resulting mapping $\underline{P} : \mathcal{K} \to \mathbb{R}$ is called a *lower prevision*.

It can be argued that $\underline{P}$ must satisfy the following rationality constraint: for all $n \in \mathbb{N}$, all $\lambda_0, \dots, \lambda_n \geq 0$, and all $X_0, \dots, X_n \in \mathcal{K}$ we must have that[1]

$$\sup\left[\sum_{i=1}^{n} \lambda_i X_i - \lambda_0 X_0\right] \geq \sum_{i=1}^{n} \lambda_i \underline{P}(X_i) - \lambda_0 \underline{P}(X_0).$$

Here and elsewhere, we denote by $\sup[X]$ the supremum value $\sup_{\omega\in\Omega} X(\omega)$ of the gamble X (and similarly for $\inf[X]$). If the lower prevision $\underline{P}$ satisfies this constraint, we say that it is *coherent*. If $\mathcal{K}$ is a linear space, e.g., when $\mathcal{K} = \mathscr{L}(\Omega)$, then $\underline{P}$ is coherent if and only if

[1]For example, take $n = 0$ and $\lambda_0 = 1$, then we find that $\underline{P}(X) \geq \inf[X]$, which means that the agent should be willing to pay at least the lowest possible reward.

(i) $\underline{P}(X) \geq \inf[X]$,

(ii) $\underline{P}(\lambda X) = \lambda \underline{P}(X)$, and

(iii) $\underline{P}(X + Y) \geq \underline{P}(X) + \underline{P}(Y)$.

for all gambles X, Y in $\mathcal{K}$ and $\lambda \geq 0$. This result can be given a simple and natural interpretation: the supremum buying prices should accept a sure gain, they should be independent of the utility scale, and finally, if we are willing to buy X for price s and Y for price t, then we should also be willing to buy $X + Y$ for price $s + t$. The following consequences of coherence will also be used in proofs further on (see [12] for details).

(i) $\overline{P}(X + Y) \leq \overline{P}(X) + \overline{P}(Y)$

(ii) $X \leq Y \implies \underline{P}(X) \leq \underline{P}(Y)$ and $\overline{P}(X) \leq \overline{P}(Y)$

(iii) $|\underline{P}(X) - \underline{P}(Y)| \leq \overline{P}(|X - Y|)$ and $\left|\overline{P}(X) - \overline{P}(Y)\right| \leq \overline{P}(|X - Y|)$

It can be shown that if $\underline{P}$ is coherent, there is always a (unique) smallest coherent extension of $\underline{P}$ from its domain $\mathcal{K}$ to $\mathscr{L}(\Omega)$. We call this extension the *natural extension* of $\underline{P}$. It is given by

$$\underline{E}(X) = \sup\left\{\inf\left[X - \sum_{i=1}^{n} \lambda_i \left[X_i - \underline{P}(X_i)\right]\right]\right\},$$

where $X \in \mathscr{L}(\Omega)$ and the supremum runs over $n \in \mathbb{N}$, $\lambda_1, \ldots, \lambda_n \geq 0$ and $X_1, \ldots, X_n$ in $\mathcal{K}$. This shows that without loss of generality, we may from now on assume that $\underline{P}$ is a coherent lower prevision defined on all of $\mathscr{L}(\Omega)$.

$\overline{P}$ denotes the conjugate *upper prevision* of $\underline{P}$, and is defined by $\overline{P}(X) = -\underline{P}(-X)$ for every $X \in \mathscr{L}(\Omega)$. $\overline{P}(X)$ represents the agent's infimum acceptable selling price for X. The difference $\overline{P}(X) - \underline{P}(X)$ measures the amount of imprecision in the agent's behavioural dispositions toward the gamble X. An *event* A is a subset of Ω. It will be identified with its indicator I_A, which is a gamble.[2] The *lower probability* $\underline{P}(A)$ is then defined as the lower prevision $\underline{P}(I_A)$ of its indicator I_A, and similarly for the upper probability $\overline{P}(A)$.

If it so happens that $\overline{P}(X) = \underline{P}(X)$ for every gamble X, then $\underline{P}$ is called a *linear prevision*, and it is denoted by P. Linear previsions are linear functionals on the linear space $\mathscr{L}(\Omega)$ that are positive and have unit norm ($P(I_\Omega) = 1$). They are the *fair prices* or *previsions* in the sense of de Finetti [6, 7]. The restriction of a linear prevision P to events is a finitely additive probability (also called a probability charge [1]) and $P(X)$ is equal to the expected value of the bounded random variable X with respect to this probability charge (see for instance [1, Theorem 4.7.4]). In this way, any Bayesian model can be considered to be a linear prevision, which is a special kind of lower prevision. The set of all linear previsions on $\mathscr{L}(\Omega)$ is denoted by $\mathscr{P}(\Omega)$.

$\mathcal{M}(\underline{P})$ will denote the set of all linear previsions that dominate $\underline{P}$ point-wise on $\mathscr{L}(\Omega)$: $\mathcal{M}(\underline{P}) = \{Q \in \mathscr{P}(\Omega) \colon Q \geq \underline{P}\}$. One can show that $\mathcal{M}(\underline{P})$ is a non-empty,

[2] I_A is the random variable that takes the value 1 on A and 0 elsewhere.

convex and compact[3] subset of $\mathscr{P}(\Omega)$, and that $\underline{P}$ is the lower envelope of $\mathcal{M}(\underline{P})$, that is,

$$\underline{P}(X) = \min_{Q \in \mathcal{M}(\underline{P})} Q(X),$$

for all $X \in \mathscr{L}(\Omega)$. This equality, and the fact that the lower envelope of any non-empty set of linear previsions is a coherent lower prevision, gives rise to what is called the *Bayesian sensitivity analysis interpretation*, or *Quasi-Bayesian interpretation* of lower previsions: specifying a coherent lower prevision is formally equivalent to specifying a non-empty, convex and compact set of linear previsions (or probability charges).

3 Previsibility

In the previous section, we have seen that within the existing framework it is possible to extend any given coherent lower prevision to the set of all bounded random variables in a natural way. We now investigate whether it can be extended still further to a larger set that includes some unbounded random variables. The first step of our investigation will be the construction of a limit procedure—approximating unbounded random variables by bounded ones—taking the necessary care to ensure that the procedure yields a unique result: we do not want our result to depend on the details of the approximation.

We begin therefore by defining the *$\underline{P}$-norm* of a gamble X by $\|X\|_{\underline{P}} = \overline{P}(|X|)$. Using the coherence of $\underline{P}$, it can be shown that $\|\cdot\|_{\underline{P}}$ is a semi-norm on $\mathscr{L}(\Omega)$. A sequence (X_n) of gambles is called *$\underline{P}$-fundamental*[4] if it is Cauchy with respect to $\|\cdot\|_{\underline{P}}$, i.e., if $\|X_n - X_m\|_{\underline{P}} \to 0$ as $n, m \to \infty$. We say that a sequence of gambles (X_n) converges *$\underline{P}$-hazily*[5] to the random variable X if for every $\epsilon > 0$ we have that

$$\lim_{n\to\infty} \overline{P}(\{\omega \in \Omega\colon\ |X(\omega) - X_n(\omega)| > \epsilon\}) = 0.$$

Observe that we do not need to impose any measurability conditions, since $\overline{P}(A)$ is defined for every $A \subseteq \Omega$. The following lemma is the basic result that will guarantee the unicity of the extension introduced in Definition 1.

Lemma 1. *If (X_n) and (Y_n) are $\underline{P}$-fundamental sequences of gambles converging $\underline{P}$-hazily to the same random variable Z, then it holds that the limits $\lim_{n\to\infty} \overline{P}(X_n)$ and $\lim_{n\to\infty} \overline{P}(Y_n)$ exist, are finite real numbers and coincide, and similarily, the limits $\lim_{n\to\infty} \underline{P}(X_n)$ and $\lim_{n\to\infty} \underline{P}(Y_n)$ exist, are finite real numbers and coincide.*

Proof. We first prove that the limits exist and are finite. This follows from the following inequalities, which are consequences of the coherence of $\underline{P}$,

$$\left|\overline{P}(X_n) - \overline{P}(X_m)\right| \le \overline{P}(|X_n - X_m|), \qquad |\underline{P}(X_n) - \underline{P}(X_m)| \le \overline{P}(|X_n - X_m|),$$
$$\left|\overline{P}(Y_n) - \overline{P}(Y_m)\right| \le \overline{P}(|Y_n - Y_m|), \qquad |\underline{P}(Y_n) - \underline{P}(Y_m)| \le \overline{P}(|Y_n - Y_m|).$$

[3] We assume in this paper that $\mathscr{P}(\Omega)$ is provided with its topology of point-wise convergence: the relativisation to $\mathscr{P}(\Omega)$ of the weak*-topology on the topological dual $\mathscr{L}(\Omega)^*$, where $\mathscr{L}(\Omega)$ is provided with the supremum norm topology.

[4] Cf. *mean fundamental* in the theory of measures.

[5] Cf. *convergence in measure* in measure theory, and *hazy convergence* [1] in the theory of finitely additive measures.

Since the right hand sides converge to zero, the left hand sides must converge to zero too. This means that $\overline{P}(X_n)$, $\underline{P}(X_n)$, $\overline{P}(Y_n)$ and $\underline{P}(Y_n)$ are Cauchy sequences. By the completeness of $\mathbb{R}$, their limits exist and are finite real numbers.

Next, we prove that $\lim_{n\to\infty}\overline{P}(X_n) = \lim_{n\to\infty}\overline{P}(Y_n)$ and $\lim_{n\to\infty}\underline{P}(X_n) = \lim_{n\to\infty}\underline{P}(Y_n)$. Let $N_n := |X_n - Y_n|$. Again by the coherence of $\underline{P}$, we have that $|\underline{P}(X_n) - \underline{P}(Y_n)| \leq \overline{P}(N_n)$ and $\left|\overline{P}(X_n) - \overline{P}(Y_n)\right| \leq \overline{P}(N_n)$. The proof is complete if we can show that $\overline{P}(N_n)$ converges to zero. For every $n \in \mathbb{N}$ and every $A \subseteq \Omega$, define $a_n(A) := \overline{P}(N_n A)$. We must prove that $\lim_{n\to\infty} a_n(\Omega) = 0$.

Every a_n is an element of the function space $\mathbb{R}^{\wp(\Omega)}$. Equip this space with the topology of uniform convergence on $\wp(\Omega)$. Note that by the completeness of $\mathbb{R}$, it follows by that $\mathbb{R}^{\wp(\Omega)}$ is complete with respect to the topology of uniform convergence on $\wp(\Omega)$ (see for instance [10, Section 19.12]). We first claim that a_n converges with respect to the topology of uniform convergence on $\wp(\Omega)$. Indeed, consider $A \subseteq \Omega$, then, using the coherence of $\underline{P}$, we find that

$$\begin{aligned}
|a_n(A) - a_m(A)| &= \left|\overline{P}\big(|X_n - Y_n|\, A\big) - \overline{P}\big(|X_m - Y_m|\, A\big)\right| \\
&\leq \overline{P}\Big(\big||X_n - Y_n| - |X_m - Y_m|\big|A\Big) \\
&\leq \overline{P}\Big(\big|(X_n - Y_n) - (X_m - Y_m)\big|A\Big) \\
&\leq \overline{P}\Big(\big|(X_n - X_m) - (Y_n - Y_m)\big|\Big) \\
&\leq \overline{P}\big(|X_n - X_m|\big) + \overline{P}\big(|Y_n - Y_m|\big).
\end{aligned}$$

Since the right hand side converges to zero independently of A, it follows that a_n is Cauchy with respect to the topology of uniform convergence on $\wp(\Omega)$. By the completeness of $\mathbb{R}^{\wp(\Omega)}$ with respect to the topology of uniform convergence on $\wp(\Omega)$, we find that a_n converges with respect to the topology of uniform convergence on $\wp(\Omega)$.

Uniform convergence implies point-wise convergence, so for every $A \subseteq \Omega$ we can define $a(A) := \lim_{n\to\infty} a_n(A)$. We must prove that $a(\Omega) = 0$. Let $\epsilon > 0$. By the convergence of a_n with respect to the topology of uniform convergence on $\wp(\Omega)$, there is an $M_\epsilon \in \mathbb{N}$ such that for all $A \subseteq \Omega$ and all $n \geq M_\epsilon$

$$|a_n(A) - a(A)| < \epsilon. \tag{1}$$

Define

$$\delta_\epsilon := \begin{cases} \epsilon / \sup N_{M_\epsilon}, & \text{if } \sup N_{M_\epsilon} > 0, \\ 1, & \text{otherwise.} \end{cases}$$

For every $A \subseteq \Omega$, if $\overline{P}(A) < \delta_\epsilon$ then $a_{M_\epsilon}(A) = \overline{P}(N_{M_\epsilon} A) \leq \sup N_{M_\epsilon} \overline{P}(A) < \epsilon$. Since $a(A) \leq |a(A) - a_{M_\epsilon}(A)| + a_{M_\epsilon}(A)$, it holds by (1) that

$$\overline{P}(A) < \delta_\epsilon \implies a(A) < 2\epsilon. \tag{2}$$

Define $B := \{\omega \in \Omega; N_{M_\epsilon}(\omega) \neq 0\}$, then $N_{M_\epsilon} \complement B = 0$. We infer that $\overline{P}(N_{M_\epsilon} \complement B) = a_{M_\epsilon}(\complement B) = 0$. From (1) it follows that $a(\complement B) < \epsilon$. We now prove that $a(\Omega) < 5\epsilon$.

(a) Consider the case that $\overline{P}(B) = 0$. Then $a(B) = \lim_{n\to\infty} \overline{P}(N_n B) = 0$ since $0 \le \overline{P}(N_n B) \le \sup N_n \overline{P}(B) = 0$ for every $n \in \mathbb{N}$. By the coherence of $\underline{P}$ it follows that $a(\Omega) \le a(B) + a(\complement B) < 0 + \epsilon < 5\epsilon$.

(b) Now consider the other case that $\overline{P}(B) > 0$. Since X_n and Y_n converge to Z $\underline{P}$-hazily, it follows easily from coherence of $\underline{P}$ that $N_n = |X_n - Y_n|$ converges $\underline{P}$-hazily to 0. This implies that there is a $K_\epsilon \ge M_\epsilon$ such that for all $n \ge K_\epsilon$

$$\overline{P}(\{\omega \in \Omega; N_n > \epsilon/\overline{P}(B)\}) < \delta_\epsilon. \tag{3}$$

Define $C := \{\omega \in \Omega; N_{K_\epsilon} \le \epsilon/\overline{P}(B)\}$. By the coherence of $\underline{P}$ we have that $a(\Omega) \le a(B \cap C) + a(B \cap \complement C) + a(\complement B)$. We now investigate each term.

(i) By (1) we have that $a(B \cap C) < a_{K_\epsilon}(B \cap C) + \epsilon$, since $K_\epsilon \ge M_\epsilon$. Since $N_{K_\epsilon}(\omega) \le \epsilon/\overline{P}(B)$ for all $\omega \in C$ and $\overline{P}(B \cap C) \le \overline{P}(B)$, we have that $a_{K_\epsilon}(B \cap C) = \overline{P}(N_{K_\epsilon}[B \cap C]) \le (\epsilon/\overline{P}(B))\overline{P}(B \cap C) \le \epsilon$. We find that $a(B \cap C) < 2\epsilon$.

(ii) We claim that $a(B \cap \complement C) < 2\epsilon$. By (3) it follows that $\overline{P}(\complement C) < \delta_\epsilon$. The claim is established using $\overline{P}(B \cap \complement C) \le \overline{P}(\complement C)$ and (2).

(iii) We already proved that $a(\complement B) < \epsilon$.

In both cases it follows that $a(\Omega) < 5\epsilon$. Since this holds for any $\epsilon > 0$, we find that $a(\Omega) = 0$. □

Observe that the proof given here uses the same techniques as its counterpart in the theory of charges (see for instance the proof of Proposition 4.4.10 in [1]).

Definition 1. A random variable Z is said to be *$\underline{P}$-previsible* if there is a $\underline{P}$-fundamental sequence (X_n) of gambles that converges $\underline{P}$-hazily to Z. We then define $\underline{P}^{\mathsf{x}}(Z) = \lim_{n\to\infty} \underline{P}(X_n)$, and (X_n) is called a *$\underline{P}$-determining sequence* for Z, or simply a *determining sequence* if there is no ambiguity regarding $\underline{P}$.

By Lemma 1, the limit $\underline{P}^{\mathsf{x}}(Z)$ is a finite real number, and is independent of the details of the determining sequence (X_n). Moreover, $\underline{P}^{\mathsf{x}}$ extends $\underline{P}$ in the mathematical sense. This follows simply from the observation that the constant sequence (X_n) defined by $X_n = X$ for each n is a determining sequence for X, whenever X is a gamble.

Proposition 1. *$\underline{P}^{\mathsf{x}}(X) = \underline{P}(X)$ for every $X \in \mathscr{L}(\Omega)$.*

The set of all $\underline{P}$-previsible random variables will be denoted by $\mathscr{L}^{\mathsf{x}}_{\underline{P}}(\Omega)$. By Proposition 1 it contains all gambles (bounded random variables). Using the coherence of $\underline{P}$, the following properties of $\mathscr{L}^{\mathsf{x}}_{\underline{P}}(\Omega)$ can be easily established (denoting the point-wise maximum by $\vee$ and the point-wise minimum by $\wedge$).

Proposition 2. *Let X and $Y \in \mathscr{L}^{\mathsf{x}}_{\underline{P}}(\Omega)$, and $a \in \mathbb{R}$. Then $X + Y$, aX, $X \vee Y$, $X \wedge Y$, and $|X| \in \mathscr{L}^{\mathsf{x}}_{\underline{P}}(\Omega)$.*

This means that $\mathscr{L}^{\mathsf{x}}_{\underline{P}}(\Omega)$ is a linear lattice with respect to the point-wise order. In particular, $|Z|$ is $\underline{P}$-previsible if Z is, and therefore we can extend the semi-norm $\|\cdot\|_{\underline{P}}$ introduced above to $\mathscr{L}^{\mathsf{x}}_{\underline{P}}(\Omega)$ through $\|Z\|_{\underline{P}} = \overline{P}^{\mathsf{x}}(|Z|)$ for all Z in $\mathscr{L}^{\mathsf{x}}_{\underline{P}}(\Omega)$. It is not difficult to show that $\|\cdot\|_{\underline{P}}$ is also a semi-norm on $\mathscr{L}^{\mathsf{x}}_{\underline{P}}(\Omega)$, using Proposition 3 below. Moreover, if (Z_n) is a sequence of $\underline{P}$-previsible random variables and $\lim_{n\to\infty} \|Z - Z_n\|_{\underline{P}} = 0$ then $\lim_{n\to\infty} \underline{P}^{\mathsf{x}}(Z_n) = \underline{P}^{\mathsf{x}}(Z)$. This shows that topologically indistinguishable random variables are also behaviourally indistinguishable, that is, they have the same extended lower (and upper) prevision.

4 Properties

4.1 Coherence

It turns out that all of the properties of coherent lower previsions listed in [12, Section 2.6.1] extend to the extension $\underline{P}^{\mathsf{x}}$. We mention the three most important ones, which also establish "coherence" of the extension: the supremum buying prices must accept sure gain, they must be independent of the utility scale, and, if we are willing to buy X for price s and Y for price t, then we should certainly be willing to buy $X+Y$ for price $s+t$.

Proposition 3. *Let X and Y in $\mathscr{L}^{\mathsf{x}}_{\underline{P}}(\Omega)$, and let $\lambda \geq 0$. It holds that*

(i) $\underline{P}^{\mathsf{x}}(X) \geq \inf[X]$

(ii) $\underline{P}^{\mathsf{x}}(\lambda X) = \lambda \underline{P}^{\mathsf{x}}(X)$

(iii) $\underline{P}^{\mathsf{x}}(X+Y) \geq \underline{P}^{\mathsf{x}}(X) + \underline{P}^{\mathsf{x}}(Y)$

Proof. The last two properties follow immediately from the coherence of $\underline{P}$ and the fact that (in)equalities are preserved when taking limits whenever both sides of the (in)equality converge. To prove the first property, let (X_n) be a determining sequence for X. Define for every $n \in \mathbb{N}$

$$X'_n(\omega) = \begin{cases} X_n(\omega), & \text{if } X_n(\omega) \geq \inf[X], \\ \inf[X], & \text{otherwise.} \end{cases}$$

Since $|X - X'_n| \leq |X - X_n|$ and $|X'_n - X'_m| \leq |X_n - X_m|$ for every n, $m \in \mathbb{N}$, it follows easily from the coherence of $\underline{P}$ that (X'_n) is also a determining sequence for X. Observe that $X'_n \geq \inf[X]$. From the coherence of $\underline{P}$ we know that $\underline{P}(X'_n) \geq \inf[X]$. Now rely on the fact that this inequality is preserved when taking the limit to conclude that $\underline{P}^{\mathsf{x}}(X) = \lim_{n\to\infty} \underline{P}(X'_n) \geq \inf[X]$. □

4.2 Increasing domain under increasing precision

It turns out that as the precision of a coherent lower prevision increases, its extension will also become more precise, and more random variables become previsible. The proof of this fact is left to the reader as a simple exercise.

Proposition 4. *If $\underline{Q}$ point-wise dominates $\underline{P}$, then $\mathscr{L}^{\mathsf{x}}_{\underline{Q}}(\Omega) \supseteq \mathscr{L}^{\mathsf{x}}_{\underline{P}}(\Omega)$, and $\underline{Q}^{\mathsf{x}}$ point-wise dominates $\underline{P}^{\mathsf{x}}$ on $\mathscr{L}^{\mathsf{x}}_{\underline{P}}(\Omega)$, i.e., $\underline{Q}^{\mathsf{x}}(X) \geq \underline{P}^{\mathsf{x}}(X)$ for all $X \in \mathscr{L}^{\mathsf{x}}_{\underline{P}}(\Omega)$.*

As an example, we consider the case in which the lower prevision $\underline{P}$ describes complete ignorance. In such a case, the supremum price we are willing to buy a gamble X for is given by the lowest possible reward we may expect from X, that is, $\inf_{\omega\in\Omega} X(\omega)$. This lower prevision is called the *vacuous lower prevision*, and we denote it by $\underline{P}_v$. The $\underline{P}_v$-norm is the supremum norm, and $\mathscr{L}^{\mathsf{x}}_{\underline{P}_v}(\Omega) = \mathscr{L}(\Omega)$. Thus the set of all vacuously previsible random variables is exactly the set of bounded random variables.

4.3 Weak*-compactness and a lower envelope theorem

Let $\mathscr{P}_{\underline{P}^{\mathsf{x}}}(\Omega)$ denote the set of all real-valued linear functionals on the linear space $\mathscr{L}_{\underline{P}}^{\mathsf{x}}(\Omega)$ that dominate $\underline{P}^{\mathsf{x}}$ point-wise. These linear functionals have all the properties of a linear prevision—they are linear and positive, and have unit norm.

Theorem 1. *$\mathscr{P}_{\underline{P}^{\mathsf{x}}}(\Omega)$ is weak*-compact.*

Proof. $\mathscr{P}_{\underline{P}^{\mathsf{x}}}(\Omega)$ is a subset of the topological dual $\mathscr{L}_{\underline{P}}^{\mathsf{x}}(\Omega)^*$. The weak*-topology on $\mathscr{L}_{\underline{P}}^{\mathsf{x}}(\Omega)^*$ is the topology of point-wise convergence, and the theorem states that $\mathscr{P}_{\underline{P}^{\mathsf{x}}}(\Omega)$ is compact as a subset of $\mathscr{L}_{\underline{P}}^{\mathsf{x}}(\Omega)^*$ with respect to this topology. Define the set

$$\mathscr{V} = \prod_{X \in \mathscr{L}_{\underline{P}}^{\mathsf{x}}(\Omega)} [\underline{P}^{\mathsf{x}}(X), \overline{P}^{\mathsf{x}}(X)].$$

Members of $\mathscr{V}$ are mappings f on $\mathscr{L}_{\underline{P}}^{\mathsf{x}}(\Omega)$ satisfying $f(X) \in [\underline{P}^{\mathsf{x}}(X), \overline{P}^{\mathsf{x}}(X)]$ for each $X \in \mathscr{L}_{\underline{P}}^{\mathsf{x}}(\Omega)$. In particular, $\mathscr{P}_{\underline{P}^{\mathsf{x}}}(\Omega)$ is a subset of $\mathscr{V}$. We equip $\mathscr{V}$ with the product topology, which is the topology of point-wise convergence.

We see that the relativisation of the weak*-topology on $\mathscr{L}_{\underline{P}}^{\mathsf{x}}(\Omega)^*$ to $\mathscr{P}_{\underline{P}^{\mathsf{x}}}(\Omega)$ is equal to the relativisation of the product topology on $\mathscr{V}$ to $\mathscr{P}_{\underline{P}^{\mathsf{x}}}(\Omega)$, since they are both topologies of point-wise convergence. Since compactness of a set is only determined by the relativisation of the topology to that set, the theorem is established if we can show that $\mathscr{P}_{\underline{P}^{\mathsf{x}}}(\Omega)$ is a compact subset of $\mathscr{V}$ with respect to the product topology on $\mathscr{V}$.

By the ultrafilter principle (i.e., Tychonov's theorem for Hausdorff spaces, see for instance [10, Section 17.22]), $\mathscr{V}$ is compact. Hence, we only need to show that $\mathscr{P}_{\underline{P}^{\mathsf{x}}}(\Omega)$ is a closed subset of $\mathscr{V}$, since any closed subset of a compact space is also compact. By its definition, $\mathscr{P}_{\underline{P}^{\mathsf{x}}}(\Omega)$ is the set of linear mappings in $\mathscr{V}$. Assume that (T_α) is a net in $\mathscr{P}_{\underline{P}^{\mathsf{x}}}(\Omega)$ that converges to $T \in \mathscr{V}$ with respect to the product topology in $\mathscr{V}$. Then T is linear. Indeed, by the point-wise convergence of (T_α) to T we have that

$$T(\lambda X) = \lim_\alpha T_\alpha(\lambda X) = \lim_\alpha \lambda T_\alpha(X) = \lambda \lim_\alpha T_\alpha(X) = \lambda T(X),$$

for any $X \in \mathscr{L}_{\underline{P}}^{\mathsf{x}}(\Omega)$ and any $\lambda \in \mathbb{R}$, and

$$\begin{aligned} T(X+Y) = \lim_\alpha T_\alpha(X+Y) &= \lim_\alpha [T_\alpha(X) + T_\alpha(Y)] \\ &= \lim_\alpha T_\alpha(X) + \lim_\alpha T_\alpha(Y) = T(X) + T(Y), \end{aligned}$$

for any $X, Y \in \mathscr{L}_{\underline{P}}^{\mathsf{x}}(\Omega)$. This shows that $T \in \mathscr{P}_{\underline{P}^{\mathsf{x}}}(\Omega)$, so $\mathscr{P}_{\underline{P}^{\mathsf{x}}}(\Omega)$ must be closed with respect to the topology of point-wise convergence. □

It is not so difficult to establish the following, quite remarkable result (where we denote by $f|_A$ the restriction of a mapping f to the subset A of its domain).

Theorem 2. *There is a canonical one-to-one correspondence between the sets $\mathcal{M}(\underline{P})$ and $\mathscr{P}_{\underline{P}^{\mathsf{x}}}(\Omega)$, given by*

$$\Phi_{\underline{P}}^{\mathsf{x}}(Q) = Q^{\mathsf{x}}|_{\mathscr{L}_{\underline{P}}^{\mathsf{x}}(\Omega)}, \qquad \Psi_{\underline{P}}^{\mathsf{x}}(R) = R|_{\mathscr{L}(\Omega)},$$

for any $Q \in \mathcal{M}(\underline{P})$ and any $R \in \mathscr{P}_{\underline{P}^{\mathsf{x}}}(\Omega)$.

Proof. It is left as an easy exercise to the reader to first verify that $Q^{\mathsf{x}}|_{\mathscr{L}^{\mathsf{x}}_{\underline{P}}(\Omega)} \in \mathscr{P}_{\underline{P}^{\mathsf{x}}}(\Omega)$ and $R|_{\mathscr{L}(\Omega)} \in \mathcal{M}(\underline{P})$ for any $Q \in \mathcal{M}(\underline{P})$ and any $R \in \mathscr{P}_{\underline{P}^{\mathsf{x}}}(\Omega)$. We still need to show that the mappings $\Phi^{\mathsf{x}}_{\underline{P}}$ and $\Psi^{\mathsf{x}}_{\underline{P}}$ are each other's inverses, that is, $\Phi^{\mathsf{x}}_{\underline{P}} \circ \Psi^{\mathsf{x}}_{\underline{P}} = 1_{\mathcal{M}(\underline{P})}$ and $\Psi^{\mathsf{x}}_{\underline{P}} \circ \Phi^{\mathsf{x}}_{\underline{P}} = 1_{\mathscr{P}_{\underline{P}^{\mathsf{x}}}(\Omega)}$.

Let $Q \in \mathcal{M}(\underline{P})$. Then, since Q^{x} is an extension of Q, it follows immediately that $\Psi^{\mathsf{x}}_{\underline{P}} \circ \Phi^{\mathsf{x}}_{\underline{P}}(Q) = \Phi^{\mathsf{x}}_{\underline{P}}(Q)|_{\mathscr{L}(\Omega)} = Q^{\mathsf{x}}|_{\mathscr{L}(\Omega)} = Q$. Conversely, let $R \in \mathscr{P}_{\underline{P}^{\mathsf{x}}}(\Omega)$. We prove that $\Phi^{\mathsf{x}}_{\underline{P}} \circ \Psi^{\mathsf{x}}_{\underline{P}}(R) = R$. Define $Q := \Psi^{\mathsf{x}}_{\underline{P}}(R) = R|_{\mathscr{L}(\Omega)}$. It is easily checked that Q is linear, positive, and has unit norm. It follows that Q is a linear prevision. Moreover, Q dominates $\underline{P}$, and hence, $Q \in \mathcal{M}(\underline{P})$. Let X be a $\underline{P}$-previsible random variable and let (X_n) be a $\underline{P}$-determining sequence for X. Then it is easily established that (X_n) also is a Q-determining sequence for X, using the fact that Q dominates $\underline{P}$. Hence,

$$\Phi^{\mathsf{x}}_{\underline{P}}(Q)(X) = Q^{\mathsf{x}}(X) = \lim_{n\to\infty} Q(X_n) = \lim_{n\to\infty} R(X_n).$$

Observe that

$$|R(X) - R(X_n)| = |R(X - X_n)| \leq R(|X - X_n|) \leq \overline{P}^{\mathsf{x}}(|X - X_n|).$$

But, since $|X_m - X_n|$ is a $\underline{P}$-determining sequence for $|X - X_n|$, it must hold that $\overline{P}^{\mathsf{x}}(|X - X_n|) = \lim_{m\to\infty} \overline{P}^{\mathsf{x}}(|X_m - X_n|)$. Hence,

$$\lim_{n\to\infty} |R(X) - R(X_n)| \leq \lim_{n,m\to\infty} \overline{P}^{\mathsf{x}}(|X_m - X_n|) = 0,$$

from which we infer that $\lim_{n\to\infty} R(X_n) = R(X)$. Thus indeed $\Phi^{\mathsf{x}}_{\underline{P}}(Q)(X) = R(X)$, for any $\underline{P}$-previsible random variable X. □

Theorem 2 leads to a number of interesting observations. Since $\mathcal{M}(\underline{P})$ is non-empty, so is $\mathscr{P}_{\underline{P}^{\mathsf{x}}}(\Omega)$. Since $\Phi^{\mathsf{x}}_{\underline{P}}$ preserves the convex structure, and $\mathcal{M}(\underline{P})$ is convex, we infer that also $\mathscr{P}_{\underline{P}^{\mathsf{x}}}(\Omega)$ is convex. Finally, Theorem 2 also leads to a very interesting lower envelope theorem for our extension, which shows that the Bayesian sensitivity analysis interpretation still holds for our extension, *in terms of linear previsions on gambles only!*

Theorem 3. *For any $\underline{P}$-previsible random variable X we have that*

$$\underline{P}^{\mathsf{x}}(X) = \min_{Q\in\mathcal{M}(\underline{P})} Q^{\mathsf{x}}(X).$$

Proof. Since $\underline{P}^{\mathsf{x}}$ is concave, and continuous with respect to the $\underline{P}$-norm, it follows from the Hahn-Banach theorem [10, HB17] that $\underline{P}^{\mathsf{x}}$ is the lower envelope of the continuous affine functions that point-wise dominate $\underline{P}^{\mathsf{x}}(X)$.

Let $T + a$ be any affine function on $\mathscr{L}^{\mathsf{x}}_{\underline{P}}(\Omega)$ which point-wise dominates $\underline{P}^{\mathsf{x}}$, that is, T linear, $a \in \mathbb{R}$ and $T + a \geq \underline{P}^{\mathsf{x}}$ (we do not assume that $T + a$ is continuous). Then we infer that $T(1) = 1$ and $a \geq 0$, since the inequality $\underline{P}(\alpha) \leq T(\alpha) + a$, which is equivalent to $a \geq \alpha(1 - T(1))$, must be satisfied for all $\alpha \in \mathbb{R}$. Secondly, we infer that T is positive. Indeed, if $X \geq 0$, then $\lambda X \geq 0$ for all $\lambda \geq 0$, and hence, $0 \leq \underline{P}(\lambda X) \leq T(\lambda X) + a$ for all $\lambda \geq 0$, or equivalently, $T(X) \geq -\frac{a}{\lambda}$ for all $\lambda \geq 0$, which can only be satisfied if $T(X) \geq 0$ (recall that we just proved that $a \geq 0$). Finally, T is continuous, since

$|T(X) - T(Y)| \leq \overline{P}^{\mathbf{x}}(|X - Y|)$ for all $X, Y \in \mathscr{L}^{\mathbf{x}}_{\underline{P}}(\Omega)$. Indeed, for all $\lambda > 0$

$$|T(X) - T(Y)| = |T(X - Y)| \leq T(|X - Y|) = \frac{1}{\lambda}T(\lambda\,|X - Y|) \\ \leq \frac{1}{\lambda}[\overline{P}^{\mathbf{x}}(\lambda\,|X - Y|) + a] = \overline{P}^{\mathbf{x}}(|X - Y|) + \frac{a}{\lambda},$$

which gives the desired inequality. So the affine functions that point-wise dominate $\underline{P}^{\mathbf{x}}$ are continuous, and are exactly of the form $T + a$, with $T \in \mathscr{P}_{\underline{P}^{\mathbf{x}}}(\Omega)$ and $a \in \mathbb{R}$, $a \geq 0$.

This shows that in taking the lower envelope, we may restrict ourselves to the (continuous) *linear* functions (i.e., we may assume that $a = 0$) that are positive, have unit norm and dominate $\underline{P}^{\mathbf{x}}$. But these functions are exactly the elements of $\mathscr{P}_{\underline{P}^{\mathbf{x}}}(\Omega)$. Hence, $\underline{P}^{\mathbf{x}}(X) = \min_{R \in \mathscr{P}_{\underline{P}^{\mathbf{x}}}(\Omega)} R(X)$. Using the one-to-one correspondence of Theorem 2,

$$\underline{P}^{\mathbf{x}}(X) = \min_{Q \in \mathcal{M}(\underline{P})} \Phi^{\mathbf{x}}_{\underline{P}}(Q)(X) = \min_{Q \in \mathcal{M}(\underline{P})} Q^{\mathbf{x}}(X). \quad \square$$

5 The linear case: charges and Dunford integrals

As our extension can be written as the lower envelope of extended linear previsions, it is of particular interest that extended linear previsions can be written as an integral. One possible choice for this integral turns out to be the Dunford integral. Before stating any results, let us first briefly review the theory of *charges*, or finitely additive measures, restricting ourselves to the essentials that are needed to understand the connection with previsibility. In particular, we shall focus on so-called probability charges on the set $\wp(\Omega)$ of all subsets of Ω as we do not need the more general theory of charges, defined on a more general field of subsets. General and detailed accounts can be found in [1] and [5].

A *probability charge* is a real-valued mapping on a field $\mathfrak{F}$ on Ω such that $\mu(\emptyset) = 0$, $\mu(\Omega) = 1$, $\mu(A) \geq 0$ whenever $A \in \mathfrak{F}$, $\mu(A) + \mu(B) = \mu(A \cup B)$ whenever $A, B \in \mathfrak{F}$ and $A \cap B = \emptyset$. From now on, $\mathfrak{F}$ is assumed to be the power set $\wp(\Omega)$. The (Dunford) integral of a simple random variable[6] $X = \sum_{i=1}^{n} a_i I_{A_i}$ ($n \in \mathbb{N}$, $a_i \in \mathbb{R}$ and $A_i \subseteq \Omega$) is then defined as $D\int X\,\mathrm{d}\mu = \sum_{i=1}^{n} a_i \mu(A_i)$.

A sequence of random variables (X_n) is said to *converge hazily* to a random variable X if for all $\epsilon > 0$ it holds that $\mu(\{\omega \in \Omega\colon\, |X_n(\omega) - X(\omega)| > \epsilon\}) \to 0$ as $n \to \infty$. X is called D-*integrable* if there is a sequence (X_n) of simple random variables that converges hazily to X, such that moreover $D\int |X_n - X_m|\,\mathrm{d}\mu \to 0$ as $n, m \to \infty$. Any such sequence is called a *determining sequence* for X. All determining sequences for X eventually have the same integral, and in this way, one defines the *Dunford integral* of X as the limit of integrals of simple random variables:

$$D\int X\,\mathrm{d}\mu = \lim_{n \to \infty} D\int X_n\,\mathrm{d}\mu.$$

The Dunford integral has all the properties that we expect from an integral, in particular, it is linear (and therefore finitely additive) and positive.

It is well known that there is a canonical one-to-one correspondence between linear previsions on $\mathscr{L}(\Omega)$ and probability charges on $\wp(\Omega)$: any linear prevision P on $\mathscr{L}(\Omega)$

[6] In this paper, a random variable which assumes only a finite number of values is called *simple*.

is uniquely determined through the probability charge μ defined by $\mu(A) = P(I_A)$. The linear prevision $P(X)$ of a gamble X is equal to the Dunford integral of X with respect to μ (see for instance [1, Theorem 4.7.4]).

But it is also well known that the set of all random variables that are D-integrable with respect to μ is much larger than the set of gambles. In fact, it is nothing but the set $\mathscr{L}_P^{\mathsf{x}}(\Omega)$ of P-previsible random variables, and it is worth nothing that it coincides with the so-called *Lebesgue space* $L_1(\Omega, \wp(\Omega), \mu)$ of those random variables X that are T_1-measurable[7] and whose absolute value $|X|$ is D-integrable. Moreover, our extension of a linear prevision P on $\mathscr{L}(\Omega)$ coincides exactly with the Dunford integral with respect to the probability charge μ. This is the subject of the following theorem.

Theorem 4. *Let P be a linear prevision on $\mathscr{L}(\Omega)$, and let μ be its restriction to the set of events $\wp(\Omega)$. Then the following statements hold.*

(i) A sequence of random variables (X_n) converges P-hazily to a random variable X if and only if (X_n) converges hazily to X with respect to μ.

(ii) A random variable X is P-previsible if and only if it is D*-integrable with respect to μ, in which case $P^{\mathsf{x}}(X) = D\int X \,\mathrm{d}\mu$.*

Proof. Compare the definition of P-hazy convergence with the definitions of hazy convergence with respect to μ to establish their equivalence.

Comparing the definitions of P-previsibility and D-integrability with respect to μ, it is immediately clear that D-integrability implies P-previsibility. To see that the converse holds too, assume that (X_n) is a P-determining sequence for X. Observe that (X_n) is a sequence of D-integrable random variables with respect to μ (recall that all gambles are D-integrable). By the definition of P-determining sequence, and the fact that P is uniquely determined by μ through the D-integral, we have that

$$\lim_{n,m\to\infty} D\int |X_n - X_m| \,\mathrm{d}\mu, = \lim_{n,m\to\infty} P(|X_n - X_m|) = 0,$$

and also that (X_n) converges hazily to X with respect to μ. From this we may conclude that X is D-integrable, and

$$\lim_{n\to\infty} D\int |X_n - X| \,\mathrm{d}\mu = 0.$$

(see [1, Theorem 4.4.20]). Again using the fact that P is uniquely determined by μ through the D-integral, we find that

$$D\int X\mathrm{d}\mu = \lim_{n\to\infty} D\int X_n\mathrm{d}\mu = \lim_{n\to\infty} P(X_n) = P^{\mathsf{x}}(X_n). \quad \square$$

6 Conclusions

The main message of this paper is that it is possible to define an extension of a coherent lower prevision to a linear space of previsible, not necessarily bounded random variables,

[7] A random variable X is called T_1-measurable with respect to μ if there is a sequence of simple random variables converging hazily to X with respect to μ

and that this extension still has properties similar to those of coherent lower previsions. Previsibility coincides with the existing notion of D-integrability when the coherent lower previsions are linear, and an extended lower prevision can be written as the lower envelope of the extensions of the dominating linear previsions of the original.

Acknowledgements

This paper presents research results of project G.0139.01 of the Fund for Scientific Research, Flanders (Belgium), and of the Belgian Programme on Interuniversity Poles of Attraction initiated by the Belgian state, Prime Minister's Office for Science, Technology and Culture. The scientific responsibility rests with the authors.

References

[1] K. Bhaskara Rao, M. Bhaskara Rao, Theory of Charges, a Study of Finitely Additive Measures, Academic Press, London, 1983.

[2] G. Choquet, Theory of capacities, Ann. Inst. Fourier 5 (1953–54) 131–295.

[3] G. de Cooman, Possibility theory I–III, International Journal of General Systems 25 (1997) 291–371.

[4] G. de Cooman, M. C. M. Troffaes, Dynamical programming for deterministic discrete-time systems with uncertain gain, 2003, submitted to the Third International Symposium on Imprecise Probabilities and Their Applications (ISIPTA '03).

[5] N. Dunford, J. T. Schwartz, Linear Operators, John Wiley & Sons, New York, 1957.

[6] B. de Finetti, Sul significato soggettivo della probabilità, Fundamenta Mathematicae 17 (1931) 298–329.

[7] B. de Finetti, Theory of Probability: a Critical Introductory Treatment, Wiley, London, 1975.

[8] I. Levi, The Enterprise of Knowledge. An Essay on Knowledge, Credal Probability, and Chance, MIT Press, Cambridge, 1983.

[9] L. V. Utkin, S. V. Gurov, Imprecise reliability for some new lifetime distribution classes, Journal of Statistical Planning and Inference 105 (1) (2002) 215–232.

[10] E. Schechter, Handbook of Analysis and Its Foundations, Academic Press, San Diego, 1997.

[11] P. Walley, T. L. Fine, Varieties of modal (classificatory) and comparative probability, Synthese 41 (1979) 321–374.

[12] P. Walley, Statistical Reasoning with Imprecise Probabilities, Chapman and Hall, London, 1991.

[13] L. A. Zadeh, Fuzzy sets as a basis for a theory of possibility, Fuzzy Sets Syst. 1 (1978) 3–28.