Is the mode a lower prevision?

Inés Couso
Dep. Statistics and O.R.
Universidad de Oviedo
couso@uniovi.es

Luciano Sánchez
Dep. Computer Sciences and A.I.
Universidad de Oviedo
luciano@uniovi.es

Abstract

We introduce the notion of mode-desirability of a gamble, that generalizes the idea of non-negativeness of the mode of a random variable. The lower and upper previsions derived from this new definition coincide with the minimum and maximum values of the set of modes of a gamble, when the credal set is a singleton, but they only bound them in the general case. The reason why the minimum and the maximum of the set of modes can not be written, in general, by means of a pair of lower and upper previsions is discussed.

Keywords. Expectation, median, mode, desirability, preference.

1 Introduction

In Decision Making Literature, several criteria of preference between random variables have been proposed within the setting of classical Probability Theory, like for instance stochastic dominance [10], dominance in the sense of expected utility [13], or statistical preference [7, 14], the last one being based on Condorcet's voting criterion ([2]). The above mentioned criteria share a commonality: the joint probability distribution induced by the pair of variables is assumed to be known in order to define each preference criterion, which is expressed in terms of it. Some generalizations of the aforementioned preference criteria have been recently reviewed ([3]) to the case where the joint distribution is not completely determined. Some of those generalizations had been previously introduced in the literature: Denoeux ([8]) generalized first-stochastic dominance to the case of belief-plausibility measures and Destercke ([9]) and Troffaes ([15]), for instance, consider several generalizations of Savage dominance criterion. We have shown that many of those preference generalizations can be expressed in terms of a general formulation that is related to the expectation of a function of both random variables, increasing in the first component and decreasing in the second one.

Differently, in Walley's setting, first hand information is expressed by means of a family of ordered pairs of variables (or "gambles"), the first one in the pair being preferred to the second one. This kind of knowledge can be equivalently represented by means of a coherent family of "desirable" gambles (those preferred to the null one). The family of desirable gambles induces a closed and convex set of linear previsions (also called a "credal set"). Each of those linear previsions is defined on the initial space and induces, for each pair of gambles, a (finitely-additive) joint probability. Thus, what is primary information in this framework is secondary information in the previous setting and vice versa. Notwithstanding, from a purely formal point of view, Walley's almost preference can be seen as a particular case of the general formula introduced in [3], if we consider the function that assigns, to each pair, the difference between both components. With those ideas in mind, we proposed in [6] a generalization of the notion of statistical preference from the setting of classical Probability Theory to the framework of Imprecise Probabilities. It leaded us naturally to a new desirability criterion that we called "signed-desirability". We say that X is signed-desirable if its sign (the gamble that takes the value 1 when X takes a positive value and -1, when it is negative) is desirable, according to Walley's framework. In [5], a set of axioms characterizing the family of signed-desirable gambles induced by a coherent set of desirable gambles is provided. Furthermore, we have found an interesting connection with the notion of median: the infimum and supremum of the set of medians of a gamble, when we range an arbitrary credal set, can be respectively expressed as the lower and upper previsions, according to this new desirability definition.

In this paper, we will propose a new desirability condition very closely related to the notion of mode. The minimum and maximum values of the family of modes of a gamble associated to a single prevision do coincide with the lower and upper previsions of this new desirability condition. However, when we consider an arbitrary credal set, those lower and upper previsions bound the set of modes, but do not necessarily coincide with their minimum and maximum values. We will explore in Section 4 the reasons why those pairs of values do not coincide in general.

2 Preliminaries

The basics on Imprecise Probabilities are assumed to be known by the reader. Notwithstanding we will introduce here the formal notation used in the rest of the paper, and specify the axioms that characterize a coherent family of desirable gambles ([16]). Those axioms have not been stable along the literature in what concerns the inclusion of the null gamble (see [4] for a detailed discussion). In this paper, we will assume it to be non-desirable.

Let Ω denote the set of outcomes of an experiment. $\mathcal{L}$ will denote the set of all gambles (bounded mappings from Ω to $\mathbb{R}$). For $X, Y \in \mathcal{L}$ let $X \geq Y$ mean that $X(\omega) \geq Y(\omega)$, $\forall\, \omega \in \Omega$ and let $X > Y$ mean that $X \geq Y$ and $X(\omega) > Y(\omega)$ for some $\omega \in \Omega$. A subset $\mathcal{D}$ of $\mathcal{L}$ is said to be a *coherent set of desirable gambles* [16] when it satisfies the following four axioms:

D1. If $X \leq 0$ then $X \notin \mathcal{D}$. *(Avoiding partial loss).*

D2. If $X > 0$, then $X \in \mathcal{D}$. *(Accepting partial gain).*

D3. If $X \in \mathcal{D}$ and $c \in \mathbb{R}^+$, then $cX \in \mathcal{D}$. *(Positive homogeneity).*

D4. If $X \in \mathcal{D}$ and $Y \in \mathcal{D}$, then $X + Y \in \mathcal{D}$. *(Addition).*

The *lower prevision induced by a set of desirable gambles* $\mathcal{D}$ is the set function $\underline{P} : \mathcal{L} \to \mathbb{R}$ defined as follows:

$$\underline{P}(X) = \sup\{c \,:\, X - c \in \mathcal{D}\}.$$

The *upper prevision induced by* $\mathcal{D}$ is the set function $\overline{P} : \mathcal{L} \to \mathbb{R}$ defined as follows:

$$\overline{P}(X) = \inf\{c \,:\, c - X \in \mathcal{D}\}.$$

The set of linear previsions induced by a coherent set of gambles $\mathcal{D}$ is defined as:

$$\mathcal{P}_{\mathcal{D}} = \{P \,:\, P(X) \geq 0 \text{ for all } X \in \mathcal{D}\}.$$

$\mathcal{P}_{\mathcal{D}}$ is always a *credal set* (a closed and convex set of linear previsions, whose restrictions to events are finitely additive probability measures). $\underline{P}$ and $\overline{P}$ are dual and they respectively coincide with the minimum and the maximum of $\mathcal{P}_{\mathcal{D}}$, which can be defined in turn, as the set of linear previsions that dominate $\underline{P}$. On the other hand, a subset $\mathcal{D}^- \subset \mathcal{L}$ satisfying Axioms D2–D4 and

D1'. If $\sup X < 0$ then $X \notin \mathcal{D}^-$. *(Avoiding sure loss).*

D5. If $X + \delta \in \mathcal{D}^-$, for all $\delta > 0$ then $X \in \mathcal{D}^-$. *(Closure).*

is called a coherent set of *almost desirable gambles.* A set of almost desirable gambles $\mathcal{D}^-$ determines a pair of lower and upper previsions, and a credal set, by means of expressions analogous to the case of desirable gambles. Conversely, a credal set univocally determines a coherent set of almost desirable gambles via the formula:

$$\mathcal{D}^-_{\mathcal{P}} = \{X \in \mathcal{L} \,:\, P(X) \geq 0,\ \forall\, P \in \mathcal{P}\}.$$

Finally, a set $\mathcal{D}^+ \subset \mathcal{L}$ is said to be a coherent set of *strict desirable gambles* if it is a coherent set of desirable gambles, and it satisfies, in addition, the following axiom:

D6. If $X \in \mathcal{D}^+$, then either $X > 0$ or $X - \delta \in \mathcal{D}^+$, for some $\delta > 0$. (Openness).

A coherent set of strict desirable gambles can be derived from a credal set as follows:

$$\mathcal{D}^+_{\mathcal{P}} = \{X \,:\, X > 0 \text{ or } P(X) > 0\ \forall\, P \in \mathcal{P}\}.$$

Let the reader notice that $\mathcal{D}^+_{\mathcal{P}}$ can be alternatively expressed in terms of the lower prevision $\underline{P}$ as follows:

$$\mathcal{D}^+_{\mathcal{P}} = \{X \,:\, X > 0 \text{ or } \underline{P}(X) > 0\}. \tag{1}$$

In Walley's theory, the notion of *preference* between two gambles is dual to the above notion of desirability: X is said to be preferred to Y when their difference $X - Y$ is desirable. Conversely, if our primary information is described by means of a partial preference ordering, we will say that X is desirable when it is preferred to the null gamble. Furthermore, there exists a formal connection between preference criteria in classical Probability literature and Walley's notion of preference: in the particular situation where the credal set associated to a preference ordering (according to Walley's view) is a singleton, $\{P\}$, Walley's almost preference of X over Y, $P(X - Y) \geq 0$, is equivalent to dominance according to the expectation, i.e., X is almost preferred to Y if and only if $E_P(X) \geq E_P(Y)$. (In the last expression, P is considered as a probability defined on the set of events, instead of a linear prevision defined in the set of gambles.)

In [3], some known notions of dominance in the (classical) probabilistic setting were reviewed, and it was shown that all of these orderings can be expressed by means of the formula $E_P[g(X,Y)] \geq 0$, where $g : \mathbb{R}^2 \to \mathbb{R}$ is increasing in the first component, and decreasing in the second one. It was also clarified that some generalizations of the above notions considered in the recent literature (see, for instance, [8, 9, 12, 15]) are very closely related to the formula $E_P[g(X,Y)] \geq 0$. This idea made possible to connect Walley's framework, where the initial information is expressed in terms of a partial ordering and the alternative setting considered in those reviewed papers, where the initial information is represented by means of a lower prevision. Therefore, we can join both frameworks and say that X is *g-preferred* to Y if $g(X,Y)$ is desirable according to Walley's framework. With this idea in mind we introduced the notion of sign-desirability in ([6]). X is said to be sign-preferred to Y if $\mathrm{sgn(X-Y)} = 1_{\mathrm{X>Y}} - 1_{\mathrm{Y>X}}$ is desirable, where 1_A denotes the indicator function of $A \subseteq \Omega$, and $X > Y$ and $Y > X$ respectively denote the subsets of Ω where X and Y satisfy each of those inequalities. According to this new preference condition, X is said to be sign-desirable when $\mathrm{sgn(X)} = 1_{\mathrm{X>0}} - 1_{\mathrm{X<0}}$ is desirable. In words, X is said to be sign-desirable when we are disposed to pay one probability currency unit if X takes a negative value in return for the gamble $1_{X>0}$ (receiving 1 unit if X takes a -strictly- positive value.). In [5] an axiomatic characterization of "coherent" sets of sign-desirable gambles is provided. The associated pair of lower and upper previsions can be defined as follows:

$$\underline{P}_S(X) = \sup\{c : X - c \text{ is strictly sign-desirable}\}$$

$$\overline{P}_S(X) = \inf\{c : c - X \text{ is strictly sign-desirable}\}.$$

We have checked in [6] that those lower and upper previsions do coincide, in fact, with the infimum and the supremum of the set of medians of X when we range the credal set associated to the initial coherent set of desirable gambles.

In this paper, we will explore the generalization of the notion of mode, and its connections with Walley's desirability theory. We will introduce a new notion of desirability, but it will not be expressed in terms of the desirability of an increasing function of the considered gamble, as it happens with the notion of sign-desirability. We will also consider the pair of lower and upper previsions of a gamble, according to the new desirability condition. The infimum of the set of modes associated to a credal set will be bounded by the lower prevision, but it will not coincide in general with it.

3 The notion of mode-desirability

Let $\mathcal{L}_F$ denote the family of "simple gambles" (those with a finite number of different possible values). Let us consider an arbitrary but fixed probability measure P on Ω. According to the classical definition, the set of modes of a gamble $X \in \mathcal{L}_F$ with a finite image $Im(X) = \{x_1, \ldots, x_n\}$ is defined as follows:

$$\begin{aligned}&Mo_P(X) =\\ &\{x_i \in Im(X) : P(X = x_j) \leq P(X = x_i),\ \forall j \neq i\} =\\ &\{x_i \in Im(X) : \nexists x_j \neq x_i \text{ with } P(X = x_j) > P(X = x_i)\} =\\ &\{x_i \in Im(X) : \nexists j \neq i \text{ s.t. } E_P(1_{X=x_j} - 1_{X=x_i}) > 0\}.\end{aligned}$$

Let us now consider the credal set, $\mathcal{P}_{\mathcal{D}}$, associated to an arbitrary coherent set of desirable gambles $\mathcal{D}$. Let $\underline{P}$ denote the induced lower prevision. A natural way to extend the classical notion of mode seems to be the following one:

$$\begin{aligned}&Mo_{\underline{P}}(X) =\\ &\{x_i \in Im(X) : \underline{P}(1_{X=x_j} - 1_{X=x_i}) \leq 0,\ \forall j \neq i\} =\\ &\{x_i \in Im(X) \quad \nexists j \neq i \text{ s.t. } \underline{P}(1_{X=x_j} - 1_{X=x_i}) > 0\}.\end{aligned}$$

We will prove the following result, in order to connect this definition with Walley's desirability framework.

Lemma 1 *Let $\underline{P}$ be the lower prevision induced by a coherent set of gambles $\mathcal{D}$. Let $\mathcal{D}^+_{\mathcal{P}}$ be the set associated set of strictly desirable gambles, according to Equation 1. Let $X \in \mathcal{L}_F$. For every $x \in Im(X)$ and all $y \in \mathbb{R}$:*

$$\underline{P}(1_{X=y} - 1_{X=x}) > 0 \quad \textit{iff} \quad 1_{X=y} - 1_{X=x} \in \mathcal{D}^+.$$

Proof: By definition, the gamble $1_{X=y} - 1_{X=x}$ is strictly desirable if and only if it is some of the following conditions are fulfilled:

$$\underline{P}(1_{X=y} - 1_{X=x}) > 0 \quad \text{or} \quad 1_{X=y} - 1_{X=x} > 0.$$

But $1_{X=y} - 1_{X=x} > 0$ implies that x does not belong to the set of outcomes of X, what is a contradiction. □

According to the above lemma, we can alternatively express the set of modes as follows:

$$\begin{aligned}&Mo_{\underline{P}}(X) =\\ &\{x_i \in Im(X) : \nexists j \neq i \text{ s.t. } 1_{X=x_j} - 1_{X=x_i} \in \mathcal{D}^+\} =\\ &\{x_i \in Im(X) : \nexists j \neq i \text{ s.t. } (1_{\{x_j\}} - 1_{\{x_i\}}) \circ X \in \mathcal{D}^+\},\end{aligned}$$

where the symbol "$\circ$" stands for the composition of functions.

Furthermore, we can skip our reference to the set of outcomes of X by taking into account the following result.

Lemma 2 *Let us consider a credal set $\mathcal{P}$, and let $\mathcal{D}^+$ denote the set of strictly desirable gambles induced by it. Let $X \in \mathcal{L}$. Then:*

1. *If $y \notin Im(X)$, and $x \in \mathbb{R}$, $1_{X=y} - 1_{X=x} \notin \mathcal{D}^+$.*
2. *$A_X^+ = \{x : \nexists y \neq x \text{ s.t. } (1_{\{y\}} - 1_{\{x\}}) \circ X \in \mathcal{D}^+\}$ is included in $Im(X)$.*

Proof:

1. If $y \notin Im(X)$, then $(1_{X=y} - 1_{X=x}) = -1_{X=x} \leq 0$. According to Axiom D1, this gamble does not belong to $\mathcal{D}^+$.
2. The second part is also straightforward: if $x \notin Im(X)$, then $(1_{\{y\}} - 1_{\{x\}}) \circ X > 0, \forall y \in Im(X)$, and therefore, the gamble $(1_{\{y\}} - 1_{\{x\}}) \circ X$ belongs to $\mathcal{D}^+$ for every $y \in Im(X) \subseteq \mathbb{R} \setminus \{x\}$. $\square$

According to the above lemma, the set of modes associated to the credal set, $Mo_{\underline{P}}(X)$, can be alternatively expressed as:

$$Mo_{\underline{P}}(X) = A_X^+ =$$
$$\{x : \nexists y \neq x \text{ s.t. } (1_{\{y\}} - 1_{\{x\}}) \circ X \in \mathcal{D}^+\}.$$

This new expression suggests us to consider the following new desirability condition. We will say that X is mode-desirable when $Mo_{\underline{P}}(X) = A_X^+$ does not contain any negative number:

Definition 1 *A gamble $X \in \mathcal{L}_F$ is said to be* mode-desirable*, if*

$$[\forall\, x < 0,\ \exists y \neq x \text{ s.t. } (1_{\{y\}} - 1_{\{x\}}) \circ X \in \mathcal{D}^+].$$

We will denote it $X \in \mathcal{D}_{Mo}$.

Remark 3.1 *There is an alternative equivalent definition for the notion of mode-desirability of simple gambles. In fact we can check that X is mode-desirable if and only if:*

$$[\forall\, x < 0,\ \exists y > x \text{ s.t. } (1_{\{y\}} - 1_{\{x\}}) \circ X \in \mathcal{D}^+].$$

One of the implications is straightforward, so we just need to check the second one: Let us suppose that $X \in \mathcal{D}_{Mo}$ and let us consider an arbitrary but fixed value $x \leq 0$. According to the definition of $\mathcal{D}_{Mo}$, there exists $y_1 \neq x$ such that $(1_{\{y_1\}} - 1_{\{x\}}) \circ X$. Furthermore, we can assure that y_1 belongs to $Im(X)$. If $y_1 > x$, the proof is finished. Otherwise, there will exist $y_2 \neq y_1$, $y_2 \in Im(X)$ such that $(1_{\{y_2\}} - 1_{\{y_1\}}) \circ X \in \mathcal{D}^+$. According to the additivity of $\mathcal{D}^+$ (Axiom D4), we can easily check that $(1_{\{y_2\}} - 1_{\{x\}}) \circ X \in \mathcal{D}^+$. According to this procedure, after a finite number of steps, $k \leq \#Im(X)$, we will get $y_{k+1} > x$ such that $(1_{y_{k+1}} - 1_{y_k}) \circ X \in \mathcal{D}^+$. Otherwise, we would need to assume that y_n is less than or equal to x, and it would lead us to a contradiction, because, there would need to exist $y \notin Im(X)$ with $(1_y - 1_{y_n}) \circ X \in \mathcal{D}^+$.

If X is mode-desirable, then, for every $x < 0$, there exists some $y \neq x$ such that we are disposed to exchange the gamble $1_{X=x}$ in return for the gamble $1_{X=y}$. The new desirability condition induces a pair of lower and upper previsions as follows:

Definition 2 *Let $\mathcal{D}$ be a coherent family of desirable gambles, and let $\mathcal{D}_{Mo}$ denote the family of mode-desirable gambles induced by it. Let $X \in \mathcal{L}_F$. The lower prevision of X is defined as follows:*

$$\underline{P}_{Mo}(X) = \sup\{c \in \mathbb{R} : X - c \in \mathcal{D}_{Mo}\}$$

Analogously, the upper prevision is:

$$\overline{P}_{Mo}(X) = \inf\{c \in \mathbb{R} : c - X \in \mathcal{D}_{Mo}\}.$$

Now we will prove that the minimum and the maximum values of the set A_X^+ do coincide with the pair of lower and upper previsions defined above. Let us first prove the following supporting result:

Lemma 3

- *The set $C = \{c : X - c \in \mathcal{D}_{Mo}\}$ can be alternatively expressed as:*

  $$\{c : [x < c \Rightarrow \exists\, y \neq x \text{ with } (1_{\{y\}} - 1_{\{x\}}) \circ X \in \mathcal{D}^+]\} =$$
  $$\{c : [x < c \Rightarrow x \notin A_X^+]\} = (-\infty, \min A_X^+].$$

- *The set $D = \{d : d - X \in \mathcal{D}_{Mo}\}$ can be alternatively written as:*

  $$\{d : [x > d \Rightarrow \exists\, y \neq x \text{ with } (1_{\{y\}} - 1_{\{x\}}) \circ X \in \mathcal{D}^+]\} =$$
  $$\{d : [x > d \Rightarrow x \notin A_X^+]\} = [\max A_X^+, \infty).$$

Proof: The proof is almost immediate, if we take into account that $1_{\{y\}} \circ (X - c) = 1_{\{y+c\}} \circ X$, and $1_{\{y\}} \circ (d - X) = 1_{\{d-y\}} \circ X\ \forall c, d, y \in \mathbb{R}$. $\square$

The next result is straightforward, according to the above lemma:

Proposition 4 *The following equalities hold: $\min A_X^+ = \underline{P}_{Mo}(X)$ and $\max A_X^+ = \overline{P}_{Mo}(X)$.*

Remark 3.2 *According to the proof of Lemma 3, the supremum of C and the infimum of D are, indeed, maximum and minimum values, respectively, and they do coincide with the minimum and the maximum of A_X^+, respectively.*

Let us now consider the set of mode values associated to the credal set:

$$Mo_{\mathcal{P}_{\mathcal{D}}}(X) = \cup_{P\in\mathcal{P}_{\mathcal{D}}}\{Mo_P(X)\}.$$

If it coincided with A^+_X, the minimum and the maximum of the family of modes associated to the credal set would coincide with the lower and upper previsions of X, according to the notion of mode-desirability. Nevertheless, those lower and upper previsions just bound, but they do not coincide in general with the minimum and maximum of the set of modes of X. More specifically, we can check that:

Proposition 5 *The set of mode values associated to the credal set $\mathcal{P}_{\mathcal{D}}$, $Mo_{\mathcal{P}_{\mathcal{D}}}(X)$ is included in A^+_X. Furthermore, if the credal set is a singleton, both sets of values do coincide.*

Proof: The set of modes can be expressed as follows:

$$Mo_{\mathcal{P}_{\mathcal{D}}}(X) = \cup_{P\in\mathcal{P}_{\mathcal{D}}}\{Mo_P(X)\} =$$

$$\cup_{P\in\mathcal{P}_{\mathcal{D}}}\{x : \forall y \neq x, P(1_{\{y\}} - 1_{\{x\}}) \circ X \leq 0\} =$$

$$\{x : \exists P \in \mathcal{P}_{\mathcal{D}} \text{ s.t. } \forall y \neq x P(1_{\{y\}} - 1_{\{x\}}) \circ X \leq 0\}.$$

On the other hand,

$$A^+_X = \{x : \forall y \neq x, \underline{P}(1_{\{y\}} - 1_{\{x\}}) \circ X) \leq 0\}.$$

According to the above expressions, and taking into account that $\underline{P}$ is the minimum of the credal set, we can easily derive the thesis of this proposition. □

According to the last results, A^+_X is a finite set containing the set of modes, $Mo_{\mathcal{P}_{\mathcal{D}}}(X)$, and included in the set of images of X. Under some additional constraints ($\mathcal{P}_{\mathcal{D}}$ being a singleton or, contrarily, expressing vacuous information, or A^+_X being included in the set of images with maximum upper probability, etc.) they do coincide. But they do not in general, as we illustrate in the following example.

Example 1 *Let Ω be a finite set with four elements, $\Omega = \{\omega_1, \omega_2, \omega_3, \omega_4\}$ and let us consider the credal set $\mathcal{P} = \{(\frac{3}{8}-\alpha, \frac{1}{8}-\frac{\alpha}{4}, \frac{1}{8}+\frac{\alpha}{4}, \frac{3}{8}+\alpha) : \alpha \in [-\frac{3}{8}, \frac{3}{8}]\}$. In the above formula, each vector of the form (p_1, p_2, p_3, p_4) represents the linear prevision P defined as:*

$$P(X) = \sum_{i=1}^{4} p_i X(\omega_i),\ \forall X \in \mathcal{L}.$$

Let $\mathcal{D}^+_{\mathcal{P}}$ denote the set of strictly desirable gambles associated to $\mathcal{P}$: $\mathcal{D}^+_{\mathcal{P}} = \{Y : Y > 0 \ \text{ or } \underline{P}(Y) > 0\}$. Let us now consider the gamble X defined as $X(\omega_i) = i,\ i = 1,2,3,4$. Let A^+_X denote the collection of numbers:

$$A^+_X = \{x : \nexists y \neq x \text{ with } (1_{\{y\}} - 1_{\{x\}}) \circ X \in \mathcal{D}^+\} =$$

$$\{i \in \{1,\ldots,4\} : \forall j \neq i,\ \underline{P}(1_{\{\omega_j\}} - 1_{\{\omega_i\}}) \leq 0\}.$$

$A^+_X = \{1,2,3,4\}$, but $Mo_{\mathcal{P}}(X) = \{1,4\}$. In order to check it, Tables 1 and 2 respectively display, for each pair (j,i), the value that the linear prevision $P_\alpha \equiv (\frac{3}{8}-\alpha, \frac{1}{8}-\frac{\alpha}{4}, \frac{1}{8}+\frac{\alpha}{4}, \frac{3}{8}+\alpha)$ and the lower prevision $\underline{P} = \min_{\alpha\in[-\frac{3}{8},\frac{3}{8}]} P_\alpha$ assign to the gamble $(1_{\{x_j\}} - 1_{\{x_i\}}) \circ X = 1_{\{\omega_j\}} - 1_{\{\omega_i\}}$.

$j \setminus i$	1	2	3	4
1	0	$\frac{1}{4}-\frac{3\alpha}{4}$	$\frac{1}{4}-\frac{5\alpha}{4}$	2α
2	$\frac{3\alpha}{4}-\frac{1}{4}$	0	$-\frac{\alpha}{2}$	$-\frac{1}{4}-\frac{5\alpha}{4}$
3	$\frac{5\alpha}{4}-\frac{1}{4}$	$\frac{\alpha}{2}$	0	$-\frac{1}{4}-\frac{3\alpha}{4}$
4	-2α	$\frac{1}{4}-\frac{5\alpha}{4}$	$\frac{1}{4}+\frac{3\alpha}{4}$	0

Table 1: It displays $P_\alpha(1_{\{\omega_j\}} - 1_{\{\omega_i\}})$, for each (j,i).

$j \setminus i$	1	2	3	4
1	0	$-\frac{1}{32}$	$-\frac{7}{32}$	$-\frac{3}{4}$
2	$-\frac{17}{32}$	0	$-\frac{3}{16}$	$-\frac{3}{8}$
3	$-\frac{23}{32}$	$-\frac{3}{16}$	0	$-\frac{1}{32}$
4	$-\frac{3}{4}$	$-\frac{7}{32}$	$-\frac{1}{32}$	0

Table 2: It displays $\underline{P}(1_{\{\omega_j\}} - 1_{\{\omega_i\}})$, for each (j,i).

None of the values in Table 2 is strictly positive, and this means that A^+_X coincides with the set of possible outcomes of the gamble X, $\{1,2,3,4\}$. On the other hand, there does not exist any $\alpha \in [-\frac{3}{8}, \frac{3}{8}]$ such that the values 2 or 3 belong to the set of modes of X associated to the linear prevision P_α, $Mo_{P_\alpha}(X)$. Thus, the set of modes associated to the credal set, $Mo_{\mathcal{P}}(X)$, is strictly included in A^+_X.

We can ask ourselves what happens if we replace, $\mathcal{D}^+$ by $\mathcal{D}$ or $\mathcal{D}^-$ in the construction of the set of values:

$$\{x : \nexists y \neq x \text{ with } (1_{\{y\}} - 1_{\{x\}}) \circ X \in \mathcal{D}^+\}.$$

Let us consider the pair of sets:

$$A_X = \{x : \nexists y \neq x \text{ with } (1_{\{y\}} - 1_{\{x\}}) \circ X \in \mathcal{D}\}$$

and

$$A^-_X = \{x : \nexists y \neq x \text{ with } (1_{\{y\}} - 1_{\{x\}}) \circ X \in \mathcal{D}^-\} =$$

$$\{x : \underline{P}((1_{\{y\}} - 1_{\{x\}}) \circ X) < 0,\ \forall y \neq x\},$$

and let us compare them with A^+_X.

Lemma 6 *$A^-_X \subseteq A_X \subseteq A^+_X$. Furthermore, if $\mathcal{P}_{\mathcal{D}}$ is a singleton, $\mathcal{P}_{\mathcal{D}} = \{P\}$, then $A^-_X = \emptyset$, unless the distribution of X is unimodal. In that case, $A^-_X = A_X = A^+_X = Mo_P(X)$.*

Proof: The first part is easy to prove if we take into account the chain of inclusions $\mathcal{D}^+ \subseteq \mathcal{D} \subseteq D^-$. Secondly, if $\mathcal{P}_{\mathcal{D}} = \{P\}$, we can easily check that x belongs to A_X^- if and only if $P(1_{X=y}) < P(1_{X=x}),\ \forall\, y \neq x$. This only happens when x is the only mode of X, with respect to the linear prevision P. $\square$

Remark 3.3 *Using expressions analogous to those considered in Lemma 3, we can easily prove that the minimum and maximum of A_X^- do respectively coincide with* $\sup\{c : X - c \in \mathcal{D}_{Mo}^-\}$ *and* $\inf\{d : d - X \in \mathcal{D}_{Mo}^-\}$*, where $\mathcal{D}_{Mo}^-$ is defined as:*

$$\{X \in \mathcal{L}_F \,:\, \forall\, x < 0\ \exists\, y \neq x \text{ s.t. } (1_{\{y\}} - 1_{\{x\}}) \circ X \in \mathcal{D}^-\}.$$

Furthermore, we have seen that A_X^- is included in A_X^+, and that the last one coincides with the set of modes, when the credal set is a singleton. We can ask ourselves whether A_X^- is, in general a subset of $Mo_{\mathcal{P}}(X)$, and therefore it approximates it from below. But we can easily check that this does not happen. In Example 1, we have shown that none of the lower previsions displayed in Table 2 was strictly positive. Furthermore, we observe that all of them are negative (except for those in the diagonal). This means that A_X^- also coincides with the whole family of possible outcomes of X, $A_X^- = \{1, 2, 3, 4\}$ and therefore, it strictly includes the set of mode values associated to the credal set.

4 What's the problem with mode-desirability?

In Walley's framework ([16]), any coherent set of gambles satisfies Axioms D2 and D4. The following property can be easily derived from both axioms:

$$Y \in \mathcal{D}, \text{ and } \ X > Y \Rightarrow X \in \mathcal{D}. \qquad (2)$$

On the other hand, the set of sign-desirable gambles induced by a coherent set of gambles $\mathcal{D}$ satisfies Axiom D2, but it does not necessarily satisfy Axiom D4. However we can easily check that it fulfills the property mentioned in Equation 2, since it is connected to $\mathcal{D}^+$ through the function $\mathrm{sgn} : \mathbb{R} \to \mathbb{R}$, that is increasing. More explicitly:

Definition 3 *Let $\mathcal{D}$ be a coherent set of desirable gambles, and let $f : \mathbb{R} \to \mathbb{R}$ be an increasing function. We will say that X is f-desirable if and only if $f(X)$ belongs to $\mathcal{D}$. We will denote it $X \in \mathcal{D}_f$.*

Lemma 7 *Let $\mathcal{D}$ be a coherent set of desirable gambles, and let $f : \mathbb{R} \to \mathbb{R}$ be an increasing function. The set of f-desirable gambles satisfies the property:*

$$X \in \mathcal{D}_f, Y > X \Rightarrow Y \in \mathcal{D}_f.$$

A "coherent" set of mode-desirable gambles does not necessarily satisfy the property considered in Equation 2 as we illustrate in Example 2:

Example 2 *Let Ω be the unit interval, and let P denote the uniform probability distribution defined on it. Let Y denote the gamble defined as follows:*

$$Y(\omega) = \begin{cases} -1 & \text{if } \ \omega \in [0, 1/3) \\ 1 & \text{if } \ \omega \in [1/3, 5/6) \\ 2 & \text{if } \ \omega \in [5/6, 1] \end{cases}$$

Y takes the values -1, 1 and 2 with respective probabilities $1/3$, $1/2$ and $1/6$. Thus, we can easily check that Y is mode-desirable, since $P(1_{\{1\}} - 1_{\{x\}} \circ Y) > 0,\ \forall\, x < 0$. Let us now consider the gamble:

$$X(\omega) = \begin{cases} -1 & \text{if } \ \omega \in [0, 1/3) \\ 1 & \text{if } \ \omega \in [1/3, 1/2) \\ 2 & \text{if } \ \omega \in [1/2, 2/3) \\ 3 & \text{if } \ \omega \in [2/3, 5/6) \\ 4 & \text{if } \ \omega \in [5/6, 1] \end{cases}$$

We clearly see that $Y \geq X$, but it is not mode-desirable. In fact, for $x = -1$ there does not exist any $y > x$ such that $P(1_{\{y\}} - 1_{\{x\}} \circ X) > 0$.

From this example, and according to Lemma 7, a "coherent" sets of mode-desirable gambles can not be expressed, in general, as the family of f-desirable gambles, according to some increasing function $f : \mathbb{R} \to \mathbb{R}$ and some coherent set of desirable gambles $\mathcal{D}$. This fact seems to be essential in relation with the properties of the lower and upper previsions derived from it, as we show below.

Lemma 8 *Let $\mathcal{D}$ be a coherent set of desirable gambles, and let us consider an increasing function $f : \mathbb{R} \to \mathbb{R}$. The set $C = \{c : f(X - c) \in \mathcal{D}\}$ satisfies the following property: $c \in C, c' \leq c \Rightarrow c' \in C$.*

Proof: Let us suppose that $c \in C$ and $c' \leq c$. By definition, $f(X - c) \in \mathcal{D}$. According to the properties of f, $f(X - c') \geq f(X - c)$ and, therefore, according to the coherence of $\mathcal{D}$, $f(X - c')$ belongs to it. $\square$

Proposition 9 *Let $\mathcal{D}$ be a coherent set of desirable gambles, and let us consider an increasing function $f : \mathbb{R} \to \mathbb{R}$. Let $\mathcal{D}_f^+$ denote the set of $f-$desirable gambles with respect to the coherent set $\mathcal{D}^+$, $\mathcal{D}_f^+ = \{X : f(X) \in \mathcal{D}^+\}$. Let us also consider, for every $P \in \mathcal{P}_{\mathcal{D}}$, the set of f-desirable gambles with respect to $\mathcal{D}_{\{P\}}^+$, i.e.: $\mathcal{D}_{f,\{P\}}^+ = \{X : f(X) > 0 \ \text{ or } P(f(X)) > 0\}$. Then:*

$$\sup\{c : X - c \in \mathcal{D}_f^+\} = \inf_{P \in \mathcal{P}_{\mathcal{D}}} \sup\{c : X - c \in \mathcal{D}_{f,\{P\}}^+\}.$$

Proof: First of all, let us take into account that $\mathcal{D}^+ \subseteq \mathcal{D}^+_{\{P\}}$, and therefore $\mathcal{D}^+_f \subseteq \mathcal{D}^+_{f,\{P\}}$, $\forall P \in \mathcal{P}$. Thus, the set $\{c : X - c \in \mathcal{D}^+_f\}$ is included in $\{c : X - c \in \mathcal{D}^+_{f,\{P\}}\}$, $\forall P \in \mathcal{P}$, and therefore

$$\sup\{c : X-c \in \mathcal{D}^+_f\} \leq \inf_{P\in\mathcal{P}_\mathcal{D}} \sup\{c : X-c \in \mathcal{D}^+_{f,\{P\}}\}.$$

Let us now prove the reverse inequality.Let c_P denote the supremum of the set $\{c : f(X-c) \in \mathcal{D}^+_{f,\{P\}}\}$ and let $c = \inf_{P\in\mathcal{P}} c_P$. Let us consider an arbitrary $c' < c$. It will suffice to check that, $c' \in \{c : X - c \in \mathcal{D}^+_f\}$. Let us consider the difference $\epsilon = c - c' > 0$. According to the definition of supremum, for every $P \in \mathcal{P}$ there exists $c'_P \in \{c : X-c \in \mathcal{D}^+_{f,\{P\}}$ such that $c_P - \epsilon < c'_P \leq c_P$. Therefore, $c' \leq \inf_{P\in\mathcal{P}} c'_P$ and thus, according to Lemma 8, $f(X - c') \in \mathcal{D}^+_{\{P\}}$, $\forall P \in \mathcal{P}$. Having into account that $\mathcal{D}^+ = \cap_{P\in\mathcal{P}}\mathcal{D}^+_{\{P\}}$, we have that $c' \in \{c : X - c \in \mathcal{D}^+_f\}$, and the result is proved. □

According to the last result, when we consider an increasing function $f : \mathbb{R} \to \mathbb{R}$, and the supremum $\sup\{c : f(X-c) \in \mathcal{D}^+_{\{P\}}\}$ coincides with some well-known parameter, $\theta_P(X)$ induced by the probability distribution P_X (like, for instance, the expectation for $f(\cdot) = \cdot$, or the infimum of the interval of medians, for $f = \text{sgn}$, the supremum $\sup\{c : f(X-c) \in \mathcal{D}^+\}$ coincides with the infimum of the values of the parameter, when we range the credal set, $\inf_{P\in\mathcal{P}_\mathcal{D}} \theta_P(X)$.

The condition of mode-desirability cannot be expressed in terms of an increasing function. According to Example 2, it is something inherent to the standard definition of mode, and it does not depend on the particular definition we have introduced in order to extend the idea of non-negativity of the mode to the Imprecise Probabilities framework. Even for the family of single-pointed credal sets, we cannot find an increasing function $f : \mathbb{R} \to \mathbb{R}$ such that $\sup\{c : f(X-c) \in \mathcal{D}^+_{\{P\}}\} = \min Mo_P(X)$, for every linear prevision, P.

5 Alternative definitions of mode desirability

As we have mentioned in the introduction, [3] reviews several classical stochastic preference criteria and shows that many of them can be written according to the general formulation:

$$X \text{ is preferred to } Y \text{ iff } E_P(g(X,Y)) \geq 0,$$

where $g : \mathbb{R}^2 \to \mathbb{R}$ is increasing in the first component and decreasing in the second one. Furthermore, in most cases, g can be expressed in terms of an increasing point-to-point function $f : \mathbb{R} \to \mathbb{R}$ as $g(x,y) = f(x) - f(y)$, $\forall (x,y) \in \mathbb{R}^2$. As we clarify in [3], some extensions of those stochastic orderings introduced in the recent literature ([6, 8, 9, 11, 15]) can be written in terms of the non-negativity of the lower prevision of $g(X,Y)$. Some others, instead, take into account the pairs of lower and upper previsions of $f(X)$ and $f(Y)$, $(\underline{E}(f(X)), \overline{E}(f(X)))$ and $(\underline{E}(f(Y)), \overline{E}(f(Y)))$. Based on both pairs, we can generate four different preference relations, that, for the sake of shortness, will be called min-max, max-max, max-min and min-min.

In Section 3, we considered the following generalization of the notion of mode:

$$Mo_{\underline{P}}(X) = \{x_i : \underline{P}(1_{X=x_j} - 1_{X=x_i}) \leq 0, \ \forall j \neq i\}.$$

Instead of the lower prevision of gambles of the form $(1_{\{x_j\}} - 1_{\{x_i\}}) \circ X$, we can alternatively consider the pairs of lower and upper previsions of the gambles $1_{\{x_j\}} \circ X$ and $1_{\{x_i\}} \circ X$ and compare them, according to the four criteria mentioned in the last paragraph. In this section we will briefly discuss these four alternative definitions.

Min-max criterion

Let $\underline{P}$ and $\overline{P}$ respectively denote the lower and upper previsions induced by a credal set $\mathcal{P}$. Let $X \in \mathcal{L}_F$ be an arbitrary simple gamble. We will define the *min-max-mode* of X with respect to $\mathcal{P}$ as the set:

$${}^M_m Mo_\mathcal{P}(X) = \{x_i : \underline{P}(1_{X=x_j}) \leq \overline{P}(1_{X=x_i}), \ \forall j \neq i\}.$$

According to the super-additivity of $\underline{P}$, and the duality between $\underline{P}$ and $\overline{P}$, the following inequality holds:

$$\underline{P}(1_{X=x_j}) - \overline{P}(1_{X=x_i}) \geq \underline{P}(1_{X=x_j}) - \overline{P}(1_{X=x_i}),$$

and therefore, we can easily check that the max-min-mode of X contains the set $Mo_{\underline{P}}(X)$, that is, in turn, a superset of the family of modes of X, when we range the credal set. Therefore, the max-min-mode is even less precise than our initial generalization of the mode.

Max-max criterion

We will define the *max-max-mode* of X with respect to $\mathcal{P}$ as follows:

$${}^M_M Mo_\mathcal{P}(X) = \{x_i : \overline{P}(1_{X=x_j}) \leq \overline{P}(1_{X=x_i}), \ \forall j \neq i\}.$$

This set is included in the set of modes of X, when we range the credal set. In fact, according to the coherence of $\overline{P}$, it is the maximum of the credal set, $\mathcal{P}$, and that means that there exists, for every $i \in$

${}^{M}_{M}Mo_{\mathcal{P}}(X)$, some $P_i \in \mathcal{P}$ that satisfies the equality $P_i(1_{X=x_i}) = \overline{P}(1_{X=x_i})$, that satisfies, by definition, the inequalities $\overline{P}(1_{X=x_i}) \geq \overline{P}(1_{X=x_j}),\ \forall j$. Thus, we get the inequalities:

$$P_i(1_{X=x_i}) = \overline{P}(1_{X=x_i}) \geq \overline{P}(1_{X=x_j}) \geq P_i(1_{X=x_j}),\ \forall j.$$

Therefore, the max-max-mode approximates the set of modes from below.

Max-min criterion

We will define the *max-min-mode* of X with respect to $\mathcal{P}$ as follows:

$${}^{m}_{M}Mo_{\mathcal{P}}(X) = \{x_i \,:\, \overline{P}(1_{X=x_j}) \leq \underline{P}(1_{X=x_i}),\ \forall j \neq i\}.$$

This set of values is clearly included in the max-max-mode, and therefore, it is a less precise approximation of the family of modes $Mo_{\mathcal{P}}(X)$.

Min-min criterion

We will define the *min-min-mode* of X with respect to $\mathcal{P}$ as the set:

$${}^{m}_{m}Mo_{\mathcal{P}}(X) = \{x_i \,:\, \underline{P}(1_{X=x_j}) \leq \underline{P}(1_{X=x_i}),\ \forall j \neq i\}.$$

$\underline{P}((1_{\{x_j\}} - 1_{\{x_i\}}) \circ X) \leq \overline{P}((1_{\{x_j\}} - 1_{\{x_i\}}) \circ X) \leq \overline{P}(1_{\{x_j\}} \circ X) - \underline{P}(1_{\{x_i\}} \circ X),\ \forall i, j.$

The above set does not necessarily include, nor is it necessarily included in the family of modes, $Mo_{\mathcal{P}}(X)$. Both sets may even be disjoint, as it happens in the following example.

Example 3 *Let us consider again the credal set of Example 1,* $\mathcal{P} = \{(\frac{3}{8} - \alpha, \frac{1}{8} - \frac{\alpha}{4}, \frac{1}{8} + \frac{\alpha}{4}, \frac{3}{8} + \alpha) \,:\, \alpha \in [-\frac{3}{8}, \frac{3}{8}]\}$. *The lower previsions of the gambles of the form* $1_{X=x_i},\ i = 1, 2, 3, 4,$ *are, respectively* 0, $\frac{1}{32}$, $\frac{1}{32}$ *and* 0. *Thus, the min-min-mode,* ${}^{m}_{m}Mo_{\mathcal{P}}(X) = \{2, 3\}$ *is the complementary of the set of modes of* X, $Mo_{\mathcal{P}}(X) = \{1, 4\}$.

6 Concluding remarks and open problems

We have introduced the notion of mode-desirability, and connected the classical notion of mode to Walley's desirability framework. The lower and upper previsions of a gamble bound, but do not necessarily coincide with the minimum and the maximum of the set of modes, when we consider an arbitrary credal set. In Section 4, we have discussed the reason why there does not seem to exist a way to express the pair of minimum and maximum values as the pair of lower and upper previsions, according to some desirability condition.

We have also studied four alternative generalizations of the notion of mode. The "min-max" approach leads to a pair of bounds that are even less precise than the lower and upper previsions induced from the notion of mode-desirability. Notwithstanding, the number of comparisons needed to calculate the outer approximation A_X^+ is greater than the number needed in order to calculate the min-max mode. It will be the expert that uses those approximations in practical problems who has to decide what is the most convenient procedure in each specific situation. On the other hand, the min-min mode does not seem to be related in general with the set of modes. Finally, the max-min and the max-max modes are included in the family of modes, the last one being the most precise of the two. In a specific problem, we can consider the outer and inner approximations of $Mo_{\mathcal{P}}(X)$ respectively derived from the notions of mode-desirability (or, alternatively, the min-max mode, when the calculation of A_X^+ is non-viable) and max-max mode. According to the notion of upper prevision, the max-max mode can be alternatively expressed as:

$$\left\{x_i \,:\, \cup_{j=1}^{n}\{d \,:\, d - 1_{X=x_j} \notin \mathcal{D}\} \subseteq \{d \,:\, d - 1_{X=x_i} \notin \mathcal{D}\}\right\}. \tag{3}$$

The max-max mode and the set A_X^+ approximate the set of bounds, respectively from below and above. At first sight, the problem of characterizing the set of modes associated to a credal set seems to be more complicated: the mode of a linear convex combination is not between the modes of both extremes. Therefore, the set of modes associated to a credal set does not seem to be easily characterized by the modes of the extremes, as it happens with other parameters, like the entropy (see [1], for instance). At least, the fact of departing from a pair of inner and outer approximations can simplify the process of characterizing the set of modes in some specific problems.

In the future, we plan to study the properties of the desirability condition that matches with the generalization of the notion of mode considered in Equation 3, as well as for the notion of mode-desirability. According to the definition introduced in this paper, a gamble is mode-desirable if and only if $A_X^+ \cap (-\infty, 0) \neq \emptyset$. The set of mode-desirable gambles does not satisfy, in general, Axiom D1 ("avoiding partial loss"). In order to overcome this inconvenient, we could have alternatively considered X to be mode-desirable if and only if $A_X^+ \cap (-\infty, 0] = \emptyset$. But this would not entail a substantial improvement, since the set of mode-desirable gambles would no longer satisfy Axiom D2 ("accepting partial gain"). We plan to study other alternatives in order to find a new definition that simultaneously satisfies both axioms.

We also plan to study necessary and sufficient conditions for a credal set $\mathcal{P}$ in order to satisfy the equality $Mo_{\mathcal{P}}(X) = A_X^+$, so that the minimum and the maximum of the set of modes do coincide with the lower and upper previsions induced by the set of mode-desirable gambles.

In the paper, we have assumed that the outcomes of the gambles were numbers, but we could easily extended this framework to a non-necessarily numerical setting. The definitions of mode-desirability and lower and upper prevision would require, anyway, the universe being an ordered set including a "neutral" element that plays the role of the value 0 in the real line.

Acknowledgements

We thank very much three anonymous reviewers for their detailed reports and insightful comments. This work has been jointly supported by the Spanish Ministry of Education and Science and the European Regional Development Fund (FEDER), under projects TIN2011-24302 and TIN2010-20900-C04.

References

[1] J. Abellán and S. Moral, Maximum of Entropy for Credal Sets. International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems 11 (2003) 587-598.

[2] M. de Condorcet, Essay sur le application de l'analyse à la probabilité des decisions rendues à la pluralité des voix, Imprimirie Royale, Paris, 1785.

[3] I. Couso and D. Dubois, An Imprecise Probability Approach to Joint Extensions of Stochastic and Interval Orderings, Lecture Notes in Artificial Intelligence / Communications in Computer and Information Science 299, Part III, p.388-399, Springer, 2012.

[4] I. Couso and S. Moral, Sets of desirable gambles: Conditioning, representation, and precise probabilities, International Journal of Approximate Reasoning, 52 (2011) 1034-1055.

[5] I. Couso, S. Moral and L. Sánchez, The median as the lower prevision with respect to sign-desirability *In preparation*, 2013.

[6] I. Couso and L. Sánchez, The behavioral meaning of the median, in *C. Borgelt et al., Eds., Combining Soft Computing and Statistical Methods in Data Analysis*, Springer, 2010.

[7] H. David, The method of paired comparisons, Griffin's Statistical Monographs & Courses, vol. 12, Charles Griffin & D. Ltd., London, 1963.

[8] T. Denoeux, Extending stochastic ordering to belief functions on the real line. Inf. Sci. 179 (2009) 1362-1376.

[9] S. Destercke, A decision rule for imprecise probabilities based on pair-wise comparison of expectation bounds, in *C. Borgelt et al., Eds., Combining Soft Computing and Statistical Methods in Data Analysis*, Springer, 2010.

[10] J. Hadar, W. Russell, Rules for Ordering Uncertain Prospects, American Economic Review 59 (1969) 25-34.

[11] I. Montes, E. Miranda, S. Montes, Stochastic dominance with imprecise information. Computational Statistics and Data Analysis, doi.org/10.1016/j.csda.2012.07.030 (In press).

[12] L. Sánchez, I. Couso, J. Casillas, Genetic Learning of Fuzzy Rules based on Low Quality Data, Fuzzy Sets and Systems, 160 (2009) 2524-2552.

[13] L.J. Savage. The Foundations of Statistics. Wiley (1954); 2nd edition, Dover Publications Inc., New York, 1972.

[14] B. De Schuymer, H. De Meyer, B. De Baets and S. Jenei, On the cycle-transitivity of the dice model, Theory and Decision 54 (2003) 261-285.

[15] M.C.M. Troffaes, Decision making under uncertainty using imprecise probabilities. Int. J. Approx. Reasoning 45 (2007) 17-29.

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