Possibilistic previsions

Gert de Cooman
Universiteit Gent, SYSTeMS, Vakgroep Elektrische Energietechniek
Technologiepark - Zwijnaarde 9, 9052 Zwijnaarde, Belgium
gert.decooman@rug.ac.be

Abstract

The paper deals with a possibilistic imprecise second-order probability model. It is argued that such models appear naturally in a number of situations. They lead to the introduction of a new type of previsions, called possibilistic previsions, which formally generalise coherent upper and lower previsions. The converse problem is also looked at: given a possibilistic prevision, under what conditions can it be generated by a second-order possibility distribution? This leads to the definition of normality, representability and natural extension of possibilistic previsions. Finally, some attention is paid to the special class of full possibilistic previsions, which can be formally related to Zadeh's fuzzy probabilities. The results have immediate applicability in decision making and statistical reasoning.

1 Introduction

Consider an unstable radioactive nucleus. Its probability of decay in a given time interval t is given by $1 - e^{-\lambda t}$, where the parameter λ is the *decay rate* of the nucleus, a constant which could in principle be determined from the physical properties of the nucleus, such as its composition, excitation, ... In practice, however, decay rates are determined experimentally, within a certain margin for error.

This is a special case of a more general kind of situation which occurs very often, especially when we are dealing with random, or chance, processes: an event (random variable), has a *true* probability (distribution), but there is not enough information to determine it unequivocally. An often used solution to this problem is the specification of a class of probability measures (or distributions), of which the true unknown probability (distribution) is a member. Evidently, the "smaller" the class, the better we know the probability. In the above example, this would amount to specifying an interval of values which we expect the true decay rate to belong to. A number of techniques have been developed for dealing with this kind of situation; see [1, 12, 17] for more details and further references.

In many cases, however, it seems difficult to draw the line between the probability measures or distributions that will be included in the class, and the ones that have to be excluded from consideration. In the case described above, for instance, the precise choice of the lower and upper bounds for the interval of possible decay rates seems to some extent arbitrary. At the same time, not all members of the class will deserve the same status, as it may be felt that central members are more to be trusted than the ones near the border.

Another example should make this clearer. It is an adaptation of the "Miss Julie takes a bet" example, due to Gärdenfors and Sahlin [8]. It runs as a *fil rouge* through the paper. Imagine a coin with a chance (or true probability) θ of landing 'heads'. A friend gives you the following information: she has loaded the coin so heavily that it will almost always fall with the same side upwards, only, she will not tell you which side. So you know that θ is very close to either 0 or 1, but you do not know which. You could model this by specifying a set of possible values for θ of the form

$$A_\alpha = \big[0, \frac{1-\alpha}{2}\big) \cup \big(\frac{1+\alpha}{2}, 1\big],$$

where α is some specific element of $[0, 1]$. The fact that these sets are symmetric with respect to the possible value $\frac{1}{2}$ reflects the symmetry in the available information.

To give another example, your friend might have told you that the chance θ is very close to $\frac{1}{2}$, and you could model this by specifying a set of possible values for θ of the form

$$B_\alpha = \big(\frac{\alpha}{2}, 1 - \frac{\alpha}{2}\big),$$

where, again, α is some specific element of $[0,1]$. As before, the symmetry of these sets with respect to $\frac{1}{2}$ reflects the symmetry in the information.

In either case, the information you have does not allow you to really choose between the different possible values of α. To a certain extent, this choice will be arbitrary. On the other hand, the smaller α, the greater A_α (or B_α), so the more you can be confident that the real θ will belong to A_α (or B_α).

The use of second-order or even higher-order probabilities [9, 10] has been suggested as a solution to this problem. Rather than a class of candidate probability measures (or distributions), a probability measure is given on the class of all probability measures relevant to the problem at hand. In the coin tossing problem described above, this would amount to specifying a probability measure on the set $[0,1]$ of all possible values for θ.

There are a number of problems with this approach, however. First of all, the class of relevant probability measures may be very large, and there may be mathematical problems associated with the specification of reasonable probability measures on such classes. At the same time, it seems a bit strange that our information would allow us to specify a unique probability measure on a large and abstract higher-order space of probability measures, given that it does not allow us to pinpoint a single probability measure on the often much more tangible lower-order spaces. For a much more detailed discussion and critique, see [15, 17, 18].

It seems more reasonable to use an imprecise probability model to describe the uncertainty about the true probability. Let me give a very sketchy outline of how this is done. For more details, see the work of Walley [17], whose behavioural account of imprecise probabilities is no doubt the most complete and encompassing to date. The notations and notions introduced here will be frequently used further on.

Consider the nonempty set Ω of all possible outcomes (or states, or worlds) associated with a given experiment E. A bounded real-valued function on Ω is called a *gamble*. The set $\mathcal{L}(\Omega)$ of all gambles is a linear space under pointwise addition of gambles and scalar multiplication with real numbers. Special gambles are (indicator functions of) events, or subsets of Ω, which only assume the values 0 and 1; and the constant gambles, which will be denoted by the unique value they assume in $\mathbb{R}$.

A positive linear functional P on $\mathcal{L}(\Omega)$ with $P(1)=1$ is called a *linear prevision*. Its restriction to events is a (finitely additive) probability measure. Conversely, a finitely additive probability can always be uniquely extended to a linear prevision on $\mathcal{L}(\Omega)$. The set of all linear previsions on $\mathcal{L}(\Omega)$ is denoted by $\mathbb{P}$.

A *lower prevision* $\underline{P}$ on a set of gambles $\mathcal{K}\subseteq\mathcal{L}(\Omega)$ is a $\mathcal{K}-\mathbb{R}$-mapping. So is an *upper prevision* $\overline{P}$ on $\mathcal{K}$. If $\mathcal{K}$ is a set of events, we speak of lower and upper *probabilities* rather than previsions. With $\underline{P}$ we may associate a set of dominating linear previsions

$$\mathcal{M}(\underline{P})=\{P\in\mathbb{P}\mid(\forall Y\in\mathcal{K})(\underline{P}(Y)\leq P(Y))\}$$

and with $\overline{P}$ a set of dominated linear previsions

$$\mathcal{M}(\overline{P})=\{P\in\mathbb{P}\mid(\forall Y\in\mathcal{K})(P(Y)\leq\overline{P}(Y))\}.$$

$\underline{P}$ *avoids sure loss* iff $\mathcal{M}(\underline{P})\neq\emptyset$. In that case, its *natural extension* $\underline{E}$ is the lower prevision on $\mathcal{L}(\Omega)$ that is the lower envelope of $\mathcal{M}(\underline{P})$:

$$\underline{E}(X)=\inf\{P(X)\mid P\in\mathcal{M}(\underline{P})\},\quad X\in\mathcal{L}(\Omega).$$

$\underline{P}$ is called *coherent* iff it coincides with $\underline{E}$ on its domain $\mathcal{K}$, or equivalently, iff it is the lower envelope of some set of linear previsions. If $\underline{P}$ avoids sure loss, $\underline{E}$ is the smallest coherent lower prevision on $\mathcal{L}(\Omega)$ that dominates $\underline{P}$ on $\mathcal{K}$. Dual definitions and results apply to upper previsions.

The lower prevision of a gamble X has the behavioural interpretation of a supremum buying price for X. Similarly, the upper prevision for X is its infimum selling price. Since buying a gamble X for a price x is the same thing as selling the gamble $-X$ for a price $-x$, we expect lower and previsions to be *conjugate*:

$$\underline{P}(-X)=-\overline{P}(X),\quad X\in\mathcal{K}\cap-\mathcal{K}.$$

A linear prevision is in particular a coherent self-conjugate lower (and upper) prevision.

A lower probability of an event A can be interpreted as the marginally highest acceptable rate for betting on the occurrence of A. An upper probability has a similar definition as one minus the marginally highest acceptable rate for betting against the occurrence of A. Thus, if both A and its set-theoretic complement co$A$ belong to $\mathcal{K}$, we have the following conjugacy relation: $\underline{P}(\text{co}A)=1-\overline{P}(A)$.

In a behavioural context, natural extension is very important: it allows us to coherently extend a coherent lower (or upper) prevision from a set of gambles $\mathcal{K}$ to all gambles in a least-committal way, i.e. with minimal behavioural implications. If a lower (or upper) prevision avoids sure loss but is not coherent, natural extension allows us make it coherent, and again in such a way that the correction carries minimal behavioural implications.

A special type of coherent upper probability is a *normal possibility measure* [5]. It is defined on the set of all events $\wp(\Omega)$ using a special $\Omega - [0,1]$-mapping π, called *possibility distribution*, as follows:

$$\overline{P}_\pi(A) = \sup_{\omega \in A} \pi(\omega), \quad A \subseteq \Omega.$$

π must be *normal* in the sense that $\overline{P}_\pi(\Omega) = \sup_{\omega\in\Omega} \pi(\omega) = 1$. For more details about possibility measures, see [2, 3, 4, 6, 7, 20].

To give an example of how the imprecise probability model can be used in the problem at hand, we return to our coin tossing problem, where the information given is that the chance θ is close to either 0 or 1. As we have seen, the available information leads us to consider the events A_α, $\alpha \in [0,1]$. It provides some evidence for the chance θ belonging to these sets, but no evidence at all that θ belongs to their complements. In other words, the given information seems to warrant our accepting bets *on* the events A_α, or *against* the events $\mathrm{co}A_\alpha$, and at higher rates the smaller α is. But it does not lead us to accept bets against the A_α, or on the $\mathrm{co}A_\alpha$. It therefore seems natural to model the available information by an upper probability assessment $\overline{P}$ on the sets $\mathrm{co}A_\alpha$ for $\alpha \in [0,1]$ – or equivalently, a lower probability assessment on the A_α – in the following way:

$$\overline{P}(\mathrm{co}A_\alpha) = f(\alpha),$$

where f is a continuous, nondecreasing transformation of $[0,1]$ with $f(1) = 1$. Such an assessment is always coherent, and it is – in a specific sense, through natural extension, see [5, 19] for more details – equivalent to an upper probability $\overline{P}$ on the class of all subsets of $[0,1]$ that is a normal possibility measure with distribution

$$\pi(\vartheta) = f(|1-2\vartheta|), \quad \vartheta \in [0,1].$$

In other words, for any $B \subseteq [0,1]$:

$$\overline{P}(B) = \sup_{\vartheta\in B} f(|1-2\vartheta|).$$

This upper probability represents the minimal behavioural implications of the assessments, and therefore seems a good model for the given information.

Similarly, the information that θ is close to $\frac{1}{2}$ leads us to consider the events B_α, $\alpha \in [0,1]$, and it seems to warrant our accepting bets against the events $\mathrm{co}B_\alpha$, and at higher rates the smaller α is. Again, it seems natural to model the available information by upper probability assessments $\overline{P}$ on the events $\mathrm{co}B_\alpha$ for $\alpha \in [0,1]$ in the following way:

$$\overline{P}(\mathrm{co}B_\alpha) = g(\alpha),$$

where g is a continuous, nondecreasing transformation of $[0,1]$ with $g(1) = 1$. Natural extension of this assessment leads to an upper probability $\overline{P}$ that is a normal possibility measure with distribution

$$\pi(\vartheta) = g(2\min\{\vartheta, 1-\vartheta\}), \quad \vartheta \in [0,1].$$

In other words, for any $B \subseteq [0,1]$:

$$\overline{P}(B) = \sup_{\vartheta\in B} g(2\min\{\vartheta, 1-\vartheta\}).$$

What makes these examples interesting is that they – and the reasoning behind them – are not atypical. The crucial point is that the available information allows us to specify in a natural way an upper (or lower) probability assessment on a class of nested sets, and this will fairly often be the case, especially if the given information is vague. A more detailed discussion of these issues can be found in [19]. Given a number of additional continuity assumptions, it can be shown that the natural extension is a possibility measure [5, 19].

For this reason, imprecise second-order probability models, where the second-order upper probability is a normal possibility measure, seem rather important. Interestingly, a thorough study of precisely this type of second-order uncertainty models is the subject of a recent paper by Walley [18]. It describes how the second-order imprecise model can, through natural extension, be converted into a first-order imprecise probability model, which can then be used as prior information for decision making and statistical reasoning, according to the general techniques described in [17].

In the rest of the paper, I discuss a number of interesting aspects of this model that are left untouched in the above-mentioned paper, and that provide additional justifications for working with second-order possibility measures (or distributions).

2 Possibilistic previsions

Assume that there is associated with the experiment E a true (at least finitely additive) probability – or equivalently, linear prevision – Q, but that we have not enough information to identify it. For the present purposes, Q may therefore be interpreted as a random variable in $\mathbb{P}$. Instead, what we have is an second-order imprecise model in the form of a normal possibility distribution π on $\mathbb{P}$.

This model can be given a slightly different form, as follows. Consider a fixed X in $\mathcal{L}(\Omega)$. The possibility distribution π on $\mathbb{P}$ and the coherent associated second-order upper probability $\overline{P}_\pi$ can be interpreted as providing information about the value which the variable Q assumes in $\mathbb{P}$. As Q takes values in $\mathbb{P}$, $Q(X)$ assumes values in $\mathbb{R}$. We may therefore transform the information π into information about which value $Q(X)$ assumes in $\mathbb{R}$ as follows: for any subset A of $\mathbb{R}$, the upper probability that $Q(X)$ assumes a value in A is given by

$$\mathfrak{P}_\pi(X) \cdot A = \overline{P}_\pi(\{P \in \mathbb{P} \mid P(X) \in A\}).$$

$\mathfrak{P}_\pi(X)$ is a normal possibility measure on $\mathbb{R}$, and is therefore completely characterised by its normal possibility distribution $\mathfrak{p}_\pi(X)$, with

$$\mathfrak{p}_\pi(X) \cdot x = \sup\{\pi(P) \mid P(X) = x\}.$$

In summary, starting with a normal second-order possibility distribution π, we have defined a mapping $\mathfrak{p}_\pi$ on $\mathcal{L}(\Omega)$ which maps any gamble X into the normal possibility distribution $\mathfrak{p}_\pi(X)$ on $\mathbb{R}$. This distribution gives us information about the true value of the prevision $Q(X)$ of X. In other words, we may associate with π a special possibilistic prevision $\mathfrak{p}_\pi$ on $\mathcal{L}(\Omega)$ in the sense of the following definition. $(\Omega, \mathcal{L}(\Omega), \mathfrak{p}_\pi)$ is called the *possibilistic prevision induced by* π.

Definition 1 *Let $\mathcal{K}$ be a subset of $\mathcal{L}(\Omega)$. A possibilistic prevision $\mathfrak{p}$ on $\mathcal{K}$ is a mapping from $\mathcal{K}$ to $[0,1]^{\mathbb{R}}$. In order to make clear what its possibility space and domain are, the possibilistic prevision $\mathfrak{p}$ will also be denoted by $(\Omega, \mathcal{K}, \mathfrak{p})$. If $\mathcal{K}$ is in particular a set of events, $\mathfrak{p}$ is called a possibilistic probability.*

It is interesting to note that for every $X \in \mathcal{K}$, $\mathfrak{p}(X)$ is a $\mathbb{R}-[0,1]$-mapping, also called a *fuzzy quantity* [13, 16]. In this purely formal sense, possibilistic previsions and probabilities are related to fuzzy probabilities [21].

What is of interest to us here, is the converse problem. Suppose that for a number of gambles X belonging to, say, a set $\mathcal{K}$, we have information about the value which the actual prevision $Q(X)$ of X assumes in $\mathbb{R}$; and that this can be represented in a natural way by an upper probability $\mathfrak{P}(X)$ that is a normal possibility measure on $\mathbb{R}$, with possibility distribution $\mathfrak{p}(X)$. In other words, we have at our disposal a possibilistic prevision $(\Omega, \mathcal{K}, \mathfrak{p})$. We have seen in the Introduction that this is not infrequently the case. A good example of such a situation is the coin tossing problem discussed in the Introduction, where $\Omega = \{h, t\}$, $\mathcal{K} = \{\{h\}\}$, and h stands for 'heads' and t for 'tails'. In the first case, where the information is that the chance θ is close to 0 or 1, we have

$$\mathfrak{p}(\{h\}) \cdot x = f(|2x-1|)$$

for $x \in [0,1]$ and zero elsewhere; when the information is that the chance θ is close to $\frac{1}{2}$, we have

$$\mathfrak{p}(\{h\}) \cdot x = g(2\min\{x, 1-x\})$$

for $x \in [0,1]$ and zero elsewhere.

The question we then want to answer is whether this possibilistic prevision $(\Omega, \mathcal{K}, \mathfrak{p})$ can be represented by a normal possibility distribution π on $\mathbb{P}$, i.e. whether there exists a possibilistic second-order model π such that $\mathfrak{p} = \mathfrak{p}_\pi$. In that case, the possibilistic prevision $\mathfrak{p}$ will be called *representable*, and π will be called a *representation* of $\mathfrak{p}$.

We will now investigate in more detail what possibilistic previsions are representable, and how, given a representable possibilistic prevision, to find suitable representations. What makes representability so interesting and desirable? It allows us to associate with (assessments in the form of) a possibilistic prevision $\mathfrak{p}$ a second-order possibility distribution. This new model can then in turn be used as a starting point for decision making and statistical reasoning, using the methods described and justified in [18].

With the possibilistic prevision $(\Omega, \mathcal{K}, \mathfrak{p})$, we may associate, for any $\alpha \in]0,1]$, the set of linear previsions

$$\mathcal{M}(\mathfrak{p}_\alpha) = \{P \in \mathbb{P} \mid (\forall Y \in \mathcal{K})(\mathfrak{p}(Y) \cdot P(Y) \geq \alpha)\}.$$

It can be proven that $\mathcal{M}(\mathfrak{p}_\alpha)$ is the *cut set at level* α of the $\mathbb{P}-[0,1]$-mapping $\mathcal{M}(\mathfrak{p})$, defined by

$$\mathcal{M}(\mathfrak{p}) \cdot P = \inf_{Y \in \mathcal{K}} \mathfrak{p}(Y) \cdot P(Y), \quad P \in \mathbb{P},$$

or in other words that

$$\mathcal{M}(\mathfrak{p}_\alpha) = \{P \in \mathbb{P} \mid \mathcal{M}(\mathfrak{p}) \cdot P \geq \alpha\}, \quad \alpha \in]0,1].$$

So we see that with any possibilistic prevision $(\Omega, \mathcal{K}, \mathfrak{p})$ we may associate a special $\mathbb{P}-[0,1]$-mapping $\mathcal{M}(\mathfrak{p})$ that will play a special part in what follows.

Definition 2 *A possibilistic prevision $(\Omega, \mathcal{K}, \mathfrak{p})$ is called* normal *iff $\mathcal{M}(\mathfrak{p})$ is a normal $\mathbb{P}-[0,1]$-mapping.*

Representable possibilistic previsions are always normal: normality is a necessary condition for representability. On the other hand, with any normal $(\Omega, \mathcal{K}, \mathfrak{p})$ we may associate a normal $\mathbb{P}-[0,1]$-mapping $\mathcal{M}(\mathfrak{p})$, and therefore also a representable possibilistic prevision.

Definition 3 *Let $(\Omega, \mathcal{K}, \mathfrak{p})$ be a normal possibilistic prevision. Its natural extension $(\Omega, \mathcal{L}(\Omega), \mathfrak{e})$ is the normal possibilistic prevision induced by $\mathcal{M}(\mathfrak{p})$, i.e. for any X in $\mathcal{L}(\Omega)$ and any x in $\mathbb{R}$:*

$$\mathfrak{e}(X) \cdot x = \sup\{\mathcal{M}(\mathfrak{p}) \cdot P \mid P \in \mathbb{P}, P(X) = x\}.$$

The natural extension $\mathfrak{e}$ of $\mathfrak{p}$ is obviously representable, and has representation $\mathcal{M}(\mathfrak{p})$. Natural extension allows us to associate with (construct from) any normal possibilistic prevision $\mathfrak{p}$, defined on a subset $\mathcal{K}$ of $\mathcal{L}(\Omega)$, a special representable possibilistic prevision defined on all gambles. Just how special this natural extension is, is made clear in the theorem below, which also answers the representability question.

Theorem 4 *Let $(\Omega, \mathcal{K}, \mathfrak{p})$ be a normal possibilistic prevision. Its natural extension $(\Omega, \mathcal{L}(\Omega), \mathfrak{e})$ is the (pointwise) greatest representable possibilistic prevision on $\mathcal{L}(\Omega)$ that is dominated by $\mathfrak{p}$ on $\mathcal{K}$; and $\mathcal{M}(\mathfrak{p}) = \mathcal{M}(\mathfrak{e})$. Moreover, the following statements are equivalent: (i) $(\Omega, \mathcal{K}, \mathfrak{p})$ is representable; (ii) $(\Omega, \mathcal{K}, \mathfrak{p})$ has representation $\mathcal{M}(\mathfrak{p})$; and (iii) $\mathfrak{p}$ coincides on its domain $\mathcal{K}$ with its natural extension $\mathfrak{e}$.*
If $(\Omega, \mathcal{K}, \mathfrak{p})$ is representable, its natural extension $(\Omega, \mathcal{L}(\Omega), \mathfrak{e})$ is the (pointwise) greatest representable possibilistic prevision on $\mathcal{L}(\Omega)$ that coincides with $\mathfrak{p}$ on $\mathcal{K}$.

If $(\Omega, \mathcal{K}, \mathfrak{p})$ is representable, $\mathcal{M}(\mathfrak{p})$ is its greatest representation, or in other words, the possibilistic second-order model that generates $\mathfrak{p}$ and is at the same time least-committal, i.e. has the most conservative behavioural implications. Natural extension can then be seen as a way to extend $\mathfrak{p}$ to $\mathcal{L}(\Omega)$ with minimal behavioural implications. If $\mathfrak{p}$ is not representable, natural extension allows us to correct $\mathfrak{p}$ on its domain $\mathcal{K}$, in such a way that it becomes representable, and this again with minimal behavioural implications.

In our coin tossing problem, the set $\mathbb{P}$ can be parametrised by the probability ϑ that the coin falls 'heads', i.e. $\mathbb{P} = \{P_\vartheta \mid \vartheta \in [0,1]\}$, where, for any gamble X on $\Omega = \{h, t\}$,

$$P_\vartheta(X) = \vartheta X(h) + (1-\vartheta)X(t).$$

Consider the case that the given information is that the chance θ is close to 0 or 1, then

$$\mathcal{M}(\mathfrak{p}) \cdot P_\vartheta = f(|1-2\vartheta|), \quad \vartheta \in [0,1].$$

If $X(t) = X(h)$, we find that

$$\mathfrak{e}(X) \cdot x = \begin{cases} 1 & \text{if } x = X(h) = X(t) \\ 0 & \text{otherwise.} \end{cases}$$

If $X(t) \neq X(h)$, then

$$\mathfrak{e}(X) \cdot x = f\big(\big|\frac{X(h)+X(t)-2x}{X(h)-X(t)}\big|\big)$$

if $\min\{X(t), X(h)\} \leq x \leq \max\{X(t), X(h)\}$ and zero otherwise.

Next, consider the case that the given information is that the chance θ is close to $\frac{1}{2}$. Here

$$\mathcal{M}(\mathfrak{p}) \cdot P_\vartheta = g(2\min\{\vartheta, 1-\vartheta\}), \vartheta \in [0,1].$$

If $X(t) = X(h)$, we find the same expression for $\mathfrak{e}(X)$ as in the first case. If $X(t) \neq X(h)$, then

$$\mathfrak{e}(X) \cdot x = g\big(2\min\big\{\big|\frac{X(h)-x}{X(h)-X(t)}\big|, \big|\frac{X(t)-x}{X(t)-X(h)}\big|\big\}\big)$$

if $\min\{X(t), X(h)\} \leq x \leq \max\{X(t), X(h)\}$ and zero otherwise.

In both cases, $\mathfrak{e}(\{h\}) = \mathfrak{p}(\{h\})$, and $\mathfrak{p}$ is therefore representable. Note also that $\mathfrak{e}(\{t\}) = \mathfrak{p}(\{h\})$, which was to be expected, given the symmetry in the available information.

On the purely formal level, the analogy with Walley's theory of lower and upper previsions [17] should now be clear; possibilistic previsions generalise lower and upper previsions, normality corresponds with avoiding sure loss, and representability with coherence; natural extension does rather the same thing in both models.

3 Full possibilistic previsions

There is a special case which deserves more attention. A normal fuzzy quantity $f\colon \mathbb{R} \to [0,1]$ whose cut sets

$$f_\alpha = \{x \in \mathbb{R} \mid f(x) \geq \alpha\}, \quad \alpha \in \,]0,1],$$

are bounded closed intervals is called a *normal bounded fuzzy closed interval* [13, 16]. Such fuzzy quantities have an important part in the theory of fuzzy numbers; see [14, 16] for a good bibliography and overview. On the other hand, a normal mapping $\pi\colon \mathbb{P} \to [0,1]$ is called *full* iff its cut sets

$$\pi_\alpha = \{P \in \mathbb{P} \mid \pi(P) \geq \alpha\}, \quad \alpha \in \,]0,1],$$

are convex, closed and therefore compact subsets of $\mathbb{P}$, where $\mathbb{P}$ is given the weak* topology, see [11, 17] for more details. Convex closed sets of linear previsions are very important in probability theory, as well as in the theory of imprecise probabilities [17].

It turns out that normal bounded fuzzy closed intervals and full $\mathbb{P} - [0,1]$-mappings are very closely related.

Definition 5 *A representable possibilistic prevision $(\Omega, \mathcal{K}, \mathfrak{p})$ is called full iff its greatest representation $\mathcal{M}(\mathfrak{p})$ is a full $\mathbb{P} - [0,1]$-mapping.*

A representable $\mathfrak{p}$ is therefore full if and only if its cut sets $\mathcal{M}(\mathfrak{p}_\alpha)$, $\alpha \in \,]0,1]$ are convex closed subsets of $\mathbb{P}$. They are moreover all nonempty.

Proposition 6 *A representable $\mathfrak{p}$ is full if and only if its values $\mathfrak{p}(X)$, $X \in \mathcal{K}$, are normal bounded fuzzy closed intervals. Conversely, the natural extension of a normal possibilistic prevision whose values are normal bounded fuzzy closed intervals is full.*

Consider a possibilistic prevision $(\Omega, \mathcal{K}, \mathfrak{p})$ whose values are normal bounded fuzzy closed intervals. Consider $\alpha \in]0,1]$ and $X \in \mathcal{K}$. Then the cut set

$$\mathfrak{p}(X)_\alpha = \{x \in \mathbb{R} \mid \mathfrak{p}(X) \cdot x \geq \alpha\}$$

of $\mathfrak{p}(X)$ is a nonempty bounded closed real interval, that is completely characterised by its smallest element

$$\underline{\mathfrak{p}}_\alpha(X) = \min\{x \in \mathbb{R} \mid \mathfrak{p}(X) \cdot x \geq \alpha\}$$

and its greatest element

$$\overline{\mathfrak{p}}_\alpha(X) = \max\{x \in \mathbb{R} \mid \mathfrak{p}(X) \cdot x \geq \alpha\},$$

since

$$\mathfrak{p}(X)_\alpha = [\underline{\mathfrak{p}}_\alpha(X), \overline{\mathfrak{p}}_\alpha(X)].$$

To put this differently, specifying such a possibilistic prevision $(\Omega, \mathcal{K}, \mathfrak{p})$ is completely equivalent to giving a family of pairs of lower and upper previsions $(\Omega, \mathcal{K}, \underline{\mathfrak{p}}_\alpha)$ and $(\Omega, \mathcal{K}, \overline{\mathfrak{p}}_\alpha)$, $\alpha \in]0,1]$. These previsions are called the *lower* and *upper cut previsions* of $\mathfrak{p}$. It turns out that the cut previsions completely determine not only the possibilistic prevision, but also its representability, and its natural extension.

Theorem 7 *Let $(\Omega, \mathcal{K}, \mathfrak{p})$ be a possibilistic prevision whose values $\mathfrak{p}(X)$, $X \in \mathcal{K}$, are normal bounded fuzzy closed intervals. Then the following statements are equivalent: (i) $\mathfrak{p}$ is full; (ii) $\mathfrak{p}$ is representable; and (iii) the cut previsions $(\underline{\mathfrak{p}}_\alpha, \overline{\mathfrak{p}}_\alpha)$, $\alpha \in]0,1]$, of $\mathfrak{p}$ form coherent pairs of lower and upper previsions.*
Assume that $\mathfrak{p}$ is normal so that its natural extension $(\Omega, \mathcal{L}(\Omega), \mathfrak{e})$ is defined. Then the cut previsions of its natural extension are the natural extension of its pair of cut previsions, or in other words, for any $\alpha \in]0,1]$ and any X in $\mathcal{L}(\Omega)$: $\underline{\mathfrak{e}}_\alpha(X) = \underline{E}_\alpha(X)$ and $\overline{\mathfrak{e}}_\alpha(X) = \overline{E}_\alpha(X)$, where $(\underline{E}_\alpha, \overline{E}_\alpha)$ is the natural extension of the pair $(\underline{\mathfrak{p}}_\alpha, \overline{\mathfrak{p}}_\alpha)$.

Proposition 8 *Let $(\Omega, \mathcal{K}, \mathfrak{p})$ be a possibilistic prevision whose values $\mathfrak{p}(X)$ are normal bounded fuzzy closed intervals that only assume the values 0 and 1. Then (i) $(\Omega, \mathcal{K}, \mathfrak{p})$ is normal if and only if the pair of cut previsions $(\underline{\mathfrak{p}}_1, \overline{\mathfrak{p}}_1)$ avoids sure loss; and (ii) $(\Omega, \mathcal{K}, \mathfrak{p})$ is representable (and therefore full) if and only if the pair of cut previsions $(\underline{\mathfrak{p}}_1, \overline{\mathfrak{p}}_1)$ is coherent.*

The values $\mathfrak{p}(X)$, $X \in \mathcal{K}$, of a full possibilistic prevision are normal bounded fuzzy closed intervals (NBFCI). In the literature [13, 14, 16] operations such as addition and multiplication on NBFCI have been proposed, which amount to doing interval arithmetic on their cut sets. Let $\mathfrak{a}$ and $\mathfrak{b}$ be NBFCI with cut sets $\mathfrak{a}_\alpha = \{x \in \mathbb{R} \mid \mathfrak{a}(x) \geq \alpha\}$ and $\mathfrak{b}_\alpha = \{x \in \mathbb{R} \mid \mathfrak{b}(x) \geq \alpha\}$, for $\alpha \in]0,1]$. These cut sets are bounded closed intervals, and we use the notation $\mathfrak{a}_\alpha = [\mathfrak{a}^\ell_\alpha, \mathfrak{a}^r_\alpha]$. The sum $\mathfrak{a} + \mathfrak{b}$ of $\mathfrak{a}$ and $\mathfrak{b}$ is the NBFCI whose cut sets are the interval sums of the cut sets of $\mathfrak{a}$ and $\mathfrak{b}$:

$$(\mathfrak{a} + \mathfrak{b})_\alpha = \mathfrak{a}_\alpha + \mathfrak{b}_\alpha$$

or in terms of the interval boundaries

$$(\mathfrak{a} + \mathfrak{b})^\ell_\alpha = \mathfrak{a}^\ell_\alpha + \mathfrak{b}^\ell_\alpha \quad \text{and} \quad (\mathfrak{a} + \mathfrak{b})^r_\alpha = \mathfrak{a}^r_\alpha + \mathfrak{b}^r_\alpha.$$

If μ is a nonnegative real number, the product $\mu\mathfrak{a}$ is the NBFCI whose cut sets are given by:

$$(\mu\mathfrak{a})_\alpha = \mu\mathfrak{a}_\alpha$$

or in terms of the interval boundaries

$$(\mu\mathfrak{a})^\ell_\alpha = \mu\mathfrak{a}^\ell_\alpha \quad \text{and} \quad (\mu\mathfrak{a})^r_\alpha = \mu\mathfrak{a}^r_\alpha.$$

If μ is negative, we have instead

$$(\mu\mathfrak{a})^\ell_\alpha = \mu\mathfrak{a}^r_\alpha \quad \text{and} \quad (\mu\mathfrak{a})^r_\alpha = \mu\mathfrak{a}^\ell_\alpha.$$

We also define two types of orderings on NBFCI. The first type, pointwise ordering, is denoted by $\leq$ and defined by $\mathfrak{a} \leq \mathfrak{b}$ iff $\mathfrak{a}(x) \leq \mathfrak{b}(x)$ for all $x \in \mathbb{R}$, or equivalently iff for all $\alpha \in]0,1]$, $\mathfrak{a}_\alpha \subseteq \mathfrak{b}_\alpha$, or alternatively,

$$\mathfrak{b}^\ell_\alpha \leq \mathfrak{a}^\ell_\alpha \quad \text{and} \quad \mathfrak{a}^r_\alpha \leq \mathfrak{b}^r_\alpha.$$

The second type is a straightforward extension of the interval ordering, and is denoted by $\preceq$. We have that $\mathfrak{a} \preceq \mathfrak{b}$ iff for all $\alpha \in]0,1]$,

$$\mathfrak{a}^\ell_\alpha \leq \mathfrak{b}^\ell_\alpha \quad \text{and} \quad \mathfrak{a}^r_\alpha \leq \mathfrak{b}^r_\alpha.$$

Full possibilistic previsions have a number of interesting properties which are very easy to formulate in terms of these orderings and operations on NBFCI. In the following results, we also identify any real number x with the NBFCI that maps x to 1 and any other real number to 0.

Proposition 9 *Let $(\Omega, \mathcal{K}, \mathfrak{p})$ be a full possibilistic prevision. Let X and Y be elements of $\mathcal{K}$ and let $\mu \in \mathbb{R}$. The following properties hold whenever the gambles involved are in $\mathcal{K}$.*
(i) $\inf X \preceq \mathfrak{p}(X) \preceq \sup X$ and in particular $\mathfrak{p}(\mu) = \mu$;
(ii) $\mathfrak{p}(\mu X) = \mu\mathfrak{p}(X)$ and therefore $\mathfrak{p}(-X) = -\mathfrak{p}(X)$;
(iii) $\mathfrak{p}(\mu + X) = \mu + \mathfrak{p}(X)$ and $\mathfrak{p}(\mu - X) = \mu - \mathfrak{p}(X)$;
(iv) $\mathfrak{p}(X + Y) \leq \mathfrak{p}(X) + \mathfrak{p}(Y)$;
(v) $X \geq Y + \mu \Rightarrow \mathfrak{p}(X) \succeq \mathfrak{p}(Y) + \mu$

When the domain $\mathcal{K}$ is a linear subspace of gambles, i.e. closed under linear combinations, it becomes very easy to characterise full possibilistic previsions.

Theorem 10 *Let $\mathcal{K}$ be a linear subspace of $\mathcal{L}(\Omega)$ and let $(\Omega, \mathcal{K}, \mathfrak{p})$ be a possibilistic prevision whose values $\mathfrak{p}(X)$, $X \in \mathcal{K}$, are normal bounded fuzzy closed intervals. Then $\mathfrak{p}$ is representable (and therefore full) if and*

only if it satisfies the axioms:
($\mathfrak{p}0$) $(\forall X \in \mathcal{K})(\mathfrak{p}(-X) = -\mathfrak{p}(X))$;
($\mathfrak{p}1$) $(\forall X \in \mathcal{K})(\mathfrak{p}(X) \preceq \sup X)$;
($\mathfrak{p}2$) $(\forall X \in \mathcal{K})(\forall \lambda > 0)(\mathfrak{p}(\lambda X) = \lambda \mathfrak{p}(X))$;
($\mathfrak{p}3$) $(\forall (X, Y) \in \mathcal{K}^2)(\mathfrak{p}(X+Y) \leq \mathfrak{p}(X) + \mathfrak{p}(Y))$.

A few remarks are in order. First, in Walley's paper about second-order possibilistic uncertainty models [18] it is also – taking into account the various correspondences – the cut previsions $\underline{\mathfrak{p}}_\alpha$ and $\overline{\mathfrak{p}}_\alpha$ and their natural extensions $\underline{\mathfrak{e}}_\alpha$ and $\overline{\mathfrak{e}}_\alpha$, $\alpha \in]0,1]$, that play an important part in turning the second-order model $\mathcal{M}(\mathfrak{p})$ into a first-order model $(\underline{P}, \overline{P})$, through the formulas

$$\underline{P}(X) = \int_0^1 \underline{\mathfrak{e}}_\alpha(X)\mathrm{d}\alpha \quad \text{and} \quad \overline{P}(X) = \int_0^1 \overline{\mathfrak{e}}_\alpha(X)\mathrm{d}\alpha,$$

for $X \in \mathcal{L}(\Omega)$, justified in [18] by the method of natural extension.

To get an idea of how this works, let us go back to the coin tossing example, where the information is that the chance θ of the coin landing heads is close to $\frac{1}{2}$. Recall that we represented this information by a the possibilistic prevision $\mathfrak{p}$ on domain $\mathcal{K} = \{\{h\}\}$ given by $\mathfrak{p}(\{h\}) \cdot x = g(2\min\{x, 1-x\})$ for $x \in [0,1]$ and zero elsewhere. This possibilistic prevision is representable, and has greatest representation $\mathcal{M}(\mathfrak{p})$ given by

$$\mathcal{M}(\mathfrak{p}) \cdot P_\vartheta = g(2\min\{\vartheta, 1-\vartheta\}), \quad \vartheta \in [0,1].$$

Note that $\mathfrak{p}$ and $\mathcal{M}(\mathfrak{p})$ are full. We define the pseudo-inverse (or Galois dual) $G\colon [0,1] \to [0,1]$ of the continuous nondecreasing mapping g as follows:

$$G(\alpha) = \min\{y \in [0,1] \mid g(y) \geq \alpha\}, \quad \alpha \in [0,1].$$

Then we have for any $\alpha \in [0,1]$ and $x \in [0,1]$ that

$$g(x) \geq \alpha \Leftrightarrow G(\alpha) \leq x.$$

The convex closed subsets $\mathcal{M}(\mathfrak{p}_\alpha)$ are determined by

$$P_\vartheta \in \mathcal{M}(\mathfrak{p}_\alpha) \Leftrightarrow \frac{1}{2}G(\alpha) \leq \vartheta \leq 1 - \frac{1}{2}G(\alpha),$$

and this yields for the cut previsions of the natural extension $\mathfrak{e}$ of $\mathfrak{p}$

$$\underline{\mathfrak{e}}_\alpha(X) = \min\{X(h), X(t)\} + \frac{1}{2}G(\alpha)\big|X(h) - X(t)\big|$$

$$\overline{\mathfrak{e}}_\alpha(X) = \max\{X(h), X(t)\} - \frac{1}{2}G(\alpha)\big|X(h) - X(t)\big|$$

where $X \in \mathcal{L}(\Omega)$ and $\alpha \in]0,1]$. The first-order previsions are consequently given by

$$\underline{P}(X) = \min\{X(h), X(t)\} + \frac{1}{2}J(g)\big|X(h) - X(t)\big|$$

$$\overline{P}(X) = \max\{X(h), X(t)\} - \frac{1}{2}J(g)\big|X(h) - X(t)\big|$$

where

$$J(g) = \int_0^1 G(\alpha)\mathrm{d}\alpha = \int_0^1 [1 - g(x)]\mathrm{d}x.$$

Note that $J(g)$ lies between 0 and 1. If g is constant, so $g(x) = g(1) = 1$ for all $x \in [0,1]$, then $J(g) = 0$ and the first-order previsions $\underline{P}$ and $\overline{P}$ are *vacuous* and model complete ignorance about the chance θ:

$$\underline{P}(X) = \min\{X(h), X(t)\}$$

$$\overline{P}(X) = \max\{X(h), X(t)\}.$$

On the other hand, we can choose g in such a way that $J(g)$ comes arbitrarily close to 1 [for instance, let $g(x) = x^\mu$, $x \in [0,1]$, $\mu > 0$ and let $\mu \to \infty$]. If we substitute $J(g) = 1$ in the above expressions for $\underline{P}(X)$ and $\overline{P}(X)$, we get

$$\underline{P}(X) = \overline{P}(X) = P_{\frac{1}{2}}(X) = \frac{1}{2}X(h) + \frac{1}{2}X(t),$$

so the model allows us to come arbitrarily close to letting $\theta = \frac{1}{2}$. The number

$$I(g) = 1 - J(g) = \int_0^1 g(x)\mathrm{d}x$$

is an indication of how imprecise the model is. In this light, I also want to point out that

$$I(g) = \overline{P}(\{h\}) - \underline{P}(\{h\}) = \overline{P}(\{t\}) - \underline{P}(\{t\})$$

and

$$\underline{P}(\{h\}) = \underline{P}(\{t\}) = \frac{1 - I(g)}{2}$$

$$\overline{P}(\{h\}) = \overline{P}(\{t\}) = \frac{1 + I(g)}{2}.$$

This first-order model can for instance be used in a decision problem where the result of any action or decision depends on the outcome of the toss. Let the gambles X_a and X_b on Ω represent the utility functions of two actions (or decisions) a and b. Assume that $\Delta(h) = X_b(h) - X_a(h) > 0$ and $\Delta(t) = X_a(t) - X_b(t) > 0$, so that the decision problem is nontrivial. Our model tells us to choose action b iff $\underline{P}(X_b - X_a) > 0$, or in other words, iff

$$\frac{\Delta(t)}{\Delta(h)} < \frac{1 - I(g)}{1 + I(g)} = \frac{\underline{P}(\{h\})}{\overline{P}(\{h\})},$$

and to choose action a iff $\underline{P}(X_a - X_b) > 0$, that is, iff

$$\frac{\Delta(h)}{\Delta(t)} < \frac{1 - I(g)}{1 + I(g)} = \frac{\underline{P}(\{h\})}{\overline{P}(\{h\})}.$$

There is *indecision*, that is, the model does not tell us which action is best, iff

$$\frac{1 - I(g)}{1 + I(g)} \leq \frac{\Delta(h)}{\Delta(t)} \leq \frac{1 + I(g)}{1 - I(g)},$$

and (at least) one of these inequalities is strict. This 'indecision condition' is more easily satisfied the more imprecise the model is, i.e. the closer $I(g)$ comes to 1. There can only be *indifference*, that is, both actions are considered to be equally good, in the limit $I(g) \to 0$ where the first-order model is precise (Bayesian), when moreover $\Delta(t) = \Delta(h)$.

The second remark deals with the implications of Theorem 7 and Proposition 8. These results tell us that the special case of full possibilistic previsions makes the formal relationship between this possibilistic model and Walley's lower and upper previsions [17] even clearer and more direct. They seem to open up a path that leads towards a general behavioural account of unreliable [8], or fuzzy [21] probabilities. An exploration of this path is the subject of current research.

Acknowledgements

This paper presents research results 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 author. Gert de Cooman is a Postdoctoral Fellow of the Fund for Scientific Research (FWO) – Flanders (Belgium). He would like to thank the FWO for funding the research reported on in this paper. He has also benefited from discussions with Peter Walley, George Klir, Dirk Aeyels and Leen Gilbert. Of course, any opinions expressed in this paper are the author's, unless explicitly stated otherwise.

References

[1] J. O. Berger. *Statistical Decision Theory and Bayesian Analysis*. Springer, New York, 1985.

[2] G. de Cooman. Possibility theory I: The measure- and integral-theoretic groundwork. *International Journal of General Systems*, 25:291–323, 1997.

[3] G. de Cooman. Possibility theory II: Conditional possibility. *International Journal of General Systems*, 25:325–351, 1997.

[4] G. de Cooman. Possibility theory III: Possibilistic independence. *International Journal of General Systems*, 25:353–371, 1997.

[5] G. de Cooman and D. Aeyels. Supremum preserving upper probabilities. submitted for publication.

[6] G. de Cooman, D. Ruan, and E. E. Kerre, editors. *Foundations and Applications of Possibility Theory – Proceedings of FAPT '95, Ghent, 13–15 December 1995*, Singapore, 1995. World Scientific.

[7] D. Dubois and H. Prade. *Possibility Theory*. Plenum Press, New York, 1988.

[8] P. Gärdenfors and N.-E. Sahlin. Unreliable probabilities, risk taking, and decision making. *Synthese*, 53:361–386, 1982.

[9] I. J. Good. Subjective probability as the measure of a non-measurable set. In E. Nagel, P. Suppes, and A. Tarski, editors, *Logic, Methodology and Philosophy of Science*, pages 319–329. Stanford University Press, Stanford, 1962.

[10] I. J. Good. Some history of the hierarchical Bayesian methodology. In J. M. Bernardo, M. H. DeGroot, D. V. Lindley, and A. F. M. Smith, editors, *Bayesian Statistics 1*, pages 489–519. Valencia University Press, Valencia, 1980.

[11] R. B. Holmes. *Geometric Functional Analysis and its Applications*. Springer, New York, 1975.

[12] P. J. Huber. *Robust Statistics*. John Wiley & Sons, New York, 1981.

[13] E. E. Kerre. Basic principles of fuzzy set theory for the representation and manipulation of imprecision and uncertainty. In E. E. Kerre, editor, *Introduction to the Basic Principles of Fuzzy Set Theory and Some of Its Applications*, pages 1–158. Communication & Cognition, Ghent, 1991.

[14] G. J. Klir. The role of constrained fuzzy arithmetic in engineering. In B. M. Ayyub, editor, *Uncertainty Analysis in Engineering and the Sciences*. Kluwer, Boston, 1997. in production.

[15] H. E. Kyburg. Higher order probabilities and intervals. *International Journal of Approximate Reasoning*, 2:195–209, 1988.

[16] M. Mareš. *Computation over Fuzzy Quantities*. CRC Press, Boca Raton, FL, 1994.

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

[18] P. Walley. Statistical inferences based on a second-order possibility distribution. *International Journal of General Systems*, 26, 1997.

[19] P. Walley and G. de Cooman. A behavioural model for linguistic uncertainty. submitted for publication.

[20] L. A. Zadeh. Fuzzy sets as a basis for a theory of possibility. *Fuzzy Sets and Systems*, 1:3–28, 1978.

[21] L. A. Zadeh. Fuzzy probabilities. *Information Processing and Management*, 20:363–372, 1984.