Extreme probability distributions of random/fuzzy sets and p-boxes

A. Bernardini
Dpt di Costruzioni e Trasporti, Università degli Studi di Padova, alberto.bernardini@unipd.it

F. Tonon
Dpt of Civil Engineering, University of Texas, tonon@mail.utexas.edu

Abstract: Uncertain information about a system variable described by a random set or an equivalent Dempster-Shafer structure on a finite space of singletons determines an infinite convex set of probability distributions, given by the convex hull of a finite set of extreme distributions. Extreme distributions allow one to evaluate (through the Choquet integral) exact upper/lower bounds of the expectation of monotonic and non-monotonic functions of uncertain variables, for example in reliability evaluation of engineering systems. The paper considers the simple case of a single variable, and details applications to random sets with nested focal elements (consonant random sets or the equivalent fuzzy set) and to p-boxes. A simple direct procedure to derive extreme distributions from a p-box is described through simple numerical examples.

Keywords: random sets, fuzzy sets, p-boxes

1 Introduction

In Civil Engineering practice, the growing need for rationally including uncertainty in engineering modelling and calculations is witnessed by the adoption of reliability-based EuroCodes or Load and Resistance Factor Design codes (Level I). More sophisticated reliability-based approaches are used in research or special practical problems (Levels II and III). This need, however, has been accompanied by the realization of the limitations that affect probabilistic modelling of uncertainty when dealing with imprecise data (Walley. 1991).

On one hand, in the enlarged ambit of a multi-valued logic, alternative models of uncertainty have been propounded that attempt to capture qualitative or ambiguous aspects of engineering models. Particularly important models are based on the idea of fuzzy sets and relations, and positive applications have been reported in the fields of automatic controls in robotics and artificial intelligence, more generally in the field of optimal decisions and approximate reasoning. Less convincing and frequently charged with leading to unrealistic or unverifiable conclusions are the tentative applications of fuzzy models in predicting or simulating objective phenomena, for example to evaluate the reliability of an engineering design or to assess the reliability of an existing engineering system.

On the other, new models of uncertainty have been formulated, based on a generalisation of the probabilistic paradigm, and in particular its objective interpretation as relative frequency of events. The main point is the considerations of “imprecise probabilities” of events or “imprecise previsions” of functions, based on the idea of bounded sets of probability distributions compatible with the available information or, alternatively, on the combination of a probability distribution (randomness) with imprecise events (set uncertainty). Because these models retain the semantics of probability theory, comparisons with probability theory are straightforward.

The subjectivist formulation of this approach (Theory of evidence, (Shafer. 1976)) is compatible with a different interpretation based on statistics of objective but imprecise events (Theory of random sets). When imprecise events are nested, it includes the notion of fuzzy set as a particular case.

After a quick review of the definitions and properties of imprecise probabilities and classification of the corresponding upper/lower bounds according to the order of Choquet capacities, the paper focuses on the theory of random sets, with particular emphasis on fuzzy sets (consonant random sets) and p-boxes (non consonant random sets that contain, as a particular case, the ordinary probability distributions). Both fuzzy sets and p-boxes are indexable-type random set, i.e the imprecise events can be ordered and univocally determined by an index varying from 0 to 1 (Alvarez. 2006). This property is very useful in applications involving numerical simulations.

With reference to a finite probability space for a single variable, the paper continues by discussing the properties of the infinite convex set of probability distributions, and of the finite set of extreme distributions generated by random sets, fuzzy sets, and p-boxes. The finite sets of extreme distributions are particularly useful in evaluating exact expectation bounds for a real-valued function of the considered variable, in the case of both monotonic and not monotonic functions.

A simple and direct procedure to derive extreme distributions from a p-box is described.

2 Imprecise probabilities and convex sets of probability distributions

2.1 Coherent upper and lower probabilities and previsions

Le us consider a finite probability space (Ω, $\mathcal{F}$, P), where $\mathcal{F}$ is the σ-algebra generated by a finite partition of Ω into elementary events (or singletons) $S = \{s_1, s_2\ldots, s_j ,\ldots s_n\}$. Hence the probability space is fully specified by the probabilities $P(s_j)$, which sum up to 1 (in the following: the “probability distribution”).

Imprecise probabilities arise when the available information does not allow one to uniquely determine a unique probability distribution. In this case, the information could be given by means of upper and/or lower probabilities, $\mu_{LOW}(T_i)$, $\mu_{UPP}(T_i)$, of some events $T_i \in \mathcal{F}$, or directly through a set of probability distributions, Ψ.

The foundation of a theory of imprecise probabilities is mainly due to the work of Peter Walley in the 1980s/90s on a new theory of probabilistic reasoning, statistical inference and decision, under uncertainty, partial information or ignorance ((Walley. 1991), or for a concise introduction (Walley. 2000)). In his work, the basic idea of upper/lower probabilities is enlarged to the more general concept of upper/lower *previsions* for a family of bounded and point-valued functions f_i: $S \rightarrow Y = \Re$. For a specific precise probability distribution $P(s_j)$, the prevision is equivalent to the linear expectation:

$$E_P[f_i] = \sum_{s_j \in S} f_i(s_j) P(s_j) \tag{2.1}$$

Since the probability of an event T_i is equal to the expectation of its indicator function (equal to: 1 if $s_j \in T_i$, 0 if $s_j \notin T_i$), upper/lower previsions generalize and hold upper/lower probabilities as a particular case.

Let us now focus on the information about the space of events in S given by upper and/or lower previsions, $E_{LOW}[f_i]$ and $E_{UPP}[f_i]$, for a family of bounded and point-valued functions f_i, $\mathcal{K}$. This is accomplished by the set, Ψ^E, of probability distributions $P(s_j)$ compatible with $E_{LOW}[f_i]$ and $E_{UPP}[f_i]$:

$$\Psi^E = \{ P : E_{LOW}[f_i] \leq E_P[f_i] \leq E_{UPP}[f_i] \quad \forall f_i \in \mathcal{K} \} \tag{2.2}$$

Ψ^E is convex and closed. One is interested in checking two basic conditions of the suggested bounds:

1. A preliminary, strong condition requires that set Ψ^E should be non-empty. If set Ψ^E is empty, it means there is something basically irrational in the suggested bounds. For example, the set Ψ^E is empty if $E_{LOW}[f_i] > \max_j f_i(s_j)$ or $E_{UPP}[f_i] < \min_j f_i(s_j)$ (for upper/lower probabilities: $\mu_{LOW}(T_i) > 1$ or $\mu_{UPP}(T_i) < 0$). In the behavioural interpretation adopted by Walley, the functions f_i are called *gambles*, and this basic condition is said to *avoid sure loss*.
2. A second, weaker but reasonable condition requires that the given bounds should be the same as the *naturally extended* expectation bounds that can be derived from Ψ^E (*coherence* according to Walley's nomenclature)

$$\begin{aligned} E_{LOW,c}[f_i] &= \min_{P \in \Psi^E} E_P[f_i] \\ E_{UPP,c}[f_i] &= \max_{P \in \Psi^E} E_P[f_i] \end{aligned} \tag{2.3}$$

In this case, one says that $E_{LOW}[f_i]$ and $E_{UPP}[f_i]$ are (lower and upper, respectively) envelopes of Ψ^E.

If the given bounds are not coherent, i.e. envelopes to Ψ^E, because they do not satisfy Eq. (2.3), the given bounds can be restricted without changing the probabilistic content of the original information, i.e. set Ψ^E. These restricted bounds, calculated by using Eq. (2.3), are called "*natural extension*" of the given bounds $E_{LOW}[f_i]$ and $E_{UPP}[f_i]$. For example if bounds are given for both function f_i and the opposite $-f_i$ coherence requires the "duality condition": $E_{UPP}[f_i] = - E_{LOW}[-f_i]$ (equivalently for upper/lower probabilities of complementary sets T_i and T_i^c : $\mu_{UPP}(T_i) = 1- \mu_{LOW}(T_i^c)$).

The applications that follow are restricted to the special case when $\mathcal{K}$ is a set of indicator functions, i.e. previsions coincide with probabilities. In this special case, there is no one-to-one correspondence between imprecise probabilities and closed convex sets of probability distributions because several closed convex sets of probability distributions could give the same imprecise probabilities. This one-to-one correspondence only holds between previsions and convex sets of probability distributions when $\mathcal{K}$ is the set of *all* functions. In other terms, imprecise probabilities are less informative than previsions.

2.2 CHOQUET CAPACITIES AND ALTERNATE CHOQUET CAPACITIES

An important criterion for classifying monotonic (with respect to inclusion) measures of sets was introduced by Choquet in his theory of *capacities* (Choquet. 1954). Given a finite set S, let $\mathcal{P}(S)$ be the power set (set of all subsets) of S. A regular monotone set function μ: $\mathcal{P}(S) \to [0, 1] \mid \mu(\emptyset) = 0$, $\mu(S)= 1$ is called 2-monotone (or a *Choquet Capacity of order k* = 2) if, given two subsets T_1 and T_2:

$$\mu(T_1 \cup T_2) \geq \mu(T_1) + \mu(T_2) - \mu(T_1 \cap T_2) \qquad (2.4)$$

The dual coherent upper probabilities ($\mu_{UPP}(T_i) = 1- \mu(T_i^c)$) are called *Alternate Choquet Capacity of order k* = 2, and satisfy the relation:

$$\mu_{UPP}(T_1 \cap T_2) \leq \mu_{UPP}(T_1) + \mu_{UPP}(T_2) - \mu_{UPP}(T_1 \cup T_2) \qquad (2.5)$$

More generally, monotone dual set functions (μ, μ_{UPP}) are k-monotone (*Choquet Capacity of order k*), and, respectively, *Alternate Choquet Capacity of order k*, if, given k subsets T_1, T_2 ….T_k:

$$\mu(T_1 \cup T_2 ... \cup T_k) \geq \sum_{\varnothing \subset K \subseteq \{1,2..k\}} (-1)^{|K|+1} \mu\left(\bigcap_{i \in K} T_i \right) \qquad (2.6)$$

$$\mu_{UPP}(T_1 \cap T_2 ... \cap T_k) \leq \sum_{\varnothing \subset K \subseteq \{1,2..k\}} (-1)^{|K|+1} \mu_{UPP}\left(\bigcup_{i \in K} T_i \right)$$

Precise probability distributions are both an Choquet Capacity and an Alternate Choquet Capacity of order $k = \infty$, that satisfy relations (2.5) and (2.6) as equalities.

Choquet and dual Alternating Choquet capacities of order $k > 1$ are coherent lower and upper probabilities respectively. Indeed, compare the above properties with the necessary conditions for coherent upper/lower probabilities:

- Monotonicity with inclusion: $T_1 \subseteq T_2 \Rightarrow \mu_{LOW}(T_1) \leq \mu_{LOW}(T_2)$; $\mu_{UPP}(T_1) \leq \mu_{UPP}(T_2)$
- Super-additivity of μ_{LOW} for disjoint sets ($T_1 \cap T_2 = \varnothing$): $\mu_{LOW}(T_1 \cup T_2) \geq \mu_{LOW}(T_1) + \mu_{LOW}(T_2)$
- Sub-additivity of μ_{UPP} for any pair of sets T_1, T_2: $\mu_{UPP}(T_1 \cup T_2) \geq \mu_{UPP}(T_1) + \mu_{UPP}(T_2)$.

Therefore, coherent super-additive lower probabilities are not necessarily Choquet capacities of order $k > 1$.

There is a strong connection between the order k and the *Möbius transform* of the set function $\mu(T)$:

$$^{\mu}m(A) = \sum (-1)^{|A-T|} \mu(T) \,|\, T \subseteq A \qquad (2.7)$$

The Möbius transform of a set function μ is a one-to-one invertible set function $^{\mu}m$: $\mathcal{P}(S) \rightarrow \Re$, and its inverse is precisely :

$$^{m}\mu(\mathrm{T}) = \sum m(A) \,|\, A \subseteq \mathrm{T}, \qquad \forall \mathrm{T} \subset \mathrm{S}; \qquad (2.8)$$

For the purposes of this study, the most interesting properties (see for example (Chateauneuf and Jaffray. 1989, Klir. 2005) are the following:

1- a set function μ is monotone if and only if:

$$m(\varnothing) = 0; \sum_{T \in \mathcal{P}(S)} {}^{\mu}m(T) = 1; \quad \forall T \in \mathcal{P}(S) : \sum_{A \subseteq T} {}^{\mu}m(A) \geq 0 \qquad (2.9)$$

and, therefore, $\forall j$: $^{\mu}m(\{s_j\}) \geq 0$.

2- If $\mu(T)$ is k-monotone and $|T|\leq k$ then ${}^{\mu}m(T)\geq 0$

3- $\mu(T)$ is ∞-monotone if and only if: $\forall\ T\in\mathcal{P}(S)$: ${}^{\mu}m(T)\geq 0$.

2.3 EXTREME DISTRIBUTIONS

For a given regular monotone set function μ, a permutation $\pi(j)$ of the indexes of the singletons in the set $S = \{s_1, s_2\ldots, s_j ,\ldots s_n\}$ defines the following probability distribution:

$$
\begin{aligned}
&P\left(s_{\pi(j)=1}\right)=\mu\left(\left\{s_{\pi(\mathrm{j})=1}\right\}\right)\\
&P\left(s_{\pi(j)=k>1}\right)=\mu\left(\left\{s_{\pi(\mathrm{j})=1},\ldots s_k\right\}\right)-\mu\left(\left\{s_{\pi(\mathrm{j})=1},\ldots s_{k-1}\right\}\right)
\end{aligned}
\tag{2.10}
$$

The |S|! possible permutations define a finite set of probability distributions, *EXT*, together with its convex hull, Ψ^{EXT}.

If the same permutation is applied to a pair (μ_{LOW}, μ_{UPP}) of dual regular monotone set functions, a pair of dual distinct probability distributions is generated, but (μ_{LOW}, μ_{UPP}) always generate the same set *EXT*.

Now, one would wonder what the relationship is between Ψ^{EXT} and the set Ψ^{μ} calculated for (μ_{LOW}, μ_{UPP}) by using (Eq. 2.2). It turns out that the two sets could be different, and satisfy the inclusion: $\Psi^{E}\subseteq\Psi^{EXT}$. Precisely:

- For coherent monotone measures (k = 1), Eq. (2.10) could generate probability distributions in *EXT* that do not satisfy the bounds in (Eq. 2.2); hence Ψ^{μ} could be strongly included in Ψ^{EXT};

- for monotone measures with $k > 1$, all probability distributions in *EXT* (and in Ψ^{EXT}) satisfy the bounds in (Eq. 2.2), and thus $\Psi^{EXT} = \Psi^{\mu}$; *EXT* coincides with the set of the extreme points (or the *profile*) of the closed convex set Ψ^{μ} .

2.4 EXPECTATION BOUNDS AND CHOQUET INTEGRALS FOR REAL VALUED FUNCTIONS

When the sets Ψ^{μ} or Ψ^{EXT} are known, or when a generic set Ψ is assigned, the upper and lower expectation bounds for any real function f: $S\rightarrow\ Y=\ \Re$ could be calculated by solving the optimization problems in Eqs (2.3) by substituting Ψ^{μ}, Ψ^{EXT}, or Ψ for Ψ^{E}, respectively. However, The expectation of a point valued function f: $S\rightarrow Y=[y_L, y_R] \subset \Re$ with CDF $F(y)$ can be calculated as follows by using the Stieltjes Integral and equivalent expressions:

$$E[y=f]=\int_{y_L}^{y_R} f\cdot dF=\left[yF\right]_{y_L}^{y_R}-\int_{y_L}^{y_R}Fdy=y_R-\int_{y_L}^{y_R}Fdy=y_L+\int_{y_L}^{y_R}(1-F)dy=$$
$$=y_L+\int_{y_L}^{y_R}P(f>\alpha)d\alpha=y_L+\int_{y_L}^{y_R}P\left({}^{\alpha}T=\left\{s\in S\mid f(s)>\alpha\right\}\right)d\alpha \tag{2.11}$$

The Choquet Integral is the direct extension of the last functional expression to a monotonic measure μ, for the ordered family of subsets ${}^{\alpha}T$, which depend on the selected function f:

$$C(f,\mu)=y_L+\int_{y_L}^{y_R}\mu({}^{\alpha}T)d\alpha \tag{2.12}$$

Indeed, the Choquet integral gives a numerical value that coincides with the expectation of the function f for a particular probability distribution. The latter distribution is obtained by the permutation leading to a monotonic (decreasing) ordering of the function values.

The expectation bounds are therefore obtained through the dual probability distributions obtained by applying Eq. (2.11) to the dual upper/lower probabilities (μ_{LOW}, μ_{UPP}). The Choquet integral determines optimal bounds with respect to the set EXT (or Ψ^{EXT}) defined in Section 2.3: hence, for general monotone measures (k = 1), it can give larger bounds than the correct bounds calculated by using the extreme points of Ψ^{μ}; on the other hand, for k > 1, the Choquet integral gives exact expectation bounds.

3 Random sets

3.1 General properties of random sets

Among the different definitions of random set, we refer here to the formalism of the Theory of Evidence, but with no particular limitation to the subjectivist emphasis of this theory. The original information is described by a family of pairs of nonempty subsets A^i ("*focal elements*") and attached $m^i = m(A^i) > 0$, i=1, 2, …n (“*probabilistic assignment*”), with the condition that the sum of m^i is equal to 1. The (total) probability of any subset T of S can therefore be bounded by means of the additivity rule. Shafer suggested the words *Belief* (*Bel*) and *Plausibility* (*Pla*) for the lower and upper bounds, respectively. Formally:

$$\begin{aligned} \forall T \subset S : \mu_{UPP}\left(T\right) = Pla(T) = \sum_i m^i \mid A^i \cap T \neq \varnothing, \\ \mu_{LOW}\left(T\right) = Bel(T) = \sum_i m^i \mid A^i \subseteq T \end{aligned} \tag{3.1}$$

Comparison with Eq. (2.8) demonstrates that *Bel* is the inverse Möbius transform of the non-negative set function m: hence *Bel* is a ∞-monotone set function, and *Pla* an Alternate Choquet capacity of order $k = \infty$. As explained in Section 2.3, Ψ^{Bel} (calculated with Eq. 2.2 for *Bel*) coincides with the set Ψ^{EXT}, where EXT (calculated with Eq. 2.11) is the set of extreme distributions that can be used to evaluate exact expectation bounds for a function of interest.

3.2 FUZZY SETS

The conclusions in Section 3.1 also apply in the particular case of a *consonant* random set; i.e. when focal elements are *nested*, and hence can be ordered in such a way that:

$$A^1 \subseteq A^2 \subseteq .... \subseteq A^n \tag{3.2}$$

Consonant random sets satisfy the relation:

$$Pla\left(T_1 \cup T_2\right) = \max\left(Pla(T_1), Pla(T_2)\right) \tag{3.3}$$

and hence (similar to classical Probability measures) they satisfy the following “*decomposability property*”: the measure of uncertainty of the union of any pair of disjointed sets depends solely on the measures of the individual sets. Therefore, in the case of a consonant random set, the point-valued *contour function* (Shafer. 1976) μ: $S \rightarrow [0, 1]$:

$$\mu(s_j) = Pla\left(\left\{s_j\right\}\right) \tag{3.4}$$

completely defines the information on the measures of any subset $T \subset S$, exactly in the same way as the probability distribution $P(s_j)$ defines, although through a different rule (the additivity rule), the probability of every subset T in the algebra generated by the singletons. Indeed:

$$Pla(T) = \max_{s_j \in T} \mu(s_j); \qquad Bel(T) = 1 - \max_{s_j \in T^c} \mu(s_j) \tag{3.5}$$

Moreover, the Möbius inversion (2.7) of the set function *Bel* allows the (nested) family of focal elements to be determined through the set function m.

More directly, let us assume:

$$\begin{aligned}
&\alpha_1 = \max_{\mathrm{j}} \left(\mu(s_j)\right) = 1 \\
&\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\
&\alpha_i = \max_{\mathrm{j}|\mu(s_j)<\alpha_{i-1}} \left(\mu(s_j)\right); \\
&\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\
&\alpha_n = \max_{\mathrm{j}|\mu(s_j)<\alpha_{n-1}} \left(\mu(s_j)\right) = \min_{\mathrm{j}} \left(\mu(s_j)\right); \\
&\alpha_{n+1} = 0
\end{aligned} \tag{3.6}$$

The family of focal elements and related probabilistic assignments (summing up to 1) are given by:

$$A^i = \left\{ s_j \in S \mid \mu(s_j) \geq \alpha_i \right\}; \qquad m^i = \alpha_i - \alpha_{i+1} \tag{3.7}$$

The number of focal elements, n, is therefore equal to the cardinality of the range of S through μ; of course this cardinality is less then or equal to $|S|$, because some singletons could map onto the same value of plausibility.

There is a narrow correspondence between consonant random sets and other decomposable measures of uncertainty: fuzzy sets and possibility distributions. This connection can clearly be envisaged using the dual representation of a fuzzy set through their α-cuts ${}^{\alpha}A$. They are classical subsets of S defined, for any selected value of membership α, by the formula:

$${}^{\alpha}A = \left\{ s \in S \mid \mu(s) \geq \alpha \right\} \tag{3.8}$$

When a fuzzy set is implicitly given through the (finite or infinite) sequence of its α-cuts $^{\alpha}A$, its membership function can be reconstructed through the equation:

$$\mu\left(s_j\right) = \max_{\alpha} \quad \min\left(\alpha, \chi_{^{\alpha}A}(s)\right) \tag{3.9}$$

where $\chi_{^{\alpha}A}(s)$ is the indicator function of the classical subset $^{\alpha}A$.

By comparing Eq. (3.8) with Eq. (3.7), it is clear that the α-cuts $^{\alpha}A$ of any given normal fuzzy set are a nested sequence of subsets of set S, and therefore the family of focal elements of an associated consonant random set: the membership function of normal fuzzy sets gives the contour function of the corresponding random sets, and the basic probabilistic assignment (for a finite sequence of α-cuts) is given by $m(A^i = {}^{\alpha_i}A) = \alpha_i - \alpha_{i+1}$.

By considering Eq. (3.5) from this point of view, the membership function of a fuzzy subset A allows measures of Plausibility and Belief to be attached to every classical subset $T \subseteq S$; this very different interpretation of a fuzzy set was recognized by Zadeh himself in 1978 (Zadeh. 1978), as the basis of a theory of *Possibilities* defined by a *possibility* distribution numerically equal to $\mu_A(s)$, and later extensively developed by other authors, in particular Dubois and Prade (Dubois and Prade. 1988).

This comparison suggests a probabilistic (objective or subjective) content of the information summarized by a fuzzy set and allows one to evaluate by means of the set *EXT* exact expectation bounds for real functions of a fuzzy variable. Although the discussion was restricted to finite discrete variables, the conclusion can be extended to continuous variables.

Example 3—1. Consider $S = \{s_1, s_2, s_3, s_4\}$, and the point-valued function $f(s_j)$ mapping to the set $Y = \{5, 20, 10, 0\}$. The fuzzy set of S is measured by the set of membership values (0, 0.1, 1, 0.1). Eqs. (3.6) give: $\alpha_1 = 1$; $\alpha_2 = 0.1$; $\alpha_3 = 0$. The associated consonant random set is defined by the set of pairs $\{ (A^1 = \{s_3\}, m^1 = 1 - 0.1 = 0.9), (A^2 = \{ s_2, s_3, s_4\}, m^2 = 0.1 - 0 = 0.1)\}$.
The permutation leading to a monotonic decreasing ordering of the function $f(s_j)$ is the following: $(\pi(s_2) = 1, \pi(s_3) = 2, \pi(s_1) = 3, \pi(s_4) = 4)$. Table 3—1 shows the corresponding dual extreme distribution according to Eq. (2.10) and the dual set functions *Pla* and *Bel*.

Table 3—1. Dual extreme distributions for Example 3—1

T	$Pla(T)$	$P_{EXT,UPP}(s)$	T^c	$Bel(T) =$ 1- $Pla(T^c)$	$P_{EXT,LOW}(s)$
$T_1 = \{s_2\}$	0.1	$P(s_2) = Pla(T_1) = 0.1$	$\{s_1, s_3, s_4\}$	0	$P(s_2) = Bel(T_1) = 0$
$T_2 = \{s_2, s_3\}$	1	$P(s_3) = Pla(T_2) - Pla(T_1) = 0.9$	$\{s_1, s_4\}$	0.9	$P(s_3) = Bel(T_2) - Bel(T_1) = 0.9$

$T_3 = \{s_2, s_3, s_1\}$	1	$P(s_1)= Pla(T_3) - Pla(T_2) = 0$		0.9	$P(s_1)= Bel(T_3) - Bel(T_2)= 0$
$T_4 = S$	1	$P(s_4)= Pla(T_4) - Pla(T_3) = 0$	$\varnothing$	1	$P(s_4)= Bel(T_4) - Bel(T_3)=0.1$

Hence:

$E_{UPP}[f] = E_{P_{EXT,UPP}}[f] = 20\text{x } 0.1 + 10 \text{ x } 0.9 = 11$; $E_{LOW}[f] = E_{P_{EXT,LOW}}[f] = 10\text{x } 0.9 + 0 \text{ x } 0.1 = 9$.

The same results can be obtained through the Choquet integral (Eq. (2.12)). For example:

$C(f, \mu_{LOW} = Bel) = 0 + Bel(\{s_2, s_3, s_1\}) \text{ x } (\Delta\alpha = f(s_1) - f(s_4)) + Bel(\{s_2, s_3\}) \text{ x } (\Delta\alpha = f(s_3) - f(s_1)) + Bel(\{s_2\}) \text{ x } (\Delta\alpha = f(s_2) - f(s_3)) = 0 + 0.9 \text{ x } (5-0) + 0.9 \text{ x } (10 - 5) + 0 \text{ x } (20 - 10) = 9$

$C(f, \mu_{UPP} = Pla) = 0 + Pla(\{s_2, s_3, s_1\}) \text{ x } (\Delta\alpha = f(s_1) - f(s_4)) + Pla(\{s_2, s_3\}) \text{ x } (\Delta\alpha = f(s_3) - f(s_1)) + Pla(\{s_2\}) \text{ x } (\Delta\alpha = f(s_2) - f(s_3)) = 0 + 1 \text{ x } (5-0) + 1 \text{ x } (10 - 5) + 0.1 \text{ x } (20 - 10) = 11$

3.3 P-BOXES

Given a finite space S, a set Ψ^F of probability distributions is implicitly defined by lower and upper bounds, $F_{LOW}(s_j)$ and $F_{UPP}(s_j)$, of the cumulative distribution functions $F(s_j)$:

$$\Psi^F = \left\{ P : F_{LOW}(s_j) \le F(s_j) = P\left(\left\{s_1, \ldots, s_j\right\}\right) \le F_{UPP}(s_j),\ j = 1 \text{ to } |S| \right\} \quad (3.10)$$

The set Ψ^F is non-empty if $F_{LOW}(s_k) \le F_{UPP}(s_j)$ for any $k \le j$.

However, coherence clearly requires stronger conditions: the bounds $F_{LOW}(s_j)$ and $F_{UPP}(s_j)$ should be non-negative, non-decreasing in j, and both must be equal to 1 for $j = |S|$ (Walley. 1991, § 4.6.6).

Explicit evaluation of set Ψ^F can be obtained by solving the constraints (3.10) for the probabilities of the singletons $P(s_j)$:

$$\begin{aligned}
&F_{LOW}(s_1) \le P(s_1) \le F_{UPP}(s_1); \qquad P(s_1) \ge 0 \\
&F_{LOW}(s_2) \le P(s_1) + P(s_2) \le F_{UPP}(s_2); \quad P(s_2) \ge 0 \\
&\ldots\ldots\ldots\ldots\ldots\ldots \\
&F_{LOW}(s_j) \le P(s_j) + \sum_{i=1}^{j-1} P(s_i) \le F_{UPP}(s_j); P(s_j) \ge 0 \\
&\ldots\ldots\ldots\ldots\ldots\ldots \\
&P(s_{j=|S|}) + \sum_{i=1}^{|S|-1} P(s_i) = 1; \quad P(s_{j=|S|}) \ge 0
\end{aligned} \tag{3.11}$$

A simple iterative procedure can be used. For example, the explicit solution of the first two constraints is shown in Figure 3—1: observe that the p-box defines 4 (case a)) or 5 (case b)) extreme points of the projection of set Ψ^F on the two-dimensional space ($P(s_1)$, $P(s_2)$).

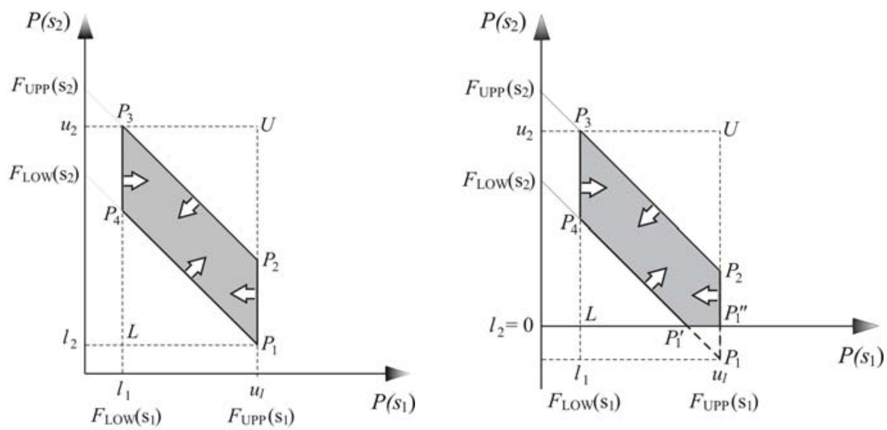


Figure 3—1. Explicit solution of the first 2 constraints in Eq. (3.11): case a): $F_{LOW}(s_2)$ -$F_{UPP}(s_1)$>0; case b: $F_{LOW}(s_2)$ -$F_{UPP}(s_1)$<0. Projection of set Ψ^F is shown hatched.

The interval bounds for the probability of the singletons are given by the intervals:

$[l_1, u_1] = [F_{LOW}(s_1), F_{UPP}(s_1)]$,

$[l_2, u_2] = [\max(0, F_{LOW}(s_2) - F_{UPP}(s_1)), F_{UPP}(s_2) - F_{LOW}(s_1)]$

However, the set Ψ^{F*} generated by the same interval probabilities thought of as being non-interactive could be much larger. Indeed, provided that the last constraint in (3.11) is satisfied, the

j

$$[l_j, u_j] = \left[\max\left(0, F_{LOW}\left(s_j\right) - F_{UPP}\left(s_{j-1}\right)\right), F_{UPP}\left(s_j\right) - F_{LOW}\left(s_{j-1}\right) \right] \quad (3.12)$$

The extreme points of the projection of set Ψ^F on the j-dimensional space ($P(s_1)$, …, $P(s_j)$) can be derived from each extreme point on the j–1-dimensional space, by considering that the sum $P(s_1)+ \ldots +P(s_j)$ must be bounded by $F_{LOW}(s_j)$ and $F_{UPP}(s_j)$.

A constructive algorithm to evaluate the extreme distributions compatible with the information given by a p-box can be obtained by selecting the set, *EXT*, corresponding to the cumulative (non-decreasing) distribution functions jumping from $F_{LOW}(s_j)$ to $F_{UPP}(s_j)$ at some points s_j and from $F_{UPP}(s_k)$ to $F_{LOW}(s_k)$ at other points s_k (or at least non-decreasing values of F, case b) in Figure 3—1). Of course, the set *EXT* contains the distribution functions corresponding to the bounds of the p-box: $P_{EXT,LOW}(s_j)= F_{LOW}(s_j) - F_{LOW}(s_{j-1})$; $P_{EXT,UPP}(s_j)= F_{UPP}(s_j) - F_{UPP}(s_{j-1})$.

The same set *EXT* (and therefore the same set $\Psi^R=\Psi^F$ of probability distributions) can be given by an equivalent random set, R, with focal elements and probabilistic assignment derived from the p-box by using a rule quite similar to the algorithm for deriving an equivalent random set from a normal fuzzy set (when the membership function is meant as a possibility distribution; see § 3.2).

Define :

$$s_j^- : F_{LOW}\left(s_j^-\right) = \lim_{\varepsilon \to 0^+} F_{LOW}\left(s_j - \varepsilon\right) = F_{LOW}\left(s_j\right) - P_{LOW}\left(s_j\right)$$

Let:

$$
\begin{aligned}
&\alpha_1 = 1 = \max_j \left(F_{LOW}\left(s_j\right)\right) = \max_j \left(F_{UPP}\left(s_j\right)\right);\\
&\alpha_2 = \max\left(\max_{j|F_{LOW}(s_j^-)<\alpha_1} \left(F_{LOW}\left(s_j\right)\right), \max_{j|F_{UPP}(s_j)<\alpha_1} \left(F_{UPP}\left(s_j\right)\right)\right);\\
&\ldots\ldots\ldots\\
&\alpha_i = \max\left(\max_{j|F_{LOW}(s_j^-)<\alpha_{i-1}} \left(F_{LOW}\left(s_j\right)\right), \max_{j|F_{UPP}(s_j)<\alpha_{i-1}} \left(F_{UPP}\left(s_j\right)\right)\right)\\
&\ldots\ldots\ldots\\
&\alpha_n = \max\left(\max_{j|F_{LOW}(s_j^-)<\alpha_{n-1}} \left(F_{LOW}\left(s_j\right)\right), \max_{j|F_{UPP}(s_j)<\alpha_{n-1}} \left(F_{UPP}\left(s_j\right)\right)\right) =\\
&= \min\left(\min_j \left(F_{LOW}\left(s_j\right)\right), \min_j \left(F_{UPP}\left(s_j\right)\right)\right)\ ; \qquad \alpha_{n+1} = 0
\end{aligned}
\tag{3.13}
$$

and define:

$$
A^i = \left\{ s_j \in S \mid F_{UPP}\left(s_j\right) \geq \alpha_i ; F_{LOW}\left(s_j^-\right) < \alpha_i \right\}; \quad m(A^i) = \alpha_i - \alpha_{i+1} \tag{3.14}
$$

Consequently:

- the lower/upper probabilities for subsets $T \subseteq S$ are Choquet capacities and Alternate Choquet capacities of order ∞ respectively (or Belief and Plausibility set functions respectively);

- the probabilistic assignment of the equivalent random set can alternatively be derived from the Belief function through the Möbius transform;

- the upper bounds u_j of the singletons (Eq. (3.12)) give the contour function of the equivalent random set R.

In (Alvarez 2006) the procedure is extended to p-boxes on infinite spaces, thus deriving equivalent random sets with infinite focal elements given by the α-cuts of the upper/lower CDFs.

Example 3—2. Let us consider $S = \{s_1, s_2, s_3, s_4\}$ and the p-box defined in the first three columns of Table 3—2. The table also displays the bounds for the singletons. The upper bounds give the contour function of the associated non-consonant random set, R. The five extreme points in the two-dimensional space ($P(s_1)$, $P(s_2)$) (case b)) determine the 10 extreme points shown in Figure

3—2a for the projection in the three-dimensional space ($P(s_1)$, $P(s_2)$, $P(s_3)$). Of course in the four-dimensional space ($P(s_1)$, $P(s_2)$, $P(s_3)$, $P(s_4)$) 10 extreme distributions are obtained when $P(s_4)=1-P(s_1)-P(s_2)-P(s_3)$. The extreme points $P_{EXT,1}$ and $P_{EXT,2}$ correspond to the cumulative distribution functions $F_{LOW}(s_j)$ and $F_{UPP}(s_j)$ respectively.

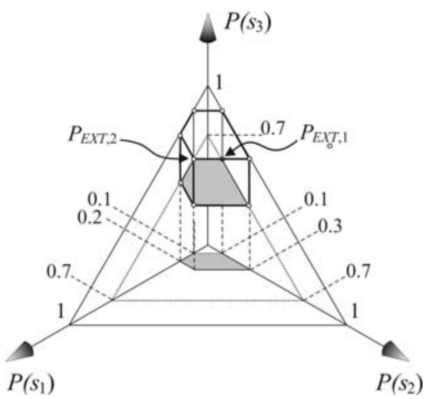
Table 3—3 presents the lower probabilities for all of the subsets in S together with their Möbius transform m, which confirms the rules given by Eqs. (3.13) and (3.14). The resulting focal elements and probabilistic assignments for R are calculated in Table 3—3 and displayed in Figure 3—2b. The random set is completely described by a stack of rectangular boxes: the width of each box identifies its (in this particular case convex, but more generally non convex) focal element along the S axis, and the height of each box is equal to its probabilistic assignment. Hence, the total height of the stack is equal to 1. The focal elements are here ordered in such a way as to obtain a stack enclosed by the cumulative upper and lower bounds of the p-box.

Table 3—2. Bounds and lower/upper CDF in Example 3-2.

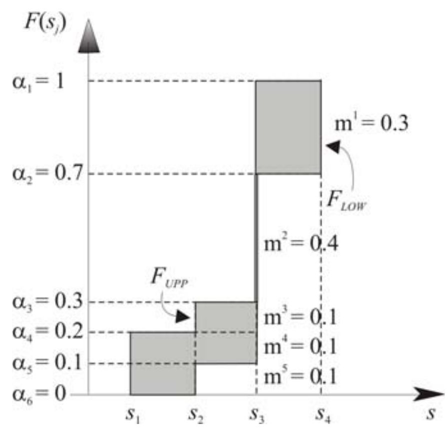
s_j	$F_{LOW}(s_j)$	$F_{UPP}(s_j)$	$F_{LOW}(s^-_j)$	$l= Bel(\{s_j\})$	$u=Pla(\{s_j\})=\mu(s_j)$
s_1	0	0.2	0	0	0.2
s_2	0.1	0.3	0	max(0, 0.1 – 0.2) = 0	0.3 – 0 = 0.3
s_3	0.7	1.0	0.1	max(0, 0.7 – 0.3) = 0.4	1.0 – 0.1 = 0.9
s_4	1.0	1.0	0.7	max(0, 1 – 1) = 0	1.0 – 0.7 = 0.3

Table 3—3. Set functions in Example 3-2.

i	$\chi_i(s_1)$	$\chi_i(s_2)$	$\chi_i(s_3)$	$\chi_i(s_4)$	$\mu_{LOW}(A^i)$	$m^i = m(A^i)$
1	1	0	0	0	0	0
2	0	1	0	0	0	0
3	0	0	1	0	0.4	0.4
4	0	0	0	1	0	0
5	1	1	0	0	0.1	0.1
6	0	1	1	0	0.5	0.5-0.4=0.1
7	0	0	1	1	0.7	0.7-0.4=0.3
8	1	0	1	0	0.4	0.4-0.4=0
9	0	1	0	1	0	0
10	1	0	0	1	0	0
11	1	1	1	0	0.7	0.7-1+0.4=0.1
12	0	1	1	1	0.8	0.8-1.2+0.4=0
13	1	0	1	1	0.7	0.7-1.1+0.4=0
14	1	1	0	1	0.1	0.1-0.1+0=0
15	1	1	1	1	1.0	1-2.3+1.7-0.4=0



a)



b)

Figure 3—2.: Example 3-2 a) extreme points in the 3-dimensional space ($P(s_1)$, $P(s_2)$, $P(s_3)$); b) equivalent random set R.

		3	
3	(0, 0.1), max (0.2, 0.3)) = 0.3	$\{s_2, s_3\}$	0.3 – 0.2 = 0.1
4	max(max(0, 0.1, 0.7), max (0.2)) = 0.2	$\{s_1, s_2, s_3\}$	0.2 – 0.1 = 0.1
5	max(max(0, 0.1)) = 0.1	$\{s_1, s_2\}$	0.1 – 0 = 0.1
6	max(max(0)) = 0		

Table 3—5. Dual extreme distributions for Example 3-2.

T	$Pla(T)$	$P_{EXT,UPP}(s)$	T^c	$Bel(T)$ = 1- $Pla(T^c)$	$P_{EXT,LOW}(s)$
$T_1 = \{s_2\}$	0.3	$P(s_2)$= 0.3	$\{s_1, s_3, s_4\}$	0	$P(s_2)$= 0
$T_2 = \{s_2, s_3\}$	1	$P(s_3)$= 0.7		0.5	$P(s_3)$= 0.5
$T_3 = \{s_2, s_3, s_1\}$	1	$P(s_1)$= 0		0.7	$P(s_1)$= 0.2
$T_4 = S$	1	$P(s_4)$= 0	$\varnothing$	1	$P(s_4)$= 0.3

Now, let us evaluate the expectation bounds for the same function considered in Example 3—1, i.e. the point-valued function $f(s_j)$ mapping onto the set Y= {5, 20, 10, 0}. The extreme distributions are identified in Table 3—5: events T are the same as events T as in Table 3—1. $Pla(T)$ and $Bel(T)$ are calculated by using m from Table 3—3. The expectation bounds are:
$E_{UPP}[f]$ = 20x 0.3 + 10 x 0.7 = 13 ; $E_{LOW}[f]$ = 10x 0.5 + 5x0.2+ 0 x 0.3 = 6

It is easy to show that the random set R determined by Eqs. (3.13) and (3.14) is not the only random set compatible with the p-box. However it must be considered as the natural extension of the information given by the p-box because the set Ψ^R determined by Eqs. (3.13) and (3.14) includes all probability distributions compatible with the p-box, and also the set Ψ^{R*} of any other random set $R*$ compatible with the p-box.

For example, when the maximum of the contour function defined by the p-box (Eq. (3.12) with $\mu(s_j)=u_j$) is equal to 1, the algorithm (3.5)-(3.6) can be used to derive a consonant random set compatible with the p-box: the focal elements are now the α-cuts of the contour function and the probabilistic assignment is again defined by the increments of α. In other words: the information given by the p-box together with additional information suggesting that the structure of the underlying random set should be consonant determine a consonant random set R' and a corresponding set $\Psi^{R'}$ of probability distributions, and of course $\Psi^{R'} \subseteq \Psi^R$.

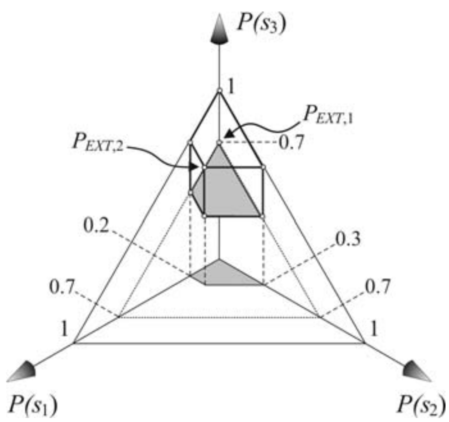
Example 3—3. Table 3—6 presents a slightly modified p-box (with respect to the p-box discussed in Example 3-2). The 8 extreme points EXT^F of set Ψ^F and the underlying non-consonant random set are shown in Figure 3—3 a) and b), respectively. The projection of Ψ^F onto the two-dimensional space ($P(s_1)$, $P(s_2)$) now contains 4 extreme points because $F_{LOW}(s_1) = F_{LOW}(s_2)$. Table 3—7 shows that the extreme distributions giving the expectation bounds are the same as in Example 3-2 (compare with Table 3—5): hence $E[f] = [6, 13]$.

Table 3—6. Reachable bounds and lower/upper CDF in Example 3—3.

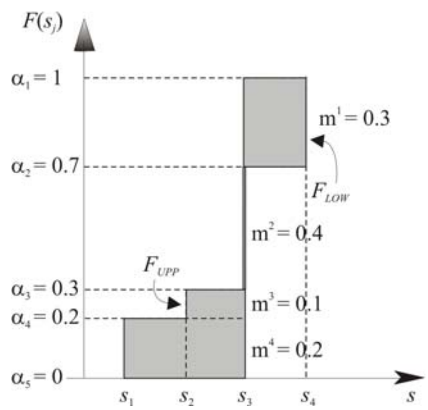
s_j	$F_{LOW}(s_j)$	$F_{UPP}(s_j)$	$l = Bel(\{s_j\})$	$u=Pla(\{s_j\})=\mu(s_j)$
s_1	0	0.2	0	0.2
s_2	0	0.3	max(0, 0 – 0.2) = 0	0.3 - 0 = 0.3
s_3	0.7	1.0	max(0, 0.7 – 0.3) = 0.4	1.0 – 0 = 1
s_4	1.0	1.0	max(0, 1 – 1) = 0	1.0 – 0.7 = 0.3

Table 3—7. Dual extreme distributions for Example 3—3.

T	$Pla(T)$	$P_{EXT,UPP}(s)$	T^c	$Bel(T)$ = 1- $Pla(T^c)$	$P_{EXT,LOW}(s)$
$T_1 = \{s_2\}$	0.3	$P(s_2)$= 0.3	$\{s_1, s_3, s_4\}$	0	$P(s_2)$= 0
$T_2 = \{s_2, s_3\}$	1	$P(s_3)$= 0.7		0.5	$P(s_3)$= 0.5
$T_3 = \{s_2, s_3, s_1\}$	1	$P(s_1)$= 0		0.7	$P(s_1)$= 0.2
$T_4 = S$	1	$P(s_4)$= 0	$\varnothing$	1	$P(s_4)$= 0.3



a)



b)

Figure 3—3. Example 3—3: a) extreme points in the 3-dimensional space ($P(s_1)$, $P(s_2)$, $P(s_3)$); b) equivalent random set.

Since now $\mu(s_3) = 1$, the contour function can be assumed to be a possibility distribution that defines a consonant random set *R'*, and corresponding set $EXT^{R'}$ of extreme distributions shown in Figure 3—4. The set $EXT^{R'}$ contains only 5 of the 8 extremes in set EXT^F. These 5 extreme points are the vertices of a pyramid with vertex in $P_{EXT,1}$ and quadrangular base on the equilateral triangle $P(s_4)= 1 - P(s_1)- P(s_2) - P(s_3) = 0$. Both $EXT^{R'}$ and EXT^F contain the extreme points $P_{EXT,1}$ and $P_{EXT,2}$, which correspond to the cumulative distribution functions $F_{LOW}(s_j)$ and $F_{UPP}(s_j)$ respectively.

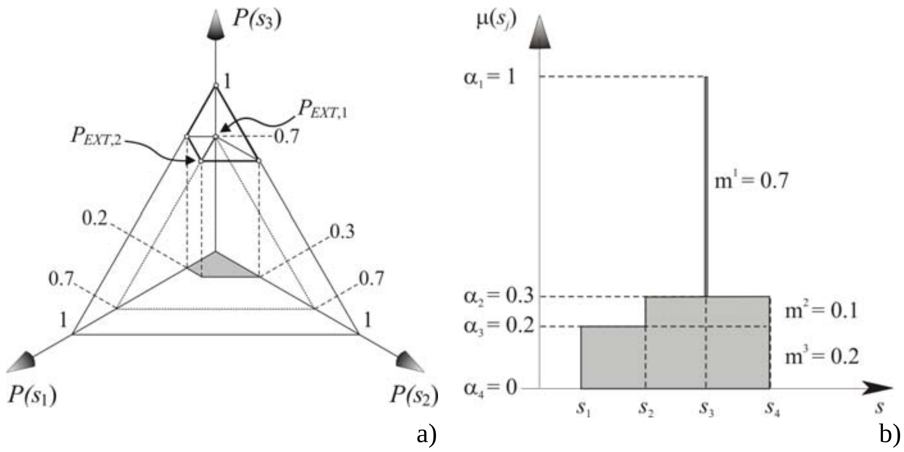


Figure 3—4. Consonant random set in Example 3—3: a) extreme points in the 3-dimensional space

Table 3—8. Dual extreme distributions for the consonant random set in Example 3—3

T	$Pla(T)$	$P_{EXT,UPP}(s)$	T^c	$Bel(T) = 1- Pla(T^c)$	$P_{EXT,LOW}(s)$
$T_1 =\{s_2\}$	0.3	$P(s_2)= 0.3$	$\{s_1, s_3, s_4\}$	0	$P(s_2)= 0$
$T_2 =\{s_2, s_3\}$	1	$P(s_3)= 0.7$		0.7	$P(s_3)= 0.7$
$T_3 =\{s_2, s_3, s_1\}$	1	$P(s_1)= 0$		0.7	$P(s_1)= 0$
$T_4 =S$	1	$P(s_4)= 0$	$\varnothing$	1	$P(s_4)= 0.3$

Hence: $E_{UPP}[f] = 13$; $E_{LOW}[f] = 10\text{x } 0.7 + 0 \text{ x } 0.3 = 7$

The same procedure (to get a consonant random set) cannot be applied to the p-box discussed in Exampe 3-2 because the contour function maximum value is equal to 0.9; however, it is easy to derive a second random set compatible with the p-box in Example 3-2 that has a nearly consonant structure: it is enough to modify the third focal element displayed in Figure 3—4 b) by taking $m^3 = 0.1$ and introducing a fourth focal element $A^4 = =\{s_1, s_2\}$, with $m^4 = 0.1$.

4 Conclusions

Random sets, which combine aleatory and set uncertainty, appear to be a powerful generalization of the classical probability theory. On the other hand, they are particular cases of a more general theory of monotone non-additive measures, Choquet capacities of different orders, coherent upper/lower probabilities, and previsions. More precisely, belief functions are coherent lower probabilities and Choquet capacities of infinite order.

The set of probability distributions compatible with the information given by a random set coincides with the natural extension of the belief/plausibility set functions, and also with the convex hull of a set of extreme distributions.

Therefore, exact bounds of the expectation of any real-valued function can be derived through the Choquet integral or equivalently by a couple of dual extreme distributions. This property seems to be very useful in engineering applications, optimal design and decision making under strong uncertainty conditions.

Fuzzy sets and p-boxes can be considered as particular indexable-type random sets, whose set of focal elements are ordered and uniquely determined by a single real number. In both the cases, simple rules can be given to derive the corresponding family of focal elements, the probabilistic assignment, and the extreme distributions of the associated random set.

Finally, the possibility of considering a hierarchy of random sets ordered by the inclusions of the corresponding sets of probability distributions has been highlighted. For example, conditions have been given to derive an included consonant random set (a fuzzy set) from the contour function of the random set corresponding to a p-box.

References

Alvarez D. A. On the calculation of the bounds of probability of events using infinite random sets. *International Journal of Approximate Reasoning*, 43(3):241-267, 2006.

Chateauneuf A.and J. Y. Jaffray. Some characterizations of lower probabilities and other monotone capacities through the use of Möbius inversion. *Mathematical Social Sciences*, 17:263-283, 1989.

Choquet G. Theory of capacities. *Annales de L'Institut Fourier*, 5:131-295, 1954.

Dubois D.and H. Prade. *Possibility theory: an approach to computerized processing of uncertainty.* Plenum Press, 1988.

Klir G. J. *Uncertainty and Information. Foundations of generalized Information Theory.* Wiley & Sons, Inc., 2005.

Shafer G. *A mathematical theory of evidence.* Princeton University Press, 1976.

Walley P. Towards a Unified Theory of Imprecise probabilities. *International Journal of Approximate Reasoning*, 24:125-148; 125, 2000.

Walley P. *Statistical reasoning with Imprecise Probabilities.* Chapman and Hall, 1991.

Zadeh L. A. Fuzzy sets as a basis for a theory of possibility. *Fuzzy Sets and Systems*, 1(1):3-28, 1978.