Extreme probability distributions of random sets, fuzzy sets and *p*-boxes

A. Bernardini

Dpt di Costruzioni e Trasporti,
Università degli Studi di Padova, Italy
Email: alberto.bernardini@unipd.it

F. Tonon*

Department of Civil Engineering,
University of Texas, USA
Email: tonon@mail.utexas.edu
*Corresponding author

Abstract: The uncertain information given by a random set on a finite space of singletons determines a set of probability distributions defined by the convex hull of a finite set of extreme distributions. After placing random sets in the context of the theory of imprecise probabilities, algorithms are given to calculate these extreme distributions, and hence exact upper/lower bounds on the expectation of functions of the uncertain variable. Detailed applications are given to consonant random sets (or their equivalent fuzzy sets) and to *p*-boxes (non-consonant random sets). A procedure is presented to calculate the random set equivalent to a *p*-box and hence to derive extreme distributions from a *p*-box. A hierarchy of non-consonant (and eventually consonant) random sets ordered by the inclusion of the corresponding sets of probability distributions can yield the same upper and lower cumulative distribution functions of the *p*-box. Simple numerical examples illustrate the presented concepts and algorithms.

Keywords: random sets; fuzzy sets; *p*-boxes.

Reference to this paper should be made as follows: Bernardini, A. and Tonon, F. (2009) 'Extreme probability distributions of random sets, fuzzy sets and *p*-boxes', *Int. J. Reliability and Safety,* Vol. 3, Nos. 1/2/3, pp.57–78.

Biographical notes: Alberto Bernardini is an Associate Professor at the Department of Costruzioni e Trasporti, Faculty of Engineering, University of Padova, Italy. His research work in the last 30 years has been developed in the fields of structural mechanics, structural optimisation, reinforced concrete (particularly high performance concrete), masonry constructions, aseismic design and repairing and retrofitting of existing building structures. His activity has combined experimental testing in the laboratory to theoretical contributions in the development of reliable models of the structural behaviour of single components and complex constructions in civil engineering. Particularly, the research has been focused on cyclic dynamic behaviour and addressed to evaluate the seismic vulnerability of buildings and infrastructures in urban areas.

Fulvio Tonon has more than 13 years of research experience in the field of geomechanics, with emphasis on rock mechanics and rock engineering, numerical modelling in geomechanics, underground excavations and tunnelling, uncertainty modelling, design process and optimisation and artificial intelligence as applied to civil engineering. He has published two books on tunnelling, 35 papers in peer-reviewed journals and 32 papers in conference proceedings. He has more than 11 years of professional experience in geotechnical and structural engineering, with projects completed in the USA, Europe and Africa.

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 realisation of the limitations that affect probabilistic modelling of uncertainty when dealing with imprecise data (Walley, 1991).

On the 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 hand, 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 consideration 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 (Helton et al., 2006).

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.

The aim of the paper is to present the second set of models of uncertainty in order to highlight their probabilistic content and to calculate expectation bounds for the system behaviour. 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 sets, i.e the imprecise events can be ordered and uniquely 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 convex set of probability distributions (that has infinite cardinality), and of the finite set of extreme distributions generated by random sets, fuzzy sets, and *p*-boxes (Ferson et al., 2003). 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.

The paper concludes with a simple and direct procedure to derive exact extreme distributions from a *p*-box by using its equivalent random set.

2 Imprecise probabilities and convex sets of probability distributions

2.1 Coherent upper and lower probabilities and previsions

Let us consider a finite probability space $(\Omega, \mathcal{F}, P)$, 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 determine a unique probability distribution. In this case, the information could be given by means of upper and/or lower probability measures, $\mu_{LOW}(T_i)$, $\mu_{UPP}(T_i)$, 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:

$$E_P[f_i] = \sum_{s_j \in S} f_i(s_j) P(s_j) \tag{1}$$

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 generalise upper/lower probabilities, which are special cases of previsions.

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 $E_{UPP}[f_i]$:

$$\Psi^E = \{P : E_{LOW}[f_i] \le E_P[f_i] \le E_{UPP}[f_i] \quad \forall f_i \in \mathcal{K}\} \tag{2}$$

Ψ^E is convex and closed. One is interested in checking two basic conditions for 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 ($E_{LOW}[f_i]$ and $E_{UPP}[f_i]$) should be the same as the *naturally extended* expectation bounds that can be derived from Ψ^E (*coherence* according to Walley's nomenclature):

$$
\begin{aligned}
E_{LOW,c}\left[f_i\right] &= \min_{P\in\Psi^E} E_P\left[f_i\right] \\
E_{UPP,c}\left[f_i\right] &= \max_{P\in\Psi^E} E_P\left[f_i\right]
\end{aligned}
\tag{3}
$$

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. they are not envelopes to Ψ^E) because they do not satisfy equation (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 equation (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* point-valued and bounded functions. In other terms, imprecise probabilities are less informative than previsions.

2.2 Choquet capacities and alternate Choquet capacities

An important criterion for classifying monotone (with respect to inclusion) measures of sets was introduced by Choquet in his theory of *capacities* (Choquet, 1954), which has been reviewed extensively in the literature, e.g. (Klir, 2006, § 4.2 and 4.3). Given a finite set *S*, let $\mathcal{P}(S)$ be the power set (set of all subsets) of *S*. A regular monotone set function $\mu: \mathcal{P}(S) \to [0, 1] \mid \mu(\varnothing) = 0, \mu(S) = 1$ is called 2-monotone (or a *Choquet Capacity of order* $k = 2$) if, given two subsets T_1 and T_2:

$$\mu(T_1 \cup T_2) \geq \mu(T_1) + \mu(T_2) - \mu(T_1 \cap T_2) \tag{4}$$

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

$$\mu_{UPP}(T_1 \cap T_2) \leq \mu_{UPP}(T_1) + \mu_{UPP}(T_2) - \mu_{UPP}(T_1 \cup T_2) \tag{5}$$

Regular monotone dual set functions not satisfying equations (4) and (5) are called 1-monotone (or *Choquet Capacity* and *Alternate Choquet Capacity of order* $k = 1$). More generally, monotone dual set functions (μ_{LOW}, μ_{UPP}) are k-monotone (*Choquet Capacity of order k*), and, respectively, *Alternate Choquet Capacity of order k*, if, given k subsets $T_1, T_2 \ldots T_k$, the following inequalities are satisfied.

$$\begin{aligned} &\mu(T_1 \cup T_2 \ldots \cup T_k) \geq \sum_{\emptyset \subset K \subseteq \{1,2..k\}} (-1)^{|K|+1} \mu\left(\bigcap_{i \in K} T_i\right) \\ &\mu_{UPP}(T_1 \cap T_2 \ldots \cap T_k) \leq \sum_{\emptyset \subset K \subseteq \{1,2..k\}} (-1)^{|K|+1} \mu_{UPP}\left(\bigcup_{i \in K} T_i\right) \end{aligned} \tag{6}$$

The following cases are notable in practice:

- A precise probability distribution is both a Choquet Capacity and an Alternate Choquet Capacity of order $k = \infty$, that satisfies relations (6) as equalities.
- Choquet and dual Alternating Choquet capacities of order $k > 1$ are coherent lower and upper probabilities, respectively. Indeed, necessary conditions for coherent upper/lower probabilities (Walley, 1991) are less restrictive; therefore, coherent lower probabilities are not necessarily Choquet capacities of order $k > 1$. See Walley (2000) and Wasserman and Kadane (1990) for some simple numerical examples.
- *Credal sets* are defined as the set of probability distributions generated by assigning probability intervals $[l_i, u_i]$ to the i-th singleton in S (Klir, 2006, § 4.3 and 5.5). Their lower bound is a 2-monotone Choquet capacity, and their upper bound is a dual Alternating Choquet capacity. They are coherent when *reachable*, i.e. when the probability intervals satisfy, for all $i \in [1, |S|]$: $\sum_{j \neq i} l_j + u_i \leq 1$ and $\sum_{j \neq i} u_j + l_i \leq 1$ (Klir, 2006, § 5.5).

There is a strong connection between the order k and a one-to-one invertible set function ${}^{\mu}m$: $\mathcal{P}(S) \to \Re$ called *Möbius transform* of the set function $\mu(T)$:

$${}^{\mu}m(A) = \sum (-1)^{|A-T|} \mu(T) \mid T \subseteq A, \tag{7}$$

whose inverse is:

$${}^{m}\mu(\mathrm{T}) = \sum m(A) \big| A \subseteq \mathrm{T}, \quad \forall \mathrm{T} \subseteq \mathrm{S}; \tag{8}$$

For the purposes of this study, the most interesting properties (see, e.g. Chateauneuf and Jaffray, 1989; Klir, 2006, § 2.3) are the following:

1 A set function μ is monotone if and only if:

$${}^{\mu}m(\emptyset) = 0; \sum_{T \in \mathcal{P}(S)} {}^{\mu}m(T) = 1; \quad \forall T \in \mathcal{P}(S): \sum_{A \subseteq T} {}^{\mu}m(A) \geq 0 \tag{9}$$

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 π(*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:

$$\begin{aligned} &P\left(s_{\pi(j)=1}\right) = \mu\left(\left\{s_{\pi(j)=1}\right\}\right) \\ &P\left(s_{\pi(j)=k>1}\right) = \mu\left(\left\{s_{\pi(j)=1}, \ldots s_k\right\}\right) - \mu\left(\left\{s_{\pi(j)=1}, \ldots s_{k-1}\right\}\right) \end{aligned} \tag{10}$$

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 both μ_{LOW} and 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 equation (2). It turns out that the two sets could be different, and that they satisfy the inclusion: $\Psi^{\mu} \subseteq \Psi^{EXT}$ (see for example (Klir, 2006, § 4.3.3) and (Walley, 2000) for simple numerical examples). Precisely:

- For coherent monotone measures with $k = 1$, equation (10) could generate probability distributions in *EXT* that do not satisfy the bounds in equation (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 equation (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} (Klir, 2006, pp.118–119).

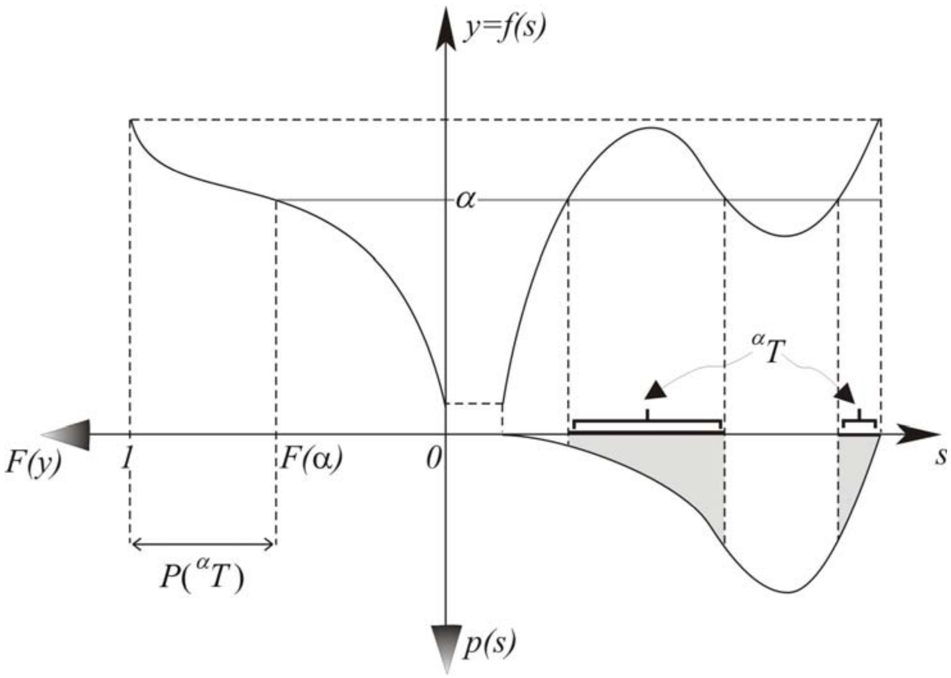
2.4 Expectation bounds and Choquet integrals for real valued functions

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: $S \to Y = \Re$ could be calculated by solving the optimisation problems in equation (3) by substituting Ψ^{μ}, Ψ^{EXT}, or Ψ for Ψ^{E}, respectively. However, the *Choquet integral* (a direct calculation based on the dual upper/lower probabilities) is generally suggested in the literature to solve the problem more easily (e.g. Klir, 2006, § 4.5).

Recall that the expectation of a point-valued bounded real function f: $S \to Y = [y_L, y_R] \subset \Re$ with CDF $F(y)$ can be calculated as follows by using the Stieltjes integral and equivalent expressions:

$$\begin{aligned} E[y = f] &= \int_{y_L}^{y_R} f \cdot dF = \int_{y_L}^{y_R} yF'(y)\,dy \overset{\substack{\text{integrate} \\ \text{by parts}}}{=} \left[yF\right]_{y_L}^{y_R} - \int_{y_L}^{y_R} Fdy \\ &= y_R - \int_{y_L}^{y_R} Fdy = y_L + \int_{y_L}^{y_R} (1 - F)dy = y_L + \int_{y_L}^{y_R} P(f > \alpha)d\alpha \\ &= y_L + \int_{y_L}^{y_R} P(^{\alpha}T = \{s \in S \mid f(s) > \alpha\})d\alpha \end{aligned} \tag{11}$$

Figure 1 Cumulative distribution function F and α–cuts for a point valued function f of a random variable s, with probability density function $p(s)$



The second equality in equation (11) can be found, for example, in (Kolmogorov and Fomin, 1970, p.364), whereas Figure 1 displays the meaning of the α-cuts ${}^{\alpha}T$ employed in the last passage. 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:

$$C(f,\mu) = y_L + \int_{y_L}^{y_R} \mu({}^{\alpha}T)d\alpha \tag{12}$$

As we have seen in Section 2.3, many distributions may be associated to a monotone measure through equation (10). 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, for finite cardinality $|S|$ of the space S, is obtained by the permutation leading to a monotonic (decreasing) ordering of the function values.

Expectation bounds are therefore calculated by applying equation (11) to the dual upper/lower measures (μ_{LOW}, μ_{UPP}). This is equivalent to the following algorithm: find the permutations for μ_{LOW} and μ_{UPP} that re-order f in a decreasing fashion; apply equation (10) to these two permutations to obtain the two sought probability distributions; calculate the expectations of f by using these two distributions. 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. Since previsions are expectations of functions [Section 2, equation (1)], the Choquet integral may be used to calculate the natural extension [right-hand sides of equation (3)] (Walley, 1991, p.130).

3 Random sets

3.1 General properties of random sets

Among the different definitions of random set (Robbins, 1944; Robbins, 1945; Matheron, 1975), we refer here to the formalism of the Theory of Evidence (Shafer, 1976; see also, Dubois and Prade, 1991), 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 $m^i = m(A^i) > 0$, $i=1, 2, \ldots n$ ('*probabilistic assignment*'), with the condition that the sum of all 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:

$$\begin{aligned} \forall T \subset S : \mu_{UPP}\left(T\right) = Pla(T) = \sum_i m^i \mid A^i \cap T \neq \varnothing, \\ \mu_{LOW}\left(T\right) = Bel(T) = \sum_i m^i \mid A^i \subseteq T \end{aligned} \tag{13}$$

Comparison with equation (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 $k = \infty$. Figure 2 illustrates the inclusions among different lower bound set functions, according to a hierarchy of decreasing generality, from coherent lower previsions to belief functions.

Since $k = \infty$, as explained in Section 2.3, Ψ^{μ} (calculated with equation 2 for $\mu_{LOW} = Bel$, $\mu_{UPP} = Pla$) coincides with the set Ψ^{EXT}, where *EXT* [calculated with equation (10)] is the set of extreme distributions that can be used to evaluate exact expectation bounds for a function of interest. From a computational viewpoint, the number of focal elements does not affect the algorithm for the expectation bounds. On the other hand, the cardinality of finite space S may affect the number of extreme probability distributions (the cardinality of *EXT*). However, if the only needed result is the expectation bounds, the algorithm does not require the calculation of the entire set of extreme distributions because it is enough to calculate the dual extreme distributions corresponding to a permutation that leads to a monotonic function.

3.2 Fuzzy sets

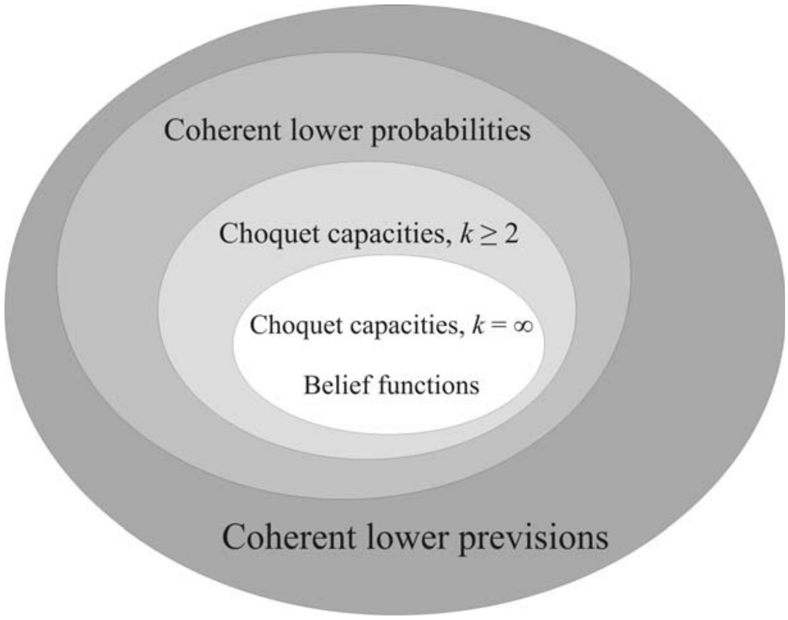
The conclusions in Section 3.1 also apply in the particular case of a *consonant* random set; i.e. when focal elements are *nested*, and hence can be ordered in such a way that:

$$A^1 \subseteq A^2 \subseteq \ldots. \subseteq A^n \tag{14}$$

Consonant random sets satisfy the relation (Klir and Yuan, 1995, p.187) for any $T_1, T_2 \subseteq S$:

$$Pla\left(T_1 \cup T_2\right) = \max\left(Pla(T_1), Pla(T_2)\right), \tag{15}$$

Figure 2 Inclusions among sets of probability distributions generated by different lower bound functions



And hence (similar to classical probability measures) they satisfy the following '*decomposability property*': the measure of uncertainty of the union of any pair of disjoint 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) $\mu: S \to [0, 1]$:

$$\mu(s_j) = Pla\left(\left\{s_j\right\}\right) \tag{16}$$

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:

$$Pla\left(T\right) = \max_{s_j \in T} \mu\left(s_j\right); \qquad Bel\left(T\right) = 1 - \max_{s_j \in T^c} \mu\left(s_j\right) \tag{17}$$

Moreover, the Möbius inversion (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:

$$
\begin{aligned}
&\alpha_1 = \max_{\mathrm{j}} \left(\mu\left(s_j\right)\right) = 1 \\
&\quad .... \\
&\alpha_i = \max_{\mathrm{j}|\mu\left(s_j\right)<\alpha_{i-1}} \left(\mu\left(s_j\right)\right); \\
&\quad .... \\
&\alpha_n = \max_{\mathrm{j}|\mu\left(s_j\right)<\alpha_{n-1}} \left(\mu\left(s_j\right)\right) = \min_{\mathrm{j}}\left(\mu\left(s_j\right)\right); \\
&\alpha_{n+1} = 0
\end{aligned}
\tag{18}
$$

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

$$A^i = \left\{s_j \in S \mid \mu\left(s_j\right) \geq \alpha_i\right\} \; ; \quad m^i = \alpha_i \text{-} \alpha_{i+1} \tag{19}$$

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 than or equal to $|S|$, because some singletons could map on to the same value of Plausibility.

For example Figure 3 displays the random set corresponding to a fuzzy subset of the space $S = \{s_1, s_2, s_3, s_4\}$ with membership $\mu(s_1) = \alpha_2$, $\mu(s_2) = \mu(s_3) = \alpha_1 = 1$, $\mu(s_4) = \alpha_3$. The random set is completely described by a stack of 3 rectangular hatched boxes: the width of each box identifies its focal element along the S-axis (a set of singletons), and the height of each box is equal to its probabilistic assignment. Hence, the total height of the stack is equal to 1.

Figure 3 Nested focal elements and probabilistic assignment of the consonant random set equivalent to a fuzzy set

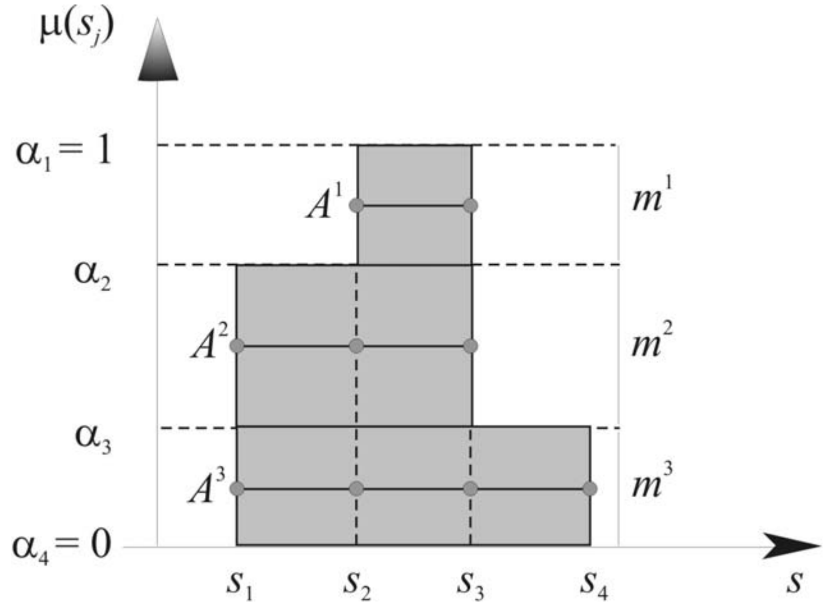
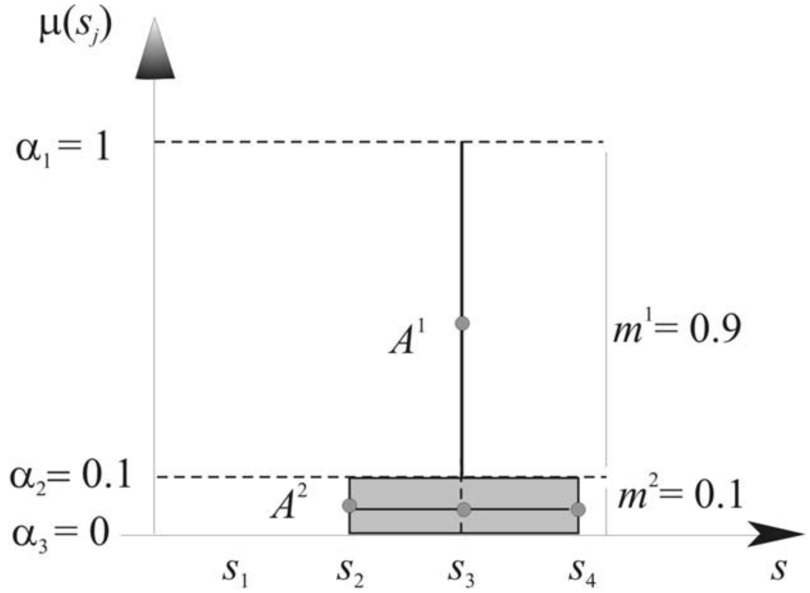


Figure 4 Nested focal elements and probabilistic assignment of the consonant random set equivalent to the fuzzy set in Example 1



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:

$$ {}^{\alpha}A = \left\{ s \in S \mid \mu(s) \geq \alpha \right\} \tag{20}$$

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:

$$ \mu\left(s_j\right) = \max_{\alpha} \quad \min\left(\alpha, \chi_{{}^{\alpha}A}(s)\right) \tag{21}$$

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

By comparing equation (20) with equation (19), 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 they are 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 $m\left(A^i = {}^{\alpha_i}A\right) = \alpha_i - \alpha_{i+1}$.

By considering equation (17) 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 recognised 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 (1988).

This comparison suggests a probabilistic (objective or subjective) content of the information summarised 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 as well. By using the interpretation of a fuzzy set as a consonant random set (Choquet capacity of infinite order in Figure 2), the following example shows how the information contained in a fuzzy set on *S* can be used to calculate upper and lower bounds on the expectation of a function defined on *S*.

Example 1: Let $S = \{s_1, s_2, s_3, s_4\}$, and consider the point-valued function $f(s_j)$ defined by the mapping: $\{s_1, s_2, s_3, s_4\} \mapsto \{5, 20, 10, 0\}$. The fuzzy set on *S* is measured by the set of membership values {0.0, 0.1, 1.0, 0.1}. Equation (18) give: $\alpha_1 = 1$; $\alpha_2 = 0.1$; $\alpha_3 = 0$. The associated consonant random set is defined by the set of pairs $\{ (A^1 = \{s_3\}, m^1 = 1.0 - 0.1 = 0.9), (A^2 = \{ s_2, s_3, s_4\}, m^2 = 0.1 - 0 = 0.1)\}$, as displayed in Figure 4.

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 1 shows the corresponding dual extreme distributions according to equation (10) and the dual set functions *Pla* and *Bel*.

Table 1 Dual extreme distributions for Example 1

T	$Pla(T)$	$P_{EXT,UPP}(s)$	$Bel(T)$	$P_{EXT,LOW}(s)$
$T_1 = \{s_2\}$	0.1	$P(s_2) = Pla(T_1) = 0.1$	0.0	$P(s_2) = Bel(T_1) = 0.0$
$T_2 = \{s_2, s_3\}$	1.0	$P(s_3) = Pla(T_2) - Pla(T_1) = 0.9$	0.9	$P(s_3) = Bel(T_2) - Bel(T_1) = 0.9$
$T_3 = \{s_2, s_3, s_1\}$	1.0	$P(s_1) = Pla(T_3) - Pla(T_2) = 0.0$	0.9	$P(s_1) = Bel(T_3) - Bel(T_2) = 0.0$
$T_4 = S$	1.0	$P(s_4) = Pla(T_4) - Pla(T_3) = 0.0$	1.0	$P(s_4) = Bel(T_4) - Bel(T_3) = 0.1$

Hence, the upper expectation is given by the expectation calculated by using $P_{EXT,UPP}$:

$$E_{UPP}[f] = E_{P_{EXT,UPP}}\left[f\right] = 20 \times 0.1 + 10 \times 0.9 + 5 \times 0.0 + 0.0 \times 0.0 = 11;$$

and the lower expectation is given by the expectation calculated by using $P_{EXT,LOW}$:

$$E_{LOW}[f] = E_{P_{EXT,LOW}}\left[f\right] = 20 \times 0.0 + 10 \times 0.9 + 5 \times 0.0 + 0 \times 0.1 = 9.$$

As explained in Section 2.4, the same results can be obtained through the Choquet integral [equation (12)]:

$$\begin{aligned} E_{UPP}[f] &= \mathrm{C}(f, \mu_{UPP} = Pla) = 0.0 + Pla(\{s_2, s_3, s_1\}) \\ &\times\left[\Delta\alpha = f(s_1) - f(s_4)\right] + Pla(\{s_2, s_3\}) \\ &\times\left[\Delta\alpha = f(s_3) - f(s_1)\right] + Pla(\{s_2\}) \\ &\times\left[\Delta\alpha = f(s_2) - f(s_3)\right] = 0.0 + 1.0 \\ &\times (5-0) + 1.0 \times (10-5) + 0.1 \times (20-10) = 11.0 \end{aligned}$$

$$E_{LOW}[f] = \mathrm{C}(f, \mu_{LOW} = Bel) = 0 + Bel(\{s_2, s_3, s_1\})$$
$$\times[\Delta\alpha = f(s_1) - f(s_4)] + Bel(\{s_2, s_3\})$$
$$\times[\Delta\alpha = f(s_3) - f(s_1)] + Bel(\{s_2\})$$
$$\times[\Delta\alpha = f(s_2) - f(s_3)] = 0.0 + 0.9 \times (5-0) + 0.9$$
$$\times(10-5) + 0.0 \times (20-10) = 9.0.$$

3.3 p-boxes

Given a finite space S, a set Ψ^F of probability distributions is implicitly defined by a p-box, i.e. lower and upper bounds, $F_{LOW}(s_j)$ and $F_{UPP}(s_j)$, on the cumulative distribution functions $F(s_j)$ (Ferson et al., 2003; Ferson and Hajagos, 2004):

$$\psi^F = \left\{P : F_{LOW}(s_j) \le F(s_j) = P\left(\left\{s_1,\ldots,s_j\right\}\right) \le F_{UPP}(s_j), j = 1 \text{ to } |S|\right\} \tag{22}$$

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 (22) for the probabilities of the singletons $P(s_j)$, i.e.:

$$\begin{aligned}
&F_{LOW}(s_1) \le P(s_1) \le F_{UPP}(s_1); \qquad P(s_1) \ge 0\\
&F_{LOW}(s_2) \le P(s_1) + P(s_2) \le F_{UPP}(s_2); \quad P(s_2) \ge 0\\
&\ldots\ldots\ldots\ldots\\
&F_{LOW}(s_j) \le P(s_j) + \sum_{i=1}^{j-1} P(s_i) \le F_{UPP}(s_j); P(s_j) \ge 0\\
&\ldots\ldots\ldots\ldots\\
&\qquad P(s_{j=|S|}) + \sum_{i=1}^{|S|-1} P(s_i) = 1; \quad P(s_{j=|S|}) \ge 0
\end{aligned} \tag{23}$$

A simple iterative procedure can be used. For example, the explicit solution of the first two constraints is shown in Figure 5: 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 $(P(s_1), P(s_2))$.

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

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

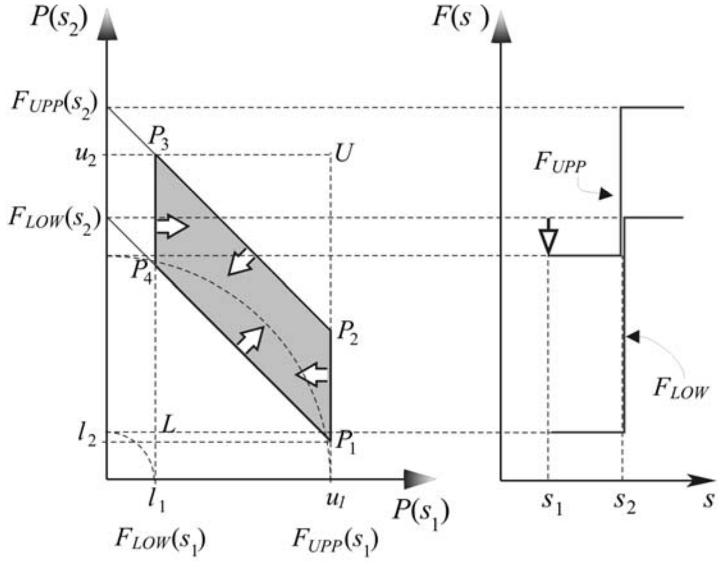
$$[l_2, u_2] = [\max(0, F_{LOW}(s_2) - F_{UPP}(s_1)), F_{UPP}(s_2) - F_{LOW}(s_1)]$$

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 (22) is satisfied, the extreme points $U=(u_1, u_2)$ and $L=(l_1, l_2)$ could be in Ψ^{F*} together with the entire Cartesian product $[l_1, u_1] \times [l_2, u_2]$.

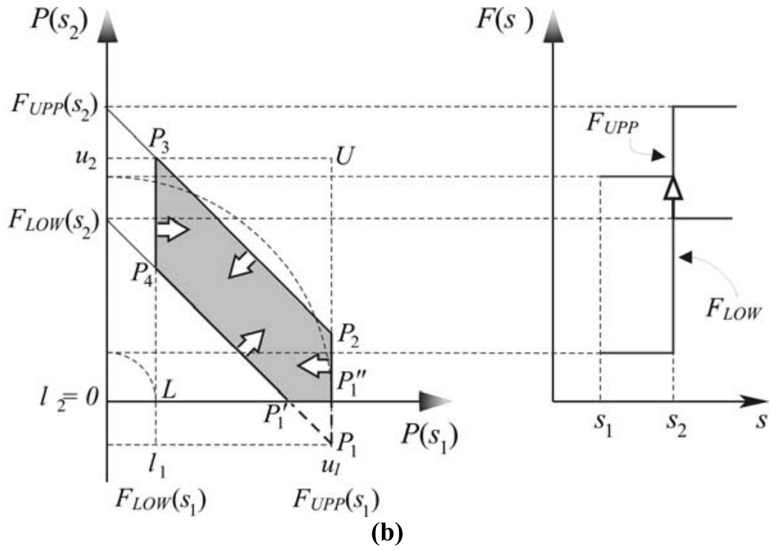
More generally, the interval probabilities for singleton $\{s_j\}$ are given by the intervals:

$$[l_j, u_j] = \left[\max\left(0, F_{LOW}(s_j) - F_{UPP}(s_{j-1})\right), F_{UPP}(s_j) - F_{LOW}(s_{j-1})\right] \tag{24}$$

Figure 5 Explicit solution of the first two constraints in equation (23): 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



(a)



(b)

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–one-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 5). 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).

With the aid of Figure 6, 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:

$$\begin{aligned}
\alpha_1 &= 1 = \max_j\left(F_{LOW}\left(s_j\right)\right) = \max_j\left(F_{UPP}\left(s_j\right)\right);\\
\alpha_2 &= \max\left(\max_{j|F_{LOW}(s_j^-)<\alpha_1}\left(F_{LOW}\left(s_j\right)\right), \max_{j|F_{UPP}(s_j)<\alpha_1}\left(F_{UPP}\left(s_j\right)\right)\right);\\
&\ldots\ldots\ldots\\
\alpha_i &= \max\left(\max_{j|F_{LOW}(s_j^-)<\alpha_{i-1}}\left(F_{LOW}\left(s_j\right)\right), \max_{j|F_{UPP}(s_j)<\alpha_{i-1}}\left(F_{UPP}\left(s_j\right)\right)\right)\\
&\ldots\ldots\ldots\\
\alpha_n &= \max\left(\max_{j|F_{LOW}(s_j^-)<\alpha_{n-1}}\left(F_{LOW}\left(s_j\right)\right), \max_{j|F_{UPP}(s_j)<\alpha_{n-1}}\left(F_{UPP}\left(s_j\right)\right)\right)\\
&= \min\left(\min_j\left(F_{LOW}\left(s_j\right)\right), \min_j\left(F_{UPP}\left(s_j\right)\right)\right)\ ; \qquad \alpha_{n+1} = 0
\end{aligned} \tag{25}$$

and define:

$$A^i = \left\{s_j \in S \middle| F_{UPP}\left(s_j\right) \geq \alpha_i ; F_{LOW}\left(s_j^-\right) < \alpha_i\right\}; \quad m\left(A^i\right) = \alpha_i - \alpha_{i+1} \tag{26}$$

Consequently:

- the lower/upper probabilities for subsets $T \subseteq S$ are Choquet capacities and Alternate Choquet capacities of infinite 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 [equation (24)] give the contour function of the equivalent random set R.

In Alvarez (2006) the procedure is extended to p-boxes on infinite spaces with general F_{upp} and F_{low}, thus deriving equivalent random sets with infinite focal elements given by the α-cuts of the upper/lower CDFs; Tonon (2008) deals with inclusion properties for discretisations of upper/lower CDFs.

Example 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 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. Figure 7a depicts the five extreme points in (projections on) the two-dimensional space $(P(s_1), P(s_2))$ (case b in Figure 5). In the three-dimensional space $(P(s_1), P(s_2), P(s_3))$, five extreme points belong to the triangle $P(s_1) + P(s_2) + P(s_3) = 0.7$ (with $P(s_4) = 0.3$), whereas five extreme points belong on the triangle $P(s_1) + P(s_2) + P(s_3) = 1$ (with $P(s_4) = 0$). 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.

Figure 6 Random set equivalent to a *p*-box in a finite discrete space

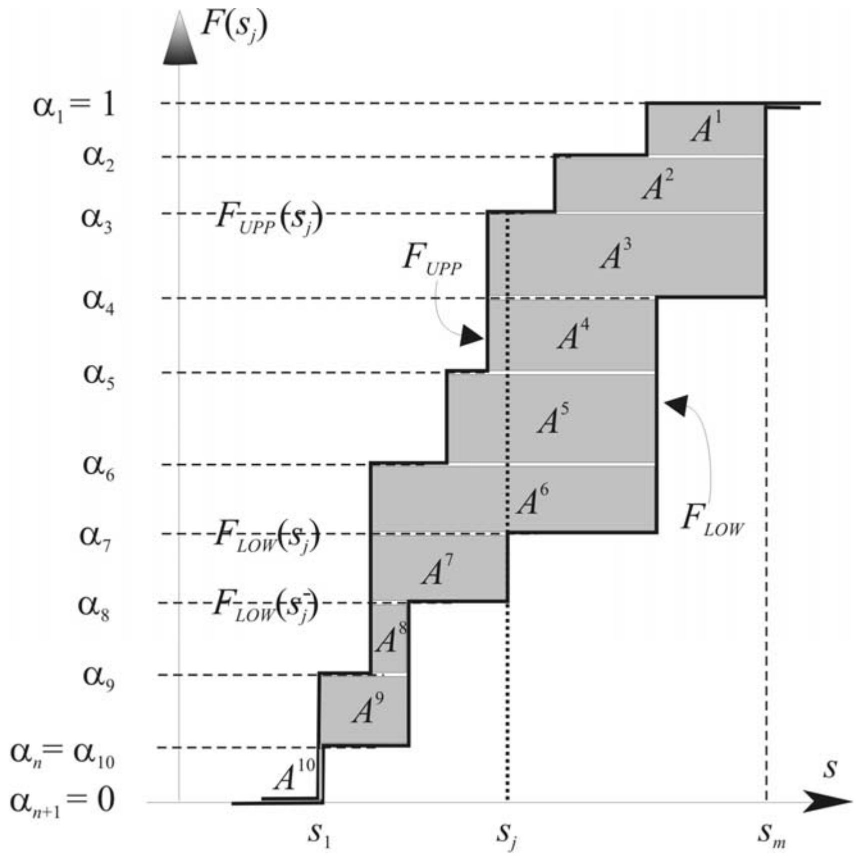


Table 3 presents the lower probabilities for all the nonempty subsets in S together with their Möbius transform m, which confirms the rules given by equations (25) and (26). The resulting focal elements and probabilistic assignments for R are calculated in Table 4 and displayed in Figure 7b. The random set is completely described by a stack of rectangular boxes, each one describing a focal element and its probabilistic assignment. 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 2 Lower/upper CDF and bounds of the singletons in Example 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	0.2	0.0	0.0	0.2
s_2	0.1	0.3	0.0	max(0.0, 0.1 – 0.2) = 0.0	0.3 – 0.0 = 0.3
s_3	0.7	1.0	0.1	max(0.0, 0.7 – 0.3) = 0.4	1.0 – 0.1 = 0.9
s_4	1.0	1.0	0.7	max(0.0, 1.0 – 1.0) = 0.0	1.0 – 0.7 = 0.3

Table 3 Set functions in Example 2; χ_i = i-th characteristic function

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	0.0
2	0	1	0	0	0.0	0.0
3	0	0	1	0	0.4	0.4
4	0	0	0	1	0.0	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 = 0.0
9	0	1	0	1	0.0	0.0
10	1	0	0	1	0.0	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 = 0.0
13	1	0	1	1	0.7	0.7 – 1.1 + 0.4 = 0.0
14	1	1	0	1	0.1	0.1 – 0.1 + 0 = 0.0
15	1	1	1	1	1.0	1 – 2.3 + 1.7 – 0.4 = 0.0

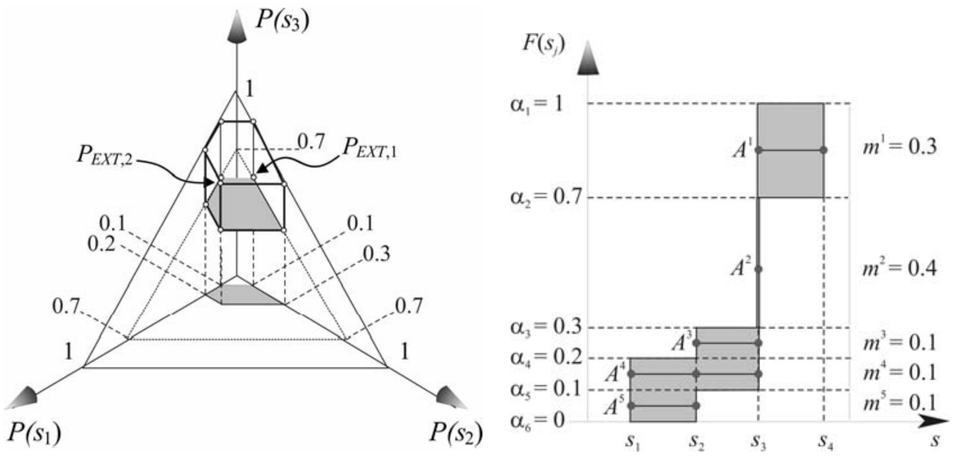
Table 4 Set functions in Example 2

i	α_i	A^i	$m^i = m(A^i)$
1	1	$\{s_3, s_4\}$	1 – 0.7 = 0.3
2	max(max(0, 0.1, 0.7), max (0.2, 0.3)) = 0.7	$\{s_3\}$	0.7 – 0.3 = 0.4
3	max(max(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 5 Dual extreme distributions for Example 2

T	$Pla(T)$	$P_{EXT,UPP}(s)$	$Bel(T)$	$P_{EXT,LOW}(s)$
$T_1 = \{s_2\}$	0.3	$P(s_2) = 0.3$	0.0	$P(s_2) = 0.0$
$T_2 = \{s_2, s_3\}$	1.0	$P(s_3) = 0.7$	0.5	$P(s_3) = 0.5$
$T_3 = \{s_2, s_3, s_1\}$	1.0	$P(s_1) = 0.0$	0.7	$P(s_1) = 0.2$
$T_4 = S$	1.0	$P(s_4) = 0.0$	1.0	$P(s_4) = 0.3$

Figure 7 Example 2 (a) 10 extreme points in the three-dimensional space ($P(s_1)$, $P(s_2)$, $P(s_3)$); $P_{EXT,1}$ and $P_{EXT,2}$ correspond to the bounds $F_{LOW}(s_j)$ and $F_{UPP}(s_j)$ of the *p*-box. (b) equivalent random set R



Now, let us evaluate the expectation bounds for the same function considered in Example 1, i.e. the point-valued function $f(s_j)$ defined by the mapping: $\{s_1, s_2, s_3, s_4\} \mapsto \{5, 20, 10, 0\}$. The extreme distributions are identified in Table 5: events T are the same as events T as in Table 1. $Pla(T)$ and $Bel(T)$ are calculated by using m from Table 3. The expectation bounds are:

$$E_{UPP}[f] = 5\times 0.0 + 20\times 0.3 + 10\times 0.7 + 0\times 0.0 = 13.0;$$
$$E_{LOW}[f] = 20\times 0.0 + 10\times 0.5 + 5\times 0.2 + 0\times 0.3 = 6.0.$$

It is easy to show that the random set R determined by equations (25) and (26) is not the only random set compatible with the *p*-box: indeed, each compatible probability distribution is a particular random set compatible with the bounds of the *p*-box (with focal elements given by singletons). However, R must be considered as the natural extension of the information given by the *p*-box because the set Ψ^R determined by equations (25) and (26) includes *all* probability distributions compatible with the *p*-box (it is the convex hull of all extreme distributions) and hence it includes the set Ψ^{R*} of any other random set $R*$ compatible with the *p*-box. Indeed, any other inner approximation of the *p*-box obtained through additional constraints reduces uncertainty and hence the set of the corresponding probability distributions.

For example, when the maximum of the contour function defined by the *p*-box [equation (24)] with $\mu(s_j)=u_j)$ is equal to 1, the algorithm (17)–(18) 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, determines a consonant random set R' and a corresponding set $\Psi^{R'}$ of probability distributions and of course $\Psi^{R'} \subseteq \Psi^R$.

Example 3: Table 6 presents a slightly modified *p*-box (with respect to the *p*-box discussed in Example 2): the value of $F_{LOW}(s_2)$ has been decreased by 0.1 to 0, thus obtaining an outer approximation of the previous *p*-box. The 8 extreme points EXT^F of set Ψ^F and the underlying non-consonant random set are shown in Figure 7 (a) and (b), respectively. The projection of Ψ^F onto the two-dimensional space ($P(s_1)$, $P(s_2)$) now contains four extreme points because $F_{LOW}(s_1) = F_{LOW}(s_2)$. Of course the set Ψ^F strongly includes the set of probability distributions in Example 2, displayed in Figure 6.

Table 7 shows that the extreme distributions, which give the same expectation bounds as in Example 2 (compare with Table 5): hence $E[f] = [6, 13]$.

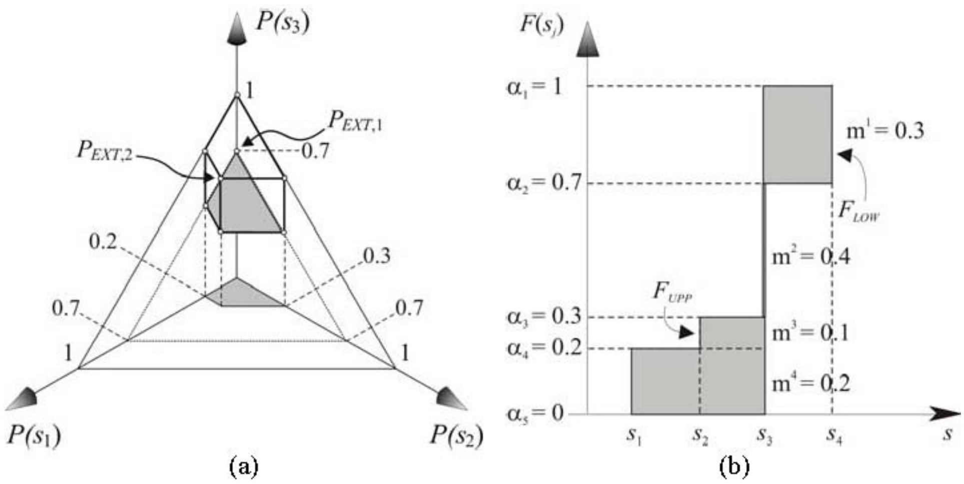
Table 6 Lower/upper CDF and reachable bounds of the singletons in Example 3

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	0.2	0.0	0.2
s_2	0.0	0.3	max(0.0, 0.0 – 0.2) = 0.0	0.3 – 0.0 = 0.3
s_3	0.7	1.0	max(0.0, 0.7 – 0.3) = 0.4	1.0 – 0.0 = 1.0
s_4	1.0	1.0	max(0.0, 1 – 1.0) = 0.0	1.0 – 0.7 = 0.3

Table 7 Dual extreme distributions for Example 3

T	$Pla(T)$	$P_{EXT,UPP}(s)$	$Bel(T)$	$P_{EXT,LOW}(s)$
$T_1 = \{s_2\}$	0.3	$P(s_2) = 0.3$	0.0	$P(s_2) = 0.0$
$T_2 = \{s_2, s_3\}$	1.0	$P(s_3) = 0.7$	0.5	$P(s_3) = 0.5$
$T_3 = \{s_2, s_3, s_1\}$	1.0	$P(s_1) = 0.0$	0.7	$P(s_1) = 0.2$
$T_4 = S$	1.0	$P(s_4) = 0.0$	1.0	$P(s_4) = 0.3$

Figure 8 Example 3: (a) eight extreme points in the three-dimensional space ($P(s_1)$, $P(s_2)$, $P(s_3)$); $P_{EXT,1}$ and $P_{EXT,2}$ correspond to the bounds $F_{LOW}(s_j)$ and $F_{UPP}(s_j)$ of the *p*-box. (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'*; its corresponding set $EXT^{R'}$ of extreme distributions is shown in Figure 9a. The set $EXT^{R'}$ contains only 5 of the 8 extremes in set EXT^{F}. These five 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. Table 8 shows the extreme distributions that give the expectation bounds of function *f* considered above. By comparing with Table 7, $P_{EXT,UPP}(s_j)$ remains unchanged, but $P_{EXT,LOW}(s_j)$ is now different and coincides with $P_{EXT,1}$, hence, again: $E_{UPP}[f] = 13.0$; while the lower bound increases to: $E_{LOW}[f] = 20 \times 0.0 + 10 \times 0.7 + 5 \times 0.0 + 0 \times 0.3 = 7.0$.

Figure 9 Consonant random set in Example 3: (a) five extreme points in the three-dimensional space; again $P_{EXT,1}$ and $P_{EXT,2}$ correspond to the bounds $F_{LOW}(s_j)$ and $F_{UPP}(s_j)$ of the *p*-box. (b) equivalent random set

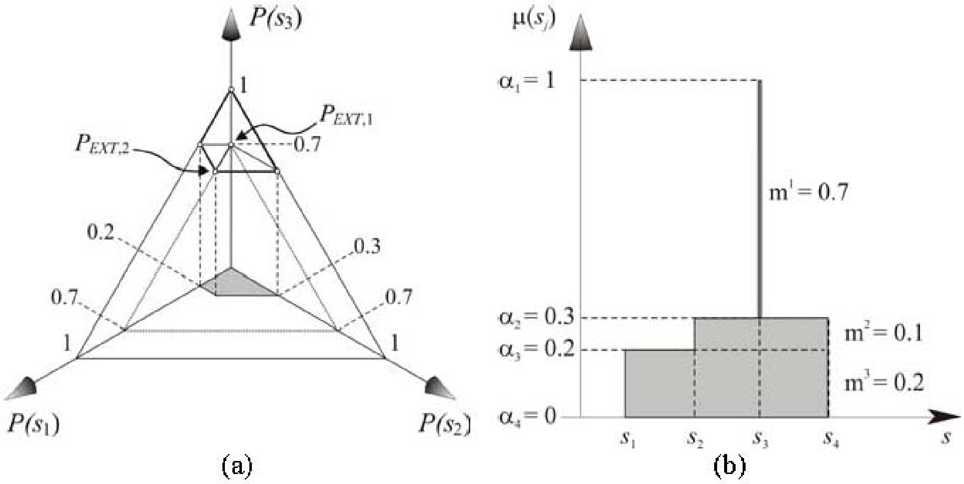


Table 8 Dual extreme distributions for the consonant random set in Example 3

T	$Pla(T)$	$P_{EXT,UPP}(s)$	$Bel(T)$	$P_{EXT,LOW}(s)$
$T_1 = \{s_2\}$	0.3	$P(s_2) = 0.3$	0.0	$P(s_2) = 0.0$
$T_2 = \{s_2, s_3\}$	1.0	$P(s_3) = 0.7$	0.7	$P(s_3) = 0.7$
$T_3 = \{s_2, s_3, s_1\}$	1.0	$P(s_1) = 0.0$	0.7	$P(s_1) = 0.0$
$T_4 = S$	1.0	$P(s_4) = 0.0$	1.0	$P(s_4) = 0.3$

The same procedure (to get a consonant random set) cannot be applied to the *p*-box discussed in Example 2 because the contour function maximum value is equal to 0.9 < 1 (Table 2, last column). Observe that the random set shown in Figure 9b gives upper and lower CDF corresponding to the bounds of the *p*-box in Example 3, and hence outer approximations of the bounds of the *p*-box in Example 2; however, the set $\Psi^{R'}$ does not include the set Ψ^{F} in Example 2.

Consonant approximations of non-consonant random sets measuring variables (for example input variables of engineering systems) are very attractive. Indeed, the equivalent fuzzy sets can be used to evaluate the response of the system, or its reliability against limit states of failure or serviceability, using the powerful and robust procedures based on the 'extension principle' for function of fuzzy variables (see for example Tonon et al., 2000). However, as shown in the above examples, consonant approximations that yield the same lower/upper CDFs or even outer approximations of the lower/upper CDFs may not guarantee inclusion of the overall convex sets of compatible probability distributions. Hence, when the behaviour of a system is described by a non-monotonic function, using such consonant approximations may overestimate the system's reliability, and thus lead to unsafe predictions. Bernardini and Tonon (2010) present conditions for the inclusion of the overall convex sets of compatible probability distributions.

4 Conclusions

Random sets, which combine aleatory and set uncertainty, appear to be a powerful generalisation 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 on 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 ordered and uniquely determined by a single real number. In both the cases, simple rules were 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. (2006) 'On the calculation of the bounds of probability of events using infinite random sets', *International Journal of Approximate Reasoning*, Vol. 43, No. 3, pp.241–267.

Bernardini, A. and Tonon, F. (2010) *Bounding Uncertainty in Civil Engineering – Theoretical Background*, Springer.

Chateauneuf, A. and Jaffray, J.Y. (1989) 'Some characterisations of lower probabilities and other monotone capacities through the use of Möbius inversion', *Mathematical Social Sciences*, Vol. 17, pp.263–283.

Choquet, G. (1954) 'Theory of capacities', *Annales de L'Institut Fourier*, Vol. 5, pp.131–295.

Dubois, D. and Prade, H.(1988) *Possibility Theory: An Approach to Computerised Processing of Uncertainty*, Plenum Press.

Dubois, D. and Prade, H.(1991) 'Random sets and fuzzy interval analysis', *Fuzzy Sets and Systems*, Vol. 42, pp.87–101.

Ferson, S., Kreinovich, V., Ginzburg, L., Myers, D.S. and Sentz, K. (2003) *Constructing Probability Boxes and Dempster-Shafer Structures*, SAND REPORT 2002–4015, Albuquerque and Livermore, Sandia National Laboratories, pp.1–143.

Ferson, S. and Hajagos, J.G. (2004) 'Arithmetic with uncertain numbers: rigorous and (often) best possible answers', *Reliability Engineering and System Safety*, Vol. 85, pp.135–152.

Helton, J., Johnson, J., Oberkampf, W. and Sallaberry, C. (2006) 'Sensitivity analysis in conjunction with evidence theory representations of epistemic uncertainty', *Reliability Engineering System and Safety*, Vol. 91, pp.1414–1434.

Klir, G.J. (2006) *Uncertainty and Information. Foundations of Generalised Information Theory*, Wiley & Sons, Inc.

Klir, G.J. and Yuan, B. (1995) *Fuzzy Sets and Fuzzy*, *Logic*, Prentice Hall, NJ.

Kolmogorov, A.N. and Fomin, S.V. (1970) *Introductory Real Analysis*, Dover Publications.

Matheron, G. (1975) *Random Sets and Integral Geometry*, Wiley & Sons, Inc., New York.

Robbins, H.E.(1944) 'On the measure of random set, I', *Annals of Mathematical Statistics*, Vol. 15, pp.70–74.

Robbins, H.E. (1945) 'On the measure of random set, II', *Annals of Mathematical Statistics*, Vol. 16, pp.342–347.

Shafer, G. (1976) *A Mathematical Theory of Evidence*, Princeton University Press.

Tonon, F. (2008) 'Some properties of a Random Set approximation to upper and lower distribution functions', *International Journal of Approximate Reasoning*, Vol. 48, pp.174–184.

Tonon, F., Bernardini, A. and Mammino, A. (2000) 'Reliability analysis of rock mass response by means of random set theory', *Reliability Engineering and System Safety*, Vol. 70, pp.263–282.

Walley, P. (1991) *Statistical Reasoning with Imprecise Probabilities*, Chapman and Hall.

Walley, P. (2000) 'Towards a unified theory of imprecise probabilities', *International Journal of Approximate Reasoning*, Vol. 24, pp.125–148.

Wasserman, L.A. and Kadane, J.B. (1990) 'Bayes's Theorem for Choquet capacities', *The Annals of Statistics*, Vol. 18, No. 3, pp.1328–1339.

Zadeh, L.A. (1978) 'Fuzzy sets as a basis for a theory of possibility', *Fuzzy Sets and Systems*, Vol. 1, No. 1, pp.3–28.