



Computing expectations with p-boxes: two views of the same problem

Lev Utkin, Sébastien Destercke

► To cite this version:

Lev Utkin, Sébastien Destercke. Computing expectations with p-boxes: two views of the same problem. The fifth international symposium on imprecise probability: theories and applications, 2007, Prague, Czech Republic. pp.435-445. irsn-00196687

HAL Id: irsn-00196687

<https://irsn.hal.science/irsn-00196687>

Submitted on 13 Dec 2007

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Preprint of the paper presented at the ISIPTA 2007 conference

Computing expectations with p-boxes: two views of the same problem

L. Utkin

Department of computer science
State Forest Technical Academy
St. Petersburg, Russia
lev.utkin@mail.ru

S. Destercke

Institut de Radioprotection
et de sûreté nucléaire (IRSN),
Cadarache, France
sebastien.destercke@irsn.fr

Abstract

Given an imprecise probabilistic model over a continuous space, computing lower (upper) expectations is often computationally hard to achieve, even in simple cases. Building tractable methods to do so is thus a crucial point in applications. In this paper, we concentrate on p-boxes (a simple and popular model), and on lower expectations computed over non-monotone functions. For various particular cases, we propose tractable methods to compute approximations or exact values of these lower expectations. We found interesting to compare two approaches: the first using general linear programming, and the second using the fact that p-boxes are special cases of random sets. We underline the complementarity of both approaches, as well as the differences.

Keywords. P-boxes, Random sets, Linear programming, Lower/upper expectation, Optimization

1 Introduction

When dealing with scarce information or with indeterminate beliefs, imprecise probability theory [11], together with lower previsions (expectations), offer a very appealing framework, for its mathematical soundness as well as for its well-defined behavioral interpretation. Nevertheless, computing lower previsions by means of the so-called natural extension when beliefs are modeled by a set of precise probability distributions on a continuous space (here, the reals) is often a very hard problem. Thus, building tractable methods to compute good approximations or exact values of such lower previsions is essential in applications.

First, let us note that the methods proposed here are "interpretation"-independent, and are valid both in Walley's behavioral theory (where the existence of an "ideal" precise distribution is not generally assumed) as well as in a more classical Bayesian sensitivity

analysis framework (where not enough information is available to precisely know the "true" probability distribution). Thus, although interpretation issue is very important, we won't deal with it in the sequel, where an "interpretation-free" vocabulary is adopted.

In this paper, we concentrate on the case when probabilistic models are p-boxes and when the function (i.e. gamble in Walley's theory) over which is computed the lower expectation is non-monotone and whose behavior is (partially) known. In other words, we propose efficient algorithms for computing lower and upper expectations of non-monotone functions of various types under the condition that the given uncertainty model is p-box.

P-boxes are one of the simplest and most popular model of sets of probability distributions, directly extending cumulative distributions used in the precise case. Although we admit that the poor expressive power of p-boxes (a price to pay for the simplicity of the model) is a limitation, we believe that they can be a good first approximation that allows for more efficient computations, and that if a decision can be taken using them, there is no reason to use a more complex model. Moreover, we should be able to efficiently compute with simple models before thinking of stepping towards more complex ones.

Although we will briefly deal with the trivial case of monotone functions, they are, as well as functions whose behavior is completely unknown, two extreme cases that will seldom be encountered in real applications (at least in "human sized" models). In most real applications, the function of interest is non-monotone but some of its characteristics are known.

Methods developed in the paper are based on two different approaches, and we found interesting to emphasize similarities and differences between these approaches, as well as how one approach can help the other: the first is based on the fact that natural extension can be viewed as a linear programming problem,

while the second use the fact that a p-box is a particular case of random set.

The next section states the problem we're going to deal with. Section 3 then explores how to compute both the unconditional and conditional interval-valued expectations of a function of one variable having one maximum. The multivariate case when the function of a set of variables has one maximum is then explored in section 4. Finally, section 5 illustrates how results could be extended to more complicated functions.

2 Problem statement

We assume that the information about a (real) variable X is (or can be) represented by some (continuous) lower $\underline{F}$ and upper $\overline{F}$ probability distributions defining the p-box $[\underline{F}, \overline{F}]$ [5]. Lower $\underline{F}$ and upper $\overline{F}$ thus define a set of precise distributions s.t.

$$\underline{F}(x) \leq F(x) \leq \overline{F}(x), \quad \forall x \in \mathbb{R}. \quad (1)$$

Given a function $h(X)$, lower ($\underline{\mathbb{E}}$) and upper ($\overline{\mathbb{E}}$) expectations over $[\underline{F}, \overline{F}]$ of $h(X)$ can be computed by means of a procedure called natural extension [11, 12], which corresponds to the following equations:

$$\underline{\mathbb{E}}(h) = \inf_{\underline{F} \leq F \leq \overline{F}} \int_{\mathbb{R}} h(x) dF, \quad \overline{\mathbb{E}}(h) = \sup_{\underline{F} \leq F \leq \overline{F}} \int_{\mathbb{R}} h(x) dF \quad (2)$$

and computing the lower (upper) expectation can be seen as finding the "optimal" distribution F reaching the infimum (supremum) in equations (2). We now detail the two generic approaches used throughout the paper.

2.1 Linear programming view

Numerically solving the above problem can be done by approximating the probability distribution function F by a set of N points $F(x_i)$, $i = 1, \dots, N$, and by translating equations (2) into the corresponding linear programming problem with N optimization variables and where constraints correspond to equation (1). Those linear programming problems are of the form

$$\underline{\mathbb{E}}^*(h) = \inf \sum_{k=1}^N h(x_k) z_k \quad \text{or} \quad \overline{\mathbb{E}}^*(h) = \sup \sum_{k=1}^N h(x_k) z_k$$

subject to

$$z_i \geq 0, \quad i = 1, \dots, N, \quad \sum_{k=1}^N z_k = 1,$$

$$\sum_{k=1}^i z_k \leq \overline{F}(x_i), \quad \sum_{k=1}^i z_k \geq \underline{F}(x_i), \quad i = 1, \dots, N.$$

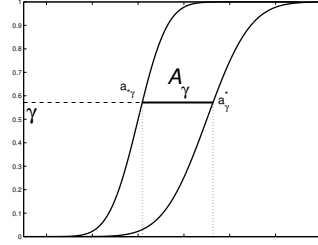


Figure 1: P-box as random set, illustration

where the z_k are the optimization variables, and objective function $\underline{\mathbb{E}}^*(h)$ ($\overline{\mathbb{E}}^*(h)$) is an approximation of the lower (upper) expectation. This way of determining the lower and upper expectations meets some computation difficulties when the value of N is rather large. Indeed, the optimization problems have N variables and $3N + 1$ constraints. On the other hand, by taking a small value of N , we take the risk of obtaining bad approximations of the exact value.

2.2 Random set view

Now that we have given a global sketch of the linear programming approach, we can detail the one using random sets. Formally, a random set is a mapping Γ from a probability space to the power set $\wp(X)$ of another space X , also called a multi-valued mapping. This mapping induces lower and upper probabilities on X [3]. Here, we shall consider the probability space $[0, 1]$ equipped with Lebesgue measure, and space $\wp(X)$ will be the measurable subsets of the real line $\mathbb{R}$.

Given the (uniformly continuous) p-box $[\underline{F}, \overline{F}]$, we will note $A_\gamma = [a_{*\gamma}, a_{\gamma}^*]$ the set s.t.

$$a_{*\gamma} := \sup\{x \in [a_{inf}, a_{sup}] : \overline{F}(x) < \gamma\} = \overline{F}^{-1}(\gamma),$$

$$a_{\gamma}^* := \inf\{x \in [a_{inf}, a_{sup}] : \underline{F}(x) > \gamma\} = \underline{F}^{-1}(\gamma),$$

By extending results from [7, 5, 4] to the continuous real line, we can conclude that the p-box $[\underline{F}, \overline{F}]$ is equivalent to the continuous random set with a uniform mass density on $[0, 1]$ and a mapping (see figure 1) s.t.

$$\Gamma(\gamma) = A_\gamma = [a_{*\gamma}, a_{\gamma}^*] \quad \gamma \in [0, 1].$$

The interest of this mapping is that it allows us to rewrite equations (2) and the Choquet integral in a

"Lebesgue" type integral, namely

$$\underline{\mathbb{E}}(h) = \int_0^1 \inf_{x \in A_\gamma} h(x) d\gamma, \quad (3)$$

$$\overline{\mathbb{E}}(h) = \int_0^1 \sup_{x \in A_\gamma} h(x) d\gamma. \quad (4)$$

Finding analytical solutions of such integrals is not easy in the general case, but approximations (either inner or outer) can be more or less easy to compute by discretizing the p-box on a finite number of levels γ_i , the main difficulty in the general case being to find the infimum or supremum of $h(X)$ for each discretized level. As in the case of linear programming, choosing too few levels γ_i or using poor heuristics can lead to bad approximations.

In both cases, it is obvious that the optimal probability distribution F providing the minimum (maximum) expectation of h depends on the form of the function h . If this form follows some typical cases, efficient solutions can be found to compute lower (upper) expectations. The simplest examples (for which solutions are well known) of such typical cases are monotone functions.

2.3 The simple case of monotone functions

Let h be a monotone function that is non-decreasing (non-increasing) in $\mathbb{R}$, then we have [12]:

$$\underline{\mathbb{E}}(h) = \int_{\mathbb{R}} h(x) d\overline{F} \quad \left(\underline{\mathbb{E}}(h) = \int_{\mathbb{R}} h(x) d\underline{F} \right), \quad (5)$$

$$\overline{\mathbb{E}}(h) = \int_{\mathbb{R}} h(x) d\underline{F} \quad \left(\overline{\mathbb{E}}(h) = \int_{\mathbb{R}} h(x) d\overline{F} \right), \quad (6)$$

and we see from (5)-(6) that lower and upper expectations are completely determined by bounding distributions $\underline{F}$ and $\overline{F}$. Using equations (3)-(4), we get the following formulas

$$\underline{\mathbb{E}}(h) = \int_0^1 h(a_{*\gamma}) d\gamma \quad \left(\underline{\mathbb{E}}(h) = \int_0^1 h(a_\gamma^*) d\gamma \right), \quad (7)$$

$$\overline{\mathbb{E}}(h) = \int_0^1 h(a_\gamma^*) d\gamma \quad \left(\overline{\mathbb{E}}(h) = \int_0^1 h(a_{*\gamma}) d\gamma \right), \quad (8)$$

which are the counterparts of equations (5)-(6). Here, expectations are totally determined by extreme values of the mappings. When h is non-monotone, equations (5)-(8) provide inner approximations of $\underline{\mathbb{E}}(h), \overline{\mathbb{E}}(h)$.

We then explore the cases where our knowledge of h can greatly improve those approximations (and even make them become exact values) without too much extra computational cost.

3 Function with one maximum - univariate case

In this section, we study the case where the function h has one maximum at point a , i.e. h is increasing (decreasing) in $(-\infty, a]$ ($[a, \infty)$). The case of h having one minimum easily follows.

3.1 Unconditional expectations

Proposition 1. *If the function h has one maximum at point $a \in \mathbb{R}$, then the upper and lower expectations of $h(X)$ on $[\underline{F}, \overline{F}]$ are*

$$\overline{\mathbb{E}}(h) = \int_{-\infty}^a h(x) d\underline{F} + h(a) [\overline{F}(a) - \underline{F}(a)] + \int_a^{\infty} h(x) d\overline{F}, \quad (9)$$

$$\underline{\mathbb{E}}(h) = \left[\int_{-\infty}^{\overline{F}^{-1}(\alpha)} h(x) d\overline{F} + \int_{\underline{F}^{-1}(\alpha)}^{\infty} h(x) d\underline{F} \right], \quad (10)$$

or, equivalently

$$\overline{\mathbb{E}}(h) = \int_0^{\underline{F}(a)} h(a_\gamma^*) d\gamma + [\overline{F}(a) - \underline{F}(a)] h(a) + \int_{\overline{F}(a)}^1 h(a_{*\gamma}) d\gamma \quad (11)$$

$$\underline{\mathbb{E}}(h) = \int_0^\alpha h(a_{*\gamma}) d\gamma + \int_\alpha^1 h(a_\gamma^*) d\gamma, \quad (12)$$

where α is one of the solution of the equation

$$h(\overline{F}^{-1}(\alpha)) = h(\underline{F}^{-1}(\alpha)). \quad (13)$$

Proof using linear programming. We assume that the function $h(x)$ is differentiable in $\mathbb{R}$ and has a finite value by $x \rightarrow \infty$. The lower and upper cumulative probability functions $\underline{F}$ and $\overline{F}$ are also differentiable. Then the following primal and dual optimization problems can be written for computing the lower expectation of the function h :

Primal problem:

Minimize $v = \int_{-\infty}^{\infty} h(x) \rho(x) dx$

subject to

$$\rho(x) \geq 0, \int_{-\infty}^{\infty} \rho(x) dx = 1,$$

$$-\int_{-\infty}^x \rho(x) dx \geq -\overline{F}(x),$$

$$\int_{-\infty}^x \rho(x) dx \geq \underline{F}(x).$$

Dual problem:

Max. $w = c_0 + \int_{-\infty}^{\infty} (-c(t) \overline{F}(t) + d(t) \underline{F}(t)) dt$

subject to

$$c_0 + \int_x^{\infty} (-c(t) + d(t)) dt \leq h(x), c_0 \in \mathbb{R},$$

$$c(x) \geq 0, d(x) \geq 0.$$

The proof that equations (9)-(10) and (13) are right then follows in three main steps:

1. We propose a feasible solution of the primal problem.
2. We then consider the feasible solution of the dual problem corresponding to the one proposed for the primal problem.
3. We show that the two solutions coincide and, therefore, according to the basic duality theorem of linear programming, these solutions are optimal ones.

First, we consider the primal problem. Let a' and a'' be real values. The function

$$\rho(x) = \begin{cases} d\bar{F}(x)/dx, & x < a' \\ 0, & a' \leq x \leq a'' \\ d\underline{F}(x)/dx, & a'' < x \end{cases}$$

is a feasible solution to the primal problem if the following conditions are respected:

$$\int_{-\infty}^{\infty} \rho(x) dx = 1,$$

which, given the above solution, can be rewritten

$$\int_{-\infty}^{a'} d\bar{F} + \int_{a''}^{\infty} d\underline{F} = 1,$$

which is equivalent to the equality

$$\bar{F}(a') = \underline{F}(a''). \quad (14)$$

We now interest ourselves in the dual problem. Let us first consider the sole constraint

$$c_0 + \int_x^{\infty} (-c(t) + d(t)) dt \leq h(x), \quad (15)$$

which is the equivalent of the primal constraint $\rho(x) \geq 0$. We then consider the following feasible solution to the dual problem as $c_0 = h(\infty)$,

$$c(x) = \begin{cases} h'(x), & x < a' \\ 0, & x \geq a' \end{cases} \quad d(x) = \begin{cases} 0, & x < a'' \\ -h'(x), & x \geq a'' \end{cases}.$$

The inequalities $c(x) \geq 0$ and $d(x) \geq 0$ are valid provided we have the inequalities $a' \leq a \leq a''$ (i.e. interval $[a', a'']$ encompasses maximum of h). By integrating $c(x)$ and $d(x)$, we get the increasing function

$$C(x) = - \int_x^{\infty} c(t) dt = \begin{cases} h(x) - h(a'), & x < a' \\ 0, & x \geq a' \end{cases}$$

and the decreasing function

$$D(x) = \int_x^{\infty} d(t) dt = \begin{cases} h(a'') - h(\infty), & x < a'' \\ h(x) - h(\infty), & x \geq a'' \end{cases}.$$

Let us rewrite condition (15) as follows:

$$c_0 + C(x) + D(x) \leq h(x). \quad (16)$$

If $x < a'$, equation (16) reads

$$c_0 + h(x) - h(a') + h(a'') - h(\infty) = h(x).$$

Hence

$$h(a'') = h(a'). \quad (17)$$

If $a' < x < a''$, we have $c_0 + h(a'') - h(\infty) \leq h(x)$ which means that for all $x \in (a', a'')$ we have $h(a'') (= h(a')) \leq h(x)$ (i.e. $h(a'')$ and a' are the minimal values of the function $h(x)$ in interval $x \in (a', a'')$.) If $x \geq a''$, then we get the trivial equality $c_0 + h(x) - h(\infty) = h(x)$. The two proposed solutions are valid iff equation (14) is valid for the primal problem and equation (17) is valid for the dual problem. To show that they are actually valid, let us consider the function

$$\varphi(\alpha) = h(\bar{F}^{-1}(\alpha)) - h(\underline{F}^{-1}(\alpha)),$$

which, being a subtraction of two continuous functions (by supposition), is continuous. Since the function h has its maximum at point $x = a$, then, by taking $\alpha = \underline{F}(a)$, we get the inequality

$$\varphi(\underline{F}(a)) = h(\bar{F}^{-1}(\underline{F}(a))) - h(a) \leq 0$$

and, by taking $\alpha = \bar{F}(a)$, we get the inequality

$$\varphi(\bar{F}(a)) = h(a) - h(\underline{F}^{-1}(\bar{F}(a))) \geq 0.$$

Consequently, there exists α in the interval $(\underline{F}(a), \bar{F}(a))$ such that $\varphi(\alpha) = 0$ (since φ is continuous). Therefore, there exist $a' = \bar{F}^{-1}(\alpha)$ and $a'' = \underline{F}^{-1}(\alpha)$ (hence, equality (14) holds) such that equality $h(a') = h(a'')$ in (17) is valid. We find the values of the objective functions

$$v_{\min} = \int_0^{a'} h(x) d\bar{F} + \int_{a''}^{\infty} h(x) d\underline{F},$$

$$w_{\max} = c_0 + \int_0^{\infty} (-c(t) \bar{F}(t) + d(t) \underline{F}(t)) dt.$$

and, by using integration by parts together with equations (14)-(17), we can show that equality $w_{\max} = v_{\min}$ holds, with α the particular solution of equation (13) for which optimum is reached, as was to be proved. $\square$

Proof using random sets. Let us now consider equations (4)-(3). Looking first at equation (4), we see that before $\gamma = \underline{F}(a)$, the supremum of h on A_γ is

$h(a_\gamma^*)$, since h is increasing between $[\infty, a]$. Between $\gamma = \underline{F}(a)$ and $\gamma = \overline{F}(a)$, the supremum of h on A_γ is $f(a)$. After $\gamma = \overline{F}(a)$, we can make the same reasoning as for the increasing part of h (except that it is now decreasing). Finally, this gives us the following formula:

$$\overline{\mathbb{E}}(h) = \int_0^{\underline{F}(a)} h(a_\gamma^*) d\gamma + \int_{\underline{F}(a)}^{\overline{F}(a)} h(a) d\gamma + \int_{\overline{F}(a)}^1 h(a_{*\gamma}) d\gamma \quad (18)$$

which is equivalent to (11). Let us now turn to the lower expectation. Before $\gamma = \underline{F}(a)$ and after $\gamma = \overline{F}(a)$, finding the infimum is again not a problem (it is respectively $h(a_{*\gamma})$ and $h(a_\gamma^*)$). Between $\gamma = \underline{F}(a)$ and $\gamma = \overline{F}(a)$, since we know that h is increasing before $x = a$ and decreasing after, infimum is either $h(a_{*\gamma})$ or $h(a_\gamma^*)$. This gives us equation

$$\underline{\mathbb{E}}h = \int_0^{\underline{F}(a)} h(a_{*\gamma}) d\gamma + \int_{\underline{F}(a)}^{\overline{F}(a)} \min(h(a_{*\gamma}), h(a_\gamma^*)) d\gamma + \int_{\overline{F}(a)}^1 h(a_\gamma^*) d\gamma \quad (19)$$

and if we use equations (14),(17) as in the first proof (reasoning used in the first proof to show that they have a solution is general, and thus applicable here), we know that there is a level α s.t. $h(\overline{F}^{-1}(\alpha)) = h(\underline{F}^{-1}(\alpha))$, and for which equation (19) reduce to equation (13). $\square$

Solutions for a function h having a minimum directly follow, due to the duality between lower and upper expectations [12] (i.e. $\underline{\mathbb{E}}(-h) = -\overline{\mathbb{E}}(h)$ and $\overline{\mathbb{E}}(-h) = -\underline{\mathbb{E}}(h)$). Of course, both proofs lead to similar formulas and, in applications, would lead to the same lower and upper expectations. Nevertheless, each view suggests a different way to solve the problem or to approximate the solution.

The proof using linear programming and the associated formulas suggest a more analytical and explicit solution, where we have to find the level α satisfying equation (14). If an analytical solution is not available, then the solution is generally approximated by scanning a larger or smaller range of possible values for α (see [10] for an example). On the other side, the proof is shorter in the case of random set, but the presence of a level α is hardly visible at first sight, and analytical results are more difficult to derive. Compared to the linear programming view, equations (11),(12),(19) suggest numerical methods based on a discretization of the levels γ rather than a heuristic search of the level α satisfying equation (14). Let us note that in the worst case, two evaluations are needed at each of the discretized levels (using equation (19)).

If the function h is symmetric about a , i.e., the equality $h(a-x) = h(a+x)$ is valid for all $x \in \mathbb{R}$, then

the value of α in (13) does not depend on h and is determined as

$$a - \overline{F}^{-1}(\alpha) = \underline{F}^{-1}(\alpha) - a.$$

Note that expressions (5),(6) can be obtained from (9),(10) by taking $a \rightarrow \infty$.

3.2 Conditional expectations

Suppose that we observe an event $B = [b_0, b_1]$. Then the lower and upper conditional expectations under condition of B can be determined as follows:

$$\begin{aligned} \underline{\mathbb{E}}(h|B) &= \inf_{\underline{F} \leq F \leq \overline{F}} \frac{\int_{\mathbb{R}} h(x) I_B(x) dF}{\int_{\mathbb{R}} I_B(x) dF}, \\ \overline{\mathbb{E}}(h|B) &= \sup_{\underline{F} \leq F \leq \overline{F}} \frac{\int_{\mathbb{R}} h(x) I_B(x) dF}{\int_{\mathbb{R}} I_B(x) dF}. \end{aligned}$$

Generally speaking, the above problems can numerically be solved by approximating the probability distribution function F by a set of N points $F(x_i)$, $i = 1, \dots, N$, and by writing linear-fractional optimization problems and then linear programming problems. Problems mentioned for the unconditional case can again occur. Figure 2 illustrates a potential optimal distribution F for which upper conditional expectation is reached (under the condition $B = [1, 8]$) when h has one maximum (which value is 5 in figure 2).

Proposition 2. *If the function h has one maximum at point $a \in \mathbb{R}$, then the upper and lower conditional expectations of $h(X)$ on $[\underline{F}, \overline{F}]$ after observing the event B are*

$$\begin{aligned} \overline{\mathbb{E}}(h|B) &= \sup_{\substack{\underline{F}(b_0) \leq \alpha \leq \overline{F}(b_0) \\ \underline{F}(b_1) \leq \beta \leq \overline{F}(b_1)}} \frac{1}{\beta - \alpha} \Psi(\alpha, \beta), \\ \underline{\mathbb{E}}(h|B) &= \inf_{\substack{\underline{F}(b_0) \leq \alpha \leq \overline{F}(b_0) \\ \underline{F}(b_1) \leq \beta \leq \overline{F}(b_1)}} \frac{1}{\beta - \alpha} \Phi(\alpha, \beta), \\ \Psi(\alpha, \beta) &= I(\alpha < \underline{F}^{-1}(a)) \int_{\underline{F}^{-1}(\alpha)}^a h(x) d\underline{F} \\ &\quad + I(\beta > \overline{F}^{-1}(a)) \int_a^{\overline{F}^{-1}(\beta)} h(x) d\overline{F} \\ &\quad + h(a) (\min(\overline{F}(a), \beta) - \max(\underline{F}(a), \alpha)) \\ &= \int_\alpha^\beta \sup_{x \in A_\gamma} h(x) d\gamma. \\ \Phi(\alpha, \beta) &= h(b_0) (\overline{F}(b_0) - \alpha) + \int_{b_0}^{\overline{F}^{-1}(\varepsilon)} h(x) d\overline{F} \\ &\quad + h(b_1) (\beta - \underline{F}(b_1)) + \int_{\underline{F}^{-1}(\varepsilon)}^{b_1} h(x) d\underline{F} \\ &= \int_\alpha^\beta \inf_{x \in A_\gamma} h(x) d\gamma. \end{aligned}$$

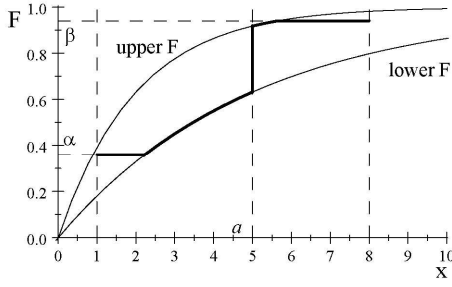


Figure 2: Optimal distribution (thick) for computing upper conditional expectation on $B = [1, 8]$

Here $I(a < b)$ is the indicator function taking 1 if $a < b$ and 0 if $a \geq b$; ε is one of the roots of the following equation:

$$h(\bar{F}^{-1}(\varepsilon)) = h(\underline{F}^{-1}(\varepsilon)). \quad (20)$$

General proof. We consider only upper expectation. We do not know how the optimal distribution function behaves outside the interval B . Therefore, we suppose that the value of the optimal distribution function at point b_0 is $F(b_0) = \alpha \in [\underline{F}(b_0), \bar{F}(b_0)]$ and its value at point b_1 is $F(b_1) = \beta \in [\underline{F}(b_1), \bar{F}(b_1)]$ (see Fig. 2). Then there holds

$$\int_{\mathbb{R}} I_B(x) dF(x) = \beta - \alpha.$$

Hence, we can write

$$\begin{aligned} \mathbb{E}(h|B) &= \sup_{\substack{\underline{F}(b_0) \leq \alpha \leq \bar{F}(b_0) \\ \underline{F}(b_1) \leq \beta \leq \bar{F}(b_1) \\ \underline{F} \leq F \leq \bar{F}}} \frac{1}{\beta - \alpha} \int_{\mathbb{R}} h(x) I_B(x) dF(x) \\ &= \sup_{\substack{\underline{F}(b_0) \leq \alpha \leq \bar{F}(b_0) \\ \underline{F}(b_1) \leq \beta \leq \bar{F}(b_1)}} \frac{1}{\beta - \alpha} \left(\sup_{\substack{\underline{F} \leq F \leq \bar{F} \\ \int_{\mathbb{R}} h(x) I_B(x) dF(x) \\ F(b_0) = \alpha \\ F(b_1) = \beta}} \int_{\mathbb{R}} h(x) I_B(x) dF(x) \right) \\ &= \sup_{\substack{\underline{F}(b_0) \leq \alpha \leq \bar{F}(b_0) \\ \underline{F}(b_1) \leq \beta \leq \bar{F}(b_1)}} \frac{1}{\beta - \alpha} \int_{\alpha}^{\beta} \sup_{x \in A_{\gamma}} h(x) d\gamma. \end{aligned} \quad (21)$$

Here $A_{\gamma} = B \cap [\underline{F}^{-1}(\gamma), \bar{F}^{-1}(\gamma)]$. By using the results obtained for the unconditional upper expectation, we can see that the integrand is equal to $\Psi(\alpha, \beta)$. The lower expectation is similarly proved, and conditional expectations when h has one minimum immediately follow. $\square$

Equation (21) shows that, as value $\beta - \alpha$ increases, so do the numerator and denominator, thus playing

opposite role in the evolution of the objective function. Hence, computing the upper (lower) conditional expectation consists in finding the values β and α s.t. any increase (decrease) in the value $\beta - \alpha$ is greater (lower) than the corresponding increase (decrease) in $\Psi(\alpha, \beta)$.

A crude algorithm to approximate the solution would be to start from the largest (tightest) interval $[\alpha, \beta]$ and then to gradually shrink (enlarge) it, evaluating each time equation (21) and retaining the highest obtained value (let us note that we can have $\bar{F}(b_0) \geq \underline{F}(b_1)$, thus the tightest interval can be void).

Another interesting point to note is that the proof takes advantage of both views, since the idea to use levels α and β comes from fractional linear programming, while the final equation (21) can be elegantly formulated by using the random set view.

4 Function with one maximum - multivariate case

Now, let h be a function from $\mathbb{R}^2 \rightarrow \mathbb{R}$ which depends on two variables X and Y . The uncertainty model becomes the following bivariate p-box:

$$\underline{F}(x, y) \leq F(x, y) \leq \bar{F}(x, y), \quad \forall (x, y) \in \mathbb{R}^2.$$

Again, we assume that h has one global maximum at point (x_0, y_0) and that $\forall z, \partial h(x, z)/\partial x = 0$ and $\partial h(z, y)/\partial y = 0$ respectively have solutions at points $x = x_0$ and $y = y_0$, making the task to find infima and suprema easier in further equations. In the next sections, we explore how we would solve the problem, under some common independence hypothesis existing in the framework of imprecise probabilities [2]. In this paper, we only provide an outline, giving general ideas and underlining the most interesting points.

In the sequel, we will consider, for the marginal random set of variable Y , the sets $B_{\kappa} = [b_{*\kappa}, b_{\kappa}^*]$ s.t.

$$\begin{aligned} b_{*\kappa} &:= \sup\{y \in [b_{inf}, b_{sup}] : \bar{F}(y) < \kappa\} = \bar{F}^{-1}(\kappa), \\ b_{\kappa}^* &:= \inf\{y \in [b_{inf}, b_{sup}] : \underline{F}(y) > \kappa\} = \underline{F}^{-1}(\kappa). \end{aligned}$$

Moreover, following Smets [9], we will note $f_X^{\mathcal{M}}$ and $f_Y^{\mathcal{M}}$ the basic belief densities corresponding to the continuous random sets of $[\underline{F}, \bar{F}]_X, [\underline{F}, \bar{F}]_Y$ when needed.

4.1 Strong Independence

In the case of strong independence, we can write

$$\mathbb{E}(h) = \inf_{\underline{F}_1 \leq F_1 \leq \bar{F}_1} \inf_{\underline{F}_2 \leq F_2 \leq \bar{F}_2} \int_{\mathbb{R}} \int_{\mathbb{R}} h(x, y) dF_1 dF_2,$$

$$\bar{\mathbb{E}}(h) = \sup_{F_1 \leq F_1 \leq \bar{F}_1} \sup_{F_2 \leq F_2 \leq \bar{F}_2} \int_{\mathbb{R}} \int_{\mathbb{R}} h(x, y) dF_1 dF_2.$$

The simplest case is when the function h can be represented as $h(X, Y) = h_1(X)h_2(Y)$. Then $\mathbb{E}(h) = \mathbb{E}(h_1) \cdot \mathbb{E}(h_2)$ and $\bar{\mathbb{E}}(h) = \bar{\mathbb{E}}(h_1) \cdot \bar{\mathbb{E}}(h_2)$. However, we consider a more complex case. Let us fix the second variable Y at point z . Denote

$$\xi(z) = \int_{\mathbb{R}} h(x, z) dF_1(x).$$

Then we have

$$\mathbb{E}(h(X, Y)) = \int_{\mathbb{R}} \xi(z) dF_2(z).$$

Let us fix variable Y to value z . Given our particular $h(X, Y)$ and Proposition 1, we have

$$\begin{aligned} \bar{\mathbb{E}}(h(X, z)) &= \sup_{F_1 \leq F_1 \leq \bar{F}_1} \xi(z) \\ &= h(x_0, z) [\bar{F}_1(x_0) - \underline{F}_1(x_0)] + \\ &\quad \int_{-\infty}^{x_0} h(x, z) d\underline{F}_1 + \int_{x_0}^{\infty} h(x, z) d\bar{F}_1. \end{aligned} \quad (22)$$

Given the assumption we've made on $h(X, Y)$ behavior, function $\xi(z)$ has a maximum at point $z = y_0$ and is monotone in intervals $(-\infty, z_0)$ and (z_0, ∞) , whatever the value of x . This implies that the optimal distribution F_2 is of the form considered in Proposition 1. Moreover, the following inequality

$$\sup_{F_2 \leq F_2 \leq \bar{F}_2} \int_{\mathbb{R}} \xi(z) dF_2(z) \geq \sup_{F_2 \leq F_2 \leq \bar{F}_2} \int_{\mathbb{R}} \hat{\xi}(z) dF_2(z)$$

holds if $\xi(z) \geq \hat{\xi}(z)$. Then it follows from the above and from the form of the optimal distribution F_2 determined in Proposition 1 that

$$\begin{aligned} \bar{\mathbb{E}}(h(X, Y)) &= \sup_{F_2 \leq F_2 \leq \bar{F}_2} \int_{\mathbb{R}} \bar{\mathbb{E}}(h(X, z)) dF_2(z) \\ &= \sup \xi(y_0) [\bar{F}_2(y_0) - \underline{F}_2(y_0)] + \\ &\quad \int_{-\infty}^{y_0} \sup \xi(z) d\underline{F}_2(z) + \int_{y_0}^{\infty} \sup \xi(z) d\bar{F}_2(z) \end{aligned}$$

and $\sup \xi(z)$ is given by equation (22). Upper expectation under strong independence can then be found in an almost explicit form. The same is not true for the lower expectation, since, relying on the first proof of Proposition 1, $\inf \xi(z)$ is obtained in this case by solving the equation

$$h(\bar{F}_1^{-1}(\alpha), z) = h(\underline{F}_1^{-1}(\alpha), z).$$

where the root α obviously depends on z . By denoting this dependency as α_z , we can nevertheless derive the

following formula

$$\begin{aligned} \mathbb{E}(h(X, Y)) &= \inf_{F_2 \leq F_2 \leq \bar{F}_2} \int_{\mathbb{R}} \mathbb{E}(h(X, z)) dF_2(z) \\ &= \int_{-\infty}^{\bar{F}_2^{-1}(\beta)} \int_{-\infty}^{\bar{F}_1^{-1}(\alpha_z)} h(x, z) d\bar{F}_1 d\bar{F}_2 \\ &\quad + \int_{-\infty}^{\bar{F}_2^{-1}(\beta)} \int_{\underline{F}_1^{-1}(\alpha_z)}^{\infty} h(x, z) d\underline{F}_1 d\bar{F}_2 \\ &\quad + \int_{\underline{F}_2^{-1}(\beta)}^{\infty} \int_{-\infty}^{\bar{F}_1^{-1}(\alpha_z)} h(x, z) d\bar{F}_1 d\underline{F}_2 \\ &\quad + \int_{\underline{F}_2^{-1}(\beta)}^{\infty} \int_{\underline{F}_1^{-1}(\alpha_z)}^{\infty} h(x, z) d\underline{F}_1 d\underline{F}_2. \end{aligned}$$

where β is a root of the equation

$$\mathbb{E}(h(X, \underline{F}_2^{-1}(\beta))) = \mathbb{E}(h(X, \bar{F}_2^{-1}(\beta))).$$

and only an approximation of such a lower bound can be found.

For the strong independence case, results rely heavily on the linear programming view and allow us to derive nice analytical formulas. Although we could set the problem in a random set framework, it would lead to numerical solutions less efficient than the one presented here (difficult problems already arise when computing lower and upper probabilities [6]).

Next cases emphasize the random set view, since this view makes solutions easier to state (especially, as could be expected, in the random set independence case).

4.2 Random set Independence

In the case of random set independence, lower and upper expectations can be computed by the following formulas:

$$\begin{aligned} \underline{\mathbb{E}}(h) &= \int_0^1 \int_0^1 \inf_{(x, y) \in [B_\kappa \times A_\gamma]} h(x, y) d\kappa d\gamma, \\ \bar{\mathbb{E}}(h) &= \int_0^1 \int_0^1 \sup_{(x, y) \in [B_\kappa \times A_\gamma]} h(x, y) d\kappa d\gamma, \end{aligned}$$

for which we can get a numerical approximation as close as we want to the exact value, by discretizing each integral. Moreover, in our particular case, evaluating $\inf h(x, y)$ or $\sup h(x, y)$ is easy.

Indeed, if h is as stated above, finding the supremum or infimum of h on $[B_\kappa \times A_\gamma]$ will often require only one computation: when $b_\kappa^* \leq y_0$ and $a_\kappa^* \leq x_0$, the supremum and infimum values are respectively on the vertices (b_κ^*, a_κ^*) and $(b_{*\kappa}, a_{*\gamma})$ of the square. when $b_{*\kappa} \leq y_0 \leq b_\kappa^*$ and $a_\gamma^* \leq x_0$, the supremum is at point

(a_{γ}^*, y_0) and the infimum is either at point $(a_{*\gamma}, b_{*\kappa})$ or $(a_{*\gamma}, b_{\kappa}^*)$. In the case where the square contains point (x_0, y_0) , this point is the supremum and the infimum is on one of the four vertices of the square. All other situations easily follow.

From a numerical standpoint, we can note that assuming random set independence is equivalent to assuming independence in a Monte-Carlo sampling scheme where each sample consists of two randomly chosen intervals A_{γ} and B_{κ} .

4.3 Unknown Interaction

Since p-boxes are special case of random sets, we can follow Fetz and Oberguggenberger [6], who show that considering unknown interaction when marginals are random sets is equivalent to consider the set of all possible joint random sets having the latter for marginals, and using results from [9] (where the extension of continuous belief functions to n -dimensional case is briefly sketched), computing lower (upper) expectation can be expressed as follows:

$$\underline{\mathbb{E}}(h) = \inf_{f_{XY}^{\mathcal{M}} \in \mathcal{J}_{XY}} \int \int \int \int \inf_{\substack{x \in [x_1, x_2] \\ y \in [y_1, y_2]}} h(x, y) Df_{XY}^{\mathcal{M}},$$

$$\overline{\mathbb{E}}(h) = \sup_{f_{XY}^{\mathcal{M}} \in \mathcal{J}_{XY}} \int \int \int \int \sup_{\substack{x \in [x_1, x_2] \\ y \in [y_1, y_2]}} h(x, y) Df_{XY}^{\mathcal{M}},$$

with

$$Df_{XY}^{\mathcal{M}} = f_{XY}^{\mathcal{M}}(x_1, x_2, y_1, y_2) dx_1 dx_2 dy_1 dy_2,$$

where $\mathcal{J}_{XY}$ is the set of all possible joint basic belief densities $f_{XY}^{\mathcal{M}}$ over $\mathbb{R}^4$ which have $f_X^{\mathcal{M}}$ and $f_Y^{\mathcal{M}}$ as their marginals.

Although the above equations are nice ways to formulate the problem, solving them analytically will be impossible in most cases. Again, the result can be approximated by approximating each p-box by a finite random set.

For instance, let us consider the two random sets $\Gamma_{\gamma}, \Gamma_{\kappa}$ approximating the p-boxes $[\underline{F}, \overline{F}]_X, [\underline{F}, \overline{F}]_Y$ with sets $A_{\gamma_i}, B_{\kappa_j}$, where $i, j = 1, \dots, n$ and where all sets have equal weights (i.e. $\gamma_i - \gamma_{i-1} = \kappa_j - \kappa_{j-1} = 1/n \forall i, j$). The problem of approximating lower expectation then comes down to finding

$$\underline{\mathbb{E}}^*(h) = \inf_{\Gamma_{\gamma}, \kappa \in \Gamma_{\gamma, \kappa}^*} \sum_{\substack{x \in A_{\gamma_i} \\ y \in B_{\kappa_j}}} h(x, y) m_{\Gamma_{\gamma}, \kappa}(A_{\gamma_i} \times B_{\kappa_j})$$

subject to

$$\begin{aligned} \sum_{j=1}^n m_{\Gamma_{\gamma}, \kappa}(A_{\gamma_i} \times B_{\kappa_j}) &= m_{\Gamma_{\gamma}}(A_{\gamma_i}), \\ \sum_{i=1}^n m_{\Gamma_{\gamma}, \kappa}(A_{\gamma_i} \times B_{\kappa_j}) &= m_{\Gamma_{\gamma}}(B_{\kappa_j}), \end{aligned}$$

where $\Gamma_{\gamma, \kappa}^*$ is the set of joint random sets having $\Gamma_{\gamma}, \Gamma_{\kappa}$ for marginals, and $m_{\Gamma_{\gamma}, \kappa}(A_{\gamma_i} \times B_{\kappa_j})$ the mass attached to the focal element $A_{\gamma_i} \times B_{\kappa_j}$. Approximation of upper expectation can be derived in a similar way (i.e. replacing the inf by sup).

Although solving the above equations is not easy, we can hope to find efficient solutions, provided we can easily evaluate $\inf h(x, y)$ on elements of the Cartesian product (we have seen that it is the case here). Also, this method can be seen as an extension of some existing methods (see [13, 8]) to functions $h(x)$ more general than indicator functions of events. Hence, we could extend some previous results concerning indicators functions to integrate some information about dependencies [1]. Another interesting thing to point out is that approximating the result in the case of unknown interaction naturally leads to a linear programming problem.

Methods given for unknown interaction and random set independence are applicable to all random sets (and only to random sets, which is a limitation compared to general linear programming), and considering special cases such as p-boxes or possibility distributions often allow the derivation of more efficient algorithms for solving the problems.

5 Function with local maxima/minima - univariate case

Now we consider a general form of the function h , i.e., the function $h(x)$ has alternate local maximum at point a_i and minimum at point b_{i-1} , $i = 1, 2, \dots$, such that

$$b_0 < a_1 < b_1 < a_2 < b_2 < \dots \quad (23)$$

Proposition 3. *If local maxima (a_i) and minima (b_i) of the function h satisfy condition (23), then the optimal distribution F for computing the lower unconditional expectation $\underline{\mathbb{E}}(h)$ has (vertical) jumps at points b_i , $i = 1, \dots$ of the size*

$$\min(\overline{F}(b_i), \alpha_{i+1}) - \max(\underline{F}(b_i), \alpha_i).$$

Between (vertical) jumps with numbers $i - 1$ and i , the optimal probability distribution function F is of

the form:

$$F(x) = \begin{cases} \bar{F}(x), & x < a' \\ \alpha_i, & a' \leq x \leq a'' \\ \underline{F}(x), & a'' < x \end{cases},$$

where α_i is the root of the equation

$$h\left(\max\left(\bar{F}^{-1}(\alpha_i), b_{i-1}\right)\right) = h\left(\min\left(\underline{F}^{-1}(\alpha_i), b_i\right)\right)$$

in interval $[F(a_i), \bar{F}(a_i)]$,

$$a' = \max\left(\bar{F}^{-1}(\alpha_i), b_{i-1}\right), \quad a'' = \min\left(\underline{F}^{-1}(\alpha_i), b_i\right).$$

The upper expectation $\bar{\mathbb{E}}(h)$ can be found from the condition $\bar{\mathbb{E}}(h) = -\underline{\mathbb{E}}(-h)$.

Proof using linear programming (brief sketch).

The first proof is based on the investigation of the following local primal and dual optimization problems for computing the lower expectation of h in finite interval $[b_0, b_1]$ where h has one maximum at point a_1 :

Primal problem:

$$\begin{aligned} v &= \int_{b_0}^{b_1} h(x) f(x) dx \rightarrow \min \text{ subject to} \\ f(x) &\geq 0, F_0 \geq 0, F_1 \geq 0, \\ -\int_{b_0}^x f(t) dt - F_0 &\geq -\bar{F}(x), \\ \int_{b_0}^x f(t) dt + F_0 &\geq \underline{F}(x), \\ -F_0 &\geq -\bar{F}(b_0), F_0 \geq \underline{F}(b_0), \\ -F_1 &\geq -\bar{F}(b_1), F_1 \geq \underline{F}(b_1), \\ \int_{b_0}^{b_1} f(t) dt + F_0 - F_1 &= 0. \end{aligned}$$

Dual problem:

$$\begin{aligned} w &= -c_0 \bar{F}(b_0) + d_0 \underline{F}(b_0) - c_1 \bar{F}(b_1) + d_1 \underline{F}(b_1) \\ &+ \int_{b_0}^{b_1} (-\bar{F}(x) c(x) + \underline{F}(x) d(x)) dx \rightarrow \max \\ \text{subject to} \\ e + \int_x^{b_1} (-c(t) + d(t)) dt &\leq h(x), \\ e - c_0 + d_0 + \int_{b_0}^{b_1} (-c(t) + d(t)) dt &\leq 0, \\ -e - c_1 + d_1 &\leq 0, \\ c(x) \geq 0, c_0 \geq 0, c_1 \geq 0, \\ d(x) \geq 0, d_0 \geq 0, d_1 \geq 0, e &\in \mathbb{R} \end{aligned}$$

All inequalities in the above primal and dual problems are valid only for $x \in [b_0, b_1]$. Results similar to those of proposition 1 can then be derived, and it is interesting to note that b_0, b_1 play similar roles to those of α, β in the conditional case. Finding the optimal distribution between each b_0, b_1 leads to four cases, depending on the situation. Figures 3.A-D illustrate these situations. The optimal F for which the lower expectation is reached is then a succession of such subcases, with a vertical jump between each of them (in figures 3.A-D, α, b_0 and b_1 are respectively equivalent to α_i, b_i and b_{i+1} of proposition 3). $\square$

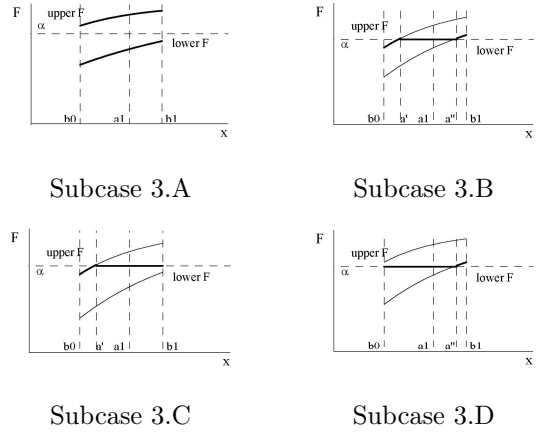


Figure 3: Subcases of piecewise optimal F

Proof using random sets (brief sketch).

For convenience, we will consider that h begins with a local minimum and ends with a local maxima a_n . Formulas when h begins/ends with a local maximum (minimum) are similar. Lower/upper expectations can be computed as follows:

$$\begin{aligned} \underline{\mathbb{E}}(h) &= \int_0^{\underline{F}(b_n)} \min_{b_i \in A_\gamma} (h(a_{*\gamma}), h(b_i), h(a_\gamma^*)) d\gamma + \int_{\underline{F}(b_n)}^1 h(a_{*\gamma}) d\gamma, \\ \bar{\mathbb{E}}(h) &= \int_0^{\underline{F}(a_1)} h(a_\gamma^*) d\gamma + \int_{\underline{F}(a_1)}^{\bar{F}(a_n)} \max_{a_i \in A_\gamma} (h(a_{*\gamma}), h(a_i), h(a_\gamma^*)) d\gamma. \end{aligned}$$

Let us explain a bit the equation for the lower expectation (details for upper one are similar). The most interesting part is the first integral. Let $B = [b_i, \dots, b_j]$ ($i \leq j$) be the set of local minima included in any particular set A_γ (B can be empty). b_{i-1} and b_{j+1} are the closest local minima outside A_γ . Let us consider the situation for which the lowest local minima $h(b_k)$ s.t. $b_k \in B$ (an empty B is a degenerated case of this one) is higher than $h(b_{i-1}), h(b_{j+1})$. As γ increases and as set A_γ evolves, various situations can happen. Either the infimum shifts from $h(a_{*\gamma})$ to $h(b_k)$ at some point (this is subcase 3.C) or it shifts from $h(b_k)$ to $h(a_\gamma^*)$ (subcase 3.D), or it shifts from $h(a_{*\gamma})$ to $h(a_\gamma^*)$ if $h(b_k)$ is too high (subcase 3.B). Subcase 3.A corresponds to the case of a local minimum b_i always dominating two other local minima (equivalent to b_0, b_1) in any set A_γ . The jumps in proposition 3 correspond to the situations where the infimum of $h(x)$ has value $h(b_k)$, either until $h(b_k) = h(a_\gamma^*)$ or until b_k is on the border of A_γ as γ increases. In the first case, it corresponds to an "horizontal" jump and to one of the root α in proposition 3, while in the latter case, the vertical jump

collapses with the upper cumulative distribution. $\square$

Similarly to figure 2, the optimal F will be a succession of vertical and horizontal jumps, sometimes following either $\overline{F}$ or $\underline{F}$ after a vertical jump has "collapsed" with $\overline{F}$ or an horizontal jump with $\underline{F}$. The proof using linear programming concentrates on "horizontal" jumps, while the proof using continuous random set emphasize vertical jumps. Again, each view suggests a different way to approximate the result.

An appealing way of formulating lower expectation is the following: let b_j $i = 1, \dots, m$ be the local minima where we have the "vertical" jumps and $\gamma_{j*}, \gamma_{j*}^*$ the associated levels on the set $[0, 1]$. Then we have

$$\underline{\mathbb{E}}(h) = \sum_{j=1}^m \left(\int_{\gamma_{(j-1)*}^*}^{\gamma_{j*}^*} \inf_{x \in A_\gamma} (h(a_\gamma^*), h(a_{*\gamma})) d\gamma + (\gamma_{(j+1)*} - \gamma_{j*}^*) h(b_j) \right). \quad (24)$$

6 Conclusions

We have considered the problem of computing lower and upper expectations on p-boxes and particular functions under two different approaches: by using linear programming and by using the fact that p-boxes are special cases of random sets. Although the two approaches try to solve identical problems, their differences suggest different ways to approximate the solutions of those problems. Moreover, some particular problems are easier to state (solve) in one approach than in the other (for example, the solutions explored in section 4). But more important than their differences is the complementarity of both approaches. Indeed, one approach can shed light on some problems shaded by the other approach (e.g. the α level of proposition 1). Another advantage of combining both approaches is the ease with which some problems are solved and the elegant formulation resulting from this combination (like in the conditional case). Let us nevertheless note that the constraint programming approach can apply to imprecise probabilities in general, while the random set approach is indeed limited to random sets.

Further works should concentrate on two directions: exploring further some ideas that were stated in the multivariate case (as well as deriving similar results for independence types not considered here), and extending the presents results to the general case of a function having many local extrema.

References

- [1] D. Berleant and J. Zhang. Using pearson correlation to improve envelopes

around the distributions of functions. *Reliable Computing*, 10(2):139–161, 2004. [<http://ifsc.ualr.edu/jdberleant/>]

- [2] I. Couso, S. Moral, and P. Walley. A survey of concepts of independence for imprecise probabilities. *Risk Decision and Policy*, 5:165–181, 2000.
- [3] A. Dempster. Upper and lower probabilities induced by a multivalued mapping. *Annals of Mathematical Statistics*, 38:325–339, 1967.
- [4] S. Destercke and D. Dubois. A unified view of some representations of imprecise probabilities. *Int. Conf. on Soft Methods in Probability and Statistics (SMPS)*, Advances in Soft Computing, pages 249–257, Bristol, 2006. Springer.
- [5] S. Ferson, L. Ginzburg, V. Kreinovich, D. Myers, and K. Sentz. Constructing probability boxes and dempster-shafer structures. Technical report, Sandia National Laboratories, 2003. [<http://www.sandia.gov/epistemic/Reports/SAND2002-4015.pdf>].
- [6] T. Fetz and M. Oberguggenberger. Propagation of uncertainty through multivariate functions in the framework of sets of probability measures. *Reliability Engineering and System Safety*, 85:73–87, 2004.
- [7] E. Kriegler and H. Held. Utilizing belief functions for the estimation of future climate change. *I. J. of Approximate Reasoning*, 39:185–209, 2005.
- [8] H. Regan, S. Ferson, and D. Berleant. Equivalence of methods for uncertainty propagation of real-valued random variables. *I. J. of Approximate Reasoning*, 36:1–30, 2004.
- [9] P. Smets. Belief functions on real numbers. *I. J. of Approximate Reasoning*, 40:181–223, 2005. [<http://iridia.ulb.ac.be/~psmets/>]
- [10] L. Utkin. Risk analysis under partial prior information and non-monotone utility functions. *I. J. of Information Technology and Decision Making*, To appear. [http://www.levvu.narod.ru/select_refern.htm]
- [11] P. Walley. *Statistical reasoning with imprecise Probabilities*. Chapman and Hall, 1991.
- [12] P. Walley. Measures of uncertainty in expert systems. *Artificial Intelligence*, 83:1–58, 1996.
- [13] R. Williamson and T. Downs. Probabilistic arithmetic : Numerical methods for calculating convolutions and dependency bounds. *I. J. of Approximate Reasoning*, 4:8–158, 1990.