Risk analysis under partial prior information and non-monotone utility functions

Lev V. Utkin
Department of Computer Science
St.Petersburg State Forest Technical Academy
Russia
lev.utkin@mail.ru

December 15, 2005

Abstract

The main objective of the paper is to investigate the risk analysis problems when a precise but unknown probability distribution of states of nature belongs to a set of continuous probability distributions restricted by some known lower and upper distributions and when utility functions are non-monotone. Methods for choosing "optimal" distributions among the set of distributions and for computing the expected utilities are proposed. Some special cases of sets of distributions, including possibility distributions, step functions, belief functions are studied under the same conditions. Various numerical examples illustrate the proposed methods.

1 Introduction

Risk can be defined in many ways and has a variety of common meanings. There exists a number of definitions of risk depending on the circumstances. The most comprehensive and exhaustive overview of methods and definitions concerning risk analysis can be found in Aven's book [3]. As pointed out by Aven, the purpose of risk analysis is to provide decision support for design and operation and, therefore, risk analysis is always part of a decision context. Moreover, probabilities are key elements in risk analysis. The traditional approach to decision analysis in the framework of expected utility theory calls for single or precise distributions of states of nature. However, we have usually only partial information about probabilities of states of nature. The information incompleteness can be modelled by means of different frameworks, including random sets, possibility theory, imprecise probabilities. It should be noted that most of these frameworks somehow or other deal with sets of possible distributions of states of nature in place of one precise distribution. Many approaches for finding optimal actions under condition of partial information have been developed [1, 2, 15, 20, 24] and every approach is directly or indirectly based on choosing some "optimal" distributions from the set of all possible distributions which are consistent with the given partial information to some extent and, then, on computing the lower and upper expected utilities with respect to these "optimal" distributions. At the same time, most approaches for decision making under partial information are developed for finite states of nature when probabilities of the states constitute discrete distributions. In this case, many decision problems can be numerically solved by using, for instance, linear programming technique. The problems become more difficult if states of nature are infinite and the corresponding probability distributions are continuous. In this case, even approximate methods can not be always applied to risk analysis. Therefore, the main objective of the paper is to investigate the decision problems when states of nature are described by sets of *continuous probability distributions* restricted by some known lower and upper distributions. It is necessary to point out that risk analysis under the above conditions does not meet any difficulties if utilities (losses) or consequences of events are monotone functions. If the condition of monotonicity holds, the "optimal" distributions for computing the lower and upper expected utilities coincide with the distributions being bounds for the set of distributions analyzed. However, this peculiarity of utility functions does not always take place. For example, Keeney and Raiffa [11] considered a *non-monotone utility function* characterizing the blood sugar. There exists some "normal" percent of the blood sugar and deviations from the "normal" percent are evidence of a certain illness. In this case, the utility function has one maximum and non-monotone. Another example is the steady-state magnification factor of a mass-spring-damper system analyzed by Oberkampf *et al* [17]. The steady-state magnification factor in this example as a function of the random frequency of the excitation acting on the mass is non-monotone and has a maximum. The list of similar examples can be continued. In case of non-monotone utility function, the procedure of choosing the "optimal" distributions and computing the expected utilities is not trivial. Therefore, methods for processing the non-monotone utility functions are studied in the paper. Moreover, some special cases of sets of distributions, including possibility distributions, step functions, belief functions are investigated. The paper is organized as follows. A general approach for determining the lower and upper expected utilities and the problem statement are given in Section 2. Sections 3, 4, 5 are devoted to studying the methods for computing the lower and upper expectations of functions which are different in type, including monotone functions, functions having one maximum or minimum and functions having one maximum and one minimum. A special type of distributions viewed as step functions and the relationship between this type of distributions and Dempster-Shafer structures are studied in Section 6. A method for computing the lower and upper expected utilities under condition that the states of nature are described by a possibility distribution is considered in Section 7. An approximate procedure for computing the lower and upper expectations of arbitrary functions is proposed in Section 8. Various numerical examples illustrate the proposed methods.

2 Problem statement

Suppose that information about random variable X, characterizing states of nature, is represented by some lower $\underline{F}$ and upper $\overline{F}$ probability distributions. This means that we do not know some "true" precise distribution F of the random variable X, but this distribution can be arbitrary and satisfies the condition

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

In other words, the lower $\underline{F}$ and upper $\overline{F}$ distributions can be viewed as bounds for a set of precise distributions F. According to [8], the pair of lower $\underline{F}$ and upper $\overline{F}$ probability distribution functions is called "p-box" for the distribution F if condition (1) is valid.

Since the precise distribution is unknown, then only lower and upper expectations (expected utilities) of a function $h(X)$ (utility function) can be found. Then the lower and upper expectations of $h(X)$ can be computed by means of a procedure called natural extension [21, 22]. According to Walley [21, 22], the natural extension in the case of known lower and upper probability distributions becomes (Choquet integrals)

$$\underline{\mathbb{E}}h = \inf_{\underline{F}\leq F\leq \overline{F}} \int_{\mathbb{R}} h(x)\mathrm{d}F(x),\ \overline{\mathbb{E}}h = \sup_{\underline{F}\leq F\leq \overline{F}} \int_{\mathbb{R}} h(x)\mathrm{d}F(x). \tag{2}$$

Here the minimum and maximum is taken over all distributions satisfying (1), i.e., there exists an "optimal" distribution among the set of distributions, which gives the lower expected utility, and there exists an "optimal" distribution which gives the upper expected utility.

Generally, the above problems can numerically be solved by approximating the probability distribution function F by a set of N points $F(x_i)$, $i = 1, ..., N$, and by writing linear programming problems with N optimization variables. The linear problems are of the form:

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

subject to

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

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

Here z_k, $k = 1, ..., N$, are the optimization variables; $\underline{\mathbb{E}}^* h$ and $\overline{\mathbb{E}}^* h$ are approximate lower and upper expectations of the function h, respectively.

However, this way of determining the lower and upper expectations meets 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 too approximate results or, sometimes, just wrong values of the lower and upper expectations. At the same time, it is obvious that the optimal probability distribution F providing the minimum or maximum of the expectation of h depends on the form of the function h, and some typical cases of functions h allows us to get solutions in a more simple way. Therefore, we study these cases below and develop the most simple ways for computing $\underline{\mathbb{E}}h$ and $\overline{\mathbb{E}}h$.

3 Monotone functions

Here we consider the most simple case when the function h is monotone. If the function h is non-decreasing in $\mathbb{R}$, then there hold [22]

$$\underline{\mathbb{E}}h = \int_{\mathbb{R}} h(x) \mathrm{d}\overline{F}(x),\quad \overline{\mathbb{E}}h = \int_{\mathbb{R}} h(x) \mathrm{d}\underline{F}(x). \tag{3}$$

If the function h is non-increasing in $\mathbb{R}$, then there hold [22]

$$\underline{\mathbb{E}}h = \int_{\mathbb{R}} h(x) \mathrm{d}\underline{F}(x),\quad \overline{\mathbb{E}}h = \int_{\mathbb{R}} h(x) \mathrm{d}\overline{F}(x). \tag{4}$$

The above expressions simplify calculation of the lower and upper expectations of monotone functions because only lower and upper distributions have to be used. It can be seen from (3)-(4) that the lower and upper expectations are completely defined by bounded distributions $\underline{F}$ and $\overline{F}$.

Example 1 *Suppose an individual has* 1000$ *to invest in a stock or a bond. The stock is a financial asset which has a variable return that has a distribution bounded by uniform distributions with expected returns of* 4% *and* 6%, *the standard deviation of* $2/\sqrt{3}$%. *The bond returns* 5% *with certainty. The individual is risk averse and the corresponding utility function over wealth is given by* $x(w) = \sqrt{w}$. *Then*

$$\underline{\mathbb{E}}_{stock}h = \frac{1}{0.06 - 0.02} \int_{0.02}^{0.06} \sqrt{1000(1+x)} \mathrm{d}x = 32.249,$$

$$\overline{\mathbb{E}}_{stock}h = \frac{1}{0.08 - 0.04} \int_{0.04}^{0.08} \sqrt{1000(1+x)} \mathrm{d}x = 32.557,$$

$$\underline{\mathbb{E}}_{bond}h = \overline{\mathbb{E}}_{bond}h = \sqrt{1000(1+0.05)} = 32.4.$$

By taking the parameter of pessimism $\gamma = 0.6$, *we get*[1]

$$\gamma \underline{\mathbb{E}}_{stock}h + (1-\gamma)\overline{\mathbb{E}}_{stock}h = 32.372 < \mathbb{E}_{bond}h = 32.4.$$

Hence, the investment in the bond is preferable with the given parameter of pessimism.

[1]The question of comparing two intervals or an interval with a value is ambiguous. In the example, we use the comparison procedure closed to Hurwicz criterion with optimism parameter $1 - \gamma$ without its detailed justification because this is not a question of the paper.

4 Functions having one maximum or minimum

In this section, we study a case when the utility function h has one maximum (minimum) at point x_0, i.e., $h(x)$ is increasing (decreasing) in $(-\infty, x_0]$ and decreasing (increasing) in $[x_0, \infty)$.

Proposition 1 *If the function h has one maximum at point x_0 in $\mathbb{R}$, the continuous lower $\underline{F}$ and upper $\overline{F}$ probability distributions of random variable X are known, then the upper and lower expectations of $h(X)$ are*

$$\overline{\mathbb{E}}h = h(x_0)\left[\overline{F}(x_0) - \underline{F}(x_0)\right] + \int_{-\infty}^{x_0} h(x)\mathrm{d}\underline{F}(x) + \int_{x_0}^{\infty} h(x)\mathrm{d}\overline{F}(x), \tag{5}$$

$$\underline{\mathbb{E}}h = \min_{\alpha\in[0,1]}\left[\int_{-\infty}^{\overline{F}^{-1}(\alpha)} h(x)\mathrm{d}\overline{F}(x) + \int_{\underline{F}^{-1}(\alpha)}^{\infty} h(x)\mathrm{d}\underline{F}(x)\right]. \tag{6}$$

Proof. The upper bound has been proposed by Kuznetsov [14] and it can be explained as follows. The function h increases in the interval $(-\infty, x_0)$ and the optimal distribution function (see Fig.1, thick curve) is the left part $(x \in (-\infty, x_0))$ of $\underline{F}(x)$, h decreases in the interval (x_0, ∞) and the optimal distribution function is the right part $(x \in (x_0, \infty))$ of $\overline{F}(x)$. The optimal function has one jump $\overline{F}(x_0) - \underline{F}(x_0)$ at point x_0. The same explanation can not be applied to the lower expectation of h. Let us prove that the optimal function corresponding to $\underline{\mathbb{E}}h$ has a form shown in Fig.2. Let us rewrite problem (2) for computing the lower expectation of the function h in the form of finite differences

$$\underline{\mathbb{E}}^{*}h = \inf\sum_{k=1}^{N} h(x_k)\rho(x_k)\triangle x_k,$$

subject to

$$\rho(x_k) \geq 0, \quad i = 1, ..., N, \ \sum_{k=1}^{N}\rho(x_k)\triangle x_k = 1,$$

$$\sum_{k=1}^{i}\rho(x_k)\triangle x_k \leq \overline{F}(x_i), \ \sum_{k=1}^{i}\rho(x_k)\triangle x_k \geq \underline{F}(x_i), \ i = 1, ..., N.$$

Here N is the numbers of points; $\rho(x_k)$ are optimization variables (points of a density). The problem has N variables and $3N+1$ constraints. This implies that N constraints are equalities. At the same time, the points $\overline{F}(x_i)$ and $\underline{F}(x_i)$ have to compose non-decreasing functions, i.e., distribution functions. It is obvious that this function has the form depicted in Fig.2 as thick. This function corresponds to the equalities $\rho(x_k) = \overline{\rho}(x_k)$ for $x_k \in (-\infty, \overline{F}^{-1}(\alpha))$, $\rho(x_k) = 0$ for $x_k \in [\overline{F}^{-1}(\alpha), \underline{F}^{-1}(\alpha)]$, $\rho(x_k) = \underline{\rho}(x_k)$ for $x_k \in (\underline{F}^{-1}(\alpha), \infty)$. Here $\underline{\rho}(x_k)$ and $\overline{\rho}(x_k)$ are points of lower and upper density functions corresponding to $\underline{F}(x_k)$ and $\overline{F}(x_k)$, respectively; α is some value from 0 to 1. Then the optimization problem for computing $\underline{\mathbb{E}}h$ can be written as (6). ■

Denote

$$\psi(\alpha) = \int_{-\infty}^{\overline{F}^{-1}(\alpha)} h(x)\mathrm{d}\overline{F}(x) + \int_{\underline{F}^{-1}(\alpha)}^{\infty} h(x)\mathrm{d}\underline{F}(x).$$

If the function $\psi(\alpha)$ can be found in the explicit form, then the minimum of the function $\psi(\alpha)$ is computed by solving the equation $\mathrm{d}\psi(\alpha)/\mathrm{d}\alpha = 0$. Otherwise, the optimization problem can be numerically solved by looking over values of $\alpha \in [0, 1]$.

Corollary 1 *The minimum over $\alpha \in [0, 1]$ in (6) is achieved at a point which is one of the solutions to the following equation:*

$$h\left(\overline{F}^{-1}(\alpha)\right) = h\left(\underline{F}^{-1}(\alpha)\right). \tag{7}$$

Moreover, the inequalities $\overline{F}^{-1}(\alpha) \leq x_0$ and $\underline{F}^{-1}(\alpha) \geq x_0$ are valid.

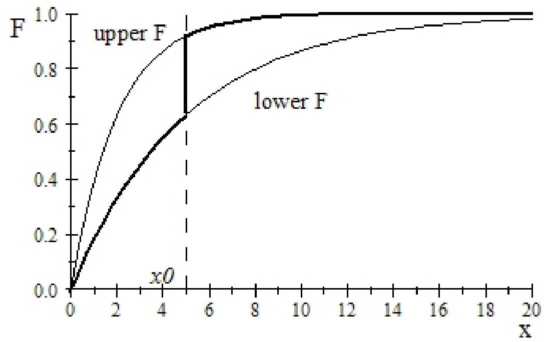

Figure 1: The optimal distribution (thick) for computing the upper expectation

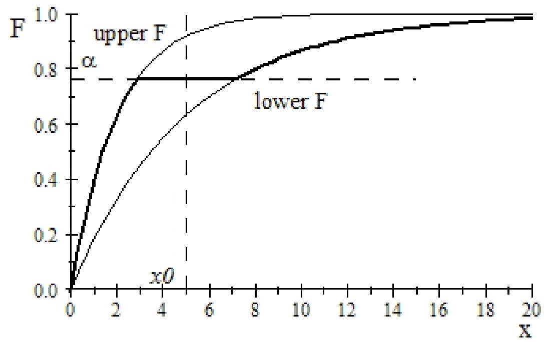

Figure 2: The optimal distribution (thick) for computing the lower expectation

Proof. It follows from (6) that

$$\frac{\mathrm{d}\psi(\alpha)}{\mathrm{d}\alpha} = \frac{\mathrm{d}\overline{F}^{-1}(\alpha)}{\mathrm{d}\alpha} h\left(\overline{F}^{-1}(\alpha)\right)\overline{\varphi}\left(\overline{F}^{-1}(\alpha)\right) - \frac{\mathrm{d}\underline{F}^{-1}(\alpha)}{\mathrm{d}\alpha} h\left(\underline{F}^{-1}(\alpha)\right)\underline{\varphi}\left(\underline{F}^{-1}(\alpha)\right).$$

Here $\underline{\varphi}(x) = \mathrm{d}\underline{F}(x)/\mathrm{d}x$ and $\overline{\varphi}(x) = \mathrm{d}\overline{F}(x)/\mathrm{d}x$. For arbitrary F and $\varphi(x) = \mathrm{d}F(x)/\mathrm{d}x$, there holds

$$\frac{\mathrm{d}F^{-1}(\alpha)}{\mathrm{d}\alpha} = \frac{1}{\varphi\left(F^{-1}(\alpha)\right)}.$$

Hence

$$\frac{\mathrm{d}\psi(\alpha)}{\mathrm{d}\alpha} = h\left(\overline{F}^{-1}(\alpha)\right) - h\left(\underline{F}^{-1}(\alpha)\right) = 0.$$

Since the function h has the maximum at point x_0, then the values $\overline{F}^{-1}(\alpha)$ and $\underline{F}^{-1}(\alpha)$ are situated around x_0. However, $\overline{F}^{-1}(\alpha) \leq \underline{F}^{-1}(\alpha)$ and, consequently, $\overline{F}^{-1}(\alpha) \leq x_0$ and $\underline{F}^{-1}(\alpha) \geq x_0$, as was to be proved. ∎

Corollary 2 *If the function h is symmetric about x_0, i.e., the equality $h(x_0 - x) = h(x_0 + x)$ is valid for all $x \in \mathbb{R}$, then the optimal value of α in (6) or in (7) does not depend on h and is determined as*

$$x_0 - \overline{F}^{-1}(\alpha) = \underline{F}^{-1}(\alpha) - x_0.$$

Proof. The proof immediately follows from (7). ∎

Corollary 3 *Expressions (3)-(4) can be obtained from (5)-(6) by taking* $x_0 \to \infty$.

Proof. Indeed, there holds $\overline{F}(x_0) - \underline{F}(x_0) \to 0$ by $x_0 \to \infty$. Moreover, it follows from Corollary 1 that $\underline{F}^{-1}(\alpha) \geq x_0 = \infty$ and optimal value of α tends to 1. Hence, $\overline{F}^{-1}(\alpha) \to \infty$ and we get (3)-(4). ∎

Corollary 3 implies that (3) and (4) can be regarded as special cases of (5)-(6) under condition $x_0 \to \infty$.

Proposition 2 *If the function h has one minimum at point x_0 in $\mathbb{R}$, the continuous lower $\underline{F}$ and upper $\overline{F}$ probability distributions of random variable X are known, then the upper $\overline{\mathbb{E}}h$ and lower $\underline{\mathbb{E}}h$ expectations of $h(X)$ are determined by means of (6) and (5), respectively.*

Proof. The proof immediately follows from equalities $\underline{\mathbb{E}}(-h) = -\overline{\mathbb{E}}h$ and $\overline{\mathbb{E}}(-h) = -\underline{\mathbb{E}}h$ [21]. ∎

Example 2 *Suppose that losses after a unit failure is the function of time $h(x) = 60 - (x-5)^2$, and it is known that the unit time to failure is governed by a distribution whose bounds are exponential distributions with a failure rate 0.2 and 0.5. Let us compute the expected losses as the expectation of h. The lower and upper distribution functions of the unit time to failure are $1 - \exp(-0.2x)$ and $1 - \exp(-0.5x)$, respectively. By taking $x_0 = 5$, we get*

$$\overline{\mathbb{E}}h = h(5)\left[\overline{F}(5) - \underline{F}(5)\right] + \int_0^5 h(x)\mathrm{d}\underline{F}(x) + \int_5^\infty h(x)\mathrm{d}\overline{F}(x)$$

$$= 60 \cdot (\exp(-0.5 \cdot 5) - \exp(-0.2 \cdot 5)) + 31.321 + 4.268 = 52.736.$$

Since $\overline{F}^{-1}(\alpha) = -2\ln(1-\alpha)$ and $\underline{F}^{-1}(\alpha) = -5\ln(1-\alpha)$, then $\underline{\mathbb{E}}h = \min_{\alpha\in[0,1]} \psi(\alpha)$, where

$$\psi(\alpha) = \int_0^{-2\ln(1-\alpha)} \left(60 - (x-5)^2\right)(0.5 \cdot \exp(-0.5x))\,\mathrm{d}x$$
$$+ \int_{-5\ln(1-\alpha)}^{\infty} \left(60 - (x-5)^2\right)(0.2 \cdot \exp(-0.2x))\,\mathrm{d}x.$$

After simplification, we get

$$\psi(\alpha) = 12\alpha + 12(1-\alpha)\ln(1-\alpha) - 21(1-\alpha)\ln^2(1-\alpha) + 35.$$

By using the equation

$$\mathrm{d}\psi(\alpha)/\mathrm{d}\alpha = 21\ln^2(1-\alpha) + 30\ln(1-\alpha) = 0,$$

we obtain that the function $\psi(\alpha)$ has the minimal value 29.745 at point $\alpha = 1 - \exp(-10/7) = 0.76$. Therefore, $\underline{\mathbb{E}}h = 29.745$. Finally, we obtain that the expected losses are in the interval $[29.745,\ 52.736]$.

It should be noted that the optimal point α can be found by using equality (7) as follows:

$$60 - (-2\ln(1-\alpha) - 5)^2 = 60 - (-5\ln(1-\alpha) - 5)^2.$$

Hence, we have two solutions $\alpha = 1 - \exp(-10/7)$ and $\alpha = 0$. Since $\overline{F}^{-1}(0) = \underline{F}^{-1}(0)$, then the second solution has to be removed. Therefore, we get $\alpha = 1 - \exp(-10/7)$.

The above example has shown that the optimization problems can explicitly be solved. However, generally, the above optimization problems are numerically solved by changing α from 0 to 1 with some accuracy and by substituting all values into the objective function.

5 Functions having one maximum and one minimum

Let us consider a more complex case when the function h has one maximum at point x_{01} and one minimum at point x_{02}, i.e., the function $h(x)$ is increasing in $(-\infty, x_{01})$, decreasing in $[x_{01}, x_{02}]$, and increasing in (x_{02}, ∞).

Proposition 3 *If the function h has one maximum at point x_{01} and one minimum at point x_{02} in $\mathbb{R}$, the continuous lower $\underline{F}$ and upper $\overline{F}$ probability distributions of random variable X are known, then the upper and lower expectations of $h(X)$ are*

$$\overline{\mathbb{E}}h = \int_{-\infty}^{x_{01}} h(x)\mathrm{d}\underline{F}(x) - h(x_{01})\underline{F}(x_{01}) + \max\left[\max_{\alpha\in[0,\overline{F}(x_{01})]}\psi_1(\alpha), \max_{\alpha\in[\overline{F}(x_{01}),1]}\psi_2(\alpha)\right],$$

where

$$\psi_1(\alpha) = \int_{\underline{F}^{-1}(\alpha)}^{\infty} h(x)\mathrm{d}\underline{F}(x) + h(x_{01})\alpha,$$

$$\psi_2(\alpha) = \int_{\underline{F}^{-1}(\alpha)}^{\infty} h(x)\mathrm{d}\underline{F}(x) + h(x_{01})\overline{F}(x_{01}) + \int_{x_{01}}^{\overline{F}^{-1}(\alpha)} h(x)\mathrm{d}\overline{F}(x),$$

and

$$\underline{\mathbb{E}}h = \int_{x_{02}}^{\infty} h(x)\mathrm{d}\overline{F}(x) + h(x_{02})\overline{F}(x_{02}) + \min\left[\min_{\alpha\in[\underline{F}(x_{02}),1]}\varphi_1(\alpha), \min_{\alpha\in[0,\underline{F}(x_{02})]}\varphi_2(\alpha)\right],$$

where

$$\varphi_1(\alpha) = \int_{-\infty}^{\overline{F}^{-1}(\alpha)} h(x)\mathrm{d}\overline{F}(x) - h(x_{02})\alpha.$$

$$\varphi_2(\alpha) = \int_{-\infty}^{\overline{F}^{-1}(\alpha)} h(x)\mathrm{d}\overline{F}(x) - h(x_{02})\underline{F}(x_{02}) + \int_{\overline{F}^{-1}(\alpha)}^{x_{02}} h(x)\mathrm{d}\underline{F}(x).$$

Proof. The proof is similar to the proof of Proposition 1. By proving the upper bound, two cases should be considered. The first case is shown in Fig.3. In this case, $\alpha < \overline{F}(x_{01})$. The second case is shown in Fig.4. In this case, $\alpha \geq \overline{F}(x_{01})$. Both cases differ by values of the optimal function in the interval $[x_{01}, \underline{F}^{-1}(\alpha)]$. Since we do not know the value α a priori, then both cases are analyzed. The lower bound is proved in the same way. ■

Example 3 *Let us return to Example 2 under condition that the loss function is of the form $h(x) = 4x^3 - 80x^2 + 508x - 940$. The function h has one maximum at point $x_{01} = 5.21$ and one minimum at point $x_{02} = 8.12$. Then*

$$\int_{0}^{5.21} h(x)\mathrm{d}\underline{F}(x) - h(5.21)\underline{F}(5.21) = -250.98,$$

$$\begin{aligned}\psi_1(\alpha) &= \int_{-5\ln(1-\alpha)}^{\infty} h(x)\mathrm{d}\underline{F}(x) + h(5.21)\alpha \\ &= 600 - 500\,(1-\alpha)\ln^2(1-\alpha) - 500\,(1-\alpha)\ln^3(1-\alpha) \\ &- 1540\,(1-\alpha)\ln(1-\alpha) - 499.16\alpha.\end{aligned}$$

Hence

$$\mathrm{d}\psi_1(\alpha)/\mathrm{d}\alpha = 500\ln^3(1-\alpha) + 2000\ln^2(1-\alpha) + 2540\ln(1-\alpha) + 1040.8.$$

The equation $\mathrm{d}\psi_1(\alpha)/\mathrm{d}\alpha = 0$ has three roots 0.852, 0.651, 0.644. However, the maximum of $\psi_1(\alpha)$ is achieved at point $\alpha = 0.852$. Hence

$$\max_{\alpha\in[\underline{F}(x_{02}),1]}\psi_1(\alpha) = \psi_1(0.852) = 856.12.$$

$$\begin{aligned}\psi_2(\alpha) &= \int_{-5\ln(1-\alpha)}^{\infty} h(x)\mathrm{d}\underline{F}(x) + h(5.21)\overline{F}(5.21) + \int_{5.21}^{-2\ln(1-\alpha)} h(x)\mathrm{d}\overline{F}(x) \\ &= 972\alpha\ln(1-\alpha) - 972\ln(1-\alpha) - 276\ln^2(1-\alpha) \\ &- 468\ln^3(1-\alpha) - 972\alpha + 276\alpha\ln^2(1-\alpha) + 468\alpha\ln^3(1-\alpha) + 987.05,\end{aligned}$$

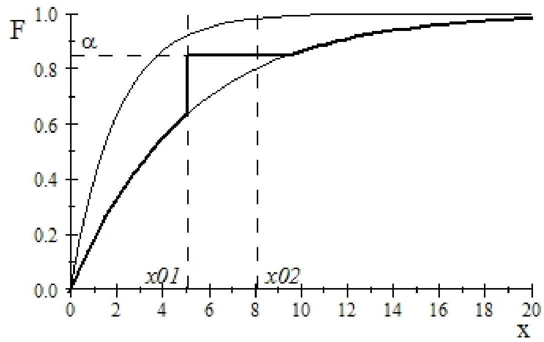

Figure 3: The optimal distribution (thick) for computing the upper expectation (the first case)

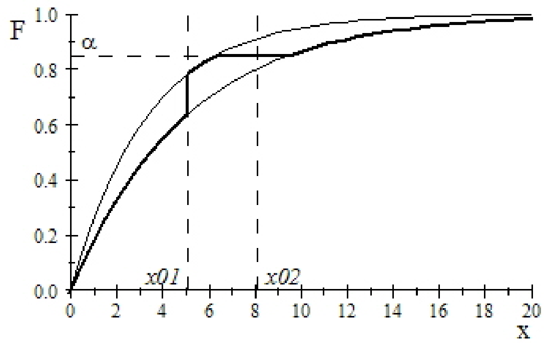

Figure 4: The optimal distribution (thick) for computing the upper expectation (the second case)

and

$$\mathrm{d}\psi_2(\alpha)/\mathrm{d}\alpha = 12\left(\ln\left(1-\alpha\right)\right)\left(39\ln^2\left(1-\alpha\right)+140\ln\left(1-\alpha\right)+127\right).$$

The equation $\mathrm{d}\psi_2(\alpha)/\mathrm{d}\alpha = 0$ *has one real root* 0. *However,* α *must belong to* $[\overline{F}(x_{01}), 1]$. *This implies that the function* $\psi_2(\alpha)$ *does not have the maximum in* $[\overline{F}(x_{01}), 1]$ *and*

$$\overline{\mathbb{E}}h = -250.98 + \psi_1(0.852) = -250.98 + 856.12 = 605.14.$$

Similarly, the lower bound $\underline{\mathbb{E}}h = -426.62$ *can similarly be computed.*

In the following, only expectations of functions having one maximum are studied because other cases of h can be derived in the same way.

It should be noted that lower and upper probability distributions are often obtained from other imprecise calculations and they are step functions in this case. This fact significantly simplifies the optimization problem because the step-wise functions produce a finite number of different intervals $[\overline{F}^{-1}(\alpha), \underline{F}^{-1}(\alpha)]$ and α takes the finite number of values.

6 P-boxes and Dempster-Shafer structures

In the previous sections, we have considered the distribution bounds of a general form. However, there is a class of distributions, which plays an important role in uncertainty modelling and representation. These distributions have the following form [13]:

$$\underline{F}(x)=\left\{\begin{array}{cc}0, & x<x_{*1}\\ \underline{F}(x_{*i}), & x_{*i}\leq x\leq x_{*i+1}\\ 1 & x_{*n}\leq x\end{array}\right.,\ \overline{F}(x)=\left\{\begin{array}{cc}0, & x<x_{1}^{*}\\ \overline{F}(x_{j+1}^{*}), & x_{j}^{*}\leq x\leq x_{j+1}^{*}\\ 1 & x_{m}^{*}\leq x\end{array}\right., \tag{8}$$

where $x_{*1}\leq...\leq x_{*n}$ and $x_1^*\leq...\leq x_m^*$ are ordered points.

In other words, the lower and upper distributions are right-continuous non-decreasing step functions from the reals into $[0,1]$ with at most n or m discontinuities located at the points $x_{*1},...,x_{*n}$ or $x_1^*,...,x_m^*$. Moreover, the functions $\underline{F}(x)$ and $\overline{F}(x)$ define a p-box for a function F if F satisfies (1). Therefore, our aim in this section is to consider lower and upper expectations of a non-monotone function of the random variable X whose lower and upper distributions can be represented in the form of (8). For simplicity, we study the expectations of the function having one maximum or minimum at point x_0.

Proposition 4 *If the function h has one maximum at point x_0 in $\mathbb{R}$ and the lower and upper distribution (8) are given, then the upper and lower expectations of $h(X)$ are*

$$\begin{aligned}\overline{\mathbb{E}}h&=h(x_0)\left[\overline{F}(x_0)-\underline{F}(x_0)\right]+\sum_{i:x_{*i}\leq x_0}h(x_{*i})\left[\underline{F}(x_{*i})-\underline{F}(x_{*i-1})\right]\\&+\sum_{j:x_j^*\geq x_0}h(x_j^*)\left[\overline{F}(x_{j+1}^*)-\overline{F}(x_j^*)\right],\end{aligned}\tag{9}$$

$$\underline{\mathbb{E}}h=\min_{\alpha\in\Phi}\left[\sum_{j:x_j^*\leq x_0}h(x_j^*)T_\alpha(x_j^*)+\sum_{i:x_{*i}\geq x_0}h(x_{*i})R_\alpha(x_{*i})\right]. \tag{10}$$

Here

$$T_\alpha(x_j^*)=\left\{\begin{array}{cc}\overline{F}(x_{j+1}^*)-\overline{F}(x_j^*), & \overline{F}(x_j^*)<\alpha\ \&\ \overline{F}(x_{j+1}^*)\leq\alpha\\ \alpha-\overline{F}(x_j^*), & \overline{F}(x_j^*)<\alpha\ \&\ \overline{F}(x_{j+1}^*)>\alpha\\ 0, & \textit{otherwise}\end{array}\right.,$$

$$R_\alpha(x_{*i})=\left\{\begin{array}{cc}\underline{F}(x_{*i})-\underline{F}(x_{*i-1}), & \underline{F}(x_{*i})>\alpha\ \&\ \underline{F}(x_{*i-1})\geq\alpha\\ \underline{F}(x_{*i})-\alpha, & \underline{F}(x_{*i})>\alpha\ \&\ \underline{F}(x_{*i-1})<\alpha\\ 0, & \textit{otherwise}\end{array}\right.,$$

$$\Phi=\left\{\overline{F}(x_{j+1}^*),\ j=1,...,m\right\}\cup\left\{\underline{F}(x_{*i}),\ i=1,...,n\right\}.$$

Proof. The upper expectation immediately follows from Proposition 1. Since the optimal density function in this case is a weighted sum of Dirac functions (with weights $\underline{F}(x_{*i})-\underline{F}(x_{*i-1})$ and $\overline{F}(x_{j+1}^*)-\overline{F}(x_j^*)$), having unit area concentrated in the immediate vicinity of some point (x_{*i} and x_{j+1}^*), then the integrals in (5) are replaced by sums. Let us consider the lower bound for the expectation. If we replace the unchanging parts of the lower distribution function corresponding to $x_{*i}\leq x\leq x_{*i+1}$ (see (8)) by lines

$$\overline{F}(x)=\frac{\varepsilon}{x_{*i+1}-x_{*i}}x+\overline{F}(x_{*i})-\frac{\varepsilon\cdot x_{*i}}{x_{*i+1}-x_{*i}},\ \varepsilon\to 0,$$

and replace the unchanging parts of the upper distribution function in the same way, then we can use Proposition 1 for computing the lower expectation. Note that α takes values from a finite set because the obtained optimal distribution is a step function. It is obvious that this set denoted by Φ consists of points $\underline{F}(x_{*i})$ and $\overline{F}(x_{j+1}^*)$. The coefficients $T_\alpha(x_j^*)$ and $R_\alpha(x_{*i})$ are nothing else but values of the weights of the density function corresponding to the optimal distribution. ■

Proposition 4 gives one of the possible ways for computing the lower and upper expectations of the non-monotone function h. However, the above result can be viewed from another point. It turns out that there is the close relationship between the above lower and upper distributions or p-boxes and belief functions in the framework of Dempster-Shafer theory. This relationship have been proved and studied by Yager [23], Ferson *et al* [8], Kriegler and Held [13].

Suppose that estimates of the random variable X are given in the form of closed intervals $A_i = [\underline{a}_i, \overline{a}_i]$, $i = 1, ..., n$. Let c_i denote the number of occurrences of the interval $A_i \subseteq \mathbb{R}$. At that, there holds $\sum_{i=1}^{n} c_i = N$. A frequency function m, called *basic probability assignment* (BPA), can be defined such that [4, 12, 18]:

$$m : 2^{\mathbb{R}} \to [0, 1],$$

$$m(\varnothing) = 1, \ \sum_{A \in \mathbb{R}} m(A) = 1.$$

According to [4], this function can be obtained as follows:

$$m(A_i) = c_i / N.$$

If $m(A_i) > 0$, i.e. A_i has occurred at least once, then A_i is called a *focal element*.

According to [18], the *belief* $Bel(A)$ and *plausibility* $Pl(A)$ measures of an event $A \subseteq \mathbb{R}$ can be defined as

$$Bel(A) = \sum_{A_i : A_i \subseteq A} m(A_i), \quad Pl(A) = \sum_{A_i : A_i \cap A \neq \varnothing} m(A_i).$$

As pointed out in [9], a belief function can formally be defined as a function satisfying axioms which can be viewed as a weakening of the Kolmogorov axioms that characterize probability functions. Therefore, it seems reasonable to understand a belief function as a generalized probability function [4] and the belief $Bel(A)$ and plausibility $Pl(A)$ measures can be regarded as lower and upper bounds for the probability of A, i.e., $Bel(A) \leq \Pr(A) \leq Pl(A)$.

The belief and plausibility functions can be considered in terms of a multivalued mapping [4]. Let $P(\omega)$ is a probability measure defined on a universal set Ω related to $\mathbb{R}$ through a multivalued mapping $G : \Omega \to 2^{\mathbb{R}}$. Then the BPA is defined as

$$m(A_i) = P(\omega_i) = c_i / N, \ \omega_i \in \Omega.$$

For each set $A_i \in \mathcal{P}(U)$, the value $m(A_i)$ expresses the probability of $\omega_i = G^{-1}(A_i)$. A *random set* is the pair $(\mathcal{F}, m)$, where $\mathcal{F}$ is the family of all N focal elements. If X_* is a subset of Ω such that $X_* = \{\omega \in \Omega : G(\omega) \subseteq A\}$, then the lower probability of A (belief function), according to Dempster's principle of inductive reasoning, is defined by $Bel(A) = P(X_*)$. If X^* is a subset of Ω such that $X^* = \{\omega \in \Omega : G(\omega) \cap A \neq \varnothing\}$, then the upper probability of A (plausibility function) is given by $Pl(A) = P(X^*)$.

On one hand, the lower and upper distributions corresponding to the available interval-valued estimates can be found as follows:

$$\underline{F}(x) = Bel((-\infty, x]) = \sum_{i : \overline{a}_i \leq x} c_i / N = \sum_{i : \overline{a}_i \leq x} m(A_i),$$

$$\overline{F}(x) = Pl((-\infty, x]) = \sum_{i : \underline{a}_i \leq x} c_i / N = \sum_{i : \underline{a}_i \leq x} m(A_i).$$

On the other hand, Ferson *et al* [8] emphasized that, by having a p-box, it is always possible to obtain from it a Dempster-Shafer structure that approximates the p-box. Moreover, Ferson *et al* [8], Kriegler and Held [13] proposed algorithms for constructing a random set from p-boxes.

Algorithm 1 (Kriegler and Held) *1. Initialize indices $k = 1$ (running over the focal elements of the random set to be constructed), $i = 1$ (running over x_{*i}), $j = 1$ (running over x_j^*). Let p_k denote the cumulative probability already accounted for in step k. Assign $p_0 = 0$.*

2. Construct focal element $A_k = (x_j^*, x_{*i}]$.

3. If $j = m$, choose arbitrary $x_{m+1}^* > x_m^*$, thus $\overline{F}(x_{m+1}^*) = 1$.

(a) $\underline{F}(x_{*i}) < \overline{F}(x_{j+1}^*)$: $m(A_k) = \underline{F}(x_{*i}) - p_{k-1}$, $p_k = \underline{F}(x_{*i})$. Raise indices $k \to k+1$, $i \to i+1$. Return to Step 2.

(b) $\underline{F}(x_{*i}) > \overline{F}(x^*_{j+1})$: $m(A_k) = \overline{F}(x^*_{j+1}) - p_{k-1}$, $p_k = \overline{F}(x^*_{j+1})$. *Raise indices* $k \to k+1$, $j \to j+1$. *Return to Step 2.*

(c) $\underline{F}(x_{*i}) = \overline{F}(x^*_{j+1})$: $m(A_k) = \overline{F}(x^*_{j+1}) - p_{k-1}$. *If* $\underline{F}(x_{*i}) = \overline{F}(x^*_{j+1}) = 1$, *stop. If* $\underline{F}(x_{*i}) = \overline{F}(x^*_{j+1}) < 1$, *set* $p_k = \overline{F}(x^*_{j+1})$. *Raise indices* $k \to k+1$, $i \to i+1$, $j \to j+1$. *Return to Step 2.*

Then Proposition 4 implies the following corollary.

Corollary 4 *If the function* h *has one maximum at point* x_0 *in* $\mathbb{R}$, *then the upper and lower expectations of* $h(X)$ *are*

$$\overline{\mathbb{E}}h = h(x_0) \sum_{i:\underline{a}_i \leq x_0 \leq \overline{a}_i} m(A_i) + \sum_{i:\overline{a}_i \leq x_0} h(\overline{a}_i) m(A_i) + \sum_{i:\underline{a}_i \geq x_0} h(\underline{a}_i) m(A_i), \tag{11}$$

$$\underline{\mathbb{E}}h = \min_{\alpha} \left[\sum_{i:\underline{a}_i \leq x_0} h(\underline{a}_i) T_\alpha(\underline{a}_i) + \sum_{i:\overline{a}_i \geq x_0} h(\overline{a}_i) R_\alpha(\overline{a}_i) \right]. \tag{12}$$

Here

$$T_\alpha(\underline{a}_i) = \begin{cases} m(A_i), & \sum_{k:\underline{a}_k \leq \underline{a}_i} m(A_k) \leq \alpha \ \& \ \sum_{k:\underline{a}_k < \underline{a}_i} m(A_k) < \alpha \\ \alpha - \sum_{k:\underline{a}_k < \underline{a}_i} m(A_k) & \sum_{k:\underline{a}_k \leq \underline{a}_i} m(A_k) > \alpha \ \& \ \sum_{k:\underline{a}_k < \underline{a}_i} m(A_k) < \alpha \\ 0, & \text{otherwise} \end{cases},$$

$$R_\alpha(\overline{a}_i) = \begin{cases} m(A_i), & \sum_{k:\overline{a}_k \leq \overline{a}_i} m(A_k) > \alpha \ \& \ \sum_{k:\overline{a}_k < \overline{a}_i} m(A_k) \geq \alpha \\ \sum_{k:\underline{a}_k \leq \underline{a}_i} m(A_k) - \alpha & \sum_{k:\overline{a}_k \leq \overline{a}_i} m(A_k) > \alpha \ \& \ \sum_{k:\overline{a}_k < \overline{a}_i} m(A_k) < \alpha \\ 0, & \text{otherwise} \end{cases},$$

$$\alpha \in \left\{ \sum_{i:\underline{a}_i \leq \underline{a}_k} m(A_i), \ k = 1, ..., n \right\} \cup \left\{ \sum_{i:\overline{a}_i{}_i \leq \overline{a}_k} m(A_i), \ k = 1, ..., n \right\}.$$

Proof. The proof immediately follows form Proposition 4 if we replace $\underline{F}(x_{*i}) - \underline{F}(x_{*i-1})$ by $m(A_i)$, $\overline{F}(x^*_{j+1}) - \overline{F}(x^*_j)$ by $m(A_j)$, x_{*i} by $\overline{a}_i$, x^*_j by $\underline{a}_j$. Moreover, there hold

$$\overline{F}(x_0) - \underline{F}(x_0) = \sum_{i:\underline{a}_i \leq x_0} m(A_i) - \sum_{i:x_0 \leq \overline{a}_i} m(A_i) = \sum_{i:\underline{a}_i \leq x_0 \leq \overline{a}_i} m(A_i),$$

$$\underline{F}(x_{*i}) = \sum_{k:\overline{a}_k \leq \overline{a}_i} m(A_k), \ \underline{F}(x_{*i-1}) = \sum_{k:\overline{a}_k < \overline{a}_i} m(A_k),$$

$$\overline{F}(x^*_{j+1}) = \sum_{k:\underline{a}_k \leq \underline{a}_i} m(A_k), \ \overline{F}(x^*_j) = \sum_{k:\underline{a}_k < \underline{a}_i} m(A_k).$$

■

It can be seen from comparison of (9)-(10) and (11)-(12) that all these expressions are dual to some extent.

At the same time, the lower and upper expectations in the framework of belief functions can be found in another way [16, 19]

$$\underline{\mathbb{E}}^{\#}h = \sum_{i=1}^{n} m(A_i) \inf_{\underline{a}_i \leq x \leq \overline{a}_i} h(x), \quad \overline{\mathbb{E}}^{\#}h = \sum_{i=1}^{n} m(A_i) \sup_{\underline{a}_i \leq x \leq \overline{a}_i} h(x). \tag{13}$$

At first sight, expressions (11)-(12) and (13) are quite different. However, the following proposition shows that (13) is nothing else than a form of (11)-(12).

Proposition 5 *Expressions (11)-(12) and (13) give the same results, i.e.,* $\overline{\mathbb{E}}h = \overline{\mathbb{E}}^{\#}h$ *and* $\underline{\mathbb{E}}h = \underline{\mathbb{E}}^{\#}h$.

Proof. Indeed, if we carefully look at (11), then we can see that

$$h(x_0)\sum_{i:\underline{a}_i\leq x_0\leq \overline{a}_i} m(A_i) = \sum_{i:x_0\in A_i} m(A_i)\sup_{\underline{a}_i\leq x\leq \overline{a}_i} h(x),$$

$$\sum_{i:\overline{a}_i\leq x_0} h(\overline{a}_i)m(A_i) = \sum_{i:\overline{a}_i\leq x_0} m(A_i)\sup_{\underline{a}_i\leq x\leq \overline{a}_i} h(x),$$

$$\sum_{i:\underline{a}_i\geq x_0} h(\underline{a}_i)m(A_i) = \sum_{i:\underline{a}_i\geq x_0} m(A_i)\sup_{\underline{a}_i\leq x\leq \overline{a}_i} h(x).$$

Therefore, there holds $\overline{\mathbb{E}}h = \overline{\mathbb{E}}^{\#}h$. The same can be said about the lower expectations. Indeed, if $\underline{a}_i \leq x_0 \leq \overline{a}_i$, then

$$\min\{h(\underline{a}_i), h(\overline{a}_i)\} = \inf_{\underline{a}_i\leq x\leq \overline{a}_i} h(x).$$

Hence

$$\min\{h(\underline{a}_i)m(A_i), h(\overline{a}_i)m(A_i)\} = m(A_i)\inf_{\underline{a}_i\leq x\leq \overline{a}_i} h(x).$$

It is easy to prove that $R_\alpha(\overline{a}_i) = 0$ if $T_\alpha(\underline{a}_i) = m(A_i)$ and $T_\alpha(\underline{a}_i) = 0$ if $R_\alpha(\overline{a}_i) = m(A_i)$. Consequently, there holds

$$\begin{aligned}&\min_\alpha\left[h(\underline{a}_i)T_\alpha(\underline{a}_i) + h(\overline{a}_i)R_\alpha(\overline{a}_i)\right]\\ &= \min\{h(\underline{a}_i)m(A_i), h(\overline{a}_i)m(A_i)\} = m(A_i)\inf_{\underline{a}_i\leq x\leq \overline{a}_i} h(x).\end{aligned}$$

The cases $T_\alpha(\underline{a}_i) = \alpha - \sum_{k:\underline{a}_k<\underline{a}_i} m(A_k)$ and $R_\alpha(\overline{a}_i) = \sum_{k:\underline{a}_k\leq\underline{a}_i} m(A_k) - \alpha$ are similarly analyzed. By extending the above sum on all intervals A_i, we get (13). ■

Example 4 *Let us return to Example 2 and suppose that the lower and upper distributions are of the form shown in Fig.5. Here* $x_{*1} = 5$, $x_{*2} = 10$, $x_{*3} = 12$, $x_1^* = 2$, $x_2^* = 4$, $x_3^* = 8$, $\underline{F}(x_{*1}) = 0.05$, $\underline{F}(x_{*2}) = 0.5$, $\underline{F}(x_{*3}) = 1$, $\overline{F}(x_1^*) = 0.1$, $\overline{F}(x_2^*) = 0.7$, $\overline{F}(x_1^*) = 1$. *The optimal distributions corresponding to the upper and lower expectations of the function* $h(x) = 60 - (x-5)^2$ *are shown as thick in Fig.7 and Fig.8, respectively. Hence, the lower and upper expectations of h are* $\underline{\mathbb{E}}h = 23.8$ *and* $\overline{\mathbb{E}}h = 57.3$, *respectively.*

By using the algorithm proposed by Kriegler and Held [13], we can write the corresponding intervals (see Fig.6) and BPAs:

$$A_1 = [2,5], A_2 = [2,10], A_3 = [4,10], A_4 = [4,12], A_5 = [8,12],$$

$$c_1 = 1, c_2 = 1, c_3 = 8, c_4 = 4, c_5 = 6,$$

$$m(A_1) = 0.05,\ m(A_2) = 0.05,\ m(A_3) = 0.4,\ m(A_4) = 0.2,\ m(A_5) = 0.3.$$

Hence, we find the lower and upper expectations of the function h by using (13) as follows:

$$\underline{\mathbb{E}}^{\#}h = 0.05h(2) + 0.05h(10) + 0.4h(10) + 0.2h(12) + 0.3h(12) = 23.8,$$

$$\overline{\mathbb{E}}^{\#}h = 0.05h(5) + 0.05h(5) + 0.4h(5) + 0.2h(5) + 0.3h(8) = 57.3$$

7 Possibility distributions

Uncertainty is often measured by possibility measures [6, 25]. A possibility measure Π on the power set of a universal set Ω is defined by

$$\Pi(A\cup B) = \max\{\Pi(A), \Pi(B)\}$$

for any subsets A and B of the power set. The function Π can be represented by a single possibility distribution π

$$\Pi(A) = \sup\{\pi(x) : x\in A\}.$$

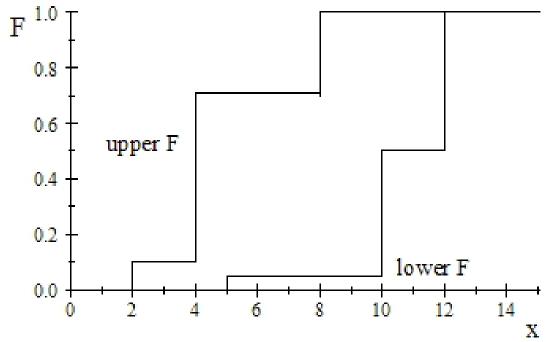

Figure 5: Lower and upper distributions

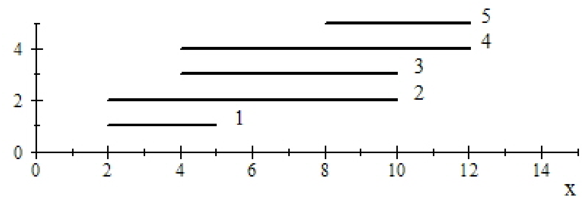

Figure 6: Interval-valued estimates

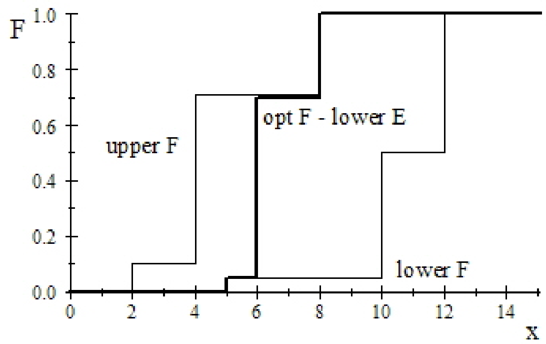

Figure 7: Relationship between the optimal distribution for computing the lower expectation and p-boxes

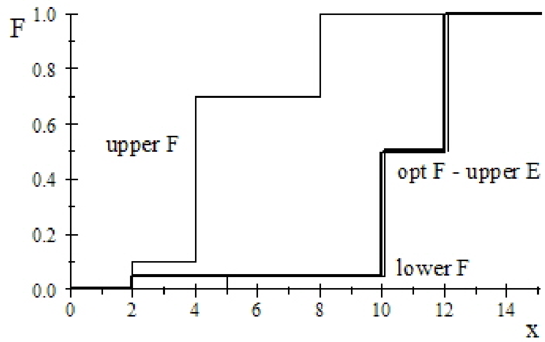

Figure 8: Relationship between the optimal distribution for computing the upper expectation and p-boxes

The possibility distribution π is normalized by the condition that there exists a $x^* \in \Omega$ with $\pi(x^*) = 1$. The possibility measure also has the following properties: $0 \leq \Pi(A) \leq 1$, $\Pi(\Omega) = 1$, $\max\{\pi(A), \Pi(A^c)\} = 1$. A possibility distribution function is analogous to a probability mass function or density function, and a possibility measure is analogous to a probability measure. We will mainly consider the case when $\Omega = \mathbb{R}$.

It is shown by Dubois and Prade [7, 22] that the possibility measure can be regarded as an upper probability measure

$$\overline{P}(A) = \Pi(A) = \sup\{\pi(x) : x \in A\}.$$

The corresponding lower probabilities are defined by

$$\underline{P}(A) = 1 - \Pi(A^c) = 1 - \sup\{\pi(x) : x \in A^c\}.$$

By taking $\Omega = \mathbb{R}$ and considering the events $A = (-\infty, x]$, we can construct the lower and upper probability distributions associated with the possibility distribution π as follows:

$$\underline{F}(x) = \underline{P}((-\infty, x]) = \begin{cases} 1 - \pi(x), & x \geq x^* \\ 0, & x < x^* \end{cases},$$

$$\overline{F}(x) = \overline{P}((-\infty, x]) = \begin{cases} \pi(x), & x \leq x^* \\ 1, & x > x^* \end{cases}.$$

Hence, the possibilistic lower and upper expectations[2] of the function h can be defined by using (2). If the function h has one maximum, then it follows from (5) that the following three cases can be studied. If $x_0 < x^*$, then

$$\overline{\mathbb{E}}h = h(x_0)\overline{F}(x_0) + \int_{x_0}^{x^*} h(x)\mathrm{d}\overline{F}(x).$$

If $x_0 > x^*$, then

$$\overline{\mathbb{E}}h = h(x_0)\left[1 - \underline{F}(x_0)\right] + \int_{x^*}^{x_0} h(x)\mathrm{d}\underline{F}(x).$$

If $x_0 = x^*$, then $\overline{\mathbb{E}}h = h(x_0)$. By taking into account the fact that

$$\pi(x) = \begin{cases} 1 - \underline{F}(x), & x \geq x^* \\ \overline{F}(x), & x < x^* \end{cases},$$

we get

$$\overline{\mathbb{E}}h = \begin{cases} h(x_0)\pi(x_0) + \int_{x_0}^{x^*} h(x)\mathrm{d}\pi(x), & x_0 < x^* \\ h(x_0) & x_0 = x^* \\ h(x_0)\pi(x_0) - \int_{x^*}^{x_0} h(x)\mathrm{d}\pi(x), & x_0 > x^* \end{cases}.$$

[2]It should be noted that another definition of the possibilistic expectations is given by Halpern and Pucella [10]. Halpern and Pucella used the well-known interpretation [5] of possibility measures as a special case of plausibility functions.

The lower expectation can be similarly obtained by using (6)

$$\underline{\mathbb{E}}h=\min_{\alpha\in[0,1]}\left[\int_{-\infty}^{\pi_{-}^{-1}(\alpha)}h(x)\mathrm{d}\pi(x)-\int_{(1-\pi_{+})^{-1}(\alpha)}^{\infty}h(x)\mathrm{d}\pi(x)\right].$$

Here π_{-} and π_{+} are the left and right branches of the possibility distribution, respectively.

Example 5 *For practical purposes, one of the most used forms of possibility distributions is the triangular form proposed by Dubois and Prade [6] and represented as*

$$\pi(x)=\left\{\begin{array}{cc}(x-a_1)/(x^*-a_1), & a_1<x\leq x^*\\ (a_2-x)/(a_2-x^*), & x^*<x\leq a_2\\ 0, & otherwise\end{array}\right..$$

Suppose $h=60-(x-5)^2$, $a_1=2$, $x^*=3$, $a_2=10$. *Then*

$$\overline{\mathbb{E}}h=\left(60-(5-5)^2\right)\frac{10-5}{10-3}+\int_{3}^{5}\left(60-(x-5)^2\right)\frac{-1}{10-3}\mathrm{d}x=59.62.$$

Note that $\pi_{-}^{-1}(\alpha)=\alpha\left(x^*-a_1\right)+a_1$ *and* $(1-\pi_{+})^{-1}(\alpha)=\alpha\left(a_2-x^*\right)+x^*$. *It follows from Corollary 1 that*

$$x_0-\left(\alpha\left(x^*-a_1\right)+a_1\right)=\alpha\left(a_2-x^*\right)+x^*-x_0.$$

Hence

$$\alpha=\frac{1}{a_1-a_2}\left(a_1+x^*-2x_0\right)=0.625,$$

$$\pi_{-}^{-1}(0.625)=2.625,\ (1-\pi_{+})^{-1}(0.625)=7.375,$$

and

$$\underline{\mathbb{E}}h=\int_{2}^{2.625}\left(60-(x-5)^2\right)\frac{1}{(3-2)}\mathrm{d}x+\int_{7.375}^{10}\left(60-(x-5)^2\right)\frac{1}{10-3}\mathrm{d}x=50.15.$$

8 Approximate computing the lower and upper expectations

If we can find the lower and upper expectations of arbitrary functions h by means of very simple expressions (13) in terms of belief functions under condition that the corresponding lower and upper probability distributions are step functions, then it makes sense to approximate arbitrary lower and upper probability distributions by step functions and to find approximate belief functions corresponding to the initial lower and upper distributions. By having the approximate belief functions, we can find the lower and upper expectations of the function h by using (13). Of course, this approach can be useful when the integrals in (6) and (5) can not be calculated in the explicit form or it is difficult to explicitly write the inverse functions $\underline{F}^{-1}(\alpha)$ and $\overline{F}^{-1}(\alpha)$. The same idea for realizing some mathematical operations with p-boxes has been proposed by Ferson *et al* [8]. The authors [8] pointed out that when the p-box has curves rather than step functions for its bounds, a discretization is necessary to produce the associated Dempster-Shafer structure, which will therefore be an approximation to the p-box.

A possible approximation to continuous lower and upper distribution functions is shown in Fig.9. This approximation was proposed by Kriegler and Held [13] and presupposes equal intervals of values of X. Another approximation (see Fig.10) was proposed by Ferson *et al* [8] and presupposes equal probability masses of the focal elements in order to provide equal BPAs of all intervals.

By using the second approximation, which produces equal BPAs, and taking M points for the discretization, we get the lower and upper expectations of the form:

$$\underline{\mathbb{E}}h=\sum_{i=1}^{M}m(A_i)\inf_{x\in A_i}h(x)=M^{-1}\sum_{i=1}^{M}\inf_{x\in A_i}h(x),$$

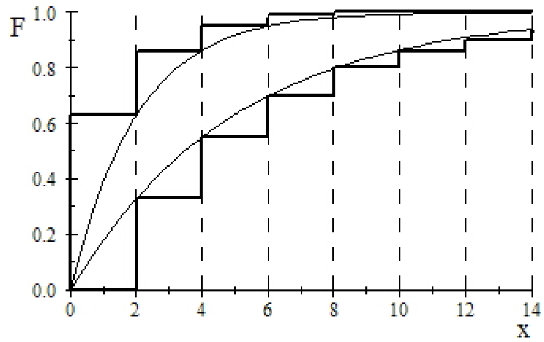

Figure 9: Approximation of lower and upper distributions by step functions

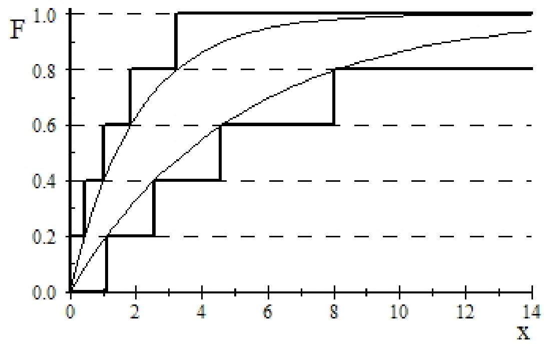

Figure 10: Approximation of lower and upper distributions by step functions

$$\overline{\mathbb{E}}h=\sum_{i=1}^{M}m(A_i)\sup_{x\in A_i}h(x)=M^{-1}\sum_{i=1}^{M}\sup_{x\in A_i}h(x).$$

Here $m(A_i)=\alpha_i-\alpha_{i-1}=1/M$, $A_i=[\overline{F}^{-1}(\alpha_{i-1}),\underline{F}^{-1}(\alpha_{i+1})]$, $\alpha_0=0$.

The above approximation does not depend on the form of the function h. Note that the same procedure can be used for computing possibilistic expectations.

9 Conclusion

Methods for computing the bounds for expected utilities under partial information about states of nature in the form of condition (1) and by non-monotone utility functions have been proposed in the paper. It has been shown that, for some types of the utility function, the lower and upper expectations can be exactly found without solving hard optimization problems. At the same time, the approximate methods can be useful and they allow us to approximately compute the lower and upper expectations also without solving hard optimization problems.

We have analyzed only decision problems defined by one random variable. However, risk analysis usually deals with a number of factors whose impact is represented by a function of random variables and these variables can be statistically independent. When the function of the random variables is monotone, risk analysis does not meet any difficulties. A direction for further work is the development of methods for risk analysis in case of a lack of monotonicity.

By viewing risk analysis in the decision context, one can see that the developed methods are used for computing only unrandomized actions (pure strategy). However, randomized actions (mixed strategy) could give sometimes better solutions to some extent. Therefore, another direction for further work could be possible extensions of proposed methods for searching optimal randomized actions.

References

[1] T. Augustin. On decision making under ambiguous prior and sampling information. In G. de Cooman, T.L. Fine, and T. Seidenfeld, editors, *Imprecise Probabilities and Their Applications. Proc. of the 2nd Int. Symposium ISIPTA'01*, pages 9–16, Ithaca, USA, June 2001. Shaker Publishing.

[2] T. Augustin. Expected utility within a generalized concept of probability - a comprehensive framework for decision making under ambiguity. *Statistical Papers*, 43:5–22, 2002.

[3] T. Aven. *Foundations of Risk Analysis: A Knowledge and Decision-Oriented Perspective*. Willey, England, 2003.

[4] A.P. Dempster. Upper and lower probabilities induced by a multi-valued mapping. *Annales of Mathematical Statistics*, 38:325–339, 1967.

[5] D. Dubois and H. Prade. On several representations of an uncertain body of evidence. In M.M. Gupta and E. Sanchez, editors, *Fuzzy Information and Decision Processes*, page 167181. North-Holland, Amsterdam, 1982.

[6] D. Dubois and H. Prade. *Possibility Theory: An Approach to Computerized Processing of Uncertainty*. Plenum Press, New York, 1988.

[7] D. Dubois and H. Prade. When upper probabilities are possibility measures. *Fuzzy Sets and Systems*, 49:65–74, 1992.

[8] S. Ferson, V. Kreinovich, L. Ginzburg, D.S. Myers, and K. Sentz. Constructing probability boxes and Dempster-Shafer structures. Report SAND2002-4015, Sandia National Laboratories, January 2003.

[9] J.Y. Halpern and R. Fagin. Two views of belief: Belief as generalized probability and belief as evidence. *Artificial Intelligence*, 54:275–317, 1992.

[10] J.Y. Halpern and R. Pucella. Reasoning about expectation. In *Proceedings of the Eighteenth Conference on Uncertainty in AI*, pages 207–215, San Francisco, July 2002. Morgan Kaufmann.

[11] R.L. Keeney and H. Raiffa. *Decisions with Multiple Objectives: Preferences and Value Tradeoffs.* John Wiley and Sons, NY, 1976.

[12] G.J. Klir and M.J. Wierman. *Uncertainty-Based Information: Elements of Generalized Information Theory.* Physica-Verlag, Heidelberg, 1998.

[13] E. Kriegler and H. Held. Utilizing belief functions for the estimation of future climate change. *International Journal of Approximate Reasoning*, 39:185–209, 2005.

[14] V. P. Kuznetsov. *Interval Statistical Models.* Radio and Communication, Moscow, 1991. in Russian.

[15] R.F. Nau. Uncertainty aversion with second-order probabilities and utilities. In G. de Cooman, T.L. Fine, and T. Seidenfeld, editors, *Imprecise Probabilities and Their Applications. Proc. of the 1st Int. Symposium ISIPTA'01*, pages 273–283, Ithaca, USA, June 2001. Shaker Publishing.

[16] H.T. Nguyen and E.A. Walker. On decision making using belief functions. In R.Y. Yager, M. Fedrizzi, and J. Kacprzyk, editors, *Advances in the Dempster-Shafer theory of evidence*, pages 311–330. Wiley, New York, 1994.

[17] W.L. Oberkampf, J.C. Helton, C.A. Joslyn, S.F. Wojtkiewicz, and S. Ferson. Challenge problems: uncertainty in system response given uncertain parameters. *Reliability Engineering and System Safety*, 85:11–19, 2004.

[18] G. Shafer. *A Mathematical Theory of Evidence.* Princeton University Press, 1976.

[19] T.M. Strat. Decision analysis using belief functions. *International Journal of Approximate Reasoning*, 4(5):391–418, 1990.

[20] L.V. Utkin and Th. Augustin. Efficient algorithms for decision making under partial prior information and general ambiguity attitudes. In T. Seidenfeld F.G. Cozman, R. Nau, editor, *Proc. of the 4th Int. Symposium on Imprecise Probabilities and Their Applications, ISIPTA'05*, pages 349–358, Pittsburgh, USA, July 2005. Carnegie Mellon University, SIPTA.

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

[22] P. Walley. Measures of uncertainty in expert systems. *Artificial Intelligence*, 83:1–58, 1996.

[23] R.R. Yager. Arithmetic and other operations on Dempster-Shafer structures. *International Journal of Man-Machine Studies*, 25:357–366, 1986.

[24] R.R. Yager. Modeling uncertainty using partial information. *Information Sciences*, 121:271–294, 1999.

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