Different Faces of the Natural Extension

Lev V. Utkin
Department of Computer Science
Forest Technical Academy,
St.Petersburg, Russia
lvu@utkin.usr.etu.spb.ru

Igor O. Kozine
Risoe National Laboratory
Systems Analysis Department,
Roskilde, Denmark
igor.kozine@risoe.dk

Abstract

The natural extension, the key concept for the construction of coherent imprecise models, can appear in different equivalent forms. Each of them has pros and cons in the context of specific applications. The use of a proper form can substantially facilitate the inference and computation of the previsions of interest.

The current paper concerns four forms of the natural extension representation. It is demonstrated that all of them are equivalent and one, discussed in the last instance, is prominent solely for gambles defined on continuous possibility sets. It is proven that the solution of the natural extension problem for continuous gambles exists on the degenerate distributions. Partial information and a characteristic to be calculated can be thought as the expectations of some real-valued functions defined on the possibility space. As they are expectations, they can be expressed as functions of probability density functions and proper real-valued functions (gambles). Each piece of the partial information acts as a constraint to the probability distributions. All together the constraints define the area of all distributions over which the interval of the desired statistical characteristic will be searched. Throughout the paper the natural extension is analyzed through the prism of reliability application.

Keywords. Imprecise probability theory, imprecise reliability, natural extension, previsions.

1 Introduction

Having tried to implement the theory of coherent imprecise previsions into practices related to reliability analysis we found that the natural extension, the key concept for the construction of coherent models, can appear in different equivalent forms. Each of them has pros and cons in the context of specific applications. The use of a proper form can substantially facilitate the inference and computation of the previsions of interest. The knowledge of different natural extension faces lets us understand better the essence of the theory and see possible ways of its development and implementation.

The current paper concerns four forms of the natural extension representation. It is demonstrated that all of them are equivalent and one, discussed in the last instance, is prominent solely for gambles defined on continuous possibility sets. It is proven that the solution of the natural extension problem for continuous gambles exists on the degenerate distributions. In order to distinguish between the different forms of the natural extension one of them is referred to as ”the primal form” (analogously to the terminology used in linear and non-linear programming); the second one is called ”Kuznetsov’s form” (first appeared in the book [5]) and it will be easy to see that this form is just a different notation for the dual optimization programming problem; the third representation is called ”Walley’s form” [9] which is probably best fitting into the behavioral interpretation; and the fourth one is referred to as ”the degenerate form”.

Informally, the problem of the natural extension for a one-dimensional case can be formulated as follows: one has some partial information on a probabilistic parameter $f(x)$ and is interested in the value of some other probabilistic characteristics. We define partial information as evidence which imposes constraints to the set of possible probability distributions without narrowing it to a single (precise) distribution. Since the given information is partial, we do not expect precise numbers but we would like to know an interval of possible values of a desired characteristic.

This definition of partial information assumes that some evidence can be precisely known but characterizing only, for example, some events on the possibility space, and, as a consequence, the probability distribution is not known completely but partially. Some evidence can be of interval-valued or comparative form. There may be direct evidence on a class of distributions etc.

We assume that partial information and a characteristic to be calculated can be thought as the expectations of some real-valued functions defined on the possibility space. As they are expectations, they can be expressed as functions of probability density functions and proper real-valued functions (gambles). Each piece of the partial information acts as a constraint to the probability distributions. All together the constraints will define the area of all distributions over which the interval of the desired statistical characteristic will be searched.

Throughout the paper the natural extension will be analyzed through the prism of reliability application highlighting our area of expertise. Nevertheless, the subject of the paper is wider and not necessarily be boiled down to a reliability analysis.

2 Primal form

Without loss of generality we consider the primal form of the natural extension only in the integral form tacitly assuming that the discrete case can be easily written.

Consider a system consisting of n components. Let $\varphi_{ij}(x_i)$ be a function of the random variable x_i, for example, the i-th component lifetime. According to [1], the system lifetime can be uniquely determined by the component lifetimes. Denote $\mathbf{X} = (x_1, ..., x_n)$. Then there exists a function $g(\mathbf{X})$ of the components lifetimes characterizing the system reliability behavior. The functions $\varphi_{ij}(x_i)$ and $g(\mathbf{X})$ can be regarded as gambles, where a gamble is a real-valued function on a possibility space whose value is uncertain [9, 5]. Suppose that partial information is represented as a set of lower and upper previsions $\underline{a}_{ij} = \underline{M}(\varphi_{ij}(x_i))$ and $\overline{a}_{ij} = \overline{M}(\varphi_{ij}(x_i))$, $i = 1, ..., n$, $j = 1, ..., m_i$. Here m_i is the number of judgements that are related to the i-th component. In this case, the natural extension in its primal form is

$$\underline{M}(g) = \min_{\mathcal{P}} \int_{\mathbf{R}_+^n} g(\mathbf{X})\rho(\mathbf{X})\mathrm{d}\mathbf{X}, \quad (1)$$

$$\overline{M}(g) = \max_{\mathcal{P}} \int_{\mathbf{R}_+^n} g(\mathbf{X})\rho(\mathbf{X})\mathrm{d}\mathbf{X}, \quad (2)$$

subject to

$$\int_{\mathbf{R}_+^n} \rho(\mathbf{X})\mathrm{d}\mathbf{X} = 1,\ \rho(\mathbf{X}) \geq 0,$$

$$\underline{a}_{ij} \leq \int_{\mathbf{R}_+^n} \varphi_{ij}(x_i)\rho(\mathbf{X})\mathrm{d}\mathbf{X} \leq \overline{a}_{ij},\ i \leq n,\ j \leq m_i. \quad (3)$$

Here the minimum and maximum are taken over the set $\mathcal{P}$ of all possible n-dimensional density functions $\{\rho(\mathbf{X})\}$ satisfying conditions (3). Throughout the paper the natural extension will be written only for the lower bound.

The natural extension in its primal form has an advantage of easy interpretation. If the distribution $\rho(\mathbf{X})$ is not known precisely, one has to solve problem (1)-(3) and seek for the upper and lower bounds of the characteristic of interest. The solution of this problem is defined on the set of possible densities $\mathcal{P}$ that are consistent with partial information expressed in the form of constraints (3). If one does not have any information on probability characteristics defined on the possibility space $\mathbf{X}$ or on the class of densities $\mathcal{P}$, then the set $\mathcal{P}$ is the largest and the solution is vacuous, i.e. $\underline{M}(g) = \inf g(\mathbf{X})$ and $\overline{M}(g) = \sup g(\mathbf{X})$. Given evidence reducing the set $\mathcal{P}$, the interval $[\underline{M}(g), \overline{M}(g)]$ becomes more narrow and different from the vacuous one.

Assume there is a comparative judgement "the prevision of a gamble $\gamma_k(\mathbf{x}_k)$ does not exceed the prevision of a gamble $\beta_k(\mathbf{y}_k)$". Here $\mathbf{x}_k$ and $\mathbf{y}_k$ are the subvectors of $\mathbf{X}$. This judgement can be written as

$$\int_{\mathbf{R}_+^n} \gamma_k(\mathbf{x}_k)\rho(\mathbf{X})\mathrm{d}\mathbf{X} \geq \int_{\mathbf{R}_+^n} \beta_k(\mathbf{y}_k)\rho(\mathbf{X})\mathrm{d}\mathbf{X}$$

and act as an additional constraint to problem (1)-(3).

It should be noted that only joint densities are used in the optimization problem (1)-(3) because in a general case we may not be aware whether the variables $x_1, ..., x_n$ are dependent or not. If it is known that the components are independent, then $\rho(\mathbf{X}) = \rho(x_1) \cdots \rho(x_n)$. Here we use the definition of independence in the sense of classical probability theory. In this case the set $\mathcal{P}$ is reduced and consists only of the densities that can be represented as a product. The optimization problem for computing a new lower prevision is of the form:

$$\underline{M}(g) = \inf_{\mathcal{P}} \int_{\mathbf{R}_+^n} g(\mathbf{X})\rho_1(x_1) \cdots \rho_n(x_n)\mathrm{d}\mathbf{X}, \quad (4)$$

subject to

$$\rho_i(x_i) \geq 0,\ \int_0^\infty \rho_i(x_i)\mathrm{d}x_i = 1,$$

$$\underline{a}_{ij} \leq \int_0^\infty \varphi_{ij}(x_i)\rho_i(x_i)\mathrm{d}x_i \leq \overline{a}_{ij},\ i \leq n,\ j \leq m_i. \quad (5)$$

In this case the problem becomes non-linear, which makes it more difficult to solve. However, some reliability problems can be easily solved by using the primal form.

Example 1 *Consider a series system consisting of n independent components. A system is called series if its lifetime x_s is given by $g(x_1, ..., x_n) = \min_{i=1,...,n} x_i$, where x_i is the i-th component lifetime [1]. Let $\underline{a}_i = \underline{M}(x_i)$ and $\overline{a}_i = \overline{M}(x_i)$ be the lower and upper mean times to failure of the i-th component, $i = 1, ..., n$. Moreover, suppose that $0 \leq x_i \leq T$, $i = 1, ..., n$. Let us find the lower mean time to failure of the system $\underline{M}(\min_{i=1,...,n} x_i)$. The optimization problem for computing new lower prevision is of* *the form:*

$$\underline{M}(g) = \inf_{\mathcal{P}} \int_{\Omega} \min(x_1, ..., x_n)\rho_1(x_1)\cdots\rho_n(x_n)\mathrm{d}\mathbf{X},$$

subject to

$$\rho_i(x_i) \geq 0,\ \int_0^T \rho_i(x_i)\mathrm{d}x_i = 1,$$

$$\underline{a}_i \leq \int_0^T x_i\rho_i(x_i)\mathrm{d}x_i \leq \overline{a}_i,\ i \leq n.$$

Here $\Omega = [0,T] \times ... \times [0,T]$.

According to [1], we can rewrite the objective function and constraints as follows:

$$\underline{M}(g) = \int_0^T \underline{F}_1(x)\cdots\underline{F}_n(x)\mathrm{d}x,$$

where $\underline{F}_i(x)$ *is the lower survivor function of the* i*-th component lifetime such that* $\int_0^T \underline{F}_i(x)\mathrm{d}x = \underline{a}_i$.

By using Chebyshev's inequality, we have

$$\int_0^T \prod_{i=1}^n \underline{F}_i(x)\mathrm{d}x \geq \frac{1}{T^{n-1}}\prod_{i=1}^n \int_0^T \underline{F}_i(x)\mathrm{d}x.$$

Hence minimally coherent solution is

$$\underline{M}(g) = \frac{1}{T^{n-1}}\prod_{i=1}^n \underline{a}_i.$$

3 Kuznetsov's and Walley's forms

Problem (1)-(3) has infinitely many variables and it can hardly be solved directly. However, we can apply some known tricks that help solve these problems. One of the pivotal tools in linear programming is the concept of duality. According to the duality theorem (see, for example, [6]) the minimum in the above-stated problem coincides with the maximum attained in the dual problem which in many cases is much easier to solve.

The transition from the primal problem to the dual one receives coverage in the literature only for the discrete case. In employing the discrete results we will demonstrate the transition to the dual problem for the continuous case.

Let us rewrite problem (1)-(3) in the form of finite difference

$$\underline{M}^*(g) = \inf_{\mathcal{P}} \sum_{K=1}^n g(\mathbf{X}^{(K)})\rho(\mathbf{X}^{(K)})\triangle\mathbf{X}_K,$$

subject to

$$\rho(\mathbf{X}^{(K)}) \geq 0,\ \sum_{K=1}^n \rho(\mathbf{X}^{(K)})\triangle\mathbf{X}_K = 1,$$

$$\begin{aligned} \underline{a}_{ij} &\leq \sum_{K=1}^n \varphi_{ij}(x_i^{(k_i)})\rho(\mathbf{X}^{(K)})\triangle\mathbf{X}_K \leq \overline{a}_{ij}, \\ i &\leq n,\ j \leq m_i. \end{aligned}$$

Here $K = (k_1, ..., k_n)$, $\mathbf{X}^{(K)} = (x_1^{(k_1)}, ..., x_n^{(k_n)})$, $\triangle\mathbf{X}_K = \triangle x_1^{(k_1)}\cdots\triangle x_n^{(k_n)}$. This is the linear programming problem with finite number of variables $\rho(\mathbf{X}^{(K)})$. Therefore, we can write the dual optimization problem for computing the lower prevision as follows:

$$\underline{M}^*(g) = \sup_{c,c_{ij},d_{ij}} \left(c + \sum_{i=1}^n \sum_{j=1}^{m_i} \left(c_{ij}\underline{a}_{ij} - d_{ij}\overline{a}_{ij}\right) \right),$$

subject to $c_{ij} \in \mathbf{R}_+, d_{ij} \in \mathbf{R}_+, c \in \mathbf{R}$ and $\forall \mathbf{X}^{(K)} \geq 0$

$$c + \sum_{i=1}^n \sum_{j=1}^{m_i} (c_{ij} - d_{ij})\varphi_{ij}(x_i^{(k_i)}) \leq g(\mathbf{X}^{(K)}).$$

As it is seen, the density functions are not variables anymore. By using the passage to limit as $\max_K \triangle\mathbf{X}_K \to 0$, we obtain the following problem:

$$\underline{M}(g) = \sup_{c,c_{ij},d_{ij}} \left(c + \sum_{i=1}^n \sum_{j=1}^{m_i} \left(c_{ij}\underline{a}_{ij} - d_{ij}\overline{a}_{ij}\right) \right), \quad (6)$$

subject to $c_{ij} \in \mathbf{R}^+$, $d_{ij} \in \mathbf{R}^+$, $c \in \mathbf{R}$ and $\forall \mathbf{X} \geq 0$

$$c + \sum_{i=1}^n \sum_{j=1}^{m_i} (c_{ij} - d_{ij})\varphi_{ij}(x_i) \leq g(\mathbf{X}). \quad (7)$$

Problem (6)-(7) can be rewritten in another form:

$$\underline{M}(g) = \sup_{c,c_{ij}} \left(c + \sum_{i=1}^n \sum_{j=1}^{m_i} c_{ij}a_{ij} \right), \quad (8)$$

subject to $c_{ij} \in \mathbf{R}$, $c \in \mathbf{R}$ and $\forall \mathbf{X} \geq 0$

$$c + \sum_{i=1}^n \sum_{j=1}^{m_i} c_{ij}\varphi_{ij}(x_i) \leq g(\mathbf{X}), \quad (9)$$

where

$$a_{ij} = \begin{cases} \underline{a}_{ij}, & c_{ij} \geq 0 \\ \overline{a}_{ij}, & c_{ij} < 0 \end{cases}.$$

Thus, we have arrived at the natural extension in Kuznetsov's form [5] which is valid for both continuous and discrete case.

The natural extension in the dual (Kuznetsov's) form of a linear optimization problem is a constructive tool and allows us to solve various applied problems. As a matter of fact, most of analytical results obtained in reliability analysis by the authors of the current paper have been inferred owing to the dual form of the natural extension. However, this representation has some limitations. For instance, independence relationships cannot be introduced to the dual form since it is valid only for linear programming problems. On the contrary, independence relationships can be easily introduced through the primal form at the cost of having a non-linear optimization problem.

Suppose only the lower previsions $\underline{M}(\varphi_{ij}(x_i)) = \underline{a}_{ij}$, $i = 1, ...n$, $j = 1, ..., m_i$, are known. Let us introduce new gambles

$$\psi_{ij}(x_i) = \varphi_{ij}(x_i) - \underline{M}(\varphi_{ij}(x_i)) = \varphi_{ij}(x_i) - \underline{a}_{ij}.$$

They are called almost desirable gambles [9] or centered gambles [5]. Then $\underline{M}(\psi_{ij}) = 0$ and we can rewrite problem (6)-(7) as follows:

$$\underline{M}(g) = \sup_{c, c_{ij}} \left(c + \sum_{i=1}^{n} \sum_{j=1}^{m_i} c_{ij} \underline{M}(\psi_{ij}) \right) = \sup c, \quad (10)$$

subject to $c_{ij} \in \mathbf{R}^+$, $c \in \mathbf{R}$ and

$$c + \sum_{i=1}^{n} \sum_{j=1}^{m_i} c_{ij} \psi_{ij}(x_i) \leq g(\mathbf{X}), \ \forall \mathbf{X} \geq 0.$$

Let us represent constraints as

$$g(\mathbf{X}) - c \geq \sum_{i=1}^{n} \sum_{j=1}^{m_i} c_{ij} \psi_{ij}(x_i), \ \forall \mathbf{X} \geq 0. \quad (11)$$

Optimization problem (10)-(11) can be written in the different form

$$\underline{M}(g) = \sup \left\{ c : g(\mathbf{X}) - c \geq \sum_{i=1}^{n} \sum_{j=1}^{m_i} c_{ij} \psi_{ij}(x_i) \right\}. \quad (12)$$

If only the upper previsions $\overline{M}(\varphi_{ij}(x_i)) = \overline{a}_{ij}$ are known, then $\forall \mathbf{X} \geq 0$,

$$\overline{M}(g) = \inf \left\{ c : g(\mathbf{X}) - c \leq \sum_{i=1}^{n} \sum_{j=1}^{m_i} c_{ij} \phi_{ij}(x_i) \right\}, \quad (13)$$

where

$$\phi_{ij}(x_i) = \overline{M}(\varphi_{ij}(x_i)) - \varphi_{ij}(x_i) = \overline{a}_{ij} - \varphi_{ij}(x_i).$$

Representations (12) and (13) are nothing else than the natural extension in Walley's form.

Example 2 *Consider a series system consisting of n components. Let $\underline{a}_i = \underline{M}(I_{[t,\infty)}(x_i))$ and $\overline{a}_i = \overline{M}(I_{[t,\infty)}(x_i))$ be the lower and upper reliabilities at time t of the i-th component, $i = 1, ..., n$. (As the reliability of a component/system within a predefined time interval $[0, t]$ is the probability that time to failure belongs to $[t, \infty)$, then it could be written as $M(I_{[t,\infty)}(x))$, where x is time to failure). There is no information about the independence of the components. Let us find the lower reliability of the system $\underline{M}(I_{[t,\infty)}(\min_{i=1,...,n} x_i))$. It follows from (8)-(9) that*

$$\underline{M}(I_{[t,\infty)}(\min_{i=1,...,n} x_i)) = \sup_{c, c_i} \left(c + \sum_{i=1}^{n} c_i a_i \right),$$

subject to $c_i \in \mathbf{R}$, $c \in \mathbf{R}$ and $\forall \mathbf{X} \geq 0$

$$c + \sum_{i=1}^{n} c_i I_{[t,\infty)}(x_i) \leq I_{[t,\infty)}(\min_{i=1,...,n} x_i),$$

It can be seen from the constraints that the set of all values of x_i can be divided into to subsets $[0, t)$ and $[t, \infty)$. If $x_i \in [0, t)$, then $I_{[t,\infty)}(x_i) = 0$ and $I_{[t,\infty)}(\min_{i=1,...,n} x_i) = 0$. Note that $I_{[t,\infty)}(\min_{i=1,...,n} x_i) = 1$ if $x_i \in [t, \infty)$ for all $i = 1, ..., n$. So, the constraints can be rewritten as follows:

$$c + \sum_{i \in J} c_i \leq 0, \ J \subset \{1, ..., n\}, \ c + \sum_{i=1}^{n} c_i \leq 1, \ c \leq 0.$$

Let us show that the optimal solution is achieved by $c_k \geq 0$, $k = 1, ..., n$. If $(c, c_1, \ldots, c_n)$ is a feasible solution and $c_k < 0$, then $(c, c_1, ..., 0_k, ..., c_n)$ is a feasible solution corresponding to a greater value of the objective function. Indeed

$$c + \sum_{i=1, i \neq k}^{n} c_i \leq 0.$$

This implies that $c_k \geq 0$ and most of the constraints result from the following ones:

$$c + \sum_{j=1, j \neq i}^{n} c_i \leq 0, \ i \leq n, \ c + \sum_{i=1}^{n} c_i \leq 1, \ c \leq 0.$$

Now we can write $0 \leq c_i \leq 1$, $i = 1, ..., n$. This implies that $c \leq -(n-1)$ if $c_i = 1$ and $c \leq 0$ if $c_i = 0$, $i = 1, ..., n$. Then

$$\underline{M}(I_{[t,\infty)}(\min_{i=1,...,n} x_i)) = \max \left(\sum_{i=1}^{n} \underline{a}_i - (n-1), 0 \right).$$

4 Degenerate form

In reliability calculations one primarily has to deal with continuous gambles, for example, time to failure $\varphi(t) = t$, i.e. the constraints in the natural extension must be satisfied for all values t of these gambles. This implies that the optimization problem has to be solved for the infinite number of constraints or variables. In some special cases it is possible to overcome this difficulty. However, generally this fact makes the optimization problems difficult to solve directly. Therefore, some findings and developments are necessary. One of the ways is to use the Dirac functions which have unit area concentrated in the immediate vicinity of some point. In this case the infinite dimensional optimization problem is reduced to a problem with a finite number of variables equal to the number of constraints (pieces of evidence) plus one. In the case of employing the Dirac functions, the optimization problem, unfortunately, becomes non-linear. It should be noted however, the natural extension for a system with independent components is a non-linear programming problem anyway. Thus, this approach is especially efficient under the condition of independent components and continuous gambles.

It should be noted that the Dirac functions have been considered in [2] as an optimal solution to the optimization problem of preventive maintenance under incomplete information. However, there were studied only some special cases of partial initial information. Here we consider a general case fitting the framework of the theory of coherent imprecise probabilities.

Theorem 1 *If an optimal solution of optimization problem (1)-(3) exists, then it can be found in the class of degenerate densities*

$$\rho^*(\mathbf{X}) = \sum_{k=1}^{N+1} c_k \delta_{\mathbf{X}_k}(\mathbf{X}),\ N = \sum_{i=1}^{n} m_i \tag{14}$$

where $\mathbf{X}_k = (x_1^{(k)}, ..., x_n^{(k)}) \in \mathbf{R}_+^n,\ c_k \in \mathbf{R}_+$. *(For the proof see Appendix.)*

By substituting the degenerate class of densities (14) into objective function (1) and constraints (3) we obtain

$$\underline{M}(g) = \inf_{c_k, \mathbf{X}_k} \sum_{k=1}^{N+1} c_k g(\mathbf{X}_k), \tag{15}$$

subject to

$$\begin{aligned} \sum_{k=1}^{N+1} c_k &= 1,\ c_k \geq 0,\ k = 1, ..., N+1, \qquad (16) \\ \underline{a}_{ij} &\leq \sum_{k=1}^{N+1} c_k \varphi_{ij}(x_i^{(k)}) \leq \overline{a}_{ij},\ i \leq n,\ j \leq m_i. \end{aligned}$$

We refer to the natural extension (15)-(16) as the degenerate form. Problem (15)-(16) is valid for a case when we are ignorant of whether the components in a system are dependent or not. For independent components the following theorem can be formulated:

Theorem 2 *If an optimal solution of optimization problem (4)-(5) exists, then it can be found in the class of degenerate densities*

$$\rho_k^*(x) = \sum_{j=1}^{m_k+1} c_j^{(k)} \delta_{x_k^{(j)}}(x),\ k = 1, ..., n, \tag{17}$$

where $x_k^{(j)} \in \mathbf{R}_+,\ c_j^{(k)} \in \mathbf{R}_+$. *(For the proof see Appendix.)*

By substituting the degenerate class of densities (17) into objective function (4) and constraints (5) we obtain

$$\underline{M}(g) = \inf_{c_j, \mathbf{X}_j} \sum_{l_1=1}^{m_1+1} ... \sum_{l_n=1}^{m_n+1} g(x_1^{(l_1)}, ..., x_n^{(l_n)}) \prod_{v=1}^{n} c_{l_v}^{(v)},$$

subject to

$$\sum_{k=1}^{m_l+1} c_k^{(l)} = 1,\ c_k^{(l)} \geq 0,\ l = 1, ..., n,$$

$$\underline{a}_{ij} \leq \sum_{l=1}^{m_i+1} \varphi_{ij}(x_i^{(l)}) c_l^{(i)} \leq \overline{a}_{ij},\ i \leq n,\ j \leq m_i.$$

The three forms, primal, dual and Walley's, have been repeatedly employed by the authors for generalizing different reliability measures, models and approaches to coherent interval-valued probabilities [3, 4, 7, 8]. The degenerate form of the natural extension has not been used until now in any reliability modelling. An example below is a finding that contributes to the generalization of reliability models and demonstrates how the degenerate form can be used and analytical expressions can be inferred.

Example 3 *on how the degenerate form of the natural extension can be used in computing the stress-strength reliability under incomplete information. Let Y be a random variable describing the strength of a system and let X be a random variable describing the stress or load placed on the system. Then the stress-strength reliability of the system is determined as $R = \Pr\{X \leq Y\}$. Assume that the probability distribution functions of the independent stress and strength are given as non-parametric functions and quantified by precise probabilities. That is*

$$\begin{aligned} \Pr\{X &\leq \alpha_i\} = p_i,\ \Pr\{Y \leq \beta_j\} = q_j, \\ i &= 1, ..., m_1,\ j = 1, ..., m_2, \end{aligned}$$

where

$$\alpha_1 \leq \alpha_2 \leq ... \leq \alpha_{m_1},\ \beta_1 \leq \beta_2 \leq ... \leq \beta_{m_2},$$

$$\begin{aligned} p_0 &= 0 \leq p_1 \leq p_2 \leq ... \leq p_{m_1} \leq p_{m_1+1} = 1, \\ q_0 &= 0 \leq q_1 \leq q_2 \leq ... \leq q_{m_2} \leq q_{m_2+1} = 1. \end{aligned}$$

We are interested in finding the lower bound $\underline{R}$ of R. There are two random variables X and Y $(n = 2)$ and $m_1 + m_2$ previsions: p_i and q_j, $i = 1, ..., m_1$, $j = 1, ..., m_2$. By using the degenerate form of the natural extension the following problem can be stated

$$\underline{R} = \inf \sum_{k=1}^{m_1+1} \sum_{j=1}^{m_2+1} I_{[0,\infty)}(y_j - x_k) c_k d_j, \tag{18}$$

subject to

$$\sum_{k=1}^{m_1+1} c_k = 1, \ \sum_{j=1}^{m_2+1} d_j = 1, \tag{19}$$

$$\begin{aligned} \sum_{i=1}^{m_1+1} I_{[0,\alpha_k]}(x_i) c_i &= p_k, \ k = 1, ..., m_1, \\ \sum_{i=1}^{m_2+1} I_{[0,\beta_j]}(y_i) d_i &= q_j, \ j = 1, ..., m_2. \end{aligned} \tag{20}$$

Here the infimum is taken over the set of variables x_i, y_j, c_i, $d_j \in \mathbf{R}_+$, $i = 1, ..., m_1$, $j = 1, ..., m_2$, satisfying constraints (19)-(20). Assume that

$$x_1 \le x_2 \le ... \le x_{m_1+1}, \ y_1 \le y_2 \le ... \le y_{m_2+1}.$$

Let us prove that for all possible k, the following conditions $x_k \in [\alpha_{k-1}, \alpha_k]$, $y_k \in [\beta_{k-1}, \beta_k]$ for optimal values of x_k and y_k are valid. Here $\alpha_0 = 0$, $\beta_0 = 0$, $\alpha_{m_1+1} = A$, $\beta_{m_2+1} = B$. In particular, $A \to \infty$, $B \to \infty$. Suppose that there are two optimal values of x_j and x_k such that $x_j \in [\alpha_{k-1}, \alpha_k]$ and $x_k \in [\alpha_{k-1}, \alpha_k]$. If $j < k$, then it follows from (20) that

$$c_1 + ... + c_{j-1} = p_{j-1}, \ c_1 + ... + c_{j-1} = p_j.$$

We have arrived at the contradiction. If $j > k$, then it follows from (20) that

$$c_1 + ... + c_j = p_k, \ c_1 + ... + c_j = p_j.$$

We have also arrived at the contradiction. Similarly, we obtain the contradiction for an arbitrary combination of optimal values belonging to the same interval. This implies that $x_k \in [\alpha_{k-1}, \alpha_k]$. The proof of the condition $y_k \in [\beta_{k-1}, \beta_k]$ can be conducted in the same way. It follows from these conditions and from (20) that

$$c_1 = p_1, \ c_1 + c_2 = p_2, ..., \sum_{i=1}^{m_1} c_i = p_{m_1},$$

$$d_1 = q_1, \ d_1 + d_2 = q_2, ..., \sum_{i=1}^{m_2} d_i = q_{m_2}.$$

Hence

$$c_k = p_k - p_{k-1}, \ d_j = q_j - q_{j-1}, \ k \le m_1, \ j \le m_2.$$

Note that the objective function (18) achieves its minimum if for all $k \le m_1+1$ and $j \le m_2+1$ there hold $I_{[0,\infty)}(y_j - x_k) = 0$. However, there exist values j and k such that $I_{[0,\infty)}(y_j - x_k) = 1$ for arbitrary values of y_j, x_k. Let $j(k)$ be a minimal index such that there hold $x_k \in [\alpha_{k-1}, \alpha_k]$, $y_{j(k)} \in [\beta_{j(k)-1}, \beta_{j(k)}]$, $\alpha_k \le \beta_{j(k)}$. Then $I_{[0,\infty)}(y_j - x_k) = 1$ for all $j \ge j(k) + 1$. This implies that

$$\begin{aligned} \underline{R} &= \sum_{k=1}^{m_1+1} \sum_{j=j(k)+1}^{m_2+1} c_k d_j \\ &= \sum_{k=1}^{m_1+1} \sum_{j=j(k)+1}^{m_2+1} (p_k - p_{k-1})(q_j - q_{j-1}). \end{aligned}$$

It follows from $\sum_{j=j(k)+1}^{m_2+1} (q_j - q_{j-1}) = 1 - q_{j(k)}$ that

$$\underline{R} = \sum_{k=1}^{m_1} (p_k - p_{k-1})(1 - q_{j(k)}).$$

5 Concluding remarks

The degenerate form of the natural extension, apart from its applied value, sheds quite some light to the understanding of what imprecise previsions really are. They are an ultimate case of coherent interval-valued previsions which can be seen as a rough modelling with ignoring evidence in some cases. When applying the natural extension, we declare that the lower and upper previsions are sought over the set of all possible probability distributions, i.e. the widest class, which is really appealing. As it is clear, the widest class includes the class of the degenerate distributions that are a very particular and practically unrealistic class of distributions. Degenerate distributions are mathematical abstractions that can hardly anyhow satisfactorily model any distribution used at least in reliability theory. Yet, it turns out that coherent previsions exist just on this class of distributions that we would really like to abstain from involving. What evidence is ignored?

For example, in reliability applications this is a fact that the probability distribution of time to failure can hardly be concentrated in a few specific points of the real line. Is not it a reason that the practitioners face the fact that interval statistical models are too much imprecise to be employed in most practical studies and widely used? The evidence that the probability distribution functions of time to failure must be smooth values differentiable at any point of the real line is ignored. In many cases another evidence is obvious: the probability distribution functions of time to failure must exceed zero at any point over an interval $[0, T]$ except for the ending points. Can these evidence be utilized through Walley's and Kuznetsov's form of the natural extension and the concept of a gamble? One would experience difficulty in doing this. Yet, this kind of evidence and many others can be utilized through the primal form, which makes us think that the natural extension in its primal form is a more general tool for extending knowledge. It gives us a possibility to involve a wider class of evidence and judgements, which may be crucial for applications and may lead to obtaining more precise previsions and reducing indeterminacy.

References

[1] R.E. Barlow and F. Proschan. *Statistical Theory of Reliability and Life Testing: Probability Models*. Holt, Rinehart and Winston, New York, 1975.

[2] E.Ju. Barzilovich and V.A. Kashtanov. *Some Mathematical Problems of the Complex System Maintenance Theory*. Sovetskoe Radio, Moscow, 1971. in Russian.

[3] S.V. Gurov and L.V. Utkin. *Reliability of Systems under Incomplete Information*. Lubavich Publ., Saint Petersburg, 1999. in Russian.

[4] I. Kozine and Y. Filimonov. Imprecise reliabilities: experiences and advances. *Reliability Engineering and System Safety*, 67:75–83, 2000.

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

[6] D.M. Simmons. *Nonlinear Programming for Operations Research*. Prentice-Hall, International Series in Management, 1975.

[7] L.V. Utkin and S.V. Gurov. Imprecise reliability of general structures. *Knowledge and Information Systems*, 1(4):459–480, 1999.

[8] L.V. Utkin and I.O. Kozine. Conditional previsions in imprecise reliability. In H.A. Abderrahim D. Ruan and P. D'Hondt, editors, *Intelligent Techniques and Soft Computing in Nuclear Science and Engineering*, pages 72–79, Bruges, Belgium, 2000. World Scientific.

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

Appendix

The proofs of Theorem 1 and 2 are based on two lemmas.

Lemma 1 *Suppose that functions $g^{(i)}(t),\ i=1,...,m$ and $g(t)$ are integrable on $[0,\infty)$. Then an optimal solution of the problem*

$$z=\max_{\Phi}\int_0^{\infty} g(t)H(t)\mathrm{d}t,$$

subject to

$$\int_0^{\infty} g^{(i)}(t)H(t)\mathrm{d}t=a_i,\ i=1,...,m,$$

can be found in a class of degenerate distributions focusing on $m+1$ points. Here $H(t)$ is a non-increasing function such that $H(t)\geq 0$, $H(0)=1$; Φ is a set of all possible functions $H(t)$ satisfying the constraints.

Proof. Let us consider a discrete optimization problem:

$$z=\max\sum_{k=0}^{n} g_k x_k,$$

subject to

$$\sum_{k=0}^{n} g_k^{(i)} x_k=a_i,\ i=1,...,m,$$

$$1=x_0\geq x_1\geq x_2\geq ...\geq x_n\geq 0.$$

Let us introduce new variables $\alpha_k=x_k-x_{k+1}$, $k=0,1,2,...,n-1$, $\alpha_n=x_n$. Hence $x_k=\alpha_k+\alpha_{k+1}+...+\alpha_n$. Now the following equivalent problem can be written:

$$z=\max\sum_{j=0}^{n}\sum_{k=0}^{j} g_k\alpha_j,$$

subject to

$$\sum_{j=0}^{n}\sum_{k=0}^{j} g_k^{(i)}\alpha_j=a_i,\ i=1,..,m,\ \alpha_k\geq 0,\ k=0,..,n.$$

This is a linear optimization problem in the canonical form. The constraints are the system of linear equations of dimension $m\times(n+1)$. It is known that an optimal solution of such problem can be found among the basic solutions for which only m components are positive and other are equal to zero. Suppose that the non-zero components are $\alpha_{k_1},\alpha_{k_2},...,\alpha_{k_m}$. Then the following equalities hold:

$$\begin{aligned} x_0 &= x_1=...=x_{k_1},\\ x_{k_1+1} &= x_{k_1+2}=...=x_{k_2},\\ &\cdots\\ x_{k_m+1} &= x_{k_m+2}=...=x_n. \end{aligned}$$

This implies that there are m jumps in the sequence $\{x_k\}$. If the last term of the sequence is $x_n>0$, then there is the $(m+1)$-th jump. Note that m is independent of n. Then the passage to the limit as $n\to\infty$ completes the proof. ■

Lemma 2 *Suppose that functions $g^{(i)}(t),\ i=1,...,m$ and $g(t)$ are integrable on $[0,\infty)$. Then an optimal solution of the problem*

$$z=\max_{\Phi}\int_0^{\infty} g(t)H(t)\mathrm{d}t,$$

subject to

$$\underline{a}_i \leq \int_0^\infty g^{(i)}(t)H(t)\mathrm{d}t \leq \overline{a}_i,\ i=1,...,m,$$

can be found in a class of degenerate distributions focusing on $m+1$ points. Here $H(t)$ is a non-increasing function such that $H(t) \geq 0$, $H(0)=1$; Φ is a set of all possible functions $H(t)$ satisfying constraints.

Proof. We write a discrete optimization problem

$$z = \max \sum_{j=0}^{n} \sum_{k=0}^{j} g_k \alpha_k,$$

subject to

$$\underline{a}_i \leq \sum_{j=0}^{n} \sum_{k=0}^{j} g_k^{(i)} \alpha_k \leq \overline{a}_i,\ i=1,...,m,$$

$$\alpha_k \geq 0,\ k=0,1,...,n,$$

as it has been done in the proof for Lemma 1. Let us rewrite m constraints in the matrix form $\underline{A} \leq G \cdot X \leq \overline{A}$, where G is a matrix with components $g_k^{(i)}$; X, $\underline{A}$ and $\overline{A}$ are vectors with components α_k, $\underline{a}_i$ and $\overline{a}_i$, respectively. Let us fix such a vector Y that $\underline{A} \leq Y \leq \overline{A}$. Then an optimization problem with constraints $G \cdot X = Y$ satisfies Lemma 1, i.e. the optimal solution has m non-zero components. Let the initial optimization problem has an optimal solution X^*. Then there exists $Y^* = GX^*$ such that $\underline{A} \leq Y^* \leq \overline{A}$. However, Lemma 1 implies that the fixed vector Y^* has m non-zero components. This completes the proof. ∎

Proof of Theorem 1. If we replace the variable t by the vector $\mathbf{X}$, then the condition of Lemma 2 is not changed and the optimal solution to the problem

$$z = \max_{\Phi} \int_{\mathbf{R}_+^n} g(\mathbf{X})H(\mathbf{X})\mathrm{d}\mathbf{X},$$

subject to

$$\underline{a}_i \leq \int_{\mathbf{R}_+^n} g^{(i)}(\mathbf{X})H(\mathbf{X})\mathrm{d}\mathbf{X} \leq \overline{a}_i,\ i \leq N,$$

is a n-dimensional distribution function $H^*(\mathbf{X})$ having $N+1$ jumps at points $\mathbf{X}_k = (x_1^{(k)},...,x_n^{(k)})$, $k \leq N$. Then

$$\rho^*(\mathbf{X}) = \frac{\partial^n H^*(\mathbf{X})}{\partial x_1 \cdots \partial x_n} = \sum_{k=1}^{N+1} c_k \delta_{\mathbf{X}_k}(\mathbf{X}),$$

where c_k is a length of the k-th jump.

Proof of Theorem 2. Let an optimal solution to the initial problem be $\rho_1^*(x_1)\cdots\rho_n^*(x_n)$. Let us fix all ρ_i^*, $i \neq k$, except ρ_k, $k \in \{1,...,n\}$. We obtain an optimal solution of the problem because the number of constraints for the problem of one unknown density ρ_k is m_k.