Merging fuzzy statistical data with imprecise prior information – application in solving complex decision problems

Olgierd Hryniewicz
Systems Research Institute, Polish Academy of Sciences
Newelska 6, 01-447 Warsaw, Poland
hryniewi@ibspan.waw.pl

ABSTRACT

Solving complex decision problems requires the usage of information from different sources. Usually this information is uncertain and statistical or probabilistic methods are needed for its processing. However, in many cases a decision maker faces not only uncertainty of a random nature but also imprecision in the description of input data that is rather of linguistic nature. Therefore, there is a need to merge uncertainties of both types into one mathematical model. In the paper we present methodology of merging information from imprecisely reported statistical data and imprecisely formulated fuzzy prior information. Moreover, we also consider the case of imprecisely defined loss functions. The proposed methodology may be considered as the application of fuzzy statistical methods for the decision making in the systems analysis.

Keywords: Bayes decision-making, imprecise information, fuzzy statistical data, possibilistic decisions

1. INTRODUCTION

Solving complex decision problems can be regarded as processing of information of a different kind coming from various sources. Objective information related to stochastic phenomena that describe the environment of the decision situation can be treated as statistical data. If only such information is available, then the complex decision problem may be reduced to a simpler problem of a statistical decision. However, existing statistical data are usually not sufficient enough to solve complex problems. A decision-maker has to rely also on information from other, non-statistical, sources. That information is usually subjective in contrast to objective statistical data. Therefore, the decision-making process must contain a sub-process of merging information from different objective and subjective sources.

The generally accepted framework for dealing with objective and subjective information is known under the title of “Bayes decision-making”. There exist numerous textbooks related to this problem, such as e.g. classical books of Raiffa and Schleifer [1] and De Groot [2]. However, in practically all popular textbooks it is assumed that both objective statistical data and additional subjective information are precisely described in terms of the theory of probability. This assumptions have been questioned by many authors who claim that epistemic vagueness of information (i.e. uncertainty due to the imprecise character of information expressed in terms of commonly used natural language) cannot be described using the same mathematical models as in the case of aleatoric uncertainty (i.e. risk due to the randomness of future events and existing statistical data). Therefore, there is a need to propose a more general approach that allows to merge information of a different type in mathematical models used for solving complex decision problems.

In the paper we present the methodology that extends the classical Bayes decision-making to the case when both linguistic and aleatoric uncertainty may be merged in one mathematical model. In the second section of the paper we present the methods for modelling imprecise (i.e. vagually described) statistical data. In the third section we generalize the well known in decision making concept of the Bayes risk, and we propose its equivalent for the case of imprecise (fuzzy) statistical data, and imprecise prior information. Finally, in the fourth section of the paper we propose a possibilistic approach to decision making when the decision model is based on both random and imprecise information.

2. MATHEMATICAL MODELS FOR IMPRECISE STATISTICAL DATA

In the analysis of statistical data related to complex problems of system analysis we often face the problem of imprecise data. In many cases such data are provided by people who are not able to present precise numbers. There are many examples of cases where such imprecise data are very common in practice. For example, in the analysis of reliability data we often face imprecisely defined data, as it has been described in Grzegorzewski and Hryniewicz [3]. In this and many other cases data are reported by people who use imprecise expressions like “about 5”, “much larger than 5, but surely smaller than 10”, etc. The attempt do describe such lack of precision in terms of probability seems to be very questionable, as these imprecise notions do not have interpretation in terms of frequencies. However, it has been noted by many authors that the fuzzy sets theory proposed by Lotfi Zadeh is especially useful for the formal description of such imprecise data. Moreover, if the imprecise data are also of a random character, then the theory of fuzzy random variables can be used for the mathematical description of imprecise statistical data.

In this paper we will use the notion of a fuzzy random variable for the description of imprecise statistical data. Before we describe this notion in a formal way, let us introduce the concept of a fuzzy number. In a more formal way, a fuzzy number can be defined as follows.

Definition 1 (Dubois and Prade [5])
The fuzzy subset A of the real line $\boldsymbol{R}$, with the membership function $\mu: \boldsymbol{R} \to [0,1]$, is a *fuzzy number* if

- is normal, i.e. there exists an element $x_0 \in \boldsymbol{R}$ such that $\mu(x_0) = 1$;
- is fuzzy convex, i.e. $\mu(\lambda x + (1-\lambda) y) \geq \mu(x) \wedge \mu(y) \quad \forall x, y \in \boldsymbol{R}$ and $\forall 0 \leq \lambda \leq 1$;
- is upper semicontinuous;
- $\text{supp}(\mu)$ is bounded.

A useful concept used for the description of fuzzy numbers is the α-cut. The α-cut, A_α, of a fuzzy number A is a non-fuzzy set defined as

$$A_\alpha = \{x \in R: \mu(x) \geq \alpha\}.$$

The family $\{A_\alpha : \alpha \in [0,1]\}$ is a set representation of the fuzzy number A. Basing on the resolution identity, we have the alternative description of fuzzy numbers:

$$\mu(x) = \sup_{\alpha \in [0,1]} \{\alpha I_{A_\alpha}(x)\},$$

where $I_{A_\alpha}(x)$ denotes the characteristic function of A_α. Definition 1 implies that every α-cut of a fuzzy number is a closed interval. Hence, we have

$$A_\alpha = [A_\alpha^L, A_\alpha^U],$$

where

$$A_\alpha^L = \inf \{x \in \mathbf{R}: \mu(x) \geq \alpha\},$$
$$A_\alpha^U = \sup \{x \in \mathbf{R}: \mu(x) \geq \alpha\}.$$

The space of all fuzzy numbers will be denoted by $\boldsymbol{F}(\boldsymbol{R})$.

A fuzzy random variable may be defined by analogy to the definition of a real-valued random variable as a mapping that assigns to a random event an imprecise fuzzy number. The notion of a fuzzy random variable has been defined independently by many authors (see [3]). In general, a fuzzy random variable X is considered as a perception of an unknown usual random variable $V: \Omega \to \mathbf{R}$, called an original of X.

Formally, a fuzzy random variable can be defined using the following definition:

Definition 2 (Grzegorzewski and Hryniewicz [3])
a mapping $X: \Omega \to \boldsymbol{F}(\boldsymbol{R})$ is called a *fuzzy random variable* if it satisfies the following properties:

(1) $\{X_\alpha(\omega): \alpha \in [0,1]\}$ is a set representation of $X(\omega)$ for all $\omega \in \Omega$,

(2) for each $\alpha \in [0,1]$ both X_α^L and X_α^U defined as

$$X_\alpha^L = X_\alpha^L(\omega) = \inf X_\alpha,$$
$$X_\alpha^U = X_\alpha^U(\omega) = \sup X_\alpha,$$

are real-valued random variables on (Ω, F, P). Let χ denotes a set of all possible originals of X. If only vague data are available, it is of course impossible to show which of the possible originals is true. Therefore, we can define a fuzzy set of χ, with a membership function $\nu: \chi \to \boldsymbol{F}(\boldsymbol{R})$ given as follows:

$$\nu(V) = \inf\{\mu_{X(\omega)}(V(\omega)): \omega \in \Omega\}$$

which corresponds to the grade of acceptability that a fixed random variable V is the original of the fuzzy random variable in question.

Fuzzy random variables have been used for the description of many practical problems where stochastic randomness is present together with fuzzy imprecision. Classical statistical methods have been also generalized to the case of the analysis of fuzzy random data.

3. BAYES RISK IN CASE OF IMPRECISE INFORMATION

There exist different methods for modeling decisions in case of imprecise data. In this paper we present a generalization of the general model proposed by Raiffa and Schlaifer [6]. The model proposed by Raiffa and Schlaifer consists of two parts: one part is dedicated to the choice of the final decision, and the second part is dedicated to the choice of the experiment whose ultimate goal is to provide the decision maker with some information about the actual state of nature. According to this model the decision maker can specify the following data defining his decision problem.

1. Space of terminal decisions (acts): $A = \{a\}$.
2. State space: $\Theta = \{\theta\}$.
3. Family of experiments: $E = \{e\}$.
4. Sample space: $X = \{x\}$.
5. Utility function: $u(\cdot,\cdot,\cdot,\cdot)$ on $E \times X \times A \times \Theta$.

The decision maker evaluates a utility $u(e, x, a, \theta)$ of making a particular experiment *e*, obtaining the result of this experiment *x*, taking a decision *a* in case when the true state of nature is *θ*. In order to find appropriate (hopefully optimal) decisions the decision maker has also to specify a joint probability measure $P_{\theta,x}(\cdot,\cdot \mid e)$ for a Cartesian product $\Theta \times X$. The knowledge of this probability measure means that we know the joint probability distribution of observing in an experiment *e* the result *z* when the *random* state of nature is described by *θ*. Knowing this joint probability distribution we can calculate some important marginal and conditional probability distributions. In particular, for a given experiment *e* we are usually interested in three distributions.

1. The marginal distribution on the state space *Θ* describing our *prior* information about possible states of nature. We assume that this distribution does not depend on *e*.
2. The conditional distribution on the sample space *X* for given state of nature *θ*.
3. The conditional distribution on the state space *Θ* for given result of the experiment *x* describing our *posterior* information about possible states of nature.

Note, that we may know only these particular distributions as their knowledge is equivalent to the knowledge of the joint probability distribution on $\Theta \times X$.

Let us consider the simplest case of the general model. when there is no experiment *e*. In such a case the only information we need is the probability distribution $\pi(\theta)$ defined on the state space Θ. We call this distribution *the prior distribution* of the parameter (parameters) describing the unknown state of nature. If we know the utility function $u(a,\theta)$ defined on $A \times \Theta$ we may calculate *the expected utility* assigned to a particular action (decision) *a*.

The basic notion used in the decision theory is the risk defined as

$$\rho(a) = \int_{\Theta} L(a,\theta)\pi(\theta)d\theta \qquad (1)$$

where $L(a,\theta)$ is the loss related to the decision (action) *a* when the state of a system is *θ*, and π(*θ*) is the probability distribution defined on the space of the all possible states that reflects our prior knowledge about the system. Optimal decision (action) can be found by the minimization of this risk. When the decision maker has an additional information about the state of nature in a form of observations $\mathbf{x} = (x_1, x_2, \ldots, x_n)$ of a random vector described by a probability distribution $f(x,\theta)$ we may calculate *the expected risk* assigned to a particular action (decision) *a* from a formula

$$\rho(a \mid \mathbf{x}) = \int_{\Theta} L(a,\theta)g(\theta \mid \mathbf{x})d\theta \qquad (2)$$

where

$$g(\theta \mid \mathbf{x}) = \frac{f(\mathbf{x} \mid \theta)\pi(\theta)}{\int_{\Theta} f(\mathbf{x} \mid \theta)\pi(\theta)d\theta} \qquad (3)$$

is the posterior distribution of the parameter θ which describes the state of nature. The procedure of finding the optimal decision is exactly the same as in the case without statistical data.

Suppose now that the prior distribution $\pi(\theta;\zeta)$ and the loss $L(a;\theta,\psi)$ are functions of parameters ζ and *ψ*, respectively, and that these parameters are known only imprecisely. Let us assume that our imprecise knowledge about possible values of ζ and ψ is represented by fuzzy sets $\tilde{\zeta}$ and $\tilde{\psi}$, respectively. A fuzzy set $\tilde{X}$ is defined using the membership function $\mu_{\tilde{X}}(x)$ which in the considered in this paper context describes the grade of possibility that a fuzzy parameter, say $\tilde{X}$, has a specified value of *x*. Each fuzzy set may be also represented by its α-cuts defined as ordinary sets

$$X^{\alpha} = \{x \in R : \mu_{\tilde{X}}(x) \geq \alpha, 0 \leq \alpha \leq 1\}$$

From the representation theorem for fuzzy sets we know that each membership function may be equivalently represented as

$$\mu_{\tilde{X}}(x) = sup\{\alpha I_{\tilde{X}^{\alpha}}(x) : \alpha \in [0,1]\}.$$

Now let us assume that imprecisely known parameters

ζ and ψ (possibly vectors) are represented by their α-contours (Cartesian products of the α-cuts), and that these α -contours are given in a form of multivariate closed intervals $\left[\zeta_L^\alpha, \zeta_U^\alpha\right]$ and $\left[\psi_L^\alpha, \psi_U^\alpha\right]$, respectively. The knowledge of these α -contours let us calculate fuzzy equivalents of the expected loss (risk). To make the presentation simple we assume that decision are based exclusively on the knowledge of the prior distribution $\pi(\theta;\zeta)$ and the loss function $L(a;\theta,\psi)$. As these function are the function of imprecise fuzzy parameters, they are also fuzzy, and may be denoted as $\tilde{\pi}(\theta;\tilde{\zeta})$ and $\tilde{L}(a;\theta;\tilde{\psi})$, respectively.

Now, let us rewrite the formula for the expected risk as

$$\tilde{\rho}(a \mid \zeta,\psi) = \int_\Theta \tilde{L}(a;\theta,\tilde{\psi})\tilde{\pi}(\theta;\tilde{\zeta})d\theta . \quad (4)$$

The risk calculated from this formula is now an imprecisely defined *fuzzy number* whose membership function may be calculated using Zadeh's extension principle.

Definition 3. Extension principle (Dubois and Prade [7])

Let X be a Cartesian product of universe $X = X_1 \times X_2 \times \cdots \times X_r$, and $A_1,\ldots,A_r$ be r fuzzy sets in $X_1,\ldots,X_r$, respectively. Let f be a mapping from $X = X_1 \times X_2 \times \cdots \times X_r$ to a universe Y such that $y = f(x_1,\ldots,x_r)$. The extension principle allows us to induce from r fuzzy sets A_i a fuzzy set B on Y through f such that

$$\mu_B(y) = \sup_{x_1,\ldots,x_r;\, y=f(x_1,\ldots,x_r)} \min\left[\mu_{A_1}(x_1),\ldots,\mu_{A_r}(x_r)\right]$$

$$\mu_B(y) = 0 \text{ if } f^{-1}(y) = \varnothing$$

When the formula (1) for the expected risk is given explicitly, then its fuzzy version (4) can be obtained by the "fuzzification" of the original non-fuzzy formula using the extension principle given above. In a general case, however, the α-cuts $\left(\rho^{\alpha,L}(a \mid \zeta,\psi), \rho^{\alpha,U}(a \mid \zeta,\psi)\right)$ of the fuzzy expected risk $\tilde{\rho}(a \mid \zeta,\psi)$ are given by the following formulae:

$$\rho^{\alpha,L}(a \mid \zeta,\psi) = \min_{\zeta,\psi \in \left(C(\tilde{\zeta})_\alpha \times C(\tilde{\psi})_\alpha\right)} \int_\Theta \tilde{L}(a;\theta,\tilde{\psi})\tilde{\pi}(\theta;\tilde{\zeta})d\theta \quad (5)$$

$$\rho^{\alpha,U}(a \mid \zeta,\psi) = \max_{\zeta,\psi \in \left(C(\tilde{\zeta})_\alpha \times C(\tilde{\psi})_\alpha\right)} \int_\Theta \tilde{L}(a;\theta,\tilde{\psi})\tilde{\pi}(\theta;\tilde{\zeta})d\theta \quad (6)$$

where $C(\tilde{\zeta})_\alpha$ and $C(\tilde{\psi})_\alpha$ are the α-contours of the fuzzy parameters $\tilde{\zeta}$ of the prior distribution $\pi(\theta;\zeta)$ and fuzzy parameters $\tilde{\psi}$ of the loss function $L(a;\theta,\psi)$, respectively.

Now, let us consider the case when the statistical data are fuzzy, and the remaining parameters of the decision model are crisp (i.e. precisely defined). In the presence of fuzzy statistical data the posterior distribution of the state variable θ can be obtained by the application of the defined above Zadeh's extension principle to the formula that describes this distribution. Let $\tilde{x}_i^\alpha = \left((\tilde{x}_i^\alpha)_L, (\tilde{x}_i^\alpha)_U\right), j = 1,\ldots,n$ be the α-cuts of the fuzzy observations $\tilde{x}_1, \tilde{x}_2,\ldots,\tilde{x}_n$. Applying the notation proposed by Fruehwirth-Schnatter [8] we denote by $C(\tilde{x})_\alpha$ the α-contour of the fuzzy sample which is equal to the Cartesian product of the α-cuts $\tilde{x}_i^\alpha, j = 1,\ldots,n$ of individual fuzzy observations. The fuzzy posterior distribution $\tilde{g}(\theta \mid \tilde{x})$ is, according to Viertl and Hule [9] given by α-contours

$$g_\alpha^L(\theta \mid \mathbf{x},\zeta) = \min_{\mathbf{x} \in C(\tilde{x})_\alpha} \frac{f(\mathbf{x} \mid \theta)\pi(\theta;\zeta)}{\eta(\mathbf{x})}, \quad (7)$$

$$g_\alpha^U(\theta \mid \mathbf{x},\zeta) = \max_{\mathbf{x} \in C(\tilde{x})_\alpha} \frac{f(\mathbf{x} \mid \theta)\pi(\theta;\zeta)}{\eta(\mathbf{x})}, \quad (8)$$

where $\eta(\mathbf{x})$ is a normalizing constant equal to the denominator of the right hand side of (3). Now, we can compute the fuzzy risk using the general methodology for integrating fuzzy functions presented in [7].

Let us denote by

$$C(\tilde{\rho})_\alpha = \left(\tilde{\rho}^{\alpha,L}(a \mid \mathbf{x},\zeta), \tilde{\rho}^{\alpha,U}(a \mid \mathbf{x},\zeta)\right)$$

the α-cut of the fuzzy risk $\tilde{\rho}(a \mid \mathbf{x})$. The lower and upper bounds of this α-cut are calculated from the following formulae:

$$\tilde{\rho}^{\alpha,L}(a \mid \mathbf{x},\zeta) = \int_\Theta L(a;\theta) g_\alpha^L(\theta \mid \mathbf{x},\zeta)d\theta \quad (9)$$

$$\tilde{\rho}^{\alpha,U}(a \mid \mathbf{x},\zeta) = \int_\Theta L(a;\theta) g_\alpha^U(\theta \mid \mathbf{x},\zeta)d\theta \quad (10)$$

Thus, we can calculate the respective fuzzy risks for all considered decisions a.

Now, let us consider the calculation of fuzzy risks when all quantities involved, i.e., loss function, prior distribution, and statistical data may be imprecisely defined. The α-cuts of the fuzzy posterior probability distribution of the parameter θ are given by the following formulae:

$$g_\alpha^L(\theta) = \min_{\mathbf{x},\zeta \in \left(C(\tilde{x})_\alpha \times C(\tilde{\zeta})_\alpha\right)} \frac{f(\mathbf{x} \mid \theta)\pi(\theta,\zeta)}{\eta(\mathbf{x},\zeta)} \qquad (11)$$

$$g_\alpha^U(\theta) = \max_{\mathbf{x},\zeta \in \left(C(\tilde{x})_\alpha \times C(\tilde{\zeta})_\alpha\right)} \frac{f(\mathbf{x} \mid \theta)\pi(\theta,\zeta)}{\eta(\mathbf{x},\zeta)} \qquad (12)$$

where $\eta(\mathbf{x},\zeta)$ is the normalizing constant. The fuzzy expected risk, $\tilde{\rho}(a \mid \mathbf{x},\zeta,\psi)$, is now defined by its α-cuts calculated from the following formulae:

$$\tilde{\rho}^{\alpha,L}(a \mid \mathbf{x},\zeta,\psi) = \int_\Theta L^{\alpha,L}(a;\theta,\psi) g_\alpha^L(\theta \mid \mathbf{x},\zeta) d\theta \qquad (13)$$

$$\tilde{\rho}^{\alpha,U}(a \mid \mathbf{x},\zeta,\psi) = \int_\Theta L^{\alpha,U}(a;\theta,\psi) g_\alpha^U(\theta \mid \mathbf{x},\zeta) d\theta \qquad (14)$$

where

$$L^{\alpha,L}(a;\theta,\psi) = \min_{\psi \in C(\tilde{\psi})_\alpha} L(a;\theta,\psi), \qquad (15)$$

$$L^{\alpha,U}(a;\theta,\psi) = \max_{\psi \in C(\tilde{\psi})_\alpha} L(a;\theta,\psi) \qquad (16)$$

are the α-cuts of the fuzzy loss function $\tilde{L}(a;\theta;\tilde{\psi})$.

4. MAKING DECISIONS WIH IMPRECISE INFORMATION – A POSSIBILISTIC APPROACH

In a classical approach a decision-maker chooses the action with the minimal expected risk. This approach cannot be directly used in the case of fuzzy risks, as there is no natural method for ordering fuzzy numbers. There exist two general ways of dealing with the problem of choosing the best solution: either to defuzzify the risks or to introduce additional measures that allow to order considered options. If the first approach is preferred we claim that the λ-average ranking method proposed by Campos and Gonzalez [4] is especially useful in decision making. Let $\tilde{X}$ be a fuzzy number (fuzzy set) described by the set of its α-cuts $\left[X_L^\alpha, X_U^\alpha\right]$, and S be an additive measure on [0,1]. Moreover, assume that the support of $\tilde{X}$ is a closed interval. The λ-average value of such a fuzzy number $\tilde{X}$ is defined by Campos and Gonzalez [4] as

$$V_S^\lambda\left(\tilde{X}\right) = \int_0^1 \left[\lambda X_U^\alpha + (1-\lambda) X_L^\alpha\right] dS(\alpha), \lambda \in [0,1]. \qquad (17)$$

In the case of continuous membership functions this integral is calculated with respect to $d\alpha$. Thus, the λ-average value of $\tilde{X}$ can be viewed as its defuzzified value. The parameter λ in the above integral is a subjective degree of the decision-maker's optimism (pessimism). In the case of fuzzy risks $\lambda = 0$ reflects his highest optimism as the minimal values of all a-cuts (representing the lowest possible risks) are taken into consideration. On the other hand, by taking $\lambda = 1$ the decision-maker demonstrates his total pessimism, as only the maximal values of all α-cuts (representing the highest possible risks) are considered. If the decision maker takes $\lambda = 0{,}5$ his attitude may described as neutral. Thus, by varying the value of λ the decision maker is able to take into account the level of his optimism (pessimism) which may arise e.g. from having some additional information that has not been reflected in the prior distribution.

When the second approach is preferred we propose to use the methodology known from the theory of possibility, namely the *Possibility of Dominance* and *Necessity of Strict Dominance* indices proposed by Dubois and Prade [5].

For two fuzzy numbers $\tilde{A}$ and $\tilde{B}$ the *Possibility of Dominance* (*PD*) index is calculated from the formula

$$PD = Poss\left(\tilde{A} \geq \tilde{B}\right) = \sup_{x,y:x \geq y} \min\left\{\mu_{\tilde{A}}(x), \mu_{\tilde{B}}(y)\right\}. \qquad (18)$$

The *PD* index gives the measure of *possibility* that the fuzzy number $\tilde{A}$ is not smaller than the fuzzy number $\tilde{B}$. Positive value of this index tells the decision maker that there exists even slightly evidence that the relation $\tilde{A} \geq \tilde{B}$ is true. The degree of *conviction* that the relation $\tilde{A} > \tilde{B}$ is true is reflected by the *Necessity of Strict Dominance* (*NSD*) index defined as

$$NSD = Ness\left(\tilde{A} > \tilde{B}\right) = 1 - \sup_{x,y:x \leq y} \min\left\{\mu_{\tilde{A}}(x), \mu_{\tilde{B}}(y)\right\} = 1 - Poss\left(\tilde{B} \geq \tilde{A}\right). \qquad (19)$$

The *NSD* index gives the measure of *necessity* that the fuzzy number $\tilde{A}$ is greater than the fuzzy number $\tilde{B}$. Positive value of this index tells the decision maker that there exists rather strong evidence that the relation $\tilde{A} > \tilde{B}$ is true. This possibilistic index, and other similar indices, may be used for choosing the best option while solving complex decision problems.

5. EXAMPLES OF APPLICATIONS

To illustrate possible applications of the proposed methodology let us consider two typical decision problems: estimation of the parameter of a probability distribution, and choosing the best from among two competing options. Both examples are simplified and have rather an illustration character.

Consider the problem of the estimation of the mean value ν of a random variable X that is distributed according to the normal distribution N(ν,σ) with the known value of the standard deviation σ. Let us assume that we have the following additional information:

a) a sample $x_1, x_2, ..., x_n$ of the random variable X is observed;

b) there exists some prior information about possible values of the parameter ν which is summarized in the form of the normal prior distribution N(γ,δ), where γ and δ are known parameters;

c) the loss function L is quadratic, i.e. proportional to the squared difference between the estimated and actual value of the parameter ν.

The considered problem has a very well known solution, see for example [1], and the Bayes decision (Bayes estimator of ν) which minimizes the posterior risk is given by a simple formula:

$$\hat{\nu}_B = \frac{\sigma^2}{\sigma^2 + n\delta^2}\gamma + \frac{n\delta^2}{\sigma^2 + n\delta^2}\overline{X} \tag{20}$$

Now, let us consider that we observe *imprecise* values of the random variable X, and each observation is described by a *fuzzy* number $\tilde{x}_i$, $i = 1,...,n$, denoted by $\left(x_{1,i}, x_{2,i}, x_{3,i}, x_{4,i}\right)$, and described by a *trapezoidal* membership function given by the following expression:

$$\mu_{\tilde{x}_i}(x) = \begin{cases} 0 & \text{if} \quad x < x_{1,i} \\ \left(x - x_{1,i}\right)/\left(x_{2,i} - x_{1,i}\right) & \text{if} \quad x_{1,i} \le x < x_{2,i} \\ 1 & \text{if} \quad x_{2,i} \le x < x_{3,i} \\ \left(x_{4,i} - x\right)/\left(x_{4,i} - x_{3,i}\right) & \text{if} \quad x_{3,i} \le x < x_{4,i} \\ 0 & \text{if} \quad x_{4,i} \le x \end{cases} \tag{21}$$

Moreover, let us assume that the parameter δ of the prior distribution is known exactly, but the parameter γ is also imprecisely defined, and is described by the following trapezoidal function:

$$\mu_{\tilde{\gamma}}(\gamma) = \begin{cases} 0 & \text{if} \quad \gamma < g_1 \\ \left(\gamma - g_1\right)/\left(g_2 - g_1\right) & \text{if} \quad g_1 \le \gamma < g_2 \\ 1 & \text{if} \quad g_2 \le \gamma < g_3 \\ \left(g_4 - \gamma\right)/\left(g_4 - g_3\right) & \text{if} \quad g_3 \le \gamma < g_4 \\ 0 & \text{if} \quad g_4 \le \gamma \end{cases} \tag{22}$$

The fuzzy Bayes estimator of the parameter ν can be found by fuzzification of (20). Simple application of the Zadeh's extension principle leads to the following result: the observed fuzzy value $\tilde{\nu}$ of the estimator of the mean value ν is also a trapezoidal fuzzy number described by the membership function

$$\mu_{\tilde{\nu}}(\nu) = \begin{cases} 0 & \text{if} \quad \nu < \nu_1 \\ \left(\gamma - \nu_1\right)/\left(\nu_2 - g_1\right) & \text{if} \quad \nu_1 \le \nu < \nu_2 \\ 1 & \text{if} \quad \nu_2 \le \nu < \nu_3 \\ \left(\nu_4 - \gamma\right)/\left(\nu_4 - \nu_3\right) & \text{if} \quad \nu_3 \le \nu < \vartheta_4 \\ 0 & \text{if} \quad \nu_4 \le \nu \end{cases} \tag{23}$$

where

$$\nu_1 = \frac{\sigma^2}{\sigma^2 + n\delta^2} g_1 + \frac{\delta^2}{\sigma^2 + n\delta^2}\sum_{i=1}^{n} x_{1,i}\ , \tag{24}$$

$$\nu_2 = \frac{\sigma^2}{\sigma^2 + n\delta^2} g_2 + \frac{\delta^2}{\sigma^2 + n\delta^2}\sum_{i=1}^{n} x_{2,i}\ , \tag{25}$$

$$\nu_3 = \frac{\sigma^2}{\sigma^2 + n\delta^2} g_3 + \frac{\delta^2}{\sigma^2 + n\delta^2}\sum_{i=1}^{n} x_{3,i}\ , \tag{26}$$

$$\nu_4 = \frac{\sigma^2}{\sigma^2 + n\delta^2} g_4 + \frac{\delta^2}{\sigma^2 + n\delta^2}\sum_{i=1}^{n} x_{4,i}\ . \tag{27}$$

It is worthy to note that in the case of imprecise values of other parameters, such as σ and δ, the result of fuzzification is not so simple, as the membership function of $\tilde{\nu}$ is no longer a trapezoidal one. However, the application of the concept of α-cuts and the extension principle let us calculate its approximation (for a finite set of α-cuts) without serious problems.

Now, let us consider the second example: the choice of the best action from among two possible actions $\{a_1,a_2\}$. Potential losses connected with the choice of both actions depend upon the value of the state variable θ. In the simplest case we may consider only two values of the state variable θ, namely θ_1 and θ_2. Suppose that there exists the following prior probability distribution over the set $\{\theta_1, \theta_2\}$: $P(\theta = \theta_1) = p, P(\theta = \theta_2) = 1 - p$.

Let us now define the loss function of the considered problem in a form of a following table:

Table 1. Loss function in a tabular form

Decision/State	θ_1	θ_2
a_1	0	w_1
a_2	w_2	0

In this simple case losses (w_1>0, w_2>0) are generated only in the case of wrong decisions.

The solution to this problem is well known in literature (for this and more complicated models see, e.g., DeGroot [2]). The expected loss (risk) connected with decision a_1 is, according to (1), equal to $\rho(a_1) = w_1(1-p)$, and the risk connected with decision a_2 is equal to $\rho(a_2) = w_2 p$. For given values of p, w_1, and w_2 we calculate both risks, and we choose the action connected with the smaller one.

Suppose now that the parameters p, w_1, and w_2 are known only imprecisely, and that they are described by fuzzy triangular numbers that have the membership function of the following general form:

$$\mu_{\tilde{y}}(x) = \begin{cases} 0 & \text{if} \quad y < y_1 \\ (y - y_1)/(y_2 - y_1) & \text{if} \quad y_1 \le y < y_2 \\ (y_3 - y)/(y_3 - y_2) & \text{if} \quad y_2 \le y < y_3 \\ 0 & \text{if} \quad y_3 \le y \end{cases} \qquad (28)$$

Let us denote this fuzzy number by a triple (y_1, y_2, y_3). For a given value of $\alpha, 0 < \alpha \le 1$, the lower limit of the respective α-cut is given as

$$y_L^\alpha = y_1 + \alpha(y_2 - y_1) \qquad (29)$$

and the upper limit is given by

$$y_u^\alpha = y_3 - \alpha(y_3 - y_2) \qquad (30)$$

The fuzzy risks connected with the considered decisions are not described by triangular fuzzy numbers. However, the limits of their α-cuts are still easy to calculate from the following formulae:

$$\rho_L^\alpha(a_1) = w_{1,L}^\alpha\left(1 - p_U^\alpha\right), \qquad (31)$$

$$\rho_U^\alpha(a_1) = w_{1,U}^\alpha\left(1 - p_L^\alpha\right), \qquad (32)$$

$$\rho_L^\alpha(a_2) = w_{2,L}^\alpha\, p_L^\alpha, \qquad (33)$$

$$\rho_U^\alpha(a_2) = w_{2,U}^\alpha\, p_U^\alpha. \qquad (34)$$

Suppose now that the actions are numbered in such a way that the following relation holds:

$$\rho_{1,U}^1(a_1) \le \rho_{2,L}^1(a_2).$$

In such a case the risk connected with action a_2 is likely to be greater than the risk connected with action a_1. Otherwise, either the risk connected with action a_1 is greater than the risk connected with action a_2 or both risk are similar, and undistinguishable due to their fuzziness.

The *NSD* index that measures the dominance of the fuzzy risk $\tilde{\rho}(a_2)$ over the fuzzy risk $\tilde{\rho}(a_1)$ can be now calculated from the following expression:

$$NSD(\tilde{\rho}(a_2) > \tilde{\rho}(a_1)) = \begin{cases} RD(a_1, a_2) & \text{if } \rho_3(a_1) > \rho_1(a_2) \\ 1 & \text{otherwise} \end{cases} \qquad (35)$$

where

$$RD(a_1, a_2) = 1 - \frac{\rho_3(a_1) - \rho_1(a_2)}{(\rho_2(a_2) - \rho_1(a_2)) + (\rho_3(a_1) - \rho_2(a_1))}, \qquad (36)$$

and

$$\rho_1(a_i) = \rho_L^0(a_i), \quad i = 1{,}2, \qquad (37)$$

$$\rho_2(a_i) = \rho_L^1(a_i), \quad i = 1{,}2, \qquad (38)$$

$$\rho_3(a_i) = \rho_U^0(a_i), \quad i = 1{,}2. \qquad (39)$$

If this value is greater than 0, we are entitled to say that the action a_1 is, to some extent, preferable to action a_2. Otherwise, there is a *possibility* that the action a_2 is preferable to the action a_1.

To give a numerical example let us assume that $\tilde{w}_1$ is described by a triangular fuzzy number (1 , 2 , 3), $\tilde{w}_2$ by (2 , 3 , 4), and $\tilde{p}$ by (0,4 , 0,5 , 0,6). From (31) – (34) and (37) – (39) we have $\rho_2(a_1) = 1$, $\rho_3(a_1) = 1{,}8$, $\rho_1(a_2) = 0{,}8$, and $\rho_2(a_2) = 1{,}5$. The *NSD* for the dominance of the risk connected with the action a_2 over the risk connected with the action a_1, calculated from (35) is equal to 0,41. Thus, there is significant evidence that the action a_1 should be preferred over the action a_2.

6. CONCLUSIONS

In the paper we have presented a general methodology for making Bayes optimal decisions when input data, i.e. parameters of the loss function, parameters of the prior distribution of the state variable, and statistical data, may be imprecisely defined. This situation frequently happens in the systems analysis of complex systems where the input information is expressed by people (experts) who use a common language. For the description of that lack of precision we use the formalism of the fuzzy sets. Therefore, the risks that are calculated in order to find optimal decisions are fuzzy. We present algorithms that are useful for the calculation of these fuzzy risks. Moreover, we present the methodology for the comparison of fuzzy risks. The theory presented in the paper is illustrated with some simple examples.

REFERENCES

[1] Raiffa H., Schleifer R. (1961) Applied Statistical Decision Theory, The M.I.T. Press, Cambridge.

[2] De Groot M.H (1970) Optimal Statistical Decisions. McGraw Hill, New York.

[3] Grzegorzewski P., Hryniewicz O.(2002) Computing with words and life data. *Int. Journ. Appl. Math. Comput. Sci* 12(3), 337-345.

[4] Campos L.M., Gonzalez A. (1989) A subjective approach for ranking fuzzy numbers. Fuzzy Sets and Systems, 29, 145-153.

[5] Dubois D., Prade H. (1983) Ranking fuzzy numbers in the setting of possibility theory, Information Sciences, 30, 184 – 244.

[6] Raiffa H., Schlaifer R. (2000) Applied Statistical Decision Theory. New York : J.Wiley (2nd Edition).

[7] Dubois, D., Prade, H. (1980) Fuzzy Sets and Systems. Theory and Applications, Academic Press New York.

[8] Fruehwirth-Schatter S. (1993) Fuzzy Bayesian inference. Fuzzy Sets and Systems, 60, 41-58

[9] Viertl R., Hule H. (1991) On Bayes' theorem for fuzzy data. Statistical papers, 32, 115-122.