Various probabilistic interpretations of fuzzy sets and their use in decision theory, classification problems and approximate reasoning

Andrey G. Bronevich

JSC "Research, Development and Planning Institute for Railway Information Technology, Automation and Telecommunication"
Nizhegorodskaya 27, building 1, 109029, Moscow, Russia
brone@mail.ru

european social fund in the czech republic
EUROPEAN UNION
MINISTRY OF EDUCATION, YOUTH AND SPORTS
OP Education for Competitiveness
INVESTMENTS IN EDUCATION DEVELOPMENT

Classification based on likelihood functions

Classification problem:
1. Patterns are vectors in $\mathbb{R}^n$ that should be classified on two classes ω_1 and ω_2.
2. Classes are described by probability densities $h(x|\omega_1)$ and $h(x|\omega_2)$.
3. It is necessary choose classifier

$$\mathbf{x} \in \omega_i \text{ if } \mathbf{x} \in \Omega_i,\ i = 1, 2,$$

where $\{\Omega_1, \Omega_2\}$ is a partition of $\mathbb{R}^n$.

that minimizes the probability of error.

Then $\{\Omega_1, \Omega_2\}$ should be chosen as a solution of the following optimization problem:

$$p(\omega_1) \int\limits_{\Omega_2} f(\mathbf{x}|\omega_1) d\mathbf{x} + p(\omega_2) \int\limits_{\Omega_1} f(\mathbf{x}|\omega_2) d\mathbf{x} \to \min,$$

where $p(\omega_1)$, $p(\omega_2) = 1 - p(\omega_1)$ are probabilities of occurrence of classes ω_1 and ω_2. The above optimization problems can be solved as follows

$$\mathbf{x} \in \Omega_1 \text{ if } L(\mathbf{x}) > \alpha,$$

where $L(\mathbf{x}) = f(\mathbf{x}|\omega_1)/f(\mathbf{x}|\omega_2)$ is a likelihood function and $\alpha = p(\omega_2)/p(\omega_1)$.

Some times probabilities $p(\omega_i)$ are not known. In this case for choosing α, we use the Neumann-Pirson criterion. Notice that we can describe all possible classifiers by sets

$$L_\alpha = \{\mathbf{x} \in \mathbb{R}^n | L(\mathbf{x}) > \alpha\}.$$

These sets can be conceived as a strict cuts of some fuzzy set. It rational to connect such sets with probabilities, for example, to choose fuzzy set F such that

$$F_{1-p} = L_\alpha \text{ if } \int_{L_\alpha} f(\mathbf{x}|\omega_1) d\mathbf{x} = p.$$

where $F_{1-p} = \{\mathbf{x} \in \mathbb{R} | F(\mathbf{x}) > 0\}$ is the strict $(1-p)$-cut of a fuzzy set F.

Generalization of the classification problem: A set $\{\omega_1, ..., \omega_N\}$ of possible elementary classes is finite and we can observe some class $A \subseteq \{\omega_1, ..., \omega_N\}$. Then it can be characterized by the density

$$f(\mathbf{x}|A) = \frac{\sum\limits_{\omega_i \in A} p(\omega_i) f(\mathbf{x}|\omega_i)}{\sum\limits_{\omega_i \in A} f(\omega_i)}.$$

If we minimize the probability of wrong classification, then we come to the following classification rule:

$$\mathbf{x} \in A \text{ if } P(A) f(\mathbf{x}|A) > P(\bar{A}) f(\mathbf{x}|\bar{A}),$$

where $P(A) = \sum\limits_{\omega_i \in A} p(\omega_i)$.

Let us introduce a function

$$f(x) = P(A)f(\mathbf{x}|A) + P(\bar{A})f(\mathbf{x}|\bar{A})$$

that is obviously the density of all possible observed patterns. Then the above classification rule is transformed to the following

$$\mathbf{x} \in A \text{ if } f(\mathbf{x}|A)/f(\mathbf{x}) > 1/(2P(A)).$$

Let us assume that we observe non-elementary classes and we can evaluate functions $f(\mathbf{x}|A)$ and $f(\mathbf{x})$. Then we can describe any class A with the help of a fuzzy set such that its strict cuts correspond to optimal classifiers.

We describe next the classification scheme based on so called statistical classes. In this scheme

1. It may be potentially infinite number of classes.
2. The observed classes are not disjoint.
3. There are may be fuzzy boundaries between classes.

X is a measurable space with a σ-algebra $\mathfrak{A}$

V is *volume measure* (an additive measure) on $\mathfrak{A}$.

Any *statistical class* is described by a probability measure P on $\mathfrak{A}$ and P has to be absolutely continuous w.r.t. V, i.e. there is a density $h(x)$ such that $P(A) = \int\limits_A h(x)dV(x)$.

Remark. If we consider previous scheme, then a measure V can be understood as

$$V(A) = \int_A f(\mathbf{x}) d\mathbf{x}$$

and probability measure P describes the occurrence of patterns in the class. In this case density

$$h(\mathbf{x}) = \frac{f(\mathbf{x}|A)}{f(\mathbf{x})}.$$

If it is hard to evaluate $h(\mathbf{x})$, then we can choose the Lebesgue measure as V for all measurable subsets of $\mathbb{R}^n$.

Each statistical class is described by the set of its minimal elements.

Let F be a statistical class. A set $A \in \mathfrak{A}$ is called *minimal* for F if it has the minimum volume among all equiprobable events, i.e

let $\mathcal{A}(p) = \{A \in \mathfrak{A} | P(A) = p\}$, then $B \in \mathcal{A}(p)$ is a minimal event if

$$V(B) = \min_{A \in \mathcal{A}(p)} V(A).$$

In some cases it is possible to describe the set of all minimal elements for the class F. For example, for the case, when P is *continuous in its values*.

P is *continuous in its values* if for any sets $A \subseteq B$ in $\mathfrak{A}$ and any $p \in [P(A), P(B)]$ there is a set $C \in \mathfrak{A}$ such that $P(C) = p$.

Definitions:

1) $A \subseteq B$ in measure V iff $V(A \backslash B) = 0$;

2) $A = B$ in measure V iff $V(A \backslash B) = V(B \backslash A) = 0$.

Theorem. Let F be a statistical class with a corresponding probability measure P and density h, and let P be continuous in its values. Then for each minimal event B there is $t \in [0, \infty)$ such that

$$\{x \in X | h(x) > t\} \subseteq B \subseteq \{x \in X | h(x) \geqslant t\}$$

in measure V.

A statistical class is called *regular* if any of its minimal event is defined uniquely in measure V by its probability.

Remark. 3 cases are possible:

1) A minimal event with a given probability $p \in [0, 1]$ does not exist;

2) It is defined uniquely;

3) It is not defined uniquely.

Theorem. Let F be a statistical class with the probability measure P and density h, and let P be continuous in its values. Then F is regular iff $P(\{x \in X | h(x) = t\}) = 0$ for any $t \in [0, \infty)$.

Remark. If conditions of above theorem are fulfilled, then any minimal event can be represented as $\{x \in X | h(x) > t\}$.

Inclusion relation for regular statistical classes

Let F_i, $i = 1, 2$, be regular statistical classes, and let $A_i(p)$ be the corresponding minimal event with probability $p \in [0, 1]$. Then by definition

1) $F_1 \subseteq F_2$ iff $A_1(p) \subseteq A_2(p)$ for all $p \in [0, 1]$;

2) $F_1 = F_2$ iff $A_1(p) = A_2(p)$ for all $p \in [0, 1]$.

Theorem. Let F_i, $i = 1, 2$, be regular statistical classes and $F_1 = F_2$. Then they generated by the same probability measure.

Local inclusion measure of regular statistical classes

Let F_1 and F_2 be regular statistical classes, then the conditional probability

$$\varphi_p(F_1 \subseteq F_2) = P_1(A_2(p)|A_1(p)) = \frac{P_1(A_2(p) \cap A_1(p))}{P_1(A_1(p))}$$

characterizes the probability of observation of class F_2 given F_1. Obviously,

$$\varphi_p(F_1 \subseteq F_2) = 1 \text{ if } F_1 \subseteq F_2.$$

Therefore, we call $\varphi_p(F_1 \subseteq F_2)$ the *local inclusion measure*.

Integral inclusion measure of regular statistical classes

The integral inclusion measure is defined as follows

$$\varphi(F_1 \subseteq F_2) = \int_0^1 w(p)\varphi_p(F_1 \subseteq F_2)dp = \int_0^1 w(p)P_1(A_1(p) \cap A_2(p))dp.$$

where w is a non-negative weight function such that $\int_0^1 w(p)dp = 1.$

In the sequel we assume that $w(p) = 2p$. Then

$$\varphi(F_1 \subseteq F_2) = 2 \int\limits_0^1 P_1(A_1(p) \cap A_2(p))dp,$$

Theorem. Let F_1 and F_2 be regular statistical classes. Then $\varphi(F_1 \subseteq F_2) = 1$ iff $F_1 \subseteq F_2$.

Let F be a regular statistical class with density $h(x)$ and corresponding probability measure P. Introduce the function

$$\mu(x) = P\left(\{y \in X | h(y) \leqslant h(x)\}\right).$$

This is the *membership function* of the statistical class F, i.e. it gives its fuzzy representation.

Proposition 1. Any regular statistical class is defined by its membership function μ uniquely. In addition, $\{x \in X | \mu(x) > 1 - p\}$ is a minimal event with probability $p \in (0, 1]$.

Corollary. Let F_1 and F_2 be regular statistical classes with membership functions μ_1 and μ_2. Then

1. $F_1 \subseteq F_2$ iff $\mu_1 \leqslant \mu_2$ in measure V;
2. $\varphi(F_1 \subseteq F_2) = 2 \int\limits_X \min\{\mu_1(x), \mu_2(x)\} dP_1$.

Remark. Formula 2 in the corollary has the following interpretation through fuzzy sets. Let F_1 and F_2 be fuzzy sets with membership functions μ_1 and μ_2. Then

$$\varphi(F_1 \subseteq F_2) = P_1(F_2|F_1) = \frac{P_1(F_1 \cap F_2)}{P_1(F_1)}.$$

If we accept probabilities of fuzzy events, proposed by L. Zadeh:

$$P_1(F_1 \cap F_2) = \int_X \min\{\mu_1(x), \mu_2(x)\} dP_1,$$

$$P_1(F_1) = \int_X \mu_1(x) dP_1 = 0.5.$$

Classification of statistical classes

In this case we have the set of etalon statistical classes $\{S_1, ..., S_n\}$. The classification of observed statistical class F consists in computing the following classifying vector:

$$(\varphi(F \subseteq S_1), ..., \varphi(F \subseteq S_n))\,.$$

The description of S_i is obtained using learning samples. It is sufficient to know only fuzzy representations of S_i.

The computation of $\varphi(F \subseteq S_i)$ is produced by the following steps:

1. Estimation of μ_F (membership function of F).

2. $\varphi(F \subseteq S_i) \approx \frac{2}{N} \sum_{i=1}^{N} \min\{\mu_F(x_i), \mu_{S_i}(x_i)\}$, where $\{x_1, ..., x_N\}$ is an independent sample from the class F.

Decision making in case of unknown utility function

Classical scheme of decision making: Assume that any decision d_i is associated with a probability measure P_i on a measurable space $(X, \mathfrak{B})$ and evaluation of d_i is based on computing the expected utility

$$u(d_i) = \int\limits_X u(x) dP_i,$$

where $u : X \to \mathbb{R}$ is the utility function.

If utility is measured in order scale

Sometimes it is hard to evaluate decisions using real numbers. In this case we assume that $u : X \to R$, where R is a linearly ordered set.

In the sequel, it is more convenient to range decisions considering probability measures on the algebra $\mathfrak{A}$ of space R. Evidently, this algebra is generated as follows

$$A \in \mathfrak{B} \Rightarrow u(A) = \{u(x)|x \in A\} \in \mathfrak{A},$$

and probability measure P on $\mathfrak{B}$ generates probability measure on $\mathfrak{A}$ in a way that sets $A \in \mathfrak{B}$ and $u(A)$ have the same probabilities.

Therefore, we will describe decisions with the help of probability measures on the measurable space $(R, \mathfrak{A})$.

We denote the order on R by $\preccurlyeq$ and use symbol $\prec$ if $r_i \preccurlyeq r_j$ and $r_i \neq r_j$.

We assume also that algebra $\mathfrak{A}$ is the minimal σ -algebra generated by sets

$$[r, +\infty) \subseteq \{x \in R | r \preccurlyeq x\}.$$

For simplicity, assume that R is a real line, then $\mathfrak{A}$ is the Borel algebra, and we can extend the order on incomes $r_i \in R$ to the order on probability measures as follows:

$$P_1 \preccurlyeq P_2 \text{ iff } P_1[r, +\infty) \leqslant P_2[r, +\infty) \text{ for all } r \in R.$$

Evidently, the function $F_i(r) = P_i[r, +\infty)$ is the cumulative distribution function and such a partial order defined on probability distributions (when $R = \mathbb{R}$) is called *stochastic dominance*.

We will next express stochastic dominance through fuzzy sets assuming that such functions are fuzzy sets.

Notation. M_{pr} is the set of all probability measures on $\mathfrak{A}$.

Property 1. Let P_1 and P_2 be probability measures on $\mathfrak{A}$, F_1 and F_2 be their corresponding cumulative distribution functions. Then $P_1 \preccurlyeq P_2$ iff $F_1 \subseteq F_2$.

Property 2. The set M_{pr} of all probability measures on $\mathfrak{A}$ is a distributive lattice, and this lattice is isomorphic to the lattice of corresponding fuzzy sets, i.e.

1) if $P_1 \wedge P_2 \in M_{pr}$ is the exact lower bound of $\{P_1, P_2\}$, then its cumulative distribution function is $F_1 \cap F_2 = \min(F_1, F_2)$;

2) if $P_1 \vee P_2 \in M_{pr}$ is the exact upper bound of $\{P_1, P_2\}$, then its cumulative distribution function is $F_1 \cup F_2 = \min(F_1, F_2)$.

Remark. Notice that the set of all cumulative distribution functions does not cover all possible fuzzy subsets of R. It contains only so called comonotone fuzzy sets.

Fuzzy sets F_1 and F_2 are called *comonotone*, if for any t_1, t_2 one of the next two inclusions

$$\{F_1 > t_1\} \supseteq \{F_1 > t_1\} \text{ or } \{F_1 > t_1\} \subseteq \{F_2 > t_2\}$$

is necessarily true.

Possibilistic inclusion

Obviously, each cumulative distribution function F_i is a normal fuzzy set, i.e. $\sup\limits_{x \in R} F_i(x) = 1$. It allows us to consider corresponding necessity and possibility measures:

$N_i(A) = \inf\limits_{x \notin A} (1 - F_i(x))$ for $A \in \mathfrak{A}$ such that $A \neq R$ $(N_i(R) = 1)$;

$\Pi_i(A) = \sup\limits_{x \in A} F_i(x)$, for $A \in \mathfrak{A}$ such that $A \neq \emptyset$.

Theorem. Let P_1 and P_2 be probability measures on $\mathfrak{A}$. Then $P_1 \preccurlyeq P_2$ if $N_2(A) \leqslant P_1(A) \leqslant \Pi_2(A)$ for all $A \in \mathfrak{A}$.

The above probabilistic interpretation allows us to use inclusion indices introduced for statistical classes for ranging decisions.

Regular case (cumulative distribution functions are continuous

Property. Let a cumulative distribution function F_i is continuous, then $P_i(\{F > 1 - p\} = p$.

Based on the above property, consider the integral inclusion index for decisions if they are described by continuous distribution functions.

$$\varphi(F_1 \subseteq F_2) = 2\int_0^1 P_1(A_1(p) \cap A_2(p))dp,$$

where $A_i(p) = \{x \in R | F_i(x) > 1 - p\}$, $p \in [0, 1]$.

Preference function

Extension of inclusion index to the irregular case is based on analyzing behavior of the *preference function*

$$\psi(F_1, F_2) = \int\limits_0^1 P_1(A_2(p))dp.$$

Obviously, for regular case:

$$\varphi(F_1 \subseteq F_2) = 2\psi(F_1, F_1 \cap F_2).$$

Properties of ψ for regular case

1. $\psi\left(F_1, F_2\right) = 1 - \psi\left(F_2, F_1\right)$.
2. $\psi\left(F_1, F_2\right) = 0.5$ if $F_1 = F_2$; $\psi\left(F_1, F_2\right) < 0.5$ if $F_1 \supset F_2$; $\psi\left(F_1, F_2\right) > 0.5$ if $F_1 \subset F_2$.
3. $\psi\left(aF_1 + (1-a)F_2, F_3\right) = a\psi\left(F_1, F_3\right) + (1-a)\psi\left(F_2, F_3\right)$, $a \in [0, 1]$.
4. $\psi\left(F_1 \subseteq F_2\right) - \psi\left(F_2 \subseteq F_1\right) = \psi\left(F_1, F_2\right) - \psi\left(F_2, F_1\right) = 1 - 2\psi\left(F_2, F_1\right)$.
5. $\psi\left(F_1 \subseteq F_2\right) = 2\psi\left(F_1 \cup F_2, F_2\right) = \psi\left(F_1, F_1 \cap F_2\right) + \psi\left(F_1 \cup F_2, F_2\right)$.

Theorem. Let the functional $\psi\left(F_1, F_2\right)$ be continuous and obey properties 1-3. In addition, $\psi\left(F_1, F_2\right) = 0$ if $F_1 \supset F_2$; $\psi\left(F_1, F_2\right) = 1$ if $F_1 \subset F_2$ for crisp sets.Then it is defined uniquely and can be computed as

$$\psi(F_1, F_2) = 0.5\left(\int\limits_0^1 P_1\left\{A_2(p)\right\}dp + 1 - \int\limits_0^1 P_2\left\{A_1(p)\right\}dp\right),$$

where P_i is a probability measure that corresponds to the cumulative distribution function F_i and $A_i(p) = \{x \in R | F_i(x) > 1 - p\}$.

To preserve property 4, using preference function we can define the inclusion index as

$$\psi\left(P_1 \preccurlyeq P_2\right) = \psi\left(F_1 \subseteq F_2\right) = \tilde{\psi}\left(F_1, F_1 \cap F_2\right) + \tilde{\psi}\left(F_1 \cup F_2, F_2\right).$$

Probabilistic interpretation of preference function

Assume that probability measures P_1 and P_2 describe random values ξ_1 and ξ_2. If we additionally assume that these variables are independent, then

$$\psi(P_1, P_2) = 0.5\left(\Pr\left\{\xi_1 \preccurlyeq \xi_2\right\} + \Pr\left\{\xi_1 \prec \xi_2\right\}\right).$$

Decision making by inclusion index

Let $D = \{d_1, d_2, ..., d_m\}$ be the set of possible decisions, and to each decision d_i corresponds a probability measure $P_i \in M_{pr}$.

Obviously, we can recover stochastic dominance, computing $\psi\left(P_1 \preccurlyeq P_2\right)$, since $P_1 \preccurlyeq P_2$ iff $\psi\left(P_1 \preccurlyeq P_2\right) = 1$.

Let us analyze what happens if we extend $\preccurlyeq$ to the relation $\Psi \subseteq D \times D$, such that $(d_i, d_j) \in \Psi$, if $\psi(P_i \preccurlyeq P_j) \geqslant \psi(P_j \preccurlyeq P_i)$.

Then we can get the following properties:

Property 1. $(d_i, d_j) \in \Psi$ iff $\tilde{\psi}(P_i, P_j) \geqslant 0.5$.

Property 2. Ψ is non-transitive in general.

Example. Let $R = \{r_1, ..., r_N\}$, $\mathfrak{A} = 2^R$, and $r_1 \prec r_2 \prec ... \prec r_N$. Then

$$\psi\left(P_1, P_2\right) =$$

$$0.5 \sum_{i=1}^{N} P_1(\{r_i\}) P_2(\{r_i\}) + \sum_{i=1}^{N-1} P_1(\{r_i\}) \sum_{j=i+1}^{N} P_2(\{r_i\}).$$

Let $N = 4$ and probability measures P_1, P_2, P_3, P_4 are defined in the following table.

	$\{r_1\}$	$\{r_2\}$	$\{r_3\}$	$\{r_4\}$
P_1	0.35	0	0.4	0.25
P_2	0.2	0.2	0.4	0.2
P_3	0.32	0	0.48	0.2

$\tilde{\psi}(P_1, P_2) = 0.5$, $\tilde{\psi}(P_2, P_3) = 0.5$,
$\tilde{\psi}(P_1, P_3) = 0.495$,

$(d_1, d_2) \in \Psi$, $(d_2, d_3) \in \Psi$, $(d_1, d_3) \notin \Psi$: Ψ is a non-transitive relation.

Logical inference in possibility theory based on upper and lower probabilities

Notation.

X is a measurable space with σ-algebra $\mathfrak{A}$;

$\tilde{A}_i$ are fuzzy subsets of X whose measurable membership functions μ_i are normal, i.e. $\sup_{x \in X} \mu_i(x) = 1$.

$\xi \in \tilde{A}_i$ is a fuzzy value.

Probabilistic interpretation

ξ is a random value, and the fuzzy set $\tilde{A}_i$ is the imprecise description of ξ. This imprecise description is given by

possibility measure: $\Pi_i(A) = \sup\limits_{x \in A} \mu_i(x)$, $A \in \mathfrak{A}$, $A \neq \emptyset$ ($\Pi_i(\emptyset) = 0$);

necessity measure: $N_i = \Pi_i^d$;

and by inequalities

$$N_i(A) \leqslant \Pr\{\xi \in A\} \leqslant \Pi_i(A),\ A \in \mathfrak{A}.$$

In other words, ξ is described by the family of probability measures

$$\Xi_i = \{P \in M_p r \,|N_i(A) \leqslant P(A) \leqslant \Pi_i(A), A \in \mathfrak{A}\}\,.$$

Conditions of propositions inconsistency

In probabilistic setting propositions $\xi \in \tilde{A}_1, ..., \xi \in \tilde{A}_m$ are inconsistent if $\bigcap_{i=1}^{m} \Xi_i \neq \emptyset$.

Notation: $A_k(p_k) = \{x \in X \,|1 - \mu_k(x) < p_k\}$ is the strict $(1 - p_k)$-cut of $\tilde{A}_k$.

Proposition 1. Propositions $\xi \in \tilde{A}_1$ and $\xi \in \tilde{A}_2$ are inconsistent iff $\exists p_1, p_2 \in [0,1]$ such that$A_1(p_1) \cap A_2(p_2) = \emptyset$, $p_1 + p_2 > 1$.

Proposition 2. Propositions $\xi \in \tilde{A}_1, ..., \xi \in \tilde{A}_m$ are inconsistent if $\exists p_1, ..., p_m \in [0,1]$ such that $A_i(p_i)$, $i = 1, ..., m$, are pairwise disjoint sets and $\sum_{i=1}^{m} p_i > 1$.

Logical inference

$\xi \in \tilde{A}_1, ..., \xi \in \tilde{A}_m \Rightarrow \xi \in \tilde{A}$ means that $\bigcap_{i=1}^{m} \Xi_i \subseteq \Xi$, where $\Xi_i,\ i = 1, ..., m,\ \Xi$ are families of probability measures that correspond to $\xi \in \tilde{A}_i$, $\xi \in \tilde{A}$.

Question: Is it possible to use usual operation $\min$, used for finding intersection of fuzzy sets? The answer is no. It is possible in case iff fuzzy sets $\tilde{A}_1, ..., \tilde{A}_m$ are comonotone. In this case the set $\bigcap_{i=1}^{m} \Xi_i$ is described by possibility distribution $\Pi(A) = \min\limits_{i} \Pi_i(A)$. In other cases we should construct other rules of inference.

Other rules of inference are based on the following lemma.

Lemma. Let $\{A_k\}_{k=1}^{n} \subseteq \mathfrak{A}$ and non-empty set $\Xi \subseteq M_{pr}$ is given by $P \in \Xi \Leftrightarrow P\{A_k\} \geqslant p_k$, $p_k \in [0,1]$, $k = 1, ..., n$. Then

$$P\left\{\bigcap_{i=1}^{n} A_i\right\} \geqslant \left(\sum_{i=1}^{n} p_i\right) - (n-1),\, P \in \Xi$$

and the above estimate cannot be improved without additional information if $\left(\sum_{i=1}^{n} p_i\right) - (n-1) > 0$ and $\overline{A_k} \cap \bigcap_{i|i \neq k} A_i \neq \emptyset$ for all $k \in \{1, ..., n\}$.

Proposition. Let $\varphi_k : [0, +\infty) \to [0, +\infty)$, $k = 1, 2$, be continuous strictly increasing functions, in addition, $\forall \lambda \in [0, +\infty) : \varphi_1(\lambda) + \varphi_2(\lambda) = \lambda$.

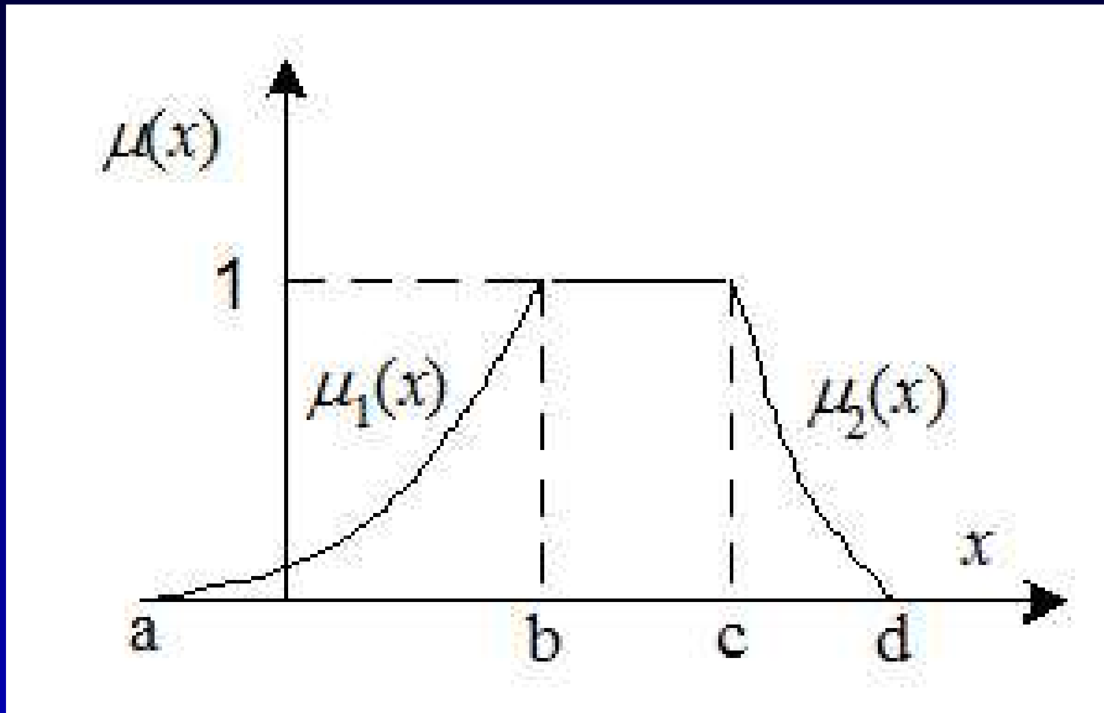
Then for consistent propositions $\xi \in \tilde{A}_1$, $\xi \in \tilde{A}_2$ is valid

$$\xi \in \tilde{A}_1, \xi \in \tilde{A}_2 \Rightarrow \xi \in \tilde{A},$$

where $\mu(x) = \ 1 \wedge \varphi_1^{-1}(\mu_1(x)) \wedge \varphi_2^{-1}(\mu_2(x))$, $x \in X$.

Example. $\mu(x) = 1 \wedge 2\mu_1(x) \wedge 2\mu_2(x)$ if $\varphi_1(\lambda) = \varphi_2(\lambda) = \lambda/2$.

Imprecision characteristics of fuzzy interval



$$\mu(x) = \begin{cases} 0, & x \leqslant a \ or \ x \geqslant d, \\ \mu_1(x), & a < x < b, \\ 1, & b \leqslant x \leqslant c, \\ \mu_2(x), & c < x < d. \end{cases}$$

μ_1 is strictly increasing on $[a, b]$;
μ_2 is strictly decreasing on $[c, d]$;
μ is continuous;

$\xi \in \tilde{A} \Leftrightarrow N_i(A) \leqslant \Pr\{\xi \in A\} \leqslant \Pi_i(A)$.

$\underline{E}\left[\tilde{A}\right] = \inf_{\xi \in \tilde{A}} E[\xi]$ is the exact lower bound of expectation.

$\overline{E}\left[\tilde{A}\right] = \sup_{\xi \in \tilde{A}} E[\xi]$ is the exact upper bound of expectation.

$\overline{D}\left[\tilde{A}\right] = \sup\limits_{\xi \in \tilde{A}} D\left[\xi\right]$ is the maximal variance.

It possible to show that $\overline{E}\left[\tilde{A}\right] - \underline{E}\left[\tilde{A}\right] = \int\limits_{-\infty}^{+\infty} \mu(x)dx$.