The Extension of Imprecise Probabilities Based on Generalized Credal Sets

Andrey G. Bronevich[1],
Igor N. Rozenberg[2]

[1] National Research University "Higher School of Economics", Moscow, Russia

[2] JSC Research, Development and Planning Institute for Railway Information Technology, Automation and Telecommunication, Moscow, Russia

8th International Conference on Soft Methods
in Probability and Statistics,
12-14 September 2016, Rome, Italy

The aim of the investigation

To develop theory of imprecise probailities that works with contradictory information when the avoiding sure condition is not fulfilled.This theory can be based on generalized credal sets introduced in

paper

Bronevich A.G., Rozenberg I.N. The generalization of the conjunctive rule for aggregating contradictory sources of information based on generalized credal sets. Proceedings of ISIPTA-15.

Remark

Contradiction in information can be caused by inconsistent assessments of a decision-maker, after combining information from contradictory sources.

Credal sets

Let X be a finite set and 2^X be the powerset of its subsets.

Definition

A family $\mathbf{P}$ of probability measures on 2^X is called a *credal set* if it is convex and closed.

Credal sets allow us to model randomness (conflict) by of probability measures and inderminacy (non-specificity)in choosing a probability measure in the credal set.

Lower previsions

Let K be the set of all functions $f : X \to \mathbb{R}$ and $K' \subseteq K$. A functional $\underline{E} : K' \to \mathbb{R}$ is called a lower prevision if values $\underline{E}(f)$ are conceived as lower estimates of expectations of random variables $f \in K$.

Notation

- M_{pr} is the set of all probability measures;
- $E_P(f) = \sum_{x \in X} f(x)P(\{x\})$ is the expectation of $f \in K$ w.r.t. $P \in M_{pr}$.

The functional $\underline{E}$ *avoids sure loss* if the set of probability measures

$$\mathbf{P} = \left\{ P \in M_{pr} | \forall f \in K' : \underline{E}(f) \leqslant E_P(f) \right\} \tag{1}$$

is not empty. In this case $\mathbf{P}$ is a credal set. In the opposite case, when $\mathbf{P} = \emptyset$, $\underline{E}$ is contradictory.

Natural extension

Let a credal set $\mathbf{P}$ corresponds to $\underline{E}$. Then the *natural extension* $\underline{E}_{\mathbf{P}} : K \to \mathbb{R}$ of $\underline{E}$ is defined by

$$\underline{E}_{\mathbf{P}}(f) = \inf \{E_P(f) | P \in \mathbf{P}\}.$$

$\underline{E}$ is called a *coherent lower prevision* if $\underline{E}_{\mathbf{P}}(f) = \underline{E}(f)$, $f \in K'$.

Upper previsions

$\overline{E} : K' \to \mathbb{R}$ is called an *upper prevision* if it gives us upper expections of random variables in K'. It is *non-contradictory* if it defines a non-empty credal set

$$\mathbf{P} = \left\{ P \in M_{pr} | \forall f \in K' : \overline{E}(f) \geqslant E_P(f) \right\}$$

and its natural extension $\overline{E}_{\mathbf{P}} : K \to \mathbb{R}$ of $\underline{E}$ on K is defined by

$$\overline{E}_{\mathbf{P}}(f) = \sup \left\{ E_P(f) | P \in \mathbf{P} \right\}.$$

If $\overline{E}_{\mathbf{P}}(f) = \overline{E}(f)$, $f \in K'$, then $\overline{E}(f)$ is called a *coherent upper prevision*.
Models based on credal sets, upper and lower coherent previsons are equivalent if X is finite.

The conjunctive rule in the theory of imprecise probabilities

Definition

The *conjunctive rule* (C-rule) for credal sets $\mathbf{P}_1, ..., \mathbf{P}_m$ is defined as

$$\mathbf{P} = \mathbf{P}_1 \cap ... \cap \mathbf{P}_m \tag{1}$$

Remark

The C-rule is defined only for the case when $\mathbf{P}$ is a non-empty set.

Let $\underline{E}_i : K' \to \mathbb{R}$, $i = 1, .., m$, be lower previsions on K'.
Then the result of the C-rule can be expressed as

$$\underline{E}(f) = \max_{i=1,...,m} \underline{E}_i(f),\ f \in K'.$$

Analogously for upper previsions the C-rule is defined using the max operation.

The case of fully contradictory sources of information

Sources of information described by $P_1, P_2 \in M_{pr}$. They are *fully contradictory* if there are disjoint sets $A, B \in 2^X$ such that $P_1(A) = P_2(B) = 1$.Fully contradictory information is described by a set function

$$\eta^d_{\langle X \rangle}(A) = \begin{cases} 1, & A \neq \emptyset, \\ 0, & A = \emptyset, \end{cases}$$

concieved as a lower probability.

Notation

- $\eta_{\langle B \rangle}(A) = \begin{cases} 1, & B \subseteq A, \\ 0, & otherwise. \end{cases}$ $\eta_{\langle B \rangle}$ is a categorical belief function.
- let μ be a monotone measure, i.e. $\mu(\emptyset) = 0$, $\mu(X) = 1$, $A \subseteq B$ implies $\mu(A) \leqslant \mu(B)$, then the dual of μ is defined as $\mu^d(A) = 1 - \mu(\bar{A})$, $A \in 2^X$.

The conjunctive rule for probability measures

If probability measures P_1 and P_2 are fully contradictory we define the conjunctive rule as $P_1 \wedge P_2 = \eta^d_{\langle X \rangle}$.
If there is no contradiction between probability measures P_1 and P_2, i.e. $P_1 = P_2 = P$. Then

$$P_1 \wedge P_2 = P.$$

Lemma

For any $P_1, P_2 \in M_{cpr}$ It is always possible to find representations

$$P_1 = (1-a)P^{(1)} + aP_1^{(2)},\ P_2 = (1-a)P^{(1)} + aP_2^{(2)},$$

such that $P^{(1)}, P_1^{(2)}, P_2^{(2)} \in M_{pr}$, $a \in [0,1]$, and $P_1^{(2)}$, $P_2^{(2)}$ are fully contradictory. These representations are defined uniquely if $a \in (0,1)$.

The conjunctive rule for probability measures

Thus, we define the C-rule for arbitrary probability measures as

$$P_1 \wedge P_2 = (1-a)P^{(1)} + a(P_1^{(2)} \wedge P_2^{(2)}).$$

The last formula can be rewritten as

$$P_1 \wedge P_2 = \sum_{i=1}^{n} \min\{P_1(\{x_i\}), P_2(\{x_i\})\}\, \eta_{\langle\{x_i\}\rangle} + a\eta^d_{\langle X\rangle}.$$

where $a = 1 - \sum_{i=1}^{n} \min\{P_1(\{x_i\}), P_2(\{x_i\})\}$.

The value a is called the *amount of contradiction* between P_1 and P_2, denoted by $Con(P_1, P_2)$.

Example. Conjunction of probability measures.

Let $X = \{x_1, x_2, x_3\}$ and probability measures P_1 and P_2 on 2^X are given in the following table.

	$\{x_1\}$	$\{x_2\}$	$\{x_3\}$	
P_1	0.3	0.2	0.5	
P_2	0.3	0.3	0.4	
$\min\{P_1, P_2\}$	0.3	0.2	0.4	$a = 0.1$

Then $P_1 \wedge P_2 = 0.3\eta_{\langle\{x_1\}\rangle} + 0.2\eta_{\langle\{x_2\}\rangle} + 0.4\eta_{\langle\{x_3\}\rangle} + 0.1\eta^d_{\langle X\rangle}$.

Notation

M_{cpr} is the set of all set functions of the type

$$P = a_0\eta^d_{\langle X\rangle} + \sum_{i=1}^{n} a_i\eta_{\langle\{x_i\}\rangle},$$

where $a_i \geqslant 0$, $i = 0, ..., n$, $\sum_{i=0}^{n} a_i = 1$.

The definition of the conjunctive rule for general measures in M_{cpr}

The interpretation of the conjunctive rule through the order $\leqslant$

Let $P_1, P_2 \in M_{pr}$. Then $P_1 \wedge P_2$ is the exact lower bound of the set $\{P \in M_{cpr} | P \geqslant P_1, P \geqslant P_2\}$.

It allows us to define the conjunctive rule for arbitrary $P_1, ..., P_m \in M_{cpr}$ as an exact lower bound of the set

$$\{P \in M_{cpr} | P \geqslant P_1, ..., P \geqslant P_m\}.$$

If any $P = a_0 \eta^d_{\langle X \rangle} + \sum_{i=1}^{n} a_i \eta_{\langle \{x_i\} \rangle}$ is identified with a point $(a_1, ..., a_n)$ in $\mathbb{R}^n$. Then $P = P_1 \wedge ... \wedge P_m$ for $P = (b_1, ..., b_n)$, $P_i = (a_1^{(i)}, ..., a_n^{(i)})$ if $b_k = \min\limits_{i=1,..,m} a_k^{(i)}$.

Generalized credal sets

Definition

A subset $\mathbf{P} \subseteq M_{cpr}$ is called an *upper generalized credal set* if

1. $P_1 \in \mathbf{P}$, $P_2 \in M_{cpr}$, $P_1 \leqslant P_2$ implies that $P_2 \in \mathbf{P}$.
   (The next two properties are essential for the most models of imprecise probabilities (cf. credal sets).)
2. if $P_1, P_2 \in \mathbf{P}$ then $aP_1 + (1-a)P_2 \in \mathbf{P}$ for any $P_1, P_2 \in \mathbf{P}$ and $a \in [0,1]$.
3. the set $\mathbf{P}$ is closed in a sense that it can be considered as a subset of Euclidian space (any $P = a_0 \eta^d_{\langle X \rangle} + \sum_{i=1}^{n} a_i \eta_{\langle \{x_i\} \rangle}$ is a point $(a_1, ..., a_n)$ in $\mathbb{R}^n$).

Generalized credal sets

Notation

$M^d_{cpr} = \{P^d | P \in M_{cpr}\}$

If we consider measures from $\bar{M}_{cpr}$ as (contradictory) lower probabilities, then any $P \in M^d_{cpr}$:

$$P = a_0 \eta_{\langle X \rangle} + \sum_{i=1}^{n} a_i \eta_{\langle \{x_i\} \rangle},$$

can be concieved as the (contradictory) upper probability.

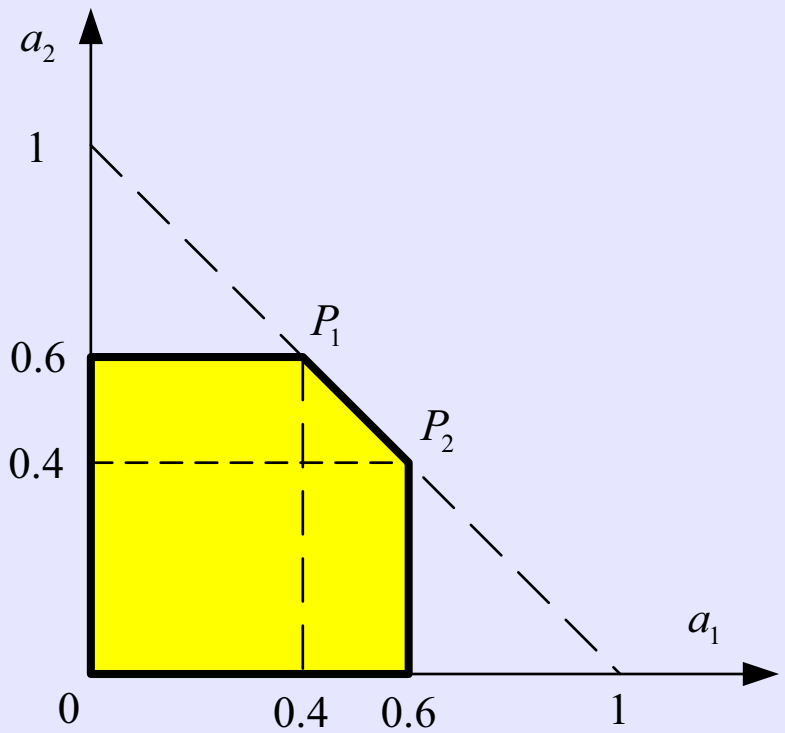
Definition

$\mathbf{P}$ is the lower generalized credal set if $\mathbf{P}^d = \left\{P^d | P \in \mathbf{P}\right\}$ is the upper generalized credal set.

The profile of a generalized credal set

Let $\mathbf{P}$ be a upper generalized credal set in M_{cpr}. A subset $profile(\mathbf{P})$ consisting of all minimal elements in $\mathbf{P}$ is called the *profile* of $\mathbf{P}$.

Example. Let $X = \{x_1, x_2\}$ and an upper generalized credal set $\mathbf{P}$ is defined by a $profile(\mathbf{P}) = \{aP_1 + (1-a)P_2 | a \in [0,1]\}$, where $P_1 = (0.4, 0.6)$, $P_2 = (0.6, 0.4)$. Then $\mathbf{P}$ can be depicted in $\mathbb{R}^2$ as



The profile of a generalized credal set

Any profile uniquely defines the corresponding credal set.

If $\mathbf{P}$ describes information without contradiction, then $profile(\mathbf{P})$ is a credal set in usual sense, i.e. $profile(\mathbf{P})$ is a set of probability measures.

Analogously, the profile of lower generalized credal set is defined.

Let $\mathbf{P}$ be a lower generalized credal set in M^d_{cpr}. A subset consisting of all maximal elements in $\mathbf{P}$ is called the profile of $\mathbf{P}$ and it is denoted by $profile(\mathbf{P})$.

Obviously, if $\mathbf{P}$ be an upper generalized credal set in $\bar{M}_{cpr}$, then $\mathbf{P}^d$ is the lower generalized credal set in $\bar{M}^d_{cpr}$ and

$$profile(\mathbf{P}^d) = profile(\mathbf{P})^d.$$

The C-rule for generalized credal sets

Definition.

Let $\mathbf{P}_1, ..., \mathbf{P}_m$ be non-empty upper generalized credal sets in M_{cpr}. Then the credal set $\mathbf{P}$ produced by the C-rule is defined as $\mathbf{P} = \mathbf{P}_1 \cap ... \cap \mathbf{P}_m$.

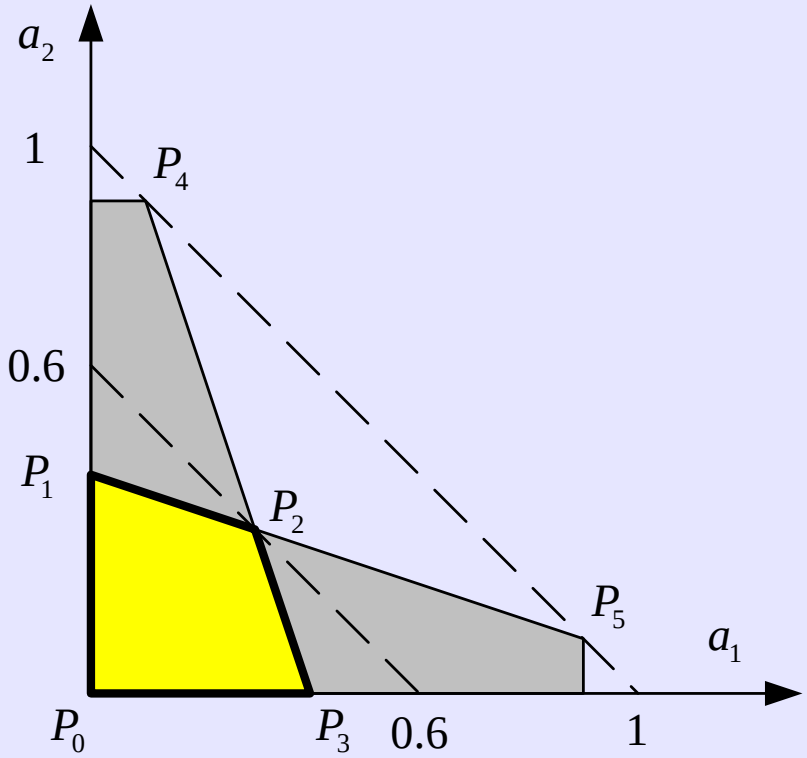
Let us observe that this definition generalizes the introduced C-rule for probability measures. Actually, let we have two credal sets $\mathbf{P}_1, \mathbf{P}_2$ in $\bar{M}_{cpr}$ with $profile(\mathbf{P}_i) \in M_{pr}$, where M_{pr} is the set of all probability measures on 2^X. Then

$$profile(\mathbf{P}_1 \cap \mathbf{P}_2) = profile(\mathbf{P}_1) \wedge profile(\mathbf{P}_2).$$

In the same way the C-rule for lower generalized credal sets are defined.

The C-rule for generalized credal sets (Example)

Let $X = \{x_1, x_2\}$ and upper generalized credal sets $\mathbf{P}_1$ and $\mathbf{P}_2$ are given by their profiles: $profile(\mathbf{P}_1) = \{aP_3 + (1-a)P_4 | a \in [0,1]\}$, $profile(\mathbf{P}_2) = \{aP_1 + (1-a)P_5 | a \in [0,1]\}$, where $P_1 = (0, 0.4)$, $P_3 = (0.4, 0)$, $P_4 = (0.1, 0.9)$, $P_5 = (0.9, 0.1)$.

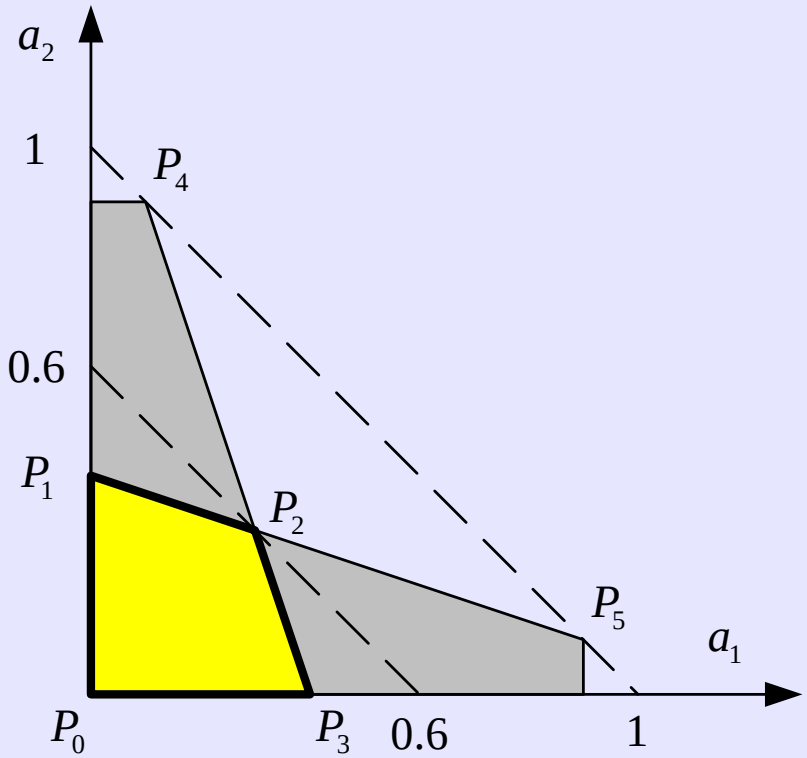


The C-rule for generalized credal sets (Example)

Then their intersection (conjunction) is the credal set $\mathbf{P}_3$ with

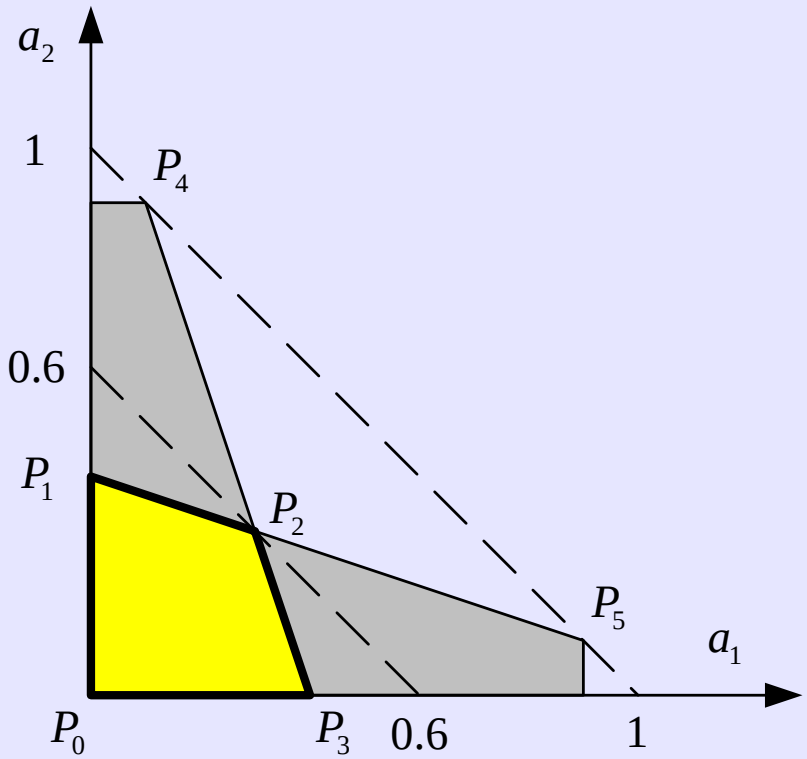
$$profile(\mathbf{P}_3) = \{aP_1 + (1-a)P_2 | a \in [0,1]\} \cup \{aP_2 + (1-a)P_3 | a \in [0,1]\}.$$

Remark $profile(\mathbf{P}_3)$ is not a convex set.



The amount of contradiction of a generalized credal set

Let $\mathbf{P}$ is a generalized credal set in M_{cpr}. Then the amount of contradiction is defined as $Con(\mathbf{P}) = \inf\{Con(P)|P \in \mathbf{P}\}$.
From previous example: $Con(\mathbf{P}_1) = Con(P_4) = 0$,
$Con(\mathbf{P}_2) = Con(P_5) = 0$, $Con(\mathbf{P}_3) = Con(P_2) = 0.4$,$P_2 = (0.3, 0.3)$,
$Con(P_2) = 1 - 0.3 - 0.3 = 0.4$.



The ways of defining generalized credal sets

Let lower probability $P \in M_{cpr}$, i.e. $P = a_0 \eta^d_{\langle X \rangle} + \sum_{i=1}^{n} a_i \eta_{\langle \{x_i\} \rangle}$ and $f : X \to \mathbb{R}$.
Then the lower expectation $\underline{E}_P(f)$ of f w.r.t. P can be computed by the Choquet integral:

$$\underline{E}_P(f) = (C) \int f dP = a_0 \max_{x \in X} f(x) + \sum_{i=1}^{n} a_i f(x_i).$$

Let $\underline{E} : K' \to \mathbb{R}$ be a lower prevision. Then it defines the upper generalized credal set

$$\mathbf{P}(\underline{E}) = \left\{ P \in \bar{M}_{cpr} | \forall f \in K' : \ \underline{E}_P(f) \geqslant \underline{E}(f) \right\}.$$

This set is not empty if $\underline{E}(f) \leqslant \max_{x \in X} f(x)$ for all $f \in K'$.

The ways of defining generalized credal sets

The same constructions can be defined for lower generalized credal sets.

If $P \in M^d_{cpr}$, i.e. $P = a_0\eta_{\langle X\rangle} + \sum\limits_{i=1}^{n} a_i\eta_{\langle\{x_i\}\rangle}$ and $f : X \to \mathbb{R}$.

Then the upper expectation $\overline{E}_P(f)$ of f w.r.t. P can be computed by the Choquet integral:

$$\overline{E}_P(f) = (C)\int f dP = a_0 \min_{x\in X} f(x) + \sum_{i=1}^{n} a_i f(x_i).$$

Let $\overline{E} : K' \to \mathbb{R}$ be an upper prevision. Then it defines the lower generalized credal set

$$\mathbf{P}(\overline{E}) = \left\{ P \in M^d_{cpr} | \forall f \in K' :\ \bar{E}_P(f) \leqslant \bar{E}(f) \right\}.$$

This set is not empty if $\overline{E}(f) \geqslant \min\limits_{x\in X} f(x)$ for all $f \in K'$.

Upper and lower previsions based on generalized credal sets

We can define the lower prevision based on non-empty upper generalized credal set $\mathbf{P}$ as

$$\underline{E}_{\mathbf{P}}(f) = \inf_{P \in \mathbf{P}} \underline{E}_P(f),\ f \in K.$$

Analogously, we can define the upper prevision based on non-empty lower generalized credal set $\mathbf{P}$ as

$$\overline{E}_{\mathbf{P}}(f) = \sup_{P \in \mathbf{P}} \bar{E}_P(f),\ f \in K.$$

Functionals $\underline{E}_{\mathbf{P}}$ and $\overline{E}_{\mathbf{P}}$ can be considered as counterparts of coherent lower und upper previsions in the theory of imprecise probabilities without contradiction.

Necessary and sufficient conditions for a fucntional to be $\overline{E}_{\mathbf{P}}$

A function $f : X \to \mathbb{R}$ is called normalized if $\min_{x \in X} f(x) = 0$.

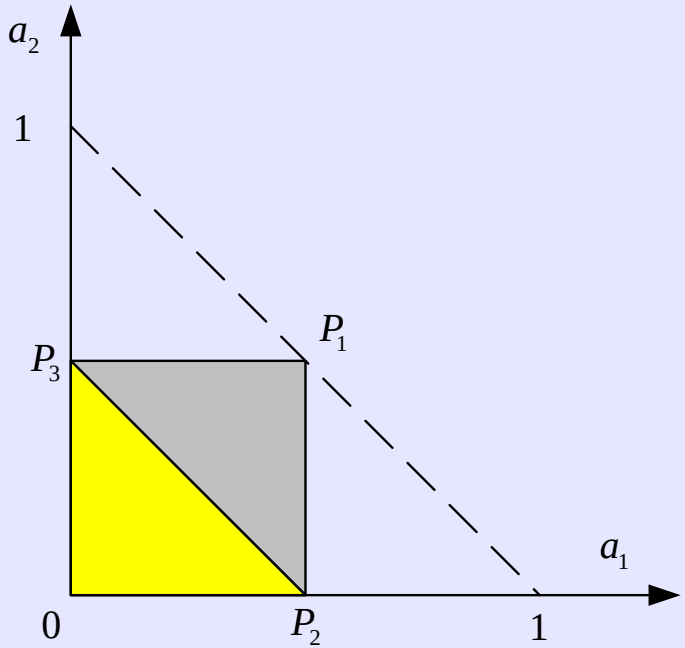
Theorem

A functional $\Phi : K \to \mathbb{R}$ coincides with $\overline{E}_{\mathbf{P}}$ on K for some credal set $\mathbf{P}$ in M^d_{cpr} iff it has the following properties:
1) $\Phi(f + a1_X) = \Phi(f) + a$ for any $f \in K$ and $a \in \mathbb{R}$;
2) $\Phi(af) = a\Phi(f)$ for any $f \in K$ and $a \geqslant 0$;
3) $\Phi(f_1) \leqslant \Phi(f_2)$ for $f_1, f_2 \in K$ if $f_1 \leqslant f_2$;
4) $\Phi(f_1) + \Phi(f_2) \geqslant \Phi(f_3)$ for any normalized functions f_1, f_2, f_3 in K such that $f_1 + f_2 = f_3$.

Remark If 4) in Theorem will be fulfilled for every possible functions f_1, f_2, f_3 in K (not necessarily normalized), then conditions 1)-4) are necessary and sufficient for the functional Φ to be a coherent lower prevision on K.

Does the functional $\overline{E}_{\mathbf{P}}$ define the underlying generalized credal set $\mathbf{P}$ uniquely?

Let $X = \{x_1, x_2\}$ and credal sets $\mathbf{P}_1$ and $\mathbf{P}_2$ in M^d_{cpr} are given by $profile(\mathbf{P}_1) = \{P_1\}$, $profile(\mathbf{P}_2) = \{aP_2 + (1-a)P_3 | a \in [0,1]\}$, where $P_1 = (0.5, 0.5)$, $P_2 = (0.5, 0)$, $P_3 = (0, 0.5)$. Then,
$\overline{E}_{\mathbf{P}_1}(f) = 0.5f(x_1) + 0.5f(x_2)$,
$\overline{E}_{\mathbf{P}_2}(f) = 0.5\max\{f(x_1), f(x_2)\} + 0.5\min\{f(x_1), f(x_2)\} = \overline{E}_{\mathbf{P}_1}(f)$.



When functionals $\overline{E}_{\mathbf{P}_1}$ and $\overline{E}_{\mathbf{P}_2}$ coincide.

Let a lower generalized credal set $\mathbf{P}$ is identified with a convex set $\mathbb{R}^n$. The *j-th projection* of $\mathbf{P}$ is

$$\Pr_j \mathbf{P} = \{\Pr_j P | P \in \mathbf{P}\},$$

where $\Pr_j P = (a_1, ..., a_{j-1}, 0, a_{j+1}, ..., a_n)$ if $P = (a_1, ..., a_n)$. Clearly, $\Pr_j \mathbf{P}$ is also a credal set in M^d_{cpr}.

Theorem

Let $\mathbf{P}_1$ and $\mathbf{P}_2$ credal sets in M^d_{cpr}. Then $\overline{E}_{\mathbf{P}_1}(f) = \overline{E}_{\mathbf{P}_2}(f)$ for all $f \in K$ iff $\Pr_j \mathbf{P}_1 = \Pr_j \mathbf{P}_2$, $j = 1, ..., n$.

Maximal generalized credal sets.

Definition

The LG-credal set $\mathbf{P}$ in M^d_{cpr} is called *maximal* if

$$\mathbf{P} = \left\{ P \in M^d_{cpr} | \forall f \in K : \bar{E}_P(f) \leqslant \bar{E}_{\mathbf{P}}(f) \right\}.$$

Theorem

Let $\mathbf{P}$ be a generalized credal set in M^d_{cpr} whose profile is an usual credal set in M_{pr}. Then the credal set $\mathbf{P}$ is maximal.

Remark

The result from the theorem means that in the theory based on generalized credal sets models based on coherent upper previsions and usual credal sets are equal.

The natural extension based on generalized credal sets

Let $\overline{E} : K' \to \mathbb{R}$ be an upper prevision, then it defines the credal set $\mathbf{P} = \left\{P \in M^d_{cpr} | \forall f \in K : \overline{E}_P(f) \leqslant \overline{E}(f)\right\}$ and $\bar{E}_{\mathbf{P}}$ can be considered as the *natural extension* of $\overline{E}$ on K. The functional $\overline{E}$ is called the *generalized coherent upper prevision* if $\overline{E}_{\mathbf{P}}(f) = \bar{E}(f)$, $f \in K'$.

Theorem

Let $\bar{E} : K' \to \mathbb{R}$ be an upper prevision functional. Then its natural extension $\bar{E}' : K \to \mathbb{R}$ based on generalized credal sets is

$$\overline{E}'\left(\underline{f}\right) = \inf\left\{\sum_k a_k \overline{E}\left(\underline{f_k}\right) + a | \sum_k a_k \underline{f_k} + a \geqslant \underline{f}, f_k \in K', a_k, a \geqslant 0\right\},$$

where $\underline{f}$, $\underline{f_k}$ are normalized functions and $\overline{E}_{\mathbf{P}}\left(\underline{f}\right) = \bar{E}_{\mathbf{P}}\left(f\right) - b$, $\overline{E}\left(\underline{f_k}\right) = \overline{E}\left(f_k\right) - b_k$, $b = \min\limits_{x \in X} f(x)$, $b_k = \min\limits_{x \in X} f_k(x)$.

Comparison of reasoning with usual and generalized credal sets

Example

Let we have two sources of information that describe possible diseases $X = \{x_1, x_2, x_3\}$ of a patient.
Source 1: $\overline{E}(1_{\{x_1,x_2\}}) = \overline{E}(1_{\{x_2,x_3\}}) = 0.5$, i.e. probabilities of events $\{x_1, x_2\}$ and $\{x_2, x_3\}$ are lower or equal to 0.5;
Source 2: $\overline{E}(1_{\{x_1,x_3\}}) = 0$, i.e. this is definitely disease x_2.
If we describe sources of information by usual credal sets, then sources 1 and 2, described by probability measures $P_1 = (0.5, 0, 0.5)$ and $P_2 = (0, 1, 0)$ are fully contradictory. If we described them by generalized credal sets $\mathbf{P}_1$ and $\mathbf{P}_2$, then $profile(\mathbf{P_1}) = \{aP_1 + (1-a)P_3 | a \in [0,1]\}$ and $profile(\mathbf{P_2}) = \{P_2\}$, where $P_2 = (0, 0.5, 0)$. Thus, $profile(\mathbf{P}_1 \cap \mathbf{P}_2) = \{P_3\}$ and we come to the conclusion that it is disease x_2 but with $Con(P_3) = 0.5$.

Thanks for you attention

brone@mail.ru
I.Rozenberg.gismps.ru