A unified view of some representation of imprecise probabilities

S. Destercke [1] D. Dubois [2]

[1] Institute of radioprotection and nuclear safety
Cadarache, France

[2] Toulouse institute of computer science
University Paul-Sabatier

SMPS 06

Outline

Family $\mathcal{P}$ of probabilities can be hard to represent (even by lower ($\underline{P}(A)$) and upper ($\overline{P}(A)$) probabilities). Simpler representations exist :

Outline

Random Sets formalism

Definition

- Multi-valued mapping from probability space to space X
- Here, mass function $m : 2^X \to [0,1]$ and $\sum_{E \subseteq X} m(E) = 1$
- A set $E \subseteq X$ is a focal set iff $m(E) > 0$
- Belief measure : $Bel(A) = \sum_{E, E \subseteq A} m(E)$
- Plausibility measure : $Pl(A) = \sum_{E, E \cap A \neq \emptyset} m(E)$

Probability family induced by random sets

$\mathcal{P}_{Bel} = \{P | \forall A \subseteq X \text{ measurable}, Bel(A) \leq P(A) \leq Pl(A)\}$

Outline

Possibility formalism

Definition

- Mapping $\pi : X \rightarrow [0,1]$ and $\exists x \in X$ s.t. $\pi(x) = 1$
- Possibility measure: $\Pi(A) = \sup_{x \in A} \pi(x)$
- Necessity measure: $N(A) = 1 - \Pi(A^c)$

Possibility and random sets

Possibility distribution = random set with nested focal elements

Probability family induced by possibility distribution

$\mathcal{P}_\pi = \{P | \forall A \subseteq X \text{ measurable}, N(A) \leq P(A) \leq \Pi(A)\}$

Outline

Outline

Generalized cumulative distribution

Usual cumulative distribution

Let *Pr* be a probability function on $\mathbb{R}$: the cumulative distribution is $F(x) = \Pr((-\infty, x])$

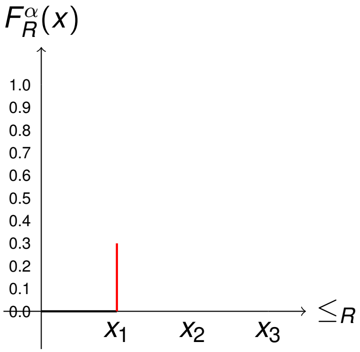
Preliminary definitions

- Let X be a finite domain of n elements and $\alpha = (\alpha_1 \ldots \alpha_n)$ a probability distribution
- R is a relation defining a complete ordering $\leq_R$ on X
- a R-downset $(x]_R$ consist of every element x_i s.t. $x_i \leq_R x$

Definition

Given a relation R, a generalized cumulative distribution is defined as $F_R^\alpha(x) = \Pr((x]_R)$.

Generalized cumulative distribution : illustration



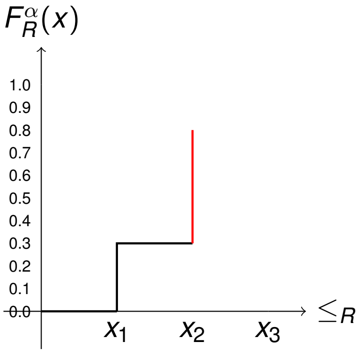
example

- $X = \{x_1, x_2, x_3\}$
- $\alpha = \{0.3, 0.5, 0.2\}$
- $R : x_i < x_j$ iff $i < j$
- $X_R = \{x_1, x_2, x_3\}$

Cumulative prob.

- $F_R^{\alpha}(x_1) = P(x_1) = 0.3$
- $F_R^{\alpha}(x_2) = P(x_1, x_2) = 0.8$
- $F_R^{\alpha}(x_3) = P(x_1, x_2, x_3) = 1$

Generalized cumulative distribution : illustration



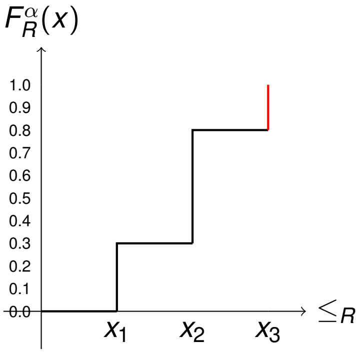
example

- $X = \{x_1, x_2, x_3\}$
- $\alpha = \{0.3, 0.5, 0.2\}$
- $R : x_i < x_j$ iff $i < j$
- $X_R = \{x_1, x_2, x_3\}$

Cumulative prob.

- $F_R^\alpha(x_1) = P(x_1) = 0.3$
- $F_R^\alpha(x_2) = P(x_1, x_2) = 0.8$
- $F_R^\alpha(x_3) = P(x_1, x_2, x_3) = 1$

Generalized cumulative distribution : illustration



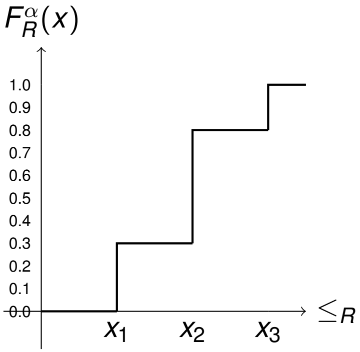
example

- $X = \{x_1, x_2, x_3\}$
- $\alpha = \{0.3, 0.5, 0.2\}$
- $R : x_i < x_j$ iff $i < j$
- $X_R = \{x_1, x_2, x_3\}$

Cumulative prob.

- $F_R^\alpha(x_1) = P(x_1) = 0.3$
- $F_R^\alpha(x_2) = P(x_1, x_2) = 0.8$
- $F_R^\alpha(x_3) = P(x_1, x_2, x_3) = 1$

Generalized cumulative distribution : illustration



example

- $X = \{x_1, x_2, x_3\}$
- $\alpha = \{0.3, 0.5, 0.2\}$
- $R : x_i < x_j$ iff $i < j$
- $X_R = \{x_1, x_2, x_3\}$

Cumulative prob.

- $F_R^\alpha(x_1) = P(x_1) = 0.3$
- $F_R^\alpha(x_2) = P(x_1, x_2) = 0.8$
- $F_R^\alpha(x_3) = P(x_1, x_2, x_3) = 1$

Generalized P-boxes : definition

Usual P-boxes

A P-box is a pair of cumulative distributions $(\underline{F}, \overline{F})$ bounding an imprecisely known distribution F ($\underline{F} \leq F \leq \overline{F}$)

Definition

Given R, a generalized p-box is a pair of gen. cumulative distributions $(F_R^\alpha(x) \leq F_R^\beta(x))$ bounding an imprecisely known distribution $F_R(x)$

Probability family induced by generalized p-box

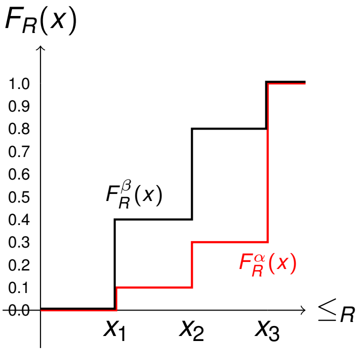
$$\mathcal{P}_{p-box} = \{P | \forall x \in X \text{ measurable}, F_R^\alpha(x) \leq F_R(x) \leq F_R^\beta(x)\}$$

Generalized P-boxes : constraint representation

- Let $A_i = (x_i]_R$ with $x_i \leq_R x_j$ iff $i < j$
- $A_1 \subset A_2 \subset \ldots \subset A_n$
- Gen. P-box can be encoded by following constraints :

$$\alpha_i \leq P(A_i) \leq \beta_i \qquad i = 1, \ldots, n$$
$$\alpha_1 \leq \alpha_2 \leq \ldots \leq \alpha_n \leq 1$$
$$\beta_1 \leq \beta_2 \leq \ldots \leq \beta_n \leq 1$$

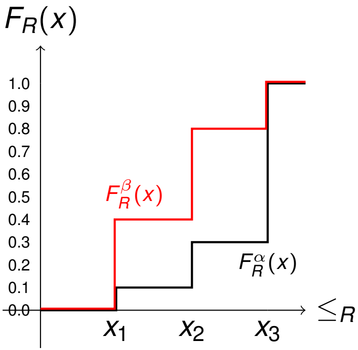
Generalized P-box : illustration



constraints

- $0.1 \leq P(A_1) = P(x_1) \leq 0.4$
- $0.3 \leq P(A_2) = P(x_1, x_2) \leq 0.8$
- $1 \leq P(A_3) = P(x_1, x_2, x_3) \leq 1$

Generalized P-box : illustration



constraints

- $0.1 \leq P(A_1) = P(x_1) \leq 0.4$
- $0.3 \leq P(A_2) = P(x_1, x_2) \leq 0.8$
- $1 \leq P(A_3) = P(x_1, x_2, x_3) \leq 1$

Outline

1. Random Sets
2. Possibility distribution
3. P-Boxes
   - Generalized P-Boxes
   - Relationships between P-Boxes and random sets
   - Relationships between P-Boxes and possibility distribution
4. Clouds

Random sets/P-boxes relation

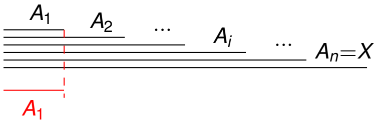
Theorem

Any generalized p-box is a special case of random set (there is a random set such that $\mathcal{P}_{Bel} = \mathcal{P}_{p-box}$)

Sketch of proof

Lower probabilities on every possible event are the same in the two cases

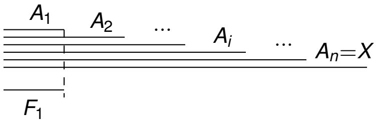
P-Box $\longrightarrow$ random set algorithm



algorithm

1. Build partition of X
2. Order α_i, β_i and rename them γ_l
3. Build focal sets E_i with weights $m(E_l) = \gamma_l - \gamma_{l-1}$

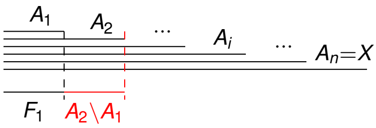
P-Box $\longrightarrow$ random set algorithm



algorithm

1. Build partition of X
2. Order α_i, β_i and rename them γ_l
3. Build focal sets E_l with weights $m(E_l) = \gamma_l - \gamma_{l-1}$

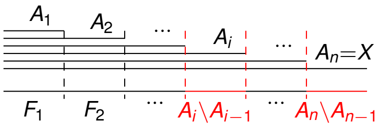
P-Box $\longrightarrow$ random set algorithm



algorithm

1. Build partition of X
2. Order α_i, β_i and rename them γ_l
3. Build focal sets E_i with weights $m(E_l) = \gamma_l - \gamma_{l-1}$

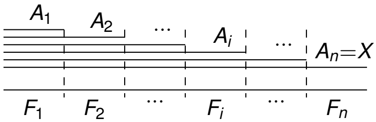
P-Box $\longrightarrow$ random set algorithm



algorithm

1. Build partition of X
2. Order α_i, β_i and rename them γ_l
3. Build focal sets E_l with weights $m(E_l) = \gamma_l - \gamma_{l-1}$

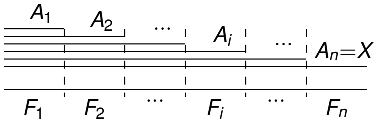
P-Box $\longrightarrow$ random set algorithm



algorithm

1. Build partition of X
2. Order α_i, β_i and rename them γ_l
3. Build focal sets E_i with weights $m(E_l) = \gamma_l - \gamma_{l-1}$

P-Box $\longrightarrow$ random set algorithm

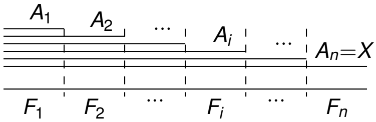


$\alpha_0{=}\beta_0{=}0{\leq}\alpha_1{\leq}\ldots{\leq}\beta_n{\leq}1{=}\beta_{n+1}{=}\alpha_{n+1}$

algorithm

1. Build partition of X
2. Order α_i, β_i and rename them γ_l
3. Build focal sets E_l with weights $m(E_l) = \gamma_l - \gamma_{l-1}$

P-Box $\longrightarrow$ random set algorithm



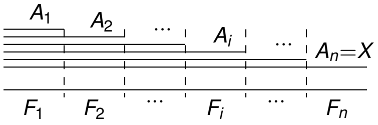
$$\alpha_0 = \beta_0 = 0 \leq \alpha_1 \leq \ldots \leq \beta_n \leq 1 = \beta_{n+1} = \alpha_{n+1}$$

$$\alpha_0 = \gamma_0 = 0 \leq \gamma_1 \leq \ldots \leq \gamma_{2n} \leq 1 = \gamma_{2n+1} = \beta_{n+1}$$

algorithm

1. Build partition of X
2. Order α_i, β_i and rename them γ_l
3. Build focal sets E_i with weights $m(E_l) = \gamma_l - \gamma_{l-1}$

P-Box $\rightarrow$ random set algorithm



$\alpha_0{=}\beta_0{=}0{\leq}\alpha_1{\leq}\ldots{\leq}\beta_n{\leq}1{=}\beta_{n+1}{=}\alpha_{n+1}$

$\alpha_0{=}\gamma_0{=}0{\leq}\gamma_1{\leq}\ldots{\leq}\gamma_{2n}{\leq}1{=}\gamma_{2n+1}{=}\beta_{n+1}$

$m(E_l) = \gamma_l - \gamma_{l-1}$

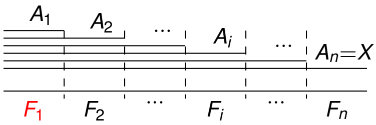
with $E_l = E_{l-1} \cup F_{i+1}$ if $\gamma_{l-1} = \alpha_i$

with $E_l = E_{l-1} \setminus F_i$ if $\gamma_{l-1} = \beta_i$

algorithm

1. Build partition of X
2. Order α_i, β_i and rename them γ_l
3. Build focal sets E_l with weights $m(E_l) = \gamma_l - \gamma_{l-1}$

P-Box $\rightarrow$ random set algorithm



$\alpha_0 = \beta_0 = 0 \leq \alpha_1 \leq \ldots \leq \beta_n \leq 1 = \beta_{n+1} = \alpha_{n+1}$

$\alpha_0 = \gamma_0 = 0 \leq \gamma_1 \leq \ldots \leq \gamma_{2n} \leq 1 = \gamma_{2n+1} = \beta_{n+1}$

$$m(E_l) = \gamma_l - \gamma_{l-1}$$

with $E_l = E_{l-1} \cup F_{i+1}$ if $\gamma_{l-1} = \alpha_i$

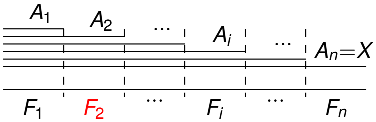
with $E_l = E_{l-1} \setminus F_i$ if $\gamma_{l-1} = \beta_i$

$$m(E_1) = \gamma_1 - \gamma_0 = \alpha_1 - \alpha_0 = m(F_1)$$

algorithm

1. Build partition of X
2. Order α_i, β_i and rename them γ_l
3. Build focal sets E_i with weights $m(E_l) = \gamma_l - \gamma_{l-1}$

P-Box $\longrightarrow$ random set algorithm



$$\alpha_0 = \beta_0 = 0 \leq \alpha_1 \leq \ldots \leq \beta_n \leq 1 = \beta_{n+1} = \alpha_{n+1}$$

$$\alpha_0 = \gamma_0 = 0 \leq \gamma_1 \leq \ldots \leq \gamma_{2n} \leq 1 = \gamma_{2n+1} = \beta_{n+1}$$

$$m(E_l) = \gamma_l - \gamma_{l-1}$$

with $E_l = E_{l-1} \cup F_{i+1}$ if $\gamma_{l-1} = \alpha_i$

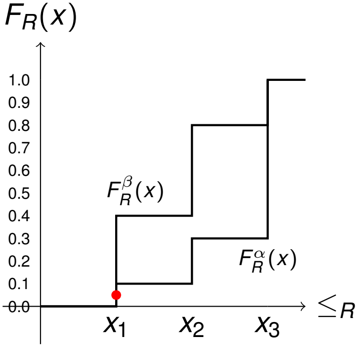
with $E_l = E_{l-1} \setminus F_i$ if $\gamma_{l-1} = \beta_i$

$$m(E_2) = \gamma_2 - \gamma_1 = \gamma_2 - \alpha_1 = m(F_1 \cup F_2)$$

algorithm

1. Build partition of X
2. Order α_i, β_i and rename them γ_l
3. Build focal sets E_i with weights $m(E_l) = \gamma_l - \gamma_{l-1}$

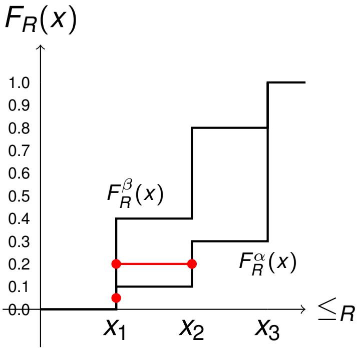
graphical representation



Random Set

- $m(E_1) = m(\{x_1\}) = 0.1$
- $m(E_2) = m(\{x_1, x_2\}) = 0.2$
- $m(E_3) = m(\{x_1, x_2, x_3\}) = 0.1$
- $m(E_4) = m(\{x_2, x_3\}) = 0.4$
- $m(E_5) = m(\{x_3\}) = 0.2$

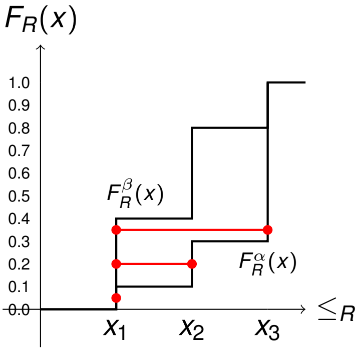
graphical representation



Random Set

- $m(E_1) = m(\{x_1\}) = 0.1$
- $m(E_2) = m(\{x_1, x_2\}) = 0.2$
- $m(E_3) = m(\{x_1, x_2, x_3\}) = 0.1$
- $m(E_4) = m(\{x_2, x_3\}) = 0.4$
- $m(E_5) = m(\{x_3\}) = 0.2$

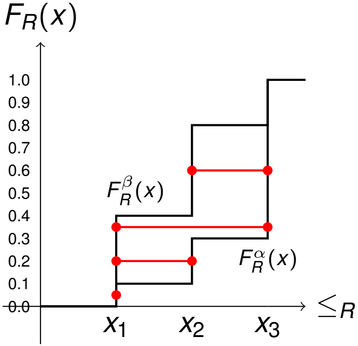
graphical representation



Random Set

- $m(E_1) = m(\{x_1\}) = 0.1$
- $m(E_2) = m(\{x_1, x_2\}) = 0.2$
- $m(E_3) = m(\{x_1, x_2, x_3\}) = 0.1$
- $m(E_4) = m(\{x_2, x_3\}) = 0.4$
- $m(E_5) = m(\{x_3\}) = 0.2$

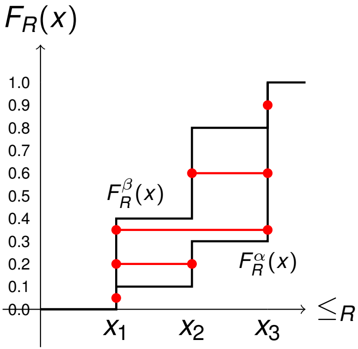
graphical representation



Random Set

- $m(E_1) = m(\{x_1\}) = 0.1$
- $m(E_2) = m(\{x_1, x_2\}) = 0.2$
- $m(E_3) = m(\{x_1, x_2, x_3\}) = 0.1$
- $m(E_4) = m(\{x_2, x_3\}) = 0.4$
- $m(E_5) = m(\{x_3\}) = 0.2$

graphical representation



Random Set

- $m(E_1) = m(\{x_1\}) = 0.1$
- $m(E_2) = m(\{x_1, x_2\}) = 0.2$
- $m(E_3) = m(\{x_1, x_2, x_3\}) = 0.1$
- $m(E_4) = m(\{x_2, x_3\}) = 0.4$
- $m(E_5) = m(\{x_3\}) = 0.2$

Outline

Generalized cumulative distribution

An upper generalized cumulative distribution $F_R(x)$ can be viewed as a possibility distribution π_R, since $\max_{x \in A} F_R(x) \geq \Pr(A)$

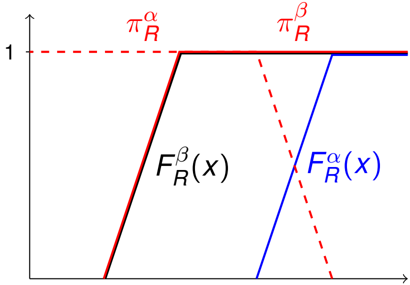
Generalized P-box

- Two cumulative distributions $F_R^{\beta}(x) \geq F_R^{\alpha}(x)$
- Upper bound $F_R^{\beta}(x)$ can be viewed as a possibility distribution $\rightarrow F_R^{\beta}(x) = \pi_R^{\beta}$
- Lower bound $F_R^{\alpha}(x)$ can be viewed as a possibility distribution $\rightarrow F_R^{\alpha}(x) = 1 - \pi_R^{\alpha}$

Probability families equivalence

We have that $\mathcal{P}_{p-box} = \mathcal{P}_{\pi_R^{\alpha}} \cap \mathcal{P}_{\pi_R^{\beta}}$

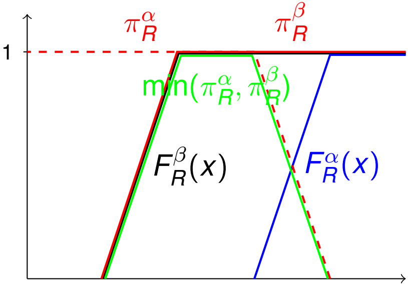
Illustration



Relations

$$(\mathcal{P}_{p-box} = \mathcal{P}_{\pi_R^\alpha} \cap \mathcal{P}_{\pi_R^\beta})$$

Illustration



Relations

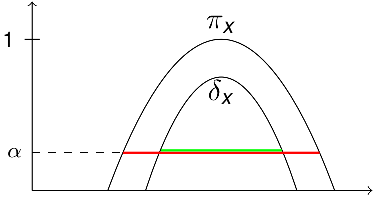
$$(\mathcal{P}_{p-box} = \mathcal{P}_{\pi_R^\alpha} \cap \mathcal{P}_{\pi_R^\beta}) \supset (\mathcal{P}_{\min(\pi_R^\alpha, \pi_R^\beta)})$$

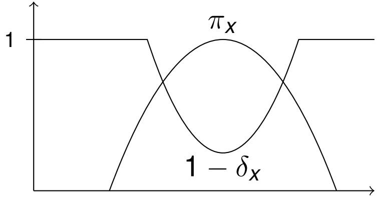
Outline

Clouds : Introduction

Definition

- A cloud Can be viewed as a pair of distributions $[\delta(x), \pi(x)]$
- A r.v. $X \in$ cloud iff $P(\delta(x) \geq \alpha) \leq 1 - \alpha \leq P(\pi(x) > \alpha)$
-
-
-

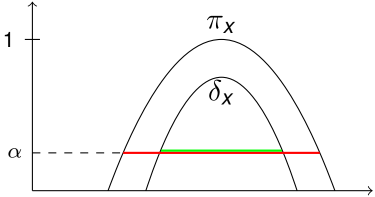


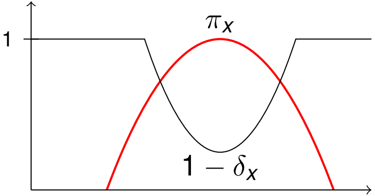


Clouds : Introduction

Definition

- A cloud Can be viewed as a pair of distributions $[\delta(x), \pi(x)]$
- A r.v. $X \in$ cloud iff $P(\delta(x) \geq \alpha) \leq 1 - \alpha \leq P(\pi(x) > \alpha)$
- π is a possibility distribution
-
-

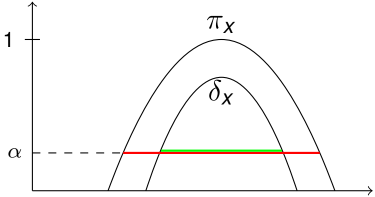


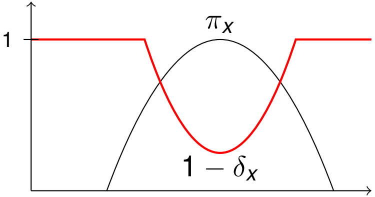


Clouds : Introduction

Definition

- A cloud Can be viewed as a pair of distributions $[\delta(x), \pi(x)]$
- A r.v. $X \in$ cloud iff $P(\delta(x) \geq \alpha) \leq 1 - \alpha \leq P(\pi(x) > \alpha)$
- π is a possibility distribution
- $1 - \delta$ is a possibility distribution
-

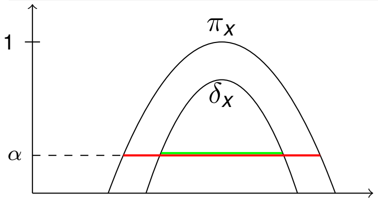


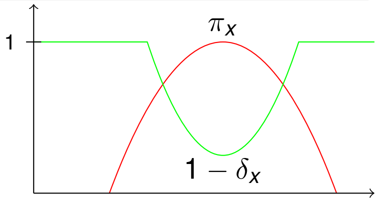


Clouds : Introduction

Definition

- A cloud Can be viewed as a pair of distributions $[\delta(x), \pi(x)]$
- A r.v. $X \in$ cloud iff $P(\delta(x) \geq \alpha) \leq 1 - \alpha \leq P(\pi(x) > \alpha)$
- π is a possibility distribution
- $1 - \delta$ is a possibility distribution
- We have that $\mathcal{P}_{cloud} = \mathcal{P}_{\pi} \cap \mathcal{P}_{1-\delta}$





Discrete clouds : formalism

Discrete clouds as collection of sets

Discrete clouds can be viewed as two set collections

- $A_1 \subseteq A_2 \subseteq \ldots \subseteq A_n \quad (\pi_x)$
- $B_1 \subseteq B_2 \subseteq \ldots \subseteq B_n \quad (\delta_x)$
- $B_i \subseteq A_i \quad (\delta_x \leq \pi_x)$

with constraints

- $P(B_i) \leq 1 - \alpha_{i+1} \leq P(A_i)$
- $1 = \alpha_1 > \alpha_2 > \ldots > \alpha_n = 0$

Relationship between clouds and generalized p-boxes

Theorem

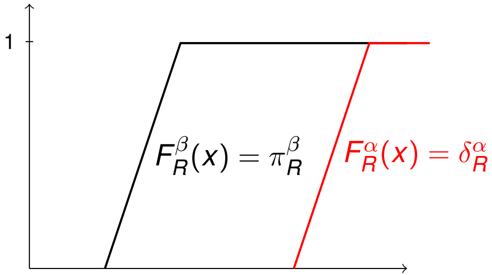
A generalized p-box is a particular case of cloud

Proof.

- $F_R^\beta(x) > F_R^\alpha(x)$
- $F_R^\beta(x) \rightarrow$ possibility distribution π_R^β
- $F_R^\alpha(x) \rightarrow$ possibility distribution δ_R^α
- Gen. P-box equivalent to the cloud $[\delta_R^\alpha, \pi_R^\beta]$

$\square$

Illustration



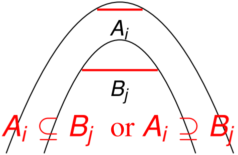
Relationship between clouds and generalized p-boxes

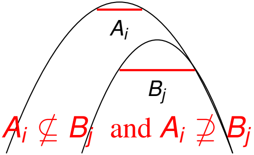
Theorem

A cloud is a gen. p-box iff the sets $\{A_i, B_i\}$ form a complete order with respect to inclusion ($\forall i, j$ $A_i \subseteq B_j$ or $A_i \supseteq B_j$)

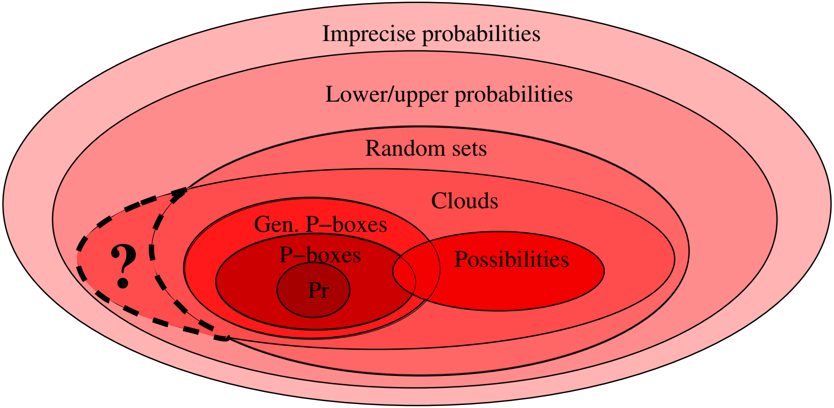
Corollary

A cloud $[\pi_1, \pi_2]$ is a generalized p-box iff π_1, π_2 are comonotonic





Graphical summary



Summary

- A gen. P-box is a special case of random set and can be represented by two possibility distributions
- Comonotonic clouds are equivalent to a gen. P-box.

- Open questions, perspectives
  - Test clouds as descriptive formalism (How to elicit them ?) and as practical representation.
  - Extending results to continuous framework and to lower/upper previsions.