Transforming probability intervals into other uncertainty models

S. Destercke [1] D. Dubois [2] and E. Chojnacki [1]

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

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

EUSFLAT 2007

Introduction

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

- Probability intervals
- Random sets
- (Generalized) P-boxes
- Possibility distributions
- Neumaier's Clouds

Comparing them in terms of their relative expressive power.

Probability intervals

Definition

Given space $X = \{x_1, \ldots, x_n\}$, probability intervals = imprecise assignments $[l_i, u_i]$ over elementary elements x_i. A collection of intervals $L = \{[l_i, u_i], i = 1, \ldots, n\}$ induces the family

$$\mathcal{P}_L = \{P | l_i \leq p(x_i) \leq u_i \ \forall x_i \in X\}$$

Properties (De Campos et al.)

- $\mathcal{P}_L$ is **non-empty** if $\sum_{i=1}^{n} l_i \leq 1 \leq \sum_{i=1}^{n} u_i$
- Bounds of $\mathcal{P}_L$ are **reachable** if $\sum_{j \neq i} l_j + u_i \leq 1$ and $\sum_{j \neq i} u_j + l_i \geq 1 \forall\ i$
- Lower/upper probabilities on events are given by $\underline{P}(A) = \max(\sum_{x_i \in A} l_i, 1 - \sum_{x_i \notin A} u_i) \quad ; \overline{P}(A) = \min(\sum_{x_i \in A} u_i, 1 - \sum_{x_i \notin A} l_i)$
- $\underline{P}$ is a 2-monotone Choquet capacity

Probability intervals vs. Random Sets

Random sets cannot capture probability families induced by probability intervals : only approximations are possible.

Existing results

- **Inner approximation:** Lemmer and Kyburg explore the problem of finding a random set *Bel* s.t. $Bel(x_i) = l_i, Pl(x_i) = u_i$ (bounds coincide on singletons and $\mathcal{P}_{Bel} \subset \mathcal{P}_L$). They show that it is possible if L is reachable, non-empty and if
$$\sum_{i=1}^{n} l_i + \sum_{i=1}^{n} u_i \geq 2$$
- **Outer approximation:** given L, Denoeux extensively explores the problem of finding the most "precise" random set *Bel* s.t. $\mathcal{P}_L \subset \mathcal{P}_{Bel}$.

Probability boxes and generalized p-boxes

P-boxes

A Cumulative distribution is a monotone function F from the reals to $[0,1]$, with $F(+\infty) = 1$, $F(-\infty) = 0$. In general, it is of the form $F(x) = Pr((-\infty, x])$ for a probability measure Pr.
A P-box is a pair of cumulative distributions $\underline{F}, \overline{F}$ with $\underline{F}(x) \leq \overline{F}(x)$ defining the family of probability functions with cumulative distributions F such that $\underline{F} \leq F \leq \overline{F}$

Generalized P-box

- A generalized cumulative distribution is a monotone function F^R from a weakly ordered space $(X, \leq_R)$ to $[0,1]$ with $F^R(\overline{x}) = 1$ ($\overline{x} =$ top of X). In general, it is of the form $F^R(x) = P(\{x_i \in X | x_i \leq_R x\})$.
- A generalized P-box is a pair of $\leq_R$-comonotone functions $\underline{F}_R, \overline{F}_R$ from X to $[0,1]$, with $\underline{F}_R(x) \leq \overline{F}_R(x)$ and $\exists\ \overline{x}$ s.t. $\overline{F}_R(\overline{x}) = 1$, $\exists\ \underline{x}$ s.t. $\underline{F}_R(\underline{x}) = 0$.

Associated probability family: $\mathcal{P}_{p-box} = \{P | \underline{F}_R(x) \leq F^R(x) \leq \overline{F}_R(x)\}$.

Generalized P-boxes and confidence sets

(Generalized) p-boxes can be viewed as upper and lower uncertainty bounds on nested confidence sets induced by the weak order R (e.g. I. Kozine, L. Utkin (I.J. of Gen. Syst., 2005)).

Finite case

- Let $A_i = \{x \in X | x \leq_R x_i\}$ 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 :

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

From Prob. Intervals to Gen. p-boxes and back

Prob. intervals $\rightarrow$ Gen. p-boxes

Given ordering R on X s.t. $x_i \leq_R x_j$ iff $i < j$ and intervals L, build the Gen. p-box

$$\underline{F}_R(x_i) = \underline{P}(A_i) = \max(\textstyle\sum_{x_i \in A_i} l_j, 1 - \sum_{x_i \notin A_i} u_j)$$
$$\overline{F}_R(x_i) = \overline{P}(A_i) = \min(\textstyle\sum_{x_i \in A_i} u_i, 1 - \sum_{x_i \notin A_i} l_i)$$

Gen. p-boxes $\rightarrow$ Prob. Intervals

If we have a gen. p-box $[\underline{F}, \overline{F}]$ defined on nested sets A_i, corresponding probability intervals are given by

$$\underline{P}(x_i) = l_i = \max(0, \underline{P}(A_i) - \overline{P}(A_{i-1}))$$
$$\overline{P}(x_i) = u_i = \overline{P}(A_i) - \underline{P}(A_{i-1})$$

Prob. Intervals and Gen. p-boxes : relations

Theorem

A probability family induced by a generalized p-box is representable by a random set

Given an initial set L of probability intervals, or an initial p-box $[\underline{F}_R, \overline{F}_R]$ over a space X, we can consider the respective transformations

$$\text{Prob. Intervals } L \rightarrow \text{ p-box } [\underline{F}'_R(x), \overline{F}'_R(x)] \rightarrow \text{ Prob. Intervals } L''$$

$$\text{p-box } [\underline{F}_R(x), \overline{F}_R(x)] \rightarrow \text{ Prob. Intervals } L' \rightarrow \text{ p-box } [\underline{F}''_R(x), \overline{F}''_R(x)]$$

we have that $\mathcal{P}_L \subseteq \mathcal{P}_{L''}$ and $\mathcal{P}_{[\underline{F}_R, \overline{F}_R]} \subseteq \mathcal{P}_{[\underline{F}''_R, \overline{F}''_R]}$.

$\Rightarrow$ Some information is lost during the transformation, due to the fact that constraints are defined on different events (i.e. singletons and nested sets)

Possibility formalism

Definition

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

Possibility and random sets

Possibility distribution = random set with nested realizations

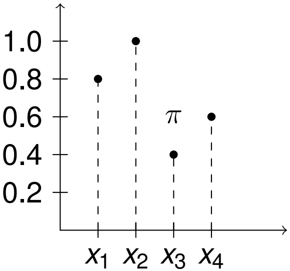
Probability family associated to possibility distribution

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

Generalized cumulative distribution $=$ possibility distribution

An **upper** cumulative distribution $\overline{F}$ bounding a probability family is such that $\max_{x \in A} \overline{F}(x) \geq \Pr(A)$ (maxitivity), and can thus be interpreted as a possibility distribution π

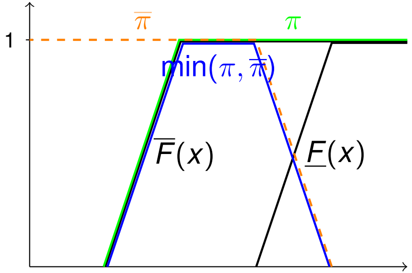
Conversely, up to a re-ordering, any possibility distribution π can be assimilated to an upper (generalized) cumulative distribution $\overline{F}$.



P-boxes as pairs of possibility distributions

If $\underline{F}(x)$ is a lower generalized cumulative distribution, we have $\min_{x \in A^c} F_*(x) \leq \Pr(A) \rightarrow \max_{x \in A^c}(1 - \underline{F}(x)) \geq \Pr(A^c)$.

- Take $\pi = \overline{F}(x)$, $\overline{\pi} = 1 - \underline{F}(x)$, we have $\mathcal{P}_{p-box} = \mathcal{P}_{\pi} \cap \mathcal{P}_{\overline{\pi}}$



$$(\mathcal{P}_{p-box} = \mathcal{P}_{\pi} \cap \mathcal{P}_{\overline{\pi}}) \supset (\mathcal{P}_{\min(\pi,\overline{\pi})})$$

From Prob. Intervals to possibility distributions

Precise probability distribution

Given a precise probability p on X, Dubois and Prade (1982) proposed to consider the complete pre-order

$$p(x_1) < p(x_2) < \ldots < p(x_j) < \ldots < p(x_n)$$

and the transformation into the possibility distribution π given by

$$\pi(x_i) = \sum_{j=1}^{i} p(x_j)$$

Imprecise probability intervals

When only probability intervals L are available, the order

$$p(x_i) \leq p(x_j) \leftrightarrow u_i \leq l_j$$

is (generally) no longer complete ($p(x_i), p(x_j)$ are incomparable if $[l_i, u_i], [l_j, u_j]$ intersect) $\Rightarrow$ need to define a method to build π s.t. $\mathcal{P}_L \subset \mathcal{P}_\pi$

From Prob. Intervals to possibility distributions

Masson and Denoeux solution

Let C_l be a complete order refining the partial order on intervals, and $\mathcal{C}$ the set of possible refinement. Masson and Denoeux propose to use the transformation used in the precise case on orders C_l and to take the possibility distribution covering all these transformations

1. For each order $C_l \in \mathcal{C}$ and each element x_i, solve

$$\pi(x_i)^{C_l} = \max_{p(x_1),\dots,p(x_n)} \sum_{\sigma_l^{-1}(j) \leq \sigma_l^{-1}(i)} p(x_j)$$

   under the constraints

$$\begin{cases} \sum_{k=1,\dots,n} p(x_k) = 1 & (p \textbf{ is a prob. distribution}) \\ l_k \leq p_k \leq u_k & (p \textbf{ is in } \mathcal{P}_L) \\ p(x_{\sigma_l(1)}) \leq p(x_{\sigma_l(2)}) \leq \dots \leq p(x_{\sigma_l(n)}) & (\textbf{order induced by refinement } C_l) \end{cases}$$

   where σ_l is the permutation of $p(x_k)$ associated with the linear extension C_l

2. Take the distribution dominating all distributions $\pi(x_i)^{C_l}$ s.t.

$$\pi(x_i) = \max_{C_l \in \mathcal{C}} \pi(x_i)^{C_l} \quad \forall i \quad (\pi(x_i) \textbf{ covers all } \pi(x_i)^{C_l})$$

From Prob. Intervals to possibility distributions

Solution using relation with p-boxes

1. Choose a particular order R on X
2. Build The upper generalized cumulative distribution $\overline{F}_R$ induced by L:

$$\overline{F}_R(x_i) = \overline{P}(A_i) = \min(\sum_{x_i \in A_i} u_i, 1 - \sum_{x_i \notin A_i} l_i)$$

3. Let $\pi(x_i) = \overline{F}(x_i)$.

- Both methods provide a guaranteed outer approximation (i.e. $\mathcal{P}_L \subset \mathcal{P}_\pi$)
- In some cases, using generalized upper cumulative distribution give a dist. π more precise than the one of Masson/Denoeux method.
- Needs a rationale to select the proper weak order R.

Neumaier's clouds : Introduction

Definition

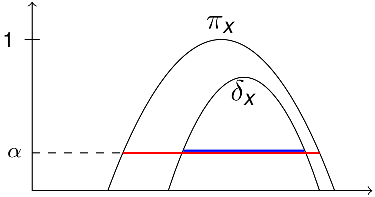
A cloud can be viewed as a pair of distributions $[\delta(x) \leq \pi(x)]$ from X to $[0,1]$ ($\equiv$ to an interval-valued fuzzy set)

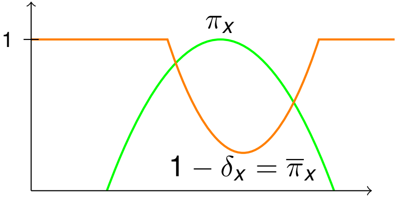
Associated probability family

- $\mathcal{P}_{cloud} = \{P | P(\{x|\delta(x) \geq \alpha\}) \leq 1 - \alpha \leq P(\{x|\pi(x) > \alpha\})$

Link with possibility distributions

- If we consider the possibility distributions $1 - \delta = \overline{\pi}$ and π, we have $\mathcal{P}_{cloud} = \mathcal{P}_{\pi} \cap \mathcal{P}_{1-\delta=\overline{\pi}}$ (Dubois & Prade 2005)





Discrete clouds : formalism

Discrete clouds as collection of sets

Discrete clouds can be viewed as two collections of confidence sets

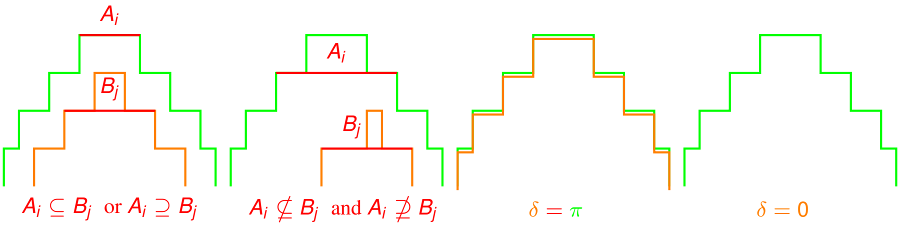
- $\emptyset = A_0 \subset A_1 \subseteq A_2 \subseteq \ldots \subseteq A_n \subset A_{n+1} = X \quad (\pi_x)$
- $\emptyset = B_0 \subset B_1 \subseteq B_2 \subseteq \ldots \subseteq B_n \subset B_{n+1} = X \quad (\delta_x)$
- $B_i \subseteq A_i \quad (\delta_x \leq \pi_x)$

with constraints

- $P(B_i) \leq 1 - \alpha_i \leq P(A_i)$
- $1 = \alpha_0 > \alpha_1 > \alpha_2 > \ldots > \alpha_n > \alpha_{n+1} = 0$

Characterizing clouds

- A cloud is said **comonotonic** if distributions δ, π are comonotone
- A cloud is said **non-comonotonic** if distributions δ, π are **not** comonotone
- A cloud is said **thin** if distributions δ, π are s.t. $\delta = \pi$
- A cloud is said **fuzzy** if distributions δ, π are s.t. $\delta = 0$ (a **fuzzy** cloud is a possibility distribution).



Comonotonic cloud | Non-comonotonic cloud | Thin cloud | Fuzzy cloud

Relating clouds to other uncertainty representations

- **Comonotonic** clouds and generalized p-boxes are equivalent representations (and both are special cases of random sets).
- **Thin** clouds define empty probability family on finite sets, infinite on infinite sets.
- On finite sets some clouds contain a single probability distribution
- **Non-comonotonic** clouds are not even 2-monotone capacities.

Transforming precise probability into cloud

Let p be a precise probability with $p_1 < p_2 < \dots p_j < \dots < p_n$, and π the poss. dist. built by Dubois and Prade method: $\pi_i = \sum_{j=1}^{i} p_j$. If we now reverse the order, we can build another distribution

$$\overline{\pi}_i = \sum_{j=i}^{n} p_j.$$

Let $\delta = 1 - \overline{\pi}$ then, (δ, π) is *the (almost **thin**) cloud containing p and only p*.

From Prob. Intervals to Clouds

Extending Masson and Denoeux solution

Again, we consider the set of complete order C_l refining the partial order on intervals L. We build distribution π with Masson and Denoeux method, and δ in the following way

1. For each order $C_l \in \mathcal{C}$ and each element x_i, solve

$$\overline{\pi}(x_i)^{C_l} = \max_{p_1,\ldots,p_n} \sum_{\sigma_l^{-1}(i) \leq \sigma_l^{-1}(j)} p_j$$
$$= 1 - \min_{p_1,\ldots,p_n} \sum_{\sigma_l^{-1}(j) < \sigma_l^{-1}(i)} p_j = 1 - \delta(x_i)^{C_l}$$

   with the same constraints as for π
2. Take the distribution dominating all distributions $\pi_{\delta_i}^{C_l}$

$$\overline{\pi}(x_i) = 1 - \delta(x_i) = \max_{C_l \in \mathcal{C}} \overline{\pi}(x_i)^{C_l} = 1 - \min_{C_l \in \mathcal{C}} \delta(x_i)^{C_l} \quad \forall i \tag{1}$$

The resulting cloud (δ, π) is s.t. $\mathcal{P}_L \subset \mathcal{P}_{\delta,\pi}$, and is more precise than π alone.

From Prob. Intervals to Clouds

Solution using relation with p-boxes

1. Choose a particular order R on X
2. Build The generalized p-box $[\underline{F}_R, \overline{F}_R]$ induced by L corresponding to this order
3. Take $\pi(x_i) = \overline{F}_R(x_i)$ and $\delta(x_i) = \underline{F}_R(x_{i-1})$ ($\delta(x_0) = 0$). The resulting cloud (δ, π) is s.t. $\mathcal{P}_L \subset \mathcal{P}_{\delta,\pi}$, and is more precise than taking $\pi = \overline{F}$ alone.

Comparison

- Both methods provide a guaranteed outer approximation (i.e. $\mathcal{P}_L \subset \mathcal{P}_{\delta,\pi}$)
- Again, using the method based on generalized p-boxes may lead to more precise results than the extension of Masson and Denoeux method.
- Needs some rationale to choose the ordering R.

Conclusions and perspective

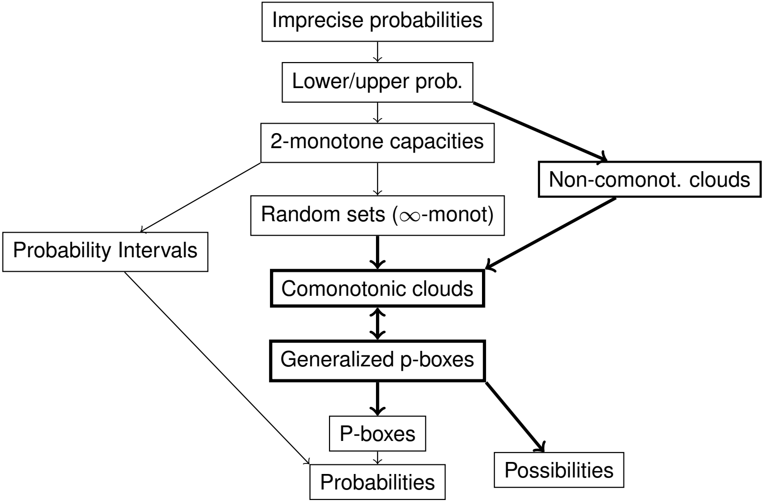
Conclusions

- We have proposed methods to transform probability intervals into p-boxes, possibility distributions or clouds that are guaranteed outer approximations
- Since probability intervals can be seen as probabilistic constraints on singletons, and the other models as constraints on nested sets, transforming one into the other either adds (inner approximation) or loses (outer approximation) information.

Perspectives

- Find or show that already proposed transformations lose (or add) a minimal amount of information (given a specific measure of information).
- Extend these results to the continuous case (probability density encompassed between a lower and upper density).

Imprecise probability representations: where is what?



General perspectives and open questions

- Study the propagation, fusion, conditioning of practical uncertainty representations (are they easy to compute, do they preserve the representation ?).
- Extend results to characterize lower/upper previsions of clouds and generalized p-boxes (in progress...)
- Explore the possible use of Gen. p-boxes, clouds (i.e. pairs of possibility distributions), probability intervals in the elicitation and linguistic assessments of imprecise probabilities.