Transforming probability intervals into other uncertainty models

Sébastien Destercke, Didier Dubois, Eric Chojnacki

▶ **To cite this version:**

HAL Id: irsn-00196700

https://irsn.hal.science/irsn-00196700

Submitted on 13 Dec 2007

Transforming probability intervals into other uncertainty models

Sébastien Destercke
Institut de radioprotection
et de sûreté nucléaire (IRSN)
DPAM, SEMIC, LIMSI
CE Cadarache
13115 Saint Paul Lez Durance
sebastien.destercke@irsn.fr
desterck@irit.fr

Didier Dubois
IRIT
Université Paul Sabatier
118 route de Narbonne
31062 Toulouse Cedex 9, France
dubois@irit.fr

Eric Chojnacki
Institut de radioprotection
et de sûreté nucléaire (IRSN)
DPAM, SEMIC, LIMSI
CE Cadarache
13115 Saint Paul Lez Durance
eric.chojnacki@irsn.fr

Abstract

Probability intervals are imprecise probability assignments over elementary events. They constitute a very convenient tool to model uncertain information : two common cases are confidence intervals on parameters of multinomial distributions built from sample data and expert opinions provided in terms of such intervals. In this paper, we study how probability intervals can be transformed into other uncertainty models such as possibility distributions, Ferson's p-boxes, random sets and Neumaier's clouds.

Keywords: Probability intervals, random sets, possibility, p-boxes, clouds.

1 Introduction

When modeling uncertainty, the theory of imprecise probabilities [17] and the so-called lower previsions formally subsume most of the existing uncertainty theories. Provided one accepts its behavioral interpretation, this theory offers a very appealing unifying and highly expressive framework. Unfortunately, computational intractability is often the price to pay for such generality and expressiveness. It is thus important to study simpler models that will often be sufficient in practice to solve a given problem tainted with uncertainty. Even when a lot of information is available (but not enough to give precise probabilities), one may want to work with approximated models (e.g. for mathematical or computational tractability considerations).

Probability intervals [1] are among these simpler models. They are easy to understand and computationally tractable. They can be, for example, confidence intervals coming from sample data or opinions given by an expert. Nevertheless, it may happen that one wishes to map the information given by probability intervals into another model, because mathematical tools proper to the latter model must be used or simply to present information in a specific way. This mapping is the object of this paper.

In section 2, we briefly recall the various formalism concerned by this paper. Section 3 then reviews some existing results relating probability intervals and random sets. The next section deals with the relations between probability intervals and p-boxes. Finally, section 5 and 6 are respectively devoted to the transformation of probability intervals into, respectively, possibility distributions and clouds (a recent model proposed by Neumaier [15]).

2 Notations and preliminaries

In the paper, we will restrict ourselves to probability families denoted $\mathcal{P}$ and defined on a finite arbitrary space X of n elements $\{x_1, \ldots, x_n\}$. We will now briefly introduce the models studied in the sequel.

2.1 Lower/upper probabilities

Lower ($\underline{P}(A)$) and upper probabilities ($\overline{P}(A)$) on events are respectively defined s.t. $\underline{P}(A) = \inf_{P \in \mathcal{P}} P(A)$ and $\overline{P}(A) = \sup_{P \in \mathcal{P}} P(A)$. We have $\mathcal{P}_{\underline{P}, \overline{P}}(A) = \{P | \forall A \subseteq X \text{ measurable}, \underline{P}(A) \leq P(A) \leq \overline{P}(A)\}$. Although this model is already a restriction from more general ones (lower/upper probabilities can be seen as projections of a family $\mathcal{P}$ on the subspace of events, and in general we have $\mathcal{P} \subseteq \mathcal{P}_{\underline{P}, \overline{P}}$), it is still fairly general and subsumes all other models studied here.

2.2 Probability intervals

Probability intervals and their properties are extensively studied in [1]. Probability intervals are defined as lower and upper bounds of probabilities restricted to singletons x_i. They can be seen as a family of intervals $L = \{[l_i, u_i], i = 1, \ldots, n\}$ defining the family $\mathcal{P}_L = \{P | l_i \leq p(x_i) \leq u_i \ \forall x_i \in X\}$.

In this paper, we will focus on reachable sets L which define non-empty families $\mathcal{P}_L$, since other ones have little interest. A set L will be called reachable if, for each x_i, we can find a probability distribution $P \in \mathcal{P}_L$ s.t. $P(x_i) = l_i$ and another one for which $P(x_i) = u_i$ (in other words, each bound can be reached by at least one distribution in the family). Non-emptiness and reachability respectively correspond to the two conditions

$$\sum_{i=1}^n l_i \leq 1 \leq \sum_{i=1}^n u_i$$

$$\sum_{j \neq i} l_j + u_i \leq 1 \text{ and } \sum_{j \neq i} u_j + l_i \geq 1 \quad \forall i$$

Given intervals L_i, general lower and upper probabilities can be computed through the simple formulas

$$\underline{P}(A) = \max(\sum_{x_i \in A} l_i, 1 - \sum_{x_i \notin A} u_i)$$
$$\overline{P}(A) = \min(\sum_{x_i \in A} u_i, 1 - \sum_{x_i \notin A} l_i)$$

and this lower (upper) probability is an order 2 monotone (alternate) Choquet capacity [2].

2.3 Random sets

Formally, a random set is a set-valued mapping from a (here finite) probability space to a set X. It induces lower and upper probabilities on X [5]. Here, we use mass functions [16] to represent random sets. A mass function m is defined by a mapping from the power set 2^X to the unit interval, s.t. $\sum_{E \subseteq X} m(E) = 1$ and $m(\emptyset) = 0$. A set E with positive mass is called a focal set. Two measures, a plausibility and a belief measure can be defined from this mass function:

Belief measure: $Bel(A) = \sum_{E, E \subseteq A} m(E)$

Plausibility measure: $Pl(A) = 1 - Bel(A^c)$

The set $\mathcal{P}_{Bel} = \{P | \forall A \subseteq X \text{ measurable}, Bel(A) \leq P(A) \leq Pl(A)\}$ is the probability family induced by the belief function.

2.4 (Generalized) P-boxes

A p-box is usually defined on the real line by a pair of cumulative distributions $[\underline{F}, \overline{F}]$, defining the probability family $\mathcal{P}_{[\underline{F}, \overline{F}]} = \{P | \underline{F}(x) \leq F(x) \leq \overline{F}(x) \quad \forall x \in \Re\}$. The notion of cumulative distribution on the real line is based on a natural ordering of numbers. In order to generalize this notion to arbitrary finite

sets, we need to define a weak order relation $\leq_R$ on this space. Given $\leq_R$, an R-downset is of the form $\{x_i : x_i \leq_R x\}$, and denoted $(x]_R$. A generalized R-cumulative distribution [7] is defined as the function $F_R : X \to [0,1]$ s.t. $F_R(x) = \Pr((x]_R)$, where $\Pr$ is a probability measure on X. We can now define a generalized p-box as a pair $[\underline{F}_R(x), \overline{F}_R(x)]$ of generalized cumulative distributions defining a probability family $\mathcal{P}_{[\underline{F}_R(x), \overline{F}_R(x)]} = \{P | \forall x, \underline{F}_R(x) \leq F_R(x) \leq \overline{F}_R(x)\}$. Generalized P-boxes can also be represented by a set of constraints

$$\alpha_i \leq P(A_i) \leq \beta_i \qquad i = 1, \ldots, n \qquad (1)$$

where $0 \leq \alpha_1 \leq \alpha_2 \leq \ldots \leq \alpha_n \leq 1$, $0 \leq \beta_1 \leq \beta_2 \leq \ldots \leq \beta_n \leq 1$ and $A_i = (x_i]_R, \forall x_i \in X$ with $x_i \leq_R x_j$ iff $i < j$ (sets A_i form a sequence of nested confidence sets $\emptyset \subset A_1 \subset A_2 \subset \ldots \subset A_n \subset X$). When $X = \Re$ and $A_i = (-\infty, x_i]$, we find back the usual definition of p-boxes.

2.5 Possibility distributions

A possibility distribution π is a mapping from X to the unit interval (hence a fuzzy set) such that $\pi(x) = 1$ for some $x \in X$. Several set-functions can be defined from them [8]:

Possibility measures: $\Pi(A) = \sup_{x \in A} \pi(x)$

Necessity measures: $N(A) = 1 - \Pi(A^c)$

Guaranteed poss. measures: $\Delta(A) = \inf_{x \in A} \pi(x)$

Possibility degrees express the extent to which an event is plausible, i.e., consistent with a possible state of the world, necessity degrees express the certainty of events and Δ-measures the extent to which all states of the world where A occurs are plausible. They apply to so-called guaranteed possibility distributions [8] generally denoted by δ.

A possibility degree can be viewed as an upper bound of a probability degree [9]. Let $\mathcal{P}_\pi = \{P, \forall A \subseteq X \text{ measurable}, P(A) \leq \Pi(A)\}$ be the set of probability measures encoded by π. A necessity measure is a special case of belief function when the focal sets are nested.

2.6 Clouds

Formally, a cloud is described by an Interval-Valued Fuzzy Set (IVF) s.t. $(0,1) \subseteq \cup_{x \in X} F(x) \subseteq [0,1]$, where $F(x)$ is an interval $[\delta(x), \pi(x)]$. A cloud is called *thin* when the two membership functions coincide ($\delta = \pi$). It is called *fuzzy* when the lower membership function δ is 0 everywhere. Let α_i be a sequence of α-cuts s.t. $1 = \alpha_0 > \alpha_1 > \alpha_2 > \ldots > \alpha_n > \alpha_{n+1} = 0$

with A_i the strong α-cut of π and B_j the α-cut of δ ($A_i = \{x_i, \pi(x_i) > \alpha_i\}$ and $B_i = \{x_i, \delta(x_i) \geq \alpha_i\}$). Then, a random variable x is in a cloud if it satisfies the constraints

$$P(B_i) \leq 1 - \alpha_i \leq P(A_i) \quad i = 1, \ldots, n. \qquad (2)$$

with $B_i \subseteq A_i$. The probability family induced by a cloud $[\delta(x), \pi(x)]$ will be noted $\mathcal{P}_{[\delta, \pi]}$.

Moreover, the following property linking clouds and possibility distributions has been shown by Dubois and Prade [11]:

Proposition 1. *A probability family $\mathcal{P}_{[\delta, \pi]}$ described by the cloud (δ, π) is equivalent to the family $\mathcal{P}_\pi \cap \mathcal{P}_{1-\delta}$ described by the two possibility distributions π and $1 - \delta$.*

3 Probability intervals and random sets

Most of the results presented in this section can be found in [1, 6], where more details can be found. There are two main approaches to build a belief function Bel from a set L_e of probability intervals.

The first one, explored in [13] by Lemmer and Kyburg, considers probability intervals as a partial specification of a belief function and consists to find a belief function Bel_1 that extends the intervals s.t.

$$Bel_1(x_i) = l_i \text{ and } Pl_1(x_i) = u_i \; \forall i \qquad (3)$$

As shown in [13], finding such a belief function is possible iff the tree following conditions hold

$$\sum_{i=1}^n l_i \leq 1 \leq \sum_{i=1}^n u_i$$

$$\sum_{j \neq i} l_j + u_i \leq 1 \text{ and } \sum_{j \neq i} u_j + l_i \geq 1 \quad \forall i$$

$$\sum_{i=1}^n l_i + \sum_{i=1}^n u_i \geq 2$$

where the two first conditions correspond to non-emptiness and reachability (which are always satisfied by supposition). Lemmer and Kyburg also provide a means to build one of the belief function satisfying constraints (3). Let us note that with this method, we have $Bel_1(A) \geq P(A)$, which imply $\mathcal{P}_{Bel_1} \subset \mathcal{P}_{L_e}$. Thus, The belief function Bel_1 is an inner approximation of the family $\mathcal{P}_{L_e}$.

The second approach, extensively explored by Denoeux [6], considers probability intervals L_e as some "most committed" information and try to find a conservative belief function Bel_2 s.t. $Bel_2(A) \leq \underline{P}_{L_e}(A) \; \forall A$. Since there exist a lot of such belief functions, Denoeux also proposes to find the belief function

Bel_2 that maximizes a given specifity criterion, in order to keep as much information as possible (in [6], this criterion is the sum of belief degrees over events). Obviously, with this approach, we have $\mathcal{P}_{L_e} \subset \mathcal{P}_{Bel_2}$, and Bel_2 is this time an outer approximation of the family $\mathcal{P}_{L_e}$. Methods given in [6] insure that $Bel_2(x_i) = l_i$, but it can be checked from the examples in [6] that $Pl_2(x_i) > u_i$, even if the three conditions required by the first approach hold.

4 Probability intervals and (Generalized) P-boxes

Given a set L of probability intervals and a meaningful ordering relation $\leq_R$ between elements x_i, one can easily build a generalized p-box $[\underline{F}, \overline{F}]_L$ from L. Given the consecutive sets $A_i = (x_i]_R, \forall x_i \in X$ and the ordering s.t. $x_i \leq_R x_j$ iff $i < j$, lower and upper generalized cumulative distributions corresponding to L are, respectively

$$\underline{F}_R(x_i) = \underline{P}(A_i) = \max(\textstyle\sum_{x_i \in A_i} l_j, 1 - \sum_{x_i \notin A_i} u_j)$$
$$\overline{F}_R(x_i) = \overline{P}(A_i) = \min(\textstyle\sum_{x_i \in A_i} u_i, 1 - \sum_{x_i \notin A_i} l_i)$$
$$(4)$$

Now, if we consider that this p-box is all the information we have, one can easily find back probability intervals L' from this information s.t.

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

and we have the following proposition

Proposition 2. *Given an initial set L of probability intervals over a space X, and given the transformations*

$$Set\ L \quad \overrightarrow{p-box} \quad [\underline{F}_R(x), \overline{F}_R(x)] \quad \overrightarrow{Interval} \quad Set\ L'$$

we have that $\mathcal{P}_L \subseteq \mathcal{P}_{L'}$.

Proof. Looking at equations given above, we can easily express $[l_i', u_i']$ in term of values l_i, u_i of the original set L, this gives us

$$l_i' = \max(0, \sum_{x_i \in A_i} l_i - \sum_{x_i \in A_{i-1}} u_i,$$
$$\sum_{x_i \in A_i} l_i + \sum_{x_i \notin A_{i-1}} l_i - 1,$$
$$1 - \sum_{x_i \notin A_i} u_i - \sum_{x_i \in A_{i-1}} u_i,$$
$$\sum_{x_i \notin A_{i-1}} l_i - \sum_{x_i \notin A_i} u_i)$$

$$= \max(0, \quad l_i + \sum_{x_i \in A_{i-1}} (l_i - u_i),$$

$$l_i + (\sum_X l_i - 1),$$

$$1 - \sum_{X \setminus x_i} u_i$$

$$l_i + \sum_{x_i \notin A_i} (l_i - u_i))$$

and, given that $u_i \geq l_i$ and that the set L is reachable and non-empty, we have that $l'_i \leq l_i$. The same procedure can be followed for the bounds u'_i, and we have $\mathcal{P}_L \subseteq \mathcal{P}_{L'}$. We also have $l'_i = l_i; u'_i = u_i \; \forall i$ only in degenerate cases of precise information or complete ignorance. $\square$

Not surprisingly, this proposition shows that transforming probability intervals into p-boxes implies a loss of information. Let us note that the same argument holds when one wants to transform a generalized p-box $[\underline{F}, \overline{F}]$ into probability intervals, and then get back a generalized p-box $[\underline{F'}, \overline{F'}]$ from these intervals using equations (4) (i.e. we have $\mathcal{P}_{[\underline{F}, \overline{F}]} \subset \mathcal{P}_{[\underline{F'}, \overline{F'}]}$).

Example 1. *Let us take the same four probability intervals as in the example given in [14], summarized in the following tabular*

	x_1	x_2	x_3	x_4
l_i	0.10	0.34	0.25	0
u_i	0.28	0.56	0.46	0.08

if we consider the order R s.t. $x_i \leq_R x_j$ iff $i \leq j$, this gives us the following generalized p-box

	$\underline{F}$	$\overline{F}$
$A_1 = \{x_1\}$	0.10	0.28
$A_2 = \{x_1, x_2\}$	0.46	0.75
$A_3 = \{x_1, x_2, x_3\}$	0.92	1
$A_4 = X$	1	1

and if we want to get back probability intervals from this sole generalized p-box, we obtain

	x_1	x_2	x_3	x_4
l'_i	0.10	0.18	0.17	0
u'_i	0.28	0.65	0.54	0.08

which is clearly less informative than the first probability intervals.

5 Probability intervals and possibility distributions

The problem of transforming a probability distribution into a quantitative possibility distribution has been

adressed by many authors (see [10] for an extended discussion about the links between probabilities and quantitative possibility theory). In this paper, we will follow the same line as Masson and Denoeux in [14], where authors study the problem of transforming interval probabilities into a possibility distribution.

Concerning the transformation of a precise probability into a possibility distribution, a first consistency principle was informally stated by Zadeh [18] as: what is probable should be possible. It was later translated by Dubois and Prade [12] as the mathematical constraint

$$P(A) \leq \Pi(A) \quad \forall A \subseteq X$$

and the possibility measure Π is said to dominate the probability measure P. The transformation of a probability into a possibility then consists of choosing a possibility measure amongst those dominating P. Dubois and prade [12] then proposed to add the following constraint

$$p(x_i) \leq p(x_j) \leftrightarrow \pi(x_i) \leq \pi(x_j)$$

and to choose the least specific possibility distribution (π' is more specific than π if $\pi' \leq \pi \; \forall x$) respecting these two constraints. They showed that the solution exists and is unique. Let us consider the order on probability masses s.t.

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

Dubois and Prade's transformation can then be formulated as

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

When working with a set $L = \{[l_i, u_i], i = 1, \ldots, n\}$ of probability intervals, the order induced on probability masses is no longer complete, and the partial order is reduced to

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

and two probabilities $p(x_i), p(x_j)$ are incomparable if intervals $[l_i, u_i], [l_j, u_j]$ intersect in some way. Let us note $\mathcal{M}$ this partial order and $\mathcal{C}$ the set of its linear extensions (a linear extension $C_l \in \mathcal{C}$ is a complete order that is compatible with the partial order $\mathcal{M}$). Given this partial order, Masson and Denoeux [14] propose the following procedure to transform the set of probability intervals into a possibility distribution:

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

$$\pi(x_i)^{C_l} = \max_{p(x_1), \ldots, p(x_n)} \sum_{\sigma_l^{-1}(j) \leq \sigma_l^{-1}(i)} p(x_j) \quad (5)$$

under the constraints

$$\begin{cases} \sum_{k=1,\ldots,n} p(x_k) = 1 \\ l_k \le p_k \le u_k \\ p(x_{\sigma_l(1)}) \le p(x_{\sigma_l(2)}) \le \ldots \le p(x_{\sigma_l(n)}) \end{cases}$$

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.

$$\pi(x_i) = \max_{C_l \in \mathcal{C}} \pi(x_i)^{C_l} \quad \forall i \qquad (6)$$

this procedure insures that the resulting possibility distribution π will dominate every probability distribution contained in $\mathcal{P}_L$. We thus have $\mathcal{P}_L \subset \mathcal{P}_\pi$. Let us note that this transformation can produce an important loss of information.

Example 2. *Taking back probability intervals of example 1, we have here three possible linear extensions $C_l \in \mathcal{C}$*

$$\begin{aligned} C_1 &= (L_4, L_1, L_3, L_2) \\ C_2 &= (L_4, L_1, L_2, L_3) \\ C_3 &= (L_4, L_3, L_1, L_2) \end{aligned}$$

using the above method on these three orders gives the following distribution (see [14] for more details)

	x_1	x_2	x_3	x_4
π	0.64	1	1	0.08

Let us note that, if this transformation produces the most specific possibility distribution dominating a precise probability distribution (i.e. $u_i = l_i$ $\forall i$ and $\sum u_i = 1$), this is no longer the case when the probability distribution is imprecisely known through intervals ($u_i \ne l_i$). To see this, let us first recall that an upper generalized R-cumulative distribution $\overline{F}_R$ can be seen as a possibility distribution dominating a probability measure, since it is a maxitive measure (i.e. we have $\max_{x \in A} \overline{F}_R(x) \ge \Pr(A), \forall A \subseteq X$). Any R-cumulative distribution dominating a family $\mathcal{P}_L$ induced by probability intervals is thus also a possibility measure dominating this family. Now, in our example, let us consider the following order $<_R$ between the four elements: $x_4 <_R x_1 <_R x_3 <_R x_2$. From this order, we can build the following R-cumulative distribution:

	x_1	x_2	x_3	x_4
$\overline{F}_R = \pi_R$	0.36	1	0.66	0.08

which still dominates $\mathcal{P}_L$ and is more specific than the distribution built by the method given in [14]. This shows that Masson and Denoeux's method tends to give conservative bounds, and thus can result in an important loss of information, which seems hard to justify. This important loss is due to the fact that Masson and Denoeux do not consider any specific ordering relation R between the elements of X, an assumption that is made in the counter-example given above.

6 Probability intervals and clouds

Let L be a set of probability intervals and $\mathcal{P}_L$ the associated family. In this section, we introduce a method, inspired from [14], which transforms intervals L into a cloud $[\delta, \pi]_L$.

From property 1, we know that families described by clouds are equivalent to the intersection of two families described by possibility distributions. Another property of clouds is that a discrete thin cloud, up to the right transformation, can represent a precise probability distribution. Let $\pi_1(x_i) = a_i$ s.t. $a_0 = 0 < a_1 < \ldots < a_n = 1$. Consider then $\pi_2(x_i) = 1 - a_{i-1} = 1 - \delta(x_i)$. Now let us consider the cloud $[\delta = 1 - \pi_2, \pi = \pi_1]$. It has been proved by Dubois and Prade [11] that the probability family $\mathcal{P}_{[\delta,\pi]}$ induced by this cloud contains a unique probability measure P s.t. $p(x_i) = a_i - a_{i-1} \forall i = 1, \ldots, n$.

Given this property, two requirements that should, in our opinion, follow any transformation of probability intervals into clouds are the following:

- If intervals L describe a precise probability distribution, then the transformation should result in the corresponding thin cloud.

- The family $\mathcal{P}_{[\delta,\pi]_L}$ should be an outer approximation of $\mathcal{P}_L$ (i.e. $\mathcal{P}_L \subset \mathcal{P}_{[\delta,\pi]_L}$)

Let us consider the distribution π built with the method described in section 5 as the upper distribution of the cloud. By reversing the inequality under the summand in equation (5), we can build another distribution π_δ in the following way:

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

$$\begin{aligned} \pi_\delta^{C_l}(x_i) &= \max_{p(x_1),\ldots,p(x_n)} \sum_{\sigma_l^{-1}(i) \le \sigma_l^{-1}(j)} p(x_j) \\ &= 1 - \min_{p(x_1),\ldots,p(x_n)} \sum_{\sigma_l^{-1}(j) < \sigma_l^{-1}(i)} p(x_j) \\ &= 1 - \delta^{C_l}(x_i) \end{aligned}$$

with the same constraints as in section 5

2. Take the distribution dominating all distributions $\pi_\delta^{C_l}(x_i)$

$$\pi_\delta(x_i) = 1 - \delta(x_i) = \max_{C_l \in \mathcal{C}} \pi_\delta^{C_l}(x_i) \quad \forall i \qquad (7)$$

And $[\delta, \pi]$ is actually a cloud ($\delta < \pi$, since $\delta(x_i)$ is a minimum computed over less intervals $L(x_i)$ than $\pi(x_i)$, which is moreover a maximum). It can be easily checked that if L is reduced to a precise probability distribution, then the above method results in the thin cloud corresponding to this distribution. We also have that $\mathcal{P}_L \subset \mathcal{P}_{\pi_\delta}$ and $\mathcal{P}_L \subset \mathcal{P}_\pi$ by construction, thus we have $\mathcal{P}_L \subset \mathcal{P}_{[\delta,\pi]_L} = \mathcal{P}_{\pi_\delta} \cap \mathcal{P}_\pi$, and our two requirements are satisfied.

Example 3. *Again, we consider the probability intervals given in example 1. From the three orders C_i given above, we can obtain the following π_δ*

C_i	$\pi_\delta(x_1)$	$\pi_\delta(x_2)$	$\pi_\delta(x_3)$	$\pi_\delta(x_4)$
1	1	0.16	0.63	1
2	1	0.9	0.46	1
3	0.75	0.5	1	1
max	1	0.9	1	1

and, finally, the following cloud

	x_1	x_2	x_3	x_4
π	0.64	1	1	0.08
δ	0	0.1	0	0

which, in this case, is only a little more informative than the upper distribution taken alone (indeed, the sole added constraint is that $p(x_2) \leq 0.9$).

As in the previous section, this method can be criticized upon the ground that an important amount of information is lost. Instead, one could take, for example, the cloud associated to the generalized p-box induced by the order $x_4 <_R x_1 <_R x_3 <_R x_2$ (since generalized p-boxes are a particular case of clouds). This would give the following distributions:

	x_1	x_2	x_3	x_4
$\overline{F}_R = \pi_R$	0.36	1	0.66	0.08
$\underline{F}_R$	0.1	1	0.44	0
δ_R	0	0.44	0.1	0

Where δ_R is $\underline{F}_R$ after a simple shift of values (the necessity of this transformation, already emphasized in [11], arises from the fact that sets A_i are strong α-cuts, while B_i are simple α-cuts). The cloud $[\delta_R, \pi_R]$ is obviously more specific than the first cloud while we still have $\mathcal{P}_L \subset \mathcal{P}_{[\delta_R, \pi_R]}$.

7 Conclusions

Probability intervals are very convenient to model uncertainty, and can be encountered in various situations. In this paper, we study how they can be transformed into other popular models of imprecise probabilities. Except for one, every method recalled or proposed here give outer approximation of the family $\mathcal{P}_L$. This corresponds to a cautious view, since there is no additional information present in the model resulting from the transformation (but some information can be lost in the transformation process).

If probability intervals are reduced to precise probability distributions, every proposed transformation result in the model corresponding to this precise probability, except for possibility distributions, which are the only model studied here that can't be seen as a generalization of classical probabilities.

Interestingly enough, most (some subclasses of clouds are not) of the studied representations in this paper are special cases of random sets, allowing one to use all the rich mathematical background of this theory as well as many computational simplifications (e.g. easiness to compute Choquet Integral).

Each of the models presented here has its own practical interest in term of expressiveness or tractability. The natural continuation of the work initiated in this paper is to extend the study made by De Campos et al. [1] to every model studied here. What becomes of random sets, possibility distributions, generalized p-boxes and clouds after fusion, marginalization, conditioning or propagation? Is it still the same kind of representation after having applied these mathematical tools, and under which assumptions? To which extends are these representations informative? Can they easily be elicited or integrated? If many results already exist for random sets and possibility distributions, few have been derived for generalized p-boxes or clouds, due to the fact that these two latter representations have only been recently proposed.

In term of behavioral interpretation of imprecise probabilities [17], conditions of non-emptiness and reachability respectively correspond to avoiding sure loss and to coherence of lower previsions. It is also interesting to note that, if we see probability intervals as a constraint satisfaction problem (CSP) [4], non-emptiness and reachability correspond to the notions of existence of a solution and of bounds consistency. Another interesting work should be to formalize links between CSP and imprecise probabilities, with the aim to study to which extent technics used in CSP can be used to solve practical problems commonly encountered when dealing with imprecise probabilities. To our knowledge, this work largely remains to be done, although CSP technics are already used to solve practical problems related to (imprecise) probabilistic reasoning, most of them being related to (credal) bayesian networks (see, for example [3]).

References

[1] L. Campos, J. Huete, and S. Moral. Probability intervals : a tool for uncertain reasoning. *I. J. of Uncertainty, Fuzziness and Knowledge-Based Systems*, 2:167–196, 1994.

[2] G. Choquet. Theory of capacities. *Annales de l'institut Fourier*, 5:131–295, 1954.

[3] R. Dechter. Bucket-elimination : a unifying framework for probabilistic inference. In E. Horvitz and F. Jensen, editors, *Proc. 12th conf. on Uncertainty in A.I.*, pages 211–219, Portland, 1996.

[4] R. Dechter. *Constraint Processing*. Morgan Kaufman, 2003.

[5] A. Dempster. Upper and lower probabilities induced by a multivalued mapping. *Annals of Mathematical Statistics*, 38:325–339, 1967.

[6] T. Denoeux. Constructing belief functions from sample data using multinomial confidence regions. *I. J. of Approximate Reasoning*, 42, 2006.

[7] S. Destercke and D. Dubois. A unified view of some representations of imprecise probabilities. In J. Lawry, E. Miranda, A. Bugarin, and S. Li, editors, *Int. Conf. on Soft Methods in Probability and Statistics (SMPS)*, Advances in Soft Computing, pages 249–257, Bristol, 2006. Springer.

[8] D. Dubois, P. Hajek, and H. Prade. Knowledge-driven versus data-driven logics. *Journal of logic, Language and information*, 9:65–89, 2000.

[9] D. Dubois and H. Prade. When upper probabilities are possibility measures. *Fuzzy Sets and Systems*, 49:65–74, 1992.

[10] D. Dubois and H. Prade. Quantitative possibility theory and its probabilistic connections. In P. Grzegorzewski, O. Hryniewicz, and M. Angeles-Gil, editors, *Soft Methods in Probability, Statistics and Data Analysis*, Advances in Soft Computing, pages 3–26. Physica Verlag, Heidelberg - Germany, 2002.

[11] D. Dubois and H. Prade. Interval-valued fuzzy sets, possibility theory and imprecise probability. In *Proceedings of International Conference in Fuzzy Logic and Technology (EUSFLAT'05)*, Barcelona, September 2005.

[12] D. Dubois, H. Prade, and S. Sandri. On possibility/probability transformations. In *Proc. of the Fourth International Fuzzy Systems Association World Congress (IFSA'91)*, pages 50–53, Brussels, Belgium, 1991.

[13] J. Lemmer and H. Kyburg. Conditions for the existence of belief functions corresponding to intervals of belief. In *Proc. 9th National Conference on A.I.*, pages 488–493, 1991.

[14] M. Masson and T. Denoeux. Inferring a possibility distribution from empirical data. *Fuzzy Sets and Systems*, 157(3):319–340, february 2006.

[15] A. Neumaier. Clouds, fuzzy sets and probability intervals. *Reliable Computing*, 10:249–272, 2004.

[16] G. Shafer. *A mathematical Theory of Evidence*. Princeton University Press, 1976.

[17] P. Walley. *Statistical reasoning with imprecise Probabilities*. Chapman and Hall, 1991.

[18] L. Zadeh. Fuzzy sets as a basis for a theory of possibility. *Fuzzy sets and systems*, 1:3–28, 1978.