IMPRECISE PROBABILITY IN RISK ANALYSIS

D. Dubois

IRIT-CNRS, Université Paul Sabatier
31062 TOULOUSE FRANCE

Outline

1. Variability vs incomplete information
2. Blending set-valued and probabilistic representations : uncertainty theories
3. Possibility theory in the landscape
4. A methodology for risk-informed decision-making
5. Some applications

Origins of uncertainty

- The variability of observed repeatable natural phenomena : « **randomness ».**
  - Coins, dice…: what about the outcome of the next throw?
- The lack of information: **incompleteness**
  - because of information is often lacking, knowledge about issues of interest is generally not perfect.
- Conflicting testimonies or reports:**inconsistency**
  - The more sources, the more likely the inconsistency

Example

- **Variability**: daily quantity of rain in Toulouse
  - May change every day
  - It can be estimated through statistical observed data.
  - Beliefs or prediction based on this data
- **Incomplete information** : Birth date of Brazil President
  - It is not a variable: it is a constant!
  - You can get the correct info somewhere, but it is not available.
  - Most people may have a rough idea (an interval), a few know precisely, some have no idea: information is subjective.
  - Statistics on birth dates of other presidents do not help much.
- **Inconsistent information :** several sources of information conflict concerning the birth date (a book, a friend, a website).

The roles of probability

Probability theory is generally used for representing two aspects:

1. **Variability:** capturing (beliefs induced by) variability through repeated observations.
2. **Incompleteness (info gaps):** directly modeling beliefs via betting behavior observation.

These two situations are not mutually exclusive.

Using a single probability distribution to represent incomplete information is not entirely satisfactory:

The betting behavior setting of Bayesian subjective probability enforces a representation of partial ignorance based on single probability distributions.

1. **Ambiguity :** In the absence of information, how can a uniform distribution tell pure randomness and ignorance apart ?
2. **Instability :** A uniform prior on $x \in [a, b]$ induces a non-uniform prior on $f(x) \in [f(a), f(b)]$ if f is increasing and non-affine : ignorance generates information???
3. **Empirical falsification**: When information is missing, decision-makers do not always choose according to a single subjective probability (Ellsberg paradox).

Motivation for going beyond probability

- Have a language that distinguishes between uncertainty due to variability from uncertainty due to lack of knowledge or missing information.
- **For describing variability: Probability distributions**
  - but information demanding, and paradoxical for ignorance
- **For representing incomplete information : Sets (intervals).**

  but a very crude representation of uncertainty
- *Find representations that allow for both aspects of uncertainty.*

Set-Valued Representations of Partial Information

- A piece of incomplete information about an ill-known quantity x is represented by a pair (x, E) where E is a set called a *disjunctive* (***epistemic***) set,
- E is a subset of *mutually exclusive* values, one of which is the real x.
- (x, E) means « *all I know is that* $x \in E$ »
  - **Intervals** E = [a, b]: incomplete numerical information
  - **Classical Logic**: incomplete symbolic information

    E = Models of a proposition stated as true.
- ***Such sets are as subjective as probabilities***

BOOLEAN POSSIBILITY THEORY

*If all **you** know is that $x \in E$ then*

- You judge **event A possible** if it is logically consistent with what you know : $A \cap E \neq \emptyset$

A Boolean possibility function : $\Pi(A) = 1$, and 0 otherwise

- **You believe event A (sure)** if it is a logical consequence of what we already know : $E \subseteq A$

A certainty (necessity) function : $N(A) = 1$, and 0 otherwise

- *This is a simple modal **epistemic** logic (KD45)*

$$N(A) = 1 - \Pi(A^c) \leq \Pi(A)$$

$$\Pi(A \cup B) = \max(\Pi(A), \Pi(B));\ N(A \cap B) = \min(N(A), N(B)).$$

WHY TWO SET-FUNCTIONS ?

- **Encoding 3 extreme epistemic states….**
  - Certainty of truth : $N(A) = 1$ (hence $\Pi(A) = 1$)
  - Certainty of falsity: $\Pi(A) = 0$ (hence $N(A) = 0$)
  - Ignorance : $\Pi(A) = 1, N(A) = 0$

….. **requires 2 Boolean variables!**

The Boolean counterpart of a subjective probability

- With one function you can only say believe A or believe not-A.

but this representation is poorly expressive (no gradation in uncertanty

Find an extended representation of uncertainty

- *Explicitly allowing for missing information (= that uses sets)*
- *More informative than pure intervals or classical logic,*
- *Less demanding and more expressive than single probability distributions*
- *Allows for addressing the issues dealt with by both standard probability, and logics for reasoning about knowledge.*

Blending intervals and probability

- Representations that may account for variability, incomplete information, and belief must combine probability and epistemic sets.
  - Sets of probabilities : imprecise probability theory
  - Random(ised) sets : Dempster-Shafer theory
  - Fuzzy sets: numerical possibility theory
- **Relaxing the probability axioms** :
  - Each event has a degree of certainty and a degree of plausibility, instead of a single degree of probability
  - When plausibility = certainty, it yields probability

A GENERAL SETTING FOR REPRESENTING GRADED CERTAINTY AND PLAUSIBILITY

- 2 set-functions Pl and Cr, with values in [0, 1], generalizing probability, possibility and necessity.
- **Conventions** :
  - Pl(A) = 0 "impossible" ;
  - Cr(A) = 1 "certain"
  - Pl(A) =1 ; Cr(A) = 0 "ignorance" **(no information)**
  - Pl(A) - Cr(A) quantifies ignorance about A
- **Postulates**
  - If $A \subseteq B$ then $Cr(A) \leq Cr(B)$ and $Pl(A) \leq Pl(B)$
  - $Cr(A) \leq Pl(A)$ "certain implies plausible"
  - $Pl(A) = 1 - Cr(A^c)$ duality certain/plausible

Imprecise probability theory

- A state of information is represented by a family $\mathcal{P}$ of probability distributions over a set X.
  - For instance an imprecise probabilistic model.
- To each event A is attached a probability interval $[P_*(A), P^*(A)]$ such that
  - $P_*(A) = \inf\{P(A), P \in \mathcal{P}\}$
  - $P^*(A) = \sup\{P(A), P \in \mathcal{P}\} = 1 - P_*(A^c)$
- $\mathcal{CP} = \{P, P(A) \geq P_*(A) \text{ for all } A\}$ is convex
- Usually $\mathcal{CP}$ is strictly contained in $\{P, P \geq P_*\}$

Random sets and evidence theory

- A family $\mathcal{F}$ of « focal » (disjunctive) non-empty sets representing
  - A collection of incomplete observations (imprecise statistics).
  - Unreliable testimonies
- A positive weighting of focal sets (a random set) :

$$\sum_{E \in \mathcal{F}} m(E) = 1 \quad (\textit{mass function})$$

- It is a randomized epistemic state where
  - m(E) = *probability*(E is the correct information)
    = *probability(only knowing"(x, E)")*

Theory of evidence

- **degree of certainty** (belief) :
  - $Bel(A) = \sum_{E_i \subseteq A,\, E_i \neq \emptyset} m(E_i)$
  - total mass of information implying the occurrence of A
  - *(probability of provability)*
- **degree of plausibility** :
  - $Pl(A) = \sum_{E_i \cap A \neq \emptyset} m(E_i) = 1 - Bel(A^c) \geq Bel(A)$
  - total mass of information consistent with A
  - *(probability of consistency)*

Possibility Theory

(Shackle, 1961, Lewis, 1973, Zadeh, 1978)

- A piece of incomplete information "$x \in E$" admits of *degrees* of possibility.
- *E is mathematically a (normalized) fuzzy set.*
- $\mu_E(s)$ = Possibility(x = s) = $\pi_x(s)$
- **Conventions:**

  $\forall s$, $\pi_x(s)$ is the degree of plausibility of x = s

  $\pi_x(s) = 0$ iff x = s is impossible, totally surprising

  $\pi_x(s) = 1$ iff x = s is normal, fully plausible, unsurprising (but no certainty)

POSSIBILITY AND NECESSITY OF AN EVENT

How confident are we that $x \in A \subset S$? (*an event A occurs*)
given a possibility distribution π for x on S

- $\Pi(A) = \max_{s \in A} \pi(s)$:
  to what extent A is consistent with π
  (= some $x \in A$ is possible)
  The degree of possibility *that* $x \in A$
- $N(A) = 1 - \Pi(A^c) = \min_{s \notin A} 1 - \pi(s)$:
  to what extent no element outside A is possible
  = to what extent π implies A
  The degree of certainty (necessity) that $x \in A$

Basic properties

$$\Pi(A \cup B) = \max(\Pi(A), \Pi(B));$$

$$N(A \cap B) = \min(N(A), N(B)).$$

Mind that most of the time :

$$\Pi(A \cap B) < \min(\Pi(A), \Pi(B));$$

$$N(A \cup B) > \max(N(A), N(B)$$

Example: Total ignorance on A and $B = A^c$

Corollary $N(A) > 0 \Rightarrow \Pi(A) = 1$

POSSIBILITY AS UPPER PROBABILITY

- Given a numerical possibility distribution π, define

  $\mathcal{P}(\pi) = \{P \mid P(A) \leq \Pi(A) \text{ for all } A\}$

  - Then, Π and N can be recovered
    - $\Pi(A) = \sup \{P(A) \mid P \in \mathcal{P}(\pi)\}$;
    - $N(A) = \inf \{P(A) \mid P \in \mathcal{P}(\pi)\}$
  - So π is a faithful representation of a special family of probability measures
- Likewise for belief functions : $\mathcal{P}(\pi) = \{P \mid P(A) \leq Pl(A), \forall A\}$
- Possibility theory corresponds to consonant belief functions
  - Nested focal sets: $m(E) > 0$ and $m(F) > 0$ imply $F \subseteq E$ or $E \subseteq F$
  - If and only if $Pl(A) = \Pi(A)$ and $Bel(A) = N(A)$.

How to build possibility distributions

(not related to linguistic fuzzy sets!!!)

- *Nested* random sets (= *consonant belief functions*)
- *Likelihood functions* (in the absence of priors).
- *Probabilistic inequalities* (Chebyshev…)
- *Confidence intervals* (moving the confidence level between 0 and 1)
- *The cumulative PDF* of P **is** a possibility distribution (accounting for all probabilities stochastically dominated by P)

LANDSCAPE OF UNCERTAINTY THEORIES

BAYESIAN/STATISTICAL PROBABILITY
Randomized points

UPPER-LOWER PROBABILITIES
Disjunctive sets of probabilities

DEMPSTER UPPER-LOWER PROBABILITIES
SHAFER-SMETS **BELIEF FUNCTIONS**
Random disjunctive sets

Quantitative Possibility theory
Fuzzy (nested disjunctive) sets

Classical logic
Disjunctive sets

A risk analysis methodology

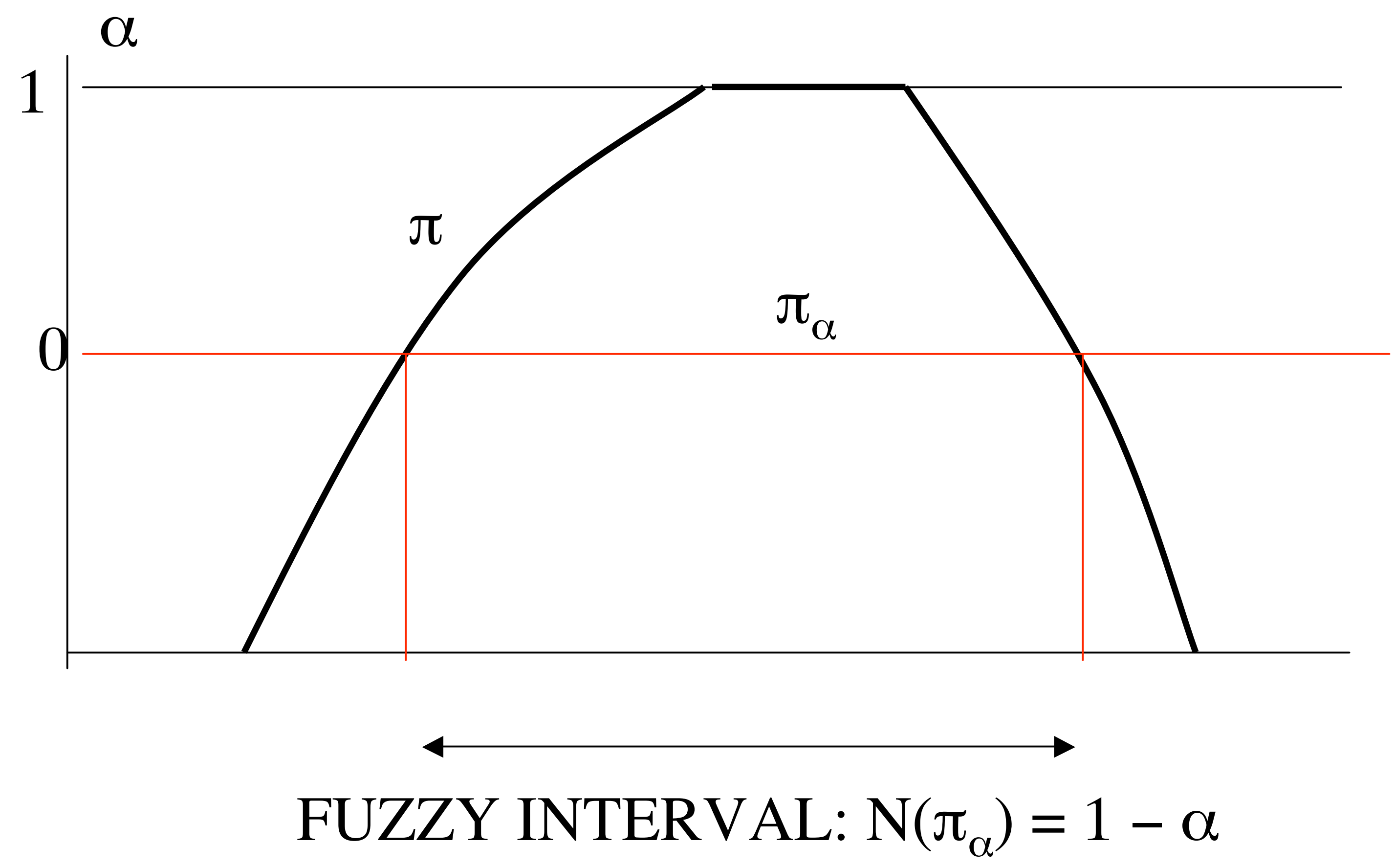
1. Information collection and representation
2. Propagation of uncertainty through a mathematical model
3. Extraction of useful information
4. Decision step

Risk analysis : *Information collection step*

- **Faithfulness principle** : choose the type of representation in agreement with the quantity of available information
  - Remain faithful to available information, including information gaps.
- Simple representations (possibility, generalized p-boxes) naturally capture expert interval information with confidence levels, quantiles, means, mode, etc.
  - If *variability and enough statistical information*: probability distributions.
  - If incomplete information on some value : interval, possibility distribution (fuzzy interval)…
  - If parameterized model with ill-known parameters : p-box
- An elicitation procedure to query an expert on available information

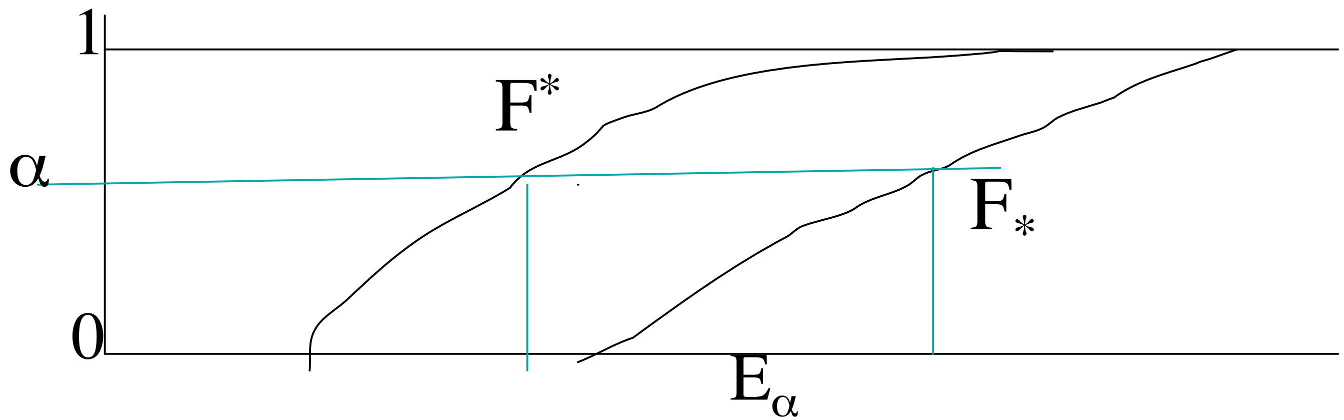
A possibility distribution can be obtained from any family of nested confidence sets :

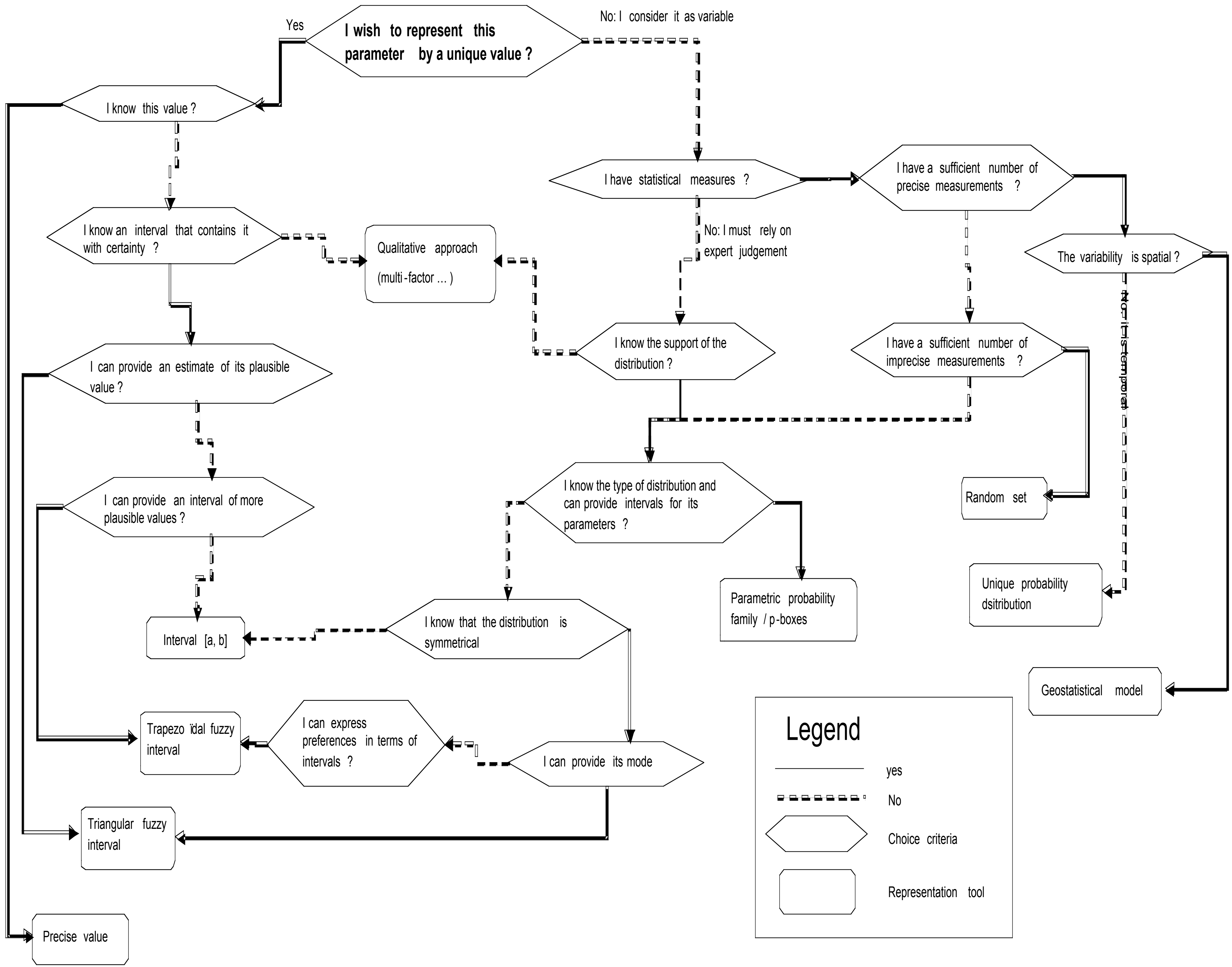
$$P(A_\alpha) \geq 1 - \alpha, \alpha \in (0, 1]$$



Probability boxes

- A set $\mathcal{P} = \{P: F^* \geq P \geq F_*\}$ induced by two cumulative disribution functions is called a **probability box (p-box)**,
- **A p-box is a special random interval whose upper and bounds induce the same ordering.**





I wish to represent this parameter by a unique value ?
Yes
No: I consider it as variable
I know this value ?
I know an interval that contains it with certainty ?
Qualitative approach (multi-factor ...)
I can provide an estimate of its plausible value ?
I can provide an interval of more plausible values ?
Interval [a, b]
Trapezoïdal fuzzy interval
Triangular fuzzy interval
Precise value
I have statistical measures ?
No: I must rely on expert judgement
I know the support of the distribution ?
I know the type of distribution and can provide intervals for its parameters ?
I know that the distribution is symmetrical
I can express preferences in terms of intervals ?
I can provide its mode
Parametric probability family / p-boxes
I have a sufficient number of precise measurements ?
I have a sufficient number of imprecise measurements ?
The variability is spatial ?
No: it is temporal
Random set
Unique probability dsitribution
Geostatistical model
Legend
yes
No
Choice criteria
Representation tool

How useful are these practical representations:

- **Cutting complexity:**
  - Convex sets of probability are very complex representations
  - Random sets are potentially exponential
  - P-boxes, possibility distributions and other extensions are linear, but still encode convex probability set, often random sets.
- **Enriching the standard probability analysis** with meta-information and capabilities for reasoning about knowledge in the risk analysis process, while remaining tractable on modern computers.

Information propagation step

- Joint Monte-Carlo and interval analysis to be carried out in the encompassing setting of random sets, with various independence assumptions.
- Distinction between epistemic (in)dependence and stochastic independence
  - Dependent sources and independent variables
  - Independent sources and variables
  - No assumption of independence (more difficult to compute)
- Simple representations cannot be preserved via propagation : general random sets are obtained.

Hybrid possibility-probability propagation

- **Formal problem**: Given a numerical function $f(x_1,\ldots x_m, y_1, \ldots y_n)$,
  - assume $x_1,\ldots x_m$ are independent random variables
  - assume $y_1, \ldots y_n$ are non-interactive possibilistic variables modelled by fuzzy intervals $F_1, \ldots F_n$
- Then $f(x_1,\ldots x_m, y_1, \ldots y_n)$ *is a fuzzy random variable*

Hybrid possibility-probability propagation

- **Computation**
  - Find N samples $a_1,\ldots a_m$ of $x_1,\ldots x_m$ using a Monte-Carlo method.
  - For each sample, compute $f(a_1,\ldots a_m, F_1, \ldots F_n)$ using fuzzy interval computation.

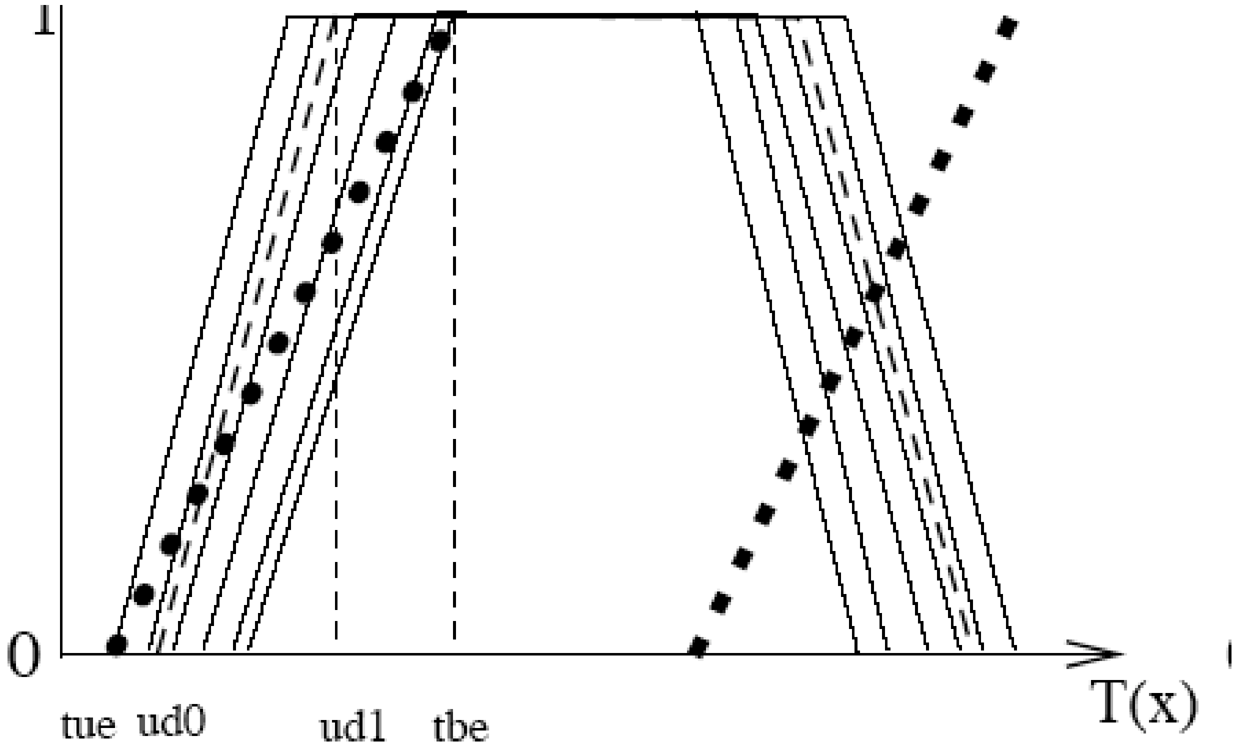
- *As the output*, we get a random fuzzy interval $\{(C_1,v_1) \ldots (C_N,v_N)\}$ *where the* C_i *are fuzzy intervals and* v_i *are frequencies*

Presentation of results:

how to interpret results?

- summarize too complex information : the kind of summary depends on the question of interest:
  - P-boxes can address questions about threshold violations ($x_{out} \geq a$??)
  - *questions of the form* $a \leq x_{out} \leq b$*?* are better addressed by possibility distributions or generalized p-boxes
  - statistical measures of trend, imprecision or variability
- **Aim**: Lay bare the resulting information gap and the resulting knowledge on the variability on the quantity of interest.

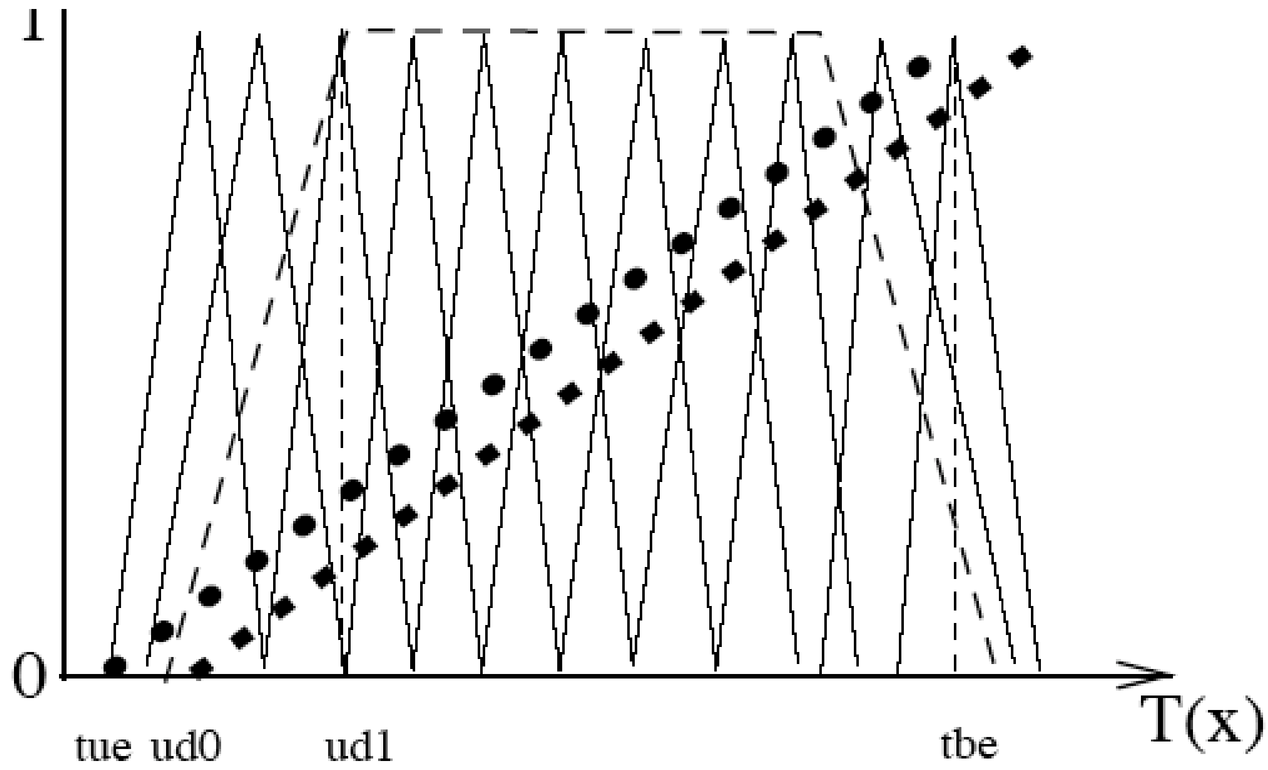
Upper and lower distributions of random fuzzy outputs



small variability of the sample

Large imprecision of each fuzzy number Fi

Upper and lower distributions of random fuzzy outputs



great variability of the sample

i Little imprecision of each fuzzy number Fi

Exploiting Random Fuzzy Intervals

Given a random fuzzy output $\{(C_1,v_1) \ldots (C_N,v_N)\}$:

- ***Average imprecision***: *compute the fuzzy average* $C = \sum_i C_i v_i$. *The average imprecision is the area under* C.
- ***Observable Variability***: *defuzzify the* C_i*'s (midpoint of the mean interval) and compute the standard deviation of these numbers*
- ***Potential Variability***: *Compute the range of the empirical variance induced by the fuzzy intervals.*
- ***Minimal and maximal average variability***: *compute interval variance of the random set with upper probability* $\sum_{I=1,N} \nu_i \Pi_i(A)$.

Example (D. Guyonnet, BRGM)

- *Generic health risk calculation for the case of the exposure of persons to a chlorinated organic solvent (1,1,2-Trichloroethane) via the consumption of contaminated drinking water.*
- The chronic carcinogenic toxicological reference value for this substance is a unit excess risk (UER), namely, a probability of excess cancer per unit daily dose of exposure.
- For a person exposed, we calculate an excess risk (ER) that is a function of the dose D absorbed by this person and of the unit excess risk.

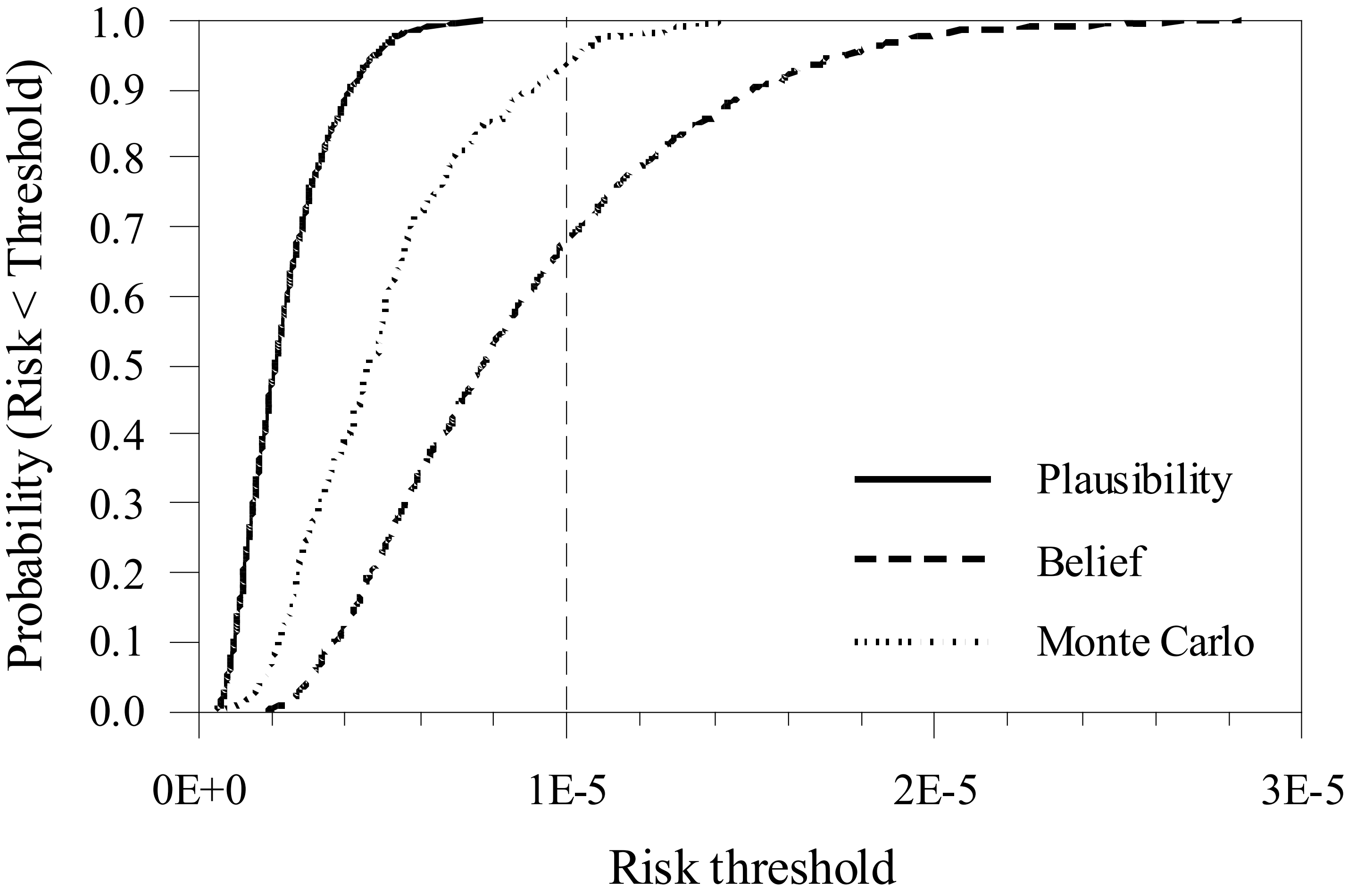
Example (D. Guyonnet, BRGM)

- $ER = D \cdot UER$
- $UER = I \cdot C \cdot FE \cdot DE / BW \cdot LE$
- where:
  - D = dose of exposure (mg pollutant absorbed, per Kg body weight and per day),
  - I = quantity of water ingested per day (L/d),
  - C = concentration of 1,1,2-trichloroethane in the drinking water (mg/L),
  - FE = exposure frequency (d/yr),
  - DE = duration of exposure (yr),
  - BW = body weight (Kg),
  - LE = life expectancy (d).

Parameter values used for the illustration

Parameter	Unit	Mode of representation	Lower limit	Mode or core	Upper limit
Concentration in water	µg/L	Probability	5	10	20
Ingestion	L/d	Fuzzy interval	1	1.5	2.5
Exposure frequency	d/year	Fuzzy interval	200	250	350
Exposure duration	Years	Probability	10	30	50
Oral slope factor	$(mg/Kg/d)^{-1}$	Fuzzy interval	2×10^{-2}	5.7×10^{-2}	10^{-1}

Results (Dubois and Guyonnet, 2010)



A decision strategy

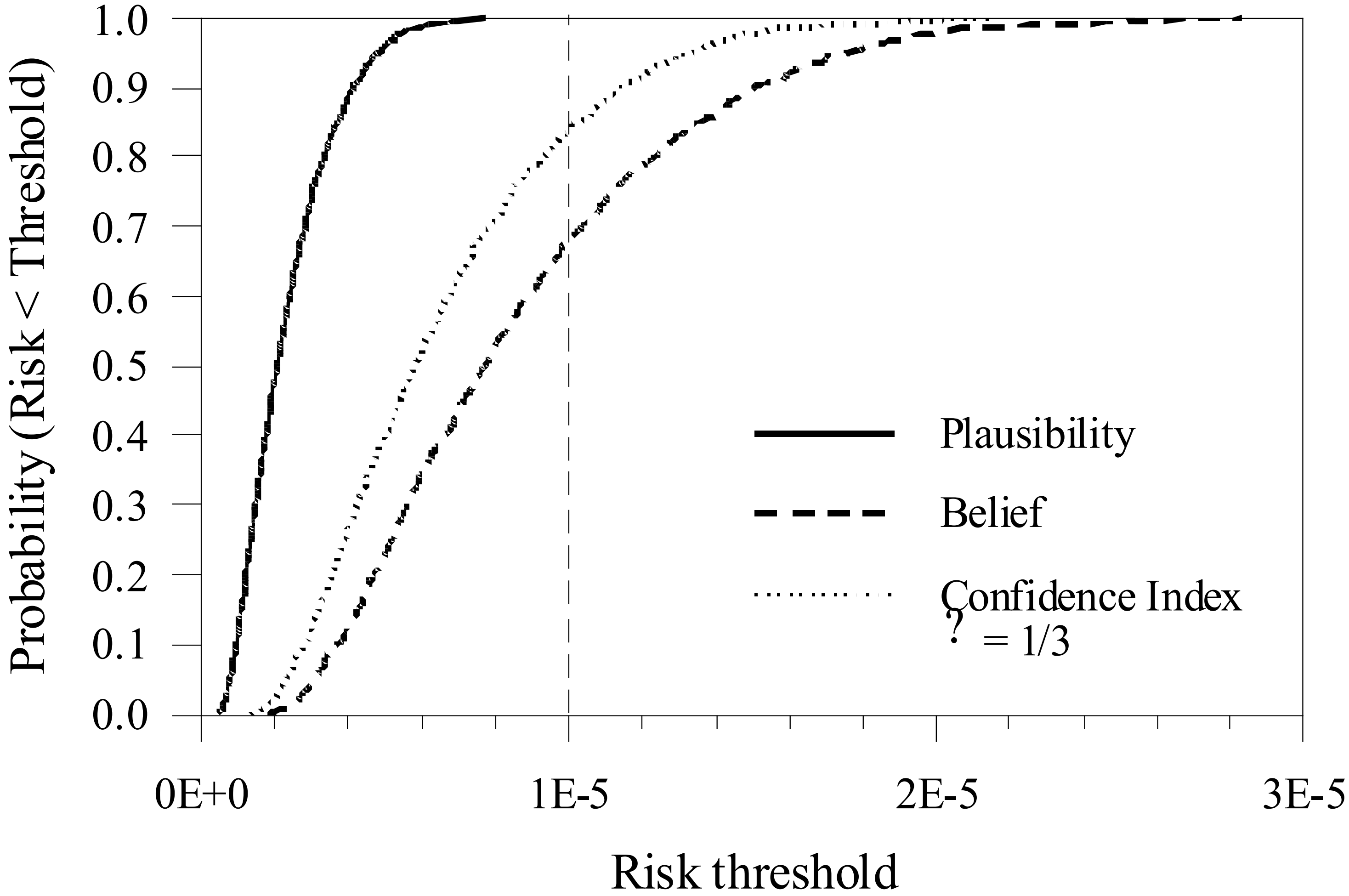
The decision is made by evaluating the probability of a risky event of the form $f(x_1, \ldots x_m, y_1, \ldots y_n) \geq \theta$,

- If the lower probability (or expectation) of the risky event is too high: *take action to circumvent the risk*.
- If the upper probability is low enough : situation not risky despite information gap.
- If the probability of the risky event is ill-known
  - Collect new information to reduce information gap and run the analysis again
  - If information collection is impossible: build up the most reasonable confidence index by consensus from the best experts (possibly being Bayesian again!)

Decision with imprecise probability: 3 attitudes

- Accept incomparability when comparing imprecise utility evaluations of decisions.
  - Pareto optimality : decisions that dominate other choices for all probability functions
  - E-admissibility : decisions that dominate other choices for at least one probability function (Walley, etc…)
- Select a single probability measure for each decision.
  - Compare lower expectations of decisions (Gilboa)
  - Generalize Hurwicz criterion to focal sets with degree of optimism (Jaffray)
- Select a single probability measure and use expected utility
  - Shapley value = pignistic transformation(SMETS)
  - By picking a probability measure that achieves a compromise between pessimistic and optimistic attitudes

Results (Dubois and Guyonnet, 2010)



Some Applications

- Child exposure to lead on an ironworks brownfield (Baudrit et al, Consoil, 2005)
- Contamination of groundwater (Baudrit et al. J. Cont. Hydrology, 2007)
- Radioactivity of cowmilk near la Hague (Baudrit & Chojnacki
- synthesis of multiple sources of information applied to nuclear computer codes (Destercke and Chojnacki, Nuclear Eng.& Design, 2008)
- Underground CO2 confinement (ANR project CRISCO2)

Diana example

- Compute the probability
  - that a mission fails
  - that a plane component fails
- From
  - statistical knowledge coming from databases reporting previous incidents
  - Diagnosis studies (fault trees, etc.)
  - Prognosis studies (duration predictions)
- In order to help in operability and maintenance decisions

Main tools

- Computing probability of faults from the knowledge of fault trees computed by Altarica software.
- Compute probabilities of risky events of interest using algorithms exploiting Markov chain models of component behavior or missions
- *But uncertainty about the probabilistic data*

Modelling uncertainties for DIANA

- **Probabilities are ill-known** : what is the impact of this imprecision on the results and the decision process?
- Computing probabilities attached to cut-sets defining disjunctions of causes of a risky event (Boolean formulas), from knowledge of probabilities of atomic events
- Study of interval-valued Markov chains
  - DO EFFICIENT ALGORITHMS IN THE COMPLETE INFORMATION CASE STILL APPLY ?

The quality of probabilistic information

- Where do basic probabilities come from ?
- Modelling imprecise probabilities : intervals, fuzzy (confidence) intervals ???
- Extreme probabilities : approximate calculations are enough ?
- Propagate imprecision from data to probabilities of events of interest.
- Question independence assumptions ?

Conclusion

- *There exists a coherent range of uncertainty theories combining interval and probability representations.*
  - Imprecise probability is the proper theoretical umbrella
  - The choice between subtheories depends on how expressive it is necessary to be in a given application.
  - There exists simple practical representations of imprecise probability
- *Many open problems, theoretical, and computational, remain.*
- *How to get this general non-dogmatic approach to uncertainty accepted by traditional statisticians?*