UNCERTAINTY THEORIES: A UNIFIED VIEW

D. Dubois

IRIT-CNRS, Université Paul Sabatier

31062 TOULOUSE FRANCE

dubois@irit.fr

Outline

1. Variability vs ignorance
2. Set-valued representations of partial ignorance
3. Blending set-valued and probabilistic representations : uncertainty theories
4. Practical representations
5. A risk analysis methodology

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).

Frequency vs. beliefs

1. **Frequencies:** capturing variability of physical phenomena through repeated observations.
2. **Belief in unique events :** due to lack of information
    1. via betting on lottery tickets for non-repeatable events
    2. by analogical reasoning using thought experiment (balls in an urn)

Probability theory used for random phenomena, and beliefs.

The connection: Degrees of belief induced on n+1th trial outcome are equated to frequencies of the n previous occurrences of a repeatable phenomenon.

What is a probability given by an expert?

- Does the expert provide
  - An (ill-known) frequency (how often an infortunate event may occur?)
  - A degree of pure belief ?
- Are frequencies and belief degrees always commensurate ?
- Often rather linguistic than numerical, then translated into numbers.

What is the expressive power of probability distributions

Bayesian credo: any state of knowledge can be represented by a unique probability distribution.

Do uniform distributions represent ignorance ?

1. **Ambiguity :** do uniform bets express knowledge of randomness or plain ignorance?
2. **Instability :** the shape of a probability distribution is not scale-invariant, while ignorance is.
3. **Empirical falsification**: When information is missing, decision-makers do not always choose according to a single subjective probability (Ellsberg paradox).

The paradox of partial ignorance

You have the same knowledge about $x > 0$ as about $y = 1/x$.

- *x in $[a, b]$ is equivalent to y in $[1/b, 1/a]$*
- *But a uniform distribution on $[a, b]$ is incompatible with a uniform distribution on $[1/b, 1/a]$*

Conclusion: uniform probability distributions do not represent ignorance.

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* contains all values of x an agent considers not impossible and represent the epistemic state of an agent.
- It is a subset of *mutually exclusive* values, one of which is the real x.
- ***Such sets are as subjective:*** E is like the support of a subjective probability function.

Set-Valued Representations of Partial Information

- (x, E) means « *all I know is that x ∈ E* »
- Examples
  - **Intervals** E = [a, b]: incomplete numerical information

    uncertainty propagation via interval analysis

  - **Classical Logic**: incomplete symbolic information

    E = Models of a proposition p (or a set thereof) believed true.

    being able or not to prove or disprove something from a knowledge base K

POSSIBILITY THEORY: Boolean beliefs

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

- **You believe event A** if A will occur in every situation x you consider possible :***A certainty (necessity) function*** (logical consequence).

  $N(A) = 1$ if $E \subseteq A$, and 0 otherwise

- You judge **event A possible** if it is not incompatible with what you know : ***A Boolean possibility function*** (logical consistency)

  $\Pi(A) = 1$, if $A \cap E \neq \emptyset$ and 0 otherwise

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

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

$$N(A) > 0 \text{ implies } \Pi(A) = 1$$

*(a simple modal **epistemic** logic)*

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* : incomplete information about probabilistic models

Find an extended representation of uncertainty

- *Explicitly allowing for missing information (= that uses sets)*
- *Distinguishes between not believing A and believing its negation*
- *More informative than pure intervals or classical logic: with grades of certainty or belief*
- *Less information demanding than single probability distributions*
- *Allows for addressing the issues dealt with by both standard probability, and logics for reasoning about knowledge.*

GRADUAL REPRESENTATIONS OF UNCERTAINTY using capacities

Family of propositions or events $\mathcal{E}$ forming a Boolean Algebra

- S, Ø are events that are certain and ever impossible respectively.

- **A confidence measure** g: a function from $\mathcal{E}$ to [0,1] such that
  - $g(\emptyset) = 0$; $g(S) = 1$
  - **monotony** : if $A \subseteq B$ (=A implies B) then $g(A) \leq g(B)$
- g(A) quantifies the confidence of an agent in proposition A.

(g is known as a Choquet capacity, or a fuzzy measure)

BASIC PROPERTIES OF CONFIDENCE MEASURES

- **$g(A \cup B) \geq \max(g(A), g(B))$;**
- **$g(A \cap B) \leq \min(g(A), g(B))$**
- It includes:
  - probability measures: $P(A \cup B) = P(A) + P(B) - P(A \cap B)$
  - possibility measures $\Pi(A \cup B) = \max(\Pi(A), \Pi(B))$
  - necessity measures $N(A \cap B) = \min(N(A), N(B))$

- ***The two latter functions do not require a numerical setting***

A GENERAL SETTING FOR REPRESENTING GRADED CERTAINTY AND PLAUSIBILITY

- 2 conjugate set-functions Pl and Cr generalizing probability P, possibility Π, and necessity N.
- **Postulates**
  - Cr and Pl are monotonic under inclusion (= capacities).
  - $Cr(A) \leq Pl(A)$ "certain implies plausible"
  - $Pl(A) = 1 - Cr(A^c)$ duality certain/plausible
  - If Pl = Cr then it is P.
- **Conventions** :
  - $Pl(A) = 0$ "impossible" ; $Cr(A) = 1$ "certain"
  - $Cr(A) = 0$ and $Pl(A) = 1$ "ignorance" **(no information)**
  - $Pl(A) - Cr(A)$ quantifies ignorance about A

Possibility Theory

(Shackle, 1961, Lewis, 1973, L.J. Cohen 1977, 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)

Improving expressivity of incomplete information representations

What about the birth date of the president?

•**partial ignorance with ordinal preferences :** May have reasons to believe that 1933 > 1932 ≡ 1934 > 1931 ≡ 1935 > 1930 > 1936 > 1929

•**Linguistic information** described by fuzzy sets: “ he is old ” : membership μ_{OLD} induces a possibility distribution on possible birth dates.

•**imprecise subjective** information summarizing opinions of one or several sources: Nested intervals E_1, E_2, …E_n with confidence levels

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 some $x \in A$ is possible

  (= to what extent A is consistent with π)

  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

then $N(A) = N(A^c) = 0$

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

A pioneer of possibility theory

- In the 1950's, **G.L.S. Shackle** called "degree of potential surprize" of an event its degree of impossibility = $1 - \Pi(A)$.

- Potential surprize is valued on a disbelief scale, namely a positive interval of the form [0, y*], where y* denotes the absolute rejection of the event to which it is assigned, and 0 means that nothing opposes to the occurrence of A.

- The degree of surprize of an event is the degree of surprize of its least surprizing realization.
- He introduces a notion of conditional possibility

Qualitative vs. quantitative possibility theories

- **Qualitative**:
  - **comparative**: A complete pre-ordering $\geq_\pi$ on U
    A well-ordered partition of U: E1 > E2 > … > En
  - **absolute:** $\pi_x(s) \in L$ = finite chain, complete lattice...
- **Quantitative**: $\pi_x(s) \in [0, 1]$, integers...

One must indicate where the numbers come from.

All theories agree on the fundamental maxitivity axiom

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

Theories diverge on the conditioning operation

POSSIBILITY AS UPPER PROBABILITY

- Given a numerical possibility distribution π, define $\mathbf{P}(\pi) = \{P \mid P(A) \leq \Pi(A) \text{ for all } A\}$

- Then, generally it holds that

  $\Pi(A) = \sup \{P(A) \mid P \in \mathbf{P}(\pi)\}$;

  $N(A) = \inf \{P(A) \mid P \in \mathbf{P}(\pi)\}$

  (coherence)
- So π is a faithful representation of a family of probability measures

LIKELIHOOD FUNCTIONS

- **Likelihood functions** $\lambda(x) = P(A| x)$ behave like possibility distributions when there is no prior on x, and $\lambda(x)$ is used as the likelihood of x.
- If $\lambda(B)$ is the likelihood that $x \in B$ then λ should be set-monotonic: $\{b\} \subseteq B$ implies $\lambda(b) \leq \lambda(B)$
- It holds that $\lambda(B) = P(A| B) \leq \max_{x \in B} P(A| x)$

It implies $\lambda(B) = \max_{x \in B} \lambda(x)$

(*But possibility degrees here are defined up to a positive multiplicative function)*

Maximum likelihood principle is in agreement with possibility theory

- The classical coin example: θ is the unknown probability of “heads”
- Within n experiments: k heads, n-k tails
- P(k heads, n-k tails | θ) = $\theta^{k} \cdot (1- \theta)^{n-k}$ is

the degree of possibility $\pi(\theta)$ that the probability of “head” is θ.

In the absence of other information the best choice is the one that maximizes $\pi(\theta)$, $\theta \in [0, 1]$

It yields $\theta = k/n$.

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

Imprecise probability theory

- A state of information is represented by a family $\mathcal{P}$ of probability distributions over a set X.
- To each event A is attached $[P_*(A), P^*(A)]$, a probability interval such that
  - $P_*(A) = \inf\{P(A), P \in \mathcal{P}\}$
  - $P^*(A) = \sup\{P(A), P \in \mathcal{P}\} = 1 - P_*(A^c)$

$$\mathcal{CP} = \{P, P(A) \geq P_*(A) \text{ for all } A\} \text{ is convex}$$

- Usually $\mathcal{CP}$ strictly contains family $\mathcal{P}$

-> The basic representation tool is a convex set of probabilities (credal set)

Frequentist view

- *Incomplete knowledge of a frequentist probabilistic model* : $\exists P \in \mathcal{P}$.
  - Expert opinion about frequencies (fractiles, intervals with confidence levels)
  - Subjective estimates of support, mode, etc. of a distribution
  - Parametric model with incomplete information on parameters (partial subjective information on mean and variance)
  - Parametric model with confidence intervals on parameters due to a small number of observations

Subjectivist view (Peter Walley)

- Expert provides for selected events A_i, $i = 1, \ldots, n$
  - $P_{low}(A)$, the highest acceptable price for buying a bet on event A winning 1 euro if A occurs
  - $P^{high}(A) = 1 - P_{low}(A^c)$ is the least acceptable price for selling this bet.
  - These prices may differ (no exchangeable bets)
- Epistemic state is modelled by the *convex probability set*
  $\mathcal{P} = \{P: P(A_i) \geq P_{low}(A_i)\ i = 1, \ldots, n\}$

- **Warning** : $P_*(A) = \inf\{P(A), P \in \mathcal{P}\}$ is a degree of belief in A but there is no unknown $P \in \mathcal{P}$

WHY REPRESENTING INFORMATION BY PROBABILITY FAMILIES ?

- In the case of generic (frequentist) information using a family of probabilistic models, rather than selecting a single one, enables to account for incompleteness and variability.
- In the case of subjective belief: distinction between
  - believing neither a proposition nor its opposite ($P_*(A)$ and $P_*(A^c)$ low)
  - and believing its negation ($P_*(A)$ low and $P_*(A^c)$ high).

Random sets and evidence theory

A probability distribution over subsets of S (a random set) :

$$\sum_{E \subseteq S} m(E) = 1 \quad (\textit{mass function}),\ m(\emptyset) = 0$$

- The family $\mathcal{F} = \{E: m(E) > 0\}$ of « focal » (disjunctive) non-empty sets represents
  - A collection of incomplete observations (imprecise statistics).
  - Unreliable testimonies
- m is a randomized epistemic state where
  - m(E) = *probability(*E is the correct epistemic state) (≠ P(E))
    = *probability(only knowing"(x in E)")*
  - m(E) is a probability mass that should be distributed among elements of E but are not by lack of information.

Theory of evidence

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

Canonical examples

- **Objectivist** : Frequentist modelling of a collection of incomplete observations (imprecise statistics) :
- **Uncertain subjective information:**
  - **Merging of unreliable testimonies** (Shafer's book) : human-originated singular information
- **Unreliable sensors** : the quality/precision of the information depends on the ill-known sensor state.

Example of uncertain evidence : Unreliable testimony (SHAFER-SMETS VIEW)

- « John tells me the president is between 60 and 70 years old, but there is some chance (*subjective* probability p) he does not know and makes it up».
  - E =[60, 70]; Prob(Knowing "$x \in E$ =[60, 70]") = 1 − p.
  - With probability p, John invents the info, **so *we know nothing*** *(Note that this is different from a lie)*.
- We get a *simple support belief function :*

$$m(E) = 1 - p \quad \text{and} \quad m(S) = p$$

- Equivalent to a possibility distribution
  - $\pi(s) = 1$ if $x \in E$ and $\pi(s) = p$ otherwise.

Example of statistical belief function: imprecise observations in an opinion poll

- **Question** : who is your preferred candidate in $\mathcal{C} = \{a, b, c, d, e, f\}$???
  - **To a population** $\Omega = \{1, \ldots, i, \ldots, n\}$ of n persons.
  - **Imprecise responses r =** « $x(i) \in E_i$ » **are allowed**
  - No opinion ($r = \mathcal{C}$) ; « left wing » $r = \{a, b, c\}$;
  - « right wing » $r = \{d, e, f\}$;
  - a moderate candidate : $r = \{c, d\}$
- **Definition of mass function**:
  - $m(E) = (1/n)\cdot\text{card}(\{i, E_i = E\})$
  - = Proportion of imprecise responses « $x(i) \in E$ »

- *The probability that a candidate in subset* $A \subseteq C$ *is elected is imprecise :*

$$\mathrm{Bel}(A) \leq P(A) \leq \mathrm{Pl}(A)$$

- **There is a fuzzy set F of potential winners**:

$$\mu_F(x) = \sum_{x \in E} m(E) = \mathrm{Pl}(\{x\}) \text{ (contour function)}$$

- $\mu_F(x)$ is an upper bound of the probability that x is elected. It gathers responses of those who *did not give up voting* for x
- Bel({x}) gathers responses of those who claim they will vote for x and no one else.

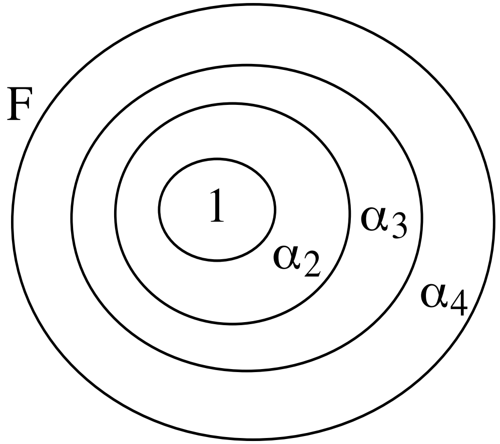
PARTICULAR CASES

- INCOMPLETE INFORMATION:

  $m(E) = 1, m(A) = 0, A \neq E$
- *TOTAL IGNORANCE :* m(S) = 1:
  - *For all* $A \neq S, \emptyset$, $Bel(A) = 0$, $Pl(A) = 1$
- PROBABILITY: if $\forall i, E_i$ = singleton $\{s_i\}$ (hence disjoint focal sets)
  - Then, *for all* A, Bel(A) = Pl(A) = P(A)
  - *Hence precise + scattered information*
- POSSIBILITY THEORY : the opposite case

  $E_1 \subseteq E_2 \subseteq E_3 \ldots \subseteq E_n$: imprecise and coherent information
  - iff $Pl(A \cup B) = \max(Pl(A), Pl(B))$, possibility measure
  - iff $Bel(A \cap B) = \min(Bel(A), Bel(B))$, necessity measure

Possibility theory case



possibility levels
$1 > \alpha_2 > \alpha_3 > \ldots > \alpha_n$

- Let $m_i = \alpha_i - \alpha_{i+1}$ then $m_1 + \ldots + m_n = 1$,
  with focal sets = cuts $A_i = \{s, \pi(s) \geq \alpha_i\}$

  A basic probability assignment (SHAFER)
- $\pi(s) = \sum_{i:\, s \in Fi} m_i$ (one point-coverage function) = Pl({s}).
- ***Only in the consonant case can m be recalculated from π***
- $Bel(A) = \sum_{Fi \subseteq A} m_i = N(A); Pl(A) = \Pi(A)$

Theory of evidence vs. imprecise probabilities

- Bel is ∞-monotone (super-additive at any order)
- Bel is a special case of lower probability
  - The set $\mathcal{P}_{bel} = \{P \geq Bel\}$ characterizes Bel:

$$Bel(A) = \inf \{P(A) \mid P(B) \geq Bel(B) \text{ for all } B\}$$

- The solution m to the set of equations $\forall A \subseteq X$

$$g(A) = \sum_{E_i \subseteq A,\, E_i \neq \emptyset} m(E_i)$$

is unique (Moebius transform)

  - **It is positive iff g is a belief function**

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

Language difficulties

- *Imprecise probability, belief functions and possibility theory use different basic tools*
  - *Imprecise probabilities: Convex probability sets (Credal sets)*
  - *Belief functions: Moebius basic probability mass*
  - *Possibility theory: Possibility distributions*
- *Concepts that make sense for credal sets, may be hard to interpret in terms of Moebius transforms or possibility distributions and conversely*

Practical representations

- Fuzzy intervals
- Probability intervals
- Probability boxes

Some are special random sets some not.

Simplified representations help us

- *cut down computation costs*
- *Facilitate elicitation*
- *summarize results in a clear way*

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)

From confidence sets to possibility distributions

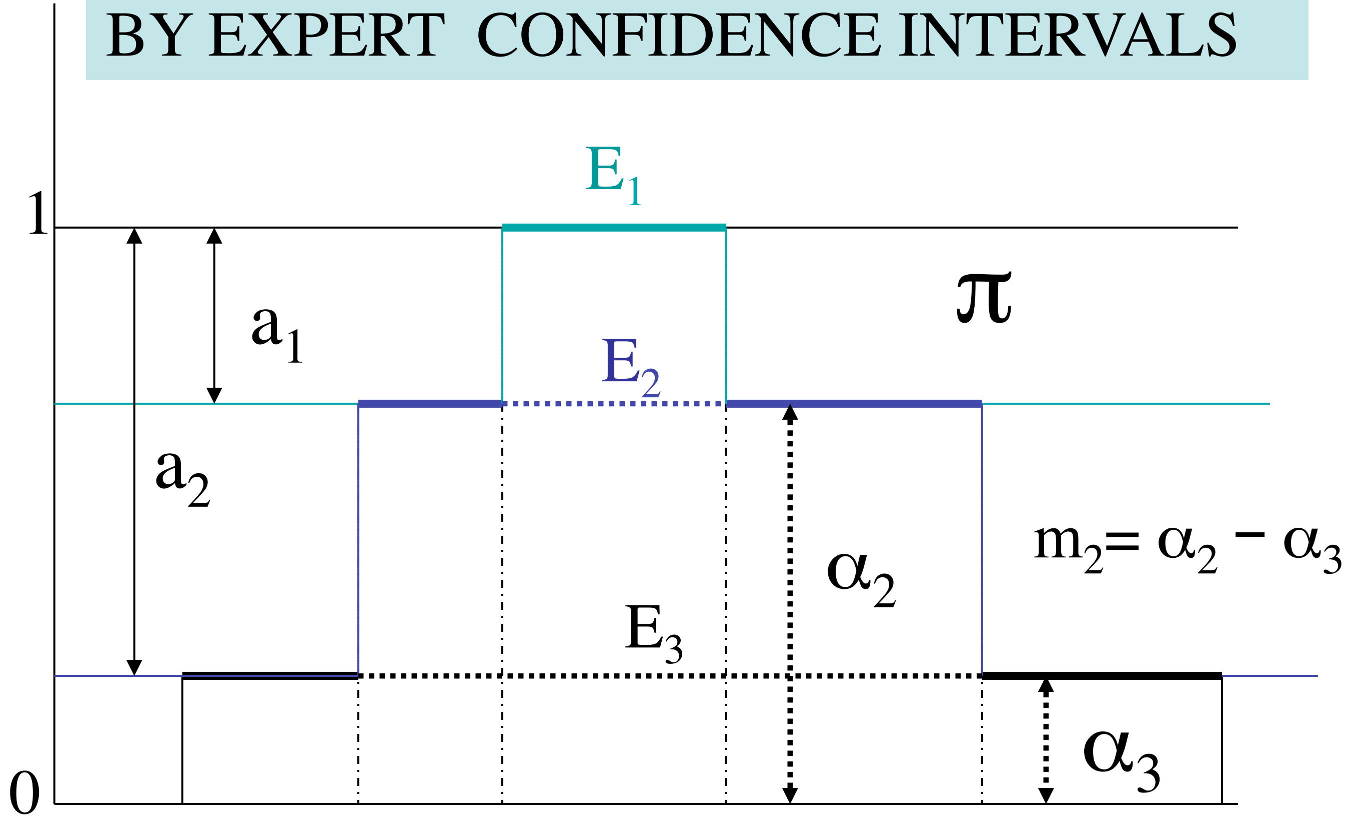
- Let $E_1, E_2, \ldots E_n$ be a nested family of sets
- A set of confidence levels $a_1, a_2, \ldots a_n$ in $[0, 1]$
- Consider the credal set

$$\mathcal{P} = \{P, P(E_i) \geq a_i, \text{ for } i = 1, \ldots n\}$$

- Then $\mathcal{P}$ is representable by means of a possibility measure with distribution

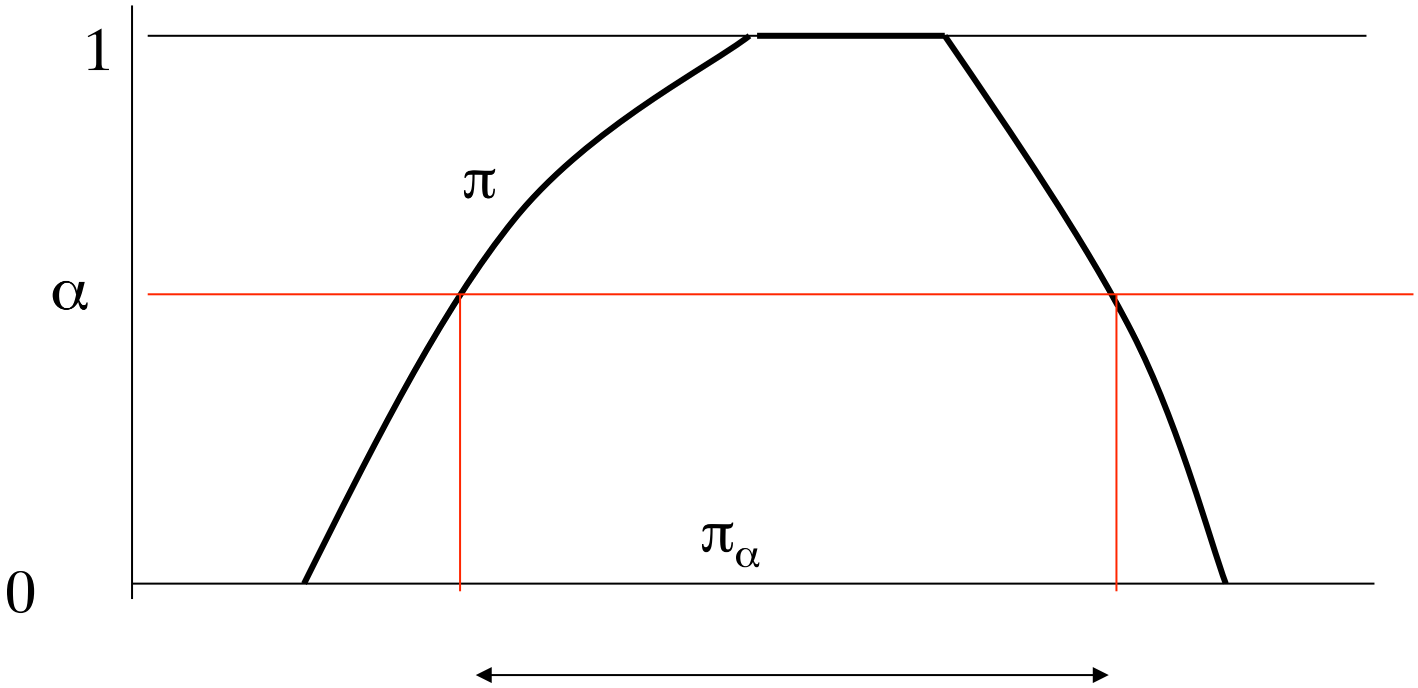
$$\pi(x) = \min_{i = 1, \ldots n} \max(\mu_{Ei}(x), 1 - a_i)$$

POSSIBILITY DISTRIBUTION INDUCED BY EXPERT CONFIDENCE INTERVALS



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

$$\mathbf{P}(\pi) = \{P \mid P(\pi_\alpha) \geq 1 - \alpha, \alpha \in (0, 1]\}$$



FUZZY INTERVAL: $N(\pi_\alpha) = 1 - \alpha$

Possibilistic view of probabilistic inequalities

They can be used for knowledge representation

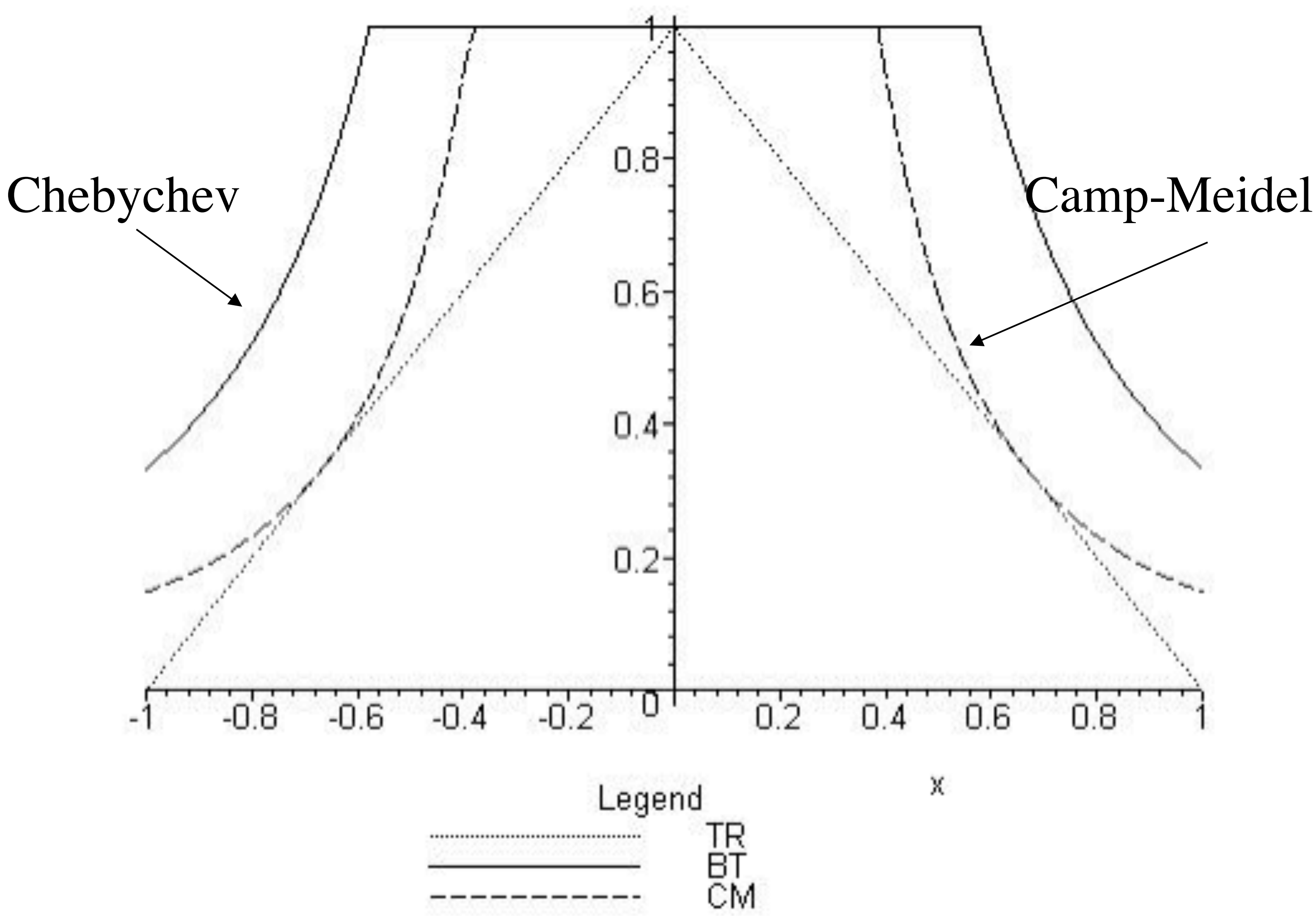
- Chebyshev inequality defines a possibility distribution that dominates *any* density with given mean and variance:

$$P(V \in [x^{mean} - k\sigma, x^{mean} + k\sigma]) \geq 1 - 1/k^2$$

is equivalent to writing

$$\pi(x^{mean} - k\sigma) = \pi(x^{mean} + k\sigma) = 1/k^2$$

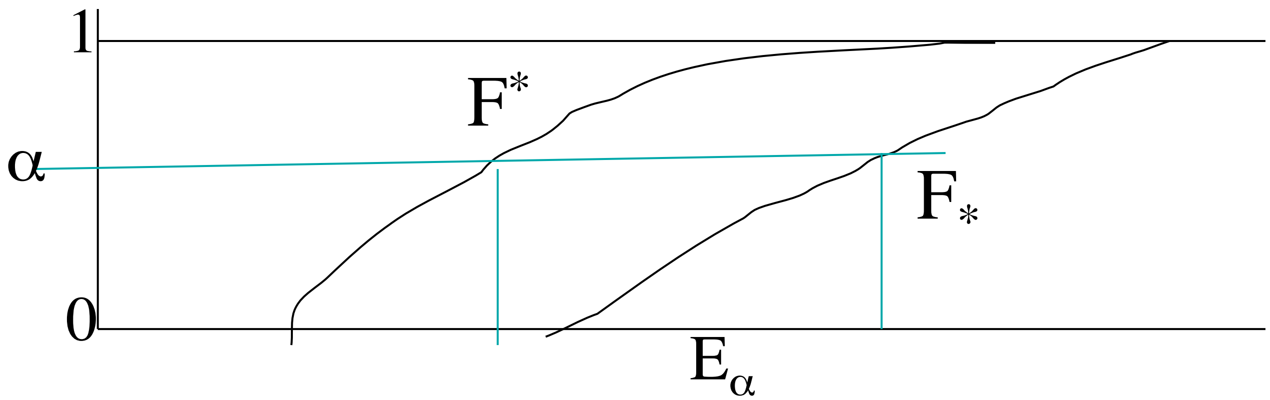
- A triangular fuzzy number (TFN) defines a possibility distribution that dominates *any* unimodal density with the same mode and bounded support as the TFN.



Chebychev
Camp-Meidel
1
0.8
0.6
0.4
0.2
0
-1
-0.8
-0.6
-0.4
-0.2
0.2
0.4
0.6
0.8
1
x
Legend
TR
BT
CM

Probability boxes

- A set $\mathcal{P}(F^*, F_*) = \{P: F^* \geq P \geq F_*\}$ 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.**



Probability boxes from possibility distributions

- *fuzzy intervals are more precise than with the corresponding pairs of PDFs:*
  - $F^*(a) = \Pi_M((-\infty, a]) = \pi(a)$ if $a \leq m_*$
    $= 1$ otherwise.
  - $F_*(a) = N_M((-\infty, a]) = 0$ if $a < m^*$
    $= 1 - \pi(a)$ otherwise
- $\mathcal{P}(\pi)$ is a proper subset of $\mathcal{P}(F^*, F_*)$: Not all P in $\mathcal{P}(F^*, F_*)$ are such that $\Pi \geq P$
- In fact you can extract a p-box from any credal set $\mathcal{P}$

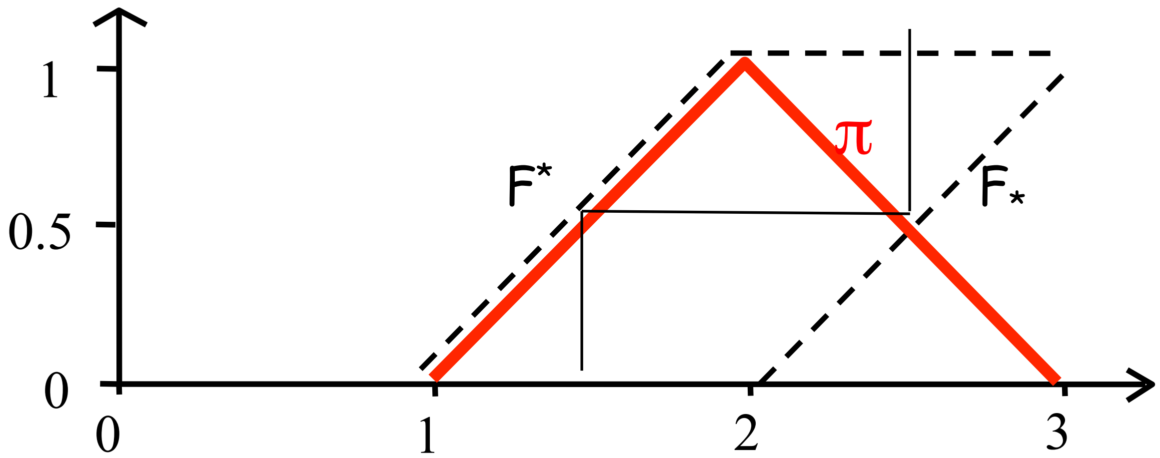
P-boxes vs. fuzzy intervals

A triangular fuzzy number with support [1, 3] and mode 2.

Let P be defined by $P(\{1.5\})=P(\{2.5\})=0.5$.

Then $F_* < F < F$, but $P \notin \mathcal{P}(\Pi)$

since $P(\{1.5, 2.5\}) = 1 > \Pi(\{1.5, 2.5\}) = 0.5$



Cumulative distributions are possibility distributions

- A cumulative distribution F is a possibility distribution generated by nested sets of the form $[x, +\infty)$, enclosing all probability distributions that stochastically dominate F.

If $\pi = F$, then

$$\mathcal{P}(\pi) = \{P: F_p \leq F\} = \{p: P([x, +\infty)) \geq 1 - F(x)\}$$

$$= \{P: P(F \geq \alpha) \geq 1 - \alpha, \alpha > 0\}$$

and we have that $P(A) \leq \sup_{x \text{ in } A} F(x)$.

Fuzzy intervals are (2-sided) cumulative distributions

- Consider a fuzzy interval π with cuts

$$\pi_\alpha = [a_\alpha, b_\alpha],\ 0 < \alpha \leq 1,\ a_1 = b_1 = m$$

It is a cumulative distribution in the sense that

$$\pi_\alpha(a_\alpha) = \pi_\alpha(b_\alpha) = P((-\infty, a_\alpha] \cup [b_\alpha, +\infty))$$

for some probability measure P with mode m.

Ordering based on distance from m.

Putting together p-boxes and fuzzy intervals

- The credal set of a p-box (F^*, F_*) is the intersection of possibilistic credal sets of $\pi^* = F^*$ and $\pi_* = 1 - F_*$:

  $$\mathcal{P}(F^*, F_*) = \{p: F_* \leq F_p \leq F^*\} = \mathcal{P}(F^*) \cap \mathcal{P}(1 - F_*)$$

  $$= \{P: P([x, +\infty)) \geq 1\text{-}F^*(x) \text{ for all } x \text{ and } P((-\infty, x]) \geq F_*(x) \text{ for all } x\}$$

  $$= \{P: P(F^* \geq \alpha) \geq 1\text{-}\alpha > P(F_* \geq \alpha) \text{ for all } 0 < \alpha \leq 1\}$$

  where $\alpha = F(x)$.

- F^*, F_* are comonotone

Generalized p-box

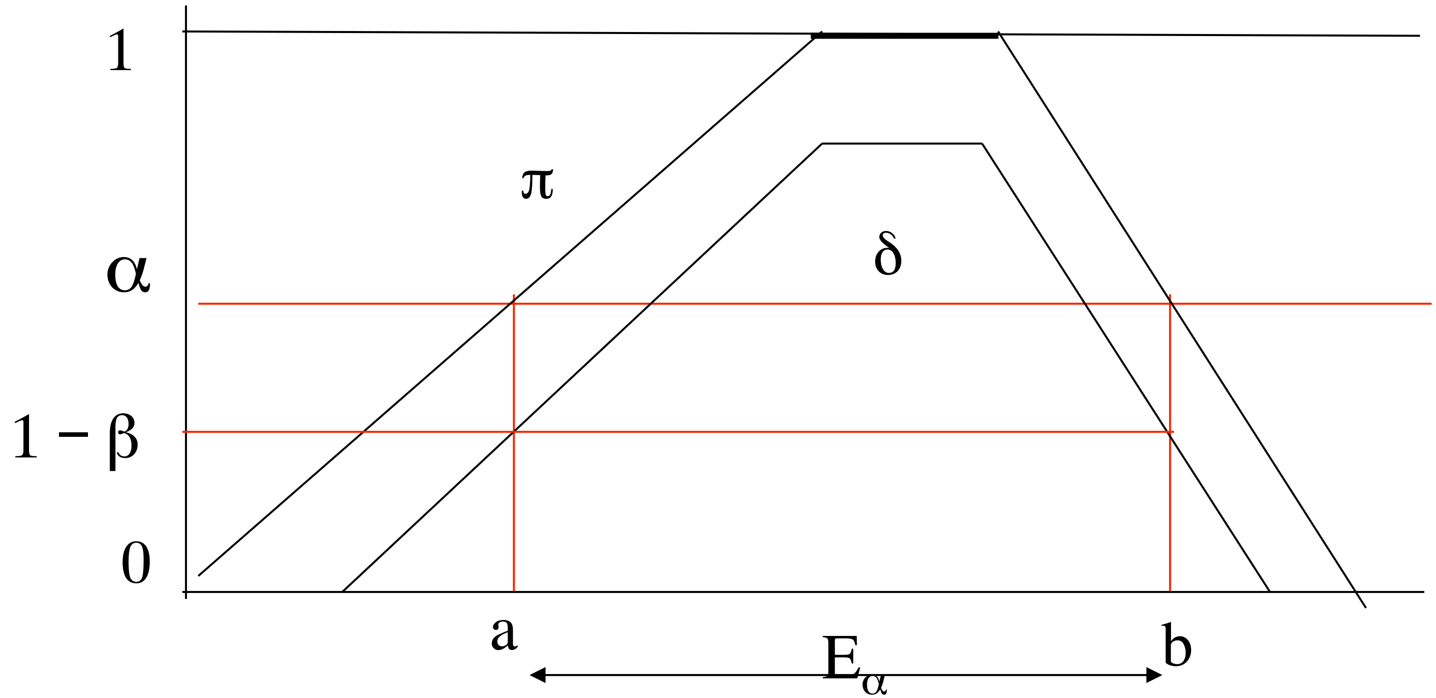
- *same construction using nested intervals and comonotone functions* $\delta \leq \pi$ *such that* $1-\delta$ *is a possibility distribution*.
- The pair (π, δ) is a generalized p-box with credal set $\mathcal{P}(\pi, \delta) = \mathcal{P}(\pi) \cap \mathcal{P}(1-\delta)$

with $\mathcal{P}(\pi) = \{P: P(\pi \geq \alpha) > 1 - \alpha, 0 < \alpha \leq 1\}$

$\mathcal{P}(1-\delta) = \{P: P(1-\delta \geq \alpha) > 1 - \alpha, 0 < \alpha \leq 1\}$

It still generates a belief function!

$$\alpha = \pi(a) = \pi(b);$$
$$\beta = 1 - \delta(a) = 1 - \delta(b) = 1 - \delta(\pi^{-1}(\alpha)).$$
$$1 - \alpha \leq P(E_\alpha) \leq \beta$$



Generalized p-box

Examples, special cases, etc.

- Nested confidence sets E_i with $a_i \leq P(E_i) \leq b_i$
- Z-numbers (Zadeh): *It is likely that I earn a lot*
- **Special cases**
  - $\pi = F^*$, $\delta = F_*$: pbox.
  - $\delta = 0$: fuzzy interval.
  - $\pi = \delta$: thin cloud (Neumaier)
- **Extension** : (π, δ) non comonotone: cloud of Neumaier (not a belief function).

From generalized p-boxes to clouds

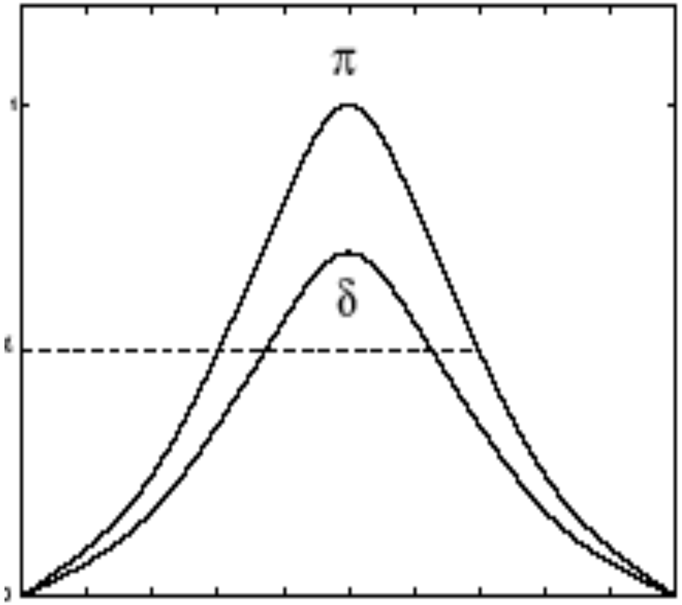


Fig 1.A Comonotonic cloud

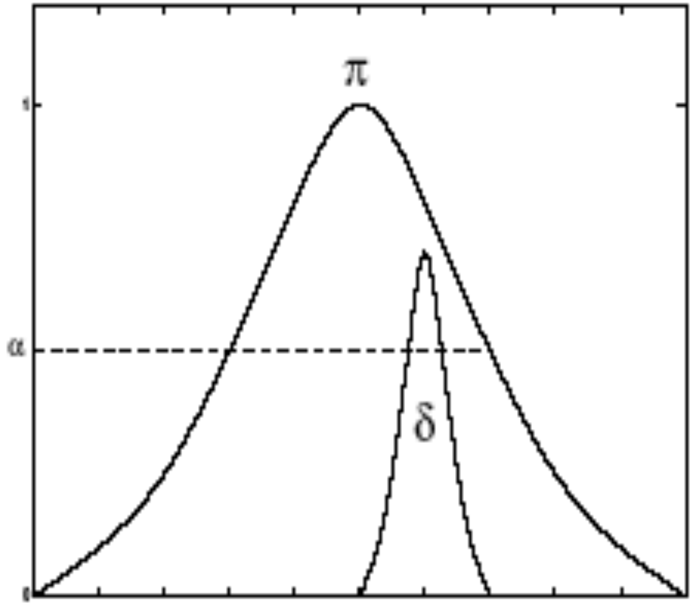


Fig 1.B Non-comonotonic cloud

How useful are these representations:

- P-boxes can address questions about threshold violations ($x \geq a$??),

  not questions of the form $a \leq x \leq b$

- The latter questions are better addressed by possibility distributions or generalized p-boxes around a specific value.

Probability intervals

- Probability intervals = a finite collection of imprecise assignments $[l_i , u_i]$ attached to elements s_i of a finite set S.
  - The collection $\{[l_i , u_i]\ i = 1,\ldots n\}$ induces the family $\mathcal{P}_L = \{P\colon l_i \leq P(\{s_i\}) \leq u_i\}$.
- Intervals $[l_i , u_i]$ can be made optimally narrow.
- Lower/upper probabilities on events are easy to compute
- P_* is a 2-monotone Choquet capacity, not a belief function.

Application to Risk Analysis

- **Formal problem**:

  Given a numerical function $f(x, y, z, \ldots)$, and some uncertain knowledge on $x, y, z, \ldots$ interval, possibilistic (π_x), probabilistic (p_y) or random set-like (v_z)… find the resulting uncertainty on $f(x, y, z, \ldots)$.

- **Application Contexts:** Evaluation of risks of potentially polluted sites for man and the environment
- Models simulate the transfer of pollutants from a source to a vulnerable target, for different scenarii of exposure.

Risk analysis methodology

- Elicitation/ data collection for inputs
- Propagation of uncertainty
- Exploitation of results
- Decision

Risk analysis methodology : elicitation

The context of uncertainty theories is versatile and lends itself to a representation of knowledge about input variables faithful to what is available.

Don't put more information than what you actually have

- sufficient statistics: probability distribution
- Ill-known parametric model: p-box
- Expert-supplied intervals: fuzzy intervals, gen p-box
- Support and mode: fuzzy interval

Risk analysis methodology : propagation

Combining Monte-Carlo and interval analysis techniques.

- Fuzzy intervals, p-boxes and generalized p-boxes are random sets amenable to Monte-Carlo methods:

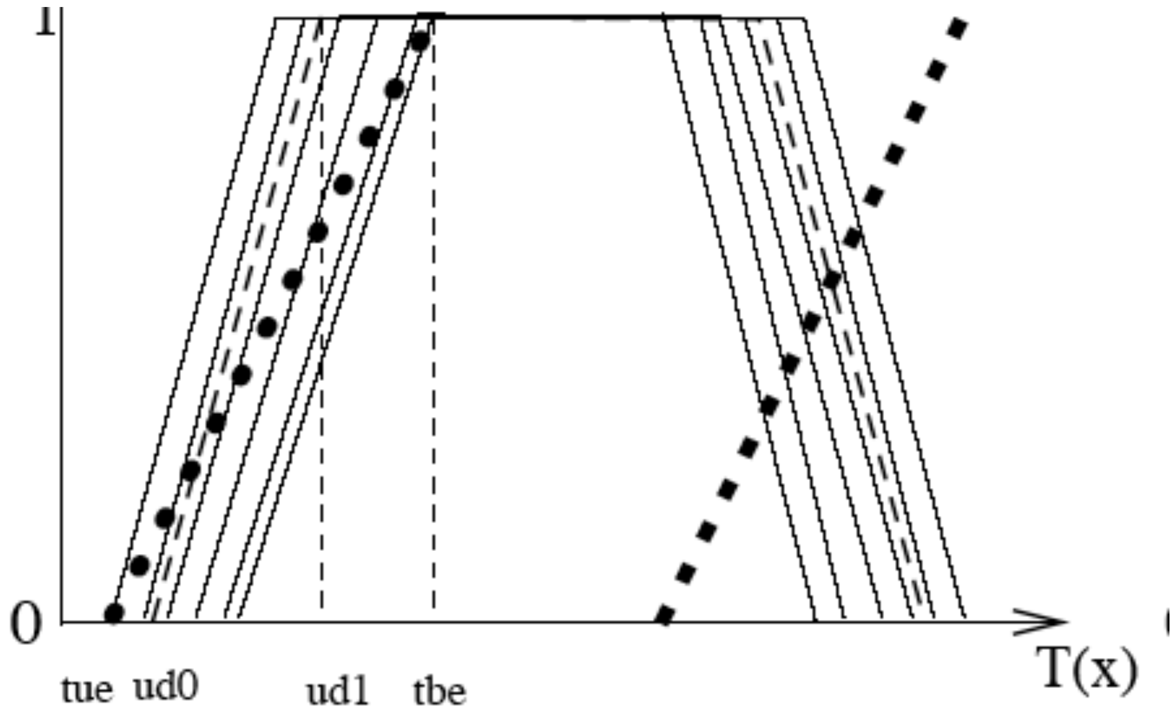
- Instead of picking values at random via the cumulative distribution, pick intervals (cuts) and perform interval analysis

Risk analysis methodology : exploitation

The result of the propagation step is a random set on the output value, that can be complex to visualize.

- We can extract suitable information
  - Imprecise mean and variance
  - Average imprecision
  - A p-box (probability of trespassing a threshold)
  - A fuzzy interval (probability of the output inside two bounds)

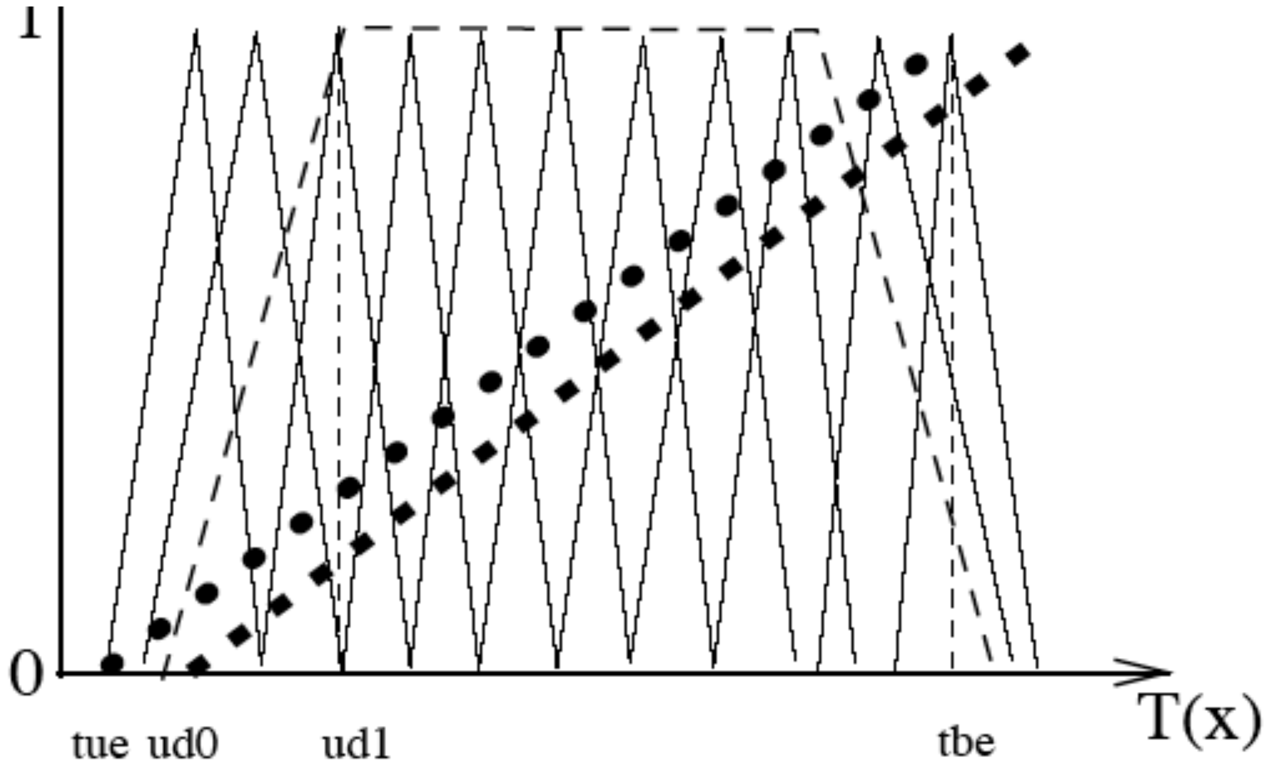
Upper and lower distributions of random fuzzy outputs



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

Decision with imprecise probability techniques

- Decisions will be evaluated by means of intervals bounded by lower and upper expected utilities:
- $V(f) = [\inf_{P \text{ in } \mathcal{P}} E(f), \sup_{P \text{ in } \mathcal{P}} E(f)]$
- We are left to compare intervals…
- **Three-way decisions**: yes, no, don't know

Decision with imprecise probability techniques

- Accept incomparability when comparing imprecise utility evaluations of decisions.

OR

- Select a single utility value that achieves a compromise between pessimistic and optimistic attitudes.
  - Compare lower expectations of decisions (Gilboa):
    $\inf_{P \text{ in } \mathcal{P}} E(f) > \inf_{P \text{ in } \mathcal{P}} E(g)$
  - Generalize Hurwicz criterion
  - Select a single probability measure (Shapley value = pignistic transformation) and use expected utility (SMETS)

Conclusion

- *There exist 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
- *Allow to explicitly encode incomplete knowledge*.
- *How to get this general non-dogmatic approach to uncertainty accepted by traditional statisticians?*

Important theoretical issues

- Comparing representations in terms of **informativeness**.
- **Conditioning** : several definitions for several purposes.
- **Independence notions**: distinguish between epistemic and objective notions.
- Find a general setting for **information fusion** operations (e.g. Dempster rule of combination).