Imprecise probabilities to propagate uncertainties: a tour

Sébastien Destercke

Heudiasyc, CNRS Compiegne, France

ETICS 2017 school

Lecture goal/content

What you will find in this talk

- Information representation seen by a non-statistician (mostly IA/engineer) guy
- Imprecise probabilities: when to (not) use it?
- Imprecise probabilities: definition and practical representation
- Merging imprecise probabilistic representations
- (In)dependence modelling and uncertainty propagation
- How (not) to decide with imprecise probabilities

What you will not find in this talk

- A deep and exhaustive study of a particular topic

Outline

Generic vs singular quantity

A quantity of interest X can be

- **Generic**, when it refers to a population, or a set of situations.

Generic quantity example

The distribution of mother tongue within French inhabitants

- **Singular**, when it refers to an individual or a peculiar situation

Singular quantity example

My own mother tongue

Ontic and epistemic information [9]

An item of information $\mathcal{I}$ possessed by an agent about X can be

- **Ontic**, if it is a faithful, perfect representation of X

Ontic information example

A set X representing the exact set of languages spoken by me
e.g.: $X = \{French, English, Spanish\}$

- **Epistemic**, if it is an imperfect representation of X

Epistemic information example

A set E containing my mother tongue
e.g., $E = \{French, Dutch, English\}$

Everything is possible

We can have

- **Ontic** information about a **singular** quantity: the hair colour of a suspect; the mother tongue of someone
- **Epistemic** information about a **singular** quantity: the result of the next dice toss; the set of possible mother tongues of someone
- **Ontic** information about a **generic** quantity: the exact distribution of pixel colours in an image
- **Epistemic** information about a **generic** quantity: an interval about the frequency of French persons higher than 1.80 m

Uncertainty definition

Uncertainty: when our information $\mathcal{I}$ about the quantity of interest X is insufficient to answer with certainty to assertions

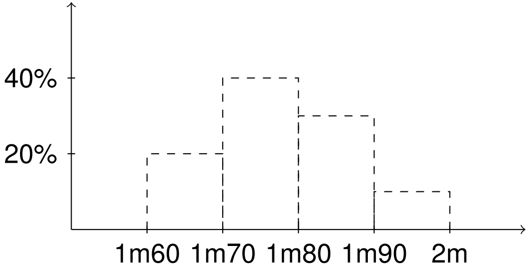
$\rightarrow$ **In this view, uncertainty is necessarily epistemic, as it reflect an imperfect knowledge of the agent.**

Can concern both:

- Singular quantity
  - items in a data-base, values of some logical variables, time before failure of **a** component
- Generic quantity
  - parameter values of classifiers/regression models/probability distributions, time before failure of component**s**, truth of a logical sentence ("birds fly")

The room example

Heights of people in a room: generic quantity



- Generic question: are 90% of people in room less than 1m80?
  $\Rightarrow$ No, with **full certainty**
- Specific question: is the last person who entered less than 1m80?
  $\Rightarrow$ Probably, with 60% chance (**uncertain answer**)

Uncertainty main origins [6, Ch. 3]

- **Variability** of a population applied to a peculiar, singular situation

Variability example

The result of one dice throw when knowing the probability of each face

- **Imprecision and incompleteness** due to partial information about the quantity S

Imprecision example

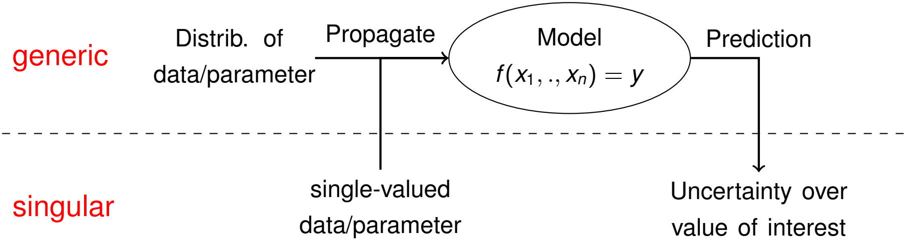
Observing limited sample of the population, describing suspect as "young", limited sensor precision

- **Conflict and unreliability** of different sources of information

Conflict example

Two redundant data base entries with different information for an attribute, two sensors giving different measurements of a quantity

Uncertainty propagation revisited



Output genericity: same as most generic input variable/parameter

Propagation: usual steps

1. **Represent**: provide an uncertainty model for $x_1, \ldots, x_n$
2. **Merge**: if multiple models given for x_i, merge into a single one
3. **Combine**: specify (in)dependencies between x_i's to get global model
4. **Propagate**: propagate to get uncertainty over y
5. **Decide**: once uncertainty on y estimated, decide on an action

Generic vs singular: why bother?

Many notions making sense for generic quantities, make no or poor sense at all for singular ones:

- frequencies and "objective" true probability
- any statistic requiring population (variance, mean, median, . . .)
- learning from samples
- stochastic independence

Mathematically equivalent notions may model something about your knowledge of the singular quantity, not about the quantity itself

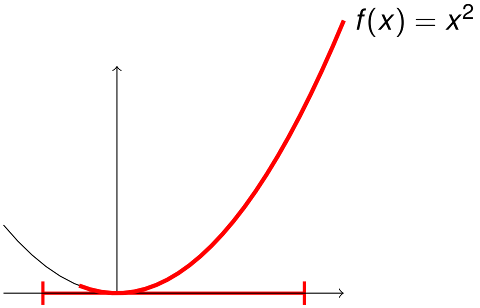
Outline

Outline

A non-probabilistic example

Assume the following:

- A function linking y and x with $f(y) = x^2$
- We want to estimate $f(y)$ but only know $x \in [-1, 6]$

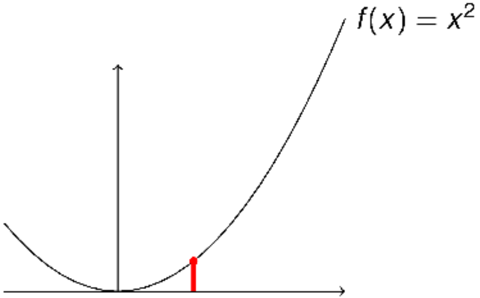


- We acknowledge our imprecise knowledge
- Our final answer is that $f(x) \in [0, 36]$

Full imprecise knowledge + "uniform" selection

Assume the following:

- A function linking y and x with $f(y) = x^2$
- We want to estimate $f(y)$ but only know $x \in [-2, 5]$

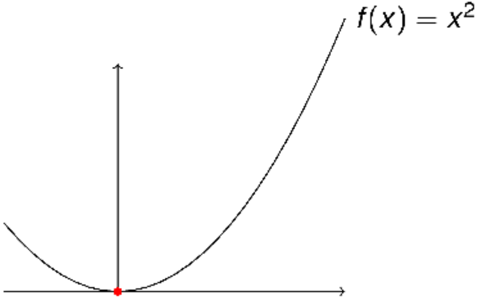


- We choose an "equiprobable" guess given x interval: $x^* = 2$
- Our final answer is that $f(x^*) = 4 \rightarrow$ is it what we want?

Full imprecise knowledge + "uniform" selection

Assume the following:

- A function linking y and x with $f(y) = x^2$
- We want to estimate $f(y)$ but only know $x \in [-2, 5]$



- We choose the worst case given x interval: $x^w = 0$
- Our final answer is $f(x^w) = 0 \rightarrow$ not easy to find? what we want?

Lesson from example

Two strategies:

1. take account of our knowledge as faithfully as possible
2. reduce it to something more manageable:
   - +: may make computations easier (not always)
   - -: selection will introduce a (possibly wanted) bias, whatever it is
   - -: "reference" point (uniform) may induce an unwanted bias

If you are fine with option 2, you can go for it. Another strategy:

3. Outer-approximate initial information for computational convenience

$$\Rightarrow$$

same remarks apply when a probability is ill-known

Outline

Imprecise probabilities for generic information

Probabilities as frequencies

$P(A) =$ frequency with which A has been observed/is observed

Imprecise probabilities as robust/sensitivity analysis models:

- "true" P only known to belong to some set $\mathcal{P}$
- $P(A)$ only known to lie in $[\underline{P}(A), \overline{P}(A)]$
- imprecise observations, limited sample, expert bounds

Eventually, with enough information, get to P or a small $\mathcal{P}$

Imprecise probabilities for singular information

Probabilities as subjective degrees

$P(A) =$ degree of belief that the true value will be in A

Imprecise probabilities as models of beliefs:

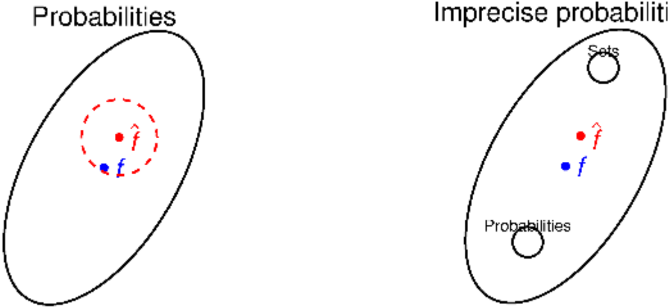
- validity of probability to model partial belief or ignorance questionable
- separate notions of certainty and plausibility to encode ignorance
- asking for a precise P very demanding
- no notion of "true" P within $\mathcal{P}$

Eventually, with enough information, get the true value

Two views of imprecise probabilities

f: true or "ideal" uncertainty model
$\hat{f}$: estimated model/representation

The robust/sensitivity view | The richer model view



- Probability sufficient in theory
- Hard to precisely obtain in practice

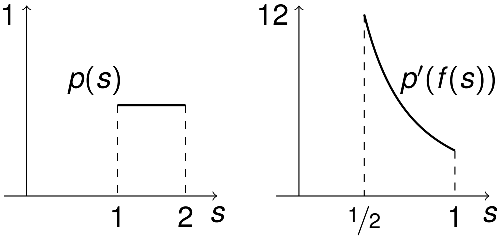
- Probabilities not universal
- Accurate modelling may require richer theory

Outline

Probability as a model of (partial) ignorance

The assumption ignorance=uniform probability has some issues

- Assume we know nothing about $S \in [1,2]$, then ignorance is $p(s) \sim \mathcal{U}[1,2]$
- Yet, if we consider the variable $1/s$, change of variable induce non-uniform probability over $[1/2, 1]$



→ "mathematically right", but model of ignorance should be insensitive to variable changes

The possibility of incomparability

Given two events A, B, whatever this event:

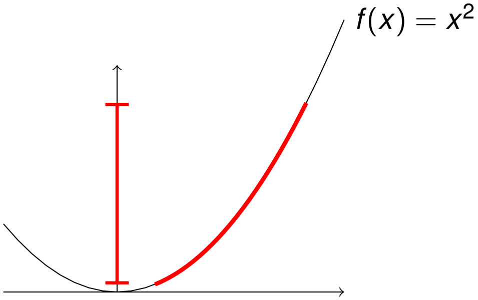
- a probabilistic model P will always output
  - $P(A) > P(B)$
  - $P(A) < P(B)$
  - $P(A) = P(B)$ (not possible for every pair A, B, though)
- in the case of $\mathcal{P}$, you can end up with

$$A >< B \text{ if } [\underline{P}(A), \overline{P}(A)] \cap [\underline{P}(B), \overline{P}(B)]$$

As a direct consequence of lack of knowledge (rather than derive it through a detour $\rightarrow$ variance/sensitivity)

Imprecision in input $\neq$ in outputs

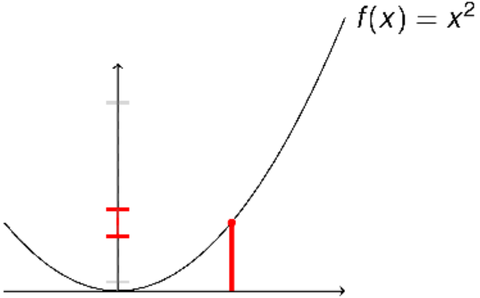
Assume x best guess is 3, $\pm 2 \Rightarrow$ get $[f(x)]$



Propagating then adding imprecision $\neq$ propagating imprecision

Imprecision in input $\neq$ in outputs

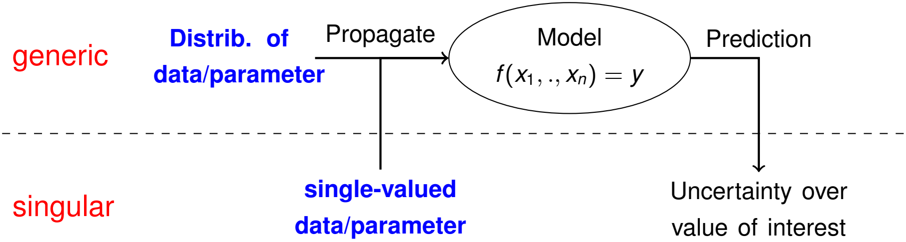
Assume x best guess is 3 $\Rightarrow$ get $f(x) \pm 2$



Propagating then adding imprecision $\neq$ propagating imprecision

Outline

Uncertainty propagation revisited



Output genericity: same as most generic input variable/parameter

Propagation: usual steps

1. **Represent: provide an uncertainty model for $x_1, \ldots, x_n$**
2. **Merge**: if multiple models given for x_i, merge into a single one
3. **Combine**: specify (in)dependencies between x_i's to get global model
4. **Propagate**: propagate to get uncertainty over y
5. **Decide**: once uncertainty on y estimated, decide on an action

Outline

Basic framework

Quantity S with possible **exclusive** states $\mathcal{S} = \{s_1, \ldots, s_n\}$

▷ $\mathcal{S}$: input variable, component state, model parameter, ...

Basic tools

A confidence degree $P : 2^{|\mathcal{S}|} \to [0, 1]$ is such that

- $P(A)$: confidence $S \in A$
- $P(\emptyset) = 0$, $P(\mathcal{S}) = 1$
- $A \subseteq B \Rightarrow P(A) \leq P(B)$

Uncertainty modelled by 2 degrees $\underline{P}, \overline{P} : 2^{|\mathcal{S}|} \to [0, 1]$:

- $\underline{P}(A) \leq \overline{P}(A)$ (monotonicity)
- $\underline{P}(A) = 1 - \overline{P}(A^c)$ (duality)

Probability

Basic tool

A probability distribution $p : \mathcal{S} \to [0, 1]$ from which

- $\underline{P}(A) = \overline{P}(A) = \mu(A) = \sum_{s \in A} p(s)$
- $P(A) = 1 - P(A^c)$: auto-dual

Main interpretations

- **Frequentist [37] :** $P(A)$= number of times A observed in a population
  - ▷ only applies to generic quantities (populations)
- **Subjectivist [24] :** $P(A)$= price for gamble giving 1 if A happens, 0 if not
  - ▷ applies to both singular and generic quantities

Sets

Basic tool

A set $E \subseteq \mathcal{S}$ with true value $S \in E$ from which

- $E \subseteq A \rightarrow \underline{P}(A) = \overline{P}(A) = 1$ (certainty truth in A)
- $E \cap A \neq \emptyset, E \cap A^c \neq \emptyset \rightarrow \underline{P}(A) = 0, \overline{P}(A) = 1$ (ignorance)
- $E \cap A = \emptyset \rightarrow \underline{P}(A) = \overline{P}(A) = 0$ (truth cannot be in A)

$\underline{P}, \overline{P}$ are binary $\rightarrow$ limited expressiveness

Classical use of sets:

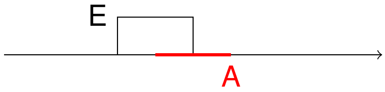
- Interval analysis [26] (E is a subset of $\mathbb{R}$)
- Propositional logic (E is the set of models of a KB)

Other cases: robust optimisation, decision under risk, ...

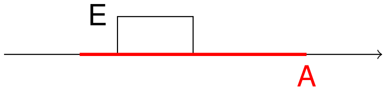
Example

Assume that it is known that pH value $E \in [4.5, 5.5]$, then

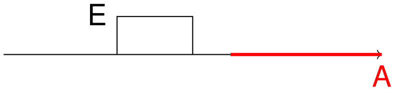
- if $A = [5, 6]$, then $\underline{P}(A) = 0, \overline{P}(A) = 1$



- if $A = [4, 7]$, then $\underline{P}(A) = \overline{P}(A) = 1$



- if $A = [6, 9]$, then $\underline{P}(A) = \overline{P}(A) = 0$



In summary

Probabilities . . .

- (+) very informative quantification (do we need it?)
- (-) need lots of information (do we have it?)
- (-) if not enough, requires a choice (do we want to do that?)
- use probabilistic calculus (convolution, stoch. independence, . . .)

Sets . . .

- (+) need very few information
- (-) very rough quantification of uncertainty (Is it sufficient for us?)
- use set calculus (interval analysis, Cartesian product, . . .)

→ Need **representations bridging these two**

Outline

Possibility distributions

Basic tool

A distribution $\pi : \mathcal{S} \to [0, 1]$, usually with s_i such that $\pi(s_i) = 1$, from which

- $\overline{P}(A) = \max_{s \in A} \pi(s)$
- $\underline{P}(A) = 1 - \overline{P}(A^c) = \min_{s \in A^c}(1 - \pi(s))$

Interval/set as special case

The set E can be modelled by the possibility distribution π_E such that

$$\pi_E(s) = \begin{cases} 1 & \text{if } s \in E \\ 0 & \text{else} \end{cases}$$

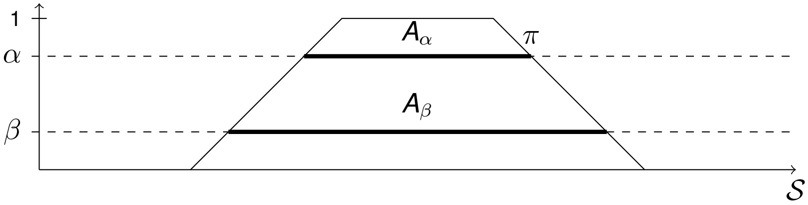
A nice characteristic: Alpha-cut [10]

Definition

$$A_\alpha = \{s \in \mathcal{S} | \pi(s) \geq \alpha\}$$

- $\underline{P}(A_\alpha) = 1 - \alpha$
- If $\beta \leq \alpha$, $A_\alpha \subseteq A_\beta$

Simulation: draw $\alpha \in [0,1]$ and associate A_α



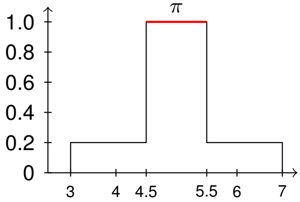
$\Rightarrow$ Possibilistic approach ideal to model **nested structures**

A basic distribution: simple support

- A set E of most plausible values
- A confidence degree $\alpha = \underline{P}(E)$
- Interesting case:
  - Expert providing most plausible values E
- Extend to multiple sets $E_1, \ldots, E_p$:
  - confidence degrees over nested sets [32]

pH value $\in [4.5, 5.5]$ with

$\alpha = 0.8$ ($\sim$ "quite probable")



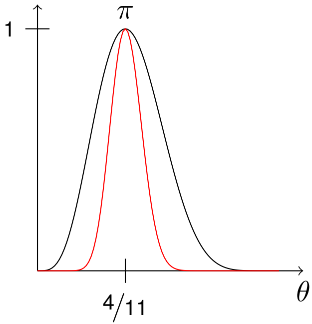
Normalized likelihood as possibilities [20] [7]

$$\pi(\theta) = {}^{\mathcal{L}(\theta|x)}/{\max_{\theta\in\Theta} \mathcal{L}(\theta|x)}$$

Binomial situation:

- θ = success probability
- x number of observed successes

- x= 4 succ. out of 11
- x= 20 succ. out of 55



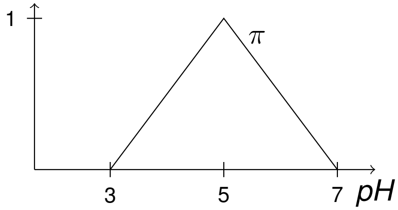
Partially specified probabilities [2] [18]

Triangular distribution: $[\underline{P}, \overline{P}]$ encompass all probabilities with

- mode/reference value M
- support domain $[a, b]$.

Getting back to pH

- $M = 5$
- $[a, b] = [3, 7]$



Other examples

- Statistical inequalities (e.g., Chebyshev inequality) [18]
- Linguistic information (fuzzy sets) [15]
- Approaches based on nested models

Possibility: limitations

For a given event A, we can only have

$$\underline{P}(A) > 0 \Rightarrow \overline{P}(A) = 1$$

$$\overline{P}(A) < 1 \Rightarrow \underline{P}(A) = 0$$

$\Rightarrow$ interval $[\underline{P}(A), \overline{P}(A)]$ either

- $[\alpha, 1]$ or
- $[0, \beta]$,

Hence cannot model any $[\underline{P}(A), \overline{P}(A)]$ with $\underline{P}(A) = \overline{P}(A)$

Possibility distributions **do not include** probabilities as special case.

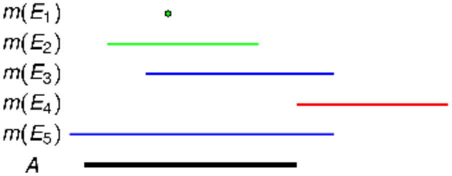
Outline

Random sets and belief functions

Basic tool

A positive distribution $m : 2^{|S|} \to [0, 1]$, with $\sum_E m(E) = 1$ and usually $m(\emptyset) = 0$, from which

- $\overline{P}(A) = \sum_{E \cap A \neq \emptyset} m(E)$
- $\underline{P}(A) = \sum_{E \subseteq A} m(E) = 1 - \overline{P}(A^c)$



$\underline{P}(A) = m(E_1) + m(E_2)$

$\overline{P}(A) = m(E_1) + m(E_2) + m(E_3) + m(E_5)$

- Mix set and probabilities by putting probability mass over sets rather than points
- Other approach: consider a (convex) set of probability masses

A characteristic of belief functions

Complete monotonicity

If $\underline{P}$ is a belief measure if and only if it satisfies the inequality

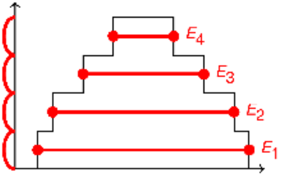
$$\underline{P}(\cup_{i=1}^{n} A_i) \geq \sum_{\mathcal{A} \subseteq \{A_1,\ldots,A_n\}} (-1)^{|\mathcal{A}|+1} \underline{P}(\cap_{A_i \in \mathcal{A}} A_i)$$

for any collection of events.
Simply the exclusion/inclusion principle with an equality

special cases

Measures $[\underline{P}, \overline{P}]$ include:

- Probability distributions: mass on atoms/singletons
- Possibility distributions: mass on nested sets



$\rightarrow$ "simplest" model including both sets and probabilities as subcases!

Frequencies of imprecise observations

Imprecise poll: "Who will win the next Wimbledon tournament?"

∘ N(adal) ∘ F(ederer) ∘ D(jokovic) ∘ M(urray) ∘ O(ther)

60 % replied $\{N, F, D\} \rightarrow m(\{N, F, D\}) = 0.6$

15 % replied "I do not know" $\{N, F, D, M, O\} \rightarrow m(\mathcal{S}) = 0.15$

10 % replied Murray $\{M\} \rightarrow m(\{M\}) = 0.1$

5 % replied others $\{O\} \rightarrow m(\{O\}) = 0.05$

...

P-box [21]

A pair $[\underline{F}, \overline{F}]$ of cumulative distributions

Bounds over events $[-\infty, x]$

- Percentiles by experts;
- Kolmogorov-Smirnov bounds;

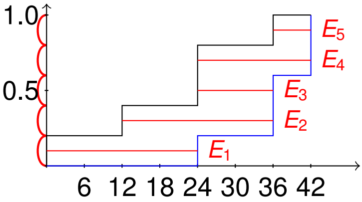
Can be extended to any pre-ordered space [17], [36] $\Rightarrow$ multivariate spaces!

Expert providing percentiles

$$0 \leq P([-\infty, 12]) \leq 0.2$$

$$0.2 \leq P([-\infty, 24]) \leq 0.4$$

$$0.6 \leq P([-\infty, 36]) \leq 0.8$$

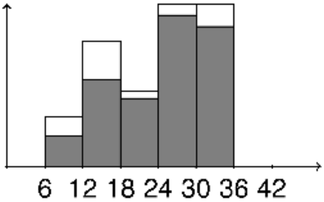


Other means to get random sets/belief functions

- Extending modal logic: probability of provability [34]
- Parameter estimation using pivotal quantities [28]
- Statistical confidence regions [16]
- Modify source information by its reliability [30]
- ...

Limits of random sets

- Not yet satisfactory extension of Bayesian/subjective approach
- Still some items of information it cannot model in a simple way, e.g.,
  - probabilistic bounds over atoms s_i (imprecise histograms, ...) [13] ;
  - comparative assessments such as $2P(B) \leq P(A)$ [29]



Outline

1. Introductory elements
2. Imprecise probabilities: use and misuse
3. Representing partial (probabilistic) knowledge
   - Basic frameworks
   - Possibility distributions
   - Random sets
   - A glimpse into probability sets
4. Merging partial (probabilistic) knowledge
5. Independence and propagation
6. Decision in presence of imprecision

Imprecise probabilities

Basic tool

A set $\mathcal{P}$ of probabilities on $\mathcal{S}$ or an equivalent representation

- $\overline{P}(A) = \sup_{P \in \mathcal{P}} P(A)$ (Upper probability)
- $\underline{P}(A) = \inf_{P \in \mathcal{P}} P(A) = 1 - \overline{P}(A^c)$ (Lower probability)

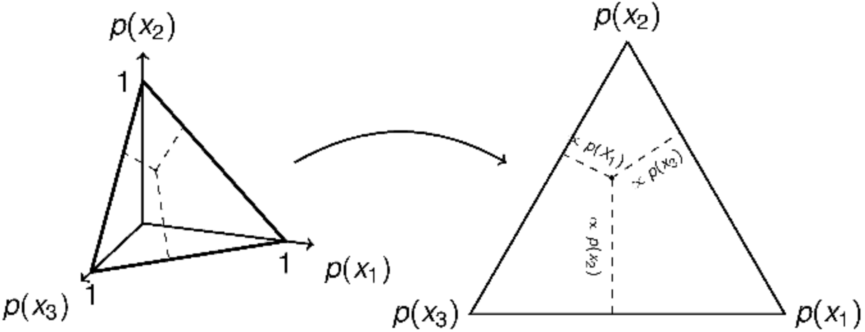
Note: lower/upper bounds on events alone cannot model any convex $\mathcal{P}$

$[\underline{P}, \overline{P}]$ as

- subjective lower and upper betting rates [38]
- bounds of an **ill-known probability measure** $P \Rightarrow \underline{P} \leq P \leq \overline{P}$ [5] [39]

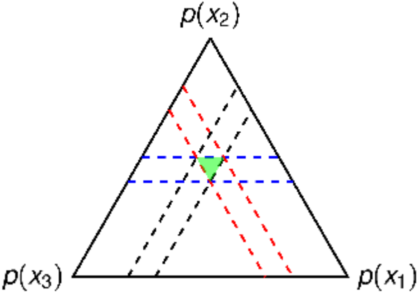
Illustrative example

$$p(x_1) = 0.2, p(x_2) = 0.5, p(x_3) = 0.3$$



Illustrative example

$p(x_1) \in [0.2, 0.3]$, $p(x_2) \in [0.4, 0.5]$, $p(x_3) = [0.2, 0.3]$



	$\{x_1\}$	$\{x_2\}$	$\{x_3\}$	$\{x_1, x_2\}$	$\{x_1, x_3\}$	$\{x_2, x_3\}$
$\underline{P}$	0.2	0.4	0.2	0.7	0.5	0.7

$\Rightarrow$ not a belief function! By computing the corresponding m, we have $m(\{x_1, x_2, x_3\}) = -0.1$

Means to get Imprecise probabilistic models

- Include all representations mentioned so far . . .
- . . . and a couple of others
  - probabilistic comparisons
  - density ratio-class
  - expectation bounds
  - . . .
- fully coherent extension of Bayesian approach

$$\mathcal{P}(\theta|x) = L(\theta|x)\mathcal{P}(\theta)$$

$\rightarrow$ often easy for "conjugate prior" [31]

- make probabilistic logic approaches imprecise [25, 14]

Example of Bayesian extension: the IDM

IDM: Imprecise Dirichlet Model

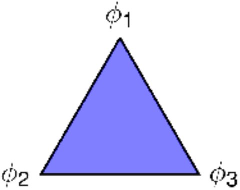
- Set of possibilities $\mathcal{X} = \{x_1, \dots, x_n\}$
- "Parameters" $\Theta = (\theta_1, \dots, \theta_n) \in [0, 1]^n$ with $\theta_i = p(x_i)$
- Observation vector $x = (a_1, \dots, a_n)$ with $a_i = \#x_i$ and $\sum_i a_i = N$
- Likelihood

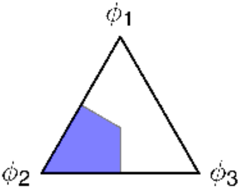
$$L(\theta|x) = P(x|\theta) = \binom{N}{x}\theta_1^{a_1}\dots\theta_n^{a_n}$$

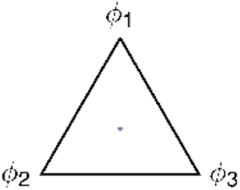
- Prior $P(\theta) \sim Diri(v\phi)$ with
  - $v \in \mathbb{R}^+$: prior strength, $\sim \#$unobserved samples ($v = 0 \to$ no strength)
  - $\phi = (\phi_1, \dots, \phi_n) \in [0, 1]^n$ with $\sum_i \phi_i = 1$: prior frequencies
- IDM: fix v, let $\phi \in \Phi$ with Φ subset of $n - 1$ unit simplex

Possible prior sets

Unknown prior	x_2 more likely (=modal value)	All results equally likely







Observation vector $x = (3, 6, 1)$ and $v = 3$

Unknown prior	x_2 more likely (=modal value)	All results equally likely
$\overline{P}(\theta_2\|x) = 9/13$	$\overline{P}(\theta_2\|x) = 9/13$	$\overline{P}(\theta_2\|x) = 7/13$
$\underline{P}(\theta_2\|x) = 6/13$	$\underline{P}(\theta_2\|x) = 7/13$	$\underline{P}(\theta_2\|x) = 7/13$

Other "imprecised" classical models

- Exponential family [31, 4]
- Bayesian Model Averaging [8]
- Gaussian process [27]
- Dirichlet process [35, 3]

A crude summary

Possibility distributions

- **+**: very simple, natural in many situations (nestedness), extend set-based approach
- **-**: at odds with probability theory, limited expressiveness

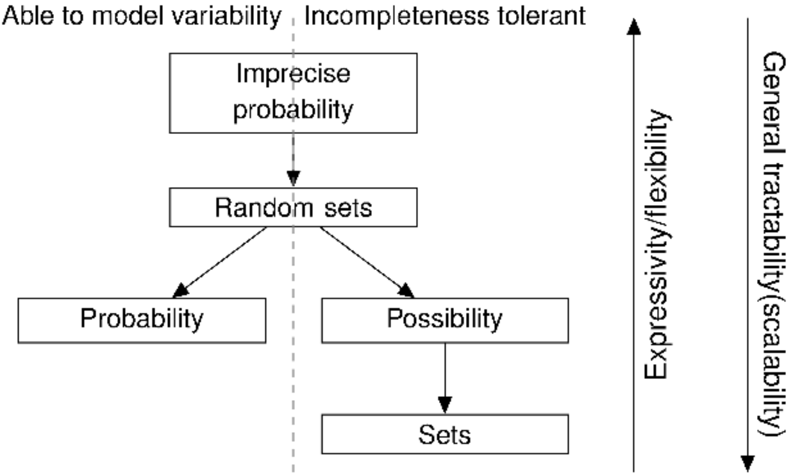
Random sets

- **+**: include probabilities and possibilities, include many models used in practice
- **-**: general models can be intractable, limited expressiveness

Imprecise probabilities

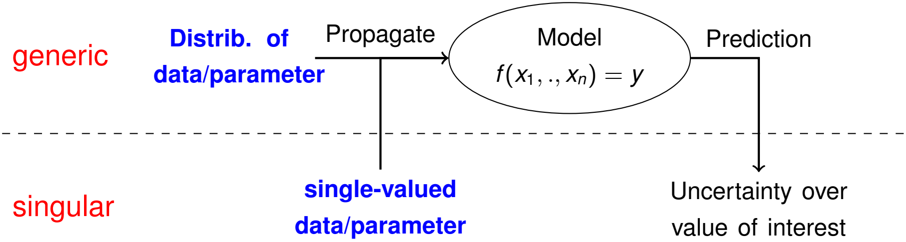
- **+**: most consistent extension of subjective probabilistic approach, very flexible
- **-**: general models can be intractable

A not completely accurate but useful picture



Outline

Uncertainty propagation revisited



Output genericity: same as most generic input variable/parameter

Propagation: usual steps

1. **Represent**: provide an uncertainty model for $x_1, \ldots, x_n$
2. **Merge**: if multiple models given for x_i, merge into a single one
3. **Combine**: specify (in)dependencies between x_i's to get global model
4. **Propagate**: propagate to get uncertainty over y
5. **Decide**: once uncertainty on y estimated, decide on an action

Merging: Definition and goals [19]

Combine items of information $\mathcal{I}_1, \ldots, \mathcal{I}_S$ on quantity $X \in \mathcal{X}$ given by S sources:

$$f(\mathcal{I}_1, \ldots, \mathcal{I}_S) = \mathcal{I}^*$$

- Usually, X assumed to have a **true**, yet unknown value in $\mathcal{X}$
- In principle $\mathcal{S}$ can be the (multi-dim) real space, finite space of elements/classes, space of functions, ...
- $\mathcal{I}_i$ and $\mathcal{I}^*$ are generally uncertainty models of the same theory (framework closeness)
- **Goal** of information merging: how to pick f to
  - Gain information from $\mathcal{I}_1, \ldots, \mathcal{I}_S$
  - Increase the reliability (trust) in my final result

The three basic fusion schemes

- Conjunction:

$$f = \cap, \quad \mathcal{I}^* = \cap_{i=1}^{S} \mathcal{I}_i$$

Assumes that all sources provide reliable information (no important conflict allowed)

- Disjunction:

$$f = \cup, \quad \mathcal{I}^* = \cup_{i=1}^{S} \mathcal{I}_i$$

Assumes that at least one source is reliable (very conservative assumption)

- (Weighted) average:

$$f = \sum w_i, \quad \mathcal{I}^* = \sum_{i=1}^{S} w_i \mathcal{I}_i$$

Assumes that most sources are ok (equivalent to counting)

Probabilities and merging

Assume that we have P_1, P_2 as opinions:

- Conjunction is impossible, as $P_1 \cap P_2$ exists only if $P_1 = P_2$ $\rightarrow$ Product $P_1 \cdot P_2$ may be considered as a surrogate to "intersection"
- Disjunction (or its convex hull) provides $P_1 \cup P_2$, not a single probability!
- Average is ok, $\alpha P_1 + (1 - \alpha) P_2$ still a probability

Sets and merging

Assume that we have E_1, E_2 as opinions:

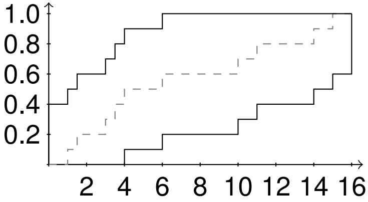
- Conjunction is possible, provided by $E_1 \cap E_2 \neq \emptyset$ (no conflict)
- Disjunction gives $E_1 \cup E_2$, again a set, possibly quite big
- Average $1/2 E_1 + 1/2 E_2$ gives a random set, not a set!

$\mathcal{P}_1$: p-box from confidence intervals

Using non-parametric Kolmogorov-Smirnov bounds.

Useful when small samples and no idea about the possible shape of the distribution (if it exists)

Example: variable $X \in [0, 16]$, observations (1; 1.5; 3; 3.5; 4; 6; 10; 11; 14; 15)



$\mathcal{P}_2$: expert imprecise percentiles

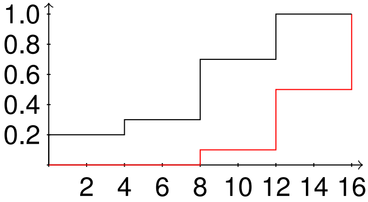
Expert providing a finite set of possible percentiles.

Exemple:

$$0 \leq P([-\infty, 4]) \leq 0.2$$

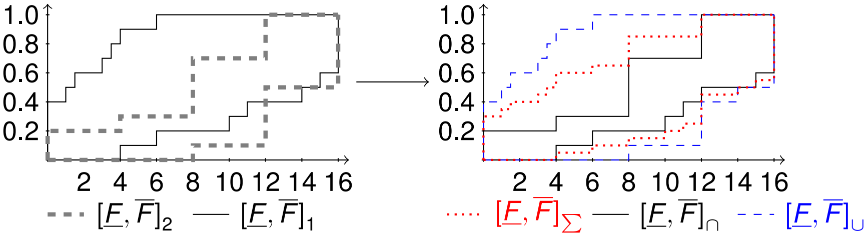
$$0.1 \leq P([-\infty, 8]) \leq 0.3$$

$$0.5 \leq P([-\infty, 12]) \leq 0.7$$



Combining those two sources

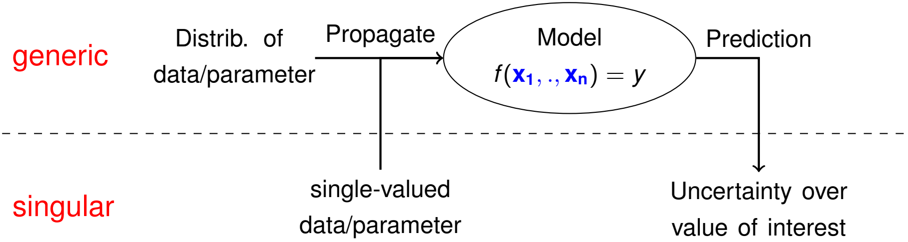
Different ways to combine this information



Outline

Outline

Uncertainty propagation revisited



Output genericity: same as most generic input variable/parameter

Propagation: usual steps

1. **Represent**: provide an uncertainty model for $x_1, \ldots, x_n$
2. **Merge**: if multiple models given for x_i, merge into a single one
3. **Combine**: specify (in)dependencies between x_i's to get global model
4. **Propagate**: propagate to get uncertainty over y
5. **Decide**: once uncertainty on y estimated, decide on an action

Independence statement=strong information

- For one X, uniformity $\neq$ lack of knowledge
- $\rightarrow$ Symmetry of knowledge $\neq$ knowledge of symmetry
- For two X, Y, independence $\neq$ lack of knowledge about interaction
- $\rightarrow$ No knowledge of interaction $\neq$ knowledge of no interaction
- Statistically speaking, stating independence requires just as much data as stating dependence

IP tools instrumental to consider sets of dependence assumptions, even when marginal distributions are well-known.

A small reliability example

- Two pumps X and Y, either functioning (x, y) or not $(\neg x, \neg y)$
- Overall system $\phi(X, Y)$ works if and only if one of the pump works (XOR):
  - no pump functioning means no pumping
  - two pumps functioning means overload

$$\phi(X, Y) = \begin{cases} 1 & \text{if } x\neg y \vee \neg xy \\ 0 & \text{else } (xy \vee \neg x\neg y) \end{cases}$$

- Probability of the system functioning is

$$\begin{aligned} P(\phi(X, Y) = 1) &= P(x\neg y \vee \neg xy) \\ &= P(x\neg y) + P(\neg xy) \end{aligned}$$

Independent case

Assume $p_X(x) = 0.7, p_Y(y) = 0.6$ and the resulting joint

	x	$\neg x$	$\sum$
y	$0.7 \cdot 0.6$	$0.3 \cdot 0.6$	0.6
$\neg y$	$0.7 \cdot 0.4$	$0.3 \cdot 0.4$	0.4
$\sum$	0.7	0.3	

$$P(\phi(X,Y) = 1) = P(x\neg y) + P(\neg xy)$$
$$= p_X(x)p_Y(\neg y) + p_X(\neg x)p_Y(y) = 0.46$$

$\rightarrow$ less chance of working than not working

Unknown dependence case: upper bound

Assume $p_X(x) = 0.7, p_Y(y) = 0.6$ and the table

	x	$\neg x$	$\sum$
y	0.3	$\min(p_X(\neg x), p_Y(y)) = 0.3$	0.6
$\neg y$	$\min(p_X(x), p_Y(\neg y)) = 0.4$	0.3	0.4
$\sum$	0.7	0.3	

$$\overline{P}(\phi(X,Y) = 1) = \max P(x\neg y) + P(\neg xy)$$
$$= 0.3 + 0.4 = 0.7$$

Unknown dependence case: lower bound

Assume $p_X(x) = 0.7, p_Y(y) = 0.6$ and the table

	x	$\neg x$	$\sum$
y	$\min(p_X(x), p_Y(y)) = 0.6$	0	0.6
$\neg y$	0.1	$\min(p_X(\neg x), p_Y(\neg y)) = 0.3$	0.4
$\sum$	0.7	0.3	

$$\begin{aligned} \underline{P}(\phi(X,Y)=1) &= \min P(x\neg y) + P(\neg xy) \\ &= 0.1 + 0 = 0.1 \end{aligned}$$

$\rightarrow [\underline{P}, \overline{P}] = [0.1, 0.7]$, incomparability of working vs not working

Partially assumed dependence: common case failure

Assume $p_X(x) = 0.7, p_Y(y) = 0.6$ and the following bounds

	x	$\neg x$	$\sum$
y	$p_X(x) \cdot p_Y(y) + \epsilon$	$p_X(\neg x) \cdot p_Y(y) - \epsilon$	0.6
$\neg y$	$p_X(x) \cdot p_Y(\neg y) - \epsilon$	$p_X(\neg x) \cdot p_Y(\neg y) + \epsilon$	0.4
$\sum$	0.7	0.3	

with $0 \leq \epsilon \leq 0.08$, mild assumption of common cause failure

$\rightarrow [\underline{P}, \overline{P}] = [0.3, 0.46]$, no change in conclusions despite imprecision

Outline

On independence and interpretation [11, 12]

Use of Independence $\sim$ facilitate computations in multi-variate problems

Meaning for two quantities X, Y to be independent, when they are:

- **generic**: in this case, X, Y associated to distributions P_X, P_Y and stochastic independence may apply $\rightarrow$ imprecise probabilities= sensitivity analysis of an "objective" concept.
- **singular**: X, Y are supposed to have one true value. Independence here is "subjective", and purely concerns beliefs, not how the values of X and Y can affect each others.

In singular case, much less clear how it should be modelled and even measured?

Two views of independence

In general, two ways to express independence of X, Y:

- Compositional (stochastic) independence:

$$P_{X,Y}(X \in A, Y \in B) = P_X(X \in A)P_Y(Y \in B)$$

  Clear if $P \simeq$frequencies, less if P=degrees of bellief

- Conditional ("epistemic") independence of Y w.r.t. X:

$$P_X(X \in A | Y \in B) = P_X(X \in A)$$

  Express that learning B about Y do not change belief about X
  - A non-symmetric notion, but with precise probabilities become symmetric
  - With precise probabilities, reduces to the first definition

When P becomes imprecise, the two notions extends in different ways.

Three different definitions

Assume I have $\mathcal{P}_X, \mathcal{P}_Y$ on finite spaces

- **Strong independence** (SI)

$$\mathcal{P}^{SI}_{XY} = \{p | p(x,y) = p(x)p(y), p(x) \in \mathcal{P}_X, p(y) \in \mathcal{P}_Y\}$$

- **Epistemic irrelevance** (IR) of X w.r.t. Y

$$\mathcal{P}^{IR}_{X \to Y} = \{p | p(x,y) = p(y|x)p(x), p(x) \in \mathcal{P}_X, p(y|x) \in \mathcal{P}_Y\}$$

  We can have $p(y|x) \neq p(y|x')$ for $x \neq x'$, and $\mathcal{P}^{IR}_{X \to Y} \neq \mathcal{P}^{IR}_{Y \to X}$

- **Random set independence**, if $\mathcal{P}_X, \mathcal{P}_Y$ representable by m_X, m_Y

$$\mathcal{P}^{RI}_{XY} = \{p | P(C) \geq \sum_{AxB \subseteq C} m_X(A)m_Y(B)\}$$

  equivalent to consider joint mass $m_XY(A \times B) = m_X(A)m_Y(B)$

Inclusion relationship

In general, we have

$$\mathcal{P}^{SI}_{XY} \subseteq \begin{cases} \mathcal{P}^{IR}_{X\to Y} \\ \mathcal{P}^{IR}_{Y\to X} \end{cases} \subseteq \mathcal{P}^{RI}_{XY}$$

Allowing to use one principle to approximate another, for example for computational convenience.

In the precise case, they all collapse to the same formal definition.

IP and robust stochastic independence

Assume now $p_X(x) \in [0.6, 0.8]$ and $p_Y(y) \in [0.5, 0.7]$

$$P(\phi(X,Y) = 1) = p_X(x)p_Y(\neg y) + p_X(\neg x)p_Y(y) = 0.46$$

$$\underline{P}(\phi(X,Y) = 1) = \overline{p}_X(x)\underline{p}_Y(\neg y) + \underline{p}_X(\neg x)\overline{p}_Y(y)$$
$$0.8 \cdot 0.3 + 0.2 \cdot 0.7 = 0.38$$

$$\overline{P}(\phi(X,Y) = 1) = \underline{p}_X(x)\overline{p}_Y(\neg y) + \overline{p}_X(\neg x)\underline{p}_Y(y)$$
$$0.6 \cdot 0.4 + 0.2 \cdot 0.5 = 0.5$$

IP and random set independence: upper case

$p_X(x) \in [0.6, 0.8]$ and $p_Y(y) \in [0.5, 0.7]$ give masses m_X, m_Y. Apply product rule to get $m_{X,Y}$

	x	$\neg x$	$x, \neg x$	$\sum$
y	0.3	0.1	0.1	0.5
$\neg y$	0.18	0.06	0.06	0.3
$y, \neg y$	0.12	0.04	0.04	0.2
$\sum$	0.6	0.2	0.2	

$$\underline{P}(\phi(X, Y) = 1) = \sum \blacksquare = 0.28$$

$$\overline{P}(\phi(X, Y) = 1) = \sum \blacksquare + \sum \square = 0.74$$

A summary so far

Imprecise probability uses:

- Model a random variable whose distribution is ill-known
- Model beliefs about a deterministic but ill-known variable
- Relax the need to specify a single (in)dependence assumption
- Rather than making one computation per hypothesis, directly compute bounds over set of hypothesis

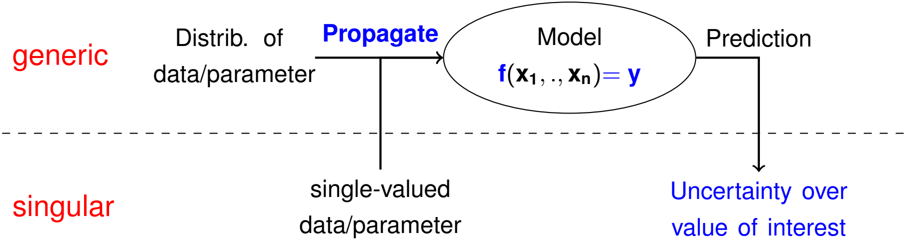
In practice, most operations can be achieved/approximated through linear optimisation, once propagation through f is done

Some warnings

- Just as for probabilities, making exact computations for complex models difficult
- Some different notions reducing to the same mathematical tools in probability (independence) may have various extensions

Outline

Uncertainty propagation revisited



Output genericity: same as most generic input variable/parameter

Propagation: usual steps

1. Represent: provide an uncertainty model for $x_1, \ldots, x_n$
2. **Merge**: if multiple models given for x_i, merge into a single one
3. **Combine**: specify (in)dependencies between x_i's to get global model
4. **Propagate: propagate to get uncertainty over** y
5. **Decide**: once uncertainty on y estimated, decide on an action

Random set propagation: principle

Assume:

- each X_i is associated to mass m_i
- random set independence holds
- only intervals receive positive mass

Propagation is equivalent to

- Take selections of intervals $E_1, \ldots, E_n$ from $m_1, \ldots, m_n$
- Compute $f(E_1, \ldots, E_n)$ (main bottleneck, as with probabilities)
- Associate product of masses $m_1(E_1) \ldots m_n(E_n)$ to $f(E_1, \ldots, E_n)$
- Make inferences from that mass

A minimal example

- two polluting elements X_1, X_2, expressed in average percentage per m^3
- experts tell that
  - $\underline{P}(X_1 \in [1,4]) = 0.8$, and in any case $X_1 \in [0,10]$
  - $\underline{P}(X_2 \in [2,5]) = 0.6$, and in any case $X_2 \in [1,8]$
- We are interested in the value $Y = X_1 + X_2$

A minimal example continued

	$m([1,4]) = 0.8$	$m([0,10]) = 0.2$
$m([2,5]) = 0.6$		
$m([1,8]) = 0.4$		

- Take selections of intervals $E_1, \ldots, E_n$ from $m_1, \ldots, m_n$

A minimal example continued

	$m([1,4]) = 0.8$	$m([0,10]) = 0.2$
$m([2,5]) = 0.6$	$m([3,9])$	$m([2,15])$
$m([1,8]) = 0.4$	$m([2,12])$	$m([1,18])$

- Take selections of intervals $E_1, \ldots, E_n$ from $m_1, \ldots, m_n$
- Compute $f(E_1, \ldots, E_n)$ (main bottleneck, as with probabilities)

A minimal example continued

	$m([1,4]) = 0.8$	$m([0,10]) = 0.2$
$m([2,5]) = 0.6$	$m([3,9]) = 0.54$	$m([2,15]) = 0.12$
$m([1,8]) = 0.4$	$m([2,12]) = 0.32$	$m([1,18]) = 0.08$

- Take selections of intervals $E_1, \ldots, E_n$ from $m_1, \ldots, m_n$
- Compute $f(E_1, \ldots, E_n)$ (main bottleneck, as with probabilities)
- Associate product of masses $m_1(E_1) \ldots m_n(E_n)$ to $f(E_1, \ldots, E_n)$

A minimal example continued

	$m([1,4]) = 0.8$	$m([0,10]) = 0.2$
$m([2,5]) = 0.6$	$m([3,9]) = 0.54$	$m([2,15]) = 0.12$
$m([1,8]) = 0.4$	$m([2,12]) = 0.32$	$m([1,18]) = 0.08$

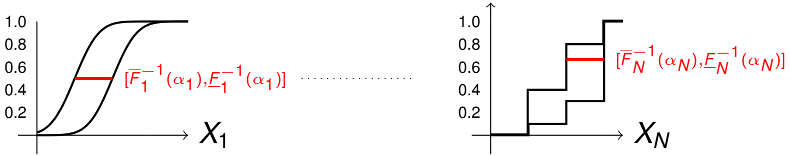
e.g.,

$$P(Y \leq 10) = P(Y \in [0,10]) \in [0.54, 1]$$

- Take selections of intervals $E_1, \ldots, E_n$ from $m_1, \ldots, m_n$
- Compute $f(E_1, \ldots, E_n)$ (main bottleneck, as with probabilities)
- Associate product of masses $m_1(E_1) \ldots m_n(E_n)$ to $f(E_1, \ldots, E_n)$
- Make inferences from that mass

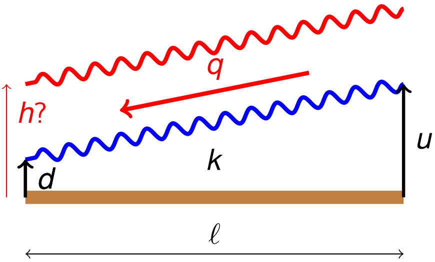
Random set propagation: practice [1, 23]

- Still assuming random set independence
- Pick $\alpha_1, \ldots, \alpha_N \in [0, 1]$ randomly from uniform



- Estimate $f([\overline{F}_1^{-1}(\alpha_1), \underline{F}_1^{-1}(\alpha_1)], \ldots, [\overline{F}_N^{-1}(\alpha_N), \underline{F}_N^{-1}(\alpha_N)])$
- Repeat R times: R imprecise samples of Y

A more involved example



- enbankment construction
- simplified model of flood levels h

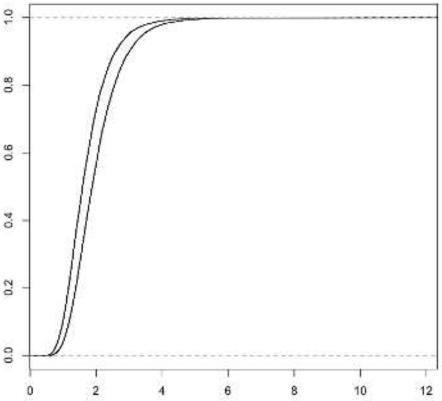
$$h(q,k,u,d) = \begin{cases} \left(\dfrac{q}{k\sqrt{\frac{u-d}{\ell}}b} \right)^{\frac{3}{5}} & \text{if } q \geq 0 \\ 0 & \text{otherwise.} \end{cases} \quad (1)$$

Parameter uncertainties

	significance	units	uncertainty	Repres.
h	flood level	m	(model)	
q	maximal flow	$m^3 s^{-1}$	Gumbel($\mu \in [1300, 1400]$, $\beta = 715$)	p-box
b	river width	m	300	value
k	Strickler coefficient	$m^{1/3} s^{-1}$	mode=30, support=[15,35]	possibility
u	upriver level	m	mode=55, support=[54,56]	possibility
d	downriver level	m	mode=50, support=[49,51]	possibility
ℓ	section length	m	6400	value

Exemple: illustration

Result on thresholding events ($P([0, x])$) for independence assumption



Random set propagation: limits and warnings

- Limited expressiveness: not all convex sets of probabilities are random sets
- Independence conservativeness: random set independence gives wider bounds than robust stochastic independence
- Dependence modelling: using copulas on distributions *m* is not equivalent to a robust applications of copulas, unless all marginal models are p-boxes (imprecise cumulative distributions)

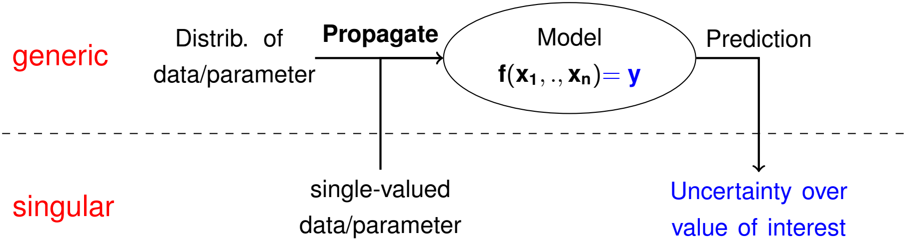
Practical application not much more difficult than for precise probabilities (similar computational bottlenecks)

Propagation and imprecise probabilities: recent trends

- Using importance sampling techniques to consider multiple (set of) probabilities with one sample [22, 40]
- Combining imprecise probability tools with surrogate models, i.e., polynomial chaos expansions [33]

Outline

Uncertainty propagation revisited



Output genericity: same as most generic input variable/parameter

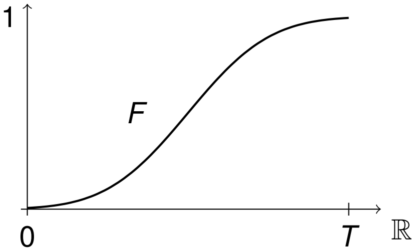
Propagation: usual steps

1. Represent: provide an uncertainty model for $x_1, \ldots, x_n$
2. **Merge**: if multiple models given for x_i, merge into a single one
3. **Combine**: specify (in)dependencies between x_i's to get global model
4. **Propagate**: propagate to get uncertainty over y
5. **Decide**: once uncertainty on y estimated, decide on an action

Outline

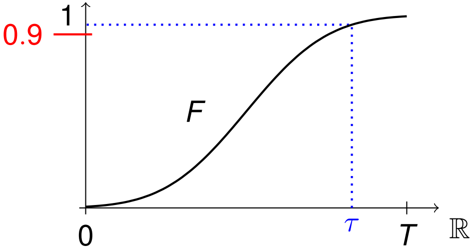
Precise case

Fix a threshold τ, decide whether $P(X \leq \tau) \geq \alpha$ with α critical level



Precise case: example

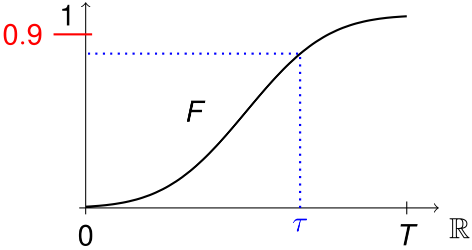
For instance, take $\alpha = 0.9$



We have $P(X \leq \tau) > \alpha$, acceptable risk

Precise case: example

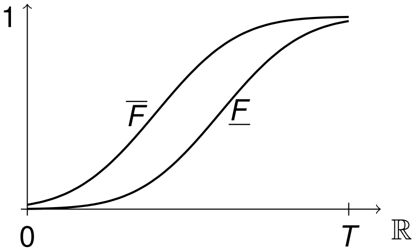
For instance, take $\alpha = 0.9$



We have $P(X \leq \tau) < \alpha$, unacceptable risk

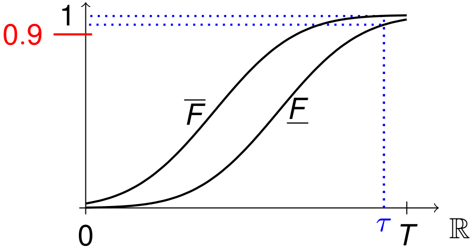
Imprecise case

Fix a threshold τ, decide whether $P(X \leq \tau) \geq \alpha$ with α critical level



Imprecise case: example

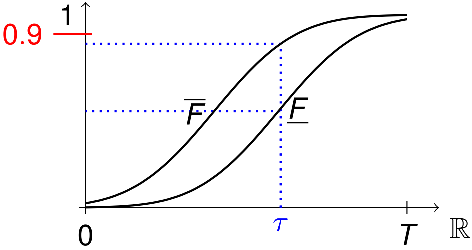
For instance, take $\alpha = 0.9$



We have $\overline{P}(X \leq \tau) > \underline{P}(X \leq \tau) > \alpha$, acceptable risk

Imprecise case: example

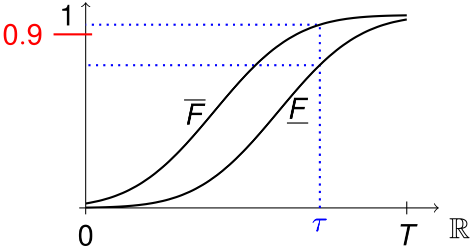
For instance, take $\alpha = 0.9$



We have $\underline{P}(X \leq \tau) < \overline{P}(X \leq \tau) < \alpha$, unacceptable risk

Imprecise case: example

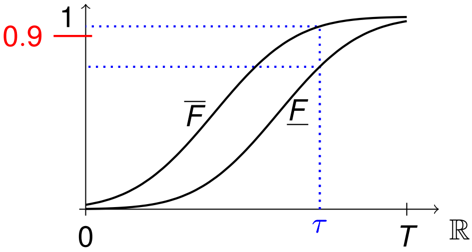
For instance, take $\alpha = 0.9$



We have $\underline{P}(X \leq \tau) < \alpha < \overline{P}(X \leq \tau)$, no clear answer

Imprecise case: three DM different attitudes

- No preference of behaviour: undecided
- Pessimistic: pick worst case $\underline{P}(X \leq \tau)$
- Optimistic: pick best case $\overline{P}(X \leq \tau)$



Key idea: decision maker attitude should be separated from the available information

Outline

Decision-making: classical frames

Given set of uncertain quantities/utilities $X_1, \dots, X_n$, compare pair-wisely

- By expectation: $X_i \succ_{\mathbb{E}} X_j$ if $\mathbb{E}(X_i) \geq \mathbb{E}(X_j)$ with

$$\mathbb{E}(X) = \sum x \cdot p(x)$$

  Pros: most recognized criteria, strong theoretical foundations

  Cons: necessitates utility values to be elicited and well-defined

  Risk decision: specific case with utility function $\mathbb{I}_{x \leq \tau}$

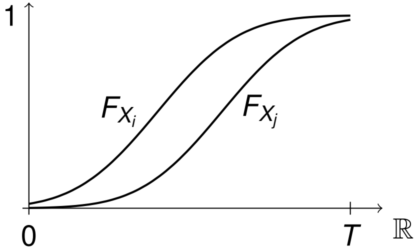
- By statistical preference: $X_i \succ_P X_j$ if $P(X > Y) > 0.5$
  Pros: only need an ordinal scale, close to the notion of median

  Cons: possible cycles ($X \succ_P Y \succ_P Z \succ_P X$), need dependencies

Decision-making: classical frames

Given set of uncertain quantities/utilities $X_1, \ldots, X_n$, compare pair-wisely

- By stochastic dominance: $X_i \succ_F X_j$ if $P(X_i \leq x) \leq P(X_j \leq x)$



Pros: if $X_i \succ_F X_j$, $g(X_i) \succ_{\mathbb{E}} g(X_j)$ for any increasing g

Cons (?): may lead to incomparability ($X_i \not\succ_F X_j$ and $X_j \not\succ_F X_i$)

Risk decision: fixing a threshold rather than all of them

Imprecise probability: rough ideas for expectations

When going imprecise, points become intervals $\rightarrow$ how to compare

$$[\underline{\mathbb{E}}(X_i), \overline{\mathbb{E}}(X_i)] \text{ and } [\underline{\mathbb{E}}(X_j), \overline{\mathbb{E}}(X_j)]?$$

Back to precise comparison (need DM attitude)

- Maximin (pessimist): $X_i \succ_{Mm} X_j$ if $\underline{\mathbb{E}}(X_i) > \underline{\mathbb{E}}(X_j)$
- Maximax (optimist): $X_i \succ_{MM} X_j$ if $\overline{\mathbb{E}}(X_i) > \overline{\mathbb{E}}(X_j)$
- Hurwicz, in-between: $X_i \succ_{\gamma} X_j$ if

$$\gamma\overline{\mathbb{E}}(X_i) + (1-\gamma)\underline{\mathbb{E}}(X_i) > \gamma\overline{\mathbb{E}}(X_j) + (1-\gamma)\underline{\mathbb{E}}(X_j)$$

Acknowledging imprecision and indecision

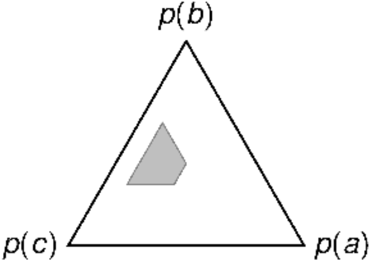
- Interval dominance: $X_i \succ_{ID} X_j$ if $\underline{\mathbb{E}}(X_i) > \overline{\mathbb{E}}(X_j)$

A small example

Imprecise probability model

Three possible states $\{a, b, c\}$

- $p(a) \in [0.1, 0.3]$
- $p(b) \in [0.3, 0.6]$
- $p(c) \in [0.3, 0.6]$

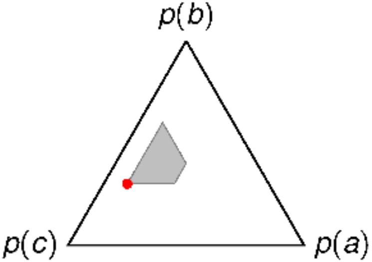


A small example

Imprecise probability model

Three possible decisions/alternatives $\{X_1, X_2, X_3\}$

$$\begin{array}{c} \\ X_1 \\ X_2 \\ X_3 \end{array} \begin{array}{c} \begin{array}{ccc} a & b & c \end{array} \\ \begin{pmatrix} 0 & 1 & 2 \\ 1 & 0.5 & 1 \\ 3 & 1 & 0 \end{pmatrix} \end{array}$$



- $\mathbb{E}(X_1) \in [1, 1.5]$
- $\mathbb{E}(X_2) \in [0.7, 0.85]$
- $\mathbb{E}(X_3) \in [0.6, 1.3]$

Summary and main messages

- Importance to know what kind of uncertainty you want to model
- Imprecise probability models useful to model lack of knowledge, or perform robustness analysis
- Not useful if you want a precise numbers, comparability of any pair of events or of any decisions, no matter what your knowledge is
- Computational burden usually higher, but not necessarily much higher than precise models

References I

[1] D. A. Alvarez.
On the calculation of the bounds of probability of events using infinite random sets.
International journal of approximate reasoning, 43(3):241–267, 2006.

[2] C. Baudrit and D. Dubois.
Practical representations of incomplete probabilistic knowledge.
Computational Statistics and Data Analysis, 51(1):86–108, 2006.

[3] A. Benavoli, F. Mangili, F. Ruggeri, and M. Zaffalon.
Imprecise dirichlet process with application to the hypothesis test on the probability that $x \leq y$.
Journal of Statistical Theory and Practice, 9(3):658–684, 2015.

[4] A. Benavoli and M. Zaffalon.
Prior near ignorance for inferences in the k-parameter exponential family.
Statistics, 49(5):1104–1140, 2015.

[5] J. O. Berger.
An overview of robust Bayesian analysis.
Test, 3:5–124, 1994.
With discussion.

[6] D. Bouyssou, D. Dubois, H. Prade, and M. Pirlot.
Decision Making Process: Concepts and Methods.
John Wiley & Sons, 2013.

[7] M. Cattaneo.
Likelihood-based statistical decisions.
In *Proc. 4th International Symposium on Imprecise Probabilities and Their Applications*, pages 107–116, 2005.

References II

[8] G. Corani and M. Zaffalon.
Credal model averaging: an extension of bayesian model averaging to imprecise probabilities.
Machine Learning and Knowledge Discovery in Databases, pages 257–271, 2008.

[9] I. Couso and D. Dubois.
Statistical reasoning with set-valued information: Ontic vs. epistemic views.
International Journal of Approximate Reasoning, 55(7):1502–1518, 2014.

[10] I. Couso, S. Montes, and P. Gil.
The necessity of the strong alpha-cuts of a fuzzy set.
Int. J. on Uncertainty, Fuzziness and Knowledge-Based Systems, 9:249–262, 2001.

[11] I. Couso and S. Moral.
Independence concepts in evidence theory.
International Journal of Approximate Reasoning, 51:748–758, 2010.

[12] I. Couso, S. Moral, and P. Walley.
A survey of concepts of independence for imprecise probabilities.
Risk Decision and Policy, 5:165–181, 2000.

[13] L. de 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.

[14] G. de Cooman.
Belief models: An order-theoretic investigation.
Ann. Math. Artif. Intell., 45(1-2):5–34, 2005.

[15] G. de Cooman and P. Walley.
A possibilistic hierarchical model for behaviour under uncertainty.
Theory and Decision, 52:327–374, 2002.

References III

[16] T. Denoeux.
Constructing belief functions from sample data using multinomial confidence regions.
I. J. of Approximate Reasoning, 42:228–252, 2006.

[17] S. Destercke, D. Dubois, and E. Chojnacki.
Unifying practical uncertainty representations: I generalized p-boxes.
Int. J. of Approximate Reasoning, 49:649–663, 2008.

[18] D. Dubois, L. Foulloy, G. Mauris, and H. Prade.
Probability-possibility transformations, triangular fuzzy sets, and probabilistic inequalities.
Reliable Computing, 10:273–297, 2004.

[19] D. Dubois, W. Liu, J. Ma, and H. Prade.
The basic principles of uncertain information fusion. an organised review of merging rules in different representation frameworks.
Information Fusion, 32:12–39, 2016.

[20] D. Dubois, S. Moral, and H. Prade.
A semantics for possibility theory based on likelihoods,.
Journal of Mathematical Analysis and Applications, 205(2):359 – 380, 1997.

[21] S. Ferson, L. Ginzburg, V. Kreinovich, D. Myers, and K. Sentz.
Constructing probability boxes and dempster-shafer structures.
Technical report, Sandia National Laboratories, 2003.

[22] T. Fetz and M. Oberguggenberger.
Imprecise random variables, random sets, and monte carlo simulation.
International Journal of Approximate Reasoning, 78:252–264, 2016.

References IV

[23] T. Fetz, M. Oberguggenberger, and S. Pittschmann.
Applications of possibility and evidence theory in civil engineering.
International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems, 8(03):295–309, 2000.

[24] B. Finetti.
Theory of probability, volume 1-2.
Wiley, NY, 1974.
Translation of 1970 book.

[25] P. Hansen, B. Jaumard, M. P. de Aragão, F. Chauny, and S. Perron.
Probabilistic satisfiability with imprecise probabilities.
Int. J. Approx. Reasoning, 24(2-3):171–189, 2000.

[26] L. Jaulin, M. Kieffer, O. Didrit, and E. Walter.
Applied Interval Analysis.
London, 2001.

[27] F. Mangili.
A prior near-ignorance gaussian process model for nonparametric regression.
International Journal of Approximate Reasoning, 78:153–171, 2016.

[28] R. Martin, J. Zhang, and C. Liu.
Dempster-shafer theory and statistical inference with weak beliefs.
Technical report, 2008.

[29] E. Miranda and S. Destercke.
Extreme points of the credal sets generated by comparative probabilities.
Journal of Mathematical Psychology, 64:44–57, 2015.

References V

[30] F. Pichon, D. Dubois, and T. Denoeux.
Relevance and truthfulness in information correction and fusion.
Int. J. Approx. Reasoning, 53(2):159–175, 2012.

[31] E. Quaeghebeur and G. de Cooman.
Imprecise probability models for inference in exponential families.
In *ISIPTA*, pages 287–296, 2005.

[32] S. Sandri, D. Dubois, and H. Kalfsbeek.
Elicitation, assessment and pooling of expert judgments using possibility theory.
IEEE Trans. on Fuzzy Systems, 3(3):313–335, August 1995.

[33] R. Schöbi and B. Sudret.
Uncertainty propagation of p-boxes using sparse polynomial chaos expansions.
Journal of Computational Physics, 339:307–327, 2017.

[34] P. Smets.
Probability of provability and belief functions.
Logique et Analyse, 133-134:177–195, 1991.

[35] S. H. Tindemans and G. Strbac.
Robust estimation of risks from small samples.
Phil. Trans. R. Soc. A, 375(2100):20160299, 2017.

[36] M. C. M. Troffaes and S. Destercke.
Probability boxes on totally preordered spaces for multivariate modelling.
Int. J. Approx. Reasoning, 52(6):767–791, 2011.

[37] R. von Mises.
Probability, Statistics and Truth.
Dover books explaining science. Dover Publications, 1981.

References VI

[38] P. Walley.
Statistical reasoning with imprecise Probabilities.
Chapman and Hall, New York, 1991.

[39] K. Weichselberger.
The theory of interval-probability as a unifying concept for uncertainty.
International Journal of Approximate Reasoning, 24(2–3):149 – 170, 2000.

[40] J. Zhang and M. D. Shields.
On the quantification and efficient propagation of imprecise probabilities resulting from small datasets.
Mechanical Systems and Signal Processing, 98:465–483, 2018.