Imprecise probability theory

Enrique Miranda and Sebastien Destercke

[1] Universidad de Oviedo, Dpto. de Estadística, Spain Heudiasyc, CNRS [2] Compiegne, France

SSIPTA 2014

An experiment

A fair coin

- Ticket worth 1€ if tails, zero else
- Put a price on that Ticket. I can sell it or buy it.

Plan

1. Precise probabilities
2. Motivation for Imprecise probabilities (IP)
3. Lower previsions and imprecise probabilities
4. Conditioning in (imprecise) probability

Objectives

- Basic of probabilities
- Interpretations of probabilities
- Geometric and constraint representations

Probabilities: an example

- $\Omega = \{C, P, T\}$:how did I come here?

Car (C)

Plane (P)

Train (T)

- $p(C) = 0.2$, $p(P) = 0.5$, $p(T) = 0.3$
- No car= $\{P, T\} = \{P\} \cup \{T\}$

$$P(\{P, T\}) = P(\{P\}) + P(\{T\}) = p(P) + p(T) = 0.8$$

Probabilities: example

- $\Omega = \{R, P, D\}$:recognizing between

Renegade (R)

Provocateur (P)

Damaged (D)

- $p(R) = 0.2$, $p(P) = 0.5$, $p(D) = 0.3$
- No danger= $\{P, D\} = \{P\} \cup \{D\}$

$$P(\{P, D\}) = P(\{P\}) + P(\{D\}) = p(P) + p(D) = 0.8$$

Probabilities: basic definitions

- Finite possibility space Ω
- Probability mass $p : \Omega \to [0, 1]$ such that
  - $p(\omega) > 0$, $\sum_{\omega \in \Omega} p(\omega) = 1$
- Probability of an event $A \subseteq \Omega$

$$P(A) = \sum_{\omega \in A} p(\omega)$$

Expectation

- X: monetary cost of a trip using ω
- $X(C) = 200$, $X(P) = 300$, $X(T) = 150$
- expected cost of my trip

$$\mathbb{E}(X) = 0.2 \cdot 200 + 0.5 \cdot 300 + 0.3 \cdot 150 = 235$$

- X: cost of non-intervention if ω is true and we let it go
- $X(R) = 500$, $X(P) = 20$, $X(D) = 100$
- expected cost

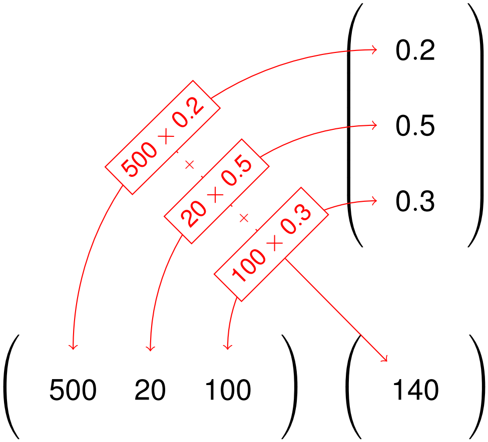
$$\mathbb{E}(X) = 0.2 \cdot 500 + 0.5 \cdot 20 + 0.3 \cdot 100 = 140$$

Expectation: definition

- $X : \Omega \to \mathbb{R}$ function or real-valued random variable
- p probability mass
- Expectation $\mathbb{E}(X)$ is the operator

$$\mathbb{E}(X) = \sum_{\omega \in \Omega} p(\omega) X(\omega)$$

Expectation: matrix multiplication form



Main interpretations

The value $P(A)$ can be given two main interpretations:

- As frequencies
- As betting prices

Frequentist probabilities

$$P(A) := \lim_{N\to\infty} \frac{\sharp(x \in A)}{N} = \lim_{N\to\infty} \frac{A \text{ happens}}{N}$$

Examples

- game of chance (loteries, poker, roulette)
- physical quantities in
  - engineering (component failure, product defect)
  - biology (patient variability)
  - economics , . . .

But. . .

Frequentist probabilities: end of story?

. . . some uncertain quantity are not repeatable/not statistical quantities:

- what's the age of the king of Sweden?
- what's the distance between you and the closest wall (in cm)? Or between this car and the next? Or the robot and the wall?
- has it rained in Edinburgh yesterday?
- when will YOUR phone fail? has THIS altimeter failed? is THIS camera not operating?

$\Rightarrow$ can we still use probability to model these uncertainties?

Betting interpretation

$P(A)$: price at which you buy/sell the ticket that

- worth 1 if A happens/is true
- worth 0 if A does not happen/is false

$P(A)$ reflects your uncertainty about A

- $P(A) = 1$ if A certain
- $P(A) = 0$ if A impossible
- The closest $P(A)$ is to 1 (0), the more certain you are that A is or will be true (false)

Combining bets: additivity

If I am willing to sell you a ticket for A, A^c, then your gain is

$$(\mathbb{I}_A - P(A)) + (\mathbb{I}_{A^c} - P(A^c)) = \underbrace{1}_{\text{You win}} - \underbrace{(P(A) + P(A^c))}_{\text{You pay}}$$

with $\mathbb{I}_A$ indicator function of A ($\mathbb{I}_A(\omega) = 1$ if $\omega \in A$, 0 else).

- if $P(A) + P(A^c) > 1$, you always lose money

  $\Rightarrow$ you will not buy it

- if $P(A) + P(A^c) < 1$, I always lose money

  $\Rightarrow$ I will not sell it

if we are *rational*, we will settle prices such that $P(A) + P(A^c) = 1$

Generic combination

You can:

- combine any bets, i.e., chose from $A_1, \ldots, A_n$
- scale them, i.e., buy/sell a number $c_i \in \mathbb{R}$ of tickets

Are they rational prices?

- Assume the following bets:

$$P(\{C\}) = 0.5, \quad P(\{C,P\}) = 0.3$$

- Assume the following bets:

$$P(\{C,P\}) = 0.5, \quad P(\{P,T\}) = 0.6, \quad P(\{C,T\}) = 0.7$$

- Assume the following bets:

$$P(\{C,P\}) = 0.5, \quad P(\{P,T\}) = 0.5$$

Generic event characterization

If $P(A_i)$, $i = 1, \dots, n$ collection of assessments, they *Avoid sure loss* if

$$\sup_{\omega \in \Omega} \sum_{i=1}^{n} c_i [\mathbb{I}_{A_i}(\omega) - P(A_i)] \geq 0$$

with $c_i \in \mathbb{R}$ numbers of bought/sold tickets.

Representation theorem

Assessments $P(A_i)$ are coherent iff there is a p s.t.

$$P(A_i) = \sum_{x \in A_i} p(x)$$

$\hookrightarrow$ any probability p can be obtained by bets on events, but this will not be true when going imprecise

- $X(\omega)$: reward if ω true
- $X(C) = 200$, $X(P) = 300$, $X(T) = 150$
- What would be the price (for selling/buying) $P(X)$ of that ticket?
- I would not sell at any price lower than 150 (why?)
- You would not buy at any price higher than 300 (why?)

Are they rational prices?

- Assume the following bets:
  - $X_1(C) = 1$, $X_1(P) = -1$, $X_1(T) = 0$, $P(X_1) = 0$
  - $X_2(C) = 0$, $X_2(P) = 1$, $X_2(T) = -1$, $P(X_2) = 0$
  - $X_3(C) = 1$, $X_3(P) = -1$, $X_3(T) = 1$, $P(X_3) = 0$
- Assume the following bets:
  - $X_1(C) = 2$, $X_1(P) = -1$, $X_1(T) = 0$, $P(X_1) = 1$
  - $X_2(C) = 1$, $X_2(P) = 1$, $X_2(T) = -1$, $P(X_2) = 0$

Generic function characterization

$P(X)$= price of ticket that gives $X(\omega)$ if ω true

If $P(X_i)$, $i = 1, \ldots, n$ collection of assessments, they *Avoid sure loss* iff

$$\sup_{\omega \in \Omega} \sum_{i=1}^{n} c_i [X_i(\omega) - P(X_i)] \geq 0$$

with $c_i \in \mathbb{R}$ numbers of bought/sold tickets.

Representation theorem

Assessments $P(X_i)$ are coherent iff there is a p s.t.

$$P(X_i) = \mathbb{E}(X_i) = \sum_{\omega \in \Omega} X_i(\omega) p(\omega)$$

that is, if they can be associated to an expectation operator

Probability as a point in the space

Probability mass= a $|\Omega|$ dimensional vector

$$p := (p(\omega_1), \ldots, p(\omega_{|\Omega|}))$$

Limited to the set $\mathbb{P}$ of all probabilities

$$p(\omega) > 0, \quad \sum_{\omega \in \Omega} p(\omega) = 1 \quad \text{and}$$

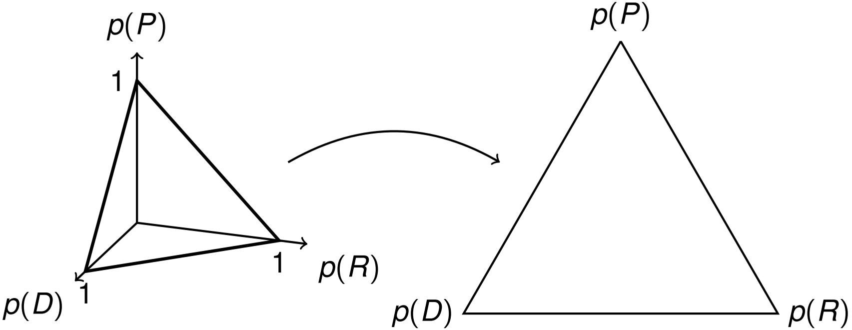
The set $\mathbb{P}$ is the $(n-1)$-unit simplex.

One price $P(X)$ gives a constraint of the type

$$\sum_{\omega \in \Omega} X(\omega) p(\omega)$$

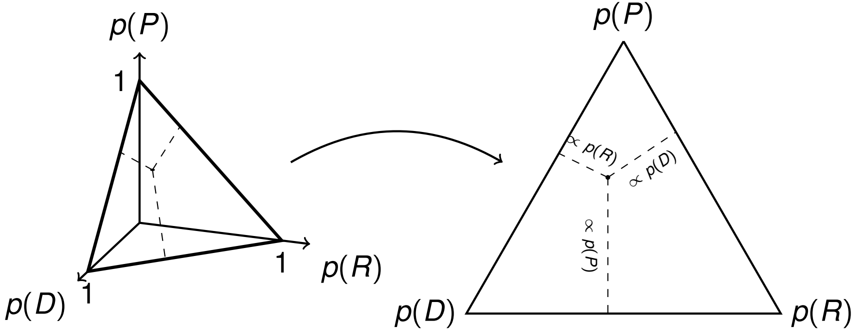
Point in $|\Omega|$ space

$$p(R) = 0.2, p(P) = 0.5, p(D) = 0.3$$



Point in $|\Omega|$ space

$$p(R) = 0.2, p(P) = 0.5, p(D) = 0.3$$

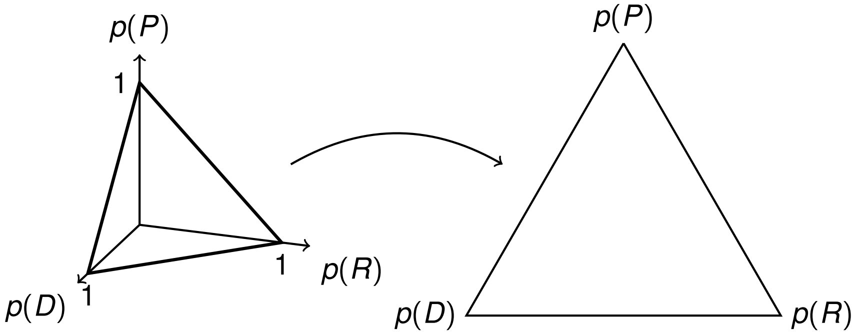


Assessments as constraints

$$p(R) = 0.2,\ p(P) = 0.5$$

$$\Downarrow$$

$$1 \cdot p(R) + 0 \cdot p(D) + 0 \cdot p(P) = 0.2$$
$$0 \cdot p(R) + 0 \cdot p(D) + 1 \cdot p(P) = 0.5$$

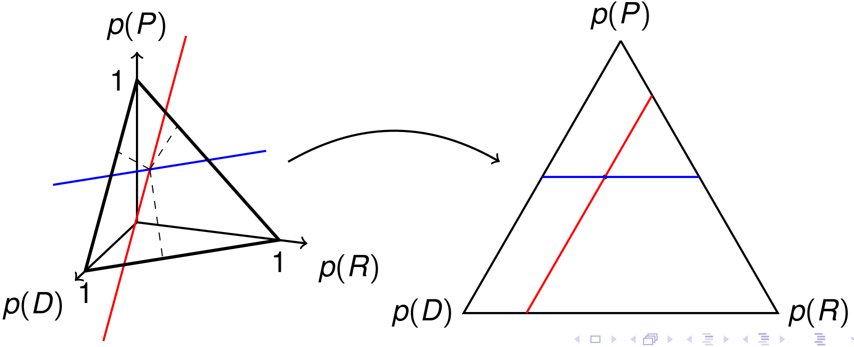


Assessments as constraints

$$p(R) = 0.2,\ p(P) = 0.5$$

$$\Downarrow$$

$$1 \cdot p(R) + 0 \cdot p(D) + 0 \cdot p(P) = 0.2$$
$$0 \cdot p(R) + 0 \cdot p(D) + 1 \cdot p(P) = 0.5$$

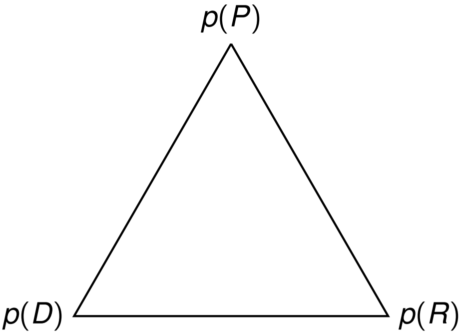


Other kinds of constraints

$$p(R)/p(P) = 0.2/0.5 = 0.4, \quad p(D)/p(P) = 0.3/0.5 = 0.6$$

$$\Downarrow$$

$$1 \cdot p(R) - 0.4 \cdot p(P) + 0 \cdot p(D) = 0$$
$$0 \cdot p(R) - 0.6 \cdot p(P) + 1 \cdot p(D) = 0$$

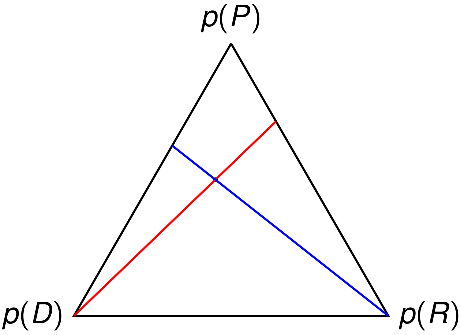


Other kinds of constraints

$$p(R)/p(P) = 0.2/0.5 = 0.4, \; p(D)/p(P) = 0.3/0.5 = 0.6$$

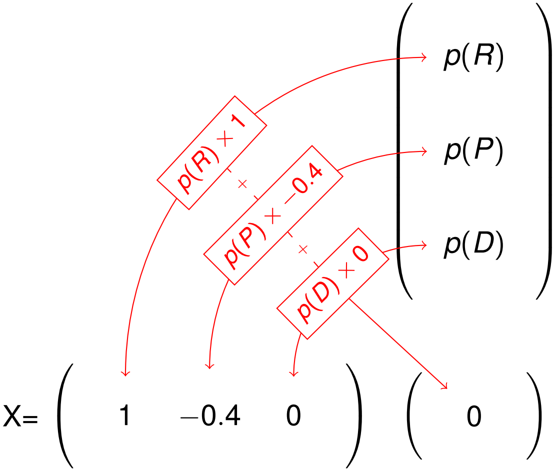
$$\Downarrow$$

$$1 \cdot p(R) - 0.4 \cdot p(P) + 0 \cdot p(D) = 0$$
$$0 \cdot p(R) - 0.6 \cdot p(P) + 1 \cdot p(D) = 0$$



Expectation as constraint

$$1 \cdot p(R) - 0.4 \cdot p(P) + 0 \cdot p(D) = 0$$



exercise

Form groups of 2/3 students, each person assessing separately price(s) for **one (or two) different** gamble among the three following ones that give 1 if events/assertions are true: Montpellier agglomeration (city + neighbouring villages) counts between

- 300K and 400K inhabitants
- 400K and 500K inhabitants
- 500K and 600K inhabitants

Compare your assessments. Discuss how you did to reach such prices? Are they rational? Can you agree on rational prices?

exercise

Represent the assessments

- $X_1(C) = 1,\ X_1(P) = -1,\ X_1(T) = 0,\ P(X_1) = 0$
- $X_2(C) = 0,\ X_2(P) = 1,\ X_2(T) = -1,\ P(X_2) = 0$
- $X_3(C) = 1,\ X_3(P) = -1,\ X_3(T) = 1,\ P(X_3) = 0$

and

- $X_1(C) = 2,\ X_1(P) = -1,\ X_1(T) = 0,\ P(X_1) = 1$
- $X_2(C) = 1,\ X_2(P) = 1,\ X_2(T) = -1,\ P(X_2) = 0$

In a geometric way, and check in this way that they are rational or not

Summary

- Two main interpretations
  - frequentist: applies to repeatable events
  - betting prices: applies to any uncertainty
- For the latter, a single price $\rightarrow$ buying=selling
- One price for a gamble X = one (linear) equality constraint on probability masses

Plan

1. Precise probabilities
2. Motivation for Imprecise probabilities (IP)
3. Lower previsions and imprecise probabilities
4. Conditioning in (imprecise) probability

Objectives

- Motivate IP through examples

Not ignorance, yet. . .

- Consider dice with six faces $\Omega = \{1, \dots, 6\}$
- You know (after 1000s of tests) that

$$P(\{1,6\}) = P(\{2,5\}) = P(\{3,4\}) = 1/3$$

- What is the probability
    - of drawing an odd number?
    - of scoring less or equal than 3?
    - of scoring less than 5?

What is the more probable?

- Spain will win 2022 world cup vs Brazil will win 2022 world cup
- We will be short of Oil first vs We will be short of drinkable water first
- In the next hour, average Wind power generation in Ireland will be above 100 MW vs below 100MW

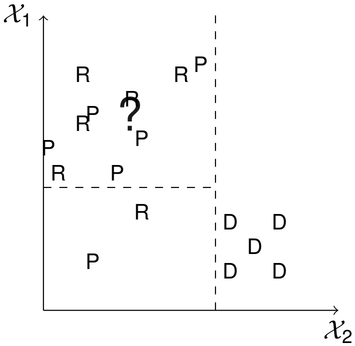
Selling/buying prices

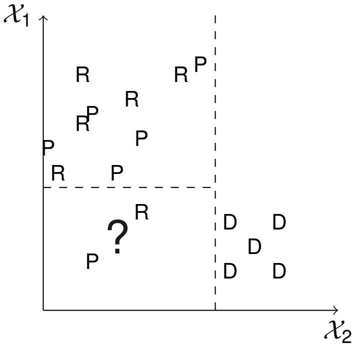
Ticket that pays 1 if the following events are true.

1. Mexico became independent in 1820
2. Mexico became independent between 1820 and 1830
3. Mexico became independent in the 19th century
4. Mexico became independent after 1500

- till what price would you be willing to buy each ticket?
- if you were to sell the ticket, what would you be your lowest price?

Classification problem





Ambiguity	Lack of information
$P(P\|?) \in [0.49, 0.51]$	$P(P\|?) \in [0.2, 0.8]$
$P(R\|?) \in [0.49, 0.51]$	$P(R\|?) \in [0.2, 0.8]$
$P(D\|?) \in [0, 0.02]$	$P(D\|?) \in [0, 0.6]$

$\Rightarrow$ difference immediate with imprecise probabilities

Applications

Situations that can happen in...

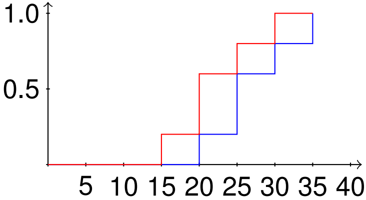
- Usual classification
- Environmental analysis
- Medical diagnosis
- Gesture recognition
- Sequence labelling
- ...

$\Rightarrow$ (machine) learning problems where knowing that we do not know is useful

Real-valued variable

What is the temperature in the room?

$$P([-\infty, 20]) = 0.2 \quad ; \quad P([-\infty, 25]) = 0.6 \quad ; \quad P([-\infty, 30]) = 0.8$$



Similar questions:

- what is the lifetime of this new component?
- what would be the core temperature in scenario B?
- how many years before the species is extinct?
- what will be the average sea level rise in 5 years?

Some final comments

- Precise assessments on some events does not mean that the probability will be uniquely defined
- Comparing probabilities requires a lot of information
- Same when providing probabilities, or unique selling/buying price, requires a lot of information

$\Rightarrow$ need to relax some assumptions, or provide a more general framework

Plan

1. Precise probabilities
2. Motivation for Imprecise probabilities (IP)
3. Lower previsions and imprecise probabilities
4. Conditioning in (imprecise) probability

Objectives

- Explain betting interpretation extension to imprecise probabilities
- Introduce avoiding sure loss and its robust interpretation
- Introduce natural extension and its robust interpretaiton
- Introduce coherence

Introduction

Next we shall see which are the most important rationality criteria and tools that allow us to work with imprecise probability models: avoiding sure loss, natural extension and coherence (in this order).

Other issues that shall be mostly tackled in other lectures, not this one, are:

- Deciding with IP models (this afternoon and Wednesday)
- Obtaining IP models from data and experts (Thursday)
- Computing with IP models (mainly tomorrow morning)

Interpretations

We shall be consider two different interpretations of IP models:

- The betting interpretation: we shall assess lower and upper betting prices on the outcomes of the experiment.
- The robust interpretation: imprecise probabilities model the lack of information about the (true) precise model of the experiment.

As we shall see later, both interpretations are formally equivalent.

Reminder

As before:

- Ω is the (finite) possible outcome space
- $X : \Omega \to \mathbb{R}$ is a *gamble* or ticket with

$$X(\omega) : \text{ reward if } \omega \text{ true}$$

$\hookrightarrow$ We shall see later why working with gambles is necessary.

Lower/upper previsions: definition

$\underline{P}(X)$: (your) maximal acceptable buying price for X

$\overline{P}(X)$: (your) minimum acceptable selling price for X

- X(R)=500, X(P)=20, X(D)=100
- till which price would you accept to **buy** the ticket?

Duality of buying and selling: example

Buying the ticket for $\underline{P}(X)$ gives a reward of

- $\underbrace{500}_{\text{you win}} - \underbrace{\underline{P}(X)}_{\text{you lose}}$ if R is true
- $20 - \underline{P}(X)$ if P is true
- $100 - \underline{P}(X)$ if D is true

which is the same as selling $-X$ for $\overline{P}(-X) = -\underline{P}(X)$

- $\underbrace{-\underline{P}(X)}_{\text{you win}} - \underbrace{(-500)}_{\text{you lose}}$ if R is true
- $-\underline{P}(X) - (-20)$ if P is true
- $-\underline{P}(X) - (-100)$ if D is true

Duality of buying and selling

For any gamble X, we do have

$$\overline{P}(X) = -\underline{P}(-X)$$

Formally, we can focus only on maximal buying prices, as any minimum selling price on X can be turned into a buying price

$\hookrightarrow$ We can only consider buying prices from a theoretical perspective (but selling prices may be useful, e.g., for elicitation or information presentation)

Lower probabilities of events

The *lower probability* of A, $\underline{P}(A)$

- $=$ lower prevision $\underline{P}(\mathbb{I}_A)$ of the indicator of A.
- $=$ supremum betting rate on A.
- $=$ measure of the *evidence* supporting A.
- $=$ measure of the strength of our *belief* in A.
- $= 1 - \overline{P}(A^c) = 1 - \overline{P}(\mathbb{I}_{A^c})$

Assessment example

- X(R)=500, X(P)=20, X(D)=100
- If I am certain that this is not a Provocateur (but nothing more), I should be willing to pay (buy) at least 100 for this, as I am sure to win at least this
- However, I should be willing to accept any selling price above 500, as I could lose that much
- Hence, $\underline{P}(X) = 100$ and $\overline{P}(X) = -\underline{P}(-X) = 500$

Are they rational **buying** prices?

- Assume the following prices:
  - $X_1(R) = 20$, $X_1(P) = -40$, $X_1(D) = 60$, $\underline{P}(X_1) = 20$
  - $X_2(R) = -20$, $X_2(P) = 40$, $X_2(D) = -60$, $\underline{P}(X_2) = -10$
- Assume the following prices:
  - $X_1(R) = 10$, $X_1(P) = -20$, $X_1(D) = 0$, $\underline{P}(X_1) = 0$
  - $X_2(R) = 0$, $X_2(P) = 10$, $X_2(D) = 0$, $\underline{P}(X_2) = 5$
- Assume the following prices:
  - $X_1(R) = 0$, $X_1(P) = 20$, $X_1(D) = -10$, $\underline{P}(X_1) = 0$
  - $X_2(R) = 20$, $X_2(P) = 0$, $X_2(D) = -10$, $\underline{P}(X_2) = 0$

Avoiding sure loss: general definition

Let $\underline{P}$ be a lower prevision defined on a (possibly infinite) set of gambles $\mathcal{K}$. It is said to *avoid sure loss* or *consistent iff*

$$\sup_{\omega\in\Omega}\sum_{i=1}^{n} c_i(X_i(\omega)-\underline{P}(X_i))\geq 0$$

for any $X_1,\ldots,X_n\in\mathcal{K}$ *and any* $c_i\in\mathbb{R}^+$*.*

That is, there is always a state of the world ω *in which we can win money, no matter how many (*c_i*) tickets* X_i *we buy of any gamble*

Representation theorem and robust interpretation

- $\underline{P}(X)$ as a lower bound of X (ill-known) expectation:

$$\underline{P}(X) \leq \sum_{\omega \in \Omega} X(\omega) p(\omega)$$

- Assessments $\underline{P}(X_1), \ldots, \underline{P}(X_n)$ avoids sure loss if and only if the set

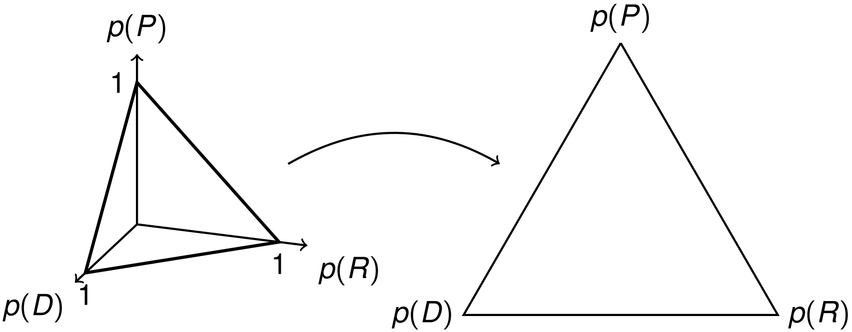
$$\mathcal{P}(\underline{P}) = \{p \in \mathbb{P} : \underline{P}(X_i) \leq \mathbb{E}(X_i),\ \forall X_i \in \mathcal{K}\}$$

  is not empty, with $\mathbb{E}$ the expectation of X_i with respect to p

- Or, in other words, if there is at least one probability mass p consistent with the assessments

Credal set example

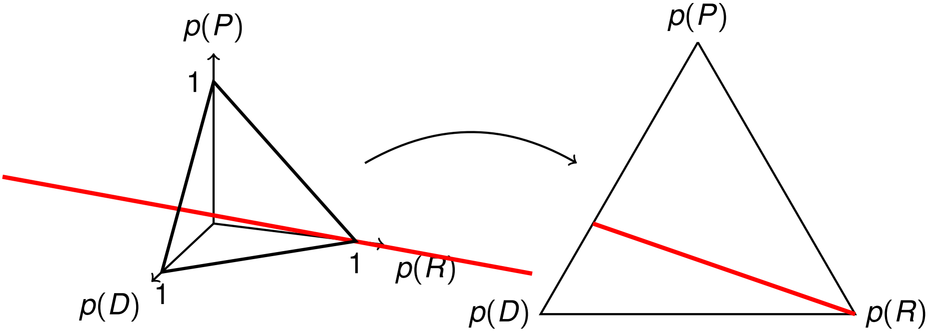
$$X_1(R) = 0,\ X_1(P) = 20,\ X_1(D) = -10,\ \underline{P}(X_1) = 0 \rightarrow$$
$$0 \leq 0p(R) + 20p(P) - 10p(D)$$



Credal set example

$$X_1(R) = 0,\ X_1(P) = 20,\ X_1(D) = -10,\ \underline{P}(X_1) = 0 \rightarrow$$
$$0 \leq 0p(R) + 20p(P) - 10p(D)$$

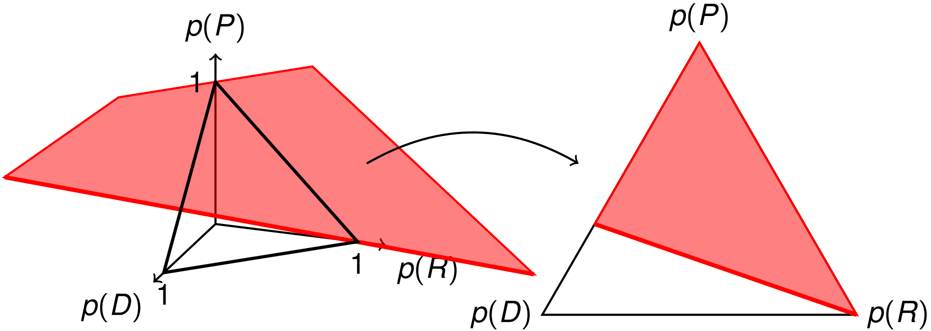
1. Line obtained by equality $0 = 20p(P) - 10p(D)$ or $2p(P) = p(D)$



Credal set example

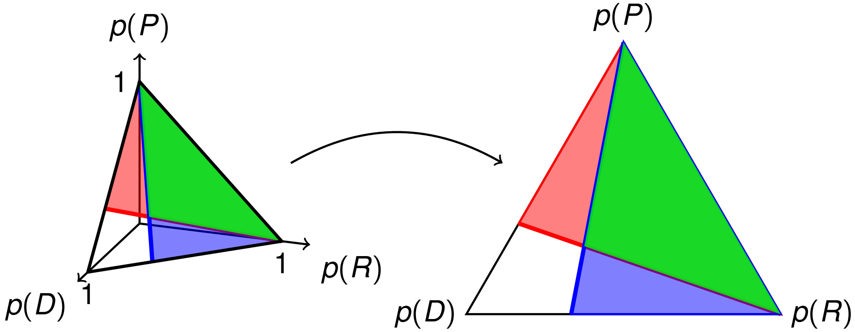
$$X_1(R) = 0,\ X_1(P) = 20,\ X_1(D) = -10,\ \underline{P}(X_1) = 0 \rightarrow$$
$$0 \leq 0p(R) + 20p(P) - 10p(D)$$

1. Line obtained by equality $0 = 20p(P) - 10p(D)$ or $2p(P) = p(D)$
2. Take side such that $0 \leq 20p(P) - 10p(D)$ or $p(D) \leq 2p(P)$



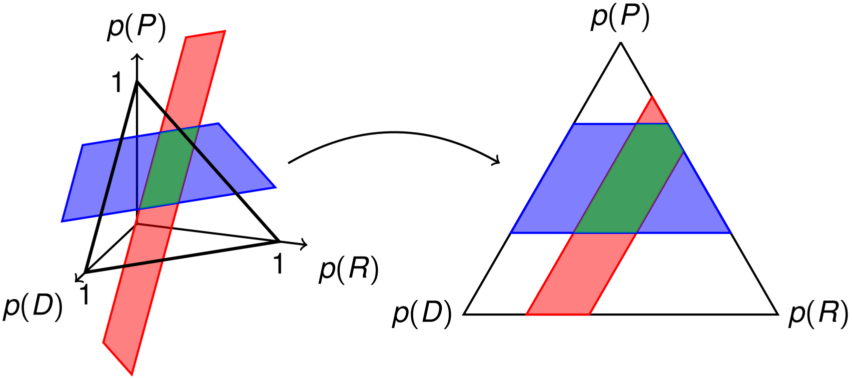
Credal set example

$$X_1(R) = 0,\ X_1(P) = 20,\ X_1(D) = -10,\ \underline{P}(X_1) = 0$$
$$X_2(R) = 20,\ X_2(P) = -10,\ X_2(D) = -10,\ \underline{P}(X_2) = 0$$
$$\mathcal{P}(\underline{P})$$



Another example

$0.2 \leq p(R) \leq 0.4$, $0.3 \leq p(P) \leq 0.7$



Alternative representation: extreme points

We can also see $\mathcal{P}(\underline{P})$ as a convex set of probability masses.
Each $p \in \mathcal{P}(\underline{P})$ a $|\Omega|$-vector, and p^* is an extreme point of $\mathcal{P}(\underline{P})$ iff

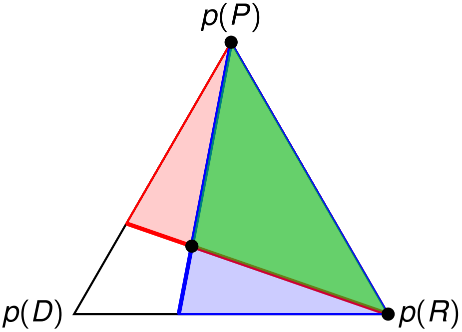
$$\nexists p_1, p_2 \in \mathcal{P}(\underline{P}) \text{ and } \lambda \in (0,1) \text{ with } p^* = \lambda p_1 + (1-\lambda) p_2$$

$\Rightarrow$ cannot be expressed as convex comb. of two other points.

$\hookrightarrow \mathcal{P}(\underline{P})$ entirely characterized by its set $\mathcal{E}(\underline{P})$ of extreme points.

Extreme points

- $p(R) = 1, p(D) = 0, p(P) = 0$
- $p(R) = 0, p(D) = 0, p(P) = 1$
- $p(R) = 0.25, p(D) = 0.5, p(P) = 0.25$



Exercise

Consider an urn with 10 balls, of which 3 are red, and the other 7 are either blue or yellow. We select one ball at random.

(a) Determine the set $\mathcal{P}$ of probabilities that represent the possible compositions of the urn.

(b) Which are the extreme points of this set?

(c) What are the lower and upper probabilities that the ball selected is blue?

Going beyond probabilities: an example

Before jumping off the wall, Humpty Dumpty tells Alice the following:

> *"I have a farm with pigs, cows and hens. There are at least as many pigs as cows and hens together, and at least as many hens as cows. How many pigs, cows and hens do I have?"*

- Which are the set of probabilities compatible with this information?
- Can we express them by using lower and upper probabilities of events?

Exercise

John is planning to bet on the winner of the Formula 1 championship. Determine the set of probabilities compatible with his beliefs, if he thinks that:

- Only one of Rosberg, Hamilton, Alonso or Vettel can win.
- The probability of Hamilton winning is at least twice as much of that of Alonso winning, and this is at least 1.5 times the probability of Vettel winning.
- Rosberg has exactly the same probability of winning than Hamilton.

Inference: natural extension

Consider the following gambles:

$$X_1(a) = 5, X_1(b) = 2, X_1(c) = -5, X_1(d) = -10$$
$$X_2(a) = 2, X_2(b) = -2, X_2(c) = 0, X_2(d) = 5$$

and assume we make the assessments $\underline{P}(X_1) = 2, \underline{P}(X_2) = 0$, that avoid sure loss. Can we deduce anything about how much should we pay for the gamble

$$Y(a) = 7, Y(b) = 4, Y(c) = -5, Y(d) = 0?$$

For instance, since $Y \geq X_1 + X_2$, we should be disposed to pay at least $\underline{P}(X_1) + \underline{P}(X_2) = 2$. But can we be more specific?

Definition

Consider a lower prevision $\underline{P}$ with domain $\mathcal{K}$, we seek to determine the consequences of the assessments in $\mathcal{K}$ on gambles outside the domain.

The *natural extension* of $\underline{P}$ to all gambles is given by

$$\underline{\mathbb{E}}(Y) := \sup\{\mu : \exists X_i \in \mathcal{K}, c_i \geq 0, i = 1, \ldots, n : \\ Y - \mu \geq \sum_{i=1}^{n} c_i(X_i(\omega) - \underline{P}(X_i))\}$$

$\underline{\mathbb{E}}(Y)$ is the supremum acceptable buying price for Y that can derived from the assessments on the gambles in the domain.

$\hookrightarrow$ If $\underline{P}$ does not avoid sure loss, then $\underline{\mathbb{E}}(Y) = +\infty$ for any gamble f.

Example

Applying this definition, we obtain that $\underline{\mathbb{E}}(Y) = 3.4$, by considering

$$Y - 3.4 \geq 1.2(X_1 - \underline{P}(X_1)).$$

Hence, the assessments $\underline{P}(X_1) = 2$, $\underline{P}(X_2) = 0$ imply that we should pay at least 3.4 for the gamble Y, but not more.

Natural extension: robustness interpretation

If assessments $\underline{P}(X_i)$ in $\mathcal{K}$ avoid sure loss , then the natural extension coincides with

$$\underline{\mathbb{E}}(Y) = \inf_{p \in \mathcal{P}(\underline{P})} \mathbb{E}(Y)$$

the lower expectation of Y taken over every $p \in \mathcal{P}(\underline{P})$

- Thus, we can give the natural extension with a *robust* interpretation.

Computing natural extension: basic methods

1. solve the linear program

$$\underline{\mathbb{E}}(Y) := \sup\{\mu : \exists X_i \in \mathcal{K}, c_i \geq 0, i = 1, \dots, n : Y - \mu \geq \sum_{i=1}^{n} c_i(X_i(\omega) - \underline{P}(X_i))\}$$

2. solve the linear program

$$\underline{\mathbb{E}}(Y) = \inf\{\sum_{\omega \in \Omega} Y(\omega)p(\omega) : \sum p(\omega) = 1, p(\omega) \geq 0, \underline{P}(X_i) \leq \sum_{\omega \in \Omega} X_i(\omega)p(\omega) \leq \overline{P}(X_i)\}$$

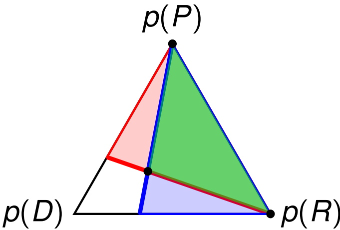
3. If you have extreme points $\mathcal{E}(\underline{P})$

$$\underline{\mathbb{E}}(Y) = \min_{p \in \mathcal{E}(\underline{P})} \mathbb{E}(Y).$$

$\hookrightarrow$ For the upper expectation, replace min/inf by max/sup.

Example

- $p(R) = 1, p(D) = 0, p(P) = 0$
- $p(R) = 0, p(D) = 0, p(P) = 1$
- $p(R) = 0.25, p(D) = 0.5, p(P) = 0.25$

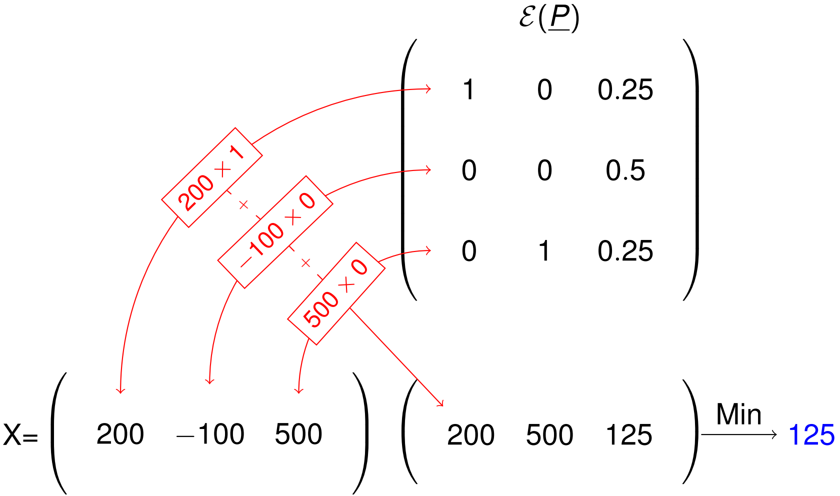


What is the natural extension of the following gamble:

$$Y(R) = 200,\ Y(D) = -100\ ,\ Y(P) = 500$$

(say you are a non-benefit organization fundraiser)

Lower expectation / natural extension: matrix



Exercise

Consider again the urn with 10 balls, of which 3 are red, and the other 7 are either blue or yellow. We select one ball at random.

(d) Assume I offer you the following gamble:

- lose 100 if you draw a red
- win 300 if you draw a yellow
- win 50 if you draw a blue

Given what you know about the urn, what would be your maximal buying price? your minimum selling?

Exercise

Consider $\Omega = \{1, 2, 3\}$, and the gambles and lower previsions given by

$$X_1(1) = 1, \quad X_1(2) = 2, \quad X_1(3) = 3$$
$$X_2(1) = 3, \quad X_2(2) = 2, \quad X_2(3) = 1$$

Assume we make the assessments $\underline{P}(X_1) = 2 = \underline{P}(X_2)$.

(a) Do these assessments avoid sure loss?

(b) Compute their natural extension on the gamble Y given by $Y(1) = 0$, $Y(2) = 1 = Y(3)$.

Going beyond probabilities: an example

Before going away, Humpty Dumpty tells Alice the following:

> *“I see you have some money with you. If you give me a coin, we will wait till an animal come out of the barn, and:*
>
> - *I will give you 1 if a pig comes out*
> - *I will give you 4 if a hen comes out*
> - *you will give me 1 if a cow comes out*
>
> *” Will Alice accept the gamble? Would she have accepted if she retained only the lower/upper probabilities?*

Exercise

Let $\underline{P}_A$ be the vacuous lower prevision relative to a set A, given by the assessment $\underline{P}_A(A) = 1$.

Prove that the natural extension $\underline{\mathbb{E}}$ of $\underline{P}_A$ is equal to the vacuous lower prevision relative to A:

$$\underline{\mathbb{E}}(X) = \underline{P}_A(X) = \inf_{\omega \in A} X(\omega),$$

for any $X \in \mathcal{L}(\Omega)$.

Exercise

John is planning to bet on the winner of the Formula 1 championship. Recall that he thnks that

- Only one of Rosberg, Hamilton, Alonso or Vettel can win.
- The probability of Hamilton winning is at least twice as much of that of Alonso winning, and this is at least 1.5 times the probability of Vettel winning.
- Rosberg has exactly the same probability of winning than Hamilton.

On his favourite betting website, they offer him a bet with reward 10 if Alonso wins, 5 if Vettel wins, and -3 if either Rosberg or Hamilton win. According to his beliefs, which are the minimum and maximum expected gains? Should he accept this bet or not?

Natural extension and coherence

We say that a lower prevision $\underline{P}$ with domain $\mathcal{K}$ is *coherent* when it coincides with its natural extension $\underline{\mathbb{E}}$ on $\mathcal{K}$.

If $\underline{P}$ avoids sure loss, then $\underline{\mathbb{E}}$ is the smallest coherent lower prevision on $\mathcal{L}(\Omega)$, the set of all possible gambles, that dominates $\underline{P}$ on $\mathcal{K}$.

Because of this, the natural extension can be regarded as the *least-committal* extension of $\underline{P}$: other coherent extensions will reflect stronger assessments than those present in $\underline{P}$.

Coherence: general definition

We can also check coherence directly, without computing the natural extension.

A lower prevision $\underline{P}$ is called *coherent* when given gambles $X_0, X_1, \dots, X_n$ in its domain and $m \in \mathbb{N}$,

$$\sup_{\omega\in\Omega}[\sum_{i=1}^{n}[X_i(\omega) - \underline{P}(X_i)] - m[X_0(\omega) - \underline{P}(X_0)]] \geq 0.$$

Otherwise, there is some $\epsilon > 0$ and ω such that

$$\sum_{i=1}^{n} X_i(\omega) - (\underline{P}(X_i) - \epsilon) < m(X_0(\omega) - \underline{P}(X_0) - \epsilon),$$

and $\underline{P}(X_0) + \epsilon$ would be an acceptable buying price for X_0.

Robust interpretation

The assessments

$$\underline{P}(X) \leq \mathbb{E}(X) \leq \overline{P}(X) \; \forall X \in \mathcal{K}$$

are coherent to iff, for any $X \in \mathcal{K}$, we have

$$\underline{\mathbb{E}}(X) = \underline{P}(X)$$

$$\overline{\mathbb{E}}(X) = \overline{P}(X)$$

In other words, they are best-possible bounds, and coincide with the lower envelope of $\mathcal{P}(\underline{P})$

Example

Consider the previously given assessments

- $X_1(R) = 0$, $X_1(P) = 20$, $X_1(D) = -10$, $\underline{P}(X_1) = 0$
- $X_2(R) = 20$, $X_2(P) = 0$, $X_2(D) = -10$, $\underline{P}(X_2) = 0$

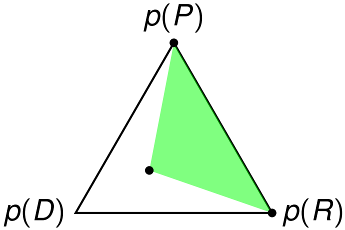
and the new assessment on

$$X_3(R) = 200, X_3(P) = 100, X_3(D) = 40$$

Given buying price is $\underline{P}(X_3) = 72$.

Are $\underline{P}(X_1), \underline{P}(X_2), \underline{P}(X_3)$ coherent?

Example



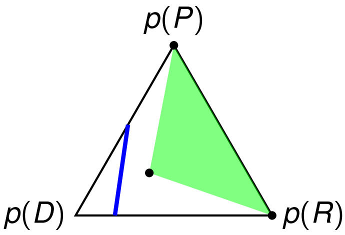
A new assessment on

$$X_3(P) = 100, X_3(D) = 40, X_3(R) = 200$$

Given buying price is $\underline{P}(X) = 72$

Are $\underline{P}(X_1), \underline{P}(X_2), \underline{P}(X_3)$ coherent?

Example



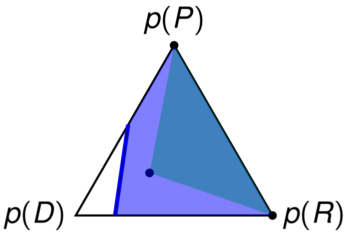
A new assessment on

$$X_3(P) = 100, X_3(D) = 40, X_3(R) = 200$$

Given buying price is $\underline{P}(X) = 72$

Are $\underline{P}(X_1), \underline{P}(X_2), \underline{P}(X_3)$ coherent?

Example



A new assessment on

$$X_3(P) = 100, X_3(D) = 40, X_3(R) = 200$$

Given buying price is $\underline{P}(X) = 72$

Are $\underline{P}(X_1), \underline{P}(X_2), \underline{P}(X_3)$ coherent?

Lower expectation / natural extension: matrix

$$\begin{pmatrix} 1 & 0 & 0.25 \\ 0 & 0 & 0.5 \\ 0 & 1 & 0.25 \end{pmatrix}$$

$$100 \times 1 + 40 \times 0 + 200 \times 0$$

$$X = \begin{pmatrix} 100 & 40 & 200 \end{pmatrix} \quad \begin{pmatrix} 100 & 200 & 95 \end{pmatrix} \xrightarrow{\text{Min}} 95$$

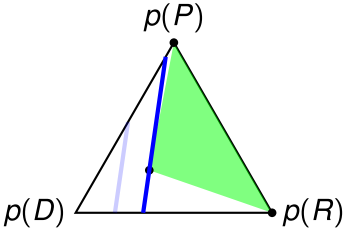
$\underline{P}(X_3)$ does not coincide with its natural extension $\underline{\mathbb{E}}(X_3)$!

Example

(corrected) assessments

- $X_1(R) = 0$, $X_1(P) = 20$, $X_1(D) = -10$, $\underline{P}(X_1) = 0$
- $X_2(R) = 20$, $X_2(P) = 0$, $X_2(D) = -10$, $\underline{P}(X_2) = 0$
- $X_3(R) = 200$, $X_3(P) = 100$, $X_3(D) = 40$, $\underline{P}(X_3) = 95$

are coherent!



Exercise

Consider the lower prevision given by:

	$X(a)$	$X(b)$	$X(c)$	$\underline{P}(X)$
X_1	2	1	0	0.5
X_2	0	1	2	1
X_3	0	1	0	1

(a) Does it avoid sure loss?

(b) Is it coherent?

Exercise

Mr. Play-it-safe is planning his upcoming holidays in the Canary Islands, and he is taking into account three possible disruptions: an unexpected illness (A), severe weather problems (B) and the unannounced visit of his mother in law (C).
He has assessed his lower and upper probabilities for these events:

	A	B	C	D
$\underline{P}$	0.05	0.05	0.2	0.5
$\overline{P}$	0.2	0.1	0.5	0.8

where *D* denotes the event 'Nothing bad happens'. He also assumes that no two disruptions can happen simultaneously.

(a) Determine the set of probabilities compatible with the assessments above.

(b) Are these lower probabilities coherent? If not, compute their natural extension.

Exercise

Consider the assessments

- $X_1(R) = 0$, $X_1(P) = 20$, $X_1(D) = -10$, $\underline{P}(X_1) = 0$
- $X_2(R) = 20$, $X_2(P) = 0$, $X_2(D) = -10$, $\underline{P}(X_2) = 0$
- $X_3(R) = 200, X_3(P) = 100, X_3(D) = 40, \underline{P}(X_3) = 120$

Answer these questions (in the order you deem fit)?

1. Do they avoid sure loss?
2. If yes, what are the extreme points of the associated probability set?
3. If yes to 1, Are they coherent?
4. If not to 3, which assessment(s) can be corrected?

Exercise

Let A be a non-empty subset of a (not necessarily finite) set Ω. Say we only know that the lower probability of A is equal to 1. This assessment is embodied through the lower prevision $\underline{P}$ defined on the singleton $\{I_A\}$ by $\underline{P}(A) = 1$. We extend it to all gambles by $\underline{P}(X) = \inf_{\omega \in A} X(\omega)$.

(a) Show that $\underline{P}$ avoids sure loss.

(b) Show that $\underline{P}$ is coherent.

Coherence on linear spaces

Suppose the domain $\mathcal{K}$ is a linear space of gambles:

- If $X, Y \in \mathcal{K}$, then $X + Y \in \mathcal{K}$.
- If $X \in \mathcal{K}, \lambda \in \mathbb{R}$, then $\lambda X \in \mathcal{K}$.

Then, $\underline{P}$ is coherent if and only if for any $X, Y \in \mathcal{K}, \lambda \geq 0$,

- $\underline{P}(X) \geq \inf X$.
- $\underline{P}(\lambda X) = \lambda \underline{P}(X)$.
- $\underline{P}(X + Y) \geq \underline{P}(X) + \underline{P}(Y)$.

Exercise

Let $\underline{P}$ be the lower prevision on $\mathcal{L}(\{1,2,3\})$ given by

$$\underline{P}(X) = \frac{\min\{X(1), X(2), X(3)\}}{2} + \frac{\max\{X(1), X(2), X(3)\}}{2}.$$

Is it coherent?

Exercise

Let $\underline{P}$ be a coherent lower prevision on $\mathcal{L}(\Omega)$, where $\Omega = \{0, 1\}$.

Prove that $\underline{P}$ is a *linear-vacuous* mixture, i.e., that there is some $\alpha \in [0, 1]$ and a linear prevision P on Ω such that

$$\underline{P} = \alpha P + (1 - \alpha)\underline{P}_{\Omega}.$$

Example: non-additive measures

As particular cases of coherent lower or upper probabilities, we have most of the models of non-additive measures existing in the literature. Let $\mu : \mathcal{P}(\Omega) \to [0, 1]$. It is called a *capacity* or *non-additive measure* when it satisfies:

1. $\mu(\emptyset) = 0, \mu(\Omega) = 1$ (normalisation).
2. $A \subseteq B \Rightarrow \mu(A) \leq \mu(B)$ (monotonicity).

Non-additive measures are used as alternative models to probability where we do not require the additivity axiom. We shall see more on fuzzy measures on Thursday.

Does the interpretation matter?

When we consider the credal set compatible with $\underline{P}$ we are giving a *robust interpretation* to our IP model:

- $\underline{P}$ describes an ill-known probability
- $\mathcal{P}(\underline{P})$ is the set of possible probabilities
- Should we know more, then we could identify the "true" probability

Always a unique probability?

Examples possibly fitting this interpretation

- Untested flipping coin
- Partially known dice example
- The Classification examples (to some extent)

But what about

- The temperature in this room ?
- The number of French people (i.e., having a french ID card/passport) today?

For those cases, it makes (more) sense to consider the *behavioural* interpretation of IP models (that also holds for other cases).

Coherent lower previsions: challenges

- Extension of the theory to *unbounded* gambles $\hookrightarrow$ M. Troffaes, G. de Cooman.
- The notion of coherence may be too weak.
- We are assuming that the utility scale is linear, which may not be reasonable in practice $\hookrightarrow$ R. Pelessoni, P. Vicig.

Lower prevision: summary and links

	Avoiding sure loss	Natural extension	Coherence
Behaviour.	$\sup_{\omega\in\Omega}\sum_{i=1}^{n} c_i (X_i(\omega) - \underline{P}(X_i)) \geq 0$	$\underline{\mathbb{E}}(Y) := \sup\{\mu : \exists X_i \in \mathcal{K}, c_i \geq 0, i = 1, \ldots, n : Y - \mu \geq \sum_{i=1}^{n} c_i(X_i(\omega) - \underline{P}(X_i))\}$	$\sup_{\omega\in\Omega} [\sum_{i=1}^{n}[X_i(\omega) - \underline{P}(X_i)] - m[X_0(\omega) - \underline{P}(X_0)]] \geq 0$
Robust	$\mathcal{P}(\underline{P}) \neq \emptyset$	$\underline{\mathbb{E}}(Y) = \inf_{p\in\mathcal{P}(\underline{P})} \mathbb{E}(Y)$	$\underline{P}(X_i) = \underline{\mathbb{E}}(X_i)$
Optim.	Feasible	Solving	Tight constraints
Logic	Consistent	Deduction	Deductively closed

Related works

- B. de Finetti.
- P. Williams.
- V. Kuznetsov.
- K. Weichselberger.
- G. Shafer and V. Vovk.

References: coherent lower previsions

The results and definitions from this part can be found in chapter 3 from:

- P. Walley, *Statistical reasoning with imprecise probabilities*. Chapman and Hall, 1991.

Additional references:

- T. Augustin, F. Coolen, M. Troffaes and G. de Cooman (eds.), *Introduction to imprecise probabilities*. Wiley, 2014.
- M. Troffaes and G. de Cooman, *Lower previsions*. Wiley, 2014.
- E. Miranda, *A survey of the theory of coherent lower previsions*. Int. J. of App. Reasoning, 48(2):628–658, 2008.

References on related works

- B. de Finetti, *Theory of Probability*. Wiley, 1974.
- V. Kuznetsov, *Interval Statistical models*. Radio and communication, 1991 (in Russian).
- K. Weischelberger, *Elementare Grundbegriffe einer allgemeineren Wahrscheinlichkeitsrechnung I: Intervallwahrscheinlichkeit als umfassendes Konzept*, Physica, Heidelberg, 2001.
- G. Shafer and V. Vovk, *Probability and finance: it's only a game!*. Wiley and Sons, 2001.
- P. Williams, *Notes on conditional previsions*. Technical Report, Univ. of Sussex, 1975.

Plan

Conditioning and conditionals

The conditional probability $P(A|B)$ can be defined

- From initial P by
$$P(A|B) = \frac{P(A \cap B)}{P(B)}$$
- Or by starting from the conditional value $P(A|B)$ and stating the relation
$$P(A|B)P(B) = P(A \cap B)$$

N.B.: the latter do not prevent $P(B) = 0$, the former do! (more than a simple rewriting)

However, for simplicity, we assume $P(B) > 0$ in this talk.

Conditional expectations

From the conditional probability $p(\cdot|B)$, we can also define a conditional expectation functional: for any gamble $X \in \mathcal{L}(\Omega)$, we have

$$P(X|B) = \frac{P(BX)}{P(B)},$$

where $P(BX)$ denotes the expectation of the function XI_B with respect to P. We can equivalently express the equation above as $P(B(X - P(X|B))) = 0$.

In our language, this conditional expectation will be called a *conditional lower prevision*. Next, we study the analogous concept in an imprecise context.

Example

- $p(R) = 0.2$, $p(P) = 0.5$, $p(D) = 0.3$
- We learn/observe $B = \{P, R\}$.
- Compute $p(R|B), p(P|B)$.
- Compute $p(X|B)$, where X is the gamble given by $X(P) = 2, X(R) = -2$.

The updated and the contingent interpretation

Consider a subset B of Ω, and a gamble X on Ω.

Under the *contingent* interpretation, $\underline{P}(X|B)$ is the supremum value of μ such that the gamble $I_B(X-\mu)$ is desirable for our subject.

We can also consider the *updated* interpretation, where $\underline{P}(X|B)$ is his supremum acceptable buying price for X, provided he later observes that the outcome of the experiment belongs to B.

Reconciling the two interpretations

Walley considers the *updating principle*: he calls a gamble X B-desirable when it is desirable provided the outcome of the experiment belongs to B.

The principle says that X is B-desirable if and only if $I_B X$ is desirable.

This relates present and future dispositions for the subject.

Conditional lower previsions

If we consider a partition $\mathcal{B}$ of Ω, we define $\underline{P}(X|\mathcal{B})$ as the gamble that takes the value $\underline{P}(X|B)$ on the elements of B. It is called a conditional lower prevision.

We shall always require that $\underline{P}(\cdot|\mathcal{B})$ is *separately coherent*, meaning that:

- For every $B \in \mathcal{B}$, the mapping $\underline{P}(\cdot|B) : \mathcal{L}(\Omega) \to \mathbb{R}$ is a coherent lower prevision.
- $\underline{P}(B|B) = 1$ for every $B \in \mathcal{B}$.

However, we need a tool to relate conditional and unconditional lower previsions.

Generalised Bayes Rule

Given a coherent lower prevision $\underline{P}$ on $\mathcal{L}(\Omega)$ and a partition $\mathcal{B}$ of Ω, our goal is to derive a conditional lower prevision $\underline{P}(\cdot|\mathcal{B})$ from $\underline{P}$.

We say that $\underline{P}, \underline{P}(\cdot|\mathcal{B})$ satisfy the *Generalised Bayes Rule* (GBR) when

$$\underline{P}(B(X - \underline{P}(X|B))) = 0 \ \forall X \in \mathcal{L}(\Omega), \forall B \in \mathcal{B}.$$

This means that if we interpret $\underline{P}(X|B)$ as the supremum buying price for X given B, we cannot combine it with those in $\underline{P}$ and get a sure loss.

Equivalent formulation

When $\underline{P}(B) > 0$, there is only one value of $\underline{P}(X|B)$ satisfying (GBR) with $\underline{P}$:

$$\underline{P}(X|B) = \inf\{P(X|B) : P \geq \underline{P}\}.$$

Thus, $\underline{P}(\cdot|B)$ is the lower envelope of the set of conditional previsions determined applying Bayes' rule on the elements of $\mathcal{P}(\underline{P})$.

Moreover, if $\mathcal{P}(\underline{P})$ has a finite number of extreme points, these can be used to determine the conditional lower prevision.

A robust interpretation

Consider a coherent lower prevision $\underline{P}$ on $\mathcal{L}(\Omega)$, let $\mathcal{B}$ be a finite partition of Ω such that $\underline{P}(B) > 0$ for every $B \in \mathcal{B}$ and define $\underline{P}(\cdot|\mathcal{B})$ by means of (GBR).

Then $\underline{P}, \underline{P}(\cdot|\mathcal{B})$ can be obtained as the lower envelopes of the set

$$\{P, P(\cdot|\mathcal{B}) : P \geq \underline{P}, P(\cdot|\mathcal{B}) \text{ derived from } P \text{ by Bayes' rule}\}.$$

Thus, the robust interpretation of coherent lower previsions carries on to the conditional case under these assumptions.

Exercise

Three horses (a,b and c) take part in a race. Our a priori lower probability for each horse being the winner is

$$\underline{P}(\{a\}) = 0.1, \qquad \underline{P}(\{b\}) = 0.25, \qquad \underline{P}(\{c\}) = 0.3,$$
$$\underline{P}(\{a, b\}) = 0.4, \qquad \underline{P}(\{a, c\}) = 0.6, \qquad \underline{P}(\{b, c\}) = 0.7.$$

There are rumors that c is not going to take part in the race due to some injury. What are the updated lower probabilities for a, b?

Exercise: the three prisoners problem

Three women, *a*, *b* and *c*, are in jail. Prisoner *a* knows that only two of the three prisoners will be executed, but she doesn't know who will be spared. She only knows that all three prisoners have equal probability $\frac{1}{3}$ of being spared.

To the warden who knows which prisoner will be spared, *a* says, "Since two out of the three will be executed, it is certain that either *b* or *c* will be. You will give me no information about my own chances if you give me the name of one man, *b* or *c*, who is going to be executed." Accepting this argument after some thinking, the warden says, "Prisoner *b* will be executed."

Does the warden's statement truly provide no information about the chance of *a* to be executed?

But things can get complicated!

The above procedure works if we consider a finite partition of the possibility space, and all the conditioning events have positive lower probability. But there are more general (and complicated) situations:

- Conditioning on sets of lower probability zero.
- Considering a partition with an infinite number of elements.
- Considering more than one partition.

Conditioning on sets of lower probability zero

If $\underline{P}(B) = 0$, there is more than one value of $\underline{P}(X|B)$ satisfying (GBR) with $\underline{P}$, so coherence is not enough. We can then consider:

- *Regular extension*: $\underline{P}(X|B) = \inf\{P(X|B) : P \geq \underline{P}, P(B) > 0\}$.
- *Natural extension*: $\underline{P}(X|B) = \inf_B X$.

They correspond to the greatest and the smallest models satisfying (GBR).

- Other approaches: *zero-layers* (Coletti/Scozzafava), *full conditional measures* (Dubins).

Infinite partitions: conglomerability

If $\mathcal{B}$ has an infinite number of elements, Walley's coherence of $\underline{P}, \underline{P}(\cdot|\mathcal{B})$ is equivalent to:

- $\underline{P}(B(X - \underline{P}(X|B))) = 0 \ \forall X, \forall B \in \mathcal{B}$. (GBR).
- $\underline{P}(\sum_{B \in \mathcal{B}}(B(X - \underline{P}(X|B)))) \geq 0 \ \forall X$ (conglomerability).

Conglomerability involves an infinite number of transactions, unlike what we have seen so far. If makes sense from a behavioral point of view but has a number of undesirable mathematical properties. Because of this, it is rejected by some authors, like de Finetti.

Conditioning on several partitions

More generally, we may want to derive more than one conditional lower prevision $\underline{P}(\cdot|\mathcal{B}_1), \ldots, \underline{P}(\cdot|\mathcal{B}_m)$ from our unconditional lower prevision $\underline{P}$. In that case, coherence can be extended in more than one manner:

- We may require that each $\underline{P}(\cdot|\mathcal{B}_j)$ is consistent with $\underline{P}$ (this is called *weak coherence*).

- We may require in addition that the assessments $\underline{P}(\cdot|\mathcal{B}_1), \ldots, \underline{P}(\cdot|\mathcal{B}_m)$ are also consistent with each other (this is called *coherence*).

These two notions are not equivalent! Moreover, the verification of coherence (which is the right notion from the behavioral point of view) is not immediate anymore.

References: conditional lower previsions

Most of the results on conditional lower previsions can be found in chapter 6 of:

- P. Walley, *Statistical reasoning with imprecise probabilities*, Chapman and Hall, 1991.

Additional references:

- E. Miranda, Fuzzy Sets and Systems, 160(9), 1286-1307, 2009.
- E. Miranda, M. Zaffalon, Artificial Intelligence, 173(1), 104-144, 2009. See also the paper at ISIPTA'09.
- E. Miranda, M. Zaffalon, G. de Cooman, Int.J. of Appr. Reasoning, 53(8), 1200-1227, 2012.
- G. Coletti and R. Scozzafava, *Probabilistic logic in a coherent setting*. Kluwer, 2002.