Coherent lower and upper previsions

and their behavioural interpretation

Gert de Cooman

gert.decooman@ugent.be

SYSTeMS research group, Ghent University

Overview

- General considerations about probability
- Epistemic probability: coherent lower previsions
- Decision making
- Conditional lower previsions

Part I

General considerations about probability

Two kinds of probabilities

- *Aleatory* probabilities
  - physical property, disposition
  - related to frequentist models
  - other names: objective, statistical or physical probability, chance

Two kinds of probabilities

- *Aleatory* probabilities
  - physical property, disposition
  - related to frequentist models
  - other names: objective, statistical or physical probability, chance
- *Epistemic* probabilities
  - model knowledge, information
  - represent strength of beliefs
  - other names: personal or subjective probability

Examples

- Weather forecast: probability of rain tomorrow
  - frequentist? physical?
  - personal probability of the forecaster, relating strength of his belief in rain tomorrow, based on his information $\Longrightarrow$ epistemic probability?

Examples

- Weather forecast: probability of rain tomorrow
  - frequentist? physical?
  - personal probability of the forecaster, relating strength of his belief in rain tomorrow, based on his information $\Longrightarrow$ epistemic probability?
- Laplace (1814): the probability that the sun won't rise tomorrow =
$$\frac{1}{1826215}$$
based on his 'knowledge': the sun has risen for 1826213 days since the day of Creation.

Examples

- Weather forecast: probability of rain tomorrow
  - frequentist? physical?
  - personal probability of the forecaster, relating strength of his belief in rain tomorrow, based on his information $\Longrightarrow$ epistemic probability?
- Laplace (1814): the probability that the sun won't rise tomorrow =
$$\frac{1}{1826215}$$
  based on his 'knowledge': the sun has risen for 1826213 days since the day of Creation.
- the distinction between aleatory–epistemic is present in probability theory from its earliest beginnings

Part II

Epistemic probability: coherent lower previsions

First observation

For many applications, we need theories to represent and reason with certain and uncertain knowledge

$$\begin{array}{rcl} \text{certain} & \longrightarrow & \text{logic} \\ \text{uncertain} & \longrightarrow & \text{probability theory} \end{array}$$

First observation

For many applications, we need theories to represent and reason with certain and uncertain knowledge

$$\begin{aligned} \text{certain} &\quad\longrightarrow\quad \text{logic} \\ \text{uncertain} &\quad\longrightarrow\quad \text{probability theory} \end{aligned}$$

One candidate: $\boxed{\text{Bayesian theory of probability}}$

First observation

For many applications, we need theories to represent and reason with certain and uncertain knowledge

$$\begin{array}{rcl} \text{certain} & \longrightarrow & \text{logic} \\ \text{uncertain} & \longrightarrow & \text{probability theory} \end{array}$$

One candidate: $\boxed{\text{Bayesian theory of probability}}$

I shall:

- argue that it is not general enough
- present the basic ideas behind a more general theory

$$\boxed{\text{imprecise probability theory (IP)}}$$

A theory of epistemic probability

Three pillars:

- how to measure epistemic probability?
- by what rules does epistemic probability abide?
- how can we use epistemic probability in reasoning, decision making, statistics . . . ?

A theory of epistemic probability

Three pillars:

- how to measure epistemic probability?
- by what rules does epistemic probability abide?
- how can we use epistemic probability in reasoning, decision making, statistics ...?

Notice that:

$$\begin{aligned} \text{1 and 2} &= \text{knowledge representation} \\ 3 &= \text{reasoning, inference} \end{aligned}$$

How to measure personal probability?

- Introspection
  - difficulty: how to convey and compare strengths of beliefs?
  - lack of a common standard

How to measure personal probability?

- Introspection
  - difficulty: how to convey and compare strengths of beliefs?
  - lack of a common standard
- belief = inclination to act
  - beliefs lead to behaviour, that can be used to measure their strength
  - special type of behaviour: accepting gambles
  - a gamble is a transaction/action/decision that yields different outcomes (utilities) in different states of the world.

Gambles – 1

- Consider a random variable X taking values in a set $\mathscr{X}$
- A gamble f is a bounded real-valued function on $\mathscr{X}$

$$f\colon \mathscr{X} \to \mathbb{R}\colon x \mapsto f(x)$$

- It can be interpreted as an uncertain reward

Gambles – 2

- Example: How did I come to Lugano?
- The random variable X is my means of transportation: by plane ($X = p$), by car ($X = c$) or by train ($X = t$)?
  - $\mathscr{X} = \{p, c, t\}$
  - $f(p) = -3,\ f(c) = 2,\ f(t) = 5$

Gambles – 2

- Example: How did I come to Lugano?
- The random variable X is my means of transportation: by plane ($X = p$), by car ($X = c$) or by train ($X = t$)?
  - $\mathscr{X} = \{p, c, t\}$
  - $f(p) = -3,\ f(c) = 2,\ f(t) = 5$
- Whether your accept this gamble or not will depend on your knowledge about how I came to Lugano
- Denote your set of desirable gambles by

$$\mathscr{D} \subseteq \mathscr{L}(\mathscr{X})$$

Modelling your uncertainty

- Accepting a gamble
  = taking a decision/action in the face of uncertainty

Modelling your uncertainty

- Accepting a gamble
  = taking a decision/action in the face of uncertainty
- Your set of desirable gambles contains the gambles that you accept
- It is a model for your uncertainty about which value X assumes (or will assume) in $\mathscr{X}$

Modelling your uncertainty

- Accepting a gamble
  = taking a decision/action in the face of uncertainty
- Your set of desirable gambles contains the gambles that you accept
- It is a model for your uncertainty about which value X assumes (or will assume) in $\mathscr{X}$
- More common models
  - (lower and upper) previsions
  - (lower and upper) probabilities
  - preference orderings
  - probability orderings
  - sets of probabilities

Desirability and rationality criteria

- Rewards are expressed in units of a linear utility scale

Desirability and rationality criteria

- Rewards are expressed in units of a linear utility scale
- Axioms: a set of desirable gambles $\mathscr{D}$ is coherent iff

D1. If $f \geq 0$ then $f \in \mathscr{D}$

D2. If $f, g \in \mathscr{D}$ then $f + g \in \mathscr{D}$

D3. If $f \in \mathscr{D}$ and $\lambda \geq 0$ then $\lambda f \in \mathscr{D}$

Desirability and rationality criteria

- Rewards are expressed in units of a linear utility scale
- Axioms: a set of desirable gambles $\mathscr{D}$ is coherent iff

D1. If $f \geq 0$ then $f \in \mathscr{D}$

D2. If $f, g \in \mathscr{D}$ then $f + g \in \mathscr{D}$

D3. If $f \in \mathscr{D}$ and $\lambda \geq 0$ then $\lambda f \in \mathscr{D}$

- Consequence: If $f \in \mathscr{D}$ and $g \geq f$ then $g \in \mathscr{D}$
- Consequence: If $f_1, \ldots, f_n \in \mathscr{D}$ and $\lambda_1, \ldots, \lambda_n \geq 0$ then $\sum_{k=1}^{n} \lambda_k f_k \in \mathscr{D}$
- A coherent set of desirable gambles is a convex cone of gambles that contains all non-negative gambles.

Definition of lower/upper prevision

Consider a gamble f on $\mathscr{X}$

- Buying f for a price μ yields a new gamble $f - \mu$
- the lower prevision $\underline{P}(f)$ of f
  = supremum acceptable price for buying f
  = supremum p such that $f - \mu$ is desirable for all $\mu < p$
  = $\sup\{\mu : f - \mu \in \mathscr{D}\}$

Definition of lower/upper prevision

Consider a gamble f on $\mathscr{X}$

- Buying f for a price μ yields a new gamble $f - \mu$
- the lower prevision $\underline{P}(f)$ of f
  - = supremum acceptable price for buying f
  - = supremum p such that $f - \mu$ is desirable for all $\mu < p$
  - = $\sup\{\mu : f - \mu \in \mathscr{D}\}$
- Selling f for a price μ yields a new gamble $\mu - f$
- the upper prevision $\overline{P}(f)$ of f
  - = infimum acceptable price for selling f
  - = infimum p such that $\mu - f$ is desirable for all $\mu > p$
  - = $\inf\{\mu : \mu - f \in \mathscr{D}\}$

Lower and upper prevision – 1

- Selling a gamble f for price μ
  = buying $-f$ for price $-\mu$:

$$\mu - f = (-f) - (-\mu)$$

Lower and upper prevision – 1

- Selling a gamble f for price μ
  = buying $-f$ for price $-\mu$:
  $$\mu - f = (-f) - (-\mu)$$

- Consequently:

$$\begin{aligned}\overline{P}(f) &= \inf\{\mu : \mu - f \in \mathscr{D}\}\\ &= \inf\{-\lambda : -f - \lambda \in \mathscr{D}\}\\ &= -\sup\{\lambda : -f - \lambda \in \mathscr{D}\}\\ &= -\underline{P}(-f)\end{aligned}$$

Lower and upper prevision – 2

$\underline{P}(f) = \sup\{\mu : f - \mu \in \mathscr{D}\}$

- if you specify a lower prevision $\underline{P}(f)$, you are committed to accepting

$$f - \underline{P}(f) + \varepsilon = f - [\underline{P}(f) - \varepsilon]$$

for all $\varepsilon > 0$ (but not necessarily for $\varepsilon = 0$).

Lower and upper prevision – 2

$\underline{P}(f) = \sup\{\mu : f - \mu \in \mathscr{D}\}$

- if you specify a lower prevision $\underline{P}(f)$, you are committed to accepting

$$f - \underline{P}(f) + \varepsilon = f - [\underline{P}(f) - \varepsilon]$$

for all $\varepsilon > 0$ (but not necessarily for $\varepsilon = 0$).

$\overline{P}(f) = \inf\{\mu : \mu - f \in \mathscr{D}\}$

- if you specify an upper prevision $\overline{P}(f)$, you are committed to accepting

$$\overline{P}(f) - f + \varepsilon = [\overline{P}(f) + \varepsilon] - f$$

for all $\varepsilon > 0$ (but not necessarily for $\varepsilon = 0$).

Precise previsions

When lower and upper prevision for f coincide:

$$\underline{P}(f) = \overline{P}(f) = P(f)$$

is called the (precise) prevision of f

- $P(f)$ is a prevision, or fair price in de Finetti's sense
- Previsions are the precise, or Bayesian, probability models

Precise previsions

When lower and upper prevision for f coincide:

$$\underline{P}(f) = \overline{P}(f) = P(f)$$

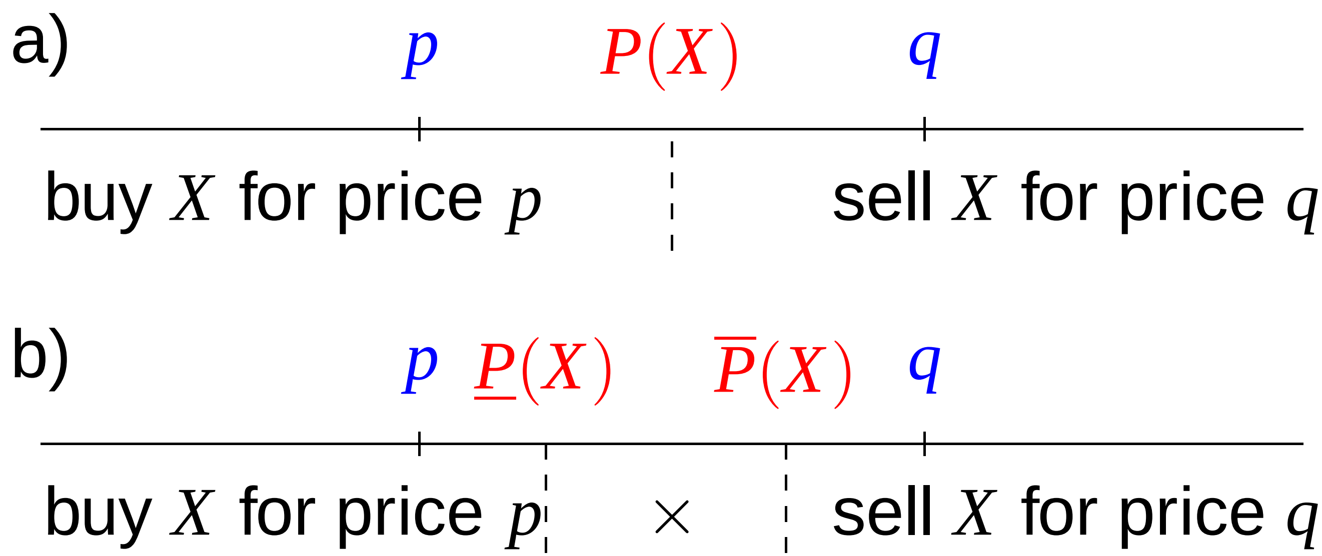
is called the (precise) prevision of f

- $P(f)$ is a prevision, or fair price in de Finetti's sense
- Previsions are the precise, or Bayesian, probability models
- if you specify a prevision $P(f)$, you are committed to accepting

$$[P(f) + \varepsilon] - f \text{ and } f - [P(f) - \varepsilon]$$

for all $\varepsilon > 0$ (but not necessarily for $\varepsilon = 0$).

Allowing for indecision



- Specifying a precise prevision $P(f)$ means that you choose, for essentially any real price p, between buying f for price p or selling f for that price
- Imprecise models allow for indecision!

Events and lower probabilities

- An event A is a subset of $\mathscr{X}$
- Example: the event $\{c,t\}$ that I did not come by plane to Lugano.

Events and lower probabilities

- An event A is a subset of $\mathscr{X}$
- Example: the event $\{c,t\}$ that I did not come by plane to Lugano.
- It can be identied with a special gamble I_A on $\mathscr{X}$

$$I_A(x) = \begin{cases} 1 & \text{if } x \in A\text{, i.e., if } A \text{ occurs} \\ 0 & \text{if } x \notin A\text{, i.e., if } A \text{ doesn't occur} \end{cases}$$

Events and lower probabilities

- An event A is a subset of $\mathscr{X}$
- Example: the event $\{c,t\}$ that I did not come by plane to Lugano.
- It can be identied with a special gamble I_A on $\mathscr{X}$

$$I_A(x) = \begin{cases} 1 & \text{if } x \in A\text{, i.e., if } A \text{ occurs} \\ 0 & \text{if } x \notin A\text{, i.e., if } A \text{ doesn't occur} \end{cases}$$

- The lower probability $\underline{P}(A)$ of A
  - = lower prevision $\underline{P}(I_A)$ of indicator I_A
  - = supremum rate for betting on A
  - = measure of evidence in favour of A
  - = measure of (strength of) belief in A

Upper probabilities

- The upper probability $\overline{P}(A)$ of A
  - = the upper prevision $\overline{P}(I_A) = \overline{P}(1 - I_{A^c}) = 1 - \underline{P}(I_{A^c})$ of the indicator I_A
  - = measures lack of evidence against A
  - = measures the plausibility of A

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

Upper probabilities

- The upper probability $\overline{P}(A)$ of A
  - = the upper prevision $\overline{P}(I_A) = \overline{P}(1 - I_{A^c}) = 1 - \underline{P}(I_{A^c})$ of the indicator I_A
  - = measures lack of evidence against A
  - = measures the plausibility of A

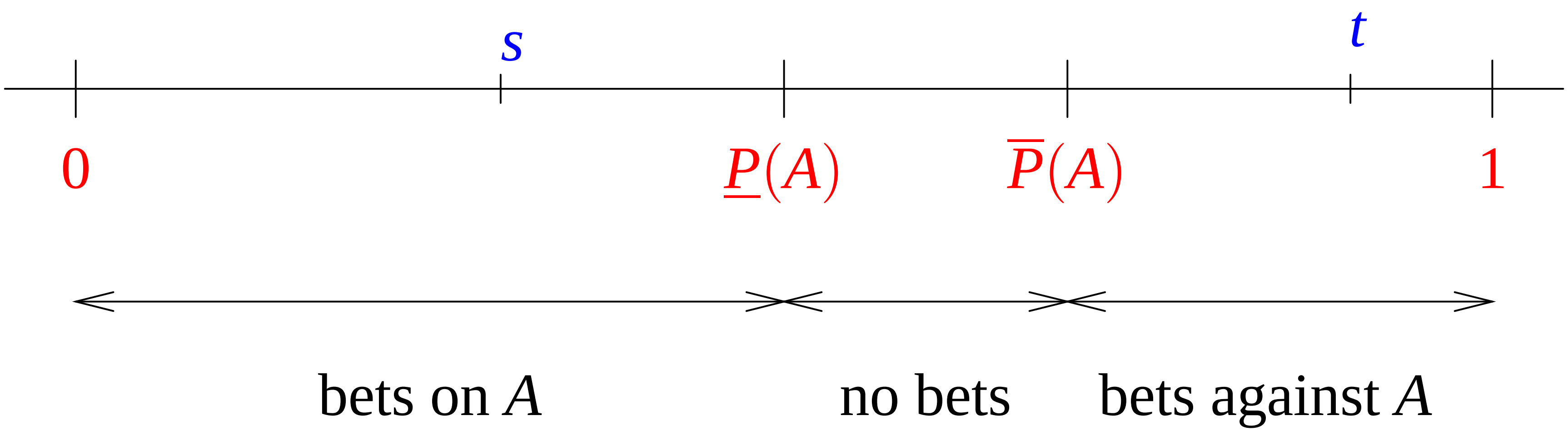
$$\overline{P}(A) = 1 - \underline{P}(A^c)$$

- This gives a behavioural interpretation to lower and upper probability

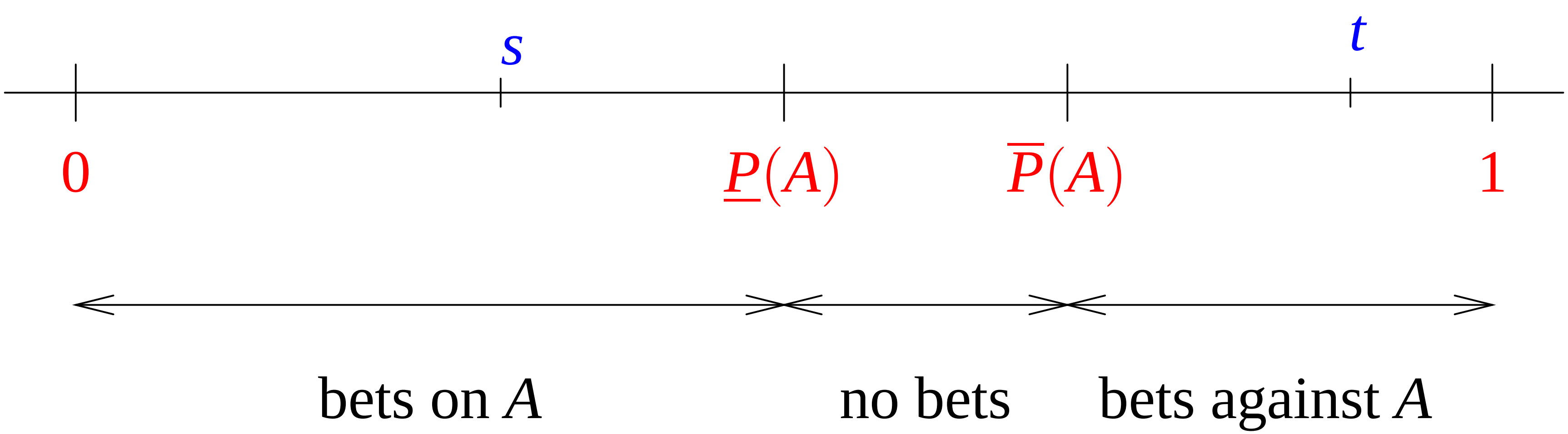
$$\text{evidence for } A \uparrow \quad \Rightarrow \quad \underline{P}(A) \uparrow$$

$$\text{evidence against } A \uparrow \quad \Rightarrow \quad \overline{P}(A) \downarrow$$

Precise probability

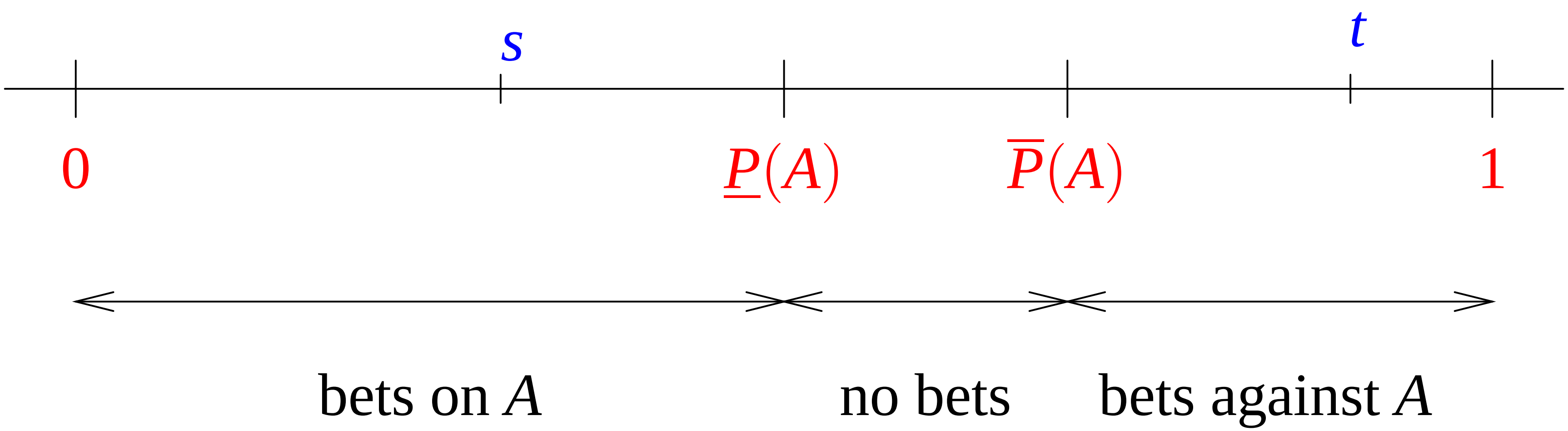


Precise probability



When $\underline{P}(A) = \overline{P}(A)$ then the common value is called the probability $P(A)$ of A

Precise probability



When $\underline{P}(A) = \overline{P}(A)$ then the common value is called the probability $P(A)$ of A

Precise probability theory

- is a special case of imprecise probability theory
- makes much more stringent assumptions about a subject's behavioural dispositions

Why imprecision?

- A forced choice will not necessarily reflect your beliefs or knowledge; it may be arbitrary

Why imprecision?

- A forced choice will not necessarily reflect your beliefs or knowledge; it may be arbitrary
- Choice is not the same as preference:
  - you may not have enough knowledge to really prefer one action over another (e.g., betting on vs. betting against A); you may be undecided

Why imprecision?

- A forced choice will not necessarily reflect your beliefs or knowledge; it may be arbitrary
- Choice is not the same as preference:
  - you may not have enough knowledge to really prefer one action over another (e.g., betting on vs. betting against A); you may be undecided
  - when forced to choose without a real preference, you will make a choice, but this will not reflect your beliefs, and it will in this sense be arbitrary

Why imprecision?

- A forced choice will not necessarily reflect your beliefs or knowledge; it may be arbitrary
- Choice is not the same as preference:
  - you may not have enough knowledge to really prefer one action over another (e.g., betting on vs. betting against A); you may be undecided
  - when forced to choose without a real preference, you will make a choice, but this will not reflect your beliefs, and it will in this sense be arbitrary
- To enforce precision emphasises choice
  To allow imprecision emphasises preference

Why imprecision?

- A forced choice will not necessarily reflect your beliefs or knowledge; it may be arbitrary
- Choice is not the same as preference:
  - you may not have enough knowledge to really prefer one action over another (e.g., betting on vs. betting against A); you may be undecided
  - when forced to choose without a real preference, you will make a choice, but this will not reflect your beliefs, and it will in this sense be arbitrary
- To enforce precision emphasises choice
  To allow imprecision emphasises preference
- You do not use a precise model automatically, but only when you have sufficient information!

Rules of epistemic probability

- Lower and upper previsions represent commitments to act/behave in certain ways
- Rules that govern lower and upper previsions reflect rationality of behaviour.

Rules of epistemic probability

- Lower and upper previsions represent commitments to act/behave in certain ways
- Rules that govern lower and upper previsions reflect rationality of behaviour.
- Your behaviour is considered to be irrational when
  - it is *harmful to yourself*: specifying betting rates such that you lose utility, whatever the outcome $\Longrightarrow$ avoiding sure loss (cf. logical consistency)

Rules of epistemic probability

- Lower and upper previsions represent commitments to act/behave in certain ways
- Rules that govern lower and upper previsions reflect rationality of behaviour.
- Your behaviour is considered to be irrational when
  - it is *harmful to yourself*: specifying betting rates such that you lose utility, whatever the outcome $\Longrightarrow$ avoiding sure loss (cf. logical consistency)
  - it is *inconsistent*: you are not fully aware of the implications of your betting rates $\Longrightarrow$ coherence (cf. logical closure)

Avoiding sure loss

- Example: two bets

on A:	$I_A - \underline{P}(A)$
on A^c:	$I_{A^c} - \underline{P}(A^c)$
together:	$1 - [\underline{P}(A) + \underline{P}(A^c)]$ should be ≥ 0

Avoiding sure loss

- Example: two bets

$$\begin{array}{ll} \text{on } A: & I_A - \underline{P}(A) \\ \text{on } A^c: & I_{A^c} - \underline{P}(A^c) \\ \hline \text{together:} & 1 - [\underline{P}(A) + \underline{P}(A^c)] \text{ should be } \geq 0 \end{array}$$

- Avoiding a sure loss therefore implies

$$\underline{P}(A) + \underline{P}(A^c) \leq 1,$$

or in other words

$$\boxed{\underline{P}(A) \leq \overline{P}(A)}$$

Avoiding sure loss: general condition

A set of gambles $\mathscr{K}$ and a lower prevision $\underline{P}$ defined for each gamble in $\mathscr{K}$.

Definition 1. $\underline{P}$ avoids sure loss if for all $n \geq 0$, $f_1, \ldots, f_n$ in $\mathscr{K}$ and for all non-negative $\lambda_1, \ldots, \lambda_n$:

$$\sup_{x \in \mathscr{X}} \left[\sum_{k=1}^{n} \lambda_k [f_k(x) - \underline{P}(f_k)] \right] \geq 0.$$

Avoiding sure loss: general condition

A set of gambles $\mathscr{K}$ and a lower prevision $\underline{P}$ defined for each gamble in $\mathscr{K}$.

Definition 1. $\underline{P}$ avoids sure loss if for all $n \geq 0$, $f_1, \dots, f_n$ in $\mathscr{K}$ and for all non-negative $\lambda_1, \dots, \lambda_n$:

$$\sup_{x \in \mathscr{X}} \left[\sum_{k=1}^{n} \lambda_k [f_k(x) - \underline{P}(f_k)] \right] \geq 0.$$

If not, then there are $\varepsilon > 0$, $n \geq 0$, $f_1, \dots, f_n$ and non-negative $\lambda_1, \dots, \lambda_n$ such that:

$$\sum_{k=1}^{n} \lambda_k [f_k - \underline{P}(f_k) + \varepsilon] \leq -\varepsilon!$$

A few consequences of avoiding sure loss

a. $\underline{P}(\emptyset) \leq 0$ and $\overline{P}(\mathscr{X}) \geq 1$

b. $A \subseteq B \Rightarrow \underline{P}(A) \leq \overline{P}(B)$

c. $\underline{P}(f) \leq \sup f$, $\overline{P}(f) \geq \inf f$ and $\underline{P}(f) \leq \overline{P}(f)$

d. $\underline{P}(\mu) \leq \mu \leq \overline{P}(\mu)$

e. $\underline{P}(f) + \underline{P}(\mu - f) \leq \mu$

f. if $f \geq g + \mu$ then $\overline{P}(f) \geq \underline{P}(g) + \mu$

g. $\underline{P}(f+g) \leq \overline{P}(f) + \overline{P}(g)$ and $\overline{P}(f+g) \geq \underline{P}(f) + \underline{P}(g)$

h. $\underline{P}(\lambda f) \leq \lambda \overline{P}(f)$ and $\overline{P}(\lambda f) \geq \lambda \underline{P}(f)$ for $\lambda \geq 0$

whenever the arguments are in the domains of $\underline{P}$ (or $\overline{P}$).

Coherence

- Example: two bets involving A and B with $A \cap B = \emptyset$

$$
\begin{array}{ll}
\text{on } A: & I_A - \underline{P}(A) \\
\text{on } B: & I_B - \underline{P}(B) \\
\hline
\text{together:} & I_{A \cup B} - [\underline{P}(A) + \underline{P}(B)]
\end{array}
$$

Coherence

- Example: two bets involving A and B with $A \cap B = \emptyset$

$$\begin{array}{ll}
\text{on } A\text{:} & I_A - \underline{P}(A) \\
\text{on } B\text{:} & I_B - \underline{P}(B) \\
\hline
\text{together:} & I_{A \cup B} - [\underline{P}(A) + \underline{P}(B)]
\end{array}$$

- Coherence implies that

$$\boxed{\underline{P}(A) + \underline{P}(B) \leq \underline{P}(A \cup B)}$$

Coherence: general condition

A set of gambles $\mathscr{K}$ and a lower prevision $\underline{P}$ defined for each gamble in $\mathscr{K}$.

Coherence: general condition

A set of gambles $\mathscr{K}$ and a lower prevision $\underline{P}$ defined for each gamble in $\mathscr{K}$.

Definition 2. $\underline{P}$ is coherent if for all $n \geq 0$, $f_0, f_1, \ldots, f_n$ in $\mathscr{K}$ and for all non-negative $\lambda_0, \lambda_1, \ldots, \lambda_n$:

$$\sup_{x \in \mathscr{X}} \left[\sum_{k=1}^{n} \lambda_k [f_k(x) - \underline{P}(f_k)] - \lambda_0 [f_0 - \underline{P}(f_0)] \right] \geq 0.$$

If not, then there are $\varepsilon > 0$, $n \geq 0$, $f_0, f_1, \ldots, f_n$ and non-negative $\lambda_1, \ldots, \lambda_n$ such that:

$$f_0(x) - [\underline{P}(f_0) + \varepsilon] \geq \sum_{k=1}^{n} \lambda_k [f_k - \underline{P}(f_k) + \varepsilon]!$$

A few consequences of coherence – 1

Whenever the arguments are in the domains of $\underline{P}$ (or $\overline{P}$):

a. $\underline{P}(\emptyset) = \overline{P}(\emptyset) = 0$ and $\underline{P}(\mathscr{X}) = \overline{P}(\mathscr{X}) = 1$

b. if $A \subseteq B$ then $\underline{P}(A) \leq \underline{P}(B)$ and $\overline{P}(A) \leq \overline{P}(B)$

c. $\inf f \leq \underline{P}(f) \leq \overline{P}(f) \leq \sup f$

d. $\underline{P}(\mu) = \overline{P}(\mu) = \mu$

e. $\underline{P}(f+\mu) = \underline{P}(f) + \mu$ and $\overline{P}(f+\mu) = \overline{P}(f) + \mu$

f. if $f \geq g + \mu$ then $\underline{P}(f) \geq \underline{P}(g) + \mu$ and $\overline{P}(f) \geq \overline{P}(g) + \mu$

g. $\underline{P}(f) + \underline{P}(g) \leq \underline{P}(f+g) \leq \underline{P}(f) + \overline{P}(g) \leq \overline{P}(f+g) \leq \overline{P}(f) + \overline{P}(g)$

h. $\underline{P}(\lambda f) = \lambda \underline{P}(f)$ and $\overline{P}(\lambda f) = \lambda \overline{P}(f)$ for $\lambda \geq 0$

A few consequences of coherence – 2

Whenever the arguments are in the domains of $\underline{P}$ (or $\overline{P}$):

a. $\underline{P}(f) \leq \underline{P}(|f|)$ and $\overline{P}(f) \leq \overline{P}(|f|)$

b. $|\underline{P}(f) - \underline{P}(g)| \leq \overline{P}(|f-g|)$ and $|\overline{P}(f) - \overline{P}(g)| \leq \overline{P}(|f-g|)$

c. if $\overline{P}(|f_n - f|) \to 0$ as $n \to \infty$ then $\underline{P}(f_n) \to \underline{P}(f)$ and $\overline{P}(f_n) \to \overline{P}(f)$

A few consequences of coherence – 2

Whenever the arguments are in the domains of $\underline{P}$ (or $\overline{P}$):

a. $\underline{P}(f) \leq \underline{P}(|f|)$ and $\overline{P}(f) \leq \overline{P}(|f|)$

b. $|\underline{P}(f) - \underline{P}(g)| \leq \overline{P}(|f-g|)$ and $|\overline{P}(f) - \overline{P}(g)| \leq \overline{P}(|f-g|)$

c. if $\overline{P}(|f_n - f|) \to 0$ as $n \to \infty$ then $\underline{P}(f_n) \to \underline{P}(f)$ and $\overline{P}(f_n) \to \overline{P}(f)$

Important consequences of c.:

- if $f_n \to f$ uniformly (i.e., $\sup|f_n - f| \to 0$) then $\underline{P}(f_n) \to \underline{P}(f)$ and $\overline{P}(f_n) \to \overline{P}(f)$
- every gamble f is a uniform limit of a sequence of simple gambles s_n, and

$$\underline{P}(f) = \lim_{n\to\infty} \underline{P}(s_n) \quad \text{and} \quad \overline{P}(f) = \lim_{n\to\infty} \overline{P}(s_n)$$

Important remarks

- The domain $\mathscr{K}$ of a lower prevision
  - is the set of those gambles for which lower prevision assessments are available
  - need not have any predefined structure
  - may contain (indicators of) events

Important remarks

- The domain $\mathscr{K}$ of a lower prevision
  - is the set of those gambles for which lower prevision assessments are available
  - need not have any predefined structure
  - may contain (indicators of) events
- a convex combination of coherent lower previsions is coherent
- a lower envelope of coherent lower previsions is coherent
- a point-wise limit (inferior) of coherent lower previsions is coherent

Coherence on a linear space

Suppose that the domain $\mathscr{K}$ is a linear space:

- if $f \in \mathscr{K}$ and $g \in \mathscr{K}$ then $f + g \in \mathscr{K}$
- if $f \in \mathscr{K}$ and $\lambda \in \mathbb{R}$ then $\lambda f \in \mathscr{K}$.

Coherence on a linear space

Suppose that the domain $\mathscr{K}$ is a linear space:

- if $f \in \mathscr{K}$ and $g \in \mathscr{K}$ then $f+g \in \mathscr{K}$
- if $f \in \mathscr{K}$ and $\lambda \in \mathbb{R}$ then $\lambda f \in \mathscr{K}$.

Theorem 1. *.Let the lower prevision $\underline{P}$ be defined on a linear space $\mathscr{K}$. Then $\underline{P}$ is coherent if and only if for all f, g in $\mathscr{K}$ and $\lambda \geq 0$,*

P1. $\underline{P}(f) \geq \inf f$ *[accepting sure gains]*

P2. $\underline{P}(\lambda f) = \lambda \underline{P}(f)$ *[positive homogeneity]*

P3. $\underline{P}(f+g) \geq \underline{P}(f) + \underline{P}(g)$ *[superlinearity]*

Precise previsions – 1

Consider a domain $\mathscr{K}'$, a lower prevision $\underline{P}$ and a conjugate upper prevision $\overline{P}$ that coincide on $\mathscr{K}'$:

$$\underline{P}(f) = \overline{P}(f) = P(f), \quad \forall f \in \mathscr{K}'$$

There is a unique way to extend this to a functional P on

$$\mathscr{K} = \mathscr{K}' \cup -\mathscr{K}', \text{ whence } \mathscr{K} = -\mathscr{K},$$

that is self-conjugate:

$$\boxed{P(f) = -P(-f), \quad \forall f \in \mathscr{K}.}$$

Precise and linear previsions

Defi nition 3. .

1. A self-conjugate P on a negation-invariant domain $\mathscr{K} = -\mathscr{K}$ is called a precise prevision. It is interpreted as a fair price.
2. A (precise) prevision is called coherent when it is coherent both as a lower and as an upper prevision.
3. We also call a coherent precise prevision a linear prevision.

- coincides with de Finetti's notion of a coherent prevision
- lineair previsions are the coherent precise probability models

Linear previsions

- a precise prevision P on $\mathscr{L}(\mathscr{X})$ is coherent iff
  - $P(\lambda f + \mu g) = \lambda P(f) + \mu P(g)$
  - if $f \geq 0$ then $P(f) \geq 0$
  - $P(\mathscr{X}) = 1$
- restriction to (indicators of) events is a finitely additive probability measure
- Let $\mathscr{P}$ denote the set of all linear previsions on $\mathscr{L}(\mathscr{X})$

Linear previsions: basic results

Theorem 2. .*Let P be a precise prevision on a linear space $\mathscr{K}$ of gambles. Then P is coherent (is a linear prevision) if and only if for all f and g in $\mathscr{K}$ and $\lambda \in \mathbb{R}$:*

LP1. $P(f) \geq \inf f$

LP2. $P(f+g) = P(f) + P(g)$

LP3. $P(\lambda f) = \lambda P(f)$.

For precise previsions, coherence and avoiding sure loss coincide!

Theorem 3. .*A precise prevision P on $\mathscr{L}(\mathscr{X})$ is coherent if and only if it avoids sure loss.*

Linear previsions: mass functions

- Assume that $\mathscr{X}$ is finite and consider a linear prevision P on $\mathscr{L}(\mathscr{X})$
- P is completely determined by its (probability) mass function

$$p(x) = P(\{x\}) = P(I_{\{x\}}), \quad \forall x \in \mathscr{X}$$

where

$$p(x) \geq 0 \quad \text{and} \quad \sum_{x \in \mathscr{X}} p(x) = 1$$

and

$$P(f) = \sum_{x \in \mathscr{X}} p(x) f(x).$$

Sets of linear previsions

Consider a lower prevision $\underline{P}$ on a set of gambles $\mathscr{K}$

- Let $\mathscr{M}(\underline{P})$ be the set of linear previsions on $\mathscr{L}(\mathscr{X})$ that dominate $\underline{P}$ on its domain $\mathscr{K}$:

$$\boxed{\mathscr{M}(\underline{P}) = \{Q \in \mathscr{P} \colon (\forall f \in \mathscr{K})(Q(f) \geq \underline{P}(f))\}.}$$

- Then avoiding sure loss is equivalent to $\mathscr{M}(\underline{P}) \neq \emptyset$.
- and coherence is equivalent to:

$$\underline{P}(f) = \min\{Q(f) \colon Q \in \mathscr{M}(\underline{P})\}, \quad \forall f \in \mathscr{K}.$$

- A lower envelope of a set of precise previsions is always a coherent lower prevision

Coherent lower/upper previsions – 1

- probability measures, previsions à la de Finetti
- 2-monotone capacities, Choquet capacities
- contamination models
- possibility and necessity measures
- belief and plausibility functions
- random set models

Coherent lower/upper previsions – 2

- reachable probability intervals
- lower and upper mass/density functions
- lower and upper cumulative distributions (p-boxes)
- (lower and upper envelopes of) credal sets
- distributions (Gaussian, Poisson, Dirichlet, multinomial, . . .) with interval-valued parameters
- robust Bayesian models
- . . .

Natural extension

Third step toward a scientific theory

= how to make the theory useful

= use the assessments to draw conclusions about other things [(conditional) events, gambles, ...]

Natural extension

Third step toward a scientific theory

= how to make the theory useful

= use the assessments to draw conclusions about other things [(conditional) events, gambles, . . .]

Problem: extend a coherent lower prevision defined on a collection of gambles to a lower prevision on all gambles (conditional events, gambles, . . .)

Natural extension

Third step toward a scientific theory

= how to make the theory useful

= use the assessments to draw conclusions about other things [(conditional) events, gambles, . . .]

Problem: extend a coherent lower prevision defined on a collection of gambles to a lower prevision on all gambles (conditional events, gambles, . . .)

Requirements:

- coherence
- as low as possible (conservative, least-committal)

$= \boxed{\text{NATURAL EXTENSION}}$

Natural extension: an example – 1

Lower probabilities $\underline{P}(A)$ and $\underline{P}(B)$ for two events A and B that are logically independent:

$$A \cap B \neq \emptyset \quad A \cap B^c \neq \emptyset \quad A^c \cap B \neq \emptyset \quad A^c \cap B^c \neq \emptyset$$

Natural extension: an example – 1

Lower probabilities $\underline{P}(A)$ and $\underline{P}(B)$ for two events A and B that are logically independent:

$$A \cap B \neq \emptyset \quad A \cap B^c \neq \emptyset \quad A^c \cap B \neq \emptyset \quad A^c \cap B^c \neq \emptyset$$

For all $\lambda \geq 0$ and $\mu \geq 0$, you accept to buy any gamble f for price α if for all x

$$f(x) - \alpha \geq \lambda [I_A(x) - \underline{P}(A)] + \mu [I_B(x) - \underline{P}(B)]$$

Natural extension: an example – 1

Lower probabilities $\underline{P}(A)$ and $\underline{P}(B)$ for two events A and B that are logically independent:

$$A \cap B \neq \emptyset \quad A \cap B^c \neq \emptyset \quad A^c \cap B \neq \emptyset \quad A^c \cap B^c \neq \emptyset$$

For all $\lambda \geq 0$ and $\mu \geq 0$, you accept to buy any gamble f for price α if for all x

$$f(x) - \alpha \geq \lambda [I_A(x) - \underline{P}(A)] + \mu [I_B(x) - \underline{P}(B)]$$

The natural extension $\underline{E}(f)$ of the assessments $\underline{P}(A)$ and $\underline{P}(B)$ to any gamble f is the highest α such that this inequality holds, over all possible choices of λ and μ.

Natural extension: an example – 2

Calculate $\underline{E}(A \cup B)$: maximise α subject to the constraints: $\lambda \geq 0$, $\mu \geq 0$, and for all x:

$$I_{A \cup B}(x) - \alpha \geq \lambda [I_A(x) - \underline{P}(A)] + \mu [I_B(x) - \underline{P}(B)]$$

or equivalently:

$$I_{A \cup B}(x) \geq \lambda I_A(x) + \mu I_B(x) + [\alpha - \lambda \underline{P}(A) - \mu \underline{P}(B)]$$

Natural extension: an example – 2

Calculate $\underline{E}(A\cup B)$: maximise α subject to the constraints: $\lambda \geq 0$, $\mu \geq 0$, and for all x:

$$I_{A\cup B}(x) - \alpha \geq \lambda[I_A(x) - \underline{P}(A)] + \mu[I_B(x) - \underline{P}(B)]$$

or equivalently:

$$I_{A\cup B}(x) \geq \lambda I_A(x) + \mu I_B(x) + [\alpha - \lambda\underline{P}(A) - \mu\underline{P}(B)]$$

and if we put $\gamma = \alpha - \lambda\underline{P}(A) - \mu\underline{P}(B)$ this is equivalent to maximising

$$\gamma + \lambda\underline{P}(A) + \mu\underline{P}(B)$$

subject to the inequalities

$$1 \geq \lambda + \mu + \gamma, \quad 1 \geq \lambda + \gamma, \quad 1 \geq \mu + \gamma, \quad 0 \geq \gamma$$

$$\lambda \geq 0, \quad \mu \geq 0$$

Natural extension: an example – 3

This is a linear programming problem, and its solution is easily seen to be:

$$\underline{E}(A \cup B) = \max\{\underline{P}(A), \underline{P}(B)\}$$

Similarly, for $f = I_{A \cap B}$ we get another linear programming problem that yields

$$\underline{E}(A \cap B) = \max\{0, \underline{P}(A) + \underline{P}(B) - 1\}$$

These are the Fréchet bounds!

Natural extension: an example – 3

This is a linear programming problem, and its solution is easily seen to be:

$$\underline{E}(A \cup B) = \max\{\underline{P}(A), \underline{P}(B)\}$$

Similarly, for $f = I_{A \cap B}$ we get another linear programming problem that yields

$$\underline{E}(A \cap B) = \max\{0, \underline{P}(A) + \underline{P}(B) - 1\}$$

These are the Fréchet bounds! Natural extension always gives the most conservative values that are still compatible with coherence and other additional assumptions made ...

Another example: set information

- Information: X assumes a value in a subset A of $\mathscr{X}$

Another example: set information

- Information: X assumes a value in a subset A of $\mathscr{X}$
- This information is represented by the vacuous lower prevision relative to A:

$$\underline{P}_A(f) = \inf_{x \in A} f(x); \quad f \in \mathscr{L}(\mathscr{X})$$

Another example: set information

- Information: X assumes a value in a subset A of $\mathscr{X}$
- This information is represented by the vacuous lower prevision relative to A:

$$\underline{P}_A(f) = \inf_{x \in A} f(x); \quad f \in \mathscr{L}(\mathscr{X})$$

- $P \in \mathscr{M}(\underline{P}_A)$ iff $P(A) = 1$
- $\underline{P}_A$ is the natural extension of the precise probability assessment '$P(A) = 1$'; also of the belief function with probability mass one on A
- Take any P such that $P(A) = 1$, then $P(f)$ is only determined up to an interval $[\underline{P}_A(f), \overline{P}_A(f)]$ according to de Finetti's fundamental theorem of probability

Natural extension: general definition

Defi nition 4. .Consider a lower prevision $\underline{P}$ on a set of gambles $\mathscr{K}$. The natural extension $\underline{E}$ of $\underline{P}$ is a lower prevision defined on the set $\mathscr{L}(\mathscr{X})$ of all gambles:

$$\boxed{\underline{E}(f) = \sup_{\substack{n \geq 0 \\ f_k \in \mathscr{K}, \lambda_k \geq 0 \\ k=1,\dots,n}} \sup \left\{ \alpha \colon f - \alpha \geq \sum_{k=1}^{n} \lambda_k [f_k - \underline{P}(f_k)] \right\}}$$

for all $f \in \mathscr{L}(\mathscr{X})$

Natural extension theorem

Theorem 4. .*Consider a lower prevision $\underline{P}$ on a set of gambles $\mathscr{K}$ that avoids sure loss. Then its natural extension $\underline{E}$ has the following properties:*

a. $\inf f \leq \underline{E}(f) \leq \sup f$

b. *$\underline{E}$ is a coherent lower prevision on $\mathscr{L}(\mathscr{X})$*

c. *$\underline{E}$ dominates $\underline{P}$ on $\mathscr{K}$: $\underline{E}(f) \geq \underline{P}(f)$ for all f in $\mathscr{K}$*

d. *$\underline{E}$ agrees with $\underline{P}$ on $\mathscr{K}$ if and only if $\underline{P}$ is coherent*

e. *$\underline{E}$ is the smallest coherent lower prevision on $\mathscr{L}(\mathscr{X})$ that dominates $\underline{P}$ on $\mathscr{K}$*

f. *if $\underline{P}$ is coherent then $\underline{E}$ is the smallest coherent extension of $\underline{P}$ to $\mathscr{L}(\mathscr{X})$*

Natural extension: sets of previsions

Consider a lower prevision $\underline{P}$ on a set of gambles $\mathscr{K}$

- If it avoids sure loss then $\mathscr{M}(\underline{P}) \neq \emptyset$ and its natural extension is given by the lower envelope of $\mathscr{M}(\underline{P})$:

$$\boxed{\underline{E}(f) = \min\{Q(f) : Q \in \mathscr{M}(\underline{P})\}, \quad \forall f \in \mathscr{L}(\mathscr{X})}$$

- $\underline{P}$ is coherent iff it coincides on its domain $\mathscr{K}$ with its natural extension
- $\mathscr{M}(\underline{P}) = \mathscr{M}(\underline{E})$

Natural extension: desirable gambles

Consider a set $\mathscr{D}$ of gambles you have judged desirable.

- What are the implications of these assessments for the desirability of other gambles?

Natural extension: desirable gambles

Consider a set $\mathscr{D}$ of gambles you have judged desirable.

- What are the implications of these assessments for the desirability of other gambles?
- The natural extension $\mathscr{E}$ of $\mathscr{D}$ is the smallest coherent set of desirable gambles that includes $\mathscr{D}$
- It is the smallest extension of $\mathscr{D}$ to a convex cone of gambles that contains all non-negative gambles.

Natural extension: special cases

Natural extension is a very powerful reasoning method. In special cases it reduces to:

- logical deduction
- belief functions via random sets
- fundamental theorem of probability/prevision
- Lebesgue integration of a probability measure
- Choquet integration of 2-monotone lower probabilities
- Bayes' rule for probability measures
- Bayesian updating of lower/upper probabilities
- robust Bayesian analysis
- first-order model from higher-order model

Three pillars

1. behavioural definition of lower/upper previsions that can be made operational
2. rationality criteria of
   - avoiding sure loss
   - coherence
3. natural extension to make the theory useful

Gambles and events – 1

- How to represent: event A is at least n times as probable as event B

Gambles and events – 1

- How to represent: event A is at least n times as probable as event B
- Set of precise previsions $\mathscr{M}$:

$$P \in \mathscr{M} \Leftrightarrow P(A) \geq nP(B) \Leftrightarrow P(I_A - nI_B) \geq 0$$

- lower previsions: $\underline{P}(I_A - nI_B) \geq 0$
- sets of desirable gambles: $I_A - nI_B + \varepsilon \in \mathscr{D}$, $\forall \varepsilon > 0$.

Gambles and events – 1

- How to represent: event A is at least n times as probable as event B
- Set of precise previsions $\mathscr{M}$:

$$P \in \mathscr{M} \Leftrightarrow P(A) \geq nP(B) \Leftrightarrow P(I_A - nI_B) \geq 0$$

- lower previsions: $\underline{P}(I_A - nI_B) \geq 0$
- sets of desirable gambles: $I_A - nI_B + \varepsilon \in \mathscr{D}$, $\forall \varepsilon > 0$.
- $I_A - nI_B$ is a gamble, generally not an indicator!
- Cannot be expressed by lower probabilities:

$$\begin{cases} \underline{P}(A) \geq \underline{P}(B), \quad \overline{P}(A) \geq \overline{P}(B) & \text{too weak} \\ \underline{P}(A) \geq \overline{P}(B) & \text{too strong} \end{cases}$$

Gambles and events – 2

- Did I come to Lugano by plane, by car or by train?

Gambles and events – 2

- Did I come to Lugano by plane, by car or by train?
- Assessments:
  - ‘not by plane’ is at least as probable as ‘by plane’
  - ‘by plane’ is at least a probable as ‘by train’
  - ‘by train’ is at least a probable as ‘by car’

Gambles and events – 2

- Did I come to Lugano by plane, by car or by train?
- Assessments:
  - ‘not by plane’ is at least as probable as ‘by plane’
  - ‘by plane’ is at least a probable as ‘by train’
  - ‘by train’ is at least a probable as ‘by car’
- Convex set $\mathscr{M}$ of probability mass functions m on $\{p,t,c\}$ such that

$$m(p) \leq \frac{1}{2}, \quad m(p) \geq m(t), \quad m(t) \geq m(c)$$

- $\mathscr{M}$ is a convex set with extreme points

$$(\frac{1}{2},\frac{1}{2},0), \quad (\frac{1}{2},\frac{1}{4},\frac{1}{4}), \quad (\frac{1}{3},\frac{1}{3},\frac{1}{3})$$

Gambles and events – 3

- the natural extension $\underline{E}$ is the lower envelope of this set

$$\underline{E}(f) = \min_{m \in \mathscr{M}} [m(p)f(p) + m(t)f(t) + m(c)f(c)]$$

Gambles and events – 3

- the natural extension $\underline{E}$ is the lower envelope of this set

$$\underline{E}(f) = \min_{m \in \mathscr{M}} [m(p)f(p) + m(t)f(t) + m(c)f(c)]$$

- The lower probabilities are completely specified by

$$\underline{E}(\{p\}) = \frac{1}{3} \qquad \overline{E}(\{p\}) = \frac{1}{2}$$

$$\underline{E}(\{t\}) = \frac{1}{4} \qquad \overline{E}(\{t\}) = \frac{1}{2}$$

$$\underline{E}(\{c\}) = 0 \qquad \overline{E}(\{c\}) = \frac{1}{3}$$

Gambles and events – 4

- the corresponding set of mass functions $\mathscr{M}^*$ is a convex set with extreme points

$$(\frac{1}{2}, \frac{1}{2}, 0), \quad (\frac{1}{2}, \frac{1}{4}, \frac{1}{4}), \quad (\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$$

$$(\frac{5}{12}, \frac{1}{4}, \frac{1}{3}), \quad (\frac{1}{3}, \frac{1}{2}, \frac{1}{6})$$

Gambles and events – 4

- the corresponding set of mass functions $\mathscr{M}^*$ is a convex set with extreme points

$$(\frac{1}{2},\frac{1}{2},0),\quad (\frac{1}{2},\frac{1}{4},\frac{1}{4}),\quad (\frac{1}{3},\frac{1}{3},\frac{1}{3})$$
$$(\frac{5}{12},\frac{1}{4},\frac{1}{3}),\quad (\frac{1}{3},\frac{1}{2},\frac{1}{6})$$

- $\mathscr{M}$ is more informative than $\mathscr{M}^*$: $\mathscr{M} \subset \mathscr{M}^*$
- with $\mathscr{M}$ we can infer that $\underline{E}(I_{\{p\}} - I_{\{t\}}) \geq 0$: 'by plane' is at least as probable as 'by train'
- with $\mathscr{M}^*$ this inference cannot be made: we lose information by restricting ourselves to lower probabilities

Gambles and events – 5

$$\begin{array}{rcl} \text{event } A & \Leftrightarrow & \text{gamble } I_A \\ \text{lower probability } \underline{P}(A) & \Leftrightarrow & \text{lower prevision } \underline{P}(I_A) \end{array}$$

Gambles and events – 5

$$\begin{aligned} \text{event } A \quad &\Leftrightarrow \quad \text{gamble } I_A \\ \text{lower probability } \underline{P}(A) \quad &\Leftrightarrow \quad \text{lower prevision } \underline{P}(I_A) \end{aligned}$$

In precise probability theory:
$\longrightarrow$ events are as expressive as gambles

In imprecise probability theory:
$\longrightarrow$ events are less expressive than gambles

And by the way

There is a natural embedding of classical propositional logic into imprecise probability theory.

set of propositions	$\longrightarrow$	lower probability
logically consistent	$\longrightarrow$	ASL
deductively closed	$\longrightarrow$	coherent
deductive closure	$\longrightarrow$	natural extension
maximal deductively closed	$\longrightarrow$	probability

No such embedding exists into precise probability theory.

Part III

Decision making

Decision making – 1

Consider an action a whose outcome (reward) depends on the actual value of X.
With such an action we can associate a reward function

$$f_a\colon \mathscr{X} \to \mathbb{R}\colon x \mapsto f_a(x)$$

Decision making – 1

Consider an action a whose outcome (reward) depends on the actual value of X.
With such an action we can associate a reward function

$$f_a \colon \mathscr{X} \to \mathbb{R} \colon x \mapsto f_a(x)$$

When do you strictly prefer action a over action b:

$$a > b \Leftrightarrow \underline{P}(f_a - f_b) > 0$$

Decision making – 1

Consider an action a whose outcome (reward) depends on the actual value of X.
With such an action we can associate a reward function

$$f_a\colon \mathscr{X} \to \mathbb{R} \colon x \mapsto f_a(x)$$

When do you strictly prefer action a over action b:

$$a > b \Leftrightarrow \underline{P}(f_a - f_b) > 0$$

You almost-prefer a over b if

$$a \geq b \Leftrightarrow \underline{P}(f_a - f_b) \geq 0$$

We identify an action a with its reward function f_a

Decision making – 2

You are indifferent between a and b if

$$a \approx b \Leftrightarrow a \geq b \text{ and } b \geq a \Leftrightarrow \underline{P}(f_a - f_b) = \overline{P}(f_a - f_b) = 0$$

Decision making – 2

You are indifferent between a and b if

$$a \approx b \Leftrightarrow a \geq b \text{ and } b \geq a \Leftrightarrow \underline{P}(f_a - f_b) = \overline{P}(f_a - f_b) = 0$$

Actions a and b are incomparable if

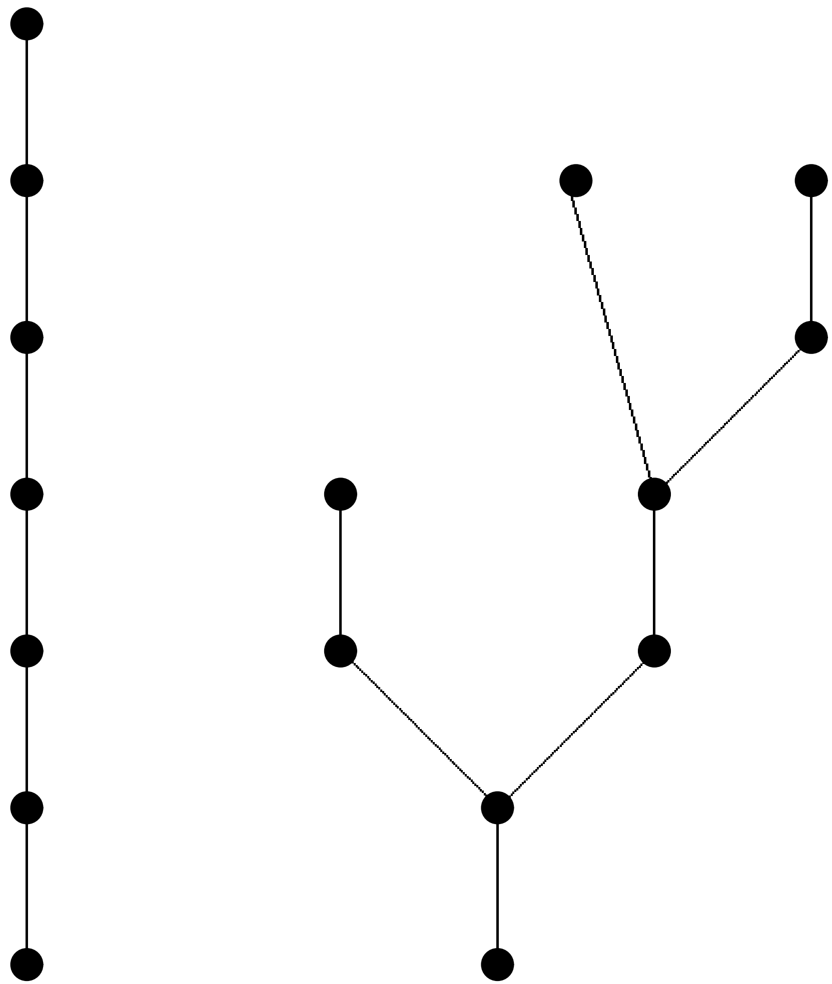
$$a \| b \Leftrightarrow a \not> b \text{ and } b \not> a \text{ and } a \not\approx b$$

- In that case there is not enough information in the model to choose between a and b: you are undecided!
- Imprecise probability models allow for indecision!
- In fact, modelling and allowing for indecision is one of the motivations for introducing imprecise probabilities

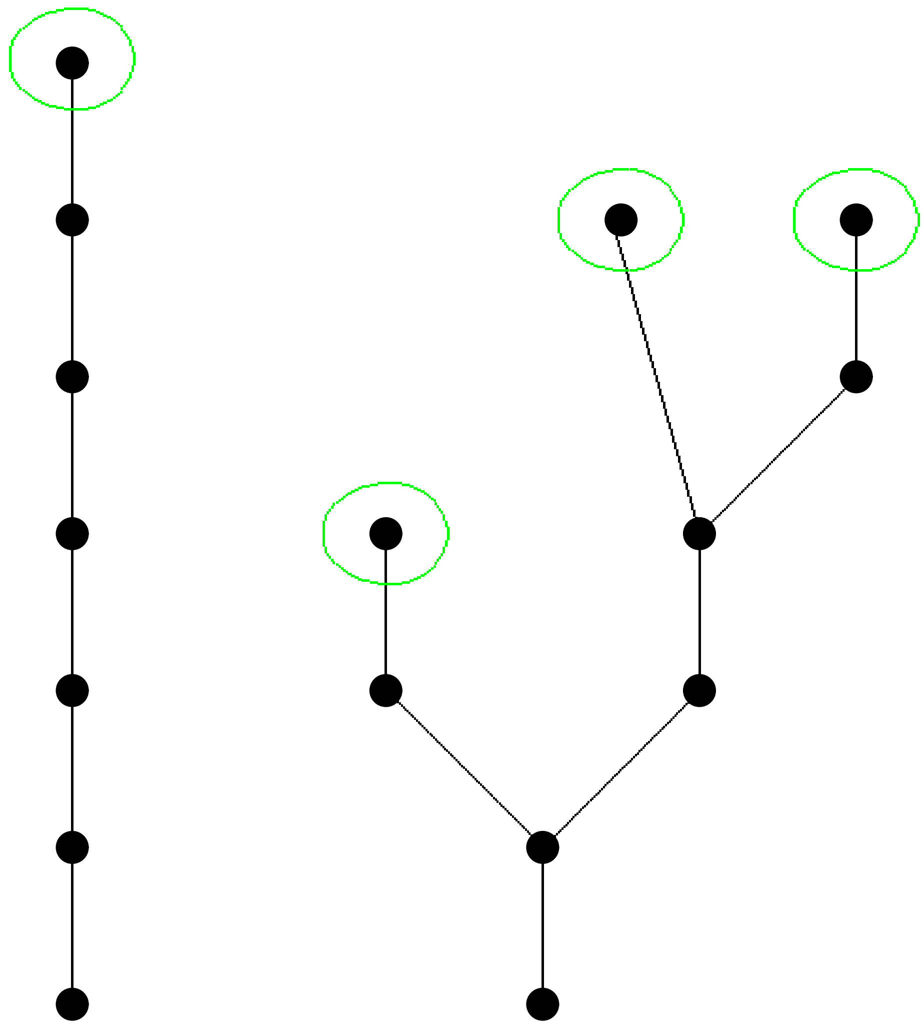
Decision making: maximal actions

- Consider a set of actions $\mathbb{A}$ and reward functions $\mathscr{K} = \{f_a : a \in \mathbb{A}\}$
- Due to the fact that certain actions may be incomparable, the actions cannot be linearly ordered, only partially!

Ordering of actions



Ordering of actions



Decision making: maximal actions

- The maximal actions a in $\mathbb{A}$ are actions that are undominated in $\mathbb{A}$:

$$(\forall b \in \mathbb{A})(b \not> a)$$

or equivalently

$$\boxed{(\forall b \in \mathbb{A})(\overline{P}(f_a - f_b) \geq 0)}$$

Decision making: maximal actions

- The maximal actions a in $\mathbb{A}$ are actions that are undominated in $\mathbb{A}$:

$$(\forall b \in \mathbb{A})(b \not> a)$$

or equivalently

$$\boxed{(\forall b \in \mathbb{A})(\overline{P}(f_a - f_b) \geq 0)}$$

- Any two maximal actions are either indifferent or incomparable!

Decision making: the precise case

- $a > b \Leftrightarrow P(f_a - f_b) > 0 \Leftrightarrow P(f_a) > P(f_b)$

Decision making: the precise case

- $a > b \Leftrightarrow P(f_a - f_b) > 0 \Leftrightarrow P(f_a) > P(f_b)$
- $a \geq b \Leftrightarrow P(f_a - f_b) \geq 0 \Leftrightarrow P(f_a) \geq P(f_b)$

Decision making: the precise case

- $a > b \Leftrightarrow P(f_a - f_b) > 0 \Leftrightarrow P(f_a) > P(f_b)$
- $a \geq b \Leftrightarrow P(f_a - f_b) \geq 0 \Leftrightarrow P(f_a) \geq P(f_b)$
- $a \approx b \Leftrightarrow P(f_a) = P(f_b)$

Decision making: the precise case

- $a > b \Leftrightarrow P(f_a - f_b) > 0 \Leftrightarrow P(f_a) > P(f_b)$
- $a \geq b \Leftrightarrow P(f_a - f_b) \geq 0 \Leftrightarrow P(f_a) \geq P(f_b)$
- $a \approx b \Leftrightarrow P(f_a) = P(f_b)$
- never $a \| b$!
- There is no indecision in precise probability models
- Whatever the available information, they always allow you a best choice between two available actions!
- Actions can always be ordered linearly, maximal actions are unique (up to indifference): they have the highest expected utility.

Decision making: sets of previsions

- $a > b \Leftrightarrow (\forall P \in \mathscr{M}(\underline{P}))(P(f_a) > P(f_b))$

Decision making: sets of previsions

- $a > b \Leftrightarrow (\forall P \in \mathscr{M}(\underline{P}))(P(f_a) > P(f_b))$
- $a \geq b \Leftrightarrow (\forall P \in \mathscr{M}(\underline{P}))(P(f_a) \geq P(f_b))$

Decision making: sets of previsions

- $a > b \Leftrightarrow (\forall P \in \mathscr{M}(\underline{P}))(P(f_a) > P(f_b))$
- $a \geq b \Leftrightarrow (\forall P \in \mathscr{M}(\underline{P}))(P(f_a) \geq P(f_b))$
- $a \approx b \Leftrightarrow (\forall P \in \mathscr{M}(\underline{P}))(P(f_a) = P(f_b))$

Decision making: sets of previsions

- $a > b \Leftrightarrow (\forall P \in \mathscr{M}(\underline{P}))(P(f_a) > P(f_b))$
- $a \geq b \Leftrightarrow (\forall P \in \mathscr{M}(\underline{P}))(P(f_a) \geq P(f_b))$
- $a \approx b \Leftrightarrow (\forall P \in \mathscr{M}(\underline{P}))(P(f_a) = P(f_b))$
- $a \| b \Leftrightarrow (\exists P \in \mathscr{M}(\underline{P}))(P(f_a) < P(f_b))$
  and $(\exists Q \in \mathscr{M}(\underline{P}))(Q(f_a) > Q(f_b))$

Decision making: sets of previsions

- $a > b \Leftrightarrow (\forall P \in \mathscr{M}(\underline{P}))(P(f_a) > P(f_b))$
- $a \geq b \Leftrightarrow (\forall P \in \mathscr{M}(\underline{P}))(P(f_a) \geq P(f_b))$
- $a \approx b \Leftrightarrow (\forall P \in \mathscr{M}(\underline{P}))(P(f_a) = P(f_b))$
- $a \| b \Leftrightarrow (\exists P \in \mathscr{M}(\underline{P}))(P(f_a) < P(f_b))$ and $(\exists Q \in \mathscr{M}(\underline{P}))(Q(f_a) > Q(f_b))$
- If $\mathscr{K}$ is convex then a is maximal if and only if there is some $P \in \mathscr{M}(\underline{P})$ such that

$$(\forall b \in \mathbb{A})(P(f_a) \geq P(f_b))$$

Part IV

Conditional lower previsions

Joint lower previsions

- Consider another random variable Y assuming values in a set $\mathscr{Y}$.
- (X,Y) assumes values in $\mathscr{X} \times \mathscr{Y}$
- Assume $\mathscr{X}$ and $\mathscr{Y}$ are finite
- Assume logical independence of X and Y
- A joint lower prevision $\underline{P}$ on $\mathscr{L}(\mathscr{X} \times \mathscr{Y})$ models beliefs about the value that X and Y assume jointly in $\mathscr{X} \times \mathscr{Y}$

Joint linear previsions

Example: A linear prevision P on $\mathscr{L}(\mathscr{X} \times \mathscr{Y})$ is completely determined by its (probability) mass function

$$p(x,y) = P(\{(x,y)\}) = P(I_{\{(x,y)\}}), \quad \forall (x,y) \in \mathscr{X} \times \mathscr{Y}$$

where

$$p(x,y) \geq 0 \quad \text{and} \quad \sum_{(x,y)\in\mathscr{X}\times\mathscr{Y}} p(x,y) = 1$$

and for all gambles $h\colon \mathscr{X} \times \mathscr{Y} \to \mathbb{R}$

$$P(h) = \sum_{(x,y)\in\mathscr{X}\times\mathscr{Y}} p(x,y)h(x,y).$$

Marginal lower previsions

- Assume you have a joint lower prevision $\underline{P}$ on $\mathscr{L}(\mathscr{X} \times \mathscr{Y})$
- You want to infer a model for the beliefs about the value that Y assumes in $\mathscr{Y}$, irrespective of the value that X assumes in $\mathscr{X}$
- Marginal lower prevision $\underline{P}_Y$ on $\mathscr{L}(\mathscr{Y})$:

$$\boxed{\underline{P}_Y(g) = \underline{P}(g') \text{ for all } g \text{ in } \mathscr{L}(\mathscr{Y})}$$

where

$$g'(x,y) = g(y) \text{ for all } x \text{ in } \mathscr{X} \text{ and } y \text{ in } \mathscr{Y}$$

- we shall identify g and g' and simply write $\underline{P}_Y(g) = \underline{P}(g)$.

Marginal linear previsions

Example: If P is a linear prevision on $\mathscr{L}(\mathscr{X} \times \mathscr{Y})$, then its Y-marginal P_Y is a linear prevision on $\mathscr{L}(\mathscr{Y})$, completely determined by its marginal mass function

$$p_Y(y) = P_Y(\{y\}) = P_Y(I_{\{y\}}), \quad \forall y \in \mathscr{Y}$$

where

$$p_Y(y) = \sum_{x \in \mathscr{X}} p(x,y)$$

and for all gambles $g\colon \mathscr{Y} \to \mathbb{R}$

$$P_Y(g) = \sum_{y \in \mathscr{Y}} p_Y(y)g(y).$$

Conditional lower previsions

- Consider a gamble h on $\mathscr{X} \times \mathscr{Y}$ and a value $y \in \mathscr{Y}$
- Conditional lower prevision $\underline{P}(h|Y=y) = \underline{P}(h|y)$ is the supremum price for buying h if the subject knew that $Y = y$.
- For each y in $\mathscr{Y}$, a value $\underline{P}(h|y)$, summarised by $\underline{P}(h|Y)$

$$\underline{P}(h|Y) \colon \mathscr{Y} \to \mathbb{R} \colon y \mapsto \underline{P}(h|y)$$

- $\underline{P}(\cdot|Y)$ is a two-place function on $\mathscr{L}(\mathscr{X} \times \mathscr{Y}) \times \mathscr{Y}$ with partial maps

$$\underline{P}(h|Y) \colon \mathscr{Y} \to \mathbb{R} \colon y \mapsto \underline{P}(h|y)$$
$$\underline{P}(\cdot|y) \colon \mathscr{L}(\mathscr{X} \times \mathscr{Y}) \to \mathbb{R} \colon h \mapsto \underline{P}(h|y)$$

Separate coherence

- Rationality criteria to be imposed on $\underline{P}(\cdot|Y)$ alone
  - for each y in $\mathscr{Y}$, $\underline{P}(\cdot|y)$ should be a coherent lower prevision on $\mathscr{L}(\mathscr{X}\times\mathscr{Y})$
  - for each y in $\mathscr{Y}$, $\underline{P}(\{y\}|Y=y)=\underline{P}(I_{\mathscr{X}\times\{y\}}|y)=1$
- Separate coherence of $\underline{P}(\cdot|Y)$
- Immediate consequence:

$$\underline{P}(h|y)=\underline{P}(h(\cdot,y)|y)$$

- $\underline{P}(\cdot|y)$ is completely determined by its values on $\mathscr{L}(\mathscr{X})$:

$$h(\cdot,y)=h'(\cdot,y)\Rightarrow\underline{P}(h|y)=\underline{P}(h'|y)$$

Joint coherence

- Assume that we have
  - a coherent joint lower prevision $\underline{P}$ on $\mathscr{L}(\mathscr{X} \times \mathscr{Y})$
  - a separately coherent conditional lower prevision $\underline{P}(\cdot|Y)$ on $\mathscr{L}(\mathscr{X})$ (or equivalently on $\mathscr{L}(\mathscr{X} \times \mathscr{Y})$)
- Then $\underline{P}$ and $\underline{P}(\cdot|Y)$ should satisfy the rationality criterion of joint coherence:

$$\boxed{\underline{P}\left(I_{\mathscr{X} \times \{y\}}[h - \underline{P}(h|y)]\right) = 0 \text{ for all } y \in \mathscr{Y},\ h \in \mathscr{L}(\mathscr{X} \times \mathscr{Y})} \quad \text{(GBR)}$$

- This criterion is also called the Generalized Bayes Rule

Joint coherence for linear previsions

- If the joint lower prevision is a linear prevision $\underline{P}$, the GBR becomes

$$P(hI_{\mathscr{X}\times\{y\}}) = \underline{P}(h|y)P(\mathscr{X}\times\{y\})$$

or equivalently,

$$\underline{P}(h|y) = \frac{P(hI_{\mathscr{X}\times\{y\}})}{P(\mathscr{X}\times\{y\})} \text{ if } P(\mathscr{X}\times\{y\}) = P_Y(\{y\}) > 0$$

- This is Bayes' rule, and $\underline{P}(\cdot|y)$ is a precise prevision with mass function

$$p(x|y) = \frac{p(x,y)}{p_Y(y)} \text{ if } p_Y(y) > 0$$

Generalised Bayes Rule

- If $\underline{P}_Y(\{y\}) = \underline{P}(\mathscr{X} \times \{y\}) > 0$ then coherence implies that $\underline{P}(h|y)$ is the unique solution of the following equation in μ:

$$\underline{P}\left(I_{\mathscr{X}\times\{y\}}[h-\mu]\right) = 0 \text{ (Generalised Bayes Rule)}$$

- Observe that also (divisive conditioning)

$$\underline{P}(h|y) = \inf\left\{\frac{Q(hI_{\mathscr{X}\times\{y\}})}{Q(\mathscr{X}\times\{y\})} : Q \in \mathscr{M}(\underline{P})\right\}$$
$$= \inf\{Q(h|y) : Q \in \mathscr{M}(\underline{P})\}$$

Regular extension

- If $\underline{P}_Y(\{y\}) = 0$ but $\overline{P}_Y(\{y\}) > 0$ then one often considers the so-called regular extension $\underline{R}(h|y)$: it is the greatest μ such that
$$\underline{P}(I_{\mathscr{X}\times\{y\}}[h-\mu]) \geq 0$$
- Observe that also
$$\underline{R}(h|y) = \inf\{Q(h|y) : Q \in \mathscr{M}(\underline{P}) \text{ and } Q(\mathscr{X}\times\{y\}) > 0\}$$
- Regular extension is the most conservative coherent extension that satisfies an additional regularity condition

Marginal extension

- Let $\underline{P}_Y$ be a coherent (marginal) lower prevision defined on some subset $\mathscr{K}$ of $\mathscr{L}(\mathscr{Y})$ and let $\underline{E}_Y$ be its natural extension to $\mathscr{L}(\mathscr{Y})$.
- Let $\underline{P}(\cdot|Y)$ be a separately coherent conditional lower prevision on some subset $\mathscr{H}$ of $\mathscr{L}(\mathscr{X})$, and let $\underline{E}(\cdot|y)$ be the natural extension of $\underline{P}(\cdot|y)$ to $\mathscr{L}(\mathscr{X})$, $y \in \mathscr{Y}$.

Theorem 5 (Marginal extension theorem). .*The smallest coherent joint lower prevision $\underline{E}$ on $\mathscr{L}(\mathscr{X} \times \mathscr{Y})$ whose $\mathscr{Y}$-marginal coincides with $\underline{P}_Y$ on $\mathscr{K}$ and that is jointly coherent with $\underline{P}(\cdot|Y)$, is given by*

$$\underline{E}(h) = \underline{E}_Y(\underline{E}(h|Y))$$

for all gambles h on $\mathscr{X} \times \mathscr{Y}$.

Marginal extension: linear previsions

- If $\underline{P}(\cdot|Y)$ is precise, then the marginal extension $\underline{E}$ is uniquely coherent
- let $P(\cdot|Y)$ be precise on $\mathscr{L}(\mathscr{X})$ and P_Y be precise on $\mathscr{L}(\mathscr{Y})$, then the marginal extension E is precise and uniquely coherent, and given by

$$E(h) = P_Y(P(h|Y))$$

or in terms of mass functions

$$p(x,y) = p_Y(y)p(x|y)$$

Marginal extension: sets of linear previsions

- Consider $\underline{P}_Y$ on $\mathscr{K} \subseteq \mathscr{L}(\mathscr{Y})$ and $\underline{P}(\cdot|Y)$ on $\mathscr{H} \subseteq \mathscr{L}(\mathscr{X})$
- The marginal extension $\underline{E}$ is the lower envelope of the set of linear previsions

  $$Q_Y(Q(\cdot|Y)) \text{ (= marginal extension of } Q_Y \text{ and } Q(\cdot|Y))$$

  where
  - Q_Y is any element of $\mathscr{M}(\underline{P}_Y)$
  - $Q(\cdot|y)$ is any element of $\mathscr{M}(\underline{P}(\cdot|y))$ for all y in $\mathscr{Y}$
- we may also restrict ourselves to sets of extreme points

Epistemic irrelevance and independence

- Y is (epistemically) irrelevant to X if additional knowledge about the value of Y does not change our beliefs about the value of X:

$$\underline{P}(f|y) = \underline{P}_X(f)$$

for gambles f on $\mathscr{X}$ and all $y \in \mathscr{Y}$

- X is (epistemically) irrelevant to Y:

$$\underline{P}(g|x) = \underline{P}_Y(g)$$

for gambles g on $\mathscr{Y}$ and all $x \in \mathscr{X}$

- X and Y are epistemically independent if Y is irrelevant to X and the other way round.

Products – 1

- Consider a coherent marginal lower prevision $\underline{P}_X$ on $\mathscr{K} \subseteq \mathscr{L}(\mathscr{X})$
- Consider a coherent marginal lower prevision $\underline{P}_Y$ on $\mathscr{H} \subseteq \mathscr{L}(\mathscr{Y})$
- a coherent product of $\underline{P}_X$ and $\underline{P}_Y$ is any coherent lower prevision on $\mathscr{L}(\mathscr{X} \times \mathscr{Y})$ whose $\mathscr{X}$-marginal coincides with $\underline{P}_X$ on $\mathscr{K}$ and whose $\mathscr{Y}$-marginal coincides with $\underline{P}_Y$ on $\mathscr{H}$

Products – 2

With increasing precision:

- natural extension: smallest coherent product of $\underline{P}_X$ and $\underline{P}_Y$
- irrelevant natural extension: smallest coherent product of $\underline{P}_X$ and $\underline{P}_Y$ such that X is epistemically irrelevant to Y
- independent natural extension: smallest coherent product of $\underline{P}_X$ and $\underline{P}_Y$ such that X and Y are epistemically independent
- type-I product: lower envelope of the products of the linear previsions in $\mathscr{M}(\underline{P}_X)$ with the linear previsions in $\mathscr{M}(\underline{P}_Y)$

Questions

?