General introduction to imprecise probability

Enrique Miranda
University of Oviedo (Spain)
mirandaenrique@uniovi.es

Overview, Part I

1. Some considerations about probability.
2. Lower previsions and probabilities.
3. Avoiding sure loss and coherence.
4. Particular cases.

Which is the goal of probability?

Probability seeks to determine the plausibility of the different outcomes of an experiment when these cannot be predicted beforehand.

- What is the probability of guessing the 6 winning numbers in the lottery?
- What is the probability of arriving in 1 hour from Newcastle airport to the center of Durham by car?
- What is the probability of having a sunny day tomorrow?

Why imprecise probabilities?

There are a number of situations where it is difficult to determine a unique probability for the different outcomes of an experiment:

- When we have little information about the experiment, and in particular when we are completely ignorant.
- When the information we have is qualitative.
- When we have computational or time limitations.
- When there are conflicts between different sources of information.

Probability: definition

A probability is a functional P on the set of outcomes of the experiment satisfying:

- $P(\emptyset) = 0, P(\mathcal{X}) = 1$.
- $A \subseteq B \Rightarrow P(A) \leq P(B)$.
- $(A_i)_{i \in I}$ pairwise disjoint $\Rightarrow P(\cup_i A_i) = \sum_i P(A_i)$.

If it satisfies the third property for finite I, it is called a finitely additive probability, and if it satisfies it for countable I, it is called a σ-additive probability.

Finitely vs. σ-additive

σ-additive probabilities are defined on σ-fields of events, and have a number of interesting continuity properties.

On the other hand, finitely additive probabilities can always be extended to $\mathcal{P}(\mathcal{X})$, without any measurability conditions.

The theory we present here is based on finitely additive probabilities.

Aleatory vs. epistemic probabilities

In some cases, the probability of an event A is a property of the event, meaning that it does not depend on the subject making the assessment. We talk then of aleatory probabilities.

However, and specially in the framework of decision making, we may need to assess probabilities that represent *our* beliefs. Hence, these may vary depending on the subject or on the amount of information he possesses at the time. We talk then of subjective probabilities.

The behavioural interpretation

One of the possible interpretations of subjective probability is the behavioural interpretation. We interpret the probability of an event A in terms of our betting behaviour: we are disposed to bet at most $P(A)$ on the event A.

If we consider the gamble I_A where we win 1 if A happens and 0 if it doesn't happen, then we accept the transaction $I_A - P(A)$, because the expected gain is

$$(1 - P(A)) * P(A) + (0 - P(A))(1 - P(A)) = 0.$$

Gambles

More generally, we can consider our betting behaviour on gambles.

A gamble is a bounded real-valued variable on $\mathcal{X}$, $f : \mathcal{X} \to \mathbb{R}$.

It represents a reward that depends on the outcome of the experiment modelled by X.

We shall denote the set of all gambles by $\mathcal{L}(\mathcal{X})$.

Example

Who shall win the next US'Open?

Consider the outcomes a=Federer, b=Nadal, c=Murray, d=Other.

$\mathcal{X} = \{a, b, c, d\}$.

Consider the gamble $f(a) = 3, f(b) = -2, f(c) = 5, f(d) = 10$.

Depending on how likely we consider each of the outcomes we will accept the gamble or not.

Betting on gambles

Consider now a gamble f on $\mathcal{X}$. We may consider the supremum value μ such that we are disposed to pay μ for f, i.e., such that the reward $f - \mu$ is desirable: it will be the expectation $E(f)$.

- For any $\mu < E(f)$, we expect to have a gain.
- For any $\mu > E(f)$, we expect to have a loss.

Buying and selling prices

We may also give money in order to get the reward: if we are disposed to pay x for the gamble f, then the gamble $f - x$ is desirable to us.

We may also sell a gamble f, meaning that if we are disposed to sell it at a price x then the gamble $x - f$ is desirable to us.

In the case of probabilities, the supremum buying price for a gamble f coincides with the infimum selling price, and we have a fair price for f.

Existence of indecision

When we don't have much information, it may be difficult (and unreasonable) to give a fair price $P(f)$: there may be some prices μ for which we would not be disposed to buy nor sell the gamble f.

In terms of desirable gambles, this means that we would be *undecided* between two gambles.

It is sometimes considered preferable to give different values $\underline{P}(f) < \overline{P}(f)$ than to give a precise (and possibly wrong) value.

Lower and upper previsions

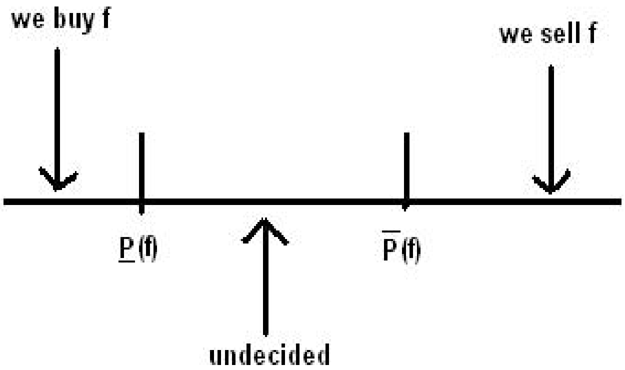
The lower prevision for a gamble f, $\underline{P}(f)$, is our supremum acceptable *buying* price for f, meaning that we are disposed to buy it for $\underline{P}(f) - \epsilon$ (or to accept the reward $f - (\underline{P}(f) - \epsilon)$) for any $\epsilon > 0$.

The upper prevision for a gamble f, $\overline{P}(f)$, is our infimum acceptable *selling* price for f, meaning that we are disposed to sell f for $\overline{P}(f) + \epsilon$ (or to accept the reward $\overline{P}(f) + \epsilon - f$) for any $\epsilon > 0$.

Example (cont.)

Consider the previous gamble
$f(a) = 3, f(b) = -2, f(c) = 5, f(d) = 10.$

- If I am certain that Nadal is not going to win US'Open, I should be disposed to accept this gamble, and even to pay as much as 3 for it. Hence, I would have $\underline{P}(f) \geq 3$.
- For the infimum selling price, if I think that the winner will be either Nadal or Federer, I should sell f for anything greater than 3, because for such prices I will always win money with the transaction. Hence, I would have $\overline{P}(f) \leq 3$.



In the precise case we have $\underline{P}(f) = \overline{P}(f)$.

Conjugacy of $\underline{P}, \overline{P}$

Under this interpretation,

$$\begin{aligned}\underline{P}(-f) &= \sup\{x : -f - x \text{ acceptable }\} \\ &= -\inf\{-x : -f - x \text{ acceptable }\} \\ &= -\inf\{y : -f + y \text{ acceptable }\} \\ &= -\overline{P}(f)\end{aligned}$$

Hence, it suffices to work with one of these two functions.

Important remark

Using this reasoning, we can determine the supremum acceptable buying prices for all gambles f in some set $\mathcal{K} \subseteq \mathcal{L}(\mathcal{X})$.

The domain $\mathcal{K}$ of $\underline{P}$:

- need not have any predefined structure.
- may contain indicators of events.

Lower probabilities of events

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

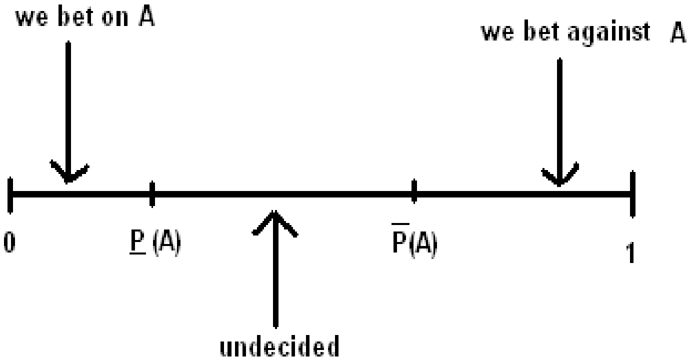
- $=$ lower prevision $\underline{P}(I_A)$ of the indicator of A.
- $=$ supremum betting rate on A.
- $=$ measure of the evidence supporting A.
- $=$ measure of the strength of our belief in A.

Upper probabilities of events

- The upper probability of A, $\overline{P}(A)$
  - $=$ upper prevision $\overline{P}(I_A)$ of the indicator of A.
  - $=$ measure of the lack of evidence against A.
  - $=$ measure of the plausibility of A.
- We have then a behavioural interpretation of upper and lower probabilities:

$$\text{evidence in favour of } A \uparrow \Rightarrow \underline{P}(A) \uparrow$$

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


we bet on A
we bet against A
0
P (A)
P(A)
1
undecided

Example(cont.)

- The lower probability we give to Nadal being the winner would be the lower prevision of I_b, where we get a reward of 1 if Nadal wins and 0 if it doesn't.
- The upper probability of Federer or Murray winning would be the upper probability of $I_{\{a,c\}}$, or, equivalently, 1 minus the lower probability of Federer and Murray not winning.

Events or gambles?

In the case of probabilities, we are indifferent between betting on events or on gambles: our betting rates on events (a probability) determine our betting rates on gambles (its expectation).

However, in the imprecise case, the lower and upper previsions for events do not determine the lower and upper previsions for gambles uniquely.

Hence, lower and upper previsions are more informative than lower and upper probabilities.

Consistency requirements

The assessments made by a lower prevision on a set of gambles should satisfy a number of consistency requirements:

- A combination of the assessments should not produce a net loss, no matter the outcome: avoiding sure loss.
- Our supremum buying price for a gamble f should not depend on our assessments for other gambles: coherence.

Avoiding sure loss

I represent my beliefs about the possible winner of US'Open saying that

$$\overline{P}(a) = 0.55, \overline{P}(b) = 0.25, \overline{P}(c) = 0.4, \overline{P}(d) = 0.1$$
$$\underline{P}(a) = 0.45, \underline{P}(b) = 0.2, \underline{P}(c) = 0.35, \underline{P}(d) = 0.05$$

where $\{a, b, c, d\} = \{Federer, Nadal, Murray, Other\}$.

This means that the gambles $I_a - 0.44$, $I_b - 0.19$, $I_c - 0.34$ and $I_d - 0.04$ are desirable for me. But if I accept all of them I get the sum

$$[I_a + I_b + I_c + I_d] - 1.01 = -0.01$$

which produces a net loss of 0.01, no matter who wins.

Avoiding sure loss: general definition

Let $\underline{P}$ be a lower prevision defined on a set of gambles $\mathcal{K}$. It avoids sure loss iff

$$\sup_{x \in \mathcal{X}} \sum_{i=1}^{n} f_i(x) - \underline{P}(f_i) \geq 0$$

for any $f_1, \dots, f_n \in \mathcal{K}$.
Otherwise, there is some $\epsilon > 0$ such that

$$\sum_{i=1}^{n} f_i - (\underline{P}(f_i) - \epsilon) < -\epsilon$$

no matter the outcome.

Consequences of avoiding sure loss

- $\underline{P}(f) \leq \sup f$.
- $\underline{P}(\mu) \leq \mu \leq \overline{P}(\mu)\ \forall \mu \in \mathbb{R}$.
- If $f \geq g + \mu$, then $\overline{P}(f) \geq \underline{P}(g) + \mu$.
- $\underline{P}(\lambda f + (1-\lambda)g) \leq \lambda \overline{P}(f) + (1-\lambda)\overline{P}(g)$.
- $\underline{P}(f+g) \leq \overline{P}(f) + \overline{P}(g)$.

Coherence

After reflecting a bit, I come up with the assessments:

$$\overline{P}(a) = 0.55, \overline{P}(b) = 0.25, \overline{P}(c) = 0.4, \overline{P}(d) = 0.1$$
$$\underline{P}(a) = 0.45, \underline{P}(b) = 0.15, \underline{P}(c) = 0.30, \underline{P}(d) = 0.05$$

These assessments avoid sure loss. However, they imply that the transaction

$$I_a - 0.44 + I_c - 0.29 + I_d - 0.04 = 0.23 - I_b$$

is acceptable for me, which means that I am disposed to bet against Nadal at a rate 0.23, smaller that $\overline{P}(b)$. This indicates that $\overline{P}(b)$ is too large.

Coherence: general definition

A lower prevision $\underline{P}$ is called coherent when given gambles $f_0, f_1, \ldots, f_n$ in its domain and $m \in \mathbb{N}$,

$$\sup_{x \in \mathcal{X}} [\sum_{i=1}^{n} [f_i(x) - \underline{P}(f_i)] - m[f_0(x) - \underline{P}(f_0)]] \geq 0.$$

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

$$\sum_{i=1}^{n} f_i - (\underline{P}(f_i) - \epsilon) < m(f_0 - \underline{P}(f_0) - \epsilon),$$

and $\underline{P}(f_0) + \epsilon$ would be an acceptable buying price for f_0.

Exercise

Consider the lower prevision given by:

	$f(1)$	$f(2)$	$f(3)$	$\underline{P}(f)$
f_1	2	1	0	0.5
f_2	0	1	2	1
f_3	0	1	0	1

(a) Does it avoid sure loss?

(b) Is it coherent?

Exercise

Let A be a non-empty subset of a (not necessarily finite) set $\mathcal{X}$. 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 $\underline{P}(f) = \inf_{x \in A} f(x)$.

(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 $f, g \in \mathcal{K}$, then $f + g \in \mathcal{K}$.
- If $f \in \mathcal{K}, \lambda \in \mathbb{R}$, then $\lambda f \in \mathcal{K}$.

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

- $\underline{P}(f) \geq \inf f$.
- $\underline{P}(\lambda f) = \lambda \underline{P}(f)$.
- $\underline{P}(f + g) \geq \underline{P}(f) + \underline{P}(g)$.

Consequences of coherence

Whenever the gambles belong to the domain of $\underline{P}, \overline{P}$:

- $\underline{P}(\emptyset) = \overline{P}(\emptyset) = 0, \underline{P}(\mathcal{X}) = \overline{P}(\mathcal{X}) = 1.$
- $A \subseteq B \Rightarrow \underline{P}(A) \leq \underline{P}(B), \overline{P}(A) \leq \overline{P}(B).$
- $\underline{P}(f) + \underline{P}(g) \leq \underline{P}(f+g) \leq \underline{P}(f) + \overline{P}(g) \leq \overline{P}(f+g) \leq \overline{P}(f) + \overline{P}(g).$
- $\underline{P}(\lambda f) = \lambda \underline{P}(f), \overline{P}(\lambda f) = \lambda \overline{P}(f)$ for $\lambda \geq 0$.
- If $f_n \to f$ uniformly, then $\underline{P}(f_n) \to \underline{P}(f)$ and $\overline{P}(f_n) \to \overline{P}(f)$.

Consequences of coherence (II)

- $\lambda\underline{P}(f) + (1-\lambda)\underline{P}(g) \leq \underline{P}(\lambda f + (1-\lambda)g)\ \forall \lambda \in [0,1]$.
- $\underline{P}(f+\mu) = \underline{P}(f) + \mu\ \forall \mu \in \mathbb{R}$.
- The lower envelope of a set of coherent lower previsions is coherent.
- A convex combination of coherent lower previsions (with the same domain) is coherent.
- The point-wise limit (inferior) of coherent lower previsions is coherent.

Exercise

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

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

Is it coherent?

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 $\mu : \mathcal{P}(\mathcal{X}) \to [0,1]$. It is called a capacity or non-additive measure when it satisfies:

1. $\mu(\emptyset) = 0, \mu(\mathcal{X}) = 1$ (normalisation).
2. $A \subseteq B \Rightarrow \mu(A) \leq \mu(B)$ (monotonicity).

Capacities are also called fuzzy measures or Choquet capacities of the 1st order. They are used as alternative models to probability where we do not require the additivity axiom.

Some types of non-additive measures

- Coherent lower probabilities.
- 2 and n-monotone capacities.
- Belief functions.
- Possibility/necessity measures.
- Convex sets of probabilities.

n-monotone capacities (Choquet, 1953)

Let $\underline{P}$ be a lower probability defined on a field of subsets $\mathcal{A}$. It is called n-monotone when

$$\underline{P}(A_1 \cup \cdots \cup A_n) \geq \sum_{\emptyset \neq I \subseteq \{1,\ldots,n\}} (-1)^{|I|+1} \underline{P}(\cap_{i \in I} A_i)$$

for any subsets $A_1, \ldots, A_n$ in $\mathcal{A}$.

Its conjugate is called n-alternating and satisfies

$$\overline{P}(A_1 \cap \cdots \cap A_n) \leq \sum_{\emptyset \neq I \subseteq \{1,\ldots,n\}} (-1)^{|I|+1} \overline{P}(\cup_{i \in I} A_i)$$

for any subsets $A_1, \ldots, A_n$ in $\mathcal{A}$.

Belief functions (Shafer, 1976)

When a lower probability $\underline{P}$ is n-monotone for every n, it is called ∞-monotone or a Choquet capacity of ∞ order. Then, for every natural number n and every family $\{A_1, \ldots, A_n\}$ of subsets of $\mathcal{X}$, it holds that

$$\underline{P}(A_1 \cup \ldots A_n) \geq \sum_{\emptyset \neq I \subseteq \{1,\ldots,n\}} (-1)^{|I|+1} \underline{P}(\cap_{i \in I} A_i).$$

When $\mathcal{X}$ is finite ∞-monotone capacities are called belief functions.

Possibility and necessity measures (Dubois and Prade, 1988)

Given a space $\mathcal{X}$, a possibility measure is a function $\Pi : \mathcal{P}(\mathcal{X}) \to [0,1]$ s.t. for every family of subsets $(A_i)_{i\in I}$ of $\mathcal{X}$,

$$\Pi(\cup_{i\in I} A_i) = \sup_{i\in I} \Pi(A_i).$$

The conjugate function of a possibility measure, given by $Nec(A) = 1 - \Pi(A^c)$, is called a necessity measure, and satisfies

$$Nec(\cap_{i\in I} A_i) = \inf_{i\in I} Nec(A_i)$$

for every family of subsets $(A_i)_{i\in I}$.

Maxitive and minitive measures

A slightly more general model than possibility measures, which is less used in practice, are the so-called maxitive measures, those upper probabilities $\overline{P} : \mathcal{P}(\mathcal{X}) \to [0,1]$ satisfying

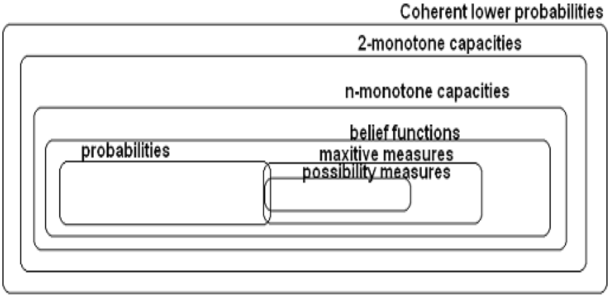
$$\overline{P}(A \cup B) = \max\{\overline{P}(A), \overline{P}(B)\}$$

para todo $A, B \subseteq \mathcal{X}$.

Their conjugate are the minitive measures, for which $\underline{P}(A \cap B) = \min\{\underline{P}(A), \underline{P}(B)\}$ for every $A, B \subseteq \mathcal{X}$.

Relationships between the definitions

The relationships between the different types of lower and upper probabilities are summarised in the following figure:



Distribution functions and p-boxes (Ferson, 2003)

We shall call a function $F : [0,1] \to [0,1]$ a distribution function when it satisfies the following two properties:

- $x \leq y \Rightarrow F(x) \leq F(y)$ (monotonicity).
- $F(1) = 1$ (normalisation).

A p-box is a pair of distribution functions, $(\underline{F}, \overline{F})$, satisfying $\underline{F}(x) \leq \overline{F}(x)$ for every $x \in [0,1]$.

The concept can be extended to arbitrary ordered spaces, producing then the so-called generalised p-boxes.

References on non-additive measures

- D. Denneberg, *Non-additive measure and integral*, 1994.
- G. Choquet, *Theory of capacities*, 1953.
- G. Shafer, *A mathematical theory of evidence*, 1976.
- G. de Cooman, M. Troffaes, E. Miranda, *n-monotone exact functionals*. J. of Math. Analysis and Applications, 347(1), 133-146, 2009.
- D. Dubois and H. Prade, *Possibility theory*, 1988.
- S. Ferson et al., Technical Report, Sandia, 2003.

Overview, Part II

1. Equivalent representations of coherent lower previsions:
   - Sets of linear previsions.
   - Sets of desirable gambles.
2. Extension to bigger domains: natural extension.
3. Related works.

Linear previsions

When $\mathcal{K} = -\mathcal{K} := \{-f : f \in \mathcal{K}\}$ and $\underline{P}(f) = \overline{P}(f)$ for all $f \in \mathcal{K}$, then $P = \underline{P} = \overline{P}$ is a called a linear or precise prevision on $\mathcal{K}$. If $\mathcal{K}$ is a linear space, this is equivalent to

- $P(f) \geq \inf f$.
- $P(f + g) = P(f) + P(g)$,

for all $f, g \in \mathcal{K}$.

These are the previsions considered by de Finetti. We shall denote by $\mathbb{P}(\mathcal{X})$ the set of all linear previsions on $\mathcal{X}$.

Linear previsions and probabilities

A linear prevision P defined on indicators of events only is a finitely additive probability.

Conversely, a linear prevision P defined on the set $\mathcal{L}(\mathcal{X})$ of all gambles is characterised by its restriction to the set of events, which is a finitely additive probability on $\mathcal{P}(\mathcal{X})$, through the expectation operator.

Coherence and precise previsions

Given a lower prevision $\underline{P}$ on $\mathcal{K}$, we can consider

$$\mathcal{M}(\underline{P}) := \{P \in \mathbb{P}(\mathcal{X}) : P(f) \geq \underline{P}(f)\ \forall f \in \mathcal{K}\}.$$

- $\underline{P}$ avoids sure loss $\iff \mathcal{M}(\underline{P}) \neq \emptyset$.
- $\underline{P}$ coherent $\iff \underline{P} = \min \mathcal{M}(\underline{P})$.

There is a 1-to-1 correspondence between coherent lower previsions and (closed and convex) sets of linear previsions.

This correspondence establishes a sensitivity analysis interpretation to coherent lower previsions.

Example (cont.)

Consider the coherent assessments:

$$\overline{P}(a) = 0.5, \overline{P}(b) = 0.2, \overline{P}(c) = 0.35, \overline{P}(d) = 0.1$$
$$\underline{P}(a) = 0.45, \underline{P}(b) = 0.15, \underline{P}(c) = 0.30, \underline{P}(d) = 0.05$$

The equivalent set of coherent previsions represents the possible models for the probabilities of each player being the winner:

$$\mathcal{M}(\underline{P}) := \{(p_a, p_b, p_c, p_d) : p_a + p_b + p_c + p_d = 1, p_a \in [0.45, 0.5], \\ p_b \in [0.15, 0.2], p_c \in [0.3, 0.35], p_d \in [0.05, 0.1]\}$$

To see that the bounds are attained, it suffices to consider the following elements of $\mathcal{M}(\underline{P})$: $(0.45, 0.15, 0.3, 0.1)$, $(0.45, 0.2, 0.3, 0.05)$, $(0.5, 0.15, 0.3, 0.05)$,$(0.45, 0.15, 0.35, 0.05)$.

Exercise

Consider an urn with 10 balls, of which 3 are red, and the other 7 are either blue or yellow.

(a) Determine the set $\mathcal{M}$ of linear previsions that represent the possible compositions of the urn.

(b) Let f be a gamble given by $f(blue) = 2, f(red) = 1, f(yellow) = -1$. Which is the lower prevision of f?

(c) Do the same for an arbitrary gamble g.

Sets of desirable gambles

Given a lower prevision $\underline{P}$, we can consider the set of gambles

$$\mathcal{D} := \{f \in \mathcal{K} : \underline{P}(f) \geq 0\}. \tag{1}$$

Conversely, given a set of gambles $\mathcal{D}$ we can define

$$\underline{P}(f) := \sup\{\mu : f - \mu \in \mathcal{D}\} \tag{2}$$

Rationality axioms for sets of desirable gambles

If we consider a set of gambles that we find desirable, there are a number of rationality requirements we can consider:

- A gamble that makes us lose money, no matter the outcome, should not be desirable, and a gamble which never makes us lose money should be desirable.
- A change of utility scale should not affect our desirability assessments.
- If two transactions are desirable, so should be their sum.

These ideas define the notion of coherence for sets of gambles.

Coherence of sets of desirable gambles

A set of desirable gambles is coherent if and only if

(D1) If $\sup f < 0$, then $f \notin \mathcal{D}$.

(D2) If $f \geq 0$, then $f \in \mathcal{D}$.

(D3) If $f, g \in \mathcal{D}$, then $f + g \in \mathcal{D}$.

(D4) If $f \in \mathcal{D}, \lambda \geq 0$, then $\lambda f \in \mathcal{D}$.

(D5) If $f + \epsilon \in \mathcal{D}$ for all $\epsilon > 0$, then $f \in \mathcal{D}$.

- If $\mathcal{D}$ is a coherent set of gambles, then the lower prevision by Eq. (2) it induces is coherent.
- Conversely, a coherent lower prevision $\underline{P}$ determines a coherent set of desirable gambles through Eq. (1) the previous formula.

Exercise

Let $\mathcal{X} = \{1, 2, 3\}$, and consider the following sets of desirable gambles:

$$\mathcal{K}_1 := \{f : f(1) + f(2) + f(3) \geq 0\}$$
$$\mathcal{K}_2 := \{f : \max\{f(1), f(2), f(3)\} \geq 0\}.$$

(a) Are $\mathcal{K}_1, \mathcal{K}_2$ coherent?

(b) If they are, which is the lower prevision they induce on the gamble f given by $f(1) = 2, f(2) = 3, f(3) = -1$?

Desirable gambles and linear previsions

Let $\mathcal{D}$ be a coherent set of desirable gambles. Then the set

$$\mathcal{M}_{\mathcal{D}} := \{P \in \mathbb{P}(\mathcal{X}) : P(f) \geq 0 \ \forall f \in \mathcal{D}\}$$

is a closed and convex set of linear previsions. Conversely, given a closed and convex set $\mathcal{M}$ of linear previsions, the set

$$\mathcal{D}_{\mathcal{M}} := \{f \in \mathcal{L}(\mathcal{X}) : P(f) \geq 0 \ \forall P \in \mathcal{M}\}$$

is a coherent set of desirable gambles.

Important remark

The sets of gambles satisfying the previous axioms are usually called almost-desirable gambles in the literature. There are other approaches, such as sets of really, strict or marginal gambles which are also interesting, and which are sometimes more informative than almost-desirable gambles.

More information about the differences between all these notions will be given in next talk.

Hence, we have three equivalent representations of our beliefs:

1. Coherent lower and upper previsions.
2. Closed and convex sets of linear previsions.
3. Coherent sets of desirable gambles,

and we can easily go from any of these formulations to the others.

Is coherence too strong?

Some critics to the property of coherence are:

- Descriptive decision theory shows that sometimes beliefs violate the notion of coherence.
- Coherent lower previsions may be difficult to assign for people not familiar with the behavioural theory of imprecise probabilities.
- Other rationality criteria may be also interesting.

Inference: natural extension

Consider the following gambles:

$$f(a) = 5, f(b) = 2, f(c) = -5, f(d) = -10$$
$$g(a) = 2, g(b) = -2, g(c) = 0, g(d) = 5,$$

and assume we make the assessments $\underline{P}(f) = 2, \underline{P}(g) = 0$. Can we deduce anything about how much should we pay for the gamble

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

using the axioms of coherence?

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

Definition

Consider a coherent 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

$$\underline{E}(f) := \sup\{\mu : \exists f_k \in \mathcal{K}, \lambda_k \geq 0, k = 1, \ldots, n : \\ f - \mu \geq \sum_{i=1}^{n} \lambda_k (f_k(x) - \underline{P}(f_k))\}$$

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

Example

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

$$h - 3.4 \geq 1.2(f - \underline{P}(f)).$$

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

Natural extension: properties

- If $\underline{P}$ does not avoid sure loss, then $\underline{E}(f) = +\infty$ for any gamble f.
- If $\underline{P}$ avoids sure loss, then $\underline{E}$ is the smallest coherent lower prevision on $\mathcal{L}(\mathcal{X})$ that dominates $\underline{P}$ on $\mathcal{K}$.
- $\underline{P}$ is coherent if and only if $\underline{E}$ coincides with $\underline{P}$ on $\mathcal{K}$.
- $\underline{E}$ is then the least-committal extension of $\underline{P}$: if there are other extensions, they reflect stronger assessments than those in $\underline{P}$.

In terms of sets of linear previsions

Given a lower prevision $\underline{P}$ and its set of dominating linear prevision $\mathcal{M}(\underline{P})$, the natural extension $\underline{E}$ of $\underline{P}$ is the lower envelope of $\mathcal{M}(\underline{P})$.

- This provides the natural extension with a sensitivity analysis interpretation.

We may then consider the previsions that dominate $\underline{P}$ on $\mathcal{K}$, extend them to $\mathcal{L}(\mathcal{X})$, and take the lower envelope to compute the natural extension.

In terms of sets of gambles

Consider a coherent set of desirable gambles $\mathcal{D}$. Its natural extension $\mathcal{E}$ is the set of gambles

$$\mathcal{E} := \{g \in \mathcal{L}(\mathcal{X}) : (\forall \delta > 0)(\exists n \geq 0, \lambda_k \in \mathbb{R}^+, f_k \in \mathcal{D}) \\ g \geq \sum_{k=1}^{n} \lambda_k f_k - \delta\}.$$

It is the smallest coherent set of desirable gambles that contains $\mathcal{D}$. It is the smallest closed convex cone including $\mathcal{D}$ and all non-negative gambles.

All these procedures of natural extension agree with one another: if we consider a coherent lower prevision $\underline{P}$, its set of desirable gambles $\mathcal{D}_{\underline{P}}$, the natural extension of this set $\mathcal{E}_{\mathcal{D}_{\underline{P}}}$ and then the coherent lower prevision associated to this set, we obtain the natural extension of $\underline{P}$.

Hence, we have three equivalent ways of representing our behavioural dispositions:

- Coherent lower previsions.
- Sets of linear previsions.
- Sets of desirable gambles.

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 $\underline{E}$ of $\underline{P}_A$ is equal to the vacuous lower prevision relative to A:

$$\underline{E}(f) = \underline{P}_A(f) = \inf_{x \in A} f(x),$$

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

Exercise

Consider $\mathcal{X} = \{1, 2, 3\}$ and let $\mathcal{D}$ be the gambles given by $\{(1, -1, 0), (0, 1, -1), (1, 0, -1)\}$.

(a) Calculate the set $\mathcal{E}$ induced by $\mathcal{D}$.

(b) Determine the set $\mathcal{M}$ and the lower prevision $\underline{P}$ induced by $\mathcal{E}$.

(c) What is the lower prevision of the gamble $(0, 1, 2)$?

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.

Related works

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

The work of de Finetti

The theory of coherent lower previsions is an extension, allowing for indecision, of earlier works by Bruno de Finetti.

There are, however, certain points of disagreement, mostly on the treatment of the problem of *updating* a coherent prevision:

- The interpretation given by de Finetti is different.
- There is an issue with the notion of *conglomerability*.

The work of Williams

A first extension of de Finetti's work allowing for imprecision was made in the 70s by Peter Williams. It derives coherent lower previsions from sets of desirable gambles. The main differences are:

- Like de Finetti, Williams does not require the property of conglomerability.
- His approach only allows to update the coherent lower previsions by means of experiments with a finite number of possible values.
- On the other hand, his approach allows for some nice mathematical properties that do not hold in general with Walley's approach.

The work of Kuznetsov

In parallel with Walley, Vladimir Kutzetsov established a theory for lower envelopes of finitely additive probability measures. Their work differ in a number of things:

- The lack of a behavioural interpretation.
- Kuznetsov allows for unbounded gambles.
- He makes, however, a number of assumptions on the domains.
- The treatment of the updating problem is also different.

See `http://www.sipta.org/` for more information.

The work of Weichselberger

Kurt Weichselberger and some of his colleagues have established a theory of interval-valued probabilities which has many things in common with the theory of coherent lower previsions. Some of the differences are:

- It is based on a frequentist, instead of subjective approach to probability.
- It makes some measurability assumptions absent from Walley's theory.
- The updating of the imprecise probabilities and the modelling of the notion of independence is also different.

The work of Shafer and Vovk

Glenn Shafer and Vladimir Vovk have connected probability and finance through a game-theoretic approach to coherent lower previsions.

Their approach is similar to the behavioural one from coherent lower previsions, but they consider a game with two players and allow for unbounded gambles.

With these ideas, they obtain laws of large numbers for probability protocols satisfying their idea of coherence.

A tighter connection with Walley's theory has been established recently by Gert de Cooman and Filip Hermans.

Some references

Unless stated otherwise, the results and definitions can be found in chapters 1-3 from:

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

Additional references:

- M. Troffaes and G. de Cooman, *Intelligent systems for information processing: for representation to application*, pages 277-288, 2003.
- M. Troffaes, *Soft methods for integrated uncertainty modelling*, pages 201-210, 2006.
- E. Miranda, *A survey of the theory of coherent lower previsions*. Int. J. of App. Reasoning, 48(2):628–658, 2008.

Further reading

- 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.
- R. Pelessoni, P. Vicig, Int. J. of Appr. Reasoning, 39(2-3), 297-319, 2005.
- P. Williams, *Notes on conditional previsions*. Technical Report, Univ. of Sussex, 1975.