An introduction to the theory of coherent lower previsions

Enrique Miranda and Gert de Cooman
Rey Juan Carlos University and Ghent University

Overview, Part I

- Some considerations about probability.
- Coherent previsions and probabilities.
- Coherent lower and upper previsions.
- Sets of desirable gambles and linear previsions.
- Natural extension.

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 30' from the airport to the center of Montpellier by car?
- What is the probability of having a sunny day tomorrow?

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.

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 Wimbledon?

Consider the outcomes a=Federer, b=Nadal, c=Djokovic, 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 I consider each of the outcomes I 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

I may also give money in order to get the reward: if I am disposed to pay x for the gamble f, then the gamble $f - x$ is desirable to me.

I may also sell a gamble f, meaning that if I am disposed to sell it at a price x then the gamble $x - f$ is desirable to me.

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 or 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

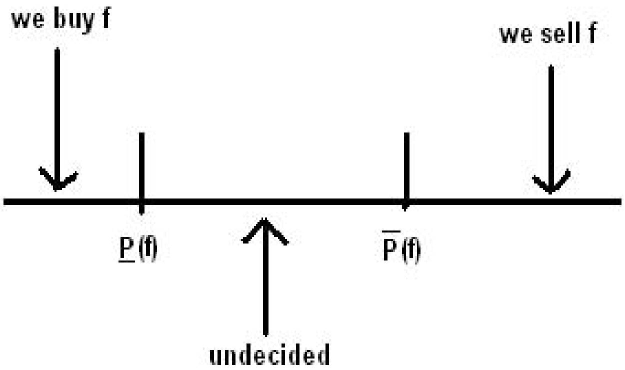
The lower prevision for a gamble f, $\underline{P}(f)$, is my supremum acceptable *buying* price for f, meaning that I am disposed to buy it for $\underline{P}(f) - \epsilon$ (or to accept the reward $f - (\underline{P}(f) - \epsilon)$) for any $\epsilon > 0$.

The upper prevision for a gamble f, $\overline{P}(f)$, is my infimum acceptable *selling* price for f, meaning that I am 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 Wimbledon, 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, I 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.

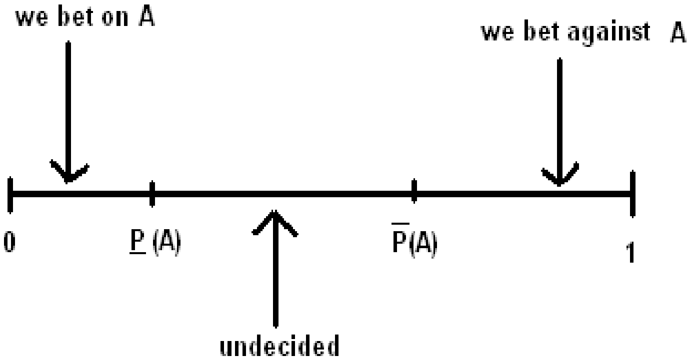
$=$ measure of the strength of our belief in A.

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 Spain wins and 0 if it doesn't.

- The upper probability of Federer or Djokovic winning would be the upper probability of $I_{\{a,c\}}$, or, equivalently, 1 minus the lower probability of Federer and Djokovic 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 Wimbledon 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 $\{a, b, c, d\} = \{Federer, Nadal, Djokovic, Other\}$.

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

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

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

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

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 $m \in \mathbb{N}$,

$$\sum_{i=1}^{n}[f_i(\omega) - \underline{P}(f_i)] \geq m[f_0(\omega) - \underline{P}(f_0)]$$

for some $\omega \in \mathcal{X}$.
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.

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.

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

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

- $\underline{P}$ 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 team 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 $\mathcal{M}(\underline{P})$: $(0.45, 0.15, 0.3, 0.1)$, $(0.45, 0.2, 0.3, 0.05)$, $(0.5, 0.15, 0.3, 0.05)$, $(0.45, 0.15, 0.35, 0.05)$.

Sets of desirable gambles

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

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

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

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

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.
- 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

- If $\sup f < 0$, then $f \notin \mathcal{D}$.
- If $f \geq 0$, then $f \in \mathcal{D}$.
- If $f, g \in \mathcal{D}$, then $f + g \in \mathcal{D}$.
- If $f \in \mathcal{D}, \lambda \geq 0$, then $\lambda f \in \mathcal{D}$.
- 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 it induces is coherent.
- Conversely, a coherent lower prevision $\underline{P}$ determines a coherent set of desirable gambles through the previous formula.

Desirable gambles an 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 a convex set $\mathcal{M}$ of linear prevision, 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.

Hence, we have three equivalent representations of our beliefs:

- Coherent lower and upper previsions.
- Closed and convex sets of linear previsions.
- 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.

Particular cases

As particular cases of coherent lower and upper previsions we have the following models:

- Probability measures.
- n-monotone and n-alternating capacities.
- Belief and plausibility measures.
- Possibility and necessity measures.

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

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

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

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

$\underline{E}(f)$ 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.

Particular cases

The natural extension coincides with some familiar extension procedures for some particular cases of coherent lower previsions:

- Lebesgue integration of a probability measure.
- Choquet integration of 2-monotone lower probabilities.
- Bayes' rule for probability measures.
- Robust Bayesian analysis.
- Logical deduction.

Related works

- B. de Finetti.
- V. Kuznetsov.
- K. Weischelberger.

Some references

- T. Fine, *Theories of probability*, Academic Press, 1973.
- B. de Finetti, *Theory of Probability*, John Wiley and Sons, 1974.
- H. Kyburg and J. Smoker (eds.), *Studies in subjective probability*, Wiley, New York, 1980.
- P. Walley, *Statistical reasoning with imprecise probabilities*, Chapman and Hall, 1991.

Overview, Part II

- Conditional lower previsions.
- Coherence of conditional and unconditionals.
- Natural and regular extensions.
- Weak and strong coherence.

Conditional lower previsions

- Definition.
- Consistency requirements.
- Natural extension.

Updating information

So far, we have assumed that all we know about the outcome of the experiment modelled by is that it belongs to a set $\mathcal{X}$.

But we may have some additional information about this outcome, for instance that it belongs to a set B.

We need to update then our assessment by means of a conditional lower prevision.

Example (cont.)

We are in the semifinals of Wimbledon, and the remaining players are Nadal, Djokovic, Roddick, and Hewitt.

For the gamble f on $\{a, b, c, d\} = \{Federer, Nadal, Djokovic, Other\}$ given by $f(a) = 5, f(b) = 2, f(c) = -5, f(d) = -10$, I had given the supremum buying price $\underline{P}(f) = 2$.

But now I should probably lower this supremum buying price, unless I am certain that Nadal will be the winner!

The updated and the contingent interpretation

Consider a subset B of $\mathcal{X}$, and a gamble f on $\mathcal{X}$.

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

We can also consider the updated interpretation, where $\underline{P}(f|B)$ is his supremum acceptable buying price for f, 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 f B-desirable when it is desirable provided the outcome of the experiment belongs to B.

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

This relates present and future dispositions for the subject.

Conditional lower previsions

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

We define

$$G(f|B) = I_B(f - \underline{P}(f|B)), G(f|\mathcal{B}) := \sum_{B \in \mathcal{B}} G(f|B) = f - \underline{P}(f|\mathcal{B}).$$

These are *desirable* gambles.

The conglomerative principle

We are using here the conglomerative principle: if a gamble f is B-desirable for all B in a partition $\mathcal{B}$, then it is desirable.

This implies that the gamble $G(f|\mathcal{B}) = \sum_{B\in\mathcal{B}} G(f|B)$ is desirable.

The condition follows from the axioms of desirability for finite partitions, but not in general.

Separate coherence

A first consistency requirement is that the updated assessments are separately coherent. This means that:

- $\underline{P}(B|B) = 1$ for any $B \in \mathcal{B}$.
- $\underline{P}(\cdot|B)$ is a coherent lower prevision.

- Consequence: $\underline{P}(\cdot|B)$ is determined by its values on B: for any $B \in \mathcal{B}$,

$$I_B h = I_B h' \Rightarrow \underline{P}(h|B) = \underline{P}(h'|B).$$

Example(cont)

The assessments
$\underline{P}(a) = 0.45, \underline{P}(b) = 0.15, \underline{P}(c) = 0.3, \underline{P}(d) = 0.05$ that I gave before the championships started, are not coherent anymore: separate coherence implies that

$$\underline{P}(a|Nadal, Djokovic, Roddick, Hewitt) = 0.$$

I should have $\underline{P}(b, c, d|Nadal, Djokovic, Roddick, Hewitt) = 1$.

Separate coherence: equivalent formulation

If the domain $\mathcal{K}$ of $\underline{P}(f|\mathcal{B})$ is a linear space that includes all constant gambles, this holds if and only if for any $\lambda \geq 0, f, g \in \mathcal{K}$ and $B \in \mathcal{B}$,

- $\underline{P}(f|B) \geq \inf_{x \in B} f(x)$
- $\underline{P}(f+g|B) \geq \underline{P}(f|B) + \underline{P}(g|B)$
- $\underline{P}(\lambda f|B) = \lambda \underline{P}(f|B)$

Unconditional lower previsions

An unconditional lower prevision can be seen as a particular case of conditional lower prevision, with respect to the trivial partition $\mathcal{B} := \{\mathcal{X}\}$.

In that case, the notion of separate coherence reduces to the coherence we saw in the unconditional case.

Conditional linear previsions

As a particular case, we also have that of conditional linear previsions. A conditional lower prevision $P(\cdot|\mathcal{B})$ with linear domain is linear when

- $P(f|B) \geq \inf_{x \in B} f(x)$
- $P(f+g|B) = P(f|B) + P(g|B)$
- $P(\lambda f|B) = \lambda P(f|B)$

for any $\lambda \geq 0, f, g \in \mathcal{K}$ and $B \in \mathcal{B}$.

Consistency with the initial assessments

Not only our updated lower previsions have to be coherent, but we need them to be coherent with the initial assessments.

For instance, if we consider a gamble f on $\{a, b, c, d\}$ given by $f(a) = -1, f(b) = 0, f(c) = 1, f(d) = 2$ and we make $\underline{P}(f) = 1.5$, it does not make sense that if we learn that the outcome of the experiment is either c or d then we make $\underline{P}(f|\{c, d\}) = 1$.

The connection between unconditional and conditional lower previsions follows from the updating principle.

Coherence of conditional and unconditional previsions

Consider an unconditional lower prevision $\underline{P}$ on a linear space of gambles $\mathcal{K}$ and a conditional lower prevision $\underline{P}(\cdot|\mathcal{B})$ with linear domain $\mathcal{H}$. They are coherent if and only if for any $f_1, f_2 \in \mathcal{K}$, $g_1, g_2 \in \mathcal{H}$ and $B \in \mathcal{B}$,

- $\sup_x [G(f_1) + G(g_1|\mathcal{B}) - G(f_2)](x) \geq 0$.
- $\sup_x [G(f_1) + G(g_1|\mathcal{B}) - G(g_2|B)](x) \geq 0$.

Interpretation

In the first condition, we require that the supremum acceptable buying price for f_2 should not be raised by considering the acceptable transactions $G(f_1), G(g_1|\mathcal{B})$.

In the second, we require that the supremum acceptable buying price for g_2, contingent in some $B \in \mathcal{B}$, should not be raised by considering the acceptable transactions $G(f_1), G(g_1|\mathcal{B})$.

A similar condition can be given for non-linear domains.

Hypotheses on the domains

This definition makes the following assumptions:

- The domains $\mathcal{K}, \mathcal{H}$ are linear spaces.
- Given $f \in \mathcal{H}$, $\underline{P}(f|\mathcal{B})$ and $I_B f$ also belong to $\mathcal{H}$ for all $B \in \mathcal{B}$.
- $\underline{P}$ is coherent and $\underline{P}(\cdot|\mathcal{B})$ is separately coherent.

The second assumption follows easily from the third, and the first can be relaxed to arbitrary domains.

Consequences of coherence

If $\underline{P}, \underline{P}(\cdot|\mathcal{B})$ are coherent, the following conditions hold whenever the involved gambles are defined:

- $\underline{P}(f) \geq \inf \underline{P}(f|\mathcal{B})$.
- $\underline{P}(f) \geq \underline{P}(\underline{P}(f|\mathcal{B}))$.
- $\underline{P}(G(f|\mathcal{B})) \geq 0$.
- $\underline{P}(f|B) \geq 0 \Rightarrow \underline{P}(B)\underline{P}(f|B) \leq \underline{P}(fI_B) \leq \overline{P}(B)\underline{P}(f|B) \leq \overline{P}(fI_B)$.

If the domain $\mathcal{K}$ of $\underline{P}$ includes the domain $\mathcal{H}$ of $\underline{P}(\cdot|\mathcal{B})$, then coherence is equivalent to:

- $\underline{P}(G(f|\mathcal{B})) \geq 0 \; \forall f \in \mathcal{H}$.
- $\underline{P}(G(f|B)) = 0 \; \forall f \in \mathcal{H}, B \in \mathcal{B}$.

The Generalised Bayes Rule

The second of these conditions is called the Generalised Bayes Rule, and can be used to derive the conditional lower prevision from $\underline{P}$.

- If $\underline{P}(B) > 0$, then $\underline{P}(f|B)$ is the unique value that satisfies the Generalised Bayes Rule.
- In that case, $\underline{P}(f|B)$ can be calculated as the lower envelope of the values $P(f|B)$, where $P \geq \underline{P}$ and $P(f|B)$ is calculated using Bayes' rule.

Example (cont.)

Given our initial coherent assessments

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

and if we know that the outcome will belong to $\{b, c, d\}$, we can update them using the envelope theorem, obtaining

$$\underline{P}(b|\{b, c, d\}) = 3/11, \underline{P}(c|\{b, c, d\}) = 6/11, \underline{P}(d|\{b, c, d\}) = 1/11.$$

Coherence in the linear case

When P and $P(\cdot|\mathcal{B})$ are linear and the partition $\mathcal{B}$ is finite, the GBR becomes

$$P(f|B) = \frac{P(fI_B)}{P(B)} \text{ if } P(B) > 0.$$

If $\mathcal{B}$ is infinite, coherence is equivalent to $P(f) = P(P(f|\mathcal{B}))$ for any gamble f, which is stronger than the GBR.

Natural extension

As in the unconditional case, we can study the behavioural consequences of the assessments given by coherent conditional and unconditional previsions. Given coherent $\underline{P}$ on $\mathcal{K}$ and $\underline{P}(\cdot|\mathcal{B})$ on $\mathcal{H}$, their natural extension is

$$\underline{E}(f) = \sup\{\mu : \exists f_1 \in \mathcal{K}, f_2 \in \mathcal{H}, f - \mu \geq G(f_1) + G(f_2|\mathcal{B})\},$$

and

$$\underline{E}(f|B) = \begin{cases} \max\{\beta : \underline{E}(I_B(f-\beta)) \geq 0\} & \text{if } \underline{E}(B) > 0 \\ \sup\{\beta : I_B(f-\beta) \geq G(g|B) \text{ for some } g \in \mathcal{H}\} & \text{otherwise} \end{cases}$$

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

Problems with the natural extension

The definition of the natural extension does not always have the properties of the natural extension from the unconditional case:

- In some cases there are no coherent extensions.
- If there are, the natural extension may be only a lower bound of the smallest coherent extensions.
- It provides the smallest coherent extensions when the partition $\mathcal{B}$ is finite.

Conglomerability

Let $\underline{P}$ be a coherent lower prevision on $\mathcal{L}(\mathcal{X})$ and let $\mathcal{B}$ be a partition of $\mathcal{X}$. $\underline{P}$ is called $\mathcal{B}$-conglomerable when given distinct sets $(B_n)_n$ in $\mathcal{B}$ for which $\underline{P}(B_n) > 0$ for all n, then

$$\underline{P}(I_{B_n} f) \geq 0 \ \forall n \Rightarrow \underline{P}(f) \geq 0.$$

This is equivalent to the existence of a conditional lower prevision $\underline{P}(\cdot|\mathcal{B})$ on $\mathcal{L}(\mathcal{X})$ which is coherent with $\underline{P}$.

Full conglomerability

A lower prevision which is $\mathcal{B}$-conglomerable for any partition $\mathcal{B}$ of $\mathcal{X}$ is called fully conglomerable.

This is rational if we admit the updating and conglomerative principles.

It is one of the points of disagreement between Walley and de Finetti's approach to conditioning.

Relationship with σ-additivity

Let P be a linear prevision on $\mathcal{L}(\mathcal{X})$ taking infinitely many different values on events. The following are equivalent:

- P is fully conglomerable.
- For any countable partition $(B_n)_n$ of $\mathcal{B}$, $\sum_n P(B_n) = 1$.

Other types of extensions

The natural extension is not the only possibility to coherently update an unconditional lower prevision. Other possibilities are:

- Regular extension.
- Marginal extension.

Regular extension

Let $\underline{P}$ be coherent on $\mathcal{L}(\mathcal{X})$, and let $\mathcal{B}$ be a partition of $\mathcal{X}$. Assume that $\overline{P}(B) > 0$ for all B. The regular extension $\underline{R}(\cdot|\mathcal{B})$ is defined by

$$\underline{R}(f|B) := \inf \left\{ \frac{P(I_B f)}{P(B)} : P \geq \underline{P}, P(B) > 0 \right\}$$

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

Properties

If $\mathcal{B}$ is finite:

- $\underline{P}, \underline{R}(\cdot|\mathcal{B})$ are coherent.
- $\underline{R}(\cdot|\mathcal{B})$ is the greatest conditional lower prevision which is coherent with $\underline{P}$.

In the infinite case the regular extension is not necessarily coherent, but gives an upper bound of any coherent extension.

Marginal extension

Let $\underline{P}$ be a coherent lower prevision on the $\mathcal{B}$-constant gambles, and $\underline{P}(\cdot|\mathcal{B})$ a separately coherent lower prevision on $\mathcal{L}(\mathcal{X})$. Their marginal extension is given by

$$\underline{M}(f) = \underline{P}(\underline{P}(f|\mathcal{B})).$$

The marginal extension is used to put together hierarchical information.

Properties

- $\underline{M}$ is the smallest coherent lower prevision which is coherent with $\underline{P}, \underline{P}(\cdot|\mathcal{B})$.
- $\underline{M}, \underline{P}(\cdot|\mathcal{B})$ are the lower envelopes of a set of dominating coherent linear previsions $\{P_\gamma, P_\gamma(\cdot|\mathcal{B}) : \gamma \in \Gamma\}$.
- The result holds for infinite spaces, and for a finite number of nested partitions.

Several conditional previsions

Can consider a number of different partitions $\mathcal{B}_1, \ldots, \mathcal{B}_n$ of $\mathcal{X}$, and separately coherent conditional lower previsions $\underline{P}_1(\cdot|\mathcal{B}_1), \ldots, \underline{P}_m(\cdot|\mathcal{B}_m)$ with linear domains $\mathcal{H}_1, \ldots, \mathcal{H}_n$.

There are several ways of generalising the notion of coherence to this case:

- Weak coherence.
- Coherence.

Weak coherence

$\underline{P}_1(\cdot|\mathcal{B}_1), \dots, \underline{P}_m(\cdot|\mathcal{B}_m)$ are weakly coherent if given $f_0, f_1, \dots, f_m$ in the domains, $B \in \mathcal{B}_j$ for some $j \in \{1, \dots, m\}$,

$$\sup_{\omega}[\sum_{i=1}^{m} G_i(f_i|\mathcal{B}_i) - G_j(f_0|B)](\omega) \geq 0.$$

If this condition does not hold, the supremum buying price for f_0 contingent on B can be raised taking into account the buying prices for other gambles.

Properties of weak coherence

- Weak coherence is equivalent to the existence of a joint lower prevision $\underline{P}$ which is coherent with each $\underline{P}_j(\cdot|\mathcal{B}_j)$.
- The smallest joint to be coherent with each of them is given by

$$\underline{P}(f) := \sup\{\alpha : \exists f_j, j = 1, \dots, m, \text{s.t.} \\ \sup_{\omega}[\sum_{j=1}^{m} G(f_j|\mathcal{B}_j) - (f - \alpha)](\omega) < 0\}.$$

But in some cases it can be too weak: the assessments

$$\begin{aligned} X_1 = 1 \Rightarrow X_2 = 2 \Rightarrow X_1 = 2, \\ X_1 = 2 \Rightarrow X_2 = 1 \Rightarrow X_1 = 1, \\ X_1 = 3 \Leftrightarrow X_2 = 3 \end{aligned}$$

can be modelled by weakly coherent conditional lower previsions.

Coherence

$\underline{P}_1(\cdot|\mathcal{B}_1), \dots, \underline{P}_m(\cdot|\mathcal{B}_m)$ are coherent if given $f_0, f_1, \dots, f_m$ in the domains $B \in \mathcal{B}_j$ for some $j \in \{1, \dots, m\}$, there is some $C \in \{B\} \cup \cup_{i=1}^m S_i(f_i)$ such that

$$\sup_{\omega \in C} [\sum_{i=1}^m G_i(f_i|\mathcal{B}_i) - G_j(f_0|B)](\omega) \geq 0.$$

where $S_i(f_i) := \{B_i \in \mathcal{B}_i : I_{B_i} f_i \neq 0\}$.

This notion is not compatible with the inconsistent assessments considered before.

Relationships between the two notions

- Coherence implies weak coherence.
- In the case of one conditional and one unconditional lower prevision, both notions are equivalent.
- Coherence is equivalent to the existence of a joint coherent with all the conditionals, *taken together*.

Cardinality of the partitions

There are some properties that hold only when all the partitions $\mathcal{B}_1, \dots, \mathcal{B}_m$ have a finite number of elements.

For instance, in that case we have *envelope theorems*:

- Weakly coherent conditional lower previsions are lower envelopes of weakly coherent conditional linear previsions.
- Coherent conditional lower previsions are lower envelopes of coherent conditional linear previsions.

Weak vs. strong coherence

If $\underline{P}_1(\cdot|\mathcal{B}_1),\dots,\underline{P}_m(\cdot|\mathcal{B}_m)$ are weakly coherent (with $\underline{P}$) but not coherent, the incoherence is caused in a set of lower probability zero.

If the conditionals are linear, the incoherence is caused in a set of upper probability zero.

If all the elements of the partition have positive lower probability, then weak and strong coherence are equivalent.

Natural extension

Given conditional $\underline{R}(\cdot|\mathcal{B}_1), \dots, \underline{R}(\cdot|\mathcal{B}_m)$ with linear domains $\mathcal{H}_1, \dots, \mathcal{H}_m$, their natural extensions to all gambles are given by

$$\underline{E}(f|B_0) := \sup\{\alpha : \exists f_i \in \mathcal{H}^i, i = 1, \dots, m, \text{ s.t.} \sup_{\omega \in C} \sum_{i=1}^{n} G(f_i|\mathcal{B}_i) - I_{B_0}(f - \alpha) < 0\}$$

for all $C \in B_0 \cup \bigcup_i S_i(f_i)$.

Properties of the natural extension

If all the partitions are finite:

- The natural extensions are the smallest coherent extensions to all gambles.
- They are the lower envelopes of all the coherent extensions.
- When one of the lower previsions is unconditional, the natural extensions may take a simpler form.

Regular extension

Given an unconditional $\underline{P}$ on $\mathcal{L}(\mathcal{X})$, we can also define the conditional lower previsions $\underline{R}(\cdot|\mathcal{B}_1), \dots, \underline{R}(\cdot|\mathcal{B}_m)$ using regular extension, provided $\overline{P}(B_i) > 0$ for all $B_i \in \mathcal{B}_i, i = 1, \dots, m$:

$$\underline{R}(f|B_i) := \inf \left\{ \frac{P(I_{B_i} f)}{P(B_i)} : P \geq \underline{P}, P(B_i) > 0 \right\}$$

for any $B_i \in \mathcal{B}_i, f \in \mathcal{L}(\mathcal{X}), i = 1, \dots, m$.

These are the greatest updated models which are coherent or weakly coherent with $\underline{P}$.

Weak vs. strong coherence (II)

Let $\underline{P}(\cdot|\mathcal{B}_1), \ldots, \underline{P}(\cdot|\mathcal{B}_m)$ be separately coherent, and assume that some of the partitions $\mathcal{B}_i$ are infinite.

- Weak or strong coherence lower conditionals may not be lower envelopes of weak or strong coherence *linear* conditionals.
- If some partition is uncountable, it is not possible that all its elements have positive lower probability, and weak and strong coherence are not equivalent.
- The difference between them is still related to conditioning on events of lower probability zero.

Natural extensions (II)

When some of the partitions are infinite:

- The natural extensions may not be coherent.
- They may not coincide with the smallest coherent extensions.
- They are a lower bound of any coherent extensions.

Regular extensions (II)

Let $\underline{R}(\cdot|\mathcal{B}_1), \dots, \underline{R}(\cdot|\mathcal{B}_m)$ be defined from some unconditional $\underline{P}$ on $\mathcal{L}(\mathcal{X})$ by regular extension. If some of the partitions are infinite:

- $\underline{R}(\cdot|\mathcal{B}_1), \dots, \underline{R}(\cdot|\mathcal{B}_m)$ may not be coherent with $\underline{P}$.
- They are an upper bound of any coherent extensions.

$\hookrightarrow$ Note that this may not apply if $\underline{P}$ is not defined on all gambles!

In terms of variables

Consider variables $\{X_1, \dots, X_n\}$, taking values in *finite* sets $\mathcal{X}_1, \dots, \mathcal{X}_n \subseteq \mathbb{R}$. Given $J \subseteq \{1, \dots, n\}$, we denote $X_J = \prod_{i \in J} X_i$ and $\mathcal{X}_J = \prod_{i \in J} \mathcal{X}_i$.

For any set of variables I, $\{\pi_I^{-1}(x) : x \in \mathcal{X}_I\}$ constitutes a partition of $\mathcal{X}^n$.

Given disjoint $O, I \subseteq \{1, \dots, n\}$, the conditional lower prevision $\underline{P}(X_O|X_I)$ represents the information that the variables in I provide about the variables in O.

In terms of variables (II)

$\underline{P}(X_O|X_I)$ will be defined in the set of gambles that depend on the value that the variables in $O \cup I$: the $\mathcal{X}_{O\cup I}$-measurable gambles. This is a subset of $\mathcal{L}(\mathcal{X}^n)$.

We interpret $\underline{P}(f|x)$ as the supremum acceptable buying price for a gamble f if we learn that X_I has taken the value x.

All the previous definitions and results can also be established under this terminology.

Related works

- P. Williams.
- G. Shafer and V. Vovk.
- G. Coletti and R. Scozzafava.

Some references

- P. Walley, *Statistical reasoning with imprecise probabilities*, Chapman and Hall, 1991.
- P. Williams, IJAR 44(3), 366–383, 2007.
- P. Walley, R. Pelessoni, P. Vicig, JSPI 126, 119–151, 2004.
- E. Miranda, IPMU'2008.

Strengths of the theory

- It is better suited for situations where the information does not allow to use a precise probability.
- It encompasses as particular cases most of the other generalisations in the literature.
- The behavioural interpretation leads naturally to decision making.
- We can also give it a sensitivity analysis interpretation.

Challenges

- Coherence may be too weak.
- Generalisation to unbounded gambles.
- Conglomerability vs. natural extension.
- Computing with coherent lower previsions may be too costly from a computational point of view.
- Develop tools to compare precise and imprecise models.
- There is not a unique notion of independence, and sometimes it can be hard to choose between the different notions.
- Some results from classical probability theory still need to be generalised.