Independent Natural Extension for Infinite Spaces

If you are not familiar with sets of desirable gambles, lower previsions, Williams-coherence, epistemic independence or independent natural extension, this poster may make little sense at first. I will do my very best to compensate with enthusiasm! If I fail, we can also simply go for a beer. In any case, the thought bubbles below may serve as a nice discussion starter.

Williams-Coherence to the Rescue!

Jasper De Bock

jasper.debock@ugent.be

Ghent University, Belgium

Modelling Uncertainty

A subject's uncertainty about a variable X that takes values x in a—**possibly infinite**—set $\mathscr{X}$ can be modelled in various ways. We consider two very general and closely connected frameworks, the latter of which **includes probabilities as a special case**.

Sets of desirable gambles. The basic idea here is to consider **the subject's attitude towards gambles** on $\mathscr{X}$, which are bounded real-valued functions f on $\mathscr{X}$ whose value $f(x)$ represents the—possibly negative—payoff for the outcome x. In particular, **we consider the gambles that she finds desirable**, in the sense that she prefers them over not betting at all. We gather all these gambles in a so-called **set of desirable gambles** $\mathscr{D}$, which is a subset of the set $\mathscr{G}(\mathscr{X})$ of all gambles.

Conditional lower previsions. Here too, the idea is to model a subject's uncertainty about X by considering her **attitude towards gambles** on $\mathscr{X}$. However, in this case, instead of considering sets of gambles, **we consider the prices at which a subject is willing to buy these gambles**. Let $\mathscr{C}(\mathscr{X})$ be the set of all pairs (f,B), where f is a gamble on $\mathscr{X}$ and B is a non-empty subset of $\mathscr{X}$—an event. A **conditional lower prevision** $\underline{P}$ on a domain $\mathscr{C} \subseteq \mathscr{C}(\mathscr{X})$ is then a map

$$\underline{P}\colon \mathscr{C} \to \overline{\mathbb{R}}\colon (f,B) \to \underline{P}(f|B),$$

where $\overline{\mathbb{R}} := \mathbb{R} \cup \{-\infty,+\infty\}$. For any (f,B) in $\mathscr{C}$, the lower prevision $\underline{P}(f|B)$ of f conditional on B is interpreted as the **subject's supremum price** μ **for buying** f, provided that the transaction is cancelled if B does not happen. In other words, $\underline{P}(f|B)$ is the supremum value of μ for which she is eager to engage in a transaction where she receives $f(x)-\mu$ if $x \in B$ and zero otherwise. If $B=\mathscr{X}$, we write $\underline{P}(f) := \underline{P}(f|\mathscr{X})$ and call $\underline{P}(f)$ the **lower prevision** of f.

Their connection. These two uncertainty frameworks are closely connected. In particular, because of their interpretation in terms of buying prices for gambles, a conditional lower previsions can easily be derived from a set of gambles $\mathscr{D}$. For every $\mathscr{D} \subseteq \mathscr{G}(\mathscr{X})$, the **corresponding conditional lower prevision** $\underline{P}_{\mathscr{D}}$ is defined by

$$\underline{P}_{\mathscr{D}}(f|B) := \sup\{\mu \in \mathbb{R}\colon [f-\mu]\mathbb{I}_B \in \mathscr{D}\}.$$

for every $(f,B) \in \mathscr{C}(\mathscr{X})$.

Coherence. For an uncertainty model to represent a rational subject's beliefs, it needs to satisfy a set of **rationality criteria**; if it does, it is called **coherent**. For a set of desirable gambles $\mathscr{D}$, coherence means that for any gambles $f,g \in \mathscr{G}(\mathscr{X})$ and any real number $\lambda > 0$:

- **D1.** if $f \geq 0$ and $f \neq 0$, then $f \in \mathscr{D}$
- **D2.** if $f \in \mathscr{D}$ then $\lambda f \in \mathscr{D}$
- **D3.** if $f,g \in \mathscr{D}$, then $f+g \in \mathscr{D}$
- **D4.** if $f \leq 0$, then $f \notin \mathscr{D}$

A conditional lower prevision $\underline{P}$ on a domain $\mathscr{C} \subseteq \mathscr{C}(\mathscr{X})$ is then said to be coherent if there is a coherent set of desirable gambles $\mathscr{D}$ on $\mathscr{X}$ such that $\underline{P}$ coincides with $\underline{P}_{\mathscr{D}}$ on $\mathscr{C}$. Equivalently, $\underline{P}$ is coherent if it satisfies the **structure-free notion of Williams-coherence** that was developed by Pelessoni and Vicig (2009).

All of this seems very abstract. Does it have any practical use?

That's weird! Shouldn't the right-hand side be unconditional?

You said that probabilities are a special case. Yeah right... how does that work?

So why is there no B_i here?

What happens if there are more than two variables?

Modelling Independence

We say that **X_1 and X_2 are independent if our uncertainty model for X_1 is not affected by conditioning on information about X_2, and vice versa**. This definition can easily be applied to a probability measure, and then yields the usual notion of independence. More generally, it can just as easily be applied to lower previsions, sets of desirable gambles, or any other type of uncertainty model, and is then referred to as **epistemic independence**.

We consider a **very general definition** of epistemic independence. In particular, for every $i \in \{1,2\}$, we consider **any set of conditioning events** $\mathscr{B}_i$ for the variable X_i, that is, any subset of the set $\mathscr{P}_\emptyset(\mathscr{X}_i)$ of all non-empty subsets of $\mathscr{X}_i$.

A coherent **conditional lower prevision** $\underline{P}$ on $\mathscr{C}(\mathscr{X}_1 \times \mathscr{X}_2)$ is then called epistemically independent if for any i and j such that $\{i,j\}=\{1,2\}$:

$$\underline{P}(f_i|B_i \cap B_j) = \underline{P}(f_i|B_i)$$

for all $(f_i,B_i) \in \mathscr{C}(\mathscr{X}_i)$ and $B_j \in \mathscr{B}_j$.

Similarly, a coherent **set of desirable gambles** $\mathscr{D}$ on $\mathscr{X}_1 \times \mathscr{X}_2$ is epistemically independent if for any i and j such that $\{i,j\}=\{1,2\}$ and for any $B_j \in \mathscr{B}_j$:

$$\operatorname{marg}_i(\mathscr{D}\rfloor B_j) = \operatorname{marg}_i(\mathscr{D}),$$

in the sense that for all $f \in \mathscr{G}(\mathscr{X}_i)$:

$$f(X_i)\mathbb{I}_{B_j}(X_j) \in \mathscr{D} \Leftrightarrow f(X_i) \in \mathscr{D},$$

where $\mathbb{I}_{B_j}$ is the indicator of B_j, defined by $\mathbb{I}_{B_j}(x_j) := 1$ if $x_j \in B_j$ and $\mathbb{I}_{B_j}(x_j) := 0$ otherwise.

Two **special cases** are particularly important. If $\mathscr{B}_1 = \mathscr{X}_1$ and $\mathscr{B}_2 = \mathscr{X}_2$, we obtain the special case of **epistemic value-independence**, which is the most conventional approach, and which is often simply called epistemic independence. If $\mathscr{B}_1 = \mathscr{P}_\emptyset(\mathscr{X}_1)$ and $\mathscr{B}_2 = \mathscr{P}_\emptyset(\mathscr{X}_2)$, we obtain what we call **epistemic subset-independence**. As we will see, the latter has superior properties.

Independent Natural Extension

For all $i \in \{1,2\}$, let $\mathscr{D}_i$ be a local coherent **set of desirable gambles** on $\mathscr{X}_i$. The **independent natural extension** of $\mathscr{D}_1$ and $\mathscr{D}_2$ is then the smallest—most conservative—epistemically independent coherent set of desirable gambles on $\mathscr{X}_1 \times \mathscr{X}_2$ that extends them, meaning that

$$(\forall i \in \{1,2\})\ \ \mathscr{D}_i = \operatorname{marg}_i(\mathscr{D}) := \{f \in \mathscr{G}(\mathscr{X}_i)\colon f(X_i) \in \mathscr{D}\}.$$

For **lower previsions**, the local models $\underline{P}_1$ and $\underline{P}_2$ are coherent conditional lower previsions on $\mathscr{C}_1 \subseteq \mathscr{C}(\mathscr{X}_1)$ and $\mathscr{C}_2 \subseteq \mathscr{C}(\mathscr{X}_2)$, respectively. The **independent natural extension** of $\underline{P}_1$ and $\underline{P}_2$ is then the smallest—most conservative—epistemically independent coherent lower prevision on $\mathscr{C}(\mathscr{X}_1 \times \mathscr{X}_2)$ that extends them, meaning that

$$(\forall i \in \{1,2\})\ \ \underline{P}_i(f_i|B_i) = \underline{P}(f_i|B_i) \text{ for all } (f_i,B_i) \in \mathscr{C}_i.$$

Existence. In both of our two frameworks, **the independent natural extension always exists**; we denote it by $\mathscr{D}_1 \otimes \mathscr{D}_2$ and $\underline{P}_1 \otimes \underline{P}_2$, respectively. For lower previsions, this result crucially depends on our use of Williams-coherence: for Walley-coherence, as shown by Miranda and Zaffalon (2015) for epistemic value-independence, this may no longer hold.

Properties. Let $\{i,j\}=\{1,2\}$ and consider any $h \in \mathscr{G}(\mathscr{X}_j)$ and $f,g \in \mathscr{G}(\mathscr{X}_i)$ such that $g \geq 0$ is **$\mathscr{B}_i$-measurable**—a technical condition that coincides with the usual notion when $\mathscr{B}_i \cup \{\emptyset\}$ is a σ-field. Then if all the terms are well-defined—if $\mathscr{C}_1$ and $\mathscr{C}_2$ are large enough—we have that

$$(\underline{P}_1 \otimes \underline{P}_2)(f+gh) = \underline{P}_i(f + g\underline{P}_j(h)).$$

As a direct consequence, we find that

$$(\underline{P}_1 \otimes \underline{P}_2)(f+h) = \underline{P}_i(f) + \underline{P}_j(h).$$

and—with $\overline{P}_i(g) := -\underline{P}_i(-g)$—that

$$(\underline{P}_1 \otimes \underline{P}_2)(gh) = \underline{P}_i(g\underline{P}_j(h)) = \begin{cases} \underline{P}_i(g)\underline{P}_j(h) & \text{if } \underline{P}_j(h) \geq 0; \\ \overline{P}_i(g)\underline{P}_j(h) & \text{if } \underline{P}_j(h) \leq 0, \end{cases}$$

known as **external additivity** and **factorisation**, respectively. Crucially, **for epistemic subset-independence**, $\mathscr{B}_i$-measurability is trivially satisfied, and **factorisation** then **always holds**.