UNIVERSITEIT GENT

FACULTY OF ENGINEERING

Structural judgements

Gert de Cooman

SYSTeMS Research Group
Ghent University
gert.decooman@UGent.be
http://users.UGent.be/~gdcooma

Fourth SIPTA School on Imprecise Probabilities
Durham, UK, 2 September 2010

Outline

Outline

General comments

Probability measure versus expectation functional

Model for a random variable X assuming values in $\mathscr{X}$:

probability $P(X \in A)$ for all events $A \subseteq \mathscr{X}$

expectation $E(f(X))$ for all gambles $f\colon \mathscr{X} \to \mathbb{R}$

probability $P(\cdot)$ and expectation $E(\cdot)$ are equally expressive

General comments

Probability measure versus expectation functional

Model for a random variable X assuming values in $\mathscr{X}$:

probability $P(X \in A)$ for all events $A \subseteq \mathscr{X}$

expectation $E(f(X))$ for all gambles $f\colon \mathscr{X} \to \mathbb{R}$

lower probability $\underline{P}(\cdot)$ is less expressive than lower expectation $\underline{E}(\cdot)$

General comments

Probability measure versus expectation functional

Model for a random variable X assuming values in $\mathscr{X}$:

probability $P(X \in A)$ for all events $A \subseteq \mathscr{X}$

expectation $E(f(X))$ for all gambles $f\colon \mathscr{X} \to \mathbb{R}$

lower probability $\underline{P}(\cdot)$ is less expressive than lower expectation $\underline{E}(\cdot)$

WHEN WORKING WITH IMPRECISE PROBABILITIES,

USE (LOWER) EXPECTATIONS AND GAMBLES

General comments

Lower expectation functional versus set of desirable gambles

Two types of imprecise-probability models:

lower expectation $\underline{P}(f(X))$ for all gambles $f\colon \mathscr{X} \to \mathbb{R}$

set of desirable gambles $\mathscr{D} \subseteq \mathscr{L}(\mathscr{X})$

with

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

General comments

Lower expectation functional versus set of desirable gambles

Two types of imprecise-probability models:

lower expectation $\underline{P}(f(X))$ for all gambles $f\colon \mathscr{X} \to \mathbb{R}$

set of desirable gambles $\mathscr{D} \subseteq \mathscr{L}(\mathscr{X})$

with

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

Working with sets of desirable gambles $\mathscr{D}$:

- is simpler, more intuitive and more elegant
- is more general and expressive
- gives a geometrical flavour to probabilistic inference
- shows that probabilistic inference is ‘logical’ inference

General comments

All the most interesting and practical aspects of probabilistic reasoning are covered by (derivable from):

D1. if $f < 0$ then $f \notin \mathscr{D}$

D2. if $f > 0$ then $f \in \mathscr{D}$

D3. if $f, g \in \mathscr{D}$ then $f + g \in \mathscr{D}$

D4. if $f \in \mathscr{D}$ then $\lambda f \in \mathscr{D}$ for all real $\lambda > 0$

Precise models correspond to the special case that the cones $\mathscr{D}$ are actually semi-spaces!

General comments

A coherent lower prevision $\underline{P}$ on $\mathscr{L}(\mathscr{X})$ has the following interesting properties:

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

(ii) $\underline{P}(\lambda f) = \lambda \underline{P}(f)$ if $\lambda \geq 0$ and $\underline{P}(\lambda f) = \lambda \overline{P}(f)$ if $\lambda \leq 0$

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

(iv) if $f \leq g$ then $\underline{P}(f) \leq \underline{P}(g)$ and $\overline{P}(f) \leq \overline{P}(g)$

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

Exercise on coherence

Problem 1

Consider a space with two elements: $\mathscr{X} = \{a, b\}$.

1. Show that any linear prevision on this space can be written as

$$P_\alpha(f) = \alpha f(a) + (1-\alpha) f(b)$$

   for some $\alpha \in [0,1]$. Actually $\alpha = P_\alpha(\{a\})$ and $1-\alpha = P_\alpha(\{b\})$.

2. Show that for any gamble f:

$$f = f(a) + [f(b) - f(a)] I_{\{b\}} = f(b) + [f(a) - f(b)] I_{\{a\}}$$

3. Show that all coherent lower previsions $\underline{P}$ on $\mathscr{L}(\mathscr{X})$ are linear-vacuous mixtures: there are α and ε in $[0,1]$ such that

$$\begin{aligned}\underline{P}(f) &= \varepsilon P_\alpha(f) + (1-\varepsilon)\min f \\ &= \varepsilon[\alpha f(a) + (1-\alpha) f(b)] + (1-\varepsilon)\min\{f(a), f(b)\}.\end{aligned}$$

   [Hint: Let $\underline{P}(\{a\}) = \varepsilon\alpha$ and $\underline{P}(\{b\}) = \varepsilon(1-\alpha)$].

Outline

Structural assessments

Local versus structural assessments

Local assessments

A subject gives values for

$$\underline{P}(f) \text{ for all gambles } f \text{ in some subset } \mathscr{A} \text{ of } \mathscr{L}(\mathscr{X})$$

Structural assessments

The model $\underline{P}$ satisfies some properties besides coherence:

1. behaves in a certain way under some transformation or operation:
   - irrelevant or independent models
   - symmetrical models
2. zero on some structurally important set of gambles

Outline

Irrelevance and independence

The basic set-up

Consider variables X_1 in $\mathscr{X}_1$ and X_2 in $\mathscr{X}_2$.

Marginal models

We have:

- a coherent model $\mathscr{D}_1$ for X_1, which is a subset of $\mathscr{L}(\mathscr{X}_1)$ and
- a coherent model $\mathscr{D}_2$ for X_2, which is a subset of $\mathscr{L}(\mathscr{X}_2)$.

Irrelevance and independence

The basic set-up

Consider variables X_1 in $\mathscr{X}_1$ and X_2 in $\mathscr{X}_2$.

Marginal models

We have:

- a coherent model $\mathscr{D}_1$ for X_1, which is a subset of $\mathscr{L}(\mathscr{X}_1)$ and
- a coherent model $\mathscr{D}_2$ for X_2, which is a subset of $\mathscr{L}(\mathscr{X}_2)$.

Joint model

We want to combine $\mathscr{D}_1$ and $\mathscr{D}_2$ into a joint model $\mathscr{D}$ which is a subset of $\mathscr{L}(\mathscr{X}_1 \times \mathscr{X}_2)$:

$$\operatorname{marg}_1(\mathscr{D}) := \mathscr{D} \cap \mathscr{L}(\mathscr{X}_1) = \mathscr{D}_1$$
$$\operatorname{marg}_2(\mathscr{D}) := \mathscr{D} \cap \mathscr{L}(\mathscr{X}_2) = \mathscr{D}_2$$

Irrelevance and independence

Natural extension

What this joint model looks like, will depend on what we know about the relation between X and Y.

No information specified

In this case the smallest joint is given by

$$\mathscr{D}_1 \boxtimes \mathscr{D}_2 := \operatorname{posi}\left(\mathscr{L}^+(\mathscr{X}_1 \times \mathscr{X}_2) \cup \mathscr{D}_1 \cup \mathscr{D}_2\right)$$

or in terms of lower previsions

$$\begin{aligned}&\underline{P}_1 \boxtimes \underline{P}_2(f)\\ := &\sup_{\substack{h_1 \in \mathscr{L}(\mathscr{X}_1)\\ h_2 \in \mathscr{L}(\mathscr{X}_2)}} \inf_{\substack{x_1 \in \mathscr{X}_1\\ x_2 \in \mathscr{X}_2}} [f(x_1,x_2) - [h_1(x_1) - \underline{P}_1(h_1)] - [h_2(x_2) - \underline{P}_2(h_2)]].\end{aligned}$$

Irrelevance and independence

Conditional models

Suppose we have a joint model $\mathscr{D} \subseteq \mathscr{L}(\mathscr{X}_1 \times \mathscr{X}_2)$ for (X_1, X_2), then we can derive the conditional models

$$\mathscr{D}\rfloor x_1 := \left\{h_2 \in \mathscr{L}(\mathscr{X}_2) \colon I_{\{x_1\}} h_2 \in \mathscr{D}\right\}, \quad x_1 \in \mathscr{X}_1$$

and

$$\mathscr{D}\rfloor x_2 := \left\{h_1 \in \mathscr{L}(\mathscr{X}_1) \colon h_1 I_{\{x_2\}} \in \mathscr{D}\right\}, \quad x_2 \in \mathscr{X}_2$$

or in terms of lower previsions:

$$\begin{aligned} \underline{P}_2(h_2|x_1) &= \sup\left\{\mu \in \mathbb{R} \colon h_2 - \mu \in \mathscr{D}\rfloor x_1\right\} \\ &= \sup\left\{\mu \in \mathbb{R} \colon I_{\{x_1\}}[h_2 - \mu] \in \mathscr{D}\right\} \end{aligned}$$

Definition

We say that X_1 is epistemically irrelevant to X_2 when learning the value $X_1 = x_1$ of X_1 does not affect our beliefs about X_2.

For a joint model $\mathscr{D}$ to express this:

$$\mathscr{D}\rfloor x_1 = \operatorname{marg}_2(\mathscr{D}) \text{ for all } x_1 \in \mathscr{X}_1$$

An irrelevant joint $\mathscr{D}$ of marginal models $\mathscr{D}_1$ and $\mathscr{D}_2$ satisfies the following structural judgements:

$$\begin{aligned}\operatorname{marg}_1(\mathscr{D}) &= \mathscr{D}_1\\ \operatorname{marg}_2(\mathscr{D}) &= \mathscr{D}_2\\ \mathscr{D}\rfloor x_1 = \operatorname{marg}_2(\mathscr{D}) &= \mathscr{D}_2 \text{ for all } x_1 \in \mathscr{X}_1.\end{aligned}$$

X_1 is epistemically irrelevant to X_2

In this case the smallest joint is given by

$$\mathscr{D}_1 \times_{1\to 2} \mathscr{D}_2 := \operatorname{posi}\left(\mathscr{L}^+(\mathscr{X}_1 \times \mathscr{X}_2) \cup \mathscr{D}_1 \cup \mathscr{A}_{1\to 2}\right)$$

$$\mathscr{A}_{1\to 2} := \operatorname{posi}\left(\left\{I_{\{x_1\}}h_2 \colon x_1 \in \mathscr{X}_1 \text{ and } h_2 \in \mathscr{D}_2\right\}\right)$$

$$= \left\{h \in \mathscr{L}_0(\mathscr{X}_1 \times \mathscr{X}_2) \colon (\forall x_1 \in \mathscr{X}_1) h(x_1,\cdot) \in \mathscr{D}_2\right\}$$

$$\underline{P}_1 \times_{1\to 2} \underline{P}_2(f)$$

$$:= \sup_{\substack{h_1 \in \mathscr{L}(\mathscr{X}_1)\\ h_2 \in \mathscr{L}(\mathscr{X}_1 \times \mathscr{X}_2)}} \inf_{\substack{x_1 \in \mathscr{X}_1\\ x_2 \in \mathscr{X}_2}} \Big[f(x_1,x_2) - [h_1(x_1) - \underline{P}_1(h_1)] - [h_2(x_1,x_2) - \underline{P}_2(h_2(x_1,\cdot))] \Big].$$

Exercise on irrelevance

Problem 2

It can be shown that

$$\underline{P}_1 \times_{1\to 2} \underline{P}_2(f) = \underline{P}_1(\underline{P}_2(f))$$

where $\underline{P}_2(f)$ is defined the gamble on $\mathscr{X}_1$ that assumes the value $\underline{P}_2(f(x_1,\cdot))$ in x_1.

1. Show by means of a counterexample that not necessarily $\underline{P}_1 \times_{1\to 2} \underline{P}_2(f) = \underline{P}_1 \times_{2\to 1} \underline{P}_2(f)$, or in other words not necessarily $\underline{P}_1(\underline{P}_2(f)) = \underline{P}_2(\underline{P}_1(f))$.

Hint: use the simplest possible case for $\mathscr{X} = \{a,b\}$ and remember Problem 1.

Definition

We say that X_1 and X_2 are epistemically independent when X_1 is epistemically irrelevant to X_2 and X_2 is epistemically irrelevant to X_1.

An independent joint $\mathscr{D}$ of marginal models $\mathscr{D}_1$ and $\mathscr{D}_2$ satisfies the following structural judgements:

$$\operatorname{marg}_1(\mathscr{D}) = \mathscr{D}_1$$
$$\operatorname{marg}_2(\mathscr{D}) = \mathscr{D}_2$$
$$\mathscr{D}\rfloor x_1 = \operatorname{marg}_2(\mathscr{D}) = \mathscr{D}_2 \text{ for all } x_1 \in \mathscr{X}_1$$
$$\mathscr{D}\rfloor x_2 = \operatorname{marg}_1(\mathscr{D}) = \mathscr{D}_1 \text{ for all } x_2 \in \mathscr{X}_2.$$

Irrelevance and independence

Independent natural extension

X_1 and X_2 are epistemically independent

In this case the smallest joint is given by

$$\mathscr{D}_1 \times \mathscr{D}_2 := \operatorname{posi}\left(\mathscr{L}^+(\mathscr{X}_1 \times \mathscr{X}_2) \cup \mathscr{A}_{2\to1} \cup \mathscr{A}_{1\to2}\right)$$

$$\mathscr{A}_{1\to2} = \{h \in \mathscr{L}_0(\mathscr{X}_1 \times \mathscr{X}_2) \colon (\forall x_1 \in \mathscr{X}_1) h(x_1,\cdot) \in \mathscr{D}_2\}$$

$$\mathscr{A}_{2\to1} = \{h \in \mathscr{L}_0(\mathscr{X}_1 \times \mathscr{X}_2) \colon (\forall x_2 \in \mathscr{X}_2) h(\cdot,x_2) \in \mathscr{D}_1\}$$

$$\underline{P}_1 \times \underline{P}_2(f)$$

$$:= \sup_{\substack{h_1 \in \mathscr{L}(\mathscr{X}_1 \times \mathscr{X}_2)\\ h_2 \in \mathscr{L}(\mathscr{X}_1 \times \mathscr{X}_2)}} \inf_{\substack{x_1 \in \mathscr{X}_1\\ x_2 \in \mathscr{X}_2}} \Big[f(x_1,x_2) - [h_1(x_1,x_2) - \underline{P}_1(h_1(\cdot,x_2))] - [h_2(x_1,x_2) - \underline{P}_2(h_2(x_1,\cdot))] \Big].$$

Irrelevance and independence

Precise products

Generally speaking, independent joints are not unique, and

$$\underline{P}_1 \times_{1\to 2} \underline{P}_2 \neq \underline{P}_1 \times_{2\to 1} \underline{P}_2$$
$$\underline{P}_1 \times_{1\to 2} \underline{P}_2 < \underline{P}_1 \times \underline{P}_2$$
$$\underline{P}_1 \times_{2\to 1} \underline{P}_2 < \underline{P}_1 \times \underline{P}_2$$

When $\underline{P}_1 = P_1$ and $\underline{P}_2 = P_2$ are precise models:

$$P_1 \times_{1\to 2} P_2 = P_1 \times_{2\to 1} P_2 = P_1 \times P_2$$

- is the only independent joint, and
- coincides with the usual independent product of probability measures.

Another independent joint of $\underline{P}_1$ and $\underline{P}_2$ is generally given by the

Strong product $\underline{P}_1 \otimes \underline{P}_2$

$$\underline{P}_1 \otimes \underline{P}_2(f) := \min\{P_1 \times P_2(f) \colon P_1 \in \mathscr{M}_1 \text{ and } P_2 \in \mathscr{M}_2\}$$

Generally speaking

$$\underline{P}_1 \times \underline{P}_2 < \underline{P}_1 \otimes \underline{P}_2.$$

Irrelevance and independence

Factorisation and external additivity

For coherent lower prevision $\underline{P}$ such that $\underline{P}_1 \times \underline{P}_2 \leq \underline{P} \leq \underline{P}_1 \otimes \underline{P}_2$:

1. $\underline{P}$ is a coherent joint of the marginals $\underline{P}_1$ and $\underline{P}_2$
2. $\underline{P}$ is factorising: for all $f_1 \in \mathscr{L}(\mathscr{X}_1)$ and all non-negative $f_2 \in \mathscr{L}(\mathscr{X}_2)$,

$$\underline{P}(f_1 f_2) = \underline{P}_2(f_2 \underline{P}_1(f_1)) = \begin{cases} \underline{P}_1(f_1)\underline{P}_2(f_2) & \text{if } \underline{P}_1(f_1) \geq 0 \\ \underline{P}_1(f_1)\overline{P}_2(f_2) & \text{if } \underline{P}_1(f_1) \leq 0 \end{cases}$$

3. $\underline{P}$ is externally additive: for all $f_1 \in \mathscr{L}(\mathscr{X}_1)$ and all $f_2 \in \mathscr{L}(\mathscr{X}_2)$,

$$\underline{P}(f_1 + f_2) = \underline{P}_1(f_1) + \underline{P}_2(f_2).$$

Exercise on independence

Problem 2

Consider belief functions $\underline{P}_1$ on $\mathscr{L}(\mathscr{X}_1)$ and $\mathscr{L}(\mathscr{X}_2)$, given by

$$\underline{P}_1(f_1) = \sum_{k=1}^{n} m_1(F_k) \min_{x_1 \in F_k} f_1(x_1)$$

$$\underline{P}_2(f_2) = \sum_{\ell=1}^{n} m_2(G_\ell) \min_{x_2 \in G_\ell} f_2(x_2)$$

and their Dempster product $\underline{P}_1 \times_D \underline{P}_2$ given by

$$\underline{P}_1 \times_D \underline{P}_2(f) = \sum_{k=1}^{n} \sum_{\ell=1}^{n} m_1(F_k) m_2(G_\ell) \min_{x_1 \in F_k} \min_{x_2 \in G_\ell} f(x_1, x_2).$$

1. Show that generally $\underline{P}_1 \times_D \underline{P}_2 \leq \underline{P}_1 \times \underline{P}_2$, so this Dempster product is generally incoherent (too conservative).

Hint: Show that $\underline{P}_1 \times_D \underline{P}_2 \leq \underline{P}_1 \times_{1\to 2} \underline{P}_2$.

Irrelevance and independence

Further references

More information can be found in:

- Gert de Cooman and Enrique Miranda.
  Irrelevance and independence for sets of desirable gambles.
  2010. In preparation.

- Gert de Cooman, Enrique Miranda, and Marco Zaffalon.
  Independent natural extension.
  2010. In preparation.

- Serafín Moral.
  Epistemic irrelevance on sets of desirable gambles.
  Annals of Mathematics and Artificial Intelligence, 45:197–214, 2005.

- Peter Walley.
  Statistical Reasoning with Imprecise Probabilities.
  Chapman and Hall, London, 1991. Chapter 9.

Outline

An example

Flipping a coin

I am going to flip a coin in the next room. *How do you model your information (beliefs) about the outcome?*

Situation A You have seen and examined the coin, and you believe it is symmetrical (not biased).

Situation B You have no information about the coin, it may be heavily loaded, it may even have two heads or two tails.

Evidence of symmetry

In Situation A, there is information that the phenomenon described is invariant under permutation of heads and tails.

Symmetry of evidence

In Situation B, your information (none) is invariant under permutation of heads and tails.

Modelling the available information

- We want a model for the available information or evidence: a belief model.
  - In Situation A, the belief model should reflect that there is evidence of symmetry.
  - In Situation B, the evidence is invariant under permutations of heads and tails, so the belief model should be invariant as well.
- Since the available information is different in both situations, the corresponding belief models should be different too!
- Belief models should be able to capture the difference between 'symmetry of evidence' and 'evidence of symmetry'.
- This is not the case for Bayesian probability models.

What are we going to do?

- Explain how to model aspects of symmetry for such coherent lower previsions
  - symmetry of evidence,
  - evidence of symmetry.
- Argue that both aspects are different in general, but coincide for precise belief models.

Being able to deal with natural symmetries is often quite useful in applications, and is of fundamental theoretical importance.

Monoids of transformations

In mathematics (geometry, topology, linear algebra)

symmetry

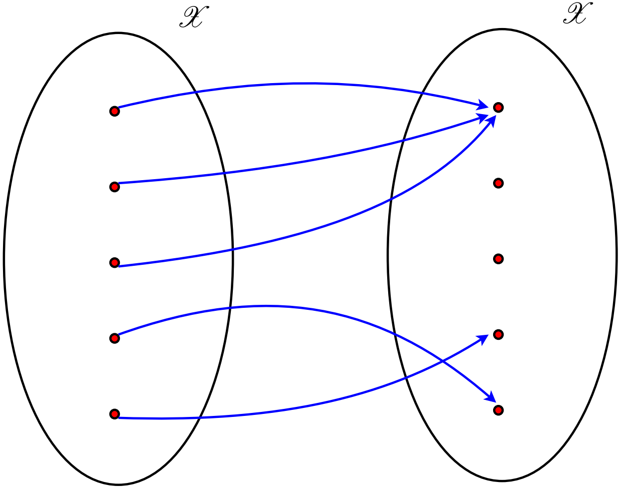
is considered to be

invariance under certain transformations.

Monoids of transformations

Transformations and permutations

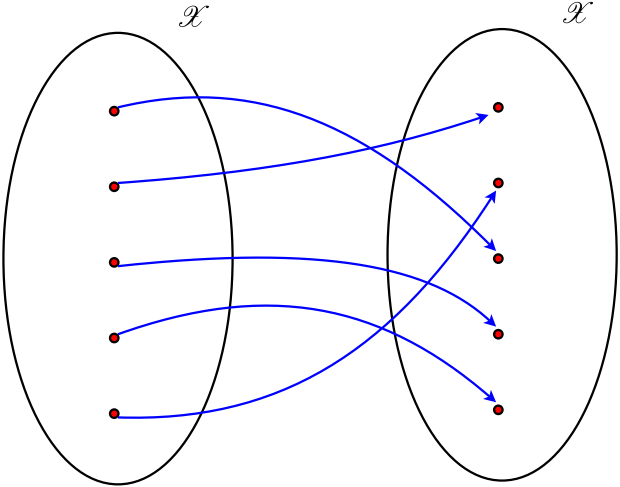
A transformation T of $\mathscr{X}$ is a map from $\mathscr{X}$ to itself, i.e., $T\colon \mathscr{X} \to \mathscr{X} : x \mapsto Tx$.



Monoids of transformations

Transformations and permutations

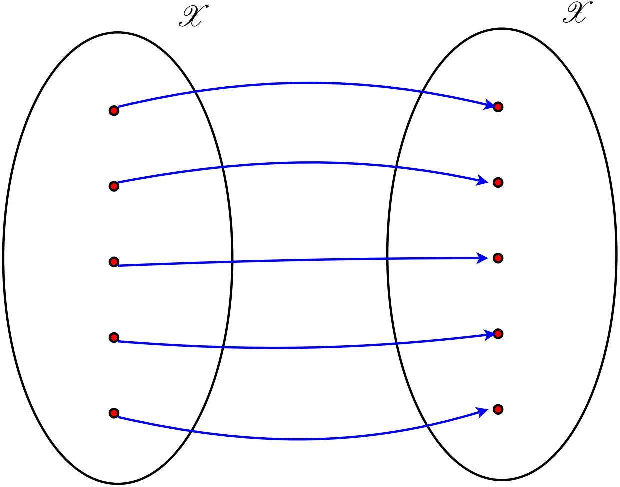
A permutation π of $\mathscr{X}$ is a transformation of $\mathscr{X}$ that is onto and one-to-one.



Examples of transformations

Identity map

The identity map $\mathrm{id}_{\mathscr{X}}$, defined by $\mathrm{id}_{\mathscr{X}} x = x$, is a permutation.



Monoids of transformations

Transformations and permutations

Consider a monoid $\mathscr{T}$ of transformations T of $\mathscr{X}$ (not necessarily permutations), i.e.,

- $\mathrm{id}_{\mathscr{X}}$ belongs to $\mathscr{T}$;
- if T and S both belong to $\mathscr{T}$ then so does $TS := T \circ S$.

Symmetry is usually expressed as invariance with respect to every transformation T in a some relevant monoid $\mathscr{T}$.

Transformations T act on elements x of $\mathscr{X}$, but we are also interested in the corresponding transformations T that act on gambles f on $\mathscr{X}$.

Lifting T to gambles

For any gamble f, define the new gamble $T^t f := f \circ T$ by lifting:

$$(T^t f)(x) := f(Tx).$$

Lifting T to lower previsions

For any lower prevision $\underline{P}$: $\mathscr{L}(\mathscr{X})$, define the new functional $T\underline{P} := \underline{P} \circ T^t$ by lifting again:

$$(T\underline{P})(f) := \underline{P}(T^t f) = \underline{P}(f \circ T).$$

Exercises on symmetry

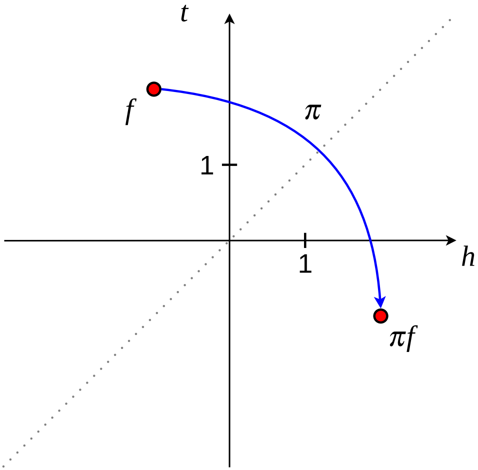
Problem 4

$\mathscr{X} = \{h, t\}$ and π is the permutation of $\mathscr{X}$ such that $\pi(h) = t$ and $\pi(t) = h$. Consider the gamble $f(h) = -1$ and $f(t) = 2$.

1. What is $\pi^t f$?
2. If P is the uniform probability on $\mathscr{X}$, then what is $\pi P(f)$?

Exercises on symmetry

Solution to Problem 4



Weak invariance of belief models

Definition

A coherent belief model is called weakly $\mathscr{T}$-invariant if the following equivalent conditions are satisfied:

W1. $T^t\mathscr{D} \subseteq \mathscr{D}$ for all $T \in \mathscr{T}$;

W2. $T\underline{P} \geq \underline{P}$ for all $T \in \mathscr{T}$;

W3. $T\mathscr{M} \subseteq \mathscr{M}$ for all $T \in \mathscr{T}$.

- A precise prevision is weakly $\mathscr{T}$-invariant iff $TP = P$, or equivalently

$$P(A) = P(T^{-1}(A))$$

for all $A \subseteq \mathscr{X}$ and all T in $\mathscr{T}$. This is the usual definition for invariance of a (probability) measure.

Weak invariance of belief models

Observations

Symmetry of evidence

Weak invariance states that belief models are symmetrical.

Existence guaranteed

There are weakly $\mathscr{T}$-invariant coherent models for any monoid $\mathscr{T}$.

The vacuous lower prevision

$$\mathscr{D}_v = \{f : f > 0\}$$

$$\underline{P}_v(f) = \inf_{x \in \mathscr{X}} f(x)$$

$$\mathscr{M}_v = \text{ the set of all precise previsions}$$

is the only coherent belief model that is weakly invariant with respect to all transformations of $\mathscr{X}$. It models complete ignorance.

Strong invariance of belief models

Definition

Evidence of symmetry

How can we model that we believe there is symmetry, characterised by a monoid $\mathscr{T}$, behind the random variable X?

Consider a gamble f and its transform $T^t f$. Because of the symmetry, you should be willing to exchange f for $T^t f$ and vice versa:

$$f - T^t f + \delta \in \mathscr{D} \quad \text{for all } \delta > 0.$$

Definition

A coherent belief model is called strongly $\mathscr{T}$-invariant if the following equivalent conditions are satisfied:

S1. $\underline{P}(f - T^t f) \geq 0$ for all $T \in \mathscr{T}$ and $f \in \mathscr{L}(\mathscr{X})$;

S2. $\underline{P}(f - T^t f) = \overline{P}(f - T^t f) = 0$ for all $T \in \mathscr{T}$ and $f \in \mathscr{L}(\mathscr{X})$;

S3. All precise previsions in $\mathscr{M}$ are (weakly) $\mathscr{T}$-invariant.

Exercises on symmetry

Problem 5

Consider $\mathscr{X} = \{h, t\}$ and the monoid (group) $\mathscr{T} = \{\mathrm{id}_{\mathscr{X}}, \pi\}$.
Observe that

$$\begin{aligned}\{f - \pi^t f : f \in \mathscr{L}(\mathscr{X})\} &= \{(f(h) - \pi^t f(h), f(t) - \pi^t f(t)) : f \in \mathscr{L}(\mathscr{X})\} \\ &= \{(f(h) - f(t), f(t) - f(h)) : f \in \mathscr{L}(\mathscr{X})\} \\ &= \{(x, -x) : x \in \mathbb{R}\}\end{aligned}$$

1. The only strongly permutation invariant belief model is the uniform precise model that assigns probability 1/2 to both h and t.

Hint: Use the solution to Problem 1.

Strong invariance of belief models

Observations

Evidence of symmetry versus symmetry of evidence

strong invariance captures 'evidence of symmetry'.

weak invariance captures 'symmetry of evidence'.

Strong invariance implies weak invariance:

$$0 = \underline{P}(T^t f - f) \leq \underline{P}(T^t f) - \underline{P}(f).$$

For precise previsions, strong and weak invariance coincide:

$$0 = P(f - T^t f) = P(f) - P(T^t f).$$

Bayesian models cannot distinguish between 'evidence of symmetry' and 'symmetry of evidence'.

Strong permutation invariance

A special case

Let $\mathscr{X}$ be a finite set and let $\mathscr{P}$ be a (finite) group of permutations π of $\mathscr{X}$, i.e. a monoid such that

- for all π in $\mathscr{P}$ there is some inverse $\varpi \in \mathscr{P}$ such that $\pi \circ \varpi = \varpi \circ \pi = \mathrm{id}_{\mathscr{X}}$.

An event $A \subseteq \mathscr{X}$ is $\mathscr{P}$-invariant if

$$\pi A = \{\pi x \colon x \in A\} = A \text{ for all } \pi \text{ in } \mathscr{P}.$$

Fact

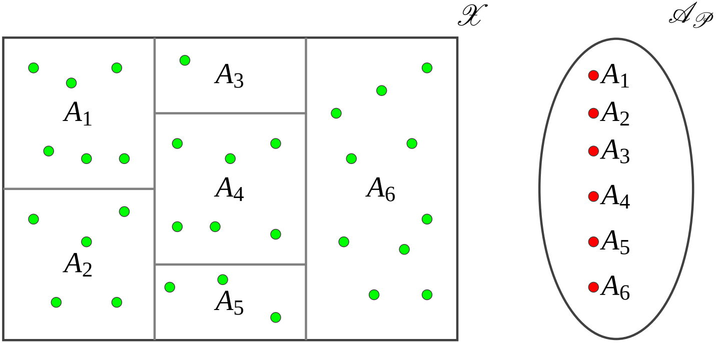
The smallest $\mathscr{P}$-invariant sets (atoms) constitute a partition of $\mathscr{X}$:

$$[x]_{\mathscr{P}} := \{\pi x \colon \pi \in \mathscr{P}\},$$

and $\mathscr{A}_{\mathscr{P}} := \{[x]_{\mathscr{P}} : x \in \mathscr{X}\}$ is the set of all $\mathscr{P}$-invariant atoms.

Strong permutation invariance

Invariant atoms



Strong permutation invariance

Fundamental theorem

Theorem

A coherent lower prevision $\underline{P}$ on $\mathscr{L}(\mathscr{X})$ is strongly $\mathscr{P}$-invariant if and only if it has the following form:

$$\underline{P}(f) = \underline{Q}(P^u(f|\cdot))$$

where $\underline{Q}$ is any coherent lower prevision on $\mathscr{L}(\mathscr{A}_{\mathscr{P}})$.

- $P^u(f|\cdot)$ is a gamble on $\mathscr{A}_{\mathscr{P}}$, whose value in any invariant atom $A \in \mathscr{A}_{\mathscr{P}}$ is given by

$$P^u(f|A) = \frac{1}{|A|} \sum_{x \in A} f(x),$$

so $P^u(\cdot|A)$ is the precise prevision whose probability mass is distributed uniformly over the atom A.

Exercises on symmetry

Problem 6

Consider a space with two elements: $\mathscr{X} = \{a, b\}$ and the set $\mathscr{P}$ of all permutations of $\mathscr{X}$.

1. What are the elements of $\mathscr{P}$?
2. What are the invariant atoms?
3. Show that all weakly $\mathscr{P}$-invariant coherent lower previsions $\underline{P}$ on $\mathscr{L}(\mathscr{X})$ are given by

$$\begin{aligned}\underline{P}(f) &= \varepsilon P_{\frac{1}{2}}(f) + (1-\varepsilon)\min f \\ &= \varepsilon \frac{f(a)+f(b)}{2} + (1-\varepsilon)\min\{f(a), f(b)\}.\end{aligned}$$

   for some ε in $[0, 1]$.
4. Use the Fundamental Theorem on Strong Permutation Invariance to show (once again) that $P_{\frac{1}{2}}$ is the only strongly $\mathscr{P}$-invariant coherent lower prevision on $\mathscr{L}(\mathscr{X})$.

Exercises on symmetry
Problem 7

Consider casting a die: $\mathscr{X} = \{1,2,3,4,5,6\}$ and suppose there is evidence of symmetry between all even outcomes, and between all odd outcomes: you have reason not to distinguish between 2, 4 and 6 on the one hand, and 1, 3 and 5 on the other hand. In other words, the invariant atoms are $\{1,3,5\}$ and $\{2,4,6\}$.

1. Characterise all the strongly invariant coherent lower previsions for this type of symmetry.
2. Characterise all the strongly invariant precise previsions for this type of symmetry.

[Hint: use the results of Problem 1, and the Fundamental Theorem on Strong Permutation Invariance]

Strong permutation invariance

More information

More information about strong invariance, with ergodicity theorems and the special case of exchangeability can be found in:

- Gert de Cooman and Enrique Miranda.
  Symmetry of models versus models of symmetry.
  In W. L. Harper and G. R. Wheeler, editors, *Probability and Inference: Essays in Honor of Henry E. Kyburg, Jr.*, pages 67–149. King's College Publications, 2007.
- Gert de Cooman and Erik Quaeghebeur.
  Exchangeability and sets of desirable gambles.
  International Journal of Approximate Reasoning, 2010. Submitted for publication. Special issue in honour of Henry E. Kyburg, Jr.
- Gert de Cooman, Erik Quaeghebeur, and Enrique Miranda.
  Exchangeable lower previsions.
  Bernoulli, 15(3):721–735, 2009.