Introduction to imprecise probabilities

SIPTA School 2024, Ghent, Belgium

Erik Quaeghebeur

Eindhoven University of Technology

12 August 2024

Overview

Kick-off (slot 1)

Classical probability theory (slot 1)

Interpretation of probability (slot 2)

Limitations of probability theory (slot 2)

Probability intervals (slot 3)

Credal sets (slot 3–4)

Acceptability & Desirability (slot 4–5)

Interval expectation & probability (slot 5–6)

Overview

Kick-off (slot 1)

- Uncertainty, representation & reasoning
- Illustration: Blood groups

Classical probability theory (slot 1)

Interpretation of probability (slot 2)

Limitations of probability theory (slot 2)

Probability intervals (slot 3)

Credal sets (slot 3–4)

Acceptability & Desirability (slot 4–5)

Interval expectation & probability (slot 5–6)

What is *uncertainty*?

There is no consensus definition of uncertainty...

> Uncertainty is any departure
> from the unachievable ideal of complete determinism.
>
> Walker et al. (2003)

> Uncertainty is the lack of certainty, a state of limited knowledge where it is impossible to exactly describe the existing state, a future outcome, or more than one possible outcome.
>
> Hubbard (2014)

How can we *represent* uncertainty?

Uncertainty representations are defined by axioms

> Any reasonable measure of belief is isomorphic to a probability distribution.
>
> Cox (1946), paraphrased

Uncertainty representations are defined by axioms

> Any reasonable measure of belief is isomorphic to a probability distribution.
>
> Cox (1946), paraphrased

Some concrete examples of representations

- probability distributions
- intervals, sets
- possibility distributions, belief functions
- sets of desirable gambles, preference orders
- lower/upper probabilities/expectations
- credal sets
- choice functions

What is *reasoning* under uncertainty?

Reasoning under uncertainty: deducing inferences or making decisions

> In each problem situation, uncertainty is associated with some purpose.
> It may, for example, be associated with prediction, retrodiction, prescription, or decision making.
>
> Klir (1995)

Reasoning under uncertainty: deducing inferences or making decisions

> In each problem situation, uncertainty is associated with some purpose.
> It may, for example, be associated with prediction, retrodiction, prescription, or decision making.
>
> Klir (1995)

Some concrete examples

- Predicting individual disease risk based on medical history
- Planning and designing of blood bank for a hospital with 100 beds
- A newspaper vendor must decide how many copies to purchase each day

Blood groups problem with expensive test

Problem setup

- A sample:

A	AB	O	B
A	B	A	O

- A disease

- Two treatments with differing effectiveness:

	A	B	AB	O
f_1	0.5	0.6	0.7	0.1
f_2	0.4	0.3	0.3	0.8

Blood groups problem with expensive test

Problem setup

- A sample:

A	AB	O	B
A	B	A	O

- A disease

- Two treatments with differing effectiveness:

	A	B	AB	O
f_1	0.5	0.6	0.7	0.1
f_2	0.4	0.3	0.3	0.8

Representation

Blood groups problem with expensive test

Problem setup

- A sample:

A	AB	O	B
A	B	A	O

- A disease

- Two treatments with differing effectiveness:

	A	B	AB	O
f_1	0.5	0.6	0.7	0.1
f_2	0.4	0.3	0.3	0.8

Representation

PMF from observed frequencies:

p_A	p_B	p_{AB}	p_O
$3/8$	$2/8$	$1/8$	$2/8$

Blood groups problem with expensive test

Problem setup

- A sample:

A	AB	O	B
A	B	A	O

- A disease
- Two treatments with differing effectiveness:

	A	B	AB	O
f_1	0.5	0.6	0.7	0.1
f_2	0.4	0.3	0.3	0.8

Representation

PMF from observed frequencies:

p_{A}	p_{B}	p_{AB}	p_{O}
$3/8$	$2/8$	$1/8$	$2/8$

Reasoning

- Probability $P(\{\text{A}, \text{O}\})$?
- Expectation $E_p(f_1)$?
- Outcome with maximal probability?
- Treatment with highest expected effectiveness?

Blood groups problem with expensive test

Problem setup

- A sample:

A	AB	O	B
A	B	A	O

- A disease

- Two treatments with differing effectiveness:

	A	B	AB	O
f_1	0.5	0.6	0.7	0.1
f_2	0.4	0.3	0.3	0.8

Representation

PMF from observed frequencies:

p_{A}	p_{B}	p_{AB}	p_{O}
$3/8$	$2/8$	$1/8$	$2/8$

Reasoning

- $P(\{\mathrm{A}, \mathrm{O}\}) = 5/8$
- $E_p(f_1) = 0.45$ (*check!*)
- Outcome A has maximal probability
- Treatment 2 has highest expected effectiveness (*check!*)

Blood groups problem with cheap test

Problem setup

- A partial information sample:

A,B	AB	O	A,B
A,B	A,B	A,B	O

- A disease

- Two treatments with differing effectiveness:

	A	B	AB	O
f_1	0.5	0.6	0.7	0.1
f_2	0.4	0.3	0.3	0.8

Blood groups problem with cheap test

Problem setup

- A partial information sample:

A,B	AB	O	A,B
A,B	A,B	A,B	O

- A disease

- Two treatments with differing effectiveness:

	A	B	AB	O
f_1	0.5	0.6	0.7	0.1
f_2	0.4	0.3	0.3	0.8

Representation

Blood groups problem with cheap test

Problem setup

- A partial information sample:

A,B	AB	O	A,B
A,B	A,B	A,B	O

- A disease

- Two treatments with differing effectiveness:

	A	B	AB	O
f_1	0.5	0.6	0.7	0.1
f_2	0.4	0.3	0.3	0.8

Representation

PMF from observed frequencies:

p_A	p_B	p_{AB}	p_O
?	?	$1/8$	$2/8$

Blood groups problem with cheap test

Problem setup

- A partial information sample:

A,B	AB	O	A,B
A,B	A,B	A,B	O

- A disease
- Two treatments with differing effectiveness:

	A	B	AB	O
f_1	0.5	0.6	0.7	0.1
f_2	0.4	0.3	0.3	0.8

Representation

PMF from observed frequencies:

p_A	p_B	p_{AB}	p_O
?	?	$1/8$	$2/8$

Reasoning

- Probability $P(\{A, O\})$?
- Expectation $E_p(f_1)$?
- Outcome with maximal probability?
- Treatment with highest expected effectiveness?

Blood groups problem with cheap test

Problem setup

- A partial information sample:

A,B	AB	O	A,B
A,B	A,B	A,B	O

- A disease

- Two treatments with differing effectiveness:

	A	B	AB	O
f_1	0.5	0.6	0.7	0.1
f_2	0.4	0.3	0.3	0.8

Representation

PMF from observed frequencies:

p_A	p_B	p_{AB}	p_O
?	?	1/8	2/8

Reasoning

- $P(\{A, O\}) = ?$
- $E_p(f_1) = ?$
- Outcome ? has maximal probability
- Treatment ? has highest expected effectiveness

We want to be able
to represent uncertainty and reason
also in situations
with partial information

Overview

Basic setup and axioms

Basic setup of **probability theory**:

- Random variable X
- Set of possible outcomes $\mathcal{X}$
- Each possible event/set $S \subseteq \mathcal{X}$
  is assigned a **probability value** $P(S)$

Basic setup and axioms

Basic setup of **probability theory**:

- Random variable X
- Finite set of possible outcomes $\mathcal{X}$
- Each possible outcome $x \in \mathcal{X}$
  is assigned a **probability value** $p_x := P(X = x) = P(\{x\})$

Basic setup and axioms

Basic setup of **probability theory**:

- Random variable X
- Finite set of possible outcomes $\mathcal{X}$
- Each possible outcome $x \in \mathcal{X}$
  is assigned a **probability value** $p_x := P(X = x) = P(\{x\})$

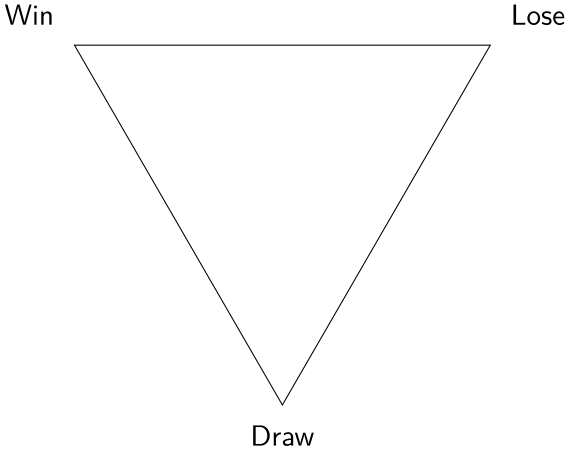
Axioms

A **probability mass function** p (and the corresponding probability measure P) must be:

1. Nonnegative: $p_x \geq 0$ for all outcomes $x \in \mathcal{X}$
2. Additive: $P(S) = \sum_{x \in S} p_x$ for all events $S \subseteq \mathcal{X}$
3. Normed: $P(\mathcal{X}) = 1$

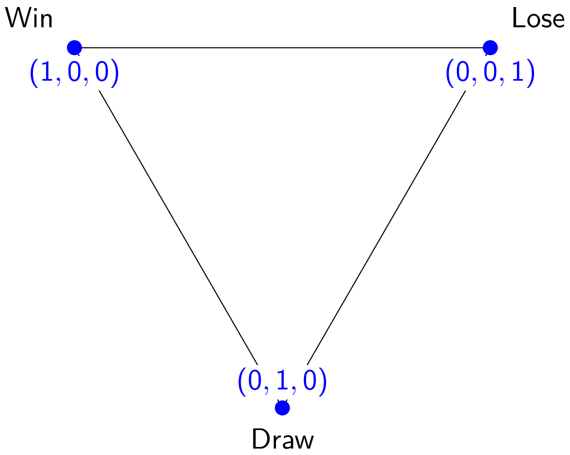
Probability simplex: visualizing probability mass functions

- $\mathcal{X} = \{\text{Win}, \text{Draw}, \text{Lose}\}$



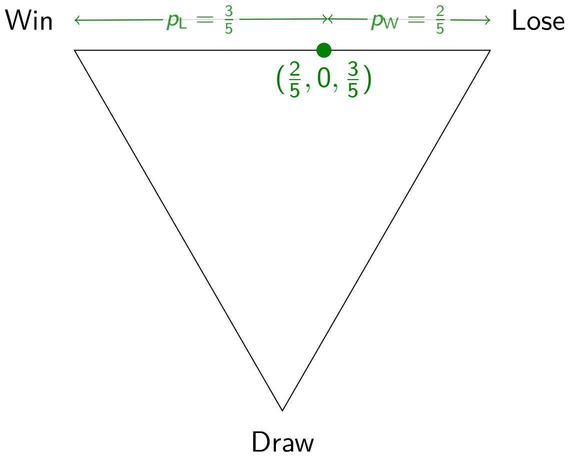
Probability simplex: visualizing probability mass functions

- $\mathcal{X} = \{\text{Win}, \text{Draw}, \text{Lose}\}$
- 'degenerate' probability mass functions (pmfs) $p = (p_\text{W}, p_\text{D}, p_\text{L})$ at the corners



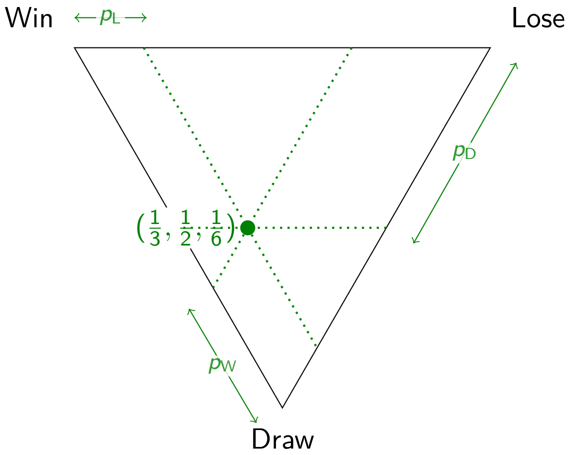
Probability simplex: visualizing probability mass functions

- $\mathcal{X} = \{\text{Win}, \text{Draw}, \text{Lose}\}$
- 'degenerate' probability mass functions (pmfs) $p = (p_\text{W}, p_\text{D}, p_\text{L})$ at the corners
- other pmfs as convex combinations thereof; values can be 'read off' as distance to opposite edge



Probability simplex: visualizing probability mass functions

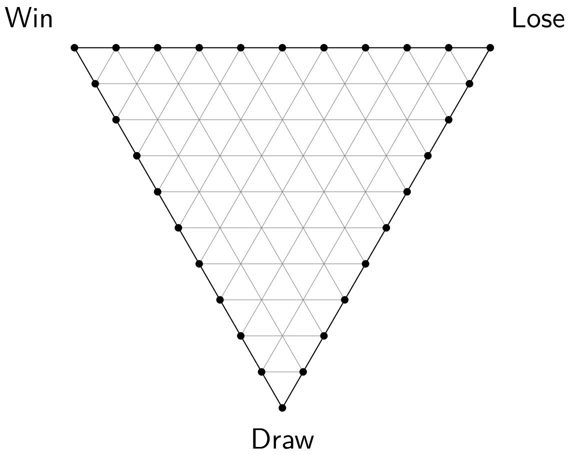
- $\mathcal{X} = \{\text{Win}, \text{Draw}, \text{Lose}\}$
- 'degenerate' probability mass functions (pmfs) $p = (p_\text{W}, p_\text{D}, p_\text{L})$ at the corners
- other pmfs as convex combinations thereof; values can be 'read off' as distance to opposite edge



Probability simplex: visualizing probability mass functions — exercise

On a gridded probability simplex on your worksheet, indicate

- The degenerate pmf corresponding to $p_L = 1$
- The pmf $(p_W, p_D, p_L) = (05, 0.5, 0)$
- The pmf $(p_W, p_D, p_L) = (0.1, 0.3, 0.6)$



Reasoning — deducing inferences and making decisions

Deducing inferences

- probability values
- expectations/previsions of real-valued functions f on $\mathcal{X}$:

$$E_p(f) := \sum_{x\in\mathcal{X}} p_x f(x)$$

Reasoning — deducing inferences and making decisions

Deducing inferences

- probability values
- expectations/previsions of real-valued functions f on $\mathcal{X}$:

$$E_p(f) := \sum_{x\in\mathcal{X}} p_x f(x)$$

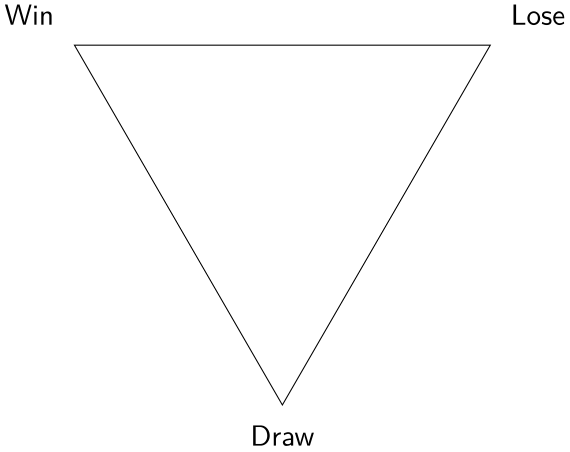
Decision making

- outcomes with maximal probability
- options minimizing/maximizing expectation

- *Topic of tomorrow morning's lecture*

Probability simplex: visualizing probabilities and expectations

- $\mathcal{X} = \{\text{Win}, \text{Draw}, \text{Lose}\}$
- Visualize probability and expectation values as the lines of pmfs for which that value is attained

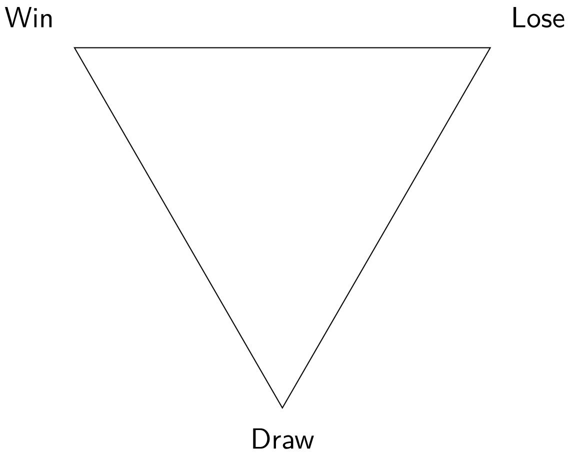


Probability simplex: visualizing probabilities and expectations

- $\mathcal{X} = \{\text{Win}, \text{Draw}, \text{Lose}\}$
- Visualize probability and expectation values as the lines of pmfs for which that value is attained

Probability example

- $A = \{\text{Win}, \text{Lose}\}$
- $P(A) = 2/3$

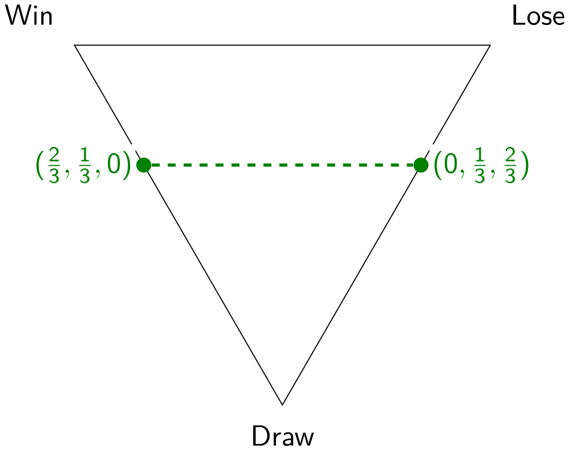


Probability simplex: visualizing probabilities and expectations

- $\mathcal{X} = \{\text{Win}, \text{Draw}, \text{Lose}\}$
- Visualize probability and expectation values as the lines of pmfs for which that value is attained

Probability example

- $A = \{\text{Win}, \text{Lose}\}$
- $P(A) = p_\text{W} + p_\text{L} = 2/3$

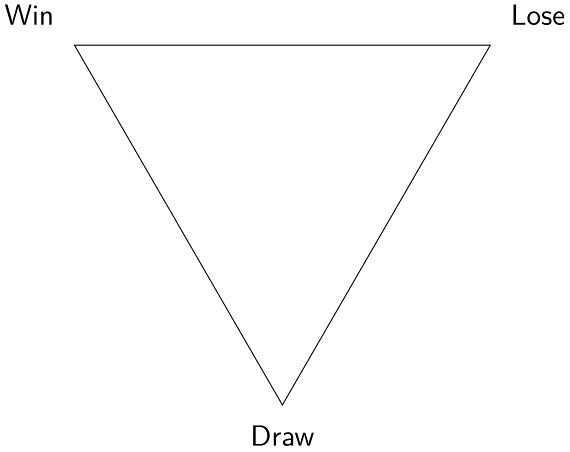


Probability simplex: visualizing probabilities and expectations

- $\mathcal{X} = \{\text{Win}, \text{Draw}, \text{Lose}\}$
- Visualize probability and expectation values as the lines of pmfs for which that value is attained

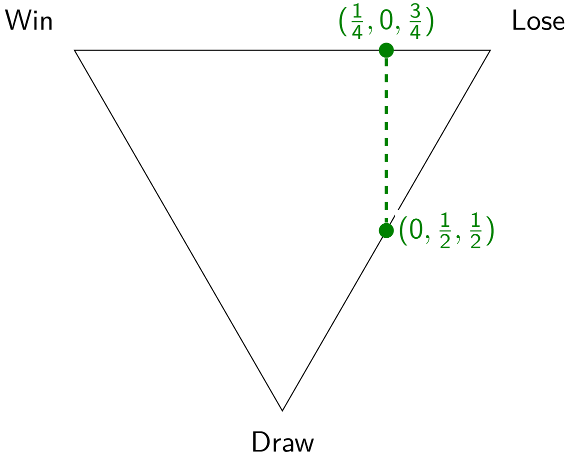
Expectation example

- $f = (f_{\text{W}}, f_{\text{D}}, f_{\text{L}}) = (1, 0, -1)$
- $E(f) = -1/2$



Probability simplex: visualizing probabilities and expectations

- $\mathcal{X} = \{\text{Win}, \text{Draw}, \text{Lose}\}$
- Visualize probability and expectation values as the lines of pmfs for which that value is attained



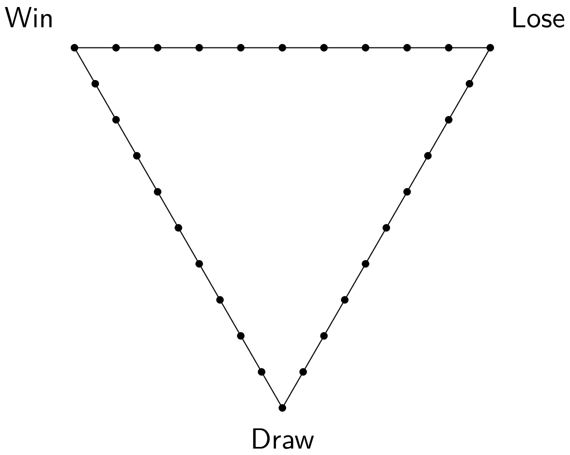
Expectation example

- $f = (f_W, f_D, f_L) = (1, 0, -1)$
- $E(f) = p_W - p_L = -1/2$

Probability simplex: visualizing probabilities and expectations — exercise

On a non-gridded probability simplex on your worksheet, indicate

- The set of pmfs for which $P(\{\mathrm{D}, \mathrm{L}\}) = 0.4$
- The set of pmfs for which $E(f) = 0$ with $f = (-1, -4, 4)$
- Is there a pmf compatible with both?
  (If yes, which?)



Learning — creating a representation

- from data, using estimation techniques (learning)
- from experts, using elicitation (asking questions)

Multivariate probability mass functions: basic setup

- Index set $N = \{1, \dots, n\}$
- Multivariate variable $\boldsymbol{X} = (X_1, X_2, \dots, X_n)$
- Set of possible outcomes $\boldsymbol{x} \in \mathcal{X} = \mathcal{X}_1 \times \mathcal{X}_2 \times \dots \times \mathcal{X}_n$
- Each possible outcome is assigned a **probability value**

Marginal probability

- Marginal probabilities are probabilities defined for events corresponding to setting some components of the variable to given values
- Let $K \subseteq N$ and

$$\boldsymbol{X}_K := (X_k : k \in K)$$

$$\boldsymbol{x}_K \in \mathcal{X}_K := \bigtimes_{k \in K} \mathcal{X}_k$$

- Marginal probabilities follow from the additivity axiom:

$$P(\boldsymbol{X}_K = \boldsymbol{x}_K) = \sum_{\substack{\boldsymbol{z} \in \mathcal{X} \\ \boldsymbol{z}_K = \boldsymbol{x}_K}} p_{\boldsymbol{z}}$$

Marginal probability example

Problem setup

- $\boldsymbol{X} = (X_1, X_2, X_3)$
- $\mathcal{X}_1 = \{\mathrm{A}, \mathrm{B}, \mathrm{C}\}$, $\mathcal{X}_2 = \{0, 1\}$, $\mathcal{X}_3 = \{+, -\}$
-

$\boldsymbol{x}$	x_1	A	A	A	A	B	B	B	B	C	C	C	C
	x_2	0	0	1	1	0	0	1	1	0	0	1	1
	x_3	$+$	$-$	$+$	$-$	$+$	$-$	$+$	$-$	$+$	$-$	$+$	$-$
$p_{\boldsymbol{x}}$		0.2	0.1	0.1	0	0.1	0	0	0.3	0	0	0	0.2

Marginal probability example

Problem setup

- $\boldsymbol{X} = (X_1, X_2, X_3)$
- $\mathcal{X}_1 = \{\mathrm{A}, \mathrm{B}, \mathrm{C}\}$, $\mathcal{X}_2 = \{0, 1\}$, $\mathcal{X}_3 = \{+, -\}$
-

	x_1	A	A	A	A	B	B	B	B	C	C	C	C
$\boldsymbol{x}$	x_2	0	0	1	1	0	0	1	1	0	0	1	1
	x_3	$+$	$-$	$+$	$-$	$+$	$-$	$+$	$-$	$+$	$-$	$+$	$-$
$p_{\boldsymbol{x}}$		0.2	0.1	0.1	0	0.1	0	0	0.3	0	0	0	0.2

Inferences

- $P(X_1 = \mathrm{B})$?
- $P(X_1 = \mathrm{A}, X_2 = 0)$?
- $P(X_1 = \mathrm{C}, X_3 = +)$?

Marginal probability example

Problem setup

- $\boldsymbol{X} = (X_1, X_2, X_3)$
- $\mathcal{X}_1 = \{\text{A}, \text{B}, \text{C}\}$, $\mathcal{X}_2 = \{0, 1\}$, $\mathcal{X}_3 = \{+, -\}$
-

	x_1	A	A	A	A	B	B	B	B	C	C	C	C
$\boldsymbol{x}$	x_2	0	0	1	1	0	0	1	1	0	0	1	1
	x_3	$+$	$-$	$+$	$-$	$+$	$-$	$+$	$-$	$+$	$-$	$+$	$-$
$p_{\boldsymbol{x}}$		0.2	0.1	0.1	0	0.1	0	0	0.3	0	0	0	0.2

Inferences

- $P(X_1 = \text{B}) = 0.4$
- $P(X_1 = \text{A}, X_2 = 0) = 0.3$
- $P(X_1 = \text{C}, X_3 = +) = 0$

Conditional probabilities

- Conditional probabilities are probabilities that hold
  - assuming some event is known to be true, or specifically
  - assuming some random variables take some given values
- Let $B \subseteq \mathcal{X}$ for which $P(B) > 0$, then for $A \subseteq \mathcal{X}$ we have $P(A|B) = \frac{P(A \cap B)}{P(B)}$
- Let $K \subseteq N$ and

$$\boldsymbol{X}_K := (X_k : k \in K)$$
$$\boldsymbol{x}_K \in \mathcal{X}_K := \bigtimes_{k \in K} \mathcal{X}_k$$

with $P(\boldsymbol{X}_K = \boldsymbol{x}_K) > 0$ then

$$P(\boldsymbol{X}_{N \setminus K} = \boldsymbol{x}_{N \setminus K} | \boldsymbol{X}_K = \boldsymbol{x}_K) = \frac{P(\boldsymbol{X}_K = \boldsymbol{x}_K, \boldsymbol{X}_{N \setminus K} = \boldsymbol{x}_{N \setminus K})}{P(\boldsymbol{X}_K = \boldsymbol{x}_K)}$$

Conditional probability example

Problem setup

- $\boldsymbol{X} = (X_1, X_2, X_3)$
- $\mathcal{X}_1 = \{\mathrm{A}, \mathrm{B}, \mathrm{C}\}$, $\mathcal{X}_2 = \{0, 1\}$, $\mathcal{X}_3 = \{+, -\}$
- Table:

	x_1	A	A	A	A	B	B	B	B	C	C	C	C
$\boldsymbol{x}$	x_2	0	0	1	1	0	0	1	1	0	0	1	1
	x_3	$+$	$-$	$+$	$-$	$+$	$-$	$+$	$-$	$+$	$-$	$+$	$-$
$p_{\boldsymbol{x}}$		0.2	0.1	0.1	0	0.1	0	0	0.3	0	0	0	0.2

Conditional probability example

Problem setup

- $\boldsymbol{X} = (X_1, X_2, X_3)$
- $\mathcal{X}_1 = \{\mathrm{A}, \mathrm{B}, \mathrm{C}\}$, $\mathcal{X}_2 = \{0, 1\}$, $\mathcal{X}_3 = \{+, -\}$
-

	x_1	A	A	A	A	B	B	B	B	C	C	C	C
$\boldsymbol{x}$	x_2	0	0	1	1	0	0	1	1	0	0	1	1
	x_3	$+$	$-$	$+$	$-$	$+$	$-$	$+$	$-$	$+$	$-$	$+$	$-$
$p_{\boldsymbol{x}}$		0.2	0.1	0.1	0	0.1	0	0	0.3	0	0	0	0.2

Inferences

- $P(X_1 = \mathrm{A} | X_2 = 0, X_3 = +)$?
- $P(X_1 = \mathrm{B} | X_2 = 0, X_3 = -)$?
- $P(X_2 = 1 | X_1 = \mathrm{C}, X_3 = +)$?

Conditional probability example

Problem setup

- $\boldsymbol{X} = (X_1, X_2, X_3)$
- $\mathcal{X}_1 = \{\mathrm{A}, \mathrm{B}, \mathrm{C}\}$, $\mathcal{X}_2 = \{0, 1\}$, $\mathcal{X}_3 = \{+, -\}$
-

	x_1	A	A	A	A	B	B	B	B	C	C	C	C
$\boldsymbol{x}$	x_2	0	0	1	1	0	0	1	1	0	0	1	1
	x_3	+	−	+	−	+	−	+	−	+	−	+	−
$p_{\boldsymbol{x}}$		0.2	0.1	0.1	0	0.1	0	0	0.3	0	0	0	0.2

Inferences

- $P(X_1 = \mathrm{A} | X_2 = 0, X_3 = +) = \frac{2}{3}$
- $P(X_1 = \mathrm{B} | X_2 = 0, X_3 = -) = 0$
- $P(X_2 = 1 | X_1 = \mathrm{C}, X_3 = +)$ is not well-defined (*why?*)

Independence

- Two random variables X_1 and X_2 are independent if their well-defined conditionals coincide with the marginals for all $\boldsymbol{x} \in \mathcal{X}$:

$$
\begin{aligned}
P(X_1 = x_1 | X_2 = x_2) &= P(X_1 = x_1) \\
P(X_2 = x_2 | X_1 = x_1) &= P(X_2 = x_2)
\end{aligned}
$$

- This is equivalent to the joint factorizing:

$$P(X_1 = x_1, X_2 = x_2) = P(X_1 = x_1)P(X_2 = x_2) \text{ for all } \boldsymbol{x} \in \mathcal{X}$$

Reasoning — deducing inferences and making decisions

- What can be done with (joint) probabilities
  can also be done with marginal and conditional probabilities

Reasoning — deducing inferences and making decisions

- What can be done with (joint) probabilities
  can also be done with marginal and conditional probabilities

Learning — creating a representation

Marginal and conditional probabilities can be either

- deduced from the learned (joint) probabilities, or
- learned directly, to together define the joint,
  possibly using independence assumptions

Multivariate probability — exercise

Consider the (joint) random variable $\boldsymbol{X} = (X_1, X_2, X_3)$ with $\mathcal{X} = \{0, 1\}^3$.
The joint probabilities are determined by the following information:

- X_1 and X_2 are independent
- It holds that $P(X_1 = 0) = 1/2$ and $P(X_2 = 0) = 1/5$
- The conditional probabilities $P(X_3 \mid X_1, X_2)$ are determined by the following table:

x_1	0	0	1	1
x_2	0	1	0	1
$P(X_3 = 0 \mid X_1 = x_1, X_2 = x_2)$	$\frac{1}{3}$	$\frac{1}{4}$	$\frac{1}{5}$	$\frac{1}{6}$

Derive that $P(X_1 = 0 \mid X_2 = 0, X_3 = 0) = \frac{5}{8}$.

As part of your calculation, you should derive and write a general expression for $P(X_1 \mid X_2, X_3)$ in terms of the (symbolic) probabilities that are given—before filling in the specific numbers.

Overview

A forecast states:

"There is an 80% probability of showers for tomorrow"

What does this mean?

(More generally, what is the meaning of probability values?)

Diversity in interpretations of probability

> There are several schools of thought regarding the interpretation of probabilities, none of them without flaws, internal contradictions, or paradoxes.
>
> De Elía & Laprise (2005)

Diversity in interpretations of probability

> There are several schools of thought regarding the interpretation of probabilities, none of them without flaws, internal contradictions, or paradoxes.
>
> De Elía & Laprise (2005)

Major interpretations

physical	*evidential*	*graded belief*
frequentist	classical	subjective
propensity	logical	

Physical

- Concerns statements about events related to physical systems
- Connected to frequency of occurrence of these events

Evidential

- A measure for the evidence supporting some (any) statement
- Typically intended to be objective

Graded belief

- A degree of belief about some (any) statement
- Typically subjective

Frequentist (physical interpretation)

An event's probability is defined as the **limit** of its **relative frequency** in many trials (the long-run probability).

Kaplan (2014)

Frequentist (physical interpretation)

An event's probability is defined as the **limit** of its **relative frequency** in many trials (the long-run probability).

Kaplan (2014)

How short or long can or should the sequence of trials be?

Frequentist (physical interpretation)

> An event's probability is defined as the **limit** of its **relative frequency** in many trials (the long-run probability).
>
> Kaplan (2014)

How short or long can or should the sequence of trials be?

Classical (evidential interpretation)

> Assuming **equally possible** cases, the probability of an event is the **ratio** of the relative number of cases favorable to it.
>
> Wikipedia (paraphrased)

Frequentist (physical interpretation)

> An event's probability is defined as the **limit** of its **relative frequency** in many trials (the long-run probability).
>
> Kaplan (2014)

How short or long can or should the sequence of trials be?

Classical (evidential interpretation)

> Assuming **equally possible** cases, the probability of an event is the **ratio** of the relative number of cases favorable to it.
>
> Wikipedia (paraphrased)

What does the assumption mean and what are its grounds?

Subjective (graded belief interpretation)

A subjective probability is anyone's **opinion** of what the probability is for an event.

Siegel & Wagner (2022)

Subjective (graded belief interpretation)

> A subjective probability is anyone's **opinion** of what the probability is for an event.
>
> Siegel & Wagner (2022)

Is it a problem if probability values are opinions?

Subjective (graded belief interpretation)

> A subjective probability is anyone's **opinion** of what the probability is for an event.
>
> Siegel & Wagner (2022)

Is it a problem if probability values are opinions?

Betting interpretation (special case of subjective interpretation)

- Probabilities are defined by the subject's betting behavior (therefore also called *behavioral interpretation*)
- Relevant for our discussion of imprecise probability theories

Betting game setup

- Two players:
  - **subject** (gambler)
  - **bookie** (proposes bets)

- **Gambles** from the subject's perspective:

$$1_x(z) = \begin{cases} 1 & \text{if } z = x \\ 0 & \text{if } z \neq x \end{cases}$$

for all $x \in \mathcal{X}$

Betting game setup

- Two players:
  - **subject** (gambler)
  - **bookie** (proposes bets)

- **Gambles** from the subject's perspective:

$$1_x(z) = \begin{cases} 1 & \text{if } z = x \\ 0 & \text{if } z \neq x \end{cases}$$

  for all $x \in \mathcal{X}$

Eliciting probabilities

- For each of the gambles 1_x, the subject offers their **fair price** p_x
- The bookie can propose to exchange one for the other, i.e., $1_x - p_x$ or $p_x - 1_x$, which the subject is committed to accept

- The p_x are the subject's probabilities

Betting game setup

- Two players:
  - **subject** (gambler)
  - **bookie** (proposes bets)
- **Gambles** from the subject's perspective:

$$1_x(z) = \begin{cases} 1 & \text{if } z = x \\ 0 & \text{if } z \neq x \end{cases}$$

for all $x \in \mathcal{X}$

Eliciting probabilities

- For each of the gambles 1_x, the subject offers their **fair price** p_x
- The bookie can propose to exchange one for the other, i.e., $1_x - p_x$ or $p_x - 1_x$, which the subject is committed to accept
- The p_x are the subject's probabilities
- Should the subject state any set of fair prices?

Coherence: deriving the axioms of probability from the betting game

- Assume the gambler specifies a negative price $p_x < 0$.
- Exchange proposed by the bookie:

$$p_x - 1_x = \begin{cases} p_x - 1 < 0 & \text{if } x \text{ occurs} \\ p_x < 0 & \text{otherwise} \end{cases}$$

- Should be unacceptable to the subject, because it implies a *sure loss*, i.e., is negative whatever occurs!
- **Nonnegativity axiom is required**

Coherence: deriving the axioms of probability from the betting game

- The bookie can choose to propose multiple bets, concerning some $S \subseteq \mathcal{X}$.
- Combined exchange proposed by the bookie:

$$\sum_{x \in S} (p_x - 1_x) = \Big(\sum_{x \in S} p_x\Big) - 1_S, \text{ with } 1_S = \sum_{x \in S} 1_x = \begin{cases} 1 & \text{if } S \text{ occurs} \\ 0 & \text{otherwise} \end{cases}$$

- The fair price $P(S)$ for the combined bet is $\sum_{x \in S} p_x$
- **Additivity axiom follows**

Coherence: deriving the axioms of probability from the betting game

- The bookie can choose to combine the bets for all $x \in \mathcal{X}$, so for the gamble $1_{\mathcal{X}} = 1$
- Possible exchanges to be proposed: $P(\mathcal{X}) - 1$ or $1 - P(\mathcal{X})$
- Exchanges do not depend on the $x \in \mathcal{X}$ that occurs
- If $P(\mathcal{X}) \neq 1$, one possible proposal would result in a *sure loss*
- **The normedness axiom is required**

Overview

- Kick-off (slot 1)
- Classical probability theory (slot 1)
- Interpretation of probability (slot 2)
- Limitations of probability theory (slot 2)
  - Aleatoric vs. epistemic uncertainty
  - Representing epistemic uncertainty
  - Rational agents vs. Real agents
  - Why go beyond probability theory?
- Probability intervals (slot 3)
- Credal sets (slot 3–4)
- Acceptability & Desirability (slot 4–5)
- Interval expectation & probability (slot 5–6)

Aleatoric uncertainty — the irreducible part

- Aleatoric uncertainty arises from *random variation*
- Additional information cannot reduce it
- Other names: *variability*, *stochasticity*, *randomness*, *chance*, *risk*

Aleatoric uncertainty — the irreducible part

- Aleatoric uncertainty arises from *random variation*
- Additional information cannot reduce it
- Other names: *variability*, *stochasticity*, *randomness*, *chance*, *risk*

Epistemic uncertainty — the reducible part

- Epistemic uncertainty arises from *a lack of knowledge* (also at inference time)
- Additional information can reduce or eliminate it
- Other names: *incertitude*, *ambiguity*, *ignorance*, *imprecision*

Aleatoric uncertainty — the irreducible part

- Aleatoric uncertainty arises from *random variation*
- Additional information cannot reduce it
- Other names: *variability*, *stochasticity*, *randomness*, *chance*, *risk*
- Sources: spatial variation, temporal fluctuations, manufacturing or genetic differences,. . .

Epistemic uncertainty — the reducible part

- Epistemic uncertainty arises from *a lack of knowledge* (also at inference time)
- Additional information can reduce or eliminate it
- Other names: *incertitude*, *ambiguity*, *ignorance*, *imprecision*
- Sources: limited sample size, mensurational limits ('measurement error'), censoring, poorly defined outcomes,. . .

Aleatoric uncertainty examples

- Outcome of toss of fair coin

Epistemic uncertainty examples

- Bias of a coin for tossing

Aleatoric uncertainty examples

- Outcome of toss of fair coin
- Decay time of a radioactive atom

Epistemic uncertainty examples

- Bias of a coin for tossing
- Weight of a proton

Aleatoric uncertainty examples

- Outcome of toss of fair coin
- Decay time of a radioactive atom
- Value of a decimal of a randomly generated number

Epistemic uncertainty examples

- Bias of a coin for tossing
- Weight of a proton
- Value of a decimal of an irrational mathematical constant

Aleatoric uncertainty examples

- Outcome of toss of fair coin
- Decay time of a radioactive atom
- Value of a decimal of a randomly generated number
- 'Noise'

Epistemic uncertainty examples

- Bias of a coin for tossing
- Weight of a proton
- Value of a decimal of an irrational mathematical constant
- Model uncertainty
  - Parameter values
  - Dependencies
  - Functional forms
  - Level of abstraction

Can probability theory differentiate between aleatoric and epistemic uncertainty?

The importance of sample size: coin-flipping example

Two different coins

Coin	S	L
Flips	2	$2 \cdot 10^6$
Heads	50%	50%
Tails	50%	50%

The importance of sample size: coin-flipping example

Two different coins

Coin	S	L
Flips	2	$2 \cdot 10^6$
Heads	50%	50%
Tails	50%	50%

Maximum likelihood estimation gives

$$p_{\text{Heads}} = p_{\text{Tails}} = \frac{1}{2}$$

for both coins

The importance of sample size: coin-flipping example

Two different coins

Coin	S	L
Flips	2	$2 \cdot 10^6$
Heads	50%	50%
Tails	50%	50%

Maximum likelihood estimation gives

$$p_{\text{Heads}} = p_{\text{Tails}} = \frac{1}{2}$$

for both coins

- What can you say about the reliability of the estimate for each coin?
- How would you communicate the probability estimates?
- How would you communicate derived inferences and decisions?

The importance of sample size: coin-flipping example

Two different coins

Coin	S	L
Flips	2	$2 \cdot 10^6$
Heads	50%	50%
Tails	50%	50%

Maximum likelihood estimation gives

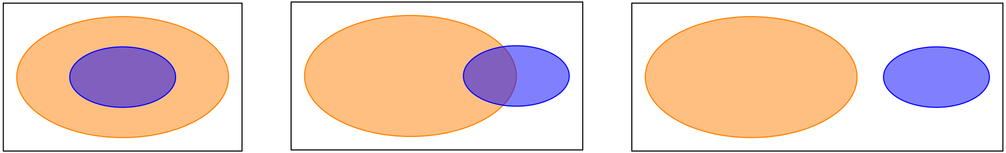
$$p_{\text{Heads}} = p_{\text{Tails}} = \frac{1}{2}$$

for both coins

- What can you say about the reliability of the estimate for each coin?
- How would you communicate the probability estimates?
- How would you communicate derived inferences and decisions?
- What about inferences depending on 1000 random variables with varying reliabilities?

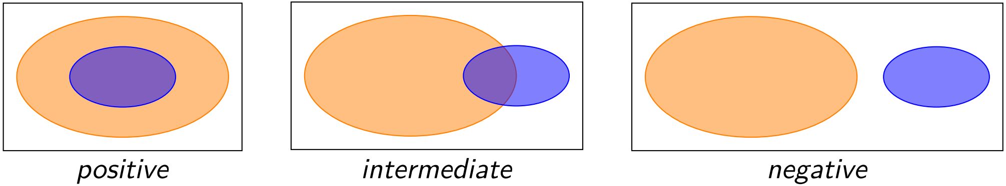
Unknown dependence

- Nature of dependence between two events A and B often not known



Unknown dependence and Fréchet's bounds

- Nature of dependence between two events A and B often not known



- Fréchet's bounds:

$$P(A \wedge B) \in \Big[\max\{0, P(A) + P(B) - 1\}, \min\{P(A), P(B)\}\Big]$$

$$P(A \vee B) \in \Big[\max\{P(A), P(B)\}, \min\{1, P(A) + P(B)\}\Big]$$

Can probability theory differentiate between aleatoric and epistemic uncertainty?

> [Aleatoric and epistemic uncertainty] must be treated *differently*;
>
> variability should be modeled as randomness
> with the methods of probability theory;
>
> incertitude should be modeled as ignorance
> with the methods of interval analysis.
>
> Ferson (2004)

> [Aleatoric and epistemic uncertainty] must be treated *differently*;
>
> variability should be modeled as randomness
> with the methods of probability theory;
>
> incertitude should be modeled as ignorance
> with the methods of interval analysis.
>
> Ferson (2004)

Interval analysis

- Computation with intervals instead of with single numbers
- Example: $[2, 4] - [3, 5] = [-3, 1]$
- Ideal is to obtain tightest bounds
- In practice often outer bounds are used for computational reasons

Rational agents

In Savage's classical account of *Subjective Expected Utility Theory*, a **'rational' agent**

- models uncertainty in a problem **using a single probability measure**
- chooses between alternatives by **maximizing expected utility**

Rational agents

In Savage's classical account of *Subjective Expected Utility Theory*, a **'rational' agent**

- models uncertainty in a problem **using a single probability measure**
- chooses between alternatives by **maximizing expected utility**

Urn example

- Urn with 20 red, 10 black, and 10 white balls
  The rational agent uses a pmf with
  $p_\mathrm{R} = 1/2$ and $p_\mathrm{B} = p_\mathrm{W} = 1/4$

Rational agents

In Savage's classical account of *Subjective Expected Utility Theory*, a **'rational' agent**

- models uncertainty in a problem **using a single probability measure**
- chooses between alternatives by **maximizing expected utility**

Urn example

- Urn with 20 red, 10 black, and 10 white balls
  The rational agent uses a pmf with
  $p_R = 1/2$ and $p_B = p_W = 1/4$
- Consider three gambles whose payoff depends on the outcome of a random draw from the urn:

	R	B	W
f_{RB}	\$100	\$100	\$0
f_{RW}	\$100	\$0	\$100
f_{BW}	\$0	\$100	\$100

  The agent is *indifferent* between f_{RB} and f_{RW}, which they *strictly prefer* to f_{BW}

Are real agents irrational?

Risk aversion

- Urn with 20 red, 20 gray balls
- Two gambles:

	R	G
f_{RG}	\$50	\$50
f_R	\$100	\$0

Are real agents irrational?

Risk aversion

- Urn with 20 red, 20 gray balls
- Two gambles:

	R	G
f_{RG}	\$50	\$50
f_{R}	\$100	\$0

- Rational agents are *indifferent* between f_{RG} and f_{R}
- Real agents *strictly prefer* f_{RG} over f_{R}

Are real agents irrational?

Risk aversion

- Urn with 20 red, 20 gray balls
- Two gambles:

	R	G
f_{RG}	\$50	\$50
f_R	\$100	\$0

- Rational agents are *indifferent* between f_{RG} and f_R
- Real agents *strictly prefer* f_{RG} over f_R

Ambiguity aversion

- Two urns; the agent must choose one:
  - E 20 black balls, 20 white balls
  - U unknown proportion of black and white balls
- One gamble:

	B	W
f_B	\$100	\$0

Are real agents irrational?

Risk aversion

- Urn with 20 red, 20 gray balls
- Two gambles:

	R	G
f_{RG}	\$50	\$50
f_R	\$100	\$0

- Rational agents are *indifferent* between f_{RG} and f_R
- Real agents *strictly prefer* f_{RG} over f_R

Ambiguity aversion

- Two urns; the agent must choose one:
  - E 20 black balls, 20 white balls
  - U unknown proportion of black and white balls
- One gamble:

	B	W
f_B	\$100	\$0

- Rational agents must choose a pmf for U; in case they choose $p_B = p_W = 1/2$ they are *indifferent* between urns E and U
- Real agents *strictly prefer* urn E over urn U

Are real agents irrational?

Ellsberg paradox

- Urn with 20 red balls and 40 white or black balls in unknown proportion
- Four gambles:

	R	B	W
f_R	\$100	\$0	\$0
f_B	\$0	\$100	\$0
$f_{RW} = f_R + f_W$	\$100	\$0	\$100
$f_{BW} = f_B + f_W$	\$0	\$100	\$100

Are real agents irrational?

Ellsberg paradox

- Urn with 20 red balls and 40 white or black balls in unknown proportion
- Four gambles:

	R	B	W
f_R	\$100	\$0	\$0
f_B	\$0	\$100	\$0
$f_{RW} = f_R + f_W$	\$100	\$0	\$100
$f_{BW} = f_B + f_W$	\$0	\$100	\$100

- Choose between f_R and f_B
- Choose between f_{RW} and f_{BW}

Are real agents irrational?

Ellsberg paradox

- Urn with 20 red balls and 40 white or black balls in unknown proportion
- Four gambles:

	R	B	W
f_R	\$100	\$0	\$0
f_B	\$0	\$100	\$0
$f_{RW} = f_R + f_W$	\$100	\$0	\$100
$f_{BW} = f_B + f_W$	\$0	\$100	\$100

- Choose between f_R and f_B

  Real agents
  strictly prefer f_R over f_B,
  for a rational agent implying
  $p_R > p_B$

- Choose between f_{RW} and f_{BW}

  Real agents
  strictly prefer f_{BW} over f_{RW},
  for a rational agent implying
  $p_R < p_B$

Why go beyond probability theory?

To be able to deal with epistemic uncertainty:

- Distinguish sample size in uncertainty representation
- Express partial or missing knowledge

To let agents act according to a less restrictive definition of rationality:

- Be able to reflect justified aversions
- Model behavior that would otherwise be paradoxical

Overview

Kick-off (slot 1)

Classical probability theory (slot 1)

Interpretation of probability (slot 2)

Limitations of probability theory (slot 2)

Probability intervals (slot 3)
- Illustration: Blood groups
- Representation
- Reasoning

Credal sets (slot 3–4)

Acceptability & Desirability (slot 4–5)

Interval expectation & probability (slot 5–6)

Blood groups problem with cheap test

Problem setup

- A partial information sample:

A,B	AB	O	A,B
A,B	A,B	A,B	O

- A disease

- Two treatments with differing effectiveness:

	A	B	AB	O
f_1	0.5	0.6	0.7	0.1
f_2	0.4	0.3	0.3	0.8

Representation

PMF from observed frequencies:

p_A	p_B	p_{AB}	p_O
?	?	$\frac{1}{8}$	$\frac{2}{8}$

Blood groups problem with cheap test

Problem setup

- A partial information sample:

A,B	AB	O	A,B
A,B	A,B	A,B	O

- A disease

- Two treatments with differing effectiveness:

	A	B	AB	O
f_1	0.5	0.6	0.7	0.1
f_2	0.4	0.3	0.3	0.8

Representation

PMF from observed frequencies:

p_A	p_B	p_{AB}	p_O
$[0, \frac{5}{8}]$	$[0, \frac{5}{8}]$	$\frac{1}{8}$	$\frac{2}{8}$

Blood groups problem with cheap test

Problem setup

- A partial information sample:

A,B	AB	O	A,B
A,B	A,B	A,B	O

- A disease

- Two treatments with differing effectiveness:

	A	B	AB	O
f_1	0.5	0.6	0.7	0.1
f_2	0.4	0.3	0.3	0.8

Representation

PMF from observed frequencies:

p_{A}	p_{B}	p_{AB}	p_{O}
$[0, \frac{5}{8}]$	$[0, \frac{5}{8}]$	$\frac{1}{8}$	$\frac{2}{8}$

Reasoning

- Bounds on $P(\{\text{A}, \text{O}\})$?
- Bounds on $E_p(f_1)$?
- Outcome with maximal lower probability?
- Treatment with highest upper expected effectiveness?

Blood groups problem with cheap test

Problem setup

- A partial information sample:

A,B	AB	O	A,B
A,B	A,B	A,B	O

- A disease

- Two treatments with differing effectiveness:

	A	B	AB	O
f_1	0.5	0.6	0.7	0.1
f_2	0.4	0.3	0.3	0.8

Representation

PMF from observed frequencies:

p_A	p_B	p_{AB}	p_O
$[0, \frac{5}{8}]$	$[0, \frac{5}{8}]$	$\frac{1}{8}$	$\frac{2}{8}$

Reasoning

- $P(\{A, O\}) \in [\frac{2}{8}, \frac{7}{8}]$
- $E_p(f_1) \in= [\frac{3.4}{8}, \frac{3.9}{8}] = [0.425, 0.4875]$
- Outcome O has maximal lower probability
- Both treatments have equal upper expected effectiveness

Basic setup, axioms, and terminology

Basic setup of **the theory of probability intervals** (Campos, Huete, Moral, 1994):

- Random variable X
- Finite set of outcomes $\mathcal{X}$
- Each outcome $x \in \mathcal{X}$ is assigned **lower and upper probability values** $(\underline{p}_x, \overline{p}_x)$

Basic setup, axioms, and terminology

Basic setup of **the theory of probability intervals** (Campos, Huete, Moral, 1994):

- Random variable X
- Finite set of outcomes $\mathcal{X}$
- Each outcome $x \in \mathcal{X}$ is assigned **lower and upper probability values** $(\underline{p}_x, \overline{p}_x)$

Axioms

A **probability interval** $(\underline{p}, \overline{p})$, a pair of lower and upper probability mass functions, must be:

1. **?**
2. **?**
3. **?**

Basic setup, axioms, and terminology

Basic setup of **the theory of probability intervals** (Campos, Huete, Moral, 1994):

- Random variable X
- Finite set of outcomes $\mathcal{X}$
- Each outcome $x \in \mathcal{X}$ is assigned **lower and upper probability values** $(\underline{p}_x, \overline{p}_x)$

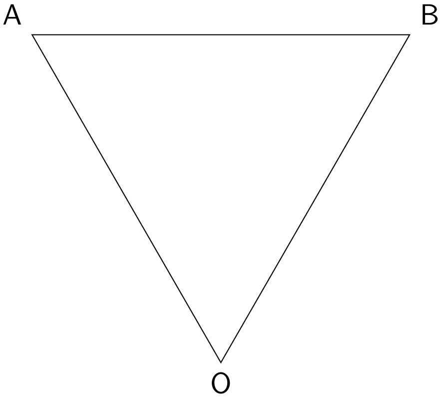
Axioms

A **probability interval** $(\underline{p}, \overline{p})$, a pair of lower and upper probability mass functions, must be:

1. Bounded: $0 \leq \underline{p}_x \leq \overline{p}_x \leq 1$ for all outcomes $x \in \mathcal{X}$
2. Proper: $\sum_{x \in \mathcal{X}} \underline{p}_x \leq 1 \leq \sum_{x \in \mathcal{X}} \overline{p}_x$
3. **?**

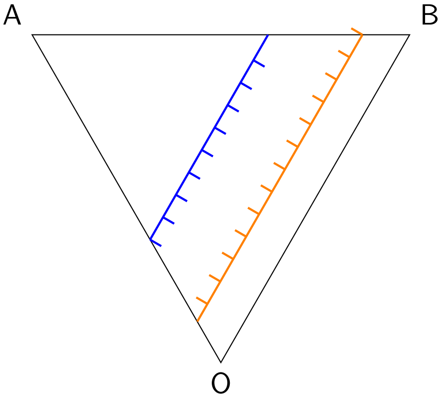
Unreachable bounds require tightening

	A	B	O
$\overline{p}$			
$\underline{p}$			



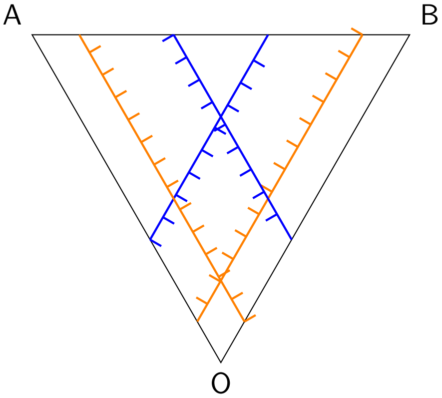
Unreachable bounds require tightening

	A	B	O
$\overline{p}$	3/8		
$\underline{p}$	1/8		



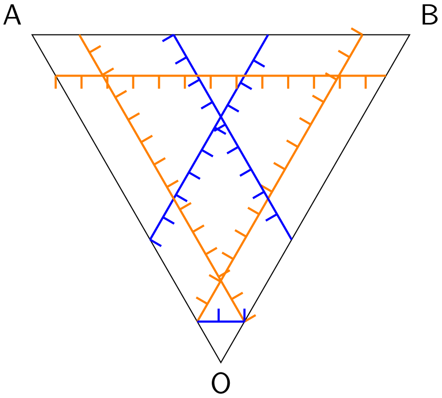
Unreachable bounds require tightening

	A	B	O
$\overline{p}$	3/8	3/8	
$\underline{p}$	1/8	1/8	



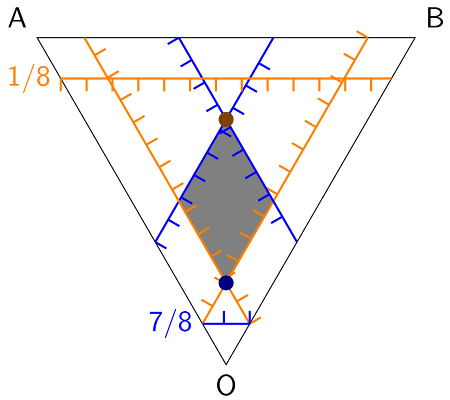
Unreachable bounds require tightening

	A	B	O
$\overline{p}$	$3/8$	$3/8$	$7/8$
$\underline{p}$	$1/8$	$1/8$	$1/8$



Unreachable bounds require tightening

	A	B	O
$\overline{p}$	3/8	3/8	7/8
$\underline{p}$	1/8	1/8	1/8



Basic setup, axioms, and terminology

Basic setup of **the theory of probability intervals** (Campos, Huete, Moral, 1994):

- Random variable X
- Finite set of outcomes $\mathcal{X}$
- Each outcome $x \in \mathcal{X}$ is assigned **lower and upper probability values** $(\underline{p}_x, \overline{p}_x)$
- Set of compatible pmfs is called the **credal set**

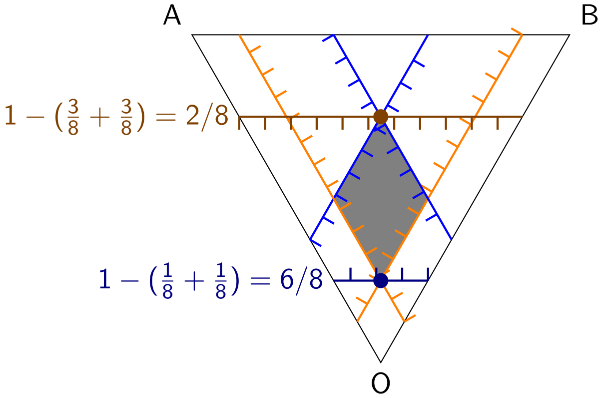
Axioms

A **probability interval** $(\underline{p}, \overline{p})$, a pair of lower and upper probability mass functions, must be:

1. Bounded: $0 \leq \underline{p}_x \leq \overline{p}_x \leq 1$ for all outcomes $x \in \mathcal{X}$
2. Proper: $\sum_{x \in \mathcal{X}} \underline{p}_x \leq 1 \leq \sum_{x \in \mathcal{X}} \overline{p}_x$
3. **?**

Unreachable bounds require tightening

	A	B	O
$\overline{p}$	3/8	3/8	6/8
$\underline{p}$	1/8	1/8	2/8



Basic setup, axioms, and terminology

Basic setup of **the theory of probability intervals** (Campos, Huete, Moral, 1994):

- Random variable X
- Finite set of outcomes $\mathcal{X}$
- Each outcome $x \in \mathcal{X}$ is assigned **lower and upper probability values** $(\underline{p}_x, \overline{p}_x)$
- Set of compatible pmfs is called the **credal set**

Axioms

A **probability interval** $(\underline{p}, \overline{p})$, a pair of lower and upper probability mass functions, must be:

1. Bounded: $0 \leq \underline{p}_x \leq \overline{p}_x \leq 1$ for all outcomes $x \in \mathcal{X}$
2. Proper: $\sum_{x \in \mathcal{X}} \underline{p}_x \leq 1 \leq \sum_{x \in \mathcal{X}} \overline{p}_x$
3. Reachable: $\underline{p}_x \geq 1 - \sum_{z \neq x} \overline{p}_z$ and $\overline{p}_x \leq 1 - \sum_{z \neq x} \underline{p}_z$ for all outcomes $x \in \mathcal{X}$

Comparison by degree of imprecision

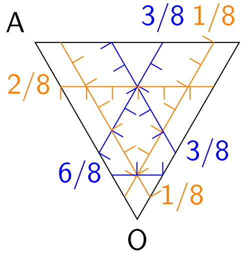
- A probability interval $(\underline{p}, \overline{p})$ is *included in* or *less imprecise than* a probability interval $(\underline{q}, \overline{q})$ if for all $x \in \mathcal{X}$

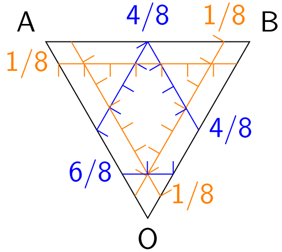
$$[\underline{p}_x, \overline{p}_x] \subseteq [\underline{q}_x, \overline{q}_x]$$

Comparison by degree of imprecision

- A probability interval $(\underline{p}, \overline{p})$ is *included in* or *less imprecise than* a probability interval $(\underline{q}, \overline{q})$ if for all $x \in \mathcal{X}$

$$[\underline{p}_x, \overline{p}_x] \subseteq [\underline{q}_x, \overline{q}_x]$$



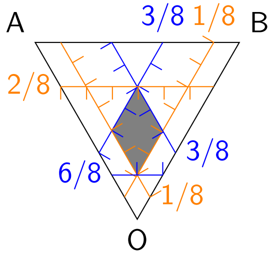


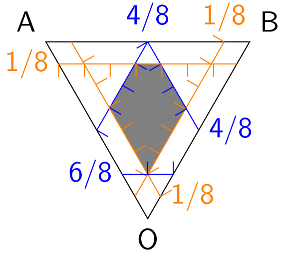
Comparison by degree of imprecision

- A probability interval $(\underline{p}, \overline{p})$ is *included in* or *less imprecise than* a probability interval $(\underline{q}, \overline{q})$ if for all $x \in \mathcal{X}$

$$[\underline{p}_x, \overline{p}_x] \subseteq [\underline{q}_x, \overline{q}_x]$$

- The corresponding credal sets will also respect the same inclusion relationship





Lower and upper probability mass sums

$$\underset{\sim}{P}(S) := \sum_{x \in S} \underline{p}_x, \qquad \widetilde{P}(S) := \sum_{x \in S} \overline{p}_x$$

Lower probability

$$\underline{P}(S) := \max\{\underset{\sim}{P}(S), 1 - \widetilde{P}(S^{\mathrm{c}})\}$$

Upper probability

$$\overline{P}(S) := \min\{\widetilde{P}(S), 1 - \underset{\sim}{P}(S^{\mathrm{c}})\}$$

Lower and upper probability mass sums

$$\underset{\sim}{P}(S) := \sum_{x \in S} \underline{p}_x, \qquad \widetilde{P}(S) := \sum_{x \in S} \overline{p}_x$$

	A	B	O	AB
$\overline{p}$	3/8	3/8	5/8	5/8
$\underline{p}$	1/8	1/8	3/8	0/8

Lower probability

$$\underline{P}(S) := \max\{\underset{\sim}{P}(S), 1 - \widetilde{P}(S^{\mathrm{c}})\}$$

Upper probability

$$\overline{P}(S) := \min\{\widetilde{P}(S), 1 - \underset{\sim}{P}(S^{\mathrm{c}})\}$$

Inferences

- Lower probability $\underline{P}(\{\mathrm{A}, \mathrm{B}\})$?
- Upper probability $\overline{P}(\{\mathrm{B}, \mathrm{O}\})$?

Lower and upper probability mass sums

$$\underset{\sim}{P}(S) := \sum_{x \in S} \underline{p}_x, \qquad \widetilde{P}(S) := \sum_{x \in S} \overline{p}_x$$

	A	B	O	AB
$\overline{p}$	3/8	3/8	5/8	5/8
$\underline{p}$	1/8	1/8	3/8	0/8

Lower probability

$$\underline{P}(S) := \max\{\underset{\sim}{P}(S), 1 - \widetilde{P}(S^c)\}$$

Upper probability

$$\overline{P}(S) := \min\{\widetilde{P}(S), 1 - \underset{\sim}{P}(S^c)\}$$

Inferences

- $\underline{P}(\{A, B\}) = 2/8$
- $\overline{P}(\{B, O\}) = 7/8$

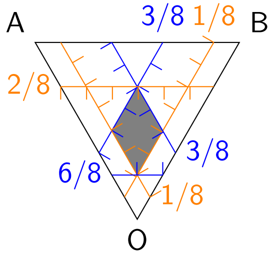
Comparison by degree of imprecision

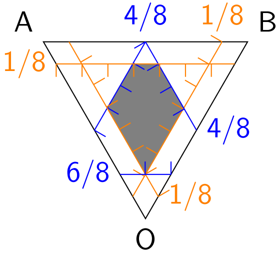
- A probability interval $(\underline{p}, \overline{p})$ is *included in* or *less imprecise than* a probability interval $(\underline{q}, \overline{q})$ if for all $x \in \mathcal{X}$

$$[\underline{p}_x, \overline{p}_x] \subseteq [\underline{q}_x, \overline{q}_x]$$

- The corresponding credal sets will also respect the same inclusion relationship
- Probability bounds will then respect the same inclusion relationship for all $S \subseteq \mathcal{X}$:

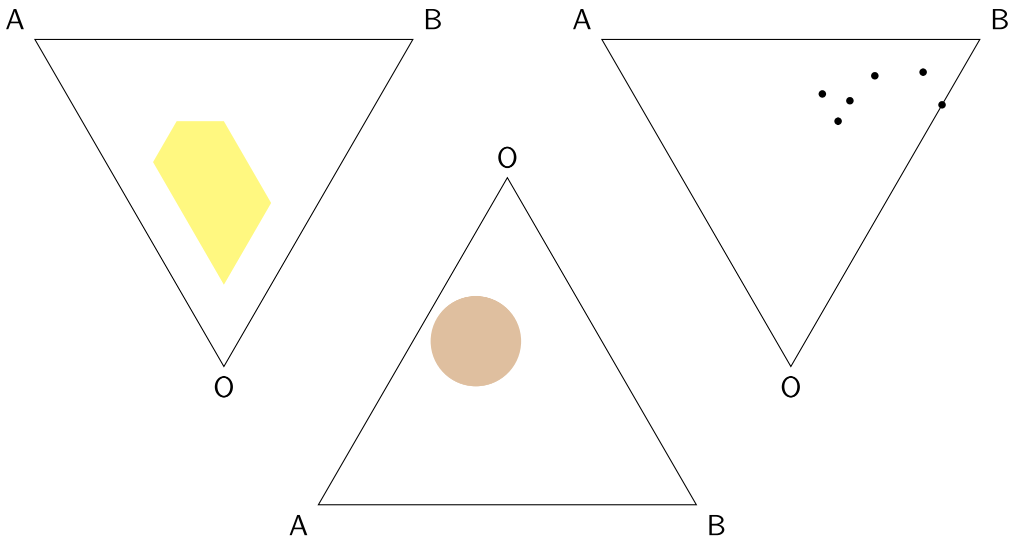
$$[\underline{P}(S), \overline{P}(S)] \subseteq [\underline{Q}(S), \overline{Q}(S)]$$





Overview

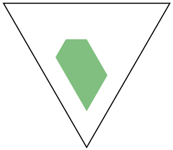
- Kick-off (slot 1)
- Classical probability theory (slot 1)
- Interpretation of probability (slot 2)
- Limitations of probability theory (slot 2)
- Probability intervals (slot 3)
- Credal sets (slot 3–4)
  - Representation
  - Reasoning
  - Multivariate credal sets
- Acceptability & Desirability (slot 4–5)
- Interval expectation & probability (slot 5–6)


A
B
O
O
A
B
A
B
O

Credal sets determined by bounds

Set of pmfs (probability measures) determined by *non-strict constraints* specified as

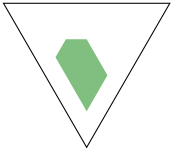
- a probability interval



Credal sets determined by bounds

Set of pmfs (probability measures) determined by *non-strict constraints* specified as

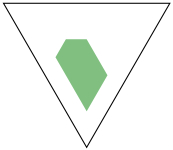
- a probability interval
- interval (lower & upper) probabilities or expectations (later)



Credal sets determined by bounds

Set of pmfs (probability measures) determined by *non-strict constraints* specified as

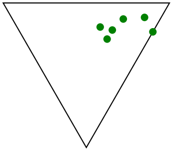
- a probability interval
- interval (lower & upper) probabilities or expectations (later)



Directly specified credal sets

Set of pmfs (probability measures) specified *directly and explicitly* as such by

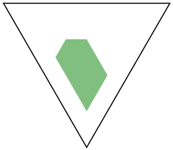
- discrete sets (‘*sets of Bayesians*’)



Credal sets determined by bounds

Set of pmfs (probability measures) determined by *non-strict constraints* specified as

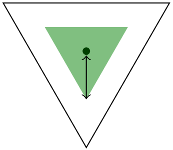
- a probability interval
- interval (lower & upper) probabilities or expectations (later)



Directly specified credal sets

Set of pmfs (probability measures) specified *directly and explicitly* as such by

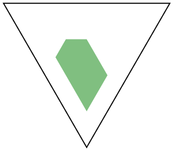
- discrete sets (‘*sets of Bayesians*’)
- neighborhoods of specific pmfs (‘*robust Bayesians*’)



Credal sets determined by bounds

Set of pmfs (probability measures) determined by *non-strict constraints* specified as

- a probability interval
- interval (lower & upper) probabilities or expectations (later)



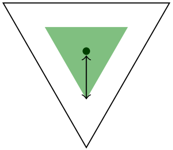
Convex and closed sets

Computationally convenient

Directly specified credal sets

Set of pmfs (probability measures) specified *directly and explicitly* as such by

- discrete sets ('*sets of Bayesians*')
- neighborhoods of specific pmfs ('*robust Bayesians*')



Generally non-convex and not closed

Can have impact on, e.g., decision rules

Basic setup, axioms, and terminology

Basic setup of **the theory of credal sets**:

- Random variable X
- Set of outcomes $\mathcal{X}$
- A *credal set* of probability mass functions for X is specified

Basic setup, axioms, and terminology

Basic setup of **the theory of credal sets**:

- Random variable X
- Set of outcomes $\mathcal{X}$
- A *credal set* of probability mass functions for X is specified
- Uses representation from classical probability as building block

Basic setup, axioms, and terminology

Basic setup of **the theory of credal sets**:

- Random variable X
- Set of outcomes $\mathcal{X}$
- A *credal set* of probability mass functions for X is specified
- Uses representation from classical probability as building block

Axioms

A **credal set** $\mathcal{C}$ must be:

1. A subset of the set of all probability mass functions for X: $\mathcal{C} \subseteq \mathcal{P}_X$
2. Non-empty: $\mathcal{C} \neq \emptyset$

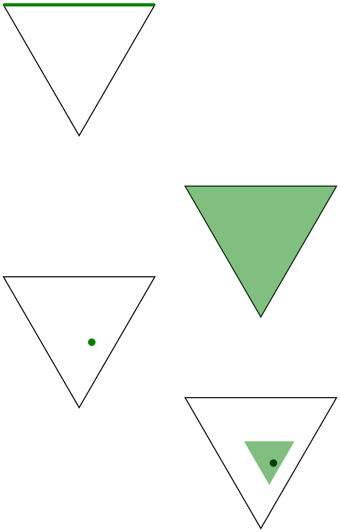
Special types of credal sets

- The vacuous credal set

$$\mathcal{P}_X^S := \{p \in \mathcal{P}_X : P(S) = 1\}$$

  relative to some event $S \subseteq \mathcal{X}$
  expresses that $X \in S$ and nothing more
- The vacuous credal set $\mathcal{P}_X^{\mathcal{X}} = \mathcal{P}_X$
  expresses complete ignorance
- Singleton credal sets $\{p\}$ correspond to
  the unique pmf p they contain
- Linear-vacuous or ε-contamination credal sets
  are a simple neighborhood model:

$$\mathcal{C}^{p,\varepsilon} := \{(1-\varepsilon)p + \varepsilon q : q \in \mathcal{P}_X\}$$



Lower & upper probability as lower & upper envelopes

$$\underline{P}(S) := \inf_{p \in \mathcal{C}} P_p(S) \qquad \overline{P}(S) := \sup_{p \in \mathcal{C}} P_p(S)$$

Lower & upper probability as lower & upper envelopes

$$\underline{P}(S) := \inf_{p \in \mathcal{C}} P_p(S) \qquad \overline{P}(S) := \sup_{p \in \mathcal{C}} P_p(S)$$

Lower & upper expectation as lower & upper envelopes

$$\underline{E}(f) := \inf_{p \in \mathcal{C}} E_p(f) \qquad \overline{E}(f) := \sup_{p \in \mathcal{C}} E_p(f)$$

Lower & upper probability as lower & upper envelopes

$$\underline{P}(S) := \inf_{p \in \mathcal{C}} P_p(S) \qquad \overline{P}(S) := \sup_{p \in \mathcal{C}} P_p(S)$$

Lower & upper expectation as lower & upper envelopes

$$\underline{E}(f) := \inf_{p \in \mathcal{C}} E_p(f) \qquad \overline{E}(f) := \sup_{p \in \mathcal{C}} E_p(f)$$

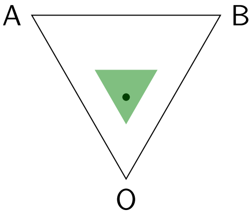
Envelope calculation as linear optimization over the credal set

$$\underline{E}(f) = \inf_{p \in \mathcal{C}} E_p(f) = \inf_{p \in \mathcal{C}} \sum_{x \in \mathcal{X}} p_x f(x)$$

For finite or closed credal sets $\mathcal{C}$, inf/sup becomes min/max

Example setup

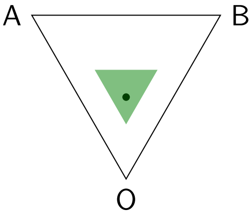
- linear-vacuous credal set $\mathcal{C}^{p,\varepsilon}$
- mixture coefficient $\varepsilon = \frac{1}{3}$
- pmf
  $p = (p_\mathrm{A}, p_\mathrm{B}, p_\mathrm{O}) = (\frac{1}{4}, \frac{1}{4}, \frac{2}{4})$



- function
  $f = (f(\mathrm{A}), f(\mathrm{B}), f(\mathrm{O})) = (1, 0, -1)$

Example setup

- linear-vacuous credal set $\mathcal{C}^{p,\varepsilon}$
- mixture coefficient $\varepsilon = \frac{1}{3}$
- pmf
  $p = (p_{\mathrm{A}}, p_{\mathrm{B}}, p_{\mathrm{O}}) = (\frac{1}{4}, \frac{1}{4}, \frac{2}{4})$



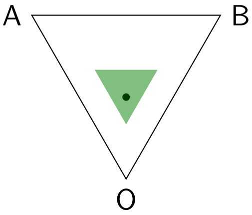
- function
  $f = (f(\mathrm{A}), f(\mathrm{B}), f(\mathrm{O})) = (1, 0, -1)$

Inferences

- Lower probability
- Lower expectation

Example setup

- linear-vacuous credal set $\mathcal{C}^{p,\varepsilon}$
- mixture coefficient $\varepsilon = \frac{1}{3}$
- pmf
  $p = (p_\mathrm{A}, p_\mathrm{B}, p_\mathrm{O}) = (\frac{1}{4}, \frac{1}{4}, \frac{2}{4})$



- function
  $f = (f(\mathrm{A}), f(\mathrm{B}), f(\mathrm{O})) = (1, 0, -1)$

Inferences

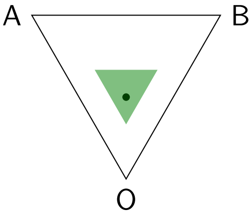
- Lower probability

$\underline{P}(\{\mathrm{B}, \mathrm{O}\})$

- Lower expectation

Example setup

- linear-vacuous credal set $\mathcal{C}^{p,\varepsilon}$
- mixture coefficient $\varepsilon = \frac{1}{3}$
- pmf
  $p = (p_\mathrm{A}, p_\mathrm{B}, p_\mathrm{O}) = (\frac{1}{4}, \frac{1}{4}, \frac{2}{4})$



- function
  $f = (f(\mathrm{A}), f(\mathrm{B}), f(\mathrm{O})) = (1, 0, -1)$

Inferences

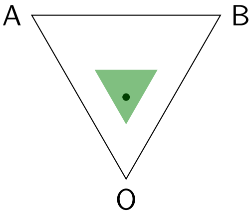
- Lower probability

$$\underline{P}(\{\mathrm{B}, \mathrm{O}\}) = \min_{q \in \mathcal{P}_X} \left(\frac{2}{3}(p_\mathrm{B} + p_\mathrm{O}) + \frac{1}{3}(q_\mathrm{B} + q_\mathrm{O})\right)$$

- Lower expectation

Example setup

- linear-vacuous credal set $\mathcal{C}^{p,\varepsilon}$
- mixture coefficient $\varepsilon = \frac{1}{3}$
- pmf
  $p = (p_A, p_B, p_O) = (\frac{1}{4}, \frac{1}{4}, \frac{2}{4})$



- function
  $f = (f(A), f(B), f(O)) = (1, 0, -1)$

Inferences

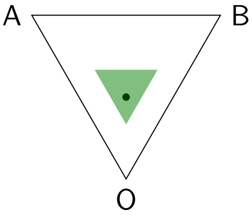
- Lower probability

$$\begin{aligned}\underline{P}(\{B, O\}) &= \min_{q \in \mathcal{P}_X} \Big(\tfrac{2}{3}(p_B + p_O) + \tfrac{1}{3}(q_B + q_O)\Big) \\ &= \tfrac{2}{3}\tfrac{3}{4} + \tfrac{1}{3}0 \\ &= \tfrac{1}{2}\end{aligned}$$

- Lower expectation

Example setup

- linear-vacuous credal set $\mathcal{C}^{p,\varepsilon}$
- mixture coefficient $\varepsilon = \frac{1}{3}$
- pmf
  $p = (p_\mathrm{A}, p_\mathrm{B}, p_\mathrm{O}) = (\frac{1}{4}, \frac{1}{4}, \frac{2}{4})$



- function
  $f = (f(\mathrm{A}), f(\mathrm{B}), f(\mathrm{O})) = (1, 0, -1)$

Inferences

- Lower probability

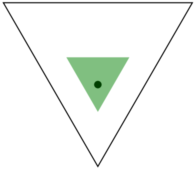
$$\underline{P}(\{\mathrm{B},\mathrm{O}\}) = \min_{q\in\mathcal{P}_X} \left(\frac{2}{3}(p_\mathrm{B} + p_\mathrm{O}) + \frac{1}{3}(q_\mathrm{B} + q_\mathrm{O})\right) = \frac{2}{3}\frac{3}{4} + \frac{1}{3}0 = \frac{1}{2}$$

- Lower expectation

$\underline{E}(f)$

Example setup

- linear-vacuous credal set $\mathcal{C}^{p,\varepsilon}$
- mixture coefficient $\varepsilon = \frac{1}{3}$
- pmf
  $p = (p_\mathrm{A}, p_\mathrm{B}, p_\mathrm{O}) = (\frac{1}{4}, \frac{1}{4}, \frac{2}{4})$



- function
  $f = (f(\mathrm{A}), f(\mathrm{B}), f(\mathrm{O})) = (1, 0, -1)$

Inferences

- Lower probability

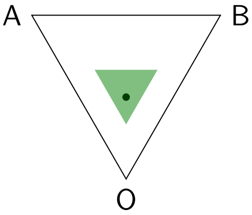
$$\underline{P}(\{\mathrm{B},\mathrm{O}\}) = \min_{q\in\mathcal{P}_X}\Big(\frac{2}{3}(p_\mathrm{B}+p_\mathrm{O}) + \frac{1}{3}(q_\mathrm{B}+q_\mathrm{O})\Big) = \frac{2}{3}\frac{3}{4} + \frac{1}{3}0 = \frac{1}{2}$$

- Lower expectation

$$\underline{E}(f) = \min_{q\in\mathcal{P}_X}\Big(\frac{2}{3}\sum_{x\in\mathcal{X}} p_x f(x) + \frac{1}{3}\sum_{x\in\mathcal{X}} q_x f(x)\Big) = \min_{q\in\mathcal{P}_X}\Big(\frac{2}{3}(p_\mathrm{A}-p_\mathrm{O}) + \frac{1}{3}(q_\mathrm{A}-q_\mathrm{O})\Big)$$

Example setup

- linear-vacuous credal set $\mathcal{C}^{p,\varepsilon}$
- mixture coefficient $\varepsilon = \frac{1}{3}$
- pmf
  $p = (p_\mathrm{A}, p_\mathrm{B}, p_\mathrm{O}) = (\frac{1}{4}, \frac{1}{4}, \frac{2}{4})$



- function
  $f = (f(\mathrm{A}), f(\mathrm{B}), f(\mathrm{O})) = (1, 0, -1)$

Inferences

- Lower probability

$$\begin{aligned}\underline{P}(\{\mathrm{B},\mathrm{O}\}) &= \min_{q\in\mathcal{P}_X}\Big(\tfrac{2}{3}(p_\mathrm{B}+p_\mathrm{O}) + \tfrac{1}{3}(q_\mathrm{B}+q_\mathrm{O})\Big) \\ &= \tfrac{2}{3}\tfrac{3}{4} + \tfrac{1}{3}0 \\ &= \tfrac{1}{2}\end{aligned}$$

- Lower expectation

$$\begin{aligned}\underline{E}(f) &= \min_{q\in\mathcal{P}_X}\Big(\tfrac{2}{3}\sum_{x\in\mathcal{X}} p_x f(x) + \tfrac{1}{3}\sum_{x\in\mathcal{X}} q_x f(x)\Big) \\ &= \min_{q\in\mathcal{P}_X}\Big(\tfrac{2}{3}(p_\mathrm{A}-p_\mathrm{O}) + \tfrac{1}{3}(q_\mathrm{A}-q_\mathrm{O})\Big) \\ &= \tfrac{2}{3}(-\tfrac{1}{4}) + \tfrac{1}{3}(-1) \\ &= -\tfrac{1}{2}\end{aligned}$$

Lower & upper probability as lower & upper envelopes

$$\underline{P}(S) := \inf_{p \in \mathcal{C}} P_p(S) \qquad \overline{P}(S) := \sup_{p \in \mathcal{C}} P_p(S)$$

Lower & upper expectation as lower & upper envelopes

$$\underline{E}(f) := \inf_{p \in \mathcal{C}} E_p(f) \qquad \overline{E}(f) := \sup_{p \in \mathcal{C}} E_p(f)$$

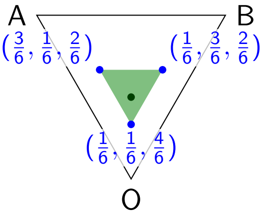
Envelope calculation as linear optimization over the credal *set's extreme points*

$$\underline{E}(f) = \inf_{p \in \mathcal{C}} E_p(f) = \inf_{p \in \mathcal{C}} \sum_{x \in \mathcal{X}} p_x f(x) = \inf_{p \in \operatorname{ext} \mathcal{C}} \sum_{x \in \mathcal{X}} p_x f(x)$$

(Extreme points exist for closed credal sets and fully characterize convex ones.)

Example setup

- linear-vacuous credal set $\mathcal{C}^{p,\varepsilon}$
- mixture coefficient $\varepsilon = \frac{1}{3}$
- pmf
  $p = (p_\mathrm{A}, p_\mathrm{B}, p_\mathrm{O}) = (\frac{1}{4}, \frac{1}{4}, \frac{2}{4})$



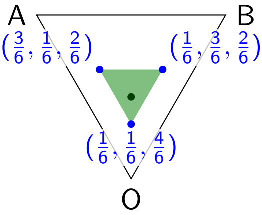
- function
  $f = (f(\mathrm{A}), f(\mathrm{B}), f(\mathrm{O})) = (1, 0, -1)$

Inferences

- Lower probability
- Lower expectation

Example setup

- linear-vacuous credal set $\mathcal{C}^{p,\varepsilon}$
- mixture coefficient $\varepsilon = \frac{1}{3}$
- pmf
  $p = (p_{\mathrm{A}}, p_{\mathrm{B}}, p_{\mathrm{O}}) = (\frac{1}{4}, \frac{1}{4}, \frac{2}{4})$



- function
  $f = (f(\mathrm{A}), f(\mathrm{B}), f(\mathrm{O})) = (1, 0, -1)$

Inferences

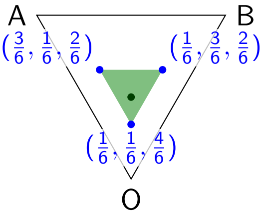
- Lower probability

$$\underline{P}(\{\mathrm{B}, \mathrm{O}\})$$

- Lower expectation

Example setup

- linear-vacuous credal set $\mathcal{C}^{p,\varepsilon}$
- mixture coefficient $\varepsilon = \frac{1}{3}$
- pmf
  $p = (p_\mathrm{A}, p_\mathrm{B}, p_\mathrm{O}) = (\frac{1}{4}, \frac{1}{4}, \frac{2}{4})$



- function
  $f = (f(\mathrm{A}), f(\mathrm{B}), f(\mathrm{O})) = (1, 0, -1)$

Inferences

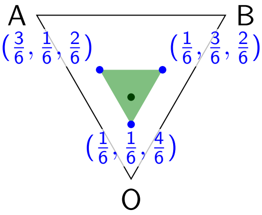
- Lower probability

$$\underline{P}(\{\mathrm{B},\mathrm{O}\}) = \min_{q \in \mathrm{ext}\,\mathcal{C}^{p,\varepsilon}} (q_\mathrm{B} + q_\mathrm{O})$$

- Lower expectation

Example setup

- linear-vacuous credal set $\mathcal{C}^{p,\varepsilon}$
- mixture coefficient $\varepsilon = \frac{1}{3}$
- pmf
  $p = (p_\mathrm{A}, p_\mathrm{B}, p_\mathrm{O}) = (\frac{1}{4}, \frac{1}{4}, \frac{2}{4})$



- function
  $f = (f(\mathrm{A}), f(\mathrm{B}), f(\mathrm{O})) = (1, 0, -1)$

Inferences

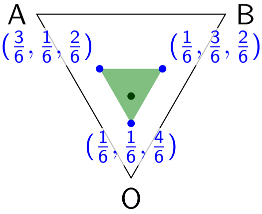
- Lower probability

$$\underline{P}(\{\mathrm{B},\mathrm{O}\}) = \min_{q \in \mathrm{ext}\,\mathcal{C}^{p,\varepsilon}} (q_\mathrm{B} + q_\mathrm{O}) = \min\{\tfrac{3}{6}, \tfrac{5}{6}, \tfrac{5}{6}\} = \tfrac{3}{6} = \tfrac{1}{2}$$

- Lower expectation

Example setup

- linear-vacuous credal set $\mathcal{C}^{p,\varepsilon}$
- mixture coefficient $\varepsilon = \frac{1}{3}$
- pmf
  $p = (p_\mathrm{A}, p_\mathrm{B}, p_\mathrm{O}) = (\frac{1}{4}, \frac{1}{4}, \frac{2}{4})$



- function
  $f = (f(\mathrm{A}), f(\mathrm{B}), f(\mathrm{O})) = (1, 0, -1)$

Inferences

- Lower probability

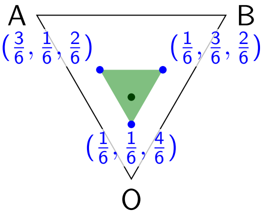
$$\underline{P}(\{\mathrm{B}, \mathrm{O}\}) = \min_{q \in \mathrm{ext}\,\mathcal{C}^{p,\varepsilon}} (q_\mathrm{B} + q_\mathrm{O}) = \min\{\tfrac{3}{6}, \tfrac{5}{6}, \tfrac{5}{6}\} = \tfrac{3}{6} = \tfrac{1}{2}$$

- Lower expectation

$$\underline{E}(f)$$

Example setup

- linear-vacuous credal set $\mathcal{C}^{p,\varepsilon}$
- mixture coefficient $\varepsilon = \frac{1}{3}$
- pmf
  $p = (p_\mathrm{A}, p_\mathrm{B}, p_\mathrm{O}) = (\frac{1}{4}, \frac{1}{4}, \frac{2}{4})$



- function
  $f = (f(\mathrm{A}), f(\mathrm{B}), f(\mathrm{O})) = (1, 0, -1)$

Inferences

- Lower probability

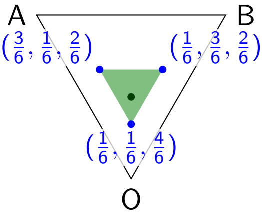
$$\underline{P}(\{\mathrm{B}, \mathrm{O}\}) = \min_{q \in \operatorname{ext} \mathcal{C}^{p,\varepsilon}} (q_\mathrm{B} + q_\mathrm{O}) = \min\{\tfrac{3}{6}, \tfrac{5}{6}, \tfrac{5}{6}\} = \tfrac{3}{6} = \tfrac{1}{2}$$

- Lower expectation

$$\underline{E}(f) = \min_{q \in \operatorname{ext} \mathcal{C}^{p,\varepsilon}} \sum_{x \in \mathcal{X}} q_x f(x) = \min_{q \in \operatorname{ext} \mathcal{C}^{p,\varepsilon}} (q_\mathrm{A} - q_\mathrm{O})$$

Example setup

- linear-vacuous credal set $\mathcal{C}^{p,\varepsilon}$
- mixture coefficient $\varepsilon = \frac{1}{3}$
- pmf
  $p = (p_\mathrm{A}, p_\mathrm{B}, p_\mathrm{O}) = (\frac{1}{4}, \frac{1}{4}, \frac{2}{4})$



- function
  $f = (f(\mathrm{A}), f(\mathrm{B}), f(\mathrm{O})) = (1, 0, -1)$

Inferences

- Lower probability

$$\underline{P}(\{\mathrm{B},\mathrm{O}\}) = \min_{q \in \mathrm{ext}\, \mathcal{C}^{p,\varepsilon}} (q_\mathrm{B} + q_\mathrm{O}) = \min\{\tfrac{3}{6}, \tfrac{5}{6}, \tfrac{5}{6}\} = \tfrac{3}{6} = \tfrac{1}{2}$$

- Lower expectation

$$\underline{E}(f) = \min_{q \in \mathrm{ext}\, \mathcal{C}^{p,\varepsilon}} \sum_{x \in \mathcal{X}} q_x f(x) = \min_{q \in \mathrm{ext}\, \mathcal{C}^{p,\varepsilon}} (q_\mathrm{A} - q_\mathrm{O}) = \min\{\tfrac{1}{6}, -\tfrac{1}{6}, -\tfrac{3}{6}\} = -\tfrac{3}{6} = -\tfrac{1}{2}$$

Multivariate credal sets: basic setup & basic idea

- Index set $N = \{1, \ldots, n\}$
- Multivariate variable $\boldsymbol{X} = (X_1, X_2, \ldots, X_n)$
- Set of possible outcomes $\boldsymbol{x} \in \mathcal{X} = \mathcal{X}_1 \times \mathcal{X}_2 \times \ldots \times \mathcal{X}_n$
- A **joint credal set** $\mathcal{C}^{\boldsymbol{X}}$ of joint probability mass functions for $\boldsymbol{X}$ is specified

Apply probabilistic operations pointwise
to the elements of the credal sets involved

Marginal credal sets

- A marginal credal set is defined for a subset of the random variables
- Let $K \subseteq N$, then
  - $\boldsymbol{X}_K := (X_k : k \in K)$ and $\boldsymbol{x}_K \in \mathcal{X}_K := \times_{k \in K} \mathcal{X}_k$
  - Notation: $p^{\boldsymbol{X}_K}$ is the $\boldsymbol{X}_K$-marginal of the joint pmf $p^{\boldsymbol{X}}$

Marginal credal sets

- A marginal credal set is defined for a subset of the random variables
- Let $K \subseteq N$, then
  - $\boldsymbol{X}_K := (X_k : k \in K)$ and $\boldsymbol{x}_K \in \mathcal{X}_K := \times_{k \in K} \mathcal{X}_k$
  - Notation: $p^{\boldsymbol{X}_K}$ is the $\boldsymbol{X}_K$-marginal of the joint pmf $p^{\boldsymbol{X}}$
- A **marginal credal set** is obtained by marginalizing each of its member pmfs:

$$\mathcal{C}^{\boldsymbol{X}_K} = \left\{ p^{\boldsymbol{X}_K} : p^{\boldsymbol{X}} \in \mathcal{C}^{\boldsymbol{X}} \right\}$$

Marginal credal sets example

Problem setup

- Two random variables X_1 and X_2 with outcome spaces $\mathcal{X}_1 = \{0, 1\}$ and $\mathcal{X}_2 = \{-, +\}$
- Joint credal set $\mathcal{C}^{\boldsymbol{X}} := \mathcal{C}^{p,\varepsilon}$ with ε unspecified and p given below (in black), together with its marginals (in green)

p	0	1	p^{X_2}
$-$	$3/9$	$1/9$	$4/9$
$+$	$2/9$	$3/9$	$5/9$
p^{X_1}	$5/9$	$4/9$	

Marginal credal sets example

Problem setup

- Two random variables X_1 and X_2 with outcome spaces $\mathcal{X}_1 = \{0, 1\}$ and $\mathcal{X}_2 = \{-, +\}$
- Joint credal set $\mathcal{C}^{\boldsymbol{X}} := \mathcal{C}^{p,\varepsilon}$ with ε unspecified and p given below (in black), together with its marginals (in green)

p	0	1	p^{X_2}
$-$	$3/9$	$1/9$	$4/9$
$+$	$2/9$	$3/9$	$5/9$
p^{X_1}	$5/9$	$4/9$	

Inference

Marginal credal set $\mathcal{C}^{X_1}$

Marginal credal sets example

Problem setup

- Two random variables X_1 and X_2 with outcome spaces $\mathcal{X}_1 = \{0, 1\}$ and $\mathcal{X}_2 = \{-, +\}$
- Joint credal set $\mathcal{C}^{\boldsymbol{X}} := \mathcal{C}^{p,\varepsilon}$ with ε unspecified and p given below (in black), together with its marginals (in green)

p	0	1	p^{X_2}
$-$	$3/9$	$1/9$	$4/9$
$+$	$2/9$	$3/9$	$5/9$
p^{X_1}	$5/9$	$4/9$	

Inference

Marginal credal set $\mathcal{C}^{X_1}$

Solution

- Marginalize element $q = (1-\varepsilon)p + \varepsilon r$ of $\mathcal{C}^{\boldsymbol{X}}$, where $r \in \mathcal{P}_{\boldsymbol{X}}$:

$$q^{X_1}_{x_1} = \sum_{x_2 \in \mathcal{X}_2} q_{(x_1,x_2)} = (1-\varepsilon)p^{X_1}_{x_1} + \varepsilon r^{X_1}_{x_1}$$

Marginal credal sets example

Problem setup

- Two random variables X_1 and X_2 with outcome spaces $\mathcal{X}_1 = \{0, 1\}$ and $\mathcal{X}_2 = \{-, +\}$
- Joint credal set $\mathcal{C}^{\boldsymbol{X}} := \mathcal{C}^{p,\varepsilon}$ with ε unspecified and p given below (in black), together with its marginals (in green)

p	0	1	p^{X_2}
$-$	$3/9$	$1/9$	$4/9$
$+$	$2/9$	$3/9$	$5/9$
p^{X_1}	$5/9$	$4/9$	

Inference

Marginal credal set $\mathcal{C}^{X_1}$

Solution

- Marginalize element $q = (1-\varepsilon)p + \varepsilon r$ of $\mathcal{C}^{\boldsymbol{X}}$, where $r \in \mathcal{P}_{\boldsymbol{X}}$:

$$q^{X_1}_{x_1} = \sum_{x_2 \in \mathcal{X}_2} q_{(x_1,x_2)} = (1-\varepsilon)p^{X_1}_{x_1} + \varepsilon r^{X_1}_{x_1}$$

- Gather all marginalized elements:

$$\mathcal{C}^{X_1} = \mathcal{C}^{p^{X_1},\varepsilon}$$

because r^{X_1} ranges over $\mathcal{P}_{X_1}$ as r ranges over $\mathcal{P}_{\boldsymbol{X}}$

Marginal credal sets example

Problem setup

- Two random variables X_1 and X_2 with outcome spaces $\mathcal{X}_1 = \{0, 1\}$ and $\mathcal{X}_2 = \{-, +\}$
- Joint credal set $\mathcal{C}^{\boldsymbol{X}} := \mathcal{C}^{p,\varepsilon}$ with ε unspecified and p given below (in black), together with its marginals (in green)

p	0	1	p^{X_2}
$-$	$3/9$	$1/9$	$4/9$
$+$	$2/9$	$3/9$	$5/9$
p^{X_1}	$5/9$	$4/9$	

Inference

Marginal credal set $\mathcal{C}^{X_1}$

Solution

- Marginalize element $q = (1-\varepsilon)p + \varepsilon r$ of $\mathcal{C}^{\boldsymbol{X}}$, where $r \in \mathcal{P}_{\boldsymbol{X}}$:

$$q^{X_1}_{x_1} = \sum_{x_2 \in \mathcal{X}_2} q_{(x_1,x_2)} = (1-\varepsilon)p^{X_1}_{x_1} + \varepsilon r^{X_1}_{x_1}$$

q	0	1
$-$	$(1-\varepsilon)3/9 + \varepsilon r_{(0,-)}$	$(1-\varepsilon)1/9 + \varepsilon r_{(1,-)}$
$+$	$(1-\varepsilon)2/9 + \varepsilon r_{(0,+)}$	$(1-\varepsilon)3/9 + \varepsilon r_{(1,+)}$
q^{X_1}	$(1-\varepsilon)5/9 + \varepsilon r^{X_1}_0$	$(1-\varepsilon)4/9 + \varepsilon r^{X_1}_1$

Conditional credal sets

- A conditional credal set is determined by
  - assuming some event is known to be true, or specifically
  - assuming some random variables take some given values: $\boldsymbol{X}_K = \boldsymbol{x}_K$, with $K \subset N$
- Notation: $p^{\boldsymbol{X}_{N \setminus K} | \boldsymbol{x}_K}$ is the $\boldsymbol{X}_K$-conditional of the joint pmf $p^{\boldsymbol{X}}$

Conditional credal sets

- A conditional credal set is determined by
  - assuming some event is known to be true, or specifically
  - assuming some random variables take some given values: $\boldsymbol{X}_K = \boldsymbol{x}_K$, with $K \subset N$

- Notation: $p^{\boldsymbol{X}_{N\setminus K}|\boldsymbol{x}_K}$ is the $\boldsymbol{X}_K$-conditional of the joint pmf $p^{\boldsymbol{X}}$

- A **conditional credal set** is obtained by conditioning each of its elements *for which this operation is defined*:

$$\mathcal{C}^{\boldsymbol{X}_{N\setminus K}|\boldsymbol{x}_K} = \left\{ p^{\boldsymbol{X}_{N\setminus K}|\boldsymbol{x}_K} : \left(p^{\boldsymbol{X}} \in \mathcal{C}^{\boldsymbol{X}} \wedge p^{\boldsymbol{X}_K}(\boldsymbol{x}_K) > 0 \right) \right\}$$

Conditional credal sets

- A conditional credal set is determined by
  - assuming some event is known to be true, or specifically
  - assuming some random variables take some given values: $\boldsymbol{X}_K = \boldsymbol{x}_K$, with $K \subset N$
- Notation: $p^{\boldsymbol{X}_{N \setminus K} | \boldsymbol{x}_K}$ is the $\boldsymbol{X}_K$-conditional of the joint pmf $p^{\boldsymbol{X}}$
- A **conditional credal set** is obtained by conditioning each of its elements *for which this operation is defined*:

$$\mathcal{C}^{\boldsymbol{X}_{N \setminus K} | \boldsymbol{x}_K} = \left\{ p^{\boldsymbol{X}_{N \setminus K} | \boldsymbol{x}_K} : \left(p^{\boldsymbol{X}} \in \mathcal{C}^{\boldsymbol{X}} \wedge p^{\boldsymbol{X}_K}(\boldsymbol{x}_K) > 0 \right) \right\}$$

This is called *conditioning by regular extension*

Conditional credal sets example

Problem setup

- Two random variables X_1 and X_2 with outcome spaces $\mathcal{X}_1 = \{0, 1\}$ and $\mathcal{X}_2 = \{-, +\}$
- Joint credal set $\mathcal{C}^{\boldsymbol{X}} := \mathcal{C}^{\Gamma}$ with elements q given below (in black), together with its marginals (in green), with $\gamma \in \Gamma = [-\frac{3}{9}, \frac{3}{9}]$

q	0	1	q^{X_2}
$-$	$3/9 - \gamma$	$1/9$	$4/9 - \gamma$
$+$	$2/9$	$3/9 + \gamma$	$5/9 + \gamma$
q^{X_1}	$5/9 - \gamma$	$4/9 + \gamma$	

Inference

Conditional credal set $\mathcal{C}^{X_1|+}$

Conditional credal sets example

Problem setup

- Two random variables X_1 and X_2 with outcome spaces $\mathcal{X}_1 = \{0, 1\}$ and $\mathcal{X}_2 = \{-, +\}$
- Joint credal set $\mathcal{C}^{\boldsymbol{X}} := \mathcal{C}^{\Gamma}$ with elements q given below (in black), together with its marginals (in green), with $\gamma \in \Gamma = [-\frac{3}{9}, \frac{3}{9}]$

q	0	1	q^{X_2}
$-$	$3/9 - \gamma$	$1/9$	$4/9 - \gamma$
$+$	$2/9$	$3/9 + \gamma$	$5/9 + \gamma$
q^{X_1}	$5/9 - \gamma$	$4/9 + \gamma$	

Inference

Conditional credal set $\mathcal{C}^{X_1|+}$

Solution

- Condition elements q of $\mathcal{C}^{\boldsymbol{X}}$:

$$q^{X_1|+} = \frac{q_{(X_1,+)}}{q^{X_2}_{+}} = \frac{(2/9, 3/9 + \gamma)}{5/9 + \gamma} = (r, 1 - r),$$

with $r = \frac{2/9}{5/9+\gamma}$

Conditional credal sets example

Problem setup

- Two random variables X_1 and X_2 with outcome spaces $\mathcal{X}_1 = \{0, 1\}$ and $\mathcal{X}_2 = \{-, +\}$
- Joint credal set $\mathcal{C}^{\boldsymbol{X}} := \mathcal{C}^\Gamma$ with elements q given below (in black), together with its marginals (in green), with $\gamma \in \Gamma = [-\frac{3}{9}, \frac{3}{9}]$

q	0	1	q^{X_2}
$-$	$3/9 - \gamma$	$1/9$	$4/9 - \gamma$
$+$	$2/9$	$3/9 + \gamma$	$5/9 + \gamma$
q^{X_1}	$5/9 - \gamma$	$4/9 + \gamma$	

Inference

Conditional credal set $\mathcal{C}^{X_1|+}$

Solution

- Condition elements q of $\mathcal{C}^{\boldsymbol{X}}$:

$$q^{X_1|+} = \frac{q_{(X_1,+)}}{q^{X_2}_+} = \frac{(2/9, 3/9 + \gamma)}{5/9 + \gamma} = (r, 1 - r),$$

with $r = \frac{2/9}{5/9+\gamma}$

- Gather all conditioned elements:

$$\mathcal{C}^{X_1|+} = \left\{(r, 1 - r) : r \in [\tfrac{1}{4}, 1]\right\}$$

Conditional credal sets example

Problem setup

- Two random variables X_1 and X_2 with outcome spaces $\mathcal{X}_1 = \{0, 1\}$ and $\mathcal{X}_2 = \{-, +\}$
- Joint credal set $\mathcal{C}^{\boldsymbol{X}} := \mathcal{C}^{\Gamma}$ with elements q given below (in black), together with its marginals (in green), with $\gamma \in \Gamma = [-\frac{3}{9}, \frac{3}{9}]$

q	0	1	q^{X_2}
$-$	$3/9 - \gamma$	$1/9$	$4/9 - \gamma$
$+$	$2/9$	$3/9 + \gamma$	$5/9 + \gamma$
q^{X_1}	$5/9 - \gamma$	$4/9 + \gamma$	

Inference

Conditional credal set $\mathcal{C}^{X_1|+}$

Solution

- Condition elements q of $\mathcal{C}^{\boldsymbol{X}}$:

$$q^{X_1|+} = \frac{q_{(X_1,+)}}{q_+^{X_2}} = \frac{(2/9, 3/9 + \gamma)}{5/9 + \gamma} = (r, 1 - r),$$

with $r = \frac{2/9}{5/9 + \gamma}$

- Gather all conditioned elements:

$$\mathcal{C}^{X_1|+} = \left\{(r, 1 - r) : r \in [\tfrac{1}{4}, 1]\right\}$$

(Exercise: show that $\mathcal{C}^{X_1|+} = \mathcal{C}^{(1,0),\frac{3}{4}}$)

Complete independence for credal sets

- Different generalizations of (stochastic) independence to credal sets are possible
- Here, we only consider the independence concept
  associated to the 'sets of Bayesians' interpretation
- Notation: $p \otimes q$ denotes the independent product of two pmfs p and q

Complete independence for credal sets

- Different generalizations of (stochastic) independence to credal sets are possible
- Here, we only consider the independence concept associated to the 'sets of Bayesians' interpretation
- Notation: $p \otimes q$ denotes the independent product of two pmfs p and q

Definition of complete independence

Consider $\boldsymbol{X} = (X_1, X_2)$; the random variables X_1 and X_2 are *completely independent* if they are stochastically independent for each of the pmfs in the joint credal set $\mathcal{C}^{\boldsymbol{X}}$:

$$\mathcal{C}^{\boldsymbol{X}} \subseteq \mathcal{C}^{X_1} \otimes \mathcal{C}^{X_2} := \left\{ p^{X_1} \otimes p^{X_2} : p^{X_1} \in \mathcal{C}^{X_1}, p^{X_2} \in \mathcal{C}^{X_2} \right\}$$

Complete independence for credal sets

- Different generalizations of (stochastic) independence to credal sets are possible
- Here, we only consider the independence concept associated to the 'sets of Bayesians' interpretation
- Notation: $p \otimes q$ denotes the independent product of two pmfs p and q

Definition of complete independence

Consider $\boldsymbol{X} = (X_1, X_2)$; the random variables X_1 and X_2 are *completely independent* if they are stochastically independent for each of the pmfs in the joint credal set $\mathcal{C}^{\boldsymbol{X}}$:

$$\mathcal{C}^{\boldsymbol{X}} \subseteq \mathcal{C}^{X_1} \otimes \mathcal{C}^{X_2} := \left\{ p^{X_1} \otimes p^{X_2} : p^{X_1} \in \mathcal{C}^{X_1}, p^{X_2} \in \mathcal{C}^{X_2} \right\}$$

Not convex even if $\mathcal{C}^{X_1}$ and $\mathcal{C}^{X_2}$ are!

Complete independence example

Problem setup

- Random variables X_1 and X_2 with outcome spaces $\mathcal{X}_1 = \{0, 1\}$ and $\mathcal{X}_2 = \{-, +\}$
- Marginal credal sets $\mathcal{C}^{X_1} := \mathcal{C}^{(\frac{1}{2}, \frac{1}{2}), \frac{1}{2}}$ and $\mathcal{C}^{X_2} := \mathcal{C}^{(\frac{1}{3}, \frac{2}{3}), \frac{1}{3}}$

Complete independence example

Problem setup

- Random variables X_1 and X_2 with outcome spaces $\mathcal{X}_1 = \{0, 1\}$ and $\mathcal{X}_2 = \{-, +\}$
- Marginal credal sets $\mathcal{C}^{X_1} := \mathcal{C}^{(\frac{1}{2},\frac{1}{2}),\frac{1}{2}}$ and $\mathcal{C}^{X_2} := \mathcal{C}^{(\frac{1}{3},\frac{2}{3}),\frac{1}{3}}$

Goal

Show that their completely independent joint credal set is not convex

Complete independence example

Problem setup

- Random variables X_1 and X_2 with outcome spaces $\mathcal{X}_1 = \{0, 1\}$ and $\mathcal{X}_2 = \{-, +\}$
- Marginal credal sets $\mathcal{C}^{X_1} := \mathcal{C}^{(\frac{1}{2},\frac{1}{2}),\frac{1}{2}}$ and $\mathcal{C}^{X_2} := \mathcal{C}^{(\frac{1}{3},\frac{2}{3}),\frac{1}{3}}$

Goal

Show that their completely independent joint credal set is not convex

Demonstration

- We construct a counterexample to convexity, i.e., we construct a convex mixture of elements of the joint that lies outside the joint because it does not factorize

Complete independence example

Problem setup

- Random variables X_1 and X_2 with outcome spaces $\mathcal{X}_1 = \{0, 1\}$ and $\mathcal{X}_2 = \{-, +\}$
- Marginal credal sets $\mathcal{C}^{X_1} := \mathcal{C}^{(\frac{1}{2},\frac{1}{2}),\frac{1}{2}}$ and $\mathcal{C}^{X_2} := \mathcal{C}^{(\frac{1}{3},\frac{2}{3}),\frac{1}{3}}$

Goal

Show that their completely independent joint credal set is not convex

Demonstration

- We construct a counterexample to convexity, i.e., we construct a convex mixture of elements of the joint that lies outside the joint because it does not factorize
- An element q of $\mathcal{C}^{\boldsymbol{X}} = \mathcal{C}^{X_1} \otimes \mathcal{C}^{X_2}$ is defined for every $r^{X_1} \in \mathcal{P}_{X_1}$ and $r^{X_2} \in \mathcal{P}_{X_2}$ by

$$q_{(x_1,x_2)} = (\tfrac{1}{2} p^{X_1}_{x_1} + \tfrac{1}{2} r^{X_1}_{x_1})(\tfrac{2}{3} p^{X_2}_{x_2} + \tfrac{1}{3} r^{X_2}_{x_2})$$

Complete independence example

Problem setup

- Random variables X_1 and X_2 with outcome spaces $\mathcal{X}_1 = \{0, 1\}$ and $\mathcal{X}_2 = \{-, +\}$
- Marginal credal sets $\mathcal{C}^{X_1} := \mathcal{C}^{(\frac{1}{2},\frac{1}{2}),\frac{1}{2}}$ and $\mathcal{C}^{X_2} := \mathcal{C}^{(\frac{1}{3},\frac{2}{3}),\frac{1}{3}}$

Goal

Show that their completely independent joint credal set is not convex

Demonstration

$r_0^{X_1} = r_-^{X_2} = 0$

	0	1	q^{X_2}
$-$			$2/9$
$+$			$7/9$
q^{X_1}	$1/4$	$3/4$	

$r_0^{X_1} = r_-^{X_2} = 1$

	0	1	q^{X_2}
$-$			$5/9$
$+$			$4/9$
q^{X_1}	$3/4$	$1/4$	

Complete independence example

Problem setup

- Random variables X_1 and X_2 with outcome spaces $\mathcal{X}_1 = \{0, 1\}$ and $\mathcal{X}_2 = \{-, +\}$
- Marginal credal sets $\mathcal{C}^{X_1} := \mathcal{C}^{(\frac{1}{2},\frac{1}{2}),\frac{1}{2}}$ and $\mathcal{C}^{X_2} := \mathcal{C}^{(\frac{1}{3},\frac{2}{3}),\frac{1}{3}}$

Goal

Show that their completely independent joint credal set is not convex

Demonstration

$r_0^{X_1} = r_-^{X_2} = 0$

	0	1	q^{X_2}
$-$	2/36	6/36	2/9
$+$	7/36	21/36	7/9
q^{X_1}	1/4	3/4	

$r_0^{X_1} = r_-^{X_2} = 1$

	0	1	q^{X_2}
$-$	15/36	5/36	5/9
$+$	12/36	4/36	4/9
q^{X_1}	3/4	1/4	

Complete independence example

Problem setup

- Random variables X_1 and X_2 with outcome spaces $\mathcal{X}_1 = \{0, 1\}$ and $\mathcal{X}_2 = \{-, +\}$
- Marginal credal sets $\mathcal{C}^{X_1} := \mathcal{C}^{(\frac{1}{2},\frac{1}{2}),\frac{1}{2}}$ and $\mathcal{C}^{X_2} := \mathcal{C}^{(\frac{1}{3},\frac{2}{3}),\frac{1}{3}}$

Goal

Show that their completely independent joint credal set is not convex

Demonstration

$r_0^{X_1} = r_-^{X_2} = 0$

$\frac{1}{2}$

	0	1	q^{X_2}
$-$	$2/36$	$6/36$	$2/9$
$+$	$7/36$	$21/36$	$7/9$
q^{X_1}	$1/4$	$3/4$	

$+\ \frac{1}{2}$

$r_0^{X_1} = r_-^{X_2} = 1$

	0	1	q^{X_2}
$-$	$15/36$	$5/36$	$5/9$
$+$	$12/36$	$4/36$	$4/9$
q^{X_1}	$3/4$	$1/4$	

$=$

	0	1	q^{X_2}
$-$	$17/72$	$11/72$	
$+$	$19/72$	$25/72$	
q^{X_1}			

Complete independence example

Problem setup

- Random variables X_1 and X_2 with outcome spaces $\mathcal{X}_1 = \{0, 1\}$ and $\mathcal{X}_2 = \{-, +\}$
- Marginal credal sets $\mathcal{C}^{X_1} := \mathcal{C}^{(\frac{1}{2},\frac{1}{2}),\frac{1}{2}}$ and $\mathcal{C}^{X_2} := \mathcal{C}^{(\frac{1}{3},\frac{2}{3}),\frac{1}{3}}$

Goal

Show that their completely independent joint credal set is not convex

Demonstration

$r_0^{X_1} = r_-^{X_2} = 0$

$\frac{1}{2}$

	0	1	q^{X_2}
$-$	$2/36$	$6/36$	$2/9$
$+$	$7/36$	$21/36$	$7/9$
q^{X_1}	$1/4$	$3/4$	

$+ \frac{1}{2}$

$r_0^{X_1} = r_-^{X_2} = 1$

	0	1	q^{X_2}
$-$	$15/36$	$5/36$	$5/9$
$+$	$12/36$	$4/36$	$4/9$
q^{X_1}	$3/4$	$1/4$	

$=$

not factorizing

	0	1	q^{X_2}
$-$	$17/72$	$11/72$	$7/18$
$+$	$19/72$	$25/72$	$11/18$
q^{X_1}	$1/2$	$1/2$	

Overview

Kick-off (slot 1)

Classical probability theory (slot 1)

Interpretation of probability (slot 2)

Limitations of probability theory (slot 2)

Probability intervals (slot 3)

Credal sets (slot 3–4)

Acceptability & Desirability (slot 4–5)

- Representation
- Reasoning
- Axiom variants
- Multivariate acceptability

Interval expectation & probability (slot 5–6)

Gambles

Earlier appearance of 'gamble':

- In the betting game, indicators 1_x and 1_S
- In urn problems, scaled indicators $f_i \propto 1_S$
- In blood groups example, effectiveness descriptions f_1, f_2, positive functions on $\mathcal{X}$

Gambles

Earlier appearance of ‘gamble’:

- In the betting game, indicators 1_x and 1_S
- In urn problems, scaled indicators $f_i \propto 1_S$
- In blood groups example, effectiveness descriptions f_1, f_2, positive functions on $\mathcal{X}$

Exchanges

- In the betting game, differences between fair prices and gambles $p_x - 1_x = p_x \cdot 1_{\mathcal{X}} - 1_x$ and $P(S) - 1_S$
- In urn problems, preferences between gambles f_i and f_j can be seen as preference between their difference $f_i - f_j$ and the zero gamble $0 = 0 \cdot 1_{\mathcal{X}}$ (status quo)
- Differences of gambles are again gambles

Gambles

Earlier appearance of 'gamble':

- In the betting game, indicators 1_x and 1_S
- In urn problems, scaled indicators $f_i \propto 1_S$
- In blood groups example, effectiveness descriptions f_1, f_2, positive functions on $\mathcal{X}$

Exchanges

- In the betting game, differences between fair prices and gambles $p_x - 1_x = p_x \cdot 1_{\mathcal{X}} - 1_x$ and $P(S) - 1_S$
- In urn problems, preferences between gambles f_i and f_j can be seen as preference between their difference $f_i - f_j$ and the zero gamble $0 = 0 \cdot 1_{\mathcal{X}}$ (status quo)
- Differences of gambles are again gambles

Linear space of gambles

(Bounded) real-valued functions in $\mathcal{L} = (\mathcal{X} \to \mathbb{R})$, as foundation for representation
(or linear subspace of functions instead)

Basic setup, axioms, and terminology

Basic setup of **acceptability**:

- Random variable X
- Finite set of outcomes $\mathcal{X}$
- Each gamble $f \in \mathcal{L}$ is acceptable $(f \in \mathcal{D})$ or not $(f \notin \mathcal{D})$

Basic setup, axioms, and terminology

Basic setup of **acceptability**:

- Random variable X
- Finite set of outcomes $\mathcal{X}$
- Each gamble $f \in \mathcal{L}$ is acceptable ($f \in \mathcal{D}$) or not ($f \notin \mathcal{D}$)

Axioms

A **set of acceptable gambles** $\mathcal{D}$ is **coherent** when:

D1. if $f \lessdot 0$, then $f \notin \mathcal{D}$ (*avoiding sure loss*)

D2. if $f \geq 0$, then $f \in \mathcal{D}$ (*accepting nonnegative gain*)

D3. if $f \in \mathcal{D}$ and $\lambda \in \mathbb{R}_{>}$, then $\lambda \cdot f \in \mathcal{D}$ (*positive scaling*)

D4. if $f, g \in \mathcal{D}$, then $f + g \in \mathcal{D}$ (*combination*)

Gamble (vector) inequalities: $f \gtrdot 0$ iff $\min f > 0$, $f \geq 0$ iff $\min f \geq 0$, $f > 0$ iff $f \geq 0$ and $f \neq 0$

Basic setup, axioms, and terminology

Basic setup of **acceptability**:

- Random variable X
- Finite set of outcomes $\mathcal{X}$
- Each gamble $f \in \mathcal{L}$ is acceptable $(f \in \mathcal{D})$ or not $(f \notin \mathcal{D})$

Axioms

A **set of acceptable gambles** $\mathcal{D}$ is **coherent** when:

D1. $\mathcal{L}_{\lessdot} \cap \mathcal{D} = \emptyset$ (*avoiding sure loss*)

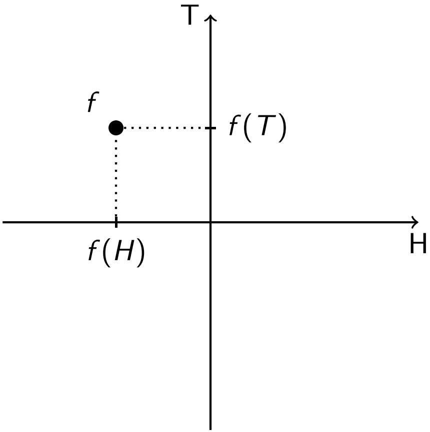
D2. $\mathcal{L}_{\geq} \subseteq \mathcal{D}$ (*accepting nonnegative gain*)

D3. $\mathbb{R}_{>} \cdot \mathcal{D} \subseteq \mathcal{D}$ (*positive scaling*)

D4. $\mathcal{D} + \mathcal{D} \subseteq \mathcal{D}$ (*combination*)

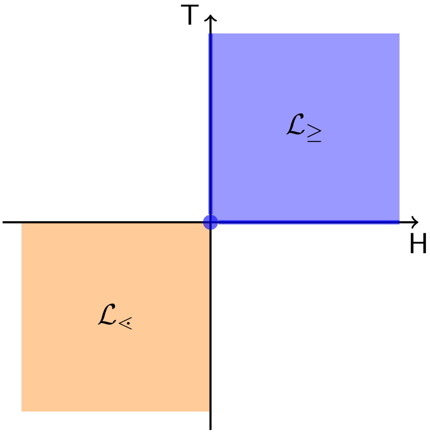
Gamble (vector) inequalities: $f \gtrdot 0$ iff $\min f > 0$, $f \geq 0$ iff $\min f \geq 0$, $f > 0$ iff $f \geq 0$ and $f \neq 0$

Visualizing gambles



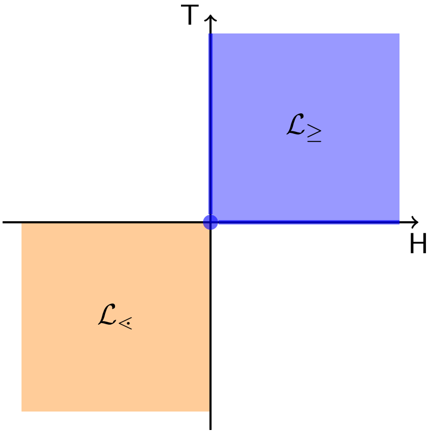
Visualizing sets of acceptable gambles

The avoiding loss and accepting gain axioms constrain $\mathcal{D}$

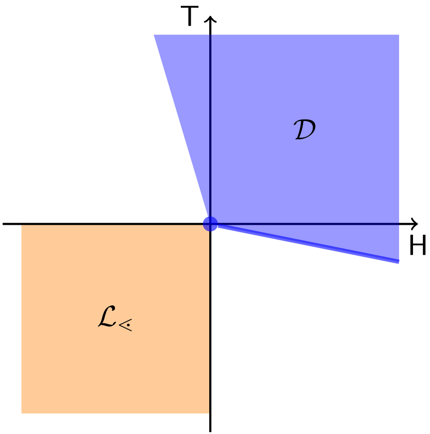


Visualizing sets of acceptable gambles

The avoiding loss and accepting gain axioms constrain $\mathcal{D}$



Coherent sets $\mathcal{D}$ are *cones* because of positive scaling and combination



From acceptability to preference

- Consider two gambles, f and g
- Depending on where their differences $f - g$ and $g - f$ lie, different preference relations hold between them

From acceptability to preference

- Consider two gambles, f and g
- Depending on where their differences $f - g$ and $g - f$ lie, different preference relations hold between them

Derived preference relations

- **Non-strict preference**: $f \succcurlyeq g$ iff $f - g \in \mathcal{D}$

From acceptability to preference

- Consider two gambles, f and g
- Depending on where their differences $f - g$ and $g - f$ lie, different preference relations hold between them

Derived preference relations

- **Non-strict preference**: $f \succcurlyeq g$ iff $f - g \in \mathcal{D}$
- **Indifference**: $f \approx g$ iff $f \succcurlyeq g$ and $g \succcurlyeq f$

From acceptability to preference

- Consider two gambles, f and g
- Depending on where their differences $f - g$ and $g - f$ lie, different preference relations hold between them

Derived preference relations

- **Non-strict preference**: $f \succcurlyeq g$ iff $f - g \in \mathcal{D}$
- **Indifference**: $f \approx g$ iff $f \succcurlyeq g$ and $g \succcurlyeq f$
- **Incomparability**: $f \asymp g$ iff $f \not\succcurlyeq g$ and $g \not\succcurlyeq f$

From acceptability to preference

- Consider two gambles, f and g
- Depending on where their differences $f - g$ and $g - f$ lie, different preference relations hold between them

Derived preference relations

- **Non-strict preference**: $f \succcurlyeq g$ iff $f - g \in \mathcal{D}$
- **Indifference**: $f \approx g$ iff $f \succcurlyeq g$ and $g \succcurlyeq f$
- **Incomparability**: $f \asymp g$ iff $f \not\succcurlyeq g$ and $g \not\succcurlyeq f$
- **Strict preference**: $f \succ g$ iff $f - g \in \mathcal{D} - \mathcal{L}_{\lessdot}$

From acceptability to preference

- Consider two gambles, f and g
- Depending on where their differences $f - g$ and $g - f$ lie, different preference relations hold between them

Derived preference relations

- **Non-strict preference**: $f \succcurlyeq g$ iff $f - g \in \mathcal{D}$
- **Indifference**: $f \approx g$ iff $f \succcurlyeq g$ and $g \succcurlyeq f$
- **Incomparability**: $f \asymp g$ iff $f \not\succcurlyeq g$ and $g \not\succcurlyeq f$
- **Strict preference**: $f \succ g$ iff $f - g \in \mathcal{D} - \mathcal{L}_{\lessdot}$

Background preference/order implied by D1 and D2

- Non-strict preference: $f \succcurlyeq g$ iff $f \geq g$
- Indifference: if $f = g$, then $f \approx g$
- Incomparability: $f \asymp g$ iff $f \not\geq g$ and $g \not\geq f$
- Strict preference: $f \succ g$ iff $f \gtrdot g$

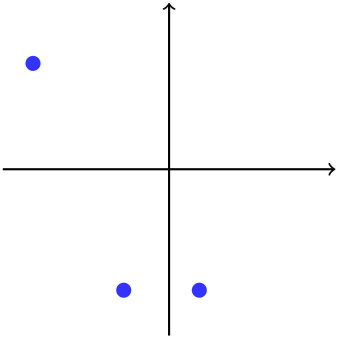
Assessments

An **assessment** $\mathcal{A}$ is a set of gambles the subject finds acceptable

Assessments

An **assessment** $\mathcal{A}$ is a set of gambles the subject finds acceptable

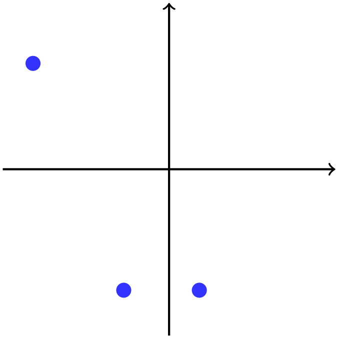
Directly specified



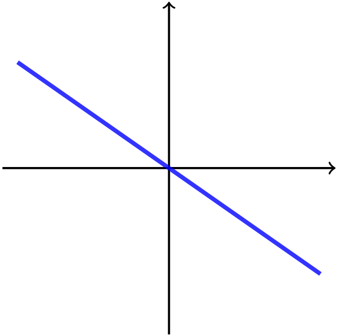
Assessments

An **assessment** $\mathcal{A}$ is a set of gambles the subject finds acceptable

Directly specified



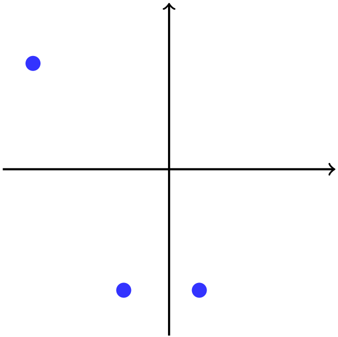
Indifference statements



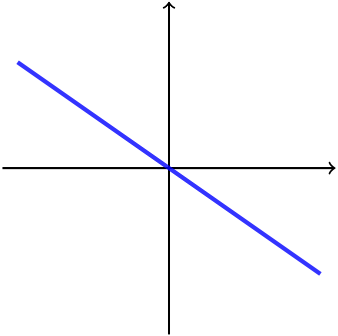
Assessments

An **assessment** $\mathcal{A}$ is a set of gambles the subject finds acceptable

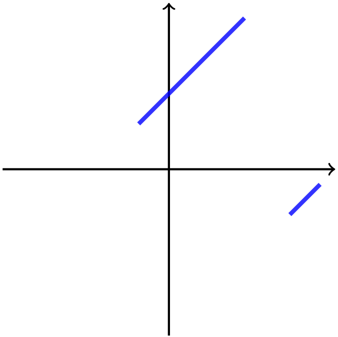
Directly specified



Indifference statements

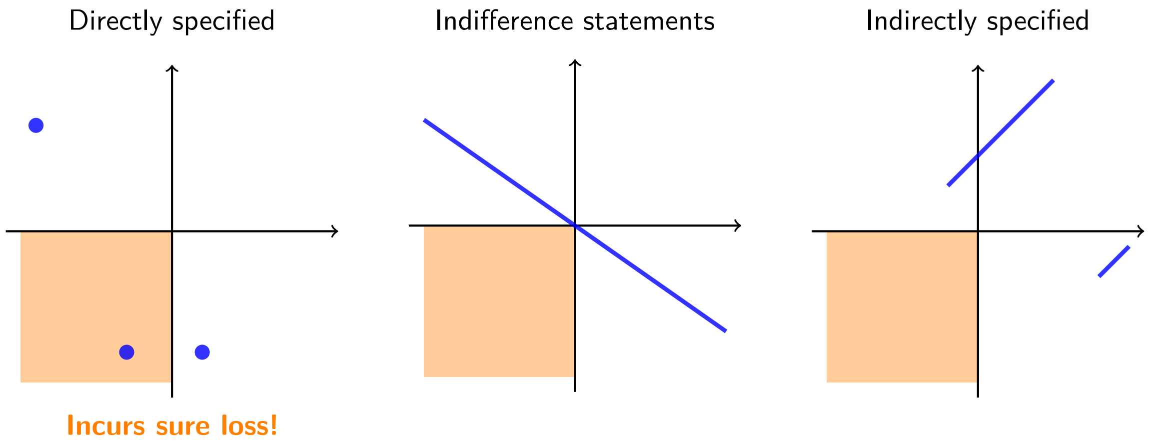


Indirectly specified



Assessments

An **assessment** $\mathcal{A}$ is a set of gambles the subject finds acceptable



Positive linear hull operation

Two of the axioms are 'generative':

- Positive scaling (D3): if $f \in \mathcal{A}$, then all gambles in $\{\lambda \cdot f : \lambda \in \mathbb{R}_{>}\}$ must also be considered acceptable
- Combination (D4): if $f, g \in \mathcal{A}$, then $f + g$ must also be considered acceptable

Positive linear hull operation

Two of the axioms are 'generative':

- Positive scaling (D3): if $f \in \mathcal{A}$, then all gambles in $\{\lambda \cdot f : \lambda \in \mathbb{R}_{>}\}$ must also be considered acceptable
- Combination (D4): if $f, g \in \mathcal{A}$, then $f + g$ must also be considered acceptable
- Together, they imply that all gambles in the *positive linear hull* (conical hull) $\operatorname{posi} \mathcal{A}$ of $\mathcal{A}$ must be considered acceptable

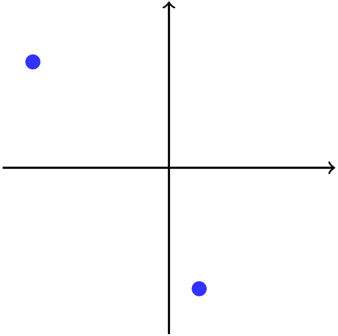
$$\operatorname{posi} \mathcal{A} := \left\{ \sum_{k=1}^{n} \lambda_k \cdot f_k : \lambda_k \in \mathbb{R}_{>}, f_k \in \mathcal{A}, n \in \mathbb{N} \right\}$$

Positive linear hull operation

Two of the axioms are 'generative':

- Positive scaling (D3): if $f \in \mathcal{A}$, then all gambles in $\{\lambda \cdot f : \lambda \in \mathbb{R}_{>}\}$ must also be considered acceptable
- Combination (D4): if $f, g \in \mathcal{A}$, then $f + g$ must also be considered acceptable
- Together, they imply that all gambles in the *positive linear hull* (conical hull) $\operatorname{posi}\mathcal{A}$ of $\mathcal{A}$ must be considered acceptable

$$\operatorname{posi}\mathcal{A} := \left\{ \sum_{k=1}^{n} \lambda_k \cdot f_k : \lambda_k \in \mathbb{R}_{>}, f_k \in \mathcal{A}, n \in \mathbb{N} \right\}$$

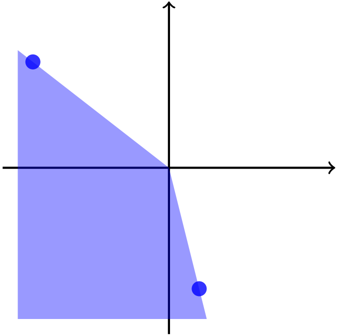


Positive linear hull operation

Two of the axioms are 'generative':

- Positive scaling (D3): if $f \in \mathcal{A}$, then all gambles in $\{\lambda \cdot f : \lambda \in \mathbb{R}_>\}$ must also be considered acceptable
- Combination (D4): if $f, g \in \mathcal{A}$, then $f + g$ must also be considered acceptable
- Together, they imply that all gambles in the *positive linear hull* (conical hull) $\operatorname{posi} \mathcal{A}$ of $\mathcal{A}$ must be considered acceptable

$$\operatorname{posi} \mathcal{A} := \left\{ \sum_{k=1}^{n} \lambda_k \cdot f_k : \lambda_k \in \mathbb{R}_>, f_k \in \mathcal{A}, n \in \mathbb{N} \right\}$$

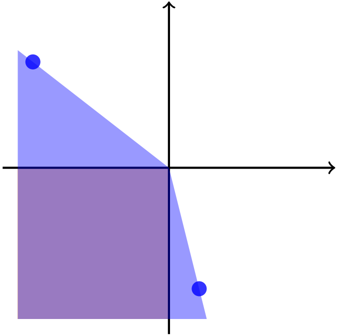


Positive linear hull operation

Two of the axioms are 'generative':

- Positive scaling (D3): if $f \in \mathcal{A}$, then all gambles in $\{\lambda \cdot f : \lambda \in \mathbb{R}_{>}\}$ must also be considered acceptable
- Combination (D4): if $f, g \in \mathcal{A}$, then $f + g$ must also be considered acceptable
- Together, they imply that all gambles in the *positive linear hull* (conical hull) $\operatorname{posi} \mathcal{A}$ of $\mathcal{A}$ must be considered acceptable

$$\operatorname{posi} \mathcal{A} := \left\{ \sum_{k=1}^{n} \lambda_k \cdot f_k : \lambda_k \in \mathbb{R}_{>}, f_k \in \mathcal{A}, n \in \mathbb{N} \right\}$$



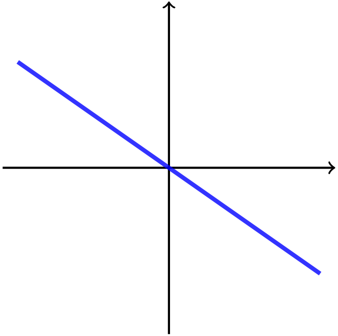
Incurs sure loss!

Positive linear hull operation

Two of the axioms are 'generative':

- Positive scaling (D3): if $f \in \mathcal{A}$, then all gambles in $\{\lambda \cdot f : \lambda \in \mathbb{R}_{>}\}$ must also be considered acceptable
- Combination (D4): if $f, g \in \mathcal{A}$, then $f + g$ must also be considered acceptable
- Together, they imply that all gambles in the *positive linear hull* (conical hull) $\operatorname{posi}\mathcal{A}$ of $\mathcal{A}$ must be considered acceptable

$$\operatorname{posi}\mathcal{A} := \left\{\sum_{k=1}^{n} \lambda_k \cdot f_k : \lambda_k \in \mathbb{R}_{>}, f_k \in \mathcal{A}, n \in \mathbb{N}\right\}$$

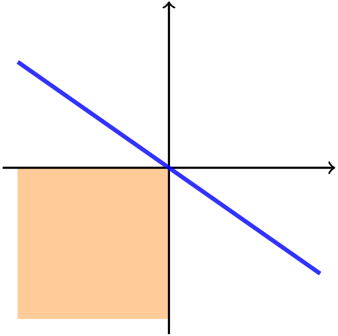


Positive linear hull operation

Two of the axioms are ‘generative’:

- Positive scaling (D3): if $f \in \mathcal{A}$, then all gambles in $\{\lambda \cdot f : \lambda \in \mathbb{R}_{>}\}$ must also be considered acceptable
- Combination (D4): if $f, g \in \mathcal{A}$, then $f + g$ must also be considered acceptable
- Together, they imply that all gambles in the *positive linear hull* (conical hull) $\operatorname{posi} \mathcal{A}$ of $\mathcal{A}$ must be considered acceptable

$$\operatorname{posi} \mathcal{A} := \left\{ \sum_{k=1}^{n} \lambda_k \cdot f_k : \lambda_k \in \mathbb{R}_{>}, f_k \in \mathcal{A}, n \in \mathbb{N} \right\}$$

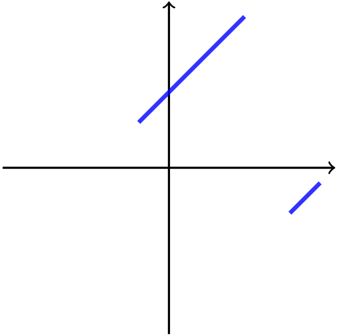


Positive linear hull operation

Two of the axioms are 'generative':

- Positive scaling (D3): if $f \in \mathcal{A}$, then all gambles in $\{\lambda \cdot f : \lambda \in \mathbb{R}_{>}\}$ must also be considered acceptable
- Combination (D4): if $f, g \in \mathcal{A}$, then $f + g$ must also be considered acceptable
- Together, they imply that all gambles in the *positive linear hull* (conical hull) $\operatorname{posi}\mathcal{A}$ of $\mathcal{A}$ must be considered acceptable

$$\operatorname{posi}\mathcal{A} := \left\{ \sum_{k=1}^{n} \lambda_k \cdot f_k : \lambda_k \in \mathbb{R}_{>}, f_k \in \mathcal{A}, n \in \mathbb{N} \right\}$$

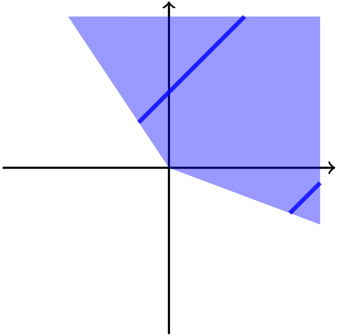


Positive linear hull operation

Two of the axioms are ‘generative’:

- Positive scaling (D3): if $f \in \mathcal{A}$, then all gambles in $\{\lambda \cdot f : \lambda \in \mathbb{R}_{>}\}$ must also be considered acceptable
- Combination (D4): if $f, g \in \mathcal{A}$, then $f + g$ must also be considered acceptable
- Together, they imply that all gambles in the *positive linear hull* (conical hull) $\operatorname{posi}\mathcal{A}$ of $\mathcal{A}$ must be considered acceptable

$$\operatorname{posi}\mathcal{A} := \left\{\sum_{k=1}^{n} \lambda_k \cdot f_k : \lambda_k \in \mathbb{R}_{>}, f_k \in \mathcal{A}, n \in \mathbb{N}\right\}$$

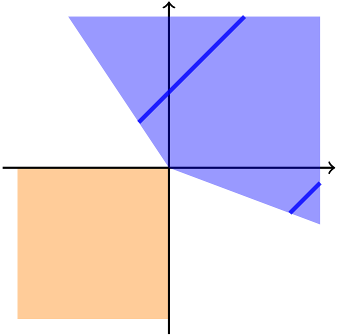


Positive linear hull operation

Two of the axioms are 'generative':

- Positive scaling (D3): if $f \in \mathcal{A}$, then all gambles in $\{\lambda \cdot f : \lambda \in \mathbb{R}_{>}\}$ must also be considered acceptable
- Combination (D4): if $f, g \in \mathcal{A}$, then $f + g$ must also be considered acceptable
- Together, they imply that all gambles in the *positive linear hull* (conical hull) $\operatorname{posi}\mathcal{A}$ of $\mathcal{A}$ must be considered acceptable

$$\operatorname{posi}\mathcal{A} := \left\{ \sum_{k=1}^{n} \lambda_k \cdot f_k : \lambda_k \in \mathbb{R}_{>}, f_k \in \mathcal{A}, n \in \mathbb{N} \right\}$$



Natural extension

- The accepting gain axiom D2 imposes a background assessment $\mathcal{L}_{\geq}$

Natural extension

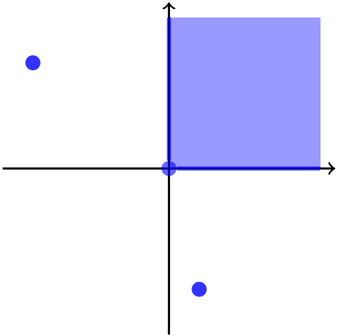
- The accepting gain axiom D2 imposes a background assessment $\mathcal{L}_{\geq}$
- The **natural extension** is the positive linear hull of the union of the background and subject's assessment:

$$\begin{aligned}\mathcal{E}(\mathcal{A}) &:= \operatorname{posi}(\mathcal{A} \cup \mathcal{L}_{\geq}) \\ &= (\operatorname{posi}\mathcal{A} + \mathcal{L}_{\geq}) \cup \mathcal{L}_{\geq}\end{aligned}$$

Natural extension

- The accepting gain axiom D2 imposes a background assessment $\mathcal{L}_{\geq}$
- The **natural extension** is the positive linear hull of the union of the background and subject's assessment:

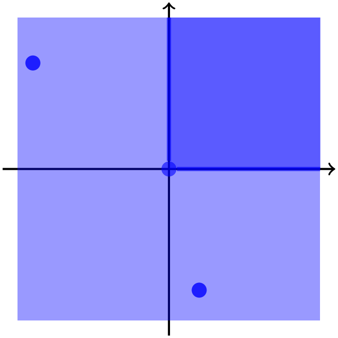
$$\begin{aligned}\mathcal{E}(\mathcal{A}) &:= \operatorname{posi}(\mathcal{A} \cup \mathcal{L}_{\geq}) \\ &= (\operatorname{posi}\mathcal{A} + \mathcal{L}_{\geq}) \cup \mathcal{L}_{\geq}\end{aligned}$$



Natural extension

- The accepting gain axiom D2 imposes a background assessment $\mathcal{L}_{\geq}$
- The **natural extension** is the positive linear hull of the union of the background and subject's assessment:

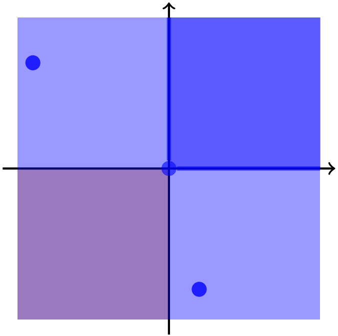
$$\begin{aligned}\mathcal{E}(\mathcal{A}) &:= \operatorname{posi}(\mathcal{A} \cup \mathcal{L}_{\geq}) \\ &= (\operatorname{posi} \mathcal{A} + \mathcal{L}_{\geq}) \cup \mathcal{L}_{\geq}\end{aligned}$$



Natural extension

- The accepting gain axiom D2 imposes a background assessment $\mathcal{L}_{\geq}$
- The **natural extension** is the positive linear hull of the union of the background and subject's assessment:

$$\begin{aligned}\mathcal{E}(\mathcal{A}) &:= \operatorname{posi}(\mathcal{A} \cup \mathcal{L}_{\geq}) \\ &= (\operatorname{posi}\mathcal{A} + \mathcal{L}_{\geq}) \cup \mathcal{L}_{\geq}\end{aligned}$$



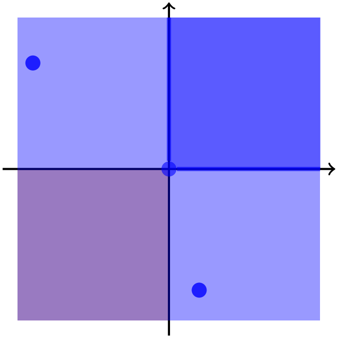
Incurs sure loss!

Natural extension

- The accepting gain axiom D2 imposes a background assessment $\mathcal{L}_{\geq}$
- The **natural extension** is the positive linear hull of the union of the background and subject's assessment:

$$\begin{aligned}\mathcal{E}(\mathcal{A}) &:= \operatorname{posi}(\mathcal{A} \cup \mathcal{L}_{\geq}) \\ &= (\operatorname{posi}\mathcal{A} + \mathcal{L}_{\geq}) \cup \mathcal{L}_{\geq}\end{aligned}$$

- The natural extension $\mathcal{E}(\mathcal{A})$ *avoids sure loss* (satisfies D1) iff $\operatorname{posi}\mathcal{A}$ avoids sure loss



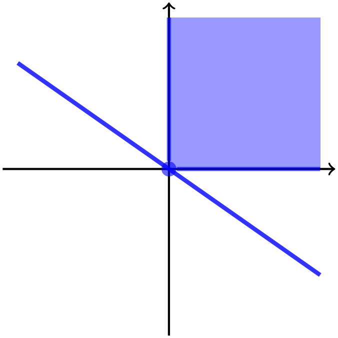
Incurs sure loss!

Natural extension

- The accepting gain axiom D2 imposes a background assessment $\mathcal{L}_{\geq}$
- The **natural extension** is the positive linear hull of the union of the background and subject's assessment:

$$\begin{aligned}\mathcal{E}(\mathcal{A}) :=&\ \operatorname{posi}(\mathcal{A} \cup \mathcal{L}_{\geq}) \\ =&\ (\operatorname{posi}\mathcal{A} + \mathcal{L}_{\geq}) \cup \mathcal{L}_{\geq}\end{aligned}$$

- The natural extension $\mathcal{E}(\mathcal{A})$ *avoids sure loss* (satisfies D1) iff $\operatorname{posi}\mathcal{A}$ avoids sure loss

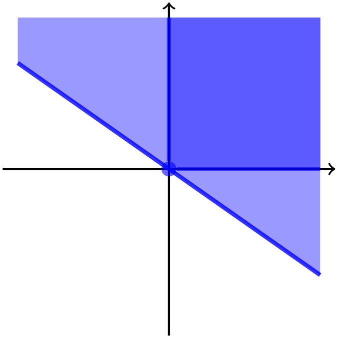


Natural extension

- The accepting gain axiom D2 imposes a background assessment $\mathcal{L}_{\geq}$
- The **natural extension** is the positive linear hull of the union of the background and subject's assessment:

$$\begin{aligned}\mathcal{E}(\mathcal{A}) &:= \operatorname{posi}(\mathcal{A} \cup \mathcal{L}_{\geq}) \\ &= (\operatorname{posi} \mathcal{A} + \mathcal{L}_{\geq}) \cup \mathcal{L}_{\geq}\end{aligned}$$

- The natural extension $\mathcal{E}(\mathcal{A})$ *avoids sure loss* (satisfies D1) iff $\operatorname{posi} \mathcal{A}$ avoids sure loss

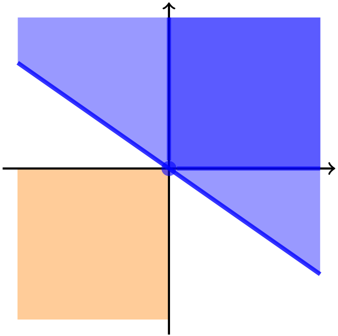


Natural extension

- The accepting gain axiom D2 imposes a background assessment $\mathcal{L}_{\geq}$
- The **natural extension** is the positive linear hull of the union of the background and subject's assessment:

$$\begin{aligned}\mathcal{E}(\mathcal{A}) &:= \operatorname{posi}(\mathcal{A} \cup \mathcal{L}_{\geq}) \\ &= (\operatorname{posi}\mathcal{A} + \mathcal{L}_{\geq}) \cup \mathcal{L}_{\geq}\end{aligned}$$

- The natural extension $\mathcal{E}(\mathcal{A})$ *avoids sure loss* (satisfies D1) iff $\operatorname{posi}\mathcal{A}$ avoids sure loss

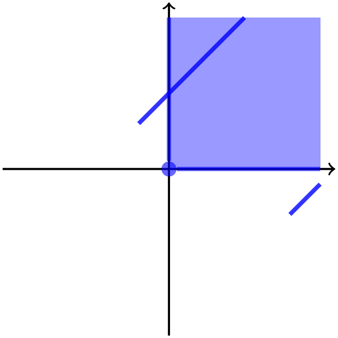


Natural extension

- The accepting gain axiom D2 imposes a background assessment $\mathcal{L}_{\geq}$
- The **natural extension** is the positive linear hull of the union of the background and subject's assessment:

$$\begin{aligned}\mathcal{E}(\mathcal{A}) &:= \operatorname{posi}(\mathcal{A} \cup \mathcal{L}_{\geq}) \\ &= (\operatorname{posi}\mathcal{A} + \mathcal{L}_{\geq}) \cup \mathcal{L}_{\geq}\end{aligned}$$

- The natural extension $\mathcal{E}(\mathcal{A})$ *avoids sure loss* (satisfies D1) iff $\operatorname{posi}\mathcal{A}$ avoids sure loss

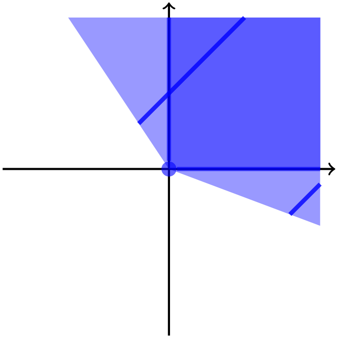


Natural extension

- The accepting gain axiom D2 imposes a background assessment $\mathcal{L}_{\geq}$
- The **natural extension** is the positive linear hull of the union of the background and subject's assessment:

$$\begin{aligned}\mathcal{E}(\mathcal{A}) &:= \operatorname{posi}(\mathcal{A} \cup \mathcal{L}_{\geq}) \\ &= (\operatorname{posi}\mathcal{A} + \mathcal{L}_{\geq}) \cup \mathcal{L}_{\geq}\end{aligned}$$

- The natural extension $\mathcal{E}(\mathcal{A})$ *avoids sure loss* (satisfies D1) iff $\operatorname{posi}\mathcal{A}$ avoids sure loss

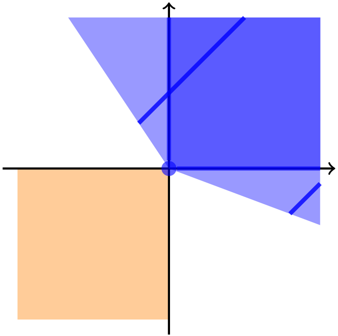


Natural extension

- The accepting gain axiom D2 imposes a background assessment $\mathcal{L}_{\geq}$
- The **natural extension** is the positive linear hull of the union of the background and subject's assessment:

$$\mathcal{E}(\mathcal{A}) := \operatorname{posi}(\mathcal{A} \cup \mathcal{L}_{\geq}) = (\operatorname{posi}\mathcal{A} + \mathcal{L}_{\geq}) \cup \mathcal{L}_{\geq}$$

- The natural extension $\mathcal{E}(\mathcal{A})$ *avoids sure loss* (satisfies D1) iff $\operatorname{posi}\mathcal{A}$ avoids sure loss



Least committal extension

Set of all coherent sets of acceptable gambles

- This set can be (partially) ordered according to set inclusion $\subset$
- If $\mathcal{D}_1 \subset \mathcal{D}_2$, then $\mathcal{D}_1$ is called less committal than $\mathcal{D}_2$
- The least committal set of acceptable gambles is $\mathcal{L}_{\geq}$

Least committal extension

Set of all coherent sets of acceptable gambles

- This set can be (partially) ordered according to set inclusion $\subset$
- If $\mathcal{D}_1 \subset \mathcal{D}_2$, then $\mathcal{D}_1$ is called less committal than $\mathcal{D}_2$
- The least committal set of acceptable gambles is $\mathcal{L}_{\geq}$

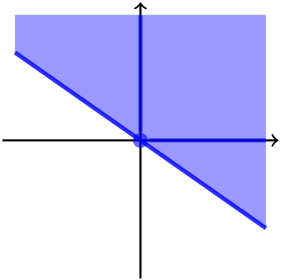
Coherent extensions of an assessment $\mathcal{A}$ that avoids sure loss

- In general, there are multiple coherent extensions of $\mathcal{A}$

Least committal extension

Set of all coherent sets of acceptable gambles

- This set can be (partially) ordered according to set inclusion $\subset$
- If $\mathcal{D}_1 \subset \mathcal{D}_2$, then $\mathcal{D}_1$ is called less committal than $\mathcal{D}_2$
- The least committal set of acceptable gambles is $\mathcal{L}_{\geq}$



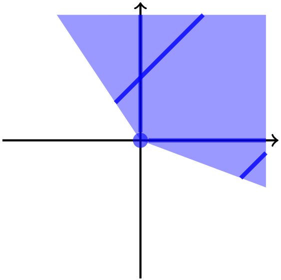
Coherent extensions of an assessment $\mathcal{A}$ that avoids sure loss

- In general, there are multiple coherent extensions of $\mathcal{A}$

Least committal extension

Set of all coherent sets of acceptable gambles

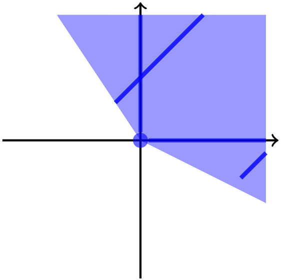
- This set can be (partially) ordered according to set inclusion $\subset$
- If $\mathcal{D}_1 \subset \mathcal{D}_2$, then $\mathcal{D}_1$ is called less committal than $\mathcal{D}_2$
- The least committal set of acceptable gambles is $\mathcal{L}_{\geq}$



Coherent extensions of an assessment $\mathcal{A}$ that avoids sure loss

- In general, there are multiple coherent extensions of $\mathcal{A}$

Least committal extension



Set of all coherent sets of acceptable gambles

- This set can be (partially) ordered according to set inclusion $\subset$
- If $\mathcal{D}_1 \subset \mathcal{D}_2$, then $\mathcal{D}_1$ is called less committal than $\mathcal{D}_2$
- The least committal set of acceptable gambles is $\mathcal{L}_{\geq}$

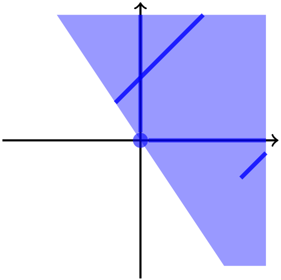
Coherent extensions of an assessment $\mathcal{A}$ that avoids sure loss

- In general, there are multiple coherent extensions of $\mathcal{A}$

Least committal extension

Set of all coherent sets of acceptable gambles

- This set can be (partially) ordered according to set inclusion $\subset$
- If $\mathcal{D}_1 \subset \mathcal{D}_2$, then $\mathcal{D}_1$ is called less committal than $\mathcal{D}_2$
- The least committal set of acceptable gambles is $\mathcal{L}_{\geq}$



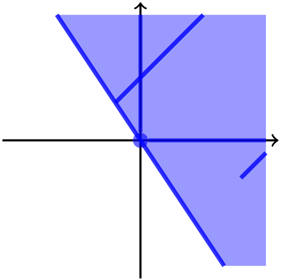
Coherent extensions of an assessment $\mathcal{A}$ that avoids sure loss

- In general, there are multiple coherent extensions of $\mathcal{A}$

Least committal extension

Set of all coherent sets of acceptable gambles

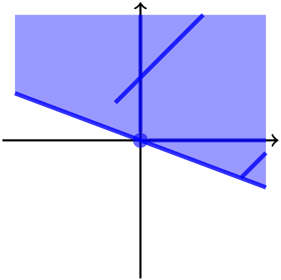
- This set can be (partially) ordered according to set inclusion $\subset$
- If $\mathcal{D}_1 \subset \mathcal{D}_2$, then $\mathcal{D}_1$ is called less committal than $\mathcal{D}_2$
- The least committal set of acceptable gambles is $\mathcal{L}_{\geq}$



Coherent extensions of an assessment $\mathcal{A}$ that avoids sure loss

- In general, there are multiple coherent extensions of $\mathcal{A}$

Least committal extension



Set of all coherent sets of acceptable gambles

- This set can be (partially) ordered according to set inclusion $\subset$
- If $\mathcal{D}_1 \subset \mathcal{D}_2$, then $\mathcal{D}_1$ is called less committal than $\mathcal{D}_2$
- The least committal set of acceptable gambles is $\mathcal{L}_{\geq}$

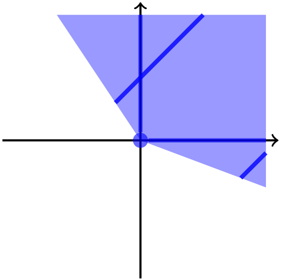
Coherent extensions of an assessment $\mathcal{A}$ that avoids sure loss

- In general, there are multiple coherent extensions of $\mathcal{A}$

Least committal extension

Set of all coherent sets of acceptable gambles

- This set can be (partially) ordered according to set inclusion $\subset$
- If $\mathcal{D}_1 \subset \mathcal{D}_2$, then $\mathcal{D}_1$ is called less committal than $\mathcal{D}_2$
- The least committal set of acceptable gambles is $\mathcal{L}_{\geq}$



Coherent extensions of an assessment $\mathcal{A}$ that avoids sure loss

- In general, there are multiple coherent extensions of $\mathcal{A}$
- The **least committal extension** is the smallest one: it adds the least commitments (it is *conservative* in that regard)
- The least committal extension is equal to the intersection of all coherent extensions
- *The least committal extension coincides with the natural extension*

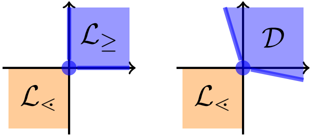
Changing how $\mathcal{D}$ is constrained

- Replacing the background sets $\mathcal{L}_{\lessdot}/\mathcal{L}_{\geq}$ in Axiom D1/D2 (extra relevant for infinite $\mathcal{X}$)
- Adding additional axioms constraining, e.g., topological closure

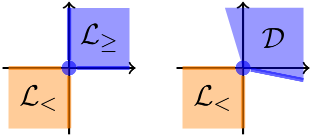
Changing how $\mathcal{D}$ is constrained

- Replacing the background sets $\mathcal{L}_{\lessdot}/\mathcal{L}_{\geq}$ in Axiom D1/D2 (extra relevant for infinite $\mathcal{X}$)
- Adding additional axioms constraining, e.g., topological closure

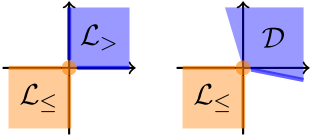
- $\mathcal{L}_{\lessdot}/\mathcal{L}_{\geq}$ and $\mathcal{D}$ closed (*almost desirability*, Walley, 1991)



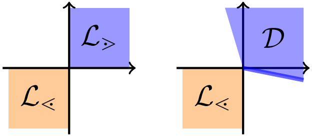
- $\mathcal{L}_{<}/\mathcal{L}_{\geq}$ (*favorable gambles*, Seidenfeld et al., 1990; Walley, 2000)



- $\mathcal{L}_{\leq}/\mathcal{L}_{>}$ (*desirability*, De Cooman, Thursday! D1 becomes $0 \notin \mathcal{D}$)



- $\mathcal{L}_{\lessdot}/\mathcal{L}_{\gtrdot}$ (*acceptable bets*, Williams, 1974)



Changing what shape $\mathcal{D}$ can take

- Replacing the generating rules in Axioms D3 and D4
- Adding additional generating rules

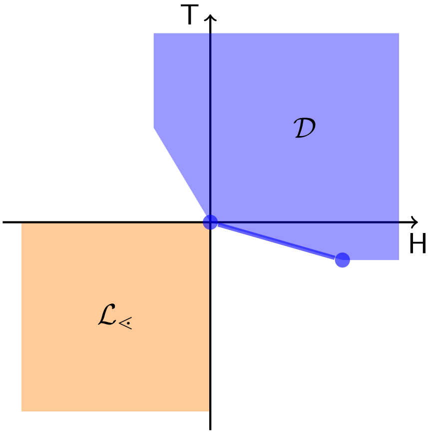
Changing what shape $\mathcal{D}$ can take

- Replacing the generating rules in Axioms D3 and D4
- Adding additional generating rules

- *Example*: replace positive scaling and combination by *convexity*,

  if $f, g \in \mathcal{D}$ and $\mu \in [0, 1]$,
  then $\mu \cdot f + (1 - \mu) \cdot g \in \mathcal{D}$

  This means replacing the *positive linear hull* posi by the *convex hull* co



Multivariate acceptability: basic setup

- Index set $N = \{1, \dots, n\}$
- Multivariate variable $\boldsymbol{X} = (X_1, X_2, \dots, X_n)$
- Set of possible outcomes $\boldsymbol{x} \in \mathcal{X} = \mathcal{X}_1 \times \mathcal{X}_2 \times \dots \times \mathcal{X}_n$
- Linear space of gambles $\mathcal{L}$ on $\mathcal{X}$
- We assume given a coherent set of acceptable gambles $\mathcal{D} \subset \mathcal{L}$

Marginal set of acceptable gambles

- A marginal set of acceptable gambles is defined on a linear subspace of $\mathcal{L}$
- We consider subspaces defined by $\mathcal{L}_K := (\mathcal{X}_K \to \mathbb{R})$, with $K \subseteq N$
- We want to know whether a gamble f on $\mathcal{X}_K$ is acceptable

Marginal set of acceptable gambles

- A marginal set of acceptable gambles is defined on a linear subspace of $\mathcal{L}$
- We consider subspaces defined by $\mathcal{L}_K := (\mathcal{X}_K \to \mathbb{R})$, with $K \subseteq N$
- We want to know whether a gamble f on $\mathcal{X}_K$ is acceptable

- For this, we lift it to its **cylindrical extension** $\uparrow_{\boldsymbol{x}_{N\setminus K}} f$ defined by

$$(\uparrow_{\boldsymbol{x}_{N\setminus K}} f)(\boldsymbol{x}) = f(\boldsymbol{x}_K)$$

  and check whether that is acceptable

Marginal set of acceptable gambles

- A marginal set of acceptable gambles is defined on a linear subspace of $\mathcal{L}$
- We consider subspaces defined by $\mathcal{L}_K := (\mathcal{X}_K \to \mathbb{R})$, with $K \subseteq N$
- We want to know whether a gamble f on $\mathcal{X}_K$ is acceptable

- For this, we lift it to its **cylindrical extension** $\uparrow_{\boldsymbol{x}_{N\setminus K}} f$ defined by

$$(\uparrow_{\boldsymbol{x}_{N\setminus K}} f)(\boldsymbol{x}) = f(\boldsymbol{x}_K)$$

  and check whether that is acceptable

- The marginal set of acceptable gambles is the inverse image of the cylindrical extension:

$$\mathcal{D}_K := \uparrow^{-1}_{\boldsymbol{x}_{N\setminus K}} \mathcal{D} = \left\{ f \in \mathcal{L}_K : \uparrow_{\boldsymbol{x}_{N\setminus K}} f \in \mathcal{D} \right\}$$

Conditional set of acceptable gambles

- A conditional set of acceptable gambles is determined by
  - assuming some event is known to be true, or specifically
  - assuming some random variables take some given values: $\boldsymbol{X}_K = \boldsymbol{x}_K$, with $K \subset N$
- We consider the linear space $\mathcal{L}_{N \setminus K} := (\mathcal{X}_{N \setminus K} \to \mathbb{R})$
- We want to know whether a gamble f on $\mathcal{X}_{N \setminus K}$ is conditionally acceptable

Conditional set of acceptable gambles

- A conditional set of acceptable gambles is determined by
  - assuming some event is known to be true, or specifically
  - assuming some random variables take some given values: $\boldsymbol{X}_K = \boldsymbol{x}_K$, with $K \subset N$
- We consider the linear space $\mathcal{L}_{N \setminus K} := (\mathcal{X}_{N \setminus K} \to \mathbb{R})$
- We want to know whether a gamble f on $\mathcal{X}_{N \setminus K}$ is conditionally acceptable
- For this, we lift it to the **called-off gamble**

$$(\uparrow_{\boldsymbol{x}_K} f) \cdot 1_{\boldsymbol{X}_K = \boldsymbol{x}_K}$$

  and check whether that is acceptable

Conditional set of acceptable gambles

- A conditional set of acceptable gambles is determined by
  - assuming some event is known to be true, or specifically
  - assuming some random variables take some given values: $\boldsymbol{X}_K = \boldsymbol{x}_K$, with $K \subset N$
- We consider the linear space $\mathcal{L}_{N \setminus K} := (\mathcal{X}_{N \setminus K} \to \mathbb{R})$
- We want to know whether a gamble f on $\mathcal{X}_{N \setminus K}$ is conditionally acceptable
- For this, we lift it to the **called-off gamble**

$$(\uparrow_{\boldsymbol{x}_K} f) \cdot 1_{\boldsymbol{X}_K = \boldsymbol{x}_K}$$

  and check whether that is acceptable

- The conditional set of acceptable gambles is the inverse image of this procedure:

$$\mathcal{D}\rfloor(\boldsymbol{X}_K = \boldsymbol{x}_K) := \left\{ f \in \mathcal{L}_{N \setminus K} : (\uparrow_{\boldsymbol{x}_K} f) \cdot 1_{\boldsymbol{X}_K = \boldsymbol{x}_K} \in \mathcal{D} \right\}$$

Epistemic irrelevance & independence

- We here consider the case $|N| = 2$
- The random variable X_2 is **epistemically irrelevant** to X_1 if the X_2-conditionals coincide with the X_1-marginal for all $x_2 \in \mathcal{X}_2$:

$$\mathcal{D}\rfloor(X_2 = x_2) = \mathcal{D}_1$$

Epistemic irrelevance & independence

- We here consider the case $|N| = 2$
- The random variable X_2 is **epistemically irrelevant** to X_1 if the X_2-conditionals coincide with the X_1-marginal for all $x_2 \in \mathcal{X}_2$:

$$\mathcal{D}\rfloor(X_2 = x_2) = \mathcal{D}_1$$

- The random variables X_1 and X_2 are **epistemically independent** if they are epistemically irrelevant to each other

Epistemic irrelevance & independence

- We here consider the case $|N| = 2$
- The random variable X_2 is **epistemically irrelevant** to X_1
  if the X_2-conditionals coincide with the X_1-marginal for all $x_2 \in \mathcal{X}_2$:

$$\mathcal{D}\rfloor(X_2 = x_2) = \mathcal{D}_1$$

- The random variables X_1 and X_2 are **epistemically independent**
  if they are epistemically irrelevant to each other
- *Structural assumptions* such as these can be combined with natural extension

Overview

Kick-off (slot 1)

Classical probability theory (slot 1)

Interpretation of probability (slot 2)

Limitations of probability theory (slot 2)

Probability intervals (slot 3)

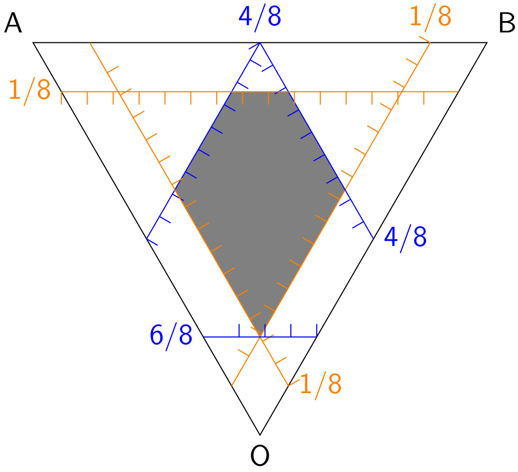
Credal sets (slot 3–4)

Acceptability & Desirability (slot 4–5)

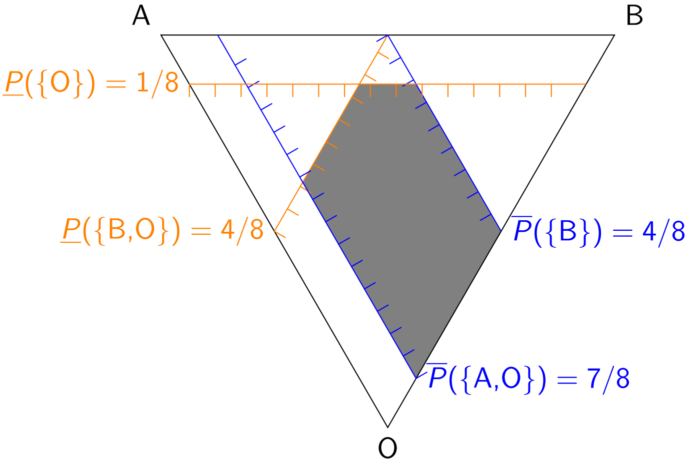
Interval expectation & probability (slot 5–6)

- Representation
- Reasoning
- Representation (continued)
- Learning
- Multivariate lower expectations

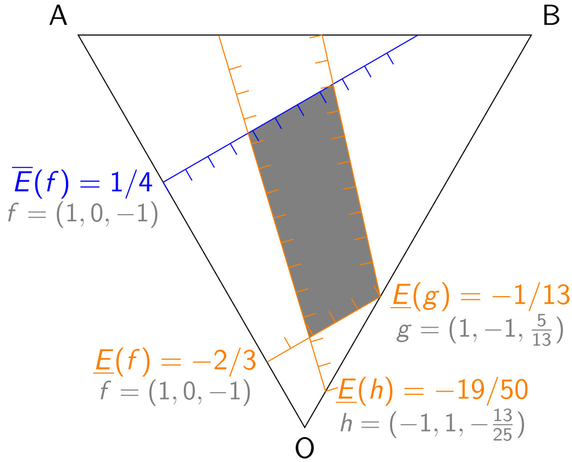
Probability interval example



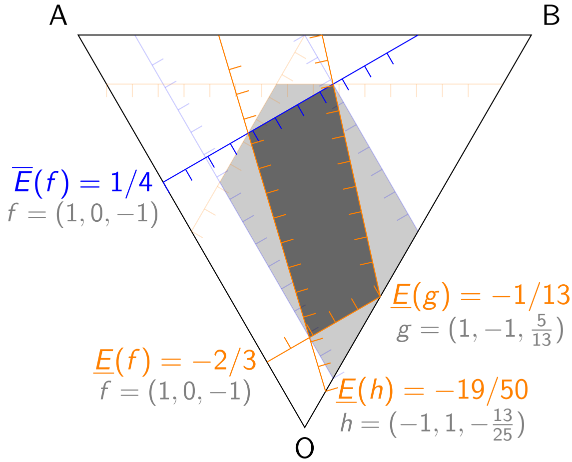
Interval probability example



Interval expectation example



Interval expectation example



Basic setup, axioms, and terminology

Basic setup of **the theory of coherent interval probability & expectation**:

- Random variable X
- Set of outcomes $\mathcal{X}$
- Set of events $\mathcal{S} \subseteq 2^{\mathcal{X}}$ or set of gambles $\mathcal{F} \subseteq \mathcal{L}$

Basic setup, axioms, and terminology

Basic setup of **the theory of coherent interval expectation**:

- Random variable X
- Set of outcomes $\mathcal{X}$
- ~~Set of events $\mathcal{S} \subseteq 2^{\mathcal{X}}$ or~~ Set of gambles $\mathcal{F} \subseteq \mathcal{L}$

Basic setup, axioms, and terminology

Basic setup of **the theory of coherent interval expectation**:

- Random variable X
- Set of outcomes $\mathcal{X}$
- ~~Set of events $\mathcal{S} \subseteq 2^{\mathcal{X}}$ or~~ Set of gambles $\mathcal{F} \subseteq \mathcal{L}$
- Each gamble $f \in \mathcal{F}$ is assigned a **lower or upper expectation value**

Basic setup, axioms, and terminology

Basic setup of **the theory of coherent interval expectation**:

- Random variable X
- Set of outcomes $\mathcal{X}$
- ~~Set of events $\mathcal{S} \subseteq 2^{\mathcal{X}}$ or~~ Set of gambles $\mathcal{F} \subseteq \mathcal{L}$
- Each gamble $f \in \mathcal{F}$ is assigned a **lower or upper expectation value**

Axioms

The **lower and upper expectation** operators $\underline{E}$ and $\overline{E}$, must be:

?

Eliciting lower & upper expectations with a relaxed betting game

- Two players:
  - *subject*: gambler, states acceptable exchanges, their '*assessment*'
  - *bookie*: chooses combination of acceptable exchanges
- *Gambles* from the subject's perspective: $f \in \mathcal{F}$

Eliciting lower & upper expectations with a relaxed betting game

- Two players:
  - *subject*: gambler, states acceptable exchanges, their '*assessment*'
  - *bookie*: chooses combination of acceptable exchanges
- *Gambles* from the subject's perspective: $f \in \mathcal{F}$

- For each f, the subject can state a *supremum acceptable buying price* $\underline{E}(f)$
- This $\underline{E}(f)$ is the subject's lower expectation for f
  (also called *lower prevision* $\underline{P}(f)$)

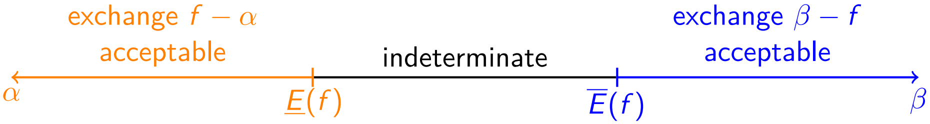
- For each f, the subject can state an *infimum acceptable selling price* $\overline{E}(f)$
- This $\overline{E}(f)$ is the subject's upper expectation for f
  (also called *upper prevision* $\overline{P}(f)$)

Eliciting lower & upper expectations with a relaxed betting game

- Two players:
  - *subject*: gambler, states acceptable exchanges, their '*assessment*'
  - *bookie*: chooses combination of acceptable exchanges
- *Gambles* from the subject's perspective: $f \in \mathcal{F}$

- For each f, the subject can state a *supremum acceptable buying price* $\underline{E}(f)$
- This $\underline{E}(f)$ is the subject's lower expectation for f

- For each f, the subject can state an *infimum acceptable selling price* $\overline{E}(f)$
- This $\overline{E}(f)$ is the subject's upper expectation for f



Assessment as a set of acceptable gambles

- The subject states $\underline{E}(f)$ for f and $\overline{E}(g)$ for g
- The corresponding *assessment* is

$$\mathcal{A} = \left\{ f - \alpha \cdot 1_{\mathcal{X}} : \alpha < \underline{E}(f) \right\} \cup \left\{ \beta \cdot 1_{\mathcal{X}} - g : \beta > \overline{E}(g) \right\}$$

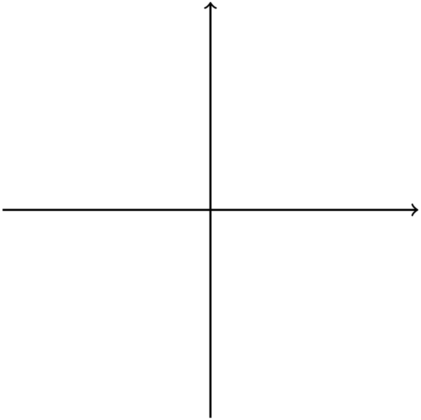
Assessment as a set of acceptable gambles

- The subject states $\underline{E}(f)$ for f and $\overline{E}(g)$ for g
- The corresponding *assessment* is

$$\mathcal{A} = \left\{ f - \alpha \cdot 1_{\mathcal{X}} : \alpha < \underline{E}(f) \right\}$$

$$\cup$$

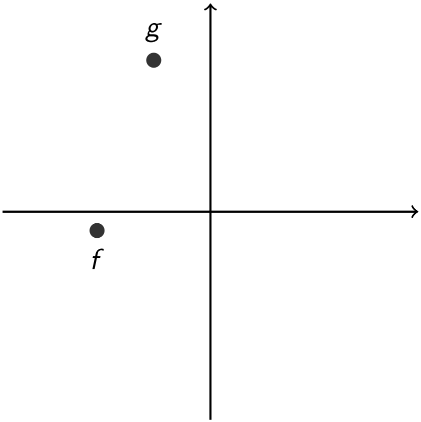
$$\left\{ \beta \cdot 1_{\mathcal{X}} - g : \beta > \overline{E}(g) \right\}$$



Assessment as a set of acceptable gambles

- The subject states $\underline{E}(f)$ for f and $\overline{E}(g)$ for g
- The corresponding *assessment* is

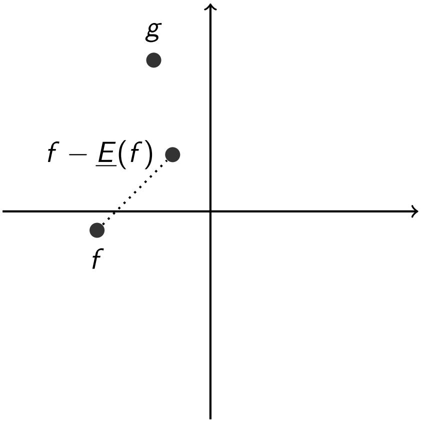
$$\mathcal{A} = \left\{ f - \alpha \cdot 1_{\mathcal{X}} : \alpha < \underline{E}(f) \right\} \cup \left\{ \beta \cdot 1_{\mathcal{X}} - g : \beta > \overline{E}(g) \right\}$$



Assessment as a set of acceptable gambles

- The subject states $\underline{E}(f)$ for f and $\overline{E}(g)$ for g
- The corresponding *assessment* is

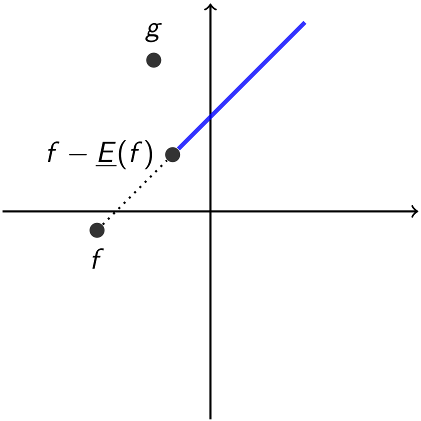
$$\mathcal{A} = \left\{ f - \alpha \cdot 1_{\mathcal{X}} : \alpha < \underline{E}(f) \right\} \cup \left\{ \beta \cdot 1_{\mathcal{X}} - g : \beta > \overline{E}(g) \right\}$$



Assessment as a set of acceptable gambles

- The subject states $\underline{E}(f)$ for f and $\overline{E}(g)$ for g
- The corresponding *assessment* is

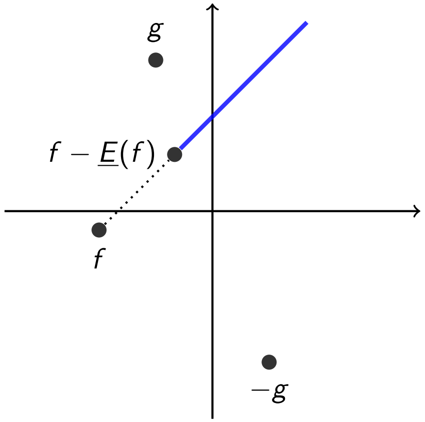
$$\mathcal{A} = \left\{ f - \alpha \cdot 1_{\mathcal{X}} : \alpha < \underline{E}(f) \right\} \cup \left\{ \beta \cdot 1_{\mathcal{X}} - g : \beta > \overline{E}(g) \right\}$$



Assessment as a set of acceptable gambles

- The subject states $\underline{E}(f)$ for f and $\overline{E}(g)$ for g
- The corresponding *assessment* is

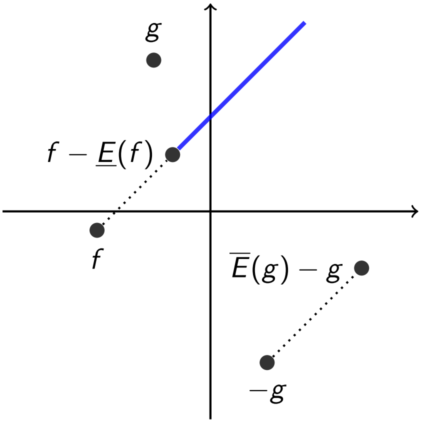
$$\mathcal{A} = \left\{ f - \alpha \cdot 1_{\mathcal{X}} : \alpha < \underline{E}(f) \right\} \cup \left\{ \beta \cdot 1_{\mathcal{X}} - g : \beta > \overline{E}(g) \right\}$$



Assessment as a set of acceptable gambles

- The subject states $\underline{E}(f)$ for f and $\overline{E}(g)$ for g
- The corresponding *assessment* is

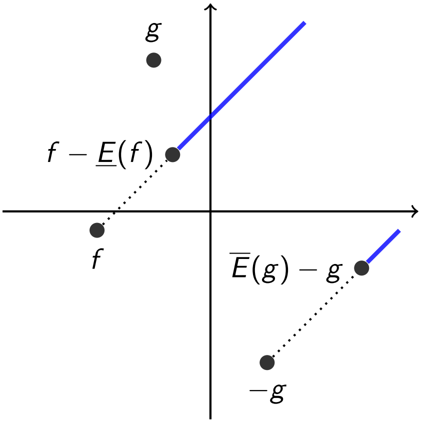
$$\mathcal{A} = \left\{ f - \alpha \cdot 1_{\mathcal{X}} : \alpha < \underline{E}(f) \right\} \cup \left\{ \beta \cdot 1_{\mathcal{X}} - g : \beta > \overline{E}(g) \right\}$$



Assessment as a set of acceptable gambles

- The subject states $\underline{E}(f)$ for f and $\overline{E}(g)$ for g
- The corresponding *assessment* is

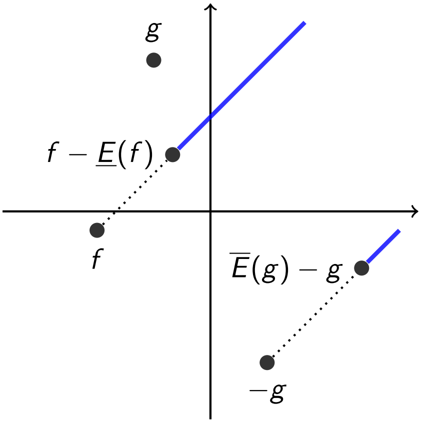
$$\mathcal{A} = \left\{ f - \alpha \cdot 1_{\mathcal{X}} : \alpha < \underline{E}(f) \right\}$$
$$\cup$$
$$\left\{ \beta \cdot 1_{\mathcal{X}} - g : \beta > \overline{E}(g) \right\}$$



Assessment as a set of acceptable gambles & marginal gambles

- The subject states $\underline{E}(f)$ for f and $\overline{E}(g)$ for g
- The corresponding *assessment* is

$$\mathcal{A} = \left\{ f - \alpha \cdot 1_{\mathcal{X}} : \alpha < \underline{E}(f) \right\} \cup \left\{ \beta \cdot 1_{\mathcal{X}} - g : \beta > \overline{E}(g) \right\}$$

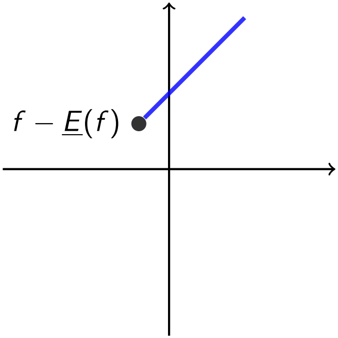


- $f - \underline{E}(f)$ and $\overline{E}(g) - g$ are called *marginal gambles*; collect them in a set $\mathcal{M}$, then

$$\mathcal{A} = \bigcup_{h \in \mathcal{M}} \{ h + \varepsilon \cdot 1_{\mathcal{X}} : \varepsilon \in \mathbb{R}_{>} \} = \mathcal{M} + \mathbb{R}_{>} \cdot 1_{\mathcal{X}}$$

Avoiding sure loss & natural extension

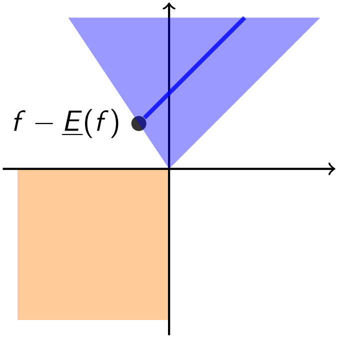
- We work in the language of acceptable gambles and later translate to the language of interval expectation



Avoiding sure loss & natural extension

- We work in the language of acceptable gambles and later translate to the language of interval expectation
- posi $\mathcal{A}$ **avoids sure loss** iff

$$\begin{aligned}\operatorname{posi}\mathcal{A}\cap\mathcal{L}_{\lessdot}\neq\emptyset &\Leftrightarrow \operatorname{posi}(\mathcal{M}+\mathbb{R}_{>}\cdot 1_{\mathcal{X}})\cap\mathcal{L}_{\lessdot}\neq\emptyset\\ &\Leftrightarrow (\operatorname{posi}\mathcal{M}+\mathbb{R}_{>}\cdot 1_{\mathcal{X}})\cap\mathcal{L}_{\lessdot}\neq\emptyset\\ &\Leftrightarrow \operatorname{posi}\mathcal{M}\cap\mathcal{L}_{\lessdot}\neq\emptyset\end{aligned}$$



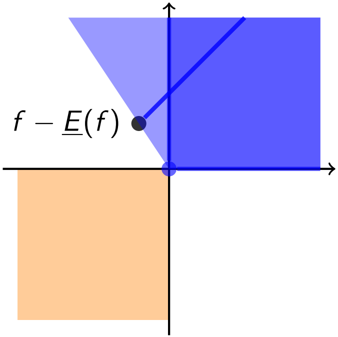
Avoiding sure loss & natural extension

- We work in the language of acceptable gambles and later translate to the language of interval expectation
- posi $\mathcal{A}$ **avoids sure loss** iff

$$\begin{aligned}\operatorname{posi}\mathcal{A}\cap\mathcal{L}_{\lessdot}\neq\emptyset &\Leftrightarrow \operatorname{posi}(\mathcal{M}+\mathbb{R}_{>}\cdot 1_{\mathcal{X}})\cap\mathcal{L}_{\lessdot}\neq\emptyset\\ &\Leftrightarrow (\operatorname{posi}\mathcal{M}+\mathbb{R}_{>}\cdot 1_{\mathcal{X}})\cap\mathcal{L}_{\lessdot}\neq\emptyset\\ &\Leftrightarrow \operatorname{posi}\mathcal{M}\cap\mathcal{L}_{\lessdot}\neq\emptyset\end{aligned}$$

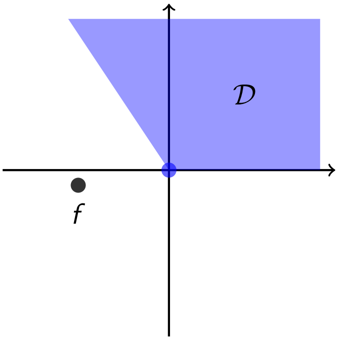
- The **natural extension** is then

$$\begin{aligned}\mathcal{E}(\mathcal{A}) &= (\operatorname{posi}\mathcal{A}+\mathcal{L}_{\geq})\cup\mathcal{L}_{\geq}\\ &= (\operatorname{posi}\mathcal{M}+\mathbb{R}_{>}\cdot 1_{\mathcal{X}}+\mathcal{L}_{\geq})\cup\mathcal{L}_{\geq}\\ &= (\operatorname{posi}\mathcal{M}+\mathcal{L}_{\gtrdot})\cup\mathcal{L}_{\geq}\end{aligned}$$



From acceptable gambles to lower & upper expectations

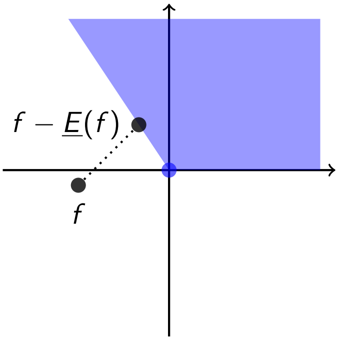
- Consider a coherent set of acceptable gambles $\mathcal{D}$ and a gamble f
- We infer lower and upper expectations for f using the betting game interpretation



From acceptable gambles to lower & upper expectations

- Consider a coherent set of acceptable gambles $\mathcal{D}$ and a gamble f
- We infer lower and upper expectations for f using the betting game interpretation
- **Lower expectation**

$$\underline{E}(f) := \sup\{\alpha \in \mathbb{R} : f - \alpha \in \mathcal{D}\}$$



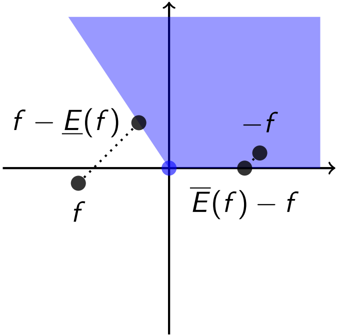
From acceptable gambles to lower & upper expectations

- Consider a coherent set of acceptable gambles $\mathcal{D}$ and a gamble f
- We infer lower and upper expectations for f using the betting game interpretation
- **Lower expectation**

$$\underline{E}(f) := \sup\{\alpha \in \mathbb{R} : f - \alpha \in \mathcal{D}\}$$

- **Upper expectation**

$$\overline{E}(f) := \inf\{\beta \in \mathbb{R} : \beta - f \in \mathcal{D}\}$$



Natural extension to lower and upper expectations

- Consider marginal gambles $\mathcal{M}$, the corresponding $\mathcal{A}$, and some gamble f
- **Lower expectation**

$$\underline{E}_*(f) := \sup\{\alpha \in \mathbb{R} : f - \alpha \in \mathcal{E}(\mathcal{A})\}$$

Natural extension to lower and upper expectations

- Consider marginal gambles $\mathcal{M}$, the corresponding $\mathcal{A}$, and some gamble f
- **Lower expectation**

$$\begin{aligned}\underline{E}_*(f) &:= \sup\{\alpha \in \mathbb{R} : f - \alpha \in \mathcal{E}(\mathcal{A})\} \\ &= \sup\left\{\alpha \in \mathbb{R} : f - \alpha \in (\operatorname{posi}\mathcal{M} + \mathcal{L}_{\gtrdot}) \cup \mathcal{L}_{\geq}\right\}\end{aligned}$$

Natural extension to lower and upper expectations

- Consider marginal gambles $\mathcal{M}$, the corresponding $\mathcal{A}$, and some gamble f

- **Lower expectation**

$$
\begin{aligned}
\underline{E}_*(f) &:= \sup\{\alpha \in \mathbb{R} : f - \alpha \in \mathcal{E}(\mathcal{A})\} \\
&= \sup\left\{\alpha \in \mathbb{R} : f - \alpha \in (\operatorname{posi}\mathcal{M} + \mathcal{L}_{\gtrdot}) \cup \mathcal{L}_{\geq}\right\} \\
&= \sup\left\{\alpha \in \mathbb{R} : (f - \alpha \gtrdot h, h \in \operatorname{posi}\mathcal{M}) \text{ or } f - \alpha \geq 0\right\}
\end{aligned}
$$

Natural extension to lower and upper expectations

- Consider marginal gambles $\mathcal{M}$, the corresponding $\mathcal{A}$, and some gamble f

- **Lower expectation**

$$\begin{aligned}
\underline{E}_*(f) &:= \sup\{\alpha \in \mathbb{R} : f - \alpha \in \mathcal{E}(\mathcal{A})\} \\
&= \sup\Big\{\alpha \in \mathbb{R} : f - \alpha \in (\operatorname{posi}\mathcal{M} + \mathcal{L}_{\gtrdot}) \cup \mathcal{L}_{\geq}\Big\} \\
&= \sup\Big\{\alpha \in \mathbb{R} : (f - \alpha \gtrdot h, h \in \operatorname{posi}\mathcal{M}) \text{ or } f - \alpha \geq 0\Big\} \\
&= \sup\Big\{\alpha \in \mathbb{R} : f - h \geq \alpha, h \in \operatorname{posi}\mathcal{M} \cup \{0\}\Big\}
\end{aligned}$$

Natural extension to lower and upper expectations

- Consider marginal gambles $\mathcal{M}$, the corresponding $\mathcal{A}$, and some gamble f
- **Lower expectation**

$$
\begin{aligned}
\underline{E}_*(f) &:= \sup\{\alpha \in \mathbb{R} : f - \alpha \in \mathcal{E}(\mathcal{A})\} \\
&= \sup\left\{\alpha \in \mathbb{R} : f - \alpha \in (\operatorname{posi}\mathcal{M} + \mathcal{L}_{\gtrdot}) \cup \mathcal{L}_{\geq}\right\} \\
&= \sup\left\{\alpha \in \mathbb{R} : (f - \alpha \gtrdot h, h \in \operatorname{posi}\mathcal{M}) \text{ or } f - \alpha \geq 0\right\} \\
&= \sup\left\{\alpha \in \mathbb{R} : f - h \geq \alpha, h \in \operatorname{posi}\mathcal{M} \cup \{0\}\right\} \\
&= \sup_{h \in \operatorname{posi}\mathcal{M} \cup \{0\}} \min(f - h)
\end{aligned}
$$

Natural extension to lower and upper expectations

- Consider marginal gambles $\mathcal{M}$, the corresponding $\mathcal{A}$, and some gamble f
- **Lower expectation**

$$
\begin{aligned}
\underline{E}_*(f) &:= \sup\{\alpha \in \mathbb{R} : f - \alpha \in \mathcal{E}(\mathcal{A})\} \\
&= \sup\left\{\alpha \in \mathbb{R} : f - \alpha \in (\operatorname{posi}\mathcal{M} + \mathcal{L}_{\gtrdot}) \cup \mathcal{L}_{\geq}\right\} \\
&= \sup\left\{\alpha \in \mathbb{R} : (f - \alpha \gtrdot h, h \in \operatorname{posi}\mathcal{M}) \text{ or } f - \alpha \geq 0\right\} \\
&= \sup\left\{\alpha \in \mathbb{R} : f - h \geq \alpha, h \in \operatorname{posi}\mathcal{M} \cup \{0\}\right\} \\
&= \sup_{h \in \operatorname{posi}\mathcal{M} \cup \{0\}} \min(f - h) \qquad \text{(linear optimization problem over a cone)}
\end{aligned}
$$

Natural extension to lower and upper expectations

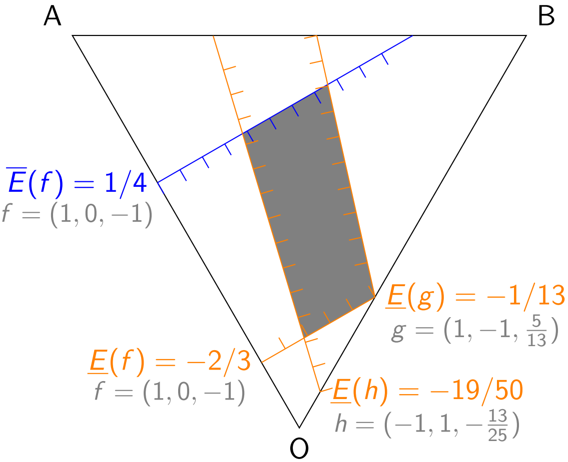
- Consider marginal gambles $\mathcal{M}$, the corresponding $\mathcal{A}$, and some gamble f
- **Lower expectation**

$$
\begin{aligned}
\underline{E}_*(f) := \;& \sup\{\alpha \in \mathbb{R} : f - \alpha \in \mathcal{E}(\mathcal{A})\} \\
=\;& \sup\left\{\alpha \in \mathbb{R} : f - \alpha \in (\operatorname{posi}\mathcal{M} + \mathcal{L}_{\gtrdot}) \cup \mathcal{L}_{\geq}\right\} \\
=\;& \sup\left\{\alpha \in \mathbb{R} : (f - \alpha \gtrdot h, h \in \operatorname{posi}\mathcal{M}) \text{ or } f - \alpha \geq 0\right\} \\
=\;& \sup\left\{\alpha \in \mathbb{R} : f - h \geq \alpha, h \in \operatorname{posi}\mathcal{M} \cup \{0\}\right\} \\
=\;& \sup_{h \in \operatorname{posi}\mathcal{M} \cup \{0\}} \min(f - h) \qquad \text{(linear optimization problem over a cone)}
\end{aligned}
$$

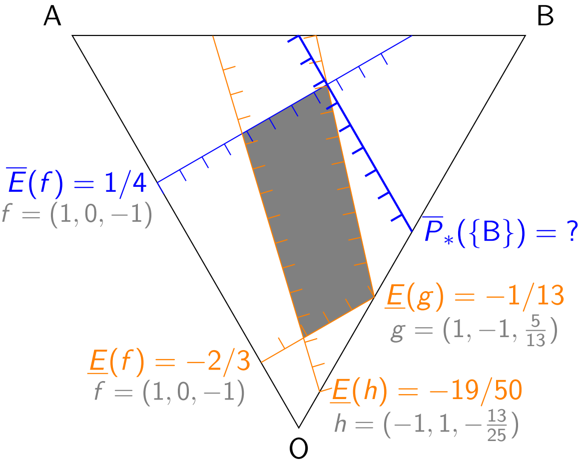
- **Upper expectation**

$$
\overline{E}_*(f) = \inf_{h \in \operatorname{posi}\mathcal{M} \cup \{0\}} \max(f + h) \qquad \text{(linear optimization problem over a cone)}
$$

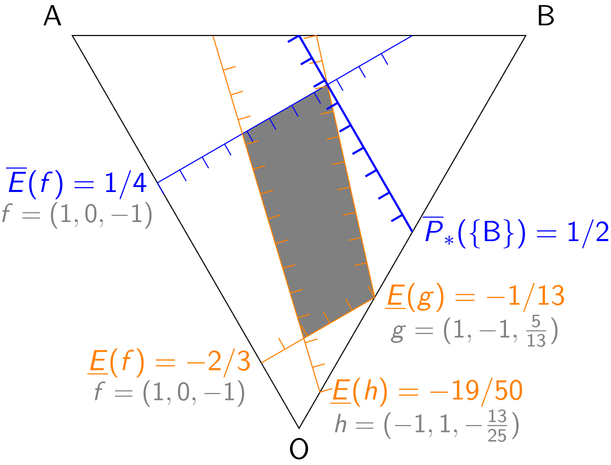
Natural extension illustration



Natural extension illustration



Natural extension illustration



Basic setup and axioms

Basic setup of **the theory of coherent interval expectations**:

- Random variable X
- Set of outcomes $\mathcal{X}$
- Set of functions $\mathcal{F} \subseteq (\mathcal{X} \to \mathbb{R})$
- Each $f \in \mathcal{F}$ is assigned a **lower or upper expectation value**, leading to a set of marginal gambles $\mathcal{M}$

Axioms

The **lower and upper expectation** operators $\underline{E}$ and $\overline{E}$ must

Basic setup and axioms

Basic setup of **the theory of coherent interval expectations**:

- Random variable X
- Set of outcomes $\mathcal{X}$
- Set of functions $\mathcal{F} \subseteq (\mathcal{X} \to \mathbb{R})$
- Each $f \in \mathcal{F}$ is assigned a **lower or upper expectation value**, leading to a set of marginal gambles $\mathcal{M}$

Axioms

The **lower and upper expectation** operators $\underline{E}$ and $\overline{E}$ must

Avoid sure loss The assessment $\mathcal{A}$ corresponding to $\mathcal{M}$ must avoid sure loss:

$$\operatorname{posi} \mathcal{A} \cap \mathcal{L}_{\lessdot} \neq \emptyset$$

Basic setup and axioms

Basic setup of **the theory of coherent interval expectations**:

- Random variable X
- Set of outcomes $\mathcal{X}$
- Set of functions $\mathcal{F} \subseteq (\mathcal{X} \to \mathbb{R})$
- Each $f \in \mathcal{F}$ is assigned a **lower or upper expectation value**, leading to a set of marginal gambles $\mathcal{M}$

Axioms

The **lower and upper expectation** operators $\underline{E}$ and $\overline{E}$ must

Avoid sure loss The assessment $\mathcal{A}$ corresponding to $\mathcal{M}$ must avoid sure loss:

$$\operatorname{posi}\mathcal{M} \cap \mathcal{L}_{\lessdot} \neq \emptyset$$

Basic setup and axioms

Basic setup of **the theory of coherent interval expectations**:

- Random variable X
- Set of outcomes $\mathcal{X}$
- Set of functions $\mathcal{F} \subseteq (\mathcal{X} \to \mathbb{R})$
- Each $f \in \mathcal{F}$ is assigned a **lower or upper expectation value**, leading to a set of marginal gambles $\mathcal{M}$

Axioms

The **lower and upper expectation** operators $\underline{E}$ and $\overline{E}$ must

Avoid sure loss The assessment $\mathcal{A}$ corresponding to $\mathcal{M}$ must avoid sure loss:

$$\inf_{h \in \operatorname{posi} \mathcal{M}} \max h \geq 0$$

Basic setup and axioms

Basic setup of **the theory of coherent interval expectations**:

- Random variable X
- Set of outcomes $\mathcal{X}$
- Set of functions $\mathcal{F} \subseteq (\mathcal{X} \to \mathbb{R})$
- Each $f \in \mathcal{F}$ is assigned a **lower or upper expectation value**, leading to a set of marginal gambles $\mathcal{M}$

Axioms

The **lower and upper expectation** operators $\underline{E}$ and $\overline{E}$ must

Avoid sure loss: The assessment $\mathcal{A}$ corresponding to $\mathcal{M}$ must avoid sure loss:

$$\inf_{h \in \operatorname{posi} \mathcal{M}} \max h \geq 0$$

Be coherent: The natural extension is identical to the specified expectation for all gambles $f \in \mathcal{F}$ (worked-out for lower expectation):

$$\underline{E}(f) = \underline{E}_*(f)$$

Basic setup and axioms

Basic setup of **the theory of coherent interval expectations**:

- Random variable X
- Set of outcomes $\mathcal{X}$
- Set of functions $\mathcal{F} \subseteq (\mathcal{X} \to \mathbb{R})$
- Each $f \in \mathcal{F}$ is assigned a **lower or upper expectation value**, leading to a set of marginal gambles $\mathcal{M}$

Axioms

The **lower and upper expectation** operators $\underline{E}$ and $\overline{E}$ must

Avoid sure loss: The assessment $\mathcal{A}$ corresponding to $\mathcal{M}$ must avoid sure loss:

$$\inf_{h \in \operatorname{posi} \mathcal{M}} \max h \geq 0$$

Be coherent: The natural extension is identical to the specified expectation for all gambles $f \in \mathcal{F}$ (worked-out for lower expectation):

$$\underline{E}(f) \geq \sup_{h \in \operatorname{posi} \mathcal{M} \cup \{0\}} \min(f - h)$$

Basic setup and axioms

Basic setup of **the theory of coherent interval expectations**:

- Random variable X
- Set of outcomes $\mathcal{X}$
- Set of functions $\mathcal{F} \subseteq (\mathcal{X} \to \mathbb{R})$
- Each $f \in \mathcal{F}$ is assigned a **lower or upper expectation value**, leading to a set of marginal gambles $\mathcal{M}$

Axioms

The **lower and upper expectation** operators $\underline{E}$ and $\overline{E}$ must

Avoid sure loss The assessment $\mathcal{A}$ corresponding to $\mathcal{M}$ must avoid sure loss:

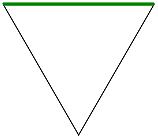
$$\inf_{h \in \operatorname{posi} \mathcal{M}} \max h \geq 0$$

Be coherent The natural extension is identical to the specified expectation for all gambles $f \in \mathcal{F}$ (worked-out for lower expectation):

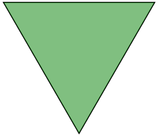
$$\inf_{h \in \operatorname{posi} \mathcal{M} \cup \{0\}} \max(h - (f - \underline{E}(f))) \geq 0$$

Special types of lower expectations

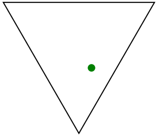
- The vacuous lower expectation $\underline{E}^S$ relative to some event $S \subseteq \mathcal{X}$ is defined by $\underline{E}^S(f) := \min_{x \in S} f(x)$ and expresses that $X \in S$ and nothing more



- The vacuous lower expectation $\underline{E}^{\mathcal{X}}$ expresses complete ignorance

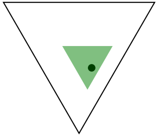


- The *linear* expectation E_p corresponding to some pmf p is also a coherent lower expectation



- The linear-vacuous lower expectation $\underline{E}_{p,\varepsilon}$ is as well:

$$\underline{E}_{p,\varepsilon}(f) := (1-\varepsilon)E_p(f) + \varepsilon \min_{x \in \mathcal{X}} f(x)$$



Some properties of coherent lower and upper expectations

Conjugacy $\overline{E}(f) = -\underline{E}(-f)$

Boundedness $\min_{x\in\mathcal{X}} f(x) \leq \underline{E}(f) \leq \overline{E}(f) \leq \max_{x\in\mathcal{X}} f(x)$

Positive homogeneity $\underline{E}(\lambda f) = \lambda \underline{E}(f)$ for all $\lambda > 0$

Some properties of coherent lower and upper expectations

Conjugacy $\overline{E}(f) = -\underline{E}(-f)$

Boundedness $\min_{x\in\mathcal{X}} f(x) \leq \underline{E}(f) \leq \overline{E}(f) \leq \max_{x\in\mathcal{X}} f(x)$

Positive homogeneity $\underline{E}(\lambda f) = \lambda \underline{E}(f)$ for all $\lambda > 0$

Monotonicity if $f \leq g$ then $\underline{E}(f) \leq \underline{E}(g)$

Constant additivity $\underline{E}(f + \mu 1_{\mathcal{X}}) = \underline{E}(f) + \mu$ for all $\mu \in \mathbb{R}$

Mixed super/sub-linearity

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

Some properties of coherent lower and upper expectations

Conjugacy $\overline{E}(f) = -\underline{E}(-f)$

Boundedness $\min_{x\in\mathcal{X}} f(x) \leq \underline{E}(f) \leq \overline{E}(f) \leq \max_{x\in\mathcal{X}} f(x)$

Positive homogeneity $\underline{E}(\lambda f) = \lambda \underline{E}(f)$ for all $\lambda > 0$

Monotonicity if $f \leq g$ then $\underline{E}(f) \leq \underline{E}(g)$

Constant additivity $\underline{E}(f + \mu 1_{\mathcal{X}}) = \underline{E}(f) + \mu$ for all $\mu \in \mathbb{R}$

Mixed super/sub-linearity

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

Convex mixtures if $\underline{E}^1$ and $\underline{E}^2$ are coherent, then so is $\varepsilon \underline{E}^1 + (1-\varepsilon)\underline{E}^2$ for all $0 \leq \varepsilon \leq 1$

Lower expectations on linear function spaces

If $\mathcal{F}$ is a linear space $(f, g \in \mathcal{F}$ then $af + bg \in \mathcal{F}$ for all $a, b \in \mathbb{R})$
then the axioms simplify:

Boundedness $\underline{E}(f) \geq \min_{x \in \mathcal{X}} f(x)$ for all $f \in \mathcal{F}$

Positive homogeneity $\underline{E}(\lambda f) = \lambda \underline{E}(f)$ for all $f \in \mathcal{F}$ and $\lambda > 0$

Super-linearity $\underline{E}(f + g) \geq \underline{E}(f) + \underline{E}(g)$ for all $f, g \in \mathcal{F}$

Lower envelopes

- Let Γ be some index set
- Consider a set $\{\underline{E}^\gamma : \gamma \in \Gamma\}$ of coherent lower expectations on $\mathcal{F}$
- Then its lower envelope $\underline{E}$ defined by $\underline{E}(f) := \inf_{\gamma\in\Gamma} \underline{E}^\gamma(f)$ on $\mathcal{F}$ is coherent

Lower envelopes

- Let Γ be some index set
- Consider a set $\{\underline{E}^\gamma : \gamma \in \Gamma\}$ of coherent lower expectations on $\mathcal{F}$
- Then its lower envelope $\underline{E}$ defined by $\underline{E}(f) := \inf_{\gamma\in\Gamma} \underline{E}^\gamma(f)$ on $\mathcal{F}$ is coherent

Lower envelopes of credal sets

- Let $\mathcal{C}$ be a non-empty credal set
- Then the lower envelope $\underline{E}^{\mathcal{C}}$ defined by $\underline{E}^{\mathcal{C}}(f) := \inf_{p\in\mathcal{C}} E_p(f)$ is coherent

Lower envelopes

- Let Γ be some index set
- Consider a set $\{\underline{E}^{\gamma} : \gamma \in \Gamma\}$ of coherent lower expectations on $\mathcal{F}$
- Then its lower envelope $\underline{E}$ defined by $\underline{E}(f) := \inf_{\gamma\in\Gamma} \underline{E}^{\gamma}(f)$ on $\mathcal{F}$ is coherent

Lower envelopes of credal sets

- Let $\mathcal{C}$ be a non-empty credal set
- Then the lower envelope $\underline{E}^{\mathcal{C}}$ defined by $\underline{E}^{\mathcal{C}}(f) := \inf_{p\in\mathcal{C}} E_p(f)$ is coherent

Credal sets from lower expectations

- Let $\underline{E}$ be a lower expectation on $\mathcal{F}$ *avoiding sure loss*
- Then the credal set $\mathcal{C}^{\underline{E}} := \{p \in \mathcal{P}_X : E_p(f) \geq \underline{E}(f) \text{ for all } f \in \mathcal{F}\}$ is non-empty

The categorical prediction task

What is the color of the next ball to be drawn from an urn?

The categorical prediction task

What is the color of the next ball to be drawn from an urn?

Setup

- Outcome space $\mathcal{X}$ (of ball colors: Red, Green, Black, White,...)
- We have $n \in \mathbb{N}$ observations $\boldsymbol{x} = (x_1, \dots, x_n)$ (sequence of colors of previous balls drawn)
- Draw inferences for or make decisions related to the next observation X_{n+1} (color of next ball drawn)

The categorical prediction task

What is the color of the next ball to be drawn from an urn?

Setup

- Outcome space $\mathcal{X}$ (of ball colors: Red, Green, Black, White, . . .)
- We have $n \in \mathbb{N}$ observations $\boldsymbol{x} = (x_1, \ldots, x_n)$
  (sequence of colors of previous balls drawn)
- Draw inferences for or make decisions related to the next observation X_{n+1}
  (color of next ball drawn)

Exchangeability assumption

- *The order of the observations is irrelevant*
  (inferences should remain the same when $(\mathrm{R}, \mathrm{G}, \mathrm{R})$ or $(\mathrm{R}, \mathrm{R}, \mathrm{G})$ has been observed)
- Observed occurrence vector $\boldsymbol{n} \in \mathbb{N}^{\mathcal{X}}$ with $n_z := \sum_{k=1}^{n} \delta_{z x_k}$ and $n_S := \sum_{x \in S} n_x$
  (if $\boldsymbol{x} = (\mathrm{R}, \mathrm{G}, \mathrm{R}, \mathrm{W})$ then $n_\mathrm{R} = 2$, $n_\mathrm{G} = 1$, $n_\mathrm{W} = 1$, and $n_{\neg\mathrm{W}} = 3$)

Learning lower expectations for categorical prediction

Idea

Add epistemic uncertainty using $s > 0$ pseudo-observations (of unknown color) and predict according to the ‘observed’ frequency

Learning lower expectations for categorical prediction

Idea

Add epistemic uncertainty using $s > 0$ pseudo-observations (of unknown color) and predict according to the 'observed' frequency

Predictive inference

- Outcome prediction: $\underline{P}_s(\{x\}|\boldsymbol{n}) = \frac{n_x}{s+n}$ and $\overline{P}_s(\{x\}|\boldsymbol{n}) = \frac{s+n_x}{s+n}$

Learning lower expectations for categorical prediction

Idea

Add epistemic uncertainty using $s > 0$ pseudo-observations (of unknown color) and predict according to the ‘observed’ frequency

Predictive inference

- Event prediction: $\underline{P}_s(S|\boldsymbol{n}) = \frac{n_S}{s+n}$ and $\overline{P}_s(S|\boldsymbol{n}) = \frac{s+n_S}{s+n}$

Learning lower expectations for categorical prediction

Idea

Add epistemic uncertainty using $s > 0$ pseudo-observations (of unknown color) and predict according to the 'observed' frequency

Predictive inference

- Event prediction: $\underline{P}_s(S|\boldsymbol{n}) = \frac{n_S}{s+n}$ and $\overline{P}_s(S|\boldsymbol{n}) = \frac{s+n_S}{s+n}$
- Lower expectation for $f \in \mathcal{L}$:

$$\underline{E}_s(f|\boldsymbol{n}) = \frac{n}{s+n} E_p(f) + \frac{s}{s+n} \min_{x\in\mathcal{X}} f(x) \quad \text{with } p_x := \frac{n_x}{n}$$

Learning lower expectations for categorical prediction

Idea

Add epistemic uncertainty using $s > 0$ pseudo-observations (of unknown color) and predict according to the 'observed' frequency

Predictive inference

- Event prediction: $\underline{P}_s(S|\boldsymbol{n}) = \frac{n_S}{s+n}$ and $\overline{P}_s(S|\boldsymbol{n}) = \frac{s+n_S}{s+n}$
- Lower expectation for $f \in \mathcal{L}$:

$$\underline{E}_s(f|\boldsymbol{n}) = \frac{n}{s+n} E_p(f) + \frac{s}{s+n} \min_{x\in\mathcal{X}} f(x) \quad \text{with } p_x := \frac{n_x}{n}$$

Properties of $\underline{E}_s(\cdot|\boldsymbol{n})$

- linear-vacuous model; vacuous for $n = 0$; more precise with more observations
- does not depend on a specific categorization $\mathcal{X}$
- immediate prediction model of the *imprecise Dirichlet-multinomial model* ID(M)M

Learning lower expectations for categorical prediction

Idea

Add epistemic uncertainty using $s > 0$ pseudo-observations (of unknown color) and predict according to the 'observed' frequency

Predictive inference

- Event prediction: $\underline{P}_s(S|\boldsymbol{n}) = \frac{n_S}{s+n}$ and $\overline{P}_s(S|\boldsymbol{n}) = \frac{s+n_S}{s+n}$
- Lower expectation for $f \in \mathcal{L}$:

$$\underline{E}_s(f|\boldsymbol{n}) = \frac{n}{s+n} E_p(f) + \frac{s}{s+n} \min_{x\in\mathcal{X}} f(x) \quad \text{with } p_x := \frac{n_x}{n}$$

Example setup

- $\boldsymbol{x} = (\mathrm{R}, \mathrm{G}, \mathrm{R}, \mathrm{W}, \mathrm{W})$
- $s = 2$

Inferences

- $\underline{P}_2(\{B, W\}|\boldsymbol{n}) = ?$
- $\overline{P}_2(\{B, W\}|\boldsymbol{n}) = ?$

Learning lower expectations for categorical prediction

Idea

Add epistemic uncertainty using $s > 0$ pseudo-observations (of unknown color) and predict according to the 'observed' frequency

Predictive inference

- Event prediction: $\underline{P}_s(S|\boldsymbol{n}) = \frac{n_S}{s+n}$ and $\overline{P}_s(S|\boldsymbol{n}) = \frac{s+n_S}{s+n}$
- Lower expectation for $f \in \mathcal{L}$:

$$\underline{E}_s(f|\boldsymbol{n}) = \frac{n}{s+n} E_p(f) + \frac{s}{s+n} \min_{x\in\mathcal{X}} f(x) \quad \text{with } p_x := \frac{n_x}{n}$$

Example setup

- $\boldsymbol{x} = (\mathrm{R}, \mathrm{G}, \mathrm{R}, \mathrm{W}, \mathrm{W})$
- $s = 2$

Inferences

- $\underline{P}_2(\{B, W\}|\boldsymbol{n}) = 2/7$
- $\overline{P}_2(\{B, W\}|\boldsymbol{n}) = 4/7$

Multivariate interval expectation: basic setup & basic idea

- Index set $N = \{1, \ldots, n\}$
- Multivariate variable $\boldsymbol{X} = (X_1, X_2, \ldots, X_n)$
- Set of possible outcomes $\boldsymbol{x} \in \mathcal{X} = \mathcal{X}_1 \times \mathcal{X}_2 \times \ldots \times \mathcal{X}_n$
- A **joint lower expectation** $\underline{E}^{\boldsymbol{X}}$ is specified on some set of functions $\mathcal{F} \subseteq \mathcal{L}$

Formulate an appropriate function on $\mathcal{L}$
whose lower expectation provides the desired inference;
calculate the lower expectation using natural extension

Marginal lower expectations

- A marginal lower expectation is defined for a subset of the random variables
- Let $K \subseteq N$, then
  - $\boldsymbol{X}_K := (X_k : k \in K)$ and $\boldsymbol{x}_K \in \mathcal{X}_K := \times_{k \in K} \mathcal{X}_k$
  - Notation: $\underline{E}^{\boldsymbol{X}_K}$ is the $\boldsymbol{X}_K$-marginal of the joint lower expectation $\underline{E}^{\boldsymbol{X}}$

Marginal lower expectations

- A marginal lower expectation is defined for a subset of the random variables
- Let $K \subseteq N$, then
  - $\boldsymbol{X}_K := (X_k : k \in K)$ and $\boldsymbol{x}_K \in \mathcal{X}_K := \times_{k \in K} \mathcal{X}_k$
  - Notation: $\underline{E}^{\boldsymbol{X}_K}$ is the $\boldsymbol{X}_K$-marginal of the joint lower expectation $\underline{E}^{\boldsymbol{X}}$
- The **marginal lower expectation** for $f \in \mathcal{L}_K$ is obtained by calculating the joint natural extension of its *cylindrical extension*:

$$\underline{E}^{\boldsymbol{X}_K}(f) := \underline{E}_*^{\boldsymbol{X}}(\uparrow_{\boldsymbol{x}_{N \setminus K}} f) \quad \text{where} \quad (\uparrow_{\boldsymbol{x}_{N \setminus K}} f)(\boldsymbol{x}) := f(\boldsymbol{x}_K)$$

Marginal lower expectation example

Example setup

- Two random variables X_1 and X_2 with outcome spaces $\mathcal{X}_1 = \{0, 1\}$ and $\mathcal{X}_2 = \{-, +\}$
- Joint lower expectation $\underline{E}^{\boldsymbol{X}} := \underline{E}_{\boldsymbol{p},\varepsilon}$ with ε unspecified and p given on the right (in black), together with its marginals (in green)

p	0	1	p^{X_2}
$-$	$3/9$	$1/9$	$4/9$
$+$	$2/9$	$3/9$	$5/9$
p^{X_1}	$5/9$	$4/9$	

Inference

Marginal lower expectation $\underline{E}^{X_1}$

Marginal lower expectation example

Example setup

- Two random variables X_1 and X_2 with outcome spaces $\mathcal{X}_1 = \{0, 1\}$ and $\mathcal{X}_2 = \{-, +\}$
- Joint lower expectation $\underline{E}^{\boldsymbol{X}} := \underline{E}_{p,\varepsilon}$ with ε unspecified and p given on the right (in black), together with its marginals (in green)

p	0	1	p^{X_2}
$-$	$3/9$	$1/9$	$4/9$
$+$	$2/9$	$3/9$	$5/9$
p^{X_1}	$5/9$	$4/9$	

Inference

Marginal lower expectation $\underline{E}^{X_1}$

Solution

$$\begin{aligned}\underline{E}_{p,\varepsilon}(\uparrow_{\mathcal{X}_2} f) &= (1-\varepsilon)E_p(\uparrow_{\mathcal{X}_2} f) + \varepsilon \min_{\boldsymbol{x}\in\mathcal{X}}(\uparrow_{\mathcal{X}_2} f)(x_1, x_2) \\ &= (1-\varepsilon)E_{p^{X_1}}(f) + \varepsilon \min_{x_1\in\mathcal{X}_1} f(x_1) \\ &= \underline{E}_{p^{X_1},\varepsilon}(f)\end{aligned}$$

Conditional lower expectations

- A conditional lower expectation is determined by
  - assuming some event is known to be true, or specifically
  - assuming some random variables take some given values: $\boldsymbol{X}_K = \boldsymbol{x}_K$, with $K \subset N$
- Notation: $\underline{E}^{\boldsymbol{X}_{N\setminus K}}(\cdot|\boldsymbol{X}_K = \boldsymbol{x}_K)$ is the $\boldsymbol{X}_{N\setminus K}$-conditional of the joint lower expectation $\underline{E}^{\boldsymbol{X}}$

Conditional lower expectations

- A conditional lower expectation is determined by
  - assuming some event is known to be true, or specifically
  - assuming some random variables take some given values: $\boldsymbol{X}_K = \boldsymbol{x}_K$, with $K \subset N$
- Notation: $\underline{E}^{\boldsymbol{X}_{N\setminus K}}(\cdot|\boldsymbol{X}_K = \boldsymbol{x}_K)$ is the $\boldsymbol{X}_{N\setminus K}$-conditional of the joint lower expectation $\underline{E}^{\boldsymbol{X}}$
- The **conditional lower expectation** for $f \in \mathcal{L}_{N\setminus K}$ is the maximum acceptable buying price for the corresponding *called-off gamble*:

$$\underline{E}^{\boldsymbol{X}_{N\setminus K}}(f|\boldsymbol{X}_K = \boldsymbol{x}_K) := \max\left\{\alpha \in \mathbb{R} : \underline{E}^{\boldsymbol{X}}_*\left(\uparrow_{\boldsymbol{x}_K}\left(f - \alpha \cdot 1_{\mathcal{X}_{N\setminus K}}\right) \cdot 1_{\boldsymbol{X}_K = \boldsymbol{x}_K}\right) \geq 0\right\}$$

  if $\underline{P}(\boldsymbol{X}_K = \boldsymbol{x}_K) > 0$, otherwise $\underline{E}^{\boldsymbol{X}_{N\setminus K}}(\cdot|\boldsymbol{X}_K = \boldsymbol{x}_K) := \underline{E}^{\mathcal{X}_{N\setminus K}}$

Conditional lower expectations

- A conditional lower expectation is determined by
  - assuming some event is known to be true, or specifically
  - assuming some random variables take some given values: $\boldsymbol{X}_K = \boldsymbol{x}_K$, with $K \subset N$
- Notation: $\underline{E}^{\boldsymbol{X}_{N\setminus K}}(\cdot|\boldsymbol{X}_K = \boldsymbol{x}_K)$ is the $\boldsymbol{X}_{N\setminus K}$-conditional of the joint lower expectation $\underline{E}^{\boldsymbol{X}}$
- The **conditional lower expectation** for $f \in \mathcal{L}_{N\setminus K}$ is the maximum acceptable buying price for the corresponding *called-off gamble*:

$$\underline{E}^{\boldsymbol{X}_{N\setminus K}}(f|\boldsymbol{X}_K = \boldsymbol{x}_K) := \max\Big\{\alpha \in \mathbb{R} : \underline{E}_*^{\boldsymbol{X}}\big(\uparrow_{\boldsymbol{x}_K}(f - \alpha \cdot 1_{\mathcal{X}_{N\setminus K}}) \cdot 1_{\boldsymbol{X}_K = \boldsymbol{x}_K}\big) \geq 0\Big\}$$

if $\underline{P}(\boldsymbol{X}_K = \boldsymbol{x}_K) > 0$, otherwise $\underline{E}^{\boldsymbol{X}_{N\setminus K}}(\cdot|\boldsymbol{X}_K = \boldsymbol{x}_K) := \underline{E}^{\mathcal{X}_{N\setminus K}}$

This is called *conditioning by natural extension* or the *generalized Bayes's rule*

Conditional lower expectation example

Example setup

- Two random variables X_1 and X_2 with outcome spaces $\mathcal{X}_1 = \{0, 1\}$ and $\mathcal{X}_2 = \{-, +\}$
- Joint lower expectation $\underline{E}^{\boldsymbol{X}} := \underline{E}_{p,\varepsilon}$ with ε unspecified and p given on the right (in black), together with its marginals (in green)

p	0	1	p^{X_2}
$-$	$3/9$	$1/9$	$4/9$
$+$	$2/9$	$3/9$	$5/9$
p^{X_1}	$5/9$	$4/9$	

Inference

Conditional lower expectation $\underline{E}^{X_1}(\cdot|X_2 = +)$

Solution (start)

$$\underline{E}^{X_1}(f|X_2 = +) = \max\Big\{\alpha \in \mathbb{R} : \underline{E}_{p,\varepsilon}\big(\uparrow_{X_2} (f - \alpha \cdot 1_{\mathcal{X}_1}) \cdot 1_{X_2=+}\big) \geq 0\Big\}$$

Conditional lower expectations example

Solution (continuation)

$$\underline{E}^{X_1}(f|X_2 = +) = \max\left\{\alpha \in \mathbb{R} : \underline{E}_{p,\varepsilon}\left(\uparrow_{X_2} (f - \alpha \cdot 1_{\mathcal{X}_1}) \cdot 1_{X_2=+}\right) \geq 0\right\}$$

Conditional lower expectations example

Solution (continuation)

$$\underline{E}^{X_1}(f|X_2=+)=\max\Big\{\alpha\in\mathbb{R}:\underline{E}_{\boldsymbol{p},\varepsilon}\big(\uparrow_{X_2}(f-\alpha\cdot 1_{\mathcal{X}_1})\cdot 1_{X_2=+}\big)\geq 0\Big\}$$

$$=\max\Big\{\alpha\in\mathbb{R}:(1-\varepsilon)E_p\big(\uparrow_{X_2}(f-\alpha\cdot 1_{\mathcal{X}_1})\cdot 1_{X_2=+}\big)+\varepsilon\min_{\boldsymbol{x}\in\mathcal{X}}(f(x_1)-\alpha)\cdot 1_{X_2=+}(x_2)\geq 0\Big\}$$

Conditional lower expectations example

Solution (continuation)

$$\underline{E}^{X_1}(f|X_2=+) = \max\left\{\alpha \in \mathbb{R} : \underline{E}_{\boldsymbol{p},\varepsilon}\left(\uparrow_{X_2}\left(f-\alpha\cdot 1_{\mathcal{X}_1}\right)\cdot 1_{X_2=+}\right) \geq 0\right\}$$

$$= \max\left\{\alpha \in \mathbb{R} : (1-\varepsilon)E_p\left(\uparrow_{X_2}\left(f-\alpha\cdot 1_{\mathcal{X}_1}\right)\cdot 1_{X_2=+}\right) + \varepsilon \min_{\boldsymbol{x}\in\mathcal{X}}\left(f(x_1)-\alpha\right)\cdot 1_{X_2=+}(x_2) \geq 0\right\}$$

$$= \max\left\{\alpha \geq \min_{x_1\in\mathcal{X}_1} f(x_1) : (1-\varepsilon)\sum_{x_1\in\mathcal{X}_1} p_{(x_1,+)}\left(f(x_1)-\alpha\right) + \varepsilon \min_{x_1\in\mathcal{X}_1}\left(f(x_1)-\alpha\right) \geq 0\right\}$$

Conditional lower expectations example

Solution (continuation)

$$\underline{E}^{X_1}(f|X_2=+) = \max\left\{\alpha\in\mathbb{R} : \underline{E}_{\boldsymbol{p},\varepsilon}\left(\uparrow_{X_2}\left(f-\alpha\cdot 1_{\mathcal{X}_1}\right)\cdot 1_{X_2=+}\right)\geq 0\right\}$$

$$= \max\left\{\alpha\in\mathbb{R} : (1-\varepsilon)E_{\boldsymbol{p}}\left(\uparrow_{X_2}\left(f-\alpha\cdot 1_{\mathcal{X}_1}\right)\cdot 1_{X_2=+}\right)+\varepsilon\min_{\boldsymbol{x}\in\mathcal{X}}\left(f(x_1)-\alpha\right)\cdot 1_{X_2=+}(x_2)\geq 0\right\}$$

$$= \max\left\{\alpha\geq\min_{x_1\in\mathcal{X}_1}f(x_1) : (1-\varepsilon)\sum_{x_1\in\mathcal{X}_1}p_{(x_1,+)}\left(f(x_1)-\alpha\right)+\varepsilon\min_{x_1\in\mathcal{X}_1}\left(f(x_1)-\alpha\right)\geq 0\right\}$$

$$= \max\left\{\alpha\geq\min_{x_1\in\mathcal{X}_1}f(x_1) : (1-\varepsilon)\tfrac{5}{9}\sum_{x_1\in\mathcal{X}_1}p_{x_1}^{X_1|+}\left(f(x_1)-\alpha\right)+\varepsilon\left(\min_{x_1\in\mathcal{X}_1}f(x_1)-\alpha\right)\geq 0\right\}$$

Conditional lower expectations example

Solution (continuation)

$$\underline{E}^{X_1}(f|X_2=+) = \max\Big\{\alpha \in \mathbb{R} : \underline{E}_{p,\varepsilon}\big(\uparrow_{X_2}(f-\alpha\cdot 1_{\mathcal{X}_1})\cdot 1_{X_2=+}\big) \geq 0\Big\}$$

$$= \max\Big\{\alpha \in \mathbb{R} : (1-\varepsilon)E_p\big(\uparrow_{X_2}(f-\alpha\cdot 1_{\mathcal{X}_1})\cdot 1_{X_2=+}\big) + \varepsilon \min_{\boldsymbol{x}\in\mathcal{X}} (f(x_1)-\alpha)\cdot 1_{X_2=+}(x_2) \geq 0\Big\}$$

$$= \max\Big\{\alpha \geq \min_{x_1\in\mathcal{X}_1} f(x_1) : (1-\varepsilon)\sum_{x_1\in\mathcal{X}_1} p_{(x_1,+)}(f(x_1)-\alpha) + \varepsilon \min_{x_1\in\mathcal{X}_1}(f(x_1)-\alpha) \geq 0\Big\}$$

$$= \max\Big\{\alpha \geq \min_{x_1\in\mathcal{X}_1} f(x_1) : (1-\varepsilon)\tfrac{5}{9}\sum_{x_1\in\mathcal{X}_1} p^{X_1|+}_{x_1}(f(x_1)-\alpha) + \varepsilon\big(\min_{x_1\in\mathcal{X}_1} f(x_1)-\alpha\big) \geq 0\Big\}$$

$$= \max\Big\{\alpha \geq \min_{x_1\in\mathcal{X}_1} f(x_1) : \alpha \leq (1-\delta)E_{p^{X_1|+}}(f) + \delta \min_{x_1\in\mathcal{X}_1} f(x_1)\Big\}$$

$$\text{with } \delta = \frac{\varepsilon}{(1-\varepsilon)\frac{5}{9}+\varepsilon}$$

Conditional lower expectations example

Solution (continuation)

$$\underline{E}^{X_1}(f|X_2=+) = \max\Big\{\alpha\in\mathbb{R} : \underline{E}_{p,\varepsilon}\big(\uparrow_{X_2}(f-\alpha\cdot 1_{\mathcal{X}_1})\cdot 1_{X_2=+}\big)\geq 0\Big\}$$

$$= \max\Big\{\alpha\in\mathbb{R} : (1-\varepsilon)E_p\big(\uparrow_{X_2}(f-\alpha\cdot 1_{\mathcal{X}_1})\cdot 1_{X_2=+}\big)+\varepsilon\min_{\boldsymbol{x}\in\mathcal{X}}(f(x_1)-\alpha)\cdot 1_{X_2=+}(x_2)\geq 0\Big\}$$

$$= \max\Big\{\alpha\geq\min_{x_1\in\mathcal{X}_1}f(x_1) : (1-\varepsilon)\sum_{x_1\in\mathcal{X}_1}p_{(x_1,+)}(f(x_1)-\alpha)+\varepsilon\min_{x_1\in\mathcal{X}_1}(f(x_1)-\alpha)\geq 0\Big\}$$

$$= \max\Big\{\alpha\geq\min_{x_1\in\mathcal{X}_1}f(x_1) : (1-\varepsilon)\tfrac{5}{9}\sum_{x_1\in\mathcal{X}_1}p^{X_1|+}_{x_1}(f(x_1)-\alpha)+\varepsilon(\min_{x_1\in\mathcal{X}_1}f(x_1)-\alpha)\geq 0\Big\}$$

$$= \max\Big\{\alpha\geq\min_{x_1\in\mathcal{X}_1}f(x_1) : \alpha\leq(1-\delta)E_{p^{X_1|+}}(f)+\delta\min_{x_1\in\mathcal{X}_1}f(x_1)\Big\}$$

$$= \underline{E}_{p^{X_1|+},\delta}(f) \qquad \text{with } \delta=\frac{\varepsilon}{(1-\varepsilon)\frac{5}{9}+\varepsilon}$$

Epistemic irrelevance & Epistemic independence

Consider $\boldsymbol{X} = (X_1, X_2)$ and a joint lower expectation $\underline{E}^{\boldsymbol{X}}$ on $(\mathcal{X} \to \mathbb{R})$

Epistemic irrelevance

- X_2 is *epistemically irrelevant* to X_1 iff $\underline{E}^{X_1}(\cdot|X_2 = x_2) := \underline{E}^{X_1}$ for all $x_2 \in \mathcal{X}_2$
- Epistemic irrelevance of X_2 to X_1 does not imply epistemic irrelevance of X_1 to X_2

Epistemic irrelevance & Epistemic independence

Consider $\boldsymbol{X} = (X_1, X_2)$ and a joint lower expectation $\underline{E}^{\boldsymbol{X}}$ on $(\mathcal{X} \to \mathbb{R})$

Epistemic irrelevance

- X_2 is *epistemically irrelevant* to X_1 iff $\underline{E}^{X_1}(\cdot|X_2 = x_2) := \underline{E}^{X_1}$ for all $x_2 \in \mathcal{X}_2$
- Epistemic irrelevance of X_2 to X_1 does not imply epistemic irrelevance of X_1 to X_2

Epistemic independence

X_1 and X_2 are called *epistemically independent* iff

- X_2 is epistemically irrelevant to X_1
- X_1 is epistemically irrelevant to X_2

Epistemically independent natural extension

Setup

- Consider two random variables X_1 and X_2
- Marginal lower expectations $\underline{E}^{X_1}$ on $\mathcal{F}_1 \subseteq \mathcal{L}_1$ and $\underline{E}^{X_2}$ on $\mathcal{F}_2 \subseteq \mathcal{L}_2$

Natural extension of epistemically independent marginals

- Creating a *joint* lower expectation from epistemically independent marginals is not done as typically with a product

Epistemically independent natural extension

Setup

- Consider two random variables X_1 and X_2
- Marginal lower expectations $\underline{E}^{X_1}$ on $\mathcal{F}_1 \subseteq \mathcal{L}_1$ and $\underline{E}^{X_2}$ on $\mathcal{F}_2 \subseteq \mathcal{L}_2$

Natural extension of epistemically independent marginals

- Creating a *joint* lower expectation from epistemically independent marginals is not done as typically with a product
- Natural extension of the implied assessment is used:

$$\underline{E}^{\boldsymbol{X}}(f) = (\underline{E}^{X_1} \boxtimes \underline{E}^{X_2})(f) := \sup_{g_1, g_2 \in \mathcal{L}} \min_{\boldsymbol{x} \in \mathcal{X}} \Big(f(\boldsymbol{x}) - \big(g_1(\boldsymbol{x}) - \underline{E}_*^{X_1}(g_1(X_1, x_2))\big) - \big(g_2(\boldsymbol{x}) - \underline{E}_*^{X_2}(g_2(x_1, X_2))\big) \Big)$$

for all $f \in \mathcal{L}$

Introduction to imprecise probabilities

SIPTA School 2024, Ghent, Belgium

Erik Quaeghebeur

Eindhoven University of Technology

12 August 2024