IP Scoring Rules: Theory

Jason Konek[a] Ben Levinstein[b]

19 August 2022

[a]Department of Philosophy
Bristol

[b]Department of Philosophy
UIUC

Which class of IP models?

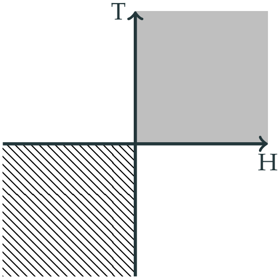
Let Ω be a finite possibility space.

A gamble $X : \Omega \to \mathbb{R}$ is an uncertain reward. We will treat them as elements $X = \langle x_1, \ldots, x_n \rangle$ of $\mathbb{R}^n$.

A set $\mathcal{D} \subseteq \mathbb{R}^n$ is a **coherent set of almost desirable gambles** if and only if it satisfies the following five axioms:

AD1. If $X < 0$ then $X \notin \mathcal{D}$ (where $X < 0 \Leftrightarrow x_i < 0$ for all $i \leq n$)

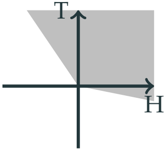
AD2. If $X \geq 0$ then $X \in \mathcal{D}$ (where $X \geq 0 \Leftrightarrow x_i \geq 0$ for all $i \leq n$)



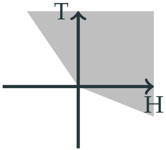
AD3. If $X \in \mathcal{D}$ and $\lambda > 0$ then $\lambda X \in \mathcal{D}$

AD4. If $X, Y \in \mathcal{D}$ then $X + Y \in \mathcal{D}$

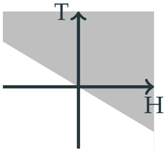
AD5. If $X + \epsilon \in \mathcal{D}$ for all $\epsilon > 0$ then $X \in \mathcal{D}$



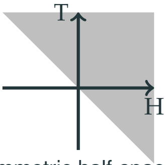
convex cone including $\mathbb{R}^2_{\geq 0}$
excluding $\mathbb{R}^2_{<0}$



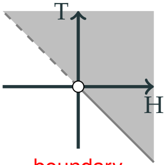
larger set:
more committal model



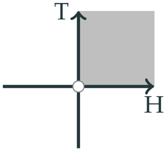
open half-space:
precise prob model



symmetric half-space:
uniform probability



boundary
infinitesimally biased



smallest coherent D_{v}
the vacuous model

Which one is *not* a coherent set of almost desirable gambles?

Correspondence Theorem [Walley, 1991, 3.8.1]:

Suppose $\mathcal{L}$ is a linear space containing constant gambles. There are 1-to-1 correspondences between the sets of models of the following types:

1. coherent lower previsions on domain $\mathcal{L}$
2. classes of almost-desirable gambles $\mathcal{D}$ that are coherent relative to $\mathcal{L}$
3. almost-preference orderings $\geq$ that are coherent relative to $\mathcal{L} \times \mathcal{L}$
4. classes of strictly desirable gambles $\mathcal{D}^+$ that are coherent relative to $\mathcal{L}$
5. strict preference orderings $>$ that are coherent relative to $\mathcal{L} \times \mathcal{L}$

Does it even make sense to talk about the accuracy of sets of almost desirable gambles?

Desirability: Behavioural Interpretation

- A gamble $g : \Omega \to \mathbb{R}$ is a random variable **that pays out in some currency/commodity C such that Your utility is linear in C.**
  - $u(£x + y) = u(£x) + u(£y)$
  - $u(£\lambda x) = \lambda u(£x)$
  - $u(x + y \text{ tickets}) = u(x \text{ tickets}) + u(y \text{ tickets})$
  - $u(\lambda x \text{ tickets}) = \lambda u(x \text{ tickets})$
- **Strict behaviourism:** Believing that g is almost-desirable ($g \in \mathcal{D}$) is nothing more than preferring $g + \epsilon$ over the status quo for all $\epsilon > 0$.

How could preferences be correct or incorrect?

> *Tis not contrary to reason to prefer the destruction of the whole world to the scratching of my finger. Tis not contrary to reason for me to chuse my total ruin, to prevent the least uneasiness of... a person wholly unknown to me. (Hume, A Treatise of Human Nature, 2.3.3.6)*

So long as Your set of almost desirable gambles is coherent, we can't sensibly talk of it being more or less "correct"

Desirability: Doxastic Interpretation

- A gamble $g : \Omega \to \mathbb{R}$ is random variable **whose outcomes are measured in a ratio or interval scale** (e.g. GBP, utility, temperature in Celsius).

	ω_1	ω_2	ω_3
g	$10°C$	$15°C$	$-20°C$
h	$-10°C$	$-20°C$	$30°C$

- **Primitivism:** Believing that g is almost-desirable ($g \in \mathcal{D}$) is a peculiarly doxastic judgment.
- You expect $g + \epsilon$ to take a positive value (for all $\epsilon > 0$), though there's no particular value you expect it to take.

Have whatever utilities you like...

IDEAL PREFERENCES **IDEAL BELIEFS**

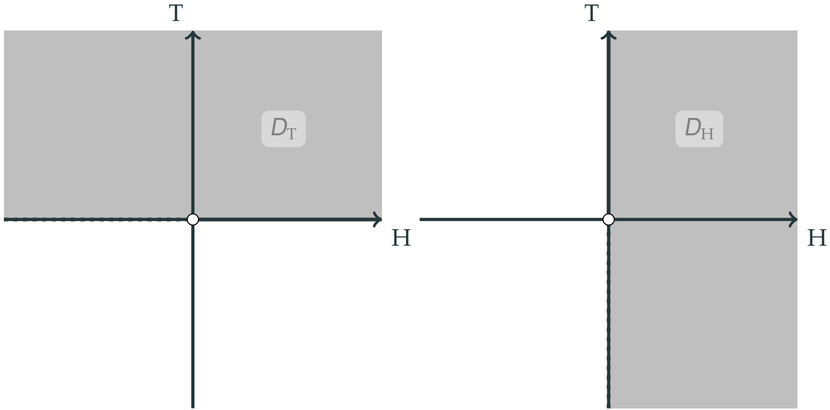
Perfection Postulate

The ideal set of almost desirable gambles if ω_i is true is given by

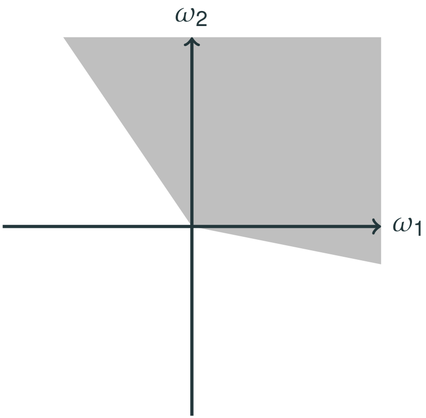
$$\mathcal{D}_i = \{X \mid x_i \geq 0\}$$

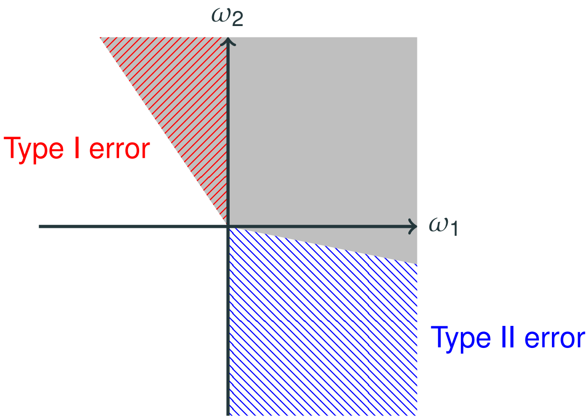
$\mathcal{D}_i$ contains all and only the gambles that are *in fact* almost desirable at ω_i (or judged to be almost desirable by God if you like).

- **Behaviourism**: $\mathcal{D}_i$ specifies preferences of a fully-informed agent at ω_i given Your utilities
- **Primitivism**: $\mathcal{D}_i$ specifies beliefs of a fully-informed agent at ω_i given a particular measurement scale


T
D_T
H
T
D_H
H

Two Ways to Fall Short of God: Type 1 and 2 Error



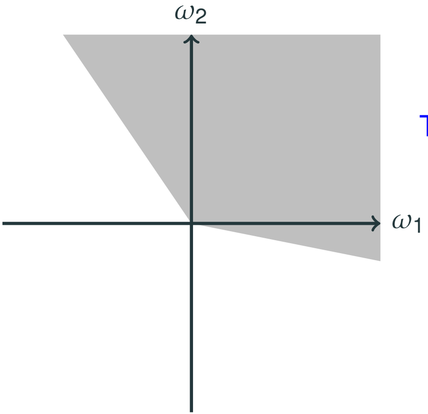


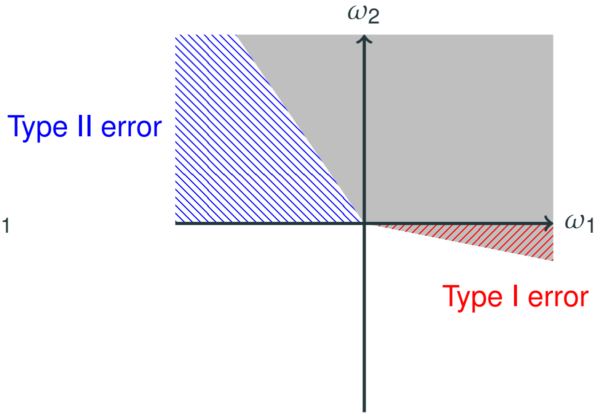
Type I and type II error for world ω_1

Two Ways to Fall Short of God: Type 1 and 2 Error

	Null Hypothesis is TRUE	Null Hypothesis is FALSE
Reject null hypothesis	Type I Error (False positive)	Correct Outcome! (True positive)
Fail to reject null hypothesis	Correct Outcome! (True negative)	Type II Error (False negative)

Two Ways to Fall Short of God: Type 1 and 2 Error





Type I and type II error for world ω_2

Why might we want to provide a real-valued measure of the accuracy of sets of almost desirable gambles?

Let $\mathbb{D}$ be the set of all $\mathcal{D} \subseteq \mathbb{R}^n$

IP scoring rules are real-valued functions

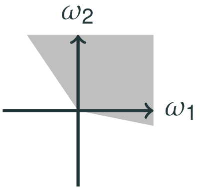
$$\mathcal{I} : \mathbb{D} \times \Omega \to \mathbb{R}_{\geq 0}$$

that measure the “accuracy” of sets of almost desirable gambles $\mathcal{D}$.

“Epistemic” or “alethic” (truth-related) loss functions.

- $\mathcal{I}(\mathcal{D}, \omega_i)$ is a function of $\mathcal{D}$’s type 1 and type 2 error at ω_i
- Type 1/2 are alethic errors: capture the two ways that $\mathcal{D}$ diverges from the “true” or “ideal” set of desirable gambles $\mathcal{D}_i$ at ω_i
- Lower penalties for strictly less type 1 and 2 error
  - Moving $\mathcal{D}$ uniformly closer to $\mathcal{D}_i$ improves (lowers) the score

Motivation



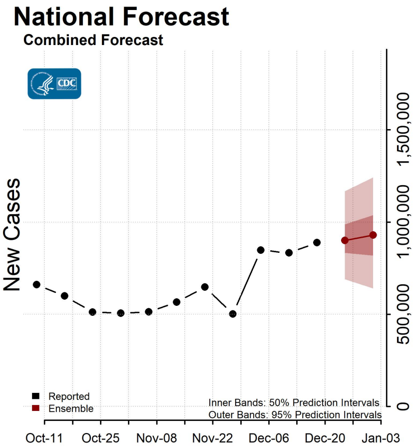
$\Rightarrow \quad \langle \mathcal{I}(\mathcal{D}, \omega_1), \mathcal{I}(\mathcal{D}, \omega_2)\rangle$

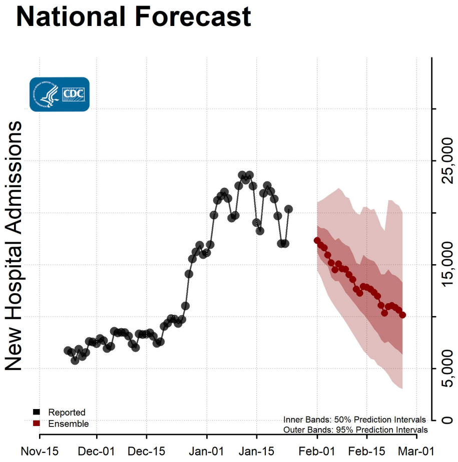
- Treat SDGs as gambles themselves
- Reason about SDGs with minimal assumptions ([de Finetti, 1974, 3.3-3.4], M.J. Schervish [2009], Joyce [2009])

$$\begin{aligned}
\mathcal{D} \notin C(\mathbb{D}) &\Leftrightarrow (\exists \mathcal{D}' \in \mathbb{D})\ \mathcal{D} <_{vacuous} \mathcal{D}' \\
&\Leftrightarrow (\exists \mathcal{D}' \in \mathbb{D})\ \langle \mathcal{I}(\mathcal{D}, \omega_1), \mathcal{I}(\mathcal{D}, \omega_2)\rangle >_{vacuous} \langle \mathcal{I}(\mathcal{D}', \omega_1), \mathcal{I}(\mathcal{D}', \omega_2)\rangle \\
&\Leftrightarrow (\exists \mathcal{D}' \in \mathbb{D})\ \mathcal{I}(\mathcal{D}, \omega_1) > \mathcal{I}(\mathcal{D}', \omega_1)\ \&\ \mathcal{I}(\mathcal{D}, \omega_2) > \mathcal{I}(\mathcal{D}', \omega_2)
\end{aligned}$$

- **Philosophy:** Epistemic justification for coherence, updating by conditioning, etc.

Forecasting: CDC COVID-19 Case & Hospitalization



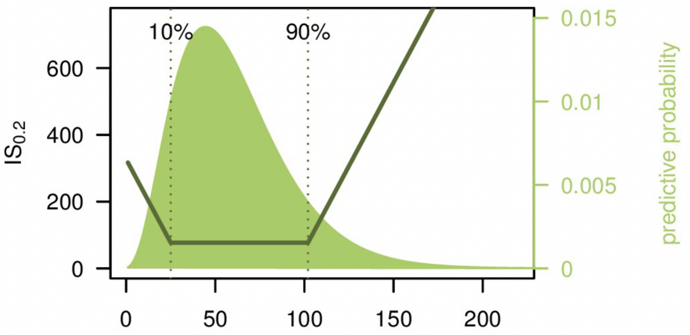


Ensemble Construction

- **April 13 - July 21, 2020:** average prediction interval; all eligible models in COVID-19 Forecast Hub
- **November 15, 2021:** Weighted ensemble forecasts of incident cases/hospitalizations/deaths and cumulative deaths

- 10 component models with best performance as measured by their Weighted Interval Score (WIS) in the 12 weeks prior to the forecast date
- Component models are assigned weights that are a function of their relative WIS during those 12 weeks
- Models with a stronger record of accuracy receiving higher weight.

Interval Score



$$IS_{0.2}(l,u,x) = (u-l) + \frac{2}{0.2}(l-x)\mathbb{1}(l>x) + \frac{2}{0.2}(x-u)\mathbb{1}(x>u)$$

Motivation

- **Aggregation**
  - Incentivise experts to report IP forecasts, evaluate aggregation procedures
- **IP Prediction Markets**
  - Traders can change market IP forecasts; receive old IP scoring rule penalty; pay IP new scoring rule penalty
- **Medicine**
  - Evaluating and improving IP expert systems
- **Artificial Intelligence**
  - Training neural net classifiers

Cautionary Note

Non-starter:

- Propose a few “intuitive” IP scoring rules with a handful of nice properties
- Compare: Absolute value score

Proper Methodology:

- Do the hard work of identifying IP scoring rules that hang together with coherence axioms, conditioning, etc.
- Otherwise we’ll incentivise...
  - ...forecasters to report incoherent IP forecasts
  - ...neural net classifiers to learn incoherent IP models

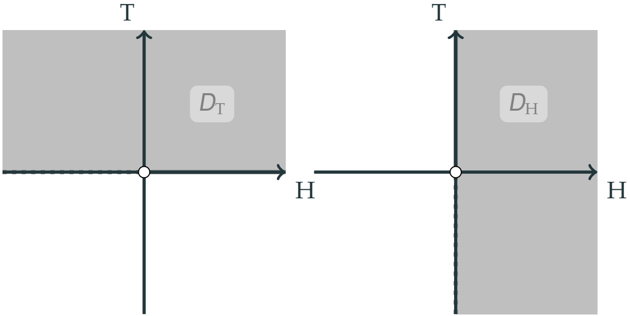
IP Scoring Rules

Perfection Postulate

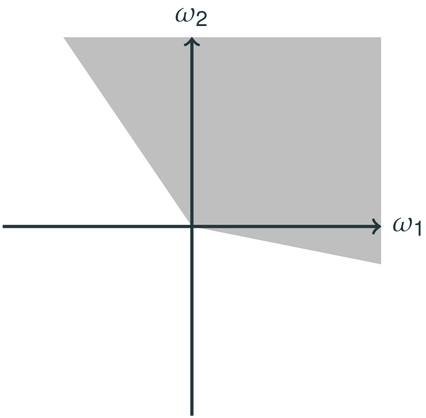
The ideal set of almost desirable gambles if ω_i is true is given by

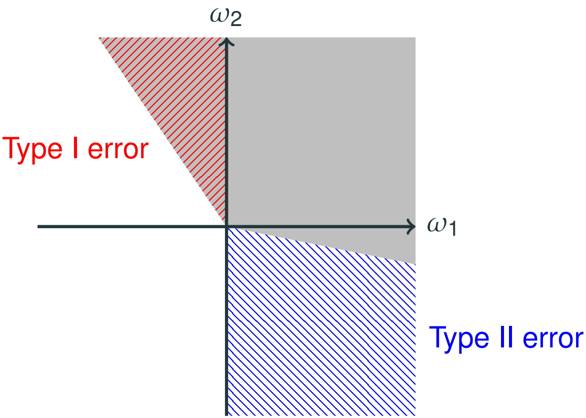
$$\mathcal{D}_i = \{X \mid x_i \geq 0\} \subseteq \mathbb{R}^n$$

$\mathcal{D}_i$ contains all and only the gambles that are *in fact* almost desirable at ω_i.



Type 1 and 2 Error





Type I and type II error at world ω_1

Type 1 Error

The type 1 error of $\mathcal{D}$ at ω_i is given by:

$$\mathcal{F}_i(\mathcal{D}) = \int_{\mathcal{D}\setminus\mathcal{D}_i} -\phi_i(x_1,\ldots,x_n)\,\mathrm{d}\mu$$

- $\phi_i : \mathbb{R}^n \to \mathbb{R}$ is a type 1/2 penalty function
- $\phi_i(X) \geq 0 \Leftrightarrow x_i \geq 0$
- For $x_i < 0$: $-\phi_i(X)$ is a **type 1** penalty for **falsely** judging X almost desirable when it is not ($x_i < 0 \Rightarrow X \notin \mathcal{D}_i$).

- $\mathrm{BIGMONEY} \in \mathcal{D}$ but in fact $\mathrm{BIGMONEY} = -2$ (so $\notin \mathcal{D}_i$)
- Type 1 error = $-\phi_i(\mathrm{BIGMONEY})$
- $\mathcal{F}_i$ averages type 1 errors.

Type 1 Error

The type 1 error of $\mathcal{D}$ at ω_i is given by:

$$\mathcal{F}_i(\mathcal{D}) = \int_{\mathcal{D} \setminus \mathcal{D}_i} -\phi_i(x_1, \ldots, x_n)\, \mathrm{d}\mu$$

- μ is a "nice" measure
  - **Domain**: Borel σ-algebra $\mathfrak{B}(\mathbb{R}^n)$ (smallest σ-algebra containing all open hypercubes)
  - **Finite**: $\mu(\mathbb{R}^n) < \infty$
  - **Absolute continuity**: μ assigns measure zero to every set with product Lebesgue measure zero

Type 2 Error

The type 2 error of $\mathcal{D}$ at ω_i is given by:

$$\mathcal{S}_i(\mathcal{D}) = \int_{\mathcal{D}_i \setminus \mathcal{D}} \phi_i(x_1, \ldots, x_n) d\mu$$

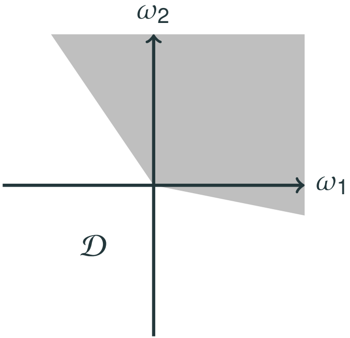
- For $x_i \geq 0$: $\phi_i(X)$ is a **type 2** penalty for failing to judge X almost desirable (staying **silent**) when it **is** ($x_i \geq 0 \Rightarrow X \in \mathcal{D}_i$).

- $\mathrm{BIGMONEY} \notin \mathcal{D}$ but in fact $\mathrm{BIGMONEY} = 50000$ (so $\in \mathcal{D}_i$)
- Type 2 error = $\phi_i(\mathrm{BIGMONEY})$
- $\mathcal{S}_i$ averages type 2 errors.

The inaccuracy (epistemic loss, epistemic disutility) of $\mathcal{D}$ at ω_i is given by:

$$\mathcal{I}_i(\mathcal{D}) = \mathcal{F}_i(\mathcal{D}) + \mathcal{S}_i(\mathcal{D}) = \int_{(\mathcal{D} \setminus \mathcal{D}_i) \cup (\mathcal{D}_i \setminus \mathcal{D})} |\phi_i(x_1, \ldots, x_n)| \, d\mu$$

Exercise 1



Suppose

- $\mathcal{I}_1(\mathcal{D}) = \mathcal{F}_1(\mathcal{D}) = \int_{\mathcal{D}\setminus\mathcal{D}_1} -\phi_1(x, y)\, \mathrm{d}\mu$
- $\mathcal{I}_2(\mathcal{D}) = \mathcal{F}_2(\mathcal{D}) = \int_{\mathcal{D}\setminus\mathcal{D}_2} -\phi_2(x, y)\, \mathrm{d}\mu$

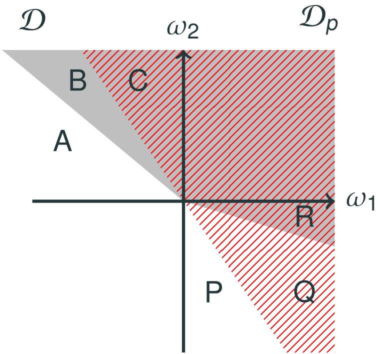
where μ assigns positive measure to every non-degenerate region. Assume $\phi_{1/2}(x, y)$ is continuous and

- $\phi_1(x, y) \geq 0 \Leftrightarrow x \geq 0$
- $\phi_2(x, y) \geq 0 \Leftrightarrow y \geq 0$

Show that $\mathcal{D}_{vac} = \mathbb{R}^2_{\geq 0}$ strictly dominates $\mathcal{D}$, *i.e.*,

- $\mathcal{I}_1(\mathcal{D}_{vac}) < \mathcal{I}_1(\mathcal{D})$
- $\mathcal{I}_2(\mathcal{D}_{vac}) < \mathcal{I}_2(\mathcal{D})$

Exercise 2



Suppose

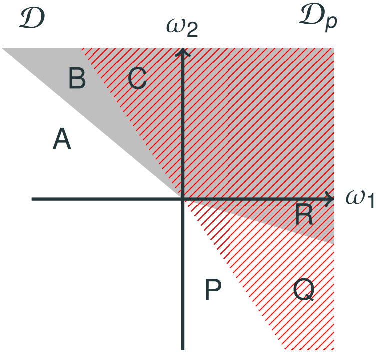
- $\mathcal{F}_1(\mathcal{D}) = \int_{\mathcal{D}\setminus\mathcal{D}_1} -x \,\mathrm{d}\mu$, $\mathcal{S}_1(\mathcal{D}) = \int_{\mathcal{D}_1\setminus\mathcal{D}} x \,\mathrm{d}\mu$
- $\mathcal{F}_2(\mathcal{D}) = \int_{\mathcal{D}\setminus\mathcal{D}_2} -y \,\mathrm{d}\mu$, $\mathcal{S}_2(\mathcal{D}) = \int_{\mathcal{D}_2\setminus\mathcal{D}} y \,\mathrm{d}\mu$

Let $\mathcal{D}_p = \{\langle x, y\rangle \mid px + (1-p)y \geq 0\}$

Fill in the following table

$\mathcal{F}_1(\mathcal{D}) = \int_{B\cup C} -x \,\mathrm{d}\mu$	$\mathcal{F}_1(\mathcal{D}_p) = \int_{C} -x \,\mathrm{d}\mu$
$\mathcal{S}_1(\mathcal{D}) = \int_{P\cup Q} x \,\mathrm{d}\mu$	$\mathcal{S}_1(\mathcal{D}_p) = \int_{P} x \,\mathrm{d}\mu$
$\mathcal{F}_2(\mathcal{D}) = \int_{R} -y \,\mathrm{d}\mu$	$\mathcal{F}_2(\mathcal{D}_p) = \int_{Q\cup R} -y \,\mathrm{d}\mu$
$\mathcal{S}_2(\mathcal{D}) = \int_{A} y \,\mathrm{d}\mu$	$\mathcal{S}_2(\mathcal{D}_p) = \int_{A\cup B} y \,\mathrm{d}\mu$

Exercise 2



Use your table to calculate

- $\mathcal{F}_1(\mathcal{D}) - \mathcal{F}_1(\mathcal{D}_p) = \int_B -x \,\mathrm{d}\mu$
- $\mathcal{S}_1(\mathcal{D}) - \mathcal{S}_1(\mathcal{D}_p) = \int_Q x \,\mathrm{d}\mu$
- $\mathcal{F}_2(\mathcal{D}) - \mathcal{F}_2(\mathcal{D}_p) = \int_Q y \,\mathrm{d}\mu$
- $\mathcal{S}_2(\mathcal{D}) - \mathcal{S}_2(\mathcal{D}_p) = \int_B -y \,\mathrm{d}\mu$

Finally show that

$$\begin{aligned} p[\mathcal{I}_1(\mathcal{D}) - \mathcal{I}_1(\mathcal{D}_p)] + (1-p)[\mathcal{I}_2(\mathcal{D}) - \mathcal{I}_2(\mathcal{D}_p)] &= \\ p[\mathcal{F}_1(\mathcal{D}) - \mathcal{F}_1(\mathcal{D}_p) + \mathcal{S}_1(\mathcal{D}) - \mathcal{S}_1(\mathcal{D}_p)] &+ \\ +(1-p)[\mathcal{F}_2(\mathcal{D}) - \mathcal{F}_2(\mathcal{D}_p) + \mathcal{S}_2(\mathcal{D}) - \mathcal{S}_2(\mathcal{D}_p)] &\geq 0 \end{aligned}$$

Type 1/2 Penalty Functions

A type 1/2 penalty function $\phi_i \colon \mathbb{R}^n \to \mathbb{R}_{\geq 0}$ is a continuous function that satisfies:

P1. $\phi_i(x_1, \ldots, x_n)$ is strictly increasing in x_i and (at least) weakly increasing in x_j for all $j \leq n$ (**Monotonicity**)

P2. $\phi_i(x_1, \ldots, x_{i-1}, 0, x_{i+1}, \ldots, x_n) = 0$ (**Zeroing marginals**)

P3. If $\lambda > 0$ then $\phi_i(\lambda X) = \lambda \phi_i(X)$ (**Positive homogeneity**)

P4. $\phi_i(X + Y) \geq \phi_i(X) + \phi_i(Y)$ (**Super-additivity**)

Status of the axioms

P1. $\phi_i(x_1, \ldots, x_n)$ is strictly increasing in x_i and (at least) weakly increasing in x_j for all $j \leq n$ (**Monotonicity**)

If g dominates h...

- ...and g and h are both almost desirable ($\in \mathcal{D}_i$), then worse to be **silent** about g (more **type 2** error)

	ω_1	ω_2
g	-1	100
h	-1	0.01

- ...and neither g nor h are almost desirable ($\notin \mathcal{D}_i$), then less bad to judge g almost desirable (less **type 1** error)

	ω_1	ω_2
g	1	-0.01
h	1	-100

Status of the axioms

P2. $\phi_i(x_1, \ldots, x_{i-1}, 0, x_{i+1}, \ldots, x_n) = 0$ (**Zeroing marginals**)

- Continuity assumption

P3. If $\lambda > 0$ then $\phi_i(\lambda x) = \lambda \phi_i(x)$ (**Positive homogeneity**)

- Guarantees that admissibility (non-dominance) is not affected by positively rescaling our underlying utility (or measurement system more generally)
  - Let $\mathcal{D}$ be Your almost-desirable weight-gambles in kg
  - Bad if $\mathcal{D}$ is admissible, but when we change units to lbs

$$\mathcal{I}_i(\mathcal{D}) = \int_{(\mathcal{D} \setminus \mathcal{D}_i) \cup (\mathcal{D}_i \setminus \mathcal{D})} |\phi_i(2.2x_1, \ldots, 2.2x_n)| \, d\mu$$

$\mathcal{D}$ is dominated

Status of the axioms

P4. $\phi_i(X+Y) \geq \phi_i(X) + \phi_i(Y)$ (**Super-additivity**)

Conservativity assumption

$$\phi_i(X) + \phi_i(-X) \leq \phi_i(0) = 0$$

Suppose $x_i > 0$.

- $\phi_i(X)$ is the type 2 error You would incur for failing to judge it desirable to accept X (as God would)
- $-\phi_i(-X)$ is the type 1 error You incur would for judging it desirable to sell X (which God would not)

Staying silent about truly desirable gambles is less bad (or no more bad) than getting it wrong about truly *un*desirable gambles

IP Scoring Rules

IP scoring rules are loss functionals of the form

$$\mathcal{I}_i(\mathcal{D}) = \mathcal{F}_i(\mathcal{D}) + \mathcal{S}_i(\mathcal{D}) = \int_{(\mathcal{D}\setminus\mathcal{D}_i)\cup(\mathcal{D}_i\setminus\mathcal{D})} |\phi_i(x_1, \ldots, x_n)| \, d\mu$$

where the $\phi_i \colon \mathbb{R}^n \to \mathbb{R}_{\geq 0}$ are continuous type 1/2 penalty functions that satisfy P1-P4.

Special Case

Exercise 3

2-parameter Penalty Function

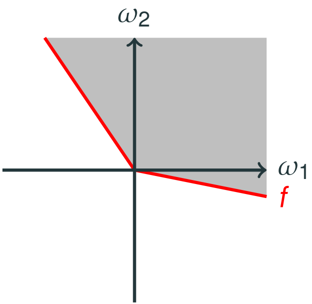
For some $\lambda \geq \gamma > 0$ and all $i \leq n$

$$\phi_i(x_1, \ldots, x_n) = \begin{cases} \lambda x_i & \text{if } x_i < 0 \\ \gamma x_i & \text{if } x_i \geq 0 \end{cases}$$

Exercise: Prove that ϕ_i satisfies P1-P4.

Coherence

Closed Convex Cones



Coherent sets of almost desirable gambles = closed convex cones

$$\mathcal{D} = \{\langle x, y \rangle \mid y \geq f(x)\}$$

In the case above, $\mathcal{D}$ is called the **epigraph** of $f : \mathbb{R} \to \mathbb{R}$.

More generally, the **epigraph** of $f : \mathbb{R}^{n-1} \to [-\infty, \infty]$ is

$$\mathcal{D} = \{\langle x_1, \ldots, x_n \rangle \mid x_n \geq f(x_1, \ldots, x_{n-1})\} \subseteq \mathbb{R}^n$$

Proposition 1

For every coherent set of almost-desirable gambles $\mathcal{D} \subseteq \mathbb{R}^n$

$$\mathcal{D} = \mathcal{D}_f$$

for some $f : \mathbb{R}^{n-1} \to [-\infty, \infty]$

We will focus on $\mathcal{D}_f$ with $f : \mathbb{R}^{n-1} \to \mathbb{R}$

Proposition 2

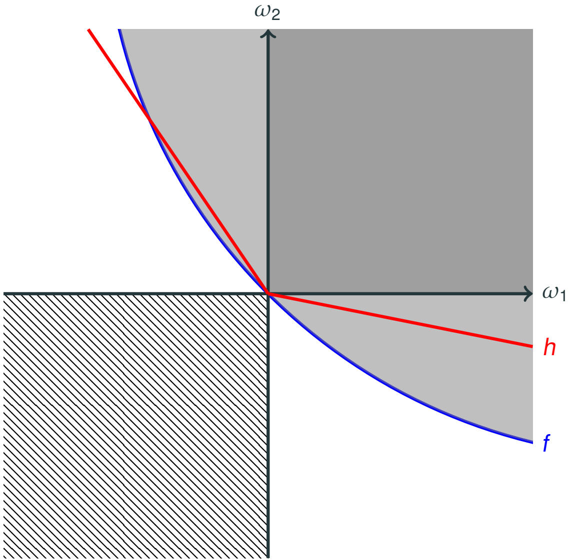
For any $f : \mathbb{R}^{n-1} \to [-\infty, \infty]$, $\mathcal{D}_f \subseteq \mathbb{R}^n$ is a coherent set of almost desirable gambles if and only if

E1. If $X \geq 0$ then $f(X) \leq 0$ (**Include Positive Orthant**)

E2. If $X \leq 0$ then $f(X) \geq 0$ (**Exclude Interior of Negative Orthant**)

E3. If $\lambda > 0$ then $f(\lambda X) = \lambda f(X)$ (**Positive homogeneity**)

E4. $f(X + Y) \leq f(X) + f(Y)$ (**Sub-additivity**)


ω_2
ω_1
h
f

An Instructive Aside: Lindley

[Lindley, 1982, Lemma 2]

If $\mathcal{I}$ is a continuously differentiable strictly proper scoring rule, then following three conditions are equivalent:

1. There are $a, b \in \mathbb{R}$ s.t.

$$\nabla_{\langle a,b\rangle}\mathcal{I}_0(x,y) = \lim_{\epsilon\to 0}\frac{1}{\epsilon}\left[\mathcal{I}_1(x+\epsilon\, a, y+\epsilon\, b) - \mathcal{I}_1(x,y)\right] < 0$$

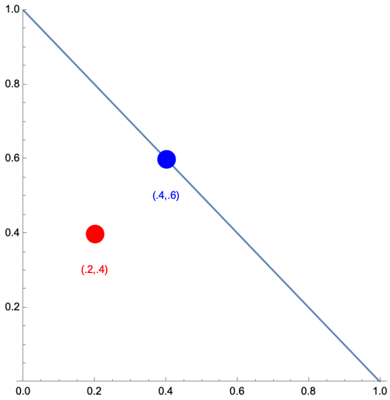
$$\nabla_{\langle a,b\rangle}\mathcal{I}_1(x,y) = \lim_{\epsilon\to 0}\frac{1}{\epsilon}\left[\mathcal{I}_0(x+\epsilon\, a, y+\epsilon\, b) - \mathcal{I}_0(x,y)\right] < 0$$

2.

$$0 \notin \operatorname{posi}\left(\{\nabla\mathcal{I}_0(x,y), \nabla\mathcal{I}_1(x,y)\}\right)$$

3. $y \neq 1 - x$

Local Dominance



Nudge $\langle 0.2, 0.4 \rangle$ toward $\langle 0.4, 0.6 \rangle$ by adding $\epsilon \langle 0.2, 0.2 \rangle$ for small $\epsilon > 0$.

Result: $\langle 0.2 + \epsilon 0.2, 0.4 + \epsilon 0.2 \rangle$

Local Dominance

$$
\begin{aligned}
\nabla_{\langle a,b\rangle}\mathcal{I}_i(x,y) &= \lim_{\epsilon\to 0}\frac{1}{\epsilon}\left[\mathcal{I}_i(x+\epsilon\, a, y+\epsilon\, b) - \mathcal{I}_i(x,y)\right] \\
&= \langle a,b\rangle\cdot\nabla\mathcal{I}_i(x,y) \\
&= a\frac{\partial\mathcal{I}_i}{\partial x}(x,y) + b\frac{\partial\mathcal{I}_i}{\partial y}(x,y)
\end{aligned}
$$

Exercise 4

$$f_1(x) = \int_x^1 (1-t)\, m(t)\, \mathrm{d}t$$
$$f_0(x) = \int_0^x t\, m(t)\, \mathrm{d}t$$
$$\mathcal{I}_1(x,y) = f_1(x) + f_0(y)$$
$$\mathcal{I}_0(x,y) = f_0(x) + f_1(y)$$

$$\frac{\partial \mathcal{I}_1}{\partial x}(x,y) = (x-1)m(x)$$
$$\frac{\partial \mathcal{I}_1}{\partial y}(x,y) = y\, m(y)$$
$$\frac{\partial \mathcal{I}_0}{\partial x}(x,y) = x\, m(x)$$
$$\frac{\partial \mathcal{I}_0}{\partial y}(x,y) = (y-1)m(y)$$

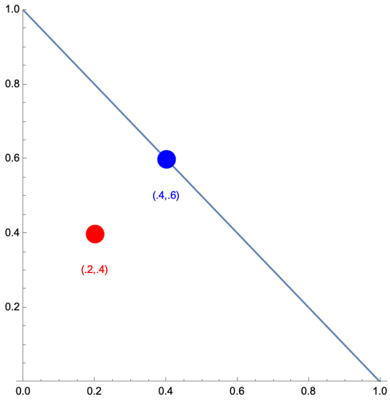
Brier score: $m(x) = 1$

Problem: Calculate $\nabla_{\langle 0.2, 0.2\rangle}\mathcal{I}_0(0.2, 0.4)$ and $\nabla_{\langle 0.2, 0.2\rangle}\mathcal{I}_1(0.2, 0.4)$

Solution:

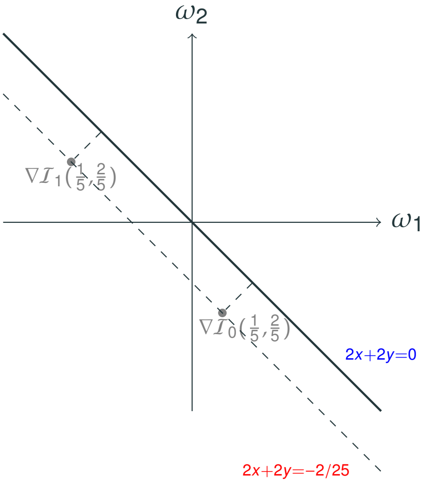
- $(0.2)(0.2) + (0.2)(0.4-1) = -0.08$
- $(0.2)(0.2-1) + (0.2)(0.4) = -0.08$

Local Dominance



Nudging $\langle 0.2, 0.4 \rangle$ toward $\langle 0.4, 0.6 \rangle$ by adding $\epsilon \langle 0.2, 0.2 \rangle$ for sufficiently small $\epsilon > 0$ is **guaranteed** to improve accuracy (*i.e.* in both ω_1 and ω_2)

Local Dominance



- There are $a, b \in \mathbb{R}$ s.t.

$$\nabla_{\langle a,b\rangle}\mathcal{I}_0(x,y) = \langle a,b\rangle \cdot \nabla\mathcal{I}_0(x,y) < 0$$

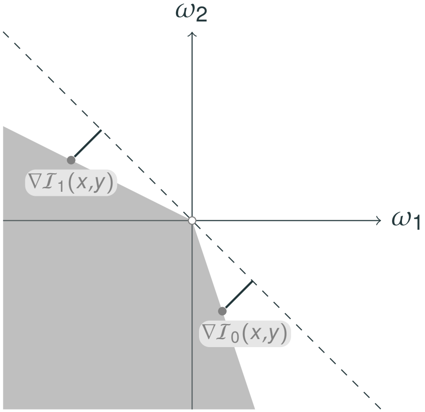
$$\nabla_{\langle a,b\rangle}\mathcal{I}_1(x,y) = \langle a,b\rangle \cdot \nabla\mathcal{I}_1(x,y) < 0$$

- $a = b = 2$
- For sufficiently small $\epsilon > 0$

$$\mathcal{I}_0(x+\epsilon 2, y+\epsilon 2) - \mathcal{I}_0(x,y) < 0$$

$$\mathcal{I}_1(x+\epsilon 2, y+\epsilon 2) - \mathcal{I}_1(x,y) < 0$$

Local Dominance: Precise



There are $a, b \in \mathbb{R}$ s.t.

$$\nabla_{\langle a,b\rangle}\mathcal{I}_0(x,y) < 0$$

$$\nabla_{\langle a,b\rangle}\mathcal{I}_1(x,y) < 0$$

iff

$$0 \notin \text{posi}\left(\{\nabla\mathcal{I}_0(x,y), \nabla\mathcal{I}_1(x,y)\}\right)$$

If $\mathcal{I}$ is continuously differentiable and strictly proper then these conditions hold iff

$$y \neq 1 - x$$

Local Dominance and Coherence

Local Dominance, Precise: There are $a, b \in \mathbb{R}$ s.t. for all $i \leq n$

$$\nabla_{\langle a,b\rangle}\mathcal{I}_i(x,y) = \langle a,b\rangle \cdot \nabla\mathcal{I}_i(x,y) = \lim_{\epsilon\to 0}\frac{1}{\epsilon}\left[\mathcal{I}_1(x+\epsilon\, a, y+\epsilon\, b) - \mathcal{I}_1(x,y)\right] < 0$$

Let $\mathcal{I}_i(f) = \mathcal{I}_i(\mathcal{D}_f)$.

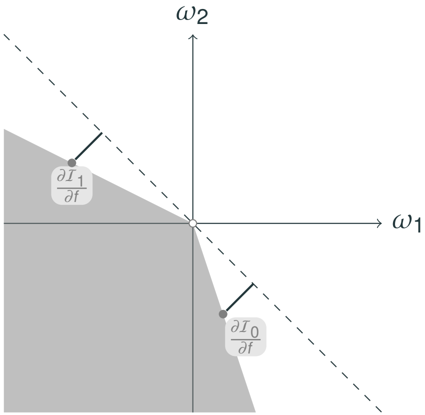
Local Dominance, Imprecise:[1] There is some $h : \mathbb{R}^{n-1} \to \mathbb{R}$ s.t. for all $i \leq n$

$$\delta\mathcal{I}_i(f,h) = \int_{\mathbb{R}^{n-1}} \frac{\partial \mathcal{I}_i}{\partial f} h \,\mathrm{d}\nu = \lim_{\epsilon\to 0}\frac{1}{\epsilon}\left[\mathcal{I}_i(f+\epsilon\, h) - \mathcal{I}_i(f)\right] < 0$$

First variation—calculus of variations analogue of directional derivative

[1] For $X \subseteq \mathbb{R}^{n-1}$, $\nu(X) = \int_{\mathbb{R}^{n-1}} m(x_1,\ldots,x_{n-1}, f(x_1,\ldots,x_{n-1}))\,\mathrm{d}x_1\ldots\mathrm{d}x_{n-1}$ where m is the Radon–Nikodym derivative of μ wrt the product Lebesgue measure.

Local Dominance: Imprecise



Under certain conditions...

There is some $h : \mathbb{R} \to \mathbb{R}$ s.t.

$$\delta \mathcal{I}_0(f, h) = \int_{\mathbb{R}} \frac{\partial \mathcal{I}_0}{\partial f} h \, \mathrm{d}\nu < 0$$

$$\delta \mathcal{I}_1(f, h) = \int_{\mathbb{R}} \frac{\partial \mathcal{I}_1}{\partial f} h \, \mathrm{d}\nu < 0$$

iff

$$0 \notin \operatorname{posi}\left(\left\{\frac{\partial \mathcal{I}_0}{\partial f}, \frac{\partial \mathcal{I}_1}{\partial f}\right\}\right)$$

iff

$$0 \notin \operatorname{posi}\left(\left\{\phi_0(x, f(x)), \phi_1(x, f(x))\right\}\right)$$

Local Dominance and Coherence

Proposition 3

If $\mathcal{I}$ is an IP scoring rule, and $\phi_i(x_1, \ldots, x_{n-1}, f(x_1, \ldots, x_{n-1})) \in \mathcal{L}^p(\nu)$ for some $p > 1$ and all $i \leq n$, then the following three conditions are equivalent:

1. There is some $h : \mathbb{R}^{n-1} \to \mathbb{R}$ s.t. for all $i \leq n$

$$\delta\mathcal{I}_i(f, h) < 0$$

2. $$0 \notin \operatorname{posi}\left(\left\{\phi_i(x_1, \ldots, x_{n-1}, f(x_1, \ldots, x_{n-1})) \mid i \leq n\right\}\right)$$

3. ???

Local Dominance and Coherence

Proposition 4

If $\mathcal{I}$ is an IP scoring rule and

$$0 \in \operatorname{posi}\left(\left\{\phi_i(x_1, \ldots, x_{n-1}, f(x_1, \ldots, x_{n-1})) \mid i \leq n\right\}\right)$$

then $\mathcal{D}_f$ is a coherent set of almost desirable gambles.

Not locally dominated $\Rightarrow$ Coherent

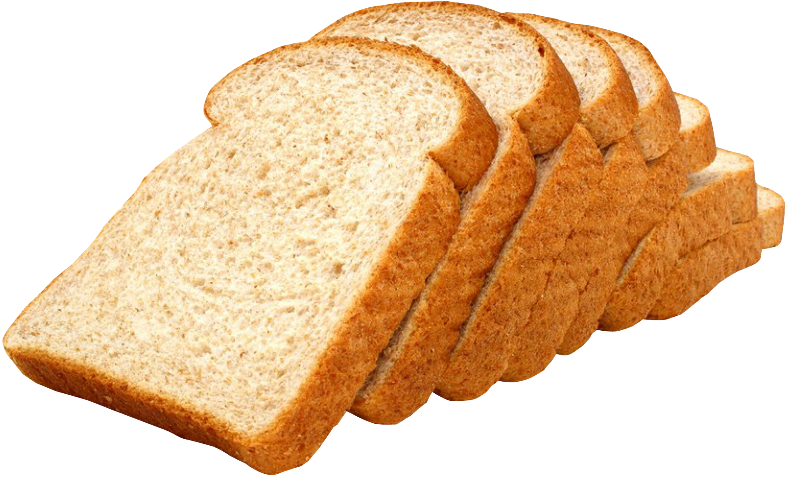
Special Case: For some $\lambda \geq \gamma > 0$, $\phi_i(x_1, x_2, x_3) = \begin{cases} \lambda x_i & \text{if } x_i < 0 \\ \gamma x_i & \text{if } x_i \geq 0 \end{cases}$

Proposition 4

The following two conditions are equivalent:

1. $0 \in \text{posi}\left(\left\{\phi_i(x_1, x_2, f(x_1, x_2)) \mid i \leq 3\right\}\right)$
2. There are $\alpha, \beta > 0$ s.t.

$$f(x_1, x_2) = \begin{cases} \frac{-\gamma(\alpha x_1 + \beta x_2)}{\lambda} & \text{if } x_1 \geq 0, x_2 \geq 0 \\ \frac{-\lambda(\alpha x_1 + \beta x_2)}{\gamma} & \text{if } x_1 < 0, x_2 < 0 \\ \frac{-(\alpha\lambda x_1 + \beta\gamma x_2)}{\gamma} & \text{if } x_1 < 0, x_2 \geq 0, \alpha\lambda x_1 + \beta\gamma x_2 < 0 \\ \frac{-(\alpha\lambda x_1 + \beta\gamma x_2)}{\lambda} & \text{if } x_1 < 0, x_2 \geq 0, \alpha\lambda x_1 + \beta\gamma x_2 \geq 0 \\ \frac{-(\alpha\gamma x_1 + \beta\lambda x_2)}{\gamma} & \text{if } x_1 \geq 0, x_2 < 0, \alpha\gamma x_1 + \beta\lambda x_2 < 0 \\ \frac{-(\alpha\gamma x_1 + \beta\lambda x_2)}{\lambda} & \text{otherwise} \end{cases}$$



See Seidenfeld et al. [2012]

Conditioning

Let $\mathcal{E} \subseteq \Omega$ with $\omega_n \in \mathcal{E}$ ($|\mathcal{E}| = m$). Let $E = \langle e_1, \ldots, e_n \rangle$ be the indicator of $\mathcal{E}$.

The set of conditional almost desirable gambles given $\mathcal{E}$ is defined by (see [Augustin et al., 2014, 1.3.3]):

$$\mathcal{D}_{\mathcal{E}} = \left\{ X\big|_{\mathcal{E}} \mid \mathbb{I}_E X = \langle e_1 x_1, \ldots, e_n x_n \rangle \in \mathcal{D} \right\} \subseteq \mathbb{R}^m$$

The conditional type 1/2 penalty functions are given by

$$\psi_i(x_1, \ldots, x_m) = \phi_i(e_1 x_1, \ldots, e_n x_n)$$

Let $h(x_1, \ldots, x_{m-1}) = f(e_1 x_1, \ldots, e_{n-1} x_{n-1})$

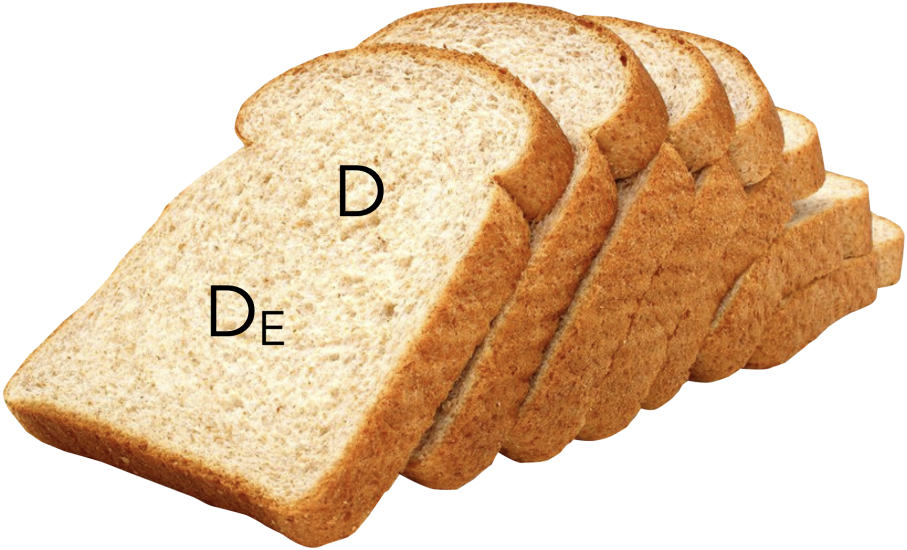
Proposition 5

If

$$0 \in \operatorname{posi}\left(\left\{ \phi_i(x_1, \ldots, x_{n-1}, f(x_1, \ldots, x_{n-1})) \mid i \leq n \right\} \right)$$

then

$$0 \in \operatorname{posi}\left(\left\{ \psi_i(x_1, \ldots, x_{n-1}, f(x_1, \ldots, x_{n-1})) \mid i \leq m \right\} \right)$$


D
DE

References

Thomas Augustin, Frank P. A. Coolen, Gert de Cooman, and Matthias C. M. Troffaes, editors. Introduction to Imprecise Probabilities. John Wiley & Sons, 2014.

Bruno de Finetti. Theory of Probability: A Critical Introductory Treatment, volume 1. John Wiley & Sons, Chichester, 1974.

James M Joyce. Accuracy and coherence: Prospects for an alethic epistemology of partial belief. In Franz Huber and Christoph Schmidt-Petri, editors, Degrees of Belief, volume 342. Springer, Dordrecht, 2009.

D. V. Lindley. Scoring rules and the inevitability of probability. International Statistical Review, 50:1–26, 1982. With discussion.

J.B. Kadane M.J. Schervish, T. Seidenfeld. Proper scoring rules dominated forecasts and coherence. Decision Analysis, 6(4):202–221, 2009.

Teddy Seidenfeld, M. J. Schervish, and J. B. Kadane. Forecasting with imprecise probabilities. International Journal of Approximate Reasoning, 53(8): 1248–1261., 2012.

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