Introduction to Lower Previsions II

Gregory Wheeler
Munich Center for Mathematical Philosophy
Ludwig Maximilian University of Munich

Progic Spring School
April 20–21, 2015

http://gregorywheeler.org/courses/progic/lp_2.pdf

Two ways of defining lower previsions w/out credal sets

Directly

- Define $\underline{P}(f)$ directly as Your supremum acceptable buying price for f
- $\underline{P}(f)$ is the highest price s s.t. for any $t < s$, You accept to pay t before observing X on the promise that You will receive $f(x)$ upon observing the event $X = x$.
- *Thus*, You are only required to consider whether you accept bounded gambles of type $f - \alpha$.
- *Conditioning is complicated*

Via Gambles

- Announce Your acceptable bounded gambles
- Enforce coherence conditions
- *Then* infer Your lower $\underline{P}(f)$ and upper $\overline{P}(f)$ previsions for any bounded gamble f.
- *Conditioning is clean(er)*

Three goals

1. *Get rid of probabilities, credal sets, and lower previsions*
   - ✓ Sets of **Acceptable Gambles**

     Williams 1977; Walley 1991, 2000; De Cooman and Troffaes 2014

2. *Propose a benchmark for evaluating a model of belief*
   - ✓ **Operationalizable**: one-sided betting
   - ✓ **Inference**: closure, marginalization, and conditionalization
   - ✓ **Unification**: unify different uncertainty models

3. *Address with some alleged problems*

Three goals

1. *Get rid of probabilities, credal sets, and lower previsions*
   - ✓ Sets of **Acceptable Gambles**
     Williams 1977; Walley 1991, 2000; De Cooman and Troffaes 2014
2. *Propose a benchmark for evaluating a model of belief*
   - ✓ **Operationalizable**: one-sided betting
   - ✓ **Inference**: closure, marginalization, and conditionalization
   - ✓ **Unification**: unify different uncertainty models
3. *Address with some alleged problems*
   - Dilation
   - Violations of Good's Principle
   - No Strictly Proper IP Scoring Rules
     - Seidenfeld, Schervish, and Kadane (2012)
       Drop quantifiability $\Rightarrow$ *lexicographic decision theory*
     - Mayo-Wilson and Wheeler (forthcoming)
       Drop strictness $\Rightarrow$ *'normal' decision theory*

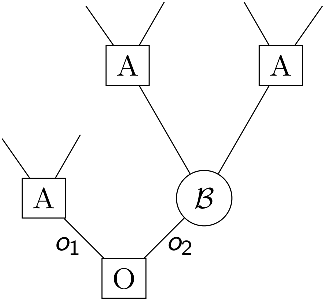
Outline

- Good's Principle
- Dilation Example
- Dilation Diagnoses
- Dilation Characterization Results (joint w/ Paul Pedersen)
- Plurality of Independence Concepts
- Rethinking Good's Principle

Good's Principle

Never turn down the offer of free information.

Good's Principle



You should delay making a terminal decision (A) between alternative courses of action if the opportunity arises to learn, at zero cost to you, the outcome of an experiment ($\mathcal{B}$) relevant to the decision

Dilation – informal

Idea: A probability estimate in E becomes less precise upon learning the outcome of an experiment, $\mathcal{B}$, **no matter how the experiment turns out.**

Defining dilation

<u>DEFN.</u> Let $(\Omega, \mathcal{A}, \mathbb{P}, \underline{P})$ be a lower probability space, let $\mathcal{B}$ be a positive measurable partition of Ω, and let E be an event.

Defining dilation

<u>Defn.</u> Let $(\Omega, \mathcal{A}, \mathbb{P}, \underline{P})$ be a lower probability space, let $\mathcal{B}$ be a positive measurable partition of Ω, and let E be an event. Say that $\mathcal{B}$ **dilates** E just in case for each $H \in \mathcal{B}$:

Defining dilation

<u>Defn.</u> Let $(\Omega, \mathcal{A}, \mathbb{P}, \underline{\mathrm{P}})$ be a lower probability space, let $\mathcal{B}$ be a positive measurable partition of Ω, and let E be an event. Say that $\mathcal{B}$ **dilates** E just in case for each $H \in \mathcal{B}$:

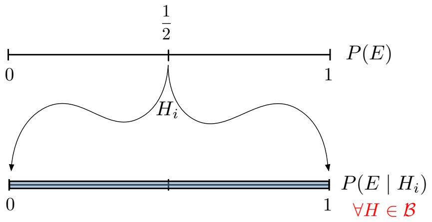
$$\underline{\mathrm{P}}(E \mid H) \ < \ \underline{\mathrm{P}}(E) \ \leq \ \overline{\mathrm{P}}(E) \ < \ \overline{\mathrm{P}}(E \mid H).$$

Defining dilation

<u>DEFN.</u> Let $(\Omega, \mathcal{A}, \mathbb{P}, \underline{\mathrm{P}})$ be a lower probability space, let $\mathcal{B}$ be a positive measurable partition of Ω, and let E be an event. Say that $\mathcal{B}$ **dilates** E just in case for each $H \in \mathcal{B}$:

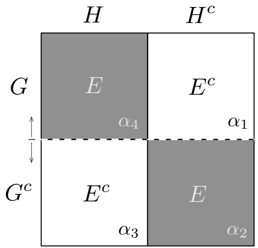
$$\underline{\mathrm{P}}(E \mid H) < \underline{\mathrm{P}}(E) \leq \overline{\mathrm{P}}(E) < \overline{\mathrm{P}}(E \mid H).$$

<u>EXAMPLE 1</u>



ip axioms

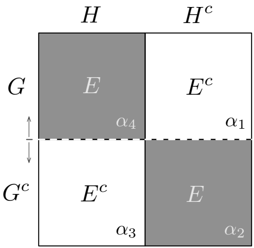
Dilation example



$\mathcal{B} = \{H, H^c\}$ outcomes of a fair coin toss,

$$\underline{P}(H) = \overline{P}(H) = \frac{1}{2} = \underline{P}(H^c) = \overline{P}(H^c) \quad (1)$$

Dilation example



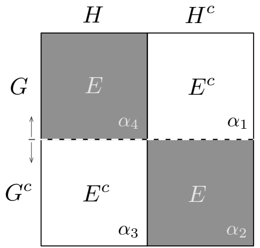
$\mathcal{B} = \{H, H^c\}$ outcomes of a fair coin toss,

$$\underline{P}(H) = \overline{P}(H) = \frac{1}{2} = \underline{P}(H^c) = \overline{P}(H^c) \quad (1)$$

G is a nearly maximal uncertain event:

$$\overline{P}(G) - \underline{P}(G) = .9 \quad (2)$$

Dilation example



$\mathcal{B} = \{H, H^c\}$ outcomes of a fair coin toss,

$$\underline{P}(H) = \overline{P}(H) = \frac{1}{2} = \underline{P}(H^c) = \overline{P}(H^c) \quad (1)$$

G is a nearly maximal uncertain event:

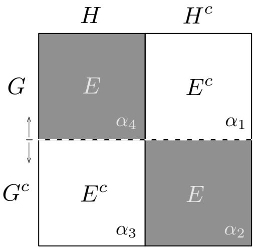
$$\overline{P}(G) - \underline{P}(G) = .9 \quad (2)$$

H and G are stochastically independent:

$$p(G \cap H) = p(G)p(H) = \frac{p(G)}{2}, \quad (3)$$

for each $p \in \mathbb{P}$.

Dilation example

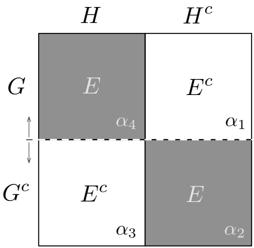


$$\underline{P}(H) = \overline{P}(H) = \underline{P}(H^c) = \overline{P}(H^c) = 1/2 \quad (1)$$

$$\overline{P}(G) - \underline{P}(G) \approx 1 \quad (2)$$

$$\forall p \in \mathbb{P},\ p(G \cap H) = p(G)p(H) = \frac{p(G)}{2} \quad (3)$$

Dilation example



$$\underline{P}(H) = \overline{P}(H) = \underline{P}(H^c) = \overline{P}(H^c) = 1/2 \tag{1}$$

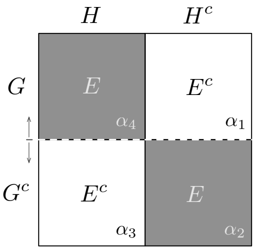
$$\overline{P}(G) - \underline{P}(G) \approx 1 \tag{2}$$

$$\forall p \in \mathbb{P},\ p(G \cap H) = p(G)p(H) = \frac{p(G)}{2} \tag{3}$$

<u>Defn.</u> $E := (G \cap H) \cup (G^c \cap H^c)$

$$\text{So, } \forall p \in \mathbb{P},\ p(E) = 1/2. \tag{4}$$

Dilation example



$$\underline{P}(H) = \overline{P}(H) = \underline{P}(H^c) = \overline{P}(H^c) = 1/2 \qquad (1)$$

$$\overline{P}(G) - \underline{P}(G) \approx 1 \qquad (2)$$

$$\forall p \in \mathbb{P},\ p(G \cap H) = p(G)p(H) = \frac{p(G)}{2} \qquad (3)$$

<u>DEFN.</u> $E := (G \cap H) \cup (G^c \cap H^c)$

$$\text{So, } \forall p \in \mathbb{P},\ p(E) = 1/2. \qquad (4)$$

Proof of (4):

$$\begin{aligned} \forall p \in \mathbb{P} : p(E) &= p(G \cap H) + p(G^c \cap H^c) && \text{(Defn. of } E\text{)} \\ &= \frac{p(G)}{2} + \frac{1 - p(G)}{2} && \text{(By Eq 3)} \\ &= \frac{p(G) + 1 - p(G)}{2} = \frac{1}{2}. \end{aligned}$$

Dilation example

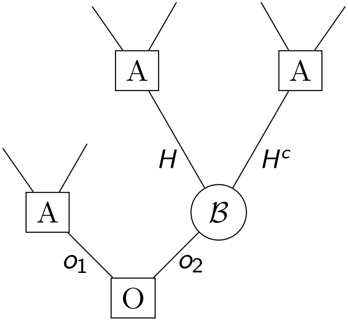
Observe: E is strictly dilated by the coin toss, $\mathcal{B} = \{H, H^c\}$.

Proof - We show that $0 \approx \underline{P}(E \mid H) < \underline{P}(E) = 1/2$.

$$
\begin{aligned}
\underline{P}(E \mid H) &= \inf \{\, p(E \mid H) \;:\; p \in \mathbb{P} \,\} && \text{(Defn.)} \\
&= \inf \left\{ \frac{p\left([(G \cap H) \cup (G^c \cap H^c)] \cap H\right)}{p(H)} : p \in \mathbb{P} \right\} \\
&= \inf \left\{ \frac{p(G \cap H)}{p(H)} : p \in \mathbb{P} \right\} \\
&= \inf \left\{ \frac{p(G)p(H)}{p(H)} : p \in \mathbb{P} \right\} && \text{(By Eq 3)} \\
&= \inf \left\{ p(G) : p \in \mathbb{P} \right\} = .1.
\end{aligned}
$$

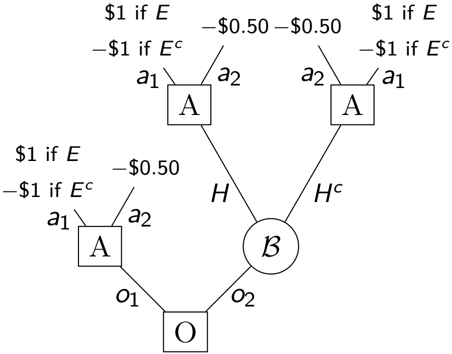
An analogous argument establishes $.9 = \overline{P}(E \mid H) > 1/2$; Finally, by the same reasoning $.1 = \underline{P}(E \mid H^c) < 1/2 < \overline{P}(E \mid H^c) = .9$. $\diamond$

Good's Principle

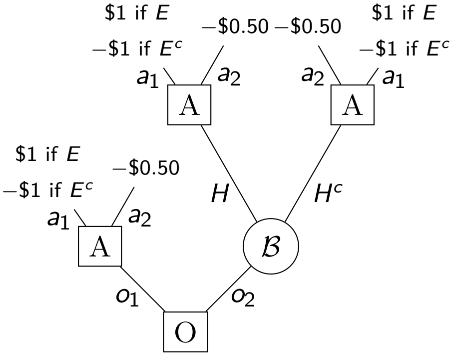


You should delay making a terminal decision (A) between alternative courses of action if the opportunity arises to learn, at zero cost to you, the outcome of an experiment ($\mathcal{B}$) relevant to the decision

Good's principle & minimum expected utility



Good's principle & minimum expected utility



$$o_1 : \; A = \begin{cases} 0 & \text{if} \quad {\color{red} a_1} \\ -\$0.50 & \text{if} \quad a_2 \end{cases} \qquad o_2 : \; A \mid \mathcal{B} = \begin{cases} -\$0.80 & \text{if} \quad a_1^{**} \\ -\$0.50 & \text{if} \quad {\color{red} a_2} \end{cases}$$

** for **some** $p \in \mathbb{P}$, i.e., $p(E \mid H^c) = .9$

Reactions to dilation

Dilation is pathological

a) Reductio of imprecise probability theory

(Elga, White, J. Williamson)

b) Spike Bayesian conditionalization

(Halpern & Grünwald, Kyburg)

Reactions to dilation

Dilation is pathological

a) Reductio of imprecise probability theory

(Elga, White, J. Williamson)

b) Spike Bayesian conditionalization

(Halpern & Grünwald, Kyburg)

Dilation is benign

c) Dilation is partially benign

(Joyce, Seidenfeld)

d) Dilation is (completely?) benign

(Pedersen & Wheeler)

Two questions to address

What is dilation?

How does it figure in failures of Good's principle?

What accounts for dilation?

Fact 1 (Seidenfeld and Wassermann 1993)
Distance from independence is a necessary condition for dilation.

What accounts for dilation?

FACT 1 (Seidenfeld and Wassermann 1993)
Distance from independence is a necessary condition for dilation.

DEFN. A simple measure of stochastic independence:

$$\mathrm{S}_p(E, F) \quad =_{df} \quad \frac{p(E \cap F)}{p(E)p(F)}.$$

NOTE. $p(E \mid F) = p(E) \;\Rightarrow\; p(F \mid E) = p(F) \;\Rightarrow\; p(E \cap F) = p(E)p(F)$

What accounts for dilation?

<u>FACTS</u> (Seidenfeld and Wassermann 1993, 1994)

- If $\mathcal{B}$ dilates E, any p realizing $\underline{P}(E \mid H)$ (any p: $\overline{P}(E \mid H)$) must make E and H negatively (positively) correlated (for every $H \in \mathcal{B}$).

  ... but vacuously true for an event if set $\underline{\mathbb{P}}(E \mid H)$ of conditional probability measures realizing lower probabilities is empty

What accounts for dilation?

<u>FACTS</u> (Seidenfeld and Wassermann 1993, 1994)

- If $\mathcal{B}$ dilates E, any p realizing $\underline{P}(E \mid H)$ (any p: $\overline{P}(E \mid H)$) must make E and H negatively (positively) correlated (for every $H \in \mathcal{B}$).

  ... but vacuously true for an event if set $\underline{\mathbb{P}}(E \mid H)$ of conditional probability measures realizing lower probabilities is empty

- *If every H in $\mathcal{B}$ is such that $\underline{P}(E)$ $(\overline{P}(E))$ is realized by some p for which E and H are negatively (pos.) correlated, then $\mathcal{B}$ dilates E.*

  ... but vacuously true for an event if the set of probability measures $\underline{\mathbb{P}}(E)$ realizing lower probabilities is empty

What accounts for dilation?

<u>Facts</u> (Seidenfeld and Wassermann 1993, 1994)

- If $\mathcal{B}$ dilates E, any p realizing $\underline{P}(E \mid H)$ (any p: $\overline{P}(E \mid H)$) must make E and H negatively (positively) correlated (for every $H \in \mathcal{B}$).

  ... but vacuously true for an event if set $\underline{\mathbb{P}}(E \mid H)$ of conditional probability measures realizing lower probabilities is empty

- *If every H in $\mathcal{B}$ is such that $\underline{P}(E)$ $(\overline{P}(E))$ is realized by some p for which E and H are negatively (pos.) correlated, then $\mathcal{B}$ dilates E.*

  ... but vacuously true for an event if the set of probability measures $\underline{\mathbb{P}}(E)$ realizing lower probabilities is empty

- *Within the convex set of supporting hyperplanes positive and negative correlation is a <u>necessary and sufficient</u> conditions for dilation.* more

What accounts for dilation?

<u>Facts</u> (Seidenfeld and Wassermann 1993, 1994)

- If $\mathcal{B}$ dilates E, any p realizing $\underline{P}(E \mid H)$ (any p: $\overline{P}(E \mid H)$) must make E and H negatively (positively) correlated (for every $H \in \mathcal{B}$).

  ... but vacuously true for an event if set $\underline{\mathbb{P}}(E \mid H)$ of conditional probability measures realizing lower probabilities is empty

- *If every H in $\mathcal{B}$ is such that $\underline{P}(E)$ ($\overline{P}(E)$) is realized by some p for which E and H are negatively (pos.) correlated, then $\mathcal{B}$ dilates E.*

  ... but vacuously true for an event if the set of probability measures $\underline{\mathbb{P}}(E)$ realizing lower probabilities is empty

- *Within the convex set of supporting hyperplanes positive and negative correlation is a <u>necessary and sufficient</u> conditions for dilation.* more

Is there a general characterization of dilation?

What accounts for dilation?

DEFN. Given a nonempty set of probabilities $\mathbb{P}$, events $E, H \in \mathcal{A}$ with $\underline{P}(H) > 0$, and $\epsilon > 0$, define:

What accounts for dilation?

<u>DEFN.</u> Given a nonempty set of probabilities $\mathbb{P}$, events $E, H \in \mathcal{A}$ with $\underline{P}(H) > 0$, and $\epsilon > 0$, define:

$$\underline{\mathbb{P}}(E|H,\epsilon) =_{df} \{p \in \mathbb{P} : |p(E|H) - \underline{P}(E|H)| < \epsilon\};$$

$$\overline{\mathbb{P}}(E|H,\epsilon) =_{df} \{p \in \mathbb{P} : |p(E|H) - \overline{P}(E|H)| < \epsilon\}.$$

Call the sets $\underline{\mathbb{P}}(E|H,\epsilon)$ and $\overline{\mathbb{P}}(E|H,\epsilon)$ <u>lower</u> and <u>upper neighborhoods</u> of E conditional on H, respectively, with <u>radius</u> ϵ.

What accounts for dilation?

<u>DEFN.</u> Given a nonempty set of probabilities $\mathbb{P}$, events E, $H \in \mathcal{A}$ with $\underline{\mathrm{P}}(H) > 0$, and $\epsilon > 0$, define:

$$\underline{\mathbb{P}}(E|H,\epsilon) \quad =_{df} \quad \{p \in \mathbb{P} : \; |p(E|H) - \underline{\mathrm{P}}(E|H)| < \epsilon\};$$

$$\overline{\mathbb{P}}(E|H,\epsilon) \quad =_{df} \quad \{p \in \mathbb{P} : \; |p(E|H) - \overline{\mathrm{P}}(E|H)| < \epsilon\}.$$

Call the sets $\underline{\mathbb{P}}(E|H,\epsilon)$ and $\overline{\mathbb{P}}(E|H,\epsilon)$ <u>lower</u> and <u>upper</u> <u>neighborhoods</u> of E conditional on H, respectively, with <u>radius</u> ϵ.

<u>IDEA:</u> A probability function is an element of a lower neighborhood of E conditional on H with radius ϵ if $p(E|H)$ is <u>within</u> ϵ of $\underline{\mathrm{P}}(E|H)$

What accounts for dilation?

<u>DEFN.</u>(Recall) Given a probability function p, a simple measure of stochastic independence:

$$\mathrm{S}_p(E, F) \quad =_{df} \quad \frac{p(E \cap F)}{p(E)p(F)}.$$

What accounts for dilation?

<u>Defn.</u>(Recall) Given a probability function p, a simple measure of stochastic independence:

$$\mathrm{S}_p(E, F) \quad =_{df} \quad \frac{p(E \cap F)}{p(E)p(F)}.$$

Given a nonempty set of probabilities $\mathbb{P}$ and events E, $F \in \mathcal{A}$, consider the following sets:

What accounts for dilation?

DEFN.(Recall) Given a probability function p, a simple measure of stochastic independence:

$$\mathrm{S}_p(E, F) \quad =_{df} \quad \frac{p(E \cap F)}{p(E)p(F)}.$$

Given a nonempty set of probabilities $\mathbb{P}$ and events $E,\ F \in \mathcal{A}$, consider the following sets:

$p \in \mathrm{S}^{+}_{\mathbb{P}}(E, F) \iff$ E and F are pos. assoc. with respect to p

$p \in \mathrm{S}^{-}_{\mathbb{P}}(E, F) \iff$ E and F are neg. assoc. with respect to p

$p \in \mathrm{I}_{\mathbb{P}}(E, F) \iff$ p is on the surface of ind. for E and F

What accounts for dilation?

DEFN.(Recall) Given a probability function p, a simple measure of stochastic independence:

$$\mathrm{S}_p(E,F) \quad =_{df} \quad \frac{p(E\cap F)}{p(E)p(F)}.$$

Given a nonempty set of probabilities $\mathbb{P}$ and events E, $F \in \mathcal{A}$, consider the following sets:

$p \in \mathrm{S}^+_{\mathbb{P}}(E,F) \iff$ E and F are pos. assoc. with respect to p

$p \in \mathrm{S}^-_{\mathbb{P}}(E,F) \iff$ E and F are neg. assoc. with respect to p

$p \in \mathrm{I}_{\mathbb{P}}(E,F) \iff$ p is on the surface of ind. for E and F

That is:

$$\mathrm{S}^+_{\mathbb{P}}(E,F) \quad =_{df} \quad \{p\in\mathbb{P} : \mathrm{S}_p(E,F) > 1\}$$

$$\mathrm{S}^-_{\mathbb{P}}(E,F) \quad =_{df} \quad \{p\in\mathbb{P} : \mathrm{S}_p(E,F) < 1\}$$

$$\mathrm{I}_{\mathbb{P}}(E,F) \quad =_{df} \quad \{p\in\mathbb{P} : \mathrm{S}_p(E,F) = 1\}.$$

What accounts for dilation?

<u>Thm.</u> (Pedersen and Wheeler 2014)
Let $(\Omega, \mathcal{A}, \mathbb{P}, \underline{P})$ be a lower probability space such that $\mathbb{P}$ is weak*-closed and convex,

What accounts for dilation?

THM. (Pedersen and Wheeler 2014)
Let $(\Omega, \mathcal{A}, \mathbb{P}, \underline{P})$ be a lower probability space such that $\mathbb{P}$ is weak*-closed and convex, let $\mathcal{B} = \{H_i : i \in I\}$ be a positive measurable partition,

What accounts for dilation?

<u>Thm.</u> (Pedersen and Wheeler 2014)
Let $(\Omega, \mathcal{A}, \mathbb{P}, \underline{P})$ be a lower probability space such that $\mathbb{P}$ is weak*-closed and convex, let $\mathcal{B} = \{H_i : i \in I\}$ be a positive measurable partition, and let $E \in \mathcal{A}$.

What accounts for dilation?

<u>THM.</u> (Pedersen and Wheeler 2014)

Let $(\Omega, \mathcal{A}, \mathbb{P}, \underline{P})$ be a lower probability space such that $\mathbb{P}$ is weak*-closed and convex, let $\mathcal{B} = \{H_i : i \in I\}$ be a positive measurable partition, and let $E \in \mathcal{A}$. Then the following are equivalent:

(i) $\mathcal{B}$ dilates E;

What accounts for dilation?

THM. (Pedersen and Wheeler 2014)
Let $(\Omega, \mathcal{A}, \mathbb{P}, \underline{P})$ be a lower probability space such that $\mathbb{P}$ is weak*-closed and convex, let $\mathcal{B} = \{H_i : i \in I\}$ be a positive measurable partition, and let $E \in \mathcal{A}$. Then the following are equivalent:

(i) $\mathcal{B}$ dilates E;

(ii) There is $(\epsilon_i)_{i \in I} \in \mathbb{R}^I_+$ such that for every $i \in I$:

$$\underline{\mathbb{P}}(E|H_i, \epsilon_i) \subseteq \mathrm{S}^-(E, H_i) \quad \text{and} \quad \overline{\mathbb{P}}(E|H_i, \epsilon_i) \subseteq \mathrm{S}^+(E, H_i);$$

What accounts for dilation?

THM. (Pedersen and Wheeler 2014)
Let $(\Omega, \mathcal{A}, \mathbb{P}, \underline{P})$ be a lower probability space such that $\mathbb{P}$ is weak*-closed and convex, let $\mathcal{B} = \{H_i : i \in I\}$ be a positive measurable partition, and let $E \in \mathcal{A}$. Then the following are equivalent:

(i) $\mathcal{B}$ dilates E;

(ii) There is $(\epsilon_i)_{i \in I} \in \mathbb{R}^I_+$ such that for every $i \in I$:

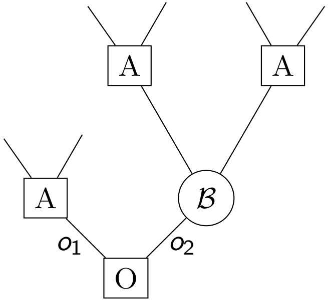
$$\underline{\mathbb{P}}(E|H_i, \epsilon_i) \subseteq \mathrm{S}^-(E, H_i) \quad \text{and} \quad \overline{\mathbb{P}}(E|H_i, \epsilon_i) \subseteq \mathrm{S}^+(E, H_i);$$

(iii) There are $(\underline{\epsilon}_i)_{i \in I}, (\overline{\epsilon}_i)_{i \in I} \in \mathbb{R}^I_+$ such that for every $i \in I$:

$$\underline{\mathbb{P}}(E|H_i, \underline{\epsilon}_i) \subseteq \mathrm{S}^-(E, H_i) \quad \text{and} \quad \overline{\mathbb{P}}(E|H_i, \overline{\epsilon}_i) \subseteq \mathrm{S}^+(E, H_i).$$

Each radius $\underline{\epsilon_i}$ may be chosen to be the unique positive min of $|p(E \mid H_i) - \underline{P}(E \mid H_i)|$ on $C_i^+ := \{p \in \mathbb{P} : S_p(E, H_i) \geq 1\}$, unique $\overline{\epsilon_i}$ with respect to C_i^-.

Good's Principle



You should delay making a terminal decision (A) between alternative courses of action if the opportunity arises to learn, at zero cost to you, the outcome of an experiment ($\mathcal{B}$) relevant to the decision

Review of probabilistic independence

Two events E and F in $\mathcal{A}$ are stochastically independent just in case:

(SI) $p(E \cap F) = p(E)p(F)$.

Review of probabilistic independence

Two events E and F in $\mathcal{A}$ are stochastically independent just in case:

(SI) $p(E \cap F) = p(E)p(F)$.

An event F is epistemically irrelevant to event E iff:

(ER) $p(E \mid F) = p(E)$, when $p(F) > 0$.

Review of probabilistic independence

Two events E and F in $\mathcal{A}$ are stochastically independent just in case:

(SI) $p(E \cap F) \;=\; p(E)p(F)$.

An event F is epistemically irrelevant to event E iff:

(ER) $p(E \mid F) \;=\; p(E)$, when $p(F) \;>\; 0$.

Two events E and F are epistemically independent just in case:

(EI) E is epistemically irrelevant to F and F is epistemically irrelevant to E.

IP independence

Our three familiar independence concepts:

($\mathbb{ER}$) Given $\underline{\mathrm{P}}(F), \underline{\mathrm{P}}(F^c) > 0$, event F is *epistemically irrelevant* to E if and only if:

1. $\underline{\mathrm{P}}(E \mid F) = \underline{\mathrm{P}}(E \mid F^c) = \underline{\mathrm{P}}(E)$;
2. $\overline{\mathrm{P}}(E \mid F) = \overline{\mathrm{P}}(E \mid F^c) = \overline{\mathrm{P}}(E)$.

For imprecise probabilities, epistemic irrelevance is not symmetric: F may be epistemically irrelevant to E without E being epistemically irrelevant to F.

IP independence

Our three familiar independence concepts:

($\mathbb{ER}$) Given $\underline{\mathrm{P}}(F), \underline{\mathrm{P}}(F^c) > 0$, event F is *epistemically irrelevant* to E if and only if:

1. $\underline{\mathrm{P}}(E \mid F) = \underline{\mathrm{P}}(E \mid F^c) = \underline{\mathrm{P}}(E)$;
2. $\overline{\mathrm{P}}(E \mid F) = \overline{\mathrm{P}}(E \mid F^c) = \overline{\mathrm{P}}(E)$.

For imprecise probabilities, epistemic irrelevance is not symmetric: F may be epistemically irrelevant to E without E being epistemically irrelevant to F.

($\mathbb{EI}$) E is *epistemically independent* of F just when both F is epistemically irrelevant to E and E is epistemically irrelevant to F.

IP independence

Our three familiar independence concepts:

($\mathbb{ER}$) Given $\underline{\mathrm{P}}(F), \underline{\mathrm{P}}(F^c) > 0$, event F is *epistemically irrelevant* to E if and only if:

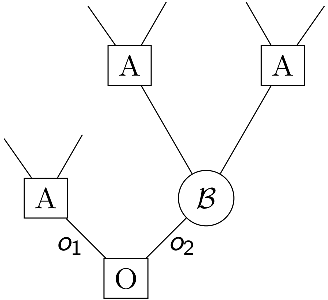
1. $\underline{\mathrm{P}}(E \mid F) = \underline{\mathrm{P}}(E \mid F^c) = \underline{\mathrm{P}}(E)$;
2. $\overline{\mathrm{P}}(E \mid F) = \overline{\mathrm{P}}(E \mid F^c) = \overline{\mathrm{P}}(E)$.

For imprecise probabilities, epistemic irrelevance is not symmetric: F may be epistemically irrelevant to E without E being epistemically irrelevant to F.

($\mathbb{EI}$) E is *epistemically independent* of F just when both F is epistemically irrelevant to E and E is epistemically irrelevant to F.

($\mathbb{SI}$) E and F are *completely stochastically independent* if and only if for all $p \in \mathbb{P}$, $p(E \cap F) = p(E)p(F)$.

Good's Principle



You *should* *delay making a terminal decision (A) between alternative courses of action if the opportunity arises to learn, at zero cost to you, the outcome of an experiment ($\mathcal{B}$) relevant to the decision*

Decision rules in IP

Assumptions in our delayed decision problem:

- You don't change your preferences
- You update by Generalized Bayesian conditionalization
- You *choose to maximize your minimum expected loss* (Γ-Maximin)

Decision rules in IP

Assumptions in our delayed decision problem:

- You don't change your preferences
- You update by Generalized Bayesian conditionalization
- You *choose to maximize your minimum expected loss* (Γ-Maximin)

But, there are other decision rules for IP:

- Γ-**maximin** – requires you to flout Good's principle

Decision rules in IP

Assumptions in our delayed decision problem:

- You don't change your preferences
- You update by Generalized Bayesian conditionalization
- You *choose to maximize your minimum expected loss* (Γ-Maximin)

But, there are other decision rules for IP:

- Γ-**maximin** – requires you to flout Good's principle
- **E-admissibility** – permits you to flout Good's principle

Decision rules in IP

Assumptions in our delayed decision problem:

- You don't change your preferences
- You update by Generalized Bayesian conditionalization
- You *choose to maximize your minimum expected loss* (Γ-Maximin)

But, there are other decision rules for IP:

- Γ-**maximin** – requires you to flout Good's principle
- **E-admissibility** – permits you to flout Good's principle
- **E-admissibility + maximize expected utility** – forbids you to flout Good's principle.

Take away

- Dilation itself is not a problem

  it dramatizes the effect of not knowing how two variables are related

Take away

- Dilation itself is not a problem

  it dramatizes the effect of not knowing how two variables are related

- In IP there are **several** logically distinct independence concepts

  cf. to conditioning on zero-probability events on a single distribution

Take away

- Dilation itself is not a problem

  it dramatizes the effect of not knowing how two variables are related

- In IP there are **several** logically distinct independence concepts

  cf. to conditioning on zero-probability events on a single distribution

- Dilation & Good's principle
  - Some decision rules violate Good's principle, others don't, and some (E admissibility) permit violations.

Take away

- Dilation itself is not a problem

  it dramatizes the effect of not knowing how two variables are related

- In IP there are **several** logically distinct independence concepts

  cf. to conditioning on zero-probability events on a single distribution

- Dilation & Good's principle
  - Some decision rules violate Good's principle, others don't, and some (E admissibility) permit violations.

- Strict Bayes & Good's principle
  - improper priors – assign 'infinite' mass to the sure event
  - Finite (but not countably additive) priors
    - *violations appear only in infinite spaces.*

Take away

- Dilation itself is not a problem

  it dramatizes the effect of not knowing how two variables are related

- In IP there are **several** logically distinct independence concepts

  cf. to conditioning on zero-probability events on a single distribution

- Dilation & Good's principle
  - Some decision rules violate Good's principle, others don't, and some (E admissibility) permit violations.

- Strict Bayes & Good's principle
  - improper priors – assign 'infinite' mass to the sure event
  - Finite (but not countably additive) priors
    - *violations appear only in infinite spaces.*

- Reflection on strict Bayesian methods

References

- Joyce, J (2010). A defense of imprecise credences in inference and decision making. *Philosophical Perspectives* 24(1), 281-323.
- Pedersen, P and G Wheeler (2014). Demystifying dilation. *Erkenntnis* 79(6): 1305–1342.
- Mayo-Wilson, C. and G. Wheeler. Scoring Imprecise Credences: A Mildly Immodest Proposal. *Philosophy and Phenomenological Research*, forthcoming.
- Seidenfeld, T and L Wasserman (1993). Dilating sets of probabilities. *The Annals of Statistics* 21:1139–54.
- Seidenfeld, T., Schervish, M., and Kadane, J. (2012). Forecasting with imprecise probabilities. *International Journal of Approximate Reasoning* 53(8), 1248–1261.
- Walley, P (1991). *Statistical Reasoning with Imprecise Probabilities*. Chapman and Hall, London.
- White, R (2010). Evidential symmetry and mushy credences. In Gendler and Hawthorne (eds.), *Oxford Studies in Epistemology*, volume 3, pp. 161-186, Oxford University Press.

$\underline{\text{Defn.}}$ A <u>probability function</u> is a real-valued function p defined on an algebra $\mathcal{A}$ over a set of states Ω satisfying the following three conditions:

(P1) $p(E) \geq 0$ for every $E \in \mathcal{A}$;

(P2) $p(\Omega) = 1$;

(P3) $p(E \cup F) = p(E) + p(F)$ ($\forall$ pairwise disjoint elements $E, F \in \mathcal{A}$).

(P4.1) $\underline{\mathrm{P}}(E) = \inf\{\, p(E) : p \in \mathbb{P} \,\}$.

(P5.1) $\overline{\mathrm{P}}(E) = \sup\{\, p(E) : p \in \mathbb{P} \,\}$.

(P4.2) $\underline{\mathrm{P}}(E \mid F) = \inf\{\, p(E \mid F) : p \in \mathbb{P} \,\}$.

(P5.2) $\overline{\mathrm{P}}(E \mid F) = \sup\{\, p(E \mid F) : p \in \mathbb{P} \,\}$.

(P6) If $\underline{\mathrm{P}} = \overline{\mathrm{P}}$, then $\{p\} = \mathbb{P}$ and $p = \underline{\mathrm{P}} = \overline{\mathrm{P}}$.

(P7) $\overline{\mathrm{P}}(E) = 1 - \underline{\mathrm{P}}(E^c)$.

(P8) $\underline{\mathrm{P}}(E \cup F) \geq \underline{\mathrm{P}}(E) + \underline{\mathrm{P}}(F) - \underline{\mathrm{P}}(E \cap F)$.

return

Proof of Pedersen and Wheeler, Thm 5.1

We first show that (i) $\iff$ (iii), and we then show that (i) $\iff$ (ii).

(i) $\Rightarrow$ (iii) Suppose that $\mathcal{B}$ dilates E. Then for each $i \in I$, $\underline{\mathrm{P}}(E|H_i) < \underline{\mathrm{P}}(E) \leq \overline{\mathrm{P}}(E) < \overline{\mathrm{P}}(E|H_i)$. For each $i \in I$, consider the real-valued function $\underline{\epsilon}_i(p) =_{df} |p(E|H_i) - \underline{\mathrm{P}}(E|H_i)|$ and the real-valued function $\mathrm{S}_{p,i}(E, H_i)$. We recall that the weak* topology on the dual space is a locally convex topological vector space with respect to which every evaluation functional f^* is a real-valued continuous linear functional on the dual space. It follows that $\underline{\epsilon}_i(p)$ and $\mathrm{S}_{p,i}(E, H_i)$ are continuous functions of p on $\mathbb{P}$ for each $i \in I$. return (continued ...)

Now let $i \in I$. By hypothesis, there is $p_1 \in \mathbb{P}$ such that $\mathrm{S}_{p_1,i}(E, H_i) > 1$, so importantly, C_i^+ is nonempty. Then since $C_i^+ =_{df} \{p \in \mathbb{P} : \mathrm{S}_p(E, H_i) \geq 1\}$ is a weak*-closed and so weak*-compact set, it follows that $\underline{\epsilon}_i$ achieves a minimum value on C_i^+ (and the set of minimizers of $\underline{\epsilon}_i$ is also compact). Choosing a minimizer $p_i \in \mathbb{P}$ of $\underline{\epsilon}_i$, we see that for every $p \in \mathbb{P}$, if $|p(E|H_i) - \underline{\mathrm{P}}(E|H_i)| < \underline{\epsilon}_i(p_i) = |p_i(E|H_i) - \underline{\mathrm{P}}(E|H_i)|$, then $\mathrm{S}_p(E, H_i) < 1$. We have accordingly shown that $\underline{\mathbb{P}}(E|H_i, \underline{\epsilon}_i(p_i)) \subseteq \mathrm{S}^-_{\mathbb{P}}(E, H_i)$.

Of course, we may suppress reference to the minimizer p_i in $\underline{\epsilon}_i(p_i)$. The other inclusion $\overline{\mathbb{P}}(E|H_i, \overline{\epsilon}_i) \subseteq \mathrm{S}^+_{\mathbb{P}}(E, H_i)$ is established by a similar argument. return (continued . . .)

(iii) $\Leftarrow$ (i) Suppose that $\mathcal{B}$ does not dilate E. Then there is $i \in I$ such that $\underline{\mathrm{P}}(E|H_i) \geq \underline{\mathrm{P}}(E)$ or $\overline{\mathrm{P}}(E) \geq \overline{\mathrm{P}}(E|H_i)$. We may assume without loss of generality that $\underline{\mathrm{P}}(E|H_i) \geq \underline{\mathrm{P}}(E)$ for some $i \in I$. First, if $\underline{\mathrm{P}}(E) \leq \overline{\mathrm{P}}(E) \leq \underline{\mathrm{P}}(E|H_i)$, then choosing a minimizer $p \in \mathbb{P}$ of $\underline{\mathrm{P}}(E|H_i)$, we see that $\mathrm{S}_p(E, H_i) \geq 1$. Second, if $\underline{\mathrm{P}}(E) < \underline{\mathrm{P}}(E|H_i) < \overline{\mathrm{P}}(E)$, then for every $\epsilon > 0$ we can find a convex combination $p \in \mathbb{P}$ of $p_0, p_1 \in \mathbb{P}$ assigning a probability to E within ϵ-distance below $\underline{\mathrm{P}}(E|H_i)$, where $\underline{\mathrm{P}}(E) \leq p_0(E) < \underline{\mathrm{P}}(E|H_i) < p_1(E) \leq \overline{\mathrm{P}}(E)$, so $\mathrm{S}_p(E, H_i) > 1$. Third, if $\underline{\mathrm{P}}(E) = \underline{\mathrm{P}}(E|H_i) < \overline{\mathrm{P}}(E)$, then choosing a minimizer $p \in \mathbb{P}$ of $\underline{\mathrm{P}}(E)$, we see that $\mathrm{S}_p(E, H_i) \geq 1$. Evidently, the conditions of the main claim cannot be jointly satisfied.

return (continued ...)

(i) $\Leftrightarrow$ (ii) On the one hand, suppose that (i) obtains. Then since (iii) accordingly obtains, define $(\epsilon_i)_{i \in I}$ by setting $\epsilon_i =_{df} \min(\underline{\epsilon}_i, \overline{\epsilon}_i)$ for each $i \in I$. Clearly the inclusions still obtain for the ϵ_i. On the other hand, if (ii) obtains, obviously by setting $\underline{\epsilon}_i =_{df} \epsilon_i$ and $\overline{\epsilon}_i =_{df} \epsilon_i$ for each $i \in I$, condition (iii) obtains and so (i) obtains. $\square$

return

2. Review of results

Here, we briefly review the results from Seidenfeld and Wasserman (1993). Given M, define $M_*(A)=\{P\in M;\ P(A)=\underline{P}(A)\}$, $M^*(A)=\{P\in M;\ P(A)=\bar{P}(A)\}$, $M_*(A\mid B)=\{P\in M; P(A\mid B)=\underline{P}(A\mid B)\}$ and $M^*(A\mid B)=\{P\in M; P(A\mid B)=\bar{P}(A\mid B)\}$. For $P\in\mathscr{P}$, the set of all probabilities, define $S_P(A, B)=P(A\cap B)/(P(A)P(B))$, $d_P(A, B)=P(A\cap B)-P(A)P(B)$, $\Sigma^+(A, B)=\{P\in\mathscr{P};\ d_P(A, B)>0\}$ and $\Sigma^-(A, B)=\{P\in\mathscr{P};\ d_P(A, B)<0\}$. The surface of independence for events A and B is defined by $\mathscr{I}(A, B)=\{P\in\mathscr{P}; d_P(A, B)=0\}$. Finally, define

$$S_0=\inf_{P\in M_*(A\mid B)} S_P(A, B), \qquad S^0=\sup_{P\in M^*(A\mid B)} S_P(A, B),$$

$$M_0=\{P\in M_*(A\mid B): S_P=S_0\}, \qquad M^0=\{P\in M^*(A\mid B); S_P=S_0\},$$

and

$$P_0=\inf_{P\in M_0} P(A), \qquad P^0=\sup_{P\in M^0} P(A).$$

Result 1. Suppose that $\mathscr{B}=\{B, B^c\}$. A necessary and sufficient condition for strict dilation is that

$$S_0<\frac{\underline{P}(A)}{P_0(A)}\leqslant 1\leqslant\frac{\bar{P}(A)}{P^0(A)}<S^0.$$

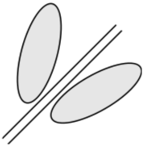
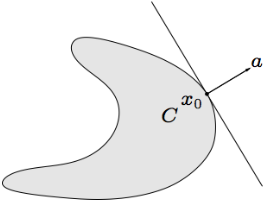


Figure 5.3. Strong separation.

Figure 5.4. These sets cannot be separated by a hyperplane.



return