Modelli Bayesiani inferenziali con multiple prior: inviluppi inferiori e superiori

Barbara Vantaggi

Dip. Scienze di Base e Applicate per l'Ingegneria
"La Sapienza" Università di Roma

Convegno in onore della Prof. Patrizia Berti

8-9 Giugno 2015

Models with multiple priors: motivations

- Statistics: elicitation of priors, partial identifiable models (models with latent variables)
- Applications: Gilboa-Schmeidler decision model, Reliability
- Fuzzy set theory

Non-additive uncertainty measures and probability

$\varphi : \mathcal{A} \to [0,1]$ s.t. $\varphi(\emptyset) = 0$, $\varphi(\Omega) = 1$ uncertainty measure:

capacity: $A \subseteq B \Rightarrow \varphi(A) \leq \varphi(B)$;

superadditive: $A \wedge B = \emptyset \Rightarrow \varphi(A \vee B) \geq \varphi(A) + \varphi(B)$;

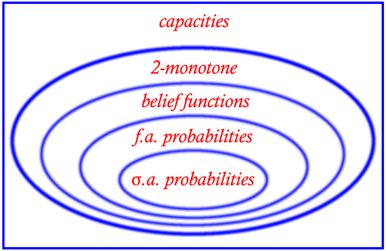
n-monotone: $\varphi\left(\bigvee_{i=1}^{n} E_i\right) \geq \sum_{\emptyset \neq I \subseteq \{1,\dots,n\}} (-1)^{|I|+1} \varphi\left(\bigwedge_{i \in I} E_i\right)$;

belief function: n-monotone for $n \in \mathbb{N}$, $n \geq 2$.

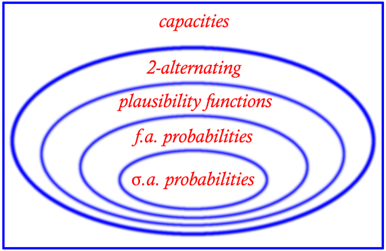
$\overline{\varphi} : \mathcal{A} \to [0,1]$, $\overline{\varphi}(A) = 1 - \varphi(A^c)$ for every $A \in \mathcal{A}$, dual measure.

Uncertainty measures

$$\varphi : \mathcal{A} \to [0, 1]$$



$$\overline{\varphi} : \mathcal{A} \to [0, 1]$$



Inner and outer measures

Let $P : \mathcal{A} \to [0,1]$ be a (finitely additive) probability and let $\mathcal{B} \supset \mathcal{A}$ be a super-algebra, consider the inner and outer measures $\underline{P}$ and $\overline{P}$, defined on any $E \in \mathcal{B}$ as

$$\begin{aligned}\underline{P}(E) &= \sup\{P(K) : K \subseteq E, K \in \mathcal{A}\}\\ \overline{P}(E) &= \inf\{P(K) : E \subseteq K, K \in \mathcal{A}\}\end{aligned}$$

The inner measure $\underline{P} : \mathcal{B} \to [0,1]$ of a probability $P : \mathcal{A} \to [0,1]$ is a belief function[1]

Any belief function on an algebra $\mathcal{A}'$ can be seen as the restriction of an inner measure of a finitely additive probability on another algebra $\mathcal{A}$

[1] [Choquet 1954]

2-monotone capacities

Any 2-monotone capacity φ (and its dual $\overline{\varphi}$) induces a not empty closed convex set of finitely additive probabilities (core)

$$\mathcal{P}_\varphi = \{\tilde{\pi} : \mathcal{A} \to [0,1] \, : \, \varphi \leq \tilde{\pi} \leq \overline{\varphi}\}$$

$$\varphi = \min \mathcal{P}_\varphi \text{ and } \overline{\varphi} = \max \mathcal{P}_\varphi$$

2-monotonicity is not always satisfied

The lower envelope of the probabilities on the Boolean algebra $\mathcal{B}$ is superadditive, but not necessarily 2-monotone.

The lower envelope of the joint probabilities related to two marginal distributions for r.v. X and Y by considering all the possible copulas is superadditive, but not 2-monotone.

Consequences of 2-monotonicity

Given an $\mathcal{A}$-continuous function[2] X, a 2-monotone capacity φ and its dual $\overline{\varphi}$ it holds [3]

$$\mathrm{C}\!\!\!\!\int X \mathrm{d}\varphi = \min \left\{ \mathrm{D}\!\!\!\!\int X \mathrm{d}\tilde{\pi} \;:\; \tilde{\pi} \in \mathcal{P}_\varphi \right\} = \min_{\tilde{\pi} \in \mathcal{P}_\varphi} \mathbb{E}_{\tilde{\pi}}(X)$$

and

$$\mathrm{C}\!\!\!\!\int X \mathrm{d}\overline{\varphi} = \max \left\{ \mathrm{D}\!\!\!\!\int X \mathrm{d}\tilde{\pi} \;:\; \tilde{\pi} \in \mathcal{P}_\varphi \right\} = \max_{\tilde{\pi} \in \mathcal{P}_\varphi} \mathbb{E}_{\tilde{\pi}}(X)$$

where

$$\mathcal{P}_\varphi = \{ \tilde{\pi} : \mathcal{A} \to [0,1] \;:\; \varphi \leq \tilde{\pi} \leq \overline{\varphi} \}$$

---

[2] $\mathcal{A}$-continuity: X is bounded and for every $t \in \mathbb{R}$ and $\epsilon > 0$ there exists $A \in \mathcal{A}$ s.t. $(X \geq t) \supseteq A \supseteq (X \geq t + \epsilon)$

[3] [Schmeidler 1986]

Choquet integral and D-integral

Choquet integral

$$\oint\!\!\!\!\!\mathbf{C}\, X \mathrm{d}\varphi = \int_{-\infty}^{0} [\varphi(X \geq t) - 1] \mathrm{d}t + \int_{0}^{+\infty} \varphi(X \geq t) \mathrm{d}t.$$

$\Rightarrow$ If φ is a finitely additive probability, and X is $\mathcal{A}$-continuous [4]

$$\oint\!\!\!\!\!\mathbf{D}\, X \mathrm{d}\varphi = \oint\!\!\!\!\!\mathbf{C}\, X \mathrm{d}\varphi$$

where the first integral is a D-integral [5]

[4] [Schmeidler 1986]
[5] [Bhaskara Rao & Bhaskara Rao 1983]

Coherent conditional probability

Definition

Given an arbitrary set $\mathcal{G} = \{E_i|H_i\}_{i\in I}$ of conditional events, an assessment $P : \mathcal{G} \to [0, 1]$ is a **coherent conditional probability** if and only if, for every $n \in \mathbb{N}$, every $E_{i_1}|H_{i_1}, \ldots, E_{i_n}|H_{i_n} \in \mathcal{G}$ and every $s_1, \ldots, s_n \in \mathbb{R}$, the random gain

$$G = \sum_{j=1}^{n} s_i(P(E_{i_j}|H_{i_j}) - |E_{i_j}|)|H_{i_j}|$$

(where $|\cdot|$ denotes the indicator of an event) is such that

$$\min_{H_0^0} G \leq 0 \leq \max_{H_0^0} G$$

where $H_0^0 = \bigvee_{j=1}^{n} H_{i_j}$.

Extension of a coherent conditional probability

Theorem (de Finetti's fundamental theorem)

$P(\cdot|\cdot)$ on $\mathcal{G}$ can be extended (generally not in a unique way) as a coherent conditional probability on $\mathcal{G}' \supset \mathcal{G}$ if and only if $P(\cdot|\cdot)$ is a coherent conditional probability.

Envelopes of coherent extensions

Set of coherent extensions

$$\mathcal{P} = \{\tilde{P}(\cdot|\cdot) \ : \ \text{coherent conditional probability on } \mathcal{G}' \text{ extending } P\}.$$

$\Rightarrow$ $\mathcal{P}$ is a compact subset of $[0,1]^{\mathcal{G}'}$ endowed with the product topology of pointwise convergence

Envelopes of coherent extensions

$$\underline{P} = \min \mathcal{P} \quad \text{and} \quad \overline{P} = \max \mathcal{P}.$$

$\Rightarrow$ $\underline{P}, \overline{P}$ are said coherent lower and upper conditional probabilities

The role of coherence in Bayesian statistics

[Regazzini 1987, Berti et al. 1991, 1994]

In the classical Bayesian setting[6]

- $\pi : \mathcal{A}_{\mathcal{L}} \to [0,1]$, finitely additive prior probability;
- $\sigma : \mathcal{A} \times \mathcal{L} \to [0,1]$, strategy s.t. for every $H_i \in \mathcal{L}$

  (S1) $\sigma(F|H_i) = 1$ if $F \wedge H = H$ for $F \in \mathcal{A}$;

  (S2) $\sigma(\cdot|H_i)$ is a finitely additive probability on $\mathcal{A}$;
- $\lambda = \sigma_{|\mathcal{A}_{\mathcal{E}} \times \mathcal{L}}$, statistical model

$\Rightarrow$ $\{\pi, \lambda\}$ and $\{\pi, \sigma\}$ is a coherent conditional probability

[6] $\mathcal{L} = \{H_i\}_{i \in I}, \mathcal{E} = \{E_j\}_{j \in J}$, partitions; $\mathcal{A}_{\mathcal{L}}, \mathcal{A}_{\mathcal{E}}$, Boolean algebras with $\langle \mathcal{L} \rangle \subseteq \mathcal{A}_{\mathcal{L}} \subseteq \langle \mathcal{L} \rangle^*$, $\langle \mathcal{E} \rangle \subseteq \mathcal{A}_{\mathcal{E}} \subseteq \langle \mathcal{E} \rangle^*$

The role of coherence in Bayesian statistics

Given a statistical model λ on $\mathcal{A}_\mathcal{E} \times \mathcal{L}$ and $\mathcal{A} = \langle \mathcal{A}_\mathcal{L} \cup \mathcal{A}_\mathcal{E} \rangle$, then there exists a unique strategy σ on $\mathcal{A} \times \mathcal{L}$ such that $\sigma_{|\mathcal{A}_\mathcal{E} \times \mathcal{L}} = \lambda$.

If $\mathcal{A} \supset \langle \mathcal{A}_\mathcal{L} \cup \mathcal{A}_\mathcal{E} \rangle$ the strategy σ on $\mathcal{A} \times \mathcal{L}$ such that $\sigma_{|\mathcal{A}_\mathcal{E} \times \mathcal{L}} = \lambda$ is not unique.

Given a strategy σ on $\mathcal{A} \times \mathcal{L}$ and a prior probability $\pi : \mathcal{A}_\mathcal{L} \to [0, 1]$, the coherent extension $\tilde{P} : \mathcal{A} \to [0, 1]$ is not unique[7].

An aim is to determine the lower and upper envelope of the coherent extensions $\tilde{P}$ of $\{\sigma, \pi\}$.

[7] Regazzini 1987

Extensions of the classical Bayesian paradigm

- Prior probability available only on some events ($\mathcal{G} \subset \mathcal{A}_{\mathcal{L}}$) or on a space different from the one where the statistical model is given ($\mathcal{A}_{\mathcal{L}'}$ with $\mathcal{L}' \neq \mathcal{L}$)
- Multiple priors in a set $\mathcal{P}_{\varphi}$, where φ is a 2-monotone prior capacity
- No prior information modelled as a vacuous 2-monotone prior capacity φ

Generalization:

1. Consider a 2-monotone prior capacity φ;
2. Make inference on more complex events than those related to the posterior probability $H|E_j \in \mathcal{A}_{\mathcal{L}} \times \mathcal{E}$, i.e., on events in $\mathcal{A} \times \mathcal{A}^0$ with $\mathcal{A} = \langle \mathcal{A}_{\mathcal{L}} \cup \mathcal{A}_{\mathcal{E}} \rangle$

Joint probability on $\mathcal{A}$

Uniqueness of the extension on $\mathcal{A}$ is lost for an infinite $\mathcal{L}$: we have a set of joint probabilities $\mathcal{P}^{\mathbf{j}}$ with $\underline{P}^{\mathbf{j}} = \min \mathcal{P}^{\mathbf{j}}$ and $\overline{P}^{\mathbf{j}} = \max \mathcal{P}^{\mathbf{j}}$

Theorem

The lower envelope $\underline{P}^{\mathbf{j}}(\cdot)$ (lower joint probability) of the of extensions of $\{\pi, \sigma\}$ on $\mathcal{A}$ is such that for every $F \in \mathcal{A}$ it holds

$$\underline{P}^{\mathbf{j}}(F) = \sup_{\mathcal{L}^{\mathcal{F}} \subseteq \mathcal{A}_{\mathcal{L}}} \left\{ \sum_{h=1}^{n} \sigma(F|H_{i_h})\pi(H_{i_h}) + \sum_{B_k \subseteq F} \pi(B_k) \right\},$$

where $\mathcal{L}^{\mathcal{F}} = \{H_{i_h}\}_{h=1}^{n} \cup \{B_k\}_{k=1}^{m} \subseteq \mathcal{A}_{\mathcal{L}}$ is a finite partition of Ω.

If $\mathcal{L}$ is countable and π is countably additive on $\mathcal{A}_{\mathcal{L}}$, then, for every $F \in \mathcal{A}$,

$$\underline{P}^{\mathbf{j}}(F) = \sum_{i=1}^{\infty} \sigma(F|H_i)\pi(H_i) = \overline{P}^{\mathbf{j}}(F)$$

is a finitely additive probability on $\mathcal{A}$.

If moreover $\sigma(\cdot|H_i)$ is countably additive on $\mathcal{A}$ for every $H_i \in \mathcal{L}$, then $\underline{P}^{\mathbf{j}}$ is countably additive.

Bayes theorem in a coherent (finitely additive) setting

$$L^{j}(F,K) = \min\{\tilde{P}(F\wedge K) : \tilde{P}(F^c\wedge K)=\overline{P}(F^c\wedge K), \tilde{P}\in\mathcal{P}^{j}\},$$
$$U^{j}(F^c,K) = \max\{\tilde{P}(F^c\wedge K) : \tilde{P}(F\wedge K)=\underline{P}(F\wedge K), \tilde{P}\in\mathcal{P}^{j}\}.$$

Theorem

The lower envelope $\underline{P}(\cdot|\cdot)$ of the extensions of $\{\pi,\sigma\}$ on $\mathcal{A}\times\mathcal{A}^0$ is such that for every $F|K\in\mathcal{A}\times\mathcal{A}^0$ with $F\wedge K\neq K$:

(i) *if $\underline{P}^{j}(K)>0$, then*

$$\underline{P}(F|K)=\min\left\{\frac{\underline{P}^{j}(F\wedge K)}{\underline{P}^{j}(F\wedge K)+U^{j}(F^c,K)},\frac{L^{j}(F,K)}{L^{j}(F,K)+\overline{P}^{j}(F^c\wedge K)}\right\};$$

(ii) *if $\underline{P}^{j}(K)=0$, then*

$$\underline{P}(F|K)=\begin{cases}\min\limits_{i\in I_2^{F|K}}\frac{\sigma(F\wedge K|H_i)}{\sigma(K|H_i)} & \text{if } I_2^{F|K}\neq\emptyset=I_3^{F|K} \text{ and } \mathrm{card} I_2^{F|K}<\aleph_0 \\ & \text{and } \sigma(K|H_i)>0 \text{ for all } i\in I_2^{F|K}, \\ 0 & \text{otherwise,}\end{cases}$$

$I_2^{F|K}=\{i\in I : H_i\wedge F\wedge K\neq\emptyset\neq H_i\wedge F^c\wedge K\}$ *and* $I_3^{F|K}=\{i\in I : H_i\wedge F\wedge K=\emptyset\neq H_i\wedge F^c\wedge K\}$.

Example: formation of a test with i questions

PROBLEM: Choose a number $i \in \mathbb{N}$ and form an urn U with i questions for each of 3 subjects E_1, E_2, E_3 for a total of $3i$ questions.

- $\mathcal{L} = \{H_i\}_{i\in\mathbb{N}}$, $\mathcal{E} = \{E_1, E_2, E_3\}$ with $H_i \wedge E_j \neq \emptyset$ for every i, j
- $\mathcal{A}_{\mathcal{L}} = \langle \mathcal{L} \rangle$, $\mathcal{A}_{\mathcal{E}} = \langle \mathcal{E} \rangle$ and $\mathcal{A} = \langle \mathcal{A}_{\mathcal{L}} \cup \mathcal{A}_{\mathcal{E}} \rangle$
- π, finitely additive prior probability defined for every $K \in \mathcal{A}_{\mathcal{L}}$ as

$$\pi(K) = \begin{cases} 0 & \text{if } K = \emptyset, \\ \sum\limits_{H_i \subseteq K} \frac{1}{2^{i+1}} & \text{if } K = \bigvee_{i\in I} H_i \text{ and } \mathrm{card} I < \aleph_0, \\ \frac{1}{2} + \sum\limits_{H_i \subseteq K} \frac{1}{2^{i+1}} & \text{otherwise,} \end{cases}$$

- λ, statistical model singled out by

$$\lambda(E_1|H_i) = \lambda(E_2|H_i) = \lambda(E_3|H_i) = \frac{1}{3} \quad \text{for } i \in \mathbb{N},$$

- consider the conditional event $F|K$ with

$$F = (E_1 \wedge (H_1 \vee H_2 \vee H_3)^c) \quad \text{and} \quad K = (E_1 \wedge (H_1 \vee H_2)^c) \vee (E_2 \wedge (H_1)^c),$$

- $\underline{P}^{\mathbf{j}}(K) = \frac{1}{8}$, $\overline{P}^{\mathbf{j}}(K) = \frac{5}{8}$, $\underline{P}^{\mathbf{j}}(F \wedge K) = \frac{1}{48}$, $\overline{P}^{\mathbf{j}}(F^c \wedge K) = \frac{29}{48}$, $L^{\mathbf{j}}(F, K) = \frac{1}{48}$ and $U^{\mathbf{j}}(F^c, K) = \frac{59}{96}$:

$$\underline{P}(F|K) = \min\left\{\frac{2}{61}, \frac{1}{30}\right\} = \frac{2}{61} \approx 3.28\%.$$

Fully $\mathcal{L}$-disintegrability

Assumption: $\sigma(F|\cdot)$ is $\mathcal{A}_{\mathcal{L}}$-continuous for every $F \in \mathcal{A}$

Fully $\mathcal{L}$-disintegrable extension on $\mathcal{A} \times \mathcal{A}^0$

A full conditional probability $\tilde{P}$ on $\mathcal{A}$ extending $\{\pi, \sigma\}$ is **fully $\mathcal{L}$-disintegrable** if, denoting with $\tilde{\pi} = \tilde{P}_{|\mathcal{A}_{\mathcal{L}} \times \mathcal{A}^0_{\mathcal{L}}}$, for every $F|K \in \mathcal{A} \times \mathcal{A}^0_{\mathcal{L}}$ it holds

$$\tilde{P}(F|K) = \mathfrak{D}\!\!\!\!\int \sigma(F|H_i)\tilde{\pi}(\mathrm{d}H_i|K).$$

Class of fully $\mathcal{L}$-disintegrable extensions

Theorem

The set $\mathcal{P}^{\mathbf{fd}}$ of fully $\mathcal{L}$-disintegrable extensions of $\{\pi, \sigma\}$ on $\mathcal{A} \times \mathcal{A}^0$ is a non-empty compact subset of $[0,1]^{\mathcal{A}\times\mathcal{A}^0}$ endowed with the product topology of pointwise convergence and has envelopes $\underline{P}^{\mathbf{fd}} = \min \mathcal{P}^{\mathbf{fd}}$ and $\overline{P}^{\mathbf{fd}} = \max \mathcal{P}^{\mathbf{fd}}$.

Envelopes of the fully $\mathcal{L}$-disintegrable extensions

$$
\begin{aligned}
L(F,K;A) &= \min\{\tilde{P}(F\wedge K|A) : \tilde{P}(F^c\wedge K|A) = \overline{P}^{\mathbf{fd}}(F^c\wedge K|A), \tilde{P}\in\mathcal{P}^{\mathbf{fd}}\},\\
U(F^c,K;A) &= \max\{\tilde{P}(F^c\wedge K|A) : \tilde{P}(F\wedge K|A) = \underline{P}^{\mathbf{fd}}(F\wedge K|A), \tilde{P}\in\mathcal{P}^{\mathbf{fd}}\}.
\end{aligned}
$$

Theorem

The lower envelope $\underline{P}^{\mathbf{fd}}(\cdot|\cdot)$ of the fully $\mathcal{L}$-disintegrable extensions of $\{\pi,\sigma\}$ on $\mathcal{A}\times\mathcal{A}^0$ is such that for every $F|K\in\mathcal{A}\times\mathcal{A}^0$ with $F\wedge K\neq K$:

(i) if $K\in\mathcal{A}^0_{\mathcal{L}}$

$$
\underline{P}^{\mathbf{fd}}(F|K) = \begin{cases} \dfrac{\oint \sigma(F\wedge K|H_i)\pi(\mathrm{d}H_i)}{\pi(K)} & \text{if } \pi(K)>0,\\[2ex] \inf\limits_{H_i\subseteq K}\sigma(K|H_i) & \text{otherwise,}\end{cases}
$$

(ii) if $K\in\mathcal{A}^0\setminus\mathcal{A}^0_{\mathcal{L}}$

$$
\underline{P}^{\mathbf{fd}}(F|K) = \begin{cases} \min\left\{\frac{\underline{P}^{\mathbf{fd}}(F\wedge K|A)}{\underline{P}^{\mathbf{fd}}(F\wedge K|A)+U(F^c,K;A)}, \frac{L(F,K;A)}{L(F,K;A)+\overline{P}^{\mathbf{fd}}(F^c\wedge K|A)}\right\} & \text{if } \exists A\in\mathcal{A}^0_{\mathcal{L}} \text{ s.t. } K\subseteq A \text{ and } \underline{P}^{\mathbf{fd}}(K|A)>0,\\[2ex] 0 & \text{otherwise.}\end{cases}
$$

Example: formation of a test with i questions (continued)

Given $\mathcal{L} = \{H_i\}_{i\in\mathbb{N}}$, $\mathcal{E} = \{E_1, E_2, E_3\}$, $f(E_j|H_i) = \frac{1}{3}$, $j = 1, 2, 3, i \in I$, for $K \in \mathcal{A}_{\mathcal{L}}$

$$\pi(K) = \begin{cases} 0 & \text{if } K = \emptyset, \\ \sum\limits_{H_i \subseteq K} \frac{1}{2^{i+1}} & \text{if } K = \bigvee_{i\in I} H_i \text{ and } \mathrm{card} I < \aleph_0, \\ \frac{1}{2} + \sum\limits_{H_i \subseteq K} \frac{1}{2^{i+1}} & \text{otherwise,} \end{cases}$$

and consider the conditional event $F|K$ with
$F = (E_1 \wedge (H_1 \vee H_2 \vee H_3)^c)$ and $K = (E_1 \wedge (H_1 \vee H_2)^c) \vee (E_2 \wedge (H_1)^c)$.

Since $\sigma(B|\cdot)$ is $\mathcal{A}_{\mathcal{L}}$-continuous for every $B \in \mathcal{A}$ and $\oint \sigma(F \wedge K|H_i)\pi(\mathrm{d}H_i) = \frac{3}{16} > 0$ it holds

$$\underline{P}^{\mathbf{fd}}(F|K) = \overline{P}^{\mathbf{fd}}(F|K) = \frac{\oint \sigma(F \wedge K|H_i)\pi(\mathrm{d}H_i)}{\oint \sigma(K|H_i)\pi(\mathrm{d}H_i)} = \frac{9}{22} \approx 40.9\%.$$

2-monotone prior capacity

$\Rightarrow$ A 2-monotone prior capacity φ (with dual $\overline{\varphi}$) on $\mathcal{A}_{\mathcal{L}}$ induces core:

- $\mathcal{P}_{\varphi} = \{\tilde{\pi} \,:\, \text{f.a. probability on } \mathcal{A}_{\mathcal{L}} \text{ s.t. } \varphi \leq \tilde{\pi} \leq \overline{\varphi}\}$
- $\mathcal{P}_{\varphi}^{\mathbf{fd}} = \{\tilde{P}^{\mathbf{fd}} \,:\, \text{fully } \mathcal{L}\text{-disintegrable f.c.p. on } \mathcal{A} \text{ extending } \{\tilde{\pi}, \sigma\},\ \tilde{\pi} \in \mathcal{P}_{\varphi}\}$

Theorem

The set $\mathcal{P}_{\varphi}^{\mathbf{fd}}$ of fully $\mathcal{L}$-disintegrable extensions of $\{\tilde{\pi}, \sigma\}$ on $\mathcal{A} \times \mathcal{A}^0$, for $\tilde{\pi} \in \mathcal{P}_{\varphi}$, is a non-empty compact subset of $[0,1]^{\mathcal{A} \times \mathcal{A}^0}$ endowed with the product topology of pointwise convergence and has envelopes $\underline{P}_{\varphi}^{\mathbf{fd}} = \min \mathcal{P}_{\varphi}^{\mathbf{fd}}$ and $\overline{P}_{\varphi}^{\mathbf{fd}} = \max \mathcal{P}_{\varphi}^{\mathbf{fd}}$.

Envelopes of fully $\mathcal{L}$-disintegrable extensions of $\{\varphi, \sigma\}$

$$L(F,K) = \min\left\{ \oint \sigma(F \wedge K)\mathrm{d}\tilde{\pi} \;:\; \oint \sigma(F^c \wedge K)\mathrm{d}\tilde{\pi} = \oint \sigma(F^c \wedge K)\mathrm{d}\overline{\varphi}, \tilde{\pi} \in \mathcal{P}_\varphi \right\}$$

$$U(F^c,K) = \max\left\{ \oint \sigma(F^c \wedge K)\mathrm{d}\tilde{\pi} \;:\; \oint \sigma(F \wedge K)\mathrm{d}\tilde{\pi} = \oint \sigma(F \wedge K)\mathrm{d}\varphi, \tilde{\pi} \in \mathcal{P}_\varphi \right\}$$

Theorem

Let φ be a 2-monotone capacity. The lower envelope $\underline{P}^{\mathbf{fd}}(\cdot|\cdot)$ of the fully $\mathcal{L}$-disintegrable extensions of $\{\tilde{\pi}, \sigma\}$ on $\mathcal{A} \times \mathcal{A}^0$, for $\tilde{\pi} \in \mathcal{P}_\varphi$, is such that for every $F|K \in \mathcal{A} \times \mathcal{A}^0$ with $F \wedge K \neq K$:

(i) if $K \in \mathcal{A}^0_{\mathcal{L}}$

$$\underline{P}^{\mathbf{fd}}(F|K) = \begin{cases} \min\left\{ \frac{\oint \sigma(F\wedge K|\cdot)\mathrm{d}\varphi}{\oint \sigma(F\wedge K|\cdot)\mathrm{d}\varphi + U(F^c,K)}, \frac{L(F,K)}{L(F,K)+\oint \sigma(F^c\wedge K|\cdot)\mathrm{d}\overline{\varphi}} \right\} & \text{if } \varphi(K) > 0, \\ \inf_{H_i \subseteq K} \sigma(K|H_i) & \text{otherwise,} \end{cases}$$

(ii) if $K \in \mathcal{A}^0 \setminus \mathcal{A}^0_{\mathcal{L}}$

$$\underline{P}^{\mathbf{fd}}(F|K) = \begin{cases} \min\left\{ \frac{\underline{P}^{\mathbf{fd}}(F\wedge K|A)}{\underline{P}^{\mathbf{fd}}(F\wedge K|A)+U(F^c,K;A)}, \frac{L(F,K;A)}{L(F,K;A)+\overline{P}^{\mathbf{fd}}(F^c\wedge K|A)} \right\} & \text{if } \exists A \in \mathcal{A}^0_{\mathcal{L}} \text{ s.t. } K \subseteq A \text{ and } \underline{P}^{\mathbf{fd}}(K|A) > 0, \\ 0 & \text{otherwise.} \end{cases}$$

Envelopes of fully $\mathcal{L}$-disintegrable extensions of $\{\varphi, \sigma\}$

$$
\begin{aligned}
L(F, K; A) &= \min\{\tilde{P}(F \wedge K|A) \,:\, \tilde{P}(F^c \wedge K|A) = \overline{P}^{\mathrm{d}}(F^c \wedge K|A), \tilde{P} \in \mathcal{P}^{\mathrm{d}}_{\varphi}\} \\
U(F^c, K; A) &= \max\{\tilde{P}(F^c \wedge K|A) \,:\, \tilde{P}(F \wedge K|A) = \underline{P}^{\mathrm{d}}(F \wedge K|A), \tilde{P} \in \mathcal{P}^{\mathrm{d}}_{\varphi}\}
\end{aligned}
$$

Theorem

Let φ be a 2-monotone capacity. The lower envelope $\underline{P}^{\mathbf{fd}}(\cdot|\cdot)$ of the fully $\mathcal{L}$-disintegrable extensions of $\{\tilde{\pi}, \sigma\}$ on $\mathcal{A} \times \mathcal{A}^0$, for $\tilde{\pi} \in \mathcal{P}_{\varphi}$, is such that for every $F|K \in \mathcal{A} \times \mathcal{A}^0$ with $F \wedge K \neq K$:

(i) *if* $K \in \mathcal{A}^0_{\mathcal{L}}$

$$
\underline{P}^{\mathbf{fd}}(F|K) = \begin{cases} \min\left\{\dfrac{\oint \sigma(F\wedge K|\cdot)\mathrm{d}\varphi}{\oint \sigma(F\wedge K|\cdot)\mathrm{d}\varphi + U(F^c, K)}, \dfrac{L(F,K)}{L(F,K) + \oint \sigma(F^c \wedge K|\cdot)\mathrm{d}\overline{\varphi}}\right\} & \text{if } \varphi(K) > 0, \\[2ex] \inf\limits_{H_i \subseteq K} \sigma(K|H_i) & \text{otherwise,} \end{cases}
$$

(ii) *if* $K \in \mathcal{A}^0 \setminus \mathcal{A}^0_{\mathcal{L}}$

$$
\underline{P}^{\mathbf{fd}}(F|K) = \begin{cases} \min\left\{\dfrac{\underline{P}^{\mathbf{fd}}(F\wedge K|A)}{\underline{P}^{\mathbf{fd}}(F\wedge K|A) + U(F^c, K;A)}, \dfrac{L(F,K;A)}{L(F,K;A) + \overline{P}^{\mathbf{fd}}(F^c \wedge K|A)}\right\} & \text{if } \exists A \in \mathcal{A}^0_{\mathcal{L}} \text{ s.t. } K \subseteq A \text{ and } \underline{P}^{\mathbf{fd}}(K|A) > 0, \\[2ex] 0 & \text{otherwise.} \end{cases}
$$

Lower posterior probability

Corollary

For a 2-monotone prior capacity φ, the lower envelope $\underline{P}_\varphi^{\mathbf{fd}}(\cdot|\cdot)$ of the fully $\mathcal{L}$-disintegrable extensions of $\{\tilde{\pi}, \sigma\}$, for $\tilde{\pi} \in \mathcal{P}_\varphi$, for every $H|E_j \in \mathcal{A}_\mathcal{L} \times \mathcal{E}$ such that $\oint \sigma(E_j|\cdot)\mathrm{d}\varphi > 0$ is

$$\underline{P}_\varphi^{\mathbf{fd}}(H|E_j) = \frac{\oint \sigma(H \wedge E_j|\cdot)\mathrm{d}\varphi}{\oint \sigma(H \wedge E_j|\cdot)\mathrm{d}\varphi + \oint \sigma(H^c \wedge E_j|\cdot)\mathrm{d}\overline{\varphi}}.$$

$\Longrightarrow$ In general, φ could be a non-2-monotone lower probability of a class of finitely additive priors $\mathcal{P}_\varphi$

Theorem

For a lower probability φ, the lower envelope $\underline{P}_\varphi^{\mathbf{fd}}(\cdot|\cdot)$ of the fully $\mathcal{L}$-disintegrable extensions of $\{\tilde{\pi}, \sigma\}$, for $\tilde{\pi} \in \mathcal{P}_\varphi$, for every $H|E_j \in \mathcal{A}_\mathcal{L} \times \mathcal{E}$ such that $\oint \sigma(E_j|\cdot)\mathrm{d}\varphi > 0$ is

$$\underline{P}_\varphi^{\mathbf{fd}}(H|E_j) \geq \frac{\oint \sigma(H \wedge E_j|\cdot)\mathrm{d}\varphi}{\oint \sigma(H \wedge E_j|\cdot)\mathrm{d}\varphi + \oint \sigma(H^c \wedge E_j|\cdot)\mathrm{d}\overline{\varphi}}.$$

Example: Nuisance parameter elimination

PROBLEM: Given a statistical model $\lambda(E|\Theta = \theta, \Gamma = \gamma)$ where Θ is the interest parameter, we want to eliminate the nuisance parameter Γ.

- **Integrated likelihood:** for a conditional prior π

$$\lambda(E|\Theta = \theta) = \int \lambda(E|\Theta = \theta, \Gamma = \gamma)\pi(\mathrm{d}(\Gamma = \gamma)|\Theta = \theta)$$

- **Profile likelihood:**

$$\hat{\lambda}(E|\Theta = \theta) = \sup_{\gamma} \lambda(E|\Theta = \theta, \Gamma = \gamma)$$

Example: Nuisance parameter elimination (1)

Consider:

- (Θ, Γ), random vector ranging in $\mathbf{\Theta} \times \mathbf{\Gamma} = \mathbb{N} \times (0,1)$
- $X = (X_1, \ldots, X_k)$, random vector ranging in $\mathbf{X} = \mathbb{N}_0^k$
- $X_i | (\Theta = \theta, \Gamma = \gamma) \sim \mathrm{Bin}(\theta, \gamma)$, for $i = 1, \ldots, k$, and independent conditionally to $(\Theta = \theta, \Gamma = \gamma)$
- $\mathcal{L} = \{H_{(\theta,\gamma)} = (\Theta = \theta, \Gamma = \gamma) : (\theta, \gamma) \in \mathbf{\Theta} \times \mathbf{\Gamma}\}$
- $\mathcal{E} = \{E_x = (X = x) : x \in \mathbf{X}\}$

$$\lambda(X = x | \Theta = \theta, \Gamma = \gamma) = \begin{cases} \left[\prod_{i=1}^{k} \binom{\theta}{x_i}\right] \gamma^{||x||_1} (1-\gamma)^{\theta k - ||x||_1}, & \text{if } \theta \geq ||x||_\infty, \\ 0 & \text{otherwise,} \end{cases}$$

Example: Nuisance parameter elimination (2)

Take:

- $\mathcal{A}_\mathcal{L} = \langle \mathcal{L} \rangle^*$ and $\mathcal{A}_\mathcal{E} = \langle \mathcal{E} \rangle$
- φ, vacuous 2-monotone capacity ($\varphi(\Omega) = 1$ and 0 otherwise) on $\mathcal{A}_\mathcal{L}$ giving rise to the class

$$\mathcal{P}^{\mathbf{p}} = \{\tilde{\pi} \ : \ \text{conditional prior on } \mathcal{A}_\mathcal{L} \times \mathcal{A}_\mathcal{L}^0\}$$

whose upper envelope $\overline{\pi}^{\mathbf{p}} = \max \mathcal{P}^{\mathbf{p}}$ is defined for $F|K \in \mathcal{A}_\mathcal{L} \times \mathcal{A}_\mathcal{L}^0$ as

$$\overline{\pi}^{\mathbf{p}}(F|K) = \begin{cases} 1 & \text{if } K \subseteq F, \\ 0 & \text{otherwise,} \end{cases}$$

GOAL

Make inference on conditional events $(X = x|(\Theta, \Gamma) \in \{\theta\} \times \boldsymbol{\Gamma})$

$\Longrightarrow$ The profile likelihood is a supremum of integrated likelihoods

$$\begin{aligned}
\hat{\lambda}(X = x|\Theta = \theta) &= \overline{P}_\varphi^{\mathbf{fd}}(X = x|(\Theta, \Gamma) \in \{\theta\} \times \boldsymbol{\Gamma}) \\
&= \oint \lambda(X = x|\Theta = \theta, \Gamma = \gamma)\overline{\pi}^{\mathbf{p}}(\mathrm{d}(\Gamma = \gamma)|\Theta = \theta) \\
&= \sup_\gamma \lambda(X = x|\Theta = \theta, \Gamma = \gamma)
\end{aligned}$$

Extensions

To study the lower and upper envelope of conglomerable extensions (instead the disintegrable one).

Open problems

- Determine which family of extensions of two marginal distributions leads to 2-monotone capacities
- The relationship between semi-copulas and the extensions of two lower probabilities on two spaces that lead to 2-monotone capacities
- Inference for multiple priors models with superadditive lower prior probabilities