Aggregation of the distortion models induced by the KL divergence and Euclidean distance

Ignacio Montes
University of Oviedo

SMPS 2024, Salzburg

Overview

Overview

Imprecise probability models

Distortion models

Aggregation of Euclidean models

Aggregation of Kullback-Leibler models

Conclusions

Coherent lower prevision

Lower prevision

Gambles: bounded real-valued functions f
Notation: set of gambles $\mathcal{L}(\mathcal{X})$
Lower prevision:

$$\begin{aligned} \underline{P}: \quad \mathcal{L}(\mathcal{X}) &\longrightarrow \mathbb{R} \\ f &\mapsto \underline{P}(f) \end{aligned}$$

Credal set

Credal set: a closed and convex set given by

$$\mathcal{M}(\underline{P}) = \{P \mid E_P(f) \geq \underline{P}(f) \ \forall f\}$$

Coherent lower previsions

Coherence: $\underline{P}$ is coherent if

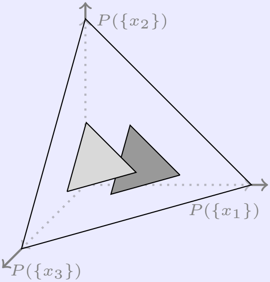
$$\underline{P}(f) = \min\{E_P(f) \mid P \in \mathcal{M}(\underline{P})\} \qquad \forall f \in \mathcal{L}(\mathcal{X})$$

Coherent lower previsions VS credal sets

$$\begin{tikzcd}[column sep=huge] \boxed{\begin{array}{c}\text{Coherent lower}\\ \text{prevision } \underline{P}\end{array}} \arrow[r, bend right, "\mathcal{M} = \{P \mid E_P(f) \geq \underline{P}(f)\ \forall f\}"'] & \boxed{\begin{array}{c}\text{Credal}\\ \text{set } \mathcal{M}\end{array}} \arrow[l, bend right, "\underline{P}(f) = \min\{E_P(f) \mid P \in \mathcal{M}\}"'] \end{tikzcd}$$

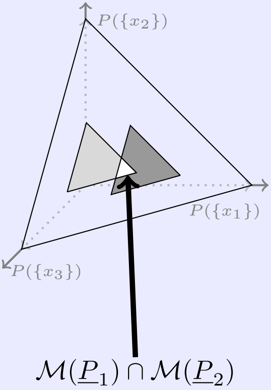
Aggregation of coherent lower previsions

Conjunction



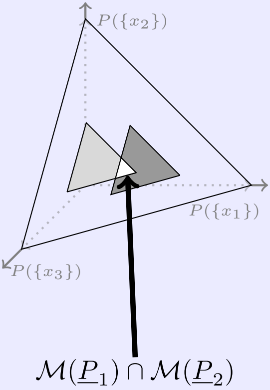
Aggregation of coherent lower previsions

Conjunction

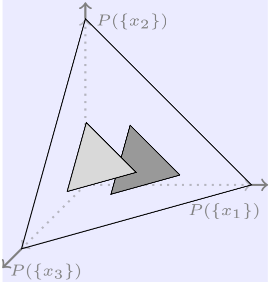


Aggregation of coherent lower previsions

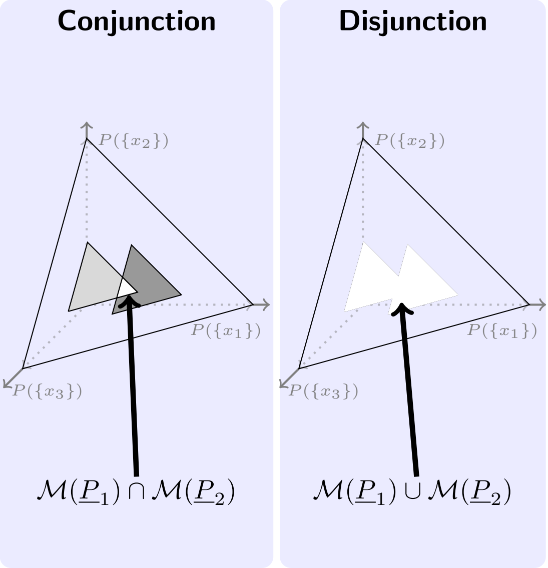
Conjunction



Disjunction

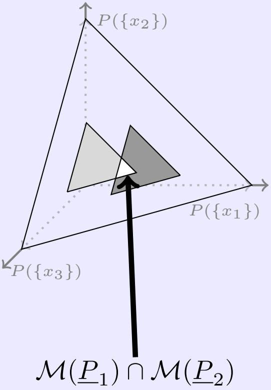


Aggregation of coherent lower previsions

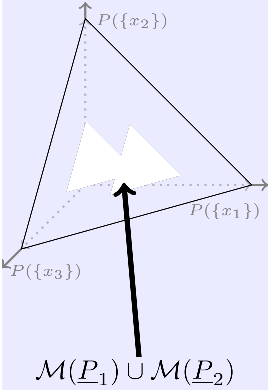


Aggregation of coherent lower previsions

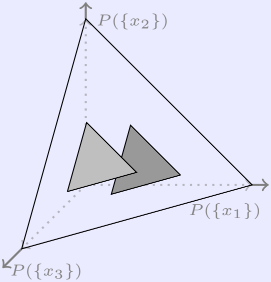
Conjunction



Disjunction

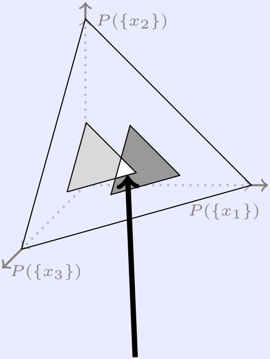


Mixture



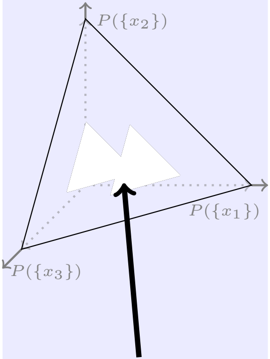
Aggregation of coherent lower previsions

Conjunction



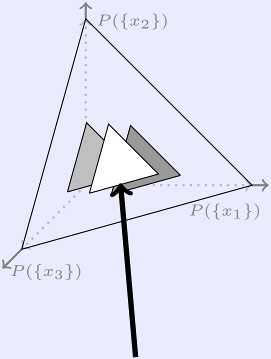
$$\mathcal{M}(\underline{P}_1) \cap \mathcal{M}(\underline{P}_2)$$

Disjunction



$$\mathcal{M}(\underline{P}_1) \cup \mathcal{M}(\underline{P}_2)$$

Mixture



$$\varepsilon\mathcal{M}(\underline{P}_1) + (1-\varepsilon)\mathcal{M}(\underline{P}_2)$$

$(\varepsilon \in (0,1))$

Overview

Distortion models (Montes et al., 2020)

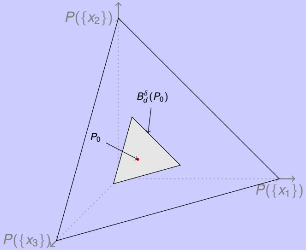
Ingredients

P_0: probability measure

δ: distortion parameter

d: distorting function

Example



The model

$$B_d^\delta(P_0) = \{P \mid d(P, P_0) \leq \delta\}$$

$$\underline{P}_d(f) = \inf\left\{E_P(f) \mid P \in B_d^\delta(P_0)\right\}$$

$$\mathcal{M}(\underline{P}_d) = \{P \mid E_P(f) \geq \underline{P}_d(f)\ \forall f\}$$

d continuous and quasi-convex

Assumptions

$\mathcal{X}$ finite

δ "small"

Particular models

Pari Mutuel Model

$$\underline{P}_{PMM}(A) = \max\{0, (1+\delta)P_0(A) - \delta\}$$

Pelessoni et al., 2010
Montes et al., 2019
Walley, 1991

Linear vacuous model

$$\underline{P}_{LV}(A) = (1-\delta)P_0(A) \qquad A \neq \mathcal{X}$$

Huber, 1981
Walley, 1991

Constant odds ratio

$$\underline{P}_{COR}(A) = \frac{(1-\delta)P_0(A)}{1-\delta P_0(A)}$$

Benavoli and Zaffalon, 2013
Walley, 1991

Total Variation model

$$\underline{P}_{TV}(A) = \max\{0, P_0(A) - \delta\} \qquad A \neq \mathcal{X}$$

Montes et al., 2020
Herron et al., 1997

Vertical barrier models

$$\underline{P}_{VB}(A) = \max\{bP_0(A) + a, 0\} \qquad \begin{array}{l} a \leq 0, b > 0 \\ a + b \in (0,1) \end{array}$$

Pelessoni et al., 2021
Corsato et al., 2019

Increasing transformation

$$\underline{P}(A) = g(P_0(A))$$

$g : [0,1] \to [0,1]$ increasing
$g(0) = 0, g(1) = 1$

Bronevich, 2005

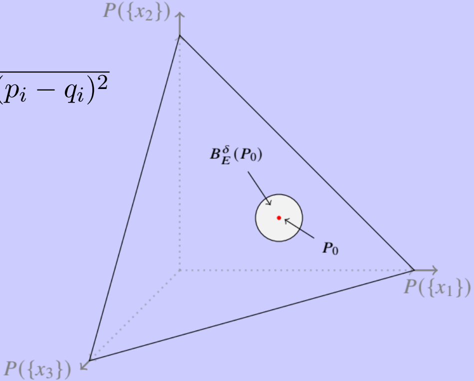
Euclidean model

Euclidean model

$$d_E(P,Q) = \sqrt{\sum_{i=1}^{n}(p_i - q_i)^2}$$

$P_0 \longrightarrow B_E^{\delta}(P_0)$

$\delta \longrightarrow B_E^{\delta}(P_0)$



Coherent lower prevision

$f \equiv (a_1, \ldots, a_n)$ $\qquad$ $\underline{P}_E(f) = E_{P_0}(f) - \delta\sqrt{n}S_f$

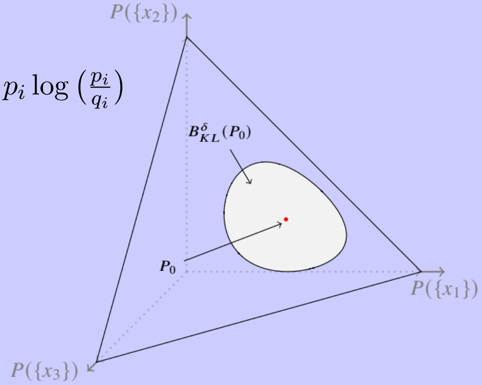
Neighbourhood models induced by the Euclidean distance and the KL divergence. Montes. Submitted.

KL-model

KL-model

$$D_{KL}(P,Q) = \sum_{i=1}^{n} p_i \log\left(\frac{p_i}{q_i}\right)$$

$$\begin{matrix} P_0 \searrow & \\ & B^{\delta}_{KL}(P_0) \\ \delta \nearrow & \end{matrix}$$



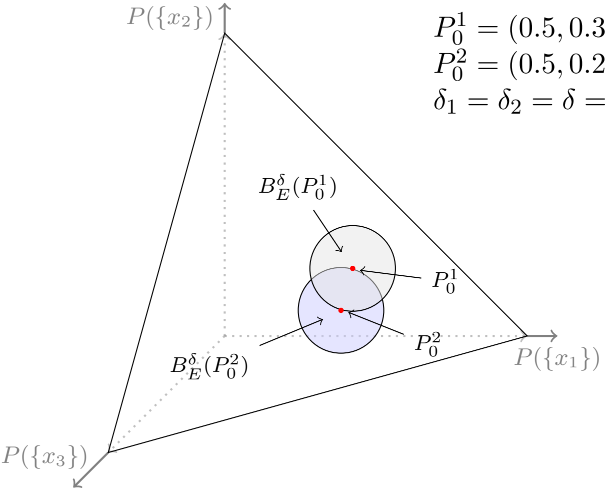
Coherent lower prevision

$$\underline{P}_{KL}(f) = \min\{E_P(f) \mid D_{KL}(P,P_0) = \delta\}$$

Neighbourhood models induced by the Euclidean distance and the KL divergence. Montes. Submitted.

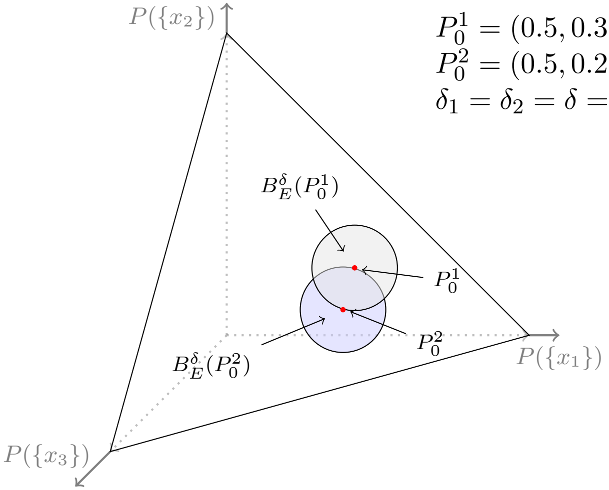
Overview

Conjunction of E-models



$$P_0^1 = (0.5, 0.3, 0.2)$$
$$P_0^2 = (0.5, 0.2, 0.3)$$
$$\delta_1 = \delta_2 = \delta = \sqrt{2}/10$$

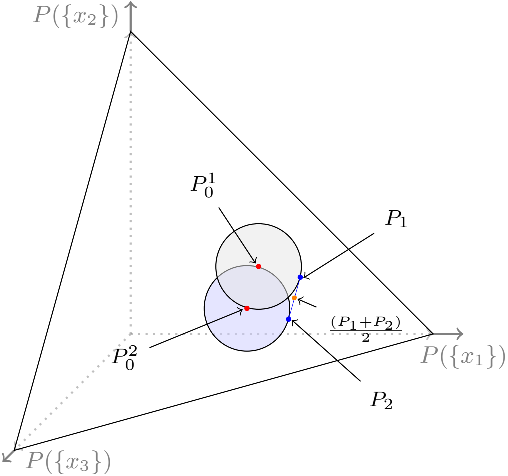
Conjunction of E-models



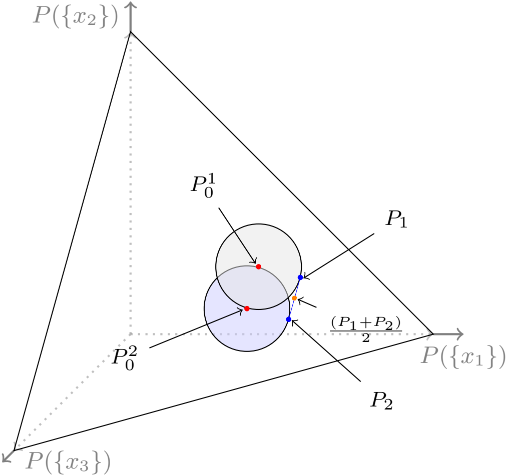
$P_0^1 = (0.5, 0.3, 0.2)$
$P_0^2 = (0.5, 0.2, 0.3)$
$\delta_1 = \delta_2 = \delta = \sqrt{2}/10$

Conclusion: the E-model is not preserved under conjunction

Disjunction of E-models

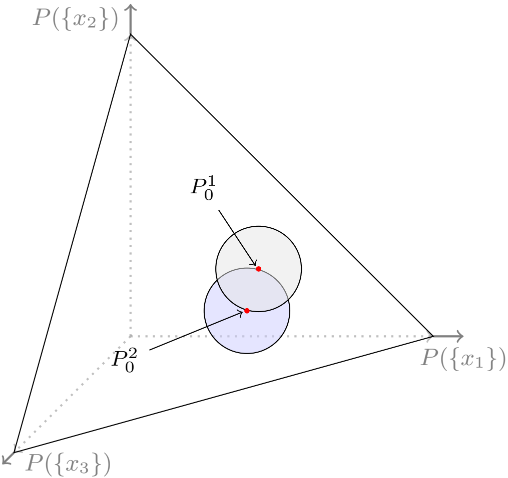


Disjunction of E-models

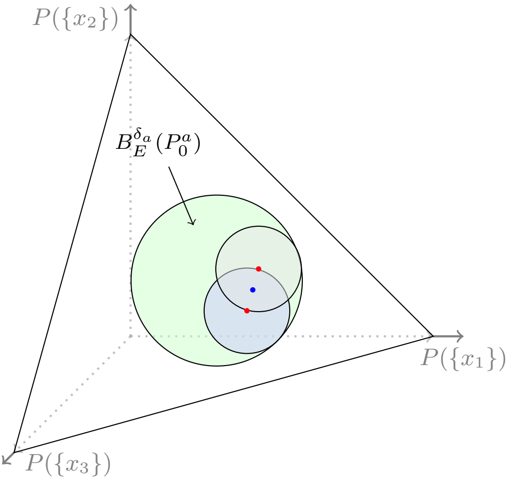


Conclusion: the E-model is not preserved under disjunction

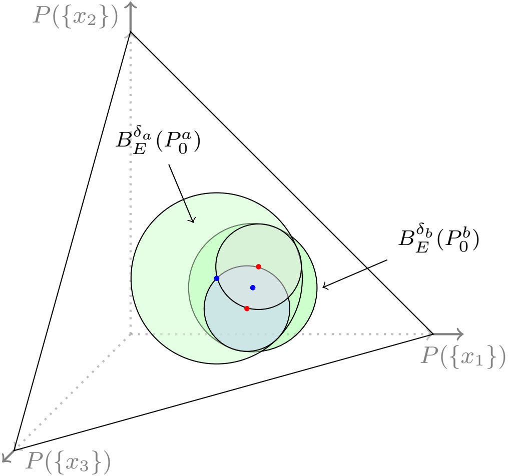
Disjunction of E-models: unique undominated OA?



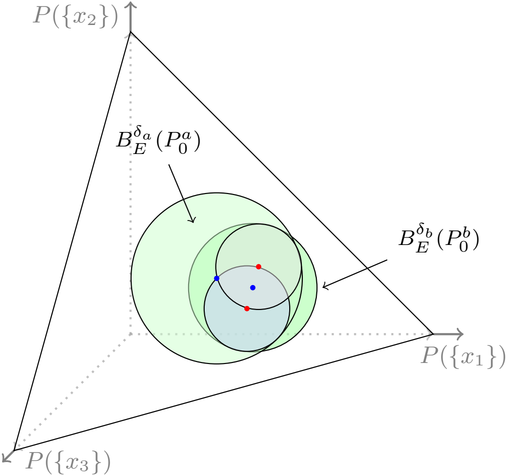
Disjunction of E-models: unique undominated OA?



Disjunction of E-models: unique undominated OA?



Disjunction of E-models: unique undominated OA?



Conclusion: there is not a unique undominated OA

Mixture of E-models

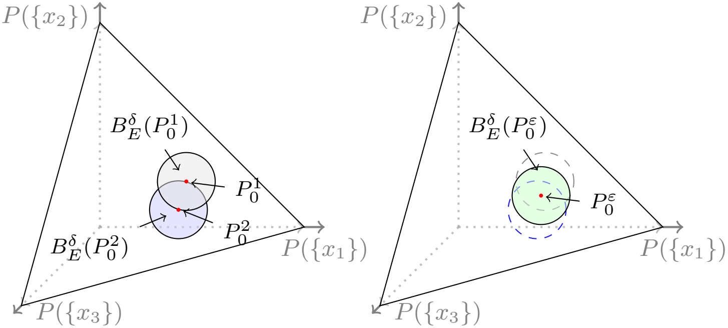
Proposition: $\varepsilon B_E^{\delta_1}(P_0^1) + (1-\varepsilon) B_E^{\delta_2}(P_0^2) = B_E^{\delta_\varepsilon}(P_0^\varepsilon)$, where:

$$\delta_\varepsilon = \varepsilon\delta_1 + (1-\varepsilon)\delta_2, \qquad P_0^\varepsilon = \varepsilon P_0^1 + (1-\varepsilon) P_0^2.$$

Mixture of E-models

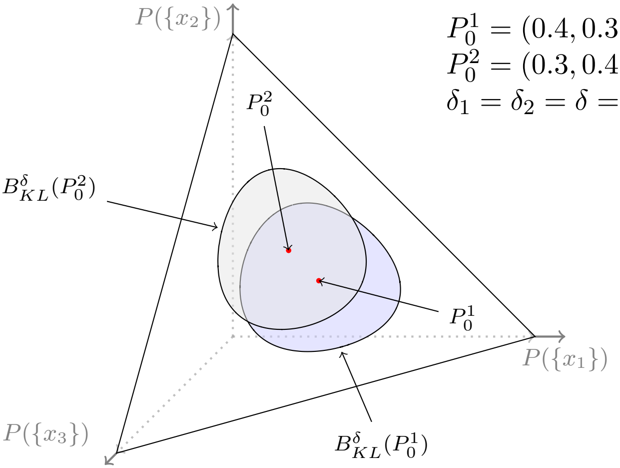
Proposition: $\varepsilon B_E^{\delta_1}(P_0^1) + (1-\varepsilon) B_E^{\delta_2}(P_0^2) = B_E^{\delta_\varepsilon}(P_0^\varepsilon)$, where:

$$\delta_\varepsilon = \varepsilon\delta_1 + (1-\varepsilon)\delta_2, \qquad P_0^\varepsilon = \varepsilon P_0^1 + (1-\varepsilon) P_0^2.$$



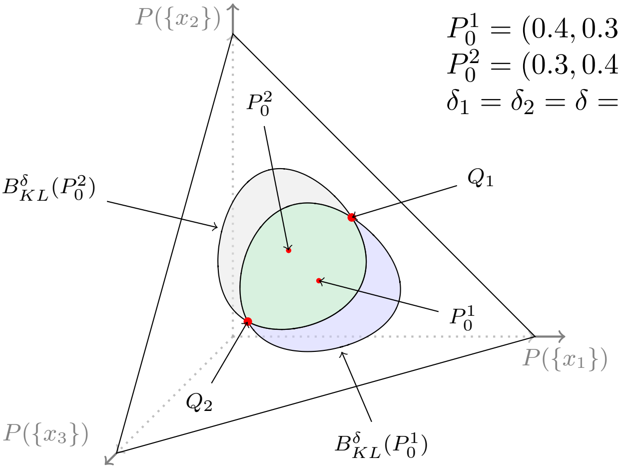
Overview

Conjunction of KL-models



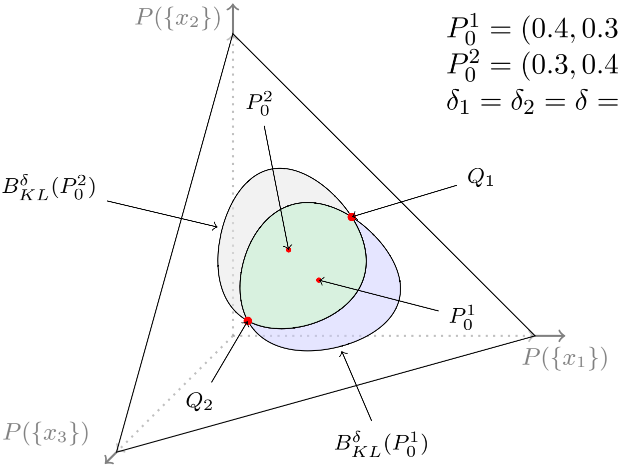
$$P_0^1 = (0.4, 0.3, 0.3)$$
$$P_0^2 = (0.3, 0.4, 0.3)$$
$$\delta_1 = \delta_2 = \delta = 0.1$$

Conjunction of KL-models



$$P_0^1 = (0.4, 0.3, 0.3)$$
$$P_0^2 = (0.3, 0.4, 0.3)$$
$$\delta_1 = \delta_2 = \delta = 0.1$$

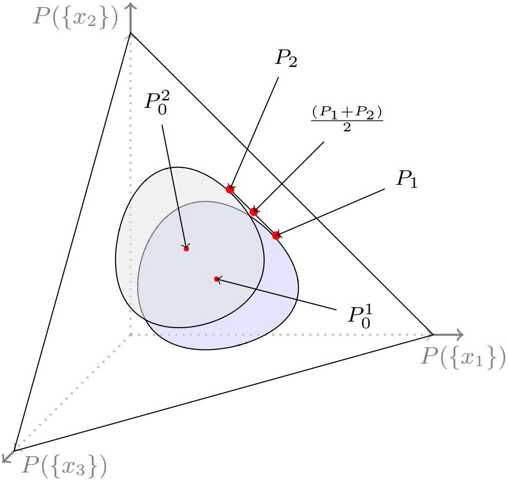
Conjunction of KL-models



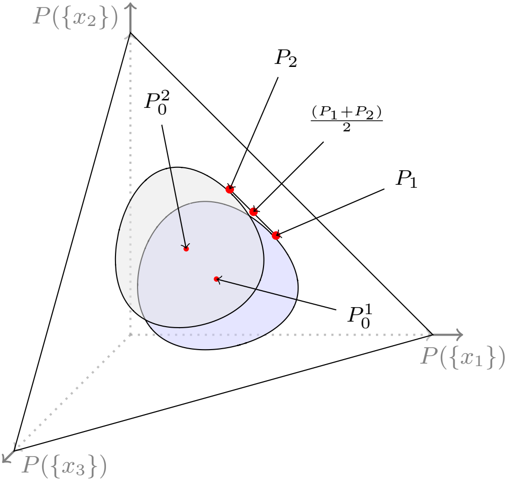
$$P_0^1 = (0.4, 0.3, 0.3)$$
$$P_0^2 = (0.3, 0.4, 0.3)$$
$$\delta_1 = \delta_2 = \delta = 0.1$$

Conclusion: the KL-model is not preserved under conjunction

Disjunction of KL-models

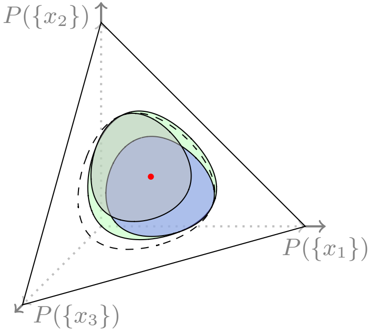


Disjunction of KL-models



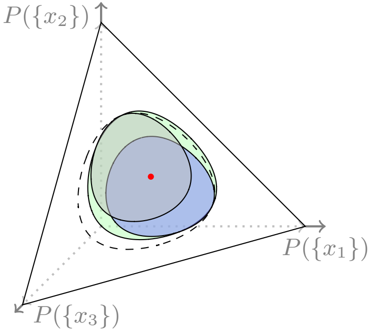
Conclusion: the KL-model is not preserved under disjunction

Disjunction of KL-models: unique undominated OA?

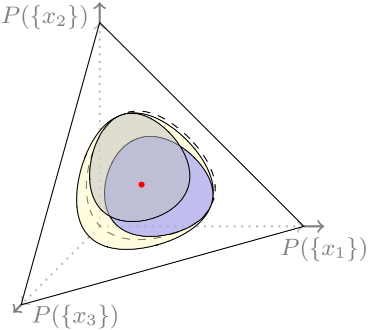


$$P_0 = (0.3552, 0.3552, 0.2896)$$
$$\delta = 0.15381$$

Disjunction of KL-models: unique undominated OA?

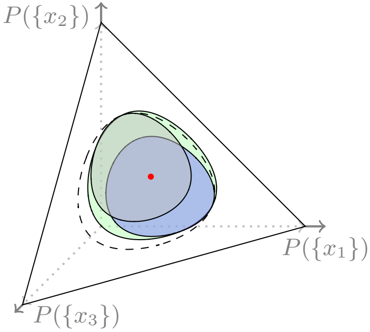


$$P_0 = (0.3552, 0.3552, 0.2896)$$
$$\delta = 0.15381$$

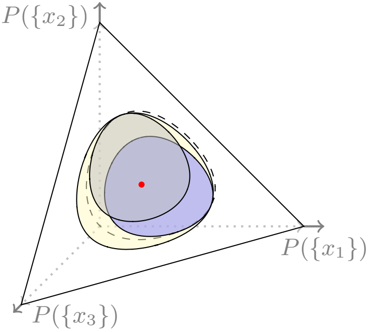


$$P_0 = (1/3, 1/3, 1/3)$$
$$\delta = 0.173594$$

Disjunction of KL-models: unique undominated OA?



$P_0 = (0.3552, 0.3552, 0.2896)$
$\delta = 0.15381$

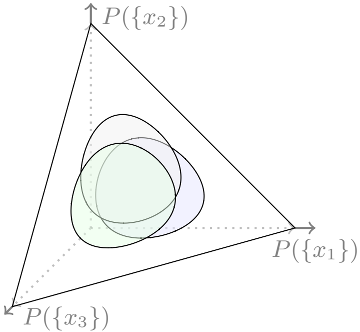


$P_0 = (1/3, 1/3, 1/3)$
$\delta = 0.173594$

Conclusion: there is not a unique undominated OA

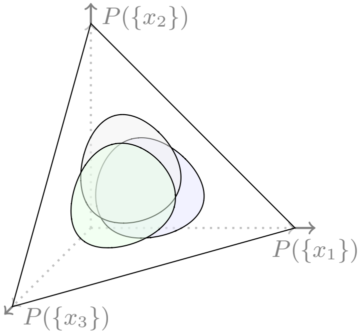
Mixture of KL-models

$$P_0^1 = (0.4, 0.3, 0.3), \;\; P_0^2 = (0.3, 0.4, 0.3), \;\; P_0^3 = (0.3, 0.3, 0.4), \;\; \delta = 0.1$$



Mixture of KL-models

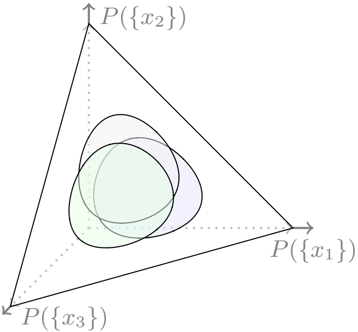
$$P_0^1 = (0.4, 0.3, 0.3), \;\; P_0^2 = (0.3, 0.4, 0.3), \;\; P_0^3 = (0.3, 0.3, 0.4), \;\; \delta = 0.1$$

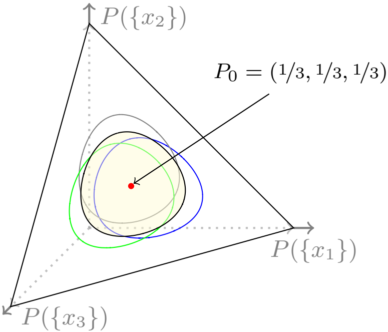


$\underline{P}(\{x_i\}) = 0.112736 \Rightarrow \delta = 0.0986 \Rightarrow \underline{P}(\{x_i, x_j\}) = 0.4499$
$\underline{P}(\{x_i, x_j\}) = 0.44956$

Mixture of KL-models

$$P_0^1 = (0.4, 0.3, 0.3),\;\; P_0^2 = (0.3, 0.4, 0.3),\;\; P_0^3 = (0.3, 0.3, 0.4),\;\; \delta = 0.1$$





$\underline{P}(\{x_i\}) = 0.112736 \Rightarrow \delta = 0.0986 \Rightarrow \underline{P}(\{x_i, x_j\}) = 0.4499$
$\underline{P}(\{x_i, x_j\}) = 0.44956$

Conclusion: the KL-model is not preserved under mixtures

Overview

Aggregation: summary

Model	Conjunction	Disjunction	Unique OA	Mixture
PMM	✓	✗	✓	✓
LV	✓	✗	✓	✓
COR	✗	✗	✗	✗
TV	✗	✗	✗	✓
VBM	✗	✗	✗	✓
Euclidean	✗	✗	✗	✓
KL	✗	✗	✗	✗

Processing distortion models: a comparative study. Destercke, Montes, Miranda. IJAR 2022.

Evaluating uncertainty with Vertical Barrier Models. Miranda, Vicig, Pelessoni. IJAR 2024.

Further research

What about combination?

- Marginalisation?
- Joint model?

More about distortion models?

- Wasserstein distance?
- Levy distance?

Applications?

- Statistical inference?
- Credal networks?

Funding

This contribution is part of grant PID2022-140585NB-I00 funded by MICIU/AEI/10.13039/501100011033 and “FEDER/UE”.