

Neighbourhood models induced by the Euclidean distance and the Kullback-Leibler divergence

Ignacio Montes
University of Oviedo

Asturias: Natural Paradise



IP group in Oviedo

IP researchers



E.Miranda



I.Montes



J.J.Salamanca



R.Pérez

Young people



J.Álvarez



D.Nieto



M.Porrón



A.Presa

Close reserchers



A.Bouchet



S.Díaz



I.Mariñas



S.Montes

Distortion models



S.Destercke



E.Miranda



P.R.Alonso

- Unifying neighbourhood and distortion models: Part I. Montes, Miranda, Destercke. IJGS 2020.
- Unifying neighbourhood and distortion models: Part II. Montes, Miranda, Destercke. IJGS 2020.
- Processing distortion models: a comparative study. Destercke, Montes, Miranda. IJAR 2022.
- Distortion models for estimating human error probabilities. Alonso, Montes, Miranda. SS 2023.

Distortion models

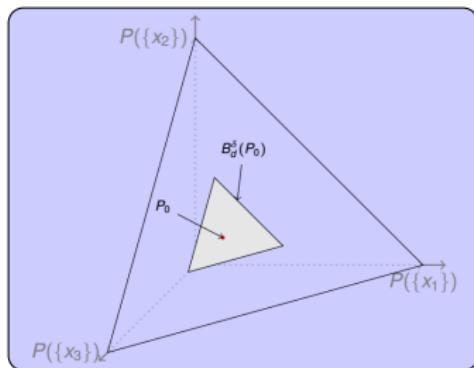
Ingredients

P_0 : probability measure

δ : distortion parameter

d : distorting function

Example



The model

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

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

d continuous and convex

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

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

$$\underline{P}_{LV}(A) = (1 - \delta)P_0(A) \quad A \neq \emptyset$$

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

$$\underline{P}_{TV}(A) = \max\{0, P_0(A) - \delta\} \quad A \neq \emptyset$$

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

Vertical barrier models

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

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

Increasing transformation

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

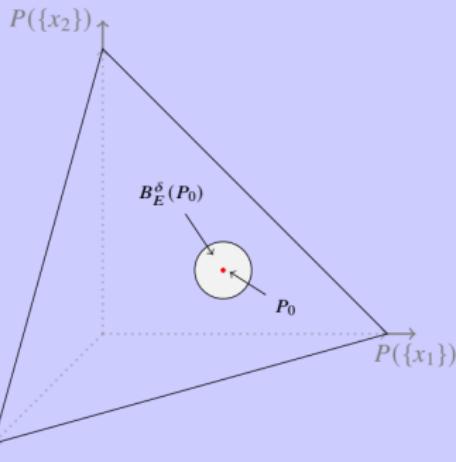
$g : [0, 1] \rightarrow [0, 1]$ increasing
 $g(0) = 0, g(1) = 1$

Bronevich, 2005

Aim of the paper

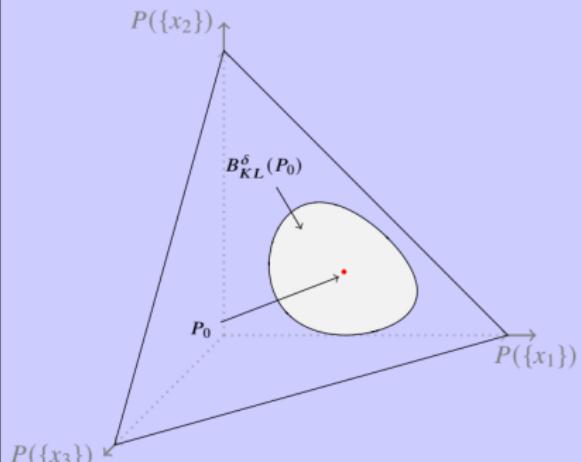
Euclidean model

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



Kullback-Leibler model

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



Euclidean model

Expression for \underline{P}_E :

$$\underline{P}_E(f) = ??$$

$$\underline{P}_E(A) = ??$$

Extreme points:

$$\text{ext}\left(B_E^\delta(P_0)\right) = ??$$

Properties:

2-monotonicity ??

Complete-monotonicity ??

Conditioning ??

KL-model

Expression for \underline{P}_{KL} :

$$\underline{P}_{KL}(f) = ??$$

$$\underline{P}_{KL}(A) = ??$$

Extreme points:

$$\text{ext}\left(B_{KL}^\delta(P_0)\right) = ??$$

Properties:

2-monotonicity ??

Complete-monotonicity ??

Conditioning ??

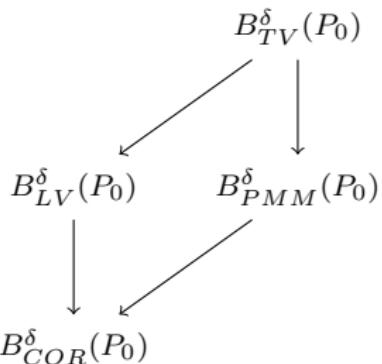
Comparison: properties

	2-monot.	∞ -monot.	Prob.interval	Conditioning
Euclidean-model...				
... on gambles	??	??	??	??
... on events	??	??	??	??
KL-model...				
... on gambles	??	??	??	??
... on events	??	??	??	??
PMM	✓	✗	✓	✓
LV	✓	✓	✓	✓
COR...				
... on gambles	✗	✗	✗	✓
... on events	✓	✓	✗	✓
TV	✓	✗	✓	✓
Vertical Barrier models	✓	✗	✗	✓

Comparison: amount of imprecision

$$B_{KL}^\delta(P_0) \quad ??$$

$$B_E^\delta(P_0) \quad ??$$



WHAT CAN I DO?

I HAVE A PROBABILISTIC MEASUREMENT MODELING AN EVENT AS A SET BUT I WANT TO ROBUSTIFY MY MODEL...

WHY DO NOT YOU USE DISTORTION MODELS?



DISTORTION MODELS

1. INGREDIENTS

P_0 : probability measure
 δ : distortion parameter
 d : distortion function

Assumption:
 $X = \{x_1, \dots, x_n\}$
 $P_0(x_i) > 0 \forall i$
 δ : "real enough"

2. THE MODEL

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

$$\underline{d}_d(P) = \inf\{P(P) \mid P \in D_d^{\delta}(P_0)\}$$

$$M_d^{\delta}(P_0) = \{P \mid P_d(P) \geq \underline{d}_d(P) \ \forall t\}$$

3. PARTICULAR CASES

3.1 Known δ -models:
 $\text{Err}_{\delta}(A) = \max\{0, (1-\delta)P_0(A) - \delta\}$
 $\text{Err}_{\delta}(A) = (1-\delta)P_0(A)$
 $\text{Err}_{\delta}(A) = (1-\delta)P_0(A)$

3.2 Distance-based models:
 $\text{Err}_{\delta}(A) = \max\{0, P_0(A) - \delta\}$

3.3 Variant barrier models:
 $\text{Err}_{\delta}(A) = \max\{bP_0(A), A + \alpha\}$

3.4 Increasing transformations:
 $\underline{d}(A) = g(P_0(A))$

δ : $0 < \delta < 1$ meaning $A + \alpha < \underline{d}(A) < 1$

WHY DO NOT YOU CONSIDER THE EUCLIDEAN DISTANCE?

BUT IF YOU USE A DISTANCE...
 YOU ARE RIGHT...



EUCLIDEAN MODEL

1. Collision distance

$$d_E(P, P_0) = \sqrt{\sum_{i=1}^n (p_i - p_{0i})^2}$$

d_E continuous and convex

2. Lower deviation

$$F = \sum_{i=1}^n \alpha_i p_{0i}$$

$$\underline{d}_E(P) = \sqrt{\sum_{i=1}^n (p_i - F)^2}$$

$$D_E(P) = P_0(F) - \alpha_F \underline{d}_E(P)$$

3. Lower probability

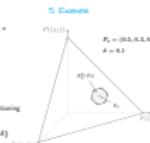
$$\underline{P}_E(A) = P_0(A) - \delta \sqrt{\frac{1-P_0(A)}{P_0(A)}}$$

4. Properties

- \underline{P}_E is not 2-monotone...
- ... but it is 2-monotone on events
- the model is not preserved under conditioning
- extreme points:

$$\text{ac}(D_E^{\delta}(P_0)) = \{P \mid d_E(P, P_0) = \delta\}$$

5. Example



WOULDN'T IT BE BETTER TO USE A DIVERGENCE?

BUT FOR COMPARING PROBABILITIES...
 NICE POINT...



KULLBACK-LEIBLER MODEL

1. KL-divergence

$$D_{KL}(P, P_0) = \sum_{i=1}^n p_i \log\left(\frac{p_i}{p_{0i}}\right)$$

D_{KL} continuous and convex

2. Lower deviation

$$F = \sum_{i=1}^n \alpha_i p_{0i}$$

$$\underline{d}_{KL}(P) = \min_{\alpha \in [0, 1]} \alpha P_0(F)$$

subject to

$$\sum_{i=1}^n \alpha p_{0i}^2 = 1, \quad \sum_{i=1}^n \alpha p_{0i}^2 \log(p_{0i}) = \delta$$

3. Lower probability

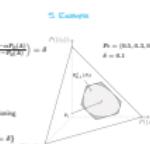
$$\underline{P}_{KL}(A) = P_0(A) - \delta \frac{1 - P_0(A)}{P_0(A)}$$

4. Properties

- \underline{P}_{KL} is not 2-monotone...
- ... but it is 2-monotone on events
- the model is not preserved under conditioning
- extreme points:

$$\text{ac}(D_{KL}^{\delta}(P_0)) = \{P \mid D_{KL}(P, P_0) = \delta\}$$

5. Example



WHICH ONE SHOULD I USE?

AMONG ALL THE MODELS...
 IT DEPENDS ON WHAT YOU ARE LOOKING FOR...



COMPARISON

1. Properties

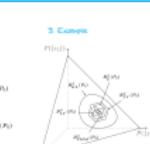
Distortion model	not	not	not
on gamblers	✓	✓	✓
on events	✓	✓	✓
No-distortion model	✓	✓	✓
on events	✓	✓	✓
KL	✓	✓	✓
COR	✓	✓	✓
TV	✓	✓	✓
Vertical barrier models	✓	✓	✓

2. Amount of information

```

graph TD
    A[P₀(x₁)] --> A₁[P₁(x₁)]
    A --> A₂[P₁(x₂)]
    A --> A₃[P₁(x₃)]
    A₁ --> A₁₁[P₂(x₁)]
    A₁ --> A₁₂[P₂(x₂)]
    A₁ --> A₁₃[P₂(x₃)]
    A₂ --> A₂₁[P₂(x₁)]
    A₂ --> A₂₂[P₂(x₂)]
    A₂ --> A₂₃[P₂(x₃)]
    A₃ --> A₃₁[P₂(x₁)]
    A₃ --> A₃₂[P₂(x₂)]
    A₃ --> A₃₃[P₂(x₃)]
    A₁₁ --> A₁₁₁[P₃(x₁)]
    A₁₁ --> A₁₁₂[P₃(x₂)]
    A₁₁ --> A₁₁₃[P₃(x₃)]
    A₁₂ --> A₁₂₁[P₃(x₁)]
    A₁₂ --> A₁₂₂[P₃(x₂)]
    A₁₂ --> A₁₂₃[P₃(x₃)]
    A₁₃ --> A₁₃₁[P₃(x₁)]
    A₁₃ --> A₁₃₂[P₃(x₂)]
    A₁₃ --> A₁₃₃[P₃(x₃)]
    A₂₁ --> A₂₁₁[P₃(x₁)]
    A₂₁ --> A₂₁₂[P₃(x₂)]
    A₂₁ --> A₂₁₃[P₃(x₃)]
    A₂₂ --> A₂₂₁[P₃(x₁)]
    A₂₂ --> A₂₂₂[P₃(x₂)]
    A₂₂ --> A₂₂₃[P₃(x₃)]
    A₂₃ --> A₂₃₁[P₃(x₁)]
    A₂₃ --> A₂₃₂[P₃(x₂)]
    A₂₃ --> A₂₃₃[P₃(x₃)]
    A₃₁ --> A₃₁₁[P₃(x₁)]
    A₃₁ --> A₃₁₂[P₃(x₂)]
    A₃₁ --> A₃₁₃[P₃(x₃)]
    A₃₂ --> A₃₂₁[P₃(x₁)]
    A₃₂ --> A₃₂₂[P₃(x₂)]
    A₃₂ --> A₃₂₃[P₃(x₃)]
    A₃₃ --> A₃₃₁[P₃(x₁)]
    A₃₃ --> A₃₃₂[P₃(x₂)]
    A₃₃ --> A₃₃₃[P₃(x₃)]
  
```

3. Example



DOES ANYBODY CARE ABOUT DISTORTION MODELS?

MANY PEOPLE!



REFERENCES

Distortion models:

- [1] Martos, Ignacio. *Measuring distortion models: A comparative study*. UAR, 2022.
- [2] Martos, Ignacio. *Defining neighbourhood and distortion models*. Part I. II. UAR, 2020.
- [3] Mikel, Pedro; Martos, Ignacio. *Probabilistic probabilities as an uncertainty model*. ISIPTA, 2018.
- [4] Peleggi, Vicig. *Zaffalon. Inference and risk measurement with the prior-exact model*. UAR, 2016.
- [5] Walley. *Statistical reasoning with imprecise Probabilities*. 2001.

Distance-based models:

- [6] Henao, Soledad; Wasserman. *Convex conditioning: better results on dilation*. ISIPTA, 1997.

Vertical barrier models:

- [7] Corato, Peleggi, Vicig. *Nearly Linear uncertainty measure*. UAR, 2019.

Transformations of P₀:

- [8] Booswijk. *On the closure of families of fuzzy measures under envelope aggregation*. FIS, 2015.

KL-divergence in P₁:

- [9] Martos, Ignacio. *Discussing imprecise probabilities. The Mathematics of the uncertain*. 2018.
- [10] Martos, Ignacio. *Model Discarding imprecise probabilities*. UAR, 2018.