Copulas, Lower Probabilities and Random Sets

How and When to apply them?

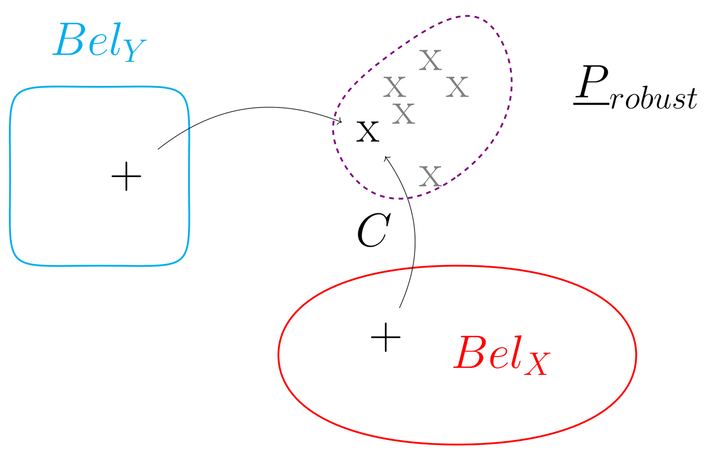
heudiasyc

Roman MALINOWSKI

Problem statement

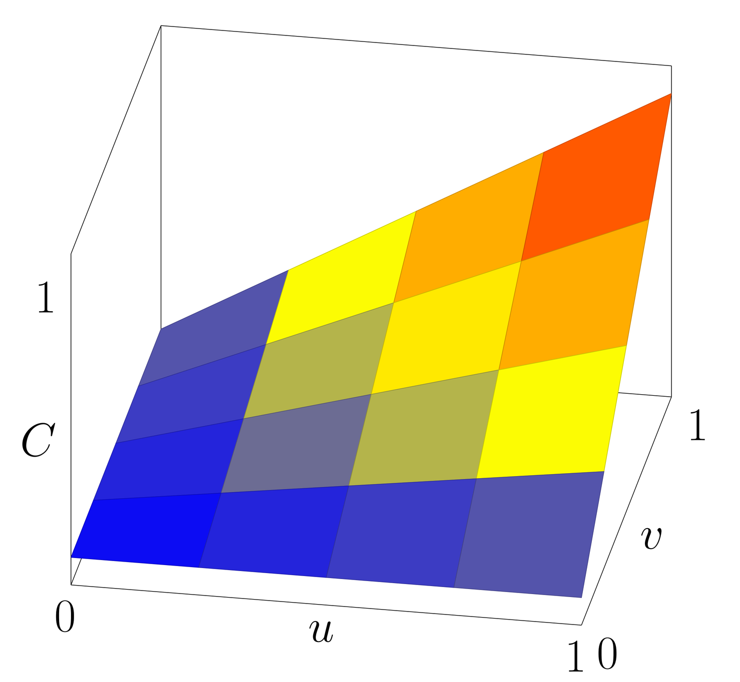
Given two discrete imprecise random variables X and Y, whose probability measures belong to credal sets $\mathcal{M}(Bel_X)$ and $\mathcal{M}(Bel_Y)$ respectively, and whose dependency is modeled by a copula C, how to join those two credal sets using the copula C?

Robust Method



We join every element of $\mathcal{M}(Bel_X)$ and $\mathcal{M}(Bel_Y)$ using the copula C and take its lower bound $\underline{P}_{robust}$ on every event. This is the lower bound of the tightest credal set $\mathcal{M}_{robust}$ containing every marginal probabilities joined by C. This is usually the optimal credal set one wishes to know, but it can be very hard to compute its lower bounds.

Copulas

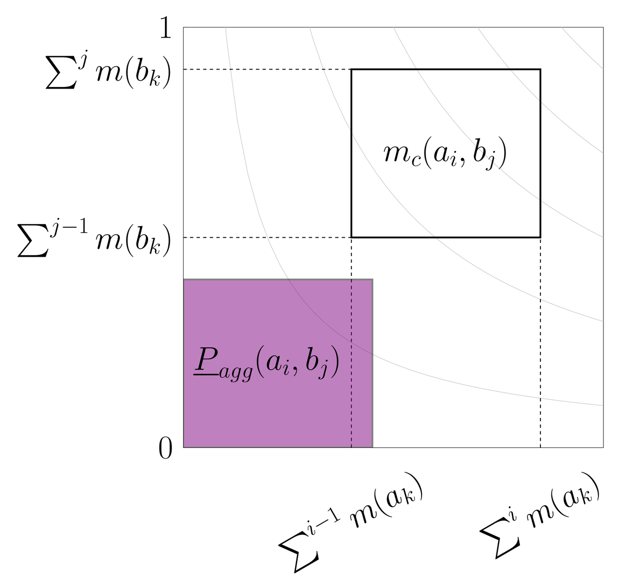


A copula is a mapping $C : [0,1]^n \to [0,1]$ which can model any dependency between n Cumulative Distribution Functions.
Sklar's theorem [3] states that any multivariate CDF G can be expressed using a unique copula C and its marginals F_i:

$$G(x_1, \ldots, x_n) = C(F_1(x_1), \ldots, F_n(x_n))$$

Conversely, any copula applied to n univariate CDFs correctly defines a n-dimension CDF.

Aggregation Method



$\underline{P}_{agg}$ is computed as:

$$\underline{P}_{agg}(.,.) = C(Bel_X(.), Bel_Y(.))$$

This method loses the "meaning" of a copula, but is very easy to compute.

Meaningful ←

$\mathcal{M}_{robust} \quad \mathcal{M}_{mass} \quad \mathcal{M}_{agg}$

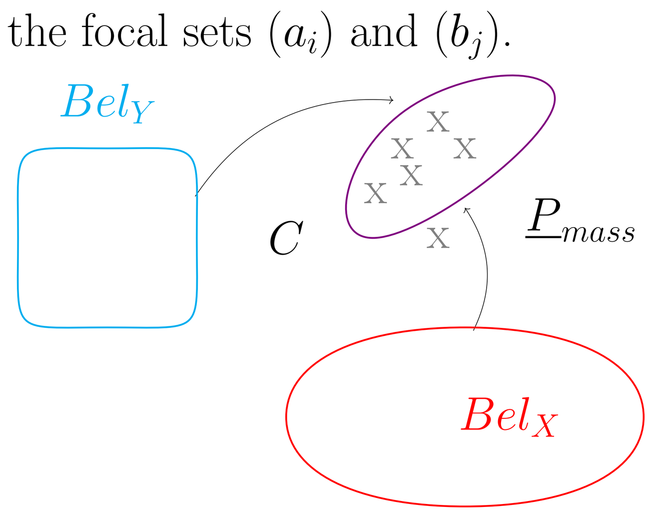
→ Easy to compute

Cumulated masses method

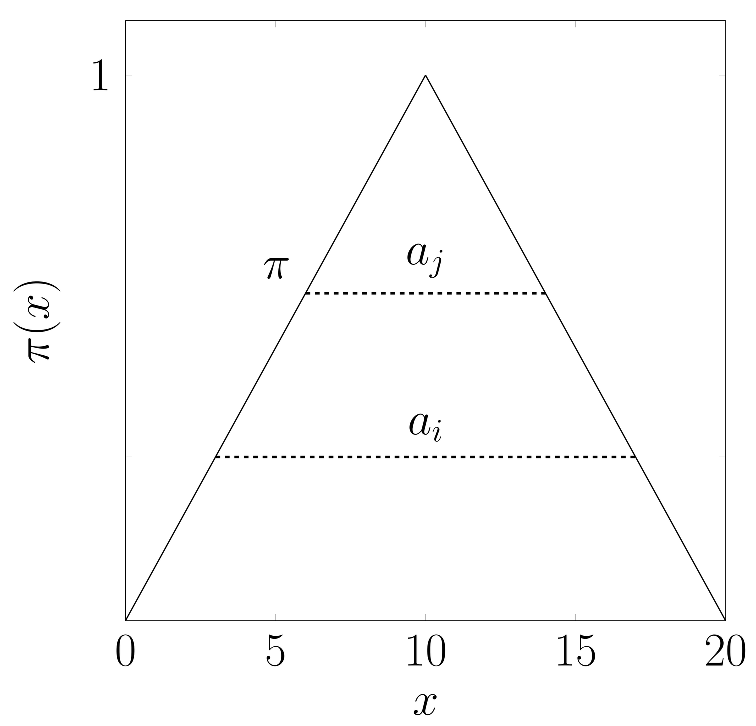
This method defines a bivariate belief function $\underline{P}_{mass}$ by joining the masses of the marginal belief functions using the H-volume of a copula [2]. In the bivariate case, writing $A_I = \sum_1^I m_X(a_i)$ and $B_J = \sum_1^J m_X(b_j)$, the H-volume is: $m_{XY}(a_I, b_J) = C(A_I, B_J) + C(A_{I-1}, B_{J-1}) - C(A_{I-1}, B_J) - C(A_I, B_{J-1})$.

The credal set resulting from $\underline{P}_{mass}$ is noted $\mathcal{M}_{mass}$. There is in general no reason for having $\mathcal{M}_{robust} = \mathcal{M}_{mass}$.

To compute the cumulated masses, it is necessary to specify an order on the focal sets (a_i) and (b_j).



Necessity functions as Marginals



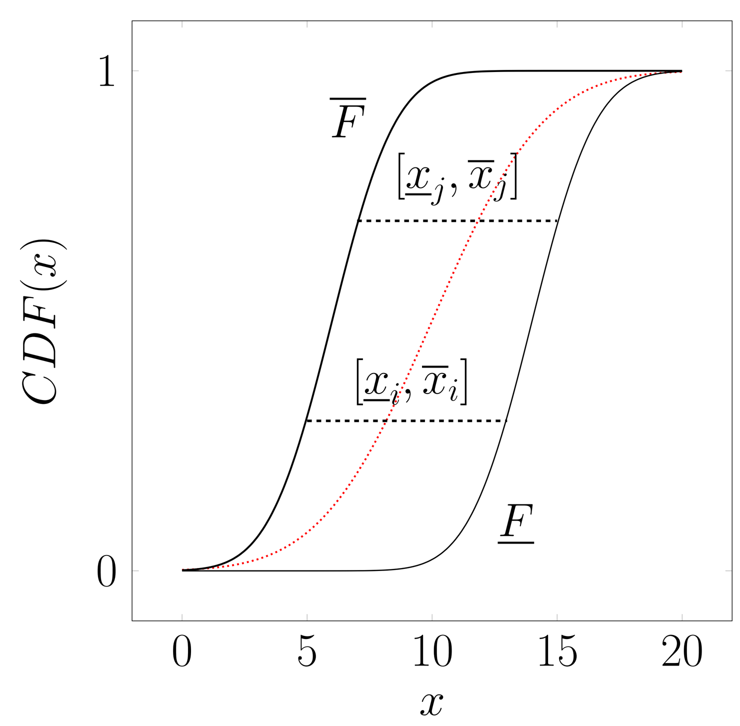
Possibility distributions are mappings $\pi : \mathcal{X} \to [0,1]$ where $\pi(x) = 1$ at least once.

Possibility distributions define special types of belief functions called Necessity functions : $Nec(A) = 1 - \sup_{x \in A^c} \pi(x)$

The focal sets of a Necessity function are the α-cuts of its possibility distribution: $a = \{x \mid \pi(x) > \alpha\}$. Focal sets form a nested family of events and we can thus define a natural order on them: $a_i \subset a_j \Leftrightarrow a_i \prec a_j$

When marginals are Necessity functions, it holds that $\underline{P}_{mass} = \underline{P}_{agg}$. But both $\mathcal{M}_{robust} \subset \mathcal{M}_{mass}$ and $\mathcal{M}_{robust} \supset \mathcal{M}_{mass}$ can happen.

P-boxes as Marginals



A P-box is a pair of cumulative distribution functions $[\underline{F}, \overline{F}]$, one dominating the other: $\underline{F} \leqslant \overline{F}$. A P-box can define a credal set $\mathcal{M} = \{F \mid \underline{F} \leqslant F \leqslant \overline{F}\}$.

The focal sets of a p-box are its α-cuts: $\{x \mid \underline{F}(x) \leqslant \alpha \leqslant \overline{F}(x)\} = [\underline{x}, \overline{x}]$, on which we can define a natural order:

$$\underline{x}_i < \underline{x}_j \text{ or } \overline{x}_i < \overline{x}_j \Leftrightarrow [\underline{x}_i, \overline{x}_i] \prec [\underline{x}_j, \overline{x}_j]$$

When marginals are p-boxes, it holds that $\underline{P}_{robust} \geqslant \underline{P}_{mass}$ and thus $\mathcal{M}_{robust} \subseteq \mathcal{M}_{mass}$. This can be proven by integrating two indicator functions (one for any F in the p-box, and one for the p-box itself) along the density copula.

Product Copula

When using the product copula (independence), it holds that [1]: $\mathcal{M}_{robust} \subseteq \mathcal{M}_{mass} \subseteq \mathcal{M}_{agg}$. Their lower envelope coincide on Cartesian product of events.

Conclusion and Perspectives

- Three methods for joining copulas and credal sets: depending on the marginals, some inclusions allow for easier computations
- Work done since submission: If the copula is directionally convex, it holds that $\underline{P}_{mass} \geqslant \underline{P}_{agg}$ and thus $\mathcal{M}_{mass} \subseteq \mathcal{M}_{agg}$
- Future work with imprecise copulas or other imprecise models (clouds)
- Finding alternative orders for Necessity functions (seems an order for $\mathcal{M}_{robust} \subseteq \mathcal{M}_{mass}$ always exist)

References

[1] Ines Couso, Serafin Moral, and Peter Walley. "A survey of concepts of independence for imprecise probabilities". In: *Risk Decision and Policy* (June 2000).

[2] Scott Ferson et al. *Dependence in probabilistic modeling, Dempster-Shafer theory, and probability bounds analysis.* Tech. rep. Oct. 2004.

[3] M. Sklar. *Fonctions de Répartition À N Dimensions Et Leurs Marges.* Université Paris 8, 1959.

Acknowledgment

This project has received financial support from the CNRS through the MITI interdisciplinary programs and from CNES

www.hds.utc.fr

CS GROUP · cnes · Alliance Sorbonne Université · utc Recherche · cnrs