Copulas, lower probabilities and random sets: how and when to apply them?

Roman Malinowski, Sébastien Destercke

Copulas, lower probabilities and random sets: how and when to apply them?

Roman Malinowski, Sébastien Destercke

Abstract A copula is an aggregation function, that can be used as a dependency model. Any multivariate distribution function can be characterized by its marginals and a copula. When introducing imprecision in the modelling of those distribution functions, different solutions are available to aggregate the univariate uncertainty representations into a multivariate one via the copula. We present some of those solutions, and discuss their respective inclusions for special cases: independence, belief functions, necessity functions and p-boxes.

1 Introduction

Credal sets, or convex sets of probability distributions, are useful tools to reason under uncertainty, especially in the presence of imprecision. They include many uncertainty models, such as belief functions, p-boxes, etc (Destercke et al., 2008). How to combine such univariate models into multivariate ones remain a very active research question, with many researchers studying how tools used in the precise setting can be extended to the imprecise one, including in particular copulas (Gray et al., 2021; Montes et al., 2015).

When considering precise probabilities, copulas have been shown to be able to model any dependency structure between probability distributions. This is no longer true when adding imprecision to probabilities (Montes et al., 2015), for various reasons, such as the fact that imprecise cumulative distributions and credal sets are no longer in one-to-one correspondence.

Roman Malinowski (✉)
CNES / CS Group / UTC, 18 avenue Edouard Belin, Toulouse e-mail: roman.malinowski@utc.fr

Sébastien Destercke (✉)
Université de Technologie de Compiègne (UTC), Avenue de Landshut Compiègne e-mail: sebastien.destercke@utc.fr

In this paper, we first provide a short review of different models of uncertainty (Section 2), and then present multiple ways of combining creal sets or their lower envelopes with a copula (Section 3). In Section 4, we finally study their relationships for various specific cases of interest.

2 Preliminaries

In this section, we present the various uncertainty models we will consider, as well as copulas, which are aggregation functions typically used to model dependencies between random variables. We will work on finite spaces.

2.1 Imprecise Models

Definition 1 Given a credal set[1] $\mathcal{M}$ over a finite space $\mathcal{X}$, we define its lower probability $\underline{P}$ as its lower envelope on all events of the power set $\mathcal{P}(\mathcal{X})$:

$$\forall A \subseteq \mathcal{X}, \underline{P}(A) = \inf\{P(A)|P \in \mathcal{M}\} \tag{1}$$

Note that although lower probabilities cannot describe any credal set, we will mostly focus on them in this paper to privilege clarity of exposure. As credal sets are quite generic, it is useful to consider simpler, more practical models. We now define such models that we will consider later on.

Definition 2 Given a finite space $\mathcal{X}$ and its power set $\mathcal{P}(\mathcal{X})$, a probability mass function (Shafer (1976)) is a mapping $\mathcal{P}(\mathcal{X}) \to [0,1]$ satisfying:

$$m(\emptyset) = 0 \text{ and } \sum_{A \subseteq \mathcal{X}} m(A) = 1 \tag{2}$$

A set A of $\mathcal{X}$ is called a *focal set* if and only if $m(A) > 0$ and we will note $\mathcal{F}_m$ the set of all focal sets. Such a probability mass function defines a Belief function $Bel : \mathcal{P}(\mathcal{X}) \to [0,1]$ and a Plausibility function $Pl : \mathcal{P}(\mathcal{X}) \to [0,1]$:

$$\forall A \subseteq \mathcal{X}, Bel(A) = \sum_{B \subseteq A} m(B) \text{ and } Pl(A) = \sum_{B \cap A \neq \emptyset} m(B) \tag{3}$$

Those two functions are conjugate as $Bel(A) = 1 - Pl(A^c)$. $Bel(A)$ can be interpreted as our belief that the truth lies in A. A belief function is a lower probability (1) inducing a credal set $\mathcal{M}(Bel) = \{P \mid \forall A \subseteq \mathcal{X}, Bel(A) \leqslant P(A)\}$.

Definition 3 A necessity measure Nec is a minitive belief function:

$$\forall A, B \subseteq \mathcal{X}, Nec(A \cap B) = min(Nec(A), Nec(B)). \tag{4}$$

[1] A convex set of probability distributions.

In a finite space, focal elements of a necessity function form a nested family of events : $\mathcal{F}_m = \{a_1 \subset ... \subset a_k\}$ (Dubois and Prade, 2009). As a specific belief function, a necessity induces a credal set $\mathcal{M}(Nec) = \{P | Nec(A) \leqslant P(A), \forall A\}$.

Definition 4 A p-box (probability-box) is a pair of two cumulative distribution functions (CDFs) $[\underline{F}, \overline{F}]$ defined on the real line s.t. $\underline{F}(x) \leqslant \overline{F}(x),\ \forall x \in \mathbb{R}$. A p-box is the extension of CDFs to imprecise probabilities. It induces a credal set, composed of all the CDFs dominating $\underline{F}$ and dominated by $\overline{F}$:

$$\mathcal{M}([\underline{F}, \overline{F}]) = \{P | \forall x \in \mathbb{R}, \underline{F}(x) \leqslant P(]-\infty, x]) \leqslant \overline{F}(x)\} \tag{5}$$

We can define 'α-levels' $C^{\alpha}_{[\underline{F},\overline{F}]}$ ($\alpha \in [0,1]$) of a p-box as intervals $[\underline{x}, \overline{x}]$ whose lower bound (resp. upper bound) is the pseudo-inverse of $\underline{F}$ (resp. $\overline{F}$) at α (Figure 1). It has been proven in Destercke et al. (2008) that to each p-box corresponds a belief function Bel s.t.:

$$\mathcal{M}([\underline{F}, \overline{F}]) = \mathcal{M}(Bel) \tag{6}$$

Additionally, the set of focal elements of a p-box is included in the α-levels $\mathcal{F}_m \subseteq C^{\alpha}_{[\underline{F},\overline{F}]}$. We will note $\mathcal{F}_m = \{[\underline{x}_k, \overline{x}_k] \mid m([\underline{x}_k, \overline{x}_k]) > 0\}$. Due to the fact that both $\underline{F}$ and $\overline{F}$ are increasing mappings, the lower and upper bounds of those intervals are ordered. Thus all focal elements $[\underline{x}_k, \overline{x}_k]$ can be ordered using the natural order:

$$\forall i, j, [\underline{x}_i, \overline{x}_i] \leqslant_{nat} [\underline{x}_j, \overline{x}_j] \Leftrightarrow \underline{x}_i \leqslant \underline{x}_j \text{ and } \overline{x}_i \leqslant \overline{x_j} \tag{7}$$

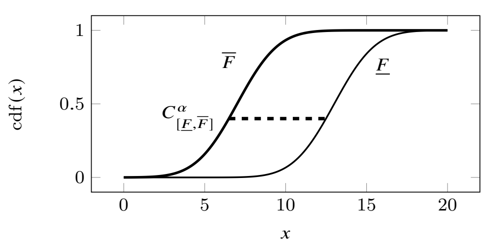

Fig. 1 p-box and one of its focal elements $[\underline{x}_i, \overline{x}_i]$

2.2 Copulas

Definition 5 A copula is a multivariate distribution function $C : [0,1]^N \to [0,1]$ whose marginals are uniform on $[0,1]$. It can be seen as a joint distribution function of N random variables. A copula verifies a number of properties, expressed here in the bivariate case. $\forall u, u', v, v' \in [0,1]^4$ s.t. $u \leqslant u', v \leqslant v'$:

$$C(u,0) = C(0,v) = 0,\ C(u,1) = u,\ C(1,v) = v \tag{8}$$

$$C(u',v') + C(u,v) - C(u,v') - C(u',v) \geqslant 0 \tag{9}$$

From (9) we have that C is a component-wise increasing mapping. There exists two copulas, the Łukasiewicz (C_L) and Minimum (C_M) copulas (also called lower and upper Fréchet–Hoeffding copulas in Nelsen (2006)), that bound all copulas C, i.e, $\forall u, v \in [0,1]$,

$$\max(0, u+v-1) \triangleq C_L(u,v) \leqslant C(u,v) \leqslant C_M(u,v) \triangleq \min(u,v) \tag{10}$$

The celebrated Sklar's theorem states that copulas, when applied to cumulative distributions of precise probabilities, can model any multivariate function, and that any multivariate function can be modelled by a copula applied to its marginals. As said earlier, this is however no longer true in the imprecise setting, hence the need to restudy how copulas can be applied to generic credal sets and lower probabilities (Montes et al., 2015).

3 Applying copulas to credal sets

There are multiple ways to apply a copula to lower probabilities. We will describe some of them: a robust method on dominated probabilities, a method on mass distributions inducing belief functions, and an aggregation method.

3.1 Robust method on dominated probabilities

Consider a copula C, and two credal sets $\mathcal{M}(\underline{P}_X)$ and $\mathcal{M}(\underline{P}_Y)$ defined over $\mathcal{P}(\mathcal{X})$ and $\mathcal{P}(\mathcal{Y})$ respectively. Applying C to every marginal $P_X \in \mathcal{M}(\underline{P}_X)$ and $P_Y \in \mathcal{M}(\underline{P}_Y)$ gives a joint lower probability such that for all $E \subseteq \mathcal{X} \times \mathcal{Y}$:

$$\underline{P}_{Robust}(E) = \inf\{P_{XY}(E) | F_{XY}(x,y) = C\left(F_X(x), F_Y(y)\right), \forall (x,y) \in \mathcal{X}\times\mathcal{Y}\} \tag{11}$$

with F_X, F_Y the CDF of P_X, P_Y and $F_{XY}(x,y) = P_{XY}(X \leqslant x, Y \leqslant y)$. Because F_{XY} is a precise CDF, it completely determines P_{XY}, allowing P_{XY} to be computed on events that are not Cartesian products. $\underline{P}_{Robust}(E)$ is then the infinimum of those probability distributions on E. Note that for defining the CDFs on finite spaces that are not subsets of $\mathbb{R}$, complete orderings on $\mathcal{X}$ and $\mathcal{Y}$ must be defined. In the following sections, we will refer to $\underline{P}_{Robust}$ as in (11) and its credal set generated with two univariate lower probabilities $\underline{P}_X$, $\underline{P}_Y$ and a copula C as:

$$\mathcal{M}_{Robust}(C, \underline{P}_X, \underline{P}_Y) = \{P_{XY} \mid \underline{P}_{Robust} \leqslant P_{XY}\} \tag{12}$$

or $\mathcal{M}_{Robust}$ [2] for short when no ambiguity can arise.

3.2 Joint masses from copulas

As mass functions inducing belief functions can be seen as probabilities over sets, one could directly apply copulas to those masses. However, in general there is no natural order over the set of focal sets $\mathcal{F}_m$ (which then play the role of atoms). To apply a copula and define a joint mass, we have to chose an arbitrary ordering $\leqslant_{arb}$ that will determine the order inside the sets of focal sets $\mathcal{F}_{m_X} = \{a_1, a_2, ..., a_n\}$ and $\mathcal{F}_{m_Y} = \{b_1, b_2, ..., b_{n'}\}$ with $a_{i-1} \leqslant_{arb} a_i\ \forall i \in [[1, n]]$ and $b_{j-1} \leqslant_{arb} b_j\ \forall j \in [[1, n']]$ (with $a_0 = b_0 = \emptyset$). The bivariate mass associated to an element (a_i, b_j) of $\mathcal{F}_{m_X} \times \mathcal{F}_{m_Y}$ can be defined using the diagonal difference as in Ferson et al. (2004) :

$$m_{XY}^C(a_i \times b_j) = C(A_X^i, B_Y^j) + C(A_X^{i-1}, B_Y^{j-1}) - C(A_X^i, B_Y^{j-1}) - C(A_X^{i-1}, B_Y^j) \quad (13)$$

with $A_X^i = \sum_{k \leqslant i} m_X(a_k)$ and $B_Y^j = \sum_{k \leqslant j} m_Y(b_k)$ being the cumulative masses over a_i and b_j (with $A_X^0 = m_X(a_0) = B_Y^0 = m_Y(b_0) = 0$). It is easy to check that the joint mass defined in (13) is a mass distribution whose focal sets form a subset of the Cartesian product of the marginal focal sets $\mathcal{F}_{m_X} \times \mathcal{F}_{m_Y}$.

This mass induces a belief function Bel_{XY} from which can be generated a credal set. In the following sections, we will refer to this credal set as:

$$\mathcal{M}_{mass}(C, Bel_X, Bel_Y) = \{P_{XY} \mid Bel_{XY} = \sum m_{XY}^C \leqslant P_{XY}\} \quad (14)$$

or $\mathcal{M}_{mass}$ for short, where m_{XY}^C is defined as in (13). However, as shows the next example, the choice of $\leqslant_{arb}$ can strongly impact the joint model.

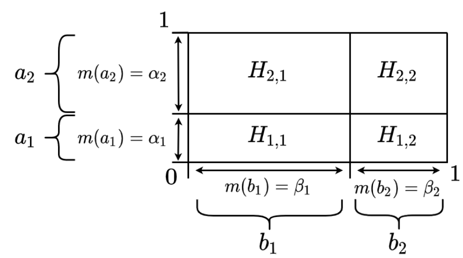

Fig. 2 Joint probability over events A and B

[2] $\mathcal{M}_{Robust}$ is the smallest credal set containing all probabilities from the marginal credal sets linked with C, but as it is convex, it can also contain probabilities linked to their marginals with a different copula C'.

Example 1 Let m_X and m_Y be mass functions defined over $\mathcal{P}(\mathcal{X})$ and $\mathcal{P}(\mathcal{Y})$, and whose sets of focal sets are $\mathcal{F}_{m_X} = \{a_1, a_2\}$ and $\mathcal{F}_{m_Y} = \{b_1, b_2\}$ (Fig 2). Assuming that $b_1 \leqslant_{arb} b_2$, consider the following $\leqslant_{arb}$ on $\mathcal{F}_{m_X}$:

- If "$a_1 \leqslant_{arb} a_2$": $m_{XY}(a_1 \times b_1) = C(\alpha_1, \beta_1)$, $m_{XY}(a_2 \times b_1) = \beta_1 - C(\alpha_1, \beta_1)$
- If "$a_1 \geqslant_{arb} a_2$": $m_{XY}(a_1 \times b_1) = \beta_1 - C(\alpha_2, \beta_1)$, $m_{XY}(a_2 \times b_1) = C(\alpha_2, \beta_1)$

Taking $C(u, v) = \min(u, v)$ and $\alpha_1 = \alpha_2 = \beta_1 = 0.5$, yields in the first case $m_{xy}(a_1 \times b_1) = 0.5$ and in the second case $m_{xy}(a_1 \times b_1) = 0$. There is, in general, no reason for two orders to have the same value of masses for the same Cartesian product of events. A notable exception is the product copula, explored below. Finally, note that such an approach is quite commonly encountered in the literature (Gray et al., 2021; Alvarez et al., 2018)

3.3 Copula applied to the lower probabilities

Another way to aggregate two lower probabilities $\underline{P}_X, \underline{P}_Y$ with a copula C is by directly applying the copula to lower probabilities:

$$\forall A \in \mathcal{X}, B \in \mathcal{Y}, \underline{P}_{agg}(A \times B) = C(\underline{P}_X(A), \underline{P}_Y(B)) \tag{15}$$

Here, we use a copula only as an aggregation operator. Note that in general we cannot expect $\underline{P}_{agg}$ to induce a non-empty $\mathcal{M}(\underline{P}_{agg})$. We can nevertheless note that in the case of the product copula, the bivariate lower probability induced by (15) induces a non empty credal set if the marginals credal sets are not empty, i.e. $\exists P_X \geqslant \underline{P}_X, \exists P_Y \geqslant \underline{P}_Y \implies \exists P_{XY} = P_X \times P_Y \geqslant \underline{P}_{agg}$ due to the fact that it is true for any precise probability. It follows that for all copulas C such that $C_\Pi(u, v) = u \times v \geqslant C(u, v)$, we have

$$\underline{P}_X \underline{P}_Y \geqslant C(\underline{P}_X, \underline{P}_Y) = \underline{P}_{agg} \implies \mathcal{M}(\underline{P}_{agg}) \neq \emptyset \tag{16}$$

In the following sections, we will refer to the credal set generated by aggregating directly two univariate lower probabilities with a copula C as:

$$\mathcal{M}_{agg}(C, \underline{P}_X, \underline{P}_Y) = \{P_{XY} \mid \underline{P}_{agg} = C(\underline{P}_X, \underline{P}_Y) \leqslant P_{XY}\} \tag{17}$$

or $\mathcal{M}_{agg}$ for short, and we will note $\underline{P}_{agg}$ its lower envelope.

We have defined three different ways to apply copulas to probability sets. In general, they will generate different, incomparable probability sets. Also, while $\mathcal{M}_{Robust}$ is probably the most well-grounded way to extend copulas, it may generate a complex set, while $\mathcal{M}_{mass}$ is guaranteed to be a belief function and $\mathcal{M}_{agg}$ is easy to compute and generate. In the next section, we study different special cases where those sets enjoy some specific relationships.

4 Special cases

This section explores what happens when considering the product copula, and when marginal models are both necessity measure or p-boxes.

Product Copula: In the case of the product copula, it is known (Couso et al., 2000) that all $\underline{P}_{robust}, \underline{P}_{mass}$ and $\underline{P}_{agg}$ factorize over Cartesian products, i.e., $\forall\ (A, B) \subseteq \mathcal{X} \times \mathcal{Y}, \underline{P}(A \times B) = \underline{P}_X(A) \times \underline{P}_Y(B)$. It is also known (Couso et al., 2000) that $\mathcal{M}_{Robust} \subseteq \mathcal{M}_{mass}$, with the inclusion being sometimes strict. Also, as $\mathcal{M}_{agg}$ is the largest credal set enjoying this factorisation properties, we necessarily have

$$\mathcal{M}_{Robust}(C_\Pi, Bel_X, Bel_Y) \subseteq \mathcal{M}_{mass}(C_\Pi, Bel_X, Bel_Y) \subseteq \mathcal{M}_{agg}(C_\Pi, Bel_X, Bel_Y)$$

therefore having strong relationships in the case of the product copula[3].

Necessity functions: Let us first compare $\mathcal{M}_{agg}$ and $\mathcal{M}_{mass}$. In the case of necessity functions, a natural ordering $\leqslant_{nat}$ between focal elements exists which corresponds to the inclusion ordering. Because of this, it holds that $\sum_{k \leqslant_{nat} i} \sum_{l \leqslant_{nat} j} m_{XY}(a_k \times b_l) = \sum_{a_k \subseteq a_i} \sum_{b_l \subseteq b_j} m_{XY}(a_k \times b_l)$ and thus computing the belief function Bel_{XY} defined in (14) on all focal elements a_i of Nec_X and b_j of Nec_Y yields:

$$Bel_{XY}(a_i \times b_j) = C\left(Nec_X(a_i), Nec_Y(b_j)\right)$$

Thus the lower envelope of $\mathcal{M}_{agg}$ and $\mathcal{M}_{mass}$ coincide on the Cartesian products of events. Since $\mathcal{M}_{agg}$ is again the largest set with this lower envelope, we do have $\mathcal{M}_{mass} \subseteq \mathcal{M}_{agg}$ in the case of necessity functions. However, as show the next example, there is no reason in general for

$$\mathcal{M}_{agg} \subseteq \mathcal{M}_{Robust} \text{ nor } \mathcal{M}_{agg} \supseteq \mathcal{M}_{Robust}$$

Example 2 Consider two necessity function Nec_X, Nec_Y defined over $\mathcal{X} = \{x_1, x_2\}$ and $\mathcal{Y} = \{y_1, y_2\}$ respectively, s.t $Nec_X(x_1) = Nec_Y(y_2) = 0$ and $Nec_X(x_2) = Nec_Y(y_1) = 0.9$. Let us first consider the Łukasiewicz copula $C(u, v) = max(0, u + v - 1)$ and events $x_2 \times y_1$. It is possible to show that on those events $\underline{P}_{Robust} = 0.9 > 0.8 = \underline{P}_{agg}$.

If we consider the Minimum copula $C(u, v) = \min(u, v)$, then taking any (P_X, P_Y) verifying $P_X(x_1) = 0.1$ and $P_Y(y_1) = 0.9$ and computing $\underline{P}_{agg}$ and $P_{XY} \in \mathcal{M}_{Robust}$ over event $x_2 \times y_1$ yields $\underline{P}_{agg}(x_2 \times y_1) = 0.9$ and $P_{XY}(x_2 \times y_1) = 0.8$. Thus it holds that on those events $\underline{P}_{agg} = 0.9 > 0.8 \geqslant \underline{P}_{Robust}$.

P-boxes When considering uncertainty represented by two p-boxes $[\underline{F}_X, \overline{F}_X]$, $[\underline{F}_Y, \overline{F}_Y]$, it does not hold in general that $\mathcal{M}_{Robust} \subseteq \mathcal{M}_{agg}$ nor $\mathcal{M}_{Robust} \supseteq$

[3] In this case, $\mathcal{M}_{mass}$ is insensitive to re-ordering, as C_Π corresponds to the uniform distribution over $[0, 1]^2$

$\mathcal{M}_{agg}$. We refer to example 2 when considering the p-boxes induced from necessity functions as in Baudrit and Dubois (2006). It however holds that (Appendix A):

$$\mathcal{M}_{Robust}(C, Bel_X, Bel_Y) \subseteq \mathcal{M}_{mass}(C, Bel_X, Bel_Y)$$

5 Conclusion

We presented different methods for joining uncertainty models with a copula: a robust method, a method based on cumulated masses, and a method using the copula as a direct aggregation operator. We showed that in the special case of the product copula, the lower probabilities coincide on Cartesian products. When using necessity functions, the aggregated lower probability coincides with the cumulated masses one on Cartesian products, and when using p-boxes we showed that the credal set obtained with the robust method is included in the credal set of the cumulated mass approach.

References

Alvarez DA, Uribe F, Hurtado JE (2018) Estimation of the lower and upper bounds on the probability of failure using subset simulation and random set theory. Mechanical Systems and Signal Processing 100:782–801

Baudrit C, Dubois D (2006) Practical representations of incomplete probabilistic knowledge. Computational Statistics & Data Analysis 51:86–108

Couso I, Moral S, Walley P (2000) A survey of concepts of independence for imprecise probabilities. Risk Decision and Policy 5:165–181

Destercke S, Dubois D, Chojnacki E (2008) Unifying practical uncertainty representations – I: Generalized p-boxes. International Journal of Approximate Reasoning 49:664–677

Dubois D, Prade H (2009) Possibility Theory, Springer New York

Ferson S, Nelsen RB, Hajagos J, Berleant DJ, Zhang J, Tucker WT, Ginzburg LR, Oberkampf WL (2004) Dependence in probabilistic modeling, Dempster-Shafer theory, and probability bounds analysis

Gray A, Hose D, De Angelis M, Hanss M, Ferson S (2021) Dependent possibilistic arithmetic using copulas. In: International Symposium on Imprecise Probability: Theories and Applications, PMLR, pp 169–179

Montes I, Miranda E, Pelessoni R, Vicig P (2015) Sklar's theorem in an imprecise setting. Fuzzy Sets and Systems 278

Nelsen RB (2006) An introduction to copulas. Springer, New York

Shafer G (1976) A mathematical theory of evidence. Princeton University Press

Appendix A

We define the pseudo-inverse of a CDF F as $F^{-1}(\alpha) = \min\{x|F(x) = \alpha\}$. In the continuous case we have

$$\begin{aligned} P_{XY}(A_X, A_Y) &= \int_{a \in A_X} \int_{b \in A_Y} dF_{XY}(a, b) \\ &= \int_{\alpha=0}^{1} \int_{\beta=0}^{1} \mathbb{I}(F_X^{-1}(\alpha) \in A_X \& F_Y^{-1}(\beta) \in A_Y) dC(\alpha, \beta) \\ Bel_{XY}(A_X, A_Y) &= \int_{\alpha=0}^{1} \int_{\beta=0}^{1} \mathbb{I}([\overline{F}_X^{-1}(\alpha), \underline{F}_X^{-1}(\alpha)] \subseteq A_X \& [\overline{F}_Y^{-1}(\beta), \underline{F}_Y^{-1}(\beta)] \subseteq A_Y) dC(\alpha, \beta) \end{aligned}$$

If a p-box $[\underline{F}, \overline{F}]$ and one of its CDF F are continuous, then we know that:

$$\forall \alpha \in [0, 1], \overline{F}^{-1}(\alpha) \leqslant F^{-1}(\alpha) \leqslant \underline{F}^{-1}(\alpha)$$

Thus under hypothesis of continuity:

$$\forall \alpha, \beta \in [0, 1], \begin{cases} [\overline{F}_X^{-1}(\alpha), \underline{F}_X^{-1}(\alpha)] \subseteq A_X \\ [\overline{F}_Y^{-1}(\beta), \underline{F}_Y^{-1}(\beta)] \subseteq A_Y \end{cases} \implies \begin{cases} F_X^{-1}(\alpha) \in A_x \\ F_Y^{-1}(\beta) \in A_Y \end{cases}$$

which implies

$$\mathbb{I}([\overline{F}_X^{-1}(\alpha), \underline{F}_X^{-1}(\alpha)] \subseteq A_X \& [\overline{F}_Y^{-1}(\beta), \underline{F}_Y^{-1}(\beta)] \subseteq A_Y) \leqslant \mathbb{I}(F_X^{-1}(\alpha) \in A_X \& F_Y^{-1}(\beta) \in A_Y)$$

and thus

$$Bel_{XY}(A_X, A_Y) \leqslant P_{XY}(A_X, A_Y)$$