# A tail-sensitive generalization of Spearman's coefficient for upper-tail dependence analysis

Miguel A. Sordo

Department of Statistics and Operation Research,
University of Cádiz

Stochastic models in biomathematics and applications
20-21 January, 2026.
University of Salerno, Fisciano (SA), Italy

Cofinanciado por la Unión Europea

MINISTERIO DE CIENCIA, INNOVACIÓN Y UNIVERSIDADES

AGENCIA ESTATAL DE INVESTIGACIÓN

UCA Universidad de Cádiz

The project PID2024-157464NB-I00 is funded by MICIU/AEI /10.13039/501100011033 and European Union

# Contents

# Introduction

- Tail dependence describes how variables behave together when extreme events occur.
- Extreme events are usually rare, but their consequences can be very serious.
- Classical correlation measures describe average behavior and may underestimate tail dependence.
- Copula functions provide a flexible way to capture this kind of dependence.

**Relevant applications:**

- **Financial risk:** simultaneous market crashes, joint asset downturns.
- **Insurance:** one natural disaster can generate many large claims.
- **Meteorology:** extreme rainfall or temperatures in multiple locations.
- **Systemic risk:** propagation of shocks across economic sectors.
- **Biomathematics:** simultaneous epidemic outbreaks in connected regions.

| | |
|---|---|
| **Event** | Earthquake in Türkiye and Syria |
| **Year** | 2023 |
| **Type** | Natural: Seismic |
| **Country / Region** | Türkiye / Syria |
| **Number of affected (X)** | 18,000,000 |
| **Total economic damage, adjusted ('000 US$) (Y)** | 42,900,000 |
| **Deaths** | 56,683 |

The following questions will be studied using real disaster data obtained from the EM-DAT International Disaster Database (CRED/UCLouvain). After cleaning, the dataset comprises 10,159 records.

**Questions to discuss:**

- Is there dependence between the number of affected (X) and total economic damage (Y)?
- How does Y behave when X takes extremely high values?


Example: affected regions

### Copulas as a tool for modeling dependence

Let $(X, Y)$ be a random vector with marginal distribution functions $F$ and $G$, and joint distribution function $H$.

A refined approach to dependence focuses on the copula of the transformed variables $(F(X), G(Y))$. These transformations are uniformly distributed on $[0, 1]$.
The joint distribution of the transformed pair is given by:

$$C(u, v) = \Pr(F(X) \leq u, G(Y) \leq v), \quad 0 < u, v < 1,$$

which defines a copula $C$, a function that captures the dependence structure independently of the marginals. This separation allows for more flexible and robust modeling of dependence.

### Spearman's Rho and Kendall's Tau

Classical nonparametric measures, such as **Spearman's** $\rho$ and **Kendall's** $\tau$, are often preferred over Pearson's correlation. These measures:

- Always exist for continuous variables.
- Equal $\pm 1$ under perfect monotonicity.
- Vanish under independence.

**Spearman's** $\rho$**:**

$$\rho_{X,Y} = 12 \int_{\mathbb{R}^2} F(x)G(y)\, dH(x,y) - 3$$
$$= 12 \int_{[0,1]^2} uv dC(u,v) - 3$$

**Kendall's** $\tau$**:**

$$\tau_{X,Y} = 4 \int_{\mathbb{R}^2} H(x,y)\, dH(x,y) - 1$$
$$= 4 \int_{[0,1]^2} C(u,v)\, dC(u,v) - 1$$

# Estimators based on ranks

**Spearman's rho** estimator can be expressed as:

$$\hat{\rho}_{X,Y} = \frac{12}{m(m+1)(m-1)} \sum_{i=1}^{m} R_i S_i - 3 \cdot \frac{m+1}{m-1}$$

where $R_i$ is the rank of $X_i$ among $X_1, \ldots, X_m$, and $S_i$ is the rank of $Y_i$ among $Y_1, \ldots, Y_m$.

**Kendall's tau** estimator is given by:

$$\hat{\tau}_{X,Y} = \frac{2}{m(m-1)} \sum_{1 \le i < j \le m} \text{sign}(X_i - X_j) \cdot \text{sign}(Y_i - Y_j)$$

**Remark**

- *$R_i$ is the number of variables in the set $X_1, \ldots, X_m$ that are less or equal than $X_i$*
- *Ranks are maximally invariant statistics of the observations under monotone transformations of the marginal distributions.*

# Traditional measures ignore tail dependence

Spearman's $\rho$ and Kendall's $\tau$ summarize **global dependence**, but they do not distinguish between **upper** and **lower** tail dependence.

This can be a handicap when the goal is to capture asymmetric relationships — for example, strong agreement in the upper tail but not in the lower tail.

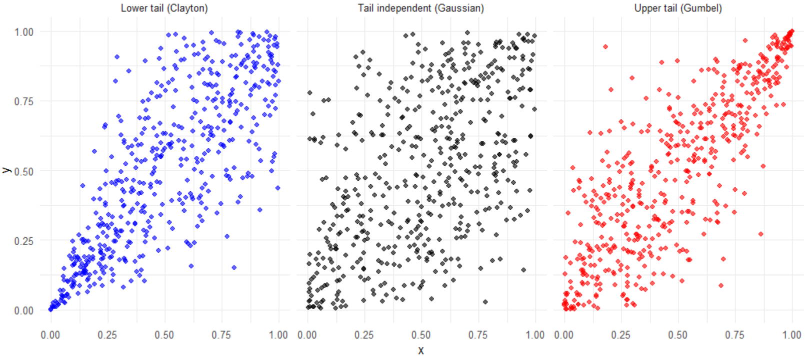



Figure: Three copulas with Spearman's $\rho = 0.6$

# Conditional Spearman's $\rho$

- **In this work, we aim to measure the impact of high values of $X$ on the behaviour of $Y$.**

- A conditional version of Spearmans'$\rho$

$$\rho_{X,Y|X>F^{-1}(p)}(u,v)$$

involves the copula of the conditional joint distribution function

$$P[X \leq x, Y \leq y | X > F^{-1}(p)]$$

which takes a complicated form.

- Its parametric and the non-parametric statistical inference are difficult (see Schmid and Schmidt, 2007).

- Blest (2000) proposed an alternative measure of rank correlation that places slightly more emphasis on agreement in the early part of a ranking:

$$\nu_n = \frac{2n+1}{n-1} - \frac{12}{n^2-n}\sum_{i=1}^{n}\left(1-\frac{R_i}{n+1}\right)^2 S_i,$$

  - It is still a measure of overall dependence.

- The upper tail dependence coefficient.
  For continuous random variables $X$ and $Y$ with marginal distribution functions $F$ and $G$, the upper tail-dependence coefficient $\lambda_U$ is defined as:

$$\begin{aligned}\lambda_U &= \lim_{p\to 1^-} \mathbb{P}\left(Y > G^{-1}(p) \mid X > F^{-1}(p)\right) \\ &= \lim_{p\to 1^-} \frac{1-2p+C(p,p)}{1-p}\end{aligned}$$

  - It only captures *asymptotic behavior*.
  - $\lambda_U$ is based on the limit as $(u,v)\to(1,1)$ along the diagonal $u=v$.

Building on this motivation:

- We introduce a flexible, Spearman-like, tail-sensitive dependence measure constructed from the ranks and concomitants of the data.
- In cases of strong tail dependence, the indices approach $\pm 1$, whereas under tail independence they remain near zero, regardless of the dependence structure elsewhere in the distribution.
- The indices are interpretable, retain the intuitive properties of Spearman's $\rho$, and possess a well-defined asymptotic distribution, enabling rigorous statistical inference.

# A copula functional for tail-sensitive dependence

Let $X_1, \dots, X_n$ be independent and identically distributed (i.i.d.) random variables with common distribution function $F$.

The $i$-th **order statistic**, denoted $X_{i:n}$, is the $i$-th smallest value among $X_1, \dots, X_n$. Its distribution function is given by:

$$F_{i:n}(x) = \beta_{i,n-i+1}(G(x)), \quad x \geq 0,$$

where

$$\beta_{i,j}(p) = \int_0^p \frac{(i+j-1)!}{(i-1)!(j-1)!}\, t^{i-1}(1-t)^{j-1}\, dt, \quad 0 \leq p \leq 1,$$

is the **incomplete beta function** with parameters $(i, j)$, as introduced by Pearson (1934).

If $F$ is absolutely continuous with density $f$, then the density of $X_{i:n}$ is:

$$f_{i:n}(x) = \frac{n!}{(i-1)!(n-i)!}\, [F(x)]^{i-1}\, [\overline{F}(x)]^{n-i}\, g(x), \quad x \geq 0.$$

- Let $(X, Y)$ be a random vector with joint distribution function $H(x, y)$, marginal distributions $F(x)$ and $G(y)$ and copula $C(u, v)$. Define the probability integral transforms:
$$U = F(X), \quad V = G(Y)$$
so that $U, V \sim \text{Uniform}(0, 1)$ and $(U, V) \sim C$.

- We denote the order statistics of $U$ by $U_{1:n} \leq U_{2:n} \leq \cdots \leq U_{n:n}$.

- For each $i = 1, \ldots, n$, the variable $V_{[i:n]}$ is defined as the $V$-component associated with $U_{i:n}$; it is called the **concomitant** of the $i$-th order statistic of $U$.

- Note that the concomitants are not necessarily ordered.

**Original values and concomitants of X**

| $X$ | $Y$ |
|---|---|
| 2.0 | 5.0 |
| 1.0 | 2.0 |
| 3.0 | 6.0 |
| 1.5 | 3.0 |
| 2.5 | 4.5 |

**Uniform transforms and concomitants of U**

| $U = F(X)$ | $V = G(Y)$ |
|---|---|
| 0.6 | 0.8 |
| 0.2 | 0.2 |
| 1.0 | 1.0 |
| 0.4 | 0.4 |
| 0.8 | 0.6 |

**Order statistics of U with concomitants V**

| $U_{i:5}$ | $V_{[i:5]}$ |
|---|---|
| 0.2 | 0.2 |
| 0.4 | 0.4 |
| 0.6 | 0.8 |
| 0.8 | 0.6 |
| 1.0 | 1.0 |

## A family of population parameters

For integers $n \geq 2$ and $1 \leq i \leq n-1$, define

$$T(i,n) = \frac{1}{n-i} \sum_{j=i+1}^{n} \mathbb{E}[V_{[j:n]}].$$

$T(i,n)$ is the expected value of the mean of the concomitants associated with the $n-i$ largest order statistics of $U$.

- The expectations $\mathbb{E}[V_{[j:n]}]$ are, in general, latent (not directly observable).

**A family of population parameters**

For integers $n \geq 2$ and $1 \leq i \leq n-1$, define

$$T(i,n) = \frac{1}{n-i} \sum_{j=i+1}^{n} \mathbb{E}[V_{[j:n]}].$$

$$\begin{aligned} T(n-1,n) &= \mathbb{E}[V_{[n:n]}] \\ T(n-2,n) &= \frac{1}{2}\,\mathbb{E}[V_{[n-1:n]} + V_{[n:n]}] \\ &\dots \\ T(1,n) &= \frac{1}{n-1}\,\mathbb{E}[V_{[2:n]} + \dots + V_{[n,n]}] \end{aligned}$$

- By changing $i$ while keeping $n$ fixed, we can look at different parts of the upper tail: a smaller $i$ includes more central values, while a larger $i$ focuses on extreme values of $X$.

**Theorem**

*Let $(X, Y)$ and $(X', Y')$ be two random vectors with copulas $C$ and $C'$ respectively. If $C(u, v) \leq C'(u, v)$ for all $u, v \in [0, 1]^2$ then*

$$T_{X,Y}(i, n) \leq T_{X',Y'}(i, n)$$

*for all $n \geq 2$ and $1 \leq i \leq n - 1$.*

# A family of rank-based tail dependence indices

**Definition**

Given $1 \leq i \leq n-1,\ n \geq 2$, we define:

$$\begin{aligned}\Phi(i,n) &= \frac{n+1}{i}\left(2T(i,n)-1\right)\\ &= \frac{2n(n+1)}{i(n-i)}\int_{[0,1]^2} C(u,v)\, d\beta_{i,n-i}(u)\, dv \ - \frac{n+1}{n-i}.\end{aligned}$$

Two notable special cases are:

- $\Phi(1,2) = 12\int_{[0,1]^2} uv\, dC(u,v) - 3$ (Spearman's $\rho$),
- $\Phi(2,3) = 2 - 12\int_{[0,1]^2} (1-u)^2\, v\, dC(u,v)$ (Blest's coefficient).

# Properties of $\Phi(i,n)$

$$\Phi(i,n) = \frac{2n(n+1)}{i(n-i)} \int_{[0,1]^2} C(u,v)\, d\beta_{i,n-i}(u)\, dv \;-\; \frac{n+1}{n-i}$$

The following theorem provides an alternative representation of the dependence index.

**Theorem**

*For integers $n \geq 2$ and $1 \leq i \leq n-1$, the tail-sensitive index $\Phi(i,n)$ can be expressed as*

$$\Phi(i,n) = \frac{\text{Cov}\left(V, \beta_{i:n-i}(U)\right)}{\text{Cov}\left(V, \beta_{i:n-i}(V)\right)}.$$

*where $U = F(X)$ and $V = G(Y)$.*

# Properties of $\Phi(i, n)$

- The index $\Phi(i, n)$ is well-defined for any pair of continuous random variables $(X, Y)$.
- It is bounded: $-1 \leq \Phi(i, n) \leq 1 \quad$ for all $(X, Y)$, $\; n \geq 2$, $\; 1 \leq i \leq n-1$.
- **Comonotonicity:**

  If any of the two following equivalent conditions hold:

  - $X = f(Z)$, $Y = g(Z)$ for some random variable $Z$ and strictly increasing functions $f$, $g$.
  - The copula of $(X, Y)$ is $C(u, v) = \min(u, v)$,

  then

$$\Phi(i, n) = 1, \quad n \geq 2, \quad 1 \leq i \leq n-1.$$

- **Counter-monotonicity:**

  If any of the two following equivalent conditions hold:

  - $X = f(Z)$, $Y = g(Z)$ with $f$ increasing and $g$ decreasing,
  - The copula of $(X, Y)$ is $C(u, v) = \max(u + v - 1, 0)$,

  then:

  $$\Phi(i, n) = -1, \quad n \geq 2, \quad 1 \leq i \leq n - 1.$$

- **Independence:**

  If $X \perp Y$ (that is, the copula of $(X, Y)$ is $C(u, v) = uv$), then

  $$\Phi(i, n) = 0, \quad n \geq 2, \quad 1 \leq i \leq n - 1.$$

# Properties of $\Phi(i, n)$

- **Positive Quadrant Dependence (PQD):** If $(X, Y)$ is positively quadrant dependent, that is, $C(u, v) > uv$ for all $u, v \in [0, 1]^2$, then:

$$\Phi(i, n) \geq 0.$$

- **Negative Quadrant Dependence (NQD):** If $(X, Y)$ is negatively quadrant dependent, that is, $C(u, v) < uv$ for all $u, v \in [0, 1]^2$, then:

$$\Phi(i, n) \leq 0.$$

- **Asymmetry:** In general,

$$\Phi_{X,Y}(i, n) \neq \Phi_{Y,X}(i, n).$$

- **Exchangeable Copula Case:** If the copula $C(u, v)$ is exchangeable, i.e., $C(u, v) = C(v, u)$, then:

$$\Phi_{X,Y}(i, n) = \Phi_{Y,X}(i, n).$$

# Consistency with concordance order

$$\Phi(i,n) = \frac{\mathrm{Cov}\left(U, \beta_{i:n-i}(V)\right)}{\mathrm{Cov}\left(U, \beta_{i:n-i}(U)\right)}, \quad 1 \leq i \leq n-1, \quad U = F(X), V = G(Y),$$

**Theorem**

*Let $(X, Y)$ and $(X', Y')$ be two random vectors with copulas $C$ and $C'$ respectively. If $C(u,v) \leq C'(u,v)$ for all $u, v \in [0,1]^2$ then*

$$\Phi_{X,Y}(i,n) \leq \Phi_{X',Y'}(i,n)$$

*for all $n \geq 2$ and $1 \leq i \leq n-1$.*

- $\Phi(i,n)$ is not a measure of concordance in the sense of Scarsini (1984). This is by design, as our focus is to interpret $X$ as a function of $Y$ in the upper tail of $Y$, rather than merely assessing whether one variable is a function of the other.

- A symmetrized version of $\Phi_{X,Y}(i,n)$ is given by

$$\frac{1}{n-i}\sum_{j=i+1}^{n}\mathbb{E}\left(U_{[j:n]}+V_{[j:n]}\right),$$

which can be interpreted as both a measure of concordance and a measure of dependence in both tails.

# Example: FGM Copula

**Example**

Let $(X, Y)$ be a random vector whose copula $C_\theta(u, v)$ belongs to the Farlie–Gumbel–Morgenstern (FGM) family, with parameter $\theta \in [-1, 1]$. Then,

$$C_\theta(u, v) = uv + \theta\, uv(1-u)(1-v).$$

It can be shown that:

$$\Phi(i, n) = \frac{\theta}{3} \quad \text{for all} \quad 1 \leq i \leq n-1, \quad n \geq 2.$$

This is consistent with the fact that the FGM copula is not able to model strong tail dependence.

**Theorem**

*Under the assumption of random sampling from a continuous distribution function H with underlying copula C, the index*

$$\hat{\Phi}_m(i,n) = \frac{2n(n+1)}{i(n-i)} \cdot \frac{1}{m} \sum_{j=1}^{m} \left( \frac{S_j}{m+1} \right) \beta_{i,n-i} \left( \frac{R_j}{m+1} \right) - \frac{n+1}{i}$$

*satisfies* $\hat{\Phi}_m(i,n) \xrightarrow[m\to\infty]{a.s.} \Phi(i,n)$. *Moreover,* $\sqrt{m} \cdot (\hat{\Phi}_m(i,n) - \Phi(i,n))$ *converges in distribution, as* $m \to \infty$, *to a normal random variable with zero mean and the same variance as*

$$\frac{2n(n+1)}{i(n-i)} \left( V\beta_{i,n-i}(U) + \int_0^1 \int_0^1 (\mathbf{1}(V \le v) - v)\, \beta_{i,n-i}(u)\, dC(u,v) + \int_0^1 \int_0^1 (\mathbf{1}(U \le u) - u) v \beta'_{i,n-i}(u)\, dC(u,v) \right)$$

*where the pair* $(U, V)$ *is distributed as* $C$.

In particular, the variance of the latter expression equals

$$\frac{n^2(n+1)^2}{3i^2(n-i)^2}\left(\sum_{k=i}^{n-1}\sum_{\ell=i}^{n-1}\binom{n-1}{k}\binom{n-1}{\ell}\frac{(k+\ell)!\,(2n-2-k-\ell)!}{(2n-1)!}-\left(1-\frac{i}{n}\right)^2\right)$$

when $U$ and $V$ are independent.

| $i$ | $n=2$ | $n=3$ | $n=4$ | $n=5$ | $n=6$ | $n=7$ | $n=8$ | $n=9$ | $n=10$ |
|---|---|---|---|---|---|---|---|---|---|
| 1 | 1.0000 | 1.0667 | 1.1905 | 1.3333 | 1.4848 | 1.6410 | 1.8000 | 1.9608 | 2.1228 |
| 2 | – | 1.0667 | 1.0119 | 1.0476 | 1.1136 | 1.1935 | 1.2808 | 1.3725 | 1.4672 |
| 3 | – | – | 1.1905 | 1.0476 | 1.0292 | 1.0526 | 1.0951 | 1.1480 | 1.2071 |
| 4 | – | – | – | 1.3333 | 1.1136 | 1.0526 | 1.0450 | 1.0620 | 1.0922 |
| 5 | – | – | – | – | 1.4848 | 1.1935 | 1.0951 | 1.0620 | 1.0588 |
| 6 | – | – | – | – | – | 1.4848 | 1.1935 | 1.0951 | 1.0588 |
| 7 | – | – | – | – | – | – | 1.4848 | 1.1935 | 1.0951 |
| 8 | – | – | – | – | – | – | – | 1.4848 | 1.1935 |
| 9 | – | – | – | – | – | – | – | – | 1.4848 |

Table: Values of the variance for some values of $i$ and $n$ when $U$ and $V$ are independent.

# Dependence index under a mixture of Gumbel and upper comonotonic copula

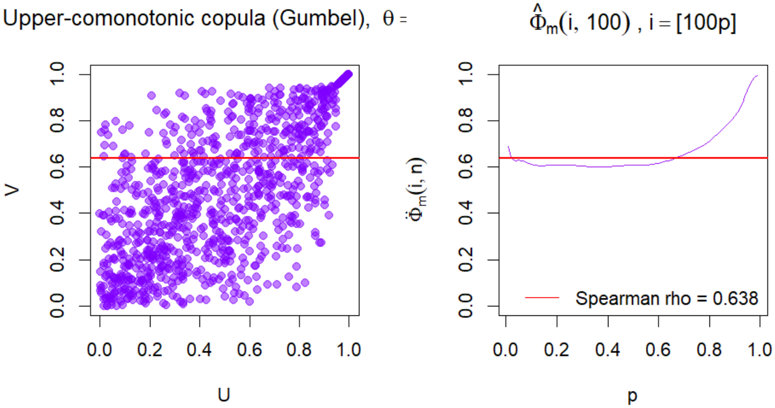
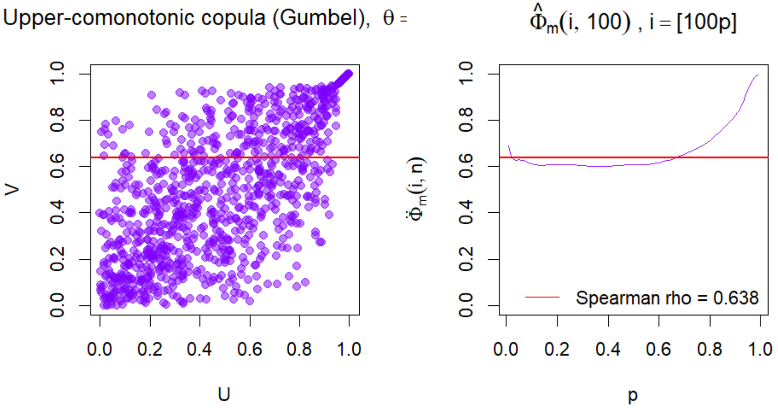



Figure: Dependence index values for a Gumbel($\theta = 2$)-upper comonotonic ($\beta = 0.95$) copula mixture under varying parameters ($n = 100, i = [100p]$). Sample size $m = 1000$




| $\rho_S$ | $\tau_K$ | $\hat{\Phi}_m(2,3)$ | $\hat{\Phi}_m(3,4)$ | $\hat{\Phi}_m(4,5)$ | $\hat{\Phi}_m(9,10)$ | $\hat{\Phi}_m(18,20)$ | $\hat{\Phi}_m(90,100)$ |
|---|---|---|---|---|---|---|---|
| 0.638 | 0.463 | 0.656 | 0.676 | 0.695 | 0.770 | 0.793 | 0.812 |

# Limiting form of tail dependence indices

**Theorem**

*Let $(X, Y)$ be a random vector with continuous marginal distribution functions $F$ and $G$, respectively, and set $(U, V) = (F(X), G(Y))$. Fix $p \in (0, 1)$ and let $i = \lfloor np \rfloor$, where $\lfloor \cdot \rfloor$ denotes the integer part. Then, as $n \to \infty$,*

$$\Phi(i, n) \longrightarrow \frac{1}{1-p}\left(\frac{2}{p}\int_0^1 C(p, v)\, dv \; - \; 1\right).$$

We denote the limiting form of the dependence index by

$$\begin{aligned} \Phi^*(p) \quad &:= \quad \frac{1}{1-p}\left(\frac{2}{p}\int_0^1 C(p, v)\, dv \; - \; 1\right), \\ &= \quad \frac{2}{p}\,\mathbb{E}[V \mid U > p] - \frac{1}{p}, \qquad p \in (0, 1). \end{aligned}$$

Under the assumption, $i/n \to p$ . The closer $p$ is to one, the more the index emphasizes the dependence between the largest values of $X$ and the corresponding values of $Y$.

**Remark**

- *For the FGM copula, we have* $\Phi^*(p) = \frac{\theta}{3}$ *for all* $p \in (0,1)$*, which coincides with* $\Phi(i,n)$ *for all* $i$ *and* $n$ *(in particular, it coincides with the Spearman's rank correlation coefficient).*

- $\Phi^*(p)$ *inherits all fundamental properties of* $\Phi(i,n)$ *and, in addition, satisfies further desirable properties.*

**Theorem**

*(a) If* $(X, Y)$ *is an upper comonotonic random vector, then there exists* $p_0 \in (0,1)$ *such that* $\Phi^*(p) = 1$ *for all* $p \in [p_0, 1)$.

*(b) If* $(X, -Y)$ *is upper comonotonic random vector, then there exists* $p_0 \in (0,1)$ *such that* $\Phi^*(p) = -1$ *for all* $p \in [p_0, 1)$.

# $\Phi^*(p)$ under tail-independence

**Definition (Tail independence in the tail of $X$)**

We say that a random vector $(X, Y)$ is tail-independent in the upper tail of $X$ if there exists $x_0 \in \mathbb{R}$ such that

$$P[X > x, Y > y] = P[X > x]P[Y > y], \quad \text{for all } \; x > x_0, y \in \mathbb{R}.$$

If the joint distribution has marginal continuous and a copula C, this is equivalent to say that there exists $\beta \in (0,1)$ such that

$$C(u, v) = uv, \quad \text{for all } u \in (0,1), \; v \in (\beta, 1).$$

- If $(X, Y)$ is tail-independent in the upper tail of $X$, then there exists $\beta \in (0,1)$ such that

$$\Phi^*(p) = 0, \; \text{for all } p \in (\beta, 1]$$

**Theorem**

*Let $p \in (0,1)$. Suppose that a random sample $(X_1, Y_1), \ldots, (X_n, Y_n)$ has been drawn from a bivariate distribution with absolutely continuous copula $C(u,v)$. Suppose that the density copula $c(u,v)$ is bounded in $(0,1) \times O_p$, where $O_p$ is an open set containing $p$. Consider the index*

$$\hat{\Phi}^*_m(p) = \frac{2}{p(1-p)} \cdot \frac{1}{m} \sum_{j=1}^{m} \left( \frac{S_j}{m+1} \right) \cdot \mathbf{1} \left[ \frac{R_j}{m+1} > p \right] - \frac{1}{p}.$$

*Then, $\sqrt{m} \cdot (\hat{\Phi}^*_m(p) - \Phi^*(p))$ converges in distribution, as $m \to \infty$, to a normal random variable with zero mean and the same variance as*

$$\frac{2}{p(1-p)} \Bigg[ V\mathbf{1}[U > p] + \int_0^1 (\mathbf{1}(V \leq v) - v)\, E[\mathbf{1}(U > p) | V = v]\, dv$$
$$+ \quad (\mathbf{1}(U \leq p) - p)\, E[V | U = p] \Bigg]$$

*where the pair $(U, V)$ is distributed as $C$.*

In particular, the variance of the latter expression is given by

$$\frac{1}{3p(1-p)}$$

when the copula $C$ exhibits *upper-tail independence* in $U$; that is, there exists some $p_0 \in (0,1)$ such that

$$C(u,v) = uv \quad \text{for all } u \in (p_0, 1),\ v \in (0,1).$$

# Numerical illustrations and simulation results

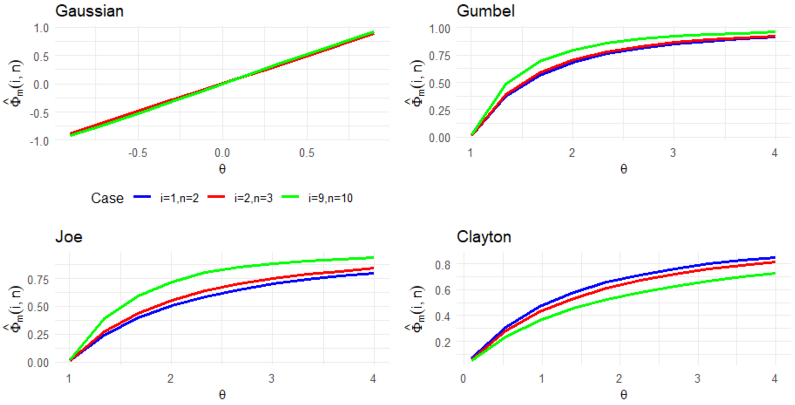



Figure: Estimated values of $\Phi_m(i,n)$ for four copulas as a function of $\theta$, shown for $i=1, n=2$ (Spearman's coefficient), $i=2, n=3$ (Blest's index) and $i=9, n=10$.

| $\theta$ | $m$ | $\Phi_{X,Y}(4,5)$ | $\overline{\hat{\Phi}}_m(4,5)$ | $\mathrm{Bias}_m(4,5)$ | $\hat{\sigma}^2_m(4,5)$ | $\mathrm{MSE}_m(4,5)$ | $m\hat{\sigma}^2_m(4,5)$ | Coverage | $p_{\text{norm}}$ |
|---|---|---|---|---|---|---|---|---|---|
| -0.9 | 100 | -0.3000 | -0.3115 | -0.0115 | 0.0099 | 0.0100 | 0.9860 | 0.9495 | 0.0004 |
| -0.9 | 500 | -0.3000 | -0.3036 | -0.0036 | 0.0021 | 0.0021 | 1.0568 | 0.9450 | 0.0208 |
| -0.9 | 1000 | -0.3000 | -0.3017 | -0.0017 | 0.0010 | 0.0010 | 1.0443 | 0.9530 | 0.0426 |
| -0.9 | 5000 | -0.3000 | -0.2998 | 0.0002 | 0.0002 | 0.0002 | 1.0632 | 0.9475 | 0.0994 |
| -0.5 | 100 | -0.1667 | -0.1820 | -0.0154 | 0.0118 | 0.0120 | 1.1761 | 0.9460 | 0.0011 |
| -0.5 | 500 | -0.1667 | -0.1697 | -0.0030 | 0.0025 | 0.0025 | 1.2510 | 0.9435 | 0.7389 |
| -0.5 | 1000 | -0.1667 | -0.1682 | -0.0015 | 0.0012 | 0.0012 | 1.1820 | 0.9535 | 0.4800 |
| -0.5 | 5000 | -0.1667 | -0.1673 | -0.0006 | 0.0003 | 0.0003 | 1.3248 | 0.9535 | 0.1204 |
| 0.0 | 100 | 0.0000 | -0.0247 | -0.0247 | 0.0124 | 0.0130 | 1.2352 | 0.9505 | 0.8406 |
| 0.0 | 500 | 0.0000 | -0.0049 | -0.0049 | 0.0027 | 0.0027 | 1.3587 | 0.9495 | 0.4113 |
| 0.0 | 1000 | 0.0000 | -0.0022 | -0.0022 | 0.0013 | 0.0013 | 1.2562 | 0.9470 | 0.6665 |
| 0.0 | 5000 | 0.0000 | -0.0008 | -0.0008 | 0.0003 | 0.0003 | 1.3207 | 0.9470 | 0.0878 |
| 0.5 | 100 | 0.1667 | 0.1400 | -0.0267 | 0.0122 | 0.0129 | 1.2191 | 0.9425 | 0.6043 |
| 0.5 | 500 | 0.1667 | 0.1620 | -0.0047 | 0.0025 | 0.0025 | 1.2263 | 0.9410 | 0.2263 |
| 0.5 | 1000 | 0.1667 | 0.1641 | -0.0026 | 0.0012 | 0.0012 | 1.1572 | 0.9540 | 0.8753 |
| 0.5 | 5000 | 0.1667 | 0.1666 | -0.0001 | 0.0003 | 0.0003 | 1.2885 | 0.9445 | 0.6366 |
| 0.9 | 100 | 0.3000 | 0.2682 | -0.0318 | 0.0103 | 0.0114 | 1.0348 | 0.9350 | 0.0000 |
| 0.9 | 500 | 0.3000 | 0.2944 | -0.0056 | 0.0021 | 0.0021 | 1.0398 | 0.9485 | 0.7629 |
| 0.9 | 1000 | 0.3000 | 0.2967 | -0.0033 | 0.0011 | 0.0011 | 1.0734 | 0.9515 | 0.0811 |
| 0.9 | 5000 | 0.3000 | 0.3000 | 0.0000 | 0.0002 | 0.0002 | 1.0876 | 0.9420 | 0.8783 |

Table: Monte Carlo results for $\hat{\Phi}_m(4,5)$, $N = 2000$ replications.

# Simulations for the family $\Phi(i, n)$

| $\theta$ | $m$ | $\Phi_{X,Y}(9,10)$ | $\overline{\hat{\Phi}}_m(9,10)$ | $\mathrm{Bias}_m(9,10)$ | $\hat{\sigma}^2_m(9,10)$ | $\mathrm{MSE}_m(9,10)$ | $m\hat{\sigma}^2_m(9,10)$ | Coverage | $p_{\text{norm}}$ |
|---|---|---|---|---|---|---|---|---|---|
| -0.9 | 100 | -0.3000 | -0.3300 | -0.0300 | 0.0146 | 0.0155 | 1.4603 | 0.9475 | 0.0021 |
| -0.9 | 500 | -0.3000 | -0.3075 | -0.0075 | 0.0032 | 0.0033 | 1.6245 | 0.9505 | 0.0269 |
| -0.9 | 1000 | -0.3000 | -0.3039 | -0.0039 | 0.0016 | 0.0016 | 1.6066 | 0.9550 | 0.1525 |
| -0.9 | 5000 | -0.3000 | -0.3001 | -0.0001 | 0.0003 | 0.0003 | 1.6571 | 0.9465 | 0.7872 |
| -0.5 | 100 | -0.1667 | -0.2035 | -0.0369 | 0.0180 | 0.0194 | 1.8045 | 0.9465 | 0.0006 |
| -0.5 | 500 | -0.1667 | -0.1734 | -0.0067 | 0.0039 | 0.0040 | 1.9741 | 0.9505 | 0.9162 |
| -0.5 | 1000 | -0.1667 | -0.1708 | -0.0042 | 0.0018 | 0.0019 | 1.8380 | 0.9540 | 0.3514 |
| -0.5 | 5000 | -0.1667 | -0.1676 | -0.0010 | 0.0004 | 0.0004 | 2.0526 | 0.9520 | 0.4052 |
| 0.0 | 100 | 0.0000 | -0.0487 | -0.0487 | 0.0184 | 0.0208 | 1.8409 | 0.9340 | 0.5243 |
| 0.0 | 500 | 0.0000 | -0.0093 | -0.0093 | 0.0043 | 0.0044 | 2.1594 | 0.9455 | 0.3831 |
| 0.0 | 1000 | 0.0000 | -0.0043 | -0.0043 | 0.0020 | 0.0020 | 1.9773 | 0.9455 | 0.5882 |
| 0.0 | 5000 | 0.0000 | -0.0016 | -0.0016 | 0.0004 | 0.0004 | 2.1400 | 0.9460 | 0.0676 |
| 0.5 | 100 | 0.1667 | 0.1112 | -0.0555 | 0.0185 | 0.0215 | 1.8469 | 0.9320 | 0.1398 |
| 0.5 | 500 | 0.1667 | 0.1567 | -0.0100 | 0.0039 | 0.0040 | 1.9579 | 0.9455 | 0.5113 |
| 0.5 | 1000 | 0.1667 | 0.1615 | -0.0052 | 0.0018 | 0.0018 | 1.8097 | 0.9505 | 0.3220 |
| 0.5 | 5000 | 0.1667 | 0.1662 | -0.0005 | 0.0004 | 0.0004 | 2.0056 | 0.9480 | 0.9468 |
| 0.9 | 100 | 0.3000 | 0.2352 | -0.0648 | 0.0154 | 0.0196 | 1.5397 | 0.9200 | 0.0000 |
| 0.9 | 500 | 0.3000 | 0.2877 | -0.0123 | 0.0031 | 0.0033 | 1.5689 | 0.9505 | 0.5107 |
| 0.9 | 1000 | 0.3000 | 0.2933 | -0.0067 | 0.0017 | 0.0017 | 1.6887 | 0.9500 | 0.4084 |
| 0.9 | 5000 | 0.3000 | 0.2993 | -0.0007 | 0.0003 | 0.0003 | 1.6716 | 0.9465 | 0.1535 |

Table: Monte Carlo results for $\hat{\Phi}_m(9, 10)$, $N = 2000$ replications.

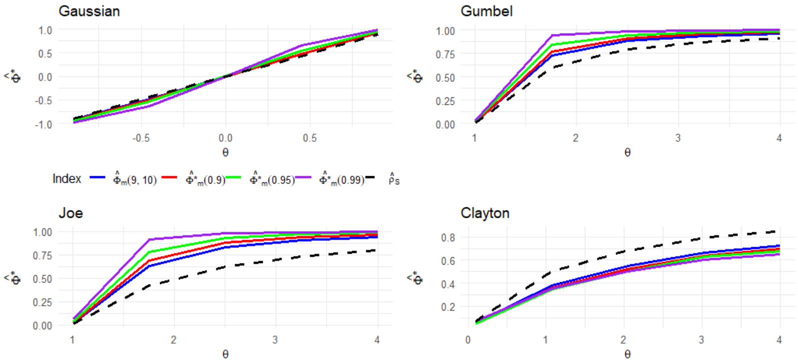



Figure: Estimated values of $\Phi^*(p)$, $\Phi(9, 10)$, and Spearman's $\rho$ for four copulas as a function of $\theta$, shown for $p = 0.9, 0.95, 0.99$.

# Simulations for the family $\Phi^*(p)$

| $\theta$ | $m$ | $\Phi^*_{X,Y}(0.9)$ | $\overline{\hat{\Phi}^*_m}(0.9)$ | $\mathrm{Bias}_m(0.9)$ | $\hat{\sigma}^2_m(0.9)$ | $\mathrm{MSE}_m(0.9)$ | $m\hat{\sigma}^2_m(0.9)$ | Coverage | $p_{\text{norm}}$ |
|---|---|---|---|---|---|---|---|---|---|
| -0.9 | 100 | -0.3000 | -0.2914 | 0.0086 | 0.0282 | 0.0282 | 2.8158 | 0.9430 | 0.0385 |
| -0.9 | 500 | -0.3000 | -0.3001 | -0.0001 | 0.0059 | 0.0059 | 2.9295 | 0.9500 | 0.0229 |
| -0.9 | 1000 | -0.3000 | -0.2999 | 0.0001 | 0.0029 | 0.0029 | 2.9009 | 0.9505 | 0.3020 |
| -0.9 | 5000 | -0.3000 | -0.2991 | 0.0009 | 0.0006 | 0.0006 | 2.9537 | 0.9500 | 0.5046 |
| -0.5 | 100 | -0.1667 | -0.1611 | 0.0055 | 0.0349 | 0.0349 | 3.4927 | 0.9520 | 0.0012 |
| -0.5 | 500 | -0.1667 | -0.1654 | 0.0012 | 0.0072 | 0.0072 | 3.5805 | 0.9445 | 0.2015 |
| -0.5 | 1000 | -0.1667 | -0.1667 | 0.0000 | 0.0033 | 0.0033 | 3.3027 | 0.9485 | 0.3066 |
| -0.5 | 5000 | -0.1667 | -0.1665 | 0.0002 | 0.0007 | 0.0007 | 3.4869 | 0.9470 | 0.9532 |
| 0.0 | 100 | 0.0000 | 0.0016 | 0.0016 | 0.0361 | 0.0361 | 3.6079 | 0.9570 | 0.2080 |
| 0.0 | 500 | 0.0000 | 0.0021 | 0.0021 | 0.0076 | 0.0076 | 3.7945 | 0.9490 | 0.8054 |
| 0.0 | 1000 | 0.0000 | 0.0008 | 0.0008 | 0.0036 | 0.0036 | 3.5868 | 0.9505 | 0.0237 |
| 0.0 | 5000 | 0.0000 | -0.0009 | -0.0009 | 0.0008 | 0.0008 | 3.8914 | 0.9475 | 0.3490 |
| 0.5 | 100 | 0.1667 | 0.1674 | 0.0007 | 0.0349 | 0.0349 | 3.4893 | 0.9450 | 0.1947 |
| 0.5 | 500 | 0.1667 | 0.1693 | 0.0026 | 0.0068 | 0.0068 | 3.3916 | 0.9510 | 0.2432 |
| 0.5 | 1000 | 0.1667 | 0.1673 | 0.0007 | 0.0033 | 0.0033 | 3.2588 | 0.9490 | 0.0471 |
| 0.5 | 5000 | 0.1667 | 0.1669 | 0.0003 | 0.0007 | 0.0007 | 3.6902 | 0.9505 | 0.7538 |
| 0.9 | 100 | 0.3000 | 0.2924 | -0.0076 | 0.0300 | 0.0300 | 2.9968 | 0.9485 | 0.0042 |
| 0.9 | 500 | 0.3000 | 0.3001 | 0.0001 | 0.0057 | 0.0057 | 2.8313 | 0.9530 | 0.7340 |
| 0.9 | 1000 | 0.3000 | 0.3001 | 0.0001 | 0.0030 | 0.0030 | 3.0070 | 0.9465 | 0.2269 |
| 0.9 | 5000 | 0.3000 | 0.3003 | 0.0003 | 0.0006 | 0.0006 | 3.0010 | 0.9540 | 0.6308 |

Table: Monte Carlo results for the new estimator $\hat{\Phi}^*_m(0.9)$, $N = 2000$ replications.

# Simulations for the family $\Phi^*(p)$

| $\theta$ | $m$ | $\Phi^*_{X,Y}(0.95)$ | $\overline{\hat{\Phi}^*_m}(0.95)$ | $\text{Bias}_m(0.95)$ | $\hat{\sigma}^2_m(0.95)$ | $\text{MSE}_m(0.95)$ | $m\hat{\sigma}^2_m(0.95)$ | Coverage | $p_{\text{norm}}$ |
|---|---|---|---|---|---|---|---|---|---|
| -0.9 | 100 | -0.3000 | -0.2914 | 0.0086 | 0.0368 | 0.0369 | 3.6796 | 0.9445 | 0.0385 |
| -0.9 | 500 | -0.3000 | -0.3000 | 0.0000 | 0.0074 | 0.0074 | 3.7070 | 0.9505 | 0.0229 |
| -0.9 | 1000 | -0.3000 | -0.2999 | 0.0001 | 0.0037 | 0.0037 | 3.7185 | 0.9500 | 0.3020 |
| -0.9 | 5000 | -0.3000 | -0.2995 | 0.0005 | 0.0007 | 0.0007 | 3.5003 | 0.9500 | 0.5046 |
| -0.5 | 100 | -0.1667 | -0.1611 | 0.0056 | 0.0436 | 0.0438 | 4.3640 | 0.9510 | 0.0012 |
| -0.5 | 500 | -0.1667 | -0.1654 | 0.0013 | 0.0087 | 0.0087 | 4.3568 | 0.9445 | 0.2015 |
| -0.5 | 1000 | -0.1667 | -0.1667 | 0.0000 | 0.0043 | 0.0043 | 4.2870 | 0.9485 | 0.3066 |
| -0.5 | 5000 | -0.1667 | -0.1665 | 0.0002 | 0.0009 | 0.0009 | 4.4652 | 0.9470 | 0.9532 |
| 0.0 | 100 | 0.0000 | 0.0016 | 0.0016 | 0.0435 | 0.0435 | 4.3478 | 0.9570 | 0.2080 |
| 0.0 | 500 | 0.0000 | 0.0021 | 0.0021 | 0.0091 | 0.0091 | 4.5450 | 0.9490 | 0.8054 |
| 0.0 | 1000 | 0.0000 | 0.0008 | 0.0008 | 0.0046 | 0.0046 | 4.5956 | 0.9505 | 0.0237 |
| 0.0 | 5000 | 0.0000 | -0.0009 | -0.0009 | 0.0010 | 0.0010 | 5.0000 | 0.9475 | 0.3490 |
| 0.5 | 100 | 0.1667 | 0.1674 | 0.0007 | 0.0435 | 0.0435 | 4.3478 | 0.9450 | 0.1947 |
| 0.5 | 500 | 0.1667 | 0.1693 | 0.0026 | 0.0085 | 0.0085 | 4.2500 | 0.9510 | 0.2432 |
| 0.5 | 1000 | 0.1667 | 0.1673 | 0.0007 | 0.0043 | 0.0043 | 4.2870 | 0.9490 | 0.0471 |
| 0.5 | 5000 | 0.1667 | 0.1669 | 0.0002 | 0.0009 | 0.0009 | 4.4652 | 0.9505 | 0.7538 |
| 0.9 | 100 | 0.3000 | 0.2924 | -0.0076 | 0.0368 | 0.0369 | 3.6796 | 0.9485 | 0.0042 |
| 0.9 | 500 | 0.3000 | 0.3001 | 0.0001 | 0.0074 | 0.0074 | 3.7070 | 0.9530 | 0.7340 |
| 0.9 | 1000 | 0.3000 | 0.3001 | 0.0001 | 0.0037 | 0.0037 | 3.7185 | 0.9465 | 0.2269 |
| 0.9 | 5000 | 0.3000 | 0.3003 | 0.0003 | 0.0007 | 0.0007 | 3.5003 | 0.9540 | 0.6308 |

Table: Monte Carlo results for the new estimator $\hat{\Phi}^*_m(0.95)$, $N = 2000$ replications.

- We illustrate the behavior of the proposed dependence indices using real disaster data obtained from the EM-DAT International Disaster Database (CRED/UCLouvain).
- The database records the occurrence and impacts of natural and technological disasters worldwide.
- For this study, we extract all natural disasters occurring globally between 2000 and 2025.
- We focus on two key impact variables:

$$X = \text{No. Affected}, \qquad Y = \text{Total Damage, Adjusted ('000 US\$)},$$

  where the latter corresponds to economic losses adjusted for inflation using the Consumer Price Index. The variable *No. Affected* reflects the number of individuals requiring immediate assistance as a direct consequence of the disaster.
- These quantities jointly describe human and economic impacts, allowing for a meaningful analysis of tail dependence.

# Application to real data

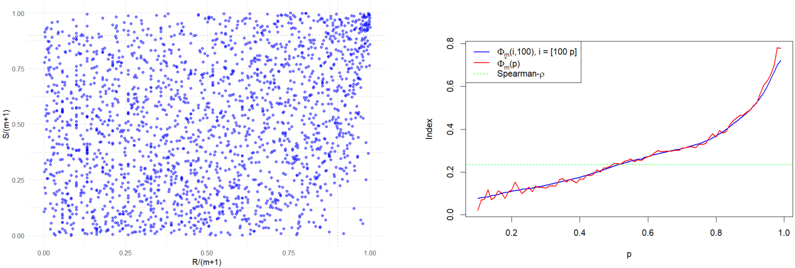
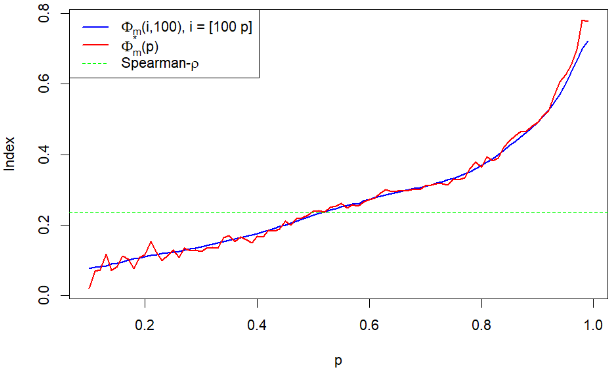



Figure: Visualization of disaster dependence: the empirical copula highlights upper-tail dependence, while the curves of the proposed indices $\Phi_m(i, n)$ (with $i = \lfloor np \rfloor$, $n = 100$) and $\Phi^*_m(p)$ illustrate tail behavior across percentiles.

# Application to real data

$$\hat{\Phi}_m(i,n) = \frac{2n(n+1)}{i(n-i)} \cdot \frac{1}{m} \sum_{j=1}^{m} \left( \frac{S_j}{m+1} \right) \beta_{i,n-i} \left( \frac{R_j}{m+1} \right) - \frac{n+1}{i}$$

$$\hat{\Phi}^*_m(p) = \frac{2}{p(1-p)} \cdot \frac{1}{m} \sum_{j=1}^{m} \left( \frac{S_j}{m+1} \right) \mathbf{1} \left[ \frac{R_j}{m+1} > p \right] - \frac{1}{p}.$$



# Conclusions

# Conclusions

- The dependence measures $\Phi(i, n)$ and $\Phi^*(p)$ capture complementary aspects of the joint distribution of $(X, Y)$.
- $\Phi(i, n)$ provides a **finite-sample, global measure**, robust across the distribution and related to classical indices like Spearman's $\rho$ and Blest's coefficient.
- $\Phi^*(p)$ focuses on **upper-tail dependence**, offering a probability-based, interpretable measure for extreme-event analysis.
- Appropriate choices of $i$, $n$, and $p$ allow **evaluation of both global and tail-specific patterns**, with $\Phi^*(p)$ emphasizing extremes as $p \to 1$.
- Simulation studies confirm:
  - $\hat{\Phi}_m(i, n)$ is robust for finite samples.
  - $\hat{\Phi}^*_m(p)$ accurately captures tail dependence for large samples, with variance scaling and coverage behaving as expected.
- Together, the two families provide a **coherent, theoretically grounded, and practically implementable framework** for assessing global and tail dependence.

- Lehmann, E. L. (1975). Nonparametrics: Statistical Methods Based on Ranks. Holden-Day, San Francisco.
- Blest, D.C. (2000). Rank Correlation: An Alternative Measure. *Australian and New Zealand Journal of Statistics*, 42(1), 101–111.
- Genest, C., Plante, J.F. (2003). On Blest's Measure of Rank Correlation. *The Canadian Journal of Statistics / La Revue Canadienne de Statistique*, 31(1), 35–52.
- Schmid, F., Schmidt, R. (2007). Multivariate conditional versions of Spearman's rho and related measures of tail dependence. *Journal of Multivariate Analysis*, 98(6), 1123–1140.
- Quesada-Molina, J.J. (1992). A generalization of an identity of Hoeffding and some applications. *Journal of the Italian Statistical Society*, 1, 405–411.
- Navarro, J., Pellerey, F., Sordo, M.A. (2021). Weak Dependence Notions and Their Mutual Relationships. *Mathematics*, 9(1), 81.
- Arnold, B.C., Balakrishnan, N., Nagaraja, H.N. (2008). Order Statistics, 2nd edition. Wiley-Interscience, Hoboken, NJ.
- Ruymgaart, F.H., Shorack, G.R., van Zwet, W.R. (1972). Asymptotic Normality of Nonparametric Tests for Independence. *Annals of Mathematical Statistics*, 43(4), 1122–1135.

- Hanbali, H., Linders, D. (2022). Monotone tail functions: definitions, properties, and application to risk-reducing strategies. *Journal of Computational and Applied Mathematics*, 416, 114484.
- Ruymgaart, F.H. (1974). Asymptotic normality of nonparametric tests for independence. *Annals of Statistics*, 2(5), 892–910.
- Yang, S.S. (1977). General distribution theory of the concomitants of order statistics. *The Annals of Statistics*, 5(5), 996–1002.
- Delforge, D., Wathelet, V., Below, R., Lanfredi Sofia, C., Tonnelier, M., van Loenhout, J.A.F., Speybroeck, N. (2024). The EM-DAT Emergency Events Database Archive. Open Data U CLouvain. `https://doi.org/10.14428/DVN/I0LTPH`.

# Thank you!

Questions?

*This work has been submitted and is currently under review.*