On tail dependence coefficients of transformed multivariate Archimedean copulas

Elena Di Bernardino,
CNAM, Paris, Département IMATH

Séminaire de statistique et économétrie
February 2015, University of Lille 3
Contents

1 Introduction

2 Transformations

3 Estimation
   - Self-nested diagonals
   - Non-parametric estimation
   - Parametric estimation
   - Real-data illustration

4 Tail dependence
   - Definition
   - Transformed Archimedean copulas
   - Estimation with given tail coefficients

5 Conclusion
The notion of *Return Period* (RP) is frequently used in *environmental sciences* for the identification of dangerous events, and provides a means for rational decision making and risk assessment.

Roughly speaking, the RP can be considered as an analogue of the “Value-at-Risk” in *Economics and Finance*, since it is used to quantify and assess the risk.

During the last years, researchers in environmental fields joined efforts to properly answer the following crucial question: “How is it possible to calculate the critical design event(s) in the multivariate case?”.

In this sense, a possible consistent theoretical framework for the calculation of the design event(s) and the associated return period(s) in a multi-dimensional environment, is proposed, e.g., by Salvadori et al. (2011), Salvadori et al. (2012), Gräler et al. (2013)

✓ Multivariate return period using the notion of *upper and lower level sets* of multivariate probability distribution $F$ and of the associated *Kendall’s measure*. 

The notion of *Return Period* (RP) is frequently used in *environmental sciences* for the identification of dangerous events, and provides a means for rational decision making and risk assessment.

Roughly speaking, the RP can be considered as an analogue of the “Value-at-Risk” in *Economics and Finance*, since it is used to quantify and assess the risk.

During the last years, researchers in environmental fields joined efforts to properly answer the following crucial question: “How is it possible to calculate the critical design event(s) in the multivariate case?”.

In this sense, a possible consistent theoretical framework for the calculation of the design event(s) and the associated return period(s) in a multi-dimensional environment, is proposed, e.g., by Salvadord et al. (2011), Salvadord et al. (2012), Gräler et al. (2013)

- Multivariate return period using the notion of *upper and lower level sets* of multivariate probability distribution $F$ and of the associated *Kendall’s measure*. 
The notion of Return Period (RP) is frequently used in environmental sciences for the identification of dangerous events, and provides a means for rational decision making and risk assessment.

Roughly speaking, the RP can be considered as an analogue of the “Value-at-Risk” in Economics and Finance, since it is used to quantify and assess the risk.

During the last years, researchers in environmental fields joined efforts to properly answer the following crucial question: “How is it possible to calculate the critical design event(s) in the multivariate case?”.

In this sense, a possible consistent theoretical framework for the calculation of the design event(s) and the associated return period(s) in a multi-dimensional environment, is proposed, e.g., by Salvadori et al.(2011), Salvadori et al.(2012), Gräler et al. (2013)

✓ Multivariate return period using the notion of upper and lower level sets of multivariate probability distribution $F$ and of the associated Kendall’s measure.
The notion of Return Period (RP) is frequently used in environmental sciences for the identification of dangerous events, and provides a means for rational decision making and risk assessment.

Roughly speaking, the RP can be considered as an analogue of the “Value-at-Risk” in Economics and Finance, since it is used to quantify and assess the risk.

During the last years, researchers in environmental fields joined efforts to properly answer the following crucial question: “How is it possible to calculate the critical design event(s) in the multivariate case?”.

In this sense, a possible consistent theoretical framework for the calculation of the design event(s) and the associated return period(s) in a multi-dimensional environment, is proposed, e.g., by Salvadore et al.(2011), Salvadore et al.(2012), Gräler et al. (2013)

✓ Multivariate return period using the notion of upper and lower level sets of multivariate probability distribution $F$ and of the associated Kendall’s measure.
The notion of Return Period (RP) is frequently used in environmental sciences for the identification of dangerous events, and provides a means for rational decision making and risk assessment.

Roughly speaking, the RP can be considered as an analogue of the “Value-at-Risk” in Economics and Finance, since it is used to quantify and assess the risk.

During the last years, researchers in environmental fields joined efforts to properly answer the following crucial question: “How is it possible to calculate the critical design event(s) in the multivariate case?”.

In this sense, a possible consistent theoretical framework for the calculation of the design event(s) and the associated return period(s) in a multi-dimensional environment, is proposed, e.g., by Salvadori et al. (2011), Salvadori et al. (2012), Gräler et al. (2013)

✓ Multivariate return period using the notion of upper and lower level sets of multivariate probability distribution F and of the associated Kendall’s measure.
Consider the (nonnegative) real-valued random vector $X = (x_1, \ldots, X_d)$ such that $X \sim F_X = C(F_{X_1}, \ldots, F_{X_d})$, with $F_X : \mathbb{R}_+^d \to [0, 1]$.

**Definition (Critical layer)**

The critical layer $\partial L(\alpha)$ associated to the multivariate distribution function $F_X$ of level $\alpha \in (0, 1)$ is defined as

$$\partial L(\alpha) = \{x \in \mathbb{R}_+^d : F_X(x) = \alpha\}.$$  

Then $\partial L(\alpha)$ is the iso-hyper-surface (with dimension $d-1$) where $F$ equals the constant value $\alpha$. The critical layer $\partial L(\alpha)$ partitions $\mathbb{R}_+^d$ into three non-overlapping and exhaustive regions:

$$\begin{cases} 
L^{<}(\alpha) & = & \{x \in \mathbb{R}_+^d : F_X(x) < \alpha\}, \\
\partial L(\alpha) & = & \text{the critical layer itself}, \\
L^{>}(\alpha) & = & \{x \in \mathbb{R}_+^d : F_X(x) > \alpha\}.
\end{cases}$$
Event of interest is of the type \( \{ X \in A \} \), where \( A \) is a non-empty Borel set in \( \mathbb{R}^d \) collecting all the values judged to be “dangerous” according to some suitable criterion.

✓ A natural choice for \( A \) is the set \( L^>(\alpha) \)
✓ Then \( \text{RP}^>(\alpha) = \frac{\Delta t}{\mathbb{P}[X \in L^>(\alpha)]} \), where \( \Delta t > 0 \) is the (deterministic) average time elapsing between \( X_k \) and \( X_{k+1} \), \( k \in \mathbb{N} \).

Then, the considered Return Period can be expressed using Kendall’s function

\[
\text{RP}^>(\alpha) = \Delta t \cdot \frac{1}{1 - K_C(\alpha)},
\]

where \( K_C(\alpha) = \mathbb{P}[X \in L^<(\alpha)] = \mathbb{P}[C(U_1, \ldots, U_d) \leq \alpha], \) for \( \alpha \in (0,1) \).
Goals and ideas

This talk aims at:

- Giving a parametric representation of the multivariate distribution $F$ of a $d$-random vector $X$ of risks using transformations
- Giving direct estimation procedure for this representation
- Giving closed parametric expressions, both for critical layers $\partial L(\alpha)$ and Return Periods $\text{RP}^>(\alpha)$
- Working on the tails by using generators exhibiting any chosen couple of lower and upper tail dependence coefficients
Goals and ideas

This talk aims at:

- Giving a parametric representation of the multivariate distribution $F$ of a $d-$random vector $\mathbf{X}$ of risks using transformations

- Giving direct estimation procedure for this representation

- Giving closed parametric expressions, both for critical layers $\partial L(\alpha)$ and Return Periods $\mathbb{RP}^>(\alpha)$

- Working on the tails by using generators exhibiting any chosen couple of lower and upper tail dependence coefficients
Goals and ideas

This talk aims at:

- Giving a parametric representation of the multivariate distribution $F$ of a $d$-random vector $X$ of risks using transformations
- Giving direct estimation procedure for this representation
- Giving closed parametric expressions, both for critical layers $\partial L(\alpha)$ and Return Periods $\text{RP}^>(\alpha)$
- Working on the tails by using generators exhibiting any chosen couple of lower and upper tail dependence coefficients
Goals and ideas

This talk aims at:

- Giving a parametric representation of the multivariate distribution $F$ of a $d-$random vector $X$ of risks using transformations
- Giving direct estimation procedure for this representation
- Giving closed parametric expressions, both for critical layers $\partial L(\alpha)$ and Return Periods $\text{RP}^>(\alpha)$
- Working on the tails by using generators exhibiting any chosen couple of lower and upper tail dependence coefficients
Goals and ideas

This talk aims at:

- Giving a parametric representation of the multivariate distribution $F$ of a $d$–random vector $X$ of risks using transformations
- Giving direct estimation procedure for this representation
- Giving closed parametric expressions, both for critical layers $\partial L(\alpha)$ and Return Periods $\text{RP}^>(\alpha)$
- Working on the tails by using generators exhibiting any chosen couple of lower and upper tail dependence coefficients
Part I. Transformations of Archimedean copulas
Let $F$ be a $d$–dimensional cdf with marginals $F_i$, $i = 1, \ldots, d$.

By Sklar’s theorem, there exists a copula function $C : [0, 1]^d \to [0, 1]$ that links the distribution $F$ with its margins:

$$F(x_1, \ldots, x_d) = C(F_1(x_1), \ldots, F_d(x_d)).$$

- $C$ is a $d$–dimensional cdf on $[0, 1]^d$ with uniform marginals.
- $C$ is unique if marginals $F_i$ are continuous.
Archimedean copulas:

\[ C(u_1, \ldots, u_d) = \phi \left( \phi^{-1}(u_1) + \ldots + \phi^{-1}(u_d) \right) \]

\( \phi : \mathbb{R}^+ \to (0, 1] \) is the generator of the Archimedean copula, \( \phi^{-1}(x) = \inf \{ s \in \mathbb{R}^+ : \phi(s) \leq x \} \) its (generalized) inverse function.

**Generator:** \( \phi \) is continuous, decreasing, \( d \)-monotone (cf. McNeil and Nešlehová, 2009), \( \phi(0) = 1, \lim_{x \to +\infty} \phi(x) = 0. \)

**One limitation:** we consider only here strict generators, i.e. \( \forall x \in \mathbb{R}^+, \phi(x) > 0. \Rightarrow \phi \) is strictly decreasing and \( \phi^{-1} \) is the regular inverse of \( \phi. \)

**Proposition (Equivalent generators, cf. Nelsen)**

*Generator* \( \phi_a(x) = \phi(ax) \) and \( \phi(x) \) lead to the same copula, \( a \in \mathbb{R} \setminus \{0\} \)

implies that one can ask \( \phi \) to be such that \( \phi(t_0) = \varphi_0 \) for an arbitrary point \( (t_0, \varphi_0) \in \mathbb{R} \times (0, 1). \)
Proposition (Transformed copula)

Copula $\tilde{C}$ depends on a continuous increasing function $T : [0, 1] \rightarrow [0, 1]$,

$$\tilde{C}(u_1, \ldots, u_d) = T(C_0(T^{-1}(u_1), \ldots, T^{-1}(u_d))).$$

and if the initial $C_0$ is a given Archimedean copula with generator $\phi_0$, then $\tilde{C}$ is an Archimedean copula with generator

$$\tilde{\phi}(x) = T \circ \phi_0(x)$$

Admissibility conditions for $\tilde{\phi}$ (d-monotonicity, cf. McNeil and Nešlehová, 2009) \implies more admissibility conditions for $T$ (e.g. using Faa Di Bruno formula).

Literature on these transformations: Durrleman et al. (2000), Charpentier (2008), Valdez and Xiao (2011).
Motivations

Possibility to transform both copulas and margins:

\[ \tilde{F}(x_1, \ldots, x_d) = \tilde{C}(\tilde{F}_1(x_1), \ldots, \tilde{F}_d(x_d)), \]

where \( \tilde{F}_i = T \circ T_i^{-1} \circ F_i \) and \( T_i : [0, 1] \rightarrow [0, 1] \) are continuous and increasing.

Why using transformations (instead of parametric multivariate distributions)?

1. **Flexible**: allows transformations composition
   - huge variety of reachable distributions (multimodal, etc.)
   - possibility to improve a fit gradually

2. **Invertible**: analytical expressions
   - for the expression of the distribution function \( F \);
   - but also for the expression of the critical layer \( \partial L(\alpha) \)

3. **Estimation facilities**
Part II. Estimation of Transformed Archimedean copulas

→ **Self-nested diagonals**
  - Non-parametric estimation
  - Parametric estimation
  - Real-data illustration
Idea: building a non-parametric estimator of the generator $\phi$ and for the transformation $T$ based on the diagonal section of the copula.

**Definition (Diagonal section of the copula)**

Consider a copula $C$ satisfying *regular conditions*. For all $u \in [0, 1]$, 

$$\delta_1(u) = C(u, \ldots, u),$$

and $\delta_{-1}$ is the inverse function of $\delta_1$ so that $\delta_1 \circ \delta_{-1} = \text{Id}$. 

**Remark:**

- The copula $C$ is not *always* uniquely determined by its diagonal (cf. Frank’s condition, Erdely et al. 2013).
- Estimation based *only* on this diagonal may fail to capture tail dependence when $\phi'(0) = -\infty$ (upper tail dependent case).
**Definitions**

**Definition (Discrete self-nested diagonals)**

Consider a copula $C$ satisfying *regular conditions*. The *discrete self-nested diagonal* of $C$ at order $k$ is the function $\delta_k$ such that for all $u \in [0,1]$, $k \in \mathbb{N}$

\[
\begin{align*}
\delta_k(u) &= \delta_1 \circ \ldots \circ \delta_1(u), \quad (k \text{ times}), \\
\delta_{-k}(u) &= \delta_{-1} \circ \ldots \circ \delta_{-1}(u), \quad (k \text{ times}), \\
\delta_0(u) &= u,
\end{align*}
\]

where $\delta_1(u) = C(u, \ldots, u)$ and $\delta_{-1}$ is the inverse function of $\delta_1$, so that $\delta_1 \circ \delta_{-1}$ is the identity function.

**Definition (Self-nested diagonals)**

Functions of a family $\{\delta_r\}_{r \in \mathbb{R}}$ are called (extended) self-nested diagonals of a copula $C$, if $\delta_k(u)$ is the discrete self-nested diagonal of $C$ at order $k$, for all $k \in \mathbb{Z}$, and if furthermore

\[
\delta_{r_1 + r_2}(u) = \delta_{r_1} \circ \delta_{r_2}(u), \quad \forall r_1, r_2 \in \mathbb{R}, \forall u \in [0,1].
\]
### Proposition (Self-nested diagonals of an Archimedean copula)

If $C$ is an Archimedean copula associated with a generator $\phi$, then the self-nested diagonals of $C$ at order $r$ is

$$\delta_r(x) = \phi(d^r \cdot \phi^{-1}(x)), \quad r \in \mathbb{R}.$$  

- If $C$ is the *Independence* copula of generator $\phi(t) = \exp(-t)$, then $\delta_r(u) = u^{(d^r)}$.
- If $C$ is a *Gumbel* copula of generator $\phi(t) = \exp(-t^{1/\theta})$, then $\delta_r(u) = u^{(d^{(r/\theta)})}$, $\theta \geq 1$.
- If $C$ is a *Clayton* copula of generator $\phi(t) = (1 + \theta t)^{-1/\theta}$, $\delta_r(u) = (1 + d^r(t^{-\theta} - 1))^{-1/\theta}$, $\theta > 0$. 

Proposition (Transformation $T$ using self-nested diagonals)

Consider Archimedean copulas $C_0$ and $\tilde{C}$ satisfying regular conditions and the associated self-nested diagonals $\delta_r$ and $\tilde{\delta}_r$, $r \in \mathbb{R}$. If $T$ is defined by $T(0) = 0$, $T(1) = 1$ and for all $x \in (0, 1)$,

$$T(x) = \tilde{\delta}_{r(x)}(y_0),$$

with $r(x)$ such that $\delta_{r(x)}(x_0) = x$,

then the transformed copula using transformation $T$ is equal to $\tilde{C}$: for all $u_1, \ldots, u_d$,

$$\tilde{C}(u_1, \ldots, u_d) = T \circ C_0(T^{-1}(u_1), \ldots, T^{-1}(u_d)),$$

where $(x_0, y_0) \in (0, 1)^2$ can be arbitrarily chosen. In the case where $C_0$ is the independence copula,

$$r(x) = \frac{1}{\ln d} \ln \left( \frac{-\ln x}{-\ln x_0} \right).$$
Generators using self-nested diagonals

**Proposition (Generator $\tilde{\phi}$ using self-nested diagonals)**

Consider an Archimedean copula $\tilde{C}$ satisfying regular conditions, and the associated self-nested copulas $\tilde{\delta}_r$, for $r \in \mathbb{R}$. Assume that the copula $\tilde{C}$ is reachable by transforming an Archimedean copula $C_0$, and denote by $\delta_r$, $r \in \mathbb{R}$, the self-nested diagonals of $C_0$ and by $\phi_0$ its generator. A generator $\tilde{\phi}$ of $\tilde{C}$ is defined for all $t \in \mathbb{R}^+$ by

$$\tilde{\phi}(t) = \tilde{\delta}_{\rho(t)}(y_0),$$

with $\rho(t)$ such that $\delta_{\rho(t)}(x_0) = \phi_0(t)$ i.e.,

$$\rho(t) = \frac{1}{\ln d} \ln \left( \frac{\frac{t}{\phi_0^{-1}(x_0)}}{\phi_0^{-1}(x_0)} \right),$$

where $(x_0, y_0) \in (0, 1)^2$ can be arbitrarily chosen. In the particular case where $C_0$ is the independent copula, then

$$\rho(t) = \frac{1}{\ln d} \ln \left( \frac{t}{-\ln x_0} \right)$$
Part II. Estimation of the transformed Archimedean copula

- Self-nested diagonals
- Non-parametric estimation
  - Parametric estimation
  - Real-data illustration
First, build a (smooth) estimator \( \hat{\delta} \) of the diagonal of the target transformed copula \( \tilde{C} \) (e.g., empirical copula, Deheuvels (1979), Fermanian et al. (2004), Omelka et al. (2009)), and denote its inverse function \( \hat{\delta}^{-1} \).

Estimators of discrete self-nested diagonal of \( \tilde{C} \) at order \( k \) are the function \( \hat{\delta}_k \) such that for all \( u \in [0, 1] \), \( k \in \mathbb{N} \)

\[
\begin{align*}
\hat{\delta}_k(u) &= \hat{\delta}_1 \circ \ldots \circ \hat{\delta}_1(u), \quad (k \text{ times}), \\
\hat{\delta}_{-k}(u) &= \hat{\delta}_{-1} \circ \ldots \circ \hat{\delta}_{-1}(u), \quad (k \text{ times}), \\
\hat{\delta}_0(u) &= u,
\end{align*}
\]
Non-parametric estimators of $\tilde{\phi}$

**Definition (Non-parametric estimation of $\tilde{\phi}$ - Case $C_0$ independent copula)**

Consider an Archimedean copula $\tilde{C}$ and associated self-nested diagonals $\tilde{\delta}_r$, for $r \in \mathbb{R}$. Denote by $\hat{\delta}_r$ the estimator of $\tilde{\delta}_r$, for $r \in \mathbb{R}$. Assume that $\tilde{\phi}(t_0) = \varphi_0$, for a given couple of values $(t_0, \varphi_0) \in \mathbb{R}^+ \setminus \{0\} \times (0,1)$. A non-parametric estimator $\hat{\phi}$ of $\tilde{\phi}$ is defined by $\hat{\phi}(0) = 1$ and for all $t \in \mathbb{R}^+ \setminus \{0\}$,

$$
\hat{\phi}(t) = \hat{\delta}_{\rho(t)}(\varphi_0),
$$

with $\rho(t) = \frac{1}{\ln \delta} \ln \left( \frac{t}{t_0} \right),
$

where $(t_0, \varphi_0) \in \mathbb{R}^+ \setminus \{0\} \times (0,1)$ can be arbitrarily chosen.

In particular, the estimator $\hat{\phi}$ of $\tilde{\phi}$ is passing through the points

$$
\{(t_k, \varphi_k)\}_{k \in \mathbb{Z}} = \{(d^k t_0, \hat{\delta}_k(\varphi_0))\}_{k \in \mathbb{Z}}
$$
Non-parametric $\hat{\phi}(t)$

Figure: Estimated versus theoretical Gumbel-generator with parameter $\theta = 3$. Size of simulated samples $n = 150$ (left) and $n = 1500$ (right). Estimated $\hat{\phi}(t) = \hat{\delta}_p(t)(y_0)$ (full line). The theoretical standardized Gumbel-generator, i.e., $\bar{\phi}(t) = \exp(-t^{1/\theta})$, is drawn using a dashed line. We force the generators to pass through the point $(t_0, \varphi_0) = (1, e^{-1})$ (black point).
Comparison with estimator of Genest et al. (2011)

Recall the $\lambda$ function, as originally introduced in Genest and Rivest (1993) for inferential purposes,

$$\lambda(u) = \phi^{-1}(u) \cdot \phi'(\phi^{-1}(u)).$$

One can easily see that this coefficient is identical for any generator belonging to the same equivalent class and on the contrary of the generator itself, does not depend on the choice of some arbitrarily point $(t_0, \varphi_0)$.

Following the same methodology as Genest and Rivest (1993), we estimate the $\lambda$ function, for our estimator and for the estimator of Genest et al. (2011).

For our estimator, we propose:

- $$\hat{\lambda}(u) = \hat{\phi}^{-1}(u) \cdot \frac{\hat{\phi}(\hat{\phi}^{-1}(u)+h)-\hat{\phi}(\hat{\phi}^{-1}(u)-h)}{2h},$$

  for a small value of $h$, $u \in (h, 1-h)$.

- $$\hat{\lambda}^*(u) = \frac{1}{\ln d} \frac{\hat{\delta}_h(u)-\hat{\delta}_{-h}(u)}{2h},$$

  (since $\lambda(u) = \frac{1}{\ln d} \frac{\partial}{\partial r} \delta_r(u)|_{r=0}$; this second simple estimator of $\lambda$ permit to avoid function inversions $\phi^{-1}$).
Black: theoretical \( \lambda \) function. Dark green dashed line: estimator by Genest et al. (2011), Black dotted line: \( \hat{\lambda}(u) \). Violet dotted-dashed line: \( \hat{\lambda}^*(u) \).
Black: theoretical $\lambda$ function. Dark green dashed line: estimator by Genest et al. (2011), Black dotted line: $\hat{\lambda}(u)$.
Violet dotted-dashed line: $\hat{\lambda}^*(u)$. 
Comparison with estimator of Genest et al. (2011)

- The violet dashed line $\hat{\lambda}^*(u)$ seems performing a little bit better.
- On tested data, no estimator seems to perform significantly better.

However, in our case, estimators relying on self-nested diagonals have several advantages among which:

✓ The facility to get generators passing through a given point, contrary to the estimator in Genest et al. (2011) which relies on the choice of a radius $r_m$. For instance, in this figures we had to find of the Genest et al.’s estimator the value $r_m = 5500$ to get a correct result.

✓ Both estimators of $T$ or $\phi$ are relying on direct analytical expressions, whereas the estimator in Genest et al. (2011) rely on a large number of root resolution procedures. Indeed in the last estimator we have to solve a triangular non-linear system containing $m$ equations. If the sample size is $n = 2000$ the value $m$ is approximately around $1200 - 1300$.

✓ Some first theoretical results on confidence bands (see slides below). Such results would probably be difficult to get with estimators relying on successive optimization procedures or root resolutions.
Definition (Non-parametric estimation of $T$ - Case $C_0$ independent copula)

Consider an Archimedean copula $\tilde{C}$ and associated self-nested diagonals $\tilde{\delta}_r$, for $r \in \mathbb{R}$. Denote by $\hat{\delta}_r$ the estimator of $\tilde{\delta}_r$, for $r \in \mathbb{R}$. A non-parametric estimator of $T$ is defined by $\hat{T}(0) = 0$, $\hat{T}(1) = 1$ and for all $x \in (0, 1)$ by

$$\hat{T}(x) = \hat{\delta}_{r(x)}(y_0),$$

with $r(x) = \frac{1}{\ln d} \ln \left( \frac{-\ln x}{-\ln x_0} \right),$

where $(x_0, y_0) \in (0, 1)^2$ can be arbitrarily chosen.
Non-parametric estimators of $T$ and $\tilde{\phi}$

Same kind of estimator for $\hat{T}$ or $\hat{\phi}$ in the general case:

$$\hat{T}(x) = \hat{\delta}_{r(x)}(y_0),$$

with $r(x)$ such that $\delta_{r(x)}(x_0) = x$,

where $(x_0, y_0) \in (0, 1)^2$ can be arbitrarily chosen.

Proposition (Theoretical confidence bands for estimators - not detailed here)

**Theoretical confidence bands on $\hat{\delta}$ $\Rightarrow$ theoretical confidence bands on $\hat{\phi}$.**

one needs the distribution of the empirical process $\hat{\delta}(u)$, $u \in [0, 1]$. 
Confidence bands for $\hat{\delta}_1$ imply confidence bands for $\hat{\phi}$

Figure: (Left) Confidence bands for $\hat{\delta}_1$ and $\hat{\delta}_{-1}$ for chosen parameters $\alpha^- = \beta^- = 1.05$, $\alpha^+ = \beta^+ = 0.9$. (Right) Resulting theoretical confidence band for $\hat{\phi}$. Here $\tilde{C}$ is a Gumbel copula of parameter $\theta = 2$, the size of generated sample is $n = 2000$. 
Part II. Estimation of the transformed Archimedean copula

- Self-nested diagonals
- Non-parametric estimation
  → Parametric estimation
- Real-data illustration
A class of parametric transformation

We take back from Bienvenüe and Rullière (2012):

**Definition (Conversion and transformation functions)**

Let \( f \) any bijective increasing function from \( \mathbb{R} \) to \( \mathbb{R} \). It is said to be a conversion function. The transformation \( T_f : [0, 1] \rightarrow [0, 1] \) is defined as

\[
T_f(u) = \begin{cases} 
0 & \text{if } u = 0, \\
\logit^{-1}(f(\logit(u))) & \text{if } 0 < u < 1, \\
1 & \text{if } u = 1.
\end{cases}
\]

*Remark:* transformations function are chosen in a way to be easily invertible \((T_f \circ T_g = T_{f \circ g}, T_f^{-1} = T_{f^{-1}})\).

- We will use (composited) hyperbolic conversion functions:

\[
f(x) = H_{m,h,p_1,p_2,\eta}(x) = m - h + (e^{p_1} + e^{p_2})\frac{x - m - h}{2} - (e^{p_1} - e^{p_2})\sqrt{\left(\frac{x - m - h}{2}\right)^2 + e^{\eta - \frac{p_1 + p_2}{2}}}
\]

with \( m, h, p_1, p_2 \in \mathbb{R} \), and one smoothing parameter \( \eta \in \mathbb{R} \).

- \( f^{-1}(x) = H^{-1}_{m,h,p_1,p_2,\eta}(x) = H_{m,-h,-p_1,-p_2,\eta}(x) \).
Idea:

- Transform an initial given copula $C_0$ using a transformation $T$,
- Transform initial given margins using transformations $T_1, \ldots, T_d$.

Estimation:

- Get non-parametric estimators for $T$
- Get non-parametric estimators for $T_1, \ldots, T_d$
- Fit all transformations $T, T_1, \ldots, T_d$ by piecewise-linear functions
- Approach piecewise linear functions by composition of hyperbolas $H_{m, h, p_1, p_2, \eta}$, cf. Bienvenüe and Rullière (2012).
Part II. Estimation of the transformed Archimedean copula

- Self-nested diagonals
- Non-parametric estimation
- Parametric estimation

→ Real-data illustration
Geyser data: transformed (parametric) bivariate density

Non-parametric $\Rightarrow$ Parametric estimation (without optimization).

Data: 272 eruptions of the **Old Faithful geyser** in Yellowstone National Park. Each observation consists of two measurements: the duration (in min) of the eruption ($X$), and the waiting time (in min) before the next eruption ($Y$).

**Figure**: Level curves of transformed density $\tilde{f}(x_1, x_2)$ and Old Faithful geyser data (red points). Left: parameter setting $\eta = -0.9$, $\eta_1 = -9$, $\eta_2 = -7.5$; (right) parameter setting $\eta = -0.9$, $\eta_1 = -4$, $\eta_2 = -4$. 
Geyser data: transformed (parametric) c.d.f. $F$ and $\partial L(\alpha)$

Figure: (Left) Transformed distribution $\tilde{F}(x_1, x_2)$ with associated transformed level curves (red curves). (Right) Black points are plotted at empirical tail probabilities on the diagonal, calculated from the empirical bivariate distribution (in log scale). Red line is $1 - \tilde{F}(x, x)$ (in log scale). The vertical dotted lines show estimate of 95% VaR (i.e., the univariate quantile in log scale).
Geyser data: transformed (parametric) marginal $F_1$

Figure: Duration (in min) of the eruption data. (Left) $\tilde{F}_1$ (red) and the empirical distribution function of eruption data (black). (Right) Black points are plotted at empirical tail probabilities calculated from empirical distribution function (in log scale). Red line is $1 - \tilde{F}_1(x)$ (in log scale). The vertical dotted lines show estimate of 95% VaR (univariate quantile) for the eruption data.
Figure: Black points are plotted at empirical tail probabilities calculated from the empirical marginal distributions (in log scale). Red line is $1 - \tilde{F}_1(x)$ (in log scale) (left) and $1 - \tilde{F}_2(x)$ (in log scale) (right). The vertical dotted lines show estimate of 95% VaR (i.e., the univariate quantiles in log scale).
Part III. Tail dependence

→ Definitions

- Transformed Archimedean copulas
- Estimation given tail coefficients
Tail problem

Need for some applications (hydrology, finance, etc.).

- Need to know joint extreme behavior of some random variables
- Need to define and capture tail dependence

Limits of our non-parametric estimator:

- Here estimation relying on the diagonal section \( \hat{\delta}_1 \),
  - The copula \( C \) is not always uniquely determined by its diagonal (cf. Frank’s condition, Erdely et al. 2013)
  - Estimation based only on this diagonal may fail to capture tail dependence when \( \phi'(0) = -\infty \).
- Not usual to capture tail behavior for non-parametric estimation (e.g. empirical cdf, our generator estimator, Genest et al., 2011 one)
Possible parametric solutions: here the transformed generator is

\[ \tilde{\phi} = T \circ \phi_0 \]

to reach a given tail dependence, possibility to work

- on the tails of the transformation \( T \)
- on the initial generator \( \phi_0 \) to be transformed

✓ e.g. \( \phi_0 \) with good tail dependence, \( T \) not changing tail dependence,

✓ e.g. \( \phi_0 \) with correct central fit, \( T \) changing only tails,

✓ etc.

We work with the lower and upper tail dependence coefficients (TDC).

**Definition (Bivariate tail dependence coefficients - Sibuya (1960))**

For \( d = 2 \), the classical bivariate upper and lower tail dependence coefficients, \( \lambda_U \) and \( \lambda_L \), are defined as

\[
\lambda_L = \lim_{u \to 0^+} \mathbb{P}[V \leq u \mid U \leq u] \quad \text{and} \quad \lambda_U = \lim_{u \to 1^-} \mathbb{P}[V > u \mid U > u].
\]
### Definition (Multivariate tail dependence coefficients: TDC)

Assume that the considered copula $C$ is the distribution of some random vector $\mathbf{U} := (U_1, \ldots, U_d)$. Denote $I = \{1, \ldots, d\}$ and consider two non-empty subsets $I_h \subset I$ and $\overline{I}_h = I \setminus I_h$ of respective cardinal $h \geq 1$ and $d - h \geq 1$. A multivariate version of classical bivariate tail dependence coefficients is given by Li (2009) and De Luca and Rivieccio (2012) (when limits exist):

$$
\lambda_{L_h, \overline{I}_h} = \lim_{u \to 0^+} \mathbb{P} \left[ U_i \leq u, i \in I_h \mid U_i \leq u, i \in \overline{I}_h \right],
$$

$$
\lambda_{U_h, \overline{I}_h} = \lim_{u \to 1^-} \mathbb{P} \left[ U_i \geq u, i \in I_h \mid U_i \geq u, i \in \overline{I}_h \right].
$$

If for all $I_h \subset I$, $\lambda_{L_h, \overline{I}_h} = 0$, (resp. $\lambda_{U_h, \overline{I}_h} = 0$) then we say $\mathbf{U}$ is lower tail independent (resp. upper tail independent).
With exchangeable r.v., depends on \( d = \text{card}(I) \) and \( h = \text{card}(I_h) \).
In particular case of Archimedean copulas, one have \( (\psi = \phi^{-1}) \):

**Proposition (Multivariate Tail Dep. Coeffs. for Archimedean copulas)**

*For Archimedean copulas the multivariate lower and upper tail dependence coefficients are respectively:*

\[
\lambda^{(h,d-h)}_L = \lim_{u \to 0^+} \frac{\psi^{-1}(d \psi(u))}{\psi^{-1}((d - h) \psi(u))},
\]

\[
\lambda^{(h,d-h)}_U = \lim_{u \to 1^-} \frac{\sum_{i=0}^{d} (-1)^i C_d^i \psi^{-1}(i \psi(u))}{\sum_{i=0}^{d-h} (-1)^i C_{d-h}^i \psi^{-1}(i \psi(u))}.
\]

Some important references on tail dependence and Archimedean copulas: Juri and Wüthrich (2003), Charpentier and Segers (2009), Durante et al. (2010).
TDC for some usual copulas

Figure: Concentration function $\lambda\text{LU}(u) = 1\{u \leq 1/2\} \lambda_L(u) + 1\{u > 1/2\} \lambda_U(u)$ for some Clayton copulas (left panel) and Gumbel copulas (right panel).
As noticed in the bivariate case by Avérous and Dortet-Bernadet (2004),

“many of the most commonly used parametric families of Archimedean copulas (such as the Clayton, Gumbel, Frank or Ali-Mikhail-Haq systems) possess strong dependence properties: they have the SI or the SD property and are ordered at least by <LTD.”

where Stochastic Increasingness (SI), Stochastically Decreasingness (SD) and Left-Tail Decreasingness (LTD) definitions are recalled in Joe (1997).

It implies in particular that for many classical Archimedean copulas (and in particular for Clayton, Gumbel, Frank or Ali-Mikhail-Haq copulas) $\lambda_{LU}$ is nondecreasing on $(0, 1/2)$ and non-increasing on $(1/2, 1)$.
Regular Variation ($\mathcal{RV}$)

At infinity:

$$f \in \mathcal{RV}_\alpha(\infty) \iff \forall s > 0, \lim_{x \to +\infty} \frac{f(sx)}{f(x)} = s^\alpha.$$ 

At zero, using $M(x) = 1 - x$ and $I(x) = 1/x$,

$$f \in \mathcal{RV}_\alpha(0) \iff f \circ I \in \mathcal{RV}_{-\alpha}(\infty) \iff \forall s > 0, \lim_{x \to 0^+} \frac{f(sx)}{f(x)} = s^\alpha.$$ 

At one,

$$f \in \mathcal{RV}_\alpha(1) \iff f \circ M \circ I \in \mathcal{RV}_{-\alpha}(\infty) \iff \forall s > 0, \lim_{x \to 0^+} \frac{f(1 - sx)}{f(1 - x)} = s^\alpha.$$
Part III. Tail dependence

- Definitions
  - Transformed Archimedean copulas
- Estimation given tail coefficients
Considered transformations

We consider transformations $T_f : [0, 1] \rightarrow [0, 1]$ such that

$$T_f(u) = \begin{cases} 
0 & \text{if } u = 0, \\
G \circ f \circ G^{-1}(u) & \text{if } 0 < u < 1, \\
1 & \text{if } u = 1,
\end{cases}$$

Why?

- The transformation $T_f$ has support $[0, 1]$.
- The function $G$ is a cdf that aims at transferring this support on the whole real line $\mathbb{R}$.
- The conversion function $f$ is any continuous bijective increasing function, $f : \mathbb{R} \rightarrow \mathbb{R}$, without bounding constraints. cf Bienvenüe and Rullière.
Considered transformations - Archimedean case

Assumption (Considered transformed generators)

Consider an initial Archimedean copula $C_0$ with generator $\phi_0$, and the associated transformed one, $\tilde{C}$, with generator

$$\tilde{\phi} = T_f \circ \phi_0,$$

i.e. on $(0, \infty)$,

$$\tilde{\phi} = G \circ f \circ G^{-1} \circ \phi_0.$$

One assumes that generators $\phi_0$ and $\tilde{\phi}$ satisfy admissibility conditions.

The transformed generator will depend on the RV properties of $f$, $\phi_0$ and $G$. 
**Assumption (Lower-tails: assumptions on $f$, $\phi_0$, $G$)**

Assume that $f$, $\phi_0$ and $G$ are continuous and differentiable functions, strictly monotone with respective proper inverse functions denoted $f^{-1}$, $\psi_0 = \phi_0^{-1}$ and $G^{-1}$. Furthermore,

i) The function $f$ **has a left asymptote** $\overline{f}(x) = ax + b$ as $x$ tends to $-\infty$, for $a \in (0, +\infty)$ and $b \in (-\infty, +\infty)$.

ii) The inverse initial generator $\psi_0$ **is regularly varying at 0** with some index $-r_0$, that is $\psi_0 \in RV_{-r_0}(0)$, with $r_0 \in [0, +\infty]$.

iii) The function $G$ is a non-defective continuous c.d.f. with support $\mathbb{R}$. The following **rate of $G$ is regularly varying** with some index $g - 1$: $m_G = G' / G \in RV_{g - 1}(-\infty)$, with $g \in (0, +\infty)$.
Assumption (Upper-tails: assumptions on $f$, $\phi_0$, $G$)

Assume that $f$, $\phi_0$ and $G$ are continuous and differentiable functions, strictly monotone with respective proper inverse functions denoted $f^{-1}$, $\psi_0 = \phi_0^{-1}$ and $G^{-1}$. Furthermore,

i) The function $f$ has a right asymptote $\bar{f}(x) = \alpha x + \beta$ as $x$ tends to $+\infty$, for $\alpha \in (0, +\infty)$ and $\beta \in (-\infty, +\infty)$.

ii) The inverse initial generator $\psi_0$ is regularly varying at 1 with some index $\rho_0$, i.e., $\psi_0 \in RV_{\rho_0}(1)$, with $\rho_0 \in [1, +\infty]$.

iii) The function $G$ is a non-defective continuous c.d.f. with support $\mathbb{R}$. The hazard rate of $G$ is regularly varying with some index $\gamma - 1$, that is $\mu_G = G'/\bar{G} \in RV_{\gamma - 1}(\infty)$, with $\bar{G} = 1 - G$ and $\gamma \in (0, +\infty)$. 


Theorem (Multivariate lower TDC of transformed Archimedean copula)

Under Lower tail assumptions,

i) the inverse transformed generator $\tilde{\psi}$ is such that

$$\tilde{\psi} \in \mathcal{RN}_{-\tilde{r}}(0), \text{ with } \tilde{r} = r_0 \cdot a^{-g}$$

ii) the associated transformed lower multivariate lower tail dependence coefficient is given by:

$$\tilde{\lambda}_L^{(h,d-h)} = \begin{cases} 
\text{see next slide}, & \text{if } \tilde{r} = 0, \\
\frac{d^{-a} r_0^{-1} (d - h)^{a} r_0^{-1}}{1}, & \text{if } \tilde{r} \in (0, +\infty), \\
1, & \text{if } \tilde{r} = +\infty.
\end{cases}$$
Theorem (Multivariate upper TDC of transformed Archimedean copula)

Under Lower tail assumptions,

i) the inverse transformed generator $\tilde{\psi}$ is such that

$$\tilde{\psi} \in \mathcal{RV}_\tilde{\rho}(1), \text{ with } \tilde{\rho} = \rho_0 \cdot \alpha^{-\gamma} \text{ and } \tilde{\rho} \in [1, +\infty]$$

ii) the associated transformed upper multivariate lower tail dependence coefficient is given by:

$$\tilde{\lambda}_U^{(h\cdot d-h)} = \begin{cases} 
\text{see next slide,} & \text{if } \tilde{\rho} = 1, \\
\frac{\sum_{i=1}^{d} C_{d}^i (-1)^i \cdot i^{\alpha \gamma} \rho_0^{-1}}{\sum_{i=1}^{d-h} C_{d-h}^i (-1)^i \cdot i^{\alpha \gamma} \rho_0^{-1}}, & \text{if } \tilde{\rho} \in (1, +\infty), \\
1, & \text{if } \tilde{\rho} = +\infty.
\end{cases}$$
Asymptotic lower tail independence When \( \psi_0 \in \mathcal{RV}_{-r_0}(0) \), with \( r_0 = 0 \),

\[
\tilde{\lambda}_L^{(h,d-h)} = \lim_{u \to 0+} \tilde{\lambda}_L^{(h,d-h)}(u) = 0
\]

- if \( \mu_{\phi_0} = \phi'_0/\phi_0 \in \mathcal{RV}_{k_0-1}(\infty) \), with \( k_0 \in [0, +\infty) \), if \( T_f \in \mathcal{RV}_{\bar{a}}(0) \) for \( \bar{a} \in (0, +\infty) \),

\[
\tilde{\lambda}_L^{(h,d-h)}(u) \in \mathcal{RV}_{\tilde{z}}(0) \text{ with } \tilde{z} = d^{k_0} - (d-h)^{k_0}.
\]

Asymptotic upper tail independence When \( \tilde{\psi} \in \mathcal{RV}_{\tilde{\rho}}(1) \), with \( \tilde{\rho} = 1 \) and if \( \tilde{\phi} \) is a \( d \) times continuously differentiable generator.

\[
\tilde{\lambda}_U^{(h,d-h)} = \lim_{u \to 1-} \tilde{\lambda}_U^{(h,d-h)}(u) = 0.
\]

- if \((-D)^d \tilde{\phi}(0)\) is finite and not zero, where \( D \) is the derivative operator,

\[
\tilde{\lambda}_U^{(h,d-h)}(u) \in \mathcal{RV}_h(1);
\]

- if \( \tilde{\psi}'(1) = 0 \) and the function \( \tilde{L}(s) := s \frac{d}{ds} \{ \frac{\tilde{\psi}(1-s)}{s} \} \) is positive and \( \tilde{L} \in \mathcal{RV}_0(0) \),

\[
\tilde{\lambda}_U^{(h,d-h)}(u) \in \mathcal{RV}_0(1).
\]
Remarks:

- \( \tilde{\lambda}^{(h,d-h)}_U(u) \in \mathcal{RV}_h(1); \)
  - This case represents a upper asymptotic independence for \( \tilde{C} \) in a rather strong sense. It is called *near independence*.

- \( \tilde{\lambda}^{(h,d-h)}_U(u) \in \mathcal{RV}_0(1) \)
  - This case represents a boundary between asymptotic independence and asymptotic dependence -case called *near asymptotic dependence*.
Part III. Tail dependence

- Definitions
- Transformed Archimedean copulas
  \textbf{Estimation given tail coefficients}
Fit with given TDC

In the bivariate case, when \( d = 2 \),

\[
\tilde{\lambda}_L^{(1,1)} = 2 - a^g r_0^{-1} \quad \text{and} \quad \tilde{\lambda}_U^{(1,1)} = 2 - 2^\alpha \gamma \cdot \rho_0^{-1}
\]

Take a transformation \( T_f(x) = G \circ f \circ G^{-1}(x) \) with \( f = H_{m, h, p_1, p_2, \eta} \),

\[
H_{m, h, p_1, p_2, \eta}(x) = m - h + \left( e^{p_1} + e^{p_2} \right) \frac{x - m - h}{2} - (e^{p_1} - e^{p_2}) \sqrt{\left( \frac{x - m - h}{2} \right)^2 + e^{-\frac{p_1 + p_2}{2}}}.
\]

So, if these tail coefficients are given, then we can easily find \( a = e^{p_1} \) and \( \alpha = e^{p_2} \) as functions of \( \tilde{\lambda}_L^{(1,1)} \) and \( \tilde{\lambda}_U^{(1,1)} \):

\[
p_1 = \frac{1}{g} \ln \left( -r_0 \frac{\ln \tilde{\lambda}_L^{(1,1)}}{\ln 2} \right) \quad \text{and} \quad p_2 = \frac{1}{\gamma} \ln \left( \rho_0 \frac{\ln(2 - \tilde{\lambda}_U^{(1,1)})}{\ln 2} \right).
\]

In the multivariate case, when \( d > 2 \), same principle, but expression of \( p_2 \) more difficult to write.
Illustration with logit transformations

- Let the initial copula $C_0$ be a Clayton copula with parameter $\theta > 0$.
- Assume that we want to obtain an arbitrarily chosen couple of target tail coefficients: $\tilde{\lambda}_L^{(1,1)} = 1/4$ and $\tilde{\lambda}_U^{(1,1)} = 3/4$.
- Let $G = \logit^{-1}$, then $g = \gamma = 1$.
- We obtain the parameters $p_1$ and $p_2$ of the hyperbolic conversion function $H$.

<table>
<thead>
<tr>
<th>chosen parameters</th>
<th>deduced parameters</th>
<th>tails coefficients</th>
</tr>
</thead>
<tbody>
<tr>
<td>$m$</td>
<td>$h$</td>
<td>$\eta$</td>
</tr>
<tr>
<td>A</td>
<td>0.5</td>
<td>0.9</td>
</tr>
<tr>
<td>B</td>
<td>0.5</td>
<td>-0.9</td>
</tr>
<tr>
<td>C</td>
<td>0.5</td>
<td>0.9</td>
</tr>
<tr>
<td>D</td>
<td>2</td>
<td>-0.9</td>
</tr>
</tbody>
</table>
Figure: Tail concentration function $\lambda_{LU}(u) = 1_{\{u \leq 1/2\}} \lambda_L(u) + 1_{\{u > 1/2\}} \lambda_U(u)$ for some transformed copulas. Chosen values: $\lambda_L = 1/4$ and $\lambda_U = 3/4$. Full line corresponds to parameters A in Table before; dashed line to B; dotted line to C; dashed-dotted line to D.
Conclusion
On the non-parametric estimation of the generator:
- Comparable performance with other estimators,
- Without solving non-linear systems of equations,
- Theoretical confidence bands.

On the parametric estimation of the generator:
- Analytical expressions for the level curves,
- Tunable number of parameters,
- Tail dependence can be chosen

Perspectives
- Estimators using other information ($\phi_\theta$ by classical estimation methods, e.g. tau’s Kendall + transformed tails)
- Using nested copulas (multimodal distribution or non-exchangeable vectors - nested copula...)
- Comparisons of our multivariate extreme measures with existing multivariate risk measures.
Papers (partly) presented

More details on **Non-parametric estimation** part:


More details on **Parametric estimation** part:


More details on **Tail dependence** part:

Thank you for your attention