“Modélisation de la dépendance et mesures de risque multidimensionnelles”

Elena Di Bernardino,

travaux sous la direction de

Véronique Maume-Deschamps, Université Lyon 1, ISFA
Clémentine Prieur, Université Joseph Fourier, Grenoble

Equipe de Statistique, IRMA, Université de Strasbourg
27 Mars 2012
Outline

1. Introduction

2. Estimating Bivariate Tail: a copula based approach

3. A multivariate extension of Value-at-Risk and Conditional Tail Expectation

4. Plug-in estimation of level sets in a non-compact setting with applications in multivariate risk theory

5. Perspectives
Modeling risks

“... when $d = 1$, concepts such as extreme values, order statistics and record values have natural definitions, but when $d > 1$, this is no longer the case as several different concepts of ordering are possible”. Resnick (1987)

Different aspects of dependence modeling with applications in risk theory:

- **Extreme Value Theory, Coles (2001)**
  - $\oplus$ Extreme regions contain hardly or no data; classical inference is difficult.
  - $\ominus$ Problem of asymptotic independence data.

- **Risk Theory, Denuit et al. (2005)**
  - $\oplus$ Provide models to study and compare risks.
  - $\ominus$ Multivariate Risks? Dependent Risks?

- **Copulae and dependence, Nelsen (1999)**
  - $\oplus$ Easy and analytical formulas.
  - $\ominus$ Dependence in extremes, comonotonic, counter-comonotonic.
Questions/Answers in this thesis:

✓ What can be considered, in a context of multidimensional portfolios, as the analogous of a “worst case” scenario and a related “tail distribution”? (multidimensional quantiles).

✓ Measures for risks with heterogeneous characteristics especially in an “external risks problem”?

✓ How to model the tail dependence, that is, the dependence between extreme events, occurring with low frequency but having a high impact?

✓ How to build measures with properties that turn to be consistent with existing properties on univariate setting?

✓ Particular attention to consistency results, asymptotic properties, illustrations of results on several real/simulated data sets.
Questions/Answers in this thesis:

✓ What can be considered, in a context of multidimensional portfolios, as the analogous of a “worst case” scenario and a related “tail distribution”? (multidimensional quantiles).

✓ Measures for risks with heterogeneous characteristics especially in an “external risks problem”?

✓ How to model the tail dependence, that is, the dependence between extreme events, occurring with low frequency but having a high impact?

✓ How to build measures with properties that turn to be consistent with existing properties on univariate setting?

✓ Particular attention to consistency results, asymptotic properties, illustrations of results on several real/simulated data sets.
Questions/Answers in this thesis:

✓ What can be considered, in a context of multidimensional portfolios, as the analogous of a “worst case” scenario and a related “tail distribution”? (multidimensional quantiles).

✓ Measures for risks with heterogeneous characteristics especially in an “external risks problem”?

✓ How to model the tail dependence, that is, the dependence between extreme events, occurring with low frequency but having a high impact?

✓ How to build measures with properties that turn to be consistent with existing properties on univariate setting?

✓ Particular attention to consistency results, asymptotic properties, illustrations of results on several real/simulated data sets.
Questions/Answers in this thesis:

✓ What can be considered, in a context of multidimensional portfolios, as the analogous of a “worst case” scenario and a related “tail distribution”? (multidimensional quantiles).

✓ Measures for risks with heterogeneous characteristics especially in an “external risks problem”?

✓ How to model the tail dependence, that is, the dependence between extreme events, occurring with low frequency but having a high impact?

✓ How to build measures with properties that turn to be consistent with existing properties on univariate setting?

✓ Particular attention to consistency results, asymptotic properties, illustrations of results on several real/simulated data sets.
Questions/Answers in this thesis:

✓ What can be considered, in a context of multidimensional portfolios, as the analogous of a “worst case” scenario and a related “tail distribution”? (multidimensional quantiles).

✓ Measures for risks with heterogeneous characteristics especially in an “external risks problem”?

✓ How to model the tail dependence, that is, the dependence between extreme events, occurring with low frequency but having a high impact?

✓ How to build measures with properties that turn to be consistent with existing properties on univariate setting?

✓ Particular attention to consistency results, asymptotic properties, illustrations of results on several real/simulated data sets.
1. **Introduction**

2. **Estimating Bivariate Tail: a copula based approach**
   - Framework
   - Estimating the tail of bivariate distributions
   - Convergence results
   - Illustrations with simulated and real data

3. **A multivariate extension of Value-at-Risk and Conditional-Tail-Expectation**

4. **Plug-in estimation of level sets in a non-compact setting with applications in multivariate risk theory**

5. **Perspectives**
Estimating $\mathbb{P}[X \leq x, Y \leq y]$, for $x, y$ large enough

**Goal**: To construct and analyze an estimator for the joint upper tail of a bivariate distribution function.

**Two challenges**: The joint tail region

1. may be so extreme that it contains no actual data points.
2. may exhibit both asymptotic dependence and asymptotic independence.

**Idea**: A general extension of the Peaks-Over-Threshold method.

**Tools**:

✓ A *two-dimensional version* of the Pickands-Balkema-de Haan Theorem.
✓ Dependence modeled by copulae and in particular by Upper Tail Dependence Copula (UTDC):

$$C_{u}^{up}(x, y) := \mathbb{P}[X \leq F_{X, u}^{-1}(x), Y \leq F_{Y, u}^{-1}(y) \mid X > u, Y > u],$$

with $F_{X, u}(x) := \mathbb{P}[X \leq x \mid X > u, Y > u]$ (see Juri & Wüthrich, 2004; Charpentier & Juri, 2006; Charpentier & Segers, 2007).
Estimating $\mathbb{P}[X \leq x, Y \leq y]$, for $x, y$ large enough

**Goal:** To construct and analyze an estimator for the joint upper tail of a bivariate distribution function.

**Two challenges:** The joint tail region

1. may be so extreme that it contains no actual data points.
2. may exhibit both asymptotic dependence and asymptotic independence.

**Idea:** A general extension of the Peaks-Over-Threshold method.

**Tools:**

- A *two-dimensional version* of the Pickands-Balkema-de Haan Theorem.
- Dependence modeled by copulae and in particular by Upper Tail Dependence Copula (UTDC):

$$C_{\upsilon}^{\text{up}}(x, y) := \mathbb{P}[X \leq F_{X, \upsilon}^{-1}(x), Y \leq F_{Y, \upsilon}^{-1}(y) | X > \upsilon, Y > \upsilon],$$

with $F_{X, \upsilon}(x) := \mathbb{P}[X \leq x | X > \upsilon, Y > \upsilon]$ (see Juri & Wüthrich, 2004; Charpentier & Juri, 2006; Charpentier & Segers, 2007).
Summary of results

✓ Construction of a two-dimensional tail estimator, study of its asymptotic properties.

✓ Using bivariate Regular Variation theory, the limit copula can be parameterized via a dependence parameter $\theta$ and some univariate functions.

Other possible approaches:

✓ Multivariate generalized Pareto distribution (Rootzen & Tajvidi, 2010) but the estimation of scaling parameters has to be addressed first.

✓ Ledford & Tawn’s models (e.g. see Beirlant et al., 2011).

✓ Based on the characterization of Resnick (1987):

$$\widehat{F}_1^*(y_1, y_2) = \exp\{-\widehat{l}_n(-\log(\widehat{F}_{Y_1}^*(y_1)), -\log(\widehat{F}_{Y_2}^*(y_2)))\},$$

(see Ledford & Tawn, 1996)

Alternative model based on regularity conditions of the copula and on a “distributional approach” to the joint tail.
Summary of results

✓ Construction of a two-dimensional tail estimator, study of its asymptotic properties.
✓ Using bivariate Regular Variation theory, the limit copula can be parameterized via a dependence parameter $\theta$ and some univariate functions.

Other possible approaches:

✓ Multivariate generalized Pareto distribution (Rootzen & Tajvidi, 2010) but the estimation of scaling parameters has to be addressed first.
✓ Ledford & Tawn’s models (e.g. see Beirlant et al., 2011).
✓ Based on the characterization of Resnick (1987):

$$\hat{F}_1^*(y_1, y_2) = \exp\{-\eta_n(-\log(\hat{F}_{Y_1}^*(y_1)), -\log(\hat{F}_{Y_2}^*(y_2)))\},$$

(see Ledford & Tawn, 1996)

Alternative model based on regularity conditions of the copula and on a “distributional approach” to the joint tail.
Summary of results

✓ Construction of a two-dimensional tail estimator, study of its asymptotic properties.

✓ Using bivariate Regular Variation theory, the limit copula can be parameterized via a dependence parameter \( \theta \) and some univariate functions.

Other possible approaches:

✓ Multivariate generalized Pareto distribution (Rootzen & Tajvidi, 2010) but the estimation of scaling parameters has to be addressed first.

✓ Ledford & Tawn’s models (e.g. see Beirlant et al., 2011).

✓ Based on the characterization of Resnick (1987):

\[
\hat{F}_1^*(y_1, y_2) = \exp\{-\hat{I}_n(-\log(\hat{F}_{Y_1}^*(y_1)), -\log(\hat{F}_{Y_2}^*(y_2)))\},
\]

(see Ledford & Tawn, 1996)

Alternative model based on regularity conditions of the copula and on a “distributional approach” to the joint tail.
Standing assumptions

✓ $X$ and $Y$ are two continuous real valued random variables, with marginal distributions, $F_X$, $F_Y$, and copula $C$.
✓ $F_X \in \text{MDA}(H_{\xi_1})$, $F_Y \in \text{MDA}(H_{\xi_2})$.
✓ $C$ as in the following Proposition:

Proposition

If \( \lim_{u \to 1} \frac{C^*(x(1-u), y(1-u))}{C^*(1-u, 1-u)} = G(x, y) \), for all $x, y > 0$, then

\[
\lim_{u \to 1} C_u^{up}(x, y) = x + y - 1 + G(g_X^{-1}(1 - x), g_Y^{-1}(1 - y)) := C^* G(x, y),
\]

where $g_X(x) := G(x, 1)$, $g_Y(y) := G(1, y)$. Moreover there is a constant $\theta > 0$ such that, for $x > 0$

\[
G(x, y) = \begin{cases} 
x^\theta g_Y\left(\frac{y}{x}\right) & \text{for } \frac{y}{x} \in [0, 1], 
y^\theta g_X\left(\frac{x}{y}\right) & \text{for } \frac{y}{x} \in (1, \infty).
\end{cases}
\]
A two dimensional Pickands-Balkema-de Haan Theorem

**Theorem**

*Under the standing assumptions,*

\[
\sup_{\mathcal{A}} \mathbb{P}[X - u \leq x, Y - F_Y^{-1}(F_X(u)) \leq y | X > u, Y > F_Y^{-1}(F_X(u))] - C^* G(1 - g_X(1 - V_{\xi_1,a_1}(u)(x)), 1 - g_Y(1 - V_{\xi_2,a_2}(F_Y^{-1}(F_X(u)))(y))) \quad (u \to x_{F_X} \to 0),
\]

\(\mathcal{A} := \{(x, y) : 0 < x \leq x_{F_X} - u, 0 < y \leq x_{F_Y} - F_Y^{-1}(F_X(u))\}, \) with \(x_{F_X} := \sup\{x \in \mathbb{R} | F_X(x) < 1\}, \) \(x_{F_Y} := \sup\{y \in \mathbb{R} | F_Y(y) < 1\}, \) \(V_{\xi_1,a_1}(\cdot) \) (resp. \(V_{\xi_2,a_2}(\cdot)\)) is the univariate GPD distribution with parameters \(\xi_1\) (resp. \(\xi_2\)) and \(a_1(\cdot)\) (resp. \(a_2(\cdot)\)) of \(X\) (resp. \(Y\)).

**Proof:** We generalize the proof by Jury & Wüthrich (2004) in the case of a symmetric copula and same marginal distributions.
Construction of the tail estimator: main steps

✓ Stable tail dependence function \( l \) (Huang, 1992):
\[
\lim_{t \to 0} \frac{1}{t} \left[ 1 - F_X(X) \leq tx \text{ or } 1 - F_Y(Y) \leq ty \right] := l(x, y).
\]

✓ The functions \( g_X, g_Y, G \) are estimated, using \( R(x, y) = x + y - l(x, y) \), by
\[
\hat{g}_X(x) = \frac{\hat{R}(x,1)}{\hat{R}(1,1)}, \quad \hat{g}_Y(x) = \frac{\hat{R}(1,y)}{\hat{R}(1,1)}, \quad \hat{G}(x,y) = \frac{\hat{R}(x,y)}{\hat{R}(1,1)}, \quad \Rightarrow \hat{\theta}
\]
(\( \hat{R} \) as in Einmahl et al., 2006).

✓ Upper tail dependence coefficient
\[
\lambda = R(1,1) \begin{cases} 
> 0, & \text{Asymptotic dependence case,} \\
= 0, & \text{Asymptotic independence case.}
\end{cases}
\]

✓ Consistency results for \( \hat{G}, \hat{g}_X \) and \( \hat{g}_Y \), both for \( \lambda > 0 \) and \( \lambda = 0 \) (second order condition for \( \lim_{t \to 0} \frac{C^*(tx,ty)}{C^*(t,t)} = G(x,y) \), see Draisma et al., 2004).

✓ We get an estimator for \( \theta \), dependence parameter.
Construction of the tail estimator: main steps

✓ Stable tail dependence function $l$ (Huang, 1992):
\[
\lim_{t \to 0} \frac{1}{t} [1 - F_X(X) \leq tx \text{ or } 1 - F_Y(Y) \leq ty] := l(x, y).
\]

✓ The functions $g_X$, $g_Y$, $G$ are estimated, using $R(x, y) = x + y - l(x, y)$, by
\[
\hat{g}_X(x) = \frac{\hat{R}(x,1)}{\hat{R}(1,1)}, \quad \hat{g}_Y(x) = \frac{\hat{R}(1,y)}{\hat{R}(1,1)}, \quad \hat{G}(x, y) = \frac{\hat{R}(x,y)}{\hat{R}(1,1)}, \quad \Rightarrow \hat{\theta}
\]
($\hat{R}$ as in Einmahl et al., 2006).

✓ Upper tail dependence coefficient
\[
\lambda = R(1, 1) \begin{cases} > 0, & \text{Asymptotic dependence case,} \\ = 0, & \text{Asymptotic independence case.} \end{cases}
\]

✓ Consistency results for $\hat{G}$, $\hat{g}_X$ and $\hat{g}_Y$, both for $\lambda > 0$ and $\lambda = 0$ (second order condition for $\lim_{t \to 0} \frac{C^*(tx,ty)}{C^*(t,t)} = G(x, y)$, see Draisma et al., 2004).

✓ We get an estimator for $\theta$, dependence parameter.
Construction of the tail estimator: main steps

✓ Stable tail dependence function \( I \) (Huang, 1992):
\[
\lim_{t \to 0} \frac{1}{t} \left[ 1 - F_X(X) \leq tx \text{ or } 1 - F_Y(Y) \leq ty \right] := I(x, y).
\]

✓ The functions \( g_X, g_Y, G \) are estimated, using \( R(x, y) = x + y - I(x, y) \), by
\[
\hat{g}_X(x) = \frac{\hat{R}(x, 1)}{\hat{R}(1, 1)}, \quad \hat{g}_Y(x) = \frac{\hat{R}(1, y)}{\hat{R}(1, 1)}, \quad \hat{G}(x, y) = \frac{\hat{R}(x, y)}{\hat{R}(1, 1)}, \quad \Rightarrow \hat{\theta}
\]
(\( \hat{R} \) as in Einmahl et al., 2006).

✓ Upper tail dependence coefficient
\[
\lambda = R(1, 1) \begin{cases} 
> 0, & \text{Asymptotic dependence case,} \\
= 0, & \text{Asymptotic independence case.}
\end{cases}
\]

✓ Consistency results for \( \hat{G}, \hat{g}_X \) and \( \hat{g}_Y \), both for \( \lambda > 0 \) and \( \lambda = 0 \) (second order condition for \( \lim_{t \to 0} \frac{C^*(tx, ty)}{C^*(t, t)} = G(x, y) \), see Draisma et al., 2004).

✓ We get an estimator for \( \theta \), dependence parameter.
A new tail estimator

✓ A “high” threshold \( u \);

✓ Define \( \widehat{u}_Y = \widehat{F}_Y^{-1}(\widehat{F}_X(u)) \), with \( \widehat{F}_X(u) \) the empirical distribution function and \( \widehat{F}_Y^{-1} \) the empirical quantile function of \( Y \).

✓ \( \widehat{k}_X, \widehat{\sigma}_X \) (resp. \( \widehat{k}_Y, \widehat{\sigma}_Y \)) the MLE based on the excesses of \( X \) (resp. \( Y \)).

✓ \( \widehat{F}_X^*(x) \) (resp. \( \widehat{F}_Y^*(y) \)) the univariate POT estimator (see McNeil, 1999):

\[
\widehat{F}_X^*(x) = (1 - \widehat{F}_X(u)) V_{\widehat{k}, \widehat{\sigma}}(x - u) + \widehat{F}_X(u), \quad \text{for } x > u.
\]

✓ Note

\[
\widehat{F}_1^*(u, y) = \exp\{ - \widehat{l}_n(- \log(\widehat{F}_X(u)), - \log(\widehat{F}_Y^*(y))) \},
\]

and

\[
\widehat{F}_2^*(x, \widehat{u}_Y) = \exp\{ - \widehat{l}_n(- \log(\widehat{F}_X^*(x)), - \log(\widehat{F}_Y(\widehat{u}_Y))) \}.
\]
A new tail estimator

✓ A “high” threshold \( u \);

✓ Define \( \hat{u}_Y = \hat{F}^{-1}_Y(\hat{F}_X(u)) \), with \( \hat{F}_X(u) \) the empirical distribution function and \( \hat{F}^{-1}_Y \) the empirical quantile function of \( Y \).

✓ \( \hat{k}_X, \hat{\sigma}_X \) (resp. \( \hat{k}_Y, \hat{\sigma}_Y \)) the MLE based on the excesses of \( X \) (resp. \( Y \)).

✓ \( \hat{F}_X^*(x) \) (resp. \( \hat{F}_Y^*(y) \)) the univariate POT estimator (see McNeil, 1999):

\[
\hat{F}_X^*(x) = (1 - \hat{F}_X(u)) V_{k_\sigma}(x - u) + \hat{F}_X(u), \quad \text{for } x > u.
\]

✓ Note

\[
\hat{F}_1^*(u, y) = \exp\{ -\hat{I}_n(-\log(\hat{F}_X(u)), -\log(\hat{F}_Y^*(y))) \},
\]

and

\[
\hat{F}_2^*(x, \hat{u}_Y) = \exp\{ -\hat{I}_n(-\log(\hat{F}_X^*(x)), -\log(\hat{F}_Y(\hat{u}_Y))) \}.
\]
A new tail estimator

✓ A “high” threshold $u$;

✓ Define $\hat{u}_Y = \hat{F}_Y^{-1}(\hat{F}_X(u))$, with $\hat{F}_X(u)$ the empirical distribution function and $\hat{F}_Y^{-1}$ the empirical quantile function of $Y$.

✓ $\hat{k}_X, \hat{\sigma}_X$ (resp. $\hat{k}_Y, \hat{\sigma}_Y$) the MLE based on the excesses of $X$ (resp. $Y$).

✓ $\hat{F}_X^*(x)$ (resp. $\hat{F}_Y^*(y)$) the univariate POT estimator (see McNeil, 1999):

\[
\hat{F}_X^*(x) = (1 - \hat{F}_X(u)) V_{\hat{k}, \hat{\sigma}}(x - u) + \hat{F}_X(u), \quad \text{for } x > u.
\]

✓ Note

\[
\hat{F}_1^*(u, y) = \exp\{-\hat{\lambda}_n(-\log(\hat{F}_X(u)), -\log(\hat{F}_Y^*(y)))\},
\]

and

\[
\hat{F}_2^*(x, \hat{u}_Y) = \exp\{-\hat{\lambda}_n(-\log(\hat{F}_X^*(x)), -\log(\hat{F}_Y(\hat{u}_Y)))\}.
\]
A new tail estimator

✓ A “high” threshold $u$;

✓ Define $\hat{u}_Y = \hat{F}^{-1}_Y(\hat{F}_X(u))$, with $\hat{F}_X(u)$ the empirical distribution function and $\hat{F}^{-1}_Y$ the empirical quantile function of $Y$.

✓ $\hat{k}_X$, $\hat{\sigma}_X$ (resp. $\hat{k}_Y$, $\hat{\sigma}_Y$) the MLE based on the excesses of $X$ (resp. $Y$).

✓ $\hat{F}^*_X(x)$ (resp. $\hat{F}^*_Y(y)$) the univariate POT estimator (see McNeil, 1999):

\[
\hat{F}^*_X(x) = (1 - \hat{F}_X(u)) V_{k,\hat{\sigma}}(x - u) + \hat{F}_X(u), \quad \text{for } x > u.
\]

✓ Note

\[
\hat{F}^*_1(u, y) = \exp\{-\hat{I}_n(- \log(\hat{F}_X(u)), - \log(\hat{F}^*_Y(y)))\},
\]

and

\[
\hat{F}^*_2(x, \hat{u}_Y) = \exp\{-\hat{I}_n(- \log(\hat{F}^*_X(x)), - \log(\hat{F}_Y(\hat{u}_Y)))\}.
\]
A new tail estimator

✓ A “high” threshold $u$;

✓ Define $\hat{u}_Y = \hat{F}_Y^{-1}(\hat{F}_X(u))$, with $\hat{F}_X(u)$ the empirical distribution function and $\hat{F}_Y^{-1}$ the empirical quantile function of $Y$.

✓ $\hat{k}_X, \hat{\sigma}_X$ (resp. $\hat{k}_Y, \hat{\sigma}_Y$) the MLE based on the excesses of $X$ (resp. $Y$).

✓ $\hat{F}_X^*(x)$ (resp. $\hat{F}_Y^*(y)$) the univariate POT estimator (see McNeil, 1999): 

$$\hat{F}_X^*(x) = (1 - \hat{F}_X(u)) V_{k, \hat{\sigma}}(x - u) + \hat{F}_X(u), \text{ for } x > u.$$ 

✓ Note

$$\hat{F}_1^*(u, y) = \exp\{-\hat{t}_n(-\log(\hat{F}_X(u)), -\log(\hat{F}_Y^*(y)))\},$$

and

$$\hat{F}_2^*(x, \hat{u}_Y) = \exp\{-\hat{t}_n(-\log(\hat{F}_X^*(x)), -\log(\hat{F}_Y(\hat{u}_Y)))\}.$$
A new tail estimator

We estimate \( F(x, y) \), for \( x > u \) and \( y > \hat{u}_Y \), by:

\[
\hat{F}^*(x, y) = \left( \frac{1}{n} \sum_{i=1}^{n} 1\{X_i > u, Y_i > \hat{u}_Y\} \right) \left( 1 - \hat{g}_X (1 - V_{\tilde{\xi}_X, \tilde{\sigma}_X} (x - u)) \right) \\
- \hat{g}_Y (1 - V_{\tilde{\xi}_Y, \tilde{\sigma}_Y} (y - \hat{u}_Y)) + \hat{G}(1 - V_{\tilde{\xi}_X, \tilde{\sigma}_X} (x - u), 1 - V_{\tilde{\xi}_Y, \tilde{\sigma}_Y} (y - \hat{u}_Y)) \\
+ \hat{F}^*_1(u, y) + \hat{F}^*_2(x, \hat{u}_Y) - \frac{1}{n} \sum_{i=1}^{n} 1\{X_i \leq u, Y_i \leq \hat{u}_Y\} ,
\]

where \( \hat{g}_X, \hat{g}_Y, \hat{G}, \hat{F}^*_1, \hat{F}^*_2 \) are new estimators.
Consistency properties for our bivariate POT estimator

\[ \lambda > 0 \] standing assumptions, first and second order conditions on the marginal laws and Smith’s hypothesis on the thresholds.

\[ \lambda = 0 \] standing assumptions, first and second order conditions on the marginal laws, second order condition on the joint distribution and Smith’s hypothesis on the thresholds.

(1) The estimators of the objects appearing in the tail representation \( (\hat{g}_X, \hat{g}_Y, \hat{G}) \) are shown to be (uniformly) consistent at certain (different) rates both in the asymptotic dependent case and in the asymptotic independent one.

(2) Results in (1) are transferred to consistency results for the joint tail \( \hat{F}^* \) and for the dependence parameter \( \hat{\theta} \).
Sensitivity of $\hat{\theta}_x$ to the sequence $k_n$

Clayton copula. Mean curve on 100 samples of size $n = 1000$.

Figure: Clayton copula with parameter 0.05: (left) estimator for $\theta$, $(k, \hat{\theta}_x)$ with $x = 0.7$, mean curve (full line) and $+/-$ the empirical standard deviation (dashed lines); (right) mean squared error for $\hat{\theta}_x$ with $x = 0.7$. 
Wave height vs Water level: asymptotic independent case

Data: recorded during 828 storm events spread over 13 years in front of the Dutch coast near the town of Petten.

Figure: (left) $\hat{\theta}_{0.91} = \hat{\theta}_{0.11}$; (right) $\hat{F}^*(5.93, 1.87)$ (full line), $\hat{F}_1^*(5.93, 1.87)$ (dashed line), with the empirical probability indicated with a horizontal line.
Introduction

Estimating Bivariate Tail: a copula based approach

A multivariate extension of Value-at-Risk and Conditional Tail Expectation
- Framework
- Multidimensional risk measures
- Properties of our measures

Plug-in estimation of level sets in a non-compact setting with applications in multivariate risk theory

Perspectives
Multivariate extension of classical univariate risk measures

**Goal:** Extension of $\text{VaR}_\alpha(X) = F_X^{-1}(\alpha)$ and $\text{CTE}_\alpha(X) = \mathbb{E}[X | X \geq \text{VaR}_\alpha(X)]$.

**Ideas:**
- Suitable definition of quantiles for multi-risk portfolios.
- To quantify risks in a much more synthetic and parsimonious way.
- To model heterogeneous risks.

**Tools:**
- $\alpha$-upper level sets $L(\alpha) = \{x \in \mathbb{R}^d_+ : F_X(x) \geq \alpha\}$ and quantile curves $\partial L(\alpha)$.
- Axiomatic approach: Artzner et al. (1999)’s properties.
- “Distributional approach” by: $K(\alpha) = \mathbb{P}[F(X) \leq \alpha]$ (Nelsen et al., 2003).
- Stochastic orders to compare risks: supermodular order, dangerousness order, Stochastic dominance order, positive dependence (Joe, 1997).
Summary of results

✓ Construction of $d$–dimensional risk measures (VaR and CTE), study axiomatic properties, behaviors in terms of risk level and dependence structure.

✓ Some analytical formulas are provided; some explicit cases are analyzed.

Other possible approaches:

✓ Multivariate Value-at-Risk as quantile curve (Embrechts & Puccetti, 2006; Nappo & Spizzichino, 2009).

✓ Several multivariate generalizations of CTE to model problems of capital allocation in a portfolio when the risks are dependent. For $i = 1, \ldots, d$

$$
\text{CTE}_{\alpha}^{\text{sum}}(X_i) = \mathbb{E}[X_i | S > Q_S(\alpha)], \quad \text{CTE}_{\alpha}^{\text{min}}(X_i) = \mathbb{E}[X_i | X_{(1)} > Q_{X_{(1)}}(\alpha)], \quad \text{CTE}_{\alpha}^{\text{max}}(X_i) = \mathbb{E}[X_i | X_{(d)} > Q_{X_{(d)}}(\alpha)],
$$

where $S = X_1 + \cdots + X_d$, $X_{(1)} = \min\{X_1, \ldots, X_d\}$ and $X_{(d)} = \max\{X_1, \ldots, X_d\}$. 
Summary of results

✓ Construction of $d$–dimensional risk measures (VaR and CTE), study axiomatic properties, behaviors in terms of risk level and dependence structure.

✓ Some analytical formulas are provided; some explicit cases are analyzed.

Other possible approaches:

✓ Multivariate Value-at-Risk as quantile curve (Embrechts & Puccetti, 2006; Nappo & Spizzichino, 2009).

✓ Several multivariate generalizations of CTE to model problems of capital allocation in a portfolio when the risks are dependent. For $i = 1, \ldots, d$

$$CTE_{\alpha}^{\text{sum}}(X_i) = \mathbb{E}[X_i \mid S > Q_S(\alpha)], \quad CTE_{\alpha}^{\min}(X_i) = \mathbb{E}[X_i \mid X_{(1)} > Q_{X_{(1)}}(\alpha)],$$

$$CTE_{\alpha}^{\max}(X_i) = \mathbb{E}[X_i \mid X_{(d)} > Q_{X_{(d)}}(\alpha)],$$

where $S = X_1 + \cdots + X_d$, $X_{(1)} = \min\{X_1, \ldots, X_d\}$ and $X_{(d)} = \max\{X_1, \ldots, X_d\}$. 
Quantile generalizations: quantile curves

Tibiletti (1993), Fernández-Ponce & Suárez-Lloréns (2002) and Belzunce et al. (2007) defined a multivariate quantile as a set of points which accumulate the same probability for a fixed orthant (called quantile curves).

Definition (Quantile curve)

For $\alpha \in (0,1)$ and a $d$–variate distribution function $F$, we define the $d$–variate quantile at probability level $\alpha$, $\partial L(\alpha) := \partial \{ F(x) \geq \alpha \}$, with $\alpha \in (0,1)$.

Remark: 1) Natural extension in dimension $d$, 2) “metric-free”, 3) for symmetric and non-symmetric distribution function, 4) provide a data segmentation of predefined size.

De Haan & Huang (1995), Chebana & Ouarda (2011) $\Rightarrow$ quantile curves to model hydrological events.
A multivariate Value-at-Risk and Conditional-Tail-Expectation

Definition

Consider a random vector $\mathbf{X}$ satisfying the regularity conditions. For $\alpha \in (0, 1)$, we define:

$$\text{VaR}_\alpha(\mathbf{X}) = \left( \begin{array}{c} \mathbb{E}[X_1 | \mathbf{X} \in \partial L(\alpha)] \\ \vdots \\ \mathbb{E}[X_d | \mathbf{X} \in \partial L(\alpha)] \end{array} \right) = \left( \begin{array}{c} \mathbb{E}[X_1 | F(\mathbf{X}) = \alpha] \\ \vdots \\ \mathbb{E}[X_d | F(\mathbf{X}) = \alpha] \end{array} \right),$$

$$\text{CTE}_\alpha(\mathbf{X}) = \left( \begin{array}{c} \mathbb{E}[X_1 | \mathbf{X} \in L(\alpha)] \\ \vdots \\ \mathbb{E}[X_d | \mathbf{X} \in L(\alpha)] \end{array} \right) = \left( \begin{array}{c} \mathbb{E}[X_1 | F(\mathbf{X}) \geq \alpha] \\ \vdots \\ \mathbb{E}[X_d | F(\mathbf{X}) \geq \alpha] \end{array} \right),$$

where $\partial L(\alpha)$ is the boundary of the $\alpha$-level set $L(\alpha)$ of $F$. 
### Properties of our measures

<table>
<thead>
<tr>
<th></th>
<th>( \text{Var}_\alpha(X) )</th>
<th>( \text{CTE}_\alpha(X) )</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Invariance properties</strong> (( c \in \mathbb{R}^d_+ )):</td>
<td>( \text{Var}<em>\alpha(cX) = c \text{Var}</em>\alpha(X) ).</td>
<td>( \text{CTE}<em>\alpha(cX) = c \text{CTE}</em>\alpha(X) ).</td>
</tr>
<tr>
<td></td>
<td>( \text{Var}<em>\alpha(c + X) = c + \text{Var}</em>\alpha(X) ).</td>
<td>( \text{CTE}<em>\alpha(c + X) = c + \text{CTE}</em>\alpha(X) ).</td>
</tr>
<tr>
<td><strong>Lower bounds</strong>:</td>
<td>( \text{Var}<em>i^\alpha(X) \geq \text{Var}</em>\alpha(X_i), \ \forall \alpha \in (0,1) ).</td>
<td>( \text{CTE}<em>i^\alpha(X) \geq \text{Var}</em>\alpha(X_i), \ \forall \alpha \in (0,1) ).</td>
</tr>
<tr>
<td></td>
<td><strong>Analytical closed-form formulas</strong> for ( \text{Var}<em>\alpha(X) ) and ( \text{CTE}</em>\alpha(X) ).</td>
<td><strong>Safety loading</strong>: ( \text{CTE}_i^\alpha(X) \geq \mathbb{E}[X_i] ). ( \text{CTE}_0(X) = \mathbb{E}[X] ).</td>
</tr>
<tr>
<td><strong>Risk level</strong></td>
<td>( \text{Var}_i^\alpha(X) ) is a non-decreasing function of ( \alpha ).</td>
<td>( \text{CTE}_i^\alpha(X) ) is a non-decreasing function of ( \alpha ).</td>
</tr>
<tr>
<td><strong>Dependence structure</strong></td>
<td>Comonotonic case: ( \text{Var}<em>i^\alpha(X) = \text{Var}</em>\alpha(X_i), \ \forall \alpha \in (0,1) ).</td>
<td>Comonotonic case: ( \text{CTE}<em>i^\alpha(X) = \text{CTE}</em>\alpha(X_i), \ \forall \alpha \in (0,1) ).</td>
</tr>
<tr>
<td></td>
<td>For a fixed copula ( C ) and ( X_i \overset{d}{=} Y_i ): ( \text{Var}<em>i^\alpha(X) = \text{Var}</em>\alpha(Y), \ \forall \alpha \in (0,1) ).</td>
<td>For a fixed copula ( C ) and ( X_i \overset{d}{=} Y_i ): ( \text{CTE}<em>i^\alpha(X) = \text{CTE}</em>\alpha(Y), \ \forall \alpha \in (0,1) ).</td>
</tr>
<tr>
<td></td>
<td>For a fixed copula ( C ) and ( X_i \overset{st}{=} Y_i ): ( \text{Var}<em>i^\alpha(X) \leq \text{Var}</em>\alpha(Y), \ \forall \alpha \in (0,1) ).</td>
<td>For a fixed copula ( C ) and ( X_i \overset{d}{=} Y_i ): ( \text{CTE}<em>i^\alpha(X) \leq \text{CTE}</em>\alpha(Y), \ \forall \alpha \in (0,1) ).</td>
</tr>
<tr>
<td>Introduction</td>
<td>Estimating Bivariate Tail</td>
<td>Multivariate VaR and CTE</td>
</tr>
<tr>
<td>--------------</td>
<td>---------------------------</td>
<td>--------------------------</td>
</tr>
</tbody>
</table>

- Framework
- Consistency results
- A possible application
- Illustrations with simulated and real data
Plug-in estimation of level sets

**Goal:** To build a consistent estimator of

\[ L(c) := \{ x \in \mathbb{R}^2_+ : F(x) \geq c \}, \quad \text{for } c \in (0,1). \]

**Idea:** We consider a plug-in approach that is \( L(c) \) is estimated by

\[ L_n(c) := \{ x \in \mathbb{R}^2_+ : F_n(x) \geq c \}, \quad \text{for } c \in (0,1), \]

where \( F_n \) is a consistent estimator of \( F \).

**Literature and background:**

- Cavalier (1997), Laloë (2009) for regression function in a compact setting.
- Cuevas et al. (2006) for general compact level sets.
Notation and preliminary results

We state consistency results with respect to two “physical proximity” criteria between sets: the **Hausdorff distance** and the **volume of the symmetric difference**.

✓ Problem of compactness property for the level sets we estimate.

Figure: (left) Hausdorff distance between sets $X$ and $Y$; (right) $\lambda(X \triangle Y)$, where $\lambda$ stands for the Lebesgue measure on $\mathbb{R}^2$ and $\triangle$ for the symmetric difference.
Notation and preliminary results

There exist $\gamma > 0$ and $A > 0$ such that, if $|t - c| \leq \gamma$ then $\forall \ T > 0$ such that $\{F = c\}^T \neq \emptyset$ and $\{F = t\}^T \neq \emptyset$,

$$d_H(\{F = c\}^T, \{F = t\}^T) \leq A |t - c|.$$  

Assumption $H$ is satisfied under mild conditions:

Proposition

Let $c \in (0, 1)$. Let $F \in \mathcal{F}$ be twice differentiable on $\mathbb{R}^2_+$. Assume there exist $r > 0$, $\lambda > 0$ such that $m^\nabla := \inf_{x \in E} \| (\nabla F)_x \| > 0$ and $M_H := \sup_{x \in E} \| (HF)_x \| < \infty$, with $E := B( \{x \in \mathbb{R}^2_+ : |F - c| \leq r \}, \lambda)$. Then

$F$ satisfies Assumption $H$, with $A = \frac{2}{m^\nabla}$.  

Consistency in Hausdorff distance

From now on we note, for $n \in \mathbb{N}^*$, and for $T > 0$,

$$\|F - F_n\|_\infty = \sup_{x \in \mathbb{R}_+^2} |F(x) - F_n(x)|, \quad \|F - F_n\|_\infty^T = \sup_{x \in [0, T]^2} |F(x) - F_n(x)|.$$ 

**Theorem (Consistency in Hausdorff distance)**

Let $c \in (0, 1)$. Let $F \in \mathcal{F}$ be twice differentiable on $\mathbb{R}_+^{2*}$. Assume that there exist $r > 0$, $\zeta > 0$ such that $m^\triangledown > 0$ and $M_H < \infty$. Let $T_1 > 0$ such that for all $t : |t - c| \leq r$, $\partial L(t)^T_{T_1} \neq \emptyset$. Let $(T_n)_{n \in \mathbb{N}^*}$ be an increasing sequence of positive values. Assume that, for each $n$, $F_n$ is continuous with probability one and that $\|F - F_n\|_\infty \to 0$, a.s. Then for $n$ large enough,

$$d_H(\partial L(c)^T_{T_n}, \partial L_n(c)^T_{T_n}) \leq 6 A \|F - F_n\|_\infty^{T_n}, \text{ a.s., where } A = \frac{2}{m^\triangledown}.$$
**L₁ consistency**

Let us introduce the following assumption:

**A1** There exist positive increasing sequences \((v_n)_{n \in \mathbb{N}^*}\) and \((T_n)_{n \in \mathbb{N}^*}\) such that

\[
\overline{v_n} \int_{[0, T_n]^2} | F - F_n |^p \lambda(dx) \overset{\mathbb{P}}{\rightarrow} 0, \quad \text{for some } 1 \leq p < \infty.
\]

**Theorem (Consistency in volume)**

Let \(c \in (0, 1)\). Let \(F \in \mathcal{F}\) be twice differentiable on \(\mathbb{R}^2_+\). Assume that there exist \(r > 0, \zeta > 0\) such that \(m^\nabla > 0\) and \(M_H < \infty\). Let \((v_n)_{n \in \mathbb{N}^*}\) and \((T_n)_{n \in \mathbb{N}^*}\) positive increasing sequences such that Assumption **A1** is satisfied and that for all \(t: |t - c| \leq r, \partial L(t)^T \neq \emptyset\). Then it holds that

\[
p_n d_\lambda(L(c)^T_n, L_n(c)^T_n) \overset{\mathbb{P}}{\rightarrow} 0,
\]

with \(p_n\) an increasing positive sequence such that \(p_n = o \left( \frac{1}{v_n^{p+1}} / T_n^{p+1} \right)\).
Estimating our bivariate $\hat{\text{CTE}}_{\alpha}(X, Y)$

Let $\hat{\text{CTE}}_{\alpha}(X, Y) = \left( \frac{\sum_{i=1}^{n} X_{i} 1\{(X_{i}, Y_{i}) \in L_n(\alpha)\}}{\sum_{i=1}^{n} 1\{(X_{i}, Y_{i}) \in L_n(\alpha)\}}, \frac{\sum_{i=1}^{n} Y_{i} 1\{(X_{i}, Y_{i}) \in L_n(\alpha)\}}{\sum_{i=1}^{n} 1\{(X_{i}, Y_{i}) \in L_n(\alpha)\}} \right)$, for $\alpha \in (0, 1)$.

**Theorem (Consistency of $\hat{\text{CTE}}_{\alpha}(X, Y)$)**

Under regularity properties of $(X, Y)$, Assumptions of Theorem Consistency volume and with the same notation, it holds that

$$
\beta_n \left| \text{CTE}_{\alpha}^{T_n}(X, Y) - \hat{\text{CTE}}_{\alpha}^{T_n}(X, Y) \right| \xrightarrow{\mathbb{P}} 0,
$$

where $\beta_n = \min\{ p_n^{\frac{r}{2(1+r)}}, a_n \}$, with $r > 0$ such that the density $f_{(X, Y)} \in L^{1+r}(\lambda)$ and $a_n = o\left(\sqrt{n}\right)$.

$T_n$ is essentially a compromise also to control $\left| \text{CTE}_{\alpha}(X, Y) - \text{CTE}_{\alpha}^{T_n}(X, Y) \right|$. 
Behavior of $\hat{\text{CTE}}_\alpha(X, Y)$ with respect to the risk level $\alpha$

$F_n$ empirical bivariate distribution function. $X \sim \mathcal{E}(1) \parallel Y \sim \mathcal{E}(2)$. $T_n = \ln(n)$.

<table>
<thead>
<tr>
<th>$\alpha$</th>
<th>0.1</th>
<th>0.24</th>
<th>0.38</th>
<th>0.52</th>
<th>0.66</th>
<th>0.88</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\hat{\sigma}$</td>
<td>(0.044, 0.022)</td>
<td>(0.069, 0.023)</td>
<td>(0.075, 0.038)</td>
<td>(0.104, 0.052)</td>
<td>(0.139, 0.071)</td>
<td>(0.251, 0.125)</td>
</tr>
<tr>
<td>RMSE</td>
<td>(0.043, 0.039)</td>
<td>(0.051, 0.042)</td>
<td>(0.044, 0.046)</td>
<td>(0.052, 0.052)</td>
<td>(0.057, 0.056)</td>
<td>(0.084, 0.082)</td>
</tr>
</tbody>
</table>

**Table**: Evolution of $\hat{\sigma}$ and RMSE in terms of $\alpha$, sample size $n = 1000$.

<table>
<thead>
<tr>
<th>$n$</th>
<th>500</th>
<th>1000</th>
<th>1500</th>
<th>2000</th>
<th>2500</th>
<th>5000</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\hat{\sigma}$</td>
<td>(0.614, 0.359)</td>
<td>(0.444, 0.308)</td>
<td>(0.431, 0.295)</td>
<td>(0.377, 0.168)</td>
<td>(0.241, 0.123)</td>
<td>(0.216, 0.121)</td>
</tr>
<tr>
<td>RMSE</td>
<td>(0.168, 0.189)</td>
<td>(0.123, 0.163)</td>
<td>(0.115, 0.161)</td>
<td>(0.099, 0.089)</td>
<td>(0.077, 0.079)</td>
<td>(0.063, 0.057)</td>
</tr>
</tbody>
</table>

**Table**: Evolution of $\hat{\sigma}$ and RMSE in terms of sample size $n$, $\alpha = 0.9$.

Here we need between 2000 and 2500 data to get the same performances as for lower levels $\alpha$. 
Estimating of level sets $L(\alpha)$ with simulated data

(a) $T_n = \ln(n)$

<table>
<thead>
<tr>
<th>$\alpha$</th>
<th>$n = 500$</th>
<th>$n = 1000$</th>
<th>$n = 2000$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.10</td>
<td>0.069</td>
<td>0.068</td>
<td>0.065</td>
</tr>
<tr>
<td>0.24</td>
<td>0.156</td>
<td>0.134</td>
<td>0.063</td>
</tr>
<tr>
<td>0.38</td>
<td>0.172</td>
<td>0.139</td>
<td>0.121</td>
</tr>
<tr>
<td>0.52</td>
<td>0.225</td>
<td>0.169</td>
<td>0.153</td>
</tr>
<tr>
<td>0.66</td>
<td>0.298</td>
<td>0.199</td>
<td>0.195</td>
</tr>
<tr>
<td>0.80</td>
<td>0.426</td>
<td>0.282</td>
<td>0.279</td>
</tr>
</tbody>
</table>

(b) $T_n = n^{0.45}$

<table>
<thead>
<tr>
<th>$\alpha$</th>
<th>$n = 500$</th>
<th>$n = 1000$</th>
<th>$n = 2000$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.10</td>
<td>0.229</td>
<td>0.248</td>
<td>0.188</td>
</tr>
<tr>
<td>0.24</td>
<td>0.361</td>
<td>0.298</td>
<td>0.209</td>
</tr>
<tr>
<td>0.38</td>
<td>0.411</td>
<td>0.357</td>
<td>0.339</td>
</tr>
<tr>
<td>0.52</td>
<td>0.734</td>
<td>0.689</td>
<td>0.743</td>
</tr>
<tr>
<td>0.66</td>
<td>0.849</td>
<td>0.752</td>
<td>0.762</td>
</tr>
<tr>
<td>0.80</td>
<td>1.071</td>
<td>1.039</td>
<td>1.124</td>
</tr>
</tbody>
</table>

Table: $X \sim \mathcal{E}(1) \perp Y \sim \mathcal{E}(2)$. Approximated $p_n \lambda(L(\alpha)^{T_n} \triangle L_n(\alpha)^{T_n})$, with $p_n = o(n^{\frac{1}{3}}/\ln(n)^{\frac{4}{3}})$.

✓ Influence of the choice of $T_n$.
✓ Here $p_n = o(n^{\frac{1}{3}}/\ln(n)^{\frac{4}{3}})$ is at least the convergence rate of $\lambda(L(\alpha)^{T_n} \triangle L_n(\alpha)^{T_n})$.
✓ Taking $T_n$ too large $\Rightarrow$ not a good approximation of $\lambda(L(\alpha)^{T_n} \triangle L_n(\alpha)^{T_n})$. 
Real data illustration

Real case: **Loss-ALAE data** in the log scale (Frees & Valdez, 1998). The data size is $n = 1500$. Let $F_n$ the empirical distribution function.

- Central Limit Theorems for $g_X$, $g_Y$, $G$, $\theta$, $\hat{F}^*$.  
- Optimal choice of $k_n$.  
- Deep comparisons of $\hat{F}^*$ with its competitors in literature: Ledford & Tawn (1996), Beirlant *et al.* (2011), etc.
Perspectives


✓ Comparisons of our multivariate CTE and VaR with existing multivariate generalizations of these measures, both theoretically and experimentally (applications on financial portfolios; micro-prudential versus macro-prudential approach, ...).

✓ Here we provide asymptotic results for a fixed level $c$. What about some uniform results?

✓ Problem of constant $A$ for $c \sim 1$.

✓ Package $R$ for level sets $L_n(c)$.

✓ A multivariate extension:

Thank you for your attention