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

http://hal.archives-ouvertes.fr/hal-00580624/fr/

Elena Di Bernardino Université Lyon 1, ISFA, Laboratoire SAF

joint works with

Areski Cousin, Université Lyon 1, ISFA, Laboratoire SAF

Thomas Laloë, Université de Nice Sophia-Antipolis

Véronique Maume-Deschamps, Université Lyon 1, ISFA, Laboratoire SAF

Clémentine Prieur, Université Joseph Fourier, Grenoble

Séminaire de Statistique, Institut de Mathématiques de Toulouse, 4 Octobre 2011
Contents

1 Univariate risk measures: VaR, CTE
   • Evaluating risks
   • Generalizations of CTE
   • Our Bivariate Conditional-Tail-Expectation

2 Bivariate risk measures
   • Bivariate Value-at-Risk
   • Bivariate Conditional-Tail-Expectation

3 Plug-in estimation
   • Estimation of level sets: literature and background
   • Notation and preliminary results
   • Consistency result in terms of the Hausdorff distance
   • $L_1$ consistency
   • Estimation of our bivariate Conditional-Tail-Expectation
   • Illustrations
Contents

1 Univariate risk measures: VaR, CTE
   - Evaluating risks
   - Generalizations of CTE
   - Our Bivariate Conditional-Tail-Expectation

2 Bivariate risk measures
   - Bivariate Value-at-Risk
   - Bivariate Conditional-Tail-Expectation

3 Plug-in estimation
   - Estimation of level sets: literature and background
   - Notation and preliminary results
   - Consistency result in terms of the Hausdorff distance
   - $L_1$ consistency
   - Estimation of our bivariate Conditional-Tail-Expectation
   - Illustrations
“A risk measure is a functional $\rho$ mapping a risk $X$ to a non-negative number $\rho[X]$, possibly infinite, representing the extra cash which has to be added to $X$ to make it acceptable”. Denuit et al. (2005), page 61.

Given an univariate continuous and strictly monotonic loss distribution function $F_X$,

$$\text{VaR}_\alpha(X) = Q_X(\alpha) = F_X^{-1}(\alpha), \quad \forall \alpha \in (0, 1).$$

Shortcoming of VaR measure:

- VaR is not a coherent risk measure (see Artzner, 1999).
- VaR does not give any information about thickness of the upper tail of $F$.

“On the other hand, the concept of VaR as a risk measure has problems for measuring extreme price movements. By definition, VaR only measures the distribution quantile, and disregards extreme loss beyond the VaR level. Thus, VaR may ignore important information regarding the tails of the underlying distributions. The BIS Committee on the Global Financial System (2000) identifies this problem as tail risk.” Yamai et Yoshiba (2002), page 182.
To overcome problems of VaR → Conditional-Tail-Expectation (CTE):

\[
CTE_\alpha(X) = \mathbb{E}[X \mid X \geq \text{VaR}_\alpha(X)] = \mathbb{E}[X \mid X \geq Q_X(\alpha)],
\]

(e.g. for properties of the \( CTE_\alpha(X) \) see Denuit et al., 2005).

Traditionally, risk measures are thought of as mappings from a set of real-valued random variables to the real numbers. However, it is often insufficient to consider a single real measure to quantify risks created by business activities, especially if the latter can be affected by other risks (see problem of solvability capital allocation for multi-branch insurance companies confronted to risks with different characteristics).

“Only in recent studies, the other dimension of systemic risk, the cross-sectional dimension, has caught attention. Because banks are interconnected, banking crises may occur simultaneously. This is regarded as a systemic risk on the cross-sectional dimension.” Zhou (2010).
Dependence and dimensional problems

- Modeling risks with qualitative evaluations are not necessarily expressed in a totally ordered universe.
- Risk measures take values in certain partially ordered cones (Cascos and Molchanov, 2007).
- Risk measures are assumed to be vector-valued (Jouini et al., 2004).
- Riskiness not only of the marginal distributions, but also of the joint distribution.

\[ \rho : (X, Y) \mapsto \left( \begin{array}{c} \rho^1[(X, Y)] \\ \rho^2[(X, Y)] \end{array} \right) \in \mathbb{R}^2_+ , \]

First question: What can be considered in a context of multidimensional portfolios as the analogous of “worst cases” or “tail distributions”?

⇒ Suitable definition of multi-dimensional quantiles.
Let $X = (X_1, ..., X_n)$ be a risk vector, where $X_i$ denotes risk (claim or loss) in subportfolio $i$, for $i = 1, \ldots, n$. In risk analysis for such a portfolio, we are not only interested in the $\text{CTE}_\alpha(X_i) = \mathbb{E}[X_i \mid X_i > Q_{X_i}(\alpha)]$ but also in

$$
\text{CTE}_\alpha(X_i \mid X_1 + \ldots + X_n) = \mathbb{E}[X_i \mid X_1 + \ldots + X_n > Q_{X_1+\ldots+X_n}(\alpha)],
$$

$$
\text{CTE}_\alpha(X_i \mid \min\{X_1, \ldots, X_n\}) = \mathbb{E}[X_i \mid \min\{X_1, \ldots, X_n\} > Q_{\min\{X_1,\ldots,X_n\}}(\alpha)],
$$

$$
\text{CTE}_\alpha(X_i \mid \max\{X_1, \ldots, X_n\}) = \mathbb{E}[X_i \mid \max\{X_1, \ldots, X_n\} > Q_{\max\{X_1,\ldots,X_n\}}(\alpha)].
$$

For Farlie-Gumbel-Morgenstern copula (Bargès et al., 2009). For elliptic distribution functions (Landsman and Valdez, 2003). For phase-type distributions (Cai and Li, 2005).

This kind of measures are often used in the literature to model problems of capital allocation in a portfolio when the risks are dependent.
Quantile generalizations: literature

✓ Relevant tool in statistics and probability. In the univariate case it is easily interpretable and calculable.

✓ Problem of the natural ordering in \( n \)-dimensions, \( n > 1 \) (Barnett, 1976).

Several attempts of multidimensional generalizations (see Serfling, 2002):

- Massé and Theodorescu (1994) defined multivariate quantiles as intersection of halfplanes with probability at least equal to \( p \).

- Koltchinskii (1997) uses the inversions of mappings.

- Einmahl and Mason (1992) introduce the quantile as

\[
U(p) := \inf \{ \lambda(C) : C \in \mathcal{C} \text{ et } P(C) \geq p \},
\]

for \( 0 < p < 1 \), with the real-value function \( \lambda \), \( P \) a probability measure on \( \mathbb{R}^n \) and \( \mathcal{C} \) a subset of Borel sets on \( \mathbb{R}^n \).

- Spatial quantile (see Kemperman, 1987). The spatial median is defined as

\[
M := \arg \min_c \int_X \| x - c \| dP(x).
\]

(Remark: importance on metric choice)
Tibiletti (1993), Fernández-Ponce and Suárez-Lloréns (2002) and Belzunce \textit{et al.} (2007) defined a multivariate quantile as a set of points which accumulate the same probability for a fixed orthant (called \textit{quantile curves}).

\textbf{Definition (Quantile curve)}

For $\alpha \in (0,1)$ and a bivariate distribution function $F$, we define the two-dimensional quantile at probability level $\alpha$, $\partial L(\alpha) \triangleq \partial \{ F(x, y) \geq \alpha \}$, with $\alpha \in (0,1)$.

Remark: Natural extension in dimension two, \textit{“metric-free”}, for symmetric and non-symmetric distribution function, provide a data segmentation of predefined size.

De Haan and Huang (1995), Chebana and Ouarda (2011) use these quantile curve to model hydrological events (with several correlated variables: flood volume, peak and duration, storm duration and intensity).
Definition (Generalization of the CTE in dimension 2)

Consider a random vector \((X, Y)\) satisfying regularity properties, with associate distribution function \(F\). Let \(L(\alpha) = \{F(x, y) \geq \alpha\}\). For \(\alpha \in (0, 1)\), we define

\[
\text{CTE}_\alpha(X, Y) = \left( \begin{array}{c} \mathbb{E}[X \mid (X, Y) \in L(\alpha)] \\ \mathbb{E}[Y \mid (X, Y) \in L(\alpha)] \end{array} \right) = \left( \begin{array}{c} \mathbb{E}[X \mid F(X, Y) \geq \alpha] \\ \mathbb{E}[Y \mid F(X, Y) \geq \alpha] \end{array} \right).
\]

Distributional approach to risk model:
Our \(\text{CTE}_\alpha(X, Y)\) does not use an aggregate variable (sum, min, max \ldots) in order to analyze the bivariate risk’s issue. Conversely \(\text{CTE}_\alpha(X, Y)\) deals with the simultaneous joint damages considering the bivariate dependence structure of data in a specific risk’s area (\(\alpha\)-level set : \(L(\alpha)\)).
Definition

Consider a random vector \((X, Y)\) satisfying the regularity conditions. For \(\alpha \in (0, 1)\) we define the bi-dimensional Value-at-Risk at probability level \(\alpha\) by

\[
\text{Var}_\alpha(X, Y) = \left( \begin{array}{c}
\mathbb{E}[X \mid (X, Y) \in \partial L(\alpha)] \\
\mathbb{E}[Y \mid (X, Y) \in \partial L(\alpha)]
\end{array} \right) = \left( \begin{array}{c}
\mathbb{E}[X \mid F(X, Y) = \alpha] \\
\mathbb{E}[Y \mid F(X, Y) = \alpha]
\end{array} \right),
\]

where \(\partial L(\alpha)\) is the boundary of the \(\alpha\)-level set of \(F\), i.e. the \(\alpha\)-quantile curve of \(F\).

In the case of Archimedean copulas with differentiable generators \(\phi\)

\[
\text{Var}_\alpha^1(X, Y) = \frac{\int_{\mathbb{R}^2} \mathcal{Q}_{X(\alpha),X,Y} \mathcal{K}_\alpha(x) dx}{K'(\alpha)},
\]

with \(F_{U,C}(u,v)(s,t) = t - \frac{\phi(t)}{\phi'(t)} + \frac{\phi(s)}{\phi'(t)}\), for \(0 < t < s < 1\).
Two different developments of research:

- Problem of estimating the level sets \( L(c) = \{ F(x) \geq c \} \), with \( c \in (0, 1) \), of an unknown distribution function \( F \) on \( \mathbb{R}_+^2 \) with a plug-in approach; in order to provide a consistent estimator for \( \text{CTE}_\alpha(X, Y) \) [Article “Plug-in estimation of level sets in a non-compact setting with applications in multivariate risk theory”, with Laloë L., Maume-Deschamps V. and Prieur C., accepted for publication in ESAIM: Probability and Statistics journal].

- Analysis of \( \text{VaR}_\alpha(X, Y) \) and \( \text{CTE}_\alpha(X, Y) \) as risk measures: study of the classical properties (monotonicity, translation invariance, positive homogeneity, ...), behavior with respect to different risk scenarios and stochastic ordering of marginals risks [Article “Some proposals about bivariate risk measures”, with Cousin A., submitted].
Univariate risk measures: VaR, CTE

- Evaluating risks
- Generalizations of CTE
- Our Bivariate Conditional-Tail-Expectation

Bivariate risk measures

- Bivariate Value-at-Risk
- Bivariate Conditional-Tail-Expectation

Plug-in estimation

- Estimation of level sets: literature and background
- Notation and preliminary results
- Consistency result in terms of the Hausdorff distance
- $L_1$ consistency
- Estimation of our bivariate Conditional-Tail-Expectation
- Illustrations
Proposition

Consider a random vector \((X, Y)\) satisfying regularity conditions. Then \(\text{VaR}_\alpha(X, Y) \geq (\text{VaR}_\alpha(X), \text{VaR}_\alpha(Y))\). In the comonotonic case \(\text{VaR}_\alpha(X, Y) = (\text{VaR}_\alpha(X), \text{VaR}_\alpha(Y))\).

Figure: Clayton copula with parameter \(\theta\); \(\theta = -0.1; \theta = 1.1; \theta = 21\).
Bivariate Value-at-Risk: properties in terms of stochastic orders

**Proposition**

Let \((X_1, Y_1)\) and \((X_2, Y_2)\) be two bivariate continuous random vectors satisfying regularity conditions and with same twice differentiable copula structure \(C\). Then

1. \(\text{VaR}^1_\alpha(X_1, Y_1) = \text{VaR}^1_\alpha(X_1, Y_2)\), for \(\alpha \in (0, 1)\);

2. If \(X_1 \preceq_{st} X_2\) then it holds that
   \(\text{VaR}^1_\alpha(X_1, Y_1) \leq \text{VaR}^1_\alpha(X_2, Y_2)\), for all \(\alpha \in (0, 1)\);

3. If \(Y_1 \preceq_{st} Y_2\) then it holds that
   \(\text{VaR}^2_\alpha(X_1, Y_1) \leq \text{VaR}^2_\alpha(X_2, Y_2)\), for all \(\alpha \in (0, 1)\);

with \(\preceq_{st}\) the stochastic dominance order.

2 and 3: Result completely analogous to the one-dimensional setting.
Bivariate Value-at-Risk: properties in terms of level risk

**Proposition**

Consider a random vector \((X, Y)\) satisfying regularity conditions, with marginal distributions \(F_X\) and \(F_Y\) and copula \(C\). Let \(U = F_X(X)\) and \(V = F_Y(Y)\). The two following assertions hold:

- If \((U, C(U, V))\) is PRD \((U | C(U, V))\) then \(\text{VaR}^1_\alpha(X, Y)\) is a non-decreasing function of \(\alpha\).

- If \((V, C(U, V))\) is PRD \((V | C(U, V))\) then \(\text{VaR}^2_\alpha(X, Y)\) is a non-decreasing function of \(\alpha\).

\((X, Y)\) is said to admit positive regression dependence with respect to \(X\), \(\text{PRD}(Y | X)\), if

\[Y | X = x_1 \preceq_{st} Y | X = x_2, \quad \forall x_1 \leq x_2.\]

✓ Archimedean case.
Proposition

Consider a random vector $(X, Y)$ satisfying the regularity conditions. For $\alpha \in (0, 1)$, $\text{CTE}_\alpha(X, Y)$ satisfies the following properties:

**Positive Homogeneity:** \( \forall (c_1, c_2) \in \mathbb{R}^2_+ \),

\[
\text{CTE}_\alpha \left( \left( \begin{array}{c} c_1 \\ c_2 \end{array} \right) \left( \begin{array}{c} X \\ Y \end{array} \right) \right) = \left( \begin{array}{c} c_1 \\ c_2 \end{array} \right) \text{CTE}_\alpha(X, Y) = \left\{ \begin{array}{c} c_1 \mathbb{E}[X | (X, Y) \in L(\alpha)] \\ c_2 \mathbb{E}[Y | (X, Y) \in L(\alpha)] \end{array} \right\}.
\]

**Translation Invariance:** \( \forall (c_1, c_2) \in \mathbb{R}^2_+ \),

\[
\text{CTE}_\alpha \left( \left( \begin{array}{c} c_1 \\ c_2 \end{array} \right) + \left( \begin{array}{c} X \\ Y \end{array} \right) \right) = \left( \begin{array}{c} c_1 \\ c_2 \end{array} \right) + \text{CTE}_\alpha(X, Y) = \left\{ \begin{array}{c} c_1 + \mathbb{E}[X | (X, Y) \in L(\alpha)] \\ c_2 + \mathbb{E}[Y | (X, Y) \in L(\alpha)] \end{array} \right\}.
\]
Bivariate Conditional-Tail-Expectation: properties in terms of stochastic orders

**Proposition**

Let \((X_1, Y_1)\) and \((X_2, Y_2)\) be two bivariate continuous random vectors satisfying regularity conditions and with same twice differentiable copula structure \(C\). Then

- \(CTE^1_\alpha(X_1, Y_1) = CTE^1_\alpha(X_1, Y_2)\), for \(\alpha \in (0, 1)\);
- If \(X_1 \preceq_D X_2\) then it holds that \(CTE^1_\alpha(X_1, Y_1) \leq CTE^1_\alpha(X_2, Y_2)\), for all \(\alpha \in (0, 1)\);
- If \(Y_1 \preceq_D Y_2\) then it holds that \(CTE^2_\alpha(X_1, Y_1) \leq CTE^2_\alpha(X_2, Y_2)\), for all \(\alpha \in (0, 1)\);

with \(\preceq_D\) the dangerousness order.

2 and 3: Result completely analogous to the one-dimensional setting.
Bivariate Conditional-Tail-Expectation: properties in terms of level risk

**Proposition**

Consider a random vector \((X, Y)\) satisfying regularity conditions. If \(\text{VaR}_{\alpha}^1(X, Y)\) is a non-decreasing function of \(\alpha\), then for \(\alpha \in (0, 1)\), \(\text{CTE}_{\alpha}^1(X, Y) \geq \text{VaR}_{\alpha}^1(X, Y)\) and \(\text{CTE}_{\alpha}(X, Y)\) is a non-decreasing function of the risk level \(\alpha\).

✓ Archimedean case:

To summarize, in the Archimedean case it holds that

\[
\text{CTE}_{\alpha}(X, Y) \geq \text{VaR}_{\alpha}(X, Y) \geq \left( \frac{\text{VaR}_{\alpha}(X)}{\text{VaR}_{\alpha}(Y)} \right), \quad \text{for } \alpha \in (0, 1),
\]

with \(\text{CTE}_{\alpha}^i(X, Y)\) and \(\text{VaR}_{\alpha}^i(X, Y)\) non-decreasing functions of \(\alpha\), for \(i = 1, 2\).
Bivariate Conditional-Tail-Expectation: properties in terms of level risk

Illustration Frank copula:

Figure: Frank copula with parameter 2 (left); with parameter $-10$ (right). Standard uniform marginals.
Contents

1 Univariate risk measures: VaR, CTE
   - Evaluating risks
   - Generalizations of CTE
   - Our Bivariate Conditional-Tail-Expectation

2 Bivariate risk measures
   - Bivariate Value-at-Risk
   - Bivariate Conditional-Tail-Expectation

3 Plug-in estimation
   - Estimation of level sets: literature and background
   - Notation and preliminary results
   - Consistency result in terms of the Hausdorff distance
   - $L_1$ consistency
   - Estimation of our bivariate Conditional-Tail-Expectation
   - Illustrations

Estimation of the level sets of a regression function in a compact setting: Cavalier (1997), Laloë (2009).

Estimation of general compact level sets: Cuevas et al. (2006).

An alternative approach, based on the geometric properties of the compact support sets: Cuevas and Fraiman (1997), Cuevas and Rodríguez-Casal (2004)
Plug-in estimation of level sets

**Goal:** We consider the problem of estimating the level sets of a bivariate distribution function $F$. More precisely our goal is to build a consistent estimator of

$$L(c) := \{F(x) \geq c\}, \quad \text{for } c \in (0, 1).$$

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

$$L_n(c) := \{F_n(x) \geq c\}, \quad \text{for } c \in (0, 1),$$

where $F_n$ is a consistent estimator of $F$. 
From a parametric formulation of $\partial L(c)$ (see Belzunce et al., 2007).

Case: Survival Clayton Copula with marginals (Burr(1), Burr(2))
We state consistency results with respect to two proximity criteria between sets: the Hausdorff distance and the volume of the symmetric difference ("physical proximity" between sets).

→ 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

Let $\mathbb{R}_+^2 := \mathbb{R}_+^2 \setminus (0, 0)$, $\mathcal{F}$ the set of continuous distribution functions $f : \mathbb{R}_+^2 \to [0, 1]$ and $F \in \mathcal{F}$.

Given an i.i.d sample $\{X_i\}_{i=1}^n$ in $\mathbb{R}_+^2$ with distribution function $F$, we denote $F_n(.) = F_n(X_1, X_2, \ldots, X_n, .)$ an estimator of $F$.

Define, for $c \in (0, 1)$, the upper $c$-level set of $F \in \mathcal{F}$ and its plug-in estimator:

$$L(c) := \{ x \in \mathbb{R}_+^2 : F(x) \geq c \}, \quad L_n(c) := \{ x \in \mathbb{R}_+^2 : F_n(x) \geq c \},$$

$$\{ F = c \} = \{ x \in \mathbb{R}_+^2 : F(x) = c \}.$$

In addition, given $T > 0$, we set

$$L(c)^T = \{ x \in [0, T]^2 : F(x) \geq c \}, \quad L_n(c)^T = \{ x \in [0, T]^2 : F_n(x) \geq c \},$$

$$\{ F = c \}^T = \{ x \in [0, T]^2 : F(x) = c \}.$$

For any $A \subset \mathbb{R}_+^2$ we note $\partial A$ its boundary.
For $r > 0$ and $\lambda > 0$, define

$$E := B\left(\{x \in \mathbb{R}^2_+ : |F - c| \leq r\}, \lambda\right),$$

$$m^\nabla := \inf_{x \in E} \|(\nabla F)_x\|, \quad M_H := \sup_{x \in E} \|(HF)_x\|,$$

- $B(x, \xi)$ is the closed ball centered on $x$ and with positive radius $\xi$,
- $(\nabla F)_x$ is the gradient of $F$ evaluated at $x$,
- $\|(HF)_x\|$ is the matrix norm induced by Euclidean distance of the Hessian matrix in $x$.

**Hausdorff distance** between $A_1$ and $A_2$ is defined by:

$$d_H(A_1, A_2) = \inf\{\varepsilon > 0 : A_1 \subset B(A_2, \varepsilon), A_2 \subset B(A_1, \varepsilon)\},$$

where $B(S, \varepsilon) = \bigcup_{x \in S} B(x, \varepsilon)$; $A_1$ and $A_2$ are compact sets in $(\mathbb{R}^2_+, d)$. 
Notation and preliminary results

**H:** 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 $\mathbf{H}$ 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 $r > 0$, $\zeta > 0$ such that $m^\triangledown > 0$ and $M_H < \infty$. Then $F$ satisfies Assumption $\mathbf{H}$, with $A = \frac{2}{m^\triangledown}$.
Under assumption of this Proposition:

- There is no plateau in the graph of $F$ for each level $t$ such that $|t - c| \leq r$.
- $\{F = t\}$ is a curve in the quadrant $\mathbb{R}^2_+$.
- From $m^{\nabla} > 0$, the plane curve $\{F = t\}$ has the following monotonic property. We consider $(x, y), (x', y') \in \{F = t\}$, if $x < x'$ then $y \geq y'$, if $y < y'$ then $x \geq x'$.
- If we suppose that each component of $(\nabla F)_x$ is greater than zero in $E$ then $\{F = t\}$ is a monotone decreasing curve in $\mathbb{R}^2_+$.
- Then we obtain $\partial L(c)^T = \{F = c\}^T = \{F = c\} \cap [0, T]^2$.

see Rossi (1976), Rodríguez-Casal (2003), Tibiletti (1993).
Consistency result in terms of the 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 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 \( t : | t - c | \leq r \), \( \partial L(t) T_1 \neq \emptyset \). 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, \( d_H(\partial L(c) T_n, \partial L_n(c) T_n) \leq 6 A \| F - F_n \|_\infty^{T_n} \), a.s., where \( A = \frac{2}{m^\triangledown} \).
Consistency result in terms of the Hausdorff distance

- \(d_H(\partial L(c)^T \hat{L}_n, \partial L_n(c)^T \hat{L}_n)\) converges to zero and the quality of our plug-in estimator is obviously related to the quality of the estimator \(F_n\).

- Note that in the case \(c\) is close to one the constant \(A = \frac{2}{m^2}\) could be large. In this case, we will need a large number of data to get a reasonable estimation.

- In order to overcome the problem of “\(F_n\) is continuous with probability one” it can be considered a smooth version of estimator (e.g. see Chaubey and Sen, 2002). However the continuity assumption will not be necessary in the following results.
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

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

**Theorem (Consistency 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 \( t : | t - c | \leq r \), \( \partial L(t)^T_1 \neq \emptyset \). Then it holds that

\[
p_n d_\lambda(L(c)^{T_n}, L_n(c)^{T_n}) \xrightarrow[n \to \infty]{\mathbb{P}} 0,
\]

with \( p_n \) an increasing positive sequence such that

\[
p_n = o \left( \frac{1}{v_n^{p+1}} / T_n^{p+1} \right).
\]
\L_1 \text{ consistency}

- Assumptions on gradient vector and Hessian matrix of \( F \), in these theorems, are satisfied for a quite large class of classical bivariate distribution functions.

- **Theorem Consistency volume** provides a convergence rate, which is closely related to the choice of the sequence \( T_n \). A sequence \( T_n \) whose divergence rate is large implies a convergence rate \( p_n \) quite slow.

- **Theorem Consistency volume** does not require continuity assumption on \( F_n \).

- Problem: optimal criterion for the choice of \( T_n \) related with the \( p \) in Assumption A1.

- Example: \( F_n \) the bivariate empirical distribution function.
If there exists \((v_n)_{n \in \mathbb{N}^*}\) such that \(v_n \| F - F_n \|_{\infty} \overset{p}{\underset{n \to \infty}{\to}} 0\). Then

\[
\forall \ p \geq 1, \ w_n \int_{[0, T_n]^2} | F - F_n |^p \lambda(dx) \overset{p}{\underset{n \to \infty}{\to}} 0, \quad \text{with} \quad w_n = \frac{v_n^p}{T_n^2}.
\]

In this case we get \(p_n = o\left(\frac{v_n^{p+1}}{T_n^{p+1}}\right)\).

⇒ Let \(F_n\) the bivariate empirical distribution function, \(v_n = o(\sqrt{n})\), with \(p = 2\), we obtain \(p_n = o\left(\frac{n^{\frac{1}{3}}}{T_n^{\frac{4}{3}}}\right)\).
Estimation of $\text{CTE}_\alpha(X, Y)$

**Definition**

Consider a random vector $(X, Y)$ with associate distribution function $F \in \mathcal{F}$. For $\alpha \in (0, 1)$, we define the estimated bivariate $\alpha$-Conditional-Tail-Expectation

$$\hat{\text{CTE}}_\alpha(X, Y) = \left( \begin{array}{c} \frac{\sum_{i=1}^{n} X_i 1\{(X_i, Y_i)\in L_n(\alpha)\}}{\sum_{i=1}^{n} 1\{(X_i, Y_i)\in L_n(\alpha)\}} \\ \frac{\sum_{i=1}^{n} Y_i 1\{(X_i, Y_i)\in L_n(\alpha)\}}{\sum_{i=1}^{n} 1\{(X_i, Y_i)\in L_n(\alpha)\}} \end{array} \right).$$

We introduce truncated versions of the theoretical and estimated $\text{CTE}_\alpha$:

$$\text{CTE}_\alpha^T(X, Y) = \mathbb{E}[(X, Y)\mid (X, Y) \in L(\alpha)^T], \quad \hat{\text{CTE}}_\alpha^T(X, Y),$$

using $L(\alpha)^T$ and $L_n(\alpha)^T$ i.e. the truncated versions $L(\alpha)$ and $L_n(\alpha)$.
Estimation of bivariate Conditional-Tail-Expectation

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

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

$$
\beta_n \left| \text{CTE}^T_n(X, Y) - \widehat{\text{CTE}}^T_n(X, Y) \right| \xrightarrow{P} 0, \quad n \to \infty
$$

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

The convergence in (1) must be interpreted as a componentwise convergence. In the case of a bounded density function $f(X, Y)$ we obtain $\beta_n = \min\{ \sqrt{p_n}, a_n \}$. 

Univariate risk measures: VaR, CTE
Bivariate risk measures
Plug-in estimation
Perspectives and References

Estimation of our bivariate Conditional-Tail-Expectation

**Estimation of \( \text{CTE}_\alpha(X, Y) \) with \( F_n \) bivariate empirical distribution function**

**Proposition**

Let \( F_n \) the bivariate empirical distribution function. Under Assumptions of “Theorem Consistency of \( \text{CTE}_\alpha(X, Y) \)” and with the same notation, it holds that

\[
\beta_n \left| \text{CTE}_{T_n}^\alpha(X, Y) - \text{CTE}_{T_n}^\alpha(X, Y) \right| \xrightarrow{P} 0 \quad \text{as} \quad n \to \infty,
\]

where \( \beta_n = o \left( n^{6(1+r)/(1+r)} / T_n^{2r/(1+r)} \right) \), with \( r > 0 \) such that the density \( f(X, Y) \in L^{1+r}(\lambda) \).

In the case of a bounded density function \( f(X, Y) \), \( \beta_n = o \left( n^{1/6} / T_n^{2/3} \right) \).
On the choice of $T_n$

Obviously it could be interesting to consider the convergence:

$$ \left| \text{CTE}_\alpha(X, Y) - \widehat{\text{CTE}}_{\alpha}^{T_n}(X, Y) \right|. $$

The convergence rate also depend on the convergence rate to zero of

$$ \left| \text{CTE}_\alpha(X, Y) - \text{CTE}^{T_n}_\alpha(X, Y) \right|, \text{ then of } \mathbb{P}[(X, Y) \in L(\alpha) \setminus L(\alpha)^{T_n}]. $$

More precisely $\left| \text{CTE}_\alpha(X, Y) - \text{CTE}^{T_n}_\alpha(X, Y) \right|$ decays to zero at least with a convergence rate $(\mathbb{P}[X \geq T_n \text{ or } Y \geq T_n])^{-1}$.

We remark that $(\mathbb{P}[X \geq T_n \text{ or } Y \geq T_n])^{-1}$ is increasing in $T_n$, whereas the speed of convergence is decreasing in $T_n$. This kind of compromise provides an illustration on how to choose $T_n$, apart from satisfying the assumptions of consistency results above.
Estimation of the level sets

Plug-in estimation of level sets: $F_n$ of the bivariate distribution function, $n = 500 : 2000$, Monte Carlo approximation (averaged on 100 iterations), $p_n = o(n^{\frac{1}{3}} / T_n^{\frac{4}{3}})$, $T_n = \ln(n)$. These results set out how $p_n = o(n^{\frac{1}{3}} / \ln(n)^{\frac{4}{3}})$ is at least the convergence rate of $\lambda (L(\alpha)^T_n \triangle L_n(\alpha)^T_n)$ in this particular case.

\[
\begin{array}{|c|c|c|c|}
\hline
\alpha & n= 500 & n= 1000 & n= 2000 \\
\hline
0.10 & 0.099 & 0.089 & 0.078 \\
0.24 & 0.226 & 0.176 & 0.075 \\
0.38 & 0.248 & 0.183 & 0.143 \\
0.52 & 0.324 & 0.223 & 0.217 \\
0.66 & 0.429 & 0.259 & 0.232 \\
0.80 & 0.613 & 0.371 & 0.332 \\
\hline
\end{array}
\]

\[
\begin{array}{|c|c|c|c|}
\hline
\alpha & n= 500 & n= 1000 & n= 2000 \\
\hline
0.10 & 0.069 & 0.068 & 0.065 \\
0.24 & 0.156 & 0.134 & 0.063 \\
0.38 & 0.172 & 0.139 & 0.121 \\
0.52 & 0.225 & 0.169 & 0.183 \\
0.66 & 0.298 & 0.199 & 0.195 \\
0.80 & 0.426 & 0.282 & 0.279 \\
\hline
\end{array}
\]

**Table:** $F(x, y) = (1 - e^{-x})(1 - e^{-2y})$. 
Plug-in estimation of level sets: $F_n$ of the bivariate distribution function, $n = 1000$, 100 simulations.

<table>
<thead>
<tr>
<th>$\alpha$</th>
<th>CTE$_\alpha(X, Y)$</th>
<th>$CTE^{\hat{T}<em>n}</em>\alpha(X, Y)$</th>
<th>$\hat{\sigma}$</th>
<th>RMSE</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.10</td>
<td>(1.255, 0.627)</td>
<td>(1.222, 0.638)</td>
<td>(0.044, 0.022)</td>
<td>(0.043, 0.039)</td>
</tr>
<tr>
<td>0.24</td>
<td>(1.521, 0.761)</td>
<td>(1.488, 0.811)</td>
<td>(0.069, 0.023)</td>
<td>(0.051, 0.042)</td>
</tr>
<tr>
<td>0.38</td>
<td>(1.792, 0.896)</td>
<td>(1.797, 0.911)</td>
<td>(0.075, 0.038)</td>
<td>(0.044, 0.046)</td>
</tr>
<tr>
<td>0.52</td>
<td>(2.102, 1.051)</td>
<td>(2.082, 1.047)</td>
<td>(0.104, 0.052)</td>
<td>(0.052, 0.052)</td>
</tr>
<tr>
<td>0.66</td>
<td>(2.492, 1.246)</td>
<td>(2.461, 1.255)</td>
<td>(0.139, 0.071)</td>
<td>(0.057, 0.056)</td>
</tr>
<tr>
<td>0.80</td>
<td>(3.061, 1.531)</td>
<td>(3.011, 1.544)</td>
<td>(0.251, 0.125)</td>
<td>(0.084, 0.082)</td>
</tr>
</tbody>
</table>

**Table:** $(X, Y)$ with independent and exponentially distributed components with parameter 1 and 2 respectively; $T_n = \ln(n)$. 

Plug-in estimation of level sets in a non-compact setting with applications in multivariate risk theory.
Estimation of $\text{CTE}_\alpha(X, Y)$: simulated data

Plug-in estimation of level sets: $F_n$ of the bivariate distribution function, $n = 1000$, 100 simulations.

<table>
<thead>
<tr>
<th>$\alpha$</th>
<th>$\text{CTE}_\alpha(X, Y)$</th>
<th>$\text{CTE}_\alpha^T(X, Y)$</th>
<th>$\hat{\sigma}$</th>
<th>$\text{RMSE}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.10</td>
<td>(1.188, 1.229)</td>
<td>(1.049, 1.192)</td>
<td>(0.032, 0.021)</td>
<td>(0.019, 0.033)</td>
</tr>
<tr>
<td>0.24</td>
<td>(1.448, 1.366)</td>
<td>(1.283, 1.379)</td>
<td>(0.053, 0.224)</td>
<td>(0.019, 0.063)</td>
</tr>
<tr>
<td>0.38</td>
<td>(1.727, 1.505)</td>
<td>(1.525, 1.471)</td>
<td>(0.046, 0.031)</td>
<td>(0.019, 0.031)</td>
</tr>
<tr>
<td>0.52</td>
<td>(2.049, 1.666)</td>
<td>(1.803, 1.625)</td>
<td>(0.058, 0.041)</td>
<td>(0.023, 0.034)</td>
</tr>
<tr>
<td>0.66</td>
<td>(2.454, 1.875)</td>
<td>(2.129, 1.823)</td>
<td>(0.071, 0.054)</td>
<td>(0.035, 0.039)</td>
</tr>
<tr>
<td>0.80</td>
<td>(3.039, 2.202)</td>
<td>(2.591, 2.144)</td>
<td>(0.111, 0.105)</td>
<td>(0.055, 0.054)</td>
</tr>
</tbody>
</table>

Table: $(X, Y)$ with Clayton copula with parameter 1, $F_X$ exponential distribution with parameter 1, $F_Y$ Burr(4, 1) distribution; $T_n = n^{0.2}$. 
Behavior with respect to the risk level $\alpha$

Plug-in estimation of level sets: $F_n$ of the bivariate distribution function, $n = 500 : 5000$, 100 simulations.

For high levels (here $\alpha = 0.9$), one needs to use large samples to get reasonable estimates of $\text{CTE}_\alpha$.

<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 $n$ for $\alpha = 0.9$; $(X, Y)$ independent and exponentially distributed components with parameter 1 and 2 respectively. $T_n = \ln(n)$, theoretical value $\text{CTE}_{0.9}(X, Y) = (3.78, 1.89)$.

In this case we need between 2000 and 2500 data to get the same performances as for lower levels.
Real case: **Loss-ALAE data** in the log scale (Frees and Valdez, 1998). The data size is $n = 1500$. Let $F_n$ the empirical distribution function.
Perspectives

- Study of $\text{CTE}_\alpha(X, Y)$ as risk measure (stochastic order for random variables and random vectors).

- In this paper we provide asymptotic results for a fixed level $c$...

- Problem of constant $A$ for $c \sim 1$: Plug in method with

$$\hat{L}_n(c) := \{(x, y) \in \mathbb{R}^2_+ : \hat{F}^*(x, y) \geq c\}, \quad \text{for } c \sim 1,$$

with $\hat{F}^*$ a “tail estimator” of $F$; Bivariate extreme value theory; first part of my PhD thesis (see de Haan et Haung, 1995).

- Provide an optimal rate of $|\text{CTE}_\alpha(X, Y) - \widehat{\text{CTE}}_\alpha^T(X, Y)|$, at least for some classes of bivariate dependence structure.

- Kendall process $K_n(t)$ in order to estimate $\text{CTE}_\alpha(X, Y)$ (see “On Kendall’s process”, Barbe et al., 1996) - Central Limit Theorem.

- Package $\mathbb{R}$ for level sets and $\text{CTE}_\alpha(X, Y)$ estimation.
References


