HAL open science

Estimating covariate functions associated to multivariate risks: a level sets approach

Elena Di Bernardino, Thomas Laloë, Rémi Servien

▶ To cite this version:

Elena Di Bernardino, Thomas Laloë, Rémi Servien. Estimating covariate functions associated to multivariate risks: a level sets approach. 2013. hal-00800461v2

HAL Id: hal-00800461

https://hal.science/hal-00800461v2

Preprint submitted on 5 Aug 2013 (v2), last revised 16 Dec 2014 (v4)

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Estimating covariate functions associated to multivariate risks : a level set approach

Elena Di Bernardino, Thomas Laloë And Rémi Servien

Conservatoire National des Arts et Métiers, Universit de Nice Sophia Antipolis, UMR 1331 Toxalim - INRA

Abstract: The aim of this paper is to study the behavior of a covariate function in a multivariate risks scenario. The first part of this paper deals with the problem of estimating the c-upper level sets $L(c) = \{F(x) \geq c\}$, with $c \in (0,1)$, of an unknown distribution function F on $\mathbb{R}_+^d$. A plug-in approach is followed. We state consistency results with respect to the volume of the symmetric difference. In the second part, we obtain the L_p-consistency, with a convergence rate, for the regression function estimate on these level sets $L(c)$. We also consider a new multivariate risk measure: the Covariate-Conditional-Tail-Expectation. We provide a consistent estimator for this measure with a convergence rate. We propose a consistent estimate when the regression cannot be estimated on the whole data set. Then, we investigate the effects of scaling data on our consistency results. All these results are proven in a non-compact setting. A complete simulation study is detailed. Finally, a real environmental application of our risk measure is provided.

Key words and phrases: multidimensional distribution function, plug-in estimation, regression function.

1 Introduction

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, especially when the risk-problem is affected by other external risk factors whose sources cannot be controlled. Note that the evaluation of an individual risk may strongly be affected by the degree of dependence amongst all risks and these risks may also be strongly heterogeneous.

For instance, several hydrological phenomena are described by two or more correlated characteristics. These dependent characteristics should be considered jointly to be more representative of the multivariate nature of the phenomenon. Consequently, probabilities of occurrence of risks cannot be estimated on the basis of univariate analysis. The multivariate hydrological risks literature mainly treated one or more of the following three elements: (1) showing the importance and explaining the usefulness of the multivariate framework, (2) fitting the appropriate multivariate distribution in order to model risks and (3) defining and studying multivariate return periods (see Chebana and Ouarda [7]), (i.e., multivariate quantile based measures of risks).

One of the most popular measures in hydrology and climate is undoubtedly the return period. A

frequency analysis in hydrology focuses on the estimation of quantities (e.g., flows or annual rainfall) corresponding to a certain return period. It is closely related to the notion of quantile which has therefore been extensively studied in dimension one. For a random variable X that represents the magnitude of an event that occurs at a given time and at a given location, the quantile of order $1-\frac{1}{T}$ expresses the magnitude of the event which is exceeded with a probability equal to $\frac{1}{T}$. T is then called the return period. In univariate risk theory the quantile is known as the *Value-at-Risk* (VaR) and is defined by

$$Q_X(c) = \inf\{x \in \mathbb{R}_+ : F_X(x) \geq c\}, \quad \text{for } c \in (0,1),$$

with F_X the univariate distribution of event X. A second important univariate risk measure, based on the quantile notion, is the *Conditional-Tail-Expectation* (CTE) defined by

$$\mathrm{CTE}_c(X) = \mathbb{E}[\, X \,|\, X > Q_X(c)\,], \quad \text{for } c \in (0,1).$$

From the years 2000 onward, much research has been devoted to risk measures and many extensions to multidimensional settings have been suggested (see, e.g., Jouini et al. [22]; Bentahar [4]; Embrechts and Puccetti [19]; Nappo and Spizzichino [24]; Ekeland et al. [17]).

For a non-negative d-dimensional risk portfolio with distribution function F, the c-upper level set of F (i.e., $L(c) = \{x \in \mathbb{R}_+^d : F(x) \geq c\}$) and its associated c-level curve (i.e., $\partial L(c) = \{x \in \mathbb{R}_+^d : F(x) = c\}$) have recently been proposed as risk measures in multivariate hydrological models. Among their many advantages, it appears that they are simple, intuitive, interpretable and probability-based (see Chebana and Ouarda [7]). de Haan and Huang [11] model a risk-problem of flood in the bivariate setting using an estimator of level curves $\partial L(c)$ of the bivariate distribution function. Furthermore, as noticed by Embrechts and Puccetti [19], $\partial L(c)$ can be viewed as a natural multivariate version of the univariate quantile. The interested reader is also referred to Tibiletti [29], Belzunce et al. [3], Nappo and Spizzichino [24].

As a starting point, in the following, we consider the multivariate version of the CTE measure, proposed by Di Bernardino et al. [15] and Cousin and Di Bernardino [8]. It is constructed as the conditional expectation of a multivariate random vector given that the latter is located in the c-upper level set of the associated multivariate distribution function. In this sense this measure is essentially based on a "multivariate distributional approach". More precisely they define, for $i = 1, \ldots, d$ and for $c \in (0,1)$,

$$\mathrm{CTE}_c^i(\mathbf{X}) = \mathbb{E}[\, X_i \,|\, \mathbf{X} \in L(c)\,], \tag{1.1}$$

where $\mathbf{X} = (X_1, \ldots, X_d)$ is a non-negative multivariate risk portfolio with distribution function F. In particular, Cousin and Di Bernardino [8] proved that properties of the multivariate Conditional-Tail-Expectation in (1.1) turn to be consistent with existing properties on univariate risk measures

(positive homogeneity, translation invariance, increasing in risk-level c, ...).

In financial econometrics literature, we are often interested to analyse the behavior of an univariate return measure (average return, skewness, ...) with respect to a set of d risk factors $\mathbf{X}$ (volatility or variance, kurtosis, ...). In other words, we consider an univariate covariate Y and a dependent multivariate vector of risk-factors $\mathbf{X}$. Furthermore, in climatology, one may be interested in how climate change over years might affect high temperatures. Multivariate examples include the study of rainfall as a covariate function represented by the geographical location. In this sense Daouia et al. [10] deal with the problem of estimating quantiles when covariate information is available.

So, the goal of this paper is the study of the behavior of a covariate Y on the level sets of a d-dimensional vector of risk-factors $\mathbf{X}$. More precisely, adapting the multivariate risk measure in (1.1), we deal with the multivariate Covariate-Conditional-Tail-Expectation (CCTE) defined by:

$$\mathrm{CCTE}_c(\mathbf{X}, Y) := \mathbb{E}[Y \,|\, \mathbf{X} \in L(c)], \tag{1.2}$$

where $c \in (0,1)$. In order to estimate this risk measure, we first need to estimate the level sets $L(c)$ associated to the d-dimensional distribution function F of $\mathbf{X}$.

Considering the level sets of a distribution function, the commonly assumed property of compactness for these sets is no more reasonable. Then, differing from the classical literature (Baíllo et al. [2]; Rigollet and Vert [26]; Cuevas et al. [9]), we need to work in a non-compact setting and this requires special attention in the statement of our problem.

Considering a consistent estimator F_n of the distribution function F, we propose a *plug-in* approach to estimate the c-upper level set $L(c)$ by

$$L_n(c) = \{x \in \mathbb{R}_+^d : F_n(x) \geq c\}, \text{ for } c \in (0,1).$$

The regularity properties of F and F_n as well as the consistency properties of F_n will be specified in the statements of our theorems. Our consistency result for $L(c)$ is stated with respect to a criterion of "physical proximity" between sets: the volume of the symmetric difference. Obviously, the convergence rate suffers from the well-known *curse of dimensionality* (see Theorem 3.1).

Using Theorem 3.1, we state L_p-consistency with a convergence rate for the estimation of the regression function

$$r(x) = \mathbb{E}[Y \,|\, \mathbf{X} = x],$$

on these level sets $L(c)$, i.e., for $x \in L(c)$, and $c \in (0,1)$ (see Theorems 4.1 and 4.2). The motivation behind the point-wise estimation of $r(x)$ for $x \in L(c)$ is an interesting problem, for different practical

problem. Indeed, this represents the expected value of a covariate Y given that a dependent multivariate vector of risk-factors $\mathbf{X}$ takes value in a specific risk area $L(c)$ (for instance $L(c)$ can represent a risk-scenario or critical-layer and so on). For the importance of $L(c)$ in the risk-management in the environmental or hydrological fields see for Chebana and Ouarda [7], Salvadori et al. [27].

Finally, we provide a consistency result for the estimation of the CCTE risk measure, if Y is completely available (see Theorem 5.1) or not (see Theorem 5.2). Furthermore, we investigate the impact of a change in the scale of data on our results. In particular this property is related to the suitable positive homogeneity property of risk measures (e.g. see Artzner et al. [1]).

Organization of the paper

The paper is organized as follows. We introduce some notation, tools and technical assumptions in Section 2. Consistency and asymptotic properties of our estimator of $L(c)$ are given in Section 3. Section 4 is devoted to the L_p-consistency of the estimation of the regression function r and Section 5 to the consistency of the CCTE's estimation. The effects of scaling data are analyzed in Section 6. Illustrations with simulated data are presented in Section 7. A real example is studied in Section 8. Section 9 summarizes and briefly mentions directions for future research. Finally, proofs are postponed to Section 10.

2 Notation and preliminaries

In this section we introduce some notations and tools which will be useful later.

Let $\mathbb{N}^* = \mathbb{N} \setminus \{0\}$, $\mathbb{R}_+^* = \mathbb{R}_+ \setminus \{0\}$ and ${\mathbb{R}_+^d}^* = \mathbb{R}_+^d \setminus \{0\}$. Let $\mathcal{F}$ be the set of continuous distribution functions $\mathbb{R}_+^d \to [0,1]$ and $\mathbf{X} := (X_1, X_2, \ldots, X_d)$ a random vector with distribution function $F \in \mathcal{F}$. Given an i.i.d sample $\{\mathbf{X}_i\}_{i=1}^n$ in $\mathbb{R}_+^d$ with distribution function $F \in \mathcal{F}$, we denote by F_n an estimator of F based on this finite sample. Let Y a random variable in J, where $J \in \mathbb{R}_+$ is supposed to be bounded. We denote by $\{(\mathbf{X}_i, Y_i)\}_{i=1,\ldots,n}$ the associated i.i.d sample.

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

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

and

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

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

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

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

Note that, in the presence of a plateau at level c, $\{F = c\}$ can be a portion of quadrant $\mathbb{R}_+^d$ instead of a set of Lebesgue measure null in $\mathbb{R}_+^d$. In the statement of our results we will require suitable conditions in order to avoid this situation.

We denote by $B(x,\rho)$ the closed ball centered on $x \in \mathbb{R}_+^d$ and with positive radius ρ. Let $B(S,\rho) = \bigcup_{x\in S} B(x,\rho)$, with S a closed set of $\mathbb{R}_+^d$. For $\kappa > 0$ and $\zeta > 0$, define

$$E = B(\{x \in \mathbb{R}_+^d : \mid F(x) - c \mid \leq \kappa\},\, \zeta),$$

and, for a twice differentiable function F,

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

where $(\nabla F)_x$ is the gradient vector of F evaluated at x and $\|(\nabla F)_x\|$ its Euclidean norm, $(HF)_x$ the Hessian matrix evaluated in x and $\|(HF)_x\|$ its matrix norm induced by the Euclidean norm.

We define :

$$||g||_p := \left(\int_{\mathbb{R}^d} \mid g(x) \mid^p \; f(x)\,\mathrm{d}x\right)^{1/p},$$

where f denotes the density function associated to the probability measure μ and

$$||g||_{p,\lambda} := \left(\int_{\mathbb{R}^d} \mid g(x) \mid^p \; \mathrm{d}x\right)^{1/p}.$$

Finally, let $r(x)$ be the regression function such that

$$r(x) := \mathbb{E}[Y \mid \mathbf{X} = x].$$

3 Estimating level sets of a multidimensional distribution function using a plug-in method

In this section, we consider the problem of estimating the level sets of a d-variate distribution function.

We consider the consistency in term of the volume (in the Lebesgue measure sense) of the symmetric difference between $L(c)^{T_n}$ and $L_n(c)^{T_n}$. This means that we define the distance between two subsets A_1 and A_2 of $\mathbb{R}_+^d$ by

$$d_\lambda(A_1, A_2) = \lambda(A_1 \triangle A_2),$$

where λ stands for the Lebesgue measure on $\mathbb{R}^d$ and $\triangle$ for the symmetric difference.

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

$$v_n \int_{[0,T_n]^d} \mid F - F_n \mid^p \ \lambda(\mathrm{d}x) \underset{n\to\infty}{\overset{\mathbb{P}}{\to}} 0,$$

for some $1 \leq p < \infty$.

We now establish our consistency result with convergence rate, in term of the volume of the symmetric difference.

Theorem 3.1. *Let $c \in (0,1)$. Let $F \in \mathcal{F}$ be a twice differentiable distribution function on $\mathbb{R}_+^{d*}$. Assume that there exist $\kappa > 0$, $\zeta > 0$ such that $m^{\triangledown} > 0$ and $M_H < \infty$. Let $T_1 > 0$ such that for all $t : \mid t - c \mid \leq \kappa$, $\partial L(t)^{T_1} \neq \emptyset$. Assume that for each n, with probability one, F_n is measurable. Let $(v_n)_{n\in\mathbb{N}^*}$ and $(T_n)_{n\in\mathbb{N}^*}$ positive increasing sequences such that Assumption* **A1** *is satisfied. Then, it holds that*

$$p_n \, d_\lambda(L(c)^{T_n}, L_n(c)^{T_n}) \underset{n\to\infty}{\overset{\mathbb{P}}{\to}} 0,$$

with p_n an increasing positive sequence such that $p_n = o\left(v_n^{\frac{1}{p+1}} / T_n^{\frac{(d-1)\,p}{p+1}}\right)$.

The proof of Theorem 3.1 is postponed to Section 10.

Remark 3.1. *Note that in the univariate case ($d = 1$) the convergence rate of Theorem 3.1 does not depend on the truncation sequence T_n. The interested reader is referred to the proof of this result in Section 10 for further details.*

Theorem 3.1 provides a convergence rate, which is closely related to the choice of the sequence T_n, for $d > 1$. Note that, as in Theorem 3 in Cuevas et al. [9], Theorem 3.1 above does not require any continuity assumption on F_n. Furthermore, we remark that a sequence T_n whose divergence rate is large, implies a convergence rate p_n quite slow. Moreover, this phenomenon is emphasized by the dimension d of the data, and we face here the well-known *curse of dimensionality*. In the following we will illustrate this aspect by giving convergence rate in the case of the empirical distribution function (see Example 1). Firstly, from Theorem 3.1 we can derive the following result.

Corollary 3.1. *Let $c \in (0,1)$. Let $F \in \mathcal{F}$ be a twice differentiable distribution function on $\mathbb{R}_+^{d*}$. Assume that there exist $\kappa > 0$, $\zeta > 0$ such that $m^{\triangledown} > 0$ and $M_H < \infty$. Let $T_1 > 0$ such that for all $t : \mid t - c \mid \leq \kappa$, $\partial L(t)^{T_1} \neq \emptyset$. Assume that for each n, with probability one, F_n is measurable. Assume that there exists a positive increasing sequence $(w_n)_{n\in\mathbb{N}^*}$ such that $w_n \, \|F - F_n\|_\infty \underset{n\to\infty}{\overset{\mathbb{P}}{\to}} 0$. Then, it holds that*

$$p_n \, d_\lambda(L(c)^{T_n}, L_n(c)^{T_n}) \underset{n\to\infty}{\overset{\mathbb{P}}{\to}} 0,$$

with p_n an increasing positive sequence such that $p_n = o\left(w_n{}^{\frac{p}{p+1}}/T_n{}^{\frac{d+(d-1)\,p}{p+1}}\right)$.

This result comes trivially from Theorem 3.1 and the fact that $w_n \|F - F_n\|_\infty \underset{n\to\infty}{\overset{\mathbb{P}}{\to}} 0$ implies

$$\forall\ p \geq 1, \quad v_n \int_{[0,T_n]^d} \mid F - F_n \mid^p \ \lambda(\mathrm{d}x) \underset{n\to\infty}{\overset{\mathbb{P}}{\to}} 0, \qquad \text{with} \qquad v_n = \frac{w_n^p}{T_n^d}.$$

Let us now present a more practical example in the case of a d-variate empirical distribution function.

Example 1. *Let F_n the d-variate empirical distribution function. Then, it holds that $w_n \|F - F_n\|_\infty \underset{n\to\infty}{\overset{\mathbb{P}}{\to}} 0$, with $w_n = o(\sqrt{n})$. From Corollary 3.1, with $p = 2$, we obtain for instance:*

$$p_n = o\left(\frac{n^{1/3}}{T_n{}^{7/3}}\right), \quad \textit{for} \ \ d = 3; \qquad\qquad p_n = o\left(\frac{n^{1/3}}{T_n{}^{10/3}}\right), \quad \textit{for} \ \ d = 4.$$

4 L_p-consistency of r_n

In this section we study the L_p-consistency of an estimator r_n of the regression function. More precisely, in the first part we provide a consistency result in terms of the L_p-distance of the absolute error between $r_n\,\mathbf{1}_{\{x\in L_n(c)^{T_n}\}}$ and $r\,\mathbf{1}_{\{x\in L(c)\}}$. In the second part we analyze the problem of a convergence rate (see Theorem 4.2).

Let us introduce the following assumption:

A2 $||F_n - F||_{p,\lambda} \overset{\mathbb{P}}{\to} 0$ and $||r_n - r||_p \overset{\mathbb{P}}{\to} 0$, for some $1 \leq p < \infty$.

Theorem 4.1. *Let $c \in (0,1)$. Let $F \in \mathcal{F}$ be a twice differentiable distribution function on $\mathbb{R}_+^{d*}$. Assume that there exist $\kappa > 0$, $\zeta > 0$ such that $m^\triangledown > 0$ and $M_H < \infty$. Assume that for each n, with probability one, F_n is measurable. Let $T_1 > 0$ such that for all $t : \mid t - c \mid \leq \kappa$, $\partial L(t)^{T_1} \neq \emptyset$. Let $(T_n)_{n\in\mathbb{N}^*}$ be an increasing sequence of positive values such that* **A2** *is verified. Then it holds that*

$$||r_n\,\mathbf{1}_{\{x\in L_n(c)^{T_n}\}} - r\,\mathbf{1}_{\{x\in L(c)\}}||_p \overset{\mathbb{P}}{\to} 0, \qquad \textit{for} \ \ n \to \infty.$$

The proof of Theorem 4.1 is postponed to Section 10. Theorem 4.1 proves that the estimate of the regression function on the truncated estimated level set converges towards the real regression function on the whole real level set.

Remark 4.1. *If $||r_n - r||_\infty \to 0$, a.s. we have that*

$$||r_n\,\mathbf{1}_{\{x\in L_n(c)\}} - r\,\mathbf{1}_{\{x\in L(c)\}}||_\infty = \sup_{\mathbb{R}^d} |r_n\,\mathbf{1}_{\{x\in L_n(c)\}} - r\,\mathbf{1}_{\{x\in L(c)\}}| \nrightarrow 0, \ \ a.s.$$

Indeed, we have $\sup_{\mathbb{R}^d} |r_n \mathbf{1}_{\{x\in L_n(c)\}} - r\,\mathbf{1}_{\{x\in L(c)\}}| = \max_{L_n\Delta L}(r(x), r_n(x))$ *which is greater than zero if* $L_n\Delta L \neq \emptyset$ *.*

Let us introduce the following assumptions :

A3 There exist positive increasing sequences $(v_{1,n})_{n\in\mathbb{N}^*}$, $(v_{2,n})_{n\in\mathbb{N}^*}$ and $(T_n)_{n\in\mathbb{N}^*}$ such that

$$v_{1,n}\,||r_n - r||_p \xrightarrow{\mathbb{P}} 0,$$

$$v_{2,n}\int_{[0,T_n]^d} \mid F - F_n \mid^p \ \lambda(\mathrm{d}x) \xrightarrow{\mathbb{P}} 0,$$

for some $1 \leq p < \infty$.

A4 The density function f of $\mathbf{X}$ is $1+\epsilon$ integrable, with $\epsilon > 0$.

Note that Assumption **A3** includes Assumption **A1** in order to control the convergence rate of the distribution function estimate. Furthermore, Assumption **A3** implies to control of the convergence rate of the regression estimate.

Theorem 4.2. *Let* $c \in (0,1)$*. Let* $F \in \mathcal{F}$ *be a twice differentiable distribution function on* $\mathbb{R}_+^{d*}$ *with an associated density* f *such that Assumption* **A4** *is satisfied. Assume that there exist* $\kappa > 0$*,* $\zeta > 0$ *such that* $m^{\triangledown} > 0$ *and* $M_H < \infty$*. Let* $T_1 > 0$ *such that for all* $t : \mid t - c \mid \leq \kappa,\ \partial L(t)^{T_1} \neq \emptyset$*. Assume that for each* n*, with probability one,* F_n *is measurable. Let* $(v_{1,n})_{n\in\mathbb{N}^*}$*,* $(v_{2,n})_{n\in\mathbb{N}^*}$ *and* $(T_n)_{n\in\mathbb{N}^*}$ *positive increasing sequences such that Assumption* **A3** *is satisfied. Then it holds that*

$$w_n||r_n\,\mathbf{1}_{\{x\in L_n(c)^{T_n}\}} - r\,\mathbf{1}_{\{x\in L(c)^{T_n}\}}||_p \xrightarrow{\mathbb{P}} 0, \qquad \textit{for } n\to\infty.$$

where $w_n = \min\{v_{1,n}, a_n\}$ *with* $a_n = o\left(v_{2,n}^{\frac{\epsilon}{(\epsilon+1)p(p+1)}}/T_n^{\frac{\epsilon(d-1)}{(\epsilon+1)(p+1)}}\right)$.

The proof of Theorem 4.2 is postponed to Section 10.

Remark 4.2. Note that we do not provide a convergence rate for $||r_n\,\mathbf{1}_{\{x\in L_n(c)^{T_n}\}} - r\,\mathbf{1}_{\{x\in L(c)\}}||_p$ because it implies the knowledge of the rate of convergence of $||r\,\mathbf{1}_{\{x\in L(c)^{T_n}\}} - r\,\mathbf{1}_{\{x\in L(c)\}}||_p = \left(\int_{L(c)\setminus[0,T_n]^d} |r(x)|^p f(x)dx\right)^{1/p}$. Indeed the rate of convergence of $\mathbb{P}[X_1 > T_n, \ldots, X_d > T_n] \to 0$, for $n\to\infty$, is unknown. In order to overcome this problem we should have to assume that the vector $\mathbf{X}$ belongs to a particular distribution class in order to know the rate of decay of its multivariate tails.

5 Covariate-Conditional-Tail-Expectation consistency

A risk measure has recently received growing attention in risk theory literature: the CTE measure. According to Artzner et al. [1], Dedu and Ciumara [12], Denuit et al. [14], for a continuous univariate

loss distribution function F_X the CTE at level $c \in [0,1]$ is defined by

$$\mathrm{CTE}_c(X) = E[\, X \mid X \geq F_X^{-1}(c)].$$

Several multivariate generalizations of the classical univariate CTE have been proposed (see for instance Cai and Li [6]). Using the same approach as Di Bernardino et al. [15], we define the multivariate CTE in such a way as to preserve both the complete information about dependence structure between $\mathbf{X}$ and Y, and the marginal behavior of each component.

Definition 5.1. *Consider $\mathbf{X}$ a random vector with distribution function F and Y a random variable. For $c \in (0,1)$, we define the theoretical multivariate c-Covariate-Conditional-Tail-Expectation as*

$$\mathrm{CCTE}_c(\mathbf{X}, Y) = \mathbb{E}\left[Y \mid \mathbf{X} \in L(c)\right].$$

Using the truncated version of the c-upper level set defined in Section 2 we also define

$$\mathrm{CCTE}_c^{T_n}(\mathbf{X}, Y) = \mathbb{E}\left[Y \mid \mathbf{X} \in L(c)^{T_n}\right].$$

In the following we define consistent estimates for the CCTE in Definition 5.1. In particular, Section 5.1 is devoted to the classical case where we have a whole sample $\{(\mathbf{X}_i, Y_i)\}_{i=1,\dots,n}$. Conversely, in Section 5.2, we investigate a rather different and more difficult case. We suppose that, for any reason, the covariable Y can not be measured on the whole sample. For these two cases we obtain a consistent estimator of the CCTE with a convergence rate.

5.1 Covariable Y is measured

In this subsection, we assume to have an i.i.d. sample $\{(\mathbf{X}_i, Y_i)\}_{i=1,\dots,n}$ (see Section 2) and we introduce the following estimate for the CCTE.

Definition 5.2. *Consider $\mathbf{X}$ a random vector with distribution function F and Y a random variable. For $c \in (0,1)$, using the truncated theoretical multivariate c-Covariate-Conditional-Tail-Expectation introduced in Definition 5.1, we define the associated estimate as*

$$\widehat{\mathrm{CCTE}}_{c,n}^{T_n}(\mathbf{X}, Y) = \mathbb{E}_n\left[Y | \mathbf{X} \in L_n(c)^{T_n}\right],$$

where $\mathbb{E}_n$ denotes the empirical version of the expected value.

Using Definition 5.2, we now establish the consistency of this estimate with a convergence rate.

Theorem 5.1. *Let $c \in (0,1)$. Let $F \in \mathcal{F}$ be a twice differentiable distribution function on $\mathbb{R}_+^{d*}$ with an associated density f such that Assumption* **A4** *is satisfied. Assume that there exist $\kappa > 0$, $\zeta > 0$ such that $m^{\triangledown} > 0$ and $M_H < \infty$. Let $T_1 > 0$ such that for all $t : \mid t - c \mid \leq \kappa$, $\partial L(t)^{T_1} \neq \emptyset$. Assume that*

for each n, with probability one, F_n is measurable. Let $(v_{1,n})_{n\in\mathbb{N}^*}$, $(v_{2,n})_{n\in\mathbb{N}^*}$ and $(T_n)_{n\in\mathbb{N}^*}$ positive increasing sequences such that Assumption **A3** *is satisfied. We have*

$$\beta_n \left| \widehat{\mathrm{CCTE}}_{c,n}^{T_n}(\mathbf{X}, Y) - \mathrm{CCTE}_c^{T_n}(\mathbf{X}, Y) \right| \xrightarrow[n\to\infty]{\mathbb{P}} 0,$$

with

$$\beta_n = \min\left(p_n^{\frac{\epsilon}{2(1+\epsilon)}}, d_n \right)$$

where $p_n = o\left(v_{2,n}^{\frac{p}{p+1}} / T_n^{\frac{d+(d-1)p}{p+1}} \right)$ *and* $d_n = o\left(\sqrt{n}\right)$.

The proof of Theorem 5.1 is postponed to Section 10.

Remark 5.1. It could be interesting to consider the convergence $\left|\widehat{\mathrm{CCTE}}_{c,n}^{T_n}(\mathbf{X}, Y) - \mathrm{CCTE}_c(\mathbf{X}, Y)\right|$. We remark that in this case the speed of convergence will also depend on the convergence rate to zero of $\left|\mathrm{CCTE}_c^{T_n}(\mathbf{X}, Y) - \mathrm{CCTE}_c(\mathbf{X}, Y)\right|$, then, in particular of $\mathbb{P}[(X, Y) \in L(c) \setminus L(c)^{T_n}]$, for $n \to \infty$. In this sense, in Section 7 we will provide an illustration on how to choose T_n, apart from satisfying the assumptions of Theorem 5.1.

Example 2. *Let F_n the d-variate empirical distribution function. Then, it holds that $v_{2,n} = o\left(\sqrt{n}\right)$. From Theorem 5.1, with $p = 2$, we obtain for instance:*

$$p_n = o\left(\frac{n^{1/3}}{T_n^{7/3}} \right), \text{ for } d = 3.$$

This gives us

$$\beta_n = \min\left(o\left(\frac{n^{1/3}}{T_n^{7/3}} \right)^{\frac{\epsilon}{2(1+\epsilon)}}, o\left(n^{1/2}\right) \right).$$

In the case of a bounded density function f we obtain

$$\beta_n = \min\left(o\left(\frac{n^{1/3}}{T_n^{7/3}} \right)^{\frac{1}{2}}, o\left(n^{1/2}\right) \right) = o\left(\frac{n^{1/6}}{T_n^{7/6}} \right).$$

5.2 Covariable Y is partially unknown

In this subsection, we deal with a more difficult case. We suppose that the covariable Y cannot be measured for all the individuals. It could happen if a measure of Y is very expensive or invasive (in some medical treatment, for example). So we have two different i.i.d. samples : $S_N^1 = \{(\mathbf{X}_i, Y_i)\}_{i=1}^N$ and $S_n^2 = \{\mathbf{X}_j\}_{j=1}^n$, with n potentially much bigger than N.

In this case we use S_N^1 to get an estimate r_N of the regression function r. Then, we apply this estimate on the sample S_n^2 in order to estimate the CCTE measure. To this aim we define:

Definition 5.3. *Consider* $\mathbf{X}$ *a random vector with distribution function* F*,* Y *a random variable and two i.i.d samples* S_N^1 *and* S_n^2*. Using the truncated version of the c-upper level set defined in Section 2, for* $c \in (0,1)$*, we define*

$$\widehat{\mathrm{CCTE}}^{\star\, T_n}_{c,n,N}(\boldsymbol{X}, Y) = \mathbb{E}_n \left[r_N(\boldsymbol{X}) | \boldsymbol{X} \in L_n(c)^{T_n} \right].$$

The following result proves the consistency of the estimate introduced in Definition 5.3.

Theorem 5.2. *Let* $c \in (0,1)$*. Let* $F \in \mathcal{F}$ *be a twice differentiable distribution function on* $\mathbb{R}_+^{d*}$ *with an associated density* f *such that Assumption* **A4** *is satisfied. Assume that there exist* $\kappa > 0$*,* $\zeta > 0$ *such that* $m^{\triangledown} > 0$ *and* $M_H < \infty$*. Let* $T_1 > 0$ *such that for all* $t : \mid t - c \mid \leq \kappa$*,* $\partial L(t)^{T_1} \neq \emptyset$*. Assume that for each* n*, with probability one,* F_n *is measurable. Let* $(v_{1,N})_{N \in \mathbb{N}^*}$*,* $(v_{2,n})_{n \in \mathbb{N}^*}$ *and* $(T_n)_{n \in \mathbb{N}^*}$ *positive increasing sequences such that Assumption* **A3** *is satisfied. Let* r *be a continuous positive regression function such that* $E\left[r(X)^2\right] < \infty$*. We have*

$$\beta_{n,N} \left| \widehat{\mathrm{CCTE}}^{\star\, T_n}_{c,n,N}(\boldsymbol{X}, Y) - \mathrm{CCTE}_c^{T_n}(\boldsymbol{X}, Y) \right| \xrightarrow{\mathbb{P}} 0, \qquad \text{for } N, n \to \infty,$$

with

$$\beta_{n,N} = \min\left(p_n^{\frac{\epsilon}{2(1+\epsilon)}}, c_N, d_n \right),$$

where $p_n = o\left(v_{2,n}^{\frac{p}{p+1}} / T_n^{\frac{d+(d-1)p}{p+1}} \right)$*,* $c_N = o\left(\mathbb{E}\left|r_N(\boldsymbol{X}) - r(\boldsymbol{X})\right|\right)$ *and* $d_n = o\left(\sqrt{n}\right)$*.*

The proof of Theorem 5.2 is postponed to Section 10. Note that we have a supplementary term (c_N) comparing to Theorem 5.1. This term controls the rate of convergence of r_N toward r.

Example 3. *Using the same context as in Example 2, taking* r_N *the kernel regression estimator of Kohler et al. [23] and assuming that* r *is* $(1, C)$*-smooth, we have*

$$c_N = o\left(N^{2/5}\right).$$

Then we obtain

$$\beta_{n,N} = \min\left(o\left(\frac{n^{1/3}}{T_n^{7/3}} \right)^{\frac{\epsilon}{2(1+\epsilon)}}, o\left(N^{2/5}\right), o\left(n^{1/2}\right) \right).$$

In the case of a bounded density function f *and as long as* $N >> \frac{n^{5/12}}{T_n^{35/12}}$*, we obtain*

$$\beta_{N,n} = \min\left(o\left(\frac{n^{1/3}}{T_n^{7/3}} \right)^{\frac{1}{2}}, o\left(N^{2/5}\right), o\left(n^{1/2}\right) \right) = o\left(\frac{n^{1/6}}{T_n^{7/6}} \right)$$

which is the same rate as in Example 2.

5.3 Estimating CCTE measure for large value of the risk level using the empirical estimator F_n

In Section 7 we provide some numerical illustrations for the estimation of the c-Covariate-Conditional-Tail-Expectation, proposed in Definition 5.2. In particular, we use the empirical estimator F_n to construct the plug-in estimation of level sets $L_n(c)$. In other words, we use an empirical version of the estimation in Definition 5.2.

The aim of this subsection is to briefly discuss the adequacy of an empirical version of estimation of Definition 5.2, in term of the risk level c. From a practical point of view, as we deal with a risk measure (CCTE) then we are interested in considering some large values of the risk level c. On the other hand, given an i.i.d sample $\{(\mathbf{X}_i, Y_i)\}_{i=1,\dots,n}$, we apply an empirical version of the estimate $\widehat{\mathrm{CCTE}}_{c,n}(\mathbf{X}, Y)$ in the case of *large but not-extreme* values of c, i.e, when some data fall in the c-upper level set $L_n(c)$.

Although in any practical application c is fixed, we have to choose c in order to obtain both some good empirical estimation of our CCTE and an interesting indicator for the associated risk phenomenon. From a mathematical point of view, to obtain this type of compromise, we have to describe the difference between high quantiles within and outside the sample. Following Embrechts et al. [18], we have that:

(a) c is a high quantiles within the sample if $c = c_n \xrightarrow[n\to\infty]{} 1$ and $n\,(1 - c_n) \xrightarrow[n\to\infty]{} \chi \in (1, +\infty]$,

(b) c is a high quantiles outside the sample if $c = c_n \xrightarrow[n\to\infty]{} 1$ and $n\,(1 - c_n) \xrightarrow[n\to\infty]{} \chi \in [0, 1)$.

The case (a) for $\chi = +\infty$ is addressed by Theorem 3.1 in Dekkers and de Haan [13]. Under these assumptions on the risk level c_n, this result basically tells us that we can just use empirical function for estimating the associated quantile. Following these consideration, in our numerical illustrations we pay attention to this compromise in the choice of the risk level c, in order to estimate high quantiles within the sample in our risk measure CCTE (see Section 7).

6 About the effects of scaling data

In this section we study the impact of a change in the scale of data with respect to the estimate $\widehat{\mathrm{CCTE}}^{T_n}_{c,n}(\mathbf{X}, Y)$. In particular, this property is related to the suitable positive homogeneity property of risk measures (e.g. see Artzner et al. [1]). In the literature, the homogeneity property of a risk measure is often motivated by a change of currency argument: the amount of required capital in order to manager risks should be independent of the currency in which it is expressed.

Given a set $A \subset \mathbb{R}_+^d$ we denote by ∂A its boundary, and by $\beta\, A$ the scaled set $\{\beta\, x, \text{ with } x \in A\}$. Suppose now that we scale our data using a scale parameter $a \in \mathbb{R}_+^*$. In our case, the scaled random vector will be $(a\, X_1, a\, X_2, \ldots, a\, X_d) := a\, \mathbf{X}$. From now on we denote $F_{a\,\mathbf{X}}$ (resp. $F_{\mathbf{X}}$) the distribution function associated to $a\, \mathbf{X}$ (resp. to $\mathbf{X}$). Using notation of Section 2, let

$$L_a(c) = \{x \in \mathbb{R}_+^d : F_{a\,\mathbf{X}}(x) \geq c\}.$$

It is easy to prove (see for instance Section 3 in Tibiletti [29]) that

$$L_a(c) = a\, L(c) \quad \text{and} \quad L_a(c)^T = aL(c)^T$$

and furthermore

$$E_a = B(\{x \in \mathbb{R}_+^d : \mid F_{a\,\mathbf{X}}(x) - c \mid \leq r\},\, \zeta) = a\, E.$$

Moreover, if we suppose $F_{n,a\mathbf{X}}(x) = F_{n,\mathbf{X}}(x/a)$, thus

$$L_{n,a}(c) = a\, L_n(c) \quad \text{and} \quad L_{n,a}(c)^T = aL_n(c)^T.$$

Define now

$$m_a^{\triangledown} = \inf_{x \in E_a} \|\nabla F_{a\,\mathbf{X}}(x)\|.$$

First, we can obtain the following lemma.

Lemma 6.1. *It holds that*

$$m_a^{\triangledown} = \frac{1}{a}\, m^{\triangledown}, \quad \forall\, a \in \mathbb{R}_+^*.$$

Furthermore, if

$$M_H = \sup_{x \in E} \|(HF_{\mathbf{X}})_x\| < +\infty \quad \textit{then} \quad M_{H,a} = \sup_{x \in a\, E} \|(HF_{a\,\mathbf{X}})_x\| < +\infty, \quad \textit{with} \;\; a \in \mathbb{R}_+^*.$$

The proof of Lemma 6.1 is postponed to Section 10.

We can now consider the effects of scaling data with respect to the volume of the symmetric difference. Corollary 6.1 is a straightforward consequence of Theorem 3.1 using the fact that $L_a^{T_n}(c) \Delta L_{n,a}^{T_n}(c) = a\left(L^{T_n}(c) \Delta L_n^{T_n}(c)\right)$, for $a \in \mathbb{R}_+^*$.

Corollary 6.1. *Assume that $F_{n,a\mathbf{X}}(x) = F_{n,\mathbf{X}}(x/a)$. Under same notations and assumptions of Theorem 3.1 it holds that*

$$p_{n,\,a}\, d_\lambda(L_a(c)^{T_n}, L_{n,\,a}(c)^{T_n}) \underset{n \to \infty}{\overset{\mathbb{P}}{\to}} 0,$$

with $p_{n,\,a}$ an increasing positive sequence such that $p_{n,\,a} = o\left(v_n^{\frac{1}{p+1}} / \left(a^d\, T_n^{\frac{(d-1)\,p}{p+1}}\right)\right)$.

Corollary 6.1 states that a change of scale of the data implies a convergence rate as

$$o\left(v_n^{\frac{1}{p+1}} / \left(a^d\, T_n^{\frac{(d-1)\,p}{p+1}}\right)\right)$$

instead of

$$o\left(v_n^{\frac{1}{p+1}}/T_n^{\frac{(d-1)p}{p+1}}\right).$$

Logically, it means that the scale factor $a \in \mathbb{R}_+^*$ impacts the volume in $\mathbb{R}^d$ with an exponent d.

Let us now investigate the effects of scaling data on Theorem 4.2 and Theorem 5.1. Let denote $r_{a\mathbf{X},bY}(x) = E[b\,Y \mid a\,\mathbf{X} = x]$ ($r_{n,a\mathbf{X},bY}$ its estimate version) and $\mathrm{CCTE}_c^{T_n}(a\mathbf{X}, bY) = \mathbb{E}\left[b\,Y \mid a\,\mathbf{X} \in L_a(c)^{T_n}\right]$ ($\widehat{\mathrm{CCTE}}_{c,n}^{T_n}(a\mathbf{X}, bY)$ its estimate version). We obtain the following result.

Theorem 6.1. *Let a and b in $\mathbb{R}_+^*$ and assume that $r_{n,a\mathbf{X},bY}(x) = b\,r_{n,\mathbf{X},Y}\left(\frac{x}{a}\right)$.*

1. *Under same hypothesis and notations of Theorem 4.2 it holds that*

$$b\,||r_{n,\mathbf{X},Y}\,\mathbf{1}_{\{x\in L_n(c)^{T_n}\}} - r_{\mathbf{X},Y}\,\mathbf{1}_{\{x\in L(c)^{T_n}\}}||_p = ||r_{n,a\mathbf{X},bY}\,\mathbf{1}_{\{x\in L_{a,n}(c)^{T_n}\}} - r_{a\mathbf{X},bY}\,\mathbf{1}_{\{x\in L_a(c)^{T_n}\}}||_p.$$

2. *Under same hypothesis and notations of Theorem 5.1 it holds that*

$$b\left|\widehat{\mathrm{CCTE}}_{c,n}^{T_n}(\mathbf{X}, Y) - \mathrm{CCTE}_c^{T_n}(\mathbf{X}, Y)\right| = \left|\widehat{\mathrm{CCTE}}_{c,n}^{T_n}(a\mathbf{X}, bY) - \mathrm{CCTE}_c^{T_n}(a\mathbf{X}, bY)\right|.$$

The proof of Theorem 6.1 is postponed to Section 10. Using Theorem 6.1 we obtain that the rates of convergence of Theorem 4.2 and Theorem 5.1 are not affected by any scaling on the data.

Remark 6.1. Note that if r_n is the classical kernel estimator, assumption $r_{n,a\mathbf{X}}(x) = r_{n,\mathbf{X}}\left(\frac{x}{a}\right)$ is not automatically satisfied. However, it can be satisfied if the scaling is also applied to the bandwidth.

7 Illustrations

In the following we consider some different simulated cases for which we illustrate the finite sample properties of our estimation of r and $\mathrm{CCTE}_c(\mathbf{X}, Y)$. In particular, we will consider an independent copula (Section 7.1), and Ali-Mikhail-Haq copula (Section 7.2). To compare the estimated results with the theoretical ones we consider cases for which we can calculate (using Maple) the explicit value of the theoretical $\mathrm{CCTE}_c(\mathbf{X}, Y)$. However, our estimator can be applied to much more general cases.

In this section we consider the kernel regression estimate proposed by Kohler et al. [23] in order to estimate r. Furthermore, the plug-in estimation of level sets, i.e., $L_n(c)$, is constructed using the empirical estimator F_n of the distribution function.

Following considerations of Section 5.3, choosing $c_n = \log(2n)/(1 + \log(n))$, the assumption of case (a) of Section 5.3 are satisfied with $\chi = +\infty$. This means that, for a sample size $n = 1000$ (resp. $n = 10000$), we could use the empirical estimation for the univariate quantile for $c < 0.96$,

(resp. $c < 0.97$). These considerations give us also an indication in terms of the sample size n and the magnitude of risk level c to use our empirical estimate $\widehat{\mathrm{CCTE}}_{c,n}(\mathbf{X}, Y)$.

7.1 Independent Copula

We consider here a bivariate independent copula with exponentially distributed marginals with parameter 1. Furthermore we choose the regression function $r(x, y) = 1/(\mathrm{e}^x \mathrm{e}^y)$, for all $x, y \geq 0$ (Figure 1.1).

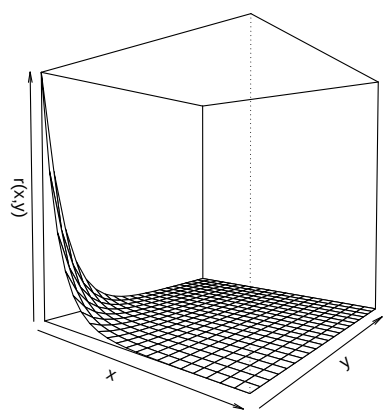

Figure 1.1: Regression function: $r(x, y) = 1/(\mathrm{e}^x \mathrm{e}^y)$ for all $x, y \geq 0$.

L_p-consistency of r_n

In this section we provide an illustration of Theorem 4.2. In particular, we denote

$$E_{L_2} := ||r_n \, \mathbf{1}_{\{x \in L_n(c)^{T_n}\}} - r \, \mathbf{1}_{\{x \in L(c)^{T_n}\}}||_2.$$

According to Remark 4.2, we take $T_n = \log(n)$ (the interested reader is also referred to Di Bernardino et al. [15]). In Table 1.1 we illustrate the mean of E_{L_2} (i.e., L_2-consistency for the estimator r_n) on $M = 50$ simulated samples, for different level of risk c and different sample size n.

As expected, the performance of the estimation increases with the size of the sample. However, the fact that the quality for the E_{L_2} error seems better for large risk level c is quite surprising. This behavior of the estimation comes from the fact that the regression r has a flat plateau for $x, y \to \infty$ (see Figure 1.1).

n	Level $c = 0.1$	Level $c = 0.25$	Level $c = 0.5$	Level $c = 0.7$	Level $c = 0.9$
n=1000, E_{L_2}	0.0567	0.0377	0.0254	0.0197	0.0131
n=5000, E_{L_2}	0.0323	0.0176	0.0072	0.0051	0.0024
n=10000, E_{L_2}	0.0256	0.0134	0.0051	0.0025	0.0016

Table 1.1: E_{L_2} for independent copula and exponentially distributed marginals with parameter 1. The regression function is $r(x, y) = 1/(\mathrm{e}^x \mathrm{e}^y)$ for all $x, y \geq 0$.

$\mathrm{CCTE}_c(\mathbf{X}, Y)$ **estimation**

In the following, we compare $\widehat{\mathrm{CCTE}}_{c,n}^{T_n}(\mathbf{X}, Y)$ with the theoretical $\mathrm{CCTE}_c(\mathbf{X}, Y)$. Following Remark 5.1, we obtain that $\left|\mathrm{CCTE}_c(\mathbf{X}, Y) - \widehat{\mathrm{CCTE}}_c^{T_n}(\mathbf{X}, Y)\right|$ decays to zero at least with a convergence rate $\beta_n = o\left(\frac{n^{1/6}}{\ln(n)^{4/6}}\right)$, with a choice of sequence $T_n = \ln(n)$. This kind of compromise provides an illustration on how to choose T_n, apart from satisfying the assumptions of Theorem 5.1.

In the following tables we denote Mean= $\overline{\widehat{\mathrm{CCTE}}_c^{T_n}}(\mathbf{X}, Y)$, i.e., the mean of $\widehat{\mathrm{CCTE}}_c^{T_n}(\mathbf{X}, Y)$ on $M = 50$ simulated samples. We denote $\widehat{\sigma}$ the empirical standard deviation

$$\widehat{\sigma} = \sqrt{\frac{1}{M-1} \sum_{j=1}^{M} \left(\widehat{\mathrm{CCTE}}_c^{T_n}(\mathbf{X}, Y)_j - \overline{\widehat{\mathrm{CCTE}}_c^{T_n}}(\mathbf{X}, Y) \right)^2}.$$

Finally, we denote RMAE the Relative Mean Absolute Error, i.e.,

$$\text{RMAE}= \frac{1}{M} \sum_{j=1}^{M} \frac{\left|\widehat{\mathrm{CCTE}}_c^{T_n}(\mathbf{X},Y)_j - \mathrm{CCTE}_c(\mathbf{X},Y)\right|}{\mathrm{CCTE}_c(\mathbf{X},Y)}.$$

The results are gathered in Table 1.2. Furthermore, in Table 1.3, we provide an illustration of the convergence rate of Theorem 5.1. In this case, we remark that $\beta_n = o\left(\frac{n^{1/6}}{\ln(n)^{4/6}}\right)$ is at least the convergence rate of this CCTE estimation.

As expected, the greater n is, the better the estimations are. Furthermore, results in Table 1.3 set out how $\beta_n = o\left(\frac{n^{1/6}}{\ln(n)^{4/6}}\right)$ is at least the convergence rate of $\left|\widehat{\mathrm{CCTE}}_{c,n}^{T_n}(\mathbf{X}, Y) - \mathrm{CCTE}_c(\mathbf{X}, Y)\right|$, in this particular case.

n		Level $c = 0.1$ $\mathrm{CCTE}_{0.1} = 0.14293$	Level $c = 0.25$ $\mathrm{CCTE}_{0.25} = 0.07927$	Level $c = 0.5$ $\mathrm{CCTE}_{0.5} = 0.02791$	Level $c = 0.7$ $\mathrm{CCTE}_{0.7} = 0.00879$	Level $c = 0.9$ $\mathrm{CCTE}_{0.9} = 0.00087$
1000	RMAE	0.0357	0.0388	0.0725	0.1187	0.2898
	Mean	0.1432	0.0798	0.0281	0.0090	0.0009
	$\widehat{\sigma}$	0.0062	0.0038	0.0026	0.0012	0.0003
5000	RMAE	0.0151	0.0194	0.0317	0.0484	0.1623
	Mean	0.1425	0.0795	0.0278	0.0087	0.0008
	$\widehat{\sigma}$	0.0027	0.0021	0.0011	0.0004	0.0002
10000	RMAE	0.0110	0.0137	0.0211	0.0453	0.0875
	Mean	0.1428	0.0792	0.0279	0.0088	0.0008
	$\widehat{\sigma}$	0.0019	0.0013	0.0007	0.0004	$9.43\,10^{-5}$

Table 1.2: Estimation of $\mathrm{CCTE}_c(\mathbf{X}, Y)$ in the case of independent copula and exponentially distributed marginals with parameter 1.

	Level $c = 0.1$ $\mathrm{CCTE}_{0.1} = 0.14293$	Level $c = 0.25$ $\mathrm{CCTE}_{0.25} = 0.07927$	Level $c = 0.5$ $\mathrm{CCTE}_{0.5} = 0.02791$	Level $c = 0.7$ $\mathrm{CCTE}_{0.7} = 0.00879$	Level $c = 0.9$ $\mathrm{CCTE}_{0.9} = 0.00087$
n=1000, RMAE	0.0311	0.0339	0.0632	0.1035	0.2526
n=5000, RMAE	0.0151	0.0191	0.0314	0.04798	0.1609
n=10000, RMAE	0.0116	0.0145	0.0223	0.0478	0.0925

Table 1.3: Approximated $\beta_n \left| \widehat{\mathrm{CCTE}}_{c,n}^{T_n}(\mathbf{X}, Y) - \mathrm{CCTE}_c(\mathbf{X}, Y) \right|$, with $\beta_n = o\left(\frac{n^{1/6}}{\ln(n)^{4/6}}\right)$, in the case of independent copula and exponentially distributed marginals with parameter 1.

7.2 Ali-Mikhail-Haq Copula

We consider here a bivariate Ali-Mikhail-Haq copula with parameter 0.5 and exponentially distributed marginals with parameter 1 (see Nelsen [25]). We take the regression function $r(x, y) = -(-2 + \mathrm{e}^{-x-y})^3$, for all $x, y \geq 0$. Conversely to previous section we now deal with dependent variables $X_1, \ldots, X_n$ and an increasing regression function r (see Figure 1.2).

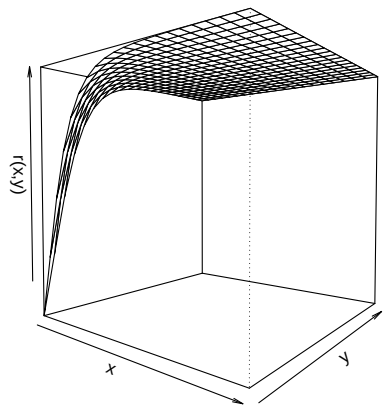

Figure 1.2: Regression function $r(x, y) = -(-2 + \mathrm{e}^{-x-y})^3$ for all $x, y \geq 0$.

L_p-consistency of r_n

In this section we provide an illustration of Theorem 4.2. As explained in Remark 5.1, we choose $T_n = n^{0.2}$. In Table 1.4 we illustrate E_{L_2} (i.e. the L_2-consistency for the estimator r_n) for different level of risk c and different sample size n.

n	Level $c = 0.1$	Level $c = 0.25$	Level $c = 0.5$	Level $c = 0.7$	Level $c = 0.9$
n=1000, E_{L_2}	0.7736	0.7980	0.7599	0.5817	0.2928
n=5000, E_{L_2}	0.4846	0.5801	0.5047	0.3975	0.1781
n=10000, E_{L_2}	0.4167	0.4883	0.4459	0.3612	0.1679

Table 1.4: E_{L_2} for Ali-Mikhail-Haq copula with parameter 0.5 and exponentially distributed marginals with parameter 1. The regression function is $r(x, y) = -(-2 + \mathrm{e}^{-x-y})^3$, for all $x, y \geq 0$.

As before, we see that the quality of the estimation is better when the size of the sample growths. Again, the presence of a plateau for $r(x, y)$ when $x, y \to \infty$ explains the good results for large risk levels c (see Figure 1.2).

$\mathrm{CCTE}_c(\mathbf{X}, Y)$ **estimation**

In the following, we compare $\widehat{\mathrm{CCTE}}_{c,n}^{T_n}(\mathbf{X}, Y)$ with the theoretical $\mathrm{CCTE}_c(\mathbf{X}, Y)$ in the case of Ali-Mikhail-Haq copula with parameter 0.5 (see Table 1.5). Furthermore, following Remark 5.1, we obtain that $\left|\,\mathrm{CCTE}_c(\mathbf{X}, Y) - \widehat{\mathrm{CCTE}}_c^{T_n}(\mathbf{X}, Y)\,\right|$ decays to zero at least with a convergence rate $\beta_n = o\left(n^{1/10}\right)$, with a choice of sequence $T_n = n^{1/5}$ (see Table 1.6). Again, our theoretical results are confirmed by these simulations.

n		Level $c = 0.1$ $\mathrm{CCTE}_{0.1} = 6.37526$	Level $c = 0.25$ $\mathrm{CCTE}_{0.25} = 7.03937$	Level $c = 0.5$ $\mathrm{CCTE}_{0.5} = 7.65131$	Level $c = 0.7$ $\mathrm{CCTE}_{0.7} = 7.89094$	Level $c = 0.9$ $\mathrm{CCTE}_{0.9} = 7.98936$
1000	RMAE	0.0081	0.0061	0.0038	0.0015	0.0004
	Mean	6.3885	7.0421	7.6544	7.8917	7.9887
	$\widehat{\sigma}$	0.0639	0.0526	0.0345	0.0147	0.0041
5000	RMAE	0.0032	0.0028	0.0015	0.0006	0.0001
	Mean	6.3761	7.0458	7.6461	7.8906	7.9891
	$\widehat{\sigma}$	0.0259	0.0249	0.0138	0.0071	0.0013
10000	RMAE	0.0021	0.0023	0.0008	0.0005	0.0001
	Mean	6.3725	7.0302	7.6527	7.8911	7.9893
	$\widehat{\sigma}$	0.0171	0.0173	0.0081	0.0047	0.0012

Table 1.5: Estimation of $\mathrm{CCTE}_c(\mathbf{X}, Y)$ in the case of Ali-Mikhail-Haq copula with parameter 0.5 and exponentially distributed marginals with parameter 1.

	Level $c = 0.1$ $\mathrm{CCTE}_{0.1} = 6.37526$	Level $c = 0.25$ $\mathrm{CCTE}_{0.25} = 7.039378$	Level $c = 0.5$ $\mathrm{CCTE}_{0.5} = 7.651305$	Level $c = 0.7$ $\mathrm{CCTE}_{0.7} = 7.890936$	Level $c = 0.9$ $\mathrm{CCTE}_{0.9} = 7.98936$
n=1000, RMAE	0.0102	0.0076	0.0033	0.0018	0.0005
n=5000, RMAE	0.0043	0.0051	0.0021	0.0007	0.0001
n=10000, RMAE	0.0028	0.0032	0.0012	0.0006	0.0001

Table 1.6: Approximated $\beta_n\left|\widehat{\mathrm{CCTE}}_{c,n}^{T_n}(\mathbf{X}, Y) - \mathrm{CCTE}_c^{T_n}(\mathbf{X}, Y)\right|$, with $\beta_n = o\left(n^{1/10}\right)$ in the case of Ali-Mikhail-Haq copula with parameter 0.5 and exponentially distributed marginals with parameter 1.

8 Real data study : waves and water levels in coastal engineering design

On coasts with high tidal ranges, or subject to high surges, both still water levels and waves can be important in assessing flood risk; their relative importance depends on location and on the type of

sea defence. The simultaneous occurrence of large waves and a high still water level is, therefore, important in estimating their combined effect on sea defences. In design of a sea defence, a key step is the estimation of the probability of failure to protect against sea conditions. It is important in engineering design to identify the combinations of sea condition variables which cause each failure. The interested reader is referred for instance to Hawkes et al. [20].

For any particular mode of failure (structural failure, excessive overtopping, ...), the regression function r is dependent on the sea condition variables. Then, in particular, at any particular time t, the overtopping covariate Y will be related to sea condition vector ($\mathbf{X}$). In the literature the sea condition variables are often represented by the significant Wave height Hm0 (X_1), the Still Water level SWL (X_2), and the Wave period Tpb (X_3), then $\mathbf{X} = (X_1, X_2, X_3)$ (see Figure 1.3).

The regression function $r(x) := \mathbb{E}[Y \,|\, \mathbf{X} = x]$ represents the relationship between the sea conditions and the overtopping at a given time t. This relationship could be complex and in some real analysis, can be represented by equations. The most advantage of the use of this regression function is to reduce a joint probability risk problem to a single covariate problem.

In this section we analyze the Wave height (Hm0), Still Water level (SWL), Wave period (Tpb) data, recorded during 828 storm events spread over 13 years in front of the Dutch coast near the town of Petten (Figure 1.4).

These data has been recently studied in the literature (for details see for instance Draisma et al. [16]). Following Tau and Dam [28], at a given time t, the principal equation used for overtopping discharge (l/m/s) Y is given by:

$$Y = a\, \mathrm{e}^{\frac{-b\,(h-SWL)}{Hm0}} \sqrt{g\,(Hm0)^3}, \tag{8.1}$$

with

$$a = 0.04\sqrt{\frac{\tan(\beta)\,L}{Hm0}}, \qquad b = \frac{7.05\sqrt{\frac{Hm0}{L}}}{\tan(\beta)}, \qquad L = \frac{g\,(Tpb)^2}{2\,\pi},$$

where

- Hm0 (m) is the wave height at the toe of the structure at time t;
- Tpb (s) is the wave period at time t. In particular the number of waves in a storm (N) can easily be computed from information about the wave period and the duration of the storm, i.e., $N =$ (duration(h)/Tpb(s))·3600;
- SWL (m) the level of the sea if it is flat, without any waves at time t;

- h (m) is the height of the costal design above SWL= 0 (see Figure 1.3);
- g (m/s^2) the gravitational acceleration (i.e., 9.8 m/s^2);
- β (rad) is the seaward slope steepness. In the following we consider the case $\tan(\beta) = 0.3$ and $\tan(\beta) = 0.6$ (see Figure 1.3).

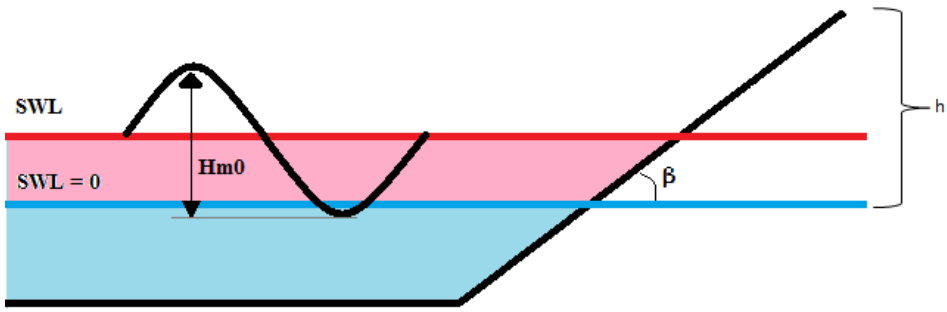

Figure 1.3: Definition of some parameters for the calculation of overtopping.

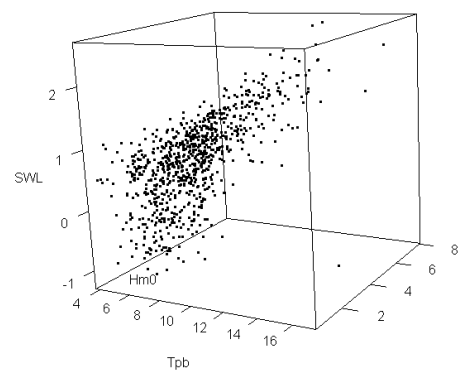

Figure 1.4: Representation of the data in three dimensions.

Note that Theorem 6.1 allow us to make any changes on the units before or after the estimation. We now estimate the mean overtopping rate (i.e., $\mathrm{CCTE}_c(\mathbf{X}, Y)$) using Equation (8.1) for Y and conditionally to the fact that the sea variable conditions $\mathbf{X}$ belong to the joint risk area $L(c)$. In particular, we consider the jointly large realisations of sea variables Hm0, SWL and Tpb for different values of risk level c. Obviously, the dangerous effects of the sea conditions will be increasing with respect to c. The results are gathered in Figure 1.5.

As expected, we obtain a decreasing behavior of mean overtopping discharge according to the height of the costal design h for a fixed risk level c. Furthermore, for a fixed height of the costal design, the higher the risk c, the higher the mean overtopping discharge. In our study we consider

both the case $\tan(\beta) = 0.3$ (Figure 1.5, left) and $\tan(\beta) = 0.6$ (Figure 1.5, right). As expected, we remark that for fixed risk level c and height of costal design h, a lower seaward slope steepness β generates a smaller mean overtopping discharge.

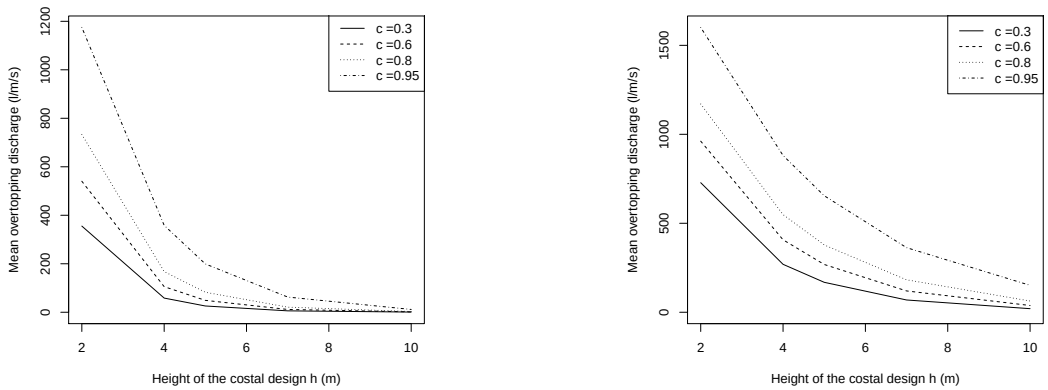

Figure 1.5: Mean overtopping discharge $\mathrm{CCTE}_c(\mathbf{X}, Y)$ in function of the height of the costal design h for different condition of risk-sea variables. We take here $\tan(\beta) = 0.3$ (left) and $\tan(\beta) = 0.6$ (right).

9 Conclusion

We propose in this paper a generalization to the estimation of the level sets of a d-variate distribution function. The non-compactness of the level sets requires special attention in the statement of the problem. The consistency results with a convergence rate are stated in term of the volume of the symmetric difference. In a second part, we analyze the problem of the estimation of a regression function on the levels sets of a d-variate distribution function and we obtain the consistency with a convergence rate in term of the L_p-distance. Then, we study a new multivariate risk measure: the Covariate-Conditional-Tail-Expectation, i.e. the Conditional-Tail-Expectation of the regression function. A consistent estimator and a rate of convergence are provided. Moreover, we analyze the impact of scaling data on our results. Our theoretical results are illustrated on a complete simulation study. We discuss a real application in the evaluation of the mean overtopping discharge conditionally to the fact that the sea variable conditions belong to some joint risk area. It highlights the importance of the parameter T_n (which solved the problem of the compactness of the level sets) as well as the curse of the dimensionality. An interesting future work could be a deeply investigation about these points, with a focus on the optimal choice for this parameter.

10 Proofs

Proof of Theorem 3.1: Under assumptions of Theorem 3.1, we can always take $T_1 > 0$ such that for all $t : |t-c| \leq \kappa$, $\partial L(t)^{T_1} \neq \emptyset$. Then for each n, for all $t : |t-c| \leq \kappa$, $\partial L(t)^{T_n}$ is a non-empty (and compact) set on $\mathbb{R}_+^d$.

We consider a positive sequence ε_n such that $\varepsilon_n \underset{n\to\infty}{\to} 0$. For each $n \geq 1$ the random sets $L(c)^{T_n} \triangle L_n(c)^{T_n}$, $Q_{\varepsilon_n} = \{x \in [0,T_n]^d : | F - F_n | \leq \varepsilon_n\}$ and $\widetilde{Q}_{\varepsilon_n} = \{x \in [0,T_n]^d : | F - F_n | > \varepsilon_n\}$ are measurable and

$$\lambda(L(c)^{T_n} \triangle L_n(c)^{T_n}) = \lambda(L(c)^{T_n} \triangle L_n(c)^{T_n} \cap Q_{\varepsilon_n}) + \lambda(L(c)^{T_n} \triangle L_n(c)^{T_n} \cap \widetilde{Q}_{\varepsilon_n}).$$

Since $L(c)^{T_n} \triangle L_n(c)^{T_n} \cap Q_{\varepsilon_n} \subset \{x \in [0,T_n]^d : c - \varepsilon_n \leq F < c + \varepsilon_n\}$ we obtain

$$\lambda(L(c)^{T_n} \triangle L_n(c)^{T_n}) \leq \lambda(\{x \in [0,T_n]^d : c - \varepsilon_n \leq F < c + \varepsilon_n\}) + \lambda(\widetilde{Q}_{\varepsilon_n}).$$

From Assumption **H** (Section 2) and Proposition 2.1 in Di Bernardino et al. [15], it follows that there exists a $\gamma > 0$ such that, if $2\,\varepsilon_n \leq \gamma$ then

$$d_H(\partial L(c+\varepsilon_n)^{T_n}, \partial L(c-\varepsilon_n)^{T_n}) \leq 2\,\varepsilon_n\,A$$

where $A = \frac{2}{m^{\nabla}}$ and d_H is the Hausdorff distance. From assumptions on first derivatives of F and Property 1 in Imlahi et al. [21], we can write

$$\lambda(\{x \in [0,T_n]^d : c - \varepsilon_n \leq F < c + \varepsilon_n\}) \leq (2\,\varepsilon_n\,A)\,d\,T_n^{d-1}.$$

Interestingly we remark that in the univariate case ($d = 1$) the Hausdorff distance between the two points $\partial L(c-\varepsilon_n)^{T_n}$ and $\partial L(c+\varepsilon_n)^{T_n}$ is also the Lebesgue measure (in dimension 1) for this interval. Then $\lambda(\{x \in [0,T_n] : c - \varepsilon_n \leq F < c + \varepsilon_n\}) \leq 2\,\varepsilon_n\,A$. This means that in this case, the result does not depend on the truncation sequence T_n.

If we now choose

$$\varepsilon_n = o\left(\frac{1}{p_n\,T_n^{d-1}}\right), \tag{10.1}$$

we obtain that, for n large enough, $2\,\varepsilon_n \leq \gamma$ and

$$p_n\,\lambda(\{x \in [0,T_n]^d : c - \varepsilon_n \leq F < c + \varepsilon_n\}) \underset{n\to\infty}{\to} 0.$$

Let us now prove that $p_n\,\lambda(\widetilde{Q}_{\varepsilon_n}) \underset{n\to\infty}{\overset{\mathbb{P}}{\to}} 0$. To this end, we write

$$p_n\,\lambda(\widetilde{Q}_{\varepsilon_n}) = p_n \int 1_{\{x\in[0,T_n]^d:\,|F-F_n|>\varepsilon_n\}}\,\lambda(\mathrm{d}x) \leq \frac{p_n}{\varepsilon_n^p} \int_{[0,T_n]^d} | F - F_n |^p\ \lambda(\mathrm{d}x).$$

Take ε_n such that

$$\varepsilon_n = \left(\frac{p_n}{v_n}\right)^{\frac{1}{p}}. \tag{10.2}$$

So, from Assumption **A1** in Section 3, we obtain $p_n\,\lambda(\widetilde{Q}_{\varepsilon_n}) \underset{n\to\infty}{\overset{\mathbb{P}}{\to}} 0$. As $p_n = o\left(v_n^{\frac{1}{p+1}}/T_n^{\frac{(d-1)\,p}{p+1}}\right)$ we can choose ε_n that satisfies (10.1) and (10.2). Hence the result. □

In the following proofs, K denotes a constant which value may change from line to line.

Proof of Theorem 4.1: We have

$$\begin{aligned}
||r\,\mathbf{1}_{\{x\in L_n(c)^{T_n}\}} - r\,\mathbf{1}_{\{x\in L(c)^{T_n}\}}||_p &= \left(\int_{\mathbb{R}^d}\left|r(x)\,\mathbf{1}_{\{x\in L_n(c)^{T_n}\}} - r(x)\,\mathbf{1}_{\{x\in L(c)^{T_n}\}}\right|^p f(x)\,\mathrm{d}x\right)^{1/p}\\
&\le \left(\int_{L(c)^{T_n}\Delta L_n(c)^{T_n}} |r(x)|^p\,f(x)\,\mathrm{d}x\right)^{1/p}\\
&\le K\,\left(\mu\left(L(c)^{T_n}\Delta L_n(c)^{T_n}\right)\right)^{1/p}\\
&\le K\left(||f||_{1+\epsilon,\lambda}\,\left(\lambda\left(L(c)^{T_n}\Delta L_n(c)^{T_n}\right)\right)^{\frac{\epsilon}{1+\epsilon}}\right)^{1/p}
\end{aligned}$$

which gives us

$$\begin{aligned}
||r_n\,\mathbf{1}_{\{x\in L_n(c)^{T_n}\}} - r\,\mathbf{1}_{\{x\in L(c)^{T_n}\}}||_p &\le ||r_n\,\mathbf{1}_{\{x\in L_n(c)^{T_n}\}} - r\,\mathbf{1}_{\{x\in L_n(c)^{T_n}\}}||_p\\
&+ ||r\,\mathbf{1}_{\{x\in L_n(c)^{T_n}\}} - r\,\mathbf{1}_{\{x\in L(c)^{T_n}\}}||_p\\
&\le ||r_n - r||_p + K\,\left(||f||_{1+\epsilon,\lambda}\,\left(\lambda\left(L(c)^{T_n}\Delta L_n(c)^{T_n}\right)\right)^{\frac{\epsilon}{1+\epsilon}}\right)^{1/p}.
\end{aligned}$$

Since

$$\begin{aligned}
||r\,\mathbf{1}_{\{x\in L(c)^{T_n}\}} - r\,\mathbf{1}_{\{x\in L(c)\}}||_p &= \left(\int_{\mathbb{R}^d}|r(x)\,\mathbf{1}_{\{x\in L(c)\cap[0,T_n]^d\}} - r(x)\,\mathbf{1}_{\{x\in L(c)\}}|^p f(x)dx\right)^{1/p}\\
&= \left(\int_{L(c)\setminus[0,T_n]^d}|r(x)|^p f(x)dx\right)^{1/p},
\end{aligned}$$

we finally get

$$\begin{aligned}
||r_n\,\mathbf{1}_{\{x\in L_n(c)^{T_n}\}} - r\,\mathbf{1}_{\{x\in L(c)\}}||_p &= ||r_n\,\mathbf{1}_{\{x\in L_n(c)^{T_n}\}} - r\,\mathbf{1}_{\{x\in L(c)^{T_n}\}} + r\,\mathbf{1}_{\{x\in L(c)^{T_n}\}} - r\,\mathbf{1}_{\{x\in L(c)\}}||_p\\
&\le ||r_n\,\mathbf{1}_{\{x\in L_n(c)^{T_n}\}} - r\,\mathbf{1}_{\{x\in L(c)^{T_n}\}}||_p + ||r\,\mathbf{1}_{\{x\in L(c)^{T_n}\}} - r\,\mathbf{1}_{\{x\in L(c)\}}||_p\\
&\le ||r_n - r||_p + K\,\left(||f||_{1+\epsilon,\lambda}\,\left(\lambda\left(L(c)^{T_n}\Delta L_n(c)^{T_n}\right)\right)^{\frac{\epsilon}{1+\epsilon}}\right)^{1/p}\\
&+ \left(\int_{L(c)\setminus[0,T_n]^d}|r(x)|^p f(x)dx\right)^{1/p}.
\end{aligned}$$

From Theorem 3.1 and assumptions of Theorem 4.1, the last inequality concludes the proof. □

Proof of Theorem 4.2: Note that the proofs of Theorem 4.2 and Theorem 3.1 are strongly related. We have

$$\begin{aligned} w_n \, ||r_n \, \mathbf{1}_{\{x \in L_n(c)^{T_n}\}} - r \, \mathbf{1}_{\{x \in L(c)^{T_n}\}}||_p &\leq w_n ||r_n \, \mathbf{1}_{\{x \in L_n(c)^{T_n}\}} - r \, \mathbf{1}_{\{x \in L_n(c)^{T_n}\}}||_p \\ &+ w_n \, ||r \, \mathbf{1}_{\{x \in L_n(c)^{T_n}\}} - r \, \mathbf{1}_{\{x \in L(c)^{T_n}\}}||_p \\ &\leq w_n \, ||r_n - r||_p + w_n \, K \, \left(\mu \left(L(c)^{T_n} \Delta L_n(c)^{T_n}\right)\right)^{1/p}. \\ &\leq w_n \, ||r_n - r||_p + w_n \, K \, \left(||f||_{1+\epsilon,\lambda} \, \left(\lambda \left(L(c)^{T_n} \Delta L_n(c)^{T_n}\right)\right)^{\frac{\epsilon}{1+\epsilon}}\right)^{1/p} \end{aligned}$$

and Theorem 3.1 concludes the proof. □

Proof of Theorem 5.1: The proof is a straightforward application of Lemma 10.1 and Lemma 10.2. □

Lemma 10.1. *Under assumptions of Theorem 5.1, we have*

$$p_n^{\frac{\epsilon}{2(1+\epsilon)}} \, | \, \mathbb{E}\left[Y \, | \, \mathbf{X} \in L(c)^{T_n}\right] - \mathbb{E}\left[Y \, | \, \mathbf{X} \in L_n(c)^{T_n}\right] | \xrightarrow[n \to \infty]{\mathbb{P}} 0,$$

with $\epsilon > 0$ such that f is $1+\epsilon$ integrable.

Proof of Lemma 10.1: Using Theorem 3.1, we obtain

$$p_n^{\frac{\epsilon}{2(1+\epsilon)}} \, |P\left[\mathbf{X} \in L(c)^{T_n} \Delta L_n(c)^{T_n}\right]| \leq p_n^{\frac{\epsilon}{2(1+\epsilon)}} \, d_\lambda \left(L(c)^{T_n}, L_n(c)^{T_n}\right)^{\frac{\epsilon}{1+\epsilon}} \, ||f||_{1+\epsilon} \xrightarrow[n \to \infty]{\mathbb{P}} 0. \tag{10.3}$$

Then we get

$$p_n^{\frac{\epsilon}{2(1+\epsilon)}} \left|P\left[\mathbf{X} \in L(c)^{T_n}\right] - P\left[\mathbf{X} \in L_n(c)^{T_n}\right]\right| \xrightarrow[n \to \infty]{\mathbb{P}} 0.$$

Note g the density of the pair $(\mathbf{X}, Y)$. Under assumptions, we also obtain

$$\begin{aligned} & p_n^{\frac{\epsilon}{2(1+\epsilon)}} \left| \int y \mathbf{1}_{\{x \in L(c)^{T_n}\}} g(x,y) \lambda(dxdy) - \int y \mathbf{1}_{\{x \in L_n(c)^{T_n}\}} g(x,y) \lambda(dxdy) \right| \\ \leq \; & p_n^{\frac{\epsilon}{2(1+\epsilon)}} \left| \int y \mathbf{1}_{\{x \in L(c)^{T_n} \Delta L_n(c)^{T_n}\}} g(x,y) \lambda(dxdy) \right| \\ \leq \; & p_n^{\frac{\epsilon}{2(1+\epsilon)}} \left(\int y^2 g(x,y) \lambda(dxdy) \right)^{1/2} \left(\int g(x,y) \mathbf{1}_{\{x \in L(c)^{T_n} \Delta L_n(c)^{T_n}\}} \lambda(dxdy) \right)^{1/2} \\ \leq \; & p_n^{\frac{\epsilon}{2(1+\epsilon)}} E\left[Y^2\right]^{\frac{1}{2}} \left(\int f(x) \mathbf{1}_{\{x \in L(c)^{T_n} \Delta L_n(c)^{T_n}\}} \lambda(dx) \right)^{1/2} \\ \leq \; & p_n^{\frac{\epsilon}{2(1+\epsilon)}} E\left[Y^2\right]^{\frac{1}{2}} \, d_\lambda \left(L(c)^{T_n}, L_n(c)^{T_n}\right)^{\frac{\epsilon}{2(1+\epsilon)}} \, ||f||_{1+\epsilon}^{1/2} \xrightarrow[n \to \infty]{\mathbb{P}} 0. \end{aligned} \tag{10.4}$$

Then

$$
\begin{aligned}
& p_n^{\frac{r}{2(1+r)}} \left|\mathbb{E}[Y|\mathbf{X}\in L(c)^{T_n}]-\mathbb{E}[Y|\mathbf{X}\in L_n(c)^{T_n}]\right| \\
= \; & p_n^{\frac{r}{2(1+r)}} \left|\int y\,\mathbf{1}_{\{x\in L(c)^{T_n}\}} g(x,y)\lambda(\mathrm{d}x\mathrm{d}y)\;\mathbb{P}[\mathbf{X}\in L(c)^{T_n}]^{-1} - \int y\,\mathbf{1}_{\{x\in L_n(c)^{T_n}\}} g(x,y)\lambda(\mathrm{d}x\mathrm{d}y)\;\mathbb{P}[\mathbf{X}\in L_n(c)^{T_n}]^{-1}\right. \\
\leq \; & \frac{p_n^{\frac{r}{2(1+r)}}}{\mathbb{P}[\mathbf{X}\in L(c)^{T_n}]\,\mathbb{P}[\mathbf{X}\in L_n(c)^{T_n}]}\Bigg(\mathbb{P}[\mathbf{X}\in L(c)^{T_n}]\left|\int y\,\mathbf{1}_{\{x\in L(c)^{T_n}\}} g(x,y)\lambda(\mathrm{d}x\mathrm{d}y)\right. \\
& \left.-\int y\,\mathbf{1}_{\{x\in L_n(c)^{T_n}\}} g(x,y)\lambda(\mathrm{d}x\mathrm{d}y)\right| + \int y\,\mathbf{1}_{\{x\in L_n(c)^{T_n}\}} g(x,y)\lambda(\mathrm{d}x\mathrm{d}y)\cdot\left|\mathbb{P}[\mathbf{X}\in L(c)^{T_n}]-\mathbb{P}[\mathbf{X}\in L_n(c)^{T_n}]\right|\Bigg).
\end{aligned}
$$

Using (10.3) and (10.4) we obtain the result. □

Lemma 10.2. *Under assumptions of Theorem 5.1, we have*

$$
d_n\left|\mathbb{E}\left[Y|\mathbf{X}\in L_n(c)^{T_n}\right]-\mathbb{E}_n\left[Y|\mathbf{X}\in L_n(c)^{T_n}\right]\right| \xrightarrow[n\to\infty]{\mathbb{P}} 0
$$

where $d_n = o\left(\sqrt{n}\right)$.

Proof of Lemma 10.2: We have

$$
d_n\left|\mathbb{E}\left[Y|\mathbf{X}\in L_n(c)^{T_n}\right]-\mathbb{E}_n\left[Y|\mathbf{X}\in L_n(c)^{T_n}\right]\right| = d_n\left|\frac{\int y\mathbf{1}_{\{x\in L_n(c)^{T_n}\}}g(x,y)\lambda(dxdy)}{P(\mathbf{X})\in L_n(c)^{T_n}} - \frac{\sum_{i=1}^n Y_i\;\mathbf{1}_{\{\mathbf{X}_i\in L_n(c)^{T_n}\}}}{\sum_{i=1}^n \mathbf{1}_{\{\mathbf{X}_i\in L_n(c)^{T_n}\}}}\right|.
$$

Under assumptions of the Lemma 10.2 and using Theorem 27.2 in Billingsley [5], we obtain that

$$
d_n\left|\mathbb{P}[\mathbf{X}\in L_n(c)^{T_n}] - \frac{1}{n}\sum_{i=1}^n \mathbf{1}_{\{\mathbf{X}_i\in L_n(c)^{T_n}\}}\right| \xrightarrow[n\to\infty]{\mathbb{P}} 0,
$$

and

$$
d_n\left|\int y\mathbf{1}_{\{x\in L_n(c)^{T_n}\}}g(x,y)\lambda(dxdy) - \frac{1}{n}\sum_{i=1}^n Y_i\,\mathbf{1}_{\{\mathbf{X}_i\in L_n(c)^{T_n}\}}\right| \xrightarrow[n\to\infty]{\mathbb{P}} 0.
$$

Hence the result. □

Proof of Theorem 5.2: The proof is a straightforward application of Lemmas 10.1, 10.3 and 10.4. □

Lemma 10.3. *Under assumptions of Theorem 5.2, we have*

$$
c_N\left|\mathbb{E}[Y\mid \boldsymbol{X}\in L_n(c)^{T_n}]-\mathbb{E}[r_N(\boldsymbol{X})\mid \boldsymbol{X}\in L_n(c)^{T_n}]\right| \xrightarrow[N\to\infty]{\mathbb{P}} 0,
$$

with $c_N = o(E|r_N(\boldsymbol{X})-r(\boldsymbol{X})|)$.

Proof of Lemma 10.3 : We have

$$
\begin{aligned}
\left|\mathbb{E}[Y\mid \mathbf{X}\in L_n(c)^{T_n}]-\mathbb{E}[r_N(\mathbf{X})\mid \mathbf{X}\in L_n(c)^{T_n}]\right| &= \frac{\left|\int_{L_n(c)^{T_n}}(r(x)-r_N(x))f(x)d\lambda(x)\right|}{\mathbb{P}[\mathbf{X}\in L_n(c)^{T_n}]} \\
&\leq K\times\int_{L_n(c)^{T_n}}|r(x)-r_N(x)|\,f(x)d\lambda(x) \\
&\leq \int_{\mathbb{R}^d}|r(x)-r_N(x)|\,f(x)d\lambda(x) = E|r_N(\mathbf{X})-r(\mathbf{X})|.
\end{aligned}
$$

Hence the result. □

Lemma 10.4. *Under assumptions of Theorem 5.2, we have*

$$d_n \left|\mathbb{E}\left[r_N(\boldsymbol{X})|\boldsymbol{X}\in L_n(c)^{T_n}\right]-\mathbb{E}_n\left[r_N(\boldsymbol{X})|\boldsymbol{X}\in L_n(c)^{T_n}\right]\right| \xrightarrow{\mathbb{P}}_{n\to\infty} 0$$

where $d_n = o\left(\sqrt{n}\right)$.

Proof of Lemma 10.4 : We have

$$d_n \left|\mathbb{E}\left[r_N(\mathbf{X})|L_n(c)^{T_n}\right]-\mathbb{E}_n\left[r_N(\mathbf{X})|L_n(c)^{T_n}\right]\right| = d_n \left|\frac{\int_{L_n(c)^{T_n}} r_N(x)f(x)\lambda(dx)}{\mathbb{P}[\mathbf{X}\in L_n(c)^{T_n}]}-\frac{\sum_{i=1}^n r_N(X_i)\mathbf{1}_{\{\mathbf{X}_i\in L_n(c)^{T_n}\}}}{\sum_{i=1}^n \mathbf{1}_{\{\mathbf{X}_i\in L_n(c)^{T_n}\}}}\right|.$$

Under assumptions of the Lemma and using Theorem 27.2 in Billingsley [5], we obtain that

$$d_n \left|\mathbb{P}[\mathbf{X}\in L_n(c)^{T_n}]-\frac{1}{n}\sum_{i=1}^n \mathbf{1}_{\{\mathbf{X}_i\in L_n(c)^{T_n}\}}\right| \xrightarrow{\mathbb{P}}_{n\to\infty} 0,$$

and

$$d_n \left|\int_{L_n(c)^{T_n}} r_N(x)f(x)\lambda(dx)-\frac{1}{n}\sum_{i=1}^n r_N(\mathbf{X}_i)\mathbf{1}_{\{\mathbf{X}_i\in L_n(c)^{T_n}\}}\right| \xrightarrow{\mathbb{P}}_{n\to\infty} 0$$

which gives us the Lemma. □

Proof of Lemma 6.1: First, we remark that

$$F_{a\mathbf{X}}(x) = F_{\mathbf{X}}\left(\frac{x}{a}\right).$$

Then, we obtain

$$\begin{aligned} m_a^{\triangledown} &= \inf_{x\in a\,E}\left\|\left(\frac{\partial}{\partial x_1}F_{\mathbf{X}}\left(\frac{x}{a}\right),\ldots,\frac{\partial}{\partial x_d}F_{\mathbf{X}}\left(\frac{x}{a}\right)\right)\right\| = \inf_{x\in a\,E}\left\|\frac{1}{a}\left(\frac{\partial F_{\mathbf{X}}}{\partial x_1}\left(\frac{x}{a}\right),\ldots,\frac{\partial F_{\mathbf{X}}}{\partial x_d}\left(\frac{x}{a}\right)\right)\right\| \\ &= \frac{1}{a}\inf_{x\in E}\left\|\left(\frac{\partial F_{\mathbf{X}}}{\partial x_1}(x),\ldots,\frac{\partial F_{\mathbf{X}}}{\partial x_d}(x)\right)\right\| = \frac{1}{a}\,m^{\triangledown}. \end{aligned}$$

Second part of Lemma 6.1 comes down from trivial calculus. □

Proof of Theorem 6.1 : We have

$$F_{a\mathbf{X}}(x) = F_{\mathbf{X}}\left(\frac{x}{a}\right),\ f_{a\mathbf{X}}(x) = \frac{1}{a}f_{\mathbf{X}}\left(\frac{x}{a}\right) \text{ and } r_{a\mathbf{X},bY}(x) = b\,r_{\mathbf{X},Y}\left(\frac{x}{a}\right).$$

Proof of 1. in Theorem 6.1: We have

$$\begin{aligned} &||r_{n,a\mathbf{X},bY}\,\mathbf{1}_{\{x\in L_{a,n}(c)^{T_n}\}} - r_{a\mathbf{X},bY}\,\mathbf{1}_{\{x\in L_a(c)^{T_n}\}}||_p \\ &= \left(\int_{\mathbb{R}^d}\left|r_{n,a\mathbf{X},bY}(x)\mathbf{1}_{\{x\in L_{a,n}(c)^{T_n}\}} - r_{a\mathbf{X},bY}(x)\mathbf{1}_{\{x\in L_a(c)^{T_n}\}}\right|^p f_{a\mathbf{X}}(x)dx\right)^{1/p} \\ &= \left(\int_{\mathbb{R}^d} b^p\left|r_{n,\mathbf{X},Y}\left(\frac{x}{a}\right)\mathbf{1}_{\{x/a\in L_n(c)^{T_n}\}} - r_{\mathbf{X},Y}\left(\frac{x}{a}\right)\mathbf{1}_{\{x/a\in L(c)^{T_n}\}}\right|^p \frac{1}{a}f_{\mathbf{X}}\left(\frac{x}{a}\right)dx\right)^{1/p} \end{aligned}$$

and taking $t = x/a$ we obtain

$$\begin{aligned} & ||r_{n,a\mathbf{X},bY}\,\mathbf{1}_{\{x\in L_{a,n}(c)^{T_n}\}} - r_{a\mathbf{X},bY}\,\mathbf{1}_{\{x\in L_a(c)^{T_n}\}}||_p \\ = & \left(\int_{\mathbb{R}^d} b^p \left|r_{n,\mathbf{X},Y}(t)\,\mathbf{1}_{\{t\in L_n(c)^{T_n}\}} - r_{\mathbf{X},Y}(t)\,\mathbf{1}_{\{t\in L(c)^{T_n}\}}\right|^p f_{\mathbf{X}}(t)\,dt\right)^{1/p} \\ = & \; b||r_{n,\mathbf{X},Y}\,\mathbf{1}_{\{x\in L_n(c)^{T_n}\}} - r_{\mathbf{X},Y}\,\mathbf{1}_{\{x\in L(c)^{T_n}\}}||_p \end{aligned}$$

Hence the result. □

Proof of 2. in Theorem 6.1: We have

$$\left|\widehat{\mathrm{CCTE}}_{c,n}^{T_n}(a\mathbf{X}, bY) - \mathrm{CCTE}_c^{T_n}(a\mathbf{X}, bY)\right| = \left|\mathbb{E}_n\left[bY|a\mathbf{X}\in L_{n,a}(c)^{T_n}\right] - \mathbb{E}\left[bY|a\mathbf{X}\in L_a(c)^{T_n}\right]\right|.$$

Using $L_a(c)^{T_n} = aL(c)^{T_n}$ and the assumptions, we obtain

$$\left|\widehat{\mathrm{CCTE}}_{c,n}^{T_n}(a\mathbf{X}, bY) - \mathrm{CCTE}_c^{T_n}(a\mathbf{X}, bY)\right| = \frac{1}{b}\left|\mathbb{E}_n\left[Y|a\mathbf{X}\in aL_n(c)^{T_n}\right] - \mathbb{E}\left[Y|a\mathbf{X}\in aL(c)^{T_n}\right]\right|.$$

Hence the result. □

Acknowledgements: This work has been partially supported by the French research national agency (ANR) under the references ANR-08BLAN-0314-01 and ANR 2011 BS01 010 01 projet Calibration. The authors thank Yannick Baraud, Roland Diel, Christine Tuleau-Malo and Patricia Reynaud-Bouret for fruitful discussions.

Bibliography

[1] Artzner, P., Delbaen, F., Eber, J.-M., and Heath, D. (1999). Coherent measures of risk. *Math. Finance*, 9(3):203–228.

[2] Baíllo, A., Cuesta-Albertos, J., and Cuevas, A. (2001). Convergence rates in nonparametric estimation of level sets. *Statistics and Probability Letters*, 53:27–35.

[3] Belzunce, F., Castaño, A., Olvera-Cervantes, A., and Suárez-Llorens, A. (2007). Quantile curves and dependence structure for bivariate distributions. *Computational Statistics & Data Analysis*, 51(10):5112–5129.

[4] Bentahar, I. (2006). Tail conditional expectation for vector-valued risks. SFB 649 Discussion Papers SFB649DP2006-029, Humboldt University, Collaborative Research Center 649.

[5] Billingsley, P. (1995). *Probability and measure*. Wiley Series in Probability and Mathematical Statistics. John Wiley & Sons Inc., New York, third edition. A Wiley-Interscience Publication.

[6] Cai, J. and Li, H. (2005). Conditional tail expectations for multivariate phase-type distributions. *Journal of Applied Probability*, 42(3):810–825.

[7] Chebana, F. and Ouarda, T. (2011). Multivariate quantiles in hydrological frequency analysis. *Environmetrics*, 22(1):63–78.

[8] Cousin, A. and Di Bernardino, E. (2012). On multivariate extensions of Value-at-Risk. *submitted*.

[9] Cuevas, A., González-Manteiga, W., and Rodríguez-Casal, A. (2006). Plug-in estimation of general level sets. *Australian & New Zealand Journal of Statistics*, 48(1):7–19.

[10] Daouia, A., Gardes, L., Girard, S., and Lekina, A. (2010). Kernel estimators of extreme level curves. *Test*, (20):311–333.

[11] de Haan, L. and Huang, X. (1995). Large quantile estimation in a multivariate setting. *J. Multivariate Anal.*, 53(2):247–263.

[12] Dedu, S. and Ciumara, R. (2010). Restricted optimal retention in stop-loss reinsurance under VaR and CTE risk measures. *Proc. Rom. Acad. Ser. A Math. Phys. Tech. Sci. Inf. Sci.*, 11(3):213–217.

[13] Dekkers, A. L. M. and de Haan, L. (1989). On the estimation of the extreme-value index and large quantile estimation. *Ann. Statist.*, 17(4):1795–1832.

[14] Denuit, M., Dhaene, J., Goovaerts, M., and Kaas, R. (2005). *Actuarial Theory for Dependent Risks*. Wiley.

[15] Di Bernardino, E., Laloë, T., Maume-Deschamps, V., and Prieur, C. (2013). Plug-in estimation of level sets in a non-compact setting with applications in multivariable risk theory. *ESAIM: Probability and Statistics*, (17):236–256.

[16] Draisma, G., Drees, H., Ferreira, A., and de Haan, L. (2004). Bivariate tail estimation: dependence in asymptotic independence. *Bernoulli*, 10(2):251–280.

[17] Ekeland, I., Galichon, A., and Henry, M. (2012). Comonotonic measures of multivariate risks. Technical report.

[18] Embrechts, P., Klüppelberg, C., and Mikosch, T. (1997). *Modelling Extremal Events: For Insurance and Finance*. Applications of mathematics, 33. Springer.

[19] Embrechts, P. and Puccetti, G. (2006). Bounds for functions of multivariate risks. *Journal of Multivariate Analysis*, 97(2):526–547.

[20] Hawkes, J., Gouldby, B. P., Tawn, J. A., and Owen, M. W. (2002). The joint probability ofwaves andwater levels in coastal engineering design. *Journal of Hydraulic Research*, 40(3):241–251.

[21] Imlahi, L., Ezzerg, M., and Chakak, A. (1999). Estimación de la curva mediana de una cópula $C(x_1, \ldots, x_n)$. *Rev. R. Acad. Cien. Exact. Fis. Nat*, 93(2):241–250.

[22] Jouini, E., Meddeb, M., and Touzi, N. (2004). Vector-valued coherent risk measures. *Finance Stoch.*, 8(4):531–552.

[23] Kohler, M., Krzyżak, A., and Walk, H. (2009). Optimal global rates of convergence for nonparametric regression with unbounded data. *J. Statist. Plann. Inference*, 139(4):1286–1296.

[24] Nappo, G. and Spizzichino, F. (2009). Kendall distributions and level sets in bivariate exchangeable survival models. *Information Sciences*, 179:2878–2890.

[25] Nelsen, R. B. (1999). *An introduction to copulas*, volume 139 of *Lecture Notes in Statistics*. Springer-Verlag, New York.

[26] Rigollet, P. and Vert, R. (2009). Optimal rates for plug-in estimators of density level sets. *Bernoulli. Official Journal of the Bernoulli Society for Mathematical Statistics and Probability*, 15(4):1154–1178.

[27] Salvadori, G., De Michele, C., and Durante, F. (2011). On the return period and design in a multivariate framework. *Hydrology and Earth System Sciences*, 15(11):3293–3305.

[28] Tau, V. and Dam, G. C. (2011). Preliminary design study. *Project group Flood Defence, Institute for Water and Environment Research*, pages 1–144.

[29] Tibiletti, L. (1993). On a new notion of multidimensional quantile. *Metron. International Journal of Statistics*, 51(3-4):77–83.

Elena Di Bernardino: CNAM, 292 rue Saint-Martin, 75141 Paris Cedex 03, France

E-mail: (elena.di_bernardino@cnam.fr)

Thomas Laloë (Corresponding Author): Université de Nice Sophia-Antipolis, CNRS, LJAD, UMR 7351, 06100 Nice, France.

E-mail: (laloe@unice.fr), Phone : +33 4 92 07 60 20

Rémi Servien: INRA, Universite de Toulouse, UMR1331 Toxalim, Research Centre in Food Toxicology, F-31027 Toulouse

E-mail: (remi.servien@toulouse.inra.fr)