HAL
open science

Modélisation de la dépendance et mesures de risque multidimensionnelles

Éléna Di Bernardino

To cite this version:

Éléna Di Bernardino. Modélisation de la dépendance et mesures de risque multidimensionnelles. Mathématiques générales [math.GM]. Université Claude Bernard - Lyon I, 2011. Français. NNT : 2011LYO10259 . tel-00838598

HAL Id: tel-00838598

https://theses.hal.science/tel-00838598

Submitted on 26 Jun 2013

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.

I.S.F.A. **École Doctorale Sciences Économiques et de Gestion**

Modélisation de la dépendance et mesures de risque multidimensionnelles

THÈSE

présentée et soutenue publiquement le 8 décembre 2011

pour l'obtention du

Doctorat de l'Université Claude Bernard Lyon I

Spécialité Mathématiques Appliquées

par

Elena Di Bernardino

Composition du jury

Rapporteurs : Wenceslao González-Manteiga, Professeur à l'Université de Saint Jacques de Compostelle, Galice, Espagne
Johan Segers, Professeur à l'Université catholique de Louvain

Examinateurs : Gérard Biau, Professeur à l'Université Paris VI
Anne-Laure Fougères, Professeur à l'Université Lyon 1
Claude Lefèvre, Professeur à l'Université Libre de Bruxelles
Véronique Maume-Deschamps, Professeur à l'Université Lyon 1 (directrice de thèse)
Clémentine Prieur, Professeur à l'Université Grenoble 1 (directrice de thèse)

Laboratoire Science Actuarielle Financière — EA 2429

`Imagination` n.f. Faculté qu'a l'être humain d'élaborer des images et des conceptions nouvelles, de trouver des solutions originales à des problèmes. Faculté d'inventer, de créer, de concevoir. *(Larousse Dictionnaire)*

`Imagination` n. The ability to form mental pictures, new ideas, images or concepts. It is a creative quality that a person has. *(Oxford Dictionary)*

`Imaginación` s.f. Facultad de la mente de representar las imágenes de las cosas reales o ideales. Facilidad para formular nuevas ideas o crear nuevos proyectos. *(Diccionario Clave)*

`Immaginazione` n.f. Facoltà dell'uomo di pensare senza regole fisse e di associare liberamente i dati dell'esperienza sensibile. Creatività libera. *(Vocabolario Zingarelli)*

Remerciements

Je tiens à remercier tout d'abord mes deux directrices de thèse Véronique Maume-Deschamps et Clémentine Prieur pour tout ce qu'elles m'ont apporté au cours de ces trois années : pour leur aide scientifique et pour leur soutien humain, pour leurs encouragements constants, et pour les nombreuses discussions que nous avons pu avoir.

Je suis très honorée que Wenceslao González-Manteiga et Johan Segers aient accepté d'être rapporteurs de ma thèse. Merci à eux pour le temps consacré, je leur en suis extrêmement reconnaissante.

Je tiens à remercier chaleureusement Gérard Biau, Anne-Laure Fougères et Claude Lefèvre, de m'avoir fait l'honneur d'accepter de faire partie du jury de cette thèse.

Merci aussi à Arthur Charpentier, Stéphane Girard, Armelle Guillou, Anne-Laure Fougères, Philippe Soulier et Stéphane Loisel pour avoir pris le temps de m'aider, de répondre à mes questions, y compris les plus naïves, et de m'avoir très souvent orienté vers des articles que je n'avais pas vu passer. Many thanks to Laurent de Haan, John P. Nolan, Ragnar Norberg, Richard Smith, Mario Wüthrich, Dominik Kortschak for help or fruitful discussions during these three years. Grazie a Luisa Tibiletti per i suoi suggerimenti sui nostri temi di ricerca comuni. Gracias a José R. León y Alberto Rodríguez-Casal para su ayuda y sus consejos.

Je souhaitais aussi remercier également mes co-auteurs. Merci à Thomas Laloë pour le séjour à Nice, le voyage ensemble en Tunisie, son temps et son aide précieuse pendant nos travaux ensemble mais aussi pendant le mois d'Aout! Merci à Areski Cousin pour avoir travaillé avec moi en s'adaptant à mon rythme : je t'en suis très reconnaissante, Ares!

Un grand merci à toutes les personnes qui composent le Laboratoire de Sciences Actuarielle et Financière (SAF) et l'ISFA en général. Un merci à Beatrice Rey-Fournier, Pierre Ribereau, Didier Rullière, Diana Dorobantu, Alexis Bienvenüe, Stéphane Loisel, Anne Eyraud-Loisel, Denis Clot, Christian Robert, Nicolas Leboisne, Corina Constantinescu. Merci à Jean Claude Augros de m'avoir accueillie à l'ISFA. Un grazie enorme a Esterina Masiello per le sue prime parole in italiano al mio arrivo in Francia nel Gennaio del 2009 (e per tutte quelle che sono seguite in tutte e due le lingue!). Une mention spéciale à Romain Biard et Samy Brias pour leur aventure à Rome et pour tout ce que nous avons partagé au cours de ces trois années.

Je ne crois pas qu'un simple merci suffirait pour remercier Abdou et Manel. Le premier, un vrai ami qui m'a accompagné durant toute ma thèse et dont les conseils ont toujours été d'une grande aide. Merci pour cette amitié, Rachiad! La deuxième, une "personne-cadeau-du-ciel" avec laquelle j'ai partagé chaque jour de ces trois ans. Elle connaît tout ce qui est dans cette thèse, mais (encore plus important) tout ce qui a gravité autour de cette thèse tout au long de ces années!

Merci à tous les doctorants de notre laboratoire. Merci au plot 1 de l'administration pour leur disponibilité.

Un grand merci à ma Prof. Geneviève pour chaque mot qu'elle m'a appris dans cette langue. Merci pour sa patience et sa constance. Merci pour tout ce que j'ai pu apprendre de vous, Geneviève, et qui va bien au-delà d'une langue ou d'une grammaire.

Un grazie immenso a Laura per ogni singolo pixel condiviso. Un grazie a Susanna, Chiara, Sara e Federica per essere rimaste sempre qui, al mio fianco. Ancora una volta so che voi cinque siete le mie compagne di strada. Un grazie a Gaetano (anche se i grazie sono solo per gli sconosciuti), a Chià per i nostri "quasi ogni giorno".

Un grazie a tutti miei amici che sono partiti in aereo, treno, macchina e quant'altro per venirmi a trovare a Lione, rendendo la mia casa un piccolo albergo, praticamente in ogni weekend di questi tre anni. Non ci sono parole per dire come questo sia stato importante.

Un grand merci à Flavie, Ombeline, Marie-lorraine et Marion pour leur soutien et leurs encouragements, à Carlotta per tutti i nostri "n'importe quoi" (e per quelli che seguiranno à Paris), à Coralie pour nos "trois-soirées-dans-une". A Mathilde et son "mais bien sûr que tu peux venir dormir chez nous à Grenoble, il ne faut même pas le demander", chaque fois que j'en ai eu besoin, lors de mes cours et séminaires à Grenoble. Merci à Stéphanie pour les mots à travers le rideau rouge, pour son accueil chaleureux et pour son aide. Merci à Madame Geneviève et Madame Peyron pour leur gentillesse.

Et encore merci (en ordre dispersé) à : Mouna, Luci, Oscar, Richi, Clod, Mari, Elenina, Mimmo, Eu, Fede, Giorgio, Giuaglia, Pulla, Mariagrazia.

Merci à Guillaume pour le "Kitch", le "rock and roll" et pour chaque "Tu ne le connais pas?". Merci de m'avoir montré que les probabilités rares existent, tout simplement.

Un grazie a Alberto Barchielli, Maria Assunta Pozio, Peter Laurence, Giovanna Nappo, Roberto Lucchetti e Viola Schiaffonati per i consigli (diretti o indiretti), per la stima e l'amicizia dimostratami in questi anni.

Un grazie ai miei quattro nonni, alle mie zie (Grazia, Alberta e Paola), a Benedetta, a zio Enzo. Un grazie per le valigie sempre colme. Un grazie a Battista per ogni cena da noi.

Un grazie ad Irene, Enza e Dino, senza i quali questa tesi non avrebbe senso.

«... Così questo paese ho creduto per molto tempo che fosse tutto il mondo.
Adesso che il mondo l'ho visto davvero e so che è fatto di tanti piccoli paesi,
non so se da ragazzo mi sbagliavo poi di molto.
Uno gira per mare e per terra,
come i giovanotti dei miei tempi andavano sulle feste dei paesi intorno [...]
Che cosa vuol dire?
Vuol dire che un paese ci vuole, non fosse che per il gusto di andarsene via.
Un paese vuol dire non essere soli, sapere che nella gente, nelle piante,
nella terra c'è qualcosa di tuo, che anche quando non ci sei resta ad aspettarti.»

Cesare Pavese, La luna e i falò (1950)

Modélisation de la dépendance et mesures de risque multidimensionnelles

Résumé

Cette thèse a pour but le développement de certains aspects de la modélisation de la dépendance dans la gestion des risques en dimension plus grande que un. Le premier chapitre de la thèse est constitué d'une introduction générale avec une présentation des principaux résultats obtenus.

Le deuxième chapitre est essentiellement constitué d'un article, coécrit avec Véronique Maume-Deschamps et Clémentine Prieur, s'intitulant *"Estimating Bivariate Tail : a copula based approach"*, actuellement soumis pour publication. Il concerne la construction d'un estimateur de la queue d'une distribution bivariée, capable de capturer les informations de la dépendance entre les deux variables aléatoires, dans le cas de dépendance asymptotique comme dans le cas d'indépendance asymptotique. La construction de cet estimateur se fonde essentiellement sur une méthode bidimensionnelle de dépassement de seuil (Peaks Over Threshold method) et donc sur une version bivariée du Théorème de Pickands-Balkema-de Haan. La modélisation de la dépendance est obtenue via la Upper Tail Dependence Copula (UTDC). Nous démontrons des propriétés de convergence pour l'estimateur ainsi construit. Nous présentons également les résultats obtenus sur plusieurs jeux de données réelles.

Le troisième chapitre repose sur un article, coécrit avec Areski Cousin : *"A multivariate extension of Value-at-Risk and Conditional-Tail-Expectation"*, soumis pour publication. Nous abordons le problème de l'extension de mesures de risque classiques, comme la *Value-at-Risk* (VaR) et la *Conditional-Tail-Expectation* (CTE), dans un cadre multidimensionnel en utilisant la fonction de Kendall multivariée. Nous nous intéressons à l'analyse des propriétés de ces mesures. En particulier nous étudions les propriétés souhaitables pour une mesure de risque comme la monotonie par rapport au niveau de risque, l'homogénéité, l'invariance par translation, le chargement de sécurité. De plus nous examinons comment ces mesures se comportent par rapport à la structure de dépendance et au changement des risques marginaux. Des formules exactes pour notre CTE et VaR sont obtenues pour certaines familles paramétriques de copules. Dans ces cas le comportement de la mesure de risque face au changement du type de dépendance et du niveau de risque est étudié analytiquement.

Enfin, dans le quatrième chapitre de la thèse, nous proposons un estimateur des courbes de niveau d'une fonction de répartition bivariée avec une méthode *plug-in*. Grâce à cet estimateur des ensembles de niveau, nous estimons ensuite une version bidimensionnelle de la CTE précédemment introduite. Nous démontrons des propriétés de convergence pour les estimateurs ainsi construits. Plusieurs exemples avec des données réelles et simulées sont proposés. Ce chapitre de la thèse est lui aussi constitué d'un article, coécrit avec Thomas Laloë, Véronique Maume-Deschamps et Clémentine Prieur, s'intitulant

"Plug-in estimation of level sets in a non-compact setting with applications in multivariate risk theory", accepté pour publication dans la revue *ESAIM : Probability and Statistics*.

AMSC : 60 E 05, 60 G 70, 62 G 05, 62 G 20, 62 H 12, 62 H 05, 91 B 30.

Mots-clés : Théorie des valeurs extrêmes, Probabilités de dépassements, Courbes de niveau pour fonctions de répartition, Théorie et gestion des risques, Copules et dépendance, Ordres stochastiques.

Thèse réalisée au sein du projet AST&Risk financée par l'ANR ;
ANR-08BLAN-0314-01.
Laboratoire Science Actuarielle Financière - Ecole ISFA - 50, avenue Tony Garnier - 69007 Lyon - France.

Dependence modeling and multidimensional risk measures

Abstract

In this PhD thesis we consider different aspects of dependence modeling with applications in multivariate risk theory. The first chapter is constituted by a general introduction.

The second chapter is essentially constituted by the article *"Estimating Bivariate Tail: a copula based approach"*, written with Véronique Maume-Deschamps and Clémentine Prieur, actually submitted for publication. It deals with the problem of estimating the tail of a bivariate distribution function in the case of asymptotic independence as well as in the case of asymptotic dependence. To this end we develop a general extension of the POT (Peaks-Over-Threshold) method, mainly based on a two-dimensional version of the Pickands-Balkema-de Haan Theorem. The dependence structure between the marginals in the upper tails is described by the Upper Tail Dependence Copula (UTDC). Then we construct a two-dimensional tail estimator and study its asymptotic properties. We also present the results on several real data sets.

The third chapter of this thesis is based on the article *"A multivariate extension of Value-at-Risk and Conditional-Tail-Expectation"*, written with Areski Cousin and submitted for publication. We propose a multivariate generalization of risk measures as *Value-at-Risk* (VaR) and *Conditional-Tail-Expectation* (CTE) and we analyze the behavior of these measures in terms of classical properties of risk measures: monotonicity, translation invariance, homogeneity, safety loading. We study the behavior of these measures with respect to different risk scenarios and stochastic ordering of marginals risks. We provide explicit formulas for the CTE and VaR for some parametric copula families. In these cases we analyze the behavior when the underlying dependence structure and the level of risk vary.

Finally in the fourth chapter we introduce a consistent procedure to estimate level sets of an unknown bivariate distribution function, using a *plug-in* approach in a non-compact setting. Furthermore we estimate a bivariate version of the *Conditional-Tail-Expectation* introduced above. Examples and illustrations with simulated and real data are presented. Also this chapter is constituted by the article *"Plug-in estimation of level sets in a non-compact setting with applications in multivariate risk theory"*, written with Thomas Laloë, Véronique Maume-Deschamps and Clémentine Prieur, accepted for publication in *ESAIM: Probability and Statistics* journal.

AMSC: 60 E 05, 60 G 70, 62 G 05, 62 G 20, 62 H 12, 62 H 05, 91 B 30.

Key Words: Extreme Value Theory, Excess probabilities, Level sets of distribution functions, Risk Theory, Copulae and dependence, Stochastic orders.

PhD thesis realized in the project AST&Risk and partially supported by the French research national agency (ANR) under the reference ANR-08BLAN-0314-01.

Laboratoire Science Actuarielle Financière - Ecole ISFA - 50, avenue Tony Garnier - 69007 Lyon - France.

Modelli per la dipendenza e misure di rischio multidimensionali

Sommario

In questo lavoro di tesi di dottorato ci interessiamo dello sviluppo e della creazione di modelli per la gestione della dipendenza nella Teoria del Rischio in dimensione più grande di uno. Il primo capitolo di questa tesi è composto da un'introduzione generale in cui vengono esposti in modo sintetico i principali risultati ottenuti.

Il secondo capitolo é essenzialmente costituito dall' articolo *"Estimating Bivariate Tail: a copula based approach"*, scritto con Véronique Maume-Deschamps e Clémentine Prieur, attualmente sottomesso per la pubblicazione. In questa parte della tesi ci interessiamo alla costruzione di uno stimatore per la coda della distribuzione di probabilità bidimensionale, capace di catturare le informazioni della dipendenza tra le due variabili aleatorie nel caso di dipendenza asintotica così come nel caso di indipendenza asintotica. Questo stimatore si basa essenzialmente su una versione bidimensionale del metodo di superamento di una soglia (Peaks Over Threshold method) e quindi su un'estensione del Teorema di Pickands-Balkema-de Haan in dimensione due. La struttura di dipendenza bidimensionale è modellata grazie alla Upper Tail Dependence Copula (UTDC). Si dimostrano risultati di consistenza asintotica per lo stimatore così costruito. In conclusione di questo capitolo forniamo alcune illustrazioni dei nostri risultati teorici utilizzando dati reali.

Il terzo capitolo si basa sull' articolo *"A multivariate extension of Value-at-Risk and Conditional-Tail-Expectation"*, scritto con Areski Cousin e attualmente sottomesso per la pubblicazione. In questo capitolo affrontiamo il problema dell'estensione di alcune misure di rischio classiche, come la *Value-at-Risk* (VaR) e la *Conditional-Tail-Expectation* (CTE), in un contesto multidimensionale, utilizzando le curve di livello di una funzione di distribuzione per un certo livello di rischio dato. Sono studiate le proprietà di queste misure di rischio in termini di omogeneità, invarianza per traslazione, monotonia rispetto al livello di rischio. Inoltre esaminiamo il comportamento di queste misure rispetto alla struttura di dipendenza e al cambiamento dei rischi marginali. Forniamo delle formule chiuse per la nostra CTE nel caso di alcune famiglie parametriche di copule. In questi casi analizziamo il comportamento della misura di rischio sia in funzione del tipo di dipendenza che del livello del rischio.

Nella quarto capitolo della tesi proponiamo una procedura consistente per stimare le curve di livello di una funzione di distribuzione bidimensionale attraverso un metodo *plug-in*. Grazie a ciò siamo in grado di fornire uno stimatore consistente anche per la versione bidimensionale della CTE introdotta nel capitolo precedente. Dimostriamo risultati di consistenza. Presentiamo infine vari studi e illustrazioni delle proprietà di consistenza di questi stimatori utilizzando sia dati simulati che dati reali. Anche questo capitolo è costituito da un articolo dal titolo *"Plug-in estimation of level sets in a non-compact setting with applications in multivariate risk theory"*, scritto con Thomas Laloë, Véronique Maume-Deschamps e Clémentine Prieur, accettato per la pubblicazione nella rivista *ESAIM: Probability and Statistics*.

AMSC: 60 E 05, 60 G 70, 62 G 05, 62 G 20, 62 H 12, 62 H 05, 91 B 30.

Parole chiave: Teoria dei valori estremi, Curve di livello di funzioni di distribuzione, Teoria e gestione del rischio, Teoria delle copule e studio della dipendenza, Ordini stocastici.

Tesi di dottorato realizzata all'interno del progetto AST&Risk e parzialmente finanziata dall'agenzia nazionale di ricerca francese (ANR) con codice di riferimento ANR-08BLAN-0314-01.

Laboratoire Science Actuarielle Financière - Ecole ISFA - 50, avenue Tony Garnier - 69007 Lyon - France.

Table des matières

Chapitre 1
Introduction générale

1.1 Présentation de la thèse 1
1.2 Description par chapitre des principaux résultats 3
1.2.1 Estimation de la queue d'une distribution bivariée : approche reposant sur les copules 3
1.2.2 Quelques généralisations de mesures de risque dans un cadre multivarié : propriétés par rapport au niveau de risque et à la structure de dépendance 18
1.2.3 Estimation de type plug-in des ensembles de niveau dans un cadre non-compact avec applications à la théorie des risques 27

Chapitre 2
Estimation de la queue d'une distribution bivariée : approche reposant sur les copules

Résumé 32
2.1 Introduction 34
2.2 On the two-dimensional Pickands-Balkema-de Haan Theorem 36
2.3 Estimating dependence structure in the bivariate framework 38
2.3.1 Asymptotic dependent case 39
2.3.2 Asymptotic independent case 41
2.4 Estimating tail distributions 43
2.4.1 Estimating the tail of univariate distributions 43
2.4.2 Estimating the tail of bivariate distributions 43
2.5 Convergence results in the univariate case 45
2.6 Convergence results in the bivariate case 47

2.6.1 Asymptotic dependent case 48
2.6.2 Asymptotic independent case 49
2.7 Illustrations with real data 50
Bibliography 58

Chapitre 3
Quelques généralisations des mesures de risque dans un cadre multivarié

Résumé 64
3.1 Basic notions and preliminaries 68
3.2 Multivariate *Value-at-Risk* 72
3.2.1 Invariance properties 76
3.2.2 Comparison of univariate and multivariate VaR 76
3.2.3 Behavior of the multivariate VaR with respect to marginal distributions 78
3.2.4 Behavior of multivariate VaR with respect to the dependence structure 78
3.2.5 Behavior of multivariate VaR with respect to risk level 80
3.3 Multivariate *Conditional-Tail-Expectation* 81
3.3.1 Invariance properties 82
3.3.2 Comparison with VaR and univariate CTE 84
3.3.3 Behavior of multivariate CTE with respect to marginal distributions 85
3.3.4 Behavior of multivariate CTE with respect to risk level 87
Bibliography 90

Chapitre 4
Estimation des ensembles de niveau et application à la théorie des risques

Résumé 96
4.1 Notation and preliminaries 100
4.2 Consistency in terms of the Hausdorff distance 101
4.3 L_1 consistency 103
4.4 Application in bivariate risk theory 104
4.5 Illustrations 108
4.5.1 Estimation of the level sets 108

4.5.2 Estimation of bivariate CTE 111

4.6 Conclusion 116

4.7 Proofs 116

Bibliography 122

Conclusion et perspectives

Annexe A
Rappels et définitions utiles

A.1 Copules 128

A.2 Fonction de répartition Burr 129

Annexe B
Simulations, illustrations et approfondissements

B.1 Lien entre les paramètres de dépendance η et θ 132

B.2 Estimation du paramètre θ, simulations 135

B.3 Estimateurs dans le cadre unidimensionnel 137

B.4 Construction de Skorohod 141

B.5 Formules explicites de la CTE bivariée pour quelques familles de copules 142

B.6 Quelques comparaisons de la CTE bivariée avec différentes mesures de risque bivariées 145

B.7 Sur la croissance de la CTE bivariée : deux cas discrets 147

B.8 Quelques formules paramétriques exactes pour une courbe quantile . . . 148

B.9 Quelques études de la CTE bivariée pour des hauts niveaux de risque . . 150

Bibliographie

Chapitre 1

Introduction générale

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

Cette thèse porte sur l'étude de certains aspects de la modélisation de la dépendance dans la gestion des risques dans un cadre multidimensionnel. Les résultats obtenus dans cette thèse s'inspirent des récents développements de la théorie des valeurs extrêmes (voir par exemple Coles, 2001), de la théorie de la gestion des risques (voir Grandell, 1991 ; Denuit *et al.*, 2005) ou encore de la théorie des copules (voir Nelsen, 1999). Tout au long de cette thèse nous avons tenté d'utiliser des outils provenant de ces différentes branches de la littérature existante, afin de proposer des modèles plus souples. Dans cette introduction, nous présentons les travaux réalisés durant cette thèse et nous les situons par rapport à la littérature déjà existante.

1.1 Présentation de la thèse

Le présent travail consiste essentiellement à étudier les liens entre la modélisation de la dépendance dans un cadre multidimensionnel et la gestion des risques grâce au développement de mesures de risques multivariées. Cette introduction générale constitue le premier chapitre de la thèse. La suite du manuscrit est organisée de la façon suivante :

Le deuxième chapitre, intitulé **Estimation de la queue d'une distribution bivariée : approche reposant sur les copules**, est essentiellement constitué d'un article, coécrit avec Véronique Maume-Deschamps et Clémentine Prieur, s'intitulant *"Estimating Bivariate Tail : a copula based approach"*, actuellement soumis pour publication. Ce chapitre est consacré à la construction et à l'étude d'un estimateur de la queue d'une distribution bivariée. Les cas de dépendance asymptotique et d'indépendance asymptotique sont étudiés. Ce travail s'inscrit dans la littérature existante sur la méthode de dépassement de seuil (Peaks Over Threshold method, voir par exemple McNeil, 1997, 1999 dans un cadre unidimensionnel, Rootzén et Tajvidi, 2006 dans un cadre multidimensionnel). Une extension bidimensionnelle est développée. Notre méthode se fonde sur une extension bivariée du Théorème de Pickands-Balkema-de Haan (Balkema et de

Haan, 1974 ; Pickands, 1975). Dans notre procédure d'estimation, la modélisation de la dépendance dans la queue de la distribution est obtenue via la Upper Tail Dependence Copula (UTDC). Nous démontrons des propriétés de convergence pour l'estimateur ainsi construit, en fournissant des vitesses de convergence de l'erreur absolue entre l'estimateur et la fonction de répartition bidimensionnelle dans les cas de la dépendance et de l'indépendance asymptotiques.

Le troisième chapitre, intitulé **Quelques généralisations de mesures de risque dans le cadre multivarié : propriétés par rapport au niveau de risque et à la structure de dépendance**, est fondé sur un article, coécrit avec Areski Cousin : *"A multivariate extension of Value-at-Risk and Conditional-Tail-Expectation"*, actuellement soumis pour publication. Dans ce chapitre, nous introduisons des généralisations multivariées de la *Value-at-Risk* (VaR) et de la *Conditional-Tail-Expectation* (CTE). En particulier nous nous sommes intéressés à l'analyse du comportement des ces mesures multivariées par rapport à deux facteurs : le changement du niveau du risque et le changement de la structure de dépendance du couple. En ce qui concerne le premier point nous mettons en lumière les caractéristiques de nos mesures en termes de propriétés classiques pour des mesures de risque (monotonie par rapport au niveau de risque, homogénéité, invariance par translation, chargement de sécurité) (voir Artzner *et al.*, 1999). En ce qui concerne le deuxième point, nous examinons comment ces mesures se comportent par rapport à la structure de copule et au changement des risques marginaux. En se plaçant dans des familles paramétriques spécifiques de copules nous arrivons à donner des formules exactes pour nos mesures et donc à analyser le comportement par rapport au changement du modèle de dépendance et du niveau de risque.

Dans le quatrième chapitre de la thèse, intitulé **Estimation de type plug-in des ensembles de niveau dans un cadre non-compact avec applications à la théorie des risques**, nous abordons d'abord le problème de l'estimation de quantiles bidimensionnels (voir Serfling, 2002) et des ensembles de niveau d'une fonction de répartition bivariée (voir Tibiletti, 1990, 1993, 1994). Grâce à une méthode *plug-in* nous proposons un estimateur des courbes quantiles. De plus la définition de la CTE introduite dans le chapitre précédent est reprise, dans un cadre bivarié, et nous proposons un estimateur pour cette mesure. Nous étudions les propriétés asymptotiques des estimateurs ainsi construits. Ce chapitre a fait l'objet d'un article intitulé *"Plug-in estimation of level sets in a non-compact setting with applications in multivariate risk theory"*, coécrit avec Thomas Laloë, Véronique Maume-Deschamps et Clémentine Prieur et accepté pour publication dans la revue *ESAIM : Probability and Statistics*.

A la fin de chaque chapitre nous proposons une section dédiée aux simulations avec des données réelles et/or simulées, pour illustrer les résultats théoriques obtenus.

Les perspectives de recherche relatives aux différents sujets abordés dans cette thèse sont proposées dans la Section *Conclusion et perspectives*, en conclusion de la thèse. Des rappels et des définitions utiles sont données dans l'*Annexe* A. Des illustrations, simulations et développements plus approfondis sont détaillés dans l'*Annexe* B.

1.2 Description par chapitre des principaux résultats

Dans cette section nous donnons une brève présentation de chaque chapitre de ce travail de thèse, en développant les points évoqués ci-dessus, ainsi que les résultats obtenus. Pour chaque chapitre, nous introduisons d'abord les résultats existants sur lesquels notre travail s'appuie, pour ensuite présenter les résultats de notre contribution.

1.2.1 Estimation de la queue d'une distribution bivariée : approche reposant sur les copules

Risques et phénomènes rares

"When simultaneously monitoring two possibly dependent, positive risks one is often interested in quantile regions with very small probability p. These extreme quantile regions contain hardly or no data and therefore statistical inference is difficult. In particular when we want to protect ourselves against a calamity that has not yet occurred, we take $p < 1/n$*, with* n *the sample size."* Einmahl and de Haan (2009), page 1.

Une grande partie de la littérature en sciences actuarielles ou en théorie du risque de ces dernières décennies gravite autour des sujets fondamentaux suivants : l'impact et l'étude des queues de distributions et des événements extrêmes, et l'étude de la dépendance entre les risques. De nombreux domaines sont concernés par ce type de problèmes. Nous pouvons citer par exemple les sciences de l'environnement (prévision de pics de pollution, de catastrophes naturelles), la santé publique (propagation des épidémies), l'assurance, ou encore l'analyse financière (par exemple la prévision de crises monétaires). Cette liste, certainement loin d'être exhaustive, d'applications possibles de la théorie du risque, nous montre l'importance de déterminer la probabilité qu'un événement rare et grave (dans un certain sens) se produise.

La théorie de valeurs extrêmes est essentiellement née pour répondre à ce genre de questions (voir par exemple Embrechts *et al.*, 1997, Chapitre 3). Elle a été développée pour estimer la probabilité d'occurrence des événements rares et pour étudier la queue de la distribution de ces phénomènes. Pour un large spectre des problématiques que la théorie des valeurs extrêmes a tenté d'explorer dans les dernières décennies nous renvoyons le lecteur aux livres de Leadbetter *et al.* (1983), Resnick (1987), Embrechts *et al.* (1997), Beirlant *et al.* (2004), Coles (2001), de Haan et Ferreira (2006), Reiss et Thomas (2007).

L'idée est la suivante : A partir d'un échantillon de données de taille n, nous voulons évaluer la probabilité d'un événement rare représenté par un ensemble extrême E. D'un point de vue statistique, les difficultés apparaissent lorsque E est tellement rare qu'aucune observation de l'échantillon ne tombe dans E. Dans ce cas nous ne possédons aucune information directe sur cet événement. Si on approche ce problème en utilisant les méthodes statistiques classiques (par exemple la fonction de répartition empirique), l'information contenue dans le reste de l'échantillon masquera les informations essentielles concernant l'événement E. Ce genre de méthode apparaît donc inapproprié dans ce contexte. Nous pouvons cependant penser, logiquement, que les informations inhérentes aux valeurs extrêmes non observées sont contenues dans les valeurs les plus

extrêmes observées dans l'échantillon. Il faudra donc sélectionner les valeurs extrêmes des données, qui sont censées contenir l'information la plus pertinente sur l'événement rare E.

La théorie des valeurs extrêmes développée jusqu'à aujourd'hui nous suggère d'effectuer cette sélection essentiellement selon deux critères : la *méthode des maxima* (voir Théorème 1.2.1 ci-dessous) et la *méthode des excès au-delà d'un seuil* (voir Théorème 1.2.2 ci-dessous). Nous commençons par présenter les résultats classiques de la théorie des valeurs extrêmes unidimensionnelle pour ensuite passer à leur généralisation en dimension plus grande, (en particulier nous nous concentrons sur le problème bidimensionnel).

Le Théorème de Fisher-Tippet (Théorème 1.2.1 ci-dessous) donne les différentes limites possibles de la loi du maximum de n variables aléatoires indépendantes et identiquement distribuées (i.i.d.), correctement normalisées.

Théorème 1.2.1 *(Embrechts* et al.*, 1997, Théorème 3.2.3) Soient $X_1, X_2, \ldots, X_n$ une suite de variables aléatoires i.i.d. avec fonction de répartition commune F. S'il existe une suite de valeurs positives $(a_n)_{n>0}$ et une suite de valeurs réelles $(b_n)_{n>0}$ telles que*

$$\lim_{n\to\infty} \mathbb{P}\left[\frac{\max\{X_1, X_2, \ldots, X_n\} - b_n}{a_n} \leq x\right] = H_k(x), \quad x \in \mathbb{R}, \tag{1.1}$$

avec $H_k(x)$ fonction de répartition non dégénérée, alors $H_k(x)$ est un élément de la famille GEV (Generalized Extreme Value Distribution ou loi de valeurs extrêmes généralisée) :

$$H_k(x) = \begin{cases} \mathrm{e}^{-(1-k\,x)^{\frac{1}{k}}}, & \text{if } k \neq 0, \\ \mathrm{e}^{-\mathrm{e}^{-x}}, & \text{if } k = 0, \end{cases} \tag{1.2}$$

où $1 - k\,x > 0$. Le paramètre $k \in \mathbb{R}$ détermine la famille à laquelle la limite appartient (Fréchet pour $k < 0$, Gumbel pour $k = 0$ et Weibull pour $k > 0$).

La fonction en (1.1) est aussi appelée représentation de Jenkinson-Von Mises. S'il existe une suite de valeurs positives $(a_n)_{n>0}$ et une suite de valeurs réelles $(b_n)_{n>0}$ telles que la limite en (1.1) est vérifiée, on dit que F appartient au domaine d'attraction pour les maxima d'une loi de valeurs extrêmes généralisée H_k, et l'on note $F \in MDA(H_k)$.

Le second critère est la méthode des excès (appelée POT, Peak Over Threshold) (voir de Haan et Rootzén, 1993 ; McNeil, 1997, 1999). Elle représente une modélisation bien analysée et étudiée dans la littérature, surtout dans un contexte unidimensionnel (voir par exemple Davison, 1984 ; Davison et Smith, 1990). En littérature unidimensionnelle cette méthode est souvent utilisée pour estimer des quantiles extrêmes, et donc des mesures de risque univariées (*Value-at-Risk*) (voir par exemple El-Aroui et Diebolt, 2002 ; McNeil et Saladin, 1997 ; McNeil, 1999 ; Einmahl *et al.*, 2006).
La méthode POT s'appuie sur le Théorème de Pickands-Balkema-de Haan (Balkema et de Haan, 1974 ; Pickands, 1975 ; Embrechts *et al.*, 1997 - Théorème 3.4.5). Ce théorème démontre que la loi des excès au-delà d'un seuil u est asymptotiquement une loi de Pareto généralisée (GPD : Generalised Pareto Distribution) de fonction de répartition :

$$V_{k,\sigma}(x) := \begin{cases} 1 - \left(1 - \frac{kx}{\sigma}\right)^{\frac{1}{k}}, & \text{if } k \neq 0,\ \sigma > 0, \\ 1 - \mathrm{e}^{\frac{-x}{\sigma}}, & \text{if } k = 0,\ \sigma > 0, \end{cases} \tag{1.3}$$

avec $x \geq 0$ si $\xi \leq 0$ ou $0 \leq x < \frac{\sigma}{\xi}$ si $\xi > 0$. Plus précisément il existe une fonction mesurable β telle que

$$\lim_{u \to x_F} \sup_{0 \leq x < x_F - u} \left| F_u(x) - V_{k,\beta(u)}(x) \right| = 0,$$

où $F_u(x) = \mathbb{P}[X - u \leq x \,|\, X > u]$ et $x_F := \sup\{x \in \mathbb{R} \,|\, F(x) < 1\}$ (le point terminal de la fonction F).
Dans le Chapitre 2 de cette thèse nous proposons une version bidimensionnelle de ce résultat en utilisant la Upper Tail Dependence Copula (UTDC, voir Définition (2.2.1)). Le théorème suivant représente le lien entre les deux approches présentées ci-dessus :

Théorème 1.2.2 *(Embrechts et al., 1997; Théorème 3.4.13(b)) Pour chaque $k \in \mathbb{R}$*

$$F \in MDA(H_k) \quad \Leftrightarrow \quad \exists\, \sigma(u) > 0 \,:\, \lim_{u \to x_F} \sup_{0 \leq x < x_F - u} \left| F_u(x) - V_{k,\sigma(u)}(x) \right| = 0,$$

avec $F_u(x) = \mathbb{P}[X - u \leq x \,|\, X > u]$, $x_F := \sup\{x \in \mathbb{R} \,|\, F(x) < 1\}$ et $V_{k,\sigma(u)}$ la distribution de Pareto généralisée introduite en (1.3).

Remarque 1.2.3 **(Conditionnellement ou inconditionnellement par rapport aux excès)** Supposons que l'on observe $X_1, \ldots, X_n$ i.i.d. avec une distribution commune F. Soit u un certain seuil fixé et N le nombre d'excès au-delà de u. Dans le Chapitre 2 de cette thèse nous considérons deux approches. La première consiste à travailler conditionnellement à N. Si n est la taille de l'échantillon et u_n le seuil associé, le nombre d'excès est m_n, avec $\lim_{n\to\infty} m_n = \infty$ et $\lim_{n\to\infty} m_n/n = 0$. La deuxième approche consiste à considérer le nombre d'excès N_n comme une variable aléatoire binomiale (ce qui est bien le cas, par exemple, dans les simulations), $N_n \sim \mathrm{Bi}(n, 1 - F(u_n))$ avec $\lim_{n\to\infty} 1 - F(u_n) = 0$ et $\lim_{n\to\infty} n(1 - F(u_n)) = \infty$. Ces considérations seront utiles par la suite, en particulier dans la Section 2.5 du Chapitre 2. Voir à ce propos les simulations dans un cadre unidimensionnel dans la Section B.3 de l'Annexe.

Evidemment une approche unidimensionnelle ne permet pas de considérer des modèles dans lesquels plusieurs risques, éventuellement dépendants entre eux, ont des interactions. Il semble donc nécessaire d'augmenter la dimension du problème et de considérer une approche multidimensionnelle (bidimensionnelle dans notre cas) du problème de la gestion des risques. Ce passage au cadre bidimensionnel change substantiellement la structure du problème. En effet une première différence essentielle vient des multiples relations d'ordre possibles en dimension supérieure à un. Une seconde vient de la nécessité de prendre en compte la dépendance entre les excès multidimensionnels.
Considérons $(X_1, Y_1), \ldots, (X_n, Y_n)$ des paires aléatoires indépendantes de fonction de répartition bivariée F. On dit que F est dans le domaine d'attraction pour les maxima d'une loi de valeurs extrêmes bivariée $H(x, y)$ (on note encore $F \in MDA(H)$) s'il existe

des suites positives $(a_{1\,n})_{n>0}$, $(a_{2\,n})_{n>0}$, et des suites réelles $(b_{1\,n})_{n>0}$ et $(b_{2\,n})_{n>0}$, telles qu'on ait :

$$\lim_{n\to\infty} \mathbb{P}\left[\frac{\max\{X_1, X_2, \ldots, X_n\} - b_{1\,n}}{a_{1\,n}} \leq x;\ \frac{\max\{Y_1, Y_2, \ldots, Y_n\} - b_{2\,n}}{a_{2\,n}} \leq y\right] = \lim_{n\to\infty} F^n(x\,a_{1\,n} + b_{1\,n};\, y\,a_{2\,n} + b_{2\,n}) = H(x, y), \quad (1.4)$$

avec $x, y \in \mathbb{R}$ et $H(x, y)$ fonction de répartition avec marges non dégénérées. La loi limite $H(x, y)$ est alors appelée loi de valeurs extrêmes bivariée et ses marges $H_1(x) = H(x, \infty)$, $H_2(y) = H(\infty, y)$ sont des lois de valeurs extrêmes univariées. Pour des suites convenablement choisies (voir de Haan et Ferreira, 2006 ; Théorème 1.1.3) il existe γ_1 et $\gamma_2 \in \mathbb{R}$ (indices des valeurs extrêmes marginales), tels que $H_1(x) = \mathrm{e}^{-(1+\gamma_1\,x)^{-1/\gamma_1}}$ et $H_2(x) = \mathrm{e}^{-(1+\gamma_2\,x)^{-1/\gamma_2}}$.
Nous introduisons maintenant une formulation du domaine d'attraction bivarié pour les maxima qui utilise la *stable tail dependence function* (voir Huang, 1992) et qui sera utile dans la suite du Chapitre 2 de cette thèse.
Une fonction de répartition bivariée F avec marges F_X et F_Y a une *stable tail dependence function* l si pour $x \geq 0$ et $y \geq 0$ la limite :

$$\lim_{t\to 0} \frac{1}{t}\,\mathbb{P}[1 - F_X(X) \leq tx \text{ or } 1 - F_Y(Y) \leq ty] := l(x, y), \quad (1.5)$$

existe.
Notons que l'existence de la limite en (1.5) est équivalente à

$$\lim_{t\to 0} \frac{1}{t}\,\mathbb{P}[1 - F_X(X) \leq tx, 1 - F_Y(Y) \leq ty] := R(x, y) = x + y - l(x, y). \quad (1.6)$$

De plus si C est la copule de la fonction de répartition F alors l'existence de la limite en (1.5) est aussi équivalente à

$$\lim_{t\to 0} \frac{1}{t}\,(1 - C(1 - tx, 1 - ty)) = l(x, y), \text{ for } x \geq 0,\ y \geq 0. \quad (1.7)$$

Donc si F_X, F_Y appartiennent au domaine d'attraction pour les maxima de deux lois de valeurs extrêmes H_1, H_2 et si (1.5) est vérifiée, alors F est dans le domaine d'attraction d'une loi de valeurs extrêmes bivariée H avec marges H_1, H_2 et avec une copule déterminée par l. La fonction l représente donc la structure de dépendance entre X et Y dans la queue de la distribution bidimensionnelle et

$$l(x, y) = -\log H\left(\frac{x^{-\gamma_1} - 1}{\gamma_1}, \frac{y^{-\gamma_2} - 1}{\gamma_2}\right), \quad (1.8)$$

où log est la fonction logarithme de base e (voir par exemple Einmahl *et al.*, 2006). On a de plus $\max\{x, y\} \leq l(x, y) \leq x + y$, pour $x \geq 0$, $y \geq 0$. La borne inférieure correspond au cas de complète dépendance asymptotique ($H(x, y) = \min\{H_1(x), H_2(y)\}$) et la borne supérieure au cas d'indépendance dans la queue ou indépendance asymptotique ($H(x, y) = H_1(x)\,H_2(y)$). De plus nous pouvons montrer que chaque loi de valeurs

extrêmes est une loi max-stable[1] et vice versa (voir de Haan et Ferreira, 2006 ; Section 6.1).

Dans la section suivante le concept d'indépendance asymptotique, essentiel dans la théorie des valeurs extrêmes multivariées, est analysé plus en détails. Plusieurs pistes pour l'étude et la description de la dépendance dans les extrêmes multidimensionnels ont été introduites : la fonction de Pickands (voir Pickands, 1981 ; Capéraà et Fougères, 2000), la mesure exposante et la mesure spectrale (voir Beirlant *et al.*, 2004, Section 8), la fonction de Sibuya (voir Sibuya, 1960 ; Geffroy, 1958) etc. Nous renvoyons le lecteur intéressé par l'analyse du lien explicite entre la fonction l et ces autres instruments mathématiques de modélisation de la dépendance extrêmale aux ouvrages de Haan et Ferreira (2006), Section 6 ou Beirlant *et al.* (2004), Section 8.

Dépendance asymptotique versus indépendance asymptotique

Comme nous l'avons souligné dans la section précédente la fonction l varie entre la dépendance asymptotique parfaite (comonotonie) et l'indépendance asymptotique. Dans ce deuxième cas nous pouvons également écrire :

$$\lambda := \lim_{u\to 1} \mathbb{P}[\, F_X^{-1}(X) > u \,|\, F_Y^{-1}(Y) > u \,] = R(1,1) = 0, \tag{1.9}$$

voir Ledford et Tawn (1996) ou Coles *et al.* (1999). La condition (1.9) représente par exemple des situations où les données montrent une association positive ou négative qui disparaît seulement progressivement pour des niveaux de plus en plus extrêmes. En d'autres termes elle décrit la situation dans laquelle la probabilité que les deux composantes soient grandes en même temps est négligeable par rapport à la probabilité qu'une composante soit grande. Notons que tel est le cas pour de nombreuses familles de copules paramétriques : Ali Mikhail-Haq, Frank, Clayton avec paramètres $a \geq 0$, Fairlie-Gumbel-Morgenstern, normale avec paramètres $\rho > 0$. Donc différentes lois classiques sont dans le domaine d'attraction de l'indépendance.

La théorie des valeurs extrêmes, avec son hypothèse d'appartenance au domaine d'attraction (voir (1.4)), nous mène à des résultats satisfaisants dans le cas de vecteurs aléatoires asymptotiquement dépendants. En revanche dans le cas de vecteurs asymptotiquement indépendants elle répond seulement en partie aux questions posées. Des conditions supplémentaires sont nécessaires pour obtenir des résultats satisfaisants.

A partir de (1.4), (1.5) et (1.8) nous pouvons écrire, dans le cas de l'indépendance asymptotique entre X et Y, que

$$\lim_{t\to\infty} t\,\mathbb{P}\left[\frac{X-b_1(t)}{a_1(t)} > x; \frac{Y-b_2(t)}{a_2(t)} > y\right] = \log H(x,y) - \log H(x,\infty) - \log H(\infty,y) = 0, \tag{1.10}$$

(voir de Haan et de Ronde, 1998, Section 6.5 ; de Haan et Ferreira, 2006, Section 8.3). Le cas d'indépendance asymptotique se présente donc comme un cas plus complexe et intéressant. A partir de la formule (1.10), nous constatons que l'approximation du

[1] Définition : Une loi H sur $\mathbb{R}^2$ est max-stable si pour tout $n \in \mathbb{N}$, il existe des suites w_n^1, w_n^2, z_n^1 et z_n^2 telles que $H^n(x,y) = H(w_n^1\, x + z_n^1,\, w_n^2\, y + z_n^2)$.

premier ordre, sur laquelle s'appuie le modèle de valeurs extrêmes multidimensionnel, n'est pas satisfaisant dans le cas d'indépendance asymptotique. Il sera nécessaire d'affiner l'hypothèse en (1.4) et de considérer des approximations du second ordre, afin de modéliser les queues jointes dans le cas de l'indépendance asymptotique (voir par exemple Fougères, 2004).

En suivant ce genre de considérations une partie du Chapitre 2 est donc consacrée au cas d'indépendance asymptotique. Nous introduirons finalement un estimateur performant dans les cas de dépendance et d'indépendance asymptotique (voir Section 2.3.2, Chapitre 2).

Estimation de la queue d'une distribution bivariée : différentes méthodes pour répondre au même problème

La plupart des modèles d'estimation des queues d'une distribution doivent affronter le problème de la dépendance-indépendance asymptotique entre les variables aléatoires en jeu (voir section précédente). Sans prétendre être exhaustifs, nous présentons dans cette section un aperçu rapide des estimateurs de queue plus connus et plus utilisés dans la littérature. Nous considérons l'hypothèse d'appartenance au domaine d'attraction d'une loi des valeurs extrêmes bivariée, présentée déjà en (1.4) :

$$\lim_{n\to\infty} F^n(a_{1\,n}x + b_{1\,n}, a_{2\,n}y + b_{2\,n}) = H(x,y), \qquad \text{pour } x,\, y \in [-\infty,\infty]. \tag{1.11}$$

En écrivant $F^n = [1 - n^{-1}(n(1-F))]^n$ et en utilisant le fait que $(1-n^{-1}x)^n \to \mathrm{e}^{-x} \in [0,1]$ si et seulement si $x_n \to x \in [0,\infty]$, pour $n \to \infty$, nous avons que la limite en (1.11) est vérifiée si et seulement si

$$\lim_{n\to\infty} n\,(1 - F(a_{1\,n}x + b_{1\,n}, a_{2\,n}y + b_{2\,n})) = -\log H(x,y), \tag{1.12}$$

avec la convention que $-\log(0) = \infty$. Grâce à la max-stabilité de H on peut réécrire la limite en (1.12) comme

$$\begin{aligned}(1 - F(a_{1\,n}x + b_{1\,n}, a_{2\,n}y + b_{2\,n})) &\sim -\log H(\alpha_{1\,n}x + \beta_{1\,n}, \alpha_{2\,n}y + \beta_{2\,n}) \\ &\sim 1 - H(\alpha_{1\,n}x + \beta_{1\,n}, \alpha_{2\,n}y + \beta_{2\,n}),\end{aligned} \tag{1.13}$$

pour $n \to \infty$ et pour (x,y) tel que $0 < H(x,y) < 1$ (voir Beirlant *et al.*, 2004 ; Section 8.3.1). Cette approximation en (1.13) est souvent utilisée comme point de départ pour l'inférence statistique de la distribution $F(x,y)$, dans la région bivariée composée par les couples (x,y) tels que $F_X(x)$ et $F_Y(y)$ soient assez proches de un. Donc en choisissant u_1, u_2 tels que $F_X(u_1)$ et $F_Y(u_2)$ soient proches de un, (1.13) nous suggère l'approximation suivante :

$$F(x,y) \approx H(a_{1\,n}^{-1}\,\alpha_{1\,n}\,x - a_{1\,n}^{-1}\,\alpha_{1\,n}\,b_{1\,n} + \beta_{1\,n}, a_{2\,n}^{-1}\,\alpha_{2\,n}\,y - a_{2\,n}^{-1}\,\alpha_{2\,n}\,b_{2\,n} + \beta_{2\,n}) := \widetilde{H}(x,y), \tag{1.14}$$

pour $x \geq u_1$, $y \geq u_2$. Puisque H et $\widetilde{H}$ diffèrent seulement par les paramètres d'échelle et de position, $\widetilde{H}$ est aussi une loi des valeurs extrêmes avec la même fonction l. L'approximation en (1.13) est à la base de la méthode nommée *censored-likelihood* (voir

Smith, 1994, Section 3 ; Ledford et Tawn, 1996, Section 2 ; Beirlant *et al.*, 2004, Section 9.4.2).

De plus la limite en (1.4) peut être réécrite en termes de copules. Il suffit de remarquer que la copule de F^n, la loi des maxima, est $C_{F^n}(u_1, u_2) = C_F^n(u_1^{1/n}, u_2^{1/n})$, pour $u_1, u_2 \in [0, 1]$. En particulier on sait que $F \in MDA(H)$ si et seulement si $F_X \in MDA(H_1)$, $F_Y \in MDA(H_2)$ et

$$\lim_{t\to\infty} C_F^t(u_1^{1/t}, u_2^{1/t}) = C_H(u_1, u_2), \quad u_1, u_2 \in [0, 1],$$

voir par exemple Berlaint *et al.* (2004), Section 8.3.2. Par la propriété de max-stabilité de H on a $C_F(u_1, u_2) \sim C_H(u_1, u_2)$, pour $u_1 \sim 1$ et $u_2 \sim 1$. En utilisant la stable tail dependence function l, on obtient l'approximation suivante pour les queues jointes de la distribution F :

$$\mathscr{F}_1(x, y) = \exp\{-l\,(-\log(F_X(x)), -\log(F_Y(y)))\}, \tag{1.15}$$

pour x et y telles que $F_X(x)$ et $F_Y(y)$ soient assez proches de un. A partir de (1.15) et en utilisant un développement du premier ordre on obtient :

$$\mathscr{F}_2(x, y) = 1 - l(1 - F_X(x), 1 - F_Y(y)), \tag{1.16}$$

pour $1 - F_X(x)$ et $1 - F_Y(y)$ suffisamment petits (voir Ledford et Tawn, 1996 ; Beirlant *et al.*, 2004 ; Section 8.3.1 et 8.3.2). Notons que l'approximation en (1.15) est préférable par rapport à (1.16) car $1 - l(\cdot, \cdot)$ sous-évalue la probabilité des excès joints (pour de plus amples détails, nous renvoyons le lecteur intéressé à l'ouvrage de Beirlant *et al.*, 2004 ; Section 8).

Une piste pour fournir une estimation des approximations en (1.15)-(1.16) consiste à utiliser l'estimateur de la fonction l (voir par exemple de Haan et de Ronde, 1998 ; Drees et Huang, 1998 ; Draisma *et al.*, 2004 ; Einmahl *et al.*, 2008). De plus les marges peuvent être estimées avec la méthode de dépassement de seuil univariée (Peaks Over Threshold method, POT, voir par exemple McNeil, 1997, 1999). Nous obtenons respectivement :

$$\widehat{\mathscr{F}_1^*}(x, y) = \exp\{-\widehat{l}(-\log(\widehat{F}_X^*(x)), -\log(\widehat{F}_Y^*(y)))\}, \tag{1.17}$$

$$\widehat{\mathscr{F}_2^*}(x, y) := 1 - \widehat{l}(1 - \widehat{F}_X^*(x), 1 - \widehat{F}_Y^*(y)), \tag{1.18}$$

pour des grandes valeurs de x et y, où $\widehat{F}_X^*(x)$ (resp. $\widehat{F}_Y^*(y)$) est l'estimateur POT de la loi marginale de X au-delà d'un certain seuil u_X (respectivement de Y au-delà d'un certain seuil u_Y). Cet estimateur POT univarié est décrit en détails dans la Section 2.4.1.

Des problèmes surgissent avec les techniques d'estimation bivariées décrites ci-dessus quand X et Y sont asymptotiquement indépendantes (voir (1.9)). En effet, dans le cas d'indépendance asymptotique ces méthodes produisent un biais significatif. Le problème est bien décrit par Ledford et Tawn (1996), début Section 5, page 177 :

"In common with all other joint tail estimation methods, the model derived in Section 2" [ici en (1.17)] *"is based on asymptotic dependence in the joint tail. Specifically a*

multivariate extreme value type dependence is assumed to hold in the region above the marginal thresholds."

Ce problème ensuite est repris par Beirlant *et al.* (2004), début Section 9.5, page 342 :

"... within the class of max-stable distributions, the only possible type of asymptotic independence is, in fact, perfect independence. This makes the class rather inappropriate for modeling data that exhibit an asymptotic independence structure."

Pour surmonter ce problème, Ledford et Tawn (1996) ont présenté un modèle dans lequel la dépendance de queue est caractérisée par un coefficient $\eta \in (0, 1]$. Dans leurs travaux la fonction de répartition de survie d'un couple aléatoire (Z_1, Z_2) avec marges Fréchet unitaires est supposé satisfaire :

$$\mathbb{P}[Z_1 > z, Z_2 > z] = \overline{F}(z_1, z_2) \sim L(z)\, \mathbb{P}[Z_1 > z]^{\frac{1}{\eta}}, \tag{1.19}$$

pour $z \to \infty$, où L est une fonction à variation lente à l'infini[2]. Le modèle en (1.19) représente le cas de dépendance asymptotique si $\eta = 1$ et $L(t) \nrightarrow 0$. Si $\eta = 1$ et $L(t) = 1$, Z_1 et Z_2 sont parfaitement dépendants ; si $\eta = 1/2$ et $L(t) = 1$ Z_1 et Z_2 sont parfaitement indépendantes. Dans le cas $\eta < 1$ (ou $\eta = 1$ et $L(t) \to 0$) on a l'indépendance asymptotique. Plus précisément :

- Indépendance asymptotique avec association positive si $1/2 < \eta < 1$;
- Indépendance asymptotique avec association négative si $0 < \eta < 1/2$;
- *Quasi-indépendance* si $\eta = 1/2$ et $L(t) \geq 1$.

Pour une étude du paramètre η pour différentes fonctions de répartition le lecteur intéressé est renvoyé à Heffernan (2000). Plusieurs méthodes d'estimation du coefficient η ont été proposées : Peng (1999), Draisma *et al.* (2004) et récemment Beirlant *et al.* (2011). Pour quelques contre-exemples du modèle de Ledford et Tawn voir par exemple Schlather (2001). De plus Ledford et Tawn ont développé différents sous-modèles à partir du modèle en (1.19) (voir par exemple Ledford et Tawn, 1997, 1998). Par exemple un sous-modèle proposé considère une approximation pour la queue du type :

$$\overline{F}(z_1, z_2) = L(z_1, z_2)\, z_1^{-c_1}\, z_2^{-c_2}, \quad \text{avec } z_1 \text{ et } z_2 \text{ simultanément grands}, \tag{1.20}$$

$c_1, c_2 > 0$, $c_1 + c_2 = 1/\eta$ et L une fonction à variation lente bivariée, i.e.

$$\lim_{t \to 0} \frac{L(tx,\, ty)}{L(t, t)} = \widetilde{g}(x, y), \tag{1.21}$$

pour une fonction $\widetilde{g}$ telle que $\widetilde{g}(cx,\, cy) = \widetilde{g}(x,\, y)$, pour chaque $c > 0$ (c'est à dire $\widetilde{g}$ est homogène d'ordre zéro). Evidemment (1.20) implique (1.19).
Comme mentionné dans la section précédente quelques auteurs (voir de Haan et de Ronde, 1998 ; Section 6.5, Draisma *et al.*, 2004 ; Section 2, de Haan et Ferreira, 2006 ; Section 7.6.) se sont intéressés récemment, inspirés par le modèle de Ledford et Tawn,

[2]Définition : Une fonction positive et mesurable L sur $(0, \infty)$ est à variation lente à l'infini si $\lim_{x \to \infty} \frac{L(t\,x)}{L(x)} = 1, \quad \forall\, t > 0$.

à renforcer l'hypothèse d'appartenance au domaine d'attraction d'une loi des valeurs extrêmes (voir condition (1.12)) avec une condition de second ordre. Avec cette hypothèse supplémentaire, la théorie des valeurs extrêmes multivariée est capable de donner une modélisation cohérente aussi dans le cas d'indépendance asymptotique. Plus précisément on suppose qu'il existe une fonction positive $A(t)$ telle que

$$\lim_{t\to\infty} \frac{t\,(1-F(t\,x,t\,y))+\log H(x,y)}{A(t)} := \widetilde{Q}(x,y), \tag{1.22}$$

avec A une fonction à variation régulière[3] d'indice $\rho \leq 0$ et $\rho = 1-1/\eta$. En supposant une condition de second ordre comme en (1.22), dans un cadre d'indépendance asymptotique, on obtient :

$$\begin{aligned}\lim_{t\to\infty} \frac{t\,\mathbb{P}\,[X>t\,x;Y>t\,y]-\log H(x,y)-x^{-1}-y^{-1}}{A(t)} &= \lim_{t\to\infty} \frac{t\,\mathbb{P}\,[X>t\,x;Y>t\,y]}{A(t)} \\ &= \widetilde{Q}(x,\infty)+\widetilde{Q}(\infty,y)-\widetilde{Q}(x,y) \neq 0. \end{aligned} \tag{1.23}$$

Nous remarquons que, grâce à la limite en (1.23), nous sommes à présent capables d'évaluer asymptotiquement une probabilité d'événements dans les queues que l'on ne pouvait auparavant prédire (voir (1.10)). De plus la condition (1.20) est satisfaite dès que l'on suppose une condition de second ordre comme en (1.22) (voir de Haan et de Ronde, 1998 ; Section 6.5). Dans notre Chapitre 2, nous suivons ce type d'approche, en utilisant une condition de second ordre, pour modéliser le problème de l'estimation de la queue dans le cas d'indépendance asymptotique.

En conclusion de cette section, nous rappelons que, durant la dernière décennie, des extensions bivariées de la méthode de dépassement de seuil (Multivariate Peak Over Threshold method) via la distribution de Pareto généralisée ont été développées dans une série d'articles par Falk et Reiss (2001, 2002, 2003a, 2003b, 2005), Kaufmann et Reiss (1995), ou dans le livre de Reiss et Thomas (voir Reiss et Thomas, 2007 ; Chapitre 13, coécrit avec Falk). Récemment une généralisation multivariée a été traitée dans Tajvidi (1996), Beirlant *et al.* (2004), Section 8.3.1, Rootzén et Tajvidi (2006) et Michel (2008).

Par exemple en suivant le récent travail de Rootzén et Tajvidi (2006) on peut développer le modèle bidimensionnel suivant. Soient $\mathbf{x} = (x,y)$ et $\mathbf{0} = (0,0)$. Soit H une loi de valeurs extrêmes bivariée et $\mathbf{u} = \{u_1(t), u_2(t)\,|\,t \in [1,\infty)\}$, une courbe partant de $\mathbf{u}(1) = (0,0)$. Soit $\sigma(\mathbf{u}(t)) = \sigma(u_1(t), u_2(t)) > (0,0)$ une fonction à valeurs dans $\mathbb{R}^2$. Si $F \in MDA(H)$ et $0 < H(0,0) < 1$ alors il existe $\mathbf{u}$ et $\sigma(\mathbf{u}(t))$ comme ci-dessus telles que

$$\begin{aligned}\lim_{t\to\infty} \mathbb{P}\left[\frac{X-u_1(t)}{\sigma_1(u_1(t))} \leq x,\ \frac{Y-u_2(t)}{\sigma_2(u_2(t))} \leq y \,\middle|\, X>u_1(t) \text{ ou } Y>u_2(t)\right] \\ = \frac{1}{-\log H(\mathbf{0})}\,\log\left(\frac{H(\mathbf{x})}{H(\min\{\mathbf{x},\mathbf{0}\})}\right), \end{aligned} \tag{1.24}$$

[3] Définition : Une fonction positive et mesurable h sur $(0,\infty)$ est à variation régulière d'indice $\alpha \in \mathbb{R}$ si $\lim_{x\to\infty} \frac{h(t\,x)}{h(x)} = t^{\alpha}, \quad \forall\, t>0$.

pour tout x, y, et vice versa (voir Rootzén et Tajvidi, 2006 ; Théorème 2.2). Ce type de résultat peut être considéré comme une possible généralisation du Théorème de Pickands-Balkema-de Haan (voir Balkema et de Haan, 1974 ; Pickands, 1975), en dimension plus grande que un.

Nous remarquons qu'en dimension plus grande que un nous retrouvons exactement la correspondance entre la loi unidimensionnelle des maxima (la loi de valeurs extrêmes ; GEV distribution) et la loi unidimensionnelle des excès au-delà du seuil (la loi de Pareto généralisée, GP distribution) : la GDP est dans une échelle logarithmique par rapport à la GEV.

A partir de (1.24), Beirlant et ses coauteurs soulignent que :

"The event $\{\boldsymbol{X} \not\leq u_n\}$ is called an exceedance over the (multivariate) threshold u_n. It entails that there is at least one coordinate variable X_j that exceeds the corresponding threshold $u_{n,j}$, although the precise coordinate where this happens remains unspecified" Beirlant *et al.*, 2004 ; Section 8.3.1, page 277.

Donc, comme l'observe Resnick dans la citation en ouverture de cette introduction, la façon dont les excès au-delà de seuils doivent être définis en dimension plus grande que un n'est pas intuitivement claire. Différents choix sont possibles et ils correspondent à des problèmes de modélisation et à des applications pratiques différentes. Par exemple une étude approfondie de la distribution bidimensionnelle quand une seule coordonnée devient grande peut être trouvée dans Fougères et Soulier (2008). Dans le Chapitre 2 de la thèse, nous nous concentrerons sur le cas où les deux coordonnées sont conjointement grandes, au-delà d'un seuil bidimensionnel (opportunément choisi).

Le rôle de la distribution de Pareto généralisée multivariée (présentée en (1.24)) dans la théorie de valeurs extrêmes est un point ouvert de la recherche de ces dernières années. Nous remarquons qu'en utilisant la distribution Pareto généralisée multivariée une estimation des paramètres d'échelle doit être développée afin de pouvoir fournir un estimateur de la queue de la distribution. Notre travail s'insère dans cette partie de la littérature : via la Upper Tail Dependence Copula (UTDC, voir Définition (2.2.1)), notre méthode réussit à séparer et donc à estimer séparément la partie de la structure de dépendance et la partie provenant de la contribution des deux marges.

A notre connaissance la procédure de POT que nous proposons dans ce papier ne peut pas être directement déduite de méthodes multivariées de dépassement de seuil proposées dans les travaux cités ci-dessus. Elle appartient cependant au même cadre d'idées et utilise une approche analogue du problème. De plus nous fournissons une estimation de la distribution bivariée $\mathbb{P}[X \leq x, Y \leq y]$, pour x et y au-delà de certains seuils, et ce type d'estimation n'est pas obtenu explicitement dans les travaux cités ci-dessus. Cependant, quelques éléments pour une possible comparaison sont examinés dans le Théorème 4.2 dans Juri et Wüthrich (2004).

Après cette introduction aux principaux aspects et problématiques de la théorie des valeurs extrêmes multidimensionnelle (brièvement résumés et sans prétention d'exhaustivité) nous sommes maintenant en mesure de décrire notre contribution à cette théorie et nos principaux résultats obtenus dans le Chapitre 2 de cette thèse. Pour une description complète nous renvoyons le lecteur directement au chapitre en question.

Estimation de la queue d'une distribution bivariée : notre proposition

Différents modèles de risque en assurance ou en finance se concentrent sur l'étude de situations dans lesquelles la structure de dépendance entre les facteurs en jeu n'est pas négligeable, comme par exemple les krachs financiers qui sont connus pour être dépendants. Le message qui provient des récents travaux (voir par exemple Frees et Valdez, 1998 ; Embrechts *et al.*, 2002 ; Embrechts *et al.*, 2003) est que l'étude de la dépendance demande plus qu'une simple approximation d'une quantité scalaire ou d'une mesure de dépendance (coefficients de dépendance dans les queues, coefficients de corrélation linéaire etc). En partant de la simple considération que la dépendance relative à un ensemble de variables aléatoires est complètement décrite par la distribution de la probabilité jointe, une approche distributionnelle du problème peut avoir de nombreux avantages. Cette approche distributionnelle consiste à modéliser la dépendance dans les queues jointes, à travers des copules appropriées. L'étude du comportement à la limite, et donc des comportements extrêmes, pour ces copules n'a reçu que peu d'attention durant la dernière décennie. Le point de départ de notre travail est essentiellement constitué des travaux de Juri et Wüthrich (2002, 2004), Wüthrich (2004), Charpentier et Juri (2006), Charpentier et Segers (2007), Javid (2009). Pour une étude dans un cas appliqué (marché financier asiatique) de la dépendance dans la queue en utilisant les copules, nous renvoyons le lecteur intéressé aux travaux de Caillault et Guégan (2005).

Nous rappelons que dans le Chapitre 2 nous nous sommes attachés à fournir un estimateur de $\mathbb{P}[X \leq x, Y \leq y]$, pour x et y plus grands qu'un seuil bidimensionnel opportunément choisi. Dans notre approche, comme dans les travaux préalablement cités, la structure bidimensionnelle de dépendance entre les composantes aléatoires continues X et Y, est modélisée via une copule C. De plus la dépendance entre X et Y dans la queue est étudiée via une copule appropriée, c'est-à-dire la Upper Tail Dependence Copula (UTDC, voir Définition (2.2.1)).

Le point de départ est l'hypothèse que le comportement dans la queue se fonde sur la condition suivante de régularité de la copule (voir condition (15) dans Juri et Wüthrich, 2004) :

$$\lim_{u \to 1} \frac{C^*(x\,(1-u),\, y\,(1-u))}{C^*(1-u,\, 1-u)} = G(x,y), \quad \forall\, x,\, y > 0, \tag{1.25}$$

où C^* est la copule de survie (voir Définition A.1.4). La fonction G représente donc, dans notre modèle une description explicite de la structure de dépendance dans la queue de F. De plus G est homogène d'ordre θ (voir de Haan *et al.*, 1984 ; Théorème 2.1) :

$$G(t\,x,\, t\,y) = t^\theta G(x,y). \tag{1.26}$$

Par ailleurs G a la forme connue suivante :

$$G(x,y) = \begin{cases} x^\theta g_Y(\frac{y}{x}) & \text{for } \frac{y}{x} \in [0,1], \\ y^\theta g_X(\frac{x}{y}) & \text{for } \frac{y}{x} \in (1,\infty), \end{cases} \tag{1.27}$$

ou $g_X(x) := G(x,1)$, $g_Y(y) := G(1,y)$, (voir Aczél, 1966 ; page 231).

Si nous supposons que la limite en (1.6) est vérifiée alors, dans le cas de dépendance asymptotique ($\lambda > 0$, voir (1.9)) on a

$$G(x,y) = \frac{R(x,y)}{R(1,1)} = \frac{R(x,y)}{\lambda},$$

et, grâce à l'homogénéité d'ordre 1 de la fonction R, on obtient $\theta = 1$ (voir la Section 2.3.1). Dans le cas d'indépendance asymptotique $\lambda = 0$ et $R(x,y) = R(1,1) = 0$, pour tout x et $y \geq 0$. Nous aurons donc besoin d'une condition de second ordre qui affine l'hypothèse (1.25) (voir condition (1.30) ci-après).

Une des justifications théoriques sur lesquelles s'appuie la construction de notre estimateur est une extension bidimensionnelle du Théorème Pickands-Balkema-de Haan (voir par exemple Balkema et de Haan, 1974 ; Pickands, 1975) utilisant la Upper Tail Dependence Copula (UTDC, voir Définition (2.2.1)).

Dans la littérature nous trouvons des extensions du Théorème de Pickands-Balkema-de Haan utilisant la Lower Tail Dependence Copula (LTDC). Nous pouvons par exemple citer les travaux de Wüthrich (2004), Charpentier et Juri (2006), Charpentier et Segers (2007). Dans le cas de la Upper Tail Dependence Copula (UTDC) on a des généralisations dans les cas de distributions marginales égales, dans le cas de copules symétriques (Juri et Wüthrich, 2004), ou dans le cas de copules Archimédiennes (Juri et Wüthrich, 2002). Notre extension (Théorème 1.2.4 ci-dessous) est plus générale dans le sens où elle requiert des hypothèses moins restrictives sur la copule C par rapport à la littérature citée ci-dessus. En utilisant la condition de variation régulière à zéro d'indice θ pour la copule de survie $C^*(t,\, t)$ présentée en (1.25) on obtient :

Théorème 1.2.4 *Soient X et Y deux variables aléatoires réelles continues, avec des marginales différentes notées F_X, F_Y, et une copule C. Supposons que $F_X \in MDA(H_{\xi_1})$, $F_Y \in MDA(H_{\xi_2})$ et que la copule C vérifie la condition* (1.25). *Alors*

$$\sup_{\mathscr{A}} \Big| \mathbb{P}\big[X - u \leq x, Y - F_Y^{-1}(F_X(u)) \leq y \big| X > u, Y > F_Y^{-1}(F_X(u))\big] \\ - C^{*\,G}\big(1 - g_X(1 - V_{\xi_1,a_1(u)}(x)), 1 - g_Y(1 - V_{\xi_2,a_2(F_Y^{-1}(F_X(u)))}(y))\big)\Big| \xrightarrow[u \to x_{F_X}]{} 0, \quad (1.28)$$

où $\mathscr{A} := \{(x,y) : 0 < x \leq x_{F_X} - u, 0 < y \leq x_{F_Y} - F_Y^{-1}(F_X(u))\}$, x_{F_X} et x_{F_Y} sont respectivement les points terminaux de F_X et F_Y, $V_{\xi_i,a_i(\cdot)}$ la GPD avec paramètres $\xi_i, a_i(\cdot)$ définie en (1.3) *et $C^{*\,G}(w,t) := w + t - 1 + G(g_X^{-1}(1-w),\, g_Y^{-1}(1-t))$ pour $w,\, t > 0$.*

Donc, en utilisant le Théorème 1.2.4, nous proposons l'estimateur suivant pour la queue de la distribution bidimensionnelle :

$$\widehat{F}^*(x,y) = \left(\frac{1}{n}\sum_{i=1}^{n} 1_{\{X_i>u,\, Y_i>\widehat{u}_Y\}}\right)\big(1 - \widehat{g}_X(1 - V_{\widehat{\xi}_X,\widehat{\sigma}_X}(x-u))$$
$$- \widehat{g}_Y(1 - V_{\widehat{\xi}_Y,\widehat{\sigma}_Y}(y-\widehat{u}_Y)) + \widehat{G}\big(1 - V_{\widehat{\xi}_X,\widehat{\sigma}_X}(x-u), 1 - V_{\widehat{\xi}_Y,\widehat{\sigma}_Y}(y-\widehat{u}_Y)\big)\big)$$
$$+ \widehat{F}_1^*(u,y) + \widehat{F}_2^*(x,\widehat{u}_Y) - \frac{1}{n}\sum_{i=1}^{n} 1_{\{X_i\le u,\, Y_i\le \widehat{u}_Y\}}, \quad (1.29)$$

où :

- $x > u$, $y > \widehat{u}_Y$, $\widehat{u}_Y := \widehat{F}_Y^{-1}(\widehat{F}_X(u))$, $\widehat{F}_X(u) = \frac{1}{n}\sum_{i=1}^{n} 1_{\{X_i\le u\}}$ et $\widehat{F}_Y^{-1}$ est la fonction quantile empirique de Y ;
- u est "assez" grand (voir Remarque 3 pour le problème du choix de u) ;
- $\widehat{F}_1^*(u,y)$ et $\widehat{F}_2^*(x,\widehat{u}_Y)$ sont définis comme en (1.17) ;
- $\widehat{g}_X$, $\widehat{g}_Y$ et $\widehat{G}$ sont des estimateurs non paramétriques consistants de g_X, g_Y et G respectivement, opportunément construits (voir Sections 2.3.1-2.3.2) ;
- V_{ξ_X,σ_X} est la loi de Pareto généralisée relative à X et V_{ξ_Y,σ_Y} relative à Y (voir (1.3)) ;
- $\widehat{\xi}_X$, $\widehat{\sigma}_X$ (resp. $\widehat{\xi}_Y$, $\widehat{\sigma}_Y$) sont les estimateurs des paramètres, selon la méthode du maximum de vraisemblance, de la loi de Pareto généralisée V_{ξ_X,σ_X} (resp. V_{ξ_Y,σ_Y}).

La construction détaillée de l'estimateur en (1.29) est proposée dans la Section 2.4.2.

En utilisant des modifications des résultats unidimensionnels de Smith (voir Smith, 1987 ; Théorèmes 3.2 et 8.1 et nos Corollaires 2.5.1 et 2.5.2), nous montrons deux théorèmes de consistance pour notre estimateur en (1.29). Plus précisément :

- Dans le cas de dépendance asymptotique nous montrons un résultat de convergence de l'écart absolu entre la fonction et l'estimateur, évalués dans la queue, avec vitesse $\sqrt{k_n}$, où $k_n \to \infty$, $k_n/n \to 0$. Pour une formulation rigoureuse de ce résultat nous renvoyons à la Section 2.6.1 (Théorème 2.6.1).
- Dans le cas de l'indépendance asymptotique, en accord avec la discussion présentée dans les sections précédentes (voir par exemple la condition (1.22)), nous avons besoin d'une hypothèse supplémentaire de second ordre, c'est-à-dire un affinement de la condition (1.25). Nous supposons :

$$\lim_{t\to 0} \frac{\frac{C^*(tx,ty)}{C^*(t,t)} - G(x,y)}{q_1(t)} := Q(x,y), \quad (1.30)$$

  pour tout $x, y \ge 0$, $x + y > 0$, où q_1 est une fonction positive et Q n'est ni une constante ni un multiple de G (voir Draisma *et al.*, 2004 ; Section 2, de Haan et Ferreira, 2006 ; Section 7.6).

Grâce à la condition (1.30) nous pouvons montrer un résultat de convergence de l'écart absolu entre la fonction et l'estimateur, évalué dans la queue, dans le cas de l'indépendance asymptotique. Dans ce cas, la vitesse de convergence sera $\sqrt{n\,q(k_n/n)}$ ou $q(t) := \mathbb{P}[1 - F_X(X) < t, 1 - F_Y(Y) < t]$, $k_n \to \infty$ et $k_n/n \to 0$. Pour plus de détails et une formulation rigoureuse de ce résultat voir la Section 2.6.2 (Théorème 2.6.2).

La robustesse de notre estimateur est comparée à celle de $\widehat{\mathscr{F}}_1^*$ dans (1.17), en utilisant des données réelles (voir la Section 2.7) qui ont déjà été étudiées dans la littérature des valeurs extrêmes multidimensionnelle (voir par exemple Beirlant *et al.*, 2011 ; Frees et Valdez 1998 ; Lescourret et Robert, 2006 ; Draisma *et al.*, 2004 ; Coles et Tawn, 1994 ; Ramos et Ledford, 2009).

Analogies et différences par rapport au modèle de Ledford et Tawn

Nous remarquons que l'hypothèse (de premier ordre) de variation régulière à zéro d'indice θ pour la fonction $q(t)$ (voir (1.25)) est en complète analogie avec le modèle de Ledford et Tawn (voir Ledford et Tawn, 1996, 1997, 1998). Par exemple dans le modèle proposé par Ledford et Tawn (1996) $q(t) = t^\theta$ (voir discussion à ce propos dans Draisma *et al.*, 2004 ; Section 2). Sous cette hypothèse la vitesse de convergence du Théorème 2.6.2, présenté ci-dessus, devient $\sqrt{n\,q(k_n/n)} = \sqrt{n^{1-\theta}\,k_n^\theta}$, avec $\theta > 1$. Ce type de résultat est en complète analogie, par exemple, avec le Théorème 2.1 de Peng (1999), le Théorème 2.1 et le Lemme 6.1 de Draisma *et al.* (2004).

En utilisant le modèle de Ledford et Tawn (1997, 1998) en (1.20), avec marges Fréchet unitaires on obtient :

$$\begin{aligned} C^*(1-u, 1-v) &\sim L\left(-1/\ln(u), -1/\ln(v)\right) \left(-1/\ln(u)\right)^{-c_1} \left(-1/\ln(v)\right)^{-c_2} \\ &\sim L\left(1/(1-u), 1/(1-v)\right) \left(1/(1-u)\right)^{-c_1} \left(1/(1-v)\right)^{-c_2}, \end{aligned} \tag{1.31}$$

où $u = \mathrm{e}^{-1/z_1}$, $v = \mathrm{e}^{-1/z_2}$ et u, $v \to 1$. Donc, sous les hypothèses du modèle de Ledford et Tawn en (1.20)-(1.21), grâce à l'approximation (1.31), la condition (1.25) devient :

$$\begin{aligned} \lim_{u\to 1} \frac{C^*(x(1-u), y(1-u))}{C^*(1-u, 1-u)} &= \lim_{u\to 1} \frac{L\left(\frac{1}{x(1-u)}, \frac{1}{y(1-v)}\right)\left(\frac{1}{x(1-u)}\right)^{\frac{-1}{2\eta}}\left(\frac{1}{y(1-v)}\right)^{\frac{-1}{2\eta}}}{L\left(\frac{1}{1-u}, \frac{1}{1-v}\right)\left(\frac{1}{1-u}\right)^{\frac{-1}{2\eta}}\left(\frac{1}{1-v}\right)^{\frac{-1}{2\eta}}} \\ &= \widetilde{g}\left(\frac{1}{x}, \frac{1}{y}\right) x^{\frac{1}{2\eta}}\, y^{\frac{1}{2\eta}}, \end{aligned} \tag{1.32}$$

avec $\widetilde{g}$ comme dans la limite en (1.21). Nous avons obtenu donc l'analogie entre les conditions (1.20)-(1.21) et les conditions (1.25)-(1.27) : $G(x, y) = \widetilde{g}\left(\frac{1}{x}, \frac{1}{y}\right) x^{\frac{1}{2\eta}}\, y^{\frac{1}{2\eta}}$ (voir Ramos et Ledford, 2009 ; Section 2).
Cela implique l'existence d'un lien explicite entre le paramètre de dépendance η du modèle dans (1.20)-(1.21) et le paramètre θ du modèle dans (1.25)-(1.27). En utilisant les propriétés de $\widetilde{g}$ et l'expression de G (voir (1.27)) on obtient $\theta = 1/\eta$. Dans la Section

4.2 dans Juri et Wüthrich (2004) ce lien entre θ et η n'est pas explicité clairement. Pour de plus amples détails et des exemples dans des cas spécifiques de fonctions de répartition nous renvoyons le lecteur intéressé à l'Annexe B.1.

De plus dans Peng (1999) ou Draisma *et al.* (2004) la condition de second ordre en (1.30) est présentée comme *un raffinement du modèle classique de Ledford et Tawn*, dans le but de fournir des résultats de consistance pour l'estimateur du paramètre de dépendance η dans le cas d'indépendance asymptotique. En complète analogie avec cette approche nous posons aussi, dans notre Chapitre 2, la condition (1.30) pour aborder le cas de l'indépendance asymptotique.

A la fin de cette section nous rapportons des considérations sur les différences entre une approximation de la queue comme en (1.20)-(1.21) et une approche centrée sur un *"Théorème de convergence UTDC"* (voir Juri et Wüthrich, 2004 ; Théorème 2.3, ou Proposition 2.2.2, dans notre version réadaptée sans supposer la symétrie pour la copule C).

Juri et Wüthrich soulignent que :

"Contrary to (51) [ici (1.19)] *Theorem 2.3* [ici Proposition 2.2.2, en version réadaptée] *describes explicitly the dependence structure in the tails (conditioned to* $\{X > u, Y > u\}$*) since we know that the limit of* C_u^{up} *equals* C^{*G}*, that is Theorem 2.3 gives the whole distributional description of the survival distribution function"*, Juri et Wüthrich (2004), page 227.

En particulier ce type de considération met en évidence la nécessité d'utiliser des instruments adéquats pour l'étude de la dépendance asymptotique sans recourir à des approximations du comportement dans la queue qui utilisent des quantités scalaires comme des coefficients de dépendance (voir Juri et Wüthrich, 2004 ; page 234). Dans ce sens la limite de la Upper Tail Dependence Copula joue un rôle fondamental dans l'approche de Juri et Wüthrich et donc dans notre Chapitre 2 (voir la Section 2.2 et en particulier la Proposition 2.2.2).

En effet cette méthode fondée sur un *"Théorème de convergence UTDC"* (voir Proposition 2.2.2) reflète une *approche distributionnelle* de la modélisation de la dépendance dans la queue et fournit une description naturelle d'événements extrêmes bivariés. Sur ce sujet le lecteur intéressé est renvoyé aux discussions dans Charpentier et Juri (2006 ; pages 563-564) et Juri et Wüthrich (2002 ; page 405). Parmi les avantages de cette approche distributionnelle on remarque qu'elle nous permet de mieux faire face au problème du manque de données, cas typique pour des événements rares (pour plus de détails voir Charpentier et Juri, 2006). De plus nous notons que notre méthode est libre du prétraitement de données. En effet nous pouvons travailler avec les échantillons originaux sans transformation sur les lois marginales.

1.2.2 Quelques généralisations de mesures de risque dans un cadre multivarié : propriétés par rapport au niveau de risque et à la structure de dépendance

Dans cette section nous abordons brièvement la problématique de l'évaluation d'un risque et la comparaison entre différents risques dans un cadre unidimensionnel. L'intérêt de considérer l'interaction et la dépendance entre les différents risques est également analysé. Nous présentons quelques tentatives existantes d'extension de ces concepts en dimension plus grande que un. Nous introduisons enfin nos propositions de mesures de risque multivariées.

Risques : comment les évaluer

Les risques étant modélisés par des variables aléatoires positives, mesurer un risque correspond à donner une fonction ρ entre l'espace de variables aléatoires et $\mathbb{R}_+$. Dans la théorie du risque univarié, énormément développée ces dernières années, les mesures de risque sont donc pensées comme unidimensionnelles : $\rho[X]$. Les mesures de risque apparaissent ainsi comme des *outils de quantification de risque*. Elles permettent d'évaluer le niveau de dangerosité d'un risque mais également de comparer des risques entre eux et de les classer selon leur niveau de dangerosité. Voici un exemple de caractérisation d'une mesure de risque :

"A risk measure is a functional ρ 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.

La quantification et la comparaison des risques peuvent ensuite être utilisées à plusieurs fins telles que l'évaluation de primes, l'allocation de capital, la détermination de risques d'un portefeuille d'assurance ou de réassurance.

Nombreuses mesures de risque ont été proposées, des plus simples aux plus compliquées. Une approche axiomatique a été développée afin de chercher la caractérisation de la notion de mesure de risque la plus "juste" (voir Artzner *et al.*, 1999). Cependant, le choix de la mesure la plus appropriée pour quantifier le risque en assurance comme en finance reste un sujet de débat. L'une des mesures de risque les plus populaires, utilisée dans la pratique, est la *Value-at-Risk* (VaR), également appelée Valeur-à-Risque. Etant donnée une fonction de distribution continue et croissante F_X d'une perte positive X, la VaR au niveau de risque α est le quantile :

$$VaR_\alpha(X) = Q_X(\alpha) = F_X^{-1}(\alpha), \quad \forall\, \alpha \in (0,1), \tag{1.33}$$

c'est-à-dire le point qui accumule une probabilité α dans la queue à gauche et $1-\alpha$ à droite.

L'utilisation de cette mesure de risque n'est pas privée d'inconvénients : voir Denuit et Charpentier (2004) ; Chapitre 5, ou Denuit *et al.* (2005), page 72 :

"Also shareholders and management should be concerned with the question "how bad is bad ?" when they want to evaluate the risk at hand in a consistent way."

Ou encore Yamai et Yoshiba (2002), page 182 :

"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.

Le problème est essentiellement que la VaR ne prend pas en compte l'importance du déficit quand la ruine survient mais seulement la fréquence du déficit. Elle ne donne qu'une information ponctuelle au quantile α de la distribution du sinistre et aucune information au-delà de ce point. Toutefois, le contexte actuariel nécessite l'étude de la queue des distributions (voir Section 1.2.1). En effet, un risque qui a une probabilité significative de causer de forts sinistres est dangereux par nature et nous avons besoin d'évaluer ce niveau de dangerosité. Ce genre de considérations est à la base des développements de mesures comme la *Conditional-Tail-Expectation* (CTE), appelée aussi Espérance de queue, ou comme la *Tail-Value-at-Risk* (TVaR).

La première mesure est définie par :

$$\mathrm{CTE}_\alpha(X) = \mathbb{E}[\, X \,|\, X \geq Q_X(\alpha)\,], \tag{1.34}$$

pour $\alpha \in (0,1)$ et $Q_X(\alpha)$ comme en (1.33) (voir Artzner *et al.*, 1999 ; Dedu et Ciumara, 2010). L'Espérance de queue consiste donc à déterminer le capital de manière à surmonter en moyenne les pires exercices.
La seconde mesure est définie par :

$$\mathrm{TVaR}_\alpha(X) = \frac{1}{1-\alpha} \int_\alpha^1 \mathrm{VaR}_u(X)\, \mathrm{d}u. \tag{1.35}$$

Elle correspond à la moyenne arithmétique des valeurs de la VaR pour les niveaux de confiance au-delà de α. Dans le cas d'une perte X continue la $\mathrm{TVaR}_\alpha(X) = \mathbb{E}[\, X \,|\, X \geq \mathrm{VaR}_\alpha(X)\,]$, ce qui est exactement la définition de la *Conditional-Tail-Expectation* au niveau α.

Un autre inconvénient de la VaR est qu'elle n'est pas une *mesure de risque cohérente* selon la définition de Artzner *et al.* (1999) (sous-additivité non nécessairement vérifiée) :

Définition 1.2.5 (Mesure de risque cohérente) *Une mesure de risque ρ est dite cohérente si, pour deux variables aléatoires X et Y elle satisfait les propriétés suivantes :*

- *Monotonie : si $\mathbb{P}[X \leq Y] = 1$, alors $\rho[X] \leq \rho[Y]$;*
- *Sous-additivité : $\rho[X+Y] \leq \rho[X] + \rho[Y]$;*
- *Homogénéité positive : $\rho[a\,X] = a\,\rho[X]$;*
- *Invariance par translation : $\rho[X+b] = \rho[X] + b$.*

En revanche la TVaR en (1.35) vérifie toutes les propriétés de la Définition 1.2.5 aussi bien pour des v.a. continues que discrètes.

D'autres propriétés pourraient également être souhaitées pour une mesure de risque. Denuit et Charpentier (2004) suggèrent par exemple :

- $\rho[X] \geq \mathbb{E}[X]$, ce qui signifie que la mesure d'un risque doit être plus grande que son espérance, on dit que ρ contient un chargement de sécurité ;
- Si $X = c$ presque sûrement, alors $\rho[X] = \rho[\,c\,] = c$, pour c scalaire, ce qui signifie que la mesure d'un montant certain est ce montant lui même ;
- Si $X \leq x_{\max}$ presque sûrement, alors $\rho[X] \leq \rho[\,x_{\max}\,]$, ce qui signifie que la mesure de risque ne peut pas dépasser le montant maximal d'un sinistre dû à ce risque.

Pour quelques utilisations pratiques des mesures en (1.33), (1.34) et (1.35) nous renvoyons par exemple à Yamai et Yoshiba (2002) ; Section 2.

Risques : comment les comparer ?

Maintenant que nous sommes capables de mesurer la dangerosité d'un risque, notre but est de formaliser de manière rigoureuse l'idée qu'un risque soit plus dangereux qu'un autre. Cela nous plonge essentiellement dans la théorie des ordres stochastiques (voir Kaas *et al.*, 2002 ; Denuit *et al.*, 2005, Chapitre 3 ; Müller et Stoyan, 2001, Chapitre 1).

L'écriture $X \preceq Y$ indique que X est moins risqué de Y et donc que $\rho[X] \leq \rho[Y]$. Soit $F_X(x)$ la fonction de répartition de X et $\overline{F}_X(x)$ la fonction survie. Il existe différents ordres stochastiques couramment utilisés dans la littérature. Par exemple nous rappelons :

Définition 1.2.6 (Dominance stochastique) *(Goovaerts et al., 1984)*
On dit qu'une variable aléatoire X est plus grande au sens de la dominance stochastique qu'une variable aléatoire Y ($X \preceq_{st} Y$) si $\overline{F}_X(x) \leq \overline{F}_Y(x)$ pour tout x.

Il y a beaucoup de conditions équivalentes pour l'ordre de la dominance stochastique (voir Wang, 1998 ou Denuit *et al.*, 2005, Section 3.3.1.1).

Définition 1.2.7 (Ordre de dangerosité) *(Goovaerts et al., 1984)*
On dit qu'une variable aléatoire X est plus grande au sens de l'ordre de dangerosité qu'une variable aléatoire Y ($X \preceq_D Y$) si $\mathbb{E}[X] \leq \mathbb{E}[Y] < \infty$ et il existe une constante c telle que $\overline{F}_X(x) \leq \overline{F}_Y(x)$ pour $x < c$ et $\overline{F}_X(x) \geq \overline{F}_Y(x)$ pour $x \geq c$.

Donc si $X \preceq_{st} Y$ ou $X \preceq_D Y$ on a $\rho[X] \leq \rho[Y]$.

Spécifiquement dans la suite nous introduisons une structure de dépendance particulière entre variables aléatoires : une structure de dépendance positive. Etant donné un couple (X, Y), la dépendance positive représente la tendance que X et Y soient grands ou petits ensemble. Nous trouvons une exposition excellente de tous les concepts de

dépendance positive dans Joe (1997). Ce n'est pas dans notre intérêt ici de présenter les différents concepts de dépendance positive. Nous nous limitons seulement à introduire la *"positive regression dependence"* qui sera ensuite utile dans le Chapitre 3 pour l'étude des propriétés de nos mesures multivariées (voir Section 3.2.5).

La *"positive regression dependence"* à été introduite par Lehmann (1966) comme :

Définition 1.2.8 ("Positive regression dependence", PRD$(Y \mid X)$**)** *Un couple* (X, Y) *admet une "positive regression dependence" par rapport à* X, PRD$(Y \mid X)$, *si*

$$[Y \mid X = x_1] \preceq_{st} [Y \mid X = x_2], \quad \forall\ x_1 \leq x_2. \tag{1.36}$$

Pour plus de détails nous renvoyons le lecteur à, par exemple, Belzunce *et al.* (2007), Section 4.

Risques et dépendance : quelques pistes pour des extensions multivariées

L'étude et la modélisation de la dépendance en finance et en assurance sont motivées par différentes raisons. Une des motivations est que les hypothèses d'indépendance, à la base des différents modèles stochastiques, sont souvent éloignées de la réalité des problèmes étudiés. De plus, dans de nombreux cas, négliger la structure de dépendance (ou mal l'évaluer) mène à une sous-estimation du risque avec des conséquences parfois dramatiques (voir par exemple Chebana et Ouarda, 2011, par rapport à l'importance de la modélisation de la dépendance entre les risques dans le domaine de l'hydrologie).

Couramment utilisés pour l'évaluation des risques, nous trouvons des coefficients scalaires de dépendance, des mesures de risque unidimensionnelles, comme par exemple la corrélation linéaire ou la *Value-at-Risk*. Récemment, différentes études ont montré l'inefficacité de ce genre d'approches, généralement limitées à des modèles trop simples, et leur incapacité à traiter des situations réelles plus sophistiquées (voir par exemple Frees et Valdez, 1998 ; Embrechts *et al.*, 2002 ; Embrechts *et al.*, 2003). Il n'est pas surprenant qu'au cours des dernières décennies, la littérature des mathématiques appliquées au risk-management et aux problèmes d'interactions et dépendance entre les risques se soit si fortement développée. Pour une ample présentation des fondements de la théorie de risque nous adressons le lecteur à Denuit *et al.* (2005), Grandell (1991), Müller et Stoyan (2001), Kaas *et al.* (2002).

Récemment une partie encore peu développée de la littérature s'est intéressée au problème de l'extension multidimensionnelle des mesures de risque univariées proposées par Artzner *et al.* (1999), en proposant des mesures qui prennent également en compte la structure de dépendance entre les risques. Nous renvoyons le lecteur intéressé à, par exemple, Müller et Stoyan (2001) (Chapitres 1 et 3) et Denuit *et al.* (2005) (Chapitre 3 et 4) pour le problème de la comparaison entre les risques par rapport à différents critères ou à Cherubini *et al.* (2004) pour des études sur les risques financiers utilisant les copules.

En particulier différentes généralisations de la *Conditional-Tail-Expectation* ont été proposées dans la littérature récente, en conditionnant par rapport au risque total ou

à quelques risques extrêmes du portefeuille. Ce genre de mesures sont souvent utilisées dans la littérature pour aborder des problèmes d'allocation de capital du portefeuille lorsque les risques sont dépendants. Sans prétendre être exhaustifs, nous rappelons par exemple :

$$\mathbb{E}[\, X_i \,|\, S > Q_S(\alpha)\,], \quad \text{avec } S := X_1 + \cdots + X_n, \tag{1.37}$$

$$\mathbb{E}[\, X_i \,|\, \min\{X_1, \ldots, X_n\} > Q_{\min\{X_1,\ldots,X_n\}}(\alpha)\,], \tag{1.38}$$

$$\mathbb{E}[\, X_i \,|\, \max\{X_1, \ldots, X_n\} > Q_{\max\{X_1,\ldots,X_n\}}(\alpha)\,], \tag{1.39}$$

pour $i = 1, \ldots, n$. Pour de plus amples détails sur ces mesures le lecteur est renvoyé par exemple à Cai et Li (2005). Pour des formules explicites de $\mathbb{E}[X_i \,|\, S > Q_S(\alpha)\,]$ dans le cas d'une copule FGM le lecteur est renvoyé à Bargès *et al.* (2009). Landsman et Valdez obtiennent la formule explicite de $\mathbb{E}[\, S \,|\, S > Q_S(\alpha)\,]$ et $\mathbb{E}[\, X_i \,|\, S > Q_S(\alpha)\,]$ pour des fonctions de répartition elliptiques (Landsman et Valdez, 2003). Cai et Li obtiennent des formules explicites des mesures ci-dessus dans le cas de fonctions de répartition *phase-type* (Cai et Li, 2005). Hürlimann propose la définition de *multivariate conditional value-at-risk vector*, pour $X = (X_1, \ldots, X_n)$, comme :

$$\mathrm{CTE}_\alpha[S \mid X] = \mathbb{E}[\, S \,|\, X_j > Q_{X_j}(\alpha), \ \text{ pour } j = 1, \ldots, n],$$

(Hürlimann, 2004).

Mesures de risque multivariées : nos propositions

En partant des considérations et des généralisations de la *Conditional-Tail-Expectation* présentées ci-dessus nous proposons maintenant des extensions des mesures de risque en dimension plus grande que un.

Les mesures de risque multivariées introduites par la suite sont des correspondances entre les ensembles de vecteurs aléatoires positifs et $\mathbb{R}_+^d$, $d \geq 2$:

$$\rho : \quad \mathbf{X} := (X_1, \ldots, X_d) \mapsto \begin{pmatrix} \rho^1[\mathbf{X}] \\ \vdots \\ \rho^d[\mathbf{X}] \end{pmatrix} \in \mathbb{R}_+^d, \tag{1.40}$$

où ρ^i indique la $i-$ème composante du vecteur $\rho[\mathbf{X}]$ en $\mathbb{R}_+^d$. Passer en dimension plus grande que un entraîne la perte d'ordre total, mais offre l'avantage d'englober plus d'informations sur les phénomènes multivariés étudiés. Notons que chaque composante du vecteur $\rho[\mathbf{X}]$ dépend des risques X_i, pour $i = 1, \ldots, d$, en même temps. La modélisation du problème est donc plus élaborée qu'une simple union de problèmes unidimensionnels.

Rappelons que la *Conditional-Tail-Expectation* univariée, $\mathrm{CTE}_\alpha(X)$, est la moyenne du risque X, sachant que ce risque est plus grand que la VaR au niveau α (voir (1.34)). Il apparaît donc naturel, pour généraliser la *Conditional-Tail-Expectation* dans un cadre

multidimensionnel, que nous ayons besoin essentiellement d'une extension de la notion de quantile en dimension plus grande que un. Nous remarquons que, même s'il existe des généralisations multivariées de la fonction quantile, l'utilisation de cet outil est essentiellement confinée au cadre unidimensionnel, où il est plus facile à calculer et à interpréter.

Quantiles univariés et extension en dimension plus grande : les courbes quantiles

L'extension multivariée de la notion de quantile a été étudiée ces dernières années, mais reste un domaine peu exploré de la littérature actuelle. Une difficulté majeure est l'absence de relation d'ordre naturel en dimension $d > 1$. Pour un article pionnier dans la littérature sur ce sujet voir Barnett (1976).

Nous rappelons en premier lieu quelques pistes suivies dans la littérature récente. Pour un résumé assez exhaustif des différentes généralisations de la fonction quantile, le lecteur est renvoyé à Serfling (2002). Par exemple, Massé et Theodorescu (1994) ont défini le quantile multivarié à travers l'intersection des demi-plans avec probabilité au moins égale à p, Koltchinskii (1997) comme une inversion de fonctions, Tukey (1975) selon le critère de minimisation des distances en utilisant les fonctions de profondeur. Einmahl et Mason (1992) introduisent une généralisation de la fonction quantile fondée sur une fonction à valeurs réelles λ, une mesure de probabilité P sur $\mathbb{R}^n$, une sous-classe des ensembles de Borel $\mathcal{C}$ et définie comme :

$$U(p) := \inf\{\lambda(C) : \quad C \in \mathcal{C} \text{ et } P(C) \geq p\}, \qquad \text{pour } 0 < p < 1. \tag{1.41}$$

En particulier la convergence de la version empirique $U_n(p)$ de $U(p)$ en (1.41) est caractérisée en termes de convergence en loi du "*processus quantile généralisé*" $U_n(p) - U(p)$, pour $0 < p < 1$, convenablement normalisé (voir aussi Molchanov, 1990).

Kemperman (1987) et Abdous et Theodorescu (1992) introduisent le concept de *quantile spatial*. Soit $\mathbf{X}$ un vecteur aléatoire d-dimensionnel avec distribution $P(\cdot)$, la médiane spatiale est définie comme :

$$M := \arg\min_{\mathbf{c}} \int_{\mathbf{X}} \| \mathbf{x} - \mathbf{c} \| \| \mathbf{x} \| \, \mathrm{d}P(\mathbf{x}). \tag{1.42}$$

Dans (1.42) la norme utilisée joue un rôle important. Par exemple dans le cas de la norme L_1 on obtient le vecteur composé par les médianes de chaque composante marginale de $\mathbf{X}$.

Tibiletti (1993), Fernández-Ponce et Suárez-Lloréns (2002), Belzunce *et al.* (2007) définissent le quantile comme l'ensemble de points qui accumulent la même probabilité pour un *orthant* fixé. Ils l'appellent courbe de niveau ou *courbe quantile*. Nous remarquons comme cette approche représente l'extension naturelle de la définition univariée (1.33). De plus cette généralisation permet une segmentation des données en groupes de tailles préfixées. Elle peut être utilisée dans le cas de distributions symétriques et non-symétriques et, à la différence des autres méthodes présentes dans la littérature, elle est *"metric-free"*, c'est à dire qu'elle n'est pas directement conditionnée par le choix de la métrique, comme dans (1.42) (voir Tibiletti, 1993 ; Sections 1 et 2). Plus précisément :

Définition 1.2.9 *Pour $\alpha \in (0,1)$ et F fonction de répartition $d-$variée, le quantile multivarié au niveau de probabilité α est la frontière de l'ensemble de niveau $L(\alpha) = \{F(\boldsymbol{x}) \geq \alpha\}$. Cette frontière est aussi appelée courbe quantile. Nous notons $Q_F(\alpha) := \partial L(\alpha)$, avec $\alpha \in (0,1)$.*

En imposant $\alpha = 1/2$ dans la Définition 1.2.9 on obtient une extension naturelle de la médiane multidimensionnelle (voir Tibiletti, 1993). Pour des détails sur les propriétés des courbes de niveau et pour des formules paramétriques de ces courbes, dans un cadre bivarié, nous renvoyons à Belzunce *et al.* (2007). De Haan et Huang proposent un modèle pour des risques extrêmes d'inondations en utilisant l'estimateur des courbes de niveau d'une fonction de répartition bidimensionnelle (voir de Haan et Huang, 1995). Chebana et Ouarda remarquent que les événements hydrologiques sont souvent caractérisés par différentes variables corrélées : volume de l'inondation, durée et magnitude, intensité de la tempête etc. Le modèle qu'ils proposent se fonde sur les courbes de niveau de la distribution multivariée associée à ces variables aléatoires (voir Chebana et Ouarda, 2011).

$\mathrm{CTE}_\alpha(\mathbf{X})$ et $\mathrm{VaR}_\alpha(\mathbf{X})$: deux mesures multivariées

Inspirés par la notion de courbe quantile introduite ci-dessus et en considérant la caractérisation multivariée de la mesure de risque en (1.40), nous proposons la mesure de risque suivante :

$$\mathrm{CTE}_\alpha(\mathbf{X}) = \mathbb{E}[\mathbf{X} \,|\, \mathbf{X} \in L(\alpha)] = \begin{pmatrix} \mathbb{E}[\, X_1 \,|\, \mathbf{X} \in L(\alpha)\,] \\ \vdots \\ \mathbb{E}[\, X_d \,|\, \mathbf{X} \in L(\alpha)\,] \end{pmatrix}, \tag{1.43}$$

avec $\alpha \in (0,1)$ et $L(\alpha) = \{\mathbf{x} \in \mathbb{R}_+^d \,:\, F(\mathbf{x}) \geq \alpha\}$ (voir Section 3.3).

Notre version multivariée de la *Conditional-Tail-Expectation*, en conditionnant le vecteur aléatoire multidimensionnel à être dans l'ensemble de niveau $L(\alpha)$, se fonde essentiellement sur une *approche distributionelle*. On définit un niveau α et donc une région de risque associée : $L(\alpha)$. Grâce au fait que l'information globale du vecteur $\mathbf{X}$ soit complètement décrite par sa fonction de répartition multidimensionnelle, $L(\alpha)$ représente une région de risque prenant en compte conjointement l'information liée aux marginales et celle liée à la structure de dépendance. De plus nous remarquons qu'à la différence des généralisations de la *Conditional-Tail-Expectation*, précédemment introduites, notre version utilise un conditionnement différent. Nous prenons en compte la structure de dépendance du vecteur à travers la fonction de répartition multivariée et non en choisissant des "risques agrégés" (comme le minimum, le maximum ou la somme). En ce sens, une caractéristique de notre mesure est son comportement par rapport à des structures de dépendance (voir par exemple Proposition 3.3.8).

Nous introduisons également une notion de VaR multivariée sur le modèle de notre CTE multidimensionnelle en (1.43). Plus précisément en utilisant la notion de courbe

quantile $\partial L(\alpha)$ (voir Définition 1.2.9), nous définissons

$$\mathrm{VaR}_\alpha(\mathbf{X}) = \mathbb{E}[\mathbf{X} \,|\, \mathbf{X} \in \partial L(\alpha)] = \begin{pmatrix} \mathbb{E}[\, X_1 \,|\, \mathbf{X} \in \partial L(\alpha)\,] \\ \vdots \\ \mathbb{E}[\, X_d \,|\, \mathbf{X} \in \partial L(\alpha)\,] \end{pmatrix}, \tag{1.44}$$

pour $\alpha \in (0,1)$ (voir Section 3.2).

En suivant la généralisation de la fonction quantile proposée dans la Définition 1.2.9 Embrechts et Puccetti (2006), Nappo et Spizzichino (2009), définissent la *Value-at-Risk* en dimension deux comme la courbe quantile $Q_F(\alpha) = \partial L(\alpha)$ (voir Section 4.4). Contrairement à cette approche notre généralisation en (1.44) quantifie la mesure de risque, en utilisant le quantile multidimensionnel $Q_F(\alpha)$, de manière plus synthétique et parcimonieuse. Nous travaillons avec un vecteur de valeurs réelles plutôt qu'avec une infinité de points. Au niveau pratique de la gestion du risque cet aspect peut représenter un avantage opérationnel de notre $\mathrm{VaR}_\alpha(\mathbf{X})$.

Nous remarquons que les deux mesures multivariées (1.43)-(1.44) sont liées à la fonction de répartition de Kendall $K(\alpha) = \mathbb{P}[F(\mathbf{X}) \leq \alpha]$. Dans la littérature elle est aussi appelée "*multivariate probability integral transformation*". Le lecteur intéressé est renvoyé aux travaux de Barbe *et al.* (1996), Capéraà *et al.* (1997), Nelsen *et al.* (2003), Genest et Boies (2003), Genest *et al.* (2006), Belzunce *et al.* (2007), Nappo et Spizzichino (2009). L'étude des propriétés de cette distribution univariée est essentielle dans l'analyse de nos mesures de risque. En particulier grâce aux résultats théoriques obtenus dans la littérature pour la fonction de répartition de Kendall, nous pouvons énoncer la plupart des propriétés de la $\mathrm{VaR}_\alpha(\mathbf{X})$ et de la $\mathrm{CTE}_\alpha(\mathbf{X})$. Pour de plus amples détails sur la fonction de répartition de Kendall nous renvoyons le lecteur directement au Chapitre 3.

Dans le Chapitre 3 nous étudions le comportement de nos mesures de risque selon les propriétés axiomatiques caractéristiques de la théorie du risque univariée (voir par exemple Artzner *et al.*, 1999).

Plus précisément au cours du Chapitre 3 nous démontrons :

- Des propriétés axiomatiques pour nos mesures comme l'homogénéité, l'invariance par translation, le chargement de sécurité (voir Sections 3.2.1 and 3.3.1) ;
- Des propretés d'invariance d' une composante de la $\mathrm{VaR}_\alpha(\mathbf{X})$ (et de la $\mathrm{CTE}_\alpha(\mathbf{X})$) par rapport au changement des lois marginales des autres composantes, si la copula reste inchangée (voir Sections 3.2.3 and 3.3.3) ;
- La cohérence de nos mesures par rapport à différents ordres stochastiques (ordre de la dominance stochastique et ordre de dangerosité) des lois marginales du vecteur (voir Propositions 3.2.7 et 3.3.9) ;
- La $\mathrm{VaR}_\alpha(\mathbf{X})$ et la $\mathrm{CTE}_\alpha(\mathbf{Y})$, dans le cas de comonotonie, sont composées par les mesures univariées associées i.e. $\mathrm{VaR}^i_\alpha(\mathbf{X}) = \mathrm{VaR}_\alpha(X_i)$ and $\mathrm{CTE}^i_\alpha(\mathbf{X}) = \mathrm{CTE}_\alpha(X_i)$, for $i = 1, \ldots, d$ (voir Propositions 3.2.4 et 3.3.6).

- La croissance de la $\mathrm{VaR}_\alpha(\mathbf{X})$ et de la $\mathrm{CTE}_\alpha(\mathbf{X})$ par rapport au niveau de risque α sous certaines conditions de *"positive regression dependence"* entre les marginales et la Kendall $F(\mathbf{X})$ (voir Sections 3.2.5- 3.3.4). Ces conditions sont toujours vérifiées pour certaines familles paramétriques de copules comme par exemple pour les copules Archimédiennes bivariées.

Dans le Chapitre 3, une attention spéciale est portée à la classe des copules Archimédiennes car dans ce cas la fonction de répartition de Kendall a une forme explicite en fonction du générateur de la copule (voir par exemple Genest et Rivest, 2001, Section 2 ou Capéraà *et al.*, 1997). Enfin nous donnons des pistes intuitives d'applications possibles de nos mesures dans le cadre de la couverture par rapport à des risques extérieurs. En effet, les mesures de risque existantes (voir (1.37)-(1.39)) sont construites pour déterminer un capital minimum de solvabilité dans un contexte micro-prudentiel. Plus précisément une institution financière doit constituer un capital réglementaire qui ne prend pas en compte les risques extérieurs qui sont corrélés à son activité. Récemment, différentes études ont montré les possibles désavantages de cette approche micro-prudentielle (voir par exemple Gauthier *et al.*, 2010 ou Zhou, 2010). Dans ce sens la littérature de ces dernières années s'intéresse à une approche plutôt macro-prudentielle dans laquelle l'évaluation du risque est faite en prenant en compte l'impact des phénomènes exogènes et liés à l'activité interne de l'institution financière. Motivés par ces considérations, les mesures de risque multivariées que nous avons proposé ci-dessus (voir (1.43)-(1.44)) intègrent dans le modèle de gestion de risque la dépendance par rapport à des risques exogènes, en tenant compte de la structure de dépendance entre $X_1, X_2, \ldots, X_d$.

1.2.3 Estimation de type plug-in des ensembles de niveau dans un cadre non-compact avec applications à la théorie des risques

Dans ce chapitre nous reprenons la définition de l'Espérance de queue introduite dans le chapitre précédent (voir (1.43)), dans un cadre bivarié, c'est-à-dire :

$$\mathrm{CTE}_\alpha(X,Y) = \mathbb{E}[(X,Y)|(X,Y) \in L(\alpha)] = \begin{pmatrix} \mathbb{E}[\,X \,|\, (X,Y) \in L(\alpha)\,] \\ \mathbb{E}[\,Y \,|\, (X,Y) \in L(\alpha)\,] \end{pmatrix},$$

avec $\alpha \in (0,1)$ et $L(\alpha) = \{(x,y) \in \mathbb{R}_+^2 \,:\, F(x,y) \geq \alpha\}$. Le but est ici d'estimer cette mesure de risque.

Estimation de $L(\alpha)$ avec une méthode *plug in*

Afin d'estimer la $\mathrm{CTE}_\alpha(X,Y)$ nous abordons d'abord le problème de l'estimation des ensembles de niveau $L(\alpha)$ d'une fonction de répartition inconnue F sur $\mathbb{R}_+^2$, pour $\alpha \in (0,1)$.

Nous suivons ici une approche de type *plug-in*. Plus précisément, étant donné un estimateur F_n de F, nous estimons $L(\alpha)$ par

$$L_n(\alpha) = \{(x,y) \in \mathbb{R}_+^2 \,:\, F_n(x,y) \geq \alpha\}, \text{ pour } \alpha \in (0,1).$$

Cette problématique nous plonge dans la littérature, assez ample, de l'estimation des ensembles de niveau d'une fonction inconnue.

Au cours des dernières décennies, différents auteurs se sont intéressés à ce type de problèmes. L'essentiel des travaux de recherche sur l'estimation d'ensembles de niveau concerne la *fonction de densité*. On peut citer par exemple les travaux de Polonik (1995), Tsybakov (1997), Baíllo *et al.* (2001), Baíllo (2003), Cadre (2006), Biau *et al.* (2007), Rigollet et Vert (2009). L'estimation des ensemble de niveau d'une *fonction de régression dans des ensembles compacts* (moins étudiée dans la littérature) a été traitée, par exemple, par Cavalier (1997), Laloë (2009), Nowak et Willett (2007). Des approches alternatives fondées sur les propriétés géométriques d'un support compact des variables aléatoires ont été présentées par Hartigan (1987), Cuevas et Fraiman (1997), Cuevas et Rodríguez-Casal (2004). Baíllo et Cuevas considèrent le problème de l'estimation d'un support étoilé compact (voir Baíllo et Cuevas, 2001). Enfin le cas d'une *fonction générique inconnue sur un ensemble compact* a été développé par Rodríguez-Casal (2003), Cuevas *et al.* (2006).

L'estimation des ensembles de niveau est notamment utile dans la mise au point de tests de multimodalité (voir par exemple Müller et Sawitzki, 1991 ; Polonik, 1995), pour repérer un comportement anormal dans un système (Baíllo *et al.*, 2000 ; Baíllo *et al.*, 2001 ; Baíllo, 2003), pour l'étude de données sphériques (Cuevas *et al.*, 2006), ou encore pour la reconnaissance de formes (Cuevas et Rodríguez-Casal, 2004 ; Polonik, 1995). Cavalier montre par ailleurs que l'estimation des ensembles de niveau d'une fonction de régression peut être utile dans l'analyse des images et l'identification des régions à plus haute intensité (Cavalier, 1997). Biau et ses coauteurs utilisent un estimateur des ensembles de niveau d'une fonction de densité pour des applications en apprentissage

(Biau *et al.*, 2007). Cai et ses coauteurs proposent une application de l'estimation des ensembles de niveau d'une fonction de densité pour des problèmes de taux de change monétaires (Cai *et al.*, 2011).

En suivant la littérature citée ci-dessus nous proposons des résultats de consistance pour notre estimateur *plug-in* $L_n(c)$ par rapport à deux critères :

1. La *distance de Hausdorff* ;

2. Le *volume de la différence symétrique.*

Ces deux critères de "proximité physique" entre les ensembles sont parmi les plus utilisés dans la littérature d'estimation des ensembles de niveau (voir par exemple Laloë, 2009, partie 3 ; Cuevas *et al.*, 2006 ; Rodríguez-Casal, 2003). C'est donc ces critères que nous avons choisi dans ce chapitre. Pour une discussion sur les caractéristiques de la distance de Hausdorff le lecteur est renvoyé à la Section 4.2. Nous donnons dans la Figure 1.1 ci-dessous une représentation de la différence symétrique entre deux ensembles.

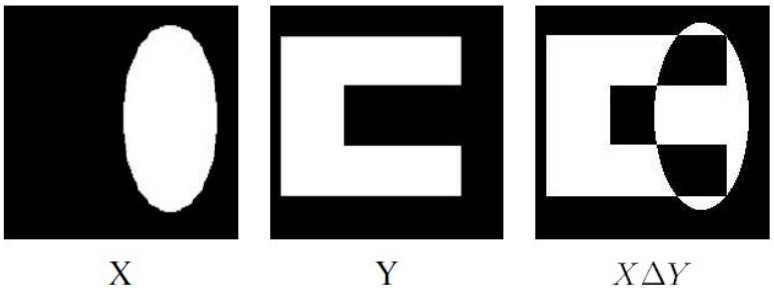

FIGURE 1.1 – Illustration de la différence symétrique $X \triangle Y$ entre deux ensembles.

Dans les travaux cités ci-dessus, une hypothèse communément requise est la compacité des ensembles de niveau à estimer. Cette hypothèse n'est pas raisonnable dans notre contexte. En effet notre problème est intrinsèquement un problème non-compact : à partir du moment où les variables aléatoires X et Y ne sont pas bornées aussi les ensembles de niveau de $F_{X,Y}$ ne sont pas bornés. De plus nous ne voulons pas borner nos risques bivariés (X, Y) car nous sommes intéressés par l'analyse et l'estimation de la mesure de risque $\mathrm{CTE}_\alpha(X, Y)$ qui évidement peut prendre des grandes valeurs. Cet aspect représente une caractéristique importante de notre travail et nécessite une attention particulière dans la formulation de nos résultats. Nous suivrons une approche en deux étapes :

1. Nous nous plaçons dans des ensembles localement tronqués et compacts afin de donner sens aux critères que nous avons choisis pour évaluer la qualité de notre estimation (distance de Hausdorff et volume de la différence symétrique).

2. Nous étendons ces ensembles pour pouvoir observer des grandes réalisations des pertes (X, Y) et donc de notre mesure $\mathrm{CTE}_\alpha(X, Y)$.

Pour ce faire, nous introduisons des ensembles de niveau tronqués :

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

avec T_n une suite croissante de valeurs positives (pour de plus amples détails voir Section 4.1).

Nous prouvons ensuite deux résultats de consistance pour $L_n(c)^{T_n}$:

- Par rapport à la distance de Hausdorff d_H :

  En supposant des conditions de différentiabilité sur la fonction bivariée F et en demandant que $\|F - F_n\|_\infty = \sup_{x \in \mathbb{R}_+^2} \mid F(x) - F_n(x) \mid \to 0, \quad p.s.$, on obtient

  $$d_H(\partial L(c)^{T_n}, \partial L_n(c)^{T_n}) = O(\|F - F_n\|_\infty), \qquad p.s.$$

  Pour une formulation rigoureuse de ce résultat voir la Section 4.2 (Théorème 4.2.1).

- Par rapport au volume de la différence symétrique d_λ :

  En supposant des conditions de différentiabilité sur la fonction bivariée F et en demandant que $v_n \int_{[0,T_n]^2} \mid F - F_n \mid^p \lambda(\mathrm{d}x) \underset{n\to\infty}{\overset{\mathbb{P}}{\to}} 0$, avec v_n une suite croissante de valeurs positives, $1 \leq p < \infty$ et λ la mesure de Lebesgue sur $\mathbb{R}^2$, on obtient

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

  avec $p_n = o\left(v_n^{\frac{1}{p+1}} / T_n^{\frac{p}{p+1}}\right)$. Pour une formulation rigoureuse de ce résultat voir la Section 4.3 (Théorème 4.3.1).

Estimation de $\mathrm{CTE}_\alpha(X,Y)$ avec une méthode *plug in*

Grâce à la consistance de l'estimateur $L_n(c)$ (Théorème 4.3.1), nous introduisons un estimateur de la mesure de risque $\mathrm{CTE}_\alpha(X,Y)$ et nous en étudions les propriétés asymptotiques.

Le point de départ est encore une procédure de troncature avec une suite T_n. On considère :

$$\mathrm{CTE}_\alpha^{T_n}(X,Y) = \mathbb{E}[(X,Y)|(X,Y) \in L(\alpha)^{T_n}],$$

et

$$\widehat{\mathrm{CTE}}_\alpha^{T_n}(X,Y) = \begin{pmatrix} \dfrac{\sum_{i=1}^n X_i 1_{\{(X_i,Y_i)\in L_n(\alpha)^{T_n}\}}}{\sum_{i=1}^n 1_{\{(X_i,Y_i)\in L_n(\alpha)^{T_n}\}}} \\ \dfrac{\sum_{i=1}^n Y_i 1_{\{(X_i,Y_i)\in L_n(\alpha)^{T_n}\}}}{\sum_{i=1}^n 1_{\{(X_i,Y_i)\in L_n(\alpha)^{T_n}\}}} \end{pmatrix},$$

où $L(\alpha)^{T_n}$ et $L_n(\alpha)^{T_n}$ sont les versions tronquées des ensembles de niveau respectivement théoriques et estimés, introduites ci-dessus.

Nous prouvons deux résultats de consistance pour $\widehat{\mathrm{CTE}}_\alpha^{T_n}(X,Y)$:

- Sous les hypothèses du Théorème 4.3.1, en supposant que les moments d'ordre deux de X et Y sont finis et que la fonction de densité $f_{(X,Y)}$ est $L^{1+r}(\lambda)$, avec $r > 0$, on obtient :

$$\beta_n \big| \operatorname{CTE}_\alpha^{T_n}(X,Y) - \widehat{\operatorname{CTE}}_\alpha^{T_n}(X,Y) \big| \underset{n\to\infty}{\overset{\mathbb{P}}{\to}} 0,$$

  avec $\beta_n = \min\{ p_n^{\frac{r}{2(1+r)}}, a_n \}$, et $a_n = o(\sqrt{n})$. Pour une formulation rigoureuse de ce résultat voir la Section 4.4 (Théorème 4.4.3).

- En particulier, si F_n est l'estimateur empirique de la fonction de répartition F, on obtient :

$$\beta_n \big| \operatorname{CTE}_\alpha^{T_n}(X,Y) - \widehat{\operatorname{CTE}}_\alpha^{T_n}(X,Y) \big| \underset{n\to\infty}{\overset{\mathbb{P}}{\to}} 0,$$

  avec $\beta_n = o\left(n^{\frac{r}{6(1+r)}} / T_n^{\frac{2r}{3(1+r)}} \right)$. Pour une formulation rigoureuse de ce résultat voir la Section 4.4 (Corollaire 4.4.4).

La possibilité d'obtenir un résultat du type :

$$w_n \big| \operatorname{CTE}_\alpha(X,Y) - \widehat{\operatorname{CTE}}_\alpha^{T_n}(X,Y) \big| \underset{n\to\infty}{\overset{\mathbb{P}}{\to}} 0, \tag{1.45}$$

avec w_n vitesse de convergence, est liée à la connaissance de la vitesse de convergence de $\big|\operatorname{CTE}_\alpha(X,Y) - \operatorname{CTE}_\alpha^{T_n}(X,Y)\big|$ et donc de la queue de la distribution $\mathbb{P}[X \geq T_n \text{ ou } Y \geq T_n]$, pour $n \to \infty$. En choisissant la suite de troncatures T_n de sorte que $\big| \operatorname{CTE}_\alpha^{T_n}(X,Y) - \widehat{\operatorname{CTE}}_\alpha^{T_n}(X,Y) \big|$ tende vers zéro plus lentement que $\mathbb{P}[X \geq T_n \text{ ou } Y \geq T_n]$ nous avons la possibilité d'étudier la convergence de l'erreur absolue en (1.45) (pour plus de détails voir la Remarque 11 dans la Section 4.4). Ce type de considérations nous donne des critères de choix pour la suite T_n. A ce propos, nous développons dans la Section 4.5.2 de nombreuses illustrations numériques sur des données simulées et réelles.

Notons que le problème du choix optimal de la suite T_n n'est pas abordé dans ce chapitre. Il s'agit cependant d'un problème intéressant qui fera l'objet d'un futur travail.

De plus nous remarquons que les Théorèmes 4.2.1, 4.3.1 et 4.4.3 sont énoncés pour un certain niveau α fixé. L'étude du comportement de notre estimateur $L_n(\alpha)$ pour $\alpha \to 1$ n'a pas été abordée dans ce Chapitre 4 et représente également une piste intéressante pour des développements futurs. Une illustration de cet aspect en utilisant des données simulées est donnée à la fin de la Section 4.5.2. Le lecteur intéressé est par ailleurs renvoyé à l'Annexe B.9.

Chapitre 2

Estimation de la queue d'une distribution bivariée : approche reposant sur les copules

Résumé

Ce chapitre est composé d'un article en anglais, intitulé *Estimating Bivariate Tail : a copula based approach*, coécrit avec Véronique Maume-Deschamps et Clémentine Prieur et actuellement soumis pour publication.

L'objet de cette première partie est l'estimation de la queue d'une distribution bivariée inconnue :

$$F(x,y) = \mathbb{P}[X \leq x, Y \leq y], \text{ pour } x \geq u_X,\ y \geq u_Y \text{ et } u_X,\ u_Y \text{ “assez” grands.}$$

Ce travail s'inscrit dans la littérature existante sur la méthode de dépassement de seuil (Peaks Over Threshold method, POT, voir par exemple McNeil 1997, 1999, dans un cadre unidimensionnel, ou Rootzén et Tajvidi, 2006, dans un cadre multidimensionnel) et en développe une extension bidimensionnelle. Notons que l'expression *“assez” grand* est un problème fondamental de la méthode POT dans le cadre unidimensionnel comme multidimensionnel (pour plus de détails voir les Section 2.5 et 2.6).

La structure de dépendance entre les marginales est modélisée par une copule. De plus la situation où toutes les composantes ont tendance à devenir grandes en même temps est modélisée par une copule de dépendance de la queue supérieure (Upper Tail Dependence Copula, UTDC) (voir Définition 2.2.1). Nous développons ensuite une version bivariée du Théorème de Pickands-Balkema-de Haan (voir Théorème 2.2.3).

Dans la Section 2.4.2 nous parcourons les étapes de la construction de notre estimateur. Dans la Section 2.5 nous partons de deux résultats unidimensionnels de Smith (voir Smith, 1987 ; Théorèmes 3.2 et 8.1) et nous réadaptons ces théorèmes de telle manière que le nombre d'excès au-delà du seuil soit aléatoire (voir Corollaires 2.5.1 et 2.5.2 dans la Section 2.5). En utilisant ces modifications nous démontrons des propriétés de convergence pour notre estimateur bivarié dans le cas de dépendance asymptotique entre les variables aléatoires (voir Théorème 2.6.1) ainsi que dans le cas d'indépendance asymptotique (voir Théorème 2.6.2). Enfin, nous proposons des études pratiques avec des données réelles, connues et utilisées dans la littérature des valeurs extrêmes multidimensionnelles (voir Section 2.7).

Présentation des approfondissements donnés en Annexe

En Annexe nous présentons des illustrations et d'autres études, inhérentes à notre procédure d'estimation de la queue, qui n'ont pas été insérées dans l'article.

- Dans la Section B.1 nous détaillons le lien qui existe entre le paramètre de dépendance η du modèle de Ledford et Tawn (1.20)-(1.21) et le paramètre θ dans le modèle de Juri et Wüthrich (2004) (voir condition (2.8)).

- Dans la Section B.2 nous proposons des simulations pour l'estimation du paramètre de dépendance θ dans des cas réels et simulés.

- Dans la Section B.3 nous présentons des illustrations du Théorème 2.5.3 et des Corollaires 2.5.1 et 2.5.2 (résultats unidimensionnels).

- Dans la Section B.4 nous présentons plus en détail l'argument utilisé dans les Corollaires 2.5.1 et 2.5.2, afin d'effectuer le déconditionnement des Théorèmes 3.2 et 8.1 de Smith (voir Smith, 1987).

2.1 Introduction

The univariate POT (Peaks-Over-Threshold) method is common for estimating extreme quantiles or tail distributions (see e.g. McNeil 1997, 1999 and references therein). A key idea of this method is that a distribution is in the domain of attraction of an extreme value distribution if and only if the distribution of excesses over high thresholds is asymptotically generalized Pareto (GPD) (e.g. Balkema and de Haan, 1974; Pickands, 1975):

$$V_{\xi,\sigma}(x) := \begin{cases} 1-\left(1-\frac{\xi x}{\sigma}\right)^{\frac{1}{\xi}}, & \text{if } \xi \neq 0,\ \sigma > 0, \\ 1-\mathrm{e}^{\frac{-x}{\sigma}}, & \text{if } \xi = 0,\ \sigma > 0, \end{cases} \tag{2.1}$$

and $x \geq 0$ for $\xi \leq 0$ or $0 \leq x < \frac{\sigma}{\xi}$ for $\xi > 0$. This univariate modeling is well understood, and has been discussed by Davison (1984), Davison and Smith (1990) and other papers of these authors.

In this paper, we are interested in the problem of fitting the joint distribution of bivariate observations exceeding high thresholds. To this end we develop a bivariate estimation procedure, mainly based on a version of the Pickands- Balkema-de Haan Theorem in dimension 2 (Theorem 2.2.3). This extension allows us to consider a two-dimensional structure of dependence between both continuous random components X and Y. This dependence is modeled via a copula C, which is supposed to be unknown. We recall here some classical bivariate threshold models, based on a characterization of the joint tail by Resnick (1987). Letting F denote the joint distribution of (Y_1, Y_2) with marginals F_j, $j = 1, 2$. Define $Z_j = -1/\log(F_j(Y_j))$, $j = 1, 2$, i.e. each Y_j is transformed to a unit Fréchet variable and $\mathbb{P}(Z_j \leq z) = \exp^{-1/z}$, for $0 < z < \infty$. Let F_* denote the joint distribution of (Z_1, Z_2), we have $F(y_1, y_2) = F_*(z_1, z_2)$. The assumption that F is in the maximum domain of attraction (MDA) of a bivariate extreme value distribution G is equivalent assuming F_* to be in the domain of attraction of a bivariate extreme value distribution G_*, where the marginals of G_* are unit Fréchet. The characterization of Resnick (1987) can be written as

$$\lim_{t\to\infty} \frac{\log(F_*(tz_1, tz_2))}{\log(F_*(t,t))} = \lim_{t\to\infty} \frac{1-F_*(tz_1, tz_2)}{1-F_*(t,t)} = \frac{\log(G_*(z_1, z_2))}{\log(G_*(1,1))}. \tag{2.2}$$

Equating the left and the right-hand terms for large t leads to the following model for the joint tail of F (see Ledford and Tawn, 1996):

$$\mathscr{F}_1(y_1, y_2) = \exp\{-l\left(-\log(F_{Y_1}(y_1)), -\log(F_{Y_2}(y_2))\right)\}, \tag{2.3}$$

for $y_j > u_j$, where u_j are high thresholds for the marginal distributions and l is the stable tail dependence function of the limiting extreme value distribution G_*. Then approximation (2.3) can be estimated by

$$\widehat{\mathscr{F}_1^*}(y_1, y_2) = \exp\{-\widehat{l}(-\log(\widehat{F}^*_{Y_1}(y_1)), -\log(\widehat{F}^*_{Y_2}(y_2)))\}, \tag{2.4}$$

for high values of y_1 and y_2, where $\widehat{F}^*_{Y_1}(y_1)$ (resp. $\widehat{F}^*_{Y_2}(y_2)$) is an estimator for the marginal tail of Y_1 (resp. Y_2). For instance $\widehat{F}^*_{Y_1}(y_1)$ (resp. $\widehat{F}^*_{Y_2}(y_2)$) comes from the univariate POT method described in Section 2.4.1. In (2.4) $\widehat{l}$ is an estimator of the

stable tail dependence function (see Drees and Huang, 1998; Draisma *et al.*, 2004; Einmahl *et al.*, 2008). For another approach, based on the estimation of the so-called univariate dependence function of Pickands (Pickands, 1981), see for instance Capéraà and Fougères (2000). Problems arise with both these bivariate techniques when (Y_1, Y_2) are asymptotically independent i.e.,

$$\lambda := \lim_{t \to 0} \mathbb{P}[F_{Y_1}^{-1}(Y_1) > 1 - t \,|\, F_{Y_2}^{-1}(Y_2) > 1 - t] = 0. \tag{2.5}$$

When the data exhibit positive or negative association that only gradually disappears at more and more extreme levels, these methods produce a significant bias. In order to overcome this problem, Ledford and Tawn (1996, 1997, 1998) introduced a model in which the tail dependence is characterized by a coefficient $\eta \in (0, 1]$. In these works the joint survival distribution function of a bivariate random vector (Z_1, Z_2) with unit Fréchet marginals is assumed to satisfy $\mathbb{P}[Z_1 > z, Z_2 > z] \sim L(z)\mathbb{P}[Z_1 > z]^{\frac{1}{\eta}}$, where L is a slowly varying function at infinity. Various methods to estimate this coefficient η are proposed in Peng (1999), Draisma *et al.* (2004), Beirlant *et al.* (2011). For some counter-examples of the Ledford and Tawn's model see Schlather (2001).
Contrarily to this approach, we propose a model based on regularity conditions of the copula and on the explicit description of the dependence structure in the joint tail (see condition in (2.8) in Proposition 2.2.2). The study of tail dependence from a distributional point of view by means of appropriate copulae has received attention in the past decade. The interested reader is referred to Juri and Wüthrich (2002, 2004), Wüthrich (2004), Charpentier and Juri (2006), Charpentier and Segers (2007), Javid (2009).
The general idea of our model is to decompose the estimation of $\mathbb{P}(X \leq x, Y \leq y)$, for x, y above some marginal thresholds u_X, u_Y, in the estimation of different bivariate regions. For the joint upper tail in $[u_X, x] \times [u_Y, y]$ we use the non parametric estimators coming from Theorem 2.2.3 (see Section 2.2). For the *lateral regions* $[-\infty, x] \times [-\infty, u_Y]$ and $[-\infty, u_X] \times [-\infty, y]$ we approximate the distribution function F using (2.3). The stability of our estimation compared to the one of $\widehat{\mathscr{F}_1^*}$ is analyzed on some real cases (Section 2.7) which have been studied in other papers (e.g. Beirlant *et al.*, 2011; Frees and Valdez 1998; Lescourret and Robert, 2006). Therefore our estimator, in a different way from the Ledford and Tawn's method, covers situations less restrictive than dependence or perfect independence above thresholds. Note also that our method is free from the pre-treatment of data because we can work directly with the original general samples without the transformation in Fréchet marginal distributions.
Finally, we recall that, in the past decade, bivariate extensions of the POT method via generalized Pareto distribution have been developed in a series of papers by Falk and Reiss (2005 and references therein) or in Reiss and Thomas (2007; Chapter 13). Recently a multivariate generalization is treated in Beirlant *et al.* (2004), Rootzén and Tajvidi (2006) and Michel (2008). The role of multivariate generalized Pareto distributions in the framework of extreme value theory is still under scrutiny. In contrast to the univariate case it is not intuitively clear, how exceedances over high thresholds are to be defined. Our paper makes a contribution to this part of recent literature. To the best of our knowledge the POT procedure we propose in this paper can not be directly deduced from POT methods proposed in works cited above. Moreover we provide an

estimation of bivariate tails such this type of estimation is not obtained in the papers cited above. However, some ingredients for a comparison are investigated in Theorem 4.2 in Juri and Wüthrich (2004).
The paper is organized as follows. In Section 2.2 we state an extension of the Pickands-Balkema-de Haan Theorem in the case of bivariate distributions with different marginals (Theorem 2.2.3). In Section 2.3 we provide a new non parametric estimator for the dependence structure of a bivariate random sample in the upper tail. In Section 2.4.2 we recall the POT procedure for univariate distributions and we use Theorem 2.2.3 in order to build a new estimator for the tail of the bivariate distribution. The study of the asymptotic properties of our estimator makes use of a convergence result in univariate case (Theorem 2.5.3) dealing with asymptotic behavior of the absolute error between the theoretical distribution function and its tail estimator. In Section 2.6 we present the consistency result of our estimator with its convergence rate both in the asymptotic dependent case (Theorem 2.6.1) and in the asymptotic independent one (Theorem 2.6.2). Examples with real data are presented in Section 2.7. Some auxiliary results and more technical proofs are postponed to the Appendix.

Remark 1 Assume we observe $X_1, \ldots, X_n$ i.i.d. with common distribution function F. If we fix some high threshold u, let N denote the number of excesses above u. In the following, two approaches will be considered. In the first one, we work conditionally on N. If n is the sample size and u_n the associated threshold, the number of excesses is m_n, with $\lim_{n\to\infty} m_n = \infty$ and $\lim_{n\to\infty} m_n/n = 0$. The second approach considers the number of excesses N_n as a binomial random variable (which is the case in the simulations), $N_n \sim \mathrm{Bi}\,(n, 1 - F(u_n))$ with $\lim_{n\to\infty} 1 - F(u_n) = 0$ and $\lim_{n\to\infty} n(1 - F(u_n)) = \infty$. Keeping in mind these considerations will be useful in the following (in particular in Section 2.5).

2.2 On the two-dimensional Pickands-Balkema-de Haan Theorem

A central one dimensional result in univariate tail estimation is the so-called Pickand-Balkema-de Haan Theorem. As our aim is the estimation of bivariate tails, we are interested in two-dimensional extensions of this theorem. Such a two dimensional generalization can be found in the literature (e.g. see Juri and Wüthrich, 2004; Wüthrich, 2004) with the assumption $F_X = F_Y$. Starting from Theorem 4.1 in Juri and Wüthrich (2004) and Theorem 3.1 in Charpentier and Juri (2006), we provide here a precise formulation and proof of a general bivariate Pickands-Balkema-de Haan Theorem (Theorem 2.2.3 below). We first introduce some notation and recall results from Juri and Wüthrich (2004) and Nelsen (1999), which we will need later.
We consider a 2-dimensional copula $C(u, v)$ and the associated survival copula $C^*(u, v)$. In a first time we assume that X and Y are uniformly distributed on $[0, 1]$. Let us fix a threshold $u \in [0, 1)$ such that $\mathbb{P}[X > u, Y > u] > 0$, i.e. such that $C^*(1 - u, 1 - u) > 0$. We consider the distribution of X and Y conditioned on $\{X > u, Y > u\}$:

$$\forall\, x \in [0,1], \quad \overline{F}_{X,u}(x) := \mathbb{P}[\, X \le x \,|\, X > u, Y > u\,] = 1 - \frac{C^*(1 - x \vee u, 1 - u)}{C^*(1 - u, 1 - u)}, \quad (2.6)$$

$$\forall\, y \in [0,1], \quad \overline{F}_{Y,u}(y) := \mathbb{P}[\,Y \le y \,|\, X > u, Y > u\,] = 1 - \frac{C^*(1-u, 1-y \vee u)}{C^*(1-u, 1-u)}. \quad (2.7)$$

Note that the continuity of the copula C implies that $\overline{F}_{X,u}$ and $\overline{F}_{Y,u}$ are also continuous.

Definition 2.2.1 Let X and Y be uniformly distributed on $[0,1]$. Assume that for a threshold $u \in [0,1)$, $C^*(1-u, 1-u) > 0$. We define the *Upper-tail dependence copula* (UTDC) at level $u \in [0,1)$ relative to the copula C by

$$C_u^{up}(x,y) := \mathbb{P}[\,X \le \overline{F}_{X,u}^{-1}(x), Y \le \overline{F}_{Y,u}^{-1}(y) \,|\, X > u, Y > u\,],$$

$\forall\,(x,y) \in [0,1]^2$, where $\overline{F}_{X,u}$, $\overline{F}_{Y,u}$ are given by (2.6)-(2.7).

Note that $\mathbb{P}[X \le x, Y \le y \,|\, X > u, Y > u]$ obviously defines a two-dimensional distribution function whose marginals are given by $\overline{F}_{X,u}$ and $\overline{F}_{Y,u}$. We remark that $C_u^{up}(x,y)$ is a copula and from the continuity of $\overline{F}_{X,u}$ and $\overline{F}_{Y,u}$ we obtain the uniqueness of C_u^{up}. Moreover, the asymptotic behavior of C_u^{up} for u around 1 describes the dependence structure of X, Y in their upper tails.
In order to provide an explicit form for $\lim_{u\to 1} C_u^{up}(x,y)$, we state Proposition 2.2.2 below, which is a modification of Theorem 3.1 in Charpentier and Juri (2006). More precisely we adapt Theorem 3.1 in Charpentier and Juri (2006) in the case of Upper-tail dependence copula, assuming that C satisfies suitable regularity condition under the direction $(1-u,\, 1-u)$ (see the limit in (2.8)). For comparisons we refer to Section 3 in Charpentier and Juri (2006).

Proposition 2.2.2 *Assume that* $\partial C^*(1-u,1-v)/\partial u < 0$ *and* $\partial C^*(1-u,1-v)/\partial v < 0$ *for all* $u, v \in [0,1)$. *Furthermore, assume that there is a positive function* G *such that*

$$\lim_{u\to 1} \frac{C^*(x\,(1-u),\, y\,(1-u))}{C^*(1-u,\, 1-u)} = G(x,y), \quad \textit{for all } x,\, y > 0. \quad (2.8)$$

Then for all $(x,y) \in [0,1]^2$

$$\lim_{u\to 1} C_u^{up}(x,y) = x + y - 1 + G(g_X^{-1}(1-x), g_Y^{-1}(1-y)) := C^{*\,G}(x,y), \quad (2.9)$$

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

$$G(x,y) = \begin{cases} x^\theta g_Y(\frac{y}{x}) & \textit{for } \frac{y}{x} \in [0,1], \\ y^\theta g_X(\frac{x}{y}) & \textit{for } \frac{y}{x} \in (1,\infty). \end{cases} \quad (2.10)$$

The proof of Proposition 2.2.2 is postponed to the Appendix. We adapt in our setting the proof of Theorem 3.1 by Charpentier and Juri (2006). Since $\partial C^*(1-u,1-v)/\partial u < 0$ and $\partial C^*(1-u,1-v)/\partial v < 0$ for all $u, v \in [0,1)$, we have $C^*(1-u,1-u) > 0$, for all $u \ge 0$, i.e. C_u^{up} is well defined for all $u \ge 0$. Then we ask that the joint survival

distribution function of X and Y, uniformly distributed on $[0,1]$, is strictly decreasing in each coordinate. As in Remark 3.2 in Charpentier and Juri (2006) one can prove that the convergence in (2.9) is uniform in $[0,1]^2$. From Proposition 2.2.2, functions G, g_X, and g_Y characterize the asymptotic behavior of the dependence structure for extremal events.

Remark 2
. We note that $C^{*G}(x,y)$ defined in (2.9) is the survival copula of the copula $C^G(x,y) := G(g_X^{-1}(x), g_Y^{-1}(y))$ and thus, in particular, is a copula (for more details see Section 3 in Charpentier and Juri, 2006).
. In the case of symmetric copula, i.e. $C(u,v) = C(v,u)$ for all u and v, the limit G in (2.8) is continuous, symmetric, with marginals $G(x,1) = G(1,x) = g(x)$, where $g : [0,\infty) \to [0,\infty)$ is a strictly increasing function and $g(x) = x^\theta g(1/x)$ for all $x \in (0,\infty)$ (for more details about properties of G in the symmetric case see Section 2 in Juri and Wüthrich, 2004).

In the univariate setting de Haan (1970) proves that $F \in MDA(H_\xi)$ is equivalent to the existence of a positive measurable function $a(\cdot)$ such that, for $1-\xi\, x > 0$ and $\xi \in \mathbb{R}$,

$$\lim_{u\to x_F} \frac{1-F(u+x\,a(u))}{1-F(u)} = \begin{cases} (1-\xi\,x)^{\frac{1}{\xi}}, & \text{if } \xi \neq 0, \\ \mathrm{e}^{-x}, & \text{if } \xi = 0, \end{cases} \tag{2.11}$$

where $x_F := \sup\{x \in \mathbb{R} \,|\, F(x) < 1\}$. It allows stating below a rigorous formulation of the two-dimensional Pickands-Balkema-de Haan Theorem in a general case.

Theorem 2.2.3 *Let X and Y be two continuous real valued random variables, with different marginal distributions, respectively F_X, F_Y, and copula C. Suppose that $F_X \in MDA(H_{\xi_1})$, $F_Y \in MDA(H_{\xi_2})$ and that C satisfies assumptions of Proposition 2.2.2. Then*

$$\sup_{\mathscr{A}} \Big| \mathbb{P}\big[X-u \le x, Y - F_Y^{-1}(F_X(u)) \le y \big| X > u, Y > F_Y^{-1}(F_X(u))\big] \\ - C^{*G}\big(1 - g_X(1 - V_{\xi_1,a_1(u)}(x)), 1 - g_Y(1 - V_{\xi_2,a_2(F_Y^{-1}(F_X(u)))}(y))\big)\Big| \xrightarrow[u\to x_{F_X}]{} 0, \tag{2.12}$$

where $V_{\xi_i,a_i(\cdot)}$ is the GPD with parameters $\xi_i, a_i(\cdot)$ defined in (2.1), $a_i(\cdot)$ is as in (2.11), for $i = 1,2$, $\mathscr{A} := \{(x,y) : 0 < x \le x_{F_X} - u, 0 < y \le x_{F_Y} - F_Y^{-1}(F_X(u))\}$, with $x_{F_X} := \sup\{x \in \mathbb{R} \,|\, F_X(x) < 1\}$, $x_{F_Y} := \sup\{y \in \mathbb{R} \,|\, F_Y(y) < 1\}$.

The proof of Theorem 2.2.3 is postponed to the Appendix.

2.3 Estimating dependence structure in the bivariate framework

It is well known that a bivariate distribution function F with continuous marginal distribution functions F_X, F_Y is said to have a *stable tail dependence function* l if for

$x \geq 0$ and $y \geq 0$ the following limit exists:

$$\lim_{t \to 0} \frac{1}{t} \mathbb{P}[1 - F_X(X) \leq tx \text{ or } 1 - F_Y(Y) \leq ty] := l(x, y), \tag{2.13}$$

or similarly

$$\lim_{t \to 0} \frac{1}{t} \mathbb{P}[1 - F_X(X) \leq tx, 1 - F_Y(Y) \leq ty] := R(x, y) = x + y - l(x, y), \tag{2.14}$$

see e.g. Huang (1992). If F_X, F_Y are in the maximum domain of attraction of two extreme value distributions H_X, H_Y and if (2.13) holds then F is in the domain of attraction of an extreme value distribution H with marginals H_X, H_Y and with copula determined by l. Furthermore (2.13) is equivalent to

$$\lim_{t \to 0} \frac{1}{t} (1 - C(1 - tx, 1 - ty)) = l(x, y), \text{ for } x \geq 0, \, y \geq 0. \tag{2.15}$$

Note that the upper tail dependence coefficient defined in (2.5) is such that $\lambda = R(1, 1)$. We introduce the non parametric estimators for l and R (see Einmahl *et al.*, 2006):

$$\widehat{l}(x, y) = \frac{1}{k_n} \sum_{i=1}^{n} 1_{\{R(X_i) > n - k_n x + 1 \text{ or } R(Y_i) > n - k_n y + 1\}}, \tag{2.16}$$

$$\widehat{R}(x, y) = \frac{1}{k_n} \sum_{i=1}^{n} 1_{\{R(X_i) > n - k_n x + 1, \, R(Y_i) > n - k_n y + 1\}}, \tag{2.17}$$

where $k_n \to \infty$, $k_n/n \to 0$ and $R(X_i)$ is the rank of X_i among $(X_1, \ldots, X_n)$, $R(Y_i)$ is the rank of Y_i among $(Y_1, \ldots, Y_n)$, for $i = 1, \ldots, n$.

2.3.1 Asymptotic dependent case

If X and Y are asymptotically dependent ($\lambda > 0$) we introduce an estimator for G, g_X and g_Y which will be used later to estimate the tail of the bivariate distribution function. Using (2.13)-(2.15), we write

$$g_X(x) = \frac{x + 1 - l(x, 1)}{2 - l(1, 1)} = \frac{R(x, 1)}{R(1, 1)}, \quad g_Y(y) = \frac{y + 1 - l(1, y)}{2 - l(1, 1)} = \frac{R(1, y)}{R(1, 1)},$$

$$G(x, y) = \frac{x + y - l(x, y)}{2 - l(1, 1)} = \frac{R(x, y)}{R(1, 1)}.$$

Using (2.10), as R is homogeneous of order one then $\theta = 1$. As $\eta \in (0, 1]$ in the Ledford and Tawn's model (see Ledford and Tawn 1996, 1997, 1998), θ describes the nature of the tail dependence, it does not depend on the marginal distribution functions.
In order to estimate g_X, g_Y and G, we use the non parametric estimator for R in (2.17) and we obtain

$$\widehat{g}_X(x) = \frac{\widehat{R}(x, 1)}{\widehat{R}(1, 1)}, \quad \widehat{g}_Y(x) = \frac{\widehat{R}(1, y)}{\widehat{R}(1, 1)}, \quad \text{and} \quad \widehat{G}(x, y) = \frac{\widehat{R}(x, y)}{\widehat{R}(1, 1)}. \tag{2.18}$$

Using (2.18) we get a non parametric estimator for θ, for $x > 0$

$$\widehat{\theta}_{\frac{y}{x}} = \begin{cases} \frac{\log(\widehat{G}(x,y)) - \log(\widehat{g}_Y(\frac{y}{x}))}{\log(x)} & \text{if } \frac{y}{x} \in [0,1], \\ \frac{\log(\widehat{G}(x,y)) - \log(\widehat{g}_X(\frac{x}{y}))}{\log(y)} & \text{if } \frac{y}{x} \in (1,\infty). \end{cases} \tag{2.19}$$

Following Remark 2, in the case of a symmetric copula, using $g_X(x) = g_Y(x) = g(x) = x^\theta g(1/x)$ for $x > 0$, we get

$$\widehat{\theta}_x = \frac{\log(\widehat{g}(x)) - \log(\widehat{g}(\frac{1}{x}))}{\log(x)}. \tag{2.20}$$

Using Theorem 2.2 in Einmahl *et al.* (2006) (see Theorem C in Appendix) we state the following consistency result for $\widehat{G}$, $\widehat{g}_X$ and $\widehat{g}_Y$:

Corollary 2.3.1 *Under assumptions of Theorem C if we have v_n such that $v_n/\sqrt{k_n} \to 0$, for $n \to \infty$, and $\lambda > 0$ we obtain*

$$v_n \sup_{0<x,y\leq 1} \left| \widehat{G}(x,y) - G(x,y) \right| \xrightarrow[n\to\infty]{\mathbb{P}} 0,$$

$$v_n \sup_{0<x\leq 1} \left| \widehat{g}_X(x) - g_X(x) \right| \xrightarrow[n\to\infty]{\mathbb{P}} 0, \; v_n \sup_{0<y\leq 1} \left| \widehat{g}_Y(y) - g_Y(y) \right| \xrightarrow[n\to\infty]{\mathbb{P}} 0,$$

with $\widehat{g}_X(x)$, $\widehat{g}_Y(y)$ and $\widehat{G}(x,y)$ as in (2.18), $k_n \to \infty$, $k_n/n \to 0$ and $k_n = o\big(n^{\frac{2\alpha}{1+2\alpha}}\big)$.

We now provide an illustration for two different copulae: survival Clayton and Logistic copulae. We remark that they are two symmetric copulae with $\lambda > 0$. In particular we observe the sensitivity of $\widehat{\theta}_x$ in (2.20) to the sequence k_n (Figure 2.1). We draw the mean curve on 100 samples of size $n = 1000$ (full line) and the empirical standard deviation (dashed lines).

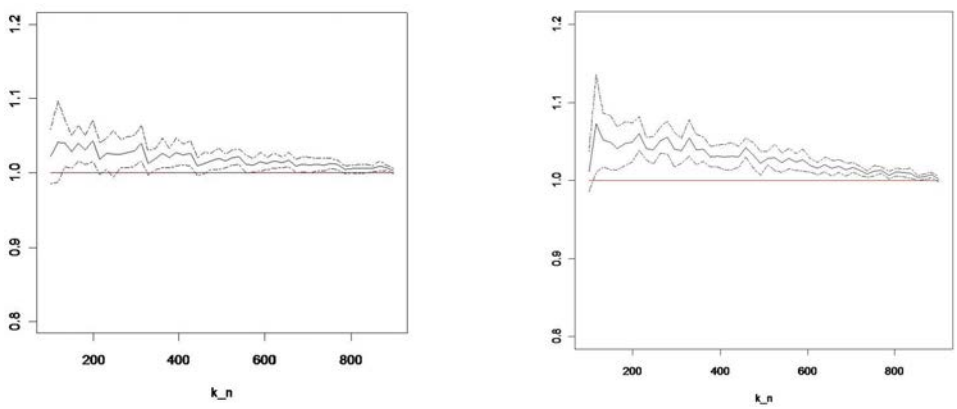

Figure 2.1: Estimator for θ, $(k, \widehat{\theta}_x)$ (left) $x = 0.07$, survival Clayton copula with parameter 1 (right) $x = 5$, Logistic copula with parameter 0.5

On simulations it seemed to us that for each value of x one could exhibit a range of values of k_n under which our estimate well behaved. In the following we fixe x for each

simulation and may vary k_n. The choice of k_n does not appear to be crucial for $\widehat{\theta}_x$. In Figure 2.2 the mean squared error for $\widehat{\theta}_x$ is calculated on 100 samples of size $n = 1000$.

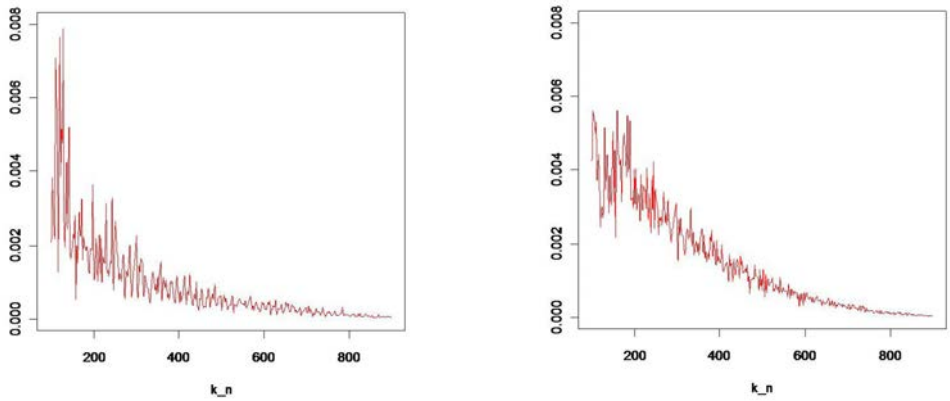

Figure 2.2: Mean squared error for $\widehat{\theta}_x$ (left) $x = 0.07$, survival Clayton copula with parameter 1 (right) $x = 5$, Logistic copula with parameter 0.5

2.3.2 Asymptotic independent case

We say that X and Y are asymptotically independent if $\lambda = R(1,1) = 0$. In terms of copula this means that $C(u,u) = 1 - 2(1-u) + o(1-u)$, for $u \to 1$. The problem, with respect to Section 2.3.1, is that $g_X(x) = \frac{R(x,1)}{R(1,1)}$ and $g_Y(y) = \frac{R(1,y)}{R(1,1)}$ have no sense as $\lambda = 0$ and $R(x,y) = x + y - l(x,y) = 0, \quad \forall\, x,\, y$.
We thus need to introduce a second-order refinement of condition in (2.8). More precisely, as in Draisma *et al.* (2004), we shall assume that:

$$\lim_{t\to 0} \frac{\frac{C^*(tx,ty)}{C^*(t,t)} - G(x,y)}{q_1(t)} := Q(x,y), \tag{2.21}$$

for all $x, y \geq 0$, $x+y > 0$, where q_1 is some positive function and Q is neither a constant nor a multiple of G. Moreover we assume that convergence in (2.21) is uniform on $\{x^2 + y^2 = 1\}$. Let $q(t) := \mathbb{P}[1 - F_X(X) < t, 1 - F_Y(Y) < t]$ and $q^{\leftarrow}$ its inverse function. Then, using (2.21), we obtain the following consistency result for $\widehat{G}$, $\widehat{g}_X$ and $\widehat{g}_Y$:

Proposition 2.3.2 *Suppose* (2.8) *and* (2.21) *hold. We assume* $\lim_{t\to 0} q(t)/t = \lambda = 0$. *Then, for a sequence* k_n *such that* $a_n := n\,q(k_n/n) \to \infty$ *(this implies* $k_n \to \infty$*),* $k_n/n \to 0$, $\sqrt{a_n}\, q_1(q^{\leftarrow}(a_n/n)) \to 0$, *it holds that*

$$\psi_n \sup_{0<x,y\leq 1} \left| \widehat{G}(x,y) - G(x,y) \right| \xrightarrow[n\to\infty]{\mathbb{P}} 0,$$

$$\psi_n \sup_{0<x\leq 1} \left|\widehat{g}_X(x) - g_X(x)\right| \xrightarrow[n\to\infty]{\mathbb{P}} 0, \quad \psi_n \sup_{0<y\leq 1} \left|\widehat{g}_Y(y) - g_Y(y)\right| \xrightarrow[n\to\infty]{\mathbb{P}} 0,$$

with $\psi_n << \sqrt{a_n}$, $\widehat{g}_X(x)$, $\widehat{g}_Y(y)$ *and* $\widehat{G}(x,y)$ *as in* (2.18).

Details of the proof are postponed to the Appendix. It is mainly based on Lemma 6.1 in Draisma *et al.* (2004).
In Proposition 2.3.3 below, by assuming some regularity properties on C, we deduce a specific form for G, g_X, g_Y and θ.

Proposition 2.3.3 *If $\lambda = 0$ and C is a twice continuously differentiable copula with the determinant of the Hessian matrix of C at $(1,1)$ different to zero, then*

$$\lim_{u\to 1} \frac{C^*(x(1-u),\, y(1-u))}{C^*(1-u, 1-u)} = x\, y, \qquad \forall\, x,\, y\, > 0,$$

$g_X(x) = g_Y(x) = x$ and $\theta = 2$.

Details of the proof will be omitted here. The main ingredient is the second-order development of copula C.
The assumptions of Proposition 2.3.3 are satisfied for a large class of asymptotic independent copulae: Ali Mikhail-Haq, Frank, Clayton with $a \geq 0$, Independent and Fairlie-Gumbel-Morgenstern copulae. An example of a non symmetric copula that satisfies the assumptions of Proposition 2.3.3 is $C(u,v) = x\,y + \frac{1}{9}(1-|2\,x-1|)\,(1-(2\,y-1)^2)$. This type of asymmetric copula is proposed in Benth and Kettler (2011) to model the evolution of price spread between electricity and gas prices.

We introduce some examples of asymptotic independent copulae that do not satisfy the assumptions of Proposition 2.3.3.
We consider the Ledford and Tawn's model (e.g. see Ledford and Tawn, 1996): $2\,u - 1 + C(1-u, 1-u) = (1-u)^{\frac{1}{\eta}} L(1-u)$, with L a slowly varying function at zero and $\eta \in (0,1]$. Then, for $\eta > 1/2$, $\lim_{u\to 1} \left(C(1,1) - C(1-u,1) - C(1,1-u) - C(1-u,1-u)\right)/(1-u)^2 = \infty$. Thus $\frac{\partial^2 C}{\partial u \partial v}$ does not exist at the point $(1,1)$. In particular this is the case of the Gaussian Copula with correlation parameter $\rho > 0$. However, from Theorem 5.3 in Juri and Wüthrich (2004), for a Gaussian Copula with $\mid \rho \mid < 1$ it holds that $\lim_{u\to 1} C_u^{up}(x,y) = x\,y$, for $(x,y) \in [0,1]^2$.
Let $C(u,v) = x\,y - \frac{1}{8}(1-|2\,x-1|)\,(1-(2\,y-1)^2)$, (for furthers details see Benth and Kettler, 2011). In this case $\frac{\partial^2 C}{\partial u \partial v}(1,1) = 0$. However we can calculate the limit in (2.8), and using (2.10) we obtain

$$G(x,y) = x\,y^2, \quad g_X(x) = x, \quad g_Y(y) = y^2, \quad \theta = 3.$$

We now provide an illustration for a Clayton copula. In particular we observe the sensitivity of $\widehat{\theta}_x$ in (2.20) to the sequence k_n (Figure 2.3). We draw the mean curve on 100 samples of size $n = 1000$ (full line) and the empirical standard deviation (dashed lines). Furthermore the mean squared error for $\widehat{\theta}_x$ is calculated on 100 samples of size $n = 1000$.

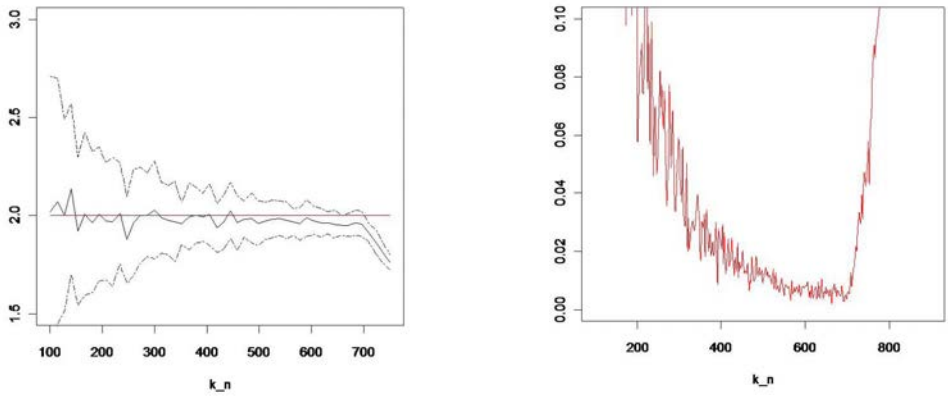

Figure 2.3: Clayton copula with parameter 0.05: (left) estimator for θ, $(k, \widehat{\theta}_x)$ with $x = 0.7$; (right) mean squared error for $\widehat{\theta}_x$ with $x = 0.7$.

2.4 Estimating tail distributions

2.4.1 Estimating the tail of univariate distributions

The estimation of the tail of bivariate distributions requires first the estimation of one-dimensional tail (McNeil, 1999; El-Aroui and Diebolt, 2002). Fix a threshold u and define $F_u(x) = \mathbb{P}[X \leq x \,|\, X > u]$. Let $X_1, X_2, \ldots$ be a sequence of i.i.d random variables with unknown distribution function F and $\widehat{F}_X(u)$ the empirical distribution function evaluated at the threshold u. Recall that the univariate tail may be estimated by

$$\widehat{F}^*(x) = (1 - \widehat{F}_X(u))V_{\widehat{\xi},\widehat{\sigma}}(x-u) + \widehat{F}_X(u), \quad \text{for } x > u, \tag{2.22}$$

where $\hat{\xi}$, $\hat{\sigma}$ are the maximum likelihood estimators (MLE) based on the excesses above u. Using (2.22) we get the estimator, proposed by Smith (1987)

$$1 - \widehat{F}^*(y) = \begin{cases} \frac{N}{n}\left(1 - \widehat{\xi}\,\frac{(y-u)}{\widehat{\sigma}}\right)^{\frac{1}{\widehat{\xi}}}, & \text{if } \widehat{\xi} \neq 0, \\ \frac{N}{n}\left(\mathrm{e}^{\frac{-(y-u)}{\widehat{\sigma}}}\right), & \text{if } \widehat{\xi} = 0, \end{cases} \tag{2.23}$$

with $u < y < \infty$ (if $\widehat{\xi} \leq 0$) or $u < y < \frac{\widehat{\sigma}}{\widehat{\xi}}$ (if $\widehat{\xi} > 0$) and N the random number of excesses above the threshold.

2.4.2 Estimating the tail of bivariate distributions

In this section we present the main construction of this paper. We propose indeed a POT procedure in order to estimate the two-dimensional distribution function $F(x, y)$. Asymptotic properties for this estimator are stated and proved in Section 2.6.
This construction generalizes the one-dimensional POT construction stated in Section 2.4.1. Let X and Y be two real valued random variables with different continuous marginal distributions F_X and F_Y. The structure of dependence between X and Y is represented by copula C.

Construction of the tail estimator:
Given a high threshold u and $u_Y := F_Y^{-1}(F_X(u))$, we introduce the distribution of excesses: $F_u(x,y) := \mathbb{P}[\,X-u \le x, Y-u_Y \le y \,|\, X>u, Y>u_Y\,]$. Using (2.3) for large value of u and $x>u$, $y>u_Y$, we can approximate $F(u,y)$ and $F(x,u_Y)$ as

$$F_1^*(u,\,y) = \mathrm{e}^{\{-l(-\log(F_X(u)),-\log(F_Y(y)))\}}, \tag{2.24}$$

$$F_2^*(x,u_Y) = \mathrm{e}^{\{-l(-\log(F_X(x)),-\log(F_Y(u_Y)))\}}, \tag{2.25}$$

where l is the stable tail dependence function defined by (2.13). We recall that behind approximations (2.24)-(2.25), in order to avoid significant bias, we suppose that the data structure is characterized by dependence (or perfect independence) in the lateral regions $[-\infty,x]\times[-\infty,u_Y]$ and $[-\infty,u_X]\times[-\infty,y]$.
From Theorem 2.2.3 we now know that, for u around x_F, we can approximate the distribution of excesses with $C^{*\,G}$. So we obtain, for $x>u$, $y>u_Y$,

$$\begin{aligned} F^*(x,y) = (\overline{F}(u,u_Y))\cdot C^{*\,G}\big(1-g_X(1-V_{\xi_X,\sigma_X}(x-u)), 1-g_Y(1-V_{\xi_Y,\sigma_Y}(y-u_Y))\big) \\ + F_1^*(u,y) + F_2^*(x,u_Y) - F(u,u_Y). \end{aligned} \tag{2.26}$$

Then, we estimate $F(u,u_Y)$ and $\overline{F}(u,u_Y)$ in (2.26) from the data $\{X_i,Y_i\}_{i=1,\dots,n}$, using the empirical distribution estimates

$$\widehat{F}(u,u_Y) = \frac{1}{n}\sum_{i=1}^n 1_{\{X_i\le u,\,Y_i\le u_Y\}}, \qquad \widehat{\overline{F}}(u,u_Y) = \frac{1}{n}\sum_{i=1}^n 1_{\{X_i>u,\,Y_i>u_Y\}}. \tag{2.27}$$

From (2.24)-(2.25) and using the non parametric estimator (2.16) we obtain

$$\widehat{F}_1^*(u,y) = \exp\{-\widehat{l}(-\log(\widehat{F}_X(u)),-\log(\widehat{F}_Y^*(y)))\}, \tag{2.28}$$

$$\widehat{F}_2^*(x,u_Y) = \exp\{-\widehat{l}(-\log(\widehat{F}_X^*(x)),-\log(\widehat{F}_Y(u_Y)))\}, \tag{2.29}$$

where $\widehat{F}_X(u)$ and $\widehat{F}_Y(u_Y)$ are the empirical univariate estimators evaluated at respective thresholds and $\widehat{F}_X^*(x)$ and $\widehat{F}_Y^*(y)$ are one-dimensional POT tail estimators of the marginal distribution functions, defined by (2.22). Now, using (2.27), (2.28) and (2.29), we can approximate $F^*(x,y)$, for $x>u$, $y>u_Y = F_Y^{-1}(F_X(u))$ and u large, by

$$\begin{aligned} \widetilde{F}^*(x,y) = \left(\frac{1}{n}\sum_{i=1}^n 1_{\{X_i>u,\,Y_i>u_Y\}}\right)\big(1-\widehat{g}_X(1-V_{\widehat{\xi}_X,\widehat{\sigma}_X}(x-u)) \\ -\widehat{g}_Y(1-V_{\widehat{\xi}_Y,\widehat{\sigma}_Y}(y-u_Y)) + \widehat{G}\big(1-V_{\widehat{\xi}_X,\widehat{\sigma}_X}(x-u), 1-V_{\widehat{\xi}_Y,\widehat{\sigma}_Y}(y-u_Y)\big)\big) \\ + \widehat{F}_1^*(u,y) + \widehat{F}_2^*(x,u_Y) - \frac{1}{n}\sum_{i=1}^n 1_{\{X_i\le u,\,Y_i\le u_Y\}}, \end{aligned} \tag{2.30}$$

where $\widehat{\xi}_X$, $\widehat{\sigma}_X$ (resp. $\widehat{\xi}_Y$, $\widehat{\sigma}_Y$) are MLE based on the excesses of X (resp. Y). Finally we remark that the second threshold in (2.30) depends on the unknown marginal distribution functions F_X and F_Y. Then, in order to compute in practice $\widetilde{F}^*(x,y)$, we

propose to estimate u_Y by $\widehat{u}_Y = \widehat{F}_Y^{-1}(\widehat{F}_X(u))$, with $\widehat{F}_X(u) = \frac{1}{n}\sum_{i=1}^n 1_{\{X_i \le u\}}$ and $\widehat{F}_Y^{-1}$ the empirical quantile function of Y.
So we obtain, from (2.30), the tail estimator for the two-dimensional distribution function for $x > u$ and $y > \widehat{u}_Y$:

$$\begin{aligned}\widehat{F}^*(x,y) = &\left(\frac{1}{n}\sum_{i=1}^n 1_{\{X_i > u,\, Y_i > \widehat{u}_Y\}}\right)\big(1 - \widehat{g}_X(1 - V_{\widehat{\xi}_X,\widehat{\sigma}_X}(x-u)) \\ &- \widehat{g}_Y(1 - V_{\widehat{\xi}_Y,\widehat{\sigma}_Y}(y - \widehat{u}_Y)) + \widehat{G}\big(1 - V_{\widehat{\xi}_X,\widehat{\sigma}_X}(x-u), 1 - V_{\widehat{\xi}_Y,\widehat{\sigma}_Y}(y-\widehat{u}_Y)\big)\big) \\ &+ \widehat{F}_1^*(u,y) + \widehat{F}_2^*(x,\widehat{u}_Y) - \frac{1}{n}\sum_{i=1}^n 1_{\{X_i \le u,\, Y_i \le \widehat{u}_Y\}}. \qquad (2.31)\end{aligned}$$

In the case with same marginal distributions we have a particular case of (2.30), with the same threshold u for X and Y, and we do not need to estimate the second threshold.

Remark 3 Note that $\widehat{F}^*(x,y)$ in (2.31), is only valid for $x > u$ and $y > \widehat{u}_Y$, when u is *large enough*. The expression large enough is a fundamental problem of the POT method. The choice of the threshold u is indeed a compromise: u has to be large for the GPD approximation to be valid, but if it is too large, the estimation of the parameters ξ_X, ξ_Y and σ_X, σ_Y will suffer from a lack of observations over the thresholds. The compromise will be explained in Sections 2.5 and 2.6.

2.5 Convergence results in the univariate case

In order to study asymptotic properties of our bivariate tail estimator we present in this section some slight modifications of one-dimensional convergence results in Smith (1987; Theorems 3.2 and 8.1). Incidentally we get asymptotic confidence intervals for the unknown theoretical univariate function $F(x)$, using Theorem 2.5.3. From now on we assume that the tail of F decays like a power function, i.e. is in the domain of attraction of Fréchet i.e. $\overline{F}(x) = x^{-\alpha}L(x)$ for some slowly varying function $L(x)$, with $\alpha > 0$. As in Smith (1987), Section 3, we shall assume that L satisfies the following condition

- SR2: $\frac{L(tx)}{L(x)} = 1 + k(t)\phi(x) + o(\phi(x)), \quad \text{as } x \to \infty, \quad \forall\, t > 0,$

where $\phi(x) > 0$ and $\phi(x) \to 0$ as $x \to \infty$. Let R_ρ be the set of $\rho-$regularly varying functions. Condition SR2 implies, excluding trivial cases, $\phi \in R_\rho$, for some $\rho \le 0$, and $k(t) = c\, h_\rho(t)$, with $h_\rho(t) = \int_1^t u^{\rho-1}\mathrm{d}u$; (for more details see Section 3 in Smith, 1987 or Goldie and Smith, 1987).
The study of the asymptotic properties of the maximum likelihood estimators of the scale and shape parameters of the generalized Pareto distribution in the POT method has received attention in the literature. For instance asymptotic normality of $\widehat{\xi}$ and $\widehat{\sigma}$, in the case of random threshold in the POT procedure is studied in depth in Drees *et al.* (2004). Smith (1987) examines a slightly different version of the MLE's that is based on the excesses over a deterministic threshold and on the second-order Condition SR2.

For details about the difference between these two approaches see, for instance, Remark 2.3 in Drees *et al.* (2004). In this paper we follow the approach proposed in Smith (1987). In particular Theorems 3.2. and 8.1. in Smith (1987) are written conditionally on $N = m_n$, with N denoting the number of excesses above the threshold. In practice we work with some deterministic threshold u and N is considered as random (see Remark 1 in Section 2.1). Therefore we give the version of Theorem 3.2 in Smith (1987) (resp. Theorem 8.1), Corollary 2.5.1 (resp. Corollary 2.5.2), unconditionally on N.

Corollary 2.5.1 *Suppose L satisfies SR2. Let n be the sample size and $u_n := \overline{f}(n)$ the threshold, such that $\overline{f}(n) \to \infty$, for $n \to \infty$. Let $N = N_n$ denote the random number of excesses of u_n. We define $\xi = -\alpha^{-1}$ and $\sigma_n = \overline{f}(n)\,\alpha^{-1}$. If*

$$n(1 - F(u_n)) \xrightarrow[n\to\infty]{} \infty, \tag{2.32}$$

$$\sqrt{n(1 - F(u_n))}\, c\, \phi(u_n) \xrightarrow[n\to\infty]{} \mu(\alpha - \rho), \tag{2.33}$$

then there exists, with probability 1, a local maximum $(\widehat{\sigma}_n, \widehat{\xi}_n)$ of the GPD log likelihood function, such that

$$\sqrt{N}\begin{pmatrix} \frac{\widehat{\sigma}_n}{\sigma_n} - 1 \\ \widehat{\xi}_n - \xi \end{pmatrix} \xrightarrow[n\to\infty]{d} \mathcal{N}\left(\begin{pmatrix} \frac{\mu(1-\xi)(1+2\xi\rho)}{1-\xi+\xi\rho} \\ \frac{\mu(1-\xi)\xi(1+\rho)}{1-\xi+\xi\rho} \end{pmatrix}; \begin{pmatrix} 2(1-\xi) & (1-\xi) \\ (1-\xi) & (1-\xi)^2 \end{pmatrix}\right).$$

Proof: If (2.32) and (2.33) hold then $N(n\,(1 - F(u_n)))^{-1} \xrightarrow[n\to\infty]{\mathbb{P}} 1$, and (3.2) in Smith (1987) holds in probability, i.e.

$$\frac{\sqrt{N}\, c\, \phi(u_n)}{\alpha - \rho} = \frac{\sqrt{N}\, c\, \phi(\overline{f}(n))}{\alpha - \rho} \xrightarrow[n\to\infty]{\mathbb{P}} \mu \in (-\infty, \infty).$$

Hence we conclude with a Skorohod-type construction of probability spaces (see Borkar, 1995, Theorem 2.2.2) on which (3.2) in Smith (1987) holds almost surely. □

Corollary 2.5.2 *Suppose all the assumptions of Corollary 2.5.1 are satisfied. Let n be the sample size, $u_n := \overline{f}(n) \to \infty$ and $z_n := f(n) \to \infty$, for $n \to \infty$, such that $(z_n)^{-s\,\rho\,\frac{\phi(u_n(z_n)^s)}{\phi(u_n)}} \to 1$, for $n \to \infty$ and $s \in [0, 1]$. Let $N = N_n$ denote the random number of excesses above u_n. If*

$$\frac{\log(z_n)}{\sqrt{n(1 - F(u_n))}} \xrightarrow[n\to\infty]{} 0, \tag{2.34}$$

then

$$\frac{\sqrt{N}}{\log(f(n))}\left[\frac{1 - \widehat{F}^*(\overline{f}(n)\, f(n))}{1 - F(\overline{f}(n)\, f(n))} - 1\right] \xrightarrow[n\to\infty]{d} \mathcal{N}(\nu, \tau^2),$$

where $\widehat{F}^*$ is as in (2.23), $\nu = 0$ if $\rho = 0$, $\nu = \frac{\mu\alpha(\alpha+1)(1+\rho)}{1+\alpha-\rho}$ for $\rho < 0$ and $\tau^2 = \alpha^2(1+\alpha)^2$.

Proof: If (2.32), (2.33) and (2.34) hold, then (8.7), (8.8) and (8.11) in Smith (1987) hold in probability, i.e

$$\frac{\log(z_n)}{\sqrt{N}} \xrightarrow[n\to\infty]{\mathbb{P}} 0, \qquad \frac{\sqrt{N}}{\log(z_n)}\left[\frac{N}{n\,(1-F(u_n))}-1\right] \xrightarrow[n\to\infty]{\mathbb{P}} 0.$$

We conclude as for Corollary 2.5.1. □
Note that, in simple cases, we often have $\phi(x) = x^\rho$; in which case $\frac{(z_n)^{-s\,\rho}\,\phi(u_n(z_n)^s)}{\phi(u_n)} \to 1$, for $n \to \infty$, is automatic satisfied. From Corollary 2.5.2 the following result can be obtained.

Theorem 2.5.3 *Assume that all the assumptions of Corollary 2.5.2 are satisfied. We use the same notation. If*

$$(z_n)^\alpha (n(1-F(u_n)))^{-1/2} \xrightarrow[n\to\infty]{} 0, \tag{2.35}$$

then

$$\frac{\sqrt{N}}{\log(f(n))\,\widehat{\overline{F}}(\overline{f}(n)\,f(n))}\left[F(\overline{f}(n)\,f(n)) - \widehat{F}^*(\overline{f}(n)\,f(n))\right] \xrightarrow[n\to\infty]{d} \mathscr{N}(\nu,\tau^2), \tag{2.36}$$

where $\widehat{\overline{F}}$ *is the univariate empirical survival function,* $\widehat{F}^*$ *is as in* (2.23), $\nu = 0$ *if* $\rho = 0$, $\nu = \frac{\mu\alpha(\alpha+1)(1+\rho)}{1+\alpha-\rho}$ *for* $\rho < 0$ *and* $\tau^2 = \alpha^2(1+\alpha)^2$.

The proof of Theorem 2.5.3 is postponed to the Appendix. As a byproduct, from (2.36) it is possible to construct in practice asymptotic confidence intervals for $F(\overline{f}(n)\,f(n))$.

2.6 Convergence results in the bivariate case

In this section we provide our main result: the consistency property of our bivariate tail estimator (2.31) with convergence rate. We consider:

Remark 4 Let n be the sample size. We choose, from Theorem 2.2.3,

$$u_{1\,n} := \overline{f}_1(n) \xrightarrow[n\to\infty]{} \infty \quad \text{threshold for } X,$$

$$u_{2\,n} := \overline{f}_2(n) = F_Y^{-1}(F_X(\overline{f}_1(n))), \xrightarrow[n\to\infty]{} \infty \quad \text{threshold for } Y.$$

Remark 5 As in Section 2.4.2 in the following we propose to estimate the second threshold $\overline{f}_2(n)$ by $\widehat{\overline{f}}_2(n) := \widehat{F}_Y^{-1}(\widehat{F}_X(\overline{f}_1(n)))$, with $\widehat{F}_X(\overline{f}_1(n)) = \frac{1}{n}\sum_{i=1}^n 1_{\{X_i \le \overline{f}_1(n)\}}$ and $\widehat{F}_Y^{-1}$ the empirical quantile function of Y.

In the following we state and prove separately our consistency result in the asymptotic dependent case (Theorem 2.6.1) and in the asymptotic independent one (Theorem 2.6.2).

2.6.1 Asymptotic dependent case

The proof of Theorem 2.6.1 below, makes use of a result by Einmahl *et al.* (2006) which specifies the asymptotic behavior of $\hat{l}(x,y)$ uniformly in $0 \leq x,\, y \leq 1$ and provides a convergence rate (see Theorem C in Appendix). More precisely in the asymptotic dependent case, using (2.18) and applying Corollary 2.3.1, we obtain the following main result:

Theorem 2.6.1 *Suppose F_X and F_Y belong to the maximum domain of attraction of Fréchet, L_X, L_Y satisfy Condition SR2. Assume that $\lambda > 0$ and that the assumptions of Theorem 2.2.3 and Corollary 2.3.1 are satisfied. If sequences $f_1(n)$, $f_2(n)$, $\overline{f}_1(n)$, $\overline{f}_2(n)$, defined by Remark 4, satisfy conditions of Theorem 2.5.3 then*

$$\left|\sqrt{k_n}(F^*(x_n,y_n) - \widetilde{F}^*(x_n,y_n))\right| \xrightarrow[n\to\infty]{\mathbb{P}} 0, \tag{2.37}$$

with $x_n = \overline{f}_1(n)f_1(n)$, $y_n = \overline{f}_2(n)f_2(n)$. Moreover if $\widehat{\overline{f}}_2(n)$ satisfies conditions of Theorem 2.5.3 in probability then

$$\left|\sqrt{k_n}(F^*(x_n,\widehat{y}_n) - \widehat{F}^*(x_n,\widehat{y}_n))\right| \xrightarrow[n\to\infty]{\mathbb{P}} 0, \tag{2.38}$$

with $\widehat{y}_n = \widehat{\overline{f}}_2(n)f_2(n)$. In (2.37)-(2.38) we have $k_n \to \infty$, $k_n/n \to 0$, $k_n/N_X \xrightarrow{\mathbb{P}} 0$, $k_n/N_Y \xrightarrow{\mathbb{P}} 0$, $k_n = o(n^{\frac{2\alpha}{1+2\alpha}})$, $\alpha > 0$.

The proof of Theorem 2.6.1 is postponed to the Appendix.

Remark 6 Let us study, on a class of examples, the assumption of Theorem 2.6.1. First if we suppose that the function $\phi(x)$ in Condition SR2 (Section 2.5) has the general form $\phi(x) = x^\rho$, with $\rho \leq 0$, then

$$(z_n)^{-s\,\rho}\,\phi(\widehat{\overline{f}}_2(n)(z_n)^s)\,/\,\phi(\widehat{\overline{f}}_2(n)) = 1, \quad \forall\ s\ \in [0,1].$$

For instance this is the case of Burr or Fréchet univariate distributions. Furthermore if we assume that F_Y belongs to the maximum domain of attraction of Fréchet (i.e. $\overline{F}_Y(y) = y^{-\alpha}L(y)$), F_Y has positive density $f_Y \in R_{-1-\alpha}$ and $\overline{f}_2(n)$ satisfies conditions in (2.32)-(2.35) then also the estimated second threshold $\widehat{\overline{f}}_2(n)$ satisfies, in probability, conditions in (2.32)-(2.35).
We remark indeed that $\widehat{F}_X(\overline{f}_1(n))$ is a high quantile within the sample (see Embrechts *et al.*, 1997), i.e. $\widehat{F}_X(\overline{f}_1(n)) \xrightarrow[n\to\infty]{\mathbb{P}} 1$ and $n(1-\widehat{F}_X(\overline{f}_1(n))) \xrightarrow[n\to\infty]{\mathbb{P}} \infty$. Then, using Theorem 6.4.14 in Embrechts *et al.* (1997) and a Skorohod-type construction of probability spaces we obtain $\widehat{\overline{f}}_2(n)\ \left(\overline{f}_2(n)\right)^{-1} \xrightarrow[n\to\infty]{\mathbb{P}} 1$.
Furthermore, using Condition SR2,
$\frac{\overline{F}_Y(\widehat{\overline{f}}_2(n))}{\overline{F}_Y(\overline{f}_2(n))} = \frac{\widehat{\overline{f}}_2(n)^{-\alpha}}{\overline{f}_2(n)^{-\alpha}}\frac{L(\widehat{\overline{f}}_2(n))}{L(\overline{f}_2(n))} = \frac{\widehat{\overline{f}}_2(n)^{-\alpha}}{\overline{f}_2(n)^{-\alpha}}\left[1 + k\left(\frac{\widehat{\overline{f}}_2(n)}{\overline{f}_2(n)}\right)\phi(\overline{f}_2(n)) + o(\phi(\overline{f}_2(n)))\right]$. Using properties of k and ϕ (see Section 2.5) we obtain $\frac{\overline{F}_Y(\widehat{\overline{f}}_2(n))}{\overline{F}_Y(\overline{f}_2(n))} \xrightarrow[n\to\infty]{\mathbb{P}} 1$.
Then $\widehat{\overline{f}}_2(n)$ satisfies, in probability, condition in (2.32):

$$n(1-F_Y(\widehat{\overline{f}}_2(n))) = \frac{\overline{F}_Y(\widehat{\overline{f}}_2(n))}{\overline{F}_Y(\overline{f}_2(n))}\, n(1-F_Y(\overline{f}_2(n))) \xrightarrow[n\to\infty]{\mathbb{P}} \infty.$$

The proof for conditions in (2.33)-(2.35) is completely analogue to that of condition in (2.32).

2.6.2 Asymptotic independent case

As noticed in Section 2.3.2 in the asymptotic independent case we need to introduce a second-order refinement of condition in (2.8). Under condition in (2.21) we obtain the following main result:

Theorem 2.6.2 *Suppose F_X and F_Y belong to the maximum domain of attraction of Fréchet, L_X, L_Y satisfy Condition SR2. Assume that the assumptions of Theorem 2.2.3, Proposition 2.3.2 and Corollary 2.7.1 are satisfied. If sequences $f_1(n)$, $f_2(n)$, $\overline{f}_1(n)$, $\overline{f}_2(n)$, defined by Remark 4, satisfy conditions of Theorem 2.5.3 then*

$$\left|\sqrt{a_n}\,(F^*(x_n,y_n)-\widetilde{F}^*(x_n,y_n))\right| \xrightarrow[n\to\infty]{\mathbb{P}} 0, \tag{2.39}$$

where $x_n = \overline{f}_1(n)f_1(n)$, $y_n = \overline{f}_2(n)f_2(n)$. Moreover if $\widehat{\overline{f}}_2(n)$ satisfies conditions of Theorem 2.5.3 in probability then

$$\left|\sqrt{a_n}\,(F^*(x_n,\widehat{y}_n)-\widehat{F}^*(x_n,\widehat{y}_n))\right| \xrightarrow[n\to\infty]{\mathbb{P}} 0, \tag{2.40}$$

with $\widehat{y}_n = \widehat{\overline{f}}_2(n)f_2(n)$. In (2.39)-(2.40) we have $a_n = n\,q(k_n/n) \to \infty$ (this implies $k_n \to \infty$), $k_n/n \to 0$, $\sqrt{a_n}\,q_1(q^{\leftarrow}(a_n/n)) \to 0$, $k_n/N_X \xrightarrow{\mathbb{P}} 0$, $k_n/N_Y \xrightarrow{\mathbb{P}} 0$, and $k_n = o(n^{\frac{2\alpha}{1+2\alpha}})$, for some $\alpha > 0$.

The proof of Theorem 2.6.2 is postponed to the Appendix.

2.7 Illustrations with real data

In this section we present four real cases (see Figures 2.4-2.5) for which we estimate bivariate tail probabilities to illustrate the finite sample properties of our estimator. We analyze the stability of our estimation compared to the one of $\widehat{\mathscr{F}}_1^*$, as well the estimation of parameter θ of these real cases.

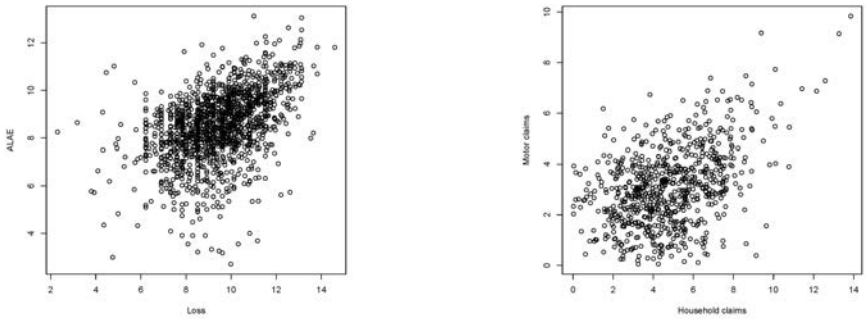

Figure 2.4: Logarithmic scale (left) ALAE versus Loss; (right) Storm damages.

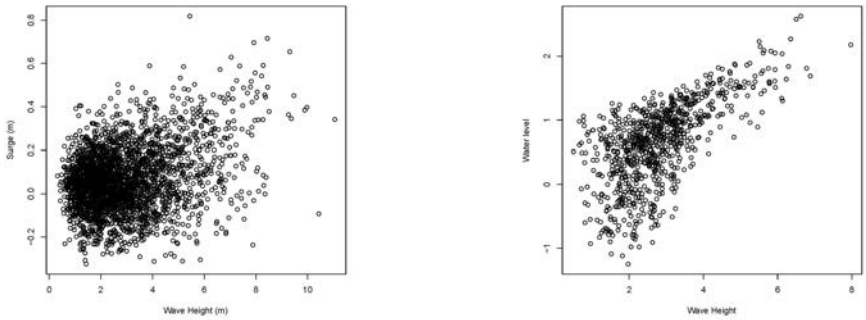

Figure 2.5: (left) Wave Height (m) versus Surge (m); (right) Wave heights versus Water level.

We consider the **Loss-ALAE data** (for details see Frees and Valdez, 1998). Each claim consists of an indemnity payment (the loss, X) and an allocated loss adjustment expense (ALAE, Y). We estimate $F(2.10^5, 10^5)$. The empirical probability, defined by (2.27), is 0.9506667 and the survival empirical probability is 0.006 (for a comparison using the Ledford and Tawn's model see Beirlant *et al.*, 2011). Figure 2.6 shows the sensitivity of $\widehat{\theta}$ and $\widehat{F}^*$ to the sequence k_n and provides a comparison with the estimator $\widehat{\mathscr{F}}_1^*$.

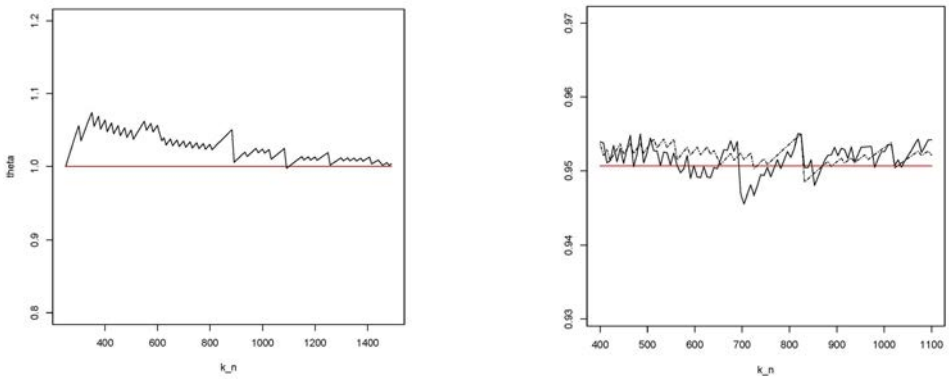

Figure 2.6: (left) $\widehat{\theta}_{0.04}$; (right) $\widehat{F}^*(2.10^5, 10^5)$ (full line), $\widehat{\mathscr{F}_1^*}(2.10^5, 10^5)$ (dashed line), with the empirical probability indicated with a horizontal line; Loss-ALAE data.

We now consider an **example from storm insurance**: aggregate claims of motor policies (Y) and aggregate claims of household policies (X) from a French insurance portfolio for 736 storm events (for a detailed description see Lescourret and Robert, 2006). We estimate $F(8000, 950)$. The empirical probability is 0.96875 and the survival empirical probability is 0.014. The stability of our estimation compared to the one of $\widehat{\mathscr{F}_1^*}$, as well the estimation of parameter θ are presented in Figure 2.7.

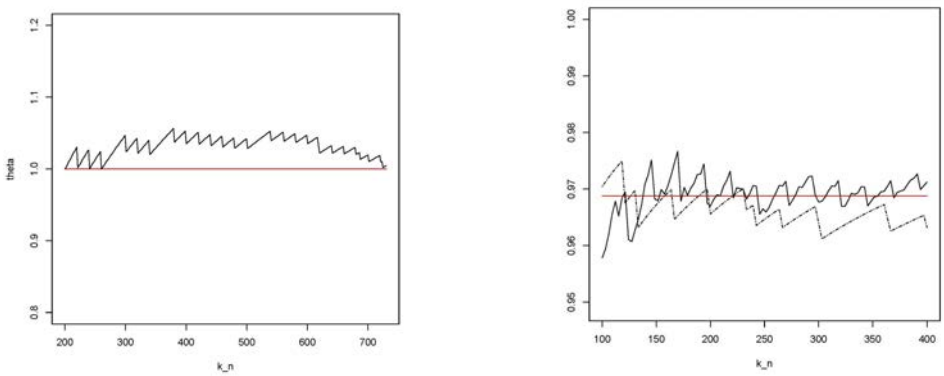

Figure 2.7: (left) $\widehat{\theta}_{0.05}$; (right) $\widehat{F}^*(8.10^3, 950)$ (full line), $\widehat{\mathscr{F}_1^*}(8.10^3, 950)$ (dashed line), with the empirical probability indicated with a horizontal line; Storm insurance data.

We study the **wave surge data** comprising 2894 bivariate events that occurred during $1971-1977$ in Cornwall (England) (for details see Coles and Tawn, 1994 or Ramos and Ledford, 2009). We estimate $F(8.32, 0.51)$. The empirical probability is 0.9903 and the survival empirical probability is 0.00069. The sensitivity of $\widehat{\theta}$ and $\widehat{F}^*$ to the sequence k_n and the estimation of θ are presented in Figure 2.8.

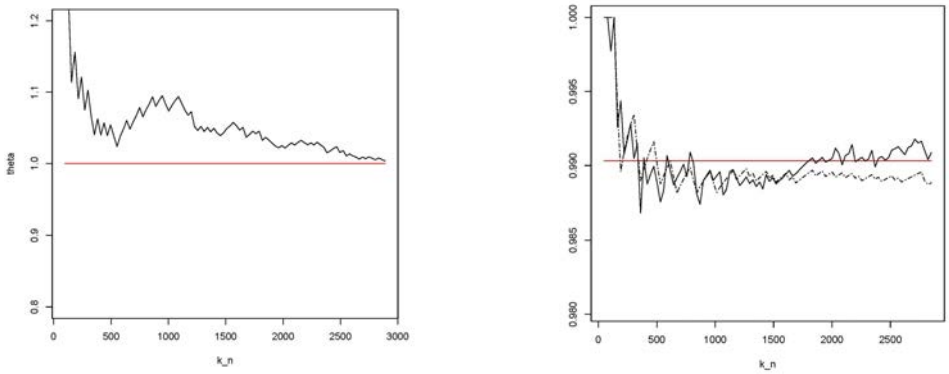

Figure 2.8: (left) $\widehat{\theta}_{0.02}$; (right) $\widehat{F}^*(8.32, 0.51)$ (full line), $\widehat{\mathscr{F}}_1^*(8.32, 0.51)$ (dashed line), with the empirical probability indicated with a horizontal line; Wave-Surge data.

Finally we analyze the **Wave height versus Water level data**, recorded during 828 storm events spread over 13 years in front of the Dutch coast near the town of Petten (for details see Draisma *et al.*, 2004). We estimate $F(5.93, 1.87)$. The empirical probability is 0.97584 and the survival empirical probability is 0.00604. The sensitivity of $\widehat{\theta}$ and $\widehat{F}^*$ to the sequence k_n and the estimation of θ are presented in Figure 2.9. From Draisma *et al.* (2004) it seems that the coefficient η of Ledford and Tawn's model is smaller than 1, then it is plausible to assume asymptotic independence between the wave heights and the water level. Analogously, in our model the estimated parameter $\widehat{\theta}$ is greater than one (see Figure 2.9).

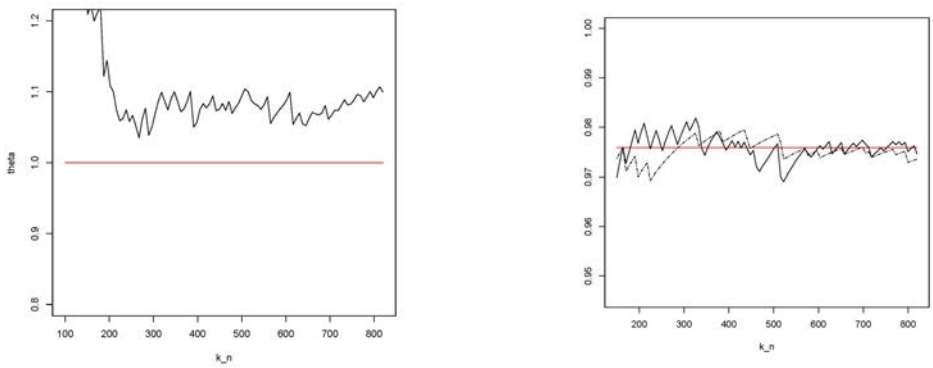

Figure 2.9: (left) $\widehat{\theta}_{\frac{0.1}{0.11}} = \widehat{\theta}_{0.91}$ as in (2.19); (right) $\widehat{F}^*(5.93, 1.87)$ (full line), $\widehat{\mathscr{F}}_1^*(5.93, 1.87)$ (dashed line), with the empirical probability indicated with a horizontal line; Wave height-Water level data.

Appendix: proofs and auxiliary results

Proof [Proposition 2.2.2]:
We know that
$C_u^{up}(x,y) = 1 - \frac{C^*(1-\overline{F}_{X,u}^{-1}(x),1-u)}{C^*(1-u,1-u)} - \frac{C^*(1-u,1-\overline{F}_{Y,u}^{-1}(y))}{C^*(1-u,1-u)} + \frac{C^*(1-\overline{F}_{X,u}^{-1}(x),1-\overline{F}_{Y,u}^{-1}(y))}{C^*(1-u,1-u)}$.

$$\text{Then} \quad \lim_{u\to 1} C_u^{up}(x,y) = \lim_{u\to 1}\left[x + y - 1 + \frac{C^*(1-\overline{F}_{X,u}^{-1}(x),\, 1-\overline{F}_{Y,u}^{-1}(y))}{C^*(1-u,\, 1-u)}\right]. \tag{2.41}$$

We introduce the following lemma.

Lemma A *(Charpentier and Juri, 2006; Lemma 6.1)*
Suppose that the random vectors (X_n, Y_n) have continuous, strictly increasing marginals and are such that $\lim_{n\to\infty}(X_n, Y_n) = (X, Y)$ in distribution for some (X, Y). Then

$$\lim_{n\to\infty} ||C_n - C||_\infty = 0,$$

where C_n and C denote the copulas of (X_n, Y_n) and (X, Y), respectively.

Let (X, Y) have distribution function C. Note that

$$\mathbb{P}\left[X > x(1-u) \,|\, X > u, Y > u\right] = \frac{C^*(1-x\,(1-u),\, 1-u)}{C^*(1-u,\, 1-u)}, \tag{2.42}$$

$$\mathbb{P}\left[Y > y(1-u) \,|\, X > u, Y > u\right] = \frac{C^*(1-u,\, 1-y\,(1-u))}{C^*(1-u,\, 1-u)}, \tag{2.43}$$

$$\mathbb{P}\left[X > x(1-u),\, Y > y(1-u) \,|\, X > u, Y > u\right] = \frac{C^*(1-x\,(1-u),\, 1-y\,(1-u))}{C^*(1-u,\, 1-u)}. \tag{2.44}$$

The distributions in (2.42)-(2.44) are respectively the survival conditional distributions of $\frac{X}{1-u}$, $\frac{Y}{1-u}$ and $\left(\frac{X}{1-u}, \frac{Y}{1-u}\right)$, given that $X > u$ and $Y > u$. Since $\partial C^*(1-u, 1-v)/\partial u < 0$ and $\partial C^*(1-u, 1-v)/\partial v < 0$, for all $u, v \in [0, 1)$, it follows that the distributions in (2.42)-(2.43) are continuous and strictly increasing.
By hypothesis, we have

$$\lim_{u\to 1} \frac{C^*(x\,(1-u),\, y\,(1-u))}{C^*(1-u,\, 1-u)} = G(x,y), \quad \text{for all } x,\, y > 0, \tag{2.45}$$

implying that the expressions in (2.42)-(2.43) respectively converge to $g_X(1-x)$ and $g_Y(1-y)$ as $u \to 1$, with $g_X(x) := G(x, 1)$, $g_Y(y) := G(1, y)$.
Since copulas are invariant under strictly increasing transformations of the underlying variables, it follows that we can use the conditional distributions in (2.42)-(2.43), instead of $\overline{F}_{X,u}$ and $\overline{F}_{Y,u}$, to construct $C_u^{up}(x,y)$. Then, from (2.41) and using Lemma

A, we have

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

As in the proof of Theorem 3.1 in Charpentier and Juri (2006), the limit in (2.45) implies that there is a $\theta > 0$ such that G is homogeneous of order θ, i.e. for all $t,\, x,\, y > 0$,

$$G(t\,x,\, t\,y) = t^\theta\, G(x,\, y). \tag{2.46}$$

By a discussion of the general solution of functional (2.46) we obtain the explicit form of G:

$$G(x,y) = \begin{cases} x^\theta g_Y(\frac{y}{x}) & \text{for } \frac{y}{x} \in [0,1], \\ y^\theta g_X(\frac{x}{y}) & \text{for } \frac{x}{y} \in (1,\infty). \end{cases}$$

For this part of the proof we refer the interested reader to Theorem 3.1 in Charpentier and Juri (2006). □

Proof [Theorem 2.2.3] :
From (2.11) we obtain the existence of $a_1(\cdot)$ and $a_2(\cdot)$ such that, for $p := u + x\,a_1(u)$ and $q := u_Y + y\,a_2(u_Y)$

$$V_{\xi_1,1}(x) = \lim_{u\to x_{F_X}} 1 - \frac{1-F_X(p)}{1-F_X(u)} = \lim_{u\to x_{F_X}} \mathbb{P}[X \le p | X > u], \tag{2.47}$$

$$V_{\xi_2,1}(y) = \lim_{u_Y\to x_{F_Y}} 1 - \frac{1-F_Y(q)}{1-F_Y(u_Y)} = \lim_{u_Y\to x_{F_Y}} \mathbb{P}[Y \le q | Y > u_Y]. \tag{2.48}$$

From $Y \overset{d}{=} F_Y^{-1}(F_X(X))$, we take $u_Y = F_Y^{-1}(F_X(u))$ and from (2.47)-(2.48), as $u \to x_{F_X}$, we have

$$1-(1-V_{\xi_1,1}(x))(1-F_X(u)) \sim F_X(u + x\,a_1(u)),$$
$$1-(1-V_{\xi_2,1}(y))(1-F_Y(F_Y^{-1}(F_X(u)))) \sim F_Y(F_Y^{-1}(F_X(u)) + y\,a_2(F_Y^{-1}(F_X(u)))).$$

Then

$$\lim_{u\to x_{F_X}} \mathbb{P}\left[\frac{X-u}{a_1(u)} > x, \frac{Y-F_Y^{-1}(F_X(u))}{a_2(F_Y^{-1}(F_X(u)))} > y \,\middle|\, X > u, Y > F_Y^{-1}(F_X(u))\right]$$
$$= \lim_{u\to x_{F_X}} \frac{C^*\big(1-F_X(u+x\,a_1(u)), 1-F_Y(F_Y^{-1}(F_X(u)) + y\,a_2(F_Y^{-1}(F_X(u))))\big)}{C^*\big(1-F_X(u), 1-F_Y(F_Y^{-1}(F_X(u)))\big)}$$
$$= \lim_{u\to x_{F_X}} \frac{C^*\big((1-V_{\xi_1,1}(x))(1-F_X(u)), (1-V_{\xi_2,1}(y))(1-F_Y(F_Y^{-1}(F_X(u))))\big)}{C^*\big(1-F_X(u), 1-F_Y(F_Y^{-1}(F_X(u)))\big)}$$
$$= \lim_{\nu\to 1} \frac{C^*\big((1-V_{\xi_1,1}(x))(1-\nu), (1-V_{\xi_2,1}(y))(1-\nu)\big)}{C^*\big(1-\nu, 1-\nu\big)}. \tag{2.49}$$

Now, if $h := (1-\xi_1\,x)^{\frac{1}{\xi_1}}$, $\xi_1 \neq 0$ or if $h := \mathrm{e}^{-x}$, $\xi_1 = 0$ then $1 - V_{\xi_1,1}(x) = V_{1,1}(h)$. So (2.49) becomes $\lim_{\nu\to 1} C^*\big(V_{1,1}(h)(1-\nu), V_{1,1}(w)(1-\nu)\big)/C^*\big(1-\nu, 1-\nu\big)$.
As C satisfies hypotheses of Proposition 2.2.2, the above limit is equal to $G(V_{1,1}(h), V_{1,1}(w)) = G(1 - V_{\xi_1,1}(x), 1 - V_{\xi_2,1}(y))$. Then

$$\lim_{u\to x_{F_X}} \mathbb{P}\left[\frac{X-u}{a_1(u)} \leq x,\ \frac{Y - F_Y^{-1}(F_X(u))}{a_2(F_Y^{-1}(F_X(u)))} \leq y \,\middle|\, X > u, Y > F_Y^{-1}(F_X(u))\right]$$
$$= C^{*\,G}\big(1 - g_X(1 - V_{\xi_1,1}(x)), 1 - g_Y(1 - V_{\xi_2,1}(y))\big). \quad (2.50)$$

Since the limit is a continuous distribution function (as $C^{*\,G}$, g and the GPD are), (2.50) can be strengthened to uniform convergence (see e.g. Embrechts *et al.* 1997, p. 552). Then (2.12) follows. □

Proof [Theorem 2.5.3]:
To begin with, we work conditionally on $N_n = m_n$. First we have to prove that

$$\widetilde{r}_{m_n}\left[F(u_{m_n}\,z_{m_n}) - \widehat{F}^*(u_{m_n}\,z_{m_n})\right] \xrightarrow[n\to\infty]{d} \mathscr{N}(\nu, \tau^2), \quad (2.51)$$

with $\widetilde{r}_{m_n} = \frac{\sqrt{m_n}}{\log(z_{m_n})}\left(\frac{1}{1-\frac{1}{n}\sum_{i=1}^n 1_{(X_i \leq u_{m_n} z_{m_n})}}\right) = \frac{\sqrt{m_n}}{\log(z_{m_n})\,\widehat{\overline{F}}(u_{m_n}\,z_{m_n})}$.

To this end we need to prove that

$$\frac{\overline{F}(u_{m_n}\,z_{m_n})}{\widehat{\overline{F}}(u_{m_n}\,z_{m_n})} \xrightarrow[n\to\infty]{\mathbb{P}} 1, \quad (2.52)$$

then, using Theorem 8.1 in Smith (1987) and the Slutsky theorem, we obtain (2.51). To prove (2.52) we use the following result:

Proposition B *(Einmahl, 1990; Corollary 1) Let a sequence of i.i.d random variables $X_1, X_2, \ldots$ from a distribution function F. We denote with $\{k_n\}_{n=1}^\infty$ an arbitrary sequence of positive numbers, such that $k_n \leq n$ and $k_n \to \infty$, $\lim_{n\to\infty} \frac{k_n}{n} = 0$. Let $\{\gamma_n\}_{n=1}^\infty$ be a sequence of positive numbers, such that $\lim_{n\to\infty} \frac{\gamma_n}{\sqrt{k_n}} = \infty$, then $\sup_{t \geq F^{-1}(1-\frac{k_n}{n})} \left(\frac{n}{\gamma_n}\right)\left|\widehat{\overline{F}}(t) - \overline{F}(t)\right| \xrightarrow[n\to\infty]{\mathbb{P}} 0$.*

We choose an arbitrary sequence $\{k_n\}_{n=1}^\infty := \{m_n\}_{n=1}^\infty$ (number of excesses on a sample of size n) such that $m_n \leq n$, $\lim_{n\to\infty} m_n = \infty$ and $\lim_{n\to\infty} \frac{m_n}{n} = 0$ (see Remark 1 in Section 2.1). We take $\{\gamma_n\}_{n=1}^\infty := \{\sqrt{m_n}\,\alpha_n\}_{n=1}^\infty$, where α_n is an arbitrary sequence of positive numbers such that $\lim_{n\to\infty} \alpha_n = \infty$.
Then, using Proposition B, we have for $u_{m_n}\,z_{m_n} \geq F^{-1}(1 - \frac{m_n}{n})$

$$\left(\frac{n}{\sqrt{m_n}\,\alpha_n}\,\overline{F}(u_{m_n}\,z_{m_n})\right)\left|\frac{\widehat{\overline{F}}(u_{m_n}\,z_{m_n}) - \overline{F}(u_{m_n}\,z_{m_n})}{\overline{F}(u_{m_n}\,z_{m_n})}\right| \xrightarrow[n\to\infty]{\mathbb{P}} 0.$$

We choose α_n such that for large n

$$\exists\, c > 0: \quad 0 < \frac{\sqrt{m_n}\,\alpha_n}{n\,\overline{F}(u_{m_n}\,z_{m_n})} \leq c. \quad (2.53)$$

In the Fréchet case we have $L(x) = x^{\alpha}\,\overline{F}(x)$, for $\alpha > 0$ and $\forall\ t > 0$, $\frac{L(tx)}{L(x)} = 1 + k(t)\phi(x) + o(\phi(x))$ for $x \to \infty$. Then, using Assumptions (8.7) and (8.8) of Theorem 8.1. in Smith (1987), we obtain

$$\frac{\overline{F}(u_{m_n}\, z_{m_n})}{\overline{F}(u_{m_n})} = z_{m_n}^{-\alpha}\left[1 + k(z_{m_n})\phi(u_{m_n}) + o(\phi(u_{m_n}))\right].$$

Hence $\frac{n\,\overline{F}(u_{m_n}\, z_{m_n})}{\sqrt{m_n}}$ is equal to $\frac{n}{\sqrt{m_n}}\,\overline{F}\,(u_{m_n})\big[z_{m_n}^{-\alpha}\big(1 + k(z_{m_n})\phi(u_{m_n}) + o(\phi(u_{m_n}))\big)\big]$ which, for n large, can be approximated by

$$\sqrt{m_n} z_{m_n}^{-\alpha}\left(1 + k(z_{m_n})\phi(u_{m_n}) + o(\phi(u_{m_n}))\right). \tag{2.54}$$

Assume now $\frac{(z_{m_n})^{\alpha}}{\sqrt{m_n}} \xrightarrow[n\to\infty]{} 0$, that is the analogue of condition in (2.35) conditionally on $N_n = m_n$. Then the properties of k and ϕ insure that the right hand side of (2.54) increases to infinity hence one can choose α_n satisfying (2.53). To conclude the proof, we use assumption (2.35) and a Skorohod type argument. □

Proof [Theorem 2.6.1]:
To prove (2.37) we first observe, using Corollary 2.3.1, Proposition 2.7.2 and the analogue of Kolmogorov-Smirnov Theorem in dimension 2 (e.g. see Dudley, 1966), that $\sqrt{k_n}\Big| C^{*G}\big(1 - g_X(1 - V_{\xi_X,\sigma_X}(\overline{f}_1(n)f_1(n) - \overline{f}_1(n))), 1 - g_Y(1 - V_{\xi_Y,\sigma_Y}(\overline{f}_2(n)f_2(n) - \overline{f}_2(n)))\big) \cdot \overline{F}(\overline{f}_1(n), \overline{f}_2(n)) - \widehat{\overline{F}}(\overline{f}_1(n), \overline{f}_2(n)) \cdot C^{*G}\Big(1 - \widehat{g}_X(1 - V_{\widehat{\xi}_X,\widehat{\sigma}_X}(\overline{f}_1(n)f_1(n) - \overline{f}_1(n))),\ 1 - \widehat{g}_Y(1 - V_{\widehat{\xi}_Y,\widehat{\sigma}_Y}(\overline{f}_2(n)f_2(n) - \overline{f}_2(n)))\Big)\Big| \xrightarrow[n\to\infty]{\mathbb{P}} 0.$
Furthermore $r_n\Big|\frac{1}{n}\sum_{i=1}^{n} 1_{\{X_i \le \overline{f}_1(n),\, Y_i \le \overline{f}_2(n)\}} - F(\overline{f}_1(n), \overline{f}_2(n))\Big| \xrightarrow[n\to\infty]{\mathbb{P}} 0$, with $r_n << \sqrt{n}$. At last using Corollary 2.3.1, Theorem 2.5.3, we obtain convergence (2.37). If $\widehat{\overline{f}}_2(n)$ satisfies conditions of Theorem 2.5.3 in probability then with the same proof structure we obtain (2.38). □

Proof [Proposition 2.3.2]: Under assumptions of Proposition 2.3.2, as in the proof of Lemma 6.1 in Draisma *et al.* (2004) we obtain

$$\sup_{0<x,y\le 1}\left|\sqrt{a_n}\left(\frac{1}{a_n}\sum_{i=1}^{n} 1_{\{R(X_i)>n-k_n x+1;\, R(Y_i)>n-k_n y+1\}} - G(x,y)\right) - W(x,y)\right| \xrightarrow[n\to\infty]{a.s.} 0,$$

where $a_n = n\,q(k_n/n)$ and $W(x,y)$ is a zero-mean gaussian process with $\mathbb{E}(W(x_1,y_1)W(x_2,y_2)) = G(x_1 \wedge x_2, y_1 \wedge y_2)$. Then, in particular

$$\psi_n \sup_{0<x,y\le 1}\left|\frac{\sum_{i=1}^{n}\frac{1}{k_n}1_{\{R(X_i)>n-k_n\,x+1;\, R(Y_i)>n-k_n\,y+1\}}}{\sum_{i=1}^{n}\frac{1}{k_n}1_{\{R(X_i)>n-k_n+1;\, R(Y_i)>n-k_n+1\}}} - G(x,y)\right|$$
$$= \psi_n \sup_{0<x,y\le 1}\left|\,\widehat{G}(x,y) - G(x,y)\,\right| \xrightarrow[n\to\infty]{\mathbb{P}} 0,$$

with $\psi_n << \sqrt{a_n} = \sqrt{n\,q(k_n/n)}$ and $\widehat{G}$ as in (2.18). Finally for the marginals g_X and g_Y we have

$$\psi_n \sup_{0<x\leq 1} \left|\widehat{g}_X(x) - g_X(x)\right| \xrightarrow[n\to\infty]{\mathbb{P}} 0, \quad \psi_n \sup_{0<y\leq 1} \left|\widehat{g}_Y(y) - g_Y(y)\right| \xrightarrow[n\to\infty]{\mathbb{P}} 0,$$

with $\widehat{g}_X$ and $\widehat{g}_Y$ as in (2.18). □

Proof [Theorem 2.6.2]:
Under assumptions of Theorem 2.6.2 and Proposition 2.3.2 we obtain asymptotic convergence results for $\widehat{G}(x,y)$, $\widehat{g}_X(x)$ and $\widehat{g}_Y(y)$, with convergence rate $\psi_n << \sqrt{n\,q(k_n/n)}$ and $\widehat{g}_X$, $\widehat{g}_Y$, $\widehat{G}$ as in (2.18).
With the same proof structure of Theorem 2.6.1, using Corollary 2.7.1 and Proposition 2.7.2 we obtain convergence (2.39). Moreover if $\widehat{\widehat{f}}_2(n)$ satisfies conditions of Theorem 2.5.3 in probability then we obtain (2.40). □

Auxiliary results

Theorem C *(Einmahl et al., 2006; Theorem 2.2) Assume that exists the limit $R(x,y)$ in* (2.14) *such that, for some $\alpha > 0$*

$$\frac{1}{t}\mathbb{P}(1 - F_X(X) \leq tx, 1 - F_Y(Y) \leq ty) - R(x,y) = O(t^\alpha), \text{ as } t \to 0, \tag{2.55}$$

uniformly for $\max(x,y) \leq 1$, $x, y \geq 0$. *Let $k_n \to \infty$, $k_n/n \to 0$ and $k_n = o\big(n^{\frac{2\alpha}{1+2\alpha}}\big)$. If $R_1(x,y) := \frac{\partial R(x,y)}{\partial x}$ and $R_2(x,y) := \frac{\partial R(x,y)}{\partial y}$ are continuous then*

$$\sup_{0<x,y\leq 1} \left|\sqrt{k_n}(\widehat{l}(x,y) - l(x,y)) + B(x,y)\right| \xrightarrow[n\to\infty]{\mathbb{P}} 0,$$

where $B(x,y) := W(x,y) - R_1(x,y)W_1(x) - R_2(x,y)W_2(y)$, with W a continuous mean zero Gaussian process on $[0,x] \times [0,y]$ with covariance structure $\mathbb{E}(W(x_1,y_1)\,W(x_2,y_2)) = R(x_1 \wedge x_2, y_1 \wedge y_2)$ and with marginal processes defined by $W_1(x) = W([0,x] \times [0,\infty])$, $W_2(y) = W([0,\infty] \times [0,y])$.

Note that (2.55) is a second-order condition quantifying the speed of convergence in (2.14) and condition $k_n = o\big(n^{\frac{2\alpha}{1+2\alpha}}\big)$ gives an upper bound on the speed with which k_n can grow to infinity. This upper bound is related to the speed of convergence in (2.55) by α. If C is a twice continuously differentiable copula on $[0,\,1]^2$ then (2.55) holds for any $\alpha \geq 1$.

Furthermore, it is easily seen that $\widehat{l}(x,y) + \widehat{R}(x,y) = \frac{\lceil k_n\,x\rceil + \lceil k_n\,y\rceil - 2}{k_n} \leq \frac{\lceil k_n\,x\rceil + \lceil k_n\,y\rceil}{k_n}$, almost surely, for each $0 < x, y \leq 1$, where $\lceil z\rceil$ is the smallest integer $\geq z$. Then under assumption of Theorem C we can easily obtain a gaussian approximation for $\widehat{R}(x,y) - R(x,y)$.
Note that the asymptotic variance of $\sqrt{k_n}(\widehat{l}(x,y) - l(x,y))$, in Theorem C, vanishes in the asymptotic independent case. Then, with $\lambda = 0$, we obtain:

Corollary 2.7.1 *Assume that, for some* $\alpha > 0$

$$\frac{1}{t}\mathbb{P}(1 - F_X(X) \leq tx, 1 - F_Y(Y) \leq ty) = O(t^{\alpha}), \text{ as } t \to 0,$$

uniformly for $\max(x, y) \leq 1$, $x, y \geq 0$. *Let* $k_n \to \infty$, $k_n/n \to 0$ *and* $k_n = o\big(n^{\frac{2\alpha}{1+2\alpha}}\big)$. *Then it holds*

$$\sup_{0<x,y\leq 1} \left|\sqrt{k_n}(\widehat{l}(x, y) - l(x, y))\right| \xrightarrow[n\to\infty]{\mathbb{P}} 0.$$

Proposition 2.7.2 *Let* $V_{\xi,\sigma}(x)$ *the Generalized Pareto Distribution (GPD) and* $\widehat{\xi}_n$, $\widehat{\sigma}_n$, *the maximum likelihood estimators of the parameters* $\xi = -\alpha^{-1} < 0$ *and* $\sigma = u_n\alpha^{-1}$, *in the case unconditionally on* N. *If all the conditions of Corollary 2.5.1 hold then*

$$p_{nx} \sup_{x\in[0,+\infty)} \left|V_{\widehat{\xi}_n,\widehat{\sigma}_n}(x) - V_{\xi,\sigma}(x)\right| \xrightarrow[n\to\infty]{\mathbb{P}} 0, \quad \text{where } \frac{p_{nx}}{\sqrt{N_x}} \xrightarrow[n\to\infty]{\mathbb{P}} 0.$$

Proof: Using Corollary 2.5.1 we obtain for each point $x \in [0, +\infty)$,

$$p_{nx}\left[V_{\widehat{\xi}_n,\widehat{\sigma}_n}(x) - V_{\xi,\sigma}(x)\right] = p_{nx}\left[\left(1 - \frac{\xi\, x}{\sigma}\right)^{\frac{1}{\xi}} - \left(1 - \frac{\widehat{\xi}_n\, x}{\widehat{\sigma}_n}\right)^{\frac{1}{\widehat{\xi}_n}}\right] \xrightarrow[n\to\infty]{\mathbb{P}} 0, \qquad (2.56)$$

where $\frac{p_{nx}}{\sqrt{N_x}} \xrightarrow[n\to\infty]{\mathbb{P}} 0$. Finally, applying a stochastic version of Polya's Theorem (see Horowitz, 2001), as $V_{\xi,\sigma}(x)$ is a continuous distribution function, the convergence in (2.56) holds uniformly on $[0, +\infty)$. □

Acknowledgements The authors thank two anonymous referees for relevant remarks and constructive comments on a previous version of the paper. They also thank Anne-Laure Fougères, Ragnar Norberg and José R. León for fruitful discussions. This work has been partially supported by the French research national agency (ANR) under the reference ANR-08BLAN-0314-01.

Bibliography

Balkema, A. A. and de Haan, L. (1974). Residual life time at great age. *Ann. Probability*, 2:792–804.

Beirlant, J., Dierckx, G., and Guillou, A. (2011). Reduced bias estimators for bivariate tail modelling. *Insurance: Mathematics and Economics*, In Press, Accepted Manuscript.

Beirlant, J., Goegebeur, Y., Teugels, J., and Segers, J. (2004). *Statistics of extremes.* Wiley Series in Probability and Statistics. John Wiley & Sons Ltd., Chichester. Theory and applications, With contributions from Daniel De Waal and Chris Ferro.

Benth, F. E. and Kettler, P. (2011). Dynamic copula models for the spark spread. *Quantitative Finance*, 11(3):407–421.

Borkar, V. S. (1995). *Probability theory.* Universitext. Springer-Verlag, New York. An advanced course.

Capéraà, P. and Fougères, A.-L. (2000). Estimation of a bivariate extreme value distribution. *Extremes*, 3(4):311–329 (2001).

Charpentier, A. and Juri, A. (2006). Limiting dependence structures for tail events, with applications to credit derivatives. *J. Appl. Probab.*, 43(2):563–586.

Charpentier, A. and Segers, J. (2007). Lower tail dependence for Archimedean copulas: characterizations and pitfalls. *Insurance Math. Econom.*, 40(3):525–532.

Coles, S. G. and Tawn, J. A. (1994). Statistical methods for multivariate extremes: an application to structural design. *Appl. Statist.*, 43:1–48. With discussion.

Davison, A. C. (1984). *A statistical model for contamination due to long-range atmospheric transport of radionuclides.* PhD thesis, Department of Mathematics, Imperial College of Science and Technology, London.

Davison, A. C. and Smith, R. L. (1990). Models for exceedances over high thresholds. *J. Roy. Statist. Soc. Ser. B*, 52(3):393–442. With discussion and a reply by the authors.

de Haan, L. (1970). *On regular variation and its application to the weak convergence of sample extremes*, volume 32 of *Mathematical Centre Tracts.* Mathematisch Centrum, Amsterdam.

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

Drees, H., Ferreira, A., and de Haan, L. (2004). On maximum likelihood estimation of the extreme value index. *Ann. Appl. Probab.*, 14(3):1179–1201.

Drees, H. and Huang, X. (1998). Best attainable rates of convergence for estimators of the stable tail dependence function. *J. Multivariate Anal.*, 64(1):25–47.

Dudley, R. M. (1966). Weak convergences of probabilities on nonseparable metric spaces and empirical measures on Euclidean spaces. *Illinois J. Math.*, 10:109–126.

Einmahl, J. H. J. (1990). The empirical distribution function as a tail estimator. *Statist. Neerlandica*, 44(2):79–82.

Einmahl, J. H. J., de Haan, L., and Li, D. (2006). Weighted approximations of tail copula processes with application to testing the bivariate extreme value condition. *Ann. Statist.*, 34(4):1987–2014.

Einmahl, J. H. J., Krajina, A., and Segers, J. (2008). A method of moments estimator of tail dependence. *Bernoulli*, 14(4):1003–1026.

El-Aroui, M.-A. and Diebolt, J. (2002). On the use of the peaks over thresholds method for estimating out-of-sample quantiles. *Comput. Statist. Data Anal.*, 39(4):453–475.

Embrechts, P., Klüppelberg, C., and Mikosch, T. (1997). *Modelling extremal events*, volume 33 of *Applications of Mathematics (New York)*. Springer-Verlag, Berlin. For insurance and finance.

Falk, M. and Reiss, R.-D. (2005). On the distribution of Pickands coordinates in bivariate EV and GP models. *J. Multivariate Anal.*, 93(2):267–295.

Frees, E. W. and Valdez, E. A. (1998). Understanding relationships using copulas. *North American Actuarial Journal*, 2(1):1–25.

Goldie, C. M. and Smith, R. L. (1987). Slow variation with remainder: theory and applications. *Quart. J. Math. Oxford Ser. (2)*, 38(149):45–71.

Horowitz, J. L. (2001). *Handbook of Econometrics*, chapter The Bootstrap, pages 3159–3228. J.J. Heckman and E.E. Leamer, Elsevier Science B.V.

Huang, X. (1992). *Statistics of bivariate extreme values*. PhD thesis, Erasmus University Rotterdam, Tinbergen Institute.

Javid, A. (2009). Limiting tail dependence copulas. *Communications in Statistics-Theory and Methods*, 38(20):3772–3781.

Juri, A. and Wüthrich, M. V. (2002). Copula convergence theorems for tail events. *Insurance Math. Econom.*, 30(3):405–420.

Juri, A. and Wüthrich, M. V. (2004). Tail dependence from a distributional point of view. *Extremes*, 6(3):213–246.

Ledford, A. W. and Tawn, J. A. (1996). Statistics for near independence in multivariate extreme values. *Biometrika*, 83(1):169–187.

Ledford, A. W. and Tawn, J. A. (1997). Modelling dependence within joint tail regions. *J. Roy. Statist. Soc. Ser. B*, 59(2):475–499.

Ledford, A. W. and Tawn, J. A. (1998). Concomitant tail behaviour for extremes. *Adv. in Appl. Probab.*, 30(1):197–215.

Lescourret, L. and Robert, C. Y. (2006). Extreme dependence of multivariate catastrophic losses. *Scandinavian Actuarial Journal*, (4):203–225.

McNeil, A. J. (1997). Estimating the tails of loss severity distributions using extreme value theory. *ASTIN Bulletin*, 27:1117–1137.

McNeil, A. J. (1999). Extreme value theory for risk managers. *preprint, ETH, Zurich.*

Michel, R. (2008). Some notes on multivariate generalized pareto distributions. *J. Multivariate Anal.*, 99(6):1288–1301.

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

Peng, L. (1999). Estimation of the coefficient of tail dependence in bivariate extremes. *Statist. Probab. Lett.*, 43(4):399–409.

Pickands, III, J. (1975). Statistical inference using extreme order statistics. *Ann. Statist.*, 3:119–131.

Pickands, III, J. (1981). Multivariate extreme value distributions. In *Proceedings of the 43rd session of the International Statistical Institute, Vol. 2 (Buenos Aires, 1981)*, volume 49, pages 859–878, 894–902. With a discussion.

Ramos, A. and Ledford, A. (2009). A new class of models for bivariate joint tails. *J. R. Stat. Soc. Ser. B Stat. Methodol.*, 71(1):219–241.

Reiss, R.-D. and Thomas, M. (2007). *Statistical analysis of extreme values with applications to insurance, finance, hydrology and other fields.* Birkhäuser Verlag, Basel, third edition. With 1 CD-ROM (Windows).

Resnick, S. I. (1987). *Extreme values, regular variation, and point processes*, volume 4 of *Applied Probability. A Series of the Applied Probability Trust.* Springer-Verlag, New York.

Rootzén, H. and Tajvidi, N. (2006). Multivariate generalized Pareto distributions. *Bernoulli*, 12(5):917–930.

Schlather, M. (2001). Examples for the coefficient of tail dependence and the domain of attraction of a bivariate extreme value distribution. *Statist. Probab. Lett.*, 53(3):325–329.

Smith, R. L. (1987). Estimating tails of probability distributions. *Ann. Statist.*, 15(3):1174–1207.

Wüthrich, M. V. (2004). Bivariate extension of the Pickands-Balkema-de Haan theorem. *Ann. Inst. H. Poincaré Probab. Statist.*, 40(1):33–41.

Chapitre 3

Quelques généralisations de mesures des risque dans un cadre multivarié : propriétés par rapport au niveau de risque et à la structure de dépendance

Résumé

Ce chapitre est composé d'un article en anglais, intitulé *A multivariate extension of Value-at-Risk and Conditional-Tail-Expectation*, coécrit avec Areski Cousin et actuellement soumis pour publication.

Nous définissons les mesures de risque multidimensionnelles, pour $\alpha \in (0,1)$ et $\mathbf{X} = (X_1, \ldots, X_d)$:

$$\mathrm{VaR}_\alpha(\mathbf{X}) = \mathbb{E}[\mathbf{X} \,|\, \mathbf{X} \in \partial L(\alpha)] = \begin{pmatrix} \mathbb{E}[\, X_1 \,|\, \mathbf{X} \in \partial L(\alpha)\,] \\ \vdots \\ \mathbb{E}[\, X_d \,|\, \mathbf{X} \in \partial L(\alpha)\,] \end{pmatrix},$$

$$\mathrm{CTE}_\alpha(\mathbf{X}) = \mathbb{E}[\mathbf{X} \,|\, \mathbf{X} \in L(\alpha)] = \begin{pmatrix} \mathbb{E}[\, X_1 \,|\, \mathbf{X} \in L(\alpha)\,] \\ \vdots \\ \mathbb{E}[\, X_d \,|\, \mathbf{X} \in L(\alpha)\,] \end{pmatrix},$$

et nous en étudions les propriétés. Plus spécifiquement nous analysons le comportement des nos mesures par rapport au changement du niveau du risque et au changement de la structure de dépendance du vecteur $\mathbf{X}$. En ce qui concerne le premier point nous mettons en lumière les caractéristiques de la CTE et de la VaR en termes de propriétés classiques pour des mesures de risque (Artzner, 1999) : monotonie par rapport au niveau de risque (voir Propositions 3.2.11 et 3.3.11) ; homogénéité et invariance par translation (voir Propositions 3.2.2 et 3.3.2) ; chargement de sécurité (voir Corollaire 3.3.15).

En ce qui concerne le deuxième point nous examinons comment la mesure se conduit par rapport au changement des risques marginaux (avec copule inchangée) (voir Sections 3.2.3 and 3.3.3). En se plaçant dans des familles spécifiques paramétriques de copules nous arrivons à donner des formules exactes pour la $\mathrm{VaR}_\alpha(\mathbf{X})$ et la $\mathrm{CTE}_\alpha(\mathbf{X})$ et donc à analyser son comportement par rapport au changement du paramètre et du niveau de risque.

Présentation des approfondissements donnés en Annexe

En Annexe nous présentons des illustrations et d'autres études, inhérentes à notre Espérance de queue multidimensionnelle (*Conditional-Tail-Expectation*, CTE), qui n'ont pas été insérées dans l'article.

- Dans la Section B.5 nous donnons d'autres formules explicites pour la $\mathrm{CTE}_\alpha(\mathbf{X})$ dans le cas de familles de copules bivariées connues.
- Dans la Section B.6 nous présentons des comparaisons entre notre $\mathrm{CTE}_\alpha(\mathbf{X})$, dans un cadre bivarié, et différentes généralisations de la *Conditional-Tail-Expectation* proposées dans la littérature des ces dernières années.
- Dans la Section B.7 nous présentons une discussion sur la croissance de la CTE dans deux cas bivariés discrets.

Introduction

During the last decades, researchers joined efforts to properly compare, quantify and manage risk. Regulators edict rules for bankers and insurers to improve their risk management and to avoid crises, not always successfully as illustrated by recent events.

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 external risk factors. Let us consider for instance the problem of solvency capital allocation for financial institutions with multi-branch businesses confronted to risks with specific characteristics. Under Basel II and Solvability II, a bottom-up approach is used to estimate a "top-level" solvency capital. This is done by using risk aggregation techniques who may capture risk mitigation or risk diversification effects. Then this global capital amount is re-allocated to each subsidiaries or activities for internal risk management purpose ("top-down approach"). Note that the solvability of each individual branch may strongly be affected by the degree of dependence amongst all branches. The interest reader is referred to Denuit *et al.* (1999). As a result, the capital allocated to each branch has to be computed in a multivariate setting where both marginal effects and dependence between risks should be captured. In this respect, the "Euler approach" (e.g., see Tasche, 2008) involving vector-valued risk measures has already been tested by risk management teams of some financial institutions.
Whereas the previous risk allocation problem only involves internal risks associated with businesses in different subsidiaries, the solvability of financial institutions could also be affected by external risks whose sources cannot be controlled. These risks may also be strongly heterogeneous in nature and difficult to diversify away. One can think for instance of systemic risk or contagion effects in a strongly interconnected system of financial companies. In this regard, micro-prudential regulation has been criticized because of its failure to limit the systemic risk within the system. This question has been dealt with recently by among others, Gauthier *et al.* (2010) and Zhou (2010) who highlights the benefit of a "macro-prudential" approach as an alternative solution to the existing "micro-prudential" one (Basel II) which does not take into account interactions between financial institutions.

In this paper, we introduce a multivariate extension of the classical univariate *Value-at-Risk* (VaR). This extension may be useful to understand how solvency capital requirement (SCR) computed for a given financial institution may be affected by the presence of additional risks. We also study the multivariate extension of the *Conditional-Tail–Expectation* (CTE), previously introduced by Di Bernardino *et al.* (2011) in a bivariate setting. Several properties have been derived. In particular, we show that these two risk measures both satisfy the positive homogeneity and the translation invariance property. Comparison results between univariate risk measures and components of multivariate risk measures are provided. We also analyze how they are impacted by a change in marginal distributions, by a change in dependence structure and by a change in risk level. Interestingly, these results turn to be consistent with existing properties on univariate risk measures. In particular, we prove that, for different Archimedean families

of copulas, an increase in dependence amongst risks tends to lower the amount of solvency capital required for these risks. At the extreme case where risks are perfectly dependent (i.e., comonotonic), each component of the multivariate VaR (resp. CTE) are equal to the corresponding univariate VaR (resp. CTE). If we think of univariate risk measures as solvency capital computed in a micro-prudential regulated system, this feature is in line with the observation made by Zhou (2010): *"When regulating a system consisting of similar institutions, or in other words, the system is highly interconnected, considering a micro-prudential regulation can be sufficient for reducing the overall systemic risk."* (Zhou, 2010).
Furthermore, our multivariate risk measures can be considered as a "fair" allocation of solvency capital with respect to individual risk-taking behavior in an economy with d interconnected financial institutions. We prove that capital required for an institution is affected both by its own marginal risk and by the degree to which its business is connected to the activity of the other institutions. However, the solvency capital required for one particular institution does not depend on the marginal risks bearing by the others. This can be considered as an invariance property with respect to a change in external risk's marginal distributions, as far as the dependence structure is being fixed.

From the years 2000 onward, much research has been devoted to risk measures and many extension to multidimensional settings have been suggested (see, e.g., Jouini *et al.*, 2004 and Bentahar, 2006). Unsurprisingly, the main difficulty regarding multivariate generalizations of quantile-based risk measures (as the VaR and the CTE) is the fact that vector preorders are, in general, partial preorders. Then, what can be considered in a context of multidimensional portfolios as the analogous of a "worst case" scenario and a related "tail distribution"? This is the first question we shall address by suggesting a suitable definition of quantiles for multi-risk portfolios. In the last decade, several attempts have been made to provide a multidimensional generalization of the univariate quantile function. For example, Massé and Theodorescu (1994) defined multivariate quantiles as half-planes and Koltchinskii (1997) provided a general treatment of multivariate quantiles as inversions of mappings. Another approach is to use geometric quantiles (see for example Chaouch *et al.*, 2009). Along with the geometric quantile, the notion of depth function has been developed in recent years to characterize the quantile of multidimensional distribution functions. We refer to Serfling (2002) for a large review on multivariate quantiles. When it turns to generalize the *Value-at-Risk* measure, Embrechts and Puccetti (2006), Nappo and Spizzichino (2009) use the notion of quantile curve (formally introduced in Section 3.2). Contrarily to the latter approach, the multivariate *Value-at-Risk* proposed in this paper quantifies multivariate risks in a more parsimonious and synthetic way. This feature can be relevant from an operational point of view.

In the literature, several generalizations of the classical univariate CTE have been proposed, mainly using as conditioning events the total risk or some univariate extreme risk in the portfolio. This kind of measures are suitable to model problems of capital allocation in a portfolio of dependent risks. More precisely, let $\mathbf{X} = (X_1, \ldots, X_d)$ be a risk vector, where, for any $i = 1, \ldots, d$, the component X_i denotes risk (usually claim

or loss) associated with subportfolio i. Then, $S = X_1 + \cdots + X_d$ corresponds to the total risk of this portfolio, $X_{(1)} = \min\{X_1, ..., X_d\}$ and $X_{(d)} = \max\{X_1, ..., X_d\}$ are the extreme risks. In capital allocation problems, we are not only interested in the "stand-alone" risk measures $\mathrm{CTE}_\alpha(X_i) = \mathbb{E}[\, X_i \,|\, X_i > Q_{X_i}(\alpha)\,]$, where $Q_{X_i}(\alpha)$ is the univariate quantile function of X_i at risk level α, but also in

$$\mathrm{CTE}_\alpha^{\mathrm{sum}}(X_i) = \mathbb{E}[\, X_i \,|\, S > Q_S(\alpha)\,], \tag{3.1}$$

$$\mathrm{CTE}_\alpha^{\min}(X_i) = \mathbb{E}[\, X_i \,|\, X_{(1)} > Q_{X_{(1)}}(\alpha)\,], \tag{3.2}$$

$$\mathrm{CTE}_\alpha^{\max}(X_i) = \mathbb{E}[\, X_i \,|\, X_{(d)} > Q_{X_{(d)}}(\alpha)\,], \tag{3.3}$$

for $i = 1, \ldots, d$. The interested reader is referred to Cai and Li (2005) for further details. For explicit formulas of $\mathrm{CTE}_\alpha^{\mathrm{sum}}(X_i)$ in the case of Fairlie-Gumbel-Morgenstern family of copulas, see Bargès *et al.* (2009). Landsman and Valdez obtain an explicit formula for $\mathrm{CTE}_\alpha^{\mathrm{sum}}(X_i)$ in the case of elliptic distribution functions (see Landsman and Valdez, 2003); Cai and Li in the case of *phase-type* distributions (see Cai and Li, 2005). Furthermore, we recall that $\mathrm{CTE}_\alpha^{\mathrm{sum}}(X_i)$ is a key tool to calculate the amount of solvency capital in the "Euler approach" (see, e.g., Tasche, 2008).

The multivariate version of *Conditional-Tail-Expectation*, studied in this paper, is essentially based on a "*distributional approach*". It is constructed as the conditional expectation of a multivariate random vector given that the latter is located in a particular set corresponding to the α-upper level set of the associated multivariate distribution function (in a bivariate setting see Di Bernardino *et al.*, 2011). More precisely we consider a non-negative multivariate random vector $\mathbf{X} = (X_1, \ldots, X_d)$ and the associated α-upper level set, i.e., $L(\alpha) = \{\mathbf{x} \in \mathbb{R}_+^d \,:\, F_{\mathbf{X}}(\mathbf{x}) \geq \alpha\}$, for $\alpha \in (0, 1)$, where $F_{\mathbf{X}}$ is the multivariate distribution function of $\mathbf{X}$. As the total information of the vector $\mathbf{X}$ is completely described by its multivariate distribution function, the set $L(\alpha)$ captures the information coming both from the marginal distributions and from the multivariate dependence structure. Then contrarily to existing generalizations of the univariate CTE presented above, our $\mathrm{CTE}_\alpha(\mathbf{X})$ does not use an arbitrary real-valued aggregate transformation (sum, min, max,...). Indeed, using an aggregate procedure between the risks can be inappropriate to measure risks with heterogeneous characteristics especially in an external risks problem. Moreover, as opposed to our proposal of multivariate risk measures, the multivariate risk measures defined in (3.1)-(3.3) do not satisfy the invariance property with respect to a change in external risk's marginal distributions. This means that for a given risk component i in the portfolio, changes in marginal distributions of the other risks lead to changes in risk measures (3.1)-(3.3) for name i, even if the dependence structure does not change.

The paper is organized as follows. In Section 3.1, we introduce some notations, tools and technical assumptions. In Section 3.2, we propose a new multivariate generalization of *Value-at-Risk*. In Section 3.3, we generalize in a multivariate setting the *Conditional-Tail-Expectation*, previously introduced by Di Bernardino *et al.* (2011) in dimension two. We study the properties of our multivariate VaR and CTE in terms of Artzner *et*

al. (1999)'s invariance properties of risk measures (see Sections 3.2.1 and 3.3.1). We also compare the components of these multivariate risk measure with the associated univariate risk measures (see Sections 3.2.2 and 3.3.2). The behavior of $\mathrm{VaR}_\alpha(\mathbf{X})$ (resp. $\mathrm{CTE}_\alpha(\mathbf{X})$) with respect to a change in marginal distributions, a change in dependence structure and a change in risk level α is discussed in Sections 3.2.3-3.2.5 (resp. Sections 3.3.3-3.3.4). Further illustrations in some Archimedean copula cases, are presented both for the multivariate *Value-at-Risk* and for the *Conditional-Tail-Expectation*. In the conclusion, we discuss possible directions for future work.

3.1 Basic notions and preliminaries

In this section, we first introduce some notation and tools which will be used later on.

Stochastic orders

From now on, let $Q_X(\alpha)$ be the univariate quantile function of a risk X at level $\alpha \in (0,1)$. More precisely, given an univariate continuous and strictly monotonic loss distribution function F_X, $Q_X(\alpha) = F_X^{-1}(\alpha)$, $\forall\, \alpha \in (0,1)$. We recall here the definition and some properties of useful univariate and multivariate stochastic orders.

Definition 3.1.1 (Stochastic dominance order) *Let X and Y be two random variables. Then X is said to be smaller than Y in stochastic dominance, denoted as $X \preceq_{st} Y$, if the inequality $Q_X(\alpha) \leq Q_Y(\alpha)$ is satisfied for all $\alpha \in (0,1)$.*

Definition 3.1.2 (Stop-loss order) *Let X and Y be two random variables. Then X is said to be smaller than Y in the stop-loss order, denoted as $X \preceq_{sl} Y$, if for all $t \in \mathbb{R}$, $\mathbb{E}[(X-t)_+] \leq \mathbb{E}[(Y-t)_+]$, with $x_+ := \max\{x, 0\}$.*

Definition 3.1.3 (Increasing convex order) *Let X and Y be two random variables. Then X is said to be smaller than Y in the increasing convex order, denoted as $X \preceq_{icx} Y$, if $\mathbb{E}[f(X)] \leq \mathbb{E}[f(Y)]$, for all non-decreasing convex function f such that the expectations exist.*

The stop-loss order and the increasing convex order are equivalent (see Theorem 1.5.7 in Müller and Stoyan, 2001). Note that stochastic dominance order implies stop-loss order. For more details about stop-loss order we refer the interested reader to Müller (1997). Moreover, a sufficient condition for the stop-loss order is the *dangerousness order relation* as stated in the following lemma.

Lemma 3.1.4 (Ohlin, 1969) *Let X and Y be random variables with finite means such that $\mathbb{E}[X] \leq \mathbb{E}[Y]$, and there exists some real number c such that $F_X(x) \leq F_Y(x)$, for all $x < c$ and $F_X(x) \geq F_Y(x)$, for all $x \geq c$. Then X precedes Y in dangerousness order, written $X \preceq_D Y$, and this implies the stop-loss order $X \preceq_{sl} Y$.*

According to Bühlmann's terminology, when $X \preceq_D Y$ the distribution function F_Y is said to be more dangerous than F_X. This terminology is essentially related to the

variability of the random variables X and Y (see Section 3.4.2.2 in Denuit et al., 2005). For further details, the reader is referred to Bühlmann *et al.* (1977).
Finally, we introduce the definition of supermodular function and supermodular order for multivariate random vectors.

Definition 3.1.5 (Supermodular function) *A function $f : \mathbb{R}^d \to \mathbb{R}$ is said to be supermodular if for any $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^d$ it satisfies*

$$f(\boldsymbol{x}) + f(\boldsymbol{y}) \leq f(\boldsymbol{x} \wedge \boldsymbol{y}) + f(\boldsymbol{x} \vee \boldsymbol{y}),$$

where the operators $\wedge$ and $\vee$ denote coordinatewise minimum and maximum respectively.

Definition 3.1.6 (Supermodular order) *Let $\boldsymbol{X}$ and $\boldsymbol{Y}$ be two $d-$dimensional random vectors such that $\mathbb{E}[f(\boldsymbol{X})] \leq \mathbb{E}[f(\boldsymbol{Y})]$, for all supermodular functions $f : \mathbb{R}^d \to \mathbb{R}$, provided the expectation exist. Then $\boldsymbol{X}$ is said to be smaller than $\boldsymbol{Y}$ with respect to the supermodular order (denoted by $\boldsymbol{X} \preceq_{sm} \boldsymbol{Y}$).*

This will be a key tool to analyze the impact of dependence on our multivariate risk measures.

Kendall distribution function

Let $\mathbf{X} = (X_1, \ldots, X_d)$ be a $d-$dimensional random vector, $d \geq 2$. As we will see later on, our study of multivariate risk measures strongly relies on the key concept of Kendall distribution function (or multivariate probability integral transformation), that is, the distribution function of the random variable $F(\mathbf{X})$, where F is the multivariate distribution of random vector $\mathbf{X}$. From now on, the Kendall distribution will be denoted by K, so that $K(\alpha) = \mathbb{P}[F(\mathbf{X}) \leq \alpha]$, for $\alpha \in [0, 1]$. We also denote by $\overline{K}(\alpha)$ the survival distribution function of $F(\mathbf{X})$, i.e., $\overline{K}(\alpha) = \mathbb{P}[F(\mathbf{X}) > \alpha]$. For more details on the multivariate probability integral transformation, the interested reader is referred to Capéraà *et al.*, (1997), Genest and Rivest (2001), Nelsen *et al.* (2003), Genest and Boies (2003), Genest *et al.* (2006) and Belzunce *et al.* (2007).

In contrast to the univariate case, it is not generally true that the distribution function K of $F(\mathbf{X})$ is uniform on $[0, 1]$, even when F is continuous. Note also that it is not possible to characterize the joint distribution F or reconstruct it from the knowledge of K alone, since the latter does not contain any information about the marginal distributions $F_{X_1}, \ldots, F_{X_d}$ (see Genest and Rivest, 2001). Indeed, as a consequence of Sklar's Theorem, the Kendall distribution only depends on the dependence structure or the copula function C associated with $\mathbf{X}$ (see Sklar, 1959). Thus, we also have $K(\alpha) = \mathbb{P}[C(\mathbf{U}) \leq \alpha]$ where $\mathbf{U} = (U_1, \ldots, U_d)$ and $U_1 = F_{X_1}(X_1), \ldots, U_d = F_{X_d}(X_d)$.
Furthermore:

- For a $d-$dimensional random vector $\mathbf{X} = (X_1, \ldots, X_d)$ with copula C, the Kendall distribution function $K(\alpha)$ is linked to the Kendall's tau correlation coefficient via: $\tau_C = \frac{2^d\ \mathbb{E}[C(\mathbf{U})]-1}{2^{d-1}-1}$, for $d \geq 2$ (see Section 5 in Genest and Rivest, 2001).

- The Kendall distribution can be obtain explicitly in the case of multivariate Archimedean copulas with generator ϕ. Table 3.1 provides the expression of Kendall distributions associated with Archimedean, independent, comonotonic and counter-monotonic $d-$dimensional random vectors (see Barbe *et al.*, 1996). Note that the Kendall distribution is uniform for comonotonic random vectors.

Copula	Kendall distribution $K(\alpha)$
Archimedean case	$\alpha + \sum_{i=1}^{d-1}(-1)^i \frac{\phi^i(\alpha)}{i!} f_{i-1}(\alpha)$
Counter-monotonic case	1
Independent case	$\alpha + \alpha \sum_{i=1}^{d-1} \left(\frac{\ln(1/\alpha)^i}{i!}\right)$
Comonotonic case	α

Table 3.1: Kendall distribution in some classical $d-$dimensional dependence structure.

In Table 3.1, $f_i(\alpha)$ stands for $\frac{\mathrm{d}^{i-1}\phi^{-1}(t)}{\mathrm{d}t^{i+1}}$, evaluated at $t = \phi(\alpha)$. For further details the interested reader is referred to Section 2 in Barbe *et al.* (1996) and Section 5 in Genest and Rivest (2001). For instance, in the bivariate case, the Kendall distribution function is equal to $\alpha - \frac{\phi(\alpha)}{\phi'(\alpha)}$, $\alpha \in (0,1)$, for Archimedean copulas with differentiable generator ϕ and $\alpha\,(1 - \ln(\alpha))$, $\alpha \in (0,1)$, for the independence copula.

- It holds that $\alpha \leq K(\alpha) \leq 1$, for all $\alpha \in (0,1)$, i.e., the graph of the Kendall distribution function is above the first diagonal (see Section 5 in Genest and Rivest, 2001). This is equivalent to state that, for any random vector $\mathbf{U}$ with copula function C and uniform marginals, $C^{\mathrm{cc}}(\mathbf{U}^{\mathrm{cc}}) \preceq_{st} C(\mathbf{U}) \preceq_{st} C^{\mathrm{c}}(\mathbf{U}^{\mathrm{c}})$ where $\mathbf{U}^{\mathrm{cc}} = (U_1^{\mathrm{cc}}, \ldots, U_d^{\mathrm{cc}})$ (resp. $\mathbf{U}^{\mathrm{c}} = (U_1^{\mathrm{c}}, \ldots, U_d^{\mathrm{c}})$) is a counter-monotonic (resp. comonotonic) random vector with copula function C^{cc} (resp. C^{c}) and uniform marginals.

This last property suggests that when the level of dependence between $X_1, \ldots, X_d$ increases, the Kendall distribution also increases in some sense. The following result, using definitions of stochastic orders described above, investigates rigorously this intuition.

Proposition 3.1.7 *Let $\boldsymbol{U} = (U_1, \ldots, U_d)$ (resp. $\boldsymbol{U}^* = (U_1^*, \ldots, U_d^*)$) be a random vector with copula C (resp. C^*) and uniform marginals.*

$$\textit{If} \;\; \boldsymbol{U} \preceq_{sm} \boldsymbol{U}^*, \;\; \textit{then} \;\; C(\boldsymbol{U}) \preceq_{sl} C^*(\boldsymbol{U}^*).$$

Proof: Trivially, $\mathbf{U} \preceq_{sm} \mathbf{U}^* \Rightarrow C(\mathbf{u}) \leq C^*(\mathbf{u})$, for all $\mathbf{u} \in [0,1]^d$ (see Section 6.3.3 in Denuit *et al.*, 2005). Let $f : [0,1] \to \mathbb{R}$ be a non-decreasing and convex function. It holds that $f(C(\mathbf{u})) \leq f(C^*(\mathbf{u}))$, for all $\mathbf{u} \in [0,1]^d$, and $\mathbb{E}[f(C(\mathbf{U}))] \leq \mathbb{E}[f(C^*(\mathbf{U}))]$. Remark that since C^* is non-decreasing and supermodular and f is non-decreasing and convex then $f \circ C^*$ is a non-decreasing and supermodular function (see Theorem 3.9.3 in Müller and Stoyan, 2001). Then, by assumptions,

$\mathbb{E}[f(C(\mathbf{U}))] \leq \mathbb{E}[f(C^*(\mathbf{U}))] \leq \mathbb{E}[f(C^*(\mathbf{U}^*))]$. This implies $C(\mathbf{U}) \preceq_{sl} C^*(\mathbf{U}^*)$. Hence the result. □

From Proposition 3.1.7, we remark that $\mathbf{U} \preceq_{sm} \mathbf{U}^*$ implies an ordering relation between corresponding Kendall's tau : $\tau_C \leq \tau_{C^*}$. Note that the supermodular order between $\mathbf{U}$ and $\mathbf{U}^*$ does not necessarily yield the stochastic dominance order between $C(\mathbf{U})$ and $C^*(\mathbf{U}^*)$ (i.e., $C(\mathbf{U}) \preceq_{st} C^*(\mathbf{U}^*)$ does not hold in general). For a bivariate counter-example, the interested reader is referred to, for instance, Capéraà *et al.* (1997) or Example 3.1 in Nelsen *et al.* (2003).

Let us now focus on some classical families of bivariate Archimedean copulas. In Table 3.2, we obtain analytical expressions of the Kendall distribution function for Gumbel, Frank, Clayton and Ali-Mikhail-Haq families.

Copula	$\theta \in$	Kendall distribution $K(\alpha, \theta)$
Gumbel	$[1, \infty)$	$\alpha\left(1 - \frac{1}{\theta}\ln\alpha\right)$
Frank	$(-\infty, \infty) \setminus \{0\}$	$\alpha + \frac{1}{\theta}\left(1 - \mathrm{e}^{\theta\alpha}\right)\ln\left(\frac{1-\mathrm{e}^{-\theta\,\alpha}}{1-\mathrm{e}^{-\theta}}\right)$
Clayton	$[-1, \infty) \setminus \{0\}$	$\alpha\left(1 + \frac{1}{\theta}\left(1 - \alpha^{\theta}\right)\right)$
Ali-Mikhail-Haq	$[-1, 1)$	$\frac{\alpha - 1 + \theta + (1-\theta+\theta\alpha)(\ln(1-\theta+\theta\,\alpha)+\ln\alpha)}{\theta - 1}$

Table 3.2: Kendall distribution in some bivariate Archimedean cases.

Remark 3.1.8 *Bivariate Clayton copulas can be extended to d-dimensional copulas with $d > 2$ as far as the underlying dependence parameter is such that $\theta > -\frac{1}{d-1}$ (see Example 4.27 in Nelsen, 1999). Frank copulas can be extended to d-dimensional copulas ($d > 2$) for $\theta > 0$ (see Example 4.24 in Nelsen, 1999).*

Note that parameter θ governs the level of dependence amongst components of the underlying random vector. Indeed, it can be shown that, for all Archimedean copulas in Table 3.2, an increase of θ yields an increase of dependence in the sense of the supermodular order, i.e., $\theta \leq \theta^* \Rightarrow \mathbf{U} \preceq_{sm} \mathbf{U}^*$ (see further examples in Joe, 1997 and Wei and Hu, 2002). Then, as a consequence of Proposition 3.1.7, the following comparison result holds

$$\theta \leq \theta^* \Rightarrow C(\mathbf{U}) \preceq_{sl} C^*(\mathbf{U}^*). \tag{3.4}$$

In fact, a stronger comparison result can be derived for Archimedean copulas of Table 3.2, as shown in the following remark.

Remark 3.1.9 *For copulas in Table 3.2, one can check that $\frac{\partial K(\alpha,\theta)}{\partial\theta} \leq 0$, for all $\alpha \in (0, 1)$. This means that, for these classical examples, the associated Kendall distributions actually increase with respect to the stochastic dominance order when the dependence parameter θ increases, i.e.,*

$$\theta \leq \theta^* \Rightarrow C(\boldsymbol{U}) \preceq_{st} C^*(\boldsymbol{U}^*). \tag{3.5}$$

In order to illustrate this property we plot in Figure 3.1 the Kendall distribution function $K(\cdot,\theta)$ for different choices of parameter θ in the bivariate Clayton copula case and in the bivariate Gumbel copula case.

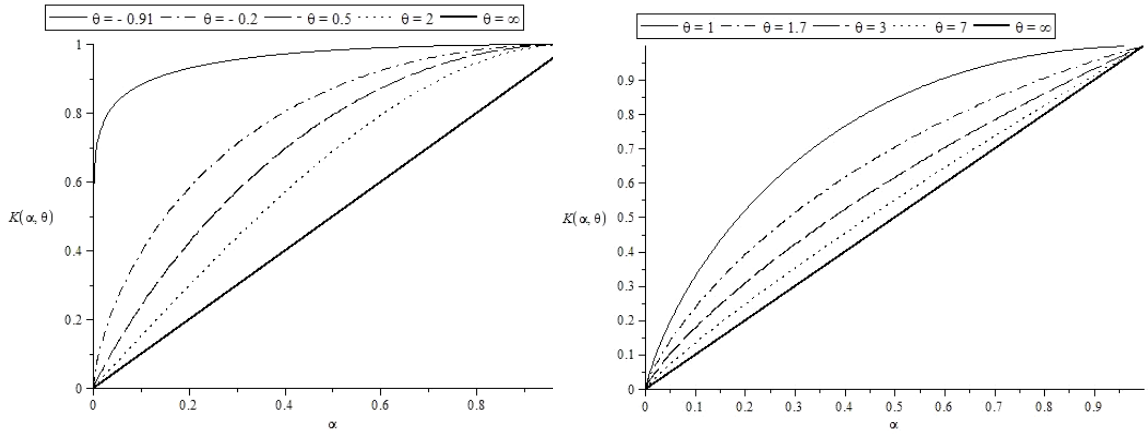

Figure 3.1: Kendall distribution $K(\cdot,\theta)$ for different values of θ in the Clayton copula case (left) and the Gumbel copula case (right). The dark full line represents the first diagonal and it corresponds to the comonotonic case.

3.2 Multivariate *Value-at-Risk*

From the usual definition in the univariate setting we know that the quantile function $Q_X(\alpha)$ provides a point which accumulates a probability α to the left tail and $1-\alpha$ to the right tail. The univariate quantile function Q_X is used in risk theory to define an univariate risk measure: the *Value-at-Risk*. This measure is defined as

$$\mathrm{VaR}_\alpha(\mathrm{X}) = \mathrm{Q_X}(\alpha), \quad \forall\ \alpha \in (0,1).$$

In the recent literature, an intuitive and immediate generalization of the VaR measure in the case of a $d-$dimensional loss distribution function F is represented by its α-quantile curves. More precisely, let $F : \mathbb{R}_+^d \to [0,1]$ be a partially increasing multivariate distribution function[4] and for any $\alpha \in (0,1)$, let $L(\alpha) = \{\mathbf{x} \in \mathbb{R}_+^d : F(\mathbf{x}) \geq \alpha\}$ be the *upper α-level set* of F. Tibiletti (1993), Embrechts and Puccetti (2006) and Nappo and Spizzichino (2009) propose to define the multivariate *Value-at-Risk* as $\partial L(\alpha)$, for $\alpha \in (0,1)$, where $\partial L(\alpha)$ denotes the boundary associated with the set $L(\alpha)$.

In the following, we will consider non-negative absolutely-continuous random vector[5] $\mathbf{X} = (X_1, \ldots, X_d)$ (with respect to Lebesgue measure λ on $\mathbb{R}^d$) with partially increasing

[4] A function $F(x_1, \ldots, x_d)$ is partially increasing on $\mathbb{R}_+^d \setminus (0,0)$ if the functions of one variable $g(\cdot) = F(x_1, \ldots, x_{j-1}, \cdot, x_{j+1}, \ldots, x_d)$ are increasing. About properties of partially increasing multivariate distribution functions we refer the interested reader to Rossi (1973), Tibiletti (1991).

[5] We restrict ourselves to $\mathbb{R}_+^d$ because, in our applications, components of $d-$dimensional vectors correspond to random losses and are then valued in $\mathbb{R}_+$.

multivariate distribution function F and such that $\mathbb{E}(X_i) < \infty$, for $i = 1, \ldots, d$. These conditions will be called *regularity conditions*.

However extensions of our results in the case of multivariate distribution function on the entire space $\mathbb{R}^d$ or in the presence of plateau in the graph of F are possible. Starting from these considerations, we introduce here a multivariate generalization of the VaR measure.

Definition 3.2.1 *Consider a random vector $\boldsymbol{X}$ satisfying the regularity conditions. For $\alpha \in (0,1)$, we define the multidimensional Value-at-Risk at probability level α by*

$$\mathrm{VaR}_\alpha(\boldsymbol{X}) = \mathbb{E}[\boldsymbol{X} \,|\, \boldsymbol{X} \in \partial L(\alpha)] = \begin{pmatrix} \mathbb{E}[\, X_1 \,|\, \boldsymbol{X} \in \partial L(\alpha)\,] \\ \vdots \\ \mathbb{E}[\, X_d \,|\, \boldsymbol{X} \in \partial L(\alpha)\,] \end{pmatrix}.$$

Analogously

$$\mathrm{VaR}_\alpha(\boldsymbol{X}) = \mathbb{E}[\boldsymbol{X} \,|\, F(\boldsymbol{X}) = \alpha] = \begin{pmatrix} \mathbb{E}[\, X_1 \,|\, F(\boldsymbol{X}) = \alpha\,] \\ \vdots \\ \mathbb{E}[\, X_d \,|\, F(\boldsymbol{X}) = \alpha\,] \end{pmatrix},$$

where $\partial L(\alpha)$ is the boundary of the α-level set of F.

From now on, we denote by $\mathrm{VaR}^1_\alpha(\mathbf{X})$, ..., $\mathrm{VaR}^d_\alpha(\mathbf{X})$ the components of the vector $\mathrm{VaR}_\alpha(\mathbf{X})$. If $\mathbf{X}$ is an exchangeable vector, then $\mathrm{VaR}^i_\alpha(\mathbf{X}) = \mathrm{VaR}^j_\alpha(\mathbf{X})$, for $i, j = 1, \ldots, d$. Furthermore, given a random variable X, since $\mathbb{E}[X \,|\, X = \mathrm{VaR}_\alpha(\mathrm{X})] = \mathrm{Q_X}(\alpha)$, for all α in $(0,1)$, Definition 3.2.1 can be viewed as a natural multivariate version of the univariate one. Moreover, the proposed generalization of *Value-at-Risk* for multivariate portfolio can be seen as a more parsimonious and synthetic measure compared to the Embrechts and Puccetti (2006)'s approach. Indeed, instead of considering the whole hyperspace $\partial L(\alpha)$ corresponding to the boundary of the α-level set of F, we only focus on the particular point in $\mathbb{R}^d_+$ that match the expectation of $\mathbf{X}$ given that $\mathbf{X}$ falls in $\partial L(\alpha)$. The latter feature could be relevant on practical grounds.

Note that, under the *regularity conditions*, $\partial L(\alpha) = \{\mathbf{x} \in \mathbb{R}^d_+ : F(\mathbf{x}) = \alpha\}$ has Lebesgue-measure zero in $\mathbb{R}^d_+$ (e.g., see Property 3 in Tibiletti, 1990). Then we make sense of Definition 3.2.1 using the limit procedure in Feller (1966), Section 3.2:

$$\begin{aligned} \mathbb{E}[\, X_i \,|\, F(\mathbf{X}) = \alpha\,] &= \lim_{h \to 0} \mathbb{E}[\, X_i \,|\, \alpha < F(\mathbf{X}) \le \alpha + h\,] \\ &= \lim_{h \to 0} \frac{\int_{Q_{X_i}(\alpha)}^{\infty} x \left(\int_\alpha^{\alpha+h} f_{(X_i, F(\mathbf{X}))}(x,y)\,\mathrm{d}y \right) \mathrm{d}x}{\int_\alpha^{\alpha+h} f_{F(\mathbf{X})}(y)\,\mathrm{d}y}, \end{aligned} \tag{3.6}$$

for $i = 1, \ldots, d$.
Dividing numerator and denominator in (3.6) by h, we obtain, as $h \to 0$

$$\mathbb{E}[X_i \,|\, F(\mathbf{X}) = \alpha] = \mathrm{VaR}^i_\alpha(\mathbf{X}) = \frac{\int_{Q_{X_i}(\alpha)}^{\infty} x\, f_{(X_i, F(\mathbf{X}))}(x, \alpha)\,\mathrm{d}x}{K'(\alpha)}, \tag{3.7}$$

for $i = 1, \ldots, d$, where $K'(\alpha) = \frac{\mathrm{d}K(\alpha)}{\mathrm{d}\alpha}$ is the Kendall distribution density function. This procedure gives a rigorous sense to our $\mathrm{VaR}_\alpha(\mathbf{X})$ in Definition 3.2.1. Remark that the existence of $f_{(X_i,F(\mathbf{X}))}$ and K' in (3.7) is guaranteed by the *regularity conditions* (for further details, see Proposition 1 in Imlahi *et al.*, 1999 or Proposition 4 in Chakak and Ezzerg, 2000).

Bivariate Archimedean copula case

In the bivariate case, the VaR introduced in Definition 3.2.1 can be computed analytically for random couple with an Archimedean copula dependence structure. Indeed, using a simple change of variable, formula (3.7) can be rewritten, for $i = 1$, as

$$\mathrm{VaR}^1_\alpha(X,Y) = \frac{\int_{Q_X(\alpha)}^{\infty} x\, f_{(U,C(U,V))}(F_X(x),\alpha)\,\mathrm{d}x}{K'(\alpha)},$$

where $f_{(U,C(U,V))}$ is the density associated with the random vector $(U, C(U,V))$. In the bivariate Archimedean case, the latter density can be obtained analytically by using a change of variable transformation[6] from (U,V) to $(U, C(U,V))$:

$$F_{(U,C(U,V))}(s,t) = t - \frac{\phi(t)}{\phi'(t)} + \frac{\phi(s)}{\phi'(t)}, \tag{3.8}$$

for $0 < t < s < 1$. The previous change of variable involves a Jacobian matrix and it can be generalized to dimension $d \geq 2$ using a multivariate transformation. Therefore, in the class of Archimedean copulas, our multivariate extension of the VaR can be computed analytically, either through a closed-form formula (see, for instance, the bivariate Clayton example below) or at least by using numerical quadratures.

Let us now consider the Clayton family of bivariate copulas. This family is interesting since it contains the counter-monotonic, the independence and the comonotonic copulas as particular cases. Let (X,Y) be a random vector distributed as a Clayton copula with parameter $\theta \geq -1$. In particular, the marginal distribution of X and Y are uniform. Then, the distribution function C_θ of (X,Y) is such that

$$C_\theta(x,y) = (\max\{x^{-\theta} + y^{-\theta} - 1, 0\})^{-1/\theta}, \quad \text{for } \theta \geq -1, \quad (x,y) \in [0,1]^2. \tag{3.9}$$

Since X and Y are exchangeable, the two components of the multivariate VaR are identical. Table 3.3 gives analytical expressions for the first (equal to the second) component of the VaR, i.e., $\mathrm{VaR}^1_{\alpha,\theta}(X,Y)$. Note that the latter can be represented as a function of the risk level α and the dependence parameter θ. For $\theta = -1$ and $\theta = \infty$ we obtain the Fréchet-Hoeffding lower and upper bounds: $W(x,y) = \max\{x+y-1, 0\}$ (counter-monotonic random variables) and $M(x,y) = \min\{x,y\}$ (comonotonic random variables) respectively. The settings $\theta = 0$ and $\theta = 1$ correspond to degenerate cases. For $\theta = 0$ we have the independence copula $\Pi(x,y) = x\,y$. For $\theta = 1$, we obtain the copula denoted by $\frac{\Pi}{\Sigma-\Pi}$ in Nelsen (1999) where $\frac{\Pi}{\Sigma-\Pi}(x,y) = \frac{x\,y}{x+y-x\,y}$.

[6] In the book by Nelsen (1999) (Corollary 4.3.5), a geometrical argument is used instead to obtain the distribution function of $(U, C(U,V))$.

Copula	θ	$\mathrm{VaR}^1_{\alpha,\theta}(X,Y)$
Clayton C_θ	$(-1,\infty)$	$\frac{\theta}{\theta-1}\frac{\alpha^\theta-\alpha}{\alpha^\theta-1}$
Counter-monotonic W	-1	$\frac{1+\alpha}{2}$
Independent Π	0	$\frac{\alpha-1}{\ln\alpha}$
$\frac{\Pi}{\Sigma-\Pi}$	1	$\frac{\alpha\ln\alpha}{\alpha-1}$
Comonotonic M	∞	α

Table 3.3: $\mathrm{VaR}^1_{\alpha,\theta}(X,Y)$ for different copula dependence structures.

Interestingly, one can readily show that $\frac{\partial \mathrm{VaR}^1_{\alpha,\theta}}{\partial\alpha} \geq 0$ and $\frac{\partial \mathrm{VaR}^1_{\alpha,\theta}}{\partial\theta} \leq 0$, for $\theta \geq -1$ and $\alpha \in (0,1)$. This proves that, for Clayton-distributed random couples, the components of our multivariate VaR are increasing functions of the risk level α and decreasing functions of the dependence parameter θ. Note also that the multivariate VaR in the comonotonic case corresponds to the vector composed of the univariate VaR associated with each component. These properties are illustrated in Figure 3.2 where $\mathrm{VaR}^1_{\alpha,\theta}(X,Y)$ is plotted as a function of the risk level α for different values of the parameter θ. Observe that an increase of the dependence parameter θ tends to lower the VaR up to the perfect dependence case where $\mathrm{VaR}^1_{\alpha,\theta}(X,Y) = \mathrm{VaR}_\alpha(X) = \alpha$. These empirical behaviors will be formally confirmed in next sections.

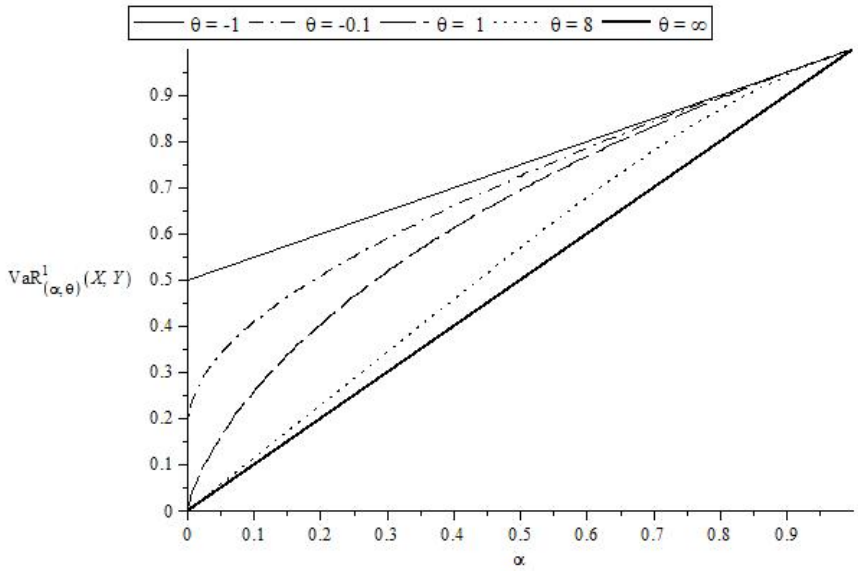

Figure 3.2: Behavior of $\mathrm{VaR}^1_{\alpha,\theta}(X,Y) = \mathrm{VaR}^2_{\alpha,\theta}(X,Y)$ with respect to risk level α for different values of dependence parameter θ. The random vector (X,Y) follows a Clayton copula distribution with parameter θ.

3.2.1 Invariance properties

Our aim in the present section is to analyze the multivariate *Value-at-Risk* introduced in Definition 3.2.1 in terms of classical invariance properties of risk measures (we refer the interested reader to Artzner *et al.*, 1999). The following proposition proves positive homogeneity and translation invariance for $\mathrm{VaR}_\alpha(\mathbf{X})$.

Proposition 3.2.2 *Consider a random vector $\boldsymbol{X}$ satisfying the regularity conditions. For $\alpha \in (0,1)$, $\mathrm{VaR}_\alpha(\boldsymbol{X})$ satisfies the following properties:*

Positive Homogeneity: $\forall\ \boldsymbol{c} \in \mathbb{R}^d_+$,

$$\mathrm{VaR}_\alpha(\boldsymbol{c}\,\boldsymbol{X}) = \boldsymbol{c}\,\mathrm{VaR}_\alpha(\boldsymbol{X}) = \begin{pmatrix} c_1\,\mathbb{E}[\,X_1 \,|\, F(\boldsymbol{X}) = \alpha\,] \\ \vdots \\ c_d\,\mathbb{E}[\,X_d \,|\, F(\boldsymbol{X}) = \alpha\,] \end{pmatrix}.$$

Translation Invariance: $\forall\ \boldsymbol{c} \in \mathbb{R}^d_+$,

$$\mathrm{VaR}_\alpha(\boldsymbol{c} + \boldsymbol{X}) = \boldsymbol{c} + \mathrm{VaR}_\alpha(\boldsymbol{X}) = \begin{pmatrix} c_1 + \mathbb{E}[\,X_1 \,|\, F(\boldsymbol{X}) = \alpha\,] \\ \vdots \\ c_d + \mathbb{E}[\,X_d \,|\, F(\boldsymbol{X}) = \alpha\,] \end{pmatrix}.$$

The proof is based on invariance properties of level sets. We refer the interested reader to Proposition 1 in Tibiletti (1993).

Proof: Let $h : \mathbb{R}^d_+ \to \mathbb{R}^d_+$, with $h(x_1, x_2, \ldots, x_d) = (h_1(x_1), h_2(x_2), \ldots, h_d(x_d))$ and $h_1, h_2, \ldots, h_d$ be increasing functions. Then we obtain

$\partial L_{h(\mathbf{X})}(\alpha) = \{\mathbf{x} \in \mathbb{R}^d_+ : F_{h(\mathbf{X})}(\mathbf{x}) = \alpha\} = \{\mathbf{x} \in \mathbb{R}^d_+ : F_{(h_1(X_1), h_2(X_2), \ldots, h_d(X_d))}(\mathbf{x}) = \alpha\}$
$= \{h(\mathbf{x}) \in \mathbb{R}^d_+ : F_{\mathbf{X}}(\mathbf{x}) = \alpha\} = \{(h_1(x_1), h_2(x_2), \ldots, h_d(x_d)) \in \mathbb{R}^d_+ : F_{\mathbf{X}}(\mathbf{x}) = \alpha\} = h(\partial L(\alpha))$.

In particular we take $h(\mathbf{x}) = \mathbf{c} + \mathbf{x}$, with $\mathbf{c} \in \mathbb{R}^d_+$. Then $\partial L_{(\mathbf{c}+\mathbf{X})}(\alpha) = \{(\mathbf{c} + \mathbf{x}) \in \mathbb{R}^d_+ : F_{\mathbf{X}}(\mathbf{x}) = \alpha\}$. Analogously with $h(\mathbf{x}) = \mathbf{c}\,\mathbf{x}$, with $\mathbf{c} \in \mathbb{R}^d_+$, $\partial L_{(\mathbf{c}\,\mathbf{X})}(\alpha) = \{(\mathbf{c}\,\mathbf{x}) \in \mathbb{R}^d_+ : F_{\mathbf{X}}(\mathbf{x}) = \alpha\}$. Then,

$$\begin{aligned} &\mathbb{E}[\,c_i + X_i \,|\, (\mathbf{c} + \mathbf{X}) \in \partial L_{(\mathbf{c}+\mathbf{X})}(\alpha)\,] = c_i + \mathbb{E}[\,X_i \,|\, h(\mathbf{X}) \in \partial L_{h(\mathbf{X})}(\alpha)\,] \\ &= c_i + \mathbb{E}[\,X_i \,|\, h(\mathbf{X}) \in h(\partial L_{\mathbf{X}})(\alpha)\,] = c_i + \mathbb{E}[\,X_i \,|\, h^{-1}(h(\mathbf{X})) \in \partial L_{\mathbf{X}}(\alpha)\,], \end{aligned}$$

for $i = 1, \ldots, d$. Hence the result. □

3.2.2 Comparison of univariate and multivariate VaR

Note that, using a change of variable, each component of the multivariate *Value-at-Risk* can be represented as an integral transformation of the associated univariate *Value-at-Risk*. Let us denote by F_{X_i} the marginal distribution functions of X_i for $i = 1, \ldots, d$ and by C the copula associated with $\mathbf{X}$. Thanks to Sklar's theorem we have $F(x_1, \ldots, x_d) = C(F_{X_1}(x_1), \ldots, F_{X_d}(x_d))$ (see Sklar, 1959). Then the random variables U_i defined by $U_i = F_{X_i}(X_i)$, for $i = 1, \ldots, d$, are uniformly distributed and their joint distribution is equal to C. Using these notations, we get:

$$\mathrm{VaR}^i_\alpha(\mathbf{X}) = \frac{1}{K'(\alpha)} \int_\alpha^1 \mathrm{VaR}_\gamma(X_i) f_{(U_i, C(\mathbf{U}))}(\gamma, \alpha)\, d\gamma, \tag{3.10}$$

for $i = 1, \ldots, d$, where $f_{(U_i, C(\mathbf{U}))}$ is the density function associated with the multivariate vector $(U_i, C(\mathbf{U}))$. The following proposition allows us to compare univariate and multivariate *Value-at-Risk*.

Proposition 3.2.3 *Consider a random vector $\boldsymbol{X}$ satisfying the regularity conditions. Assume that its multivariate distribution function F is a quasi concave*[7]*. Then, for all $\alpha \in (0,1)$, the following inequality holds*

$$\mathrm{VaR}^i_\alpha(\boldsymbol{X}) \geq \mathrm{VaR}_\alpha(X_i), \tag{3.11}$$

for $i = 1, \ldots, d$.

Proof: Let $\alpha \in (0,1)$. From the definition of the accumulated probability, it is easy to show $\partial L(\alpha)$ is inferiorly bounded by the marginal univariate quantile functions. Moreover, recall that $L(\alpha)$ is a convex set in $\mathbb{R}^d_+$ from the quasi concavity of F (see Section 2 in Tibiletti, 1995). Then, for all $\mathbf{x} = (x_1, \ldots, x_d) \in \partial L(\alpha)$, $x_1 \geq \mathrm{VaR}_\alpha(X_1), \cdots, x_d \geq \mathrm{VaR}_\alpha(X_d)$ and trivially, $\mathrm{VaR}^i_\alpha(\mathbf{X})$ is greater than $\mathrm{VaR}_\alpha(X_i)$, for $i = 1, \ldots, d$. Hence the result. □

Proposition 3.2.3 states that the multivariate $\mathrm{VaR}_\alpha(\mathbf{X})$ is a more conservative measure than the vector composed with the univariate α-*Value-at-Risk* of marginals. Furthermore, we can prove that the previous lower bound in (3.11) is reached for comonotonic random vectors.

Proposition 3.2.4 *Consider a comonotonic non-negative random vector $\boldsymbol{X}$. Then, for all $\alpha \in (0,1)$, it holds that*

$$\mathrm{VaR}^i_\alpha(\boldsymbol{X}) = \mathrm{VaR}_\alpha(X_i),$$

for $i = 1, \ldots, d$.

Proof: Let $\alpha \in (0,1)$. If $\mathbf{X} = (X_1, \ldots, X_d)$ is a comonotonic non-negative random vector then there exist a random variable Z and d increasing functions $g_1, \ldots, g_d$ such that $\mathbf{X}$ is equal to $(g_1(Z), \ldots, g_d(Z))$ in distribution. So the set $\{(x_1, \ldots, x_d) : F(x_1, \ldots, x_d) = \alpha\}$ becomes $\{(x_1, \ldots, x_d) : \min\{g_1^{-1}(x_1), \ldots, g_d^{-1}(x_d)\} = Q_Z(\alpha)\}$, where Q_Z is the quantile function of Z. Then, trivially, $\mathrm{VaR}^i_\alpha(\mathbf{X}) = \mathbb{E}[\, X_i \,|\, F(\mathbf{X}) = \alpha\,] = Q_{X_i}(\alpha)$, for $i = 1, \ldots, d$. Hence the result. □

Remark 3.2.5 *Using* (3.8) *for bivariate independent random couple (X, Y), formula* (3.10) *becomes*

$$VaR^1_\alpha(X, Y) = \frac{1}{-\ln(\alpha)} \int_\alpha^1 \frac{VaR_\gamma(X)}{\gamma}\, d\gamma,$$

then, obviously, in this case the X-related component only depends on the marginal behavior of X.

[7] A function F is quasi concave if the upper level sets of F are convex sets. Tibiletti (1995) points out families of distribution functions which satisfy the property of quasi concavity. For instance all the Archimedean copulas are quasi concave functions (see Nelsen, 1999).

3.2.3 Behavior of the multivariate VaR with respect to marginal distributions

In this section we study the behavior of our risk measure with respect to a variation of marginals. Results presented below provide natural multivariate extensions of some classical results in the univariate setting (see, e.g., Denuit and Charpentier, 2004).

Proposition 3.2.6 *Let $\boldsymbol{X}$ and $\boldsymbol{Y}$ be two d–dimensional continuous random vectors satisfying the regularity conditions and with the same copula structure C. If $X_i \stackrel{d}{=} Y_i$ then it holds that*

$$\mathrm{VaR}_\alpha^i(\boldsymbol{X}) = \mathrm{VaR}_\alpha^i(\boldsymbol{Y}), \qquad \textit{for all } \alpha \in (0,1).$$

The proof of the previous Proposition directly comes down from formula (3.10). From Proposition 3.2.6, we remark that, for a fixed copula structure C, the i-th component $\mathrm{VaR}_\alpha^i(\mathbf{X})$ does not depend on marginal distributions of the other components j with $j \neq i$.

In order to derive the next result, we use the definitions of stochastic orders presented in Section 3.1.

Proposition 3.2.7 *Let $\boldsymbol{X}$ and $\boldsymbol{Y}$ be two d–dimensional continuous random vectors satisfying the regularity conditions and with the same copula structure C. If $X_i \preceq_{st} Y_i$ then it holds that*

$$\mathrm{VaR}_\alpha^i(\boldsymbol{X}) \leq \mathrm{VaR}_\alpha^i(\boldsymbol{Y}), \qquad \textit{for all } \alpha \in (0,1).$$

The proof comes down from formula (3.10) and Definition 3.1.1. Note that, the result in Proposition 3.2.7 is consistent with the one-dimensional setting (see Section 3.3.1 in Denuit *et al.*, 2005). Indeed, as in dimension one, a stochastic order on the coordinates implies an order relation on the coordinates of *Value-at-Risk*.

As a result, in an economy with several interconnected financial institutions, capital required for one particular institution is affected by its own marginal risk. But, for a fixed dependence structure, the solvency capital required for this specific institution does not depend on marginal risks bearing by the others. Then, our multivariate VaR implies a "fair" allocation of solvency capital with respect to individual risk-taking behavior. In other words, individual financial institutions may not have to pay more for risky business activities undertook by the others.

3.2.4 Behavior of multivariate VaR with respect to the dependence structure

In this section we study the behavior of our risk measure with respect to a variation of the dependence structure, with unchanged marginal distributions.

Proposition 3.2.8 *Let $\boldsymbol{X}$ and $\boldsymbol{X}^*$ be two d–dimensional continuous random vectors satisfying the regularity conditions and with the same margins F_{X_i} and $F_{X_i^*}$, for $i = 1, \ldots, d$, and let C (resp. C^*) denote the copula function associated with $\boldsymbol{X}$ (resp. $\boldsymbol{X}^*$). Let $U_i = F_{X_i}(X_i)$, $U_i^* = F_{X_i^*}(X_i^*)$, $\boldsymbol{U} = (U_1, \ldots, U_d)$ and $\boldsymbol{U}^* = (U_1^*, \ldots, U_d^*)$.*

$$\textit{If}\ \ [U_i \,|\, C(\boldsymbol{U}) = \alpha] \preceq_{st} [U_i^* \,|\, C^*(\boldsymbol{U}^*) = \alpha] \ \ \textit{then}\ \ \mathrm{VaR}_\alpha^i(\boldsymbol{X}) \leq \mathrm{VaR}_\alpha^i(\boldsymbol{X}^*).$$

Proof: Let $U_1 \stackrel{d}{=} [U_i \,|\, C(\mathbf{U}) = \alpha]$ and $U_2 \stackrel{d}{=} [U_i^* \,|\, C^*(\mathbf{U}^*) = \alpha]$. We recall that $U_1 \preceq_{st} U_2$ if and only if $\mathbb{E}[f(U_1)] \leq \mathbb{E}[f(U_2)]$, for all non-decreasing function f, such that the expectations exist (see Denuit *et al.*, 2005; Proposition 3.3.14). We now choose $f(u) = Q_{X_i}(u)$, for $u \in (0,1)$. Then we obtain

$$\mathbb{E}[\, Q_{X_i}(U_i) \,|\, C(\mathbf{U}) = \alpha \,] \leq \mathbb{E}[\, Q_{X_i}(U_i^*) \,|\, C^*(\mathbf{U}^*) = \alpha \,],$$

But the right-hand side of the previous inequality is equal to $\mathbb{E}[\, Q_{X_i^*}(U_i^*) \,|\, C^*(\mathbf{U}^*) = \alpha \,]$ since X_i and X_i^* have the same distribution. Finally, from formula (3.10) we obtain $\mathrm{VaR}_\alpha^i(\mathbf{X}) \leq \mathrm{VaR}_\alpha^i(\mathbf{X}^*)$. Hence the result. □

We now provide an illustration of Proposition 3.2.8 in the case of bivariate Archimedean copulas.

Example 3.2.9 We consider a bivariate Archimedean copula with differentiable generator ϕ. In this case, using formula (3.8) we obtain

$$\mathbb{P}[U > u \,|\, C(U,V) = \alpha] = \frac{\phi\,(u)}{\phi\,(\alpha)}, \quad \text{for} \ \ 0 < \alpha < u < 1. \tag{3.12}$$

Let C_θ and C_{θ^*} be two Archimedean copula with dependence parameter θ and θ^*. We consider that (X,Y) and (X^*,Y^*) are two bivariate continuous random vectors with the same marginals F_X and F_Y and with copula C_θ and C_{θ^*} respectively. Let $U = F_X(X)$, $V = F_Y(Y)$, $U^* = F_X(X^*)$ and $V^* = F_Y(Y^*)$. It can be shown that an increase of dependence ($\theta \leq \theta^*$) yields a decrease of the bivariate VaR ($\mathrm{VaR}_\alpha^1(X^*,Y^*) \leq \mathrm{VaR}_\alpha^1(X,Y)$). Using formula (3.12), the condition $[U^* \,|\, C_\theta(U^*,V^*) = \alpha] \preceq_{st} [U \,|\, C_\theta(U,V) = \alpha]$ is equivalent to

$$\frac{\phi_{\theta^*}\,(u)}{\phi_{\theta^*}\,(\alpha)} \leq \frac{\phi_\theta\,(u)}{\phi_\theta\,(\alpha)}, \quad \text{for} \ \ 0 < \alpha < u < 1. \tag{3.13}$$

Interestingly, the previous condition is satisfied for all Archimedean families for which the relation $\theta \leq \theta^* \Rightarrow C_\theta(U,V) \preceq_{st} C_{\theta^*}(U^*,V^*)$ holds (see, e.g., copulas in Table 3.2). Indeed, starting from the expression of Kendall distribution function for Archimedean copulas (see Table 3.1), one can readily show that the previous relation yields (3.13). As a result, from Proposition 3.2.8, we have

$$\theta \leq \theta^* \Rightarrow \mathrm{VaR}_\alpha^1(X^*,Y^*) \leq \mathrm{VaR}_\alpha^1(X,Y). \tag{3.14}$$

Then, for copulas in Table 3.2, the multivariate VaR is non-increasing with respect to the dependence parameter θ (coordinate by coordinate). In particular, this means that limit behaviors of dependence parameters are associated with bounds for our multivariate risk measure in the case of Archimedean copula. For instance, if we denote by $\mathrm{VaR}_{(\alpha,\theta)}^1(X,Y)$ the first component of the bivariate VaR for a vector (X,Y) with a Clayton dependence structure with parameter θ, we get from Table 3.3 the following comparison result:

$$\mathrm{VaR}_{(\alpha,\infty)}^1(X,Y) \leq \mathrm{VaR}_{(\alpha,\theta)}^1(X,Y) \leq \mathrm{VaR}_{(\alpha,-1)}^1(X,Y), \tag{3.15}$$

for all $\alpha \in (0,1)$ and all $\theta \in (-1,\infty)$. Note that the lower bound corresponds to comonotonic random variables, so that $\mathrm{VaR}^1_{(\alpha,\infty)}(X,Y) = \mathrm{VaR}_\alpha(X) = \alpha$ for random variables X, Y with uniform margins (see Table 3.3). The upper bound corresponds to counter-monotonic random variables, so that $\mathrm{VaR}^1_{(\alpha,-1)}(X,Y) = \frac{1+\alpha}{2}$ for random variables X, Y with uniform margins, which turns to be also equal to $\mathrm{CTE}_\alpha(X)$ in that case, where CTE stands for the univariate *Conditional-Tail-Expectation* defined in Section 3.3 (see (3.17)).

3.2.5 Behavior of multivariate VaR with respect to risk level

In order to study the behavior of the multivariate *Value-at-Risk* with respect to risk level α, we need to introduce the *positive regression dependence* concept. For a bivariate random vector (X,Y) we mean by positive dependence that X and Y are likely to be large or to be small together. An excellent presentation of positive dependence concepts can be found in Chapter 2 of the book by Joe (1997). A positive dependence concept that will be useful in the following was called by Lehmann (1966) *positive regression dependence* (PRD) but most of the authors use the term *stochastically increasing* (SI) (see Nelsen, 1999; Section 5.2.3).

Definition 3.2.10 (Positive regression dependence) *A bivariate random vector (X,Y) is said to admit positive regression dependence with respect to X,* $\mathrm{PRD}(Y\,|\,X)$, *if*

$$[Y\,|\,X = x_1] \preceq_{st} [Y\,|\,X = x_2], \quad \forall\, x_1 \le x_2. \tag{3.16}$$

Clearly condition in (3.16) is a positive dependence notion (see Section 2.1.2 in Joe, 1997). In the following, a few examples are used to illustrate the dependence concept in Definition 3.2.10.

- Let $f(x_1,x_2) = \frac{1}{(2\pi)(1-\rho^2)^{1/2}}\,\mathrm{e}^{\frac{-\frac{1}{2}(x_1^2+x_2^2-2\rho\,x_1\,x_2)}{(1-\rho^2)}}$, be the bivariate normal density with $\rho \in (-1,1)$. Then if $\rho > 0$ we obtain $\mathrm{PRD}(X_2\,|\,X_1)$ (see Example 2.1 in Joe, 1997).

- We consider the bivariate copula (4.2.6) in Table 4.1 in Nelsen (1999): $C(u,v) = 1 - [(1-u)^\theta + (1-v)^\theta - (1-u)^\theta\,(1-v)^\theta]^{\frac{1}{\theta}}$, with $\theta \in [1,\infty)$. Then $\mathrm{PRD}(V\,|\,U)$, for each $\theta \in [1,\infty)$ (see Example 2.3 in Joe, 1997).

From Definition 3.2.10, it is straightforward to derive the following result.

Proposition 3.2.11 *Consider a d–dimensional random vector $\boldsymbol{X}$, satisfying the regularity conditions, with marginal distributions F_{X_i}, for $i = 1,\ldots,d$, and copula C. Let $U_i = F_{X_i}(X_i)$ and $\boldsymbol{U} = (U_1,\ldots,U_d)$. Then it holds that :*

If $(U_i, C(\boldsymbol{U}))$ is PRD$(U_i\,|\,C(\boldsymbol{U}))$ then $\mathrm{VaR}^i_\alpha(\boldsymbol{X})$ is a non-decreasing function of α.

The proof of Proposition 3.2.11 essentially comes down from formula (3.10) and Definition 3.2.10.

Proof: We have $[U_i\,|\,C(\mathbf{U}) = \alpha_1] \preceq_{st} [U_i\,|\,C(\mathbf{U}) = \alpha_2]$, $\forall\,\alpha_1 \le \alpha_2$. As in the proof of Proposition 3.2.8,

$$\mathbb{E}[\,Q_{X_i}(U_i)\,|\,C(\mathbf{U}) = \alpha_1\,] \leq \mathbb{E}[\,Q_{X_i}(U_i)\,|\,C(\mathbf{U}) = \alpha_2\,].$$

Then $\mathrm{VaR}^i_{\alpha_1}(\mathbf{X}) \leq \mathrm{VaR}^i_{\alpha_2}(\mathbf{X})$, for $\alpha_1 \leq \alpha_2$. Then $\mathrm{VaR}^i_{\alpha}(\mathbf{X})$ is a non-decreasing function of α. $\square$

Note that behavior of the multivariate VaR with respect to a change in the risk level does not depend on marginal distributions of $\mathbf{X}$.

Here is an illustration of Proposition 3.2.11 in the bivariate Archimedean case.

Remark 3.2.12 *Note that each component of VaR is increasing with respect to α for all random couples with an Archimedean copula dependence structure. Indeed, from formula* (3.12)*, it is straightforward to show that $\mathbb{P}[U > u\,|\,C(U,V) = \alpha]$ is a non-decreasing function of α, for all u, then PRD$(U\,|\,C(U,V))$ and PRD$(V\,|\,C(U,V))$ properties are satisfied in the bivariate Archimedean case.*

3.3 Multivariate *Conditional-Tail-Expectation*

As well as in the univariate case, the multivariate VaR at a predetermined level α does not give any information about the thickness of the upper tail of the distribution function. This is a considerable shortcoming of the VaR measure because in practice we are not only concerned with the frequency of the default but also with the severity of loss in case of default. In order to overcome this problem, another risk measure has recently received growing attentions in the insurance and finance literature: the *Conditional-Tail-Expectation* (CTE). Following Artzner *et al.* (1999) and Dedu and Ciumara (2010), for a continuous loss distribution function F_X the CTE at level α is defined by

$$\mathrm{CTE}_\alpha(X) = \mathbb{E}[\,X\,|\,X \geq \mathrm{VaR}_\alpha(X)\,], \tag{3.17}$$

where $\mathrm{VaR}_\alpha(X)$ is the univariate *Value-at-Risk* introduced above. For a comprehensive treatment and for references to the extensive literature on $\mathrm{VaR}_\alpha(X)$ and $\mathrm{CTE}_\alpha(X)$ one may refer to Denuit *et al.* (2005).

In the following, we propose a multivariate generalization of the bivariate *Conditional-Tail-Expectation*, previously introduced by Di Bernardino *et al.* (2011).

Definition 3.3.1 *Consider a d–dimensional random vector $\boldsymbol{X}$ satisfying the regularity conditions. For $\alpha \in (0,1)$, we define the multivariate α-Conditional-Tail-Expectation by*

$$\mathrm{CTE}_\alpha(\boldsymbol{X}) = \mathbb{E}[\boldsymbol{X}\,|\,\boldsymbol{X} \in L(\alpha)] = \begin{pmatrix} \mathbb{E}[\,X_1\,|\,\boldsymbol{X} \in L(\alpha)\,] \\ \vdots \\ \mathbb{E}[\,X_d\,|\,\boldsymbol{X} \in L(\alpha)\,] \end{pmatrix}.$$

Analogously,

$$\mathrm{CTE}_\alpha(\boldsymbol{X}) = \mathbb{E}[\boldsymbol{X}\,|\,F(\boldsymbol{X}) \geq \alpha] = \begin{pmatrix} \mathbb{E}[\,X_1\,|\,F(\boldsymbol{X}) \geq \alpha\,] \\ \vdots \\ \mathbb{E}[\,X_d\,|\,F(\boldsymbol{X}) \geq \alpha\,] \end{pmatrix}.$$

From now on, we denote by $\mathrm{CTE}^1_\alpha(\mathbf{X})$, ..., $\mathrm{CTE}^d_\alpha(\mathbf{X})$ the components of the vector $\mathrm{CTE}_\alpha(\mathbf{X})$. Note that this multivariate *Conditional-Tail-Expectation* is a natural extension of the univariate one. Moreover, if $\mathbf{X}$ is an exchangeable vector, $\mathrm{CTE}^i_\alpha(\mathbf{X}) = \mathrm{CTE}^j_\alpha(\mathbf{X})$, for $i, j = 1, \ldots, d$.

From Definition 3.3.1 and formula (3.7) we write, for $\alpha \in (0,1)$

$$\mathrm{CTE}^i_\alpha(\mathbf{X}) = \int_{\mathbf{x}\in L(\alpha)} x_i \frac{f_{\mathbf{X}}(x_1,\ldots,x_d)}{\overline{K}(\alpha)} \mathrm{d}x_1 \cdots \mathrm{d}x_d$$
$$= \frac{\int_\alpha^1 \left(\int_{Q_{X_i}(\gamma)}^\infty x f_{(X_i,F(\mathbf{X}))}(x,\gamma)\mathrm{d}x\right)\mathrm{d}\gamma}{\int_\alpha^1 K'(\gamma)\mathrm{d}\gamma}. \quad (3.18)$$

Formula (3.18) will be useful in Proposition 3.3.5 and Corollary 3.3.11 below.

Bivariate Archimedean copula case

We consider here a random couple (X, Y) which follows a Clayton copula distribution with parameter $\theta \geq -1$ as in (4.4). In that case, X and Y are uniformly distributed. We obtain in Table 3.4 a closed-form expression for the multivariate CTE in that case.

Copula	θ	$\mathrm{CTE}^1_{\alpha,\theta}(X,Y)$
Clayton C_θ	$(-1,\infty)$	$\frac{1}{2}\,\frac{\theta}{\theta-1}\,\frac{\theta-1-\alpha^2(1+\theta)+2\,\alpha^{1+\theta}}{\theta-\alpha(1+\theta)+\alpha^{1+\theta}}$
Counter-monotonic W	-1	$\frac{1}{4}\,\frac{1-\alpha^2+2\,\ln\alpha}{1-\alpha+\ln\alpha}$
Independent Π	0	$\frac{1}{2}\,\frac{(1-\alpha)^2}{1-\alpha+\alpha\ln\alpha}$
$\frac{\Pi}{\Sigma-\Pi}$	1	$\frac{1}{2}\,\frac{1+\alpha^2(2\,\ln\alpha-1)}{(1-\alpha)^2}$
Comonotonic M	∞	$\frac{1+\alpha}{2}$

Table 3.4: $\mathrm{CTE}^1_{\alpha,\theta}(X,Y)$ for different copula dependence structures.

Interestingly, one can readily show that $\frac{\partial \mathrm{CTE}^1_{\alpha,\theta}}{\partial\alpha} \geq 0$ and $\frac{\partial \mathrm{CTE}^1_{\alpha,\theta}}{\partial\theta} \leq 0$, for $\theta \geq -1$ and $\alpha \in (0,1)$. This proves that, for Clayton-distributed random vectors, the components of our CTE are increasing functions of the risk level α and decreasing functions of the dependence parameter θ. Note also that, in the comonotonic case, our CTE corresponds to the vector composed of the univariate CTE associated with each component. These properties are illustrated in Figure 3.3 where $\mathrm{CTE}^1_{\alpha,\theta}(X,Y)$ is plotted as a function of the risk level α for different values of θ. Note that we obtained exactly the same feature for the multivariate VaR (see Figure 3.2). The previous empirical behaviors will be formally investigated in next sections.

3.3.1 Invariance properties

As in Section 3.2.1, we analyze here the multivariate *Conditional-Tail-Expectation* in terms of classical invariance properties of risk measures (we refer the interested reader

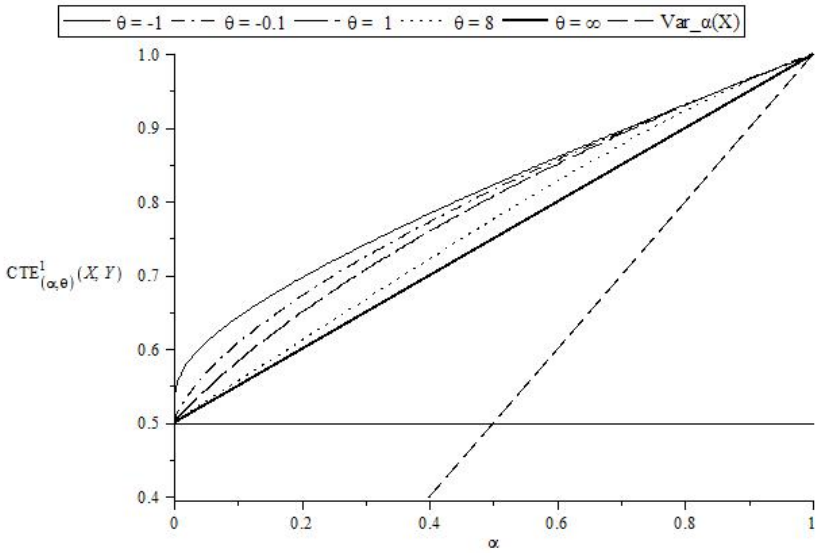

Figure 3.3: Behavior of $\mathrm{CTE}^1_{\alpha,\theta}(X,Y) = \mathrm{CTE}^2_{\alpha,\theta}(X,Y)$ with respect to risk level α for different values of dependence parameter θ. The random vector (X,Y) follows a Clayton copula distribution with parameter θ. The horizontal line corresponds to $\mathbb{E}[\,X\,] = \frac{1}{2}$.

to Artzner *et al.*, 1999). The following proposition proves positive homogeneity and translation invariance for $\mathrm{CTE}_\alpha(\mathbf{X})$.

Proposition 3.3.2 *Consider a random vector $\boldsymbol{X}$ satisfying the regularity conditions. For $\alpha \in (0,1)$, $CTE_\alpha(\boldsymbol{X})$ satisfies the following properties:*

Positive Homogeneity: $\quad \forall\ \boldsymbol{c} \in \mathbb{R}^d_+$,

$$\mathrm{CTE}_\alpha(\boldsymbol{c}\,\boldsymbol{X}) = \boldsymbol{c}\,\mathrm{CTE}_\alpha(\boldsymbol{X}) = \begin{pmatrix} c_1\,\mathbb{E}[\,X_1 \,|\, F(\boldsymbol{X}) \geq \alpha\,] \\ \vdots \\ c_d\,\mathbb{E}[\,X_d \,|\, F(\boldsymbol{X}) \geq \alpha\,] \end{pmatrix}.$$

Translation Invariance: $\quad \forall\ \boldsymbol{c} \in \mathbb{R}^d_+$,

$$\mathrm{CTE}_\alpha(\boldsymbol{c} + \boldsymbol{X}) = \boldsymbol{c} + \mathrm{CTE}_\alpha(\boldsymbol{X}) = \begin{pmatrix} c_1 + \mathbb{E}[\,X_1 \,|\, F(\boldsymbol{X}) \geq \alpha\,] \\ \vdots \\ c_d + \mathbb{E}[\,X_d \,|\, F(\boldsymbol{X}) \geq \alpha\,] \end{pmatrix}.$$

Arguments completely analogous to those in the proof of Proposition 3.2.2 are used to prove Proposition 3.3.2.

Remark 3.3.3 *For $\alpha = 0$, using both the definition of $\mathrm{CTE}_\alpha(\boldsymbol{X})$ and the definition of the α-upper level set $L(\alpha)$, we obtain*

$$\mathrm{CTE}_0(\boldsymbol{X}) = \begin{pmatrix} \mathbb{E}[\,X_1\,] \\ \vdots \\ \mathbb{E}[\,X_d\,] \end{pmatrix} = \mathbb{E}[\boldsymbol{X}].$$

Then, as in the univariate case, the multivariate CTE is equal to the expected value of the underlying random vector for $\alpha = 0$.

3.3.2 Comparison with VaR and univariate CTE

In the following, we prove a comparison result between $\mathrm{CTE}_\alpha(\mathbf{X})$ and univariate *Value-at-Risk* (see Proposition 3.3.4).

Proposition 3.3.4 *Consider a d–dimensional random vector $\boldsymbol{X}$ satisfying the regularity conditions. Assume that its multivariate distribution function F is quasi concave. Then, for any $i = 1, \ldots, d$, the following inequality holds:*

$$\mathrm{CTE}^i_\alpha(\boldsymbol{X}) \geq \mathrm{VaR}_\alpha(X_i), \qquad \textit{for all } \alpha \in (0,1).$$

The proof is completely analogous to the proof of Proposition 3.2.3.

Let us now compare multivariate CTE (see Definition 3.3.1) with multivariate VaR (see Definition 3.2.1).

Proposition 3.3.5 *Consider a d–dimensional random vector $\boldsymbol{X}$ satisfying the regularity conditions. If $VaR^i_\alpha(\boldsymbol{X})$ is a non-decreasing function of α, then it holds that*

$$CTE^i_\alpha(\boldsymbol{X}) \geq VaR^i_\alpha(\boldsymbol{X}), \qquad \textit{for all } \alpha \in (0,1).$$

Proof: Let us remark that, as in the univariate case, the multivariate *Conditional-Tail-Expectation* can be represented as an integral transformation of the multivariate *Value-at-Risk*:

$$\mathrm{CTE}^i_\alpha(\mathbf{X}) = \frac{1}{\overline{K}(\alpha)} \int_\alpha^1 K'(\gamma) \mathrm{VaR}^i_\gamma(\mathbf{X}) d\gamma. \tag{3.19}$$

Since K' is non-negative and $\mathrm{VaR}^i_\gamma(\mathbf{X})$ is assumed to be non-decreasing with respect to γ, we obtain

$$\mathrm{CTE}^i_\alpha(\mathbf{X}) \geq \tfrac{1}{\overline{K}(\alpha)} \textstyle\int_\alpha^1 K'(\gamma) \mathrm{VaR}^i_\alpha(\mathbf{X}) d\gamma = \tfrac{K(1)-K(\alpha)}{\overline{K}(\alpha)} \mathrm{VaR}^i_\alpha(\mathbf{X}) = \mathrm{VaR}^i_\alpha(\mathbf{X}),$$

for $\alpha \in (0,1)$. □

Analogously to Proposition 3.2.4, we can prove that, for comonotonic random vectors, components of the multivariate CTE are equal to univariate CTE of the corresponding marginals.

Proposition 3.3.6 *Consider a comonotonic non-negative d–dimensional random vector $\boldsymbol{X}$. Then, for all $\alpha \in (0,1)$, it holds that*

$$\mathrm{CTE}^i_\alpha(\boldsymbol{X}) = \mathrm{CTE}_\alpha(X_i),$$

for $i = 1, \ldots, d$.

Proof: Let $\alpha \in (0,1)$. If $\mathbf{X} = (X_1, \ldots, X_d)$ is a comonotonic non-negative random vector then there exist a random variable Z and d increasing functions $g_1, \ldots, g_d$ such that $\mathbf{X}$ is equal to $(g_1(Z), \ldots, g_d(Z))$ in distribution. So the set $\{(x_1, \ldots, x_d) : F(x_1, \ldots, x_d) \geq \alpha\}$ becomes $\{(x_1, \ldots, x_d) : \min\{g_1^{-1}(x_1), \ldots, g_d^{-1}(x_d)\} \geq Q_Z(\alpha)\}$. Then, trivially, $\mathrm{CTE}_\alpha^i(\mathbf{X}) = \mathbb{E}[X_i \,|\, X_i \geq Q_{X_i}(\alpha)]$, for $i = 1, \ldots, d$. Hence the result. □

Remark 3.3.7 *Consider a random couple (X, Y) with standard uniform marginals and independent copula. For all $\alpha \in (0,1)$, it holds that*

$$\mathrm{CTE}_\alpha^1(X, Y) \geq \mathrm{CTE}_\alpha(X), \tag{3.20}$$

$$\mathrm{CTE}_\alpha^2(X, Y) \geq \mathrm{CTE}_\alpha(Y). \tag{3.21}$$

Indeed $\mathrm{CTE}_\alpha^1(X, Y) - \mathrm{CTE}_\alpha(X) = -\frac{1}{2}\frac{\alpha\,(-2\,\alpha+\alpha\,\ln\alpha+2+\ln\alpha)}{-\alpha+\alpha\,\ln\alpha+1} \geq 0$, *for all* $\alpha \in (0,1)$. *Then in this particular case, coordinate by coordinate,* $\mathrm{CTE}_\alpha(X, Y)$ *is a more conservative measure than the usual univariate* CTE *of marginals.*

3.3.3 Behavior of multivariate CTE with respect to marginal distributions

In this section we study the behavior of the multivariate *Conditional-Tail-Expectation* in Definition 3.3.1 with respect to a variation of marginals. Results presented below provide a natural multivariate extension of results in the univariate setting (see, e.g., Denuit and Charpentier, 2004).

Analogously to Proposition 3.2.6, we can state the following result.

Proposition 3.3.8 *Let $\boldsymbol{X}$ and $\boldsymbol{Y}$ be two d–dimensional random vector satisfying the regularity conditions and with the same copula structure C. If $X_i \stackrel{d}{=} Y_i$, then it holds that*

$$\mathrm{CTE}_\alpha^i(\boldsymbol{X}) = \mathrm{CTE}_\alpha^i(\boldsymbol{Y}), \qquad \textit{for all } \alpha \in (0,1).$$

Proof: Let $F_1(\mathbf{x}) = F_1(x_1, \ldots, x_d) = C(F_{X_1}(x_1), \ldots, F_{X_d}(x_d))$. We recall that

$$\mathrm{CTE}_\alpha^i(\mathbf{X}) = \int_{\mathbf{x} \in L(\alpha)} x_i \, \frac{f_{\mathbf{X}}(x_1, \ldots, x_d)}{\overline{K}(\alpha)} \, \mathrm{d}x_1 \cdots \mathrm{d}x_d.$$

We remark that $\overline{K}(\alpha)$ depends solely on the copula structure (for further details see Theorem 9 in Belzunce *et al.*, 2007). Then, for a fixed copula C, $\overline{K}(\alpha)$ is invariant to a change of marginals (see Section 3.1). Let $\mathbf{u} := (u_1, \ldots, u_d)$. Using a change of variables, we obtain:

$$\mathrm{CTE}_\alpha^1(\mathbf{X}) = \left(\int_{\mathbf{u}:\, C(\mathbf{u}) \geq \alpha} Q_{X_i}(u_i) \frac{\partial^d}{\partial_1 \cdots \partial_d} C(\mathbf{u}) \, \mathrm{d}\mathbf{u} \right) \frac{1}{\overline{K}(\alpha)},$$

for $\alpha \in (0,1)$. Hence the result. $\square$

In the following we analyze how the $\mathrm{CTE}_\alpha(\mathbf{X})$ behaves when the marginal random variables increase with respect to some particular stochastic orders (see, e.g., Section 3.3 in Denuit *et al.*, 2005). Using Definitions 3.1.1-3.1.2 and Lemma 3.1.4 in Section 3.1 we can state the following result.

Proposition 3.3.9 *Let $\boldsymbol{X}$ and $\boldsymbol{Y}$ be two d–dimensional random vector satisfying the regularity conditions and with the same copula structure C. If $X_i \preceq_D Y_i$ then it holds that*

$$\mathrm{CTE}^i_\alpha(\boldsymbol{X}) \leq \mathrm{CTE}^i_\alpha(\boldsymbol{Y}), \qquad \textit{for all } \alpha \in (0,1). \tag{3.22}$$

Proof: Let us consider the function Δ defined by

$$\begin{aligned}\alpha \mapsto \Delta(\alpha) &= \overline{K}(\alpha)\left(\mathrm{CTE}^i_\alpha(\mathbf{Y}) - \mathrm{CTE}^i_\alpha(\mathbf{X})\right)\\ &= \left(\int_{\mathbf{u}:\, C(\mathbf{u})\geq\alpha} (Q_{Y_i}(u_i) - Q_{X_i}(u_i))\frac{\partial^d}{\partial_1\ldots\partial_d}C(\mathbf{u})\,\mathrm{d}\mathbf{u}\right)\frac{1}{\overline{K}(\alpha)},\end{aligned}$$

for $\alpha \in (0,1)$. Since $X_i \preceq_D Y_i$ yields $\mathbb{E}[X_i] \leq \mathbb{E}[Y_i]$ then $\Delta(0) = \mathbb{E}[Y_i] - \mathbb{E}[X_i] \geq 0$ and obviously $\lim_{t\to 1}\Delta(t) = 0$. Furthermore, from the definition of the dangerousness order relation, we observe that there exists some real number c such that $Q_{X_i}(t) \geq Q_{Y_i}(t)$, for all $t \in (0, F_{X_i}(c))$ and $Q_{X_i}(t) \leq Q_{Y_i}(t)$, for all $t \in (F_{X_i}(c), 1)$. So Δ first increases on $(0, F_{X_i}(c))$ and then decreases on $(F_{X_i}(c), 1)$. It thus remains non-negative and ensures that $\mathrm{CTE}^i_\alpha(\mathbf{X}) \leq \mathrm{CTE}^i_\alpha(\mathbf{Y})$, for all $\alpha \in (0,1)$. Hence the result. $\square$

From Proposition 3.3.9, if we analyze the marginal behavior without varying the dependence structure we have that $\mathrm{CTE}_\alpha(\mathbf{X})$ associated with a more dangerous marginal distribution will be greater. This means that this measure induces a greater coverage in the case of more dangerous risk (in the sense of the dangerousness order). Note that if $X_i \preceq_D Y_i$ for all $i = 1,\ldots,d$, we obtain $\mathrm{CTE}_\alpha(\mathbf{X}) \leq \mathrm{CTE}_\alpha(\mathbf{Y})$, for all $\alpha \in (0,1)$.

In the following example we provide an illustration of Propositions 3.3.8 and 3.3.9.

Example 3.3.10 We consider a bivariate Clayton copula with parameter 1 and five different bivariate random vectors (X, Y_i), for $i = 1,\ldots,5$. Let $X \sim \mathrm{Exp}(1)$, $Y_1 \sim \mathrm{Exp}(2)$, $Y_2 \sim \mathrm{Burr}(2,1)$, $Y_3 \sim \mathrm{Exp}(1)$, $Y_4 \sim$ Fréchet(4) and $Y_5 \sim \mathrm{Burr}(4,1)$. We assume the same copula structure for all vectors. We calculate $\mathrm{CTE}_\alpha(X, Y_i)$, for $i = 1,\ldots,5$. The results are gathered in Table 3.5. As proved in Proposition 3.3.8 we obtain an invariant property on the first coordinate of all $\mathrm{CTE}_\alpha(X, Y_i)$, for $i = 1,\ldots,5$.
Remark that $\mathrm{CTE}_\alpha(X, Y_3)$ is a case of an exchangeable continuous random vector then we have $\mathrm{CTE}^1_\alpha(X, Y_3) = \mathrm{CTE}^2_\alpha(X, Y_3)$. Furthermore, as $Q_X(\alpha) = 2\,Q_{Y_1}(\alpha)$, then $\mathrm{CTE}^1_\alpha(X, Y_1) = 2\,\mathrm{CTE}^2_\alpha(X, Y_1)$, for $\alpha \in (0,1)$ (see Table 3.5).

Since $Y_1 \preceq_D Y_5 \preceq_D Y_4 \preceq_D Y_2$, then, for any level $\alpha \in (0,1)$,

$$\mathrm{CTE}^2_\alpha(X, Y_1) \leq \mathrm{CTE}^2_\alpha(X, Y_5) \leq \mathrm{CTE}^2_\alpha(X, Y_4) \leq \mathrm{CTE}^2_\alpha(X, Y_2). \ .$$

Notice that the relation $\preceq_D$ is not transitive, since if F and G cross once, and G and H cross once, then F and H may cross twice. However, in the simple cases considered in Example 3.3.10, this situation does not happen and therefore the transitive relation is valid. Conversely Y_3, Y_4 and Y_3, Y_5 are not ordered in dangerousness sense, and also $(\mathrm{CTE}^2_\alpha(X, Y_3), \mathrm{CTE}^2_\alpha(X, Y_4))$ and $(\mathrm{CTE}^2_\alpha(X, Y_3),\ \mathrm{CTE}^2_\alpha(X, Y_4))$ are not ordered for any level $\alpha \in (0, 1)$. As before we can also verify that $\mathrm{CTE}_\alpha(X, Y_1) \leq \mathrm{CTE}_\alpha(X, Y_3) \leq \mathrm{CTE}_\alpha(X, Y_2)$, for any level $\alpha \in (0, 1)$.

α	$\mathrm{CTE}_\alpha(X, Y_1)$	$\mathrm{CTE}_\alpha(X, Y_2)$	$\mathrm{CTE}_\alpha(X, Y_3)$	$\mathrm{CTE}_\alpha(X, Y_4)$	$\mathrm{CTE}_\alpha(X, Y_5)$
0.10	(1.188, 0.594)	(1.188, 1.838)	(1.188, 1.188)	(1.188, 1.315)	(1.188, 1.229)
0.24	(1.449, 0.724)	(1.449, 2.218)	(1.449, 1.449)	(1.449, 1.431)	(1.449, 1.366)
0.38	(1.727, 0.864)	(1.727, 2.661)	(1.727, 1.727)	(1.727, 1.555)	(1.727, 1.506)
0.52	(2.049, 1.025)	(2.049, 3.235)	(2.049, 2.049)	(2.049, 1.704)	(2.049, 1.667)
0.66	(2.454, 1.227)	(2.454, 4.074)	(2.454, 2.454)	(2.454, 1.902)	(2.454, 1.876)
0.80	(3.039, 1.519)	(3.039, 5.591)	(3.039, 3.039)	(3.039, 2.219)	(3.039, 2.202)
0.90	(3.768, 1.884)	(3.768, 8.175)	(3.768, 3.768)	(3.768, 2.675)	(3.768, 2.665)
0.99	(6.102, 3.059)	(6.102, 26.59)	(6.102, 6.102)	(6.102, 4.813)	(6.102, 4.811)

Table 3.5: $\mathrm{CTE}_\alpha(X, Y_i)$, for $i = 1, \ldots, 5$, with the same copula Clayton copula with parameter 1, $X \sim \mathrm{Exp}(1)$ and $Y_1 \sim \mathrm{Exp}(2)$; $Y_2 \sim \mathrm{Burr}(2, 1)$; $Y_3 \sim \mathrm{Exp}(1)$; $Y_4 \sim$ Fréchet(4); $Y_5 \sim \mathrm{Burr}(4, 1)$.

3.3.4 Behavior of multivariate CTE with respect to risk level

As in Section 3.2.5 we study the behavior of the multivariate *Conditional-Tail-Expectation* with respect to risk level α. Also in this case the relationship between $\mathrm{CTE}_\alpha(\mathbf{X})$ and the level α will be connected to some positive dependence concepts.

Corollary 3.3.11 *Consider a d–dimensional random vector $\boldsymbol{X}$ satisfying assumptions of Proposition 3.3.5, then $\mathrm{CTE}^i_\alpha(\boldsymbol{X})$ is a non-decreasing function of risk level α.*

Proof: Let us consider the i-th coordinate $\mathrm{CTE}^i_\alpha(\mathbf{X})$. From (3.19) we have

$$\frac{d}{d\alpha}\mathrm{CTE}^i_\alpha(\mathbf{X}) = \frac{K'(\alpha)}{\overline{K}(\alpha)}\left[\mathrm{CTE}^i_\alpha(\mathbf{X}) - \mathrm{VaR}^i_\alpha(\mathbf{X})\right].$$

Using Proposition 3.3.5 the latter expression is non-negative for any level $\alpha \in (0, 1)$. □

Corollary 3.3.11 is consistant with the univariate risk measure theory. Indeed in the continuous univariate framework, the *Conditional-Tail-Expectation* is non-decreasing with respect to α (see, e.g., Denuit and Charpentier, 2004).

Proofs of Proposition 3.3.5 and Corollary 3.3.11 use direct calculations. We get below an alternative approach based to the positive dependence concepts (see Joe, 1997).

Definition 3.3.12 (Right-tail increasing) *A bivariate random vector* (X, Y) *is right-tail increasing with respect to* X, $\mathrm{RTI}(Y \mid X)$, *if*

$$[Y \mid X > x_1] \preceq_{st} [Y \mid X > x_2], \quad \forall\, x_1 \leq x_2. \tag{3.23}$$

Remark that in (3.23), Y is more likely to take large values if X increases.

Definition 3.3.13 (Right conditional increasing) *A bivariate random vector* (X, Y) *is right conditional increasing with respect to* X, $\mathrm{RCI}(Y \mid X)$, *if*

$$[Y \mid X = x] \preceq_{st} [Y \mid X > x], \quad \forall\, x.$$

For more details on these positive dependence concepts, the reader is referred to Belzunce *et al.* (2007). In particular in the case of absolutely continuous random vector (X, Y) it holds that

$$\mathrm{PRD}(Y \mid X) \Rightarrow \mathrm{RTI}(Y \mid X) \Longleftrightarrow \mathrm{RCI}(Y \mid X), \tag{3.24}$$

(see Lemma 15 in Belzunce *et al.*, 2007). Then, using Definitions 3.3.12-3.3.13, we can alternatively prove Proposition 3.3.5 and Corollary 3.3.11.

Remark 3.3.14 *We note that, if* (X, Y) *has one of the positive dependence properties mentioned above, then all random vectors with the same copula than* (X, Y) *have also the same property (see Nelsen, 1999; Corollary 5.2.11). For instance, if* $\mathrm{PRD}(Y \mid X)$ *then also* $\mathrm{PRD}(a_2(Y) \mid a_1(X))$, *for some increasing functions* a_1 *and* a_2.

In the univariate setting the *Conditional-Tail-Expectation* contains a safety loading i.e., $\mathrm{CTE}_\alpha(\mathrm{X}) \geq \mathbb{E}[\mathrm{X}]$, $\forall\, \alpha \in (0, 1)$ (see Section 2.4.3.3 in Denuit *et al.*, 2005). The safety loading should cover the fluctuations of loss experience. Corollary 3.3.15 below provides a similar property also for our multivariate CTE.

Corollary 3.3.15 *Under assumptions of Corollary 3.3.11, for all* $\alpha \in (0, 1)$, *it holds that*

$$\mathrm{CTE}^i_\alpha(\boldsymbol{X}) \geq \mathbb{E}[X_i],$$

Remark 3.3.16 *To summarize, under assumptions of Proposition 3.2.3 and Proposition 3.2.11 it holds that*

$$\mathrm{CTE}^i_\alpha(\boldsymbol{X}) \geq \mathrm{VaR}^i_\alpha(\boldsymbol{X}) \geq \mathrm{VaR}_\alpha(X_i),$$

for $\alpha \in (0, 1)$, *with* $CTE^i_\alpha(X, Y)$ *and* $VaR^i_\alpha(X, Y)$ *non-decreasing functions of* α. *We remark that this non-decreasing property of our measures depends only on the dependence structure and not on the marginal distribution functions. Furthermore, from Remark 3.2.12 we know that, in the bivariate Archimedean class of copulas, these assumptions are automatically satisfied.*

In the following we provide illustrations of Remark 3.3.16 in some particular bivariate Archimedean cases: Frank and Ali-Mikhail-Haq copulas (see Figures 3.4 and 3.5).

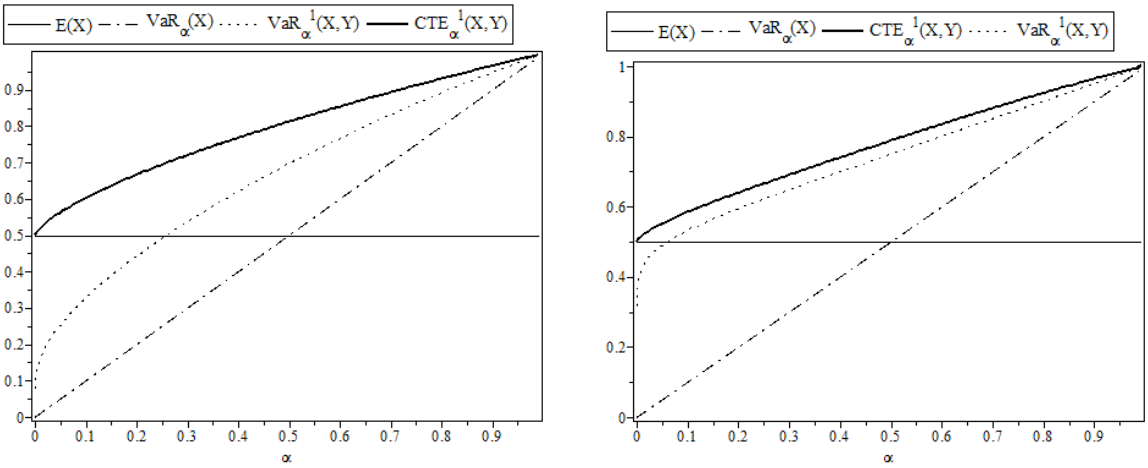

Figure 3.4: Frank copula with standard uniform marginals, parameter $\theta = 2$ (left), parameter $\theta = -10$ (right).

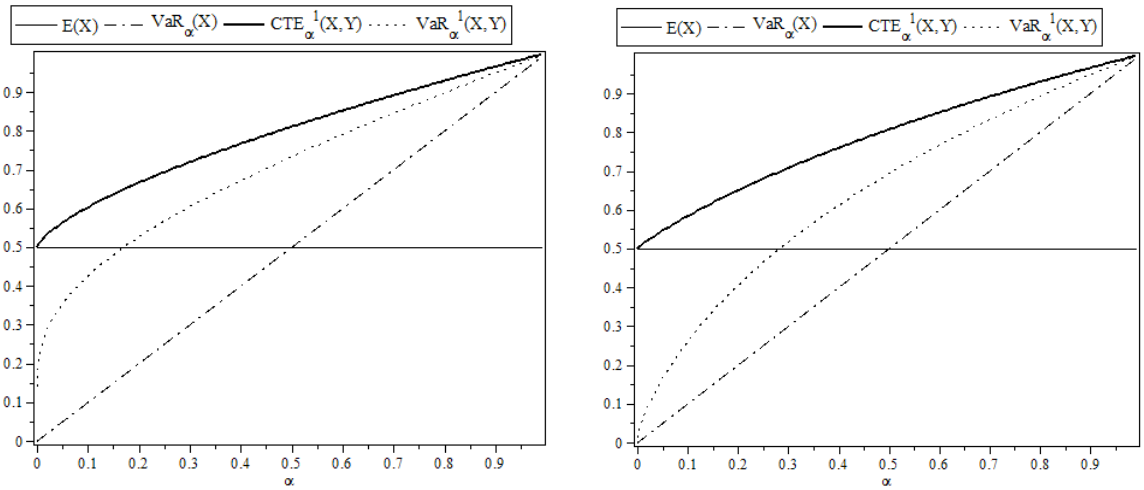

Figure 3.5: Ali-Mikhail-Haq copula with standard uniform marginals, parameter $\theta = -0.7$ (left), parameter $\theta = 0.99$ (right).

Conclusion and perspectives

In this paper, we proposed an extension of the classical Value-at-Risk and Conditional-Tail-Expectation risk measures for continuous random vectors. As in the Embrechts and Puccetti (2006)'s approach, the introduced risk measures are based on multivariate generalization of quantiles but they are able to quantify risks in a much more parsimonious and synthetic way: the risk of a d-dimensional portfolio is evaluated by a point in $\mathbb{R}_+^d$. Multivariate risk measures may be useful for some applications where risks are heterogeneous in nature or because they cannot be diversify away by an aggregation procedure.

We analyzed our multivariate risk measures in several directions. Interestingly, we showed that many properties satisfied by the univariate VaR and CTE can be translated to our proposed multivariate versions under some conditions. In particular, the proposed VaR and CTE both satisfy the positive homogeneity and the translation invariance property which are parts of the classical axiomatic properties of Artzner *et al.*, 1999. Thanks to the theory of stochastic orders, we also analyzed the effect of some risk perturbations on these measures. In the same vein as for the univariate VaR and CTE, we proved that an increase of marginal risks yield an increase of our multivariate VaR and CTE. We also gave the condition under which an increase of the risk level tends to increase components of the multivariate extensions and we show that these conditions are satisfied for bivariate families of Archimedean copulas. We also study the effect of dependence between risks on individual contribution to the multivariate risk measure and we prove that for different bivariate families of Archimedean copulas, an increase of the dependence parameter tends to lower the components of the VaR and the CTE. At the extreme case where risks are perfectly dependent or comonotonic, our multivariate risk measures is equal to the vector composed of univariate risk measures associated with each component. This feature is in line with the observation made by Zhou (2010): *"When regulating a system consisting of similar institutions, or in other words, the system is highly interconnected, considering a micro-prudential regulation can be sufficient for reducing the overall systemic risk."* (Zhou, 2010).
In a future perspective, it should be interesting to discuss the possible extensions of our risk measures to the case of discrete distribution functions, using "discrete level sets" as multivariate definitions of quantiles. For further details the reader is referred, for instance, to Laurent, 2003. Another subject of future work should be to compare our multivariate *Conditional-Tail-Expectation* and *Value-at-Risk* with existing multivariate generalizations of these measures, both theoretically and experimentally.

Acknowledgements: The authors thank Véronique Maume-Deschamps and Clémentine Prieur for their comments and help. This work has been partially supported by the French research national agency (ANR) under the reference ANR-08BLAN-0314-01. Part of this work also benefit from the support of the MIRACCLE-GICC project.

Bibliography

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

Barbe, P., Genest, C., Ghoudi, K., and Rémillard, B. (1996). On Kendall's process. *J. Multivariate Anal.*, 58(2):197–229.

Bargès, M., Cossette, H., and Marceau, E. (2009). TVaR-based capital allocation with copulas. *Insurance Math. Econom.*, 45(3):348–361.

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.

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

Bühlmann, H., Gagliardi, B., Gerber, H., and Straub, E. (1977). Some inequalities for stop-loss premiums. *Astin Bull.*, 9(2):75–83.

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

Capéraà, P., Fougères, A.-L., and Genest, C. (1997). A stochastic ordering based on a decomposition of Kendall's tau. In *Distributions with given marginals and moment problems (Prague, 1996)*, pages 81–86. Kluwer Acad. Publ., Dordrecht.

Chakak, A. and Ezzerg, M. (2000). Bivariate contours of copula. *Comm. Statist. Simulation Comput.*, 29(1):175–185.

Chaouch, M., Gannoun, A., and Saracco, J. (2009). Estimation de quantiles géométriques conditionnels et non conditionnels. *J. SFdS*, 150(2):1–27.

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.

Denuit, M. and Charpentier, A. (2004). *Mathématiques de l'assurance non-vie. Tome 1: Principes fondamentaux de thorie du risque*. Economica.

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

Denuit, M., Lefèvre, C., and Mesfioui, M. (1999). A class of bivariate stochastic orderings, with applications in actuarial sciences. *Insurance Math. Econom.*, 24(1-2):31–50. 1st IME Conference (Amsterdam, 1997).

Di Bernardino, E., Laloë, T., Maume-Deschamps, V., and Prieur, C. (2011). Plug-in estimation of level sets in a non-compact setting with applications in multivariate risk theory. *to appear in ESAIM Probability and Statistic.*

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

Feller, W. (1966). *An introduction to probability theory and its applications. Vol. II.* John Wiley & Sons Inc., New York.

Gauthier, C., Lehar, A., and Souissi, M. (2010). Macroprudential regulation and systemic capital requirements. Working Papers 10-4, Bank of Canada.

Genest, C. and Boies, J.-C. (2003). Detecting dependence with kendall plots. *The American Statistician*, 57:275–284.

Genest, C., Quessy, J.-F., and Rémillard, B. (2006). Goodness-of-fit procedures for copula models based on the probability integral transformation. *Scandinavian Journal of Statistics*, 33:337–366.

Genest, C. and Rivest, L.-P. (2001). On the multivariate probability integral transformation. *Statist. Probab. Lett.*, 53(4):391–399.

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.

Joe, H. (1997). *Multivariate models and dependence concepts*, volume 73 of *Monographs on Statistics and Applied Probability*. Chapman & Hall, London.

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

Koltchinskii, V. I. (1997). M-estimation, convexity and quantiles. *Ann. Statist.*, 25(2):435–477.

Landsman, Z. M. and Valdez, E. A. (2003). Tail conditional expectations for elliptical distributions. *N. Am. Actuar. J.*, 7(4):55–71.

Laurent, J.-P. (2003). Sensitivity analysis of risk measures for discrete distributions. *Working paper*.

Lehmann, E. L. (1966). Some concepts of dependence. *Ann. Math. Statist.*, 37:1137–1153.

Massé, J.-C. and Theodorescu, R. (1994). Halfplane trimming for bivariate distributions. *J. Multivariate Anal.*, 48(2):188–202.

Müller, A. (1997). Stop-loss order for portfolios of dependent risks. *Insurance Math. Econom.*, 21(3):219–223.

Müller, A. and Stoyan, D. (2001). *Comparison methodes for stochastic models and risks*. Wiley.

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

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

Nelsen, R. B., Quesada-Molinab, J. J., Rodríguez-Lallenac, J. A., and Úbeda-Floresc, M. (2003). Kendall distribution functions. *Statistics and Probability Letters*, 65:263–268.

Ohlin, J. (1969). On a class of measures of dispersion with application to optimal reinsurance. *ASTIN Bulletin*, 5:249–266.

Rossi, C. (1973). Sulle curve di livello di una superficie di ripartizione in due variabili. *Giornale dell'Istituto Italiano degli Attuari*, 36:87–108.

Serfling, R. (2002). Quantile functions for multivariate analysis: approaches and applications. *Statist. Neerlandica*, 56(2):214–232. Special issue: Frontier research in theoretical statistics, 2000 (Eindhoven).

Sklar, M. (1959). Fonctions de répartition à n dimensions et leurs marges. *Publ. Inst. Statist. Univ. Paris*, 8:229–231.

Tasche, D. (2008). Capital allocation to business units and sub-portfolios: the Euler Principle. Quantitative finance papers, arXiv.org.

Tibiletti, L. (1990). Gli insiemi di livello delle funzioni di ripartizione n-dimensionali: un'applicazione statistica (Level sets of n-dimensional distribution functions: a statistic application). *Quaderni dell'Istituto di Matematica Finanziaria dell'università di Torino: terza serie*, 56:1–11.

Tibiletti, L. (1991). Sulla quasi concavità delle funzioni di ripartizione n-dimensionali - on quasi-concavity of n-dimensional distribution functions. In *Atti del XV convegno A.M.A.S.E.S.*, pages 503–515.

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

Tibiletti, L. (1995). Quasi-concavity property of multivariate distribution functions. *Ratio Mathematica*, 9:27–36.

Wei, G. and Hu, T. (2002). Supermodular dependence ordering on a class of multivariate copulas. *Statist. Probab. Lett.*, 57(4):375–385.

Zhou, C. (2010). Why the micro-prudential regulation fails? The impact on systemic risk by imposing a capital requirement. Dnb working papers, Netherlands Central Bank, Research Department.

Appendix: Gumbel copula case

We consider Gumbel bivariate family of copulas:

$$C_\theta(x,y) = \mathrm{e}^{-\left((-\ln x)^\theta + (-\ln y)^\theta\right)^{\frac{1}{\theta}}},$$

for $\theta \in [1,\infty)$, $(x,y) \in [0,1]^2$ (e.g. see Section 3.3.1 in Nelsen, 1999) and X, Y standard uniform marginals. For $\theta = 1$ we have the independent copula $C_1(x,y) = \Pi(x,y) = x\,y$; for $\theta = \infty$ the Fréchet bound $M(x,y) = \min\{x,y\}$ (comonotonic random variables). In this case we get

$$\mathrm{CTE}_\alpha(X,Y) = \begin{pmatrix} t(\alpha,\theta) \\ t(\alpha,\theta) \end{pmatrix},$$

with

$$t(\alpha,\theta) = \frac{1}{2}\,\frac{\ln\alpha(\theta+\theta^2-\alpha^2\theta-\alpha^2\theta^2)+2\,\alpha^{\frac{3}{2}}(-\ln\alpha)^{-\frac{\theta}{2}+2}\mathrm{WM}\left(\frac{1}{2}\,\theta,\frac{1}{2}\,\theta+\frac{1}{2},-\ln\alpha\right)+2\,\alpha^2(\ln\alpha)^2+2\,\alpha^2(\ln\alpha)^2\theta}{(1+\theta)\ln\alpha(\theta-\alpha\,\theta+\alpha\,\ln\alpha)},$$

where WM is the Whittaker function $\mathrm{WM}(\mu,\nu,z) = \exp^{-\frac{1}{2}z}\,z^{\frac{1}{2}+\nu}\mathrm{Hpy}(\frac{1}{2}+\nu-\mu, 1+2\nu; z)$ and Hpy is the Kummer's confluent hypergeometric function.

Remark that X and Y are exchangeable, then obviously $\mathrm{CTE}^1_{\alpha,\theta}(X,Y) = \mathrm{CTE}^2_{\alpha,\theta}(X,Y)$. Furthermore $\frac{\partial \mathrm{CTE}^1_{\alpha,\theta}}{\partial\alpha} \geq 0$ and $\frac{\partial \mathrm{CTE}^1_{\alpha,\theta}}{\partial\theta} \leq 0$, for $\theta \geq 1$ and $\alpha \in (0,1)$. In analogy with Figure 3.3, in Figure 3.6 we propose a graphical illustration of $\mathrm{CTE}^1_{\alpha,\theta}(X,Y)$, in Gumbel copula case, for different values of the dependence parameter θ and with respect to risk level α.

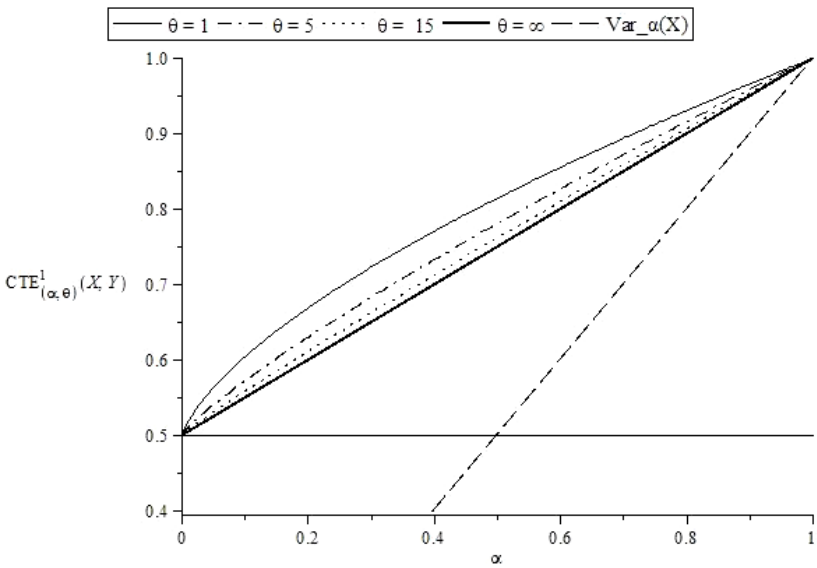

Figure 3.6: Behavior of $\mathrm{CTE}^1_{\alpha,\theta}(X,Y) = \mathrm{CTE}^2_{\alpha,\theta}(X,Y)$ with respect to risk level α for different values of dependence parameter θ. The random vector (X,Y) follows a Gumbel copula distribution with parameter θ. The horizontal line corresponds to $\mathbb{E}[\,X\,] = \frac{1}{2}$.

Chapitre 4

Estimation de type plug-in des ensembles de niveau dans un cadre non-compact avec applications à la théorie des risques

Résumé

Ce chapitre est composé d'un article en anglais, intitulé *Plug-in estimation of level sets in a non-compact setting with applications in multivariate risk theory*, coécrit avec Thomas Laloë, Véronique Maume-Deschamps et Clémentine Prieur et accepté pour publication dans la revue *ESAIM : Probability and Statistics*.

Ce chapitre est subdivisé en deux parties. Tout d'abord nous nous intéressons à l'estimation des ensembles de niveau d'une fonction de distribution inconnue F sur $\mathbb{R}_+^2$, c'est-à-dire des ensembles du type

$$L(c) = \{(x,y) \in \mathbb{R}_+^2 : F(x,y) \geq c\}, \text{ avec } c \in (0,1).$$

Pour ce faire nous suivons une méthode de type *plug-in*. C'est-à-dire étant donné un estimateur F_n de F, nous estimons $L(c)$ par $L_n(c) = \{(x,y) \in \mathbb{R}_+^2 : F_n(x,y) \geq c\}$. Notre cadre de travail nous impose de travailler avec des ensembles de niveau non compacts. Ce point représente un aspect délicat et intéressant (voir Hypothèse **H** et Proposition 4.1.1). Nous proposons des résultats de consistance pour notre estimateur plug-in par rapport à deux critères :

- la distance Hausdorff (voir Théorème 4.2.1) ;
- le volume de la différence symétrique (voir Théorème 4.3.1).

Dans la seconde partie de ce chapitre nous nous intéressons à une application de notre estimateur $L_n(c)$ dans le cadre de la théorie du risque multidimensionnelle. En particulier nous reprenons la version bivariée de l'Espérance de queue (*Conditional-Tail-Expectation*, CTE) en conditionnant le vecteur aléatoire bidimensionnel à être dans l'ensemble de niveau $L(c)$, introduite dans le chapitre précédent :

$$\mathrm{CTE}_\alpha(X,Y) = \mathbb{E}[(X,Y) \,|\, (X,Y) \in L(\alpha)] = \begin{pmatrix} \mathbb{E}[\,X \,|\, (X,Y) \in L(\alpha)\,] \\ \mathbb{E}[\,Y \,|\, (X,Y) \in L(\alpha)\,]. \end{pmatrix}.$$

Nous donnons une procédure consistante pour estimer cette mesure (voir Théorème 4.4.3). En particulier un résultat de consistance, dans le cas où l'estimateur F_n est l'estimateur empirique de la fonction de distribution F, est proposé (voir Corollaire 4.4.4).

Nous présentons des exemples avec des données simulées et réelles qui illustrent nos résultats théoriques (voir Sections 4.5.1 et 4.5.2).

Présentation des approfondissements donnés en Annexe

En Annexe nous présentons des illustrations et d'autres études, inhérentes à notre procédure d'estimation des ensembles de niveau, qui n'ont pas été insérées dans l'article.

- La frontière des ensembles de niveau $\partial L(c)$ est appelée dans la littérature *courbe quantile* (voir par exemple Embrechts et Puccetti, 2006 ou Nappo et Spizzichino, 2009). Dans la Section B.8 nous proposons quelques formes paramétriques exactes et quelques graphiques afin de fournir des illustrations de $\partial L(c)$ dans des cas bivariés spécifiques.

- Un des développements les plus intéressants et assez naturel de ce Chapitre 4 est l'étude du comportement de la $\mathrm{CTE}_\alpha(\mathrm{X},\mathrm{Y})$ pour des valeurs "extrêmes" du niveau de risque. En particulier il serait intéressant de considérer l'estimation de la $\mathrm{CTE}_\alpha(\mathrm{X},\mathrm{Y})$ en utilisant un estimateur de la queue de la distribution. Dans cette optique, dans la Section B.9, nous proposons quelque estimations, sur des données réelles et simulées, de la $\mathrm{CTE}_\alpha(\mathrm{X},\mathrm{Y})$ pour des hauts niveaux de risque en utilisant notre estimateur pour la queue $\widehat{F}^*(x,y)$ présenté dans le Chapitre 2 (voir (2.31)). Ces considérations représentent des pistes de recherche ouvertes pour de futurs travaux.

Introduction

The Conditional Tail Expectation (CTE) is a commonly used tool in the univariate risk theory. Given a random variable X with continuous and strictly monotonic distribution function F_X, the CTE at level α is defined by

$$\mathrm{CTE}_\alpha(X) = \mathbb{E}[\, X \mid X \geq Q_X(\alpha)\,], \tag{4.1}$$

with $\alpha \in (0,1)$ and $Q_X(\alpha)$ the univariate quantile function of X at level α. In other words the $\mathrm{CTE}_\alpha(X)$ is defined by the expectation value of X conditionally to the fact that X belongs to a certain one-dimensional level set of F_X. For further details about this univariate risk measure we refer the reader e.g. to Denuit *et al.* (2005).

In this paper we are interested to introduce a new bivariate generalization of CTE in (4.1) and to provide a consistent procedure to estimate it. Consequently we essentially need a generalization of the notion of quantile in dimension higher than one. Therefore we remark that the quantile function is a relevant tool in statistics and probability, but its practical use is mostly confined to the univariate case where it is easily interpretable and calculable. The bivariate (or multivariate) characterization has received some attention however, it is much less common, due to a number of theoretical and practical reasons. In particular the main difficulty is that there does not exist a natural ordering in n-dimensions, for $n > 1$. For this reason, in the last decade several attempts have been made to provide a multidimensional generalization of the univariate quantile function. We refer to Serfling (2002) for a complete review on the topic. For example Massé and Theodorescu (1994) defined multivariate quantiles as halfplanes and Koltchinskii (1997) provided a general treatment of multivariate quantiles as inversions of mappings.

As a possible generalization of the univariate quantile function, Tibiletti (1993), Fernández-Ponce and Suárez-Lloréns (2002) and Belzunce *et al.* (2007) defined a multivariate quantile as a set of points which accumulate the same probability for a fixed orthant. They called it level curves or quantile curves. In this sense de Haan and Huang (1995) model a risk-problem of flood in the bivariate setting using an estimator of level curves of bivariate distribution function. In this paper we will follow this approach in order to propose a new risk measure which generalizes the CTE in (4.1). To this end we need to estimate the level sets of a bivariate distribution function F. This leads us into the general field of level sets estimation.
The problem of estimating level sets of an unknown function (for instance of a density function and more recently a regression function) has received attention recently. In particular the estimation of density level sets has been studied in Polonik (1997), Tsybakov (1997), Baíllo *et al.* (2001), Baíllo (1999), Cadre (2006), Rigollet and Vert (2009), Biau *et al.* (2007). The estimation of regression level sets in a compact setting has been analyzed in Cavalier (1997), Laloë (2009). An alternative approach, based on the geometric properties of the compact support sets, has been presented by Hartigan (1987), Cuevas and Fraiman (1997), Cuevas and Rodríguez-Casal (2004). The problem of estimating general level sets under compactness assumptions has been discussed by

Cuevas *et al.* (2006). The asymptotic behavior of minimum volume sets and of a generalized quantile process is analyzed by Polonik (1997).

Considering the level sets of a bivariate distribution function, the commonly assumed property of compactness for these sets is no more reasonable. Then, differing from the literature cited above, we need to work in a non-compact setting and this requires special attention in the statement of our problem.

Our general approach will be the following: first, we provide a consistent estimator of the level set

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

Therefore we consider a *plug-in* approach (e.g. see Baíllo *et al.* (2001); Rigollet and Vert (2009); Cuevas *et al.* (2006)), 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. 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.

In order to provide a consistency results we define two proximity criteria between sets. A standard choice is the volume of the symmetric difference. Another natural criterion is given by the Hausdorff distance that corresponds to an intuitive notion of "physical proximity" between sets. Our consistency properties are stated with respect to these two criteria under reasonable assumptions on F and F_n (Theorems 4.2.1 - 4.3.1). These results are based on a slight modification of Proposition 3.1 in the PhD Thesis of Rodríguez-Casal (2003) (Proposition 4.1.1).

We then introduce our new bivariate version of the Conditional Tail Expectation by conditioning the two-dimensional random vector to be in the level set $L(c)$. This new risk measure is based on the bivariate Value-at-Risk proposed in Embrechts and Puccetti (2006) and Nappo and Spizzichino (2009) and it is different from the existing bivariate generalizations of the univariate Conditional Tail Expectation (e.g. see Cai and Li, 2005). Indeed we do not use any aggregate risk in the portfolio in order to analyze the bivariate risk's issue. Our bivariate version of the CTE deals with the simultaneous joint damages considering the dependence structure of data in a bivariate specific risk's area ($L(c)$). We propose an estimator for this new risk measure using plug-in approach for level sets and provide a consistency result (Theorem 4.4.3).

The paper is organized as follows. We introduce some notation, tools and technical assumptions in Section 4.1. Consistency and asymptotic properties of our estimator of $L(c)$ are given in Sections 4.2 and 4.3. In Section 4.4 we introduce a new bivariate generalization of CTE in (4.1) and we provide a consistent estimator for this risk measure. Illustrations with simulated and real data are presented in Section 4.5. Section 4.6 summarizes and briefly mentions directions for future research. Finally, some auxiliary results and more technical proofs are postponed to Section 4.7.

4.1 Notation and preliminaries

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

Let $\mathbb{N}^* = \mathbb{N} \setminus \{0\}$, ${\mathbb{R}_+^2}^* = \mathbb{R}_+^2 \setminus (0,0)$, $\mathcal{F}$ the set of continuous distribution functions $\mathbb{R}_+^2 \to [0,1]$ 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 by $F_n(\cdot) = F_n(X_1, X_2, \ldots, X_n, \cdot)$ an estimator of F based on the finite sample $(X_1, X_2, \ldots, X_n)$. We restrict ourselves to $\mathbb{R}_+^2$ for convenience but the following results are completely adaptable in $\mathbb{R}^2$. This choice is motivated essentially by our application in risk theory proposed in Section 4.4, where random variables will be losses then defined in a positive support.

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

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

and

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

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

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

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

Furthermore, for any $A \subset \mathbb{R}_+^2$ we denote by ∂A its boundary.

Note that $\{F = c\}$ can be a portion of quadrant $\mathbb{R}_+^2$ instead of a set of Lebesgue measure null in $\mathbb{R}_+^2$ (that is the presence of plateau at level c). In the following we will introduce suitable conditions in order to avoid the presence of plateau in the graph of F.

In the metric space $(\mathbb{R}_+^2, d)$, where d stands for the Euclidean distance, we denote by $B(x,\rho)$ the closed ball centered on x and with positive radius ρ. Let $B(S,\rho) = \bigcup_{x \in S} B(x,\rho)$, with S a closed set of $\mathbb{R}_+^2$. For $r > 0$ and $\zeta > 0$, define

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

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.

For sake of completeness, we recall that if A_1 and A_2 are compact sets in $(\mathbb{R}_+^2, d)$, the Hausdorff distance between A_1 and A_2 is defined by

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

or equivalently by

$$d_H(A_1, A_2) = \max\left\{\sup_{x\in A_1} d(x, A_2), \sup_{x\in A_2} d(x, A_1)\right\},$$

where $d(x, A_2) = \inf_{y\in A_2} \| x - y \|$.

The above expression is well-defined even when A_1 and A_2 are just closed (not necessarily compact) sets but in this case the value $d_H(A_1, A_2)$ could be infinity. Then in our setting, in order to avoid these situations we need a truncated version of Assumption T in Cuevas *et al.* (2006) or Tsybakov (1997). More precisely we introduce the following assumption.

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 **H** is satisfied under mild conditions. Proposition 4.1.1 below is a slight modification of Proposition 3.1 in the PhD Thesis of Rodríguez-Casal (2003) in order to deal with non-compact sets.

Proposition 4.1.1 *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* **H**, *with $A = \frac{2}{m^\triangledown}$.*

The proof is postponed to Section 4.7.

Remark 7 Under assumptions of Proposition 4.1.1, F is continuous and $m^\triangledown > 0$, there is no plateau in the graph of F for each level t such that $|t - c| \leq r$. Furthermore from Theorem 1 in Rossi (1976) we know that each half-line in $\mathbb{R}^2_+$, parallel to one of the axis, meets $\{F = t\}$ in a connected set of points. As a consequence we obtain that $\{F = t\}$ is a curve in the quadrant $\mathbb{R}^2_+$. In particular, from $m^\triangledown > 0$, for a fixed x we have to consider all corresponding values y in a specific interval (we refer the interested reader to Remark 2.1 in Rossi, 1976 or Remark 3.1 in Rossi, 1973). In this case the plane curve $\{F = t\}$ is not the graph of a function but it 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 $x \geq x'$. In particular 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 $\mathbb{R}^2_+$. Finally under assumptions of Proposition 4.1.1 we obtain $\partial L(c)^T = \{F = c\}^T = \{F = c\} \cap [0, T]^2$ (we refer for details to Theorem 3.2 in Rodríguez-Casal, 2003).

4.2 Consistency in terms of the Hausdorff distance

In this section we study the consistency properties of $L_n(c)^T$ with respect to the Hausdorff distance between $\partial L_n(c)^T$ and $\partial L(c)^T$. The Hausdorff distance corresponds to an

intuitive notion of "physical proximity" between sets (see Figure 4.1(left) for a graphical illustration). However the metric d_H is not always completely successful in capturing the shape properties: two sets can be very close in d_H and still have quite different shapes. For instance in Figure 4.1(right) we have X, the outside ball, and Y the set of balls inside X. The distance $d_H(X, Y)$ is small even if X and Y are quite different in many important aspects.

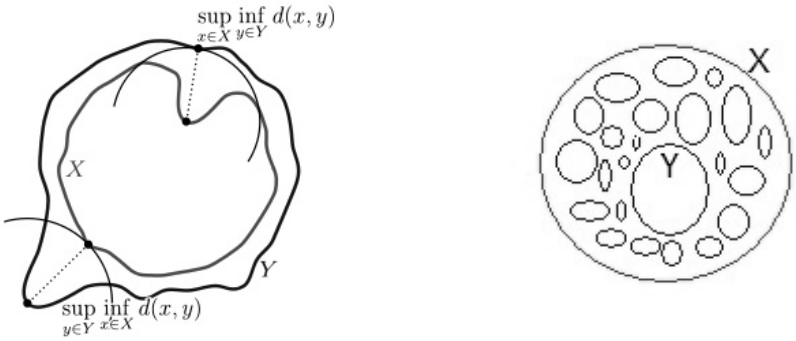

Figure 4.1: (left) Hausdorff distance between sets; (right) particular case where $d_H(X, Y)$ does not capture the shape properties of X and Y.

In order to avoid these situations, following Cuevas and Rodríguez-Casal (2004) and Cuevas *et al.* (2006), a way to reinforce the notion of visual proximity between two sets provided by d_H is to impose the proximity of the respective boundaries. We will follow this approach in Theorem 4.2.1 below.

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

$$\|F - F_n\|_\infty = \sup_{x \in \mathbb{R}_+^2} \mid F(x) - F_n(x) \mid,$$

and for $T > 0$

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

The following result can be considered an adapted version of Theorem 2 in Cuevas *et al.* (2006).

Theorem 4.2.1 *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 $t : \mid t - c \mid \leq r$, $\partial L(t)^{T_1} \neq \emptyset$. Let $(T_n)_{n\in\mathbb{N}^*}$ be an increasing sequence of positive values. Assume that, for each n and for almost all samples of size n, F_n is a continuous function and that*

$$\|F - F_n\|_\infty \to 0, \quad a.s. \tag{4.2}$$

Then

$$d_H(\partial L(c)^{T_n}, \partial L_n(c)^{T_n}) = O(\|F - F_n\|_\infty), \qquad a.s.$$

The proof is postponed to Section 4.7. Under assumptions of Theorem 4.2.1, $d_H(\partial L(c)^{T_n}, \partial L_n(c)^{T_n})$ converges to zero and the quality of our plug-in estimator is obviously related to the quality of the estimator F_n.

Remark 8 Theorem 4.2.1 provides an asymptotic result for a fixed level c. In particular following the proof of Theorem 4.2.1 (postponed to Section 4.7) we remark that, for n large enough,

$$d_H(\partial L(c)^{T_n}, \partial L_n(c)^{T_n}) \leq 6\, A\, \|F - F_n\|_\infty^{T_n}, \quad a.s.,$$

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

Note that the empirical distribution estimator does not satisfy continuity assumption imposed in Theorem 4.2.1. However, in order to overcome this problem it can be considered a smooth version of this estimator (e.g. see Chaubey and Sen, 2002), which satisfies convergence in (4.2). However we remark that the continuity assumption, imposed in Theorem 4.2.1, will not be necessary in the following results.

4.3 L_1 consistency

The previous section was devoted to the consistency of L_n in terms of the Hausdorff distance. We consider now another consistency criterion: the consistency 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 $\mathbb{R}_2^+$ by

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

where λ stands for the Lebesgue measure on $\mathbb{R}^2$ 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

$$v_n \int_{[0,T_n]^2} \mid F - F_n \mid^p \ \lambda(\mathrm{d}x) \underset{n\to\infty}{\overset{\mathbb{P}}{\to}} 0,$$

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

We now establish our consistency result in terms of the volume of the symmetric difference. We can interpret the following theorem as a generalization of Theorem 3 in Cuevas *et al.* (2006) (Section 3) in the case of non-compact level sets. In addition, we provide also a convergence rate for the symmetric difference between $L(c)^{T_n}$ and $L_n(c)^{T_n}$.

Theorem 4.3.1 *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 $M_H < \infty$. 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 and that for all $t : \mid t - c \mid \leq r$, $\partial L(t)^{T_1} \neq \emptyset$. Then it holds that*

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

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

The proof is postponed to Section 4.7. Theorem 4.3.1 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. Note that, as in Theorem 3 in Cuevas *et al.* (2006), Theorem 4.3.1 above does not require any continuity assumption on F_n.

Remark 9 Assumptions on gradient vector and Hessian matrix of F, in Theorems 4.2.1 and 4.3.1, are satisfied for a quite large class of classical bivariate distribution functions; for instance independent copula and exponential marginals, Farlie-Gumbel-Morgenstern (FGM), Clayton or Survival Clayton copulas and Burr marginals.

We now suppose that there exists a positive increasing sequences $(v_n)_{n\in\mathbb{N}^*}$ such that $v_n \, \|F - F_n\|_\infty \underset{n\to\infty}{\overset{\mathbb{P}}{\to}} 0$. Then

$$\forall \; p \geq 1, \quad w_n \int_{[0,T_n]^2} \mid F - F_n \mid^p \; \lambda(\mathrm{d}x) \underset{n\to\infty}{\overset{\mathbb{P}}{\to}} 0, \qquad \text{with} \qquad w_n = \frac{v_n^p}{T_n^2}.$$

In this case, from Theorem 4.3.1 we get

$$p_n = o(w_n^{\frac{1}{p+1}} / T_n^{\frac{p}{p+1}}) = o(v_n^{\frac{p}{p+1}} / T_n^{\frac{p+2}{p+1}}). \tag{4.3}$$

From these considerations we can introduce the following remark.

Remark 10 Let F_n the bivariate empirical distribution function. Then it holds that $v_n \, \|F - F_n\|_\infty \underset{n\to\infty}{\overset{\mathbb{P}}{\to}} 0$, with $v_n = o(\sqrt{n})$. From Theorem 4.3.1 and using the convergence rate in (4.3) with $p = 2$, we obtain $p_n = o(n^{\frac{1}{3}} / T_n^{\frac{4}{3}})$. For instance, for $F(x,y) = (1 - \exp^{-x})\,(1 - \exp^{-2\,y})$ and using the bivariate empirical distribution function F_n we obtain $p_n = o(n^{\frac{1}{3}} / \ln(n)^{\frac{4}{3}})$, with $p = 2$ and $T_n = \ln(n)$.

4.4 Application in bivariate risk theory

From the usual definition in the univariate setting we know that the quantile function Q_X provides a point which accumulates a probability α to the left tail and $1 - \alpha$ to the

right tail. More precisely, given an univariate continuous and strictly monotonic loss distribution function F_X,

$$Q_X(\alpha) = F_X^{-1}(\alpha), \quad \forall\, \alpha \in (0,1). \tag{4.4}$$

The notion of univariate quantile function Q_X is used in risk theory to define an univariate measure of risk: the Value-at-Risk (VaR). This measure is defined as

$$\mathrm{VaR}_\alpha(X) = Q_X(\alpha), \quad \forall\, \alpha \in (0,1). \tag{4.5}$$

As presented in the introduction of this paper, following the general ideas of Embrechts and Puccetti (2006) and Nappo and Spizzichino (2009) an intuitive generalization of the VaR measure in the case of a bidimensional loss distribution function F is given by its α-quantile curves. More precisely:

Definition 4.4.1 *For $\alpha \in (0,1)$ and $F \in \mathcal{F}$, the bidimensional Value-at-Risk at probability level α is the boundary of its α-level set, i.e. $Q_\alpha(F) = \partial L(\alpha)$.*

As pointed out in Tibiletti (1993), imposing $\alpha = \frac{1}{2}$ in Definition 4.4.1, we get a natural extension of the bidimensional median. For details about a parametric formulation of the quantile curve $\partial L(\alpha)$ see Belzunce *et al.* (2007). For details about its properties see for instance Fernández-Ponce and Suárez-Lloréns (2002) (and references therein) and Nappo and Spizzichino (2009).

Then, using a bivariate estimator F_n as in Sections 4.2 - 4.3, we can define our plug-in estimator of the bivariate Value-at-Risk by

$$Q_\alpha(F_n) = \partial L_n(\alpha), \quad \text{for } \alpha \in (0,1).$$

Moreover, under assumptions of Theorem 4.2.1, we obtain a consistency result, with respect to the Hausdorff distance, for the $Q_\alpha(F_n)$ on the quadrant $\mathbb{R}_+^2$ i.e.

$$d_H(Q_\alpha(F)^{T_n}, Q_\alpha(F_n)^{T_n}) = O(\|F - F_n\|_\infty), \quad a.s.$$

As in the univariate case, the two-dimensional VaR at a predetermined level α does not give any information about thickness of the upper tail of the distribution function. This is a considerable shortcoming of VaR measure because in practice we are not only concerned with frequency of the default but also with the severity of loss in case of default. In other words we are interested to analyze the behavior of X and Y not only on the boundary but also in the whole α-level set.

In dimension one, in order to overcome this problem, another risk measure has recently received growing attention in insurance and finance literature: Conditional Tail Expectation (CTE). Following Artzner *et al.* (1999) and Dedu and Ciumara (2010), for a continuous loss distribution function F_X the CTE at level α is defined by

$$\mathrm{CTE}_\alpha(X) = \mathbb{E}[\, X \,|\, X \geq \mathrm{VaR}_\alpha(X)\,],$$

where $\mathrm{VaR}_\alpha(X)$ is the univariate VaR in (4.5). For a comprehensive treatment and for references to the extensive literature on $\mathrm{VaR}_\alpha(X)$ and $\mathrm{CTE}_\alpha(X)$ one may refer to Denuit *et al.* (2005).

Several bivariate generalizations of the classical univariate CTE have been proposed; mainly using as conditioning events the total risk or some univariate aggregate extreme risk in the portfolio (for instance the sum, the max or the min between two risks):

$$\mathbb{E}[\,(X,Y)\,|\,X+Y > Q_{X+Y}(\alpha)\,], \quad \mathbb{E}[\,(X,Y)\,|\,\min\{X,Y\} > Q_{\min\{X,Y\}}(\alpha)\,]$$

$$\text{and} \quad \mathbb{E}[\,(X,Y)\,|\,\max\{X,Y\} > Q_{\max\{X,Y\}}(\alpha)\,].$$

The interested reader is referred to Cai and Li (2005) for further details. Using the same approach of Definition 4.4.1, we are interested to preserve the complete information about dependence structure between X and Y and the bi-dimensional aspect of our problem. To this end, starting from these general considerations, we propose to study a new bivariate version of the Conditional Tail Expectation (Definition 4.4.2 below).

Let us first introduce the following assumption:

A2: X and Y are positive random variables fulfilling $E(X^2) < \infty$ and $E(Y^2) < \infty$. The random vector (X,Y) has a λ-density function f such that $\int f^{1+r}\mathrm{d}\lambda < \infty$ for some $r > 0$.

Definition 4.4.2 *Consider a random vector (X,Y) satisfying Assumption* **A2**, *with associate distribution function $F \in \mathcal{F}$. For $\alpha \in (0,1)$, we define*

1. *the bivariate α-Conditional Tail Expectation*

$$\mathrm{CTE}_\alpha(X,Y) = \mathbb{E}[(X,Y)|(X,Y) \in L(\alpha)] = \begin{pmatrix} \mathbb{E}[\,X\,|\,(X,Y) \in L(\alpha)\,] \\ \mathbb{E}[\,Y\,|\,(X,Y) \in L(\alpha)\,] \end{pmatrix}.$$

2. *the estimated bivariate α-Conditional Tail Expectation*

$$\widehat{\mathrm{CTE}}_\alpha(X,Y) = \begin{pmatrix} \dfrac{\sum_{i=1}^n X_i 1_{\{(X_i,Y_i)\in L_n(\alpha)\}}}{\sum_{i=1}^n 1_{\{(X_i,Y_i)\in L_n(\alpha)\}}} \\ \dfrac{\sum_{i=1}^n Y_i 1_{\{(X_i,Y_i)\in L_n(\alpha)\}}}{\sum_{i=1}^n 1_{\{(X_i,Y_i)\in L_n(\alpha)\}}} \end{pmatrix}. \tag{4.6}$$

Note that this bivariate Conditional Tail Expectation is a natural extension of the univariate one. Moreover, if X and Y are identically distributed with a symmetric copula then $\mathbb{E}[\,X\,|\,(X,Y) \in L(\alpha)\,] = \mathbb{E}[\,Y\,|\,(X,Y) \in L(\alpha)\,]$.

Contrarily to the generalizations of the univariate CTE above, our $\mathrm{CTE}_\alpha(X,Y)$ does not use an aggregate variable (sum, min, max ...) in order to analyze the bivariate risk's issue. By contrast, using a geometric approach, $\mathrm{CTE}_\alpha(X,Y)$ rather deals with the simultaneous joint damages considering the bivariate dependence structure of data in a specific risk's area $(L(\alpha))$.

Let $\alpha \in (0,1)$ and $F \in \mathcal{F}$. We introduce truncated versions of the theoretical and estimated CTE_α:

$$\mathrm{CTE}_\alpha^T(X,Y) = \mathbb{E}[(X,Y)|(X,Y) \in L(\alpha)^T],$$

$$\widehat{\mathrm{CTE}}_\alpha^T(X,Y) = \begin{pmatrix} \dfrac{\sum_{i=1}^n X_i 1_{\{(X_i,Y_i)\in L_n(\alpha)^T\}}}{\sum_{i=1}^n 1_{\{(X_i,Y_i)\in L_n(\alpha)^T\}}} \\ \dfrac{\sum_{i=1}^n Y_i 1_{\{(X_i,Y_i)\in L_n(\alpha)^T\}}}{\sum_{i=1}^n 1_{\{(X_i,Y_i)\in L_n(\alpha)^T\}}} \end{pmatrix},$$

where $L(\alpha)^T$ and $L_n(\alpha)^T$ are the truncated versions of theoretical and estimated upper α-level set defined in Section 4.1.

Theorem 4.4.3 *Under Assumption* **A2**, *Assumptions of Theorem 4.3.1 and with the notation of Theorem 4.3.1, it holds that*

$$\beta_n \left| \mathrm{CTE}_\alpha^{T_n}(X,Y) - \widehat{\mathrm{CTE}}_\alpha^{T_n}(X,Y) \right| \underset{n\to\infty}{\overset{\mathbb{P}}{\to}} 0, \tag{4.7}$$

where $\beta_n = \min\{ p_n^{\frac{r}{2(1+r)}}, a_n \}$, *with* $r > 0$ *satisfying Assumption* **A2** *and* $a_n = o(\sqrt{n})$.

The convergence in (4.7) must be interpreted as a componentwise convergence. In the case of a bounded density function $f_{(X,Y)}$ we obtain $\beta_n = \min\{\sqrt{p_n}, a_n\}$.

Using Theorem 4.4.3 and Remark 10, we obtain the following result in the case of the bivariate empirical distribution function.

Corollary 4.4.4 *Let* F_n *the bivariate empirical distribution function. Under Assumptions Theorem 4.4.3 and with the notation of Theorem 4.4.3, it holds that*

$$\beta_n \left| \mathrm{CTE}_\alpha^{T_n}(X,Y) - \widehat{\mathrm{CTE}}_\alpha^{T_n}(X,Y) \right| \underset{n\to\infty}{\overset{\mathbb{P}}{\to}} 0,$$

where $\beta_n = o\left(n^{\frac{r}{6(1+r)}} / T_n^{\frac{2r}{3(1+r)}} \right)$, *with* $r > 0$ *satisfying Assumption* **A2**.

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

Remark 11 Starting from Corollary 4.4.4, it could be interesting to consider the convergence $\left|\,\mathrm{CTE}_\alpha(X,Y) - \widehat{\mathrm{CTE}}_\alpha^{T_n}(X,Y)\,\right|$. We remark that in this case the speed of convergence will also depend on the convergence rate to zero of $\left|\,\mathrm{CTE}_\alpha(X,Y) - \mathrm{CTE}_\alpha^{T_n}(X,Y)\,\right|$, then, in particular of $\mathbb{P}[(X,Y) \in L(\alpha) \setminus L(\alpha)^{T_n}]$ for $n \to \infty$. More precisely $\left|\,\mathrm{CTE}_\alpha(X,Y) - \mathrm{CTE}_\alpha^{T_n}(X,Y)\,\right|$ decays to zero at least with a convergence rate $(\mathbb{P}[X \geq T_n \text{ or } Y \geq T_n])^{-1}$. We now provide an illustration (in a simple case) on how to choose T_n in order to consider the convergence $\left|\,\mathrm{CTE}_\alpha(X,Y) - \widehat{\mathrm{CTE}}_\alpha^{T_n}(X,Y)\,\right|$. Let $F(x,y) = (1-e^{-x})(1-e^{-y})$, i.e. independent and exponentially distributed marginals with parameter 1, with bounded density function in $\mathbb{R}_+^2$. Let F_n the bivariate empirical distribution function. In this case all assumptions of Corollary 4.4.4 are satisfied. Then we obtain that $\left|\,\mathrm{CTE}_\alpha(X,Y) - \widehat{\mathrm{CTE}}_\alpha^{T_n}(X,Y)\,\right|$ decays to zero at least with a convergence rate $\beta_n = o(n^{\frac{1}{6}}/\ln(n)^{\frac{2}{3}})$, with a choice of sequence $T_n = \ln(n)$. 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 in Section 4.3 (see Section 4.5.2 for several numerical illustrations of this aspect).

4.5 Illustrations

4.5.1 Estimation of the level sets

In this section we confront our estimator of level sets of the distribution function with various simulated samples. The plug-in estimation of level sets is constructed using the empirical estimator F_n. We consider two distribution functions which satisfy assumptions of Theorem 4.3.1: independent copula with exponential marginals and Survival Clayton copula with Burr marginals.

Following Remark 10 we consider $T_n = \ln(n)$ and we obtain $p_n = o(n^{\frac{1}{3}}/\ln(n)^{\frac{4}{3}})$. We also take $T_n = n^{0.45}$. This second sequence provides an illustration for the case of T_n too large (see Tables 4.1(b)-4.2(b) and 4.4(b)-4.5(b)).

We take a random grid of 10000 points in $[0,T_n]^2$ and a distribution with independent and exponentially distributed marginals with parameter 1 and 2 respectively. We provide a Monte Carlo approximation for $\lambda(L(\alpha)^{T_n} \triangle L_n(\alpha)^{T_n})$ (averaged on 100 iterations), for different values of α and n. The results are gathered in Table 4.1.

In order to illustrate the influence of the choice of T_n, in Table 4.2 we show $p_n\,\lambda(L(\alpha)^{T_n} \triangle L_n(\alpha)^{T_n})$, with $p_n = o(n^{\frac{1}{3}}/\ln(n)^{\frac{4}{3}})$ in the case $T_n = \ln(n)$ (Table 4.2(a)) and $T_n = n^{0.45}$ (Table 4.2(b)). Results in Table 4.2(a) 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.
This illustrates not only the consistency property but also the speed of convergence of our plug-in estimator. As expected in Table 4.2(b) we see that taking T_n too large, we can not expect to get a good approximation of $\lambda(L(\alpha)^{T_n} \triangle L_n(\alpha)^{T_n})$.

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

α	n= 500	n= 1000	n= 2000
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

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

α	n= 500	n= 1000	n= 2000
0.10	0.331	0.326	0.223
0.24	0.519	0.391	0.249
0.38	0.591	0.469	0.396
0.52	1.057	0.906	0.881
0.66	1.222	0.989	0.904
0.80	1.541	1.367	1.334

Table 4.1: Distribution with independent and exponentially distributed marginals with parameter 1 and 2 respectively. Approximated $\lambda(L(\alpha)^{T_n} \triangle L_n(\alpha)^{T_n})$.

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

α	n= 500	n= 1000	n= 2000
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

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

α	n= 500	n= 1000	n= 2000
0.10	0.229	0.248	0.188
0.24	0.361	0.298	0.209
0.38	0.411	0.357	0.339
0.52	0.734	0.689	0.743
0.66	0.849	0.752	0.762
0.80	1.071	1.039	1.124

Table 4.2: Distribution with independent and exponentially distributed marginals with parameter 1 and 2 respectively. Approximated $p_n\,\lambda(L(\alpha)^{T_n} \triangle L_n(\alpha)^{T_n})$, with $p_n = o(n^{\frac{1}{3}}/\ln(n)^{\frac{4}{3}})$.

In Tables 4.3(a) and 4.3(b) we calculate the theoretical $\lambda(L(\alpha)^{T_n})$, for $T_n = \ln(n)$ and $T_n = n^{0.45}$ respectively, and for different values of n and α. These quantities provide an illustration of the truncation of level sets $L(\alpha)$, with respect to T_n and α.

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

α	n= 500	n= 1000	n= 2000
0.10	37.472	46.459	56.405
0.24	35.777	44.588	54.358
0.38	33.837	42.436	51.995
0.52	31.505	39.837	49.129
0.66	28.513	36.487	45.421
0.80	24.201	31.621	40.003

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

α	n= 500	n= 1000	n= 2000
0.10	142.205	248.516	433.872
0.24	139.039	244.378	428.449
0.38	135.328	239.495	422.021
0.52	130.771	233.464	414.045
0.66	124.779	225.488	403.451
0.80	115.851	213.505	387.436

Table 4.3: Theoretical $\lambda(L(\alpha)^{T_n})$, for different values of n and α. Distribution with independent and exponentially distributed marginals with parameter 1 and 2 respectively.

Table 4.4 provides a Monte Carlo approximation for $\lambda(L(\alpha)^{T_n} \triangle L_n(\alpha)^{T_n})$ (averaged on 100 iterations), in the case of a Survival Clayton copula with parameter 1 and Burr$(2,1)$

marginals. In Table 4.5 we show $p_n\,\lambda(L(\alpha)^{T_n} \triangle L_n(\alpha)^{T_n})$, with $p_n = o(n^{\frac{1}{3}}/\ln(n)^{\frac{4}{3}})$ in the case $T_n = \ln(n)$ (Table 4.5(a)) and $T_n = n^{0.45}$ (Table 4.5(b)). We obtain results that are analogous to those Table 4.2.

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

α	n= 500	n= 1000	n= 2000
0.10	0.208	0.171	0.126
0.24	0.252	0.241	0.149
0.38	0.339	0.288	0.132
0.52	0.349	0.335	0.245
0.66	0.499	0.364	0.353
0.80	0.771	0.644	0.605

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

α	n= 500	n= 1000	n= 2000
0.10	0.697	0.633	0.536
0.24	0.893	0.872	0.809
0.38	0.971	0.911	0.879
0.52	1.001	0.982	1.229
0.66	1.569	1.522	1.413
0.80	2.377	2.269	2.175

Table 4.4: Distribution with Survival Clayton copula with parameter 1 and Burr$(2,1)$ marginals. Approximated $\lambda(L(\alpha)^{T_n} \triangle L_n(\alpha)^{T_n})$.

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

α	n= 500	n= 1000	n= 2000
0.10	0.144	0.129	0.106
0.24	0.175	0.183	0.126
0.38	0.235	0.219	0.111
0.52	0.242	0.254	0.206
0.66	0.347	0.277	0.297
0.80	0.535	0.489	0.501

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

α	n= 500	n= 1000	n= 2000
0.10	0.484	0.481	0.452
0.24	0.620	0.663	0.682
0.38	0.674	0.692	0.741
0.52	0.695	0.746	1.036
0.66	1.089	1.157	1.191
0.80	1.651	1.725	1.834

Table 4.5: Distribution with Survival Clayton copula with parameter 1 and Burr$(2,1)$ marginals. Approximated $p_n\,\lambda(L(\alpha)^{T_n} \triangle L_n(\alpha)^{T_n})$, with $p_n = o(n^{\frac{1}{3}}/\ln(n)^{\frac{4}{3}})$.

As expected, the greater n is, the better the estimations are (see Tables 4.1(a) and 4.4(a)). Moreover we note that for big values of α we need more data to get a good estimation of the level sets. This may come from the fact that when α grows the gradient of the distribution function decreases to zero and the constant A grows significantly (see proof of Theorem 4.3.1 in Section 4.7).

Remark 12 As remarked in Section 4.3 the choice of the sequence T_n is essentially a compromise: if T_n is large we have a quite slow convergence rate p_n for $\lambda(L(\alpha)^{T_n} \triangle L_n(\alpha)^{T_n})$ and if T_n is quite small we obtain a faster convergence rate. However choosing T_n too small we substantially truncate our level sets. This aspect can be problematic in the case c is close to one. Finally the theoretical optimal criterion for the choice of sequence T_n in Theorem 4.3.1 is an open interesting problem that we leave for a future work.

4.5.2 Estimation of bivariate CTE

Simulated data

In order to evaluate the performance of our estimator we present here some simulated cases of estimation of $\mathrm{CTE}_\alpha(X,Y)$, for different values of level α. To compare the estimated results with the theoretical ones we consider cases for which we can calculate (with Maple) the explicit value of the theoretical $\mathrm{CTE}_\alpha(X,Y)$. However our estimator can be applied to much more general cases. By being able to get a good estimator for the level sets, we can expect a good estimator for $\mathrm{CTE}_\alpha(X,Y)$.

In the following we compare $\widehat{\mathrm{CTE}}_\alpha^{T_n}(X,Y)$ with the theoretical $\mathrm{CTE}_\alpha(X,Y)$, for different distributions: Independent copula with exponentially distributed marginals; Clayton copula with parameter 1, with exponential and Burr$(4,1)$ univariate marginals. The plug-in estimation of level sets is constructed using the empirical estimator F_n of the bivariate distribution function, with $n = 1000$.

In the following we denote $\overline{\widehat{\mathrm{CTE}}_\alpha^{T_n}}(X,Y) = \left(\overline{\widehat{\mathrm{CTE}}_\alpha^{T_n,1}}(X,Y), \overline{\widehat{\mathrm{CTE}}_\alpha^{T_n,2}}(X,Y)\right)$ the mean (coordinate by coordinate) of $\widehat{\mathrm{CTE}}_\alpha^{T_n}(X,Y)$ on 100 simulations. We denote $\widehat{\sigma} = (\widehat{\sigma}_1, \widehat{\sigma}_2)$ the empirical standard deviation (coordinate by coordinate) with

$$\widehat{\sigma}_1 = \sqrt{\frac{1}{99}\sum_{j=1}^{100}\left(\widehat{\mathrm{CTE}}_\alpha^{T_n,1}(X,Y)_j - \overline{\widehat{\mathrm{CTE}}_\alpha^{T_n,1}}(X,Y)\right)^2}$$

relatives to the first coordinate (resp. $\widehat{\sigma}_2$ relatives to the second one).

We denote $\mathrm{RMSE} = (\mathrm{RMSE}_1, \mathrm{RMSE}_2)$ the relative mean square error (coordinate by coordinate) with

$$\mathrm{RMSE}_1 = \sqrt{\frac{1}{100}\sum_{j=1}^{100}\left(\frac{\widehat{\mathrm{CTE}}_\alpha^{T_n,1}(X,Y)_j - \mathrm{CTE}_\alpha^{T_n,1}(X,Y)}{\mathrm{CTE}_\alpha^{T_n,1}(X,Y)}\right)^2}$$

relatives to the first coordinate of $\mathrm{CTE}_\alpha^{T_n}(X,Y)$ (resp. RMSE_2 relatives to the second one).

For the Independent copula with exponentially distributed marginals, from Corollary 4.4.4 and Remark 11, we take $T_n = \ln(n)$ (see Tables 4.6 and 4.8). This choice provides a convergence rate $\beta_n = o(n^{\frac{1}{6}}/\ln(n)^{\frac{2}{3}})$. Furthermore, as in Section 4.5.1, we consider the case with $T_n = n^{0.45}$. This second sequence provides an illustration for the case of T_n too large (see Tables 4.7 and 4.9). Following the same approach of Remark 11, in the case of Clayton copula with parameter 1, with exponential and Burr$(4,1)$ univariate marginals we have $\beta_n = o(n^{\frac{1}{6}}/(n^{0.2})^{\frac{2}{3}})$, with a sequence $T_n = n^{0.2}$ (see Table 4.10). Also in this case, we show the behavior of our estimator for $T_n = n^{0.45}$ (see Table 4.11).

α	$\mathrm{CTE}_\alpha(X,Y)$	$\overline{\widehat{\mathrm{CTE}_\alpha}^{T_n}}(X,Y)$	$\widehat{\sigma}$	RMSE
0.10	(0.627, 0.627)	(0.585, 0.637)	(0.022, 0.021)	(0.035, 0.039)
0.24	(0.761, 0.761)	(0.766, 0.759)	(0.031, 0.031)	(0.041, 0.046)
0.38	(0.896, 0.896)	(0.864, 0.917)	(0.036, 0.037)	(0.053, 0.049)
0.52	(1.051, 1.051)	(1.027, 1.071)	(0.051, 0.053)	(0.055, 0.052)
0.66	(1.246, 1.246)	(1.252, 1.238)	(0.075, 0.067)	(0.063, 0.065)
0.80	(1.531, 1.531)	(1.497, 1.549)	(0.123, 0.117)	(0.083, 0.077)

Table 4.6: (X,Y) with independent and exponentially distributed components with parameter 2; $T_n = \ln(n)$.

α	$\mathrm{CTE}_\alpha(X,Y)$	$\overline{\widehat{\mathrm{CTE}_\alpha}^{T_n}}(X,Y)$	$\widehat{\sigma}$	RMSE
0.10	(0.627, 0.627)	(0.603, 0.656)	(0.031, 0.031)	(0.062, 0.068)
0.24	(0.761, 0.761)	(0.774, 0.731)	(0.061, 0.071)	(0.082, 0.130)
0.38	(0.896, 0.896)	(0.927, 0.871)	(0.072, 0.076)	(0.087, 0.119)
0.52	(1.051, 1.051)	(1.086, 1.128)	(0.094, 0.082)	(0.095, 0.107)
0.66	(1.246, 1.246)	(1.281, 1.322)	(0.127, 0.101)	(0.102, 0.101)
0.80	(1.531, 1.531)	(1.545, 1.611)	(0.157, 0.161)	(0.105, 0.117)

Table 4.7: (X,Y) with independent and exponentially distributed components with parameter 2; $T_n = n^{0.45}$.

α	$\mathrm{CTE}_\alpha(X,Y)$	$\overline{\widehat{\mathrm{CTE}_\alpha}^{T_n}}(X,Y)$	$\widehat{\sigma}$	RMSE
0.10	(1.255, 0.627)	(1.222, 0.638)	(0.044, 0.022)	(0.043, 0.039)
0.24	(1.521, 0.761)	(1.488, 0.811)	(0.069, 0.023)	(0.051, 0.042)
0.38	(1.792, 0.896)	(1.797, 0.911)	(0.075, 0.038)	(0.044, 0.046)
0.52	(2.102, 1.051)	(2.082, 1.047)	(0.104, 0.052)	(0.052, 0.052)
0.66	(2.492, 1.246)	(2.461, 1.255)	(0.139, 0.071)	(0.057, 0.056)
0.80	(3.061, 1.531)	(3.011, 1.544)	(0.251, 0.125)	(0.084, 0.082)

Table 4.8: (X,Y) with independent and exponentially distributed components with parameter 1 and 2 respectively; $T_n = \ln(n)$.

α	$\mathrm{CTE}_\alpha(X,Y)$	$\overline{\widehat{\mathrm{CTE}_\alpha}^{T_n}}(X,Y)$	$\widehat{\sigma}$	RMSE
0.10	(1.255, 0.627)	(1.253, 0.624)	(0.061, 0.043)	(0.051, 0.054)
0.24	(1.521, 0.761)	(1.514, 0.803)	(0.074, 0.049)	(0.048, 0.075)
0.38	(1.792, 0.896)	(1.793, 0.948)	(0.096, 0.055)	(0.053, 0.084)
0.52	(2.102, 1.051)	(2.087, 1.111)	(0.118, 0.076)	(0.056, 0.092)
0.66	(2.492, 1.246)	(2.497, 1.311)	(0.169, 0.108)	(0.068, 0.101)
0.80	(3.061, 1.531)	(3.056, 1.602)	(0.313, 0.153)	(0.102, 0.111)

Table 4.9: (X,Y) with independent and exponentially distributed components with parameter 1 and 2 respectively; $T_n = n^{0.45}$.

α	$\mathrm{CTE}_\alpha(X,Y)$	$\overline{\widehat{\mathrm{CTE}}_\alpha}^{T_n}(X,Y)$	$\widehat{\sigma}$	RMSE
0.10	(1.188, 1.229)	(1.049, 1.192)	(0.032, 0.021)	(0.019, 0.033)
0.24	(1.448, 1.366)	(1.283, 1.379)	(0.053, 0.224)	(0.019, 0.063)
0.38	(1.727, 1.505)	(1.525, 1.471)	(0.046, 0.031)	(0.019, 0.031)
0.52	(2.049, 1.666)	(1.803, 1.625)	(0.058, 0.041)	(0.023, 0.034)
0.66	(2.454, 1.875)	(2.129, 1.823)	(0.071, 0.054)	(0.035, 0.039)
0.80	(3.039, 2.202)	(2.591, 2.144)	(0.111, 0.105)	(0.055, 0.054)

Table 4.10: (X,Y) with Clayton copula with parameter 1, F_X exponential distribution with parameter 1, F_Y Burr$(4,1)$ distribution; $T_n = n^{0.2}$.

α	$\mathrm{CTE}_\alpha(X,Y)$	$\overline{\widehat{\mathrm{CTE}}_\alpha}^{T_n}(X,Y)$	$\widehat{\sigma}$	RMSE
0.10	(1.188, 1.229)	(1.189, 1.238)	(0.061, 0.035)	(0.039, 0.029)
0.24	(1.448, 1.366)	(1.462, 1.365)	(0.065, 0.037)	(0.046, 0.031)
0.38	(1.727, 1.505)	(1.751, 1.536)	(0.082, 0.046)	(0.049, 0.037)
0.52	(2.049, 1.666)	(2.063, 1.713)	(0.091, 0.061)	(0.051, 0.045)
0.66	(2.454, 1.875)	(2.457, 1.951)	(0.117, 0.104)	(0.057, 0.068)
0.80	(3.039, 2.202)	(3.037, 2.322)	(0.192, 0.165)	(0.063, 0.108)

Table 4.11: (X,Y) with Clayton copula with parameter 1, F_X exponential distribution with parameter 1, F_Y Burr$(4,1)$ distribution; $T_n = n^{0.45}$.

Following we show that for high levels (here $\alpha = 0.9$), one needs to use large samples to get reasonable estimates of CTE_α. We consider for the purpose $(X,\ Y)$ independent and exponentially distributed with respective parameters 1 and 2. The theoretical value is $\mathrm{CTE}_{0.9}(X,Y) = (3.78, 1.89)$. We take $T_n = \ln(n)$ (see Table 4.12). In this case we need between 2000 and 2500 data to get the same performances as for lower level (see Table 4.8 above). Furthermore we consider the case $T_n = n^{0.45}$ (see Table 4.13). In this case we need between 2500 and 5000 data to get the same performances as for lower level (see Table 4.9 above).

n	500	1000	1500	2000	2500	5000
$\widehat{\sigma}$	(0.614, 0.359)	(0.444, 0.308)	(0.431, 0.295)	(0.377, 0.168)	(0.241, 0.123)	(0.216, 0.121)
RMSE	(0.168, 0.189)	(0.123, 0.163)	(0.115, 0.161)	(0.099, 0.089)	(0.077, 0.079)	(0.063, 0.057)

Table 4.12: Evolution of $\widehat{\sigma}$ 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. $T_n = \ln(n)$.

n	500	1000	1500	2000	2500	5000
$\widehat{\sigma}$	(0.919, 0.419)	(0.568, 0.319)	(0.511, 0.294)	(0.382, 0.239)	(0.348, 0.223)	(0.307, 0.151)
RMSE	(0.242, 0.221)	(0.151, 0.172)	(0.133, 0.165)	(0.101, 0.144)	(0.093, 0.129)	(0.092, 0.108)

Table 4.13: Evolution of $\widehat{\sigma}$ 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. $T_n = n^{0.45}$.

The bad properties of the estimate for sample sizes less than 2000 (in the case of $T_n = \ln(n)$) can be explained by the fact that for high levels, the constant A is large (see proof of Theorem 4.3.1 in Section 4.7), but also by the fact that for large values of α the empirical estimate F_n of F may not be the most appropriate.

Real data

We consider here the estimation of CTE_α in a real case: **Loss-ALAE data** in the log scale (for details see Frees and Valdez, 1998). Each claim consists of an indemnity payment (the loss, X) and an allocated loss adjustment expense (ALAE, Y). Examples of ALAE are the fees paid to outside attorneys, experts, and investigators used to defend claims.

The data size is $n = 1500$. Let $T_n = n^{0.4}$. Again our estimator is constructed using the empirical estimator F_n. Results about the estimation of $\mathrm{CTE}_\alpha^{T_n}$ are gathered in Table 4.14. In Figure 4.2 we represent data, estimated level sets and estimated bivariate Conditional Tail Expectation for several values of α.

α	0.10	0.24	0.38
$\widehat{\mathrm{CTE}}_\alpha^{T_n}$	(9.937, 9.252)	(10.361, 9.566)	(10.731, 9.728)
α	0.52	0.66	0.80
$\widehat{\mathrm{CTE}}_\alpha^{T_n}$	(11.096, 10.011)	(11.518, 10.315)	(12.057, 10.758)

Table 4.14: $\widehat{\mathrm{CTE}}_\alpha^{T_n}$ for Loss-ALAE data in log scale, with different values of level α.

In this real setting the estimation of CTE_α can be used in order to quantify the mean of the Loss (resp. ALAE) in the log scale conditionally to the fact that Loss and ALAE data belong jointly to the specific risk's area $L(\alpha)$.

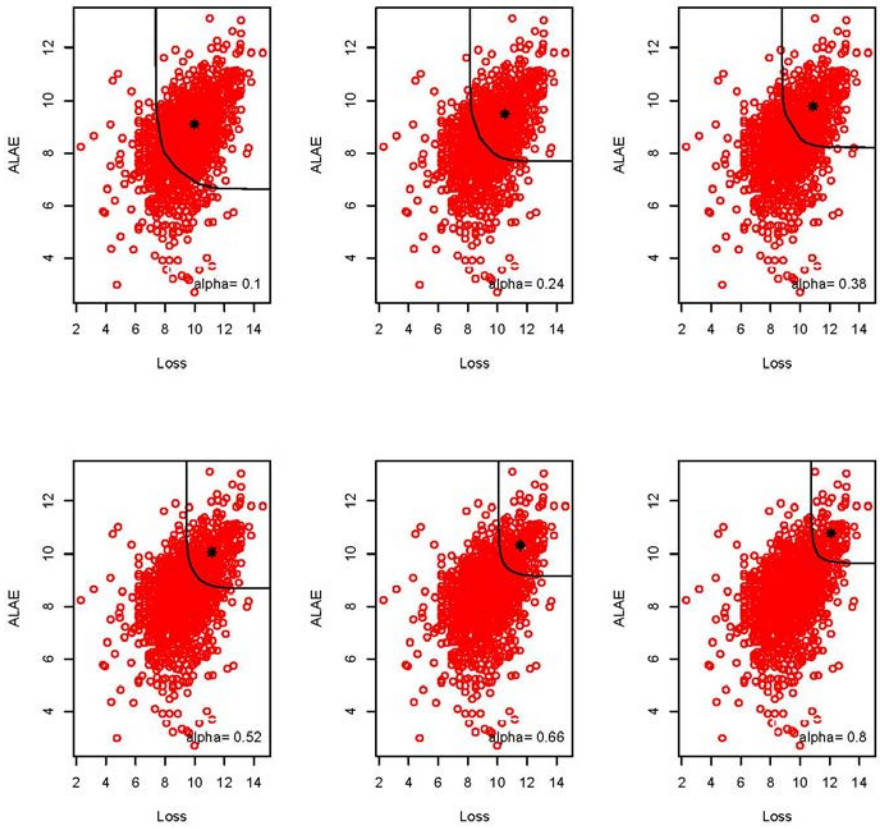

Figure 4.2: Loss-ALAE data in log scale: boundary of estimated level sets (line) and $\widehat{\mathrm{CTE}}_{\alpha}^{T_n}$ (star) for different values of α.

4.6 Conclusion

In this paper we have provided convergence results for the plug-in estimator of the level sets of an unknown distribution function on $\mathbb{R}_+^2$ in terms of Hausdorff distance and volume of the symmetric difference. In this setting we have proposed and estimated a new bivariate risk measure: $\mathrm{CTE}_\alpha(X,Y)$. A future work comparing our bivariate Conditional Tail Expectation with existing bivariate generalizations of the CTE in terms of classical properties (monotonicity, translation invariance, homogeneity, ...), dependence structure, behavior with respect to different risk scenarios, is in preparation.

4.7 Proofs

Proof of Proposition 4.1.1

Take $T > 0$ such that for all $t : \mid t - c \mid \leq r$, $\{F = t\}^T \neq \emptyset$, (from assumptions of Proposition 4.1.1 we know that such an T there exists).

Let $x \in \{z \in [0,T]^2 : \mid F(z) - c \mid \leq r\}$. Define for $\overline{\lambda} \in \mathbb{R}$

$$y_{\overline{\lambda}} \equiv y_{\overline{\lambda},x} = x + \overline{\lambda}\frac{(\nabla F)_x}{\|(\nabla F)_x\|},$$

so $\|y_{\overline{\lambda}} - x\| = \mid \overline{\lambda} \mid$. From the differentiability properties of F and using Taylor's formula we have, for $\mid \overline{\lambda} \mid < \zeta$

$$F(y_{\overline{\lambda}}) = F(x) + (\nabla F)_x^T(y_{\overline{\lambda}} - x) + \frac{1}{2}(y_{\overline{\lambda}} - x)^T(HF)_{\overline{x}}(y_{\overline{\lambda}} - x),$$

with $\overline{x}$ a point on the segment between x and $y_{\overline{\lambda}}$. So

$$F(y_{\overline{\lambda}}) = F(x) + \overline{\lambda}\|(\nabla F)_x\| + \frac{\overline{\lambda}^2}{2\|(\nabla F)_x\|^2}(\nabla F)_x^T(HF)_{\overline{x}}(\nabla F)_x.$$

By the Cauchy Schwarz Inequality we obtain

$$F(y_{\overline{\lambda}}) \geq F(x) + \overline{\lambda}\|(\nabla F)_x\| - \frac{\overline{\lambda}^2}{2\|(\nabla F)_x\|}\|(HF)_{\overline{x}}(\nabla F)_x\|$$

and

$$F(y_{\overline{\lambda}}) \leq F(x) + \overline{\lambda}\|(\nabla F)_x\| + \frac{\overline{\lambda}^2}{2\|(\nabla F)_x\|}\|(HF)_{\overline{x}}(\nabla F)_x\|.$$

Since $\|(HF)_{\overline{x}}(\nabla F)_x\| \leq \|(HF)_{\overline{x}}\|\|(\nabla F)_x\|$, we have

$$F(x) + \overline{\lambda}\|(\nabla F)_x\| - \frac{\overline{\lambda}^2}{2}\|(HF)_{\overline{x}}\| \leq F(y_{\overline{\lambda}}) \leq F(x) + \overline{\lambda}\|(\nabla F)_x\| + \frac{\overline{\lambda}^2}{2}\|(HF)_{\overline{x}}\|.$$

Since $\overline{x} \in E$ and $M_H < \infty$ we obtain

$$F(x) + \overline{\lambda}\|(\nabla F)_x\| - \frac{\overline{\lambda}^2}{2}M_H \leq F(y_{\overline{\lambda}}) \leq F(x) + \overline{\lambda}\|(\nabla F)_x\| + \frac{\overline{\lambda}^2}{2}M_H. \tag{4.8}$$

For $0 < \overline{\lambda} < \zeta$, we have from the left side of (4.8)

$$F(y_{\overline{\lambda}}) \geq F(x) + \overline{\lambda}\|(\nabla F)_x\| - \frac{\overline{\lambda}^2}{2} M_H \geq F(x) + \overline{\lambda} m^{\triangledown} - \frac{\overline{\lambda}^2}{2} M_H.$$

We assume now that $M_H > 0$ (the case $M_H = 0$ is trivial).

For $x \in \{z \in [0,T]^2 : \mid F(z) - c \mid \leq r\}$ and $0 < \overline{\lambda} < \zeta \wedge \frac{m^{\triangledown}}{M_H}$ we have then

$$F(y_{\overline{\lambda}}) \geq F(x) + \frac{\overline{\lambda}}{2} m^{\triangledown}, \tag{4.9}$$

Similarly using the right side of (4.8) we obtain for $0 < \overline{\lambda} < \zeta \wedge \frac{m^{\triangledown}}{M_H}$ and $x \in \{z \in [0,T]^2 : \mid F(z) - c \mid \leq r\}$,

$$F(y_{-\overline{\lambda}}) \leq F(x) - \frac{\overline{\lambda}}{2} m^{\triangledown}. \tag{4.10}$$

Define

$$\gamma = \left(\frac{m^{\triangledown}}{4} \left(\zeta \wedge \frac{m^{\triangledown}}{M_H} \right) \right) \wedge r > 0.$$

Suppose that $t = c + \varepsilon$, $0 < \varepsilon \leq \gamma$. Let $x \in [0,T]^2$ such that $F(x) = t = c + \varepsilon$ then $x \in \{z \in [0,T]^2 : \mid F(z) - c \mid \leq r\}$. Take now

$$0 < \overline{\lambda} = \frac{2\varepsilon}{m^{\triangledown}} < \zeta \wedge \frac{m^{\triangledown}}{M_H}.$$

We obtain from (4.10)

$$F(y_{-\overline{\lambda}}) \leq F(x) - \frac{\overline{\lambda}}{2} m^{\triangledown} = c + \varepsilon - \varepsilon = c.$$

From the continuity property of F, we deduce that there exists y between x and $y_{-\overline{\lambda}}$ such that $F(y) = c$ and we have

$$\|x - y\| \leq \|x - y_{-\overline{\lambda}}\| = \mid \overline{\lambda} \mid = \frac{2\varepsilon}{m^{\triangledown}} = \frac{2}{m^{\triangledown}} \mid t - c \mid .$$

So we have proved that

$$\sup_{x \in \{F = t\}^T} d(x, \{F = c\}^T) \leq \frac{2}{m^{\triangledown}} \mid t - c \mid .$$

Similarly, take $x \in [0,T]^2$ such that $F(x) = c$ and use (4.9) to obtain

$$\sup_{x \in \{F = c\}^T} d(x, \{F = t\}^T) \leq \frac{2}{m^{\triangledown}} \mid t - c \mid .$$

The proof in case $t < c$ is completely analogous. So F satisfies Assumption **H** (see Section 4.1) with $A = \frac{2}{m^{\triangledown}}$. $\square$

Proof of Theorem 4.2.1

Under assumptions of Theorem 4.2.1, we can always take $T_1 > 0$ such that for all $t : |t - c| \leq r,\ \partial L(t)^{T_1} \neq \emptyset$.
Then for each n, for all $t : |t - c| \leq r,\ \partial L(t)^{T_n}$ is a non-empty (and compact) set on $\mathbb{R}_+^2$.
In each $[0, T_n]^2$, from Proposition 4.1.1, Assumption **H** (Section 4.1) is satisfied with

$$\gamma = \left(\frac{m^\triangledown}{4}\left(\zeta \wedge \frac{m^\triangledown}{M_H}\right)\right) \wedge r > 0 \quad \text{and} \quad A = \frac{2}{m^\triangledown}.$$

First we have to find a bound for $\sup_{x \in \partial L(c)^{T_n}} d(x, \partial L_n(c)^{T_n})$.

Take $x \in \partial L(c)^{T_n}$ and define $\varepsilon_n = 2\,\|F - F_n\|_\infty^{T_n}$. Using $\|F - F_n\|_\infty \to 0$, *a.s.*, for $n \to \infty$, then $\varepsilon_n \to 0$, *a.s.*, for $n \to \infty$. So with probability one there exists n_0 such that $\forall\, n \geq n_0$, $\varepsilon_n \leq \gamma$.

Since for all $t : |t - c| \leq r\ \ \partial L(t)^{T_n} \neq \emptyset$, from Assumption **H**, there exist $u_n \equiv u_x^{\varepsilon_n}$ and $l_n \equiv l_x^{\varepsilon_n}$ in $[0, T_n]^2$ such that

$$F(u_n) = c + \varepsilon_n; \quad d(x, u_n) \leq A\,\varepsilon_n,$$

$$F(l_n) = c - \varepsilon_n; \quad d(x, l_n) \leq A\,\varepsilon_n.$$

Suppose now $\|F - F_n\|_\infty^{T_n} > 0$ (the other case is a trivial one). In this case

$$F_n(u_n) = c + \varepsilon_n + F_n(u_n) - F(u_n) \geq c + \varepsilon_n - \|F - F_n\|_\infty^{T_n} = c + 2\|F - F_n\|_\infty^{T_n} - \|F - F_n\|_\infty^{T_n} > c,$$

and in a similar way we can prove that $F_n(l_n) < c$.

As $F_n(l_n) < c$ and $F_n(u_n) > c$, with u_n and l_n in $[0, T_n]^2$, there exists $z_n \in \partial L_n(c)^{T_n} \cap B(u_n, d(u_n, l_n))$ with

$$\begin{aligned} d(z_n, x) \leq d(z_n, u_n) + d(u_n, x) \leq d(u_n, l_n) + d(u_n, x) &\leq d(u_n, x) + d(x, l_n) + d(u_n, x) \\ &\leq 3\,A\,\varepsilon_n = 6\,A\|F - F_n\|_\infty^{T_n}. \end{aligned}$$

Hence, for $n \geq n_0$

$$\sup_{x \in \partial L(c)^{T_n}} d(x, \partial L_n(c)^{T_n}) \leq\ 6\,A\,\|F - F_n\|_\infty^{T_n}.$$

Let us now bound $\sup_{x \in \partial L_n(c)^{T_n}} d(x, \partial L(c)^{T_n})$.

Take $x \in \partial L_n(c)^{T_n}$. From the *a.s.* continuity of F_n we obtain $F_n(x) = c$, *a.s.*, so

$$|\ F(x) - c\ | \leq |\ F(x) - F_n(x)\ | \leq\ \|F - F_n\|_\infty^{T_n}\ \leq\ \varepsilon_n, \quad a.s.$$

Remember that $\forall\ n \geq n_0$, $\varepsilon_n \leq \gamma$, *a.s.* Then from Assumption **H** $d(x, \partial L(c)^{T_n}) \leq\ A\ |\ F(x) - c\ | \leq\ A\,\|F - F_n\|_\infty^{T_n}$. We can conclude that with probability one, for $n \geq n_0$

$$\sup_{x \in \partial L_n(c)^{T_n}} d(x, \partial L(c)^{T_n}) \leq\ A\,\|F - F_n\|_\infty^{T_n}.$$

We obtain for $n \geq n_0$, $d_H(\partial L(c)^{T_n}, \partial L_n(c)^{T_n}) \leq 6\,A\,\|F - F_n\|_\infty^{T_n}$, then

$$d_H(\partial L(c)^{T_n}, \partial L_n(c)^{T_n}) = O(\|F - F_n\|_\infty^{T_n}), \quad a.s.$$

Hence the result. $\square$

Proof of Theorem 4.3.1

Under assumptions of Theorem 4.3.1, we can always take $T_1 > 0$ such that for all $t : |t - c| \leq r$, $\partial L(t)^{T_1} \neq \emptyset$. Then for each n, for all $t : |t - c| \leq r$, $\partial L(t)^{T_n}$ is a non-empty (and compact) set on $\mathbb{R}_+^2$.
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 $L(c)^{T_n} \triangle L_n(c)^{T_n}$, $Q_{\varepsilon_n} = \{x \in [0, T_n]^2 : \mid F - F_n \mid \leq \varepsilon_n\}$ and $\widetilde{Q}_{\varepsilon_n} = \{x \in [0, T_n]^2 : \mid F - F_n \mid > \varepsilon_n\}$ are measurable and

$$\lambda(L(c)^{T_n} \triangle L_n(c)^{T_n}) = \lambda(L(c)^{T_n} \triangle L_n(c)^{T_n} \cap Q_{\varepsilon_n}) + \lambda(L(c)^{T_n} \triangle L_n(c)^{T_n} \cap \widetilde{Q}_{\varepsilon_n}). \quad (4.11)$$

Since $L(c)^{T_n} \triangle L_n(c)^{T_n} \cap Q_{\varepsilon_n} \subset \{x \in [0, T_n]^2 : c - \varepsilon_n \leq F < c + \varepsilon_n\}$ we obtain

$$\lambda(L(c)^{T_n} \triangle L_n(c)^{T_n}) \leq \lambda(\{x \in [0, T_n]^2 : c - \varepsilon_n \leq F < c + \varepsilon_n\}) + \lambda(\widetilde{Q}_{\varepsilon_n}).$$

From Assumption **H** (Section 4.1) and Proposition 4.1.1, if $2\,\varepsilon_n \leq \gamma$ then

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

So we can write

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

If we now choose

$$\varepsilon_n = o\left(\frac{1}{p_n\,T_n}\right) \quad (4.12)$$

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

$$p_n\,\lambda(\{x \in [0, T_n]^2 : c - \varepsilon_n \leq F < c + \varepsilon_n\}) \underset{n\to\infty}{\to} 0.$$

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

$$p_n\,\lambda(\widetilde{Q}_{\varepsilon_n}) = p_n \int 1_{\{x \in [0,T_n]^2 : |F - F_n| > \varepsilon_n\}}\,\lambda(\mathrm{d}x) \leq \frac{p_n}{\varepsilon_n^p} \int_{[0,T_n]^2} \mid F - F_n \mid^p \lambda(\mathrm{d}x). \quad (4.13)$$

Take ε_n such that

$$\varepsilon_n = \left(\frac{p_n}{v_n}\right)^{\frac{1}{p}}. \quad (4.14)$$

Then, from Assumption **A1**, in Section 4.3, we obtain $p_n\,\lambda(\widetilde{Q}_{\varepsilon_n}) \underset{n\to\infty}{\overset{\mathbb{P}}{\to}} 0$. As $p_n = o\left(v_n^{\frac{1}{p+1}} / T_n^{\frac{p}{p+1}}\right)$ we can choose ε_n that satisfies (4.12) and (4.14). Hence the result. $\square$

Proof of Theorem 4.4.3

We only prove the result for the first coordinate of $\mathrm{CTE}_\alpha(X,Y)$ (the proof is similar for the second one). First we remark that, in the following, probabilities involving events which depend on $L_n(\alpha)^{T_n}$, are conditional probabilities to the sample (X_i, Y_i), for $i = 1, \dots n$. We will denote $\mathbb{P}[(X,Y) \in L_n(\alpha)^{T_n}]$ purely for reasons of notational convenience.

We introduce these two preliminary results (Lemma 4.7.1 and 4.7.2):

Lemma 4.7.1 *Under Assumption* **A2**, *Assumptions of Theorem 4.3.1 and with the notation of Theorem 4.3.1, it holds that*

$$p_n^{\frac{r}{2(1+r)}} \left| \mathbb{E}[\, X \mid (X,Y) \in L(\alpha)^{T_n} \,] - \mathbb{E}[\, X \mid (X,Y) \in L_n(\alpha)^{T_n} \,] \right| \underset{n\to\infty}{\overset{\mathbb{P}}{\to}} 0,$$

with $r > 0$ satisfying Assumption **A2**.

Proof of Lemma 4.7.1

From Assumption **A2** and Theorem 4.3.1 we obtain

$$p_n^{\frac{r}{2(1+r)}} \left|\mathbb{P}[(X,Y) \in L(\alpha)^{T_n} \triangle L_n(\alpha)^{T_n}]\right| \leq p_n^{\frac{r}{2(1+r)}} \, d_\lambda\big(L(\alpha)^{T_n}, L_n(\alpha)^{T_n}\big)^{\frac{r}{1+r}} \parallel f \parallel_{1+r} \underset{n\to\infty}{\overset{\mathbb{P}}{\to}} 0. \tag{4.15}$$

As a straightforward consequence we find

$$p_n^{\frac{r}{2(1+r)}} \left|\mathbb{P}[(X,Y) \in L(\alpha)^{T_n}] - \mathbb{P}[(X,Y) \in L_n(\alpha)^{T_n}]\right| \underset{n\to\infty}{\overset{\mathbb{P}}{\to}} 0.$$

Using Assumption **A2** we also obtain

$$\begin{aligned} p_n^{\frac{r}{2(1+r)}} &\left| \int_{L(\alpha)^{T_n}} x\, f_{X,Y}(x,y)\lambda(\mathrm{d}x\,\mathrm{d}y) - \int_{L_n(\alpha)^{T_n}} x\, f_{X,Y}(x,y)\lambda(\mathrm{d}x\,\mathrm{d}y) \right| \\ &\leq p_n^{\frac{r}{2(1+r)}} \, \mathbb{E}[X^2]^{\frac{1}{2}} \, d_\lambda\big(L(\alpha)^{T_n}, L_n(\alpha)^{T_n}\big)^{\frac{r}{2(1+r)}} \parallel f \parallel_{1+r}^{\frac{1}{2}} \underset{n\to\infty}{\overset{\mathbb{P}}{\to}} 0. \end{aligned} \tag{4.16}$$

Then

$$\begin{aligned} p_n^{\frac{r}{2(1+r)}} &\left|\mathbb{E}[X|(X,Y) \in L(\alpha)^{T_n}] - \mathbb{E}[X|(X,Y) \in L_n(\alpha)^{T_n}]\right| \\ &= p_n^{\frac{r}{2(1+r)}} \left| \int_{L(\alpha)^{T_n}} x\, f_{X,Y}(x,y)\lambda(\mathrm{d}x\,\mathrm{d}y)\ \mathbb{P}[(X,Y) \in L(\alpha)^{T_n}]^{-1} \right. \\ &\qquad \left. - \int_{L_n(\alpha)^{T_n}} x\, f_{X,Y}(x,y)\lambda(\mathrm{d}x\,\mathrm{d}y)\ \mathbb{P}[(X,Y) \in L_n(\alpha)^{T_n}]^{-1} \right| \\ &\leq \frac{p_n^{\frac{r}{2(1+r)}}}{\mathbb{P}[(X,Y) \in L(\alpha)^{T_n}]\,\mathbb{P}[(X,Y) \in L_n(\alpha)^{T_n}]} \Bigg(\mathbb{P}[(X,Y) \in L(\alpha)^{T_n}] \\ &\qquad \left| \int_{L(\alpha)^{T_n}} x\, f_{X,Y}(x,y)\lambda(\mathrm{d}x\,\mathrm{d}y) - \int_{L_n(\alpha)^{T_n}} x\, f_{X,Y}(x,y)\lambda(\mathrm{d}x\,\mathrm{d}y) \right| \\ &+ \int_{L(\alpha)^{T_n}} x\, f_{X,Y}(x,y)\lambda(\mathrm{d}x\,\mathrm{d}y) \left| \mathbb{P}[(X,Y) \in L(\alpha)^{T_n}] - \mathbb{P}[(X,Y) \in L_n(\alpha)^{T_n}] \right| \Bigg). \end{aligned} \tag{4.17}$$

Using (4.15)-(4.16) we obtain that (4.17) converges to zero in probability, for $n \to \infty$. Hence the result. $\square$

Lemma 4.7.2 *Under Assumption* **A2**, *Assumptions of Theorem 4.3.1 and with the notation of Theorem 4.3.1, it holds that*

$$a_n \left| \mathbb{E}[\, X \mid (X,Y) \in L_n(\alpha)^{T_n} \,] - \frac{\sum_{i=1}^n X_i 1_{\{(X_i,Y_i)\in L_n(\alpha)^{T_n}\}}}{\sum_{i=1}^n 1_{\{(X_i,Y_i)\in L_n(\alpha)^{T_n}\}}} \right| \underset{n\to\infty}{\overset{\mathbb{P}}{\to}} 0,$$

for $a_n = o\,(\sqrt{n})$.

Proof of Lemma 4.7.2
We can write

$$\begin{aligned} a_n &\left| \mathbb{E}[X|(X,Y) \in L_n(\alpha)^{T_n}] - \frac{\displaystyle\sum_{i=1}^n X_i 1_{\{(X_i,Y_i)\in L_n(\alpha)^{T_n}\}}}{\displaystyle\sum_{i=1}^n 1_{\{(X_i,Y_i)\in L_n(\alpha)^{T_n}\}}} \right| \\ &= a_n \left| \frac{\displaystyle\int_{L_n(\alpha)^{T_n}} x\, f_{X,Y}(x,y)\lambda(\mathrm{d}x\,\mathrm{d}y)}{\mathbb{P}[(X,Y) \in L_n(\alpha)^{T_n}]} - \frac{\displaystyle\sum_{i=1}^n X_i 1_{\{(X_i,Y_i)\in L_n(\alpha)^{T_n}\}}}{\displaystyle\sum_{i=1}^n 1_{\{(X_i,Y_i)\in L_n(\alpha)^{T_n}\}}} \right|. \end{aligned}$$

Under assumptions of Lemma 4.7.2, from the central limit theorem for triangular arrays (e.g. Theorem 27.2 in Billingsley, 1995) we obtain

$$a_n \left| \mathbb{P}[(X,Y) \in L_n(\alpha)^{T_n}] - \frac{1}{n}\sum_{i=1}^n 1_{\{(X_i,Y_i)\in L_n(\alpha)^{T_n}\}} \right| \underset{n\to\infty}{\overset{\mathbb{P}}{\to}} 0,$$

$$a_n \left| \int_{L_n(\alpha)^{T_n}} x\, f_{X,Y}(x,y)\lambda(\mathrm{d}x\,\mathrm{d}y) - \frac{1}{n}\sum_{i=1}^n X_i 1_{\{(X_i,Y_i)\in L_n(\alpha)^{T_n}\}} \right| \underset{n\to\infty}{\overset{\mathbb{P}}{\to}} 0.$$

Hence the result. □
Then to prove Theorem 4.4.3 we can write (4.7) as

$$\begin{aligned} \beta_n &\left| \mathbb{E}[X|(X,Y) \in L(\alpha)^{T_n}] - \frac{\displaystyle\sum_{i=1}^n X_i 1_{\{(X_i,Y_i)\in L_n(\alpha)^{T_n}\}}}{\displaystyle\sum_{i=1}^n 1_{\{(X_i,Y_i)\in L_n(\alpha)^{T_n}\}}} \right| \\ &\le \beta_n \left| \mathbb{E}[X|(X,Y) \in L(\alpha)^{T_n}] - \mathbb{E}[X|(X,Y) \in L_n(\alpha)^{T_n}] \right| \\ &\quad + \beta_n \left| \mathbb{E}[X|(X,Y) \in L_n(\alpha)^{T_n}] - \frac{\displaystyle\sum_{i=1}^n X_i 1_{\{(X_i,Y_i)\in L_n(\alpha)^{T_n}\}}}{\displaystyle\sum_{i=1}^n 1_{\{(X_i,Y_i)\in L_n(\alpha)^{T_n}\}}} \right|. \end{aligned}$$

The result is a straightforward application of Lemma 4.7.1 and Lemma 4.7.2. □

Acknowledgements: The authors thank two anonymous referees for relevant remarks and constructive comments on a previous version of the paper. The authors thank Alberto Rodríguez-Casal for his fruitful comments. They thank also Stéphane Loisel and Ragnar Norberg for helpful discussions. This work has been partially supported by the French research national agency (ANR) under the reference ANR-08BLAN-0314-01. The fourth author thanks the BNP Paribas Cardif Chair "Management de la modélisation" for financial support. The views expressed in this document are the author's own and do not necessarily reflect those endorsed by BNP Paribas Cardif.

Bibliography

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

Baíllo, A. (2003). Total error in a plug-in estimator of level sets. *Statist. Probab. Lett.*, 65(4):411–417.

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

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.

Biau, G., Cadre, B., and Pelletier, B. (2007). A graph-based estimator of the number of clusters. *ESAIM. Probability and Statistics*, 11:272–280.

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.

Cadre, B. (2006). Kernel estimation of density level sets. *J. Multivariate Anal.*, 97(4):999–1023.

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

Cavalier, L. (1997). Nonparametric estimation of regression level sets. *Statistics. A Journal of Theoretical and Applied Statistics*, 29(2):131–160.

Chaubey, Y. P. and Sen, P. K. (2002). Smooth estimation of multivariate survival and density functions. *J. Statist. Plann. Inference*, 103(1-2):361–376. C. R. Rao 80th birthday felicitation volume, Part I.

Cuevas, A. and Fraiman, R. (1997). A plug-in approach to support estimation. *The Annals of Statistics*, 25(6):2300–2312.

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.

Cuevas, A. and Rodríguez-Casal, A. (2004). On boundary estimation. *Advances in Applied Probability*, 36(2):340–354.

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

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.

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

Fernández-Ponce, J. M. and Suárez-Lloréns, A. (2002). Central regions for bivariate distributions. *Austrian Journal of Statistic*, 31(2–3):141–156.

Frees, E. W. and Valdez, E. A. (1998). Understanding relationships using copulas. *North American Actuarial Journal*, 2(1):1–25.

Hartigan, J. A. (1987). Estimation of a convex density contour in two dimensions. *J. Amer. Statist. Assoc.*, 82(397):267–270.

Koltchinskii, V. I. (1997). M-estimation, convexity and quantiles. *Ann. Statist.*, 25(2):435–477.

Laloë, T. (2009). *Sur Quelques Problèmes d'Apprentissage Supervisé et Non Supervisé*. PhD thesis, University Montpellier II.

Massé, J.-C. and Theodorescu, R. (1994). Halfplane trimming for bivariate distributions. *J. Multivariate Anal.*, 48(2):188–202.

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

Polonik, W. (1995). Measuring mass concentrations and estimating density contour clusters—an excess mass approach. *The Annals of Statistics*, 23(3):855–881.

Polonik, W. (1997). Minimum volume sets and generalized quantile processes. *Stochastic Process. Appl.*, 69(1):1–24.

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.

Rodríguez-Casal, A. (2003). *Estimacíon de conjuntos y sus fronteras. Un enfoque geometrico*. PhD thesis, University of Santiago de Compostela.

Rossi, C. (1973). Sulle curve di livello di una superficie di ripartizione in due variabili; on level curves of two dimensional distribution function. *Giornale dell'Istituto Italiano degli Attuari*, 36:87–108.

Rossi, C. (1976). Proprietà geometriche delle superficie di ripartizione. *Rend. Mat. (6)*, 9(4):725–736 (1977).

Serfling, R. (2002). Quantile functions for multivariate analysis: approaches and applications. *Statist. Neerlandica*, 56(2):214–232. Special issue: Frontier research in theoretical statistics, 2000 (Eindhoven).

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

Tsybakov, A. B. (1997). On nonparametric estimation of density level sets. *The Annals of Statistics*, 25(3):948–969.

Conclusion et perspectives

Dans ce travail de thèse, nous avons apporté des contributions supplémentaires aux thèmes de la modélisation et de l'estimation des risques en dimension plus grande que un. Tout au long de cette thèse, nous avons porté une attention particulière aux problématiques liées à la structure de dépendance entre les risques.

Conclusions

Le premier chapitre de la thèse est constitué d'une introduction générale avec une présentation des principaux résultats obtenus.

Dans le deuxième chapitre, nous avons proposé un estimateur consistant de la queue d'une distribution bivariée, en utilisant l'étude de la limite de la Upper Tail Dependence Copula (UTDC). En particulier nous nous sommes intéressés au problème de l'indépendance asymptotique.

Dans le troisième chapitre, nous avons abordé la problématique de l'extension de mesures de risque classiques dans un cadre multivarié : *Value-at-Risk* (VaR) et *Conditional-Tail-Expectation* (CTE). Les résultats obtenus dans cette partie ont montré que nos mesures multidimensionnelles étendent certaines propriétés vérifiées en dimension un, par rapport au changement du niveau du risque et au changement de la structure de dépendance du vecteur. Une attention particulière a été portée aux notions d'ordre stochastique et de Kendall multidimensionnelle.

Le quatrième chapitre de cette thèse a apporté des éléments de réponse au problème de l'estimation de la *Conditional-Tail-Expectation* (CTE), introduite dans la section précédente, dans un cadre bivarié. Grâce à une méthode *plug-in* pour les ensembles de niveau de la fonction de répartition bivariée, cette approche nous a permis de construire un estimateur de la CTE consistant et facile à calculer.

Perspectives

Les perspectives de ces travaux sont nombreuses. Elles peuvent être essentiellement fondées sur des développements théoriques des résultats présentés dans cette thèse ou bien sur la mise au point dans des procédures d'évaluation de risque dans des contextes d'assurance, finance ou encore environnementaux et climatiques. Nous essayons ici d'explorer davantage dans le détail quelques pistes possibles.

En ce qui concerne le deuxième chapitre plusieurs perspectives de recherche sont envisageables. Une piste consiste à obtenir des convergences en loi, en partant de résultats comme le Théorème C, pour les estimateurs du paramètre de dépendance θ (voir (2.19)), pour les fonctions g_X, g_Y et G (voir Section 2.3.1) et pour notre estimateur de la queue F^*. Un autre développement possible de ce travail est l'optimisation de l'estimateur par rapport au choix de la suite auxiliaire k_n présente dans la procédure d'estimation (voir Section 2.3) ou encore une comparaison plus approfondie avec d'autres estimateurs de la queue de distribution déjà existants dans la littérature. Afin de pouvoir utiliser notre estimateur de la queue F^* dans la méthode *plug-in* pour les ensembles de niveau proposée dans le quatrième chapitre de cette thèse, il serait également intéressant de démontrer, par exemple, des résultats de consistance en norme uniforme pour F^*. En effet l'utilisation d'un estimateur de la queue de la distribution représente une approche possible afin d'améliorer l'estimation de la $\mathrm{CTE}_\alpha(X,Y)$ pour des hautes valeurs du niveau de risque α.

Le deuxième travail de cette thèse offre un certain nombre de perspectives intéressantes. Il serait bien d'utiliser ces mesures multivariées dans des études pratiques en assurance ou en finance, pour analyser leur comportement avec des données réelles et leur possible utilisation dans la gestion des risques. En particulier elles pourraient être étudiées dans un contexte macro-prudentiel pour tester les intuitions présentées dans le Chapitre 3 et de mieux explorer les résultats théoriques obtenus. Un autre développement possible est la comparaison théorique entre notre $\mathrm{CTE}_\alpha(\mathbf{X})$ et les autre mesures de risque qui généralisent la CTE en dimension plus grande que un (voir Introduction du Chapitre 3). Notamment il serait intéressant d'évaluer leur comportement par rapport à différents scenarios de risque et leurs possibles applications dans différents cas réels de couverture du risque.

En ce qui concerne le travail sur l'estimation de la CTE, il serait intéressant de donner la vitesse optimale (au moins aux constantes près) de $\left|\,\mathrm{CTE}_\alpha(X,Y) - \widehat{\mathrm{CTE}}_\alpha^{T_n}(X,Y)\,\right|$ pour des classe de loi bivariées données. Une autre approche envisageable pour l'estimation de la CTE serait d'estimer directement la Kendall plutôt que la fonction de répartition bivariée, toujours par méthode *plug-in*. Il serait intéressant d'approfondir cette piste et de confronter les deux approches. Sur l'estimation de la Kendall, un travail de référence est par exemple Barbe *et al.* (1996). Comme il a été dit précédemment, il serait intéressant aussi de développer un estimateur de ces mesures en utilisant un estimateur de la queue de la distribution. Dans ce sens des travaux de référence sont par exemple : de Haan et Haung (1995), Einmahl et de Haan (2009), Cai *et al.* (2011). Enfin, parmi les perspectives à court terme du dernier travail de cette thèse, il y a la mise en place d'un package `R` pour l'estimation des ensembles de niveau d'une fonction de répartition bivariée selon la méthode *plug-in* présentée dans le Chapitre 4.

Annexe A

Rappels et définitions utiles

Cet appendice propose des rappels, notions et outils mathématiques utiles dans la lecture de cette thèse.

A.1 Copules

Une notion fondamentale dans toute cette thèse est celle de *copule*. Cette notion a été introduite par Sklar (1959). Les définitions ci-dessous sont présentées en dimension deux par simplicité. L'extension en dimension plus grande est immédiate (voir par exemple Nelsen, 1999 ; Section 2.10).

Définition A.1.1 (Copule) *Une copule C bivariée est une fonction de répartition bidimensionnelle sur $[0,1]^2$, dont les lois marginales sont uniformes sur $[0,1]$.*

Pour un examen exhaustif des propriétés de copules voir par exemple Nelsen (1999), Chapitres 2 et 3. On parle aussi de copule associée à une fonction de répartition ou, par abus de langage, à un vecteur aléatoire :

Définition A.1.2 *Soit $X = (X_1, X_2)$ un vecteur aléatoire de $\mathbb{R}^2$ dont les marges sont continues (les fonctions de répartitions marginales $F_i(x) = \mathbb{P}(X_i \leq x)$ sont continues). La copule associée à X est la fonction de répartition du vecteur $(F_1(X_1), F_2(X_2))$, i.e. $C(u_1, u_2) = \mathbb{P}[F_1(X_1) \leq u_1, F_2(X_2) \leq u_2]$ avec $(u_1, u_2) \in [0,1]^2$.*

Les copules selon la terminologie retenue par Sklar (1959) sont alors des "*fonctions de dépendance*". Elles ne dépendent que de la structure de dépendance des composantes du vecteur X, et elles sont en particulier invariantes par transformation croissante des marges (voir Nelsen, 1999 ; Théorème 2.4.3). Ainsi, le vecteur (X_1, X_2) a la même copule que $(\psi(X_1), \psi(X_2))$ pour des fonctions ψ croissantes. De plus, ce sont des fonctions qui "couplent" les lois marginales, au sens où

$$\mathbb{P}[X_1 \leq x_1, X_2 \leq x_2] = C(F_1(x_1), F_2(x_2)). \tag{A.1}$$

Si F_1 et F_2 sont continues alors il existe une unique copule C qui vérifie A.1 (Théorème de Sklar, voir par exemple Nelsen, 1999 ; Théorème 2.3.3).
Pour chaque copula C et $(u, v) \in [0,1]^2$ on a

$$\max\{u + v - 1, 0\} \leq C(u, v) \leq \min\{u, v\}. \tag{A.2}$$

Nous notons $W(u, v) := \max\{u+v-1, 0\}$, le *Fréchet-Hoeffding lower bound* et $M(u, v) := \min\{u, v\}$, le *Fréchet-Hoeffding upper bound*. Deux variables aléatoires avec copula M sont comonotones et avec copula W anti-comonotones (voir Nelsen, 1999 ; Section 2.5). Dans le cas de deux variables aléatoires indépendant on a :

Théorème A.1.3 *(Nelsen, 1999; Théorème 2.4.2) Soit $X = (X_1, X_2)$ un vecteur aléatoire de $\mathbb{R}^2$ dont les marges sont continues. Alors X et Y sont indépendants si et seulement si $C(u, v) = \Pi(u, v) = u\,v$, pour $(u, v) \in [0,1]^2$.*

Définition A.1.4 (Copule de survie) *Soit* (U_1, U_2) *un vecteur aléatoire de fonction de répartition la copule* C*, alors la fonction de répartition de* $(1 - U_1, 1 - U_2)$ *est une copule appelée copule de survie associée à* C*, et elle est notée* C^**.*

Il faut faire attention à ne pas confondre la copule de survie C^* avec la fonction de répartition survie du vecteur aléatoire (U_1, U_2) car :

$$\mathbb{P}[U_1 > u_1, U_2 > u_2] = 1 - u_1 - u_2 + C(u_1, u_2) = C^*(1 - u_1, 1 - u_2),$$

pour $(u_1, u_2) \in [0, 1]^2$. Dans le Chapitre 2 nous nous intéressons à une région de l'espace où toutes les composantes du couple (X,Y) sont grandes en même temps, ce qui conduit à étudier la copule dans la région proche du coin $(1, 1)$. Donc cette notion de copule de survie permettra d'alléger les notations, en notant que l'étude de C est équivalente à l'étude de C^* en $(0, 0)$ (voir par exemple la condition en (2.8)).

Une classe importante de copules, qui est traité dans cette thèse, est celle des copules Archimédiennes, presentées ci-dessous dans un cadre bivarié :

Définition A.1.5 (Copule Archimédienne) *Considérons une fonction* $\phi : [0, 1] \to \mathbb{R}_+$ *avec dérivées première et seconde en* $(0, 1)$ *satisfaisant :*

$$\phi(1) = 0, \ \phi'(\tau) < 0 \ \ et \ \ \phi''(\tau) > 0, \ \forall \ \tau \in (0, 1). \tag{A.3}$$

Chaque fonction ϕ *qui vérifie* (A.3) *génère une copule bivariée donnée par :*

$$C_\phi(u_1, u_2) = \phi^{\leftarrow}(\phi(u_1) + \phi(u_2)), \quad pour \ \ u_1, u_2 \in [0, 1]$$

Les copules comme C_ϕ*, avec* ϕ *comme en* (A.3)*, sont appelées copules Archimédiennes et la fonction* ϕ *est appelée générateur de la copule.*

Si le générateur est un outil naturel pour caractériser les copules, d'autres fonctions peuvent aussi être utilisées, en particulier la fonction de répartition de Kendall K, définie comme la fonction de répartition de la variable $C(X_1, X_2)$, où (X_1, X_2) est un vecteur aléatoire de fonction de répartition C, i.e.

$$K(t) = \mathbb{P}[C(X_1, X_2) \leq t], \quad \text{pour } t \in [0, 1].$$

Si C est une copule Archimédienne alors la notion de générateur et celle de fonction de répartition de Kendall sont liées. En effets si C a un générateur ϕ, alors $K(t) = t - \frac{\phi(t)}{\phi'(t)}$, où ϕ' désigne la dérivée de ϕ sur $[0, 1)$. La notion de fonction de répartition de Kendall est utilisée dans le Chapitre 3 de cette thèse, dans un cadre multivarié. Pour de plus amples détails, nous renvoyons le lecteur intéressé à la Section 3.1.

A.2 Fonction de répartition Burr

Dans la théorie des probabilités, la statistique et l'économétrie, la fonction de répartition *Burr Type XII* ou simplement la fonction de répartition *Burr* est une distribution

de probabilité continue pour une variable aléatoire non-négative. Elle est connue aussi comme la distribution Singh-Maddala ou parfois appelée "generalized log-logistic distribution". Cette distribution est souvent utilisée dans les modèles de gestion du risque, autant que fonction de répartition à queues épaisses.

La distribution Burr a fonction de répartition de la forme

$$F(x; c, k) = 1 - (1 + x^c)^{-k} \tag{A.4}$$

et fonction de densité

$$f(x; c, k) = c\,k\,\frac{x^{c-1}}{(1 + x^c)^{k+1}}. \tag{A.5}$$

Dans la Figure A.1 on observe la fonction de répartition et de densité Burr pour différents valeurs de c et k.

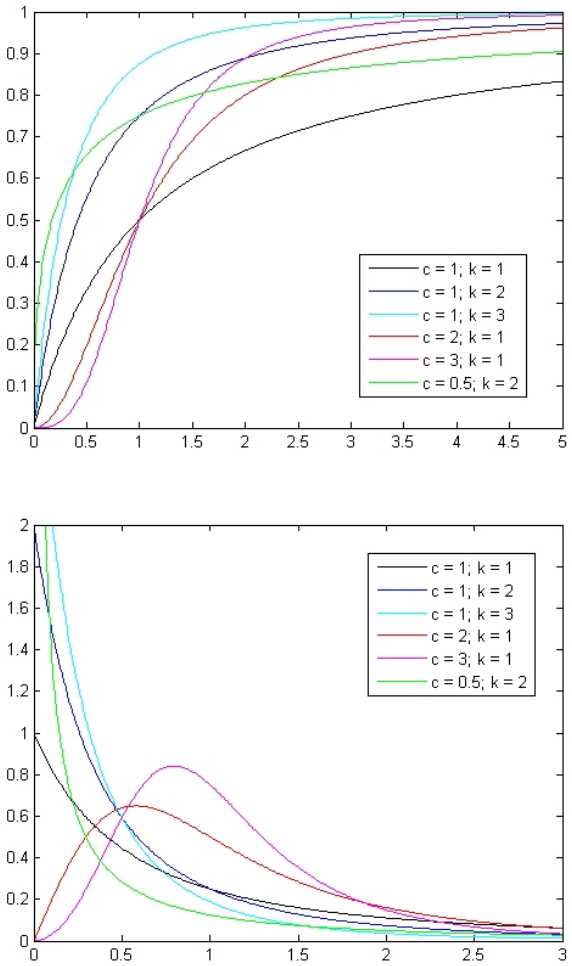

FIGURE A.1 – (en haut) Fonction de répartition Burr ; (en bas) Fonction de densité Burr, pour différents valeurs de c et k.

Annexe B

Simulations, illustrations et approfondissements

B.1 Lien entre les paramètres de dépendance η et θ

Dans cette section nous détaillons le lien qui existe entre le paramètre de dépendance du modèle de Ledford et Tawn (1.20)-(1.21) et le paramètre θ dans le modèle de Juri et Wüthrich (2004) (voir (1.25)-(1.27)).

Sous les hypothèses du modèle de Ledford et Tawn en (1.20)-(1.21), et en utilisant l'expression (1.31), la condition (1.25) proposée par Juri et Wüthrich (2004), devient :

$$\lim_{u\to 1} \frac{C^*(x\,(1-u),\, y\,(1-u))}{C^*(1-u,\, 1-u)} = G(x,y) = \widetilde{g}\left(\frac{1}{x},\, \frac{1}{y}\right) x^{\frac{1}{2\,\eta}}\, y^{\frac{1}{2\,\eta}},$$

pour $x, y \geq 0$, $\widetilde{g}$ étant définie en (1.21), et avec $\widetilde{g}(cx,\, cy) = \widetilde{g}(x,\, y)$ pour chaque $c > 0$ ($\widetilde{g}$ est homogène d'ordre zéro). En utilisant les propriétés de G et $\widetilde{g}$ on obtient

$$g_X(x) = G(x,1) = \widetilde{g}\left(\frac{1}{x},\, 1\right) x^{\frac{1}{2\,\eta}} = \widetilde{g}\,(1,\, x)\; x^{\frac{1}{2\,\eta}}.$$

De plus en utilisant la forme de $G(x,y)$:

$$G(x,y) = \begin{cases} x^{\theta} g_Y(\frac{y}{x}) & \text{for } \frac{y}{x} \in [0,1], \\ y^{\theta} g_X(\frac{x}{y}) & \text{for } \frac{y}{x} \in (1,\infty), \end{cases}$$

(voir Proposition 2.2.2) on obtient, par exemple, pour $\frac{y}{x} \in (1,\infty)$,

$$G(x,y) = \widetilde{g}\left(\frac{y}{x},\, 1\right) x^{\frac{1}{2\,\eta}}\, y^{\frac{1}{2\,\eta}} = y^{\theta} g_X\left(\frac{x}{y}\right) = y^{\theta}\, G\left(\frac{x}{y},\, 1\right) = y^{\theta}\, \widetilde{g}\left(\frac{y}{x},\, 1\right) \left(\frac{x}{y}\right)^{\frac{1}{2\,\eta}}.$$

Nous avons donc montré qu'un modèle du type (1.20)-(1.21) (pour une certaine valeur de η) vérifie un modèle du type (1.25)-(1.27) avec $\theta = 1/\eta$ (voir Juri et Wüthrich, 2004 ; Section 4.2)

Nous allons à présent illustrer le lien entre θ et η dans des cas spécifiques :

Cas Clayton Survie

Nous considérons une fonction de distribution avec marges Fréchet unitaires et une copule Clayton survie de paramètre $a > 0$:

$$F(z_1, z_2) = F(z_1) + F(z_2) - 1 + (\overline{F}(z_1)^{-1/a} + \overline{F}(z_2)^{-1/a} - 1)^{-a}, \qquad \text{avec } a > 0.$$

Dans ce cas le modèle de Ledford et Tawn (1.20)-(1.21) est vérifié avec $\eta = 1$ et

$$L(z_1, z_2) = (w\,(1-w))^{\frac{-1}{2}}\, (w^{\frac{-1}{a}} + (1-w)^{\frac{-1}{a}})^{-a},$$

pour $w = \frac{z_1}{z_1+z_2}$ (voir Ledford et Tawn, 1997). Par exemple pour $a = 1$ on obtient :

$$\lim_{t\to 0} \frac{L(tx,\, ty)}{L(t,t)} = \frac{2\,\sqrt{xy}}{x+y} = \widetilde{g}(x,y),$$

et, en utilisant (1.32),

$$\lim_{u\to 1}\frac{C^*(x\,(1-u),\,y\,(1-u))}{C^*(1-u,\,1-u)}=\lim_{u\to 1}\frac{L\big(\frac{1}{x\,(1-u)},\frac{1}{y\,(1-v)}\big)\,\big(\frac{1}{x\,(1-u)}\big)^{\frac{-1}{2}}\,\big(\frac{1}{y\,(1-v)}\big)^{\frac{-1}{2}}}{L\big(\frac{1}{1-u},\frac{1}{1-v}\big)\,\big(\frac{1}{1-u}\big)^{\frac{-1}{2}}\,\big(\frac{1}{1-v}\big)^{\frac{-1}{2}}}$$
$$=\widetilde{g}\left(\frac{1}{x},\,\frac{1}{y}\right)x^{\frac{1}{2}}\,y^{\frac{1}{2}}=\frac{2\,x\,y\,x^{\frac{-1}{2}}\,y^{\frac{-1}{2}}}{x+y}\sqrt{y}\,\sqrt{y}=\frac{2\,x\,y}{x+y}=G(x,y),$$

donc $\theta=\eta=1$ (cas de dépendance asymptotique).

Cas Fairlie-Gumbel-Morgenstern

Nous considérons une fonction de distribution avec marges Fréchet unitaires et une copule Fairlie-Gumbel-Morgenstern de paramètre $a\in(-1,1]$:

$$F(z_1,z_2)=F(z_1)\,F(z_2)\,(1+a\,\overline{F}(z_1)\,\overline{F}(z_2)),\qquad \text{avec } a\in(-1,1].$$

Dans ce cas le modèle de Ledford et Tawn (1.20)-(1.21) est vérifié avec $\eta=1/2$ et

$$L(z_1,z_2)=1+a,$$

(voir Ledford et Tawn, 1997). On obtient :

$$\lim_{t\to 0}\frac{L(tx,\,ty)}{L(t,t)}=1=\widetilde{g}(x,y),$$

et, en utilisant (1.32),

$$\lim_{u\to 1}\frac{C^*(x\,(1-u),\,y\,(1-u))}{C^*(1-u,\,1-u)}=\widetilde{g}\left(\frac{1}{x},\,\frac{1}{y}\right)x^{\frac{1}{2\,\eta}}\,y^{\frac{1}{2\,\eta}}=x\,y=G(x,y),$$

donc $\theta=1/\eta=2$ (cas *quasi-indépendance*, voir Section 1.2).

Cas gaussien

Nous considérons une fonction de distribution normale avec marges Fréchet unitaires et un paramètre de corrélation $|\rho|<1$.
Dans ce cas le modèle de Ledford et Tawn (1.20)-(1.21) est vérifié avec $\eta=(1+\rho)/2$ et

$$L(z_1,z_2)=\frac{\left((z_1\,z_2)^{2\,\frac{\rho\sqrt{\ln(z_1)\,\ln(z_2)}}{\ln(z_2\,z_1)-1}}\,(\ln(z_1))^{1-\rho\sqrt{\frac{\ln(z_2)}{\ln(z_1)}}}\,(\ln(z_2))^{1-\rho\sqrt{\frac{\ln(z_1)}{\ln(z_2)}}}\,(4\,\pi)^{\rho\left(2\,\rho-\sqrt{\frac{\ln(z_2)}{\ln(z_1)}}-\sqrt{\frac{\ln(z_1)}{\ln(z_2)}}\right)}\right)^{\left(2-2\,\rho^2\right)^{-1}}}{(1-\rho^2)^{-3/2}\left((1+\rho^2)\sqrt{\ln(z_1)\,\ln(z_2)}-\rho\,\ln(z_2\,z_1)\right)},$$

(voir Ledford et Tawn, 1997). On obtient :

$$\lim_{t\to 0}\frac{L(tx,\,ty)}{L(t,t)}=1=\widetilde{g}(x,y),$$

$$\lim_{u\to 1}\frac{C^*(x\,(1-u),\,y\,(1-u))}{C^*(1-u,\,1-u)}=\widetilde{g}\left(\frac{1}{x},\,\frac{1}{y}\right)x^{\frac{1}{2\,\eta}}\,y^{\frac{1}{2\,\eta}}=x^{\frac{1}{1+\rho}}\,y^{\frac{1}{1+\rho}}=G(x,y),$$

donc $\theta = 1/\eta = 2/(1+\rho)$ (cas d'indépendance asymptotique, avec association positive si $\rho > 0$ et association négative si $\rho < 0$) et

$$\lim_{u \to 1} C_u^{up}(x, y) = x\, y = C^{*\,G}(x, y), \tag{B.1}$$

(voir (2.9)).
Comme Juri et Wüthrich le remarquent dans leur article (voir Juri et Wüthrich, 2004, page 233) les coefficients de dépendance sont souvent dépendants des facteurs d'échelle et donc leur interprétation est plus difficile. Par exemple dans le cas gaussien ci-dessus nous avons :

$$\lambda = 0, \quad \eta = (1+\rho)/2, \quad \theta = 2/(1+\rho), \quad \chi := 2\,\eta - 1 = \rho,$$

(voir (59)-(62) dans Juri et Wüthrich, 2004). En revanche la limite en (B.1) apparaît plus claire puisqu'elle met en évidence la structure de dépendance de la copule gaussienne, avec la présence d'une dépendance faible qui disparaît asymptotiquement dans les queues de la distribution.

B.2 Estimation du paramètre θ, simulations

Cette section peut être considérée comme un appendice de la Section 2.3.2. Ici nous proposons plusieurs exemples pour l'estimation du paramètre θ à partir de données simulées et réelles.

Estimation du paramètre θ dans le cas de la copule indépendante :
Les hypothèses de la Proposition 2.3.3 sont satisfaites, par exemple, dans le cas des copules Ali Mikhail-Haq, Frank, Clayton avec $a \geq 0$, Indépendante et Fairlie-Gumbel-Morgenstern. La Figure B.1 nous montre le comportement de l'estimateur de θ dans le cas de la copule indépendante : la courbe moyenne (la ligne pleine), la courbe moyenne $+/-$ l'écart-type empirique (la ligne pointillée) et l'erreur quadratique moyenne sur 100 échantillons de taille $n = 1000$. Rappelons que dans ce cas $\theta = 2$.

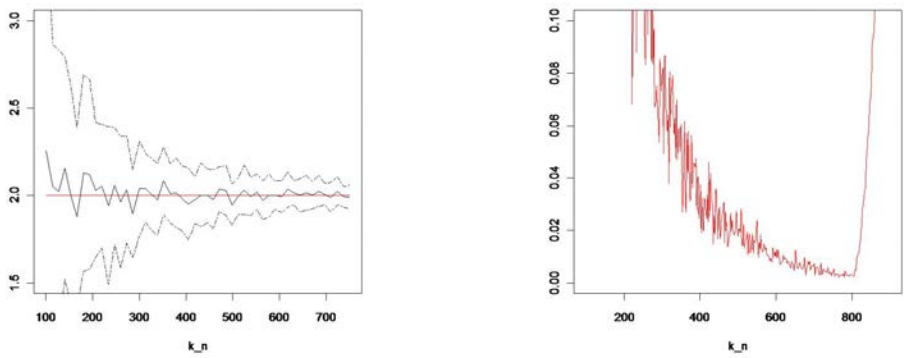

FIGURE B.1 – Copule Indépendante : (gauche) Estimateur pour θ, $(k, \widehat{\theta}_x)$ avec $x = 0.8$, courbe moyenne (la ligne pleine), la courbe moyenne $+/-$ l'écart-type empirique (la ligne pointillée) ; (droite) erreur quadratique moyenne pour $\widehat{\theta}_x$, avec $x = 0.8$.

Estimation du paramètre θ dans le cas $\theta = 3$:
Nous présentons un exemple de copule qui ne satisfait pas les hypothèses de la Proposition 2.3.3. Nous considérons la copule asymptotiquement indépendante ($\lambda = 0$) et non-symétrique :

$$C(u, v) = x\,y - \frac{1}{8}(1 - |2\,x - 1|)\,(1 - (2\,y - 1)^2), \tag{B.2}$$

pour plus de détails voir Benth et Kettler (2011). Dans ce cas $\frac{\partial^2 C}{\partial u \partial v}(1, 1) = 0$ et nous obtenons

$$G(x, y) = x\,y^2, \quad g_X(x) = x, \quad g_Y(y) = y^2 \quad \text{et} \quad \theta = 3.$$

La Figure B.2 nous montre le comportement de l'estimateur de θ : la courbe moyenne (la ligne pleine), l'écart-type empirique (la ligne pointillée) et l'erreur quadratique moyenne sur 100 échantillons de taille $n = 1000$. Nous renvoyons le lecteur à l'article de Benth et Kettler (voir Benth et Kettler 2011, Section 4) pour la procédure de simulation d'un

échantillon avec copule (B.2).

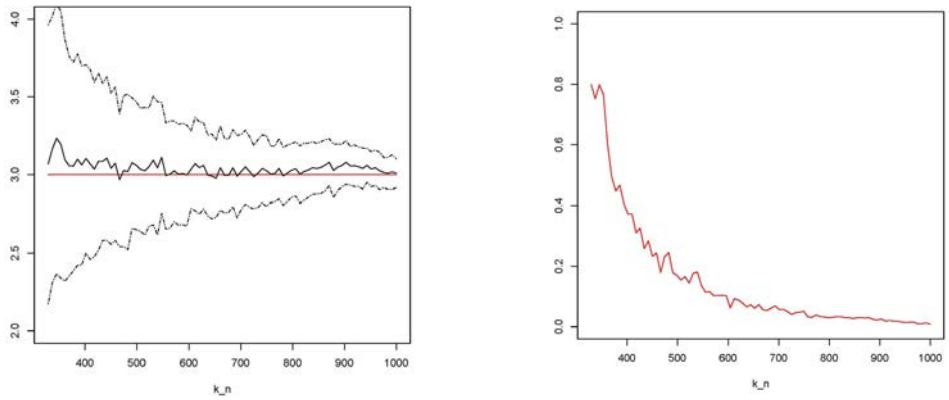

FIGURE B.2 – Copule (B.2) : (gauche) Estimateur pour θ, $(k, \widehat{\theta}_{\frac{y}{x}})$, courbe moyenne (la ligne pleine), la courbe moyenne $+/-$ l'écart-type empirique (la ligne pointillée) ; (droite) erreur quadratique moyenne pour $\widehat{\theta}_{\frac{y}{x}}$, avec $x = 0.4$ et $y = 0.65$.

Estimation de θ avec les données réelles NHANES

Nous considérons l'estimation du coefficient de dépendance θ pour des données réelles NHANES (National Health and Nutrition Examination Survey). X représente la hauteur en cm et Y le poids en kg. La taille de l'échantillon est $n = 3500$. Les données sont représentées en Figure B.3 (gauche). Dans la Figure B.3 (droite) nous proposons l'estimation de θ. Nous remarquons que $\theta \approx 1$ pratiquement sur toute la fenêtre des valeurs possibles de la séquence k_n. En effet nous pouvons observer que les données montrent une dépendance asymptotique.

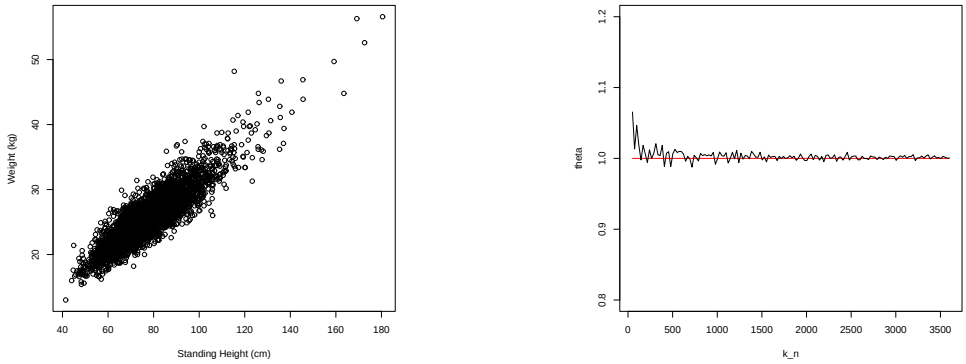

FIGURE B.3 – (gauche) Données NHANES (Hauteur vs Poids) ; (droite) Estimateur pour θ, $(k, \widehat{\theta}_x)$ avec $x = 0.07$.

B.3 Estimateurs dans le cadre unidimensionnel

Dans cette section nous présentons des illustrations du Théorème 2.5.3 et des Corollaires 2.5.1 et 2.5.2 (nos versions déconditionnées des Théorème 3.2 et 8.1 de Smith, 1987) pour la normalité des estimateurs de vraisemblance des paramètres d'une loi de Pareto généralisée et la convergence de l'estimateur POT unidimensionnel $\widehat{F}$ (voir Section 2.4.1). Nous utilisons ici les notations de la Section 2.5.

Estimateurs de vraisemblance des paramètres d'une loi de Pareto généralisée :
Nous supposons que F est une fonction Burr de la forme :

$$F_X(x) = 1 - \frac{1}{1+x}, \qquad \text{pour } x \geq 0. \tag{B.3}$$

Cette loi est dans le domaine d'attraction de Fréchet, Φ_α, avec $\alpha = 1$ et $L(x) = \frac{x}{1+x}$. Nous remarquons que $L(x)$ est à variation lente à l'infini et $\frac{L(tx)}{L(x)} = 1 + k(t)\frac{1}{x} + o(\frac{1}{x})$, $\forall\, t > 0$, pour $x \to \infty$. Donc $\rho = -1$ et $k(t) = c\, h_{-1}(t)$, avec $c = 1$ et $h_{-1}(t) = 1 - \frac{1}{t}$.
Nous proposons de considérer une suite de seuils $u_n = \frac{\sqrt[3]{n}}{3}$, Nous remarquons alors que :

- $u_n = \frac{\sqrt[3]{n}}{3} \xrightarrow{n\to\infty} \infty$, et $n(1 - F(\frac{\sqrt[3]{n}}{3})) \xrightarrow[n\to\infty]{} \infty$ (voir Condition (2.32) du Corollaire 2.5.1) ;
- $\frac{1}{2}\sqrt{n(1 - F(\frac{\sqrt[3]{n}}{3}))}, \frac{3}{\sqrt[3]{n}} \xrightarrow[n\to\infty]{} \mu = \frac{3\sqrt{3}}{2} \approx 2.59$ (voir Condition (2.33) du Corollaire 2.5.1).

Soient $\xi = -1$ et $\sigma_n = \frac{\sqrt[3]{n}}{3}$. Alors par le Corollaire 2.5.1 on obtient :

$$\sqrt{N_n}\begin{pmatrix} \frac{\widehat{\sigma}_n}{\frac{\sqrt[3]{n}}{3}} - 1 \\ \widehat{k}_n + 1 \end{pmatrix} \xrightarrow{\mathscr{L}} \mathscr{N}\left(\begin{bmatrix} 3\sqrt{3} \\ 0 \end{bmatrix}; \begin{pmatrix} 4 & 2 \\ 2 & 4 \end{pmatrix}\right), \tag{B.4}$$

où $3\sqrt{3} \approx 5.19$.

Nous calculons la moyenne et l'écart-type pour un échantillon de taille n, (moyennés sur 200 itérations), des deux coordonnées :

$\sqrt{N_n}\left(\widehat{\sigma}/(\frac{\sqrt[3]{n}}{3}) - 1\right)$	moyenne	écart-type
$n = 5000$	4.77	2.17
$n = 10000$	4.98	2.11
$n = 100000$	5.19	2.02

$\sqrt{N_n}\left(\widehat{k} + 1\right)$	moyenne	écart-type
$n = 5000$	-0.04	2.04
$n = 10000$	0.02	2.02
$n = 100000$	0.004	2.01

Dans les Figures B.4 et B.5 nous proposons quelques illustrations graphiques du résultat (B.4)

FIGURE B.4 – Histogramme pour $\sqrt{N_n}\left(\frac{\widehat{\sigma}_n}{\frac{\sqrt[3]{n}}{3}}-1\right)$ (gauche) et $\sqrt{N_n}\left(\widehat{k}_n+1\right)$ (droite) ; avec $n = 10000$, 200 itérations.

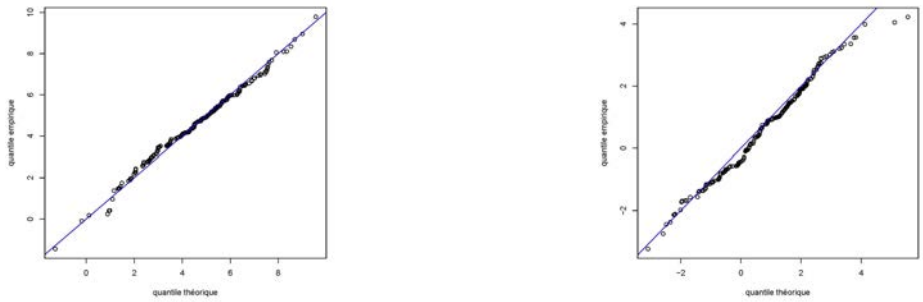

FIGURE B.5 – QQplot pour $\sqrt{N_n}\left(\frac{\widehat{\sigma}_n}{\frac{\sqrt[3]{n}}{3}}-1\right)$(gauche) et $\sqrt{N_n}\left(\widehat{k}_n+1\right)$ (droite) ; avec $n = 10000$, 200 itérations.

Estimateur POT unidimensionnel :
Pour illustrer le Corollaire 2.5.2 et le Théorème 2.5.3 nous choisissons une suite $z_n := \frac{\log(\sqrt[3]{n})}{2}$. Nous remarquons que :

- $z_n := \frac{\log(\sqrt[3]{n})}{2} \xrightarrow{n\to\infty} \infty$, et $\frac{\log(z_n)}{\sqrt{n(1-F(u_n))}} \xrightarrow[n\to\infty]{} 0$ (voir Condition (2.34) du Corollaire 2.5.2) ;
- $z_n^{-s\,\rho}\frac{\phi(v_n z_n^s)}{\phi(v_v)} = 1$ pour $s \in [0,1]$ (voir Corollaire 2.5.2) ;
- $\frac{(z_n)}{\sqrt{n(1-F(u_n))}} \xrightarrow[n\to\infty]{} 0$ (voir Condition (2.35) du Théorème 2.5.3).

Alors, grâce au Corollaire 2.5.2 on a :

$$\mathscr{Z} := r_n\left[F\left(\frac{\sqrt[3]{n}}{3}\,\frac{\log(\sqrt[3]{n})}{2}\right) - \widehat{F}\left(\frac{\sqrt[3]{n}}{3}\,\frac{\log(\sqrt[3]{n})}{2}\right)\right] \xrightarrow[n\to\infty]{\mathscr{L}} \mathscr{N}(\nu = 0, \tau^2 = 4), \quad \text{(B.5)}$$

où

$$r_n = \frac{\sqrt{N_n}}{\log\left(\frac{\log(\sqrt[3]{n})}{2}\right)}\left(\frac{1}{1-F\left(\frac{\sqrt[3]{n}}{3}\,\frac{\log(\sqrt[3]{n})}{2}\right)}\right). \quad \text{(B.6)}$$

De plus, en utilisant le Théorème 2.5.3 on obtient :

$$\widetilde{\mathscr{Z}} := \widetilde{r}_n \left[F\left(\frac{\sqrt[3]{n}}{3} \frac{\log(\sqrt[3]{n})}{2} \right) - \widehat{F}\left(\frac{\sqrt[3]{n}}{3} \frac{\log(\sqrt[3]{n})}{2} \right) \right] \xrightarrow[n\to\infty]{\mathscr{L}} \mathscr{N}(\nu = 0, \tau^2 = 4), \quad \text{(B.7)}$$

où

$$\widetilde{r}_n = \frac{\sqrt{N_n}}{\log\left(\frac{\log(\sqrt[3]{n})}{2}\right)} \left(\frac{1}{1 - \frac{1}{n}\sum_{i=1}^{n} 1_{\left\{X_i \leq \frac{\sqrt[3]{n}}{3} \frac{\log(\sqrt[3]{n})}{2}\right\}}} \right). \quad \text{(B.8)}$$

Nous proposons une brève exploration graphique (Figures B.6 et B.7) des résultats (B.5) et (B.7). Nous prenons $n = 10000$ et 200 itérations.

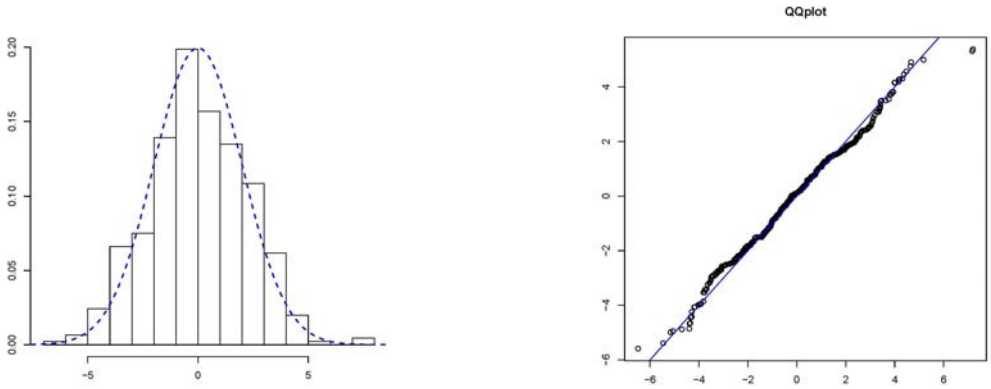

FIGURE B.6 – Histogramme pour $\mathscr{Z}$ (gauche), QQplot for $\mathscr{Z}$ (droite), avec $n = 10000$, 200 itérations.

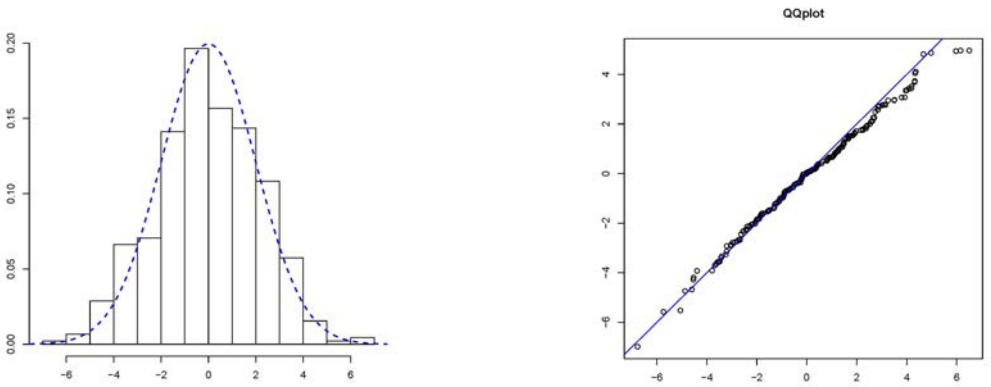

FIGURE B.7 – Histogramme pour $\widetilde{\mathscr{Z}}$ (gauche), QQplot for $\widetilde{\mathscr{Z}}$ (droite), avec $n = 10000$, 200 itérations.

Nous calculons la moyenne et l'écart-type pour un échantillon de taille n, (moyennées sur 200 itérations), de $\mathscr{Z}$ et $\widetilde{\mathscr{Z}}$:

$\mathscr{Z}$	moyenne	écart-type
$n = 5000$	0.15	2.96
$n = 10000$	0.08	2.52
$n = 100000$	0.07	2.05

$\widetilde{\mathscr{Z}}$	moyenne	écart-type
$n = 5000$	0.14	2.95
$n = 10000$	0.09	2.51
$n = 100000$	0.03	2.06

A partir du Théorème 2.5.3 nous pouvons obtenir aussi des intervalles de confiance asymptotique pour $F(u_n\, z_n)$, tels que, en choisissant $\alpha = 0.05$, on obtienne :

$$\mathbb{P}\left[F\left(\frac{\sqrt[3]{n}}{3}\,\frac{\log(\sqrt[3]{n})}{2}\right) \in \left[\widehat{F}\left(\frac{\sqrt[3]{n}}{3}\,\frac{\log(\sqrt[3]{n})}{2}\right) \pm\, \tau\,\frac{z_\alpha}{\widetilde{r}_n}\right]\right] \xrightarrow[n\to\infty]{} 1 - 2\,\alpha = 0.9. \qquad \text{(B.9)}$$

Si nous calculons le nombre de fois, dans les 200 échantillons simulés, où la valeur théorique

$$F\left(\tfrac{\sqrt[3]{n}}{3}\,\tfrac{\log(\sqrt[3]{n})}{2}\right)$$

tombe dans la bande de confiance, nous obtenons que pour $n = 5000$ elle y tombe approximativement 77% des fois, pour $n = 10000$ le 85% et pour $n = 100000$ le 89%.

B.4 Construction de Skorohod

Les Corollaires 2.5.1 et 2.5.2 représentent une version déconditionnée des Théorèmes 3.2 et 8.1 dans Smith (1987). En effet dans les Théorèmes 3.2 et 8.1 de Smith le seuil et le nombre des excès au-delà du seuil sont considérés déterministes, c'est-à-dire des fonctions de n qui divergent avec n. Les résultats dans Smith (1987) sont écrits conditionnellement au nombre d'excès N. Cependant dans les applications le seuil sera supposé déterministe et le nombre d'excès au-delà du seuil sera donc aléatoire. Pour cette raison nous avons besoin de déconditionner les résultats de Smith.

Nous parcourons plus en détail le raisonnement dans le cas du déconditionnement du Théorème 3.2 de Smith (voir Corollaire 2.5.1). Le déconditionnement du Théorème 8.1 de Smith est complètement analogue (voir Corollaire 2.5.2).

L'argument utilisé et repris de l'article de Smith (Smith, 1987) est le suivant. Nous partons du résultat du Théorème 3.2 de Smith, écrit pour N non aléatoire. On suppose maintenant que $N_n \sim \mathrm{Bi}\,(n, 1 - F(u_n))$ avec $\lim_{n\to\infty} 1 - F(u_n) = 0$ et $\lim_{n\to\infty} n(1 - F(u_n)) = \infty$. En utilisant les hypothèses (2.32) et (2.33) du Corollaire 2.5.1, que nous rappelons ici :

$$n(1 - F(u_n)) \xrightarrow[n\to\infty]{} \infty, \quad \sqrt{n(1 - F(u_n))}c\,\phi(u_n) \xrightarrow[n\to\infty]{} \mu(\alpha - \rho),$$

et comme

$$\frac{N_n}{n(1 - F(u_n))} \xrightarrow[n\to\infty]{\mathbb{P}} 1,$$

on a facilement que la Condition (3.2) dans Smith est vraie en probabilité, c'est à dire :

$$\sqrt{N_n}\,c\,\phi(u_n) \xrightarrow[n\to\infty]{\mathbb{P}} \mu(\alpha - \rho). \tag{B.10}$$

Maintenant nous utilisons un argument *"construction de type Skorohod"* (voir Smith, 1987, page 1180). Nous pouvons grâce au Théorème de Skorohod (par exemple voir Théorème 2.2.2 dans Borkar, 1995) nous ramener à un espace de probabilité dans lequel la Condition (3.2) de Smith (c'est à dire la condition en (B.10)) converge presque sûrement vers 1. Plus précisément :

Nous notons $\mathcal{L}(Z) = \mathcal{L}(W)$ si et seulement si Z et W ont la même loi. Etant donné un espace de probabilité $(\Omega, \mathcal{F}, \mathbb{P})$ et une convergence en probabilité, dans cet espace, des éléments $K_1 := \sqrt{N_1}\,c\,\phi(u_1)$, $K_2 := \sqrt{N_2}\,c\,\phi(u_2)$, ..., $K_n := \sqrt{N_n}\,c\,\phi(u_n)$ vers $K := \mu(\alpha - \rho)$ (voir (B.10)) alors il existe un espace de probabilité $(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbb{P}})$ et des éléments $\tilde{K_1}$, $\tilde{K_2}$, ... $\tilde{K_n}$, $\tilde{K}$ tels que $\mathcal{L}(\tilde{K_1}) = \mathcal{L}(K_1)$, $\mathcal{L}(\tilde{K_2}) = \mathcal{L}(K_2)$, ... $\mathcal{L}(\tilde{K_n}) = \mathcal{L}(K_n)$, $\mathcal{L}(\tilde{K}) = \mathcal{L}(K)$ et

$$\tilde{K_n} \xrightarrow[n\to\infty]{p.s.} \tilde{K}.$$

En se plaçant alors dans ce nouvel espace de probabilité $(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbb{P}})$ nous avons un théorème de convergence en loi pour les estimateurs des paramètres, selon la méthode du maximum de vraisemblance, d'une loi de Pareto généralisée, conditionnellement à un évènement presque certain (c'est-à-dire que la convergence en (B.10) sera une convergence presque sûre dans $(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbb{P}})$). Cet argument permet alors de déconditionner le Théorème 3.2 de Smith et d'obtenir notre Corollaire 2.5.1.

B.5 Formules explicites de la CTE bivariée pour quelques familles de copules

Les résultats présentés dans cette section sont obtenus avec l'aide du logiciel Maple 15. Nous rappelons ici l'expression de notre *Conditional-Tail-Expectation*, dans un cadre bivarié :

$$\mathrm{CTE}_\alpha(X,Y) = \begin{pmatrix} \mathbb{E}[\,X \,|\, (X,Y) \in L(\alpha)\,] \\ \mathbb{E}[\,Y \,|\, (X,Y) \in L(\alpha)\,] \end{pmatrix},$$

avec $\alpha \in (0,1)$ et $L(\alpha) = \{(x,y) \in \mathbb{R}_+^2 \ : \ F(x,y) \geq \alpha\}$. Nous notons

$$\mathrm{CTE}_\alpha^1(X,Y) = \mathbb{E}[\,X \,|\, (X,Y) \in L(\alpha)\,] \ \text{ et } \ \mathrm{CTE}_\alpha^2(X,Y) = \mathbb{E}[\,Y \,|\, (X,Y) \in L(\alpha)\,].$$

Dans la Section 4.5.2 nous avons déjà utilisé des formules explicites de la $\mathrm{CTE}_\alpha(\mathrm{X},\mathrm{Y})$ pour évaluer la qualité de notre estimateur $\widehat{\mathrm{CTE}_\alpha}(X,Y)$. Ci-dessous nous présentons des formules explicites dans des autres cas (différents par rapport à la Section 4.5.2) pour illustrer comme notre mesure est calculable pour différentes fonctions de répartition.

Copule de Frank avec marges uniformes

Nous considérons la copule :

$$C_\theta(x,y) = -\frac{1}{\theta}\ln\left(1 + \frac{(\mathrm{e}^{-x\,\theta}-1)\,(\mathrm{e}^{-y\,\theta}-1)}{\mathrm{e}^{-\theta}-1}\right), \ \text{ pour } \theta \in (-\infty,+\infty] \setminus \{0\},$$

et $(x,y) \in [0,1]^2$. Pour $\theta = 0$ on obtient la copule indépendante $C_0(x,y) = x\,y$; $C_{-\infty}(x,y) = W$ et $C_{+\infty}(x,y) = M$.
Dans le cas $\theta > 0$, nous donnons donc l'expression exacte de notre CTE :

$$\mathrm{CTE}_\alpha^1(X,Y) = \mathrm{CTE}_\alpha^2(X,Y) = t(\alpha,\theta),$$

$$\text{avec } t(\alpha,\theta) = \frac{\theta\left(-\frac{1}{2}\,\alpha^2\,\mathrm{e}^{\alpha\,\theta} - \frac{\mathrm{e}^{\alpha\,\theta}}{\theta^2}\,\mathrm{dilog}(\mathrm{e}^{\alpha\,\theta}) + \frac{1}{\theta^2}\,\mathrm{dilog}(\mathrm{e}^{\alpha\,\theta}) + \frac{1}{2}\,\mathrm{e}^{\alpha\,\theta} + \frac{\mathrm{e}^{\alpha\,\theta}}{\theta^2}\,\mathrm{dilog}(\mathrm{e}^{\theta}) - \frac{1}{\theta^2}\,\mathrm{dilog}(\mathrm{e}^{\theta})\right)}{\theta - \alpha\,\theta - \ln\left(\frac{\mathrm{e}^{-\alpha\,\theta}-1}{\mathrm{e}^{-\theta}-1}\right) + \ln\left(\frac{\mathrm{e}^{-\alpha\,\theta}-1}{\mathrm{e}^{-\theta}-1}\right)\mathrm{e}^{\alpha\,\theta}},$$

et $\mathrm{dilog}(x) = \int_1^x \ln(t)/(1-t)\mathrm{d}t$.

Copule Indépendante avec marges uniformes

Nous considérons la copule :

$$C(x,y) = x\,y, \qquad \text{pour } \ (x,y) \in [0,1]^2.$$

Dans ce cas on obtient :

$$\mathrm{CTE}_\alpha^1(X,Y) = \mathrm{CTE}_\alpha^2(X,Y) = \frac{-\alpha + \frac{1}{2} + \frac{1}{2}\,\alpha^2}{-\alpha + \alpha\,\ln(\alpha) + 1}.$$

Copule Indépendante avec marges exponentielles

Nous considérons une copule indépendante avec marges $X \sim \mathrm{Exp}(2)$, $Y \sim \mathrm{Exp}(2)$. Dans ce cas on obtient :

$$\mathrm{CTE}^1_\alpha(X,Y) = \mathrm{CTE}^2_\alpha(X,Y) = \frac{\frac{1}{2}\,\ln(-\alpha+1)(\alpha-1)-\frac{1}{2}(\alpha-1)+1/2\,\alpha\,\mathrm{dilog}(-\alpha+1)-\frac{1}{12}\,\alpha\,\pi^2}{\alpha\,\ln(\alpha)-\alpha+1},$$

avec $\mathrm{dilog}(x) = \int_1^x \ln(t)/(1-t)\mathrm{d}t$.

Copule Indépendante avec marges Burr

Nous considérons une copule indépendante avec marges $X \sim \mathrm{Burr}(2,1)$, $Y \sim \mathrm{Burr}(2,1)$. Dans ce cas on obtient :

$$\mathrm{CTE}^1_\alpha(X,Y) = \mathrm{CTE}^2_\alpha(X,Y) = \frac{\arctan\left(\frac{\sqrt{-(\alpha-1)\,\alpha}}{\alpha-1}\right)-2\,\arctan\left(\frac{\sqrt{-(\alpha-1)\,\alpha}}{\alpha-1}\right)\alpha+\sqrt{-(\alpha-1)\alpha}-\pi\,\alpha+\frac{1}{2}\,\pi}{\alpha\,\ln(\alpha)-\alpha+1}.$$

Copule de Gumbel-Barnett avec marges uniformes

Nous considérons la copule :

$$C(x,y) = x\,y\,\mathrm{e}^{-\theta\ln(x)\,\ln(y)}, \qquad \text{pour } \theta \in (0,1] \quad \text{et} \quad (x,y) \in [0,1]^2,$$

voir copule (4.2.9) dans le Tableau 4.1 de Nelsen (1999). Pour $\theta = 0$ on obtient $C_0(x,y) = x\,y$. Nous donnons donc l'expression exacte de notre CTE dans ce cas :

$$\mathrm{CTE}^1_\alpha(X,Y) = \mathrm{CTE}^2_\alpha(X,Y) = g(\alpha,\theta),$$

avec $g(\alpha,\theta) = \Bigg(-\alpha\,\mathrm{e}^{\frac{1}{\theta}}\mathrm{Ei}\left(1,-\ln(\alpha)+\frac{1}{\theta}\right)\ln(\alpha) + \frac{\alpha\,\mathrm{e}^{\frac{1}{\theta}}\mathrm{Ei}\left(1,-\ln(\alpha)+\frac{1}{\theta}\right)}{\theta} - 1/2\,\alpha^2$

$$+\alpha\,\mathrm{e}^{\frac{1}{\theta}}\mathrm{Ei}\left(1,\frac{1}{\theta}\right)\ln(\alpha)-\frac{\alpha\,\mathrm{e}^{\frac{1}{\theta}}\mathrm{Ei}\left(1,\frac{1}{\theta}\right)}{\theta}+1/2\Bigg)\left(\alpha\,\ln(\alpha)\ln(1-\theta\,\ln(\alpha))+1-\alpha-\frac{\alpha\,\ln(1-\theta\,\ln(\alpha))}{\theta}\right)^{-1},$$

et $\mathrm{Ei}(a,z) = z^{a-1}\,\Gamma(1-a,z)$ (Γ est la fonction Gamma).

Copule de Clayton avec marges exponentielle et Burr

Nous considérons une fonction de répartition avec copule de Clayton de paramètre 1 et marges $X \sim \mathrm{Exp}(1)$, $Y \sim \mathrm{Burr}(2,1)$:

$$F(x,y) = \frac{(-1+\mathrm{e}^{-x})\,y^2}{-y^2-1+\mathrm{e}^{-x}}, \qquad \text{pour } x,\ y \geq 0.$$

Nous donnons donc l'expression exacte de notre CTE dans ce cas, coordonnée par coordonnée :

$$\mathrm{CTE}^1_\alpha(X,Y) = \frac{\alpha^2\,\ln(\alpha)-\alpha^2\,\ln(-\alpha+1)-\ln(-\alpha+1)-\alpha+2\,\ln(-\alpha+1)\,\alpha+1}{\alpha^2-2\,\alpha+1},$$

$$\mathrm{CTE}^2_\alpha(X,Y) = \frac{1}{2}\,\frac{\left(4\,\alpha^3-6\,\alpha^2+2\,\alpha+2\,\arctan\left(\frac{\sqrt{(-\alpha+1)\alpha}}{\alpha-1}\right)\sqrt{(-\alpha+1)\alpha}+\pi\,\sqrt{(-\alpha+1)\alpha}\right)(\alpha^2-2\,\alpha+1)}{\sqrt{(-\alpha+1)\alpha}}.$$

Copule de Clayton avec marges exponentielles

Nous considérons une fonction de répartition avec copule de Clayton de paramètre 1 et marges $X \sim \mathrm{Exp}(2)$, $Y \sim \mathrm{Exp}(2)$:

$$F(x,y) = -\frac{(-1+\mathrm{e}^{-2\,x})\,(-1+\mathrm{e}^{-2\,y})}{-1+\mathrm{e}^{-2\,x-2\,y}}, \qquad \text{pour} x,\ y \geq 0.$$

On obtient :

$$\begin{aligned}\mathrm{CTE}^1_\alpha(X,Y) &= \mathrm{CTE}^2_\alpha(X,Y)\\ &= \frac{\frac{1}{2}+(1-\frac{1}{2})\,\ln(-\alpha+1)\,\alpha+\frac{1}{2}\,\alpha^2\,\ln(\sqrt{-\alpha+1}+1)-\frac{1}{2}\,\ln(-\alpha+1)\,\alpha^2+\frac{1}{2}\,\alpha^2\,\ln(1-\sqrt{-\alpha+1})-\frac{1}{2}\,\alpha}{1-2\,\alpha+\alpha^2}.\end{aligned}$$

Illustration invariance pour translation et homogénéité de la CTE

Nous proposons une illustration de la Proposition 3.3.2. Nous considérons une copule de Clayton de paramètre 1 avec marges $X \sim \mathrm{Exp}(1)$, $Y \sim \mathrm{Exp}(2)$. On obtient :

$$\mathrm{CTE}_\alpha\left(\begin{pmatrix}4\\12\end{pmatrix}\begin{pmatrix}X\\Y\end{pmatrix}\right) = \begin{cases}\frac{4-4\alpha+8\ln(-\alpha+1)\alpha-4\ln(-\alpha+1)\alpha^2-4\ln(-\alpha+1)+4\alpha^2\ln(\alpha)}{1-2\alpha+\alpha^2}\\ \frac{-6\ln(-\alpha+1)\alpha^2+6+12\ln(-\alpha+1)\alpha-6\alpha+6\alpha^2\ln(\alpha)-6\ln(-\alpha+1)}{1-2\alpha+\alpha^2}\end{cases}$$

$$= \begin{cases}4\ \frac{\alpha^2\ln(\alpha)-\ln(-\alpha+1)\alpha^2-\ln(-\alpha+1)-\alpha+2\ln(-\alpha+1)\alpha+1}{1-2\alpha+\alpha^2}\\ 12\ \frac{-\frac{1}{2}\ln(-\alpha+1)\alpha^2+\frac{1}{2}+\ln(-\alpha+1)\alpha-\frac{1}{2}\alpha+\frac{1}{2}\alpha^2\ln(\alpha)-\frac{1}{2}\ln(-\alpha+1)}{1-2\alpha+\alpha^2}\end{cases} = \begin{pmatrix}4\\12\end{pmatrix}\mathrm{CTE}_\alpha(\mathrm{X},\mathrm{Y}).$$

B.6 Quelques comparaisons de la CTE bivariée avec différentes mesures de risque bivariées

Récemment différentes généralisations de la *Conditional-Tail-Expectation* ont été proposées dans la littérature (voir la Section 1.2.2 dans l'Introduction). Pour de plus amples détails, nous renvoyons le lecteur intéressé à Cai et Li, (2005), Bargès *et al.* (2009), Landsman et Valdez (2003), Hürlimann (2004). Nous rappelons par exemple, dans un cadre bivarié :

- $\mathrm{CTE}_\alpha(X,Y)^{\mathrm{sum}} = \mathbb{E}[\,(X,Y)\,|\,X+Y > Q_{X+Y}(\alpha)\,]$,
- $\mathrm{CTE}_\alpha(X,Y)^{\min} = \mathbb{E}[\,(X,Y)\,|\,\min\{X,Y\} > Q_{\min\{X,Y\}}(\alpha)\,]$,
- $\mathrm{CTE}_\alpha(X,Y)^{\max} = \mathbb{E}[\,(X,Y)\,|\,\max\{X,Y\} > Q_{\max\{X,Y\}}(\alpha)\,]$.

Nous présentons ici quelques comparaisons entre ces différentes mesures dans des cas simples. Une étude plus détaillée est laissée pour des travaux futurs. Soit (X,Y) un couple aléatoire avec marginales exponentielles de paramètre 2 et indépendantes . Tout d'abord nous proposons une illustration du comportement des ensembles : $\{(X,Y) \in L(\alpha)\}$, $\{\min\{X,Y\} > Q_{\min\{X,Y\}}(\alpha)\}$, $\{X > Q_X(\alpha)\}$ et $\{X+Y > Q_{X+Y}(\alpha)\}$, pour $\alpha = 0.9$ (voir Figure B.8).

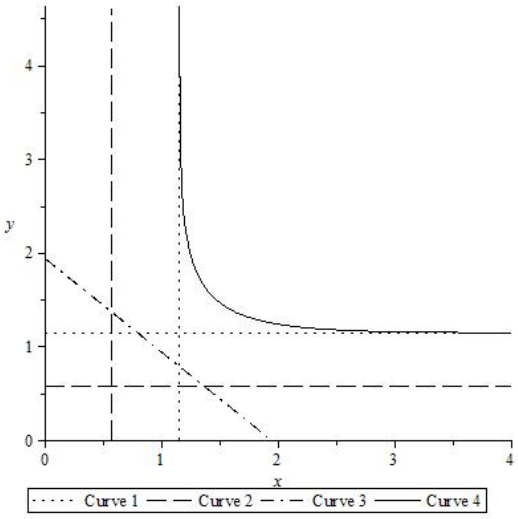

FIGURE B.8 – Courbe 1 : $x = Q_X(0.9)$ et $y = Q_X(0.9)$; Courbe 2 : $x = Q_{\min\{X,Y\}}(0.9)$ et d $y = Q_{\min\{X,Y\}}(0.9)$; Courbe 3 : $y = Q_{X+Y}(0.9) - x$; Courbe 4 : $\partial L(0.9)$. (X,Y) avec marginales exponentielles de paramètre 2 et indépendantes.

Dans ce cas notre mesure $\mathrm{CTE}_\alpha(X,Y)$ est donnée par :

$$\mathrm{CTE}^1_\alpha(X,Y)\frac{\frac{1}{2}\,\ln(-\alpha+1)(\alpha-1)-\frac{1}{2}\,(\alpha-1)+\frac{1}{2}\,\alpha\,\mathrm{dilog}(-\alpha+1)-\frac{1}{12}\,\alpha\,\pi^2}{\alpha\,\ln(\alpha)-\alpha+1},$$

avec $\mathrm{dilog}(x) = \int_1^x \ln(t)/(1-t)\mathrm{d}t$. En utilisant les mesures ci-dessus, on obtient pour $\alpha \in (0,1)$,

$$\mathrm{CTE}_\alpha(X) = \mathbb{E}[\,X \,|\, X > Q_X(\alpha)\,] = \tfrac{1}{2} - \tfrac{1}{2}\,\ln(-\alpha+1)\,,$$

$$\mathrm{CTE}^1_\alpha(X,Y)^{\mathrm{sum}} = \mathbb{E}[\,X \,|\, X+Y > Q_{X+Y}(\alpha)\,] = -\tfrac{1}{4}\,\frac{1+\left(\mathrm{LambertW}\left(-1,\mathrm{e}^{-1}(\alpha-1)\right)\right)^2}{\mathrm{LambertW}(-1,\mathrm{e}^{-1}(\alpha-1))}\,,$$

$$\mathrm{CTE}^1_\alpha(X,Y)^{\min} = \mathbb{E}[\,X \,|\, \min\{X,Y\} > Q_{\min\{X,Y\}}(\alpha)\,] = \tfrac{1}{2} - \tfrac{1}{4}\,\ln(-\alpha+1)\,,$$

avec LambertW la fonction de Lambert W telle que $\mathrm{LambertW}(x)\,\mathrm{e}^{\mathrm{LambertW}(x)} = x$, (voir Figure B.9 (gauche) pour une illustration graphique).
Nous remarquons que notre région de risque $L(\alpha)$ est un *orthant* supérieur de $\mathbb{R}^2_+$; elle est incluse dans les autres régions considérées (voir Figure B.8). Cela signifie que dans $L(\alpha)$ nous considérons des dommages simultanés grands de (X,Y) et que la sévérité de ces pertes est plus grande que dans les autres régions. Dans la Figure B.9 (droite) nous représentons $\mathrm{CTE}_\alpha(X,Y)$, $\mathrm{CTE}_\alpha(X,Y)^{\mathrm{sum}}$, $\mathrm{CTE}_\alpha(X,Y)^{\min}$ et $(\mathrm{CTE}_\alpha(X), \mathrm{CTE}_\alpha(X))$ pour $\alpha = 0.9$. Pour différents niveau de α nous observons le même genre de comportement.

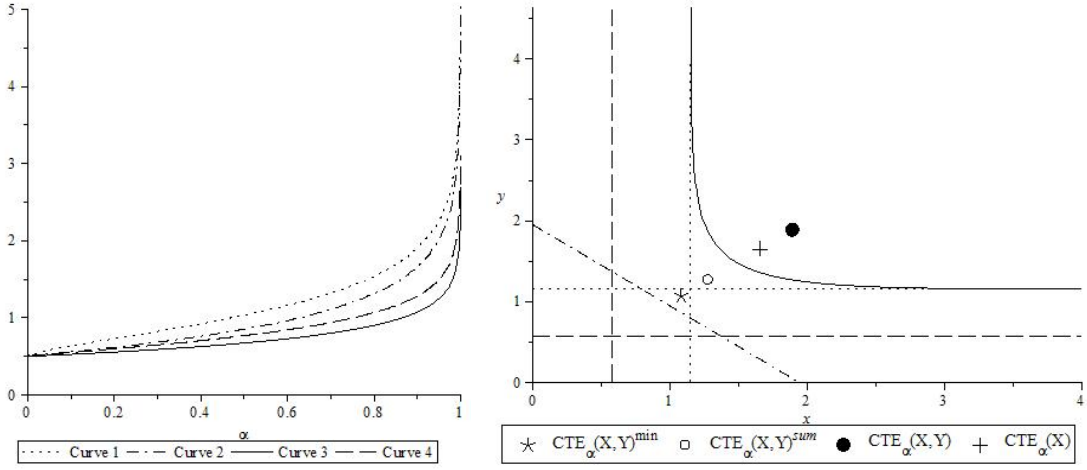

FIGURE B.9 – (X,Y) avec marginales exponentielles de paramètre 2 et indépendantes. (gauche) Courbe 1 : $\mathrm{CTE}^1_\alpha(X,Y) = \mathbb{E}[\,X \,|\, (X,Y) \in L(\alpha)\,]$; Courbe 2 : $\mathbb{E}[\,X \,|\, X > Q_X(\alpha)\,]$; Courbe 3 : $\mathbb{E}[\,X \,|\, \min\{X,Y\} > Q_{\min\{X,Y\}}(\alpha)\,]$; Courbe 4 : $\mathbb{E}[\,X \,|\, X+Y > Q_{X+Y}(\alpha)\,]$; (droite)) différentes mesures pour $\alpha = 0.9$.

Dans le cas comonotone $\mathbb{P}[(X,Y) \in L(\alpha)] = 1-\alpha$, pour $\alpha \in (0,1)$, nous avons :

Proposition B.6.1 *Soit (X,Y) un couple aléatoire comonotone avec $Y = g(X)$ et g une fonction strictement croissante sur $\mathbb{R}_+$. Alors, $\forall\ \alpha \in (0,1)$, on a*

$$\mathrm{CTE}_\alpha(X,Y) = \mathrm{CTE}_\alpha(X,Y)^{\mathrm{sum}} = \mathrm{CTE}_\alpha(X,Y)^{\min} = \mathrm{CTE}_\alpha(X,Y)^{\max}.$$

La preuve découle de la propriété d'additivité des quantiles dans le cas d'un couple comonotone : $Q_{X+Y}(\alpha) = Q_X(\alpha) + Q_Y(\alpha)$, for $\alpha \in (0,1)$ (voir Kaas *et al.*, 2002 ; Théorème 7).

B.7 Sur la croissance de la CTE bivariée : deux cas discrets

Nous présentons deux exemples bivariés discrets où notre mesure, $\mathrm{CTE}_\alpha(\mathrm{X},\mathrm{Y})$, n'est pas toujours croissante par rapport au niveau de risque α.

Exemple B.7.1 On considère la distribution discrète bivariée :

		X	
	$\mathbb{P}_{X,Y}$	0	1
	0	2/8	2/8
Y	1	3/8	1/8

Soit $S = \{(0,0),(0,1),(1,0),(1,1)\}$ l'ensemble des valeurs possibles pour (X,Y) et $L(\alpha) = \{(x,y) \in S \mid \mathbb{P}(X \leq x, Y \leq y) \geq \alpha\}$. Alors on a

α	$L(\alpha)$	$\mathrm{CTE}_\alpha(\mathrm{X},\mathrm{Y})$	$\mathrm{CTE}_\alpha(\mathrm{X})$	$\mathrm{CTE}_\alpha(\mathrm{Y})$
$0 \leq \alpha \leq 1/4$	S	$(3/8, 1/2)$	$3/8$	$1/2$
$1/4 < \alpha \leq 1/2$	$\{(0,1),(1,0),(1,1)\}$	$(3/5, 4/5)$	$3/8$	$1/2$
$1/2 < \alpha \leq 5/8$	$\{(0,1),(1,1)\}$	$(1/4, 1)$	$3/8$	1
$5/8 < \alpha \leq 1$	$\{(1,1)\}$	$(1,1)$	1	1

Si α_1 et α_2 sont choisi tels que $1/4 < \alpha_1 \leq 1/2$ et $1/2 < \alpha_2 \leq 5/8$, alors $\alpha_1 < \alpha_2$ et $\mathrm{CTE}^1_{\alpha_1} = 3/5 > 1/4 = \mathrm{CTE}^1_{\alpha_2}$. De plus, le dernier exemple montre aussi que dans ce cas discret la $\mathrm{CTE}_\alpha(\mathrm{X},\mathrm{Y})$ n'est pas toujours plus grande que le vecteur composé par les CTE univariées $(\mathrm{CTE}_\alpha(\mathrm{X}), \mathrm{CTE}_\alpha(\mathrm{Y}))$ car pour $\alpha = 5/8$, $\mathrm{CTE}^1_\alpha(\mathrm{X},\mathrm{Y}) < \mathrm{CTE}_\alpha(\mathrm{X})$.

Exemple B.7.2 On considère la distribution discrète bivariée :

			X	
	$\mathbb{P}_{X,Y}$	5	10	15
	1	0	0	2/9
Y	2	1/6	2/9	1/6
	3	1/9	1/9	0

Soit $\alpha_1 = 5/18 \simeq 0.277$. Alors $\mathrm{CTE}_{\alpha_1}(\mathrm{X},\mathrm{Y}) = (115/11,\, 26/11) \simeq (10.45,\, 2.36)$. Soit $\alpha_2 = 7/18 \simeq 0.388$, avec $\alpha_2 > \alpha_1$. Alors on a $\mathrm{CTE}_{\alpha_2}(\mathrm{X},\mathrm{Y}) = (35/3,\, 20/9) \simeq (11.67,\, 2.22)$. Nous pouvons conclure comme dans l'Exemple B.7.1.

B.8 Quelques formules paramétriques exactes pour une courbe quantile

Dans la littérature la frontière des ensemble des niveau $\partial L(\gamma)$, pour $\gamma \in (0,1)$ est appelle courbe quantile. En particulier elle a été étudiée autant que généralisation multidimensionnelle de la *Value-at-Risk* par Embrechts et Puccetti (2006), Nappo et Spizzichino (2009) (voir Section 1.2.2). Belzunce et ses coauteurs proposent une formulation paramétrique pour cette courbe (voir Belzunce *et al.*, 2007). Dans cette section nous nous intéressons au calcul exact de cette expression paramétrique dans des cas bivariés spécifiques.

Nous avons deux risques (par exemple deux pertes) différents X et Y avec une copule *Clayton Survie*. Avec ce type de copule nous avons une corrélation faible dans la queue inférieure, mais forte dans la queue supérieure, (c'est-à-dire que les grandes pertes sont fortement corrélées). Soient $F_X \neq F_Y$, deux distributions Burr avec paramètres différents. Plus précisément on considère :

$$F_X(x) = 1 - (1+x^2)^{-1}, \quad F_Y(y) = 1 - (1+y)^{-1},$$
$$F_{X,Y}(x,y) = 1 - (1+x^2)^{-1} - (1+y)^{-1}. + (1+x^2+y)^{-1}$$

En suivant Belzunce *et al.* (2007), la courbe quantile peut être paramétrée de la façon suivante :

$$\partial L(\gamma) = \{(x,y) \in \mathbb{R}^2 : F(x,y) = \gamma\} = \left\{ \left(Q_X(u), Q_{Y|X \leq Q_X(u)}\left(\frac{\gamma}{u}\right)\right) : u > \gamma \right\}$$
$$= \left\{ \sqrt{\tfrac{u}{1-u}};\ \tfrac{1}{2}\ \frac{\sqrt{2}\sqrt{\left(\gamma - \frac{\gamma}{u} + 1 - u\right)\left(-\gamma + 2\frac{\gamma}{u} + u - 2 + \sqrt{\gamma^2 - 2\gamma u + 4\gamma - 4\frac{\gamma}{u} + u^2 - 4u + 4}\right)}}{\gamma - \frac{\gamma}{u} + 1 - u} : u > \gamma \right\},$$

avec Q_X la fonction quantile de F_X et $Q_{Y|X \leq Q_X(u)}$ la fonction quantile de $F_{Y|X \leq Q_X(u)}$. Dans la Figure B.10 on représente $\partial L(\gamma)$ pour $\gamma = 0.8$ et $\gamma = 0.99$.

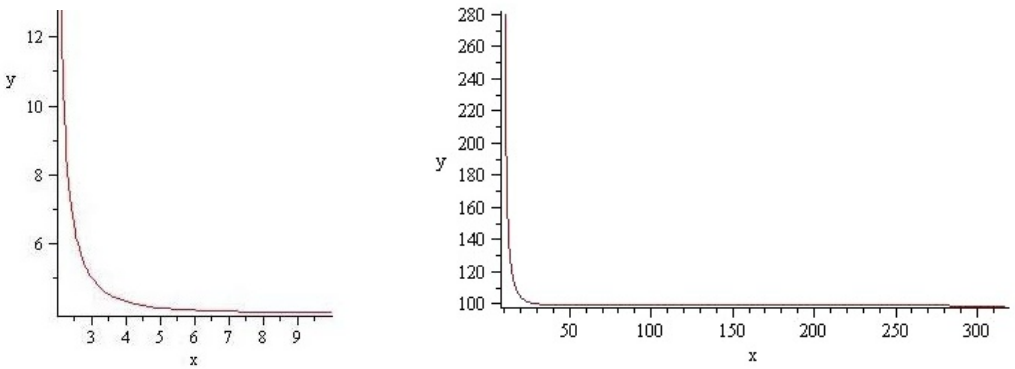

FIGURE B.10 – $\partial L(\gamma)$ pour $\gamma = 0.8$ (gauche) et $\gamma = 0.99$ (droite) ; cas $F_X \neq F_Y$.

Nous notons que les quantiles univariés $Q_X(\gamma)$ et $Q_Y(\gamma)$ des marginales sont les asymptotes de la courbe de niveau. Pour $\gamma = 0.8$ nous obtenons $(Q_X(0.8), Q_Y(0.8)) = (2,4)$

et pour $\gamma = 0.99 \quad (Q_X(0.99), Q_Y(0.99)) = (9.95, 99)$.

Maintenant nous considérons le cas $F_X = F_Y$:

$$F_X(x) = 1 - (1+x)^{-1}, \quad F_Y(y) = 1 - (1+y)^{-1},$$
$$F_{X,Y}(x,y) = 1 - (1+x)^{-1} - (1+y)^{-1} + [1+x+y]^{-1}.$$

Dans ce cas on obtient :

$$\partial L(\gamma) = \left\{ \frac{u}{1-u}; -\frac{1}{2} \frac{\left(\gamma - u - 2\,\frac{\gamma}{u} + 2 - \sqrt{\gamma^2 - 2\,\gamma\,u + 4\,\gamma + u^2 - 4\,u - 4\,\frac{\gamma}{u} + 4}\right)}{\left(\gamma - \frac{\gamma}{u} - u + 1\right)} : \ u > \gamma \right\}.$$

Dans la Figure (B.11) on représente $\partial L(\gamma)$ pour $\gamma = 0.8$ et $\gamma = 0.95$.

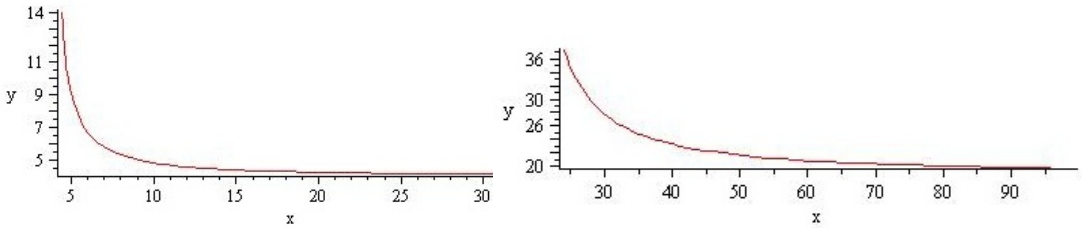

FIGURE B.11 – $\partial L(\gamma)$ pour $\gamma = 0.8$ (gauche) et $\gamma = 0.95$ (droite) ; cas $F_X = F_Y$.

Nous notons que les quantiles univariés $Q_X(\gamma)$ et $Q_Y(\gamma)$ des marginales sont les asymptotes de la courbe de niveau. Pour $\gamma = 0.8$ on a $(Q_X(0.8), Q_Y(0.8)) = (4, 4)$ et pour $\gamma = 0.95$ $(Q_X(0.95), Q_Y(0.95)) = (19, 19)$ (voir Figure B.11).

B.9 Quelques études de la CTE bivariée pour des hauts niveaux de risque

Dans cette section nous proposons quelques illustrations de la *Conditional-Tail-Expectation* bivariée, $\mathrm{CTE}_\alpha(\mathrm{X},\mathrm{Y})$, pour des hautes valeurs du niveau de risque α. Nous ne prétendons pas envisager cette problématique de manière rigoureuse. Nous voulons plutôt donner une illustration d'une approche possible. L'estimation est faite en utilisant une méthode plug-in (exactement comme dans le Chapitre 4) avec un estimateur pour la queue de la distribution.

En particulier dans les trois exemples qui suivent l'estimateur choisi est notre estimateur $\widehat{F}^*$ présenté dans le Chapitre 2 (voir (2.31)). Plus précisément nous définissons

$$\widehat{L}_n(c) := \{(x,y) \in \mathbb{R}_+^2 : \ \widehat{F}^*(x,y) \geq c\}, \qquad \text{for } c \in (0,1),$$

et donc

$$\widehat{\mathrm{CTE}}_{\alpha,\widehat{F}^*}(X,Y) = \begin{pmatrix} \dfrac{\sum_{i=1}^n X_i 1_{\{(X_i,Y_i)\in\widehat{L}_n(\alpha)\}}}{\sum_{i=1}^n 1_{\{(X_i,Y_i)\in\widehat{L}_n(\alpha)\}}} \\ \dfrac{\sum_{i=1}^n Y_i 1_{\{(X_i,Y_i)\in\widehat{L}_n(\alpha)\}}}{\sum_{i=1}^n 1_{\{(X_i,Y_i)\in\widehat{L}_n(\alpha)\}}} \end{pmatrix}. \tag{B.11}$$

Le comportement de $\widehat{\mathrm{CTE}}_{\alpha,\widehat{F}^*}(X,Y)$ et la possibilité d'étendre les résultats du Chapitre 4 au cas de méthode plug-in avec $\widehat{F}^*(x,y)$ sont des sujets de recherche non traités dans cette thèse et laissés pour des travaux futurs.

Dans la Figure B.12 nous représentons le $\widehat{\mathrm{CTE}}_{\alpha,\widehat{F}^*}(X,Y)$ pour une copule Clayton Survie avec paramètre 1 et marges Burr$(2,1)$.

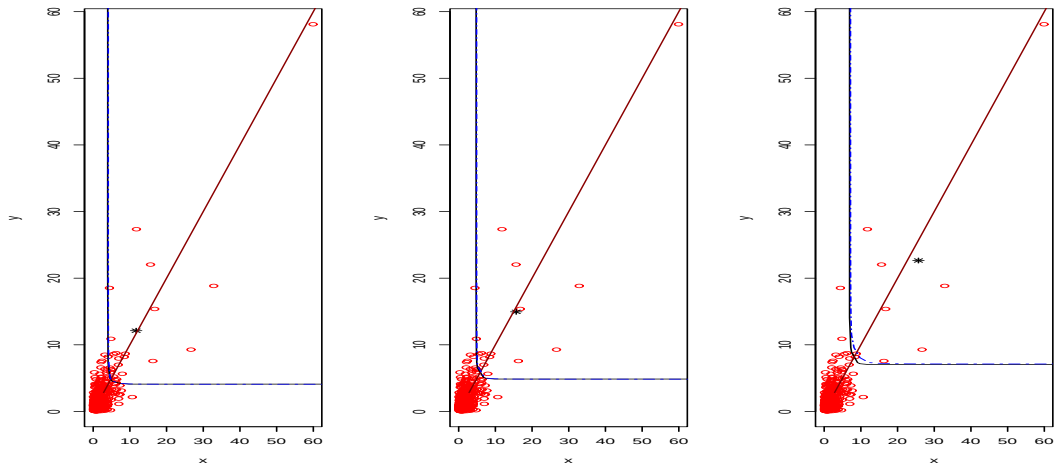

FIGURE B.12 – Copule Clayton Survie avec paramètre 1 et marges Burr$(2,1)$. Etoile : $\widehat{\mathrm{CTE}}_{\alpha,\widehat{F}^*}(X,Y)$; courbe bleu : $F(x,y) = \alpha$; courbe noire : $\widehat{F}^*(x,y) = \alpha$; avec $\alpha = 0.93$, 0.955, 0.98 (de gauche vers droite) ; droite $y = x$.

Nous remarquons que dans ce cas $\mathrm{CTE}_\alpha^1(\mathrm{X},\mathrm{Y}) = \mathrm{CTE}_\alpha^2(\mathrm{X},\mathrm{Y})$. Ainsi nous nous attendons à ce que la CTE estimée soit également sur la bissectrice du quadrant positif.

Dans les Figures B.13-B.14 nous représentons la $\widehat{\mathrm{CTE}}_{\alpha,\widehat{F}^*}(X,Y)$ dans deux cas réels : les données Loss-ALAE et les données Wave height-Water level data (Hauteurs de vague-Niveau de l'eau). Ces données ont été analysées dans la Section 2.7 (voir Figures 2.6-2.9).

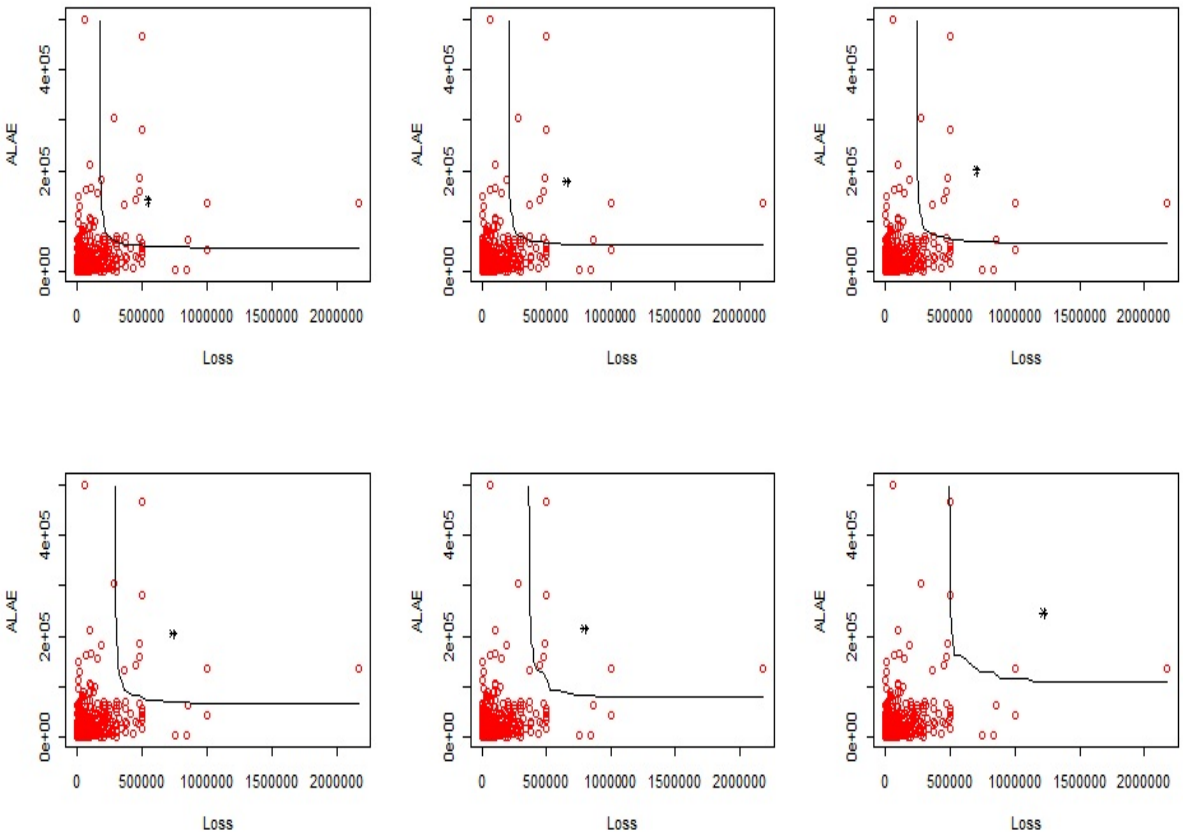

FIGURE B.13 – Données Loss-ALAE, $n = 1500$. Etoile : $\widehat{\mathrm{CTE}}_{\alpha,\widehat{F}^*}(X,Y)$; courbe noire : $\widehat{F}^*(x,y) = \alpha$; avec $\alpha = 0.952$, 0.959, 0.967, 0.975, 0.985, 0.991 (de gauche vers droite, du haut vers le bas).

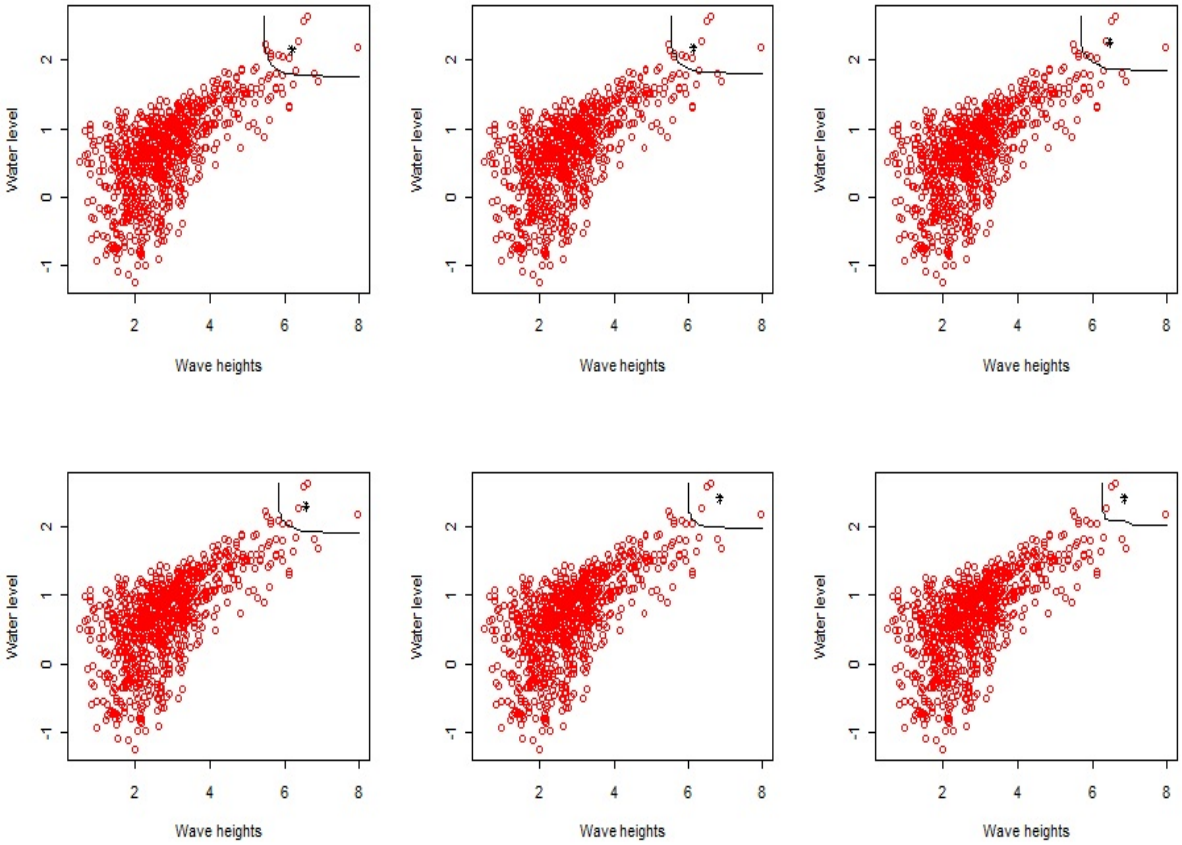

FIGURE B.14 – Données Wave height-Water level data, $n = 1500$. Etoile : $\widehat{\mathrm{CTE}}_{\alpha,\widehat{F}^*}(X,Y)$; courbe noire : $\widehat{F}^*(x,y) = \alpha$, avec $\alpha = 0.971, 0.974, 0.978, 0.983, 0.987, 0.992$ (de gauche vers droite, du haut vers le bas).

Bibliographie

Abdous, B. et Theodorescu, R. (1992). Note on the spatial quantile of a random vector. *Statist. Probab. Lett.*, 13(4):333–336.

Aczél, J. (1966). *Lectures on functional equations and their applications.* Mathematics in Science and Engineering, Vol. 19. Academic Press, New York. Translated by Scripta Technica, Inc. Supplemented by the author. Edited by Hansjorg Oser.

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

Baíllo, A. (2003). Total error in a plug-in estimator of level sets. *Statist. Probab. Lett.*, 65(4):411–417.

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

Baíllo, A. et Cuevas, A. (2001). On the estimation of a star-shaped set. *Adv. in Appl. Probab.*, 33(4):717–726.

Baíllo, A., Cuevas, A. et Justel, A. (2000). Set estimation and nonparametric detection. *Canad. J. Statist.*, 28(4):765–782.

Balkema, A. A. et de Haan, L. (1974). Residual life time at great age. *Ann. Probability*, 2:792–804.

Barbe, P., Genest, C., Ghoudi, K. et Rémillard, B. (1996). On Kendall's process. *J. Multivariate Anal.*, 58(2):197–229.

Bargès, M., Cossette, H. et Marceau, E. (2009). TVaR-based capital allocation with copulas. *Insurance Math. Econom.*, 45(3):348–361.

Barnett, V. (1976). The ordering of multivariate data. *J. Roy. Statist. Soc. Ser. A*, 139(3):318–355.

Beirlant, J., Dierckx, G. et Guillou, A. (2011). Reduced bias estimators for bivariate tail modelling. *Insurance : Mathematics and Economics*, In Press, Accepted Manuscript.

Beirlant, J., Goegebeur, Y., Teugels, J. et Segers, J. (2004). *Statistics of extremes*. Wiley Series in Probability and Statistics. John Wiley & Sons Ltd., Chichester. Theory and applications, With contributions from Daniel De Waal and Chris Ferro.

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

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

Benth, F. E. et Kettler, P. (2011). Dynamic copula models for the spark spread. *Quantitative Finance*, 11(3):407–421.

Biau, G., Cadre, B. et Pelletier, B. (2007). A graph-based estimator of the number of clusters. *ESAIM. Probability and Statistics*, 11:272–280.

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

Borkar, V. S. (1995). *Probability theory*. Universitext. Springer-Verlag, New York. An advanced course.

Bühlmann, H., Gagliardi, B., Gerber, H. et Straub, E. (1977). Some inequalities for stop-loss premiums. *Astin Bull.*, 9(2):75–83.

Cadre, B. (2006). Kernel estimation of density level sets. *J. Multivariate Anal.*, 97(4):999–1023.

Cai, J., Einmahl, J. et de Haan, L. (2011). Estimation of extreme risk regions under multivariate regular variation. Open access publications from tilburg university, Tilburg University.

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

Caillault, C. et Guégan, D. (2005). Empirical estimation of tail dependence using copulas : application to Asian markets. *Quant. Finance*, 5(5):489–501.

Capéraà, P. et Fougères, A.-L. (2000). Estimation of a bivariate extreme value distribution. *Extremes*, 3(4):311–329 (2001).

Capéraà, P., Fougères, A.-L. et Genest, C. (1997). A stochastic ordering based on a decomposition of Kendall's tau. *In Distributions with given marginals and moment problems (Prague, 1996)*, pages 81–86. Kluwer Acad. Publ., Dordrecht.

Cavalier, L. (1997). Nonparametric estimation of regression level sets. *Statistics. A Journal of Theoretical and Applied Statistics*, 29(2):131–160.

Chakak, A. et Ezzerg, M. (2000). Bivariate contours of copula. *Comm. Statist. Simulation Comput.*, 29(1):175–185.

Chaouch, M., Gannoun, A. et Saracco, J. (2009). Estimation de quantiles géométriques conditionnels et non conditionnels. *J. SFdS*, 150(2):1–27.

Charpentier, A. et Juri, A. (2006). Limiting dependence structures for tail events, with applications to credit derivatives. *J. Appl. Probab.*, 43(2):563–586.

Charpentier, A. et Segers, J. (2007). Lower tail dependence for Archimedean copulas : characterizations and pitfalls. *Insurance Math. Econom.*, 40(3):525–532.

Chaubey, Y. P. et Sen, P. K. (2002). Smooth estimation of multivariate survival and density functions. *J. Statist. Plann. Inference*, 103(1-2):361–376. C. R. Rao 80th birthday felicitation volume, Part I.

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

Cherubini, U., Luciano, E. et Vecchiato, W. (2004). *Copula Methods in Finance.* Wiley, New York.

Coles, S. (2001). *An introduction to statistical modeling of extreme values.* Springer Series in Statistics. Springer-Verlag London Ltd., London.

Coles, S., Heffernan, J. et Tawn, J. (1999). Dependence Measures for Extreme Value Analyses. *Extremes*, 2(4):339–365.

Coles, S. G. et Tawn, J. A. (1994). Statistical methods for multivariate extremes : an application to structural design. *Appl. Statist.*, 43:1–48. With discussion.

Cuevas, A. et Fraiman, R. (1997). A plug-in approach to support estimation. *The Annals of Statistics*, 25(6):2300–2312.

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

Cuevas, A. et Rodríguez-Casal, A. (2004). On boundary estimation. *Advances in Applied Probability*, 36(2):340–354.

Davison, A. C. (1984). *A statistical model for contamination due to long-range atmospheric transport of radionuclides.* Thèse de doctorat, Department of Mathematics, Imperial College of Science and Technology, London.

Davison, A. C. et Smith, R. L. (1990). Models for exceedances over high thresholds. *J. Roy. Statist. Soc. Ser. B*, 52(3):393–442. With discussion and a reply by the authors.

de Haan, L. (1970). *On regular variation and its application to the weak convergence of sample extremes*, volume 32 de *Mathematical Centre Tracts*. Mathematisch Centrum, Amsterdam.

de Haan, L. et de Ronde, J. (1998). Sea and wind : multivariate extremes at work. *Extremes. Statistical Theory and Applications in Science, Engineering and Economics*, 1(1):7–45.

de Haan, L. et Ferreira, A. (2006). *Extreme value theory*. Springer Series in Operations Research and Financial Engineering. Springer, New York. An introduction.

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

de Haan, L., Omey, E. et Resnick, S. (1984). Domains of attraction and regular variation in $\mathbf{R}^d$. *J. Multivariate Anal.*, 14(1):17–33.

de Haan, L. et Rootzén, H. (1993). On the estimation of high quantiles. *J. Statist. Plann. Inference*, 35(1):1–13.

Dedu, S. et 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.

Denuit, M. et Charpentier, A. (2004). *Mathématiques de l'assurance non-vie. Tome 1 : Principes fondamentaux de thorie du risque*. Economica.

Denuit, M., Dhaene, J., Goovaerts, M. et Kaas, R. (2005). *Actuarial Theory for Dependent Risks*. Wiley.

Denuit, M., Lefèvre, C. et Mesfioui, M. (1999). A class of bivariate stochastic orderings, with applications in actuarial sciences. *Insurance Math. Econom.*, 24(1-2):31–50. 1st IME Conference (Amsterdam, 1997).

Di Bernardino, E., Laloë, T., Maume-Deschamps, V. et Prieur, C. (2011). Plug-in estimation of level sets in a non-compact setting with applications in multivariate risk theory. *to appear in ESAIM Probability and Statistic*.

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

Drees, H., Ferreira, A. et de Haan, L. (2004). On maximum likelihood estimation of the extreme value index. *Ann. Appl. Probab.*, 14(3):1179–1201.

Drees, H. et Huang, X. (1998). Best attainable rates of convergence for estimators of the stable tail dependence function. *J. Multivariate Anal.*, 64(1):25–47.

Dudley, R. M. (1966). Weak convergences of probabilities on nonseparable metric spaces and empirical measures on Euclidean spaces. *Illinois J. Math.*, 10:109–126.

Einmahl, J. H. J. (1990). The empirical distribution function as a tail estimator. *Statist. Neerlandica*, 44(2):79–82.

Einmahl, J. H. J. et de Haan, L. (2009). Estimating extreme bivariate quantile regions. Discussion Paper 2009-29, Tilburg University, Center for Economic Research.

Einmahl, J. H. J., de Haan, L. et Li, D. (2006). Weighted approximations of tail copula processes with application to testing the bivariate extreme value condition. *Ann. Statist.*, 34(4):1987–2014.

Einmahl, J. H. J., Krajina, A. et Segers, J. (2008). A method of moments estimator of tail dependence. *Bernoulli*, 14(4):1003–1026.

Einmahl, J. H. J. et Mason, D. M. (1992). Generalized quantile processes. *Ann. Statist.*, 20(2):1062–1078.

El-Aroui, M.-A. et Diebolt, J. (2002). On the use of the peaks over thresholds method for estimating out-of-sample quantiles. *Comput. Statist. Data Anal.*, 39(4): 453–475.

Embrechts, P., Höing, A. et Juri, A. (2003). Using copulae to bound the value-at-risk for functions of dependent risks. *Finance Stoch.*, 7(2):145–167.

Embrechts, P., Klüppelberg, C. et Mikosch, T. (1997). *Modelling extremal events*, volume 33 de *Applications of Mathematics (New York)*. Springer-Verlag, Berlin. For insurance and finance.

Embrechts, P., McNeil, A. J. et Straumann, D. (2002). Correlation and dependence in risk management : properties and pitfalls. *In Risk management : value at risk and beyond (Cambridge, 1998)*, pages 176–223. Cambridge Univ. Press, Cambridge.

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

Falk, M. et Reiss, R. (2003a). Efficient estimation of the canonical dependence function. *Extremes*, 6(1):61–82.

Falk, M. et Reiss, R.-D. (2001). Estimation of canonical dependence parameters in a class of bivariate peaks-over-threshold models. *Statist. Probab. Lett.*, 52(3):233–242.

Falk, M. et Reiss, R. D. (2002). A characterization of the rate of convergence in bivariate extreme value models. *Statist. Probab. Lett.*, 59(4):341–351.

Falk, M. et Reiss, R.-D. (2003b). Efficient estimators and LAN in canonical bivariate POT models. *J. Multivariate Anal.*, 84(1):190–207.

FALK, M. et REISS, R.-D. (2005). On the distribution of Pickands coordinates in bivariate EV and GP models. *J. Multivariate Anal.*, 93(2):267–295.

FELLER, W. (1966). *An introduction to probability theory and its applications. Vol. II.* John Wiley & Sons Inc., New York.

FERNÁNDEZ-PONCE, J. M. et SUÁREZ-LLORÉNS, A. (2002). Central regions for bivariate distributions. *Austrian Journal of Statistic*, 31(2–3):141–156.

FOUGÈRES, A.-L. (2004). *Multivariate extremes, In Extreme Values in Finance, Telecommunications, and the Environment.* B. Finkenstädt and H. Rootzén (eds), Boca Raton, FL : Chapman & Hall/CRC.

FOUGÈRES, A.-L. et SOULIER, P. (2008). Modeling and estimation of conditional excesses. Rapport technique arXiv :0806.2426.

FREES, E. W. et VALDEZ, E. A. (1998). Understanding relationships using copulas. *North American Actuarial Journal*, 2(1):1–25.

GAUTHIER, C., LEHAR, A. et SOUISSI, M. (2010). Macroprudential regulation and systemic capital requirements. Working Papers 10-4, Bank of Canada.

GEFFROY, J. (1958). Contribution à la théorie des valeurs extrêmes. *Publ. Inst. Statist. Univ. Paris*, 7(3/4):37–121.

GENEST, C. et BOIES, J.-C. (2003). Detecting dependence with kendall plots. *The American Statistician*, 57:275–284.

GENEST, C., QUESSY, J.-F. et RÉMILLARD, B. (2006). Goodness-of-fit procedures for copula models based on the probability integral transformation. *Scandinavian Journal of Statistics*, 33:337–366.

GENEST, C. et RIVEST, L.-P. (2001). On the multivariate probability integral transformation. *Statist. Probab. Lett.*, 53(4):391–399.

GOLDIE, C. M. et SMITH, R. L. (1987). Slow variation with remainder : theory and applications. *Quart. J. Math. Oxford Ser. (2)*, 38(149):45–71.

GOOVAERTS, M. J., DE VYLDER, F. et HAEZENDONCK, J. (1984). *Insurance premiums.* North-Holland Publishing Co., Amsterdam. Theory and applications.

GRANDELL, J. (1991). *Aspects of risk theory.* Springer Series in Statistics : Probability and its Applications. Springer-Verlag, New York.

HARTIGAN, J. A. (1987). Estimation of a convex density contour in two dimensions. *J. Amer. Statist. Assoc.*, 82(397):267–270.

HEFFERNAN, J. E. (2000). A directory of coefficients of tail dependence. *Extremes*, 3:279–290.

Horowitz, J. L. (2001). *Handbook of Econometrics*, chapitre The Bootstrap, pages 3159–3228. J.J. Heckman and E.E. Leamer, Elsevier Science B.V.

Huang, X. (1992). *Statistics of bivariate extreme values*. Thèse de doctorat, Erasmus University Rotterdam, Tinbergen Institute.

Hürlann, W. (2004). Multivariate Fréchet copulas and conditional value-at-risk. *Int. J. Math. Math. Sci.*, (5-8):345–364.

Imlahi, L., Ezzerg, M. et 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.

Javid, A. (2009). Limiting tail dependence copulas. *Communications in Statistics-Theory and Methods*, 38(20):3772–3781.

Joe, H. (1997). *Multivariate models and dependence concepts*, volume 73 de *Monographs on Statistics and Applied Probability*. Chapman & Hall, London.

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

Juri, A. et Wüthrich, M. V. (2002). Copula convergence theorems for tail events. *Insurance Math. Econom.*, 30(3):405–420.

Juri, A. et Wüthrich, M. V. (2004). Tail dependence from a distributional point of view. *Extremes*, 6(3):213–246.

Kaas, R., Dhaene, J., Vyncke, D., Goovaerts, M. J. et Denuit, M. (2002). A simple geometric proof that comonotonic risks have the convex-largest sum. *Astin Bull.*, 32(1):71–80.

Kaufmann, E. et Reiss, R.-D. (1995). Approximation rates for multivariate exceedances. *J. Statist. Plann. Inference*, 45(1-2):235–245. Extreme value theory and applications (Villeneuve d'Ascq, 1992).

Kemperman, J. H. B. (1987). The median of a finite measure on a Banach space. *In Statistical data analysis based on the L_1-norm and related methods (Neuchâtel, 1987)*, pages 217–230. North-Holland, Amsterdam.

Koltchinskii, V. I. (1997). M-estimation, convexity and quantiles. *Ann. Statist.*, 25(2):435–477.

Laloë, T. (2009). *Sur Quelques Problèmes d'Apprentissage Supervisé et Non Supervisé*. Thèse de doctorat, University Montpellier II.

Landsman, Z. M. et Valdez, E. A. (2003). Tail conditional expectations for elliptical distributions. *N. Am. Actuar. J.*, 7(4):55–71.

Laurent, J.-P. (2003). Sensitivity analysis of risk measures for discrete distributions. *Working paper*.

LEADBETTER, M. R., LINDGREN, G. et ROOTZÉN, H. (1983). *Extremes and related properties of random sequences and processes.* Springer Series in Statistics. Springer-Verlag, New York.

LEDFORD, A. W. et TAWN, J. A. (1996). Statistics for near independence in multivariate extreme values. *Biometrika*, 83(1):169–187.

LEDFORD, A. W. et TAWN, J. A. (1997). Modelling dependence within joint tail regions. *J. Roy. Statist. Soc. Ser. B*, 59(2):475–499.

LEDFORD, A. W. et TAWN, J. A. (1998). Concomitant tail behaviour for extremes. *Adv. in Appl. Probab.*, 30(1):197–215.

LEHMANN, E. L. (1966). Some concepts of dependence. *Ann. Math. Statist.*, 37:1137–1153.

LESCOURRET, L. et ROBERT, C. Y. (2006). Extreme dependence of multivariate catastrophic losses. *Scandinavian Actuarial Journal*, (4):203–225.

MASSÉ, J.-C. et THEODORESCU, R. (1994). Halfplane trimming for bivariate distributions. *J. Multivariate Anal.*, 48(2):188–202.

MCNEIL, A. J. (1997). Estimating the tails of loss severity distributions using extreme value theory. *ASTIN Bulletin*, 27:1117–1137.

MCNEIL, A. J. (1999). Extreme value theory for risk managers. *preprint, ETH, Zurich.*

MCNEIL, A. J. et SALADIN, T. (1997). The peaks over thresholds method for estimating high quantiles of loss distributions. *ASTIN Colloquium.*

MICHEL, R. (2008). Some notes on multivariate generalized pareto distributions. *J. Multivariate Anal.*, 99(6):1288–1301.

MOLCHANOV, I. (1990). Empirical estimation of distribution quantiles of random closed sets. *Theory of Probability and Its Applications*, 35:594–600.

MÜLLER, A. (1997). Stop-loss order for portfolios of dependent risks. *Insurance Math. Econom.*, 21(3):219–223.

MÜLLER, A. et STOYAN, D. (2001). *Comparison methodes for stochastic models and risks.* Wiley.

MÜLLER, D. W. et SAWITZKI, G. (1991). Excess mass estimates and tests for multimodality. *J. Amer. Statist. Assoc.*, 86(415):738–746.

NAPPO, G. et SPIZZICHINO, F. (2009). Kendall distributions and level sets in bivariate exchangeable survival models. *Information Sciences*, 179:2878–2890.

NELSEN, R. B. (1999). *An introduction to copulas*, volume 139 de *Lecture Notes in Statistics.* Springer-Verlag, New York.

Nelsen, R. B., Quesada-Molinab, J. J., Rodríguez-Lallenac, J. A. et Úbeda-Floresc, M. (2003). Kendall distribution functions. *Statistics and Probability Letters*, 65:263–268.

Nowak, R. D. et Willett, R. M. (2007). Minimax optimal level-set estimation. *IEEE Transactions on Image Processing*, 16(12):2965–2979.

Ohlin, J. (1969). On a class of measures of dispersion with application to optimal reinsurance. *ASTIN Bulletin*, 5:249–266.

Peng, L. (1999). Estimation of the coefficient of tail dependence in bivariate extremes. *Statist. Probab. Lett.*, 43(4):399–409.

Pickands, III, J. (1975). Statistical inference using extreme order statistics. *Ann. Statist.*, 3:119–131.

Pickands, III, J. (1981). Multivariate extreme value distributions. *In Proceedings of the 43rd session of the International Statistical Institute, Vol. 2 (Buenos Aires, 1981)*, volume 49, pages 859–878, 894–902. With a discussion.

Polonik, W. (1995). Measuring mass concentrations and estimating density contour clusters—an excess mass approach. *The Annals of Statistics*, 23(3):855–881.

Polonik, W. (1997). Minimum volume sets and generalized quantile processes. *Stochastic Process. Appl.*, 69(1):1–24.

Ramos, A. et Ledford, A. (2009). A new class of models for bivariate joint tails. *J. R. Stat. Soc. Ser. B Stat. Methodol.*, 71(1):219–241.

Reiss, R.-D. et Thomas, M. (2007). *Statistical analysis of extreme values with applications to insurance, finance, hydrology and other fields.* Birkhäuser Verlag, Basel, third édition. With 1 CD-ROM (Windows).

Resnick, S. I. (1987). *Extreme values, regular variation, and point processes*, volume 4 de *Applied Probability. A Series of the Applied Probability Trust.* Springer-Verlag, New York.

Rigollet, P. et 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.

Rodríguez-Casal, A. (2003). *Estimacíon de conjuntos y sus fronteras. Un enfoque geometrico.* Thèse de doctorat, University of Santiago de Compostela.

Rootzén, H. et Tajvidi, N. (2006). Multivariate generalized Pareto distributions. *Bernoulli*, 12(5):917–930.

Rossi, C. (1973). Sulle curve di livello di una superficie di ripartizione in due variabili ; on level curves of two dimensional distribution function. *Giornale dell'Istituto Italiano degli Attuari*, 36:87–108.

Rossi, C. (1976). Proprietà geometriche delle superficie di ripartizione. *Rend. Mat. (6)*, 9(4):725–736 (1977).

Schlather, M. (2001). Examples for the coefficient of tail dependence and the domain of attraction of a bivariate extreme value distribution. *Statist. Probab. Lett.*, 53(3): 325–329.

Serfling, R. (2002). Quantile functions for multivariate analysis : approaches and applications. *Statist. Neerlandica*, 56(2):214–232. Special issue : Frontier research in theoretical statistics, 2000 (Eindhoven).

Sibuya, M. (1960). Bivariate extreme statistics. I. *Ann. Inst. Statist. Math. Tokyo*, 11:195–210.

Sklar, M. (1959). Fonctions de répartition à n dimensions et leurs marges. *Publ. Inst. Statist. Univ. Paris*, 8:229–231.

Smith, R. L. (1987). Estimating tails of probability distributions. *Ann. Statist.*, 15(3): 1174–1207.

Smith, R. L. (1994). *Multivariate threshold methods.* Kluwer.

Tajvidi, N. (1996). *Characterisation and Some Statistical Aspects of Univariate and Multivariate Generalized Pareto Distributions.* Thèse de doctorat, Chalmers, Göteborg, Department of Mathematics.

Tasche, D. (2008). Capital allocation to business units and sub-portfolios : the Euler Principle. Quantitative finance papers, arXiv.org.

Tibiletti, L. (1990). Gli insiemi di livello delle funzioni di ripartizione n-dimensionali : un'applicazione statistica (Level sets of n-dimensional distribution functions : a statistic application). *Quaderni dell'Istituto di Matematica Finanziaria dell'università di Torino : terza serie*, 56:1–11.

Tibiletti, L. (1991). Sulla quasi concavita delle funzioni di ripartizione n-dimensionali - on quasi-concavity of n-dimensional distribution functions. *In Atti del XV convegno A.M.A.S.E.S.*, pages 503–515.

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

Tibiletti, L. (1994). Connectedness, arcwise-connectedness and convexity for levels sets of multidimensional distribution functions. *Ratio Mathematica*, 8:21–27.

Tibiletti, L. (1995). Quasi-concavity property of multivariate distribution functions. *Ratio Mathematica*, 9:27–36.

Tsybakov, A. B. (1997). On nonparametric estimation of density level sets. *The Annals of Statistics*, 25(3):948–969.

Tukey, J. W. (1975). Mathematics and the picturing of data. *In Proceedings of the International Congress of Mathematicians (Vancouver, B. C., 1974), Vol. 2*, pages 523–531. Canad. Math. Congress, Montreal, Que.

Wang, S. (1998). An actuarial index of the right-tail risk. *N. Am. Actuar. J.*, 2(2):88–101.

Wei, G. et Hu, T. (2002). Supermodular dependence ordering on a class of multivariate copulas. *Statist. Probab. Lett.*, 57(4):375–385.

Wüthrich, M. V. (2004). Bivariate extension of the Pickands-Balkema-de Haan theorem. *Ann. Inst. H. Poincaré Probab. Statist.*, 40(1):33–41.

Yamai, Y. et Yoshiba, T. (2002). Comparative analyses of expected shortfall and value-at-risk. *J. Oper. Res. Soc. Japan*, 45(4):490–506. Theory, methodology and applications in financial engineering.

Zhou, C. (2010). Why the micro-prudential regulation fails ? The impact on systemic risk by imposing a capital requirement. Dnb working papers, Netherlands Central Bank, Research Department.