

République Algérienne Démocratique et Populaire  
Ministère de l'Enseignement Supérieur et de la Recherche Scientifique  
**UNIVERSITÉ MOHAMED KHIDER-BISKRA**  
Faculté des Sciences Exactes et des Sciences de la Nature et de la Vie  
Département de Mathématiques



## THESE

Présentée en vue de l'obtention du diplôme de  
**DOCTORAT EN SCIENCES EN MATHÉMATIQUES**

Option : Statistique

Présentée par

**BETTEKA Samah**

Titre :

---

# LES VALEURS EXTREMES BIVARIEES

---

Sous la direction de Pr. **BRAHIMI Brahim**

Membres du Comité d'Examen :

MERAGHNI	Djamel	Pr.	U de Biskra	Président
BENATIA	Fatah	M.C.(A)	U de Biskra	Examineur
BENCHABEN	Abbes	M.C.(A)	U de Guelma	Examineur
HADJI	Mohamed Lakhdar	M.C.(A)	U de Annaba	Examineur
ZAGHDOUDI	Halim	M.C.(A)	U de Annaba	Examineur

**2017**

بِسْمِ اللَّهِ الرَّحْمَنِ الرَّحِيمِ

« رَبِّ أَوْزِعْنِي أَنْ أَشْكُرَ

نِعْمَتِكَ الَّتِي أَنْعَمْتَ عَلَيَّ

وَعَلَى وَالِدَيَّ وَأَنْ أَعْمَلَ صَالِحًا

تَرْضَاهُ وَأَدْخِلْنِي بِرَحْمَتِكَ فِي

عِبَادِكَ الصَّالِحِينَ »

صدق الله العظيم

سورة النمل « 19 »

# Dédicace

Je dédie cette thèse

A mes mes très chers parents,

A mon cher mari,

A mes enfants : AMMAR, ABDELMALEK & DJOUMANA,

A mon cher oncle Boukhalfa CHAOUCH,

A mon beau-père, Ammari ABDELHAFID.

# Remerciements

**A**vant tout, je tiens à remercier ALLAH le tout puissant qui m'a donné la santé, le courage et la patience pour arriver à ce jour.

*Je tiens tout d'abord à remercier sincèrement mon directeur de thèse Monsieur le professeur Brahimi Brahim, à l'université de Biskra d'avoir bien assuré la direction et l'encadrement de cette thèse. Merci pour votre gentillesse, votre patience, votre simplicité et vos précieux conseils. Vraiment merci beaucoup pour la deuxième chance que vous me donnez pour continuer mon parcours scientifique. J'ai beaucoup apprécié travailler à vos côtés tant sur le plan scientifique que sur le plan humain.*

*J'adresse mes remerciements sincères et chaleureux à Monsieur Meraghni djamel Professeur à l'université de Biskra qui me fait l'honneur de présider le jury de soutenance. Je lui adresse mes sentiments les plus respectueux.*

*Je suis extrêmement reconnaissant aux membres du jury qui m'ont fait l'honneur d'accepter d'examiner et d'évaluer ce travail. Mes sincères remerciements et ma gratitude vont à Monsieur Benatia Fatah d'université de Biskra d'avoir accepté de faire partie du jury de soutenance. Merci infiniment à Messieurs Hadji Mohamed lakhdar et Zaghdoudi Halim d'université de Annaba et Monsieur Benchaben Abbas d'université de Guelma d'avoir accepté de rapporter cette thèse et d'avoir accepté de faire le déplacement malgré leurs multiples responsabilités.*

*Mes plus vifs remerciements s'adressent à Monsieur Necir Abdelhakim, Professeur à l'université de Biskra, pour sa disponibilité sa rigueur scientifique pour sa gentillesse et son humanité.*

*Je remercie Madame Messaci Fatiha, Professeur à l'université de Constantine, pour ses qualités humaines et pour avoir lu attentivement mon travail et avoir suggéré des remarques qui ont améliorés la thèse malgré son emploi du temps chargé.*

*Je me permets de remercier mes parents et mon cher époux pour leur soutien moral et leur encouragement tout au long de mes études.*

*Je voudrais également remercier mon amie et collègue à l'université de Biskra Sana Benameur qui trouve toujours les bons mots pour m'encourager.*

*Merci à tous les membres du département de Mathématiques d'université Mohamed khider, Biskra et d'université des frères Mentouri, Constantine, ainsi que tout le cadre professoral et administratif de l'institut de vétérinaire El khroub.*

*Enfin un grand merci à toute personne qui a contribué de près ou de loin à l'élaboration de ce travail.*

*Samah Bettaka*

# Table des matières

Dédicace	i
Remerciements	ii
Table des Matières	iv
Table des Figures	viii
Liste des Tableaux	xi
Notation et abréviations	xii
Introduction	1
<b>1 Étude des Valeurs Extrêmes Univariées</b>	<b>5</b>
1.1 Introduction . . . . .	5
1.2 Statistique d'ordre . . . . .	6
1.3 Loi des valeurs extrêmes . . . . .	10
1.3.1 Distribution des valeurs extrêmes généralisée . . . . .	13
1.3.2 Variations régulières . . . . .	15
1.3.3 Domaines d'attraction . . . . .	19
1.3.4 Conditions de Von Mises . . . . .	22
1.4 Distributions des dépassements d'un seuil . . . . .	23

1.4.1	Détermination du seuil . . . . .	25
<b>2</b>	<b>Estimation de l'indice des valeurs extrêmes et de quantiles extrêmes</b>	<b>27</b>
2.1	Introduction . . . . .	27
2.2	Analyse exploratoire des données . . . . .	28
2.2.1	Probabilité et Quantile Plots . . . . .	28
2.2.2	Pareto quantile plot . . . . .	29
2.2.3	Quantile plot généralisé . . . . .	30
2.2.4	Exponentiel quantile plot . . . . .	31
2.2.5	Mean Excess Plot . . . . .	32
2.3	Estimateurs de l'indice des valeurs extrêmes . . . . .	33
2.3.1	Estimateur de Pickands . . . . .	34
2.3.2	Estimateur de Hill . . . . .	35
2.3.3	Estimateur du moment . . . . .	37
2.3.4	Estimateur de type noyau . . . . .	38
2.4	Choix du nombre optimal de statistiques d'ordre extrêmes . . . . .	40
2.4.1	Approche de Hall et Welsh . . . . .	41
2.4.2	Approche de Bootstrap . . . . .	41
2.4.3	Approche séquentielle . . . . .	43
2.4.4	Choix automatique de paramètre de lissage . . . . .	44
2.5	Estimation des quantiles extrêmes . . . . .	46
<b>3</b>	<b>Copules et dépendance</b>	<b>49</b>
3.1	Introduction . . . . .	49
3.2	Définitions et propriétés de base . . . . .	50
3.3	Théorème de Sklar . . . . .	52
3.4	Bornes de Fréchet-Hoeffding . . . . .	54
3.5	Copule et variables aléatoires . . . . .	55

3.6	Copule de survie . . . . .	56
3.7	Familles de copules . . . . .	57
3.7.1	Copules elliptiques . . . . .	57
3.7.2	Copules Archimédiennes . . . . .	60
3.7.3	Copule de valeur extrême . . . . .	65
3.8	Mesures d'association . . . . .	67
3.8.1	Corrélations linéaires . . . . .	67
3.8.2	Mesures de concordance . . . . .	68
3.8.3	Tau de Kendall et Rho de Spearman . . . . .	69
3.8.4	Mesures de dépendance . . . . .	71
3.8.5	Dépendance de queue . . . . .	72
<b>4</b>	<b>Nombre optimal des statistiques d'ordre extrêmes utilisées dans l'esti-</b>	
	<b>mation du coefficient de dépendance de queue</b>	<b>78</b>
4.1	Introduction et motivation . . . . .	79
4.1.1	Distribution de Farlie Gumbel Morgenstern . . . . .	82
4.1.2	Distribution normale bivariée . . . . .	83
4.2	Estimation du coefficient de dépendance de queue . . . . .	83
4.2.1	Estimation des paramètres EPD . . . . .	85
4.3	Sélection optimale de la fraction d'échantillon . . . . .	87
4.3.1	Approche graphique . . . . .	87
4.3.2	Erreur moyenne quadratique . . . . .	88
4.3.3	Approche de Reiss et Thomas . . . . .	89
4.3.4	Approche de Cheng et Peng . . . . .	89
4.4	Simulation . . . . .	91
4.4.1	Données réelles . . . . .	96
4.5	Conclusion . . . . .	99



**Bibliographie**

**102**

# Table des figures

1.1	(a) : Densité de Gumbel ( $\gamma = 0$ noir), Fréchet ( $\gamma = 1$ bleu) et Weibull ( $\gamma = -1$ rouge). (b) : fonction de distribution correspondante. . . . .	12
1.2	Observations $X_1, \dots, X_{11}$ et excès $Y_1, \dots, Y_6$ au-delà du seuil $u$ . . . . .	24
1.3	(a) Densité de la loi de Paréto généralisé standard avec le paramètre de forme 0 (noir), 1 (bleu) et -1 (rouge). (b) : fonction de distribution correspondante. . . . .	25
2.1	qq-plot pour les réclamations d'assurance incendie danoises vs (a) : Quantiles exponentiels, (b) Quantiles gaussiennes . . . . .	29
2.2	Pareto-quantile plot de : (a) distribution Uniforme standard et (b) distribution Pareto standard, basé sur 4000 observations. . . . .	30
2.3	Quantile plot généralisé de : (a) distribution exponentielle et (b) distribution de Burr, basé sur 4000 observations. . . . .	31
2.4	ME-plot des données simulées à partir d'une (a) : distribution exponentielle basée sur 4000 observations et (b) : de données de Réclamations danoises .	33
2.5	Estimateur de Pickands, en fonction du nombre des extrêmes (en trait plein) avec l'intervalle de confiance 95% (lignes tirées), pour l'IVE de la distribution Pareto ( $\gamma = 1$ ) basé sur 300 échantillons de 4000 observations. . . . .	35
2.6	Estimateur de Hill, en fonction du nombre des extrêmes (en trait plein) avec l'intervalle de confiance 95% (lignes tirées), pour l'IVE de la distribution Fréchet ( $\gamma = 1$ ) basé sur 300 échantillons de 4000 observations. . . . .	36

2.7	Estimateur de Moment, en fonction du nombre des extrêmes (en trait plein) avec l'intervalle de confiance 95% (lignes tirées), pour la distribution de Gumbel(0) basée sur 300 échantillons de 4000 observations. . . . .	38
3.1	Représentation de la copule produit (gauche) et densité correspondante (droite). . . . .	53
3.2	Densité (gauche), contour plot (droite) d'une copule Gaussian bivariée avec $\rho = 0.5$ . . . . .	59
3.3	Densité (gauche) et contour plot (droite) d'une t-copule pour $\rho = 0.7$ et $\nu = 3$ . . . . .	60
3.4	Densité (gauche) et contour plot (droite) de la copule de Frank bivariée avec $\theta = 10$ . . . . .	63
3.5	Densité (gauche) et contour plot (droite) de la copule de Clayton bivariée avec $\theta = 10$ . . . . .	64
3.6	Densité (gauche) et contour plot (droite) de la copule de Gumbel bivariée avec $\theta = 2$ . . . . .	65
4.1	Estimateur de $\eta$ pour la distribution Farlie-Gumbel-Morgenstern, basée sur 500 échantillons de taille 1000 (a) estimateurs de Hill, (b) estimateurs de Moment. La ligne horizontale correspond à la vraie valeur de $\eta$ et la ligne verticale correspond $k_{opt}$ . . . . .	88
4.2	Copule de Farlie Gumbel Morgenstern avec $\alpha = -1$ : (a) estimateurs pour $\eta$ , (b) MSE : (ligne complète) Estimateur de Hill, (ligne pointillée) Estimateur de moment, basé sur 1000 simulations. La ligne horizontale correspond à la vraie valeur de $\eta$ . . . . .	92

4.3	Copule de Farlie Gumbel Morgenstern avec $\alpha = 0.75$ : (a) estimateurs pour $\eta$ , (b) MSE : (ligne complète) Estimateur linéaire , (ligne pointillée) Estimateur de Hill, basé sur 1000 échantillons. La ligne horizontale correspond à la vraie valeur de $\eta$ . . . . .	93
4.4	Copule Normale bivariéé avec $\rho = -0.5$ : (a) estimateurs pour $\eta$ , (b) MSE : (ligne complète) Estimateur de Hill, (ligne pointillée) Estimateur de moment, basé sur 1000 échantillons. La ligne horizontale correspond à la vraie valeur de $\eta$ . . . . .	93
4.5	L'estimateur de Hill de la distribution FGM avec $\alpha = 0.75$ , basé sur 1000 simulations. La ligne horizontale représente la valeur réelle de $\eta$ , alors que les lignes verticales correspondent au nombre optimal d'extrêmes de Cheng & Peng (solide) et Reiss & Thomas (pointillés). . . . .	96
4.6	Données de loss-ALAE : (a) l'estimateur de Hill pour $\eta$ , (b) MSE. . . . .	97
4.7	Données de loss-ALAE : (a) l'estimateur de Moment pour $\eta$ , (b) MSE. . . . .	98
4.8	Données de loss-ALAE : (a) l'estimateur linéaire pour $\eta$ , (b) MSE. . . . .	98
4.9	Données de loss-ALAE : (a) estimateurs pour $\eta$ , (b) MSE : l' estimateur linéaire (Ligne complète), l'estimateur de Hill (ligne en pointillé). . . . .	99

# Liste des tableaux

1.1	Domaines d'attraction des lois usuelles. . . . .	21
3.1	Tau de Kendall et Rho de Spearman de quelques copules. . . . .	71
3.2	Dépendance de la queue de certaines familles des copules. . . . .	75
4.1	Résultats de simulation de la procédure de sélection le seuil pour l'estimateur de Hill et l'estimateur de Moment. . . . .	94
4.2	Résultats de simulation de la procédure de sélection de $k$ pour l'estimateur de Hill par la méthode Cheng et Peng. . . . .	95
4.3	Résultats de simulation de la procédure de sélection de $k$ pour l'estimateur de Hill par la méthode de Reiss et Thomas . . . . .	95
4.4	Nombre optimal de statistiques d'ordre extrêmes utilisées dans le calcul de l'estimation de Hill de la distribution des FGM (0,75), basée sur 1000 observations. . . . .	95
4.5	Nombres optimaux de statistiques d'ordre supérieur utilisées dans le calcul des estimations de coefficient de dépendance de la queue pour les données de loss-ALAE. . . . .	97

# Notation et abréviations

$\xrightarrow{d}$	: convergence en distribution
$\xrightarrow{P}$	: convergence en probabilité
$\xrightarrow{p.s.}$	: convergence presque sûre
$H_\gamma$	: distribution des valeurs extrêmes (EVD)
GEVD	: distribution des valeurs extrêmes généralisée
GPD	: distribution de Paréto généralisée
$\mathcal{DA}(H_\gamma)$	: domaine d'attraction de $H_\gamma$
$(\Omega, F, P)$	: espace de probabilité
MSE	: erreur moyenne quadratique
EMV	: estimation par maximum de vraisemblance
POT	: Peaks-Over-Threshold
$\stackrel{D}{=}$	: égale en distribution
$F$	: fonction de répartition
$\bar{F}$	: fonction de survie
$Q$	: fonction inverse généralisée de F
$F_n$	: fonction de répartition empirique
$Q_n$	: fonction des quantiles empiriques
$1_A$	: fonction indicatrice de l'événement $A$
$l$	: fonction à variation lente
al	: autre
IVE	: indice des valeurs extrêmes
$\mathcal{N}(\mu, \sigma^2)$	: loi normale d'espérance $\mu$ et de variance $\sigma^2$

$k$	: nombre de statistique d'ordre extrêmes
$\#$	: nombre de
TCL	: théorème centrale limite
$x_F$	: point terminal de $F$
s.o	: statistique d'ordre
TVE	: théorie des valeurs extrêmes
GEV	: valeurs extrêmes généralisé
$Cov(X, Y)$	: covariance entre X et Y
$E(X)$	: espérance mathématique de X
$Var(X)$	: variance de la variable aléatoire X
$ \cdot $	: valeur absolue
$\sim$	: $f(x) \sim g(x)$ quand $x \rightarrow x_0$ : $f(x)/g(x) \rightarrow 1$ , quand $x \rightarrow x_0$
$o_{\mathbb{P}}(\cdot)$	: Converge vers 0 en probabilité
$O_{\mathbb{P}}(\cdot)$	: Être borné en probabilité
$\mathbb{R}$	: ensemble des réels
$\mathbb{N}$	: ensemble des entiers naturels
$\mathbb{N}^*$	: ensemble des entiers naturels non nuls
$X_{n,n}$	: maximum de $X_1, \dots, X_n$
$X_{k,n}$	: $k^{\text{ème}}$ statistique d'ordre
$:=$	: égalité par définition

# Introduction générale

La théorie des valeurs extrêmes (TVE) ou "Extreme Value Theory" en anglais, est une branche de la théorie des probabilités et des statistiques mathématiques qui se concentre sur l'analyse et l'inférence sur les événements extrêmes, c'est-à-dire les événements avec très faible probabilité d'occurrence. Les événements extrêmes revêtent une grande importance dans presque tous les domaines de la science et de la technologie, car ils peuvent s'avérer catastrophiques et donc très coûteux. Cette motivation vient souvent de finance et d'assurance deux des domaines d'application les plus populaires de TVE. Pour la littérature concernant la théorie des valeurs extrêmes nous référons aux ouvrages de Embrechts et al (1997) [51], Beirlant et al. (2005) [9], Reiss et Thomas (2007) [119] et de Hann et Ferreira (2006) [79].

L'étude des copules est un domaine en pleine croissance. La construction et les propriétés de copules ont fait l'objet d'études assez abondantes au cours des quinze dernières années. Hutchinson et Lai (1990) ont été parmi les premiers auteurs qui ont popularisé l'étude des copules. Nelsen (1999) [108] a présenté un traitement de copulas bivariée, tandis que Joe (1997) a consacré un chapitre de son livre à multivariable copules. D'autres mises à jour faisant autorité sur les copules sont données dans Nelsen (2006) [109]. Les méthodes de copule ont de nombreuses applications importantes voir Cherubini et al. (2004) [23] et Embrechts et al. (2003) [54].

Que sont les copules ? En bref, les copules sont des fonctions qui "Couple" des distributions multivariées à leur distribution marginale unidimensionnelle et de façon équivalente, les



copules sont des distributions multivariées dont les marginaux sont uniformes sur l'intervalle  $[0, 1]$ . Dans cette thèse, nous limitons attention aux copulas bivariées. Fisher (1997) a donné deux raisons majeures pour expliquer pourquoi les copules sont intéressantes aux statisticiens : Tout d'abord, comme un moyen d'étudier les mesures de dépendance sans échelle, et deuxièmement, comme point de départ pour construire des familles de variables bivariées.

Les copules constituent une part importante de l'étude des dépendance entre deux variables puisqu'elles permettent de séparer l'effet de la dépendance par rapport aux effets des distributions marginales. De nombreux auteurs ont étudié les constructions de distributions bivariées marginals.

Les concepts de dépendance de queue sont des outils standards pour décrire la dépendance extrême entre les variables aléatoires. Les coefficients de dépendance de queue (supérieur et inférieur) mesurent la probabilité d'occurrence de valeurs extrêmes (très grande ou petite) pour une variable aléatoire étant donné que l'autre suppose une valeur extrême aussi. Ces mesures de dépendance peuvent être exprimées via des copules de vecteurs aléatoires qui capturent les propriétés de la distribution conjointe qui sont invariantes à l'échelle.

Il existe différents estimateurs du coefficient de dépendance de queue dans un cadre paramétrique, semi-paramétrique et non paramétrique, voir Peng (1999) [113], Draisma et al. (2004) [43] et Beirlant et Vandewalle (2002) [5], Heffernan et Tawn (2004) [85] et Beirlant et al. (2011) [12]. Les estimateurs proposés dépendent largement du nombre de statistiques d'ordre extrêmes observées.

Notre premier but dans cette thèse est de donner un aperçu structuré et une compréhension des principaux thèmes de la théorie de la valeur extrême, à la fois univariée et multivariée et de modéliser la structure de dépendance entre les données extrêmes en utilisant la copule et le deuxième but c'est discuter, comparer et essayer d'améliorer la performance de l'estimateur du coefficient de dépendance de queue par la sélection optimale de nombre de statistiques d'ordre extrêmes. Les méthodes statistiques les plus avancées pour les

événements extrêmes seront étudiées à la fois du côté théorique et du côté des applications. Pour ce dernier, le logiciel d'analyse statistique **R** sera utilisé.

La présente thèse est une synthèse des travaux de recherches concernant la théorie des valeurs extrêmes. Elle est constituée de quatre chapitres dont l'introduction générale. Le travail présenté s'articule autour le plan du travail suivant :

**Chapitre 1 :** Dans ce chapitre nous présentons les principaux résultats de la théorie des valeurs extrêmes univariée, nous rappelons la statistique d'ordre, qui est très utile en théorie des valeurs extrêmes. Nous étudions le comportement des valeurs extrêmes d'un échantillon de variables aléatoires unidimensionnelles, nous donnons des résultats limites sur la distribution du maximum de l'échantillon et sur la répartition des dépassements au-delà d'un seuil élevé. Cette étude faisant appel à la notion de fonctions à variations régulières qui permet de caractériser les domaines d'attraction. Deux théorèmes sont essentiels et importants dans ce chapitre sont de Fisher-Tippet et de Balkema-de Haan-Pickands.

**Chapitre 2 :** Dans ce chapitre, nous présentons les estimateurs de l'indice des valeurs extrêmes et des quantiles extrêmes les plus utilisés avec une attention particulière à l'estimation semi-paramétrique (Hill, Moment, Pickands parmi autres). Dans l'estimation de l'IVE on distingue deux modèles d'estimation : la première est basée sur la distribution des valeurs extrêmes généralisée et la deuxième est basée sur la distribution de Pareto généralisée. Nous concentrons dans cette thèse sur l'estimateur de Hill qui est l'estimateur le plus populaire pour l'indice de queue positif de plus dans la majorité des applications réelles les ensembles de données correspondent aux distributions appartenant au domaine d'attraction de Fréchet.

**Chapitre 3 :** Les copules sont utilisées pour modéliser la relation de dépendance entre deux ou plusieurs variables aléatoires. Elles servent à établir une fonction de répartition multivariée à partir des fonctions de répartition marginales. Dans ce chapitre, nous introduisons quelques notions de base sur les concepts de copules. Il existe un grand nombre de familles de copules qui conduisent à des structures de dépendance positive ou négative

très variées, nous donnons dans ce chapitre les familles les plus importantes des copules paramétriques : la copule elliptique, la copule archimédienne et la copule des valeurs extrêmes. Aussi, nous donnons les mesures d'association. Des ouvrages de référence sur la théorie des copules sont Joe (1997) [89] et Nelsen (2006) [109]. Des applications en gestion des risques et actuariat sont proposées dans les livres, Nous nous référons à McNeil et al (2005) [101] Denuit et al (2005) [37].

**Chapitre 4 :** Dans ce chapitre, nous nous intéressons au choix optimal du nombre de statistiques d'ordre extrêmes impliqué dans le calcul de l'estimateur du coefficient de dépendance de queue, puisque cette estimation est nécessairement liée du ce nombre alors nous avons proposé dans ce chapitre certaines des méthodes pour choisir la valeur de  $k$  qui minimise l'erreur moyenne quadratique (MSE) de l'estimateur, puisque cette quantité équilibre deux effets : le biais et la variance. Le problème du choix de la valeur optimale des statistiques d'ordre supérieur a reçu beaucoup d'attention dans la littérature voir Danielsson & al. (2001) [28], Draisma & al. (1999) [42], Drees & Kaufmann (1998) [47], Neves & Fraga (2004) [110] et Cheng & Peng (2001) [22]. Ce Chapitre correspond à l'article "Optimal number of upper order statistics used in estimation for the coefficient of tail dependence", publié dans Journal Afrika Statistika, 12 (1), 1095-1108 (2017).

# Chapitre 1

## Étude des Valeurs Extrêmes

### Univariées

#### 1.1 Introduction

**D**epuis quelques années, la théorie des valeurs extrêmes a reçu beaucoup d'attention de nombreux statisticiens, ingénieurs et scientifiques tant le champ d'application qu'elle touche est vaste : Hydrologie, biologie, ingénierie, météorologie, gestion de l'environnement, finance, assurance, etc.

Dans ce chapitre, Nous proposons un aperçu des éléments théoriques essentiels de la théorie des valeurs extrêmes. On commence par la statistique d'ordre, qui est très utile en TVE et nous étudions le comportement des valeurs extrêmes d'un échantillon de variables aléatoires unidimensionnelles, nous donnons des critères pour que la limite en loi du maximum suive telle loi des valeurs extrêmes. Ces critères faisant appel à la notion de fonctions à variations régulières, on rappelle au préalable la définition de telles fonctions et on en donne quelques propriétés. Pour plus de détails sur ce chapitre, nous referons aux ouvrages de référence sur la théorie et les applications des valeurs extrêmes suivantes : Embrechts et al (1997) [51], Beirlant et al. (2006) [8], Reiss et Thomas (2007) [119], de Hann et Ferreira

(2006) [79] et Coles (2001) [25].

## 1.2 Statistique d'ordre

Les statistiques d'ordre sont très utiles dans la théorie des valeurs extrêmes, ils fournissent des informations sur la distribution de queue. Dans cette section, nous allons résumer certains de leurs résultats. Pour une étude approfondie, sur les statistiques d'ordre nous suggérons l'ouvrage d'Arnold et al [1].

**Définition 1.2.1 (Statistique d'ordre)** *La statistique d'ordre de l'échantillon  $(X_1, \dots, X_n)$  est le réarrangement croissant, notée par  $(X_{1,n}, \dots, X_{n,n})$  de sorte que  $X_{1,n} \leq X_{2,n} \leq \dots \leq X_{n,n}$ . En particulier, la variable aléatoire  $X_{i,n}$ ,  $i = 1, \dots, n$  désigne la  $i^{\text{ème}}$  statistique d'ordre.*

Les statistiques d'ordre extrêmes sont définies comme termes du maximum et du minimum de l'échantillon  $(X_1, \dots, X_n)$ . La variable  $X_{1,n}$  est la plus petite statistique d'ordre et  $X_{n,n}$  est la plus grande statistique d'ordre sont définies par :

$$X_{1,n} := \min(X_1, \dots, X_n) \quad \text{et} \quad X_{n,n} := \max(X_1, \dots, X_n). \quad (1.1)$$

- On s'intéresse dans cette thèse au comportement de la statistique d'ordre  $X_{n,n}$ .
- On peut convertir les résultats sur le maximum pour appliquer au minimum en utilisant la règle triviale suivante :

$$\min(X_1, \dots, X_n) = -\max(-X_1, \dots, -X_n).$$

**Proposition 1.2.1 (Fonction de répartition et densité de la  $i^{\text{ème}}$  statistique d'ordre)**

- La fonction de distribution de  $X_{i,n}$  pour  $i = 1, \dots, n$  est donnée par :

$$F_{X_{i,n}}(x) = \sum_{r=i}^n C_n^r [F(x)]^r [1 - F(x)]^{n-r}, \quad x \in \mathbb{R}.$$

- La Fonction de densité de  $X_{i,n}$  pour  $i = 1, \dots, n$  est présentée sous la forme suivante :

$$f_{X_{i,n}}(x) = \frac{n!}{(i-1)!(n-i)!} (F(x)^{i-1}) (1-F(x))^{i-k} f(x), \quad x \in \mathbb{R}.$$

**Démonstration.** Voir Reiss (1989) [118], Lemme 1.3.1, page 20. □

**Corollaire 1.2.1** Si la loi de  $X_1$  possède une densité  $f$ , alors la statistique d'ordre  $(X_{1,n}, \dots, X_{n,n})$  possède la densité suivante

$$f_{X_{1,n}, \dots, X_{n,n}}(x_1, x_2, \dots, x_n) = n! \prod_{i=1}^n f(x_i), \quad -\infty < x_1 < \dots < x_n < \infty.$$

**Corollaire 1.2.2 (Lois du maximum et du minimum)** Soit  $(X_1, \dots, X_n)$  une suite de variables aléatoires indépendantes et identiquement distribuées (i.i.d) définie sur l'espace  $(\Omega, \mathcal{B})$ , d'une densité commune  $f$  et d'une fonction de distribution  $F$  alors on a

- (a) La fonction de répartition de  $X_{1,n}$  est donnée par

$$F_{X_{1,n}}(x) = 1 - [1 - F_X(x)]^n, \quad x \in \mathbb{R}.$$

- (b) La fonction de répartition de  $X_{n,n}$  est donnée par

$$F_{X_{n,n}}(x) = [F_X(x)]^n, \quad x \in \mathbb{R}.$$

- (c) La fonction de densité de  $X_{1,n}$  est donnée par

$$f_{X_{1,n}}(x) = n [1 - F_X(x)]^{n-1} f(x), \quad x \in \mathbb{R}.$$

- (d) La fonction de densité de  $X_{n,n}$  est donnée par

$$f_{X_{n,n}}(x) = n [F_X(x)]^{n-1} f(x), \quad x \in \mathbb{R}.$$

(e) La fonction de densité conjointe de  $X_{1,n}$  et  $X_{n,n}$

$$f_{1,n}(x, y) = n(n-1)(F(y) - F(x))^{n-2} f(x)f(y), \quad -\infty < x < y < \infty.$$

**Lemme 1.2.1** La variable aléatoire  $Y_k = F(X_{k,n})$  suit une loi bêta de paramètre  $(k, n - k + 1)$ .

**Définition 1.2.2 (Distribution empirique)** La fonction de répartition empirique de l'échantillon  $(X_1, \dots, X_n)$  est évaluée à l'aide des statistiques d'ordre comme suit :

$$F_n(x) := \begin{cases} 0 & \text{si } x < X_{1,n} \\ \frac{i-1}{n} & \text{si } X_{i-1,n} \leq x < X_{i,n}, \quad 2 \leq i \leq n \\ 1 & \text{si } x \geq X_{n,n}. \end{cases}$$

**Définition 1.2.3 (Fonctions quantile et quantile de queue)** La fonction quantile de la fonction de distribution  $F$  est la fonction inverse généralisée de  $F$  définie par :

$$Q(s) := F^{\leftarrow}(s) = \inf \{x \in \mathbb{R} : F(x) \geq s\}, \quad 0 < s < 1, \quad (1.2)$$

Dans la théorie des extrêmes, une fonction notée  $U$  et appelée fonction quantile de queue, est définie par :

$$U(t) := Q(1 - 1/t) = (1/\bar{F})^{\leftarrow}(t), \quad 1 < t < \infty. \quad (1.3)$$

**Définition 1.2.4 (Fonctions empiriques de quantile et de quantile de queue)** La fonction quantile empirique de l'échantillon  $(X_1, X_2, \dots, X_n)$  est définie par :

$$Q_n(s) := F_n^{\leftarrow}(s) = \inf \{x \in \mathbb{R} : F_n(x) \geq s\}, \quad 0 < s < 1. \quad (1.4)$$

La fonction empirique de quantile de queue correspondante est

$$U_n(t) := Q_n(1 - 1/t), \quad 1 < t < \infty. \quad (1.5)$$

$Q_n$  Peut être exprimée comme une fonction simple des statistiques d'ordre relatives à l'échantillon  $(X_1, X_2, \dots, X_n)$  et nous avons

$$Q_n(s) := \begin{cases} X_{i,n} & \text{si } \frac{(i-1)}{n} < s \leq \frac{i}{n} \\ X_{[np]+1,n} & \text{si } 0 < s \leq 1. \end{cases} \quad (1.6)$$

Notons que pour  $0 < p < 1$ ,  $X_{[np]+1,n}$  est le quantile d'échantillon de l'ordre  $p$ .

**Théorème 1.2.1 (Glivenko-Cantelli, 1933)**

$$\sup_{x \in \mathbb{R}} |F_n(x) - F(x)| \xrightarrow{p.s} 0 \quad \text{quand } n \rightarrow \infty.$$

**Définition 1.2.5 (Point extrême)** On note par  $x_F$  (resp.  $x_F^*$ ) le point extrême supérieur (resp. inférieur) de la distribution  $F$  (i.e. la plus grande valeur possible pour  $X_{k,n}$  peut prendre la valeur  $+\infty$  (resp.  $-\infty$ )) au sens où

$$x_F := \sup\{x : F(x) < 1\} \leq \infty, \quad (1.7)$$

respectivement

$$x_F^* := \inf\{x : F(x) > 0\}.$$

**Proposition 1.2.2 (Limite de  $X_{n,n}$ )**  $X_{n,n} \xrightarrow{p.s} x_F$  quand  $n \rightarrow \infty$ .

D'autre part, le comportement asymptotique de la loi de  $X_{n,n}$  est donné, dans certaines conditions sur la queue de distribution.



**Lemme 1.2.2 (Transformation quantile)** Soit  $X_1, \dots, X_n$  des variables aléatoires indépendantes et de fonction de répartition  $F$ . Soit  $U_1, \dots, U_n$  des variables aléatoires indépendantes de loi uniforme standard. Alors

(a) Pour toute fonction de distribution  $F$ , on a

$$X_{i,n} \stackrel{d}{=} F^{\leftarrow}(U_{i,n}), \quad i = 1, \dots, n. \quad (1.8)$$

(b) Lorsque  $F$  est continue, on a

$$F(X_{i,n}) \stackrel{d}{=} U_{i,n}, \quad i = 1, \dots, n. \quad (1.9)$$

**Démonstration.** Voir Reiss (1989) [118], Théorème 1.2.5, page 17. □

### 1.3 Loi des valeurs extrêmes

Quand on s'intéresse à la partie centrale d'un échantillon, le résultat nécessaire est le théorème centrale limite<sup>1</sup> qui donne la loi asymptotique de la somme des observations. D'autre part, si l'on veut étudier les valeurs extrêmes de cet échantillon, ce théorème est inutile car la fonction de distribution de maximum est dégénérée<sup>2</sup>. Donc nous utilisons un résultat qui établit le maximum asymptotique de l'échantillon  $X_{n,n}$  d'une façon analogue au théorème central limite, la théorie des valeurs extrêmes montre qu'il existe des suites

---

<sup>1</sup>

**Théorème 1.3.1 (TCL)** Si  $(X_1, \dots, X_n)$  est une suite de variable aléatoire définie sur le même espace de probabilité de variance  $\sigma^2$  finie et de moyenne  $\mu$  alors

$$\frac{1}{\sqrt{n}} \left( \frac{S_n - n\mu}{\sigma} \right) \xrightarrow{d} \mathcal{N}(0, 1) \quad \text{quand } n \rightarrow \infty.$$

<sup>2</sup>La distribution asymptotique du maximum, déterminée en faisant tendre  $n$  vers l'infini, donne une loi dégénérée. En effet

$$\lim_{n \rightarrow \infty} F_{X_{n,n}} = \lim_{n \rightarrow \infty} [F(x)]^n = \begin{cases} 1 & \text{si } F(x) = 1, \\ 0 & \text{si } F(x) < 1. \end{cases}$$

$\{a_n\}$  et  $\{b_n\}$   $n \in \mathbb{N}$ , avec  $a_n > 0$  et  $b_n \in \mathbb{R}$ , telles que

$$\lim_{n \rightarrow \infty} P \left( \frac{X_{n,n} - b_n}{a_n} \leq x \right) = \lim_{n \rightarrow \infty} F^n(a_n x + b_n) = H(x), \quad x \in \mathbb{R}, \quad (1.10)$$

avec  $H$  est une distribution non dégénérée.

Puisque les fonctions de distribution de valeur extrême sont continues sur  $\mathbb{R}$ , l'hypothèse 1.10 est équivalente à l'hypothèse de convergence faible suivante :

$$\frac{X_{n,n} - b_n}{a_n} \xrightarrow{d} H \quad \text{quand } n \rightarrow \infty, \quad (1.11)$$

$a_n$  et  $b_n$  sont appelées constantes de normalisation.

Un des résultats fondamentaux de la TVE est le théorème suivant qu'a été énoncé pour la première fois par Fisher et Tippett en (1928), puis démontré par Gnedenko en (1943) ensuite il était repris et simplifié par de Haan en (1970).

**Théorème 1.3.2 (Fisher et Tippett (1928), Gnedenko (1943))** *Soit  $(X_n)_n$  une suite de variables aléatoire (i.i.d), s'il existe un réel  $\gamma$  et deux suites réelles  $(a_n)$  et  $(b_n)$ ,  $n \in \mathbb{N}$ , avec  $a_n > 0$  et  $b_n \in \mathbb{R}$  telles que 1.11 satisfait. Pour tout  $x$ , alors  $H$  est du même type que l'une des fonctions suivantes :*

**Gumbel (type I) :**

$$H_0(x) = \Lambda(x) = \exp[-\exp(-x)], \quad x \in \mathbb{R}.$$

**Fréchet (type II) :**

$$H_\gamma(x) = \Phi_\gamma(x) = \begin{cases} 0, & x \leq 0, \\ \exp(-(x)^{-1/\gamma}), & x > 0, \end{cases} \quad \text{avec } \gamma > 0.$$

**Weibull (type III) :**

$$H_\gamma(x) = \Psi_\gamma(x) = \begin{cases} \exp(-(-x)^{-\gamma}), & x \leq 0, \\ 1, & x > 0, \end{cases} \quad \text{avec } \gamma < 0.$$

Une preuve détaillée de ce théorème peut être trouvée dans Embrechts et al (1997) [51] section 3.2, page 122.

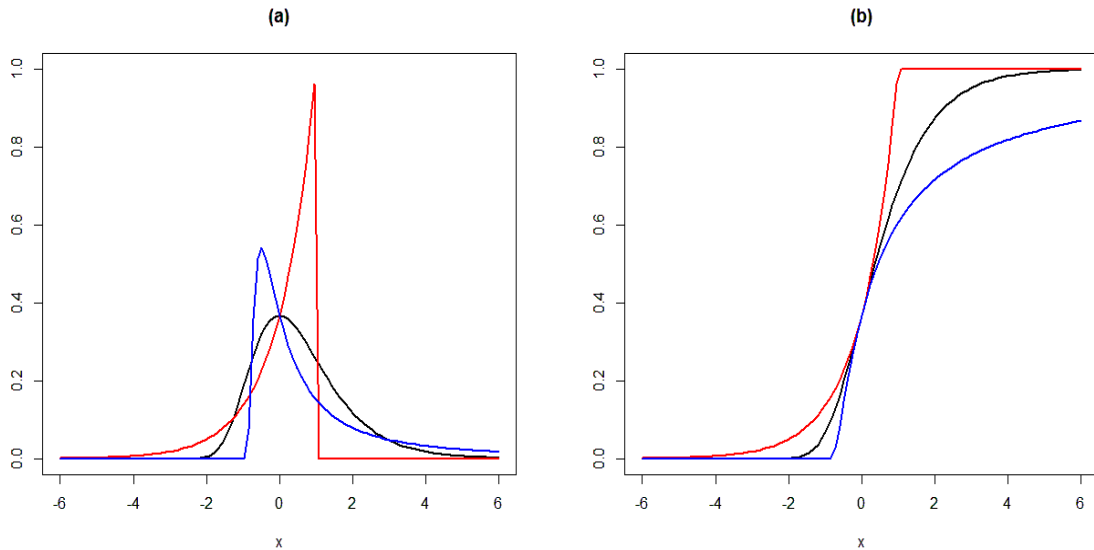


FIG. 1.1 – (a) : Densité de Gumbel ( $\gamma = 0$  noir), Fréchet ( $\gamma = 1$  bleu) et Weibull ( $\gamma = -1$  rouge). (b) : fonction de distribution correspondante.

**Définition 1.3.1** *Les trois fonctions de distribution du théorème 1.3.2 s'appellent les distributions standard ou traditionnelle des valeurs extrêmes.*

**Proposition 1.3.1 (Relation entre  $\Lambda$ ,  $\Phi_\gamma$  et  $\Psi_\gamma$ )** *Soit  $Y$  une variable aléatoire positive ( $Y > 0$ ) alors les affirmations suivantes sont équivalentes :*

- (a)  $Y \sim \Phi_\gamma$ .
- (b)  $\ln Y^\gamma \sim \Lambda$ .
- (c)  $-Y^{-1} \sim \Psi_\gamma$ .

### 1.3.1 Distribution des valeurs extrêmes généralisée

Pour faciliter le travail avec les trois distributions limites, Jenkinson-Von Mises a donné une représentation qui a obtenu en introduisant les paramètres de localisation  $\mu$  et de dispersion  $\sigma$  dans la paramétrisation des distributions extrêmes, on obtient la forme la plus générale de la distribution des valeurs extrêmes, notée GEVD (Generalized Extreme Value distribution). Elle est simplement une reparamétrisation des distributions apparaissant dans le théorème 1.3.2

**Définition 1.3.2 (GEVD)** *la distribution généralisée des valeurs extrêmes a la forme suivante :*

$$H_{\mu,\sigma,\gamma}(x) := \begin{cases} \exp \left\{ - \left[ 1 + \gamma \left( \frac{x - \mu}{\sigma} \right) \right]^{-1/\gamma} \right\} & \text{pour } 1 + \frac{\gamma}{\sigma}(x - \mu) > 0, \\ \exp \left( - \exp \left( - \frac{x - \mu}{\sigma} \right) \right) & \text{pour } \gamma = 0, x \in \mathbb{R}, \end{cases} \quad (1.12)$$

avec  $\mu \in \mathbb{R}$  et  $\sigma > 0$ .

En remplaçant  $\left( \frac{x - \mu}{\sigma} \right)$  par  $x$  on obtient la forme standard de la GEVD :

$$H_{\gamma}(x) = \begin{cases} \exp \left\{ - (1 + \gamma x)^{-1/\gamma} \right\} & \text{pour } \gamma \neq 0, 1 + \gamma x > 0, \\ \exp \{ - \exp(-x) \} & \text{pour } \gamma = 0, x \in \mathbb{R}. \end{cases} \quad (1.13)$$

Où  $\gamma$  est le paramètres de forme que l'on appelle indice des valeurs extrêmes (IVE) ou indice de queue.

La fonction de densité standard correspondante  $h_{\mu,\sigma,\gamma}$  est :

$$h_{\gamma}(x) = \begin{cases} H_{\gamma}(x)(1 + \gamma x)^{-1/\gamma-1} & \text{si } \gamma \neq 0, 1 + \gamma x > 0, \\ \exp(-x - e^{-x}) & \text{si } \gamma = 0, x \in \mathbb{R}. \end{cases}$$

Le comportement des extrêmes est alors contrôlé par le signe et la valeur de  $\gamma$ . Suivant le signe du paramètre de forme, on définit trois types de GEVD :

- $\gamma = 0$ , lois à queue légère (distribution de Gumbel).
- $\gamma > 0$ , lois à queue lourde (distribution de Fréchet).
- $\gamma < 0$ , distribution bornée (distribution de Weibull).

Les lois des valeurs extrêmes généralisées correspondent à une translation et un changement d'échelle près aux lois des valeurs extrêmes. Voir la proposition suivante :

**Proposition 1.3.2** *On a, les relations suivantes entre  $\Lambda$ ,  $\Phi_\gamma$  et  $\Psi_\gamma$  en terme de  $H_{\mu,\sigma,\gamma}$*

$$H_{1, \frac{1}{\gamma}, \frac{1}{\gamma}}(\gamma(x-1)) = \Phi_\gamma(x) \quad \text{si } \gamma > 0.$$

$$H_{0, \frac{1}{\gamma}, \frac{-1}{\gamma}}(\gamma(x+1)) = \Psi_\gamma(x) \quad \text{si } \gamma < 0.$$

$$H_{0,1,0}(x) = \Lambda(x) \quad \text{si } \gamma = 0.$$

Le choix des constantes de normalisation ( $\{a_n\}$  et  $\{b_n\}$ ) dans 1.11) n'est pas unique parce que l'unicité de la limite de  $H$  est seulement jusqu'à affiner les transformations. Réellement, c'est une conséquence du théorème de Khintchine. Les choix les plus communs sont accusés dans le théorème suivant.

**Théorème 1.3.3 (Khintchine)** *Soit  $\{G_n, n \in \mathbb{N}\}$  une suite de fonctions de distributions et  $G$  une fonction de distribution non dégénérée. Pour  $x \in \mathbb{R}$ , soit  $\{a_n\}$  et  $\{b_n\}$ ,  $n \in \mathbb{N}$ , avec  $a_n > 0$  et  $b_n \in \mathbb{R}$  tels que*

$$\lim_{n \rightarrow \infty} G_n(a_n x + b_n) = G(x).$$

*Alors, pour certaine fonction de distribution non dégénérée  $G^*$  et suites  $\{\alpha_n\}$  et  $\{\beta_n\}$ ,  $n \in \mathbb{N}$ , avec  $\alpha_n > 0$  et  $\beta_n \in \mathbb{R}$ ,*

$$\lim_{n \rightarrow \infty} G_n(\alpha_n x + \beta_n) = G^*(x).$$

*ssi*

$$\frac{\alpha_n}{a_n} \rightarrow A > 0 \quad \text{et} \quad \frac{\beta_n - b_n}{a_n} \rightarrow B \in \mathbb{R}, \quad \text{quand } n \rightarrow \infty.$$

De plus,

$$G^*(x) = G(Ax + B), \quad x \in \mathbb{R}.$$

Pour plus d'exemples de suites de normalisation de chaque loi se référer à Embrechts et al (1997) [51] page 145.

### 1.3.2 Variations régulières

La théorie des fonctions à variations régulières apparaissent dans un grand nombre d'applications dans l'ensemble des mathématiques, c'est un outil essentiel pour traiter les queues lourdes, la dépendance à longue portée et les domaines d'attraction. Les fonctions à variation régulière sont ces fonctions qui se comportent asymptotiquement comme des fonctions de pouvoir (puissance). On donne d'abord quelques généralités, ensuite on établit des propriétés fondamentales qui sont importantes dans la suite.

#### Définition et propriétés

**Définition 1.3.3** Une fonction mesurable  $h : \mathbb{R}^+ \rightarrow \mathbb{R}^+$  est à variations régulière d'indice  $\rho \in \mathbb{R}$  (notation  $h \in \mathcal{RV}_\rho$ ) si pour  $x > 0$ ,

$$\lim_{t \rightarrow +\infty} \frac{h(tx)}{h(t)} = x^\rho. \quad (1.14)$$

Nous appelons  $\rho$  l'indice de variation régulière.

- Si  $\rho = 0$ , on dit que  $h$  est une fonction à variation lente à l'infini et notée  $l$ .
- Si  $h \in \mathcal{RV}_\rho$ , alors  $h(x)/x^\rho \in \mathcal{RV}_0$ , et on pose  $l(x) = h(x)/x^\rho$ , on voit qu'il est toujours possible de représenter une fonction à variation régulière d'indice  $\rho$  comme  $x^\rho l(x)$ .

Les deux résultats les plus importants sur les fonctions appartenant à  $\mathcal{RV}_0$  sont donnés dans le théorème suivant dû à Karamata.

**Théorème 1.3.4 (Convergence Uniforme)** *Si  $l \in \mathcal{RV}_0$ , alors la convergence*

$$\lim_{x \rightarrow +\infty} \frac{l(tx)}{l(x)} = 1.$$

*est uniforme pour  $t \in [a, b]$  où  $0 < a < b < 1$ .*

**Théorème 1.3.5 (Représentation de Karamata)**

(i)  *$l$  est une fonction à variation lente ssi pour tout  $x > 0$ ,*

$$l(x) = c(x) \exp \left\{ \int_1^x t^{-1} \varepsilon(t) dt \right\},$$

*où  $c$  et  $\varepsilon$  sont des fonctions positives telles que*

$$\lim_{x \rightarrow \infty} c(x) = c \in ]0, +\infty[ \quad \text{et} \quad \lim_{t \rightarrow \infty} \varepsilon(t) = 0. \quad (1.15)$$

(ii) *Une fonction  $h : \mathbb{R}^+ \rightarrow \mathbb{R}^+$  à variation régulière d'indice  $\rho$  ssi  $h$  a la représentation suivante*

$$h(x) = c(x) \exp \left\{ \int_1^x t^{-1} \rho(t) dt \right\}, \quad (1.16)$$

*où  $c(\cdot)$  vérifie 1.15 et  $\lim_{t \rightarrow \infty} \rho(t) = \rho$ .*

**Démonstration.** Voir Resnick (2007) [121], Corollaire 2.1, page 29. □

### Remarques

- Si  $h$  est une fonction à variation régulière d'indice  $\rho > 0$ , alors  $h^\leftarrow(x)$  est une fonction à variation régulière d'indice  $1/\rho$ .
- Si  $h$  est une fonction à variations régulière d'indice  $\rho < 0$ , alors  $h^\leftarrow(1/x)$  est une fonction à variation régulière d'indice  $-1/\rho$ .

### Condition de variation régulière du second ordre

Nous définissons tout d'abord l'hypothèse du second ordre qui renforce la condition de variation régulière de la queue de distribution  $\bar{F}$ , avec un rappel sur la condition à variation régulière du premier ordre dans le cas des distributions à queue lourde.

**Proposition 1.3.3 (condition à variation régulière du premier ordre)** *les affirmations suivantes sont équivalentes :*

(a)  $F$  à queue lourde

$$F \in \mathcal{DA}(\Phi_{1/\gamma}), \quad \gamma > 0.$$

(b)  $\bar{F}$  à variation régulière à l'infini d'indice  $-1/\gamma$

$$\lim_{t \rightarrow \infty} \frac{1 - F(tx)}{1 - F(t)} = x^{-1/\gamma}, \quad x > 0. \quad (1.17)$$

(c)  $Q(1 - s)$  à variation régulière à zéro d'indice  $-\gamma$

$$\lim_{s \rightarrow 0} \frac{Q(1 - tx)}{Q(1 - t)} = x^{-\gamma}, \quad x > 0.$$

(d)  $U$  à variation régulière à l'infini d'indice  $\gamma$

$$\lim_{t \rightarrow \infty} \frac{U(tx)}{U(t)} = x^\gamma, \quad x > 0. \quad (1.18)$$

**Définition 1.3.4 (Hypothèse à variation régulière du second ordre)** *Nous disons que (la queue de)  $F \in \mathcal{DA}(\Phi_{1/\gamma})$   $\gamma > 0$ , est à variation régulière du second ordre à l'infini s'elle satisfait l'une des conditions (équivalentes) suivantes :*

(a) *Il existe un certain paramètre  $\rho \leq 0$  et une fonction  $A^*$ , tendant à zéro et ne pas change de signe au voisinage de  $\infty$ , de sorte que pour tout  $x > 0$*

$$\lim_{t \rightarrow \infty} \frac{(1 - F(tx))/(1 - F(t)) - x^{-1/\gamma}}{A^*(t)} = x^{-1/\gamma} \frac{x^\rho - 1}{\rho}. \quad (1.19)$$



(b) *Il existe un certain paramètre  $\rho \leq 0$  et une fonction  $A^{**}$ , tendant à zéro et ne pas change de signe au voisinage de  $\infty$ , de sorte que pour tout  $x > 0$*

$$\lim_{s \rightarrow 0} \frac{Q(1 - sx)/Q(1 - s) - x^{-\gamma}}{A^{**}(s)} = x^{-\gamma} \frac{x^\rho - 1}{\rho}. \quad (1.20)$$

(c) *Il existe un certain paramètre  $\rho \leq 0$  et une fonction  $A$ , tendant à zéro et ne pas change de signe au voisinage de  $\infty$ , de sorte que pour tout  $x > 0$*

$$\lim_{t \rightarrow \infty} \frac{U(tx)/U(t) - x^\gamma}{A(t)} = x^\gamma \frac{x^\rho - 1}{\rho}. \quad (1.21)$$

Si  $\rho = 0$ , interpréter  $x^\rho - 1/\rho$  comme  $\log x$ .

$A$ ,  $A^*$  et  $A^{**}$  sont des fonctions qui varient régulièrement avec  $A^*(t) = A(1/(1 - F(t)))$  et  $A^{**}(s) = A(1/s)$ . Leur rôle est de contrôler la vitesse de convergence dans 1.19, 1.20 et 1.21 respectivement. Plus précisément, nous avons  $A \in \mathcal{RV}_\rho$ ,  $A^* \in \mathcal{RV}_{\rho/\gamma}$  et  $A^{**} \in \mathcal{RV}_{-\rho}^0$ . Les relations ci-dessus peuvent être reformulées respectivement :

$$\lim_{t \rightarrow \infty} \frac{\log(1 - F(tx)) - \log(1 - F(t)) + (1/\gamma) \log x}{A^*(t)} = \frac{x^\rho - 1}{\rho}. \quad (1.22)$$

$$\lim_{s \rightarrow 0} \frac{\log Q(1 - sx) - \log Q(1 - s) + \gamma \log(x)}{A^{**}(s)} = \frac{x^\rho - 1}{\rho}.$$

$$\lim_{t \rightarrow \infty} \frac{\log U(tx) - \log U(t) - \gamma \log(x)}{A(t)} = \frac{x^\rho - 1}{\rho}. \quad (1.23)$$

### Classe de Hall de la fonction de distribution

Cette classe est présentée par Hall et Welsh en 1985 [81] et elle est mentionnée par "le modèle de Hall". Ce modèle est un exemple de distributions à queues lourdes qui satisfont la condition du second ordre où

$$F(x) = 1 - cx^{-1/\gamma}(1 + dx^{-\rho/\gamma} + o(x^{-\rho/\gamma})) \quad \text{quand } x \rightarrow \infty, \quad (1.24)$$

pour  $\gamma > 0$ ,  $\rho \leq 0$ ,  $c > 0$ ,  $d \in \mathbb{R} \setminus \{0\}$ .

La relation 1.24 peut être reformulée en termes de fonctions  $Q$  et  $U$  comme suit :

$$Q(1 - s) = c^\gamma s^{-\gamma} (1 + \gamma dc^\rho s^{-\rho} + o(s^{-\rho})) \quad \text{quand } s \rightarrow \infty.$$

et

$$U(t) = c^\gamma t^\gamma (1 + \gamma dc^\rho t^\rho + o(t^\rho)) \quad \text{quand } t \rightarrow \infty.$$

Cette sous-classe contient la plupart des distributions à queue lourde comme les distributions Pareto, Burr, Fréchet t-Student, Cauchy, Pareto généralisée,  $\alpha$ -stable, log-logistic et log-gamma, ...etc. Elle joue un rôle important dans les estimateurs de l'indice de queue positif (voir le chapitre 2).

### 1.3.3 Domaines d'attraction

Dans cette partie de la section nous donnerons des conditions nécessaires et suffisantes, utilisé pour décider auquel domaine d'attraction (Gumbel, Fréchet ou Weibull) une fonction de distribution spécifique appartient.

**Définition 1.3.5 (Domaines d'attraction)** *Si  $F$  vérifie le Théorème 1.3.2, on dit alors que  $F$  appartient au domaine d'attraction de  $H_\gamma$ , dénotée par  $F \in \mathcal{DA}(H_\gamma)$ .*

**Proposition 1.3.4 (Caractérisation de  $\mathcal{DA}(H)$ )** *La fonction de distribution  $F$  appartient au domaine d'attraction de la distribution de valeur extrême  $H$  avec des constantes de normalisation  $a_n$  et  $b_n \in \mathbb{R}$  ssi*

$$\lim_{n \rightarrow \infty} n \overline{F}(b_n + a_n x) = -\log H(x), \quad x \in \mathbb{R}. \quad (1.25)$$

Lorsque  $H(x) = 0$  la limite est interprétée comme  $\infty$ .

**Théorème 1.3.6 (Domaine d'attraction de  $\Phi_\gamma$ )** *La fonction de distribution  $F$  appartient au domaine d'attraction de  $\Phi_\gamma$ ,  $\gamma > 0$  ssi  $1 - F(x) = x^{-1/\gamma}l(x)$  où  $l$  est une fonction à variation lente. Si  $F \in DA(\Phi_\gamma)$  alors  $a_n^{-1}X_{n,n} \xrightarrow{d} \Phi_\gamma$ , où les constantes de normalisation  $a_n = Q(1 - 1/n)$  et  $b_n = 0$ .*

**Démonstration.** Voir Embrechts et al (1997) [51], Théorème 3.3.7, page 131. □

**Théorème 1.3.7 (Domaine d'attraction de  $\Psi_\gamma$ )**  *$F \in \mathcal{DA}(\Psi_\gamma)$  avec  $\gamma < 0$  ssi  $x_F < +\infty$  et la fonction  $(1 - F^*)$  est à variation régulière d'indice  $1/\gamma$  (i.e.  $1 - F(x) = (x_F - x)^{-1/\gamma} [l(x_F - x)^{-1}]$ ) avec*

$$F^*(x) = \begin{cases} F(x_F - x^{-1}) & x > 0, \\ 0 & x \leq 0. \end{cases}$$

*Dans ce domaine d'attraction les suites de normalisation  $a_n$  et  $b_n$  sont déterminées comme suit :*

$$a_n = x_F - F^{\leftarrow}\left(1 - \frac{1}{n}\right) \quad \text{et} \quad b_n = x_F.$$

**Théorème 1.3.8 (Domaine d'attraction de  $\Lambda_\gamma$ )**  *$F \in \mathcal{DA}(\Lambda_\gamma)$ ,  $\gamma > 0$  ssi il existe  $z < x_F$  tel que  $F$  a une représentation suivante :*

$$\bar{F}(x) = c(x) \exp \left\{ \int_z^x \frac{g(t)}{a(t)} dt \right\}, \quad z < x < x_F \leq \infty,$$

*où  $c$  et  $g$  sont des fonctions mesurables satisfaisant  $c(x) \rightarrow c > 0$ ,  $g(x) \rightarrow 1$  quand  $x \rightarrow x_F$  et  $a(x)$  est une fonction positive absolument continue par rapport à la mesure de Lebesgue avec densité  $a'(x)$  ayant  $\lim_{x \rightarrow x_F^+} a'(x) = 0$ . Avec représentation ?? on peut choisir*

$$b_n = Q(1 - 1/n) \quad \text{et} \quad a_n = a(b_n).$$

Domaine d'attraction	Fréchet $\gamma > 0$	Gumbel $\gamma = 0$	Weibull $\gamma < 0$
Loi	Burr Fréchet Loggamma Loglogistic Pareto	Gamma Gumbel Logistique Lognormale Weibull	Uniforme ReverseBurr

TAB. 1.1 – Domaines d'attraction des lois usuelles.

Comme des constantes de normalisation. Un choix possible pour la fonction  $a$  est

$$a(x) = \frac{1}{\overline{F}(x)} \int_x^{x_F} \overline{F}(t) dt, \quad x < x_F. \quad (1.26)$$

Le tableau 1.1 donne différents exemples de distributions standard dans les trois domaines d'attraction.

**Proposition 1.3.5 (Caractérisations de  $\mathcal{D}(H_\gamma)$ )** Pour  $\gamma \in \mathbb{R}$  les affirmations suivantes sont équivalentes.

(a)  $F \in \mathcal{D}(H_\gamma)$

(b) Pour certaine fonction positive  $b$

$$\lim_{t \rightarrow x_F} \frac{\overline{F}(t + xb(t))}{\overline{F}(t)} = \begin{cases} (1 + \gamma x)^{-1/\gamma} & \text{si } \gamma \neq 0, \\ e^{-x} & \text{si } \gamma = 0, \end{cases} \quad (1.27)$$

pour tout  $x > 0$  avec  $(1 + \gamma x) > 0$ .

(c) Pour certaine fonction positive  $\tilde{a}$

$$\lim_{s \rightarrow 0} \frac{Q(1 - sx) - Q(1 - s)}{\tilde{a}(s)} = \begin{cases} \frac{x^{-\gamma} - 1}{\gamma} & \text{si } \gamma \neq 0, \\ \log x & \text{si } \gamma = 0, \end{cases}$$

pour  $x > 0$ .

(d) Pour certain fonction positive  $a(t) = \tilde{a}(1/t)$

$$\lim_{t \rightarrow 0} \frac{U(tx) - U(t)}{a(t)} = \begin{cases} \frac{x^\gamma - 1}{\gamma} & \text{si } \gamma \neq 0, \\ \log x & \text{si } \gamma = 0, \end{cases}$$

pour  $x > 0$ .

Les deux dernières affirmations sont respectivement équivalentes à :

$$\lim_{s \rightarrow 0} \frac{Q(1 - sx) - Q(1 - s)}{Q(1 - sy) - Q(1 - s)} = \begin{cases} \frac{x^{-\gamma} - 1}{y^{-\gamma} - 1} & \text{si } \gamma \neq 0, \\ \frac{\log x}{\log y} & \text{si } \gamma = 0, \end{cases}$$

$$\lim_{t \rightarrow 0} \frac{U(tx) - U(t)}{U(ty) - U(t)} = \begin{cases} \frac{x^\gamma - 1}{y^\gamma - 1} & \text{si } \gamma \neq 0, \\ \frac{\log x}{\log y} & \text{si } \gamma = 0, \end{cases} \quad (1.28)$$

pour  $x, y > 0, y \neq 1$

### 1.3.4 Conditions de Von Mises

Von Mises (1936) a proposé quelques conditions simples et utiles pour les fonctions de répartition ayant une densité, ces conditions permettant de déterminer à quel des trois domaines d'attraction appartient la fonction de distribution correspondante  $F$ .

Soit  $F$  une fonction de distribution absolument continue dans l'intervalle  $]x_1, x_F[$ , de densité  $f$ . Alors les conditions suffisantes pour que  $F$  appartient à l'un des trois domaines d'attraction sont données dans le théorème suivant :

#### **Théorème 1.3.9 (Conditions de Von Mises)**

(i) Si  $f$  admet une dérivée  $f'$  négative pour tout  $x \in ]x_1, x_F[$ ,  $f(x) = 0$  pour  $x \geq x_F$  et

$$\lim_{t \rightarrow x_F} \frac{f'(t)(1 - F(t))}{(f(t))^2} = -1,$$

alors  $F \in DA(\Lambda)$ .

(ii) Si  $f(x) > 0$  pour tout  $x \in ]x_1, \infty[$  et pour  $\gamma > 0$

$$\lim_{t \rightarrow \infty} \frac{tf(t)}{1 - F(t)} = \gamma,$$

alors  $F \in D(\Phi_\gamma)$ .

(iii) Si  $f(x) > 0$  pour tout  $x \in ]x_1, x_F[$ ,  $f(x) = 0$  pour tout  $x > x_F$  et pour  $\gamma > 0$

$$\lim_{t \rightarrow x_F^+} \frac{(x_F - t)f(t)}{1 - F(t)} = \gamma,$$

alors  $F \in DA(\Psi_\gamma)$ .

## 1.4 Distributions des dépassements d'un seuil

L'approche basée sur la GEVD a été critiquée dans la mesure où l'utilisation d'un seul maximum conduit à une perte d'information contenue dans les autres grandes valeurs de l'échantillon. Pour surmonter ce problème, la méthode des excès au-delà d'un seuil (Peaks-Over-Threshold, POT) a été introduite par Pickands en (1975), elle consiste à observer les valeurs des réalisations qui excèdent un certain seuil élevé (déterministe). i.e choisit un seuil suffisamment élevé et étudier les excès au-delà de ce seuil (les différences entre ces observations et le seuil appelées excès). Voir la Figure 1.2

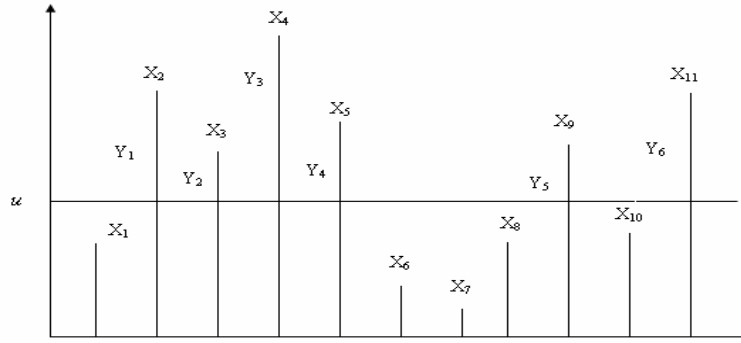


FIG. 1.2 – Observations  $X_1, \dots, X_{11}$  et excès  $Y_1, \dots, Y_6$  au-delà du seuil  $u$ .

**Définition 1.4.1 (Distribution de Pareto généralisée)** *La loi de Pareto généralisée notée GPD (generalized Pareto distribution), pour  $\sigma > 0$ ,  $\gamma \in \mathbb{R}$  est définie par :*

$$G_{\gamma, \mu, \sigma}(x) := \begin{cases} 1 - \left(1 + \gamma \left(\frac{x - \mu}{\sigma}\right)\right)^{-1/\gamma} & \text{si } \gamma \neq 0, \\ 1 - \exp\left(-\frac{x - \mu}{\sigma}\right) & \text{si } \gamma = 0, \end{cases}$$

où

$$\begin{cases} x \geq 0 & \text{si } \gamma \geq 0, \\ x \in [\mu, \mu - \sigma/\gamma] & \text{si } \gamma < 0. \end{cases}$$

Le GPD standard correspond au cas  $\mu = 0$  et  $\sigma = 1$ .

Le GPD avec deux paramètres regroupe trois distributions selon les valeurs du paramètre de forme. Lorsque  $\gamma > 0$ , c'est la loi Pareto, lorsque  $\gamma < 0$ , nous avons la loi Bêta et  $\gamma = 0$  donne la loi exponentielle.

**Théorème 1.4.1 (Balkema-de Haan-Pickands))** *Ce théorème signifie que si  $F$  appartient à l'un des trois domaines d'attraction de la loi limite des extrêmes (Fréchet, Gumbel ou Weibull), alors il existe une fonction de répartition des excès au-delà de  $u$ ,*

noté  $F_u$  qui peut être uniformément approchée par une loi de Pareto généralisée (GPD) telle que :

$$\lim_{u \rightarrow x_F^+} \sup_{0 < y < x_F - u} |F_u(y) - G_{\gamma, \sigma(u)}(y)| = 0. \quad (1.29)$$

où  $\sigma$  est une fonction strictement positive.

Cette considération théorique suggère que, lorsque nous avons des données issues d'une distribution inconnue, il est possible d'approximer la distribution au-delà d'un certain seuil (assez grand) par une distribution de Pareto généralisée.

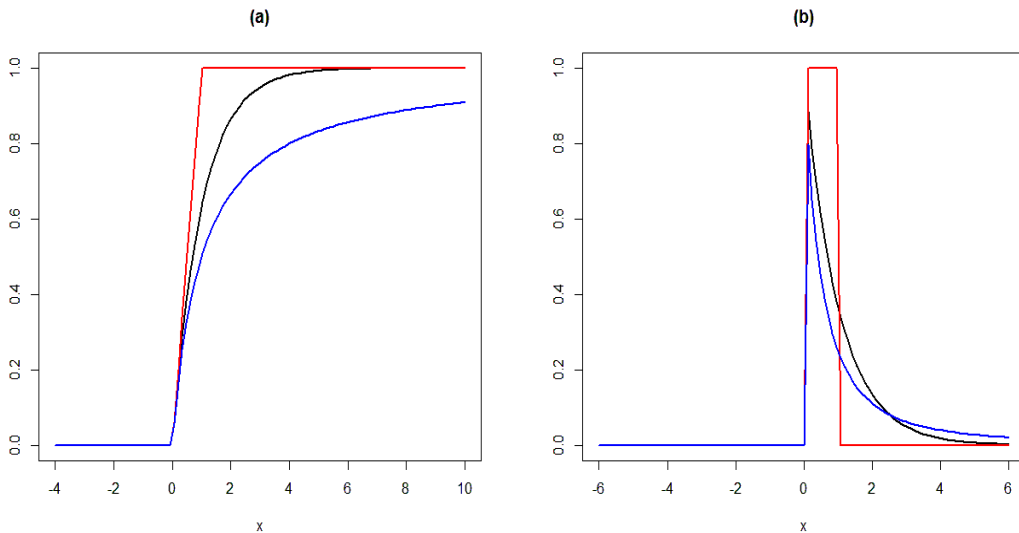


FIG. 1.3 – (a) Densité de la loi de Pareto généralisée standard avec le paramètre de forme 0 (noir), 1 (bleu) et -1 (rouge). (b) : fonction de distribution correspondante.

### 1.4.1 Détermination du seuil

Une variable aléatoire suit une loi GPD est stable par rapport au seuil  $u > 0$ , alors pour  $x > 0$ , on a  $P(X - u \leq x | X > u) = G_{\gamma, \sigma + \gamma(u - \mu)}(x)$  sera noté  $G_{\gamma, \sigma(u)}(x)$ .

- Si  $\mu = 0$  alors  $\sigma(u) = \sigma + \gamma u$ .



- Si  $\mu \neq 0$  alors  $\sigma(u) = \sigma + \gamma(u - \mu)$ .

L'approche des excès au-delà d'un seuil suppose l'existence d'un seuil élevé approprié pour lequel l'approximation du théorème Balkema-de Haan-Pickands est bonne, il faut construire un échantillon adéquat et il faut trouver le seuil  $u$  de sélection des données extrêmes suffisamment grand pour obtenir des estimations précises des paramètres de forme et d'échelle. Le choix du seuil est choisi de façon à faire à un compromis entre le biais et la variance. En effet, une valeur trop élevée plus de biais de l'estimateur est réduit ce qui donne un meilleur modèle et par ailleurs, la variance de l'estimateur est réduit car plus de données participent à l'estimation. Par conséquent, le seuil doit être choisi de manière optimale. Un outil graphique d'utilisation immédiate est disponible pour cette raison. C'est l'ME-plot défini dans le chapitre 2.

# Chapitre 2

## Estimation de l'indice des valeurs extrêmes et de quantiles extrêmes

### 2.1 Introduction

De nombreux auteurs se sont intéressés à l'estimation de l'indice des valeurs extrêmes  $\gamma$  et des quantiles extrêmes pour des lois à queue lourde, voir Hill (1975) [86], Pickands (1975) [116], de Haan et Resnick (1980) [78], Mason (1982) [98], Csörgö et al (1985) [19], Hall et Welsh (1985) [81], Dekkers et al (1989) [39], Davis et Resnick (1989) [31]. Il existe deux approches différentes pour cette estimation le premier c'est l'approche d'EVT, est basés sur la GEVD et le second connu comme approche de POT, utilise la GPD. Dans les deux cas, il existe plusieurs méthodes d'estimation, la plus connue c'est la méthode du maximum de vraisemblance.

Dans ce chapitre, nous nous concentrons sur certains des estimateurs avec leurs propriétés asymptotiques, on donne aussi quelques estimateurs pour des quantiles extrêmes. Mais on commence premièrement par quelques techniques exploratoires pour les extrêmes.

## 2.2 Analyse exploratoire des données

Les outils graphiques jouent un rôle important dans l'analyse statistique. Différents types graphiques sont utilisés dans la plupart des analyses de données. La courbe Quantile-Quantile (QQ) et la courbe Mean Excess (ME) sont parmi les graphiques les plus populaires utilisées dans l'analyse des données à queue lourde i.e. obtenir une indication de la lourdeur des queues. Dans cette section, nous examinerons quelques méthodes pratiques pour étudier les propriétés extrêmes d'un ensemble de données.

### 2.2.1 Probabilité et Quantile Plots

Supposons que  $\hat{F}$  l'estimateur de la fonction de répartition  $F$  a été obtenu. La probabilité probabilité plot (pp-plot) et le quantile quantile plot (qq-plot) peuvent fournir une estimation graphique à la fonction de distribution ajustée  $\hat{F}$ .

**Définition 2.2.1 (PP-plot)** *Le graphe*

$$\left\{ \left( \hat{F}(X_{i,n}), \frac{i}{n+1} \right) : i = 1, \dots, n \right\},$$

*est appelé la probabilité probabilité plot (PP-plot).*

**Définition 2.2.2 (QQ-plot)** *Le graphe*

$$\left\{ \left( X_{i,n}, \hat{F}^{\leftarrow} \left( \frac{i}{n+1} \right) \right) : i = 1, \dots, n \right\},$$

*est appelé le quantile quantile plot (QQ-plot).*

$X_{i,n}$  est le quantile empirique d'ordre  $(i/(n+1))$  de la fonction de distribution  $F$  tandis que  $\hat{F}^{\leftarrow}(i/(n+1))$  est son estimateur.

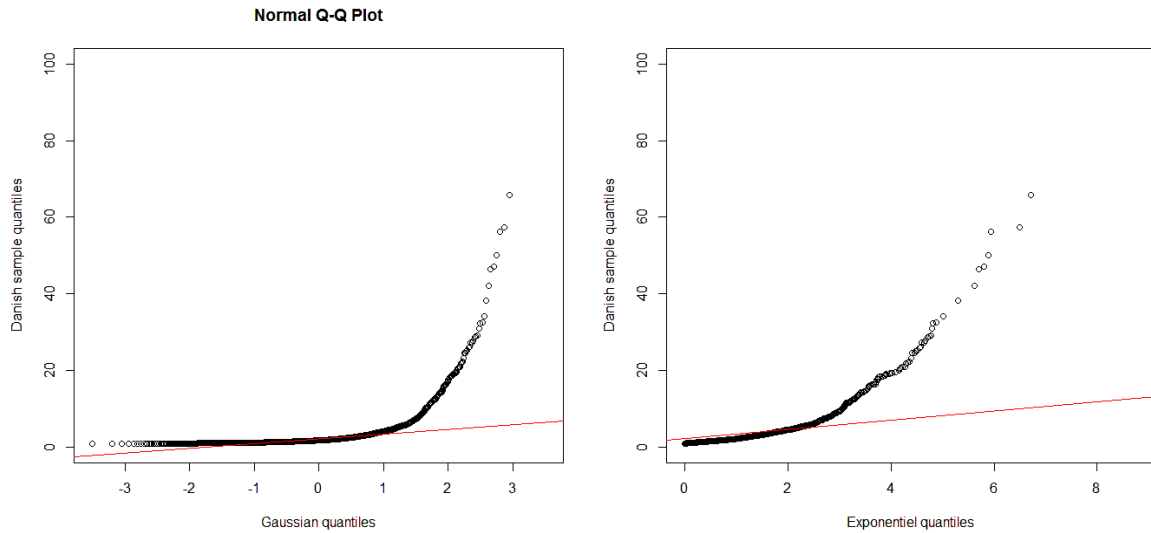


FIG. 2.1 – qq-plot pour les réclamations d'assurance incendie danoises vs (a) : Quantiles exponentiels, (b) Quantiles gaussiennes .

Le qq-plot est un graphique qui oppose les quantiles de la distribution empirique aux quantiles de la distribution théorique envisagée. Si l'échantillon provient bien de cette distribution théorique, alors le qq-plot sera linéaire.

## 2.2.2 Pareto quantile plot

Un outil graphique pour vérifier le comportement des distributions du type Pareto est le graphe "Pareto quantile plot". Comme les variables aléatoires distribuées par Pareto sont transformées logarithmiquement, l'hypothèse d'un comportement strict de Pareto peut être vérifiée en examinant une courbe de quantile exponentielle basée sur les données transformées log, conduisant à un graphe de quantiles de Pareto.

**Définition 2.2.3 (Pareto quantile plot)** *le graphe*

$$\left\{ \log \left( \frac{n+1}{i} \right), \log(X_{n-i+1,n}) : i = 1, \dots, n \right\},$$

*est appelé pareto quantile plot.*

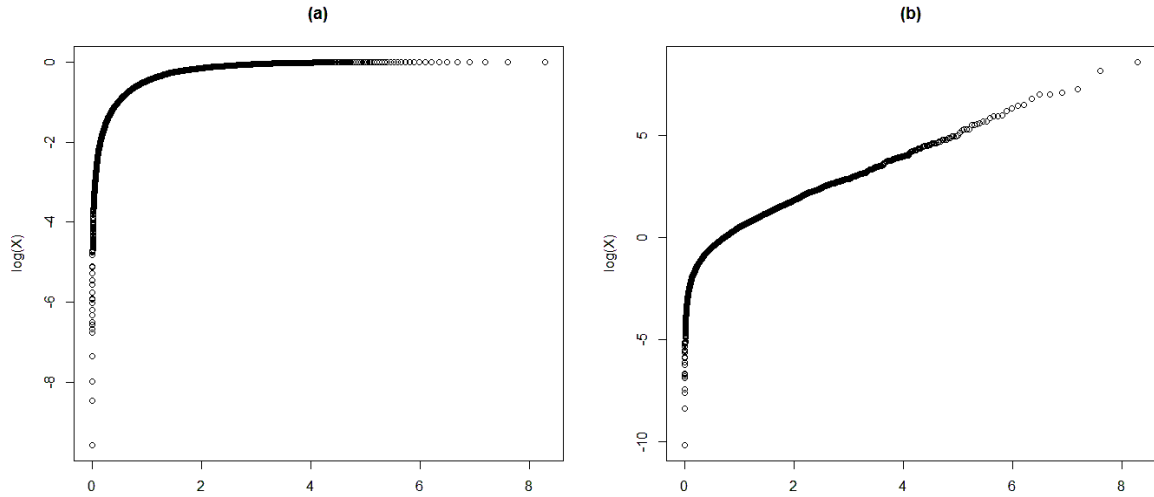


FIG. 2.2 – Pareto-quantile plot de : (a) distribution Uniforme standard et (b) distribution Pareto standard, basé sur 4000 observations.

dans le cas de distributions de type Pareto, ce graphe sera approximativement linéaire de pente  $\gamma$ .

### 2.2.3 Quantile plot généralisé

Une approche pour éviter le choix a priori du domaine d'attraction a été proposée par Beirlant et al (1996). Il consiste à utiliser un "carré de quantile généralisé", elle est définie par :

**Définition 2.2.4 (Quantile plot généralisé)** *le graphe*

$$\left\{ \left( \log \frac{n+1}{j}, \log UH_{j,n} \right) : j = 1, \dots, n \right\},$$

où

$$UH_{j,n} = X_{n-j,n} \left( j^{-1} \sum_{i=1}^j \log X_{n-i+1,n} - \log X_{n-j,n} \right).$$

Suivant le graphe, on peut déduire dans quel domaine d'attraction on se situe : si les points extrêmes on voit apparaître une droite de pente positive alors on est dans le domaine de Fréchet, s'il est constant, alors on est dans le domaine de Gumbel et le cas d'une décroissance linéaire signifie que l'on appartient au domaine de Weibull.

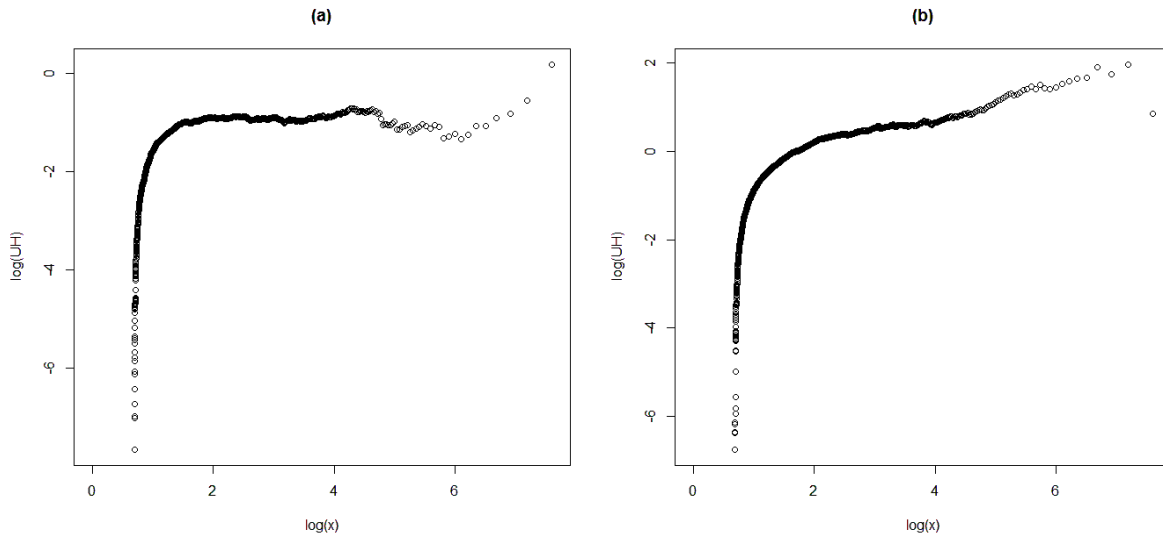


FIG. 2.3 – Quantile plot généralisé de : (a) distribution exponentielle et (b) distribution de Burr, basé sur 4000 observations.

### 2.2.4 Exponentiel quantile plot

L'exponentiel quantile plot est la représentation des quantiles de la distribution empirique sur l'axe des  $X$  contre les quantiles de la fonction de distribution exponentielle sur l'axe des  $Y$ .

Le graphe est l'ensemble des points tel que :

$$\left\{ \left( X_{k,n}, G_{0,1}^{\leftarrow} \left( \frac{n-k+1}{n+1} \right) \right), k = 1, \dots, n \right\},$$

$X_{k,n}$  : Représente le  $k^{\text{ième}}$  ordre statistique et  $G_{0,1}^-$  est la fonction inverse de la distribution exponentielle.

Si  $\gamma = 0$  ce graphe semble être approximativement linéaire.

### 2.2.5 Mean Excess Plot

**Définition 2.2.5 (Fonction moyenne des excès )** *La fonction moyenne des excès  $e$  est l'espérance conditionnelle de  $X - u$  sachant que  $X > u$ , elle a la forme suivante :*

$$e(u) = E[X - u | X > u].$$

**Définition 2.2.6 (Fonction moyenne des excès empirique)** *La fonction moyenne des excès empirique (sample mean excess function)  $e_n$  est la somme des excès au-dessus du seuil  $u$  divisé par le nombre  $N_u$  de données qui excèdent  $u$ .*

$$\hat{e}(u) = e_n(u) := \frac{1}{\bar{F}_n(u)} \int_u^\infty \bar{F}_n(x) dx = \frac{1}{N_u} \sum_{i=1}^n (X_i - u) I_{\{X_i > u\}}, \quad (2.1)$$

avec

$$N_u = \# \{i \in \{1, \dots, n\} : X_i > u\}.$$

Un test graphique pour le comportement de queue peut être effectué par *mean-excess plot*, ce dernier consiste à représenter  $\hat{e}(u)$  en fonction de  $u$ .

**Définition 2.2.7 (ME-plot)** *Le mean excess plot (ME-plot) est donné par le tracé des points suivant*

$$\{(X_{k,n}, e_n(X_{k,n})) : k = 1, \dots, n\}.$$

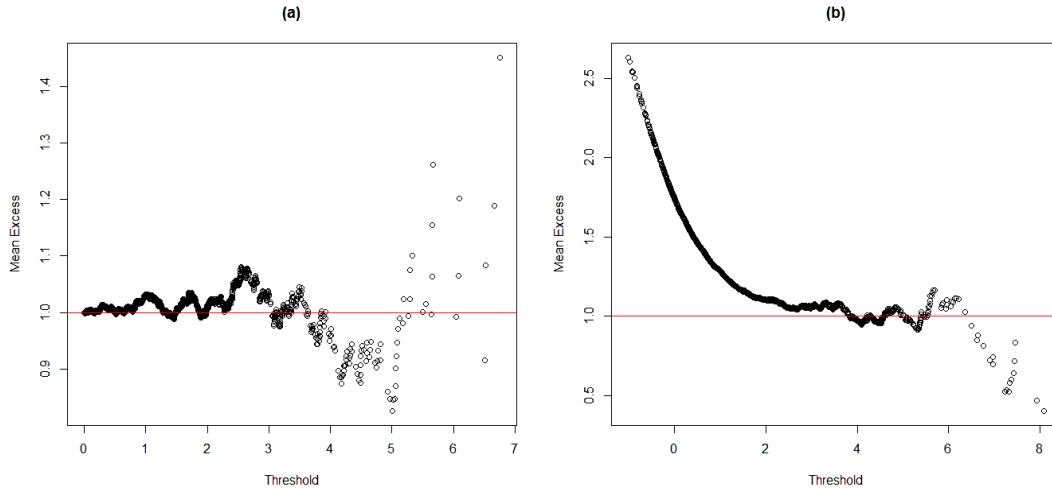


FIG. 2.4 – ME-plot des données simulées à partir d'une (a) : distribution exponentielle basée sur 4000 observations et (b) : de données de Réclamations danoises .

Le graphe ME-plot est méthode graphique utilisé principalement pour distinguer les modèles à queue lourde et les modèles à queue légère, dans le cas de GPD le ME-plot est linéaire voir la figure 2.4.

### 2.3 Estimateurs de l'indice des valeurs extrêmes

La connaissance de l'indice de queue  $\gamma$  est un élément important car il contrôle l'épaisseur (lourdeur) de la queue de distribution. Dans cette section, nous donnons quelques estimateurs les plus connus avec leurs propriétés asymptotiques.



### 2.3.1 Estimateur de Pickands

L'estimateur de Pickands (1975) [116] est le premier estimateur suggéré pour le paramètre  $\gamma \in \mathbb{R}$  et il est donné par la formule suivante

$$\hat{\gamma}_n^{(P)} := (\log 2)^{-1} \log \frac{X_{n-k,n} - X_{n-2k,n}}{X_{n-2k,n} - X_{n-4k,n}}. \quad (2.2)$$

L'auteur démontre la consistance faible de son estimateur. La convergence forte et la normalité asymptotique ont été démontrées par Dekkers et de Haan [40].

**Théorème 2.3.1 (Propriétés asymptotiques de  $\hat{\gamma}_n^{(P)}$ )** *Supposons que  $F \in \mathcal{DA}(H_\gamma)$ ,  $\gamma \in \mathbb{R}$ ,  $k \rightarrow \infty$  et  $k/n \rightarrow 0$  quand  $n \rightarrow \infty$  :*

(a) Consistance faible :

$$\hat{\gamma}_n^{(P)} \xrightarrow{\mathbb{P}} \gamma \quad \text{quand } n \rightarrow \infty.$$

(b) Consistance forte : si  $k/\log \log n \rightarrow \infty$  quand  $n \rightarrow \infty$ , donc

$$\hat{\gamma}_n^{(P)} \xrightarrow{p.s.} \gamma \quad \text{quand } n \rightarrow \infty.$$

(c) Normalité asymptotique : sous certaines conditions sur  $k$  et  $F$  on a :

$$\sqrt{k} \left( \hat{\gamma}_n^{(P)} - \gamma \right) \xrightarrow{d} \mathcal{N}(0, \eta^2), \quad \text{quand } n \rightarrow \infty,$$

où

$$\eta^2 := \frac{\gamma^2(2^{2\gamma+1} + 1)}{(2(2^\gamma - 1) \log 2)^2}.$$

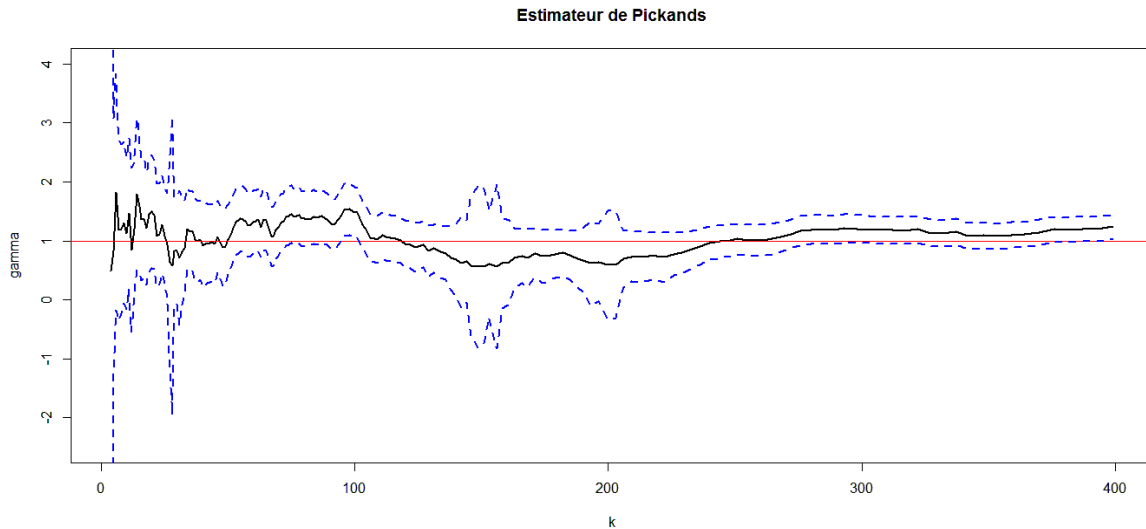


FIG. 2.5 – Estimateur de Pickands, en fonction du nombre des extrêmes (en trait plein) avec l'intervalle de confiance 95% (lignes tirées), pour l'IVE de la distribution Pareto ( $\gamma = 1$ ) basé sur 300 échantillons de 4000 observations.

### 2.3.2 Estimateur de Hill

L'estimateur le plus connu pour l'IVE positif est l'estimateur proposé par Hill (1975) [86] et il est défini par :

$$\hat{\gamma}_n^{(H)} := \frac{1}{k} \sum_{i=1}^k \log X_{n-i+1,n} - \log X_{n-k,n}, \quad (2.3)$$

avec  $X_{1,n}, \dots, X_{n,n}$  les statistiques d'ordre associées à l'échantillon  $X_1, \dots, X_n$ .

Le comportement statistique et les propriétés de l'estimateur de Hill ont été étudiés par de nombreux auteurs. Mason (1982) [98] a démontré la consistance faible et Deheuvels, Haeusler et Mason (1988) [34] ont établi la consistance forte. La normalité asymptotique est due entre autres à Davis et Resnick (1984) [30], Csörgö et Mason (1985) [19], Haeusler et Teugels (1985) [83]. On trouvera l'heuristique de la construction de cet estimateur dans l'article fondateur de Hill [86] ou dans le livre de de Haan et Ferreira (2006) [79].

**Théorème 2.3.2 (Propriétés asymptotiques de  $\hat{\gamma}_n^{(H)}$ )** Supposons que  $F \in \mathcal{DA}(\Phi_\gamma)$ ,  $\gamma > 0$ ,  $1 \leq k < n$ ,  $k \rightarrow \infty$  et  $k/n \rightarrow 0$  quand  $n \rightarrow \infty$  :

(a) *Convergence faible* :

$$\hat{\gamma}_n^{(H)} \xrightarrow{\mathbb{P}} \gamma \quad \text{quand } n \rightarrow \infty.$$

(b) *Convergence forte* : Si  $k/\log \log n \rightarrow \infty$  et  $k_n/n \rightarrow 0$

$$\hat{\gamma}_n^{(H)} \xrightarrow{p.s.} \gamma \quad \text{quand } n \rightarrow \infty.$$

(c) *Normalité asymptotique* : Supposons que  $F$  satisfait la condition (1.8) : Si  $\sqrt{k}A(n/k) \rightarrow \infty$  quand  $n \rightarrow \infty$  alors

$$\sqrt{k}(\hat{\gamma}_n^{(H)} - \gamma) \stackrel{d}{\sim} \mathcal{N}\left(\frac{\lambda}{1-\rho}, \gamma^2\right).$$

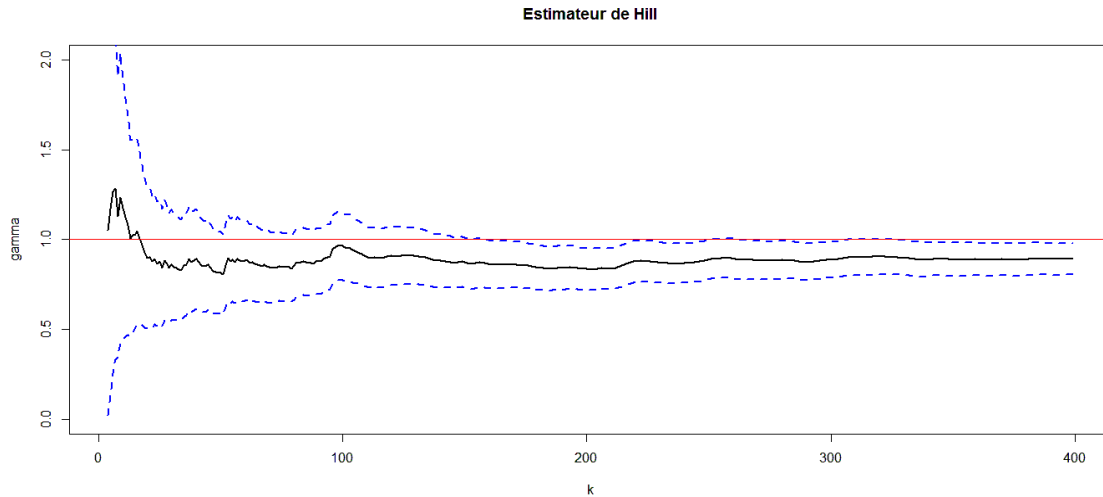


FIG. 2.6 – Estimateur de Hill, en fonction du nombre des extrêmes (en trait plein) avec l'intervalle de confiance 95% (lignes tirées), pour l'IVE de la distribution Frécher ( $\gamma = 1$ ) basé sur 300 échantillons de 4000 observations.

### 2.3.3 Estimateur du moment

Cet estimateur introduit par Dekkers et al (1989) [39] est une généralisation de l'estimateur de Hill, valable pour tout  $\gamma \in \mathbb{R}$  :

$$\hat{\gamma}_n^{(M)} := M_1 + 1 - \frac{1}{2} \left( 1 - \frac{\left( M_n^{(1)} \right)^2}{M_n^{(2)}} \right)^{-1}, \quad (\text{mom})$$

où

$$M_n^{(r)} := \frac{1}{k} \sum_{i=0}^k (\log X_{n-i+1,n} - \log X_{n-k,n})^r; \quad r = 1, 2. \quad (2.4)$$

**Théorème 2.3.3 (Propriétés asymptotiques de  $\hat{\gamma}_n^{(M)}$ )** *Supposons que  $F \in \mathcal{DA}(H_\gamma)$ ,  $\gamma \in \mathbb{R}$ ,  $k \rightarrow \infty$  et  $k/n \rightarrow 0$  quand  $n \rightarrow \infty$  :*

(a) *Consistance faible :*

$$\hat{\gamma}_n^{(M)} \xrightarrow{\mathbb{P}} \gamma \quad \text{quand } n \rightarrow \infty.$$

(b) *Consistance forte : Si  $k/(\log n)^\delta \rightarrow \infty$  quand  $n \rightarrow \infty$  avec  $\delta > 0$ , alors*

$$\hat{\gamma}_n^{(M)} \xrightarrow{p.s.} \gamma \quad \text{quand } n \rightarrow \infty.$$

(c) *Normalité asymptotique : (voir le théorème 3.1 et le corollaire 3.2 de [16])*

$$\sqrt{k} \left( \hat{\gamma}_n^{(M)} - \gamma \right) \xrightarrow{d} \mathcal{N}(0, \eta^2) \quad \text{quand } n \rightarrow \infty,$$

où

$$\eta^2 := \begin{cases} 1 + \gamma^2, & \gamma \geq 0, \\ (1 - \gamma^2)(1 - 2\gamma) \left( 4 - 8 \frac{1 - 2\gamma}{1 - 3\gamma} + \frac{(5 - 11\gamma)(1 - 2\gamma)}{(1 - 3\gamma)(1 - 4\gamma)} \right), & \gamma < 0. \end{cases}$$

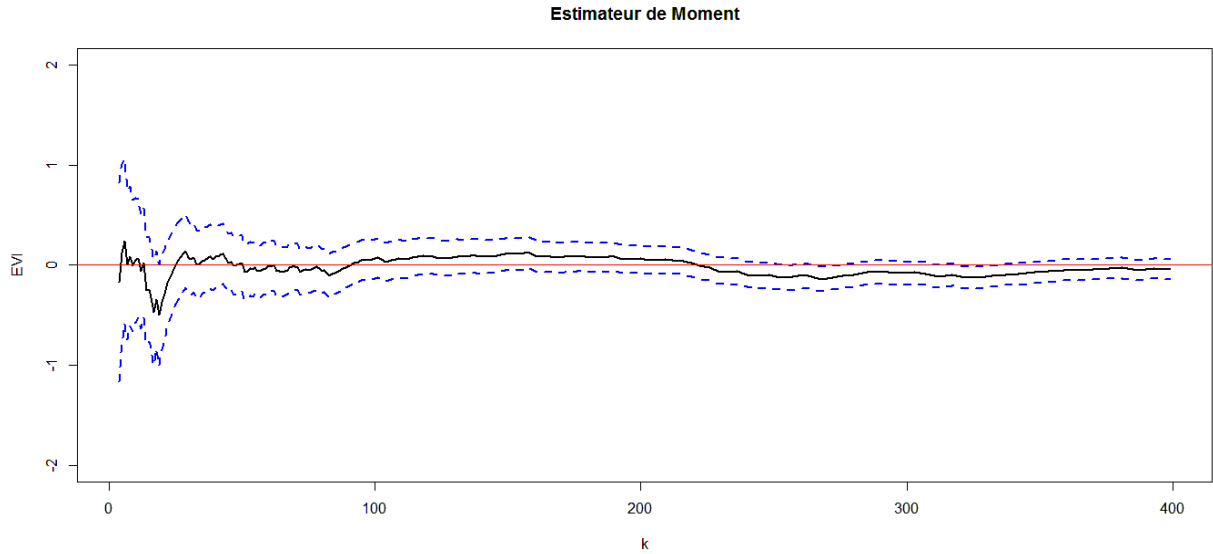


FIG. 2.7 – Estimateur de Moment, en fonction du nombre des extrêmes (en trait plein) avec l'intervalle de confiance 95% (lignes tirées), pour la distribution de Gumbel(0) basée sur 300 échantillons de 4000 observations.

### 2.3.4 Estimateur de type noyau

Csörgö, Deheuvels et Mason en 1985 [19] ont proposé une version plus lisse pour l'estimateur de Hill (dénotée par  $\hat{\gamma}_{n,h}^K$ ) et prouvaient sa consistance et normalité asymptotique. Cet estimateur (CDM) utilisé seulement dans le cas où  $\gamma > 0$ .

$$\hat{\gamma}_{n,h}^K = \sum_{i=1}^{n-1} \left( \frac{i}{nh} \right) K \left( \frac{i}{nh} \right) \log(X_{n-i+1,n}/X_{n-i,n}) / \int_0^{1/h} K(u) du \quad (2.5)$$

tel que  $\alpha > 0$ ,  $h > 0$  et

$$K(u) := h^{-1}K(u/h),$$

avec  $u \in ]0, 1]$  et  $h > 0$  est appelé paramètre de lissage ou fenêtre (*Bandwidth* en Anglais).

Où  $K(\cdot)$  une fonction noyau vérifiant les conditions suivantes :

**(CK1)**  $K(s) \geq 0$  pour  $0 < s < \infty$ .

(CK2) est une fonction non-négative et continue à droit dans  $]0, \infty[$ .

$$(CK3) \int_0^1 K(u) du = 1.$$

$$(CK4) \int_0^1 u^{-1/2} K(u) du < \infty.$$

Dans l'estimateur de type noyau le paramètre de lissage  $h$  joue un rôle semblable comme nombre de statistiques d'ordre  $k$  dans les estimateurs mentionnés auparavant ( $h$  représente la proportion de statistiques de l'ordre supérieur utilisée). On note que l'estimateur de Hill était un cas particulier où la fonction uniforme  $K_h = I_{[0,1]}$  et  $h = k/n$  dans la formule 2.5 Groeneboom et al (2003) [76] ont introduit une modification sur l'estimateur de CDM valable pour tous les indices de queue réels, de sorte que  $F$  est différentiable vérifiant la condition de von Mises.

L'estimateur  $\hat{\gamma}_{n,h}^{(GLW)}$  est défini comme suit pour  $\gamma \in \mathbb{R}$  :

$$\hat{\gamma}_{n,h}^{(GLW)} = \hat{\gamma}_{n,h}^{(CDM)} - 1 + \frac{\hat{q}_{n,h}^{(2)}}{\hat{q}_{n,h}^{(1)}}, \quad (2.6)$$

telle que

$$\hat{\gamma}_{n,h}^{(CDM)} = \sum_{i=1}^{n-1} \frac{i}{n} K_h \left( \frac{i}{n} \right) \log(X_{n-i+1,n}/X_{n-i,n}). \quad (2.7)$$

$$\hat{q}_{n,h}^{(1)} := \sum_{i=1}^{n-1} \left( \frac{i}{n} \right)^\alpha K_h \left( \frac{i}{n} \right) \log(X_{n-i+1,n}/X_{n-i,n}), \quad (2.8)$$

$$\hat{q}_{n,h}^{(2)} = \sum_{i=1}^{n-1} \frac{d}{du} [u^{\alpha+1} K_h(u)]_{u=i/n} (\log(X_{n-i+1,n}/X_{n-i,n})), \quad (2.9)$$

avec  $K_h(u) = K(u/h)/h$  et  $\alpha > 0$ . Ici  $K$  est une fonction noyau nulle dans l'intervalle  $]0, 1[$  et satisfait  $K$ ,  $K'$  et  $K''$  bornés et pour tout  $\alpha > 1/2$ ,  $\int_0^1 u^{\alpha-1} K(u) du \neq 0$ .

Les propriétés de la consistance et la normalité asymptotique de l'estimateur  $\hat{\gamma}_n^{(GLW)}$  ont été établies dans [76]. Plus récemment, Necir [106] a proposé une loi fonctionnelle du logarithme itéré pour cet estimateur et a prouvé sa consistance forte.

## 2.4 Choix du nombre optimal de statistiques d'ordre extrêmes

L'objectif principal de la statistique des extrêmes est la prédiction des événements rares et le problème principal c'est que les estimateurs de l'indice des valeurs extrêmes sont basés essentiellement sur le nombre de plus grandes observations de l'échantillon  $k$  ou sur les excès sur un niveau élevé  $u$ , le problème concernant le choix de ce nombre n'est pas facile à manipuler.

Sélectionner une bonne valeur de  $k$  est une tâche sensible. Lorsque  $k$  est petit la variance de l'estimateur est grande et l'utilisation de  $k$  introduit un grand biais. L'équilibrage de ces composants (la variance et le biais) est important dans les applications de la théorie des valeurs extrêmes, parce que cela réduit l'erreur moyenne quadratique.

nous avons introduit certaines des méthodes proposées pour choisir ce nombre optimal de statistiques d'ordre  $k_{opt}$ . Certaines de ces procédures seront examinées. Malgré le fait que les méthodes proposées peuvent être appliquées (voir chapitre 4).

Le processus de choisir le  $k$  optimal est difficile par le fait que le dernier ne dépend pas exclusivement de la taille de l'échantillon et de l'IVE mais il dépend également des paramètres inconnus caractérisant  $F$  (le paramètre de deuxième ordre  $\rho$  parmi autres de la fonction de distribution  $F$ ). Pour surmonter cet obstacle, une grande variété d'algorithmes et de procédures adaptatives ont proposé de calculer  $\hat{k}_{opt}$  pour  $k_{opt}$  dans le sens

$$\frac{\hat{k}_{opt}}{k_{opt}} \xrightarrow{p} 1 \quad \text{quand } n \rightarrow \infty. \quad (2.10)$$

L'estimation correspondante  $\hat{\gamma}_n(\hat{k}_{opt})$  est aussi asymptotiquement efficace que  $\hat{\gamma}_n(k_{opt})$ .

### 2.4.1 Approche de Hall et Welsh

Hall et Welsh (1985) [81] ont prouvé que l'erreur moyenne quadratique asymptotique de l'estimateur de Hill est minimale pour

$$k_{opt} \sim \left( \frac{c^{2\rho} (\rho + 1)^2}{2d^2 \rho^3} \right)^{1/(2\rho+1)} n^{2\rho/(2\rho+1)}. \quad (2.11)$$

Si la fonction de répartition  $F$  satisfaisant la classe de Hall 1.24.

Comme les paramètres  $\rho > 0$ ,  $c > 0$  et  $d \neq 0$  sont inconnus, ce résultat ne peut pas être utilisé directement pour déterminer  $k_{opt}$ .

L'erreur moyenne quadratique MSE est liée par le paramètre du second ordre  $\rho$  donc on ne peut pas être utilisé dans les questions pratiques. Pour cette raison, nous présentons la méthode bootstrap.

### 2.4.2 Approche de Bootstrap

En 1990 Hall [82] a utilisé la méthode du bootstrap dans l'estimation de l'indice de queue pour le nombre de statistiques d'ordre extrêmes. Son fondement théorique est minimisé MSE de l'IVE.

Hall<sup>1</sup> a précisé la taille  $n_1$  de l'échantillon bootstrap qui doit être inférieure à la taille de l'échantillon total  $n$ . Et il avait prouvé que l'estimation n'était pas fiable quand  $n_1 = n$ .

Nous avons besoin de retirer un sous-échantillon  $\mathcal{X}_{n_1}^* = \{X_1^*, \dots, X_{n_1}^*\}$  du l'échantillon total  $\mathcal{X}_n = \{X_1, \dots, X_n\}$  ( $n_1 \ll n$ ) et  $\mathcal{X}_{n_1}^*$  est appelé sous-échantillon bootstrap. Nous avons utilisé  $X_{1,n_1} \leq X_{2,n_1} \leq \dots \leq X_{n_1,n_1}$  les statistiques d'ordre de  $\mathcal{X}_{n_1}^*$  et nous avons défini :

$$\gamma_{n_1}^*(k_1) := \frac{1}{k_1} \sum_{i=1}^{k_1} \log X_{n_1-i+1,n_1}^* - \log X_{n_1-k_1,n_1}^*.$$

Nous déterminons  $k$  et  $k_1$  en minimisant

---

<sup>1</sup>Hall (1990) utilise également la même idée de choisir le paramètre de lissage dans les procédures d'estimation de type noyau.



$$MSE(n_1, k_1) = E((\gamma_{n_1}^*(k_1) - \gamma_n(k))^2 | \mathcal{X}_n). \quad (2.12)$$

Et la relation est  $k = k_1 \left(\frac{n}{n_1}\right)^\mu$ . En fait, Hall avait supposé la relation de pouvoir de  $k$  et  $n$  était  $k = cn^\mu$ , ( $0 < \mu < 1$ ) et le rapport de pouvoir  $n_1$  et  $n$  était  $n_1 = n^\beta$ . Il a suggéré que  $\mu$  et  $\beta$  étaient  $1/2$ . Cette méthode dépend de  $\gamma$  et  $\rho$  et MSE est liée à  $k$ .

### Méthode du Bootstrap proposée par Danielsson

Danielsson et al (1997) [27] ont amélioré la méthode bootstrap qui est proposée par Hall. Ils ont proposé une nouvelle méthode que l'estimateur de Hill  $\hat{\gamma}_n^{(H)}(k)$  a été remplacée par une nouvelle estimateur  $M_n(k)$ , où

$$M_n(k) = \frac{1}{k} \sum_{i=1}^k (\log X_{i,n} - \log X_{k+1,n})^2. \quad (2.13)$$

Nous avons prouvé  $M_n(k)/2\hat{\gamma}_n(k)$  converge en probabilité vers  $\gamma$  quand  $k \rightarrow \infty$ . Il a aussi équilibré la variance et le biais.

Les statistiques  $M_n(k)/2\hat{\gamma}_n(k) - \gamma_n(k)$  et  $\hat{\gamma}_n(k) - \gamma$  ont la même propriété asymptotique que leur moyenne asymptotique.. Sous quelques situations, il peut recevoir la même valeur de  $k$  pour minimiser  $MSE$  et  $E_\infty(M_n(k) - 2(\hat{\gamma}_n(k))^2)^2$ .

Selon le sous-échantillon bootstrap  $\mathcal{X}_{n_1}^*$ , Nous sélectionnons les statistiques pour déterminer  $k_1$  en minimisant  $Q(n_1, k_1)$ .

$$Q(n_1, k_1) = E((M_{n_1}^*(k_1) - 2(\gamma_{n_1}^*(k_1))^2)^2 | \mathcal{X}_n). \quad (2.14)$$

Afin de déterminer  $k$  nous avons besoin d'un autre sous-échantillon bootstrap  $\mathcal{X}_{n_2}^*$  où ( $n_2 = n_1^2/n$ ) pour déterminer  $k_2$  par la même méthode de détermination  $k_1$ . Ensuite, nous pouvons calculer  $k$  avec le rapport de  $k_1$  et  $k_2$ .

Nous expliquons, étape par étape, comment mettre en œuvre la procédure :

**Etape 1 :** Sélectionner un échantillon de taille  $n$  à partir d'une série de rapport de retour comme échantillon naissant et arranger par convenable.

**Etape 2 :** Choisir d'abord  $n_1$  et  $k_1$ . Echantillons aléatoires de taille  $n_1$ ,  $m$  fois de la taille naissante d'échantillon  $n_1$  ( $n_1 = O(n^{1-\varepsilon})$ ,  $0 \leq \varepsilon \leq 1$ ) et déterminer  $n_1$  et  $k_1$  en réduisant au minimum avec la méthode du bootstrap

$$Q(n_1, k_1) = E((M_{n_1}(k_1) - 2(\gamma_{n_1}(k_1))^2) \setminus \mathcal{X}_n) = \frac{1}{m} \sum_{i=1}^m (M_{n_1}^*(k_1) - 2(\gamma_{n_1}^*(k_1))^2)^2.$$

**Etape 3 :** Sélectionner  $k_2$ . Nous supposons  $n_2 = n_1^2/n$  et répéter l'étape (2) pour déterminer  $k_2$ .

**Etape 4 :** Déterminer  $\hat{k}$ . Nous pouvons calculer  $\hat{k}_{opt}$  avec

$$\hat{k}_{opt} = \frac{k_1^2}{k_2} \left( \frac{(\log k_1)^2}{(2 \log n_1 - \log k_1)^2} \right)^{(\log n_1 - \log k_1(n_1)) / \log(n_1)}. \quad (2.15)$$

### 2.4.3 Approche séquentielle

Drees et Kaufmaan (1998) [47] ont proposé une approche séquentielle pour construire un estimateur consistant de  $k_{opt}$  sans aucune connaissance préalable sur la fonction de distribution  $F$ .

Le  $\hat{k}_{opt}$  proposé est défini sur la base des temps d'arrêt suivant  $\tilde{k}(r_n)$  et  $\tilde{k}(r_n^\zeta)$  avec  $0 < \zeta < 1$  pour une suite d'estimateur de Hill par :

$$\tilde{k}(r_n) := \min\{k \in \{2, \dots, n\} : \max_{2 \leq i \leq k} i^{1/2} \left| \hat{\gamma}_{i,n}^{(H)} - \hat{\gamma}_{k,n}^{(H)} \right| > r_n\}, \quad (2.16)$$

où les seuils  $r_n$  constituent une suite qui est de plus grand ordre que  $(\log \log n)^{1/2}$  mais de plus petit ordre que  $n^{1/2}$ .

Drees et Kaufmann ont mentionné dans leurs résultats de simulation, que pour beaucoup de distributions, les meilleurs résultats sont obtenus si  $\rho$  est fixé (généralement on prend

$\rho_0$  égale  $-1$ ) au lieu d'être estimé.

Pour fixer  $\rho_0 = -1$ , la méthode de Drees et Kaufmann peut être résumée dans les trois étapes suivantes :

**Étape 1 :** Poser  $r_n = 2.5 \times n^{0.25} \times \tilde{\gamma}_n^{(H)}$ , avec  $\tilde{\gamma}_n^{(H)} := \hat{\gamma}_{2\sqrt{n},n}^{(H)}$ .

**Étape 2 :** Obtenir  $\tilde{k}(r_n)$ . Si la condition  $\max_{i < k} i^{1/2} \left| \hat{\gamma}_{i,n}^{(H)} - \hat{\gamma}_{k,n}^{(H)} \right| > r_n$  est satisfait pour certains  $k$  tel que  $k_n(r_n)$  bien définis, puis passer à l'étape 3. Sinon, assigner  $0.9 \times r_n$  à  $r_n$  et répéter l'étape 2.

**Étape 3 :** Pour  $\xi \in (0, 1)$  (en particulier  $\xi = 0.7$ ), déterminer

$$\hat{k}_{opt} := \left[ \frac{1}{3} \left( 2 \left( \tilde{\gamma}_n^{(H)} \right)^2 \right)^{1/3} \left( \frac{\tilde{k}(r_n^\xi)}{\left( \tilde{k}(r_n) \right)^\xi} \right)^{1/(1-\xi)} \right],$$

où  $[x]$  dénote le plus grand entier inférieur ou égal à  $x$ .

L'avantage découvert de la méthode de Drees et Kaufmann (1998) est que le paramètre du second ordre  $\rho$  peut être fixé à une valeur fixe  $\rho_0$ . Le théorème 1 de Drees et Kaufmann (1998) ont assuré que  $\hat{k}_{opt}$  est un estimateur convergent de  $k_{opt}$  si la fonction de distribution sous-jacente satisfait la condition Hall. Malheureusement,  $\hat{k}_{opt}$  n'est pas toujours bien défini, car la procédure peut revenir  $\hat{k}_{opt} < 2$ .

#### 2.4.4 Choix automatique de paramètre de lissage

Pour l'estimateur de type noyau, le problème devient une question de choix de paramètre de lissage optimale  $h_{opt}$  pour laquelle une estimation consistante  $\hat{h}_{opt}$  doit être calculé de manière adaptative. Groeneboom et al [76] fondent leurs estimateurs  $\hat{\gamma}_n^{(W1)}$  et  $\hat{\gamma}_n^{(W2)}$  sur les noyaux suivants :

$$K_1(u) = \frac{15}{8}(1 - u^2)^2 I_{[0,1]}(u)$$

et

$$K_2(u) = \frac{35}{16}(1 - u^2)^3 I_{[0,1]}(u).$$

L'algorithme de cinq étapes menant au  $\hat{h}_{opt}$  est comme suit :

**Etape 1 :** Sélectionner un échantillon de bootstrap  $(X_1^*, \dots, X_{n_1}^*)$  de taille  $(n_1 \ll n)$  de l'échantillon initial  $(X_1, \dots, X_n)$  et calculer  $\hat{\gamma}_{n_1, h}^{(K_1^*)}$  et  $\hat{\gamma}_{n_1, h}^{(K_2^*)}$  en termes des statistiques d'ordre  $X_{1, n_1}^* \leq \dots \leq X_{n_1, n_1}^*$  concernant l'échantillon bootstrap. Ensuite calculer

$$\delta_{n_1, h}^* := \gamma_{n_1, h}^{(W_1^*)} - \gamma_{n_1, h}^{(W_2^*)}.$$

**Etape 2 :** Répéter l'étape 1,  $r$  fois de façon indépendante. Calculer

$$\widehat{MSE}^*(\delta_{n_1, h}^*) := \frac{1}{r} \sum_{i=1}^r (\delta_{n_1, h, i}^*)^2.$$

qu'est l'estimation de MSE du bootstrap de  $\delta_{n_1, h}^*$ .

**Etape 3 :** Calculer

$$h^*(n_1) := \arg \min_h \widehat{MSE}^*(\delta_{n_1, h}^*)$$

Dans la pratique, on pourrait calculer  $\widehat{MSE}^*(\delta_{n_1, h}^*)$  sur une grille des valeurs du  $h_i$ , et prendre pour  $h^*(n_1)$  minimiser les valeurs de  $\widehat{MSE}^*(\delta_{n_1, h_i}^*)$ .

**Etape 4 :** Répéter les étapes 1-3 indépendamment avec  $n_2 = [n_1^2/n]$  au lieu  $n_1$  : soit

$$h^*(n_2) := \arg \min_h \widehat{MSE}^*(\delta_{n_2, h}^*).$$

**Etape 5 :** Estimer la fenêtre optimale  $h_{opt}$  par

$$\hat{h}_{opt} := c(h^*(n_1), h^*(n_2)) \frac{(h^*(n_1))^2}{h^*(n_2)},$$

où  $c(h_1, h_2)$  est une fonction de  $h_1$  et  $h_2$  selon les noyaux  $K_1$  et  $K_2$  et les tailles de l'échantillon  $n_1$  et  $n_2$  respectivement.

**Remarque 2.4.1** *Il existe des autres méthodes proposées pour la détermination du nombre de statistiques d'ordre extrêmes définies et appliquées dans le chapitre 4.*

## 2.5 Estimation des quantiles extrêmes

Dans cette section, nous discutons sur l'estimation des quantiles extrêmes qui joue un rôle important dans nombreuses applications comme l'assurance, la finance, l'hydrologie la gestion des risques où il est crucial d'évaluer de manière adéquate le risque d'une grande perte qui se produit très rarement.

On considère  $X_1, X_2, \dots, X_n$  la réalisation de  $n$  variables aléatoires réelles indépendantes et de fonction de répartition commune  $F$  supposée continue.

Les estimateurs des quantiles extrêmes pour les distributions à queue lourde, i.e. des quantiles d'ordre  $(1 - p)$  noté  $x_p$  avec  $0 < p < 1$  et  $p$  tendant à zéro lorsque la taille de l'échantillon augmente, donc  $x_p$  de la fonction de distribution  $F$  est défini comme étant la solution de l'équation :

$$1 - F(x) = p.$$

En utilisant les fonctions présentées dans les définitions 1.2 et 1.3, on définit les quantiles extrêmes par :

$$x_p = F^{\leftarrow}(1 - p) = Q(1 - p) = U(1/p), \quad \text{quand } p \rightarrow 0. \quad (2.17)$$

La difficulté principale dans l'estimation de quantile extrême est due au fait que lorsque  $p$  est très petit proche de 0, le point  $x_p$  est strictement supérieur à l'observation maximale de

l'échantillon. On ne peut pas, comme pour l'estimation de quantiles "classiques", inverser tout simplement la fonction de répartition empirique. Comme nous utilisons la théorie asymptotique,  $p$  doit dépendre de la taille de l'échantillon  $n$  i.e.  $p = p_n$ .

On recherche un estimateur de  $x_p$  lorsque  $np_n$  admet une limite  $c \in [0, 1]$  quand  $n$  tend vers l'infini, on a deux cas possibles pour  $x_p$ , à l'intérieur et à l'extérieur de l'échantillon. L'estimation est à l'intérieur de l'échantillon si  $c > 1$ , et à l'extérieur si  $c < 1$ .

Pour la première situation, nous appliquons le résultat 1.6 avec  $s = 1-p = 1-(i-1)/n$  pour  $i = 2, \dots, n$  et nous obtenons

$$Q_n \left( 1 - \frac{i-1}{n} \right) = X_{n-i+1, n}.$$

Dans le second cas, nous devons déduire au-delà des limites de l'échantillon en extrapolant à partir de quantiles intermédiaires. Évidemment, cela ne peut se faire sans un certain type d'information sur les queues de la distribution. Une modélisation précise des queues de distribution est alors nécessaire. L'estimation des quantiles extrêmes est directement liée à l'estimation de l'IVE, on pourrait s'attendre à trouver dans la littérature autant d'estimateurs de quantile que d'estimateurs d'indice de queue.

Les intervalles de confiance pour les estimations de quantile sont facilement construits puisque les estimateurs proposés sont asymptotiquement normaux et les points d'extrémité au cas où ils sont finis sont estimés comme des quantiles d'ordre 1.

Il existe plusieurs approches pour estimer les quantiles extrêmes. Celle basée sur la théorie des valeurs extrêmes (Gnedenko, 1943) est donnée par :

$$\hat{x}_p = \hat{b}_n + \hat{a}_n \frac{(np)^{-\hat{\gamma}} - 1}{\hat{\gamma}}.$$

Par la méthode des excès et l'approximation de la loi GPD (théorème Pickands-Belkema-de Haan) à savoir :

$$\hat{x}_p = \left( \frac{k}{np} \right)^{\hat{\gamma}} \frac{\hat{\sigma}}{\hat{\gamma}} + u.$$

Nous définissons maintenant les grands estimateurs de quantile qui sont associées aux estimateurs semi-paramétriques de l'IVE présenté dans la Section 2.3.

Weissman (1978) [132] a donné l'estimation des quantiles extrêmes pour chacune des trois distributions standard des valeurs extrêmes de Fisher-Tippett. Pour la classe de Fréchet ( $\gamma > 0$ ), l'estimateur de quantile d'ordre  $(1-p)$  du type Weissman prend la forme suivante :

$$\hat{x}_p^{(W)} := X_{n-k,n} \left( \frac{k}{np} \right)^{\hat{\gamma}_n}, \quad (2.18)$$

- L'estimateur classique de quantile de Weissman pour l'estimateur de Hill  $\hat{\gamma}_n^{(H)}$  est défini par :

$$\hat{x}_p^{(H)} := X_{n-k,n} \left( \frac{k}{np} \right)^{\hat{\gamma}_n^{(H)}},$$

- L'estimateur  $\hat{x}_p^{(P)}$  du quantile d'ordre  $(1-p)$  lié au l'estimateur de Pickands  $\hat{x}_n^{(P)}$  est de la forme suivante :

$$\hat{x}_p^{(P)} := X_{n-k+1,n} + \frac{(np/k)^{-\hat{\gamma}_n^{(P)}} - 1}{1 - 2^{-\hat{\gamma}_n^{(P)}}} (X_{n-k+1,n} - X_{n-2k+1,n}).$$

- De la même façon, le quantile d'ordre  $(1-p)$  est estimé sur la base de l'estimateur du moment  $\hat{x}_p^{(M)}$  par

$$\hat{x}_p^{(M)} := X_{n-k,n} + \left( \frac{X_{n-k,n} M_n^{(1)}}{\varphi(\hat{\gamma}_n^{(M)})} \right) \frac{(np/k)^{-\hat{\gamma}_n^{(M)}} - 1}{\hat{\gamma}_n^{(M)}},$$

où

$$\varphi(\gamma) := \begin{cases} 1, & \gamma \geq 0, \\ 1/(1-\gamma), & \gamma > 0. \end{cases}$$

# Chapitre 3

## Copules et dépendance

### 3.1 Introduction

Le mot "copule" vient du mot latin "copūlae" qui désigne le lien ou la connexion entre les parties. Ce mot a été adopté dans la statistique pour désigner une classe de fonctions permettant de construire et de modéliser les distributions multivariées de manière simple. Bien que les termes corrélation et dépendance soient souvent utilisés.

Le problème de la modélisation des structures de dépendance est que cette caractéristique ne ressort pas toujours de la fonction de répartition conjointe considérée. Il serait utile de séparer les propriétés statistiques de chaque variable de leur structure de dépendance.

Ce chapitre présente le concept de copula, qui s'avérera utile pour l'analyse de la dépendance de queue. Il est organisé comme suit. La section 3.1 introduit la notion de copule avec des définitions de base et un théorème le plus important concernant copule est indiqué : le théorème de Sklar (1959), qui illustre en outre l'origine du mot copula. La section 3.2 présente les familles importantes de copules avec leurs propriétés et dans la section 3.3 nous donnons les mesures de dépendance et examinons leur lien avec les copules.



## 3.2 Définitions et propriétés de base

La notion de copule a été présentée par A. Sklar en 1959 [127], au moment de répondre à une question soulevée par M. Fréchet sur les liens entre une fonction de probabilité conjointe multidimensionnelle et ses marges. Au début, les copules ont été principalement utilisées dans le développement de la théorie probabiliste des espaces métriques. Plus tard, elles ont été utilisées à définir des mesures non paramétriques de dépendance entre variables aléatoires, elles ont commencé à jouer un rôle important dans la probabilité et statistique.

Tout d'abord, nous donnons la définition de la distribution générale multivariée et ensuite on va étudier les propriétés spéciales de la copule.

**Définition 3.2.1 (fonction de répartition conjointe)** *Soit un vecteur aléatoire  $X = (X_1, \dots, X_d)$  de dimension  $d$ . La fonction de répartition conjointe de  $X$  est la fonction de  $d$ -variables définie par :*

$$H(x_1, \dots, x_d) := P(X_1 \leq x_1, \dots, X_d \leq x_d). \quad (3.1)$$

1.  $H$  est continue à droite,
2.  $\lim_{x_i \rightarrow -\infty} H(x_1, \dots, x_d) = 0$ , pour  $i = 1, \dots, d$
3.  $\lim_{x_i \rightarrow +\infty} H(x_1, \dots, x_d) = 1$
4. pour tout  $(a_1, \dots, a_d), (b_1, \dots, b_d) \in [0, 1]^d$  avec  $a_i \leq b_i$  nous avons :

$$\sum_{i_1=1}^2 \sum_{i_2=1}^2 (-1)^{i_1+\dots+i_d} H(x_{1_{i_1}}, \dots, x_{d_{i_d}}) \geq 0,$$

où  $x_{j1} = a_j$  et  $x_{j2} = b_j$  pour tout  $j$  et  $x_{jk} \in [0, 1]$  pour tout  $j$  et  $k$ .

Si  $H$  a des dérivées d'ordre  $d$ , alors cette condition est équivalente à

$$\frac{\partial^d H}{\partial x_1 \dots \partial x_d} \geq 0.$$

Par l'utilisation d'une transformation intégrale de probabilité. Séparer la fonction de distribution conjointe  $H$  de ses fonctions de distribution marginales et une partie qui décrit la structure de dépendance. Cela est établi par

$$\begin{aligned} H(x_1, \dots, x_d) &= P[X_1 \leq x_1, \dots, X_d \leq x_d] \\ &= P[F_1(X_1) \leq F_1(x_1), \dots, F_d(X_d) \leq F_d(x_d)] \\ &= C(F(x_1), \dots, F(x_d)), \end{aligned}$$

Les fonctions de copule deviennent des outils populaires et flexibles pour modéliser la dépendance multivariée entre les risques. En général, une copule est une distribution multivariée  $d$ -dimensionnelle sur  $[0, 1]^d$  avec des distributions marginales uniformes.

**Définition 3.2.2 (Copula)** *Une copule  $d$ -dimensionnelle  $C : [0, 1]^d \rightarrow [0, 1]$  est une fonction de répartition multivariée ayant des lois marginales uniforme sur  $[0, 1]$*

$$C(u_1, \dots, u_d) = P(U_1 \leq u_1, \dots, U_d \leq u_d).$$

**Définition 3.2.3** *Une copule  $d$ -dimensionnelle est une fonction de  $[0, 1]^d$  dans  $[0, 1]$  ayant les propriétés suivantes :*

- (i) Pour tout  $u \in [0, 1]^d$ , on a :  $C(u) = 0$ , si au moins une coordonnée de  $u$  est égale à 0,
- (ii) Pour tout  $u_i \in [0, 1]$  et pour  $i = 1, \dots, d$ , on a  $C(1, \dots, 1, u_i, 1, \dots, 1) = u_i$ ,
- (iii)  $C$  est  $d$ -croissante, i.e. pour tout  $(a_1, \dots, a_d), (b_1, \dots, b_d) \in [0, 1]^d$  avec  $a_i \leq b_i$  on a :

$$\sum_{i_1=1}^2 \sum_{i_2=1}^2 \dots \sum_{i_d=1}^2 (-1)^{i_1 + \dots + i_d} C(u_{1_{i_1}}, \dots, u_{d_{i_d}}) \geq 0,$$

où  $u_{j1} = a_j$  et  $u_{jk} = b_j$  pour tout  $j$  et  $u_{jk} \in [0, 1]$  pour tout  $j$  et  $k$ .

**Proposition 3.2.1 (Continuité uniforme)** *Une copule  $C$  est uniformément continue sur son domaine. En particulier, pour tout  $u$  et  $v$  dans  $[0, 1]^d$ , on a :*

$$|C(u_1, \dots, u_d) - C(v_1, \dots, v_d)| \leq \sum_{i=1}^d |u_i - v_i|.$$

**Proposition 3.2.2** *Soit  $C$  une  $d$ -copule. Pour presque tous  $(u_1, \dots, u_d) \in [0, 1]^d$ , la dérivée partielle de  $C$  par rapport à  $u_i$  existe et :*

$$0 \leq \frac{\partial C}{\partial u_i}(u_1, \dots, u_d) \leq 1.$$

### 3.3 Théorème de Sklar

Le théorème bien connu de Sklar publié en 1959, est un outil de base de la théorie des copules et c'est le fondement de la plupart, des applications de cette théorie aux statistiques. Il dit qu'on peut toujours trouver une copule unique  $C(\cdot)$  pour une distribution conjointe avec des marginales continues données. Ce théorème montre le rôle que jouent les copules dans la relation entre les fonctions de distribution multivariée et leurs fonctions marginales univariées.

**Théorème 3.3.1 (Théorème de Sklar)** *Soit  $F$  une fonction de répartition  $d$ -dimensionnelle de fonctions de répartition marginales  $F_1, \dots, F_d$  alors il existe une copule  $C$  telle que, pour tout  $x \in \mathbb{R}^d$  ;*

$$H(x_1, \dots, x_d) = C(F_1(x_1), \dots, F_d(x_d)). \quad (3.2)$$

*Si les fonctions  $F_1, \dots, F_d$  sont continues, alors  $C$  est unique. Réciproquement, si  $C$  est une copule et  $F_1, \dots, F_d$  sont les fonctions de répartition, alors la fonction  $F$  définie par l'équation 3.2 est une fonction de répartition conjointe des marges  $F_1, \dots, F_d$ .*

**Corollaire 3.3.1** *Si  $F$  est une fonction de répartition conjointe de marges  $F_1, \dots, F_d$  de fonctions inverses respectivement  $F_1^{-1}, \dots, F_d^{-1}$ , alors pour tout  $(u_1, \dots, u_d)$  appartient au domaine de  $C$  on a :*

$$C(u_1, \dots, u_d) = H(F_1^{-1}(u_1), \dots, F_d^{-1}(u_d)).$$

Si  $X \sim F$  et  $F$  est continue alors  $(F_1(x_1), \dots, F_d(x_d)) \sim C$  et si  $U \sim C$  alors :

$$(F_1^{-1}(u_1), \dots, F_d^{-1}(u_d)) \sim F.$$

Ce résultat est une conséquence directe du théorème et il est important pour la construction de copule de distributions multivariées.

**Corollaire 3.3.2 (Densité de copule)** *La relation entre la fonction de densité multivariée  $f(x_1, \dots, x_d)$  et la densité de copule  $c$ , est donnée par :*

$$f(x_1, \dots, x_d) = c(F_1(x_1), \dots, F_d(x_d)) \prod_{i=1}^d f_i(x_i), \quad (3.3)$$

où  $f_1(x_1), \dots, f_d(x_d)$  sont les densités marginales.

**Définition 3.3.1 (Copule d'indépendance)** *la fonction  $\Pi : [0, 1]^d \rightarrow [0, 1]$  donnée par :*

$$\Pi(u_1, \dots, u_d) = u_1 \times \dots \times u_d. \quad (3.4)$$

*est dite copule d'indépendance.*

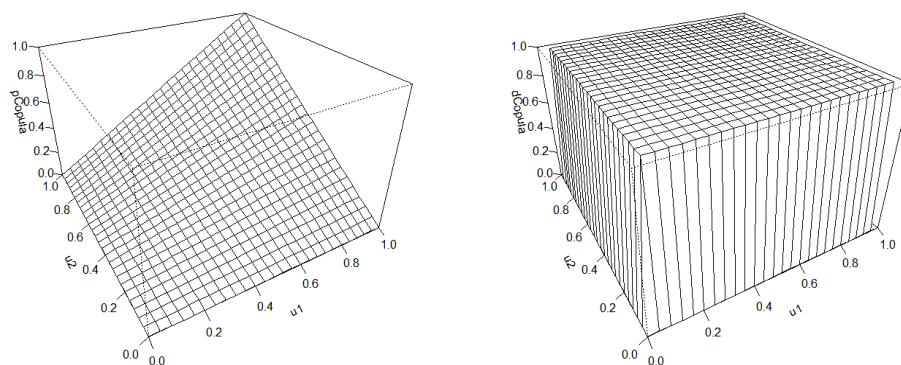


FIG. 3.1 – Représentation de la copule produit (gauche) et densité correspondante (droite).

### 3.4 Bornes de Fréchet-Hoeffding

Hoeffding en (1940) a donné une formulation explicite de la déclaration "il ya une dépendance fonctionnelle entre les variables aléatoires  $X$  et  $Y$ "

Soit les fonctions  $M$  et  $W$  définies sur l'intervalle  $[0, 1]^d$  comme suit :

La borne supérieure de Fréchet  $M(u_1, \dots, u_d)$  est

$$M(u_1, \dots, u_d) = \min \{u_1, \dots, u_d\} \quad (3.5)$$

La borne inférieure de Fréchet  $W(u_1, \dots, u_d)$  est

$$W(u_1, \dots, u_d) = \max \left\{ \sum_{i=1}^d u_i + 1 - d, 0 \right\}. \quad (3.6)$$

**Théorème 3.4.1 (Bornes de Fréchet-Hoeffding)** *Étant donné une  $d$ -copules  $C$ , pour tout  $u \in [0, 1]$ , on a :*

$$W(u_1, \dots, u_d) \leq C(u_1, \dots, u_d) \leq M(u_1, \dots, u_d).$$

**Remarque 3.4.1** *La limite supérieure de Fréchet-Hoeffding représente la plus forte forme de dépendance que plusieurs variables aléatoires peuvent présenter. En fait, ce n'est rien d'autre que la copule de comonotonie quand  $n = 2$ . De même, lorsque  $n = 2$ , la limite inférieure de Fréchet-Hoeffding n'est rien d'autre que la copule de contremonotonie.*

**Remarque 3.4.2** *Une famille de copules qui comprend  $W$ ,  $\Pi$  et  $M$  est considérée comme complète. La copule complète à deux paramètres donnée par la formule 3.7 est due à Fréchet (1958) :*

$$C_{\alpha,\beta} = \alpha W(u, v) + \beta M(u, v) + (1 - \alpha - \beta)\Pi(u, v), \quad (3.7)$$

*appelée la copule de Fréchet.*

### 3.5 Copule et variables aléatoires

Soit  $X_1, \dots, X_d$  variables aléatoires avec des fonctions de distribution continues  $F_1, \dots, F_d$  respectivement, et  $H$  la fonction de distribution conjointe. Alors  $(X_1, \dots, X_d)$  a une copule unique  $C$ . La représentation standard de la copule de la distribution du vecteur aléatoire  $(X_1, \dots, X_d)$  devient alors :

$$H(x_1, \dots, x_d) = P(X_1 \leq x_1, \dots, X_d \leq x_d) = C(F_1(x_1), \dots, F_d(x_d)).$$

Les transformations  $X_i \rightarrow F_i(X_i)$  utilisés dans la représentation ci-dessus habituellement dénommés les transformations de probabilité-intégrale et constituent un outil standard dans la méthode de simulation.

**Proposition 3.5.1 (Propriété d'invariance)** *Soit  $(X_1, \dots, X_d)$  un vecteur de variables aléatoires continues, de fonctions de répartition  $F$  associée à une copule  $C$  et  $(T_1, \dots, T_d)$  est une suite de fonctions strictement croissantes. Alors, la fonction de répartition jointe du vecteur aléatoire  $(T_1(X_1), \dots, T_d(X_d))$  est aussi associée la même copule  $C$ .*

**Proposition 3.5.2 (Transformée distributive)** *Soit  $U$  la transformée distributive de  $X$  comme définie dans 1.9 alors*

$$U \stackrel{d}{=} U[0, 1] \text{ et } X = F^{-1}(U) \quad (p.s.).$$

**Définition 3.5.1** *Les composantes du vecteur aléatoire  $X = (X_1, \dots, X_d)$  sont indépendantes ssi*

$$H(x_1, \dots, x_d) = F_1(x_1) \times \dots \times F_d(x_d).$$

**Théorème 3.5.1** *Soit  $(X_1, \dots, X_d)$  un vecteur de variables aléatoires continues avec une copule  $C$ ,  $(X_1, \dots, X_d)$  sont indépendants ssi  $C = \Pi^d$ .*

### 3.6 Copule de survie

Si on remplace  $C$  par  $\hat{C}$ ,  $u$  par  $1 - u$ , et  $v$  par  $1 - v$  dans la formule de la copule, on déplace effectivement l'origine du système de coordonnées de  $(0, 0)$  à  $(1, 1)$  et on obtient la mesure des variantes dans le sens inverse. Bien que cela soit une procédure banale, les deux distributions sont considérées comme distinctes, car les résultats de leur adaptation aux données sont différents (à moins qu'il n'y ait de symétrie). La copule  $\hat{C}$  ainsi obtenue est appelée copule de survie ou copule complémentaire satisfaisant

$$\hat{C}(u, v) = u + v - 1 + C(1 - u, 1 - v), \quad (3.8)$$

et

$$\overline{H}(x, y) = \hat{C}(\overline{F}(x), \overline{G}(y)). \quad (3.9)$$

Il est clair que  $\hat{C}$  est une copule qui "couple" la fonction de survie commune  $\overline{H}$  aux fonctions de survie marginale univariées d'une manière complètement analogue à la façon dont une copule relie la distribution conjointe à ses marges.

**Définition 3.6.1** Soit  $\overline{C}$  la fonction de survie conjointe de deux variables uniformes dont la distribution conjointe est la copule  $C$ . On a la relation suivante :

$$\overline{C}(1 - u, 1 - v) = 1 - u - v + C(u, v) = \hat{C}(u, v). \quad (3.10)$$

**Exemple 3.6.1** Considérons la distribution bivariée de Pareto considérée dans Hutchinson et Lai (1990). Soit  $X$  et  $Y$  une paire de variables aléatoires dont la fonction de survie

conjointe est donnée par

$$\overline{H}_\theta(x, y) = \begin{cases} (1 + x + y)^{-\theta}, & x \geq 0, y \geq 0 \\ (1 + x)^{-\theta}, & x \geq 0, y < 0 \\ (1 + y)^{-\theta}, & x < 0, y \geq 0 \\ 1, & x < 0, y < 0 \end{cases},$$

où  $\theta > 0$ . Les fonctions de survie marginales sont  $\overline{F}_1(x) = (1 + x)^{-\theta}$  et  $\overline{F}_2(y) = (1 + y)^{-\theta}$ .

Alors la copule de survie est :

$$\hat{C}_\theta(u, v) = (u^{-1/\theta} + v^{-1/\theta} - 1)^{-\theta}.$$

## 3.7 Familles de copules

Il existe nombreuses familles de copules qui diffèrent dans le détail de la dépendance qu'elles représentent. Une famille en général a plusieurs paramètres qui se rapportent à la force et la forme de la dépendance. Certaines familles de copules sont décrites ci-dessous.

### 3.7.1 Copules elliptiques

Les copules elliptiques dérivent de distributions elliptiques multivariées, elles jouent un rôle important dans les applications financières. Ici, Nous donnons les deux exemples les plus importants, la copule Gaussienne et celle de Student. Par construction, ces deux copules sont proches dans leur partie centrale, et se rapprochent dans leur queue seulement lorsque le nombre de degrés de liberté de la copule de Student augmente. En conséquence, il est parfois difficile de les distinguer, même avec des tests sensibles.

Par le théorème 2.3.7 de Nelsen (2006) [109] la copule bivariée est elliptique a une symétrie



de réflexion (symétrie radiale) ssi

$$C(u, v) = C(v, u) \text{ et } C(u, v) = \overline{C}(u, v) = u + v - 1 + C(1 - u, 1 - v), \quad u, v \in [0, 1].$$

De plus les copules elliptiques n'ont pas d'expressions de forme fermée.

**Définition 3.7.1 (Distribution elliptique)** Soit  $X$  un vecteur aléatoire  $d$ -dimensionnel et, pour certains  $\mu \in \mathbb{R}^n$ , une matrice symétrique définie non négatif  $\Sigma \in \mathbb{R}^{d \times d}$ , La fonction caractéristique  $\varphi_{X-\mu}$  de  $X - \mu$  est une fonction de la forme quadratique  $t^T \Sigma t$ , i.e.

$$\varphi_{X-\mu}(t) = \phi(t^T \Sigma t),$$

On dit que  $X$  a une distribution elliptique avec des paramètres  $\mu$ ,  $\Sigma$  et  $\phi$ .

### Copule Normale (Gaussienne)

La copule gaussienne c'est la distribution la plus populaire dans les applications, notamment en hydrologie et en finance. Soit  $\Phi$  la fonction standard de distribution normale univariée et  $\Psi$  la fonction de distribution normale bivariée avec un paramètre de corrélation  $\rho$ .

La copule gaussienne bivariée est donnée par

$$C_\rho(u, v) = \Psi(\Phi^{-1}(u), \Phi^{-1}(v)). \quad (3.11)$$

On a de façon explicite

$$C_\rho(u, v) = \frac{1}{2\pi\sqrt{1-\rho^2}} \int_{-\infty}^{\Phi^{-1}(u)} \int_{-\infty}^{\Phi^{-1}(v)} \exp\left[\frac{-(s^2 - 2\rho st + t^2)}{2(1-\rho^2)}\right] ds dt. \quad (3.12)$$

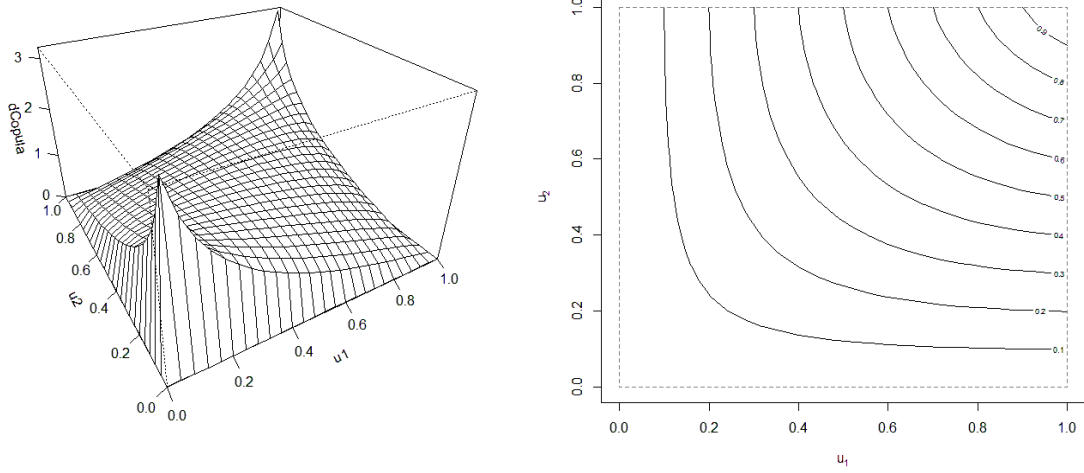


FIG. 3.2 – Densité (gauche), contour plot (droite) d’une copule Gaussian bivariée avec  $\rho = 0.5$ .

### t-copula

La t-copule d-dimensionnelle est définie par :

$$C_{\Sigma, \nu}^t(u_1, \dots, u_d) = t_{\Sigma, \nu}(t_{\nu}^{-1}(u_1), \dots, t_{\nu}^{-1}(u_d)),$$

où  $t_{\Sigma, \nu}$  est la fonction de distribution t-dimensionnelle multivariée standard à  $\nu$  degrés de liberté et  $\Sigma$  matrice de corrélation.

**Définition 3.7.2 (t-copula multivariée)** La t-copule d-dimensionnelle avec une matrice de corrélation linéaire  $\Sigma$  et degrés de liberté  $\nu$  est défini comme suit :

$$C_{\Sigma, \nu}^t(u_1, \dots, u_d) = \int_{-\infty}^{t_{\nu}^{-1}(u_1)} \dots \int_{-\infty}^{t_{\nu}^{-1}(u_d)} \frac{\Gamma((\nu + d)/2)(1 + w^T \Sigma^{-1} w / \nu)^{-(\nu + d)/2}}{|\Sigma|^{1/2} \Gamma(\nu/2)(\nu \Pi)^{d/2}} dw_1 \dots dw_d,$$

où  $|\Sigma|$  représente le déterminant de  $\Sigma$ ,  $w = (w_1, \dots, w_d)$  et  $\Gamma(\cdot)$  est la fonction Gamma.

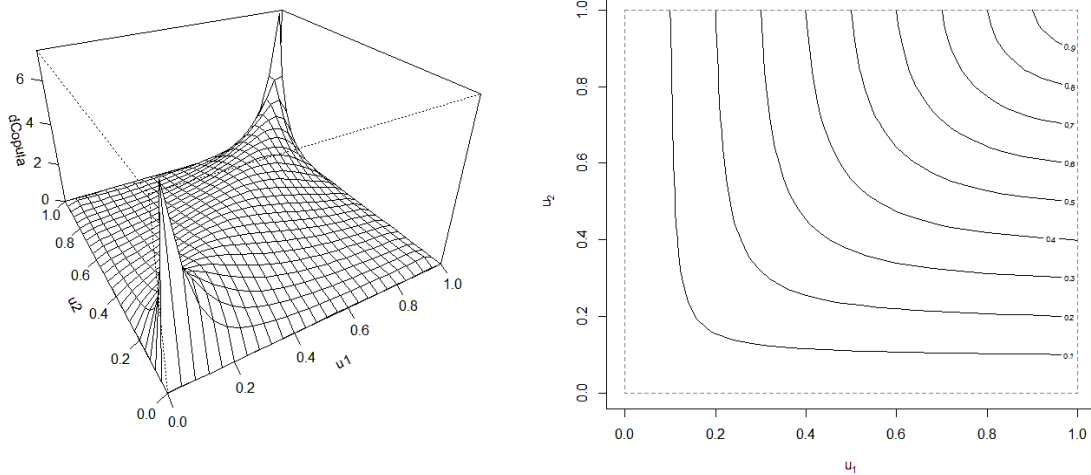


FIG. 3.3 – Densité (gauche) et contour plot (droite) d’une t-copule pour  $\rho = 0.7$  et  $\nu = 3$ .

### 3.7.2 Copules Archimédiennes

L’importance de cette classe de copules réside dans le fait qu’elle englobe très grand nombre de copules tout en jouissant d’un certain nombre de propriétés intéressantes. De plus, comme l’ont souligné Frees et Valdez [63], un grand nombre de modèles développés pour tenir compte de la dépendance entre diverses sources de risques dans la théorie de l’assurance conduisent à des copules archimédienne.

Les avantages de cette classe sont :

- Facilité dans la construction.
- Grande variété de familles de copules appartenant à cette classe.
- Possibilité de réduction de l’étude d’une copule multivariée a une seule fonction univariée.

**Définition 3.7.3** Soit  $\varphi$  une fonction continue, strictement décroissante, de  $[0, 1]$  dans  $[0, \infty]$  telle que  $\varphi(1) = 0$ . Le pseudo-inverse de  $\varphi$  est la fonction  $\varphi^{[-1]}$ , avec le domaine de

définition  $[0, \infty]$  à valeurs dans  $[0, 1]$  est donné par

$$\varphi^{[-1]}(t) = \begin{cases} \varphi^{-1}(t) & \text{si } 0 \leq t \leq \varphi(0) \\ 0 & \text{si } \varphi(0) \leq t \leq \infty \end{cases}. \quad (3.13)$$

On note que  $\varphi^{[-1]}$  est continue, non-croissante sur  $[0, \infty]$ , et strictement décroissante sur  $[0, \varphi(0)]$ . D'autre part, on a

$$\varphi(\varphi^{[-1]}(t)) = \begin{cases} t & \text{si } 0 \leq t \leq \varphi(0) \\ \varphi(0) & \text{si } \varphi(0) \leq t \leq \infty \end{cases}.$$

Ainsi que si  $\varphi(0) = \infty$ , alors  $\varphi^{[-1]}(t) = \varphi^{-1}(t)$ .

**Théorème 3.7.1** Soit  $\varphi$  une fonction continue, strictement décroissante, de  $[0, 1]$  dans  $[0, \infty]$  telle que  $\varphi(1) = 0$ . Soit  $\varphi^{[-1]}$  le pseudo-inverse de  $\varphi$ . Soit  $C$  la fonction définie sur  $[0, 1]^2$  à valeurs dans  $[0, 1]$ , donnée par

$$C(u, v) = \varphi^{[-1]}(\varphi(u) + \varphi(v)). \quad (3.14)$$

Alors  $C$  est une copule ssi la pseudo-inverse de  $\varphi$  est une fonction convexe.

**Démonstration.** Voir Nelsen (2006) [109], page 111. □

La fonction  $\varphi$  est appelée générateur de la copule. Si  $\varphi(0) = \infty$ , on dit alors que  $\varphi$  c'est un générateur stricte et  $C(u, v) = \varphi^{-1}(\varphi(u) + \varphi(v))$  est dite une copule archimédienne stricte.

**Théorème 3.7.2** Soit  $C$  une copule archimédienne de générateur  $\varphi$  alors

1. la symétrie i.e.  $C(u, v) = C(v, u)$ , pour tout  $u, v \in [0, 1]$ .
2. l'associativité i.e.  $C(C(u, v), w) = C(u, C(v, w))$ , pour tout  $u, v, w \in [0, 1]$ .

**Exemples de Copules Archimédienne**

**Famille de Frank** La famille de Frank donnée par :

$$C(u, v) = -\frac{1}{\theta} \ln \left\{ 1 + \frac{(e^{-\theta}(u-1)e^{-\theta}(v-1))}{e^{-\theta} - 1} \right\}, \quad \theta \in \mathbb{R} \setminus \{0\}, \quad (3.15)$$

Cette famille de copules stricte a un générateur

$$\varphi(t) = -\ln \frac{e^{-\theta t} - 1}{e^{-\theta} - 1}.$$

La copule de Frank est populaire pour plusieurs raisons. Tout d'abord, contrairement à d'autres copules, elle permet une dépendance négative entre les marginales. Deuxièmement, la dépendance est symétrique dans les deux queues, semblable aux copules Gaussiennes et Student. Troisièmement, elle est complète en ce sens que les deux bornes Fréchet sont incluses dans la gamme de la dépendance permise. En conséquence, la copule de Frank peut être utilisée en théorie pour modéliser des résultats ayant une forte dépendance positive ou négative. La dépendance dans les queues de la copule de Frank tend relativement faible par rapport à la copule gaussienne, et la plus forte dépendance est centrée au milieu de la distribution, ce qui suggère que la copule de Frank est la plus appropriée pour les données qui présentent une faible dépendance de queue.

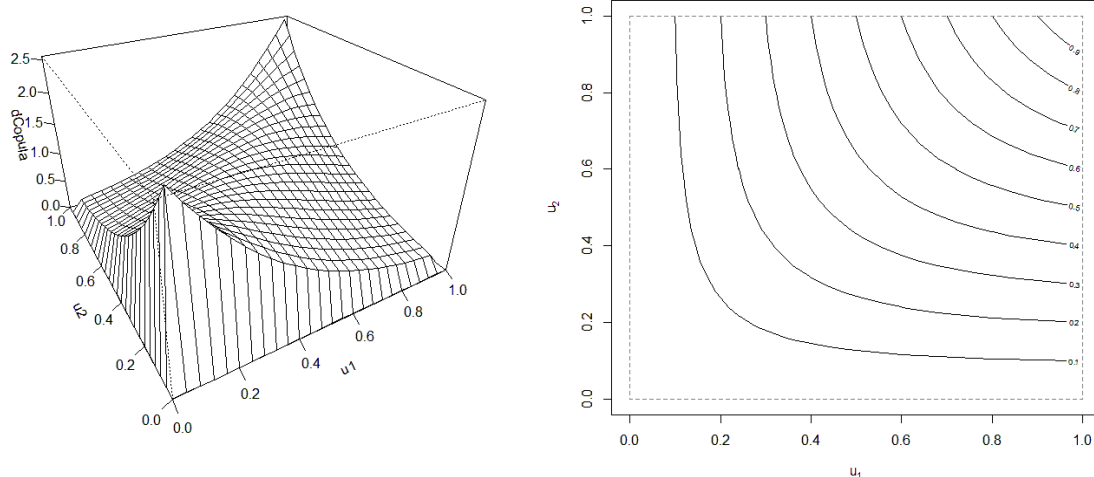


FIG. 3.4 – Densité (gauche) et contour plot (droite) de la copule de Frank bivariée avec  $\theta = 10$ .

**Famille de Clayton** La copule de Clayton (1978), également connue sous le nom de copule de Cook et Johnson (1981), initialement étudiée par Kimeldorf et Sampson en (1975), elle prend la forme suivante :

$$C(u, v) = (u^{-\theta} + v^{-\theta} - 1)^{-1/\theta}, \quad \theta \in [-1, \infty \setminus \{0\}], \quad (3.16)$$

avec le paramètre de dépendance restreint sur la région  $]0, 1[$ . Le cas limite  $\theta = 0$  correspond au cas indépendant, i.e.  $C = \Pi^2$ . En particulier, cette famille est positivement ordonnée et ses membres sont absolument continus pour  $\theta > 0$ .

Cette famille de copule stricte a le générateur suivant

$$\varphi^{-1}(s) = (\theta s + 1)^{-1/\theta}.$$

Il s'ensuit que

$$\varphi(t) = (t^{-\theta} - 1)/\theta.$$

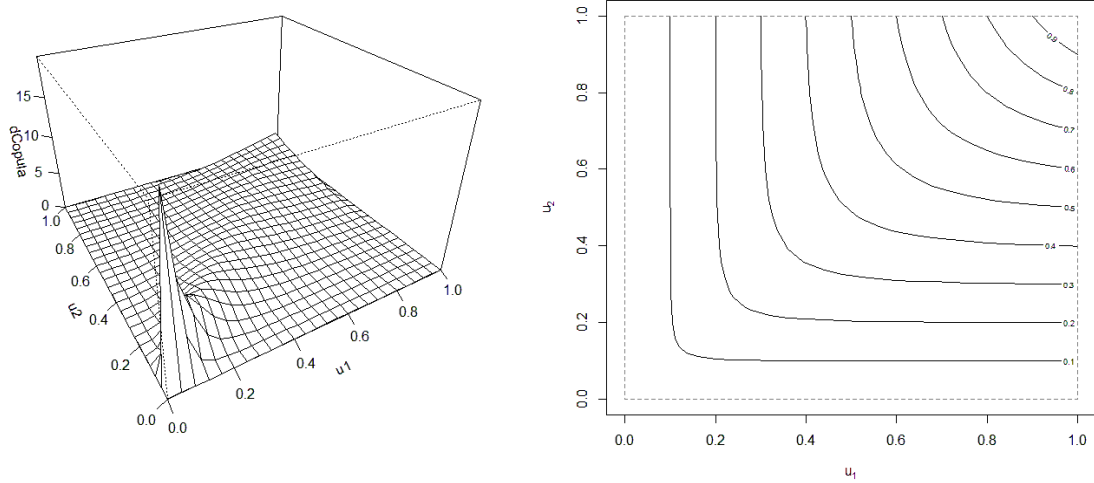


FIG. 3.5 – Densité (gauche) et contour plot (droite) de la copule de Clayton bivariée avec  $\theta = 10$ .

**Copule de Gumbel** La copule de Gumbel (ou modèle logistique) (Gumbel, 1960) prend la forme :

$$C(u, v) = \exp \left\{ - \left[ (-\ln u)^\theta + (-\ln v)^\theta \right]^{1/\theta} \right\}. \quad (3.17)$$

Le paramètre de dépendance est limité à l'intervalle  $[1, \infty]$ . Les valeurs de 1 et  $\infty$  correspondent à l'indépendance et à la limite supérieure de Fréchet, mais cette copule n'atteint pas la borne inférieure de Fréchet pour aucune valeur de  $\theta$ . Les copules de Gumbel sont archimédiennes strictes avec le générateur suivant :  $\varphi(t) = (-\ln t)^\theta$ .

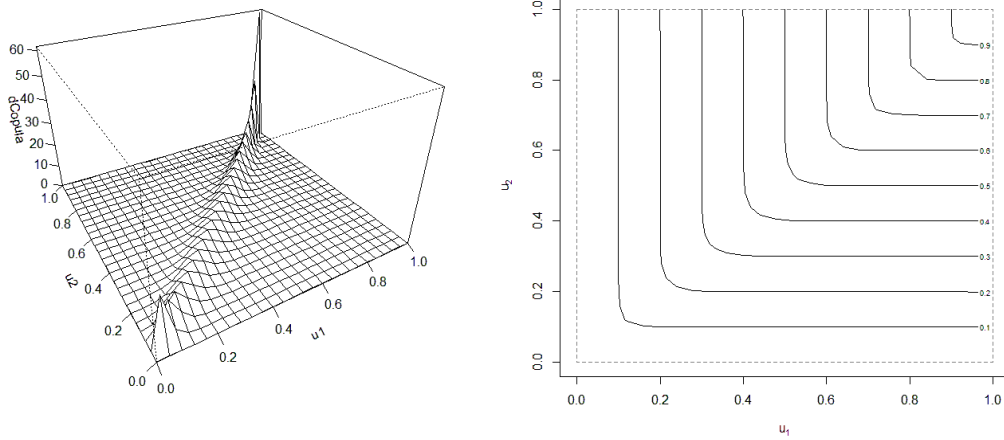


FIG. 3.6 – Densité (gauche) et contour plot (droite) de la copule de Gumbel bivariée avec  $\theta = 2$ .

### 3.7.3 Copule de valeur extrême

Une autre famille de copules qui est d'usage commun est celle des copules de valeur extrême. Ces copules sont dérivées de la structure de dépendance des distributions multivariées des valeurs extrêmes généralisée, qui fournissent les distributions limites des maximums composites de vecteurs aléatoires  $d$ -dimensionnel, après une normalisation appropriée.

Soit  $(X_1, Y_1), (X_2, Y_2), \dots, (X_n, Y_n)$  des couples i.i.d de variables aléatoires avec une copule commune  $C$ . On note  $M_n$  le vecteur des maximums, composante par composantes,

$$M_n = (\max \{X_1, \dots, X_n\}, \dots, \max \{Y_1, \dots, Y_n\})$$

Soit  $C_{(n)}$  la copule associée à  $M_n$ . D'après le théorème 3.3.1 de Nelsen (2006), on a :

$$C_{(n)}(u, v) = C^n(u^{1/n}, v^{1/n}), \quad 0 \leq u, v \leq 1. \quad (3.18)$$



La limite de la suite  $\{C_{(n)}\}$  conduit à la notion suivante de copule de valeur extrême.

**Définition 3.7.4 (Copule de valeur extrême)** *Une copule  $C^*$  est une copule de valeur extrême s'il existe une copule  $C$  telle que*

$$C^*(u, v) = \lim_{n \rightarrow \infty} C^n(u^{1/n}, v^{1/n}),$$

pour tout  $u, v \in [0, 1]$ , en outre on dit que  $C$  appartenir au domaine d'attraction de  $C^*$ .

**Théorème 3.7.3** *Une copule est max-stable ssi elle est une copule de valeur extrême.*

Nous présentons maintenant la représentation de Pickands en (1981) pour construire des copules de valeur extrême (ou équivalent, max-stable) qui est maintenant largement utilisée. Il assume des marges exponentielles standard alors les fonctions de survie de  $X$  et  $Y$  sont  $\bar{F}(x) = e^{-x}$ ,  $x > 0$  et  $\bar{F}(y) = e^{-y}$ ,  $y > 0$  donc la fonction de survie conjointe est donnée par

$$\bar{H}(x, y) = \exp \left[ -(x + y)A \left( \frac{x}{x + y} \right) \right].$$

Pour une fonction continue, non négative  $A : [0, 1] \rightarrow [1/2, 1]$  définie par

$$A(t) = -\ln C(e^{-(1-t)}, e^{-t}),$$

$A$  doit satisfaire les conditions suivantes :

1.  $A(0) = A(1) = 1$ ,
2.  $\max\{t, 1 - t\} \leq A(t) \leq 1$ ,
3.  $A$  est convexe.

La copule de valeur extrême a la forme suivante :

$$C(u, v) = \exp \left\{ \ln(uv)A \left( \frac{\ln v}{\ln(uv)} \right) \right\}. \quad (3.19)$$

Pour un choix approprié de la fonction  $A$  (appelée la fonction de dépendance de la copule de valeur extrême).

**Exemple 3.7.1 (Copule de Gumbel–Hougaard)**

$$C(u, v) = \exp \left( - [(-\log u)^\alpha + (-\log v)^\alpha]^{1/\alpha} \right).$$

La copule de Gumbel–Hougaard est aussi un copula archimédienne, en fait, il n’y a aucune autre copule archimédienne qui est aussi une copule de valeur extrême.

**Exemple 3.7.2 (copule de Marshall et Olkin )**

$$C(u, v) = uv \min(u^{-\alpha}, v^{-\beta}) = \min v(uv^{1-\beta}, u^{1-\alpha}v).$$

### 3.8 Mesures d’association

Les copules offrent un moyen utile d’étudier et de mesurer l’association entre les variables aléatoires. Dans ce qui suit, nous donnons une définition formelle de la concordance et certaines de ses propriétés et présentons les mesures de concordance les plus cruciales qui sont pertinentes pour cette thèse.

#### 3.8.1 Corrélations linéaires

Le coefficient de corrélation linéaire (ou corrélation de Pearson) est une mesure d’association pour deux variables aléatoires  $X$  et  $Y$ , qui capte le degré de relation linéaire. Il est défini comme suit :

**Définition 3.8.1** *Pour un couple aléatoire  $(X, Y)$  ayant des marginaux avec des écarts finis, le coefficient de corrélation de Pearson  $r_P$  est défini par*

$$r_P(X, Y) = \frac{\text{Cov}(X, Y)}{\sqrt{\text{Var}(X) \cdot \text{Var}(Y)}},$$

où  $Cov(X, Y)$  est la covariance de  $X$  et  $Y$  défini par

$$Cov(X, Y) := E[(X - E(X))(Y - E(Y))] = E(XY) - E(X)E(Y).$$

La connaissance du coefficient  $r_P(X, Y)$  est équivalente à celle du coefficient  $\beta$  de la régression linéaire  $Y = \beta X + \varepsilon$ , où  $\varepsilon$  est le résidu qui est linéairement non corrélé avec  $X$ , en effet  $r_P = \beta \sqrt{Var(X)/Var(Y)}$ .

**Remarque 3.8.1** *Les variances de  $X$  et  $Y$  doivent être finies pour le coefficient de corrélation linéaire de Pearson à définir. C'est un inconvénient pour les mesures de dépendance et cause des problèmes lorsque l'on travaille avec des distributions à queue lourde (distributions de variance infinies).*

### 3.8.2 Mesures de concordance

Deux variables aléatoires sont dites concordantes si les grandes valeurs de l'un viennent avec les grandes valeurs de l'autre et les petites valeurs d'une viennent avec petites valeurs de l'autre. Deux échantillons  $(x_i, y_i)$  et  $(x_j, y_j)$  d'un vecteur aléatoire continu  $(X, Y)$  sont concordants si  $(x_i - y_i)(x_j - y_j) > 0$ . Cela signifie que soit  $x_i < x_j$  et  $y_i < y_j$  ou  $x_i > x_j$  et  $y_i > y_j$ . De même, nous disons que ils sont discordant si  $(x_i - y_i)(x_j - y_j) < 0$  i.e  $x_i < x_j$  et  $y_i > y_j$  ou  $x_i > x_j$  et  $y_i < y_j$ .

**Définition 3.8.2 (Mesure de concordance)** *Une fonction  $\delta$  définie sur des paires de variables aléatoires est une mesure de concordance si pour chaque paire  $X$  et  $Y$  de variables aléatoires elle satisfait les propriétés suivantes :*

1.  $\delta_{X,Y}$  est bien défini pour chaque paire  $X, Y$  de variables aléatoires continues.
2.  $-1 \leq \delta_{X,Y} \leq 1$ .
3.  $\delta_{X,Y} = \delta_{Y,X}$ .
4.  $\delta_{X,Y} = \delta_{C^n} = 0$  si  $X$  et  $Y$  sont indépendants.

5.  $\delta_{-X,Y} = \delta_{X,-Y} = -\delta_{X,Y}$ .
6. Pour toutes les paires de copules telles que  $C_1 < C_2$ , alors  $\delta_{C_1} \leq \delta_{C_2}$ .
7. Si  $(X_n, Y_n)$  a la copule  $C_n$  et  $(C_n)_{n \geq 1}$  converge vers le point  $C$ , alors  $\lim_{n \rightarrow \infty} \delta_{C_n} = \delta_C$ .

### 3.8.3 Tau de Kendall et Rho de Spearman

Le tau de Kendall et le rho de Spearman sont des mesures de concordance jouant un rôle important dans les statistiques non paramétriques. Ils nous fournissent des alternatives viables au coefficient de corrélation linéaire comme mesure d'association pour les distributions non unipolaires.

**Définition 3.8.3** *Le tau de Kendall correspond à la différence entre la probabilité de la concordance et de la probabilité de discordance pour deux vecteurs  $(X_1, Y_1)$  et  $(X_2, Y_2)$ , il est défini par :*

$$\tau = P((X_1 - X_2)(Y_1 - Y_2) > 0) - P((X_1 - X_2)(Y_1 - Y_2) < 0). \quad (3.20)$$

**Théorème 3.8.1** *Soit  $(X, Y)$  un vecteur de variables aléatoires continues de copule  $C$ . Alors le tau de Kendall  $\tau$  est exprimée en termes de copule par la formule suivante :*

$$\tau = 4 \int_0^1 \int_0^1 C(u_1, u_2) dC(u_1, u_2) - 1 \quad (3.21)$$

Soit  $C$  une copule archimédienne générée par  $\varphi$ . Genest et MacKay (1986) ont montré que

$$\tau = 4E(C(U, V)) - 1 = 4 \int_0^1 \frac{\varphi(t)}{\varphi'(t)} dt. \quad (3.22)$$

**Définition 3.8.4** *Le rho de Spearman est exprimé en termes de probabilités de concordance et discordance. Plus précisément, étant donné trois vecteurs i.i.d.  $(X_1, Y_1)$ ,  $(X_2, Y_2)$  et*

$(X_3, Y_3)$ , il est défini par

$$\rho = 3 [P((X_1 - X_2)(Y_1 - Y_3) > 0) - P((X_1 - X_2)(Y_1 - Y_3) < 0)].$$

**Théorème 3.8.2** Soit  $(X, Y)$  un vecteur de variables aléatoires continues avec copule  $C$ .

Le rho de Spearman  $\rho$  est exprimée en termes de copule comme suit :

$$\rho = 12 \int_0^1 \int_0^1 C(u_1, u_2) du_1 du_2 - 3.$$

**Remarque 3.8.2** Si  $X \sim F$  et  $Y \sim G$  et si  $U_1 = F(X)$  et  $U_2 = G(Y)$  alors

$$\begin{aligned} \rho &= 12 \int \int_{[0,1]^2} u_1 u_2 dC(u_1, u_2) - 3 \\ &= 12E[U_1 U_2] - 3 \\ &= \frac{E[U_1 U_2] - 1/4}{1/12} \\ &= \frac{Cov(U_1, U_2)}{\sqrt{Var(U_1)} \sqrt{Var(U_2)}} \\ &= r_p(F(X), G(Y)). \end{aligned}$$

Ainsi le rho de Spearman est le coefficient de corrélation linéaire entre les transformées intégrales de X et Y, ( $F(X)$  et  $G(Y)$ ). Pour cette raison, il est souvent mesuré de corrélation de rang. Parce que le tau Kendall et le rho de Spearman satisfont les propriétés de concordance et si nous rappelons les limites de Fréchet-Hoeffding, il s'ensuit que par exemple, la concordance maximale est atteinte lorsque en utilisant la copule maximale i.e  $C = \Pi \Rightarrow \rho = \tau = 1$ . L'inverse est également vraie, de sorte que

$$\rho = \tau = -1 \quad \text{ssi} \quad C = W$$

$$\rho = \tau = 1 \quad \text{ssi} \quad C = M.$$

On donne dans le tableau 3.1 quelques exemples de mesures d'association.

Ici,  $D_1(\theta) = \theta^{-1} \int_0^\theta \frac{x}{e^x - 1} dx$  représente la première fonction de Debye.

Copule	Domaine de $\theta$	Tau de Kendall	Rho de Spearman
<b>Produit</b>	<i>N.A</i>	0	0
<b>FGM</b>	$-1 \leq \theta \leq 1$	$\frac{2}{9}\theta$	$\frac{1}{3}\theta$
<b>Clayton</b>	$0 \leq \theta \leq \infty$	$\frac{\theta}{\theta+2}$	Pas de forme fermée
<b>Frank</b>	$-\infty \leq \theta \leq \infty$	$1 - \frac{4(1-D_1(\theta))}{\theta}$	$t - \frac{1-e^{-\theta}}{\theta} \log \left( \frac{1-e^{-\theta}}{1-e^{-\theta t}} \right)$
<b>Normale</b>	$-1 \leq \theta \leq 1$	$\frac{2}{\pi} \arcsin(\theta)$	$\frac{6}{\pi} \arcsin(\theta/2)$

TAB. 3.1 – Tau de Kendall et Rho de Spearman de quelques copules.

### 3.8.4 Mesures de dépendance

Dans cette partie, nous présentons la notion de mesure de la dépendance. Une mesure de dépendance donne une indication sur la mesure dans laquelle deux variables sont mutuellement liées. L'idée clé est d'évaluer une distance entre la copule associée à une paire donnée de variables aléatoires et la copule du produit.

**Définition 3.8.5 (Mesures de dépendance)** *Une fonction  $\delta$  définie sur des variables aléatoires paires est une mesure de dépendance si elle satisfait les conditions dans la définition de concordance, avec les propriétés suivantes :*

1.  $0 = \delta_{C_{\Pi}} \leq \delta_C \leq \delta_W = 1$ .
2.  $\delta_{X_1, X_2} = \delta_{C_{\pi}} = 0$  ssi  $X_1$  et  $X_2$  sont indépendants.

Les mesures de dépendance les plus populaires sont définies comme suit :

**Définition 3.8.6 (Phi de Hoeffding)** *le Phi de Hoeffding pour un vecteur de variables aléatoires continues  $(X, Y)$  avec la copule  $C$  est donné par :*

$$\phi = 3 \sqrt{10 \int \int_{[0,1]^2} (C(u_1, u_2) - u_1 u_2)^2 du_1 du_2.}$$

**Définition 3.8.7 (Sigma de Schweizer-Wolff)** *Sigma de Schweizer-Wolff pour un vecteur de variables aléatoires continues  $(X, Y)$  avec la copule  $C$  est donné par :*

$$\sigma = 12 \sqrt{\int_{[0,1]^2} |C(u_1, u_2) - u_1 u_2| du_1 du_2}. \quad (3.23)$$

### 3.8.5 Dépendance de queue

Le concept de dépendance de queue est lié à la quantité de dépendance dans la queue du quadrant supérieur droite ou la queue du quadrant inférieur gauche d'une distribution bivariable. C'est un concept qui est pertinent pour l'étude de la dépendance entre les valeurs extrêmes. Il s'avère que la dépendance de la queue entre deux variables aléatoires continues  $X$  et  $Y$  est une propriété de copule et donc la quantité de dépendance de queue est invariante sous des transformations strictement croissantes de  $X$  et  $Y$ .

La dépendance de queue est une mesure locale contrairement aux mesures de concordance qui sont des mesures sur toute la distribution. Le coefficient de dépendance de queue joue un rôle important dans la théorie financière puisqu'il mesure la probabilité que l'actif  $X$  subisse une perte plus grande que  $x_p$ , associée au quantile  $p$  tendant vers zéro, conditionnée à la réalisation d'une perte de l'actif  $Y$  plus grande que  $y_p$  associée au même quantile  $p$ . Voir Joe (1997) [89]. Des discussions détaillées sur la dépendance de la queue peuvent être trouvées chez McNeil et al ([101], section 5.2.3).

Soit  $X$  et  $Y$  deux variables aléatoire avec une fonction de distribution bivariable  $F$  et des marginales continues  $F_1$  et  $F_2$  et soit  $u$  une valeur de seuil, alors le coefficient de dépendance de queue supérieure  $\lambda_U$  est défini par.

**Définition 3.8.8** *Le coefficient de dépendance de queue supérieure (upper tail dependence coefficient) de  $(X, Y)$  est*

$$\lambda_U = \lim_{u \rightarrow 1^+} P(X > F_1^{-1}(u) \mid Y > F_2^{-1}(u)). \quad (3.24)$$

Le coefficient de dépendance de queue inférieure  $\lambda_L$  (lower tail dependence coefficient) est défini de la même manière, alors

$$\lambda_L = \lim_{u \rightarrow 0^-} P(X \leq F_1^{-1}(u) \setminus Y \leq F_2^{-1}(u)), \quad (3.25)$$

À condition que les limites  $\lambda_U$  et  $\lambda_L$  existent et appartenant à l'intervalle  $[0, 1]$ . Si  $\lambda_U \in ]0, 1]$ , on dit que  $X$  et  $Y$  sont asymptotiquement dépendants au niveau supérieur de la queue de distribution, si  $\lambda_U = 0$  on dit que  $X$  et  $Y$  sont asymptotiquement indépendants au niveau supérieur de la queue de distribution.

Comme  $P((X > F_1^{-1}(u) \setminus Y > F_2^{-1}(u)))$  peut être écrit comme suit :

$$\frac{1 - P[X \leq F_1^{-1}(u)] - P[Y \leq F_2^{-1}(u)] + P[X \leq F_1^{-1}(u), Y \leq F_2^{-1}(u)]}{1 - P[X \leq F_1^{-1}(u)]},$$

Une définition alternative et équivalente (pour les variables aléatoires continues), dont il est vu que le concept de dépendance de queue est en effet une propriété de copule, est la suivante qui se trouve dans Joe (1997) [89], page. 33.

**Proposition 3.8.1** *Le coefficient de dépendance de queue supérieure peut être écrit en terme de copula :*

$$\lambda_U = \lim_{u \rightarrow 1^+} \frac{1 - 2u + C(u, u)}{1 - u}, \quad (3.26)$$

où  $C$  une copule bivariée. La dépendance de queue inférieure peut être définie de manière similaire que  $\lambda_U$ .

$$\lambda_L = \lim_{u \rightarrow 0^-} \frac{C(u, u)}{u} \quad (3.27)$$

Si les limites existes, alors on dit que  $C$  a une dépendance de queue supérieure (inférieure) si  $\lambda_U \in ]0, 1]$  (respectivement  $\lambda_L \in ]0, 1]$ ) et l'indépendance de queue supérieure (inférieure) si  $\lambda_U = 0$  (respectivement  $\lambda_L = 0$ ).



**Exemple 3.8.1** Pour la couple  $M$ , nous obtenons

$$\begin{aligned}\lambda_U &= \lim_{u \rightarrow 1^+} \frac{1 - 2u + C(u, u)}{1 - u} \\ &= \lim_{u \rightarrow 1^+} \frac{1 - u}{1 - 2u + \min(u, u)} \\ &= \lim_{u \rightarrow 1^+} \frac{1 - u}{1 - 2u + u} \\ &= 1.\end{aligned}$$

et

$$\begin{aligned}\lambda_L &= \lim_{u \rightarrow 1^+} \frac{\min(u, u)}{u} \\ &= \lim_{u \rightarrow 1^+} \frac{u}{u} \\ &= 1.\end{aligned}$$

**Exemple 3.8.2** Nous voulons calculer  $\lambda_U$  pour la copule gumbel donnée par la fomule

3.17. Ce serait facile si nous considérons la règle de l'Hopital :  $\lim_{u \rightarrow a} \frac{f(\theta)}{g(\theta)} = \lim_{u \rightarrow a} \frac{f'(\theta)}{g'(\theta)}$ .

Alors

$$\begin{aligned}\lambda_U &= \lim_{u \rightarrow 1^+} \frac{1 - 2u + C(u, u)}{1 - u} \\ &= \lim_{u \rightarrow 1^+} \frac{1 - 2u + \exp \left[ - \left\{ (-\ln u)^\theta + (-\ln u)^\theta \right\}^{1/\theta} \right]}{1 - u} \\ &= \lim_{u \rightarrow 1^+} \frac{1 - u}{1 - 2u + \exp \left[ - \left\{ 2(-\ln u)^\theta \right\}^{1/\theta} \right]} \\ &= \lim_{u \rightarrow 1^+} \frac{-2 - \frac{2u(-\ln u)^\theta \exp \left[ - \left\{ 2(-\ln u)^\theta \right\}^{1/\theta} \right]}{u \ln u}}{-1} \\ &= \lim_{u \rightarrow 1^+} \frac{2 - 2^\theta}{-1} \\ &= 2 - 2^\theta.\end{aligned}$$

Puisque  $\lambda_U \in ]0, 1]$  pour  $\theta > 1$ , Les variable aléatoires avec cette copule sont asymptotiquement dépendants dans la partie supérieure de la queue.

**Remarque 3.8.3** Dans cette thèse on s'intéresse au coefficient de dépendance de queue supérieure mais il faut mentionner que le coefficient de dépendance de queue supérieure de  $C$  est le coefficient de la queue inférieure de  $\bar{C}$  : de même, le coefficient de dépendance de la queue inférieure de  $C$  est le coefficient de la dépendance de queue supérieure de  $\bar{C}$  puisque  $\bar{C}(v, v) = 1 - 2v + C(v, v)$ .

Les expressions des coefficients de dépendance de queue pour quelques distributions bivariées sont présentées dans le tableau 3.2.

Famille	$\lambda_L$	$\lambda_U$
Fréchet	$\alpha$	$\alpha$
Clayton	$2^{-1/\theta}$	0
Marshall and Olkin	0	$\min(\alpha, \beta)$
Placket	0	0
Gumbel	0	$2 - 2^\theta$
Frank	0	0

TAB. 3.2 – Dépendance de la queue de certaines familles des copules.

**Remarque 3.8.4** 1. Pour la copule gaussienne bivariée, définie par la formule (3.13) on a :  $\lambda_U = \lambda_L = 0$ , si le coefficient de corrélation de la copule  $\rho < 1$ .

2. Pour une copule archimédienne, les paramètres de dépendance de queue peuvent être exprimé en termes de limites impliquant le générateur et son inverse.

### Estimation du coefficient de dépendance de queue

Le coefficient de dépendance de queue  $\eta$ , a été jugé très utile pour distinguer la dépendance asymptotique ou l'indépendance asymptotique et dans la dernière classe, entre une association positive, une indépendance ou une association négative. Cela pose le problème de l'estimation de  $\eta$ , le sujet de cette section, particulièrement pertinent.

Supposons  $X_1, \dots, X_n$  vecteurs aléatoires bivariés i.i.d avec une fonction de distribution F et copule C. Nous supposons des fonctions de distribution marginales continues  $F_i$ ,  $i = 1, 2$ . Les tests de dépendance de queue ou l'indépendance de queue sont donnés par exemple dans Ledford et Tawn (1996) ou Draisma et al (2004).

Les méthodes d'estimation paramétriques pour le coefficient de dépendance de queue sont efficaces si le modèle de distribution envisagé est vrai, mais ils présentent des estimations biaisées dans le cas où le modèle sous-jacent est différent.

Les procédures d'estimation non paramétriques évitent ce type d'erreur de modèle mais viennent avec une plus grande variance d'estimation. Les estimations du coefficient de dépendance de queue sont basées sur :

- (a) Une distribution spécifique ou une famille de distributions.
- (b) Une copule spécifique ou une famille de copules.
- (c) Un modèle non paramétrique.

Il y a beaucoup d'études qui ont essayé de présenter des estimateurs non paramétriques : certains sont basés sur l'estimation empirique de la copule de queue voir Schmidt et Stadtmüller (2006) [124], pour cette méthode il n'y a pas d'hypothèses paramétriques concernant les copules et les fonctions de distribution marginales.

Soit  $C_n$  la copule empirique définie par :

$$C_n(u, v) = F_n(F_{1n}^{-1}(u), F_{2n}^{-1}(v)),$$

avec  $F_n$  et  $F_{in}$  sont les fonctions de distribution empirique correspondant à  $F$  et  $F_i$ ,  $i = 1, 2$  respectivement.

Les premiers estimateurs sont basés sur les formules 3.24 et 3.25 sont :

$$\hat{\lambda}_{U,k,n} = \frac{n}{k} C_n \left( \left(1 - \frac{k}{n}, 1\right) \times \left(1 - \frac{k}{n}, 1\right) \right). \quad (3.28)$$

Et

$$\hat{\lambda}_{L,k,n} = \frac{n}{k} C_n \left( \frac{k}{n}, \frac{k}{n} \right).$$

où  $k = k(n) \rightarrow \infty$  et  $k/n \rightarrow 0$  quand  $n \rightarrow \infty$ , et la première expression dans 3.28 doit être compris comme la mesure de copule empirique dans l'intervalle  $(1 - k/n, 1) \times (1 - k/n, 1)$ ,

et

$$C_n = \frac{1}{n} \sum_{i=1}^n 1_{\{R_{1j}/n \leq u, R_{2j}/n \leq v\}}.$$

avec  $R_{1j}$  et  $R_{2j}$  sont les rangs dans les blocs de maxima de  $X_n$  et  $Y_n$ .

En utilisant la notation de Coles et al. (1999) [24] suivante :

$$\frac{1 - 2v + C(v, v)}{1 - v} = 2 - \frac{1 - C(v, v)}{1 - v} \underset{v \rightarrow 1^+}{\sim} 2 - \frac{\log C(v, v)}{\log v}.$$

Alors la deuxième forme de l'estimateur de  $\lambda$  est donnée par :

$$\hat{\lambda}_{U,k,n}^{(2)} = 2 - \frac{\log C_n \left( \frac{n-k}{n}, \frac{n-k}{n} \right)}{\log \left( \frac{n-k}{n} \right)}, 0 < k < n,$$

Le troisième estimateur de  $\lambda$  a été présenté par Joe et al. (1992) [?] et elle est donné par :

$$\hat{\lambda}_{U,k,n}^{(3)} = 2 - \frac{1 - C_n \left( \frac{n-k}{n}, \frac{n-k}{n} \right)}{1 - \left( \frac{n-k}{n} \right)}, \quad 0 < k \leq n,$$

La normalité asymptotique et la forte consistance de cet estimateur ont été discutées dans Schmidt et Stadtmüller (2006).

Une autre approche d'estimation non paramétrique de la dépendance de queue est qui consiste en la transformation des variables  $X$  et  $Y$  en des variables aléatoires de Fréchet suivant le résultat de Ledford et Tawn (1996, 1998) voir chapitre 4.

Les propriétés statistiques des estimateurs de ce coefficient ont basés sur l'ensemble des observations ou sur des données extrêmes et ils ont été étudiées par plusieurs auteurs en utilisant des techniques issues de la théorie des valeurs extrêmes. (voir le chapitre 4 de cette thèse).

# Chapitre 4

## Nombre optimal des statistiques d'ordre extrêmes utilisées dans l'estimation du coefficient de dépendance de queue

**B**eirlant et al (2011) [12] ont introduit un estimateur à bias-réduit pour le coefficient de dépendance de queue et pour la probabilité de queue bivariée dans la statistique de valeur extrême bivariée. Dans le présent chapitre, nous nous intéressons au problème du choix du nombre de statistiques d'ordre extrême des observations bivariées dépassant les seuils élevés, nous exposons différentes méthodes pour déterminer ce nombre. L'efficacité de nos méthodes est illustrée par une étude de simulation et par un ensemble de données réelles.

**Mots-clés :** Coefficient de dépendance de la queue ; Bias-réduit ; Distribution de Pareto étendue ; Probabilité de queue ; Copule ; Estimateur de Hill ; Estimateur de moment.

## 4.1 Introduction et motivation

Dans le cadre classique de la théorie des valeurs extrêmes bivariées, les procédures pour estimer la probabilité d'un événement extrême ne sont pas applicables si les maximums des observations sont asymptotiquement indépendants. Pour résoudre ce problème, Ledford et Tawn (1996) [96] ont proposé un sous-modèle dans lequel la dépendance de la queue est caractérisée par un paramètre additionnel appelé coefficient de dépendance de queue noté  $\eta$  et satisfaisant  $\eta \in (0, 1]$ , et les plus grandes valeurs de cela indiquent une dépendance extrême forte asymptotiques, il s'est avéré être un outil utile pour décrire le comportement de queue d'une distribution bivariée (pas nécessairement, fonction de distribution F max-stable).

La structure de dépendance entre les marges de Fréchet  $(Z_1, Z_2)$  est analysé en utilisant le modèle de *Ledford et Tawn (1997)* suivant :

$$P(Z_1 > z_1, Z_2 > z_2) = z_1^{-c_1} z_2^{-c_2} \mathcal{L}(z_1, z_2) \text{ avec } c_1, c_2 > 0, \quad (4.1)$$

où  $\eta = (c_1 + c_2)^{-1}$  et la fonction  $\mathcal{L}$  est une fonction à variation lente bivariée, i.e.

$$\mathcal{L}(z_1, z_2) \sim g_1(z_1, z_2) (1 + g_2(z_1, z_2) z_1^{\rho_1} z_2^{\rho_2}) \text{ quand } z_1, z_2 \rightarrow \infty, \quad (4.2)$$

Une extension de la condition de Hall univariée au cas bivarié pourrait être faite en remplaçant les constantes dans la classe de Hall univariée par des fonctions homogènes de l'ordre 0, ceci mène à la condition de Hall bivariée.

La fonction  $g$  est homogène d'ordre 0, cela signifie par définition que  $g(tz_1, tz_2) = g(z_1, z_2)$ .

Cela implique qu'il existe une fonction  $g^*$  telle que

$$g(z_1, z_2) = g^*(z_1/(z_1 + z_2)).$$

Le modèle 4.1 représente le cas de dépendance asymptotique si  $\eta = 1$  et  $\mathcal{L}(t) \rightarrow 0$ . Si

$\eta = 1$  et  $\mathcal{L}(t) = 1$ ,  $Z_1$  et  $Z_2$  sont parfaitement dépendants; si  $\eta = 1/2$  et  $\mathcal{L}(t) = 1$   $Z_1$  et  $Z_2$  sont parfaitement indépendantes. Dans le cas  $\eta < 1$  (ou  $\eta = 1$  et  $\mathcal{L}(t) \rightarrow 0$ ) 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  $\mathcal{L}(t) \geq 1$ .

Pour certains exemples du modèle de Ledford et de Tawn, voir Ledford et Tawn (1997) [97].

Différentes méthodes pour estimer le coefficient de dépendance de la queue sont proposées voir Peng (1999) [113], Draisma et al (2004) [43] et Beirlant et Vandewalle (2002) [5] pour n'en nommer que quelques uns..

Ledford et Tawn ont proposé d'abord de standardiser la distribution marginale à Fréchet unitaire en utilisant soit les distributions marginales empiriques (i.e. en utilisant les rangs des composantes), puis estimer  $\eta$  comme paramètre de forme par un estimateur classique pour l'indice des valeurs extrêmes, par exemple l'estimateur de Hill [86] ou l'estimateur de moment.

Goegebeur et Guillou (2013) [72] ont obtenu un estimateur asymptotiquement sans biais en prenant une somme correctement pondérée de deux estimateurs biaisés pour  $\eta$ . Dutanga et al. (2014) [48] ont introduit un estimateur robuste et asymptotiquement sans biais pour  $\eta$ , cet estimateur est obtenu en ajustant un modèle de second ordre aux données au moyen du critère de divergence de puissance de densité minimale.

Dans l'article de Beirlant et al. (2011) [12], un estimateur asymptotiquement non biaisé pour  $\eta$  a été proposé, basé sur l'ajustement de la distribution de Pareto étendue avec la méthode du maximum de vraisemblance des variables aléatoires correctement transformées. Dans cette thèse, nous utilisons cet estimateur pour le coefficient de dépendance de queue.

Ce chapitre est organisé comme suit. Dans la section 4.2 nous présentons l'estimation de la queue bivariée et l'estimation du coefficient de dépendance de queue. Dans la section 4.3, nous avons introduit certaines des méthodes proposées pour équilibrer entre le biais et la variance afin d'obtenir un nombre optimal  $k_{opt}$  de statistique d'ordre. Dans la section 4.4, nous effectuons une étude de simulation pour comparer trois estimateurs différents du coefficient de dépendance de queue avec une attention particulière sur leur erreurs moyennes quadratiques.

Draisma et al (2004) [43] ont interprété une extension du modèle de Ledford et de Tawn comme une condition de variation régulière bivariée du second ordre (SO), cette condition SO n'est pas trop restrictive et elle est couramment utilisée dans les problèmes d'estimation impliquant un cas bivarié.

**Condition de variation régulière bivariée du second ordre :** Soit  $(X, Y)$  un vecteur aléatoire de la fonction de répartition conjointe  $F$  et de fonctions de distributions marginales continues  $F_X$  et  $F_Y$  tel que

$$\lim_{t \rightarrow 0^-} q_1(t)^{-1} \left( \frac{P(1 - F_X(X) < tx \text{ et } 1 - F_Y(Y) < ty)}{q(t)} - c(x, y) \right) =: c_1(x, y) \quad (4.3)$$

existe pour toutes  $x, y \geq 0$  avec  $x + y > 0$ , une fonction positive  $q$  et une fonction  $q_1$  tendant à zéro quand  $t \rightarrow 0^-$ , et  $c_1$  une fonction ni constante ni multiple de  $c$ . En outre, nous supposons que la convergence est uniforme sur  $\{(x, y) \in [0, \infty[^2 | x^2 + y^2 = 1\}$ , telle que  $c_1$  est continue et  $c_1(x, x) = x^{1/\eta}(x^\tau - 1)/\tau_1$ .

Essentiellement, la relation 4.3 est une condition de variation régulière de second ordre pour la fonction  $Q$  définie par  $Q(x, y) := P(1 - F_X(X) < x \text{ et } 1 - F_Y(Y) < y)$ . La fonction  $(x, y) \rightarrow Q(1 - x, 1 - y)$  est parfois appelée la copule de survie. Il s'ensuit que la fonction  $q(t) := P(1 - F_X(X) < t, 1 - F_Y(Y) < t)$ . Dans le papier de Ledford et Tawn (1996)  $q(t)$  est égal à  $t^{1/\eta}$ . Les fonctions  $q$  et  $|q_1|$  varient régulièrement en zéro avec l'indice  $1/\eta$  et  $\tau \geq 0$ , respectivement. La fonction  $c$  homogène d'ordre  $1/\eta$ , i.e.



$$c(tx, ty) = t^{1/\eta}c(x, y).$$

**Démonstration.** Voir de Hann et Ferreira (2006) [79], théorème 7.6.1, page 276.

Le paramètre  $\eta$  a la même signification que dans Ledford et Tawn (1996) et la condition (4.3) est similaire à la condition (4.2).

Pour plus d'informations sur la condition de second ordre, voir de Haan et Ferreira (2006), Section 7.6 et Draisma et al (2001).

En outre, ils supposent que  $l := \lim_{t \rightarrow 0^-} q(t)/t$  existe. Cette condition est toujours satisfaite si  $\eta < 1$  et  $\tau > 0$  de plus,  $l = 0$  si  $\eta < 1$  et  $l > 0$  si les marginales sont asymptotiquement dépendant.

**Lemme 4.1.1** *Le modèle 4.1 satisfait la condition SO. Beaucoup ont utilisé la fonction de répartition conjointe satisfaisant le modèle 4.1). Notons que ce modèle est en fait une condition de la fonction copule C. En effet, on vérifie facilement que*

$$P(1 - F_X(X) < x, 1 - F_Y(Y) < y) = x + y - 1 + C(1 - x, 1 - y).$$

Nous donnons maintenant quelques exemples de distributions satisfaisant la condition SO.

### 4.1.1 Distribution de Farlie Gumbel Morgenstern

La fonction copule de Farlie Gumbel Morgenstern est donnée par :

$$C(u, v) = uv[1 + \alpha(1 - u)(1 - v)], \quad (u, v) \in [0, 1]^2,$$

où  $\alpha \in [-1, 1]$ . Des calculs simples conduisent à

$$P(1 - F_X(X) < tx, 1 - F_Y(Y) < ty) = t^2xy[1 + \alpha - \alpha t(x + y) + \alpha t^2xy].$$

Dans le cas où  $\alpha \in [-1, 1]$

$$\frac{P(1 - F_X(X) < tx, 1 - F_Y(Y) < ty)}{P(1 - F_X(X) < t, 1 - F_Y(Y) < t)} = xy \left[ 1 - \frac{\alpha t}{1 + \alpha}(x + y - 2) + O(t^2) \right],$$

de lequel un on vérifie facilement que la condition SO est satisfaite avec  $\eta = 0.5$ ,  $c(x, y) = xy$ ,  $c_1(x, y) = xy(x + y - 2)/2$ ,  $q_1(t) = -2\alpha t/(1 + \alpha)$ , donc  $\tau = 1$ .

Pour le cas  $\alpha = -1$

$$\frac{P(1 - F_X(X) < tx, 1 - F_Y(Y) < ty)}{P(1 - F_X(X) < t, 1 - F_Y(Y) < t)} = xy \left[ \frac{x + y}{2} + \frac{1}{4}(x + y - 2xy) + O(t^2) \right],$$

Et donc la condition SO est satisfaite avec  $\eta = 1/3$ ,  $c(x, y) = xy(x + y)/2$ ,  $c_1(x, y) = xy(2xy - x - y)/2$ ,  $q_1(t) = -t/2$ , alors  $\tau = 1$ .

### 4.1.2 Distribution normale bivariée

La distribution normale bivariée avec une moyenne nulle, variance vaut 1 et coefficient de corrélation  $\rho \in ]-1, 1[$  satisfait la condition SO avec  $\eta = (1 + \rho)/2$ . Nous considérons  $\rho = -0.5$  pour  $\eta = 0.25$ . Nous nous référons à Ledford and Tawn (1996) [96] et Draisma et al (2004) [43] pour plus de détails.

## 4.2 Estimation du coefficient de dépendance de queue

Beirlant et al (2009) [11] ont introduit la distribution de Pareto étendue comme modèle approximatif pour les excès relatifs sur un seuil. Nous utilisons l'estimation de la queue bivariée suggérée par Beirlant et al (2011) et nommée estimateur à biais réduit, cette estimation pour le coefficient de dépendance de queue basée sur la condition de second ordre et elle est résumé comme suit :

$$P(Z_1 > z_1, Z_2 > z_2) = P(\min(Z_1, Z_2 w/(1 - w)) > z_1) =: P(Y_w > z_1).$$

Où  $w = \frac{z_1}{z_1 + z_2}$  est un paramètre du réglage.

Sous 4.2, nous avons :

$$P(Y_w > z_1) = z_1^{-1/\eta} C_w \left[ 1 + \frac{1}{\eta} \delta_w(z_1) \right], \quad (4.4)$$

où

$$C_w = \left( \frac{w}{1-w} \right)^{c_2} g_1^*(w),$$

et

$$\delta_w = \eta g_2^*(w) \left( \frac{w}{1-w} \right)^{-\rho_2} z_1^\tau (1 + o(1)),$$

avec  $tz = z_1$  et  $t$  est un seuil convenablement choisi.

$$\begin{aligned} P(Z_1 > z_1, Z_2 > z_2) &= \frac{P(Y_w > tz)}{P(Y_w > t)} P(Y_w > t) \\ &= \left( z^{-1/\eta} \frac{1 + \delta_w(tz)/\eta}{1 + \delta_w(t)/\eta} \right) P(Y_w > t). \end{aligned}$$

Comme démontré dans Beirlant et al (2009), sous 4.4,  $z^{-1/\eta} \frac{1 + \delta_w(tz)/\eta}{1 + \delta_w(t)/\eta}$  peut être se approché par

$$\{z(1 - \delta_w(t) - \delta_w(t) z^\tau)\}^{-1/\eta} =: \overline{G}_{\eta, \delta_w, \tau}(z) \quad (4.5)$$

Le modèle  $\overline{G}_{\eta, \delta_w, \tau}$  a été introduit par Beirlant et al (2009) [11], et nommé par la distribution de Pareto étendue (EPD). Il s'ensuit que

$$P(Z_1 > z_1, Z_2 > z_2) = \overline{G}_{\eta, \delta_w, \tau}(z_1/t) P(Y_w > t) + o(|\delta_w(t)| P(Y_w(t))).$$

Il est bien connu que le choix du seuil  $t$  est un problème difficile même si les estimateurs ont une forme explicite, le nombre d'excès sur  $t$  doit être suffisamment grand pour rendre l'inférence possible.

Le choix d'un seuil de données adaptatif  $t = Y_{n-k,n}$  dans la séquence ordonnée  $Y_{1,n} \leq Y_{2,n} \leq \dots \leq Y_{n,n}$  des valeurs observées de  $Y$  peut garantir les deux critères en choisissant  $k \rightarrow \infty$  et  $k/n \rightarrow 0$  quand  $n \rightarrow \infty$ .

Pour plus de détails sur le choix du nombre de statistiques d'ordre extrême, voir Cheng et Peng (2001) [22], Neves and Fraga (2004) [110], Danielsson et al (2001) [28], Dekkers et al (1993) [41], Draisma et al (1999) [42] et Drees and Kaufmann (1998) [47].

**Définition 4.2.1 (EPD)** *EPD avec le vecteur de paramètre  $(\eta, \delta, \tau)$  dans la gamme  $\tau < 0 < \eta$  et  $\delta > \max(-1, 1/\tau)$  est définie par sa fonction de distribution suivante*

$$\bar{G}_{\eta, \delta, \tau}(z) = \begin{cases} 1 - \{z(1 + \delta z)^\tau\}^{-1/\eta} & \text{si } z > 1 \\ 0 & \text{si } z \leq 1. \end{cases} \quad (4.6)$$

### 4.2.1 Estimation des paramètres EPD

L'estimation des distributions de queue bivariées nécessite d'abord l'estimation des paramètres EPD, voir Beirlant et al (2009) [11]. Depuis

$$P(Y_w > tz \setminus Y_w > t) = \bar{G}_{\eta, \delta_w, \tau}(z) = o(|\delta_w(t)|).$$

L'estimation des paramètres EPD est basée sur les excès relatifs  $Y/t$  lorsque  $Y > t$ . Comme indiqué dans Beirlant et al (2009), la distribution asymptotique de l'estimateur de queue ne dépendra pas de la distribution asymptotique de l'estimateur de  $\tau$ . Par conséquent, le paramètre du second ordre inconnu sera remplacé par un estimateur cohérent qui n'influera pas sur les distributions asymptotiques des autres estimateurs. En pratique, le paramètre  $\tau$  est souvent remplacé par  $-1$ .

#### Estimation du maximum de vraisemblance (MLE)

Les estimateurs de  $\eta$  et  $\delta$  seront trouvés en maximisant une approximation de la vraisemblance EPD étant donné l'échantillon de  $k$  excès relatifs  $Y_{n-k+i,n}/Y_{n-k,n}$ ,  $i = 1, \dots, k$ . La log-vraisemblance de l'EPD est donnée par :

$$\log L(z) = -\log \eta - (1/\eta + 1) \log z - (1/\eta + 1) \log(1 + \delta - \delta z^\tau) + \log(1 + \delta - \delta z^\tau - \delta \tau z^\tau).$$

Les dérivées partielles

$$\begin{aligned}\hat{\eta}_B &= \frac{\log L(z)}{\delta\eta} = -\frac{1}{\eta} + \frac{1}{\eta^2} \log x + \frac{1}{\eta^2} (1 + \delta - \delta z^\tau) \\ \hat{\delta}_B &= \frac{\log L(z)}{\delta\delta} = -(1/\eta + 1) \frac{1 - z^\tau}{1 + \delta - \delta z^\tau} + \frac{1 - z^\tau - \tau z^\tau}{1 + \delta - \delta z^\tau - \delta\tau z^\tau}.\end{aligned}$$

définie

$$\begin{aligned}\hat{\eta}_B &= H_{k,n} = \frac{1}{k} \sum_{j=1}^k \log(Y_{n-k+j,n}/Y_{n-k,n}), \\ \hat{\delta}_B &= E_{k,n}(s) = \frac{1}{k} \sum_{j=1}^k \log(Y_{n-k+j,n}/Y_{n-k,n})^s, \quad s \leq 0.\end{aligned}\tag{4.7}$$

Notons que  $H_{k,n}$  est l'estimateur de Hill [86].

Un autre estimateur asymptotiquement équivalent pour  $\eta$  trouvé dans Beirlant et al (2009) où les auteurs ont proposé des estimateurs simplifiés pour le modèle basé sur les fonctions EPD qui sont linéarisées dans  $\delta$  par conséquent

$$\begin{aligned}\hat{\delta}_L &= H_{k,n}(1 - 2\hat{\beta})(1 - \hat{\beta})^3\hat{\beta}^{-4} \left( \frac{1}{k} \sum_{j=1}^k (Y_{n-k+j,n}/Y_{n-k,n})^{\hat{\beta}/H_{k,n}} - \frac{1}{1-\hat{\beta}} \right), \\ \hat{\eta}_L &= H_{k,n} - \hat{\delta} \frac{\hat{\beta}}{1-\hat{\beta}},\end{aligned}\tag{4.8}$$

où  $\hat{\beta}$  est une suite d'estimateurs faiblement cohérente pour  $\beta = \eta\tau$ . Beirlant et al (2009) ont utilisé l'estimateur introduit dans Fraga et al. (2003) [60].

### Estimation de la probabilité rare

Estimant  $P(Y_w > t)$  par la proportion empirique  $k/n$ , on obtient alors

$$\hat{p} = \frac{k}{n} \overline{G}_{\hat{\eta}, \hat{\delta}, \hat{\tau}}(z_1/Y_{n-k,n}).\tag{4.9}$$

Lorsque l'on omet la partie de second ordre du modèle ( $\delta_w$  et  $\tau$ ) on obtient un estimateur classique de Weissman (1978) [132] pour la probabilité de queue bivariee :

$$\hat{p}_W = \frac{k}{n} (z_1/Y_{n-k,n})^{-1/\hat{\eta}}.\tag{4.10}$$

### 4.3 Sélection optimale de la fraction d'échantillon

Dans cette section, nous commençons par la méthode graphique pour la sélection optimale de la fraction d'échantillon et nous notons que lorsque  $k$  est petit, la variance est grande et l'utilisation d'une grande valeur de  $k$  introduit un biais important dans l'estimation. Ensuite, la valeur optimale  $k$  peut être obtenue par minimisation de l'erreur quadratique moyenne de l'estimateur considéré. Récemment, des procédures adaptatives pour la sélection automatisée de la fraction d'échantillon optimale proposée pour calculer  $\hat{k}_{opt}$  pour  $k_{opt}$  dans le sens  $\hat{k}_{opt}/k_{opt} \rightarrow 1$  quand  $n \rightarrow \infty$  comme les méthodes de Bootstrap voir [28], [42] et les procédures séquentielle voir [47]. Nous allons appliquer la méthode de Reiss et Thomas et la méthode de Cheng et Peng. Certaines simulations sont discutées dans la section 4 avec une comparaison entre les deux dernières méthodes, dans laquelle nous rappelons le critère sur  $k = k_n$ .

$$k \rightarrow \infty, 1 \leq k \leq n \text{ et } k/n \rightarrow 0 \text{ quand } n \rightarrow \infty, \quad (4.11)$$

#### 4.3.1 Approche graphique

L'approche graphique est une méthode universelle et doit être appliquée avant toutes les recherches numériques. Elle traite notre problème en sélectionnant un certain nombre de statistiques d'ordre extrêmes. Cette méthode consiste à dessiner les points de coordonnées

$$\{(k, \hat{\eta}(k)) : k = 1, \dots, n\},$$

pour faire un choix optimal de  $k$ . La sélection de  $k_{opt}$  est illustrée graphiquement dans la figure 4.1.

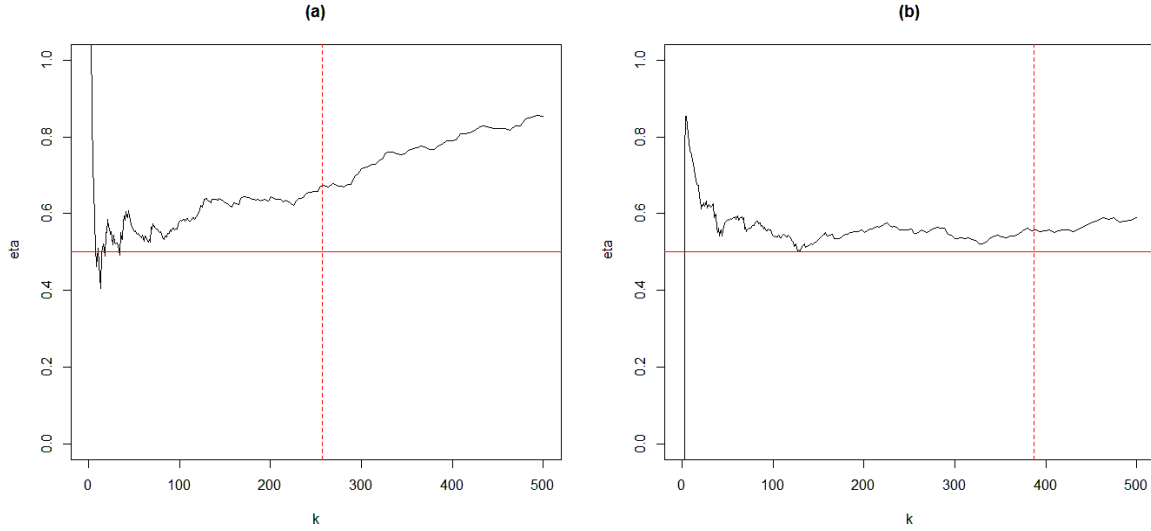


FIG. 4.1 – Estimateur de  $\eta$  pour la distribution Farlie-Gumbel-Morgenstern, basée sur 500 échantillons de taille 1000 (a) estimateurs de Hill, (b) estimateurs de Moment. La ligne horizontale correspond à la vraie valeur de  $\eta$  et la ligne verticale correspond  $k_{opt}$ .

### 4.3.2 Erreur moyenne quadratique

L'erreur moyenne quadratique est très utile pour comparer plusieurs estimateurs, notamment lorsque l'un d'eux est biaisé. Si les deux estimateurs à comparer sont sans biais l'estimateur le plus efficace est simplement celui qui a la variance la plus petite., Le  $MSE$  d'un estimateur du coefficient de dépendance de queue  $\hat{\eta}$  est défini par :

$$MSE(\hat{\eta}) := E_{\infty}(\hat{\eta} - \eta)^2,$$

où  $E_{\infty}$  dénote l'espérance en ce qui concerne la distribution de la limite.

Il est nécessaire de faire un compromis entre le biais et la variance. Il semble raisonnable que le MSE permet de réduire au minimum pour une estimation plus précise possible. Le choix optimal de  $k$ , noté  $k_{opt}$ , correspond à la plus petite  $MSE$ , i.e.

$$k_{opt} := \arg \min_k MSE(\hat{\eta}). \quad (4.12)$$

### 4.3.3 Approche de Reiss et Thomas

Reiss et Thomas (1997) [110] ont basé leur approche de choisir le nombre adéquat de plus grandes observations sur un moyen de minimiser la distance résumant un terme de pénalité. Dans un certain sens, ce coefficient est prévu pour être plus sévère en ce qui concerne des estimations de  $\eta$  avec l'origine dans les observations prises plus loin de la queue réelle.

Ils proposent une manière automatique de choisir  $\hat{k}_{opt}$  en minimisant

$$\frac{1}{k-1} \sum_{i \leq k} i^\beta |\hat{\eta}_n(i) - \text{median}(\hat{\eta}_1, \dots, \hat{\eta}_n)|, \quad 0 \leq \beta \leq 1/2. \quad (4.13)$$

Reiss et Thomas ont aussi suggéré de minimiser la modification de (4.13) :

$$\hat{k}_{opt} = \arg \min_k \frac{1}{k-1} \sum_{i < k} i^\beta (\hat{\eta}_i - \hat{\eta}_k)^2, \quad 0 \leq \beta \leq 1/2. \quad (4.14)$$

En ce qui concerne la connaissance de  $\beta$ , pas de méthodologie concernant la spécification a été disponible, mais il faut trouver des bons ensembles de poids dans lesquels  $\hat{k}_{opt} = \hat{k}_{opt}(\beta)$  capable de localiser l'estimation raisonnables de  $\eta$ .

### 4.3.4 Approche de Cheng et Peng

#### **Proposition 4.3.1 (Approximation normale asymptotique de l'estimateur de Hill)**

Supposer que 1.24 est satisfaite et  $k \rightarrow \infty$ ,  $k/n \rightarrow 0$  on a

$$\sqrt{k} (\hat{\gamma}_n^{(H)} - \gamma) \xrightarrow{d} \mathcal{N}(0, \gamma^2) \quad \text{quand } n \rightarrow \infty,$$

ssi  $k = o(n^{-2\rho/(1-2\rho)})$ .

Une constatation surprenante est que l'ordre de précision pour une couverture optimale de la face d'un intervalle de confiance dépend du signe de la variation régulière du second



ordre. Ainsi, pour  $0 < \alpha < 1$  les intervalles de confiance unilatéral et bilatéral pour  $\gamma$  de niveau  $(1 - \alpha)$  sont respectivement

$$I_1(\alpha) = \left[ 0, \hat{\gamma}_n^{(H)} + \frac{z_\alpha \hat{\gamma}_n^{(H)}}{\sqrt{k}} \right]$$

et

$$I_2(\alpha) = \left[ \hat{\gamma}_n^{(H)} - z_{\alpha/2} \frac{\hat{\gamma}_n^{(H)}}{\sqrt{k}}, \hat{\gamma}_n^{(H)} + z_{\alpha/2} \frac{\hat{\gamma}_n^{(H)}}{\sqrt{k}} \right],$$

où  $z_w$  ( $0 \leq w \leq 1$ ) est défini par  $P(\mathcal{N}(0, 1) \leq z_w) = 1 - w$ .  $z_w$  est le quantile d'ordre  $(1 - w)$  de la distribution normale standard.

Nous donnons la précision de la couverture des intervalles de confiance  $I_1(\alpha)$  et  $I_2(\alpha)$  et le choix optimal théorique de fraction de l'échantillon  $k$  pour  $I_1(\alpha)$  dans le sens de minimiser l'erreur absolue de couverture.

**Théorème 4.3.1** *Supposons que 1.24 satisfait et  $k \rightarrow \infty$ ,  $k/n \rightarrow 0$ . Alors*

$$P(\gamma \in I_1(\alpha)) = \alpha - \phi(z_\alpha) \left\{ \frac{1 + 2z_\alpha^2}{3\sqrt{k}} - \frac{\rho dc^\rho}{(1 - \rho)} \sqrt{k} \left(\frac{n}{k}\right)^\rho \right\} + o\left(\frac{1}{\sqrt{k}} + \sqrt{k} \left(\frac{n}{k}\right)^\rho\right) \quad (4.15)$$

et

$$P(\gamma \in I_2(\alpha)) = \alpha + o\left(\frac{1}{\sqrt{k}} + \sqrt{k} \left(\frac{n}{k}\right)^\rho\right). \quad (4.16)$$

où  $\phi$  désigne la fonction de densité normale standard.

Par conséquent la valeur optimale de  $k$  qui minimise l'erreur absolue de la couverture pour  $I_1(\alpha)$  est

$$k_{opt} := \begin{cases} \left( \frac{(1 + 2z_\alpha^2)(1 - \rho)}{-3dc^\rho \rho(1 - 2\rho)} \right)^{1/(1-\rho)} n^{-\rho/(1-\rho)} & \text{si } d > 0, \\ \left( \frac{(1 + 2z_\alpha^2)(1 - \rho)}{3dc^\rho \rho} \right)^{1/(1-\rho)} n^{-\rho/(1-\rho)} & \text{si } d < 0, \end{cases} \quad (4.17)$$

qui satisfait automatiquement la condition  $k = o(n^{-2\rho/(1-2\rho)})$  dans la proposition 4. 3.1.

En outre, la précision de la couverture optimale pour  $I_1(\alpha)$  est

$$P(\gamma \in I_1(\alpha)) = \begin{cases} \alpha - 2 \left\{ \frac{(1-\rho)(1+2z_\alpha^2)}{3(1-2\rho)} \right\}^{(1-2\rho)/(2(1-2\rho))} \left\{ \frac{-d\rho}{c^{-\rho}} \right\} \\ \quad \times \phi(z_\alpha) n^{\rho/(2(1-2\rho))} (1 + o(1)) & \text{si } d > 0 \\ \alpha + o(n^{\rho/(2(1-2\rho))}) & \text{si } d < 0. \end{cases} \quad (4.18)$$

Cheng et peng ont proposé un estimateur plug-in (numéric) pour la fraction de l'échantillon optimale en se concentrant sur un intervalle de confiance unilatéral (puisque  $k$  dépend de quelques quantités inconnues), donc

$$\hat{k}_{opt} := \begin{cases} \left( \frac{(1+2z_\alpha^2)}{3\hat{\delta}(1+2\hat{\rho})} \right)^{1/(1+\hat{\rho})} n^{\hat{\rho}/(1+\hat{\rho})} & \text{si } \hat{\delta} > 0, \\ \left( \frac{(1+2z_\alpha^2)}{-3\hat{\delta}} \right)^{1/(1-\hat{\rho})} n^{-\hat{\rho}/(1-\hat{\rho})} & \text{si } \hat{\delta} < 0, \end{cases}$$

où

$$\hat{\rho} := -\log \left( \left| \frac{M_n^{(2)}(n/(2\sqrt{\log n})) - 2 \left\{ M_n^{(1)}(n/(2\sqrt{\log n})) \right\}^2}{M_n^{(2)}(n/\sqrt{\log n}) - 2 \left\{ M_n^{(1)}(n/\sqrt{\log n}) \right\}^2} \right| \right) / \log 2,$$

et

$$\hat{\delta} := (1 + \hat{\rho}) (\log n)^{\hat{\rho}/2} \frac{M_n^{(2)}(n/\sqrt{\log n}) - 2 \left\{ M_n^{(1)}(n/\sqrt{\log n}) \right\}^2}{2\hat{\rho} \left\{ M_n^{(1)}(n/\sqrt{\log n}) \right\}^2},$$

avec  $M_n^{(2)}$  et  $M_n^{(1)}$  sont l'estimateur de Moment et l'estimateur de Hill respectivement.

## 4.4 Simulation

Dans l'expérience de simulation, nous comparons le nombre optimal de statistiques d'ordre supérieur utilisées dans l'estimation du coefficient de dépendance de queue (au sens de

l'erreur moyenne quadratique) obtenu avec deux méthodes listées dans la section 4.3.

Nous examinons le comportement de plusieurs estimateurs (l'estimateur linéaire donné dans la formule (4.8), l'estimateur de Hill (voir l'équation 2.3) et l'estimateur de Moment (voir l'équation ??), pour différentes distributions, avec les marginales de Fréchet. Pour chacune des distributions, nous avons généré 500 échantillons de taille 1000 et sélectionner le nombre de statistiques d'ordre supérieur utilisées dans l'estimation et nous calculons les estimateurs de  $\eta$ .

Le paramètre  $w$  a été introduit pour estimer les probabilités de queue dans les régions conjointe, mais peu de pertinence pratique pour l'estimation de  $\eta$ , donc nous le fixons à 0,5.

Premièrement, nous avons effectué des simulations à partir d'une distribution Farlie-Gumbel-Morgenstern (FGM) avec les marges de Fréchet ( $\alpha = -1$ ), et de la distribution Normale avec les marges de Fréchet ( $\rho = -0.5$ ).

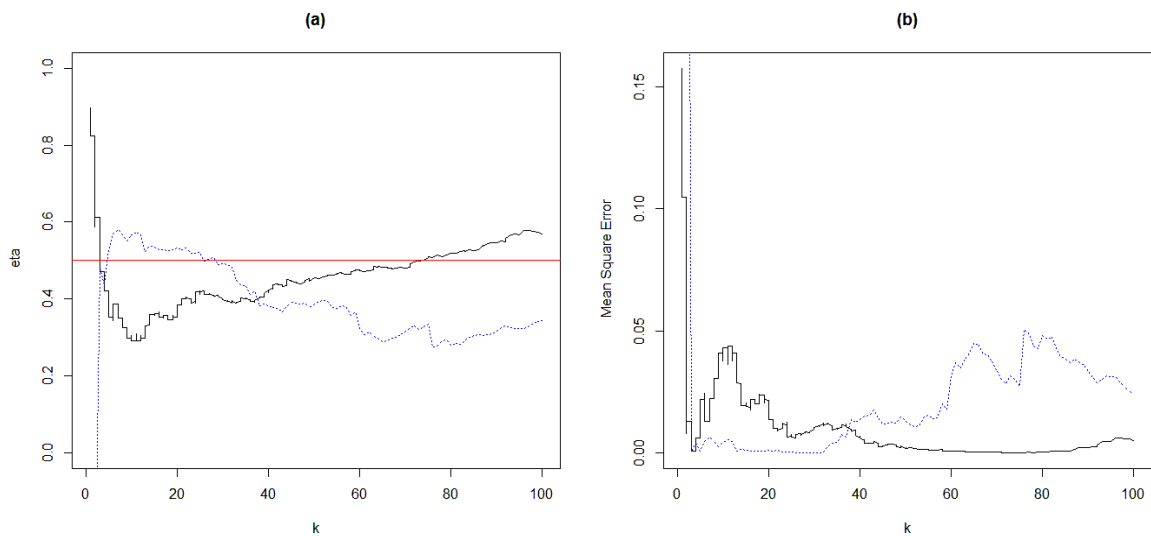


FIG. 4.2 – Copule de Farlie Gumbel Morgenstern avec  $\alpha = -1$  : (a) estimateurs pour  $\eta$ , (b) MSE : (ligne complète) Estimateur de Hill, (ligne pointillée) Estimateur de moment, basé sur 1000 simulations. La ligne horizontale correspond à la vraie valeur de  $\eta$ .

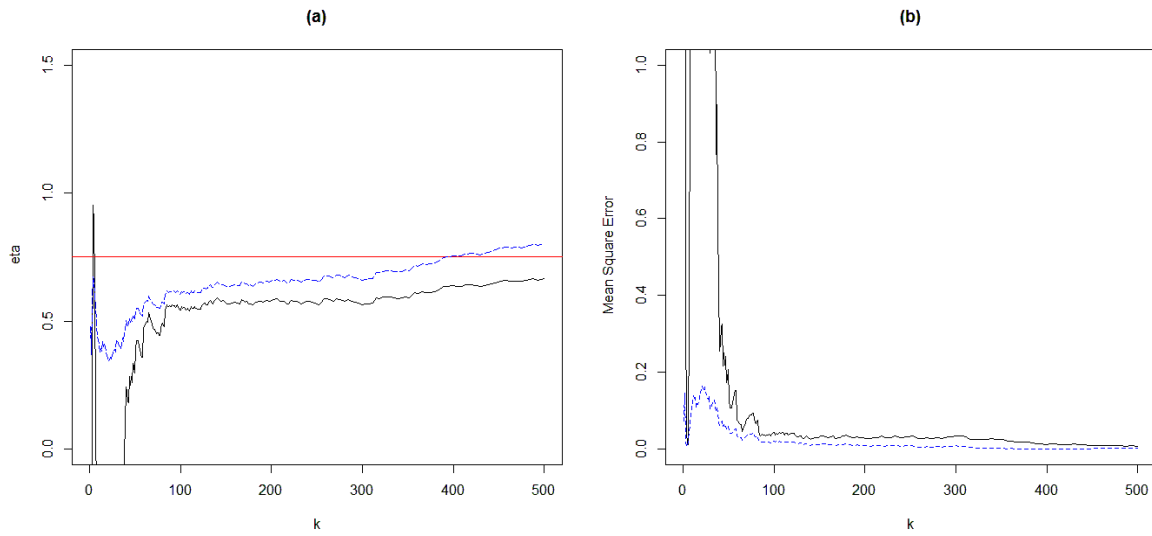


FIG. 4.3 – Copule de Farlie Gumbel Morgenstern avec  $\alpha = 0.75$  : (a) estimateurs pour  $\eta$ , (b) MSE : (ligne complète) Estimateur linéaire, (ligne pointillée) Estimateur de Hill, basé sur 1000 échantillons. La ligne horizontale correspond à la vraie valeur de  $\eta$ .

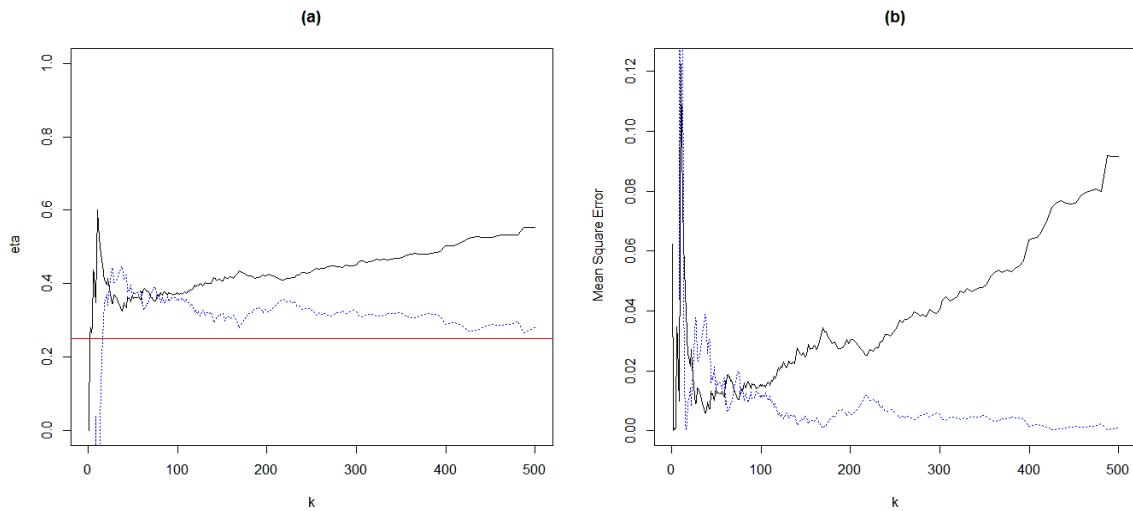


FIG. 4.4 – Copule Normale bivarié avec  $\rho = -0.5$  : (a) estimateurs pour  $\eta$ , (b) MSE : (ligne complète) Estimateur de Hill, (ligne pointillée) Estimateur de moment, basé sur 1000 échantillons. La ligne horizontale correspond à la vraie valeur de  $\eta$ .

Ensuite, nous examinons le comportement de l'échantillon des estimateurs pour différentes distributions. Pour chacune des distributions, nous calculons le MSE, la vraie valeur de  $\eta$

et  $k_{opt}$ . Dans le Tableau 4.1 nous présentons quelques résultats de simulation comme FGM ( $\alpha = 0.6$ ), Normal bivariée ( $\rho = -0.5$ ) et Frank ( $\alpha = 1$ ) basé sur 100 échantillons simulés de tailles  $n = 300, 500$  et  $1000$ .

<b>Distribution FGM</b>						
	$\eta$	$\hat{\eta}_H$ $K_{opt}$	$MSE$		$\hat{\eta}_M$ $K_{opt}$	$MSE$
$n = 300$	0.5008	179	$6.2821 \times 10^{-7}$	0.5001	193	$1.1189 \times 10^{-8}$
$n = 500$	0.4945	257	$3.0081 \times 10^{-7}$	0.4993	387	$4.2669 \times 10^{-7}$
$n = 1000$	0.5008	294	$6.8282 \times 10^{-7}$	0.4996	957	$1.3123 \times 10^{-7}$
<b>Distribution Normale Bivariée</b>						
$n = 300$	0.2505	143	$2.8621 \times 10^{-6}$	0.2497	193	$5.2328 \times 10^{-7}$
$n = 500$	0.2486	229	$1.9785 \times 10^{-6}$	0.2494	390	$2.9245 \times 10^{-7}$
$n = 1000$	0.2521	336	$4.7927 \times 10^{-6}$	0.2521	890	$4.6871 \times 10^{-7}$
<b>Distribution de Frank</b>						
$n = 300$	0.4999	185	$2.6563 \times 10^{-7}$	0.4998	289	$2.6915 \times 10^{-8}$
$n = 500$	0.4976	321	$5.6776 \times 10^{-7}$	0.5001	346	$3.7154 \times 10^{-8}$
$n = 1000$	0.5080	501	$6.4849 \times 10^{-7}$	0.4993	691	$3.9494 \times 10^{-8}$

TAB. 4.1 – Résultats de simulation de la procédure de sélection le seuil pour l'estimateur de Hill et l'estimateur de Moment.

Dans le Tableau 4.2 nous présentons quelques résultats de simulation basés sur la procédure simplifiée de la méthode de Cheng & Peng et la méthode de Reiss & Thomas de la distribution la FGM bivariée ( $\alpha = -0.25$  et  $0.75$ ) et la distribution normale bivariée ( $\rho = -0.5$  et  $0.5$ ) avec les marges de Fréchet, basée sur 1000 échantillons simulés de taille  $n = 500$ . Les résultats de comparaison sont présentés dans les tableaux 4.2 et 4.3 et illustrés dans la Figure 4.5.

Nous appliquons la méthode de Reiss & Thomas pour l'estimateur Moment. Donc pour  $\hat{\eta}_M < 0$  nous prenons  $\beta = 0.35$  dans la version (4.14). Si  $\hat{\eta}_M \geq 0$ , sélectionnons  $\beta = 0.4$  dans la version (4.13). Voir le Tableau 4.3

En appliquant des méthodes adaptatives pour sélectionner une bonne valeur des statistiques d'ordre extrême (Cheng & Peng et Reiss & Thomas), nous sélectionnons la fraction

Distribution	$MSE$	$\hat{K}_{opt}$	$\hat{\eta}_H$
FGM $\alpha = 0.6$	$6.8282 \times 10^{-7}$	51	0.4784
FGM $\alpha = -0.25$	$2.1494 \times 10^{-7}$	38	0.4906
Bivariate normal $\rho = -0.5$	$1.7963 \times 10^{-7}$	25	0.2442
Bivariate normal $\rho = 0.25$	$4.7452 \times 10^{-7}$	60	0.7049

TAB. 4.2 – Résultats de simulation de la procédure de sélection de k pour l'estimateur de Hill par la méthode Cheng et Peng.

Distribution	$MSE$	$\hat{K}_{opt}$	$\hat{\eta}_M$
FGM ( $\alpha = 0.6$ )	$1.1667 \times 10^{-4}$	201	0.5532
FGM ( $\alpha = -0.25$ )	$1.1047 \times 10^{-4}$	260	0.2498
Bivariate normal ( $\rho = -0.5$ )	$1.1446 \times 10^{-4}$	246	0.1846
Bivariate normal ( $\rho = 0.5$ )	$1.1524 \times 10^{-7}$	335	0.7254

TAB. 4.3 – Résultats de simulation de la procédure de sélection de k pour l'estimateur de Hill par la méthode de Reiss et Thomas .

d'échantillon optimale  $k_{opt}$ , i.e. le nombre optimal d'extrêmes supérieurs utilisés dans le calcul de l'estimation du coefficient de dépendance de queue, les résultats sont résumés dans le tableau 4.4 et à partir de cet tableau, on peut dire que les deux méthodes appliquées pour déterminer k donnent des résultats très proches de la vraie valeur de  $\eta$ , mais la méthode de Cheng & Peng est plus rapide que l'autre méthode. La méthode de Reiss & Thomas est longue puisqu'elle est basée sur le calcul de la médiane.

Approche	# des extrêmes	% des extrêmes	Estimation/vraie $\eta$
Cheng et Peng	90	9%	0.5092/0.5
Reiss et Thomas	131	13.1%	0.4679/0.5

TAB. 4.4 – Nombre optimal de statistiques d'ordre extrêmes utilisées dans le calcul de l'estimation de Hill de la distribution des FGM (0,75), basée sur 1000 observations.

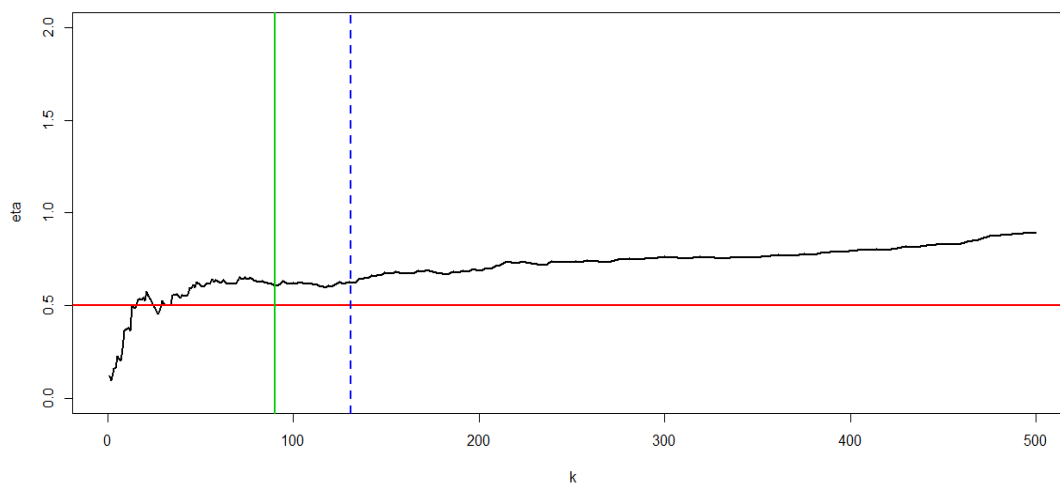


FIG. 4.5 – L'estimateur de Hill de la distribution FGM avec  $\alpha = 0.75$ , basé sur 1000 simulations. La ligne horizontale représente la valeur réelle de  $\eta$ , alors que les lignes verticales correspondent au nombre optimal d'extrêmes de Cheng & Peng (solide) et Reiss & Thomas (pointillés).

#### 4.4.1 Données réelles

Nous considérons ici l'estimation pour  $\eta$  dans le cas réel : Les données Loss-ALAE contient 1500 réclamations générales choisies aléatoirement à partir du bureau des services d'assurance. Chaque réclamation comprend un paiement d'indemnité (Loss,  $X$ ) et une charge d'ajustement de perte attribuée (ALAE,  $Y$ ), des exemples d'ALAE sont les honoraires versés aux avocats, aux experts et aux enquêteurs extérieurs utilisés pour défendre les réclamations.(pour plus de détails voir Frees and Valdez, 1998 [63]).

D'abord les données bivariées sont transformées en marges de Fréchet  $(Z_1, Z_2)$  en utilisant les fonctions de répartition empiriques  $\hat{F}_1(x)$ ,  $\hat{F}_2(x)$  comme suit:

$$Z_1 = -1/\log \hat{F}_1(X), \quad Z_2 = -1/\log \hat{F}_2(Y).$$

Les Figures 4.5 (a), 4.6 (a) et 4.7 (a), nous donnons des différents estimateurs pour  $\eta$  et nous notons que la signification de  $k_{opt}$  est différente de celle des estimateurs  $\hat{\eta}_H$ ,  $\hat{\eta}_M$  et  $\hat{\eta}_L$  respectivement.

Nous comparons deux estimateurs, Hill et Moment dans les données Loss-ALAE, les résultats dans les figures 4.9 (a) et leurs MSE (b) 4.9.

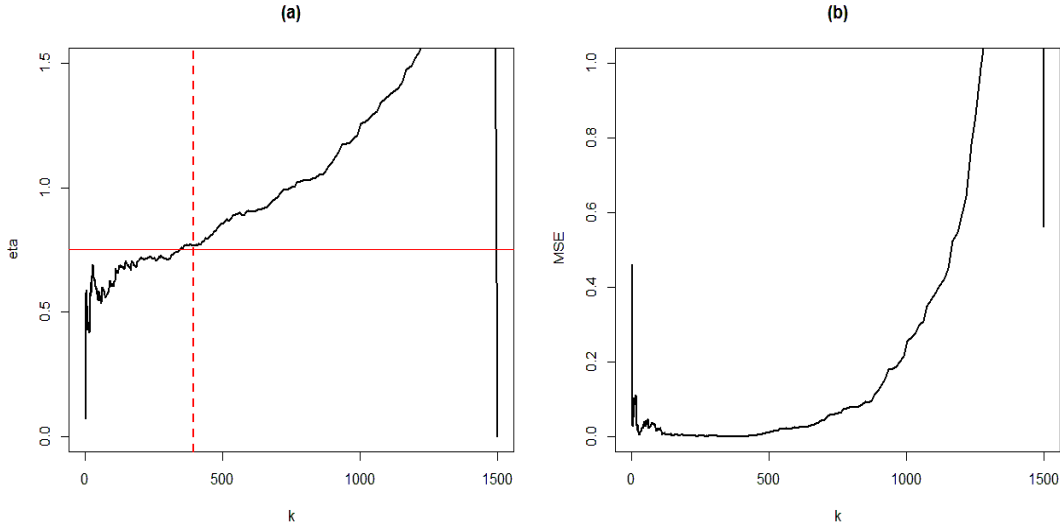


FIG. 4.6 – Données de loss-ALAE : (a) l'estimateur de Hill pour  $\eta$ , (b) MSE.

Nous illustrons la performance de  $k_{opt}$  dans une petite étude de simulation, les résultats résumés dans le tableau 4.5 où nous présentons également le pourcentage des extrêmes et le MSE.

Estimateur	# des extrêmes	% des extrêmes	Estimation	MSE
$\hat{\eta}_H$	419	27.93%	0.8019	$3.8823 \times 10^{-6}$
$\hat{\eta}_M$	484	32.26%	0.7990	$9.0713 \times 10^{-7}$
$\hat{\eta}_L$	454	30.26%	0.8052	$2.7151 \times 10^{-5}$

TAB. 4.5 – Nombres optimaux de statistiques d'ordre supérieur utilisées dans le calcul des estimations de coefficient de dépendance de la queue pour les données de loss-ALAE.



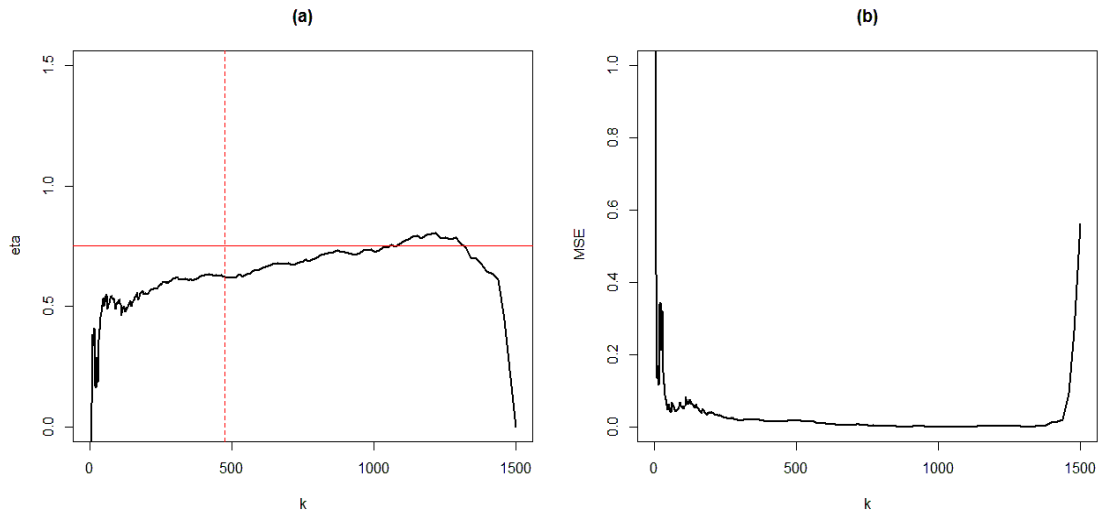


FIG. 4.7 – Données de loss-ALAE : (a) l'estimateur de Moment pour  $\eta$ , (b) MSE.

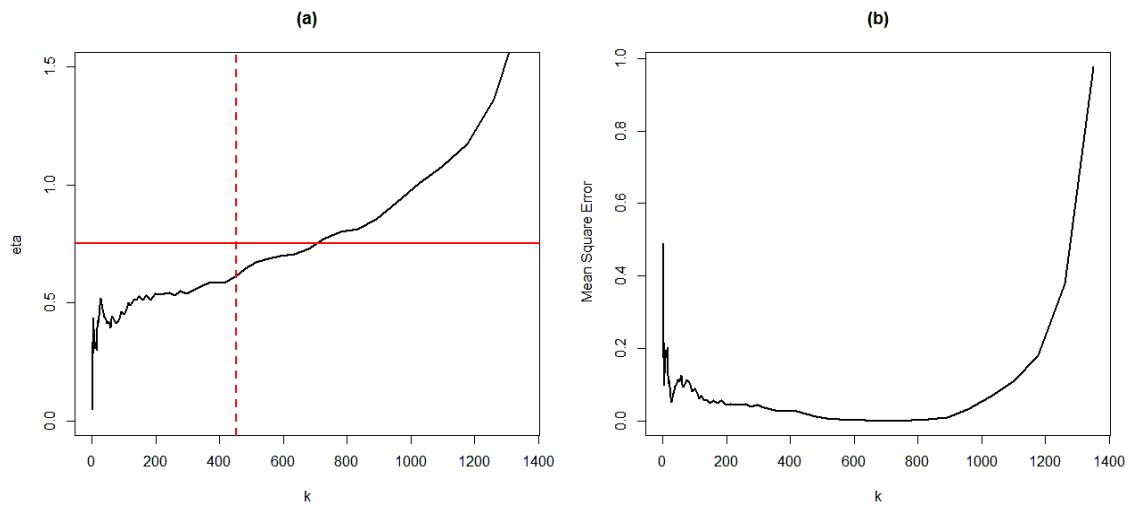


FIG. 4.8 – Données de loss-ALAE : (a) l'estimateur linéaire pour  $\eta$ , (b) MSE.

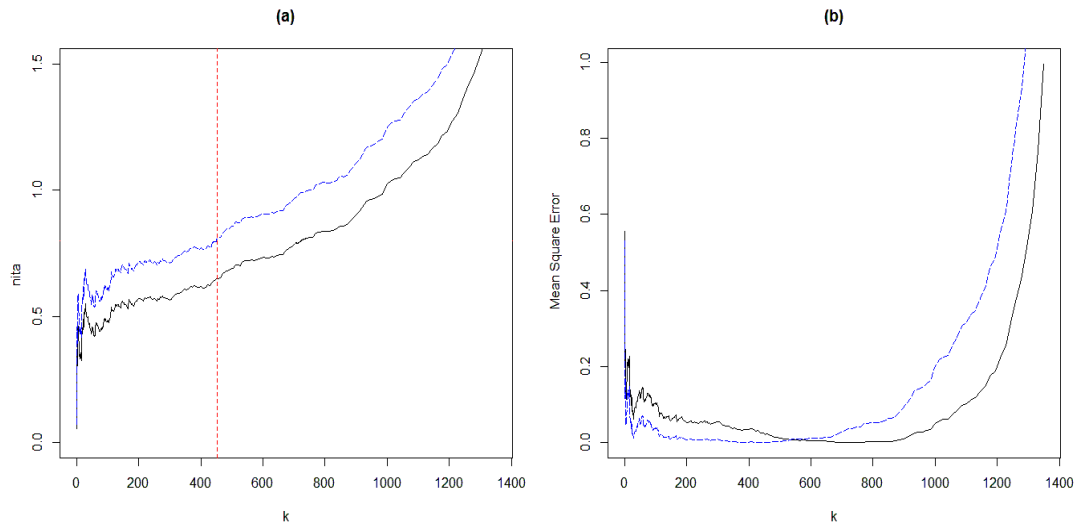


FIG. 4.9 – Données de loss-ALAE : (a) estimateurs pour  $\eta$ , (b) MSE : l'estimateur linéaire (Ligne complète), l'estimateur de Hill (ligne en pointillé).

## 4.5 Conclusion

Dans ce chapitre, nous avons introduit l'estimation du coefficient de dépendance de queue dans les statistiques de valeur extrême bivariable. Nous avons comparé trois estimateurs différents de ce coefficient. Nous nous concentrons sur la sélection des statistiques d'ordre supérieur utilisées dans l'estimation. Nous notons que la plupart des méthodes sont basées sur la minimisation de l'erreur moyenne quadratique asymptotique. Nous avons décrit deux méthodes de sélection adaptative pour l'estimateur de Hill et l'estimateur de Moment et avons les appliquées à un exemple pratique, les simulations indiquent qu'en général les résultats sont assez similaires.

## Conclusion générale

Dans cette thèse, on s'est intéressé à un problème récent en théorie des valeurs extrêmes univariées et multivariées c'est le choix de la valeur optimale de statistiques d'ordre extrêmes. Ce nombre est très important dans l'estimation de l'indice de queue dans le cas univarié et l'estimation du coefficient de dépendance de queue dans le cas bivarié. Plusieurs chercheurs ont soulevé ce problème par nombreuses méthodes listées dans ce travail.

Cette thèse se décompose en quatre chapitres, nous avons rappelé dans le premier chapitre quelques concepts fondamentaux sur la statistique des valeurs extrêmes univariées.

Dans le deuxième chapitre nous avons présenté quelques estimateurs de l'indice des valeurs extrêmes avec leurs propriétés asymptotiques et enfin nous avons exposés quelques méthodes adaptatives pour déterminer le nombre de statistiques d'ordre extrêmes.

Au troisième chapitre on parle sur la copule qu'est utilisée comme un moyen de séparer les distributions marginales de la dépendance et la dépendance de queue pour quantifier la probabilité d'événements conjoints dans la queue d'une distribution. Comme on l'a vu, les copules peuvent être utilisées pour modéliser la dépendance de queue. Un concept théorique important utilisé tout au long de cette thèse est la théorie des valeurs extrêmes bivariées, qui tente précisément d'analyser le comportement de la queue d'une distribution bivariée.

Au quatrième chapitre, nous avons introduit l'estimation du coefficient de dépendance de queue qu'est importante pour la théorie des valeurs extrêmes bivariées. La fraction

des statistiques d'ordre supérieur est impliquée dans l'estimation est importante dans la pratique, nous basons dans notre thèse sur l'estimateur à biais réduit proposé par Beirlant et al (2011). Ce dernier basé sur un modèle flexible proposé par Ledford et Tawn (1996,1997).

Pour le coefficient de dépendance de queue et leur sensibilité sur le nombre  $k$  de statistiques d'ordre supérieurs utilisés dans l'estimation. Aucune règle stricte et rapide n'existe pour affronter ce problème. Grosso modo le choix meilleur du  $k$  provient de la compétition entre le biais et la variance. L'arbitrage entre les deux effets contraires se fait usuellement en calculant l'erreur moyenne quadratique de l'estimateur. Habituellement, le scientifique décide subjectivement sur le nombre  $k$  à utiliser, en regardant les graphiques appropriés. Nous avons utilisé dans cette thèse des procédures adaptatives pour déterminer ce nombre. Des applications sur des données aussi bien réelles que simulées permettront d'illustrer les résultats obtenus. Un autre développement possible de ce travail est l'optimisation des estimateurs par rapport au choix de  $k$  ou encore une comparaison plus approfondie avec d'autres estimateurs de la queue de distribution déjà existants dans la littérature.

# Bibliographie

- [1] Arnold, B.C., Balakrishnan, N. et Nagaraja, H.N. (1992). A First Course in Order Statistics. Wiley, New York.
- [2] Beirlant, J., Vynckier, P. et Teugels, J.L. (1996, a). Tail index estimation, Pareto quantile plots, and regression diagnostics. *J. Amer. Statist. Assoc.*, 91, 1659-1667.
- [3] Beirlant, J., Vynckier, P. et Teugels, J.L. (1996, b). Excess function and estimation of the extreme-value index. *Bernoulli* 2, 293-318.
- [4] Beirlant, J., Goegebeur, Y. et Matthys, G. (1999). Tail Index Estimation and an Exponential Regression Model. *Extremes* 2, 177-200.
- [5] Beirlant, J. et Vandewalle, B., (2002). Some comments on the estimation of a dependence index in bivariate extreme value in statistics. *Statist. Probab. Lett.* 60, 265–278.
- [6] Beirlant, J., Matthys, G. (2003). Estimating the extreme value index and high quantiles
- [7] with exponential regression models. *Statistica Sinica*, 13 :853–880.
- [8] Beirlant, J., Goegebeur, Y., Segers, J., Teugels, J., (2006). *Statistics of Extremes – Theory and Applications*. Wiley.
- [9] Beirlant, J., Dierckx, G. et Guillou, A. (2005). Estimation of the extreme-value index and generalized quantile plots. *Bernoulli* 11, 949–970.
- [10] Beirlant, J., Bouquiaux, C. et Werker, B.J.M. (2006). Semiparametric Lower Bounds for Tail Index Estimation. *Journal of Statistical Planning and Inference* 136, 705-729.

- 
- [11] Beirlant, J., Joossens, E., Segers, J., (2009). Second-order refined peaks-over-threshold modelling for heavy-tailed distribution. *Journal of Statistical Planning and Inference*. 139, 2800–2815.
- [12] Beirlant J., Dierckx G. and Guillou A., (2011). Reduced bias estimators in joint tail modelling. *Insurance : Mathematics and Economics*. 49, 18–26.
- [13] Betteka, S. (2010). Détermination du nombre de statistiques d'ordre extrêmes. Mémoire de Magister, Université de Biskra, Algérie.
- [14] Betteka, S., and Brahim, B. (2017). Optimal number of upper order statistics used in estimation for the coefficient of tail dependence. *Journal Afrika Statistika*, 12 (1), 1095-1108.
- [15] Billingsley, P. (1995) : Probability and Measure, 3rd edition. Wiley, New York.
- [16] Brahim, B. (2010). Statistics of Bivariate Extreme Values. Thèse de doctorat, Université de Biskra, Algérie.
- [17] Brahim, B., Meraghni, D., and Necir, A. (2015). Gaussian approximation to the extreme value index estimator of a heavy-tailed distribution under random censoring. *Math. Methods Statist.*, 24(4), 266-279.
- [18] Brahim, B., Meraghni, D., Necir, A., and Yahia, D. (2013). A bias-reduced estimator for the mean of a heavy-tailed distribution with an infinite second moment. *J. Statist. Plann. Inference*, 143(6), 1064-1081.
- [19] Csörgö, S., Deheuvels, P. et Mason, D. (1985). Kernel Estimates of the Tail Index of a Distribution. *Annals of Statistics* 13, 1050-1077.
- [20] Cai, Z., Chen, X., Fan, Y., Wang, X., 2008. Selection of Copulas with Applications in Finance. Unpublished manuscript, University of North Carolina at Charlotte.
- [21] Cheng, S. and Pan, J., 1998. Asymptotic expansions of estimators for the tail index with applications. *Scand. J. Statist.* 25, 717–728.

- 
- [22] Cheng, S. and Peng, L., (2001). Confidence intervals for the tail index. *Bernoulli*. 7, 751-760.
- [23] Cherubini, U., E. Luciano, and W. Vecchiato, (2004). Copula methods in finance. Wiley.
- [24] Coles, S., Heffernan, J. and Tawn, J. (1999). Dependence measures for extreme value analysis. *Extremes* 2, 339-366.
- [25] Coles, S., 2001. An Introduction to Statistical Modeling of Extreme Values. Springer Series in Statistics, Springer New York.
- [26] Csörgö, S. et Viharos, L. (1998). Estimating the Tail Index. *Asymptotic Methods in Probability and Statistics*. B. Szyszkowicz, ed., North Holland, Amsterdam, 833-881.
- [27] Danielsson, J. et de Vries, C.G. (1997). Tail Index and Quantile Estimation with Very High Frequency Data. *Journal of Empirical Finance* 4, 241-257.
- [28] Danielsson, J., de Haan, L., Peng, L. et de Vries, C.G., (2001). Using a Bootstrap Method to Choose the Sample Fraction in Tail Index Estimation. *Journal of Multivariate Analysis*. 76, 226-248.
- [29] Davison, A.C. et Smith R.L. (1990). Models for Exceedances over High Thresholds (with discussion). *Journal of the Royal Statistical Society, Series B* 52, 393-442.
- [30] Davis, R., and Resnick, S. (1984). Tail estimates motivated by extreme value theory. *Ann. Statist.*, 1467-1487.
- [31] Davis, R.A., Resnick, S.I (1989). Basic properties and prediction of max-arma processes. *Adv. in Appl. Probab.*, 21 :781–803.
- [32] Deheuvels, P., 1979. La fonction de dépendance empirique et ses propriétés. *Acad. Roy. Belg. Bull. Cl. Sci.* 65, 274-292.
- [33] Deheuvels, P., 1981b. A nonparametric test for independence. *Publ. Inst. Statist. Univ. Paris.*, 26, 29-50.

- 
- [34] Deheuvels, P., Häeusler, E., and Mason, D. M. (1988). Almost sure convergence of the Hill estimator. *Math. Proc. Cambridge Philos. Soc.*, 104(02), 371-381.
- [35] Deheuvels, P., Tiago de Oliveira J., 1989. On the Non-Parametric Estimation of the Bivariate Extreme-Value Distributions *Statistics & Probability Letters* 8, 315-323.
- [36] Delmas, J.F. et Jourdain , B. (2006). *Modèles aléatoires. Applications aux sciences de l'ingénieur et du vivant.* Springer.
- [37] Denuit, M., Lambert, P., 2005. Constraints on concordance measures in bivariate discrete data. *Journal of Multivariate Analysis* 93, 40-57.
- [38] de Wolf, P.P. (1999). *Estimating the Extreme Value Index.* Thèse de doctorat, Delft University of Technology.
- [39] Dekkers, A.L.M. et de Haan, L. (1989). On the Estimation of the Extreme-Value Index and Large Quantile Estimation. *Annals of Statistics* 17, 1795-1832.
- [40] Dekkers, A. L. M., Einmahl, J. H. J. & de Haan, L., 1989. A moment estimator for the index of an extreme-value distribution. *Ann. Statist.* 17, 1833–1855.
- [41] Dekkers, A.L.M. et de Haan, L., 1993. Optimal choice of sample fraction in extreme-value estimation. *J. Multivariate Anal.* 47, 173–195.
- [42] Draisma, G., de Haan, L., Peng, L. et Ferreira, T.T., 1999. A Bootstrap-Based Method to Achieve Optimality in Estimating the Extereme-Value Index. *Extremes.* 2, 367-404.
- [43] Draisma G., Drees H., Ferreira A. and de Haan L., 2004. Bivariate tail estimation : dependence in asymptotic independence, *Bernoulli.* 10, 251-280.
- [44] Drees, H. (1998a). On smooth statistical tail functionals. *Scand. J. Statist.* 25, 187-210.
- [45] Drees, H. (1998b). A general class of estimators of the extreme value index. *J. Statist. Plann. Inference* 66, 95-112.



- 
- [46] Drees, H., Ferreira, A. and de Haan, L. (2002). On maximum likelihood estimation of the extreme value index. To appear in *Ann. Appl. Probab.*
- [47] Drees, H. et Kaufmann, E., 1998. Selection of the Optimal Sample Fraction in Univariate Extreme Value Estimation. *Stochastic Processes and their Applications.* 75 , 149-195.
- [48] Dutanga, C., Goegebeurb, Y., Armelle Guillo, A., 2014. Robust and bias-corrected estimation of the coefficient of tail dependence. *Insurance : Mathematics and Economics.* 57, 46-57.
- [49] Einmahl, J. H., Fils-Villetard, A., and Guillo, A. (2008). Statistics of extremes under random censoring. *Bernoulli*, 14(1), 207-227.
- [50] Einmahl, J.H.J. and Koning, A.J. (1992). Limit theorems for a general weighted process under random censoring. *Canad. J. Statist.*, 20, 77-89.
- [51] Embrechts P., Klüppelberg C., Mikosh T. (1997). *Modelling Extremal Events for Insurance and Finance*, Springer, Berlin.
- [52] Embrechts, P., 2000. Extreme Value Theory : Potential and Limitations as an Integrated Risk Management Tool Derivatives Use, Trading & Regulation 6, 449-456 .
- [53] Embrechts, P., McNeil, A., and Straumann D., 2002. Correlation and Dependence in Risk Management : Properties and Pitfalls To appear in *Risk Management : Value at Risk and Beyond*. Ed. by M., Dempster and H., K., Moffatt, Cambridge University Press, 176-223.
- [54] Embrechts, P., Hoeing, A., Juri., A., 2003. Using Copulae to bound the Value at-Risk for functions of dependent risks *Finance & Stochastics* 7, 145-167
- [55] Fama, E. F.(1965). The behavior of stock-market prices. *J. Business*, 38(1), 34-105.
- [56] Feller, W. (1971). *An Introduction to Probability Theory and its Applications*, 2nd edition. Wiley, New York.

- 
- [57] Fermanian, J.D. and Scaillet, O., 2004. Some Statistical Pitfalls in Copula Modeling for Financial Applications. FAME Working Paper, 108.
- [58] Fisher, R.A. et Tippett, L.H.C. (1928). Limiting Forms of the Frequency Distribution of the Largest or Smallest Member of a Sample. *Proceedings of the Cambridge Philosophical Society* **24**, 180-190.
- [59] Fraga Alves, M.I. (1995). Estimation of the Tail Parameter in the Domain of Attraction of an Extremal Distribution. *Journal of Statistical Planning and inference* **45**, 143-173.
- [60] Fraga Alves, M.I., Gomes, M.I., de Haan, L., 2003. A new class of semi-parametric estimators of the second order parameter. *Portugaliae Mathematica*. **60**, 193–214.
- [61] Frank, M., 1979. On the simultaneous associativity of  $F(x, y)$  and  $x+y-F(x, y)$ . *Aequationes Mathematicae*, **19**, 194-226.
- [62] Frahm, G., Junker, M. and Schmidt, R. (2005). Estimating the tail-dependence coefficient : Properties and Pit-falls. *Insurance : Mathematics and Economics*. **37**, 80-100.
- [63] Frees, E.W., Valdez, E.A., (1998). Understanding relationships using copulas. *North American Actuarial Journal*. **2**, 1–15.
- [64] Fréchet, M., 1951. Sur les tableaux de corrélation dont les marges sont données. *Annales de l'Université de Lyon. Sciences Mathématiques et Astronomie*, **14**, 53-77.
- [65] Fréchet, M., 1957. Les tableaux de corrélation dont les marges et des bornes sont données. *Annales de l'Université de Lyon. Sciences Mathématiques et Astronomie*, **20**, 13-31.
- [66] Gardes, L., and Girard, S. (2013). Estimation de quantiles extrêmes pour les lois à queue de type Weibull : une synthèse bibliographique. *J. de la Société Française de Statistique*, **154(2)**, 98-118.

- 
- [67] Geluk, J., De Haan, L., Resnick, S., and Stărică, C. (1997). Second-order regular variation, convolution and the central limit theorem. *Stochastic Process. Appl.*, 69(2), 139-159.
- [68] Genest, C., MacKay, R., 1986b. Copules archimédiennes et familles de lois bidimensionnelles dont les marges sont données. *The Canadian Journal of Statistics*, 14, 145-159.
- [69] Genest, C., Rivest L.P., 1993. Statistical inference procedures for bivariate Archimedean copulas. *Journal of the American Statistical Association*, 88, 1034-1043.
- [70] Genest, C., Boies, J.C., 2003. Detecting Dependence with Kendall Plots, *American Statistical Association*.
- [71] Genest, C., Gho Guillou. A. et Willems. P. (2006). Application de la théorie des valeurs extrêmes en hydrologie. *Statistique Appliquée*, LIV (2), 5-31.
- [72] Goegebeur, Y., Guillou, A., 2013. Asymptotically unbiased estimation of the coefficient of tail dependence. *Scand. J. Stat.* 40, 174–189.
- [73] Gnedenko, B. (1943). Sur la distribution limite du terme maximum d'une série aléatoire. *Ann. Math.*, 423-453.
- [74] Gomes, M. I., Henriques-Rodrigues, L., and Miranda, M. C. (2011), Reduced-bias location-invariant extreme value index estimation : a simulation study, *Communications in Statistics-Simulation and Computation*, 40, 424-447.
- [75] Gomes, M. I., Henriques-Rodrigues, L., Vandewalle, B., and Viseu, C. (2008), A heuristic adaptive choice of the threshold for bias-corrected Hill estimators, *Journal of Statistical Computation and Simulation*, 78, 133-150.
- [76] Groeneboom, P., Lopuhaä, H.P. et de Wolf, P.P. (2003). Kernel Estimators for the Extreme Value Index. *Annals of Statistics* 31, 1956-1995.
- [77] de Haan, L. (1976). Sample extremes : an elementary introduction. *Stat. Neerl.*, 30(4), 161-172.

- 
- [78] de Haan, L., Resnick, S.I. (1980). A simple asymptotic estimate for the index of a stable distribution. *Journal of the Royal Stat. Soc. B*, 42 :83–87.
- [79] de Haan. L et Ferreira. A. (2006). *Extreme value theory : an introduction*. Springer.
- [80] Hall, P. (1982). On some simple estimates of an exponent of regular variation. *J. R. Stat. Soc.*, 37-42.
- [81] Hall, P. et Welsh, A.H. (1985). Adaptive Estimates of Parameters of Regular Variation. *Annals of Statistics* **13**, 331-341.
- [82] Hall, P. (1990). Using the Bootstrap to Estimate Mean Squared Error and Select Smoothing Parameter in Nonparametric Problems. *Journal of Multivariate Analysis* 32, 177-203.
- [83] Häeusler, E., and Teugels, J. L. (1985). On asymptotic normality of Hill's estimator for the exponent of regular variation. *Ann. Statist.*, 743-756.
- [84] Heffernan, J. E. (2000). A directory of coefficients of tail dependence. *Extremes* 3, 279–290.
- [85] Heffernan, J. E. & Tawn, J. A. (2004). A conditional approach for multivariate extreme values. *J. Roy. Statist. Soc. Ser. B Statist. Methodol.* 66, 497–546.
- [86] Hill, B.M., 1975. A simple general approach to inference about the tail of a distribution. *Ann.Statist.* **3** , 1163–1174.
- [87] Hosking, J. et Wallis, J. (1987). Parameter and Quantile Estimation for the Generalized Pareto Distribution. *Technometrics* 29, 339-349.
- [88] Jenkinson, A. F. (1955). The frequency distribution of the annual maximum (or minimum) values of meteorological elements. *Quarterly J. R. Methodol. Soc.*, 81(348), 158-171.
- [89] Joe, H. (1997). *Multivariate Models and Dependence Concepts*. Chapman & Hall, London.

- 
- [90] Joe, H. (2011). Tail dependence in vine copulae. In : Kurowicka, D., Joe, H. (Eds.), Dependence Modeling. World Sci. Publ, Hackensack, NJ, pp. 165–187.
- [91] Juri, A., Wüthrich, M.V. (2002). Copula convergence theorems for tail events. Insurance : Mathematics and Economics 24, 139–148.
- [92] Johansson, J. (2003). Estimating the mean of heavy-tailed distributions. Extremes, 6(2), 91-109.
- [93] Kotz, S., Nadarajah, S., 2004. Multivariate t distributions and their applications. Cambridge University Press.
- [94] Klugman, S.A. (1992). Bayesian Statistics in Actuarial Science. With Emphasis on Credibility. In : Huebner International Series on Risk, Insurance, and Economic Security, Kluwer Academic Publishers, Boston, MA.
- [95] Klugman, S.A., Parsa, R. (1999). Fitting bivariate loss distributions with copulas. Insurance Math. Econom. 24, 139–148.
- [96] Ledford, A.W., Tawn, J.A., (1996). Statistics for near independence in multivariate extreme values. *Biometrika*. 83, 169-187.
- [97] Ledford, A.W., Tawn, J.A., (1997). Modelling dependence within joint tail regions. *Journal of the Royal Statistical Society : Series B* 59, 475–499.
- [98] Mason, D.M., 1982. Laws of large numbers for sums of extreme values. *Ann. Probab.* 10, 756-764.
- [99] Marshall, A., 1996. Copulas, marginals and joint distributions. In Distributions with Fixed Marginals and Related Topics. Ed. by L. Rüschendorf, B. Schweizer, and M. Taylor, pp. 213-222. Institute of Mathematical Statistics, Hayward, CA.
- [100] Marshall, A., and Olkin I., 1967. A multivariate exponential distribution. *Journal of the American Statistical Association*, 62, 30-44.
- [101] McNeil, A., Frey, R. et Embrechts, P. (2005) Quantitative Risk management. Princeton Press.

- 
- [102] Meraghni, D. (2008). Modelling distribution tails. Thèse de doctorat, Université de Biskra, Algérie.
- [103] Matthys, G. and Beirlant, J. (2000), Adaptive threshold selection in tail index estimation, in *Extremes and Integrated Risk Management*, ed. Embrechts, P., London, UK : Risk Books, pp. 37-49.
- [104] Matthys, G. et Beirlant, J. (2003). Estimating the Extreme Value Index and High Quantiles with Exponential Regression Models. *Statistica Sinica* 13, 853-880.
- [105] Nagaraja, H. N. et David, H. A. (2003). *Order Statistics*. Third Edition. Wiley.
- [106] Necir, A. (2006). A Functional Law of the Iterated Logarithm for Kernel-type Estimators of the Tail Index. *Journal of Statistical Planning and Inference* 136, 780-802.
- [107] Nelsen R.B., (1991). Copulas and association. In : Dall’Aglia G, Kotz S, Salinetti G (eds) *Advances in Probability Distributions with Given Marginals*. Kluwer, Dordrecht, 51-74.
- [108] Nelsen, R.B., (1999). *An introduction to Copulas.*, Lecture Notes in Statistics, 139, Springer Verlag, New-York.
- [109] Nelsen, R.B., (2006). *An Introduction to Copulas*, second ed. Springer, New York.
- [110] Neves, C. et Fraga Alves, M.I., (2004). Reiss and Thomas’ Automatic Selection of the Number of Extremes. *Computational Statistics and Data Analysis*. 47, 689–704.
- [111] Neves, M. M., Gomes, M. I., Figueiredo, F., and Gomes, D. P. (2014), \Modeling extreme events : Sample fraction adaptive choice in parameter estimation, *Journal of Statistical Theory and Practic*.
- [112] Oakes, D., (1982). A model for association in bivariate survival data. *J. Roy. Statist. Soc. Ser.*, 414-422
- [113] Peng, L., (1999). Estimation of the coefficient of tail dependence in bivariate extremes. *Statistics and Probability Letters*. 43, 399–409.

- 
- [114] Peng, L. (2010). A practical way for estimating tail dependence functions. *Statist. Sinica* 20, 365–378.
- [115] Joe, H., 1993. Parametric families of bivariate distributions with given marginals. *Journal of Multivariate Analysis* 46, 262-282.
- [116] Pickands, J. (1975). Statistical inference using extreme order statistics. *Ann. Statist.* 3, 119–131.
- [117] Poon, S.H., Rockinger, M., Tawn, J., 2004. Extreme-value dependence in financial markets : diagnostics, models, and financial implications. *Rev. Financial Studies* 17, 581–610.
- [118] Reiss, R.D. (1989). *Approximate distributions of order statistics*. Springer, New York.
- [119] Reiss, R.D. et Thomas, M. (2007). *Statistical Analysis of Extreme Values with Applications to Insurance, Finance, Hydrology and Other Fields*. Birkhäuser, Basel.
- [120] Resnick, S., 1987. *Extreme values, Regular variation and Point processes*. Springer, New York.
- [121] Resnick, S., 2007. *Heavy-Tail phenomena. Probabilistic and statistical modeling*. Springer, New York.
- [122] Rootzen, H. et Tajvidi, N. (1996). Extreme Value Statistics and Wind Storm Losses : a Case Study, *Scandinavian Actuarial Journal*, 70-94.
- [123] Saporta, G. (1990). *Probabilités, Analyse des Données et Statistique*. Editions Technip, Paris.
- [124] Schmidt, R., Stadtmüller, U. (2006). Non parametric estimation of tail dependence. *The Scandinavian Journal of Statistics.* 33, 307-335.
- [125] Segers, J. (2001). Residual estimators. *J. Statist. Plann. Inference*, 98(1), 15-27.
- [126] Sklar, A., 1996. Random Variables, Distribution Functions, and Copulas - a Personal Look Backward and Forward published in *Distributions with Fixed Marginals and*

- Related Topics, edited by L. Rüschendorf, B. Schweizer, and M.D. Taylor, Institute of Mathematical Statistics, Hayward, CA, 1-14.
- [127] Sklar, A., 1959. Fonctions de répartition à  $n$  dimensions et leurs marges, *Publ. Inst. Statist. Univ. Paris* 8, 229-231.
- [128] Smith, R.L. (1987), Estimating tails of probability distributions, *Annals of Statistics* 15, No. 3, 1174–1207.
- [129] Sousa, B.C. (2002). A Contribution to the Estimation of the Tail Index of Heavy-Tailed Distributions. Thèse de doctorat, Université de Michigan, [www.utstat.toronto.edu/desousa](http://www.utstat.toronto.edu/desousa).
- [130] Tassi, Ph. et Legait, S. (1990). Théorie des probabilités en vue des applications statistiques. Edition Technip.
- [131] Varadhan, R., Gilbert, P.D. (2009). BB : An R Package for Solving a Large System of Nonlinear Equations and for Optimizing a High-Dimensional Nonlinear Objective Function, *J. Statistical Software*, 32 :4, <http://www.jstatsoft.org/v32/i04/>.
- [132] Weissman, I., 1978. Estimation of parameters and large quantiles based on the  $k$  largest observations. *Journal of the American Statistical Association*. 73, 812–815.
- [133] Yun, S. (2002). On a Generalized Pickands Estimator of the Extreme Value Index. *Journal of Statistical Planning and Inference* 102, 389-409.



# Résumé

Le concept de dépendance de queue fournit une description de la dépendance au niveau des queues de distribution, très intéressante pour étudier la survenance simultanée de valeurs extrêmes. La motivation principale de notre thèse de doctorat est améliorer la performance de l'estimateur du coefficient de dépendance de la queue proposé par Beirlant et al (2011) par un choix optimal de statistique de d'ordre extrêmes utilisé dans l'estimation. Dans cette thèse, nous exposons les différentes méthodes de détermination de ce nombre. Grâce à la simulation, nous avons appliqué deux méthodes, à savoir Reiss & Tomas et Cheng & peng, sur l'ensemble des données réelles qui illustreraient les résultats obtenus.

**Mots-clés :** Valeur extrême, Copule, Estimateur de Hill, Point terminal, Variation régulière, Domaine d'attraction, paramètre du second ordre, Erreur moyenne quadratique, Queues lourdes, Indice de queue, Fraction de l'échantillon, Coefficient de dépendance de la queue, Bias-réduit, Distribution de Pareto étendue.

# Abstract

The concept of tail dependence provides a description of the dependence at the tail distribution level, which is very interesting for studying the simultaneous occurrence of extreme values. The main motivation of our PhD thesis is to improve the performance of the tail dependency coefficient estimator proposed by Beirlant et al (2011) by an optimal choice of extreme order statistics used in the estimation. In this thesis, we present the different methods for determining this number. Through simulation, we applied two methods, namely Reiss & Tomas, and Cheng & peng , on the real data set that would illustrate the results obtained.

**Key words** : Extreme value, Copula, Hill Estimator, Upper endpoint, Regular Variation, Domain of attraction, Second order parameter, Bootstrap, Mean Square Error, Heavy tails, Tail index, Fraction sample, Coefficient of tail dependence, Bias reduction, Extended Pareto distribution.

## ملخص

يوفر مفهوم اعتماد الذيل وصفا للاعتماد على مستوى التبعية من توزيع الذيل و هذا امر مثير للاهتمام لدراسة تزامن حدوث القيم المتطرفة. الهدف الرئيسي من اطروحة الدكتوراه هو تحسين اداء مقدر معامل تبعية الذيل المقترح من طرف بيرلاننت و اخرون (2011) عن طريق تحديد العدد الأمثل للاحصاءات الترتيب الذي هو عنصر مهم جدا في تقدير قيمة معامل تبعية الذيل. في هذا الصدد قمنا في هذه الاطروحة بسرد وتطبيق عدة أساليب مختلفة لتحديد هذا العدد المهم. من خلال المحاكاة قمنا بتطبيق طريقتين مختلفتين هما طريقة ريس و طوماس و طريقة شانغ و بينغ على مجموعة البيانات الحقيقية التي من شأنها أن توضح النتائج التي تم الحصول عليها.

**الكلمات المفتاحية:** القيمة القصوى، دالة الارتباط، مقدر هيل، نقطة النهاية العليا، الاختلاف المضبوط، منطقة جذب، إحصائيات الترتيب، معامل الدرجة الثانية، بوتسراب، الخطأ التربيعي المتوسط، الذيل الثقيلة، مؤشر الذيل، جزء من العينة، توزيع باريتو الموسع.