

Université de Montréal

**Correction d'estimateurs de la fonction de Pickands et
estimateur bayésien**

par

Kevin Chalifoux

Département de mathématiques et de statistique
Faculté des arts et des sciences

Mémoire présenté en vue de l'obtention du grade de
Maître ès sciences (M.Sc.)
en statistique

26 janvier 2023

Université de Montréal

Faculté des arts et des sciences

Ce mémoire intitulé

**Correction d'estimateurs de la fonction
de Pickands et estimateur bayésien**

présenté par

Kevin Chalifoux

a été évalué par un jury composé des personnes suivantes :

Pr. Martin Bilodeau

(président-rapporteur)

Pr. François Perron

(directeur de recherche)

Pr. Maciej Augustyniak

(membre du jury)

Résumé

Faire l'estimation d'une copule de valeurs extrêmes bivariée revient à estimer A, sa fonction de Pickands qui lui est associée. Cette fonction $A : [0,1] \rightarrow [0,1]$ doit satisfaire certaines contraintes :

$$\max\{1-t, t\} \leq A(t) \leq 1, \quad t \in [0,1]$$

A est convexe.

Plusieurs estimateurs ont été proposés pour estimer cette fonction A, mais peu respectent ses contraintes imposées. La contribution principale de ce mémoire est d'introduire une technique simple de correction d'estimateurs de la fonction de Pickands de sorte à ce que les estimateurs corrigés respectent les contraintes exigées. La correction proposée utilise une nouvelle propriété du vecteur aléatoire bivarié à valeurs extrêmes, combinée avec l'enveloppe convexe de l'estimateur obtenu pour garantir le respect des contraintes de la fonction A.

La seconde contribution de ce mémoire est de présenter un estimateur bayésien non paramétrique de la fonction de Pickands basé sur la forme introduite par Capéraà *et al.* (1997). L'estimateur utilise les processus de Dirichlet pour estimer la fonction de répartition d'une transformation du vecteur aléatoire bivarié à valeurs extrêmes.

Des analyses par simulations sont produites sur un ensemble d'estimateurs pour mesurer la performance de la correction et de l'estimateur bayésien proposés, sur un ensemble de 18 distributions de valeurs extrêmes bivariées. La correction améliore l'erreur quadratique moyenne sur l'ensemble des niveaux. L'estimateur bayésien proposé obtient l'erreur quadratique moyenne minimale pour les estimateurs considérés.

Mots clés : Copules, valeurs extrêmes, fonctions de Pickands, estimation avec contraintes, estimation bayésienne.

Abstract

Estimating a bivariate extreme-value copula is equivalent to estimating A, its associated Pickands function. This function $A: [0,1] \rightarrow [0,1]$ must satisfy some constraints :

$$\max\{1-t, t\} \leq A(t) \leq 1, \quad t \in [0,1]$$

A is convex.

Many estimators have been proposed to estimate A, but few satisfy the imposed constraints. The main contribution of this thesis is the introduction of a simple correction technique for Pickands function estimators so that the corrected estimators respect the required constraints. The proposed correction uses a new property of the extreme-value random vector and the convex hull of the obtained estimator to guaranty the respect of the Pickands function constraints.

The second contribution of this thesis is to present a nonparametric bayesian estimator of the Pickands function based on the form introduced by Capéraà, Fougères and Genest (1997). The estimator uses Dirichlet processes to estimate the cumulative distribution function of a transformation of the extreme-value bivariate vector.

Analysis by simulations and a comparison with popular estimators provide a measure of performance for the proposed correction and bayesian estimator. The analysis is done on 18 bivariate extreme-value distributions. The correction reduces the mean square error on all distributions. The bayesian estimator has the lowest mean square error of all the considered estimators.

Keywords: Copula, Extreme Value, Pickands Function, Constraint Estimation, Bayesian Estimation.

Table des matières

Résumé	5
Abstract	7
Liste des tableaux	11
Table des figures	13
Liste des sigles et des abréviations	15
Remerciements	17
Introduction	19
Chapitre 1. Théorie des valeurs extrêmes.....	23
Théorème des valeurs extrêmes	23
Chapitre 2. Copules de valeurs extrêmes	27
Copules	27
Théorème de Sklar	28
Bornes de Fréchet-Hoeffding	31
Copules de valeurs extrêmes	31
Chapitre 3. Estimateurs et corrections non paramétriques	37
Estimateurs	38
Estimateur de Pickands (1981)	38
Estimateur de Deheuvels (1991)	39
Estimateur de Capéraà, Fougères et Genest (1997)	40
Estimateur de Hall-Tajvidi (2000)	43
Correction d'estimateurs	45
Plus grand minorant convexe	45
Spline définie dans \mathcal{A}	46

Conclusion	48
Chapitre 4. Correction convexe et estimateur bayésien	49
Correction par centralisation et enveloppe convexe	49
Propriété de la variable aléatoire Z	49
Enveloppe convexe	50
Correction par centralisation et enveloppe convexe	51
Estimateur bayésien	54
Contexte bayésien	54
Processus de Dirichlet	56
Estimateur bayésien	57
Simulations	60
Conclusion	65
Bibliographie.....	67
Annexe 1 : Preuve univariée d'équivalence du max-stable	69
Annexe 2 : Calculs de Capéraà, Fougères et Genest (1997)	71
Annexe 3 : Algorithme de l'estimateur bayésien	75

Liste des tableaux

4.1	Description des distributions à partir desquelles les échantillons de la simulation ont étées générées.	62
4.2	Description des erreurs quadratiques moyennes des estimateurs et des corrections appliqués sur les échantillons générés, selon la taille de l'échantillon.	64

Table des figures

1.1	Densités de valeurs extrêmes.....	26
2.1	Données selon différentes forces de dépendance.....	29
2.2	Données avec faible dépendance (Copule mixte ($\theta = 0.1, \phi = 0$)).....	34
2.3	Données avec forte dépendance (Copule logistique ($\theta = 0.95, \phi = 0.75, r = 10$))....	35
3.1	Jeu de données de 30 observations.....	37
3.2	Estimation par l'estimateur de Pickands.....	39
3.3	Estimation par l'estimateur de Deheuvels.....	40
3.4	Estimation par l'estimateur de CFG.....	44
3.5	Estimation par l'estimateur de Hall-Tajvidi.....	45
3.6	Exemple de correction par le PGMC.....	46
3.7	Correction d'estimateurs par FGS.....	48
4.1	Estimateur de CFG corrigé.....	53
4.2	Fonctions de Pickands a priori.....	58
4.3	Estimation par l'estimateur bayésien.....	61
4.4	Copules logistiques utilisées.....	62
4.5	Copules mixtes utilisées.....	63

Liste des sigles et des abréviations

Estimateur de CFG	Estimateur de Capéraà, Fougères et Genest (1997).
Correction FGS	Correction de Fils-Villetard, Guillou et Segers (2008).
i.i.d.	[Variables aléatoires] indépendantes et identiquement distribuées.
TVE	Théorie des valeurs extrêmes.
MCMC	Monte-Carlo par chaînes de Markov.
PGMC	Plus grand minorant convexe.
s.à.	Sujet à (Contraintes dans le contexte d'optimisation).

Remerciements

Mes remerciements les plus sincères sont adressés à mon directeur de recherche, Pr. François Perron. C'est grâce à son aide, à ses conseils et à son soutien financier que ce mémoire peut être publié. Je lui suis aussi reconnaissant pour les échanges de fin de rencontre que nous avons eus et les conseils qu'il m'a partagés pour mes objectifs suivant la publication de ce mémoire.

Je tiens également à remercier le personnel du département de mathématiques et de statistique de l'Université de Montréal pour leur disponibilité, pour la qualité de leur enseignement et pour les ressources nécessaires à la réussite qu'ils ont fournies.

Finalement, je tiens à donner une mention spéciale aux personnes près de moi qui m'ont supporté au cours des dernières années. Plus précisément, à mon père, Daniel, à ma mère, Josée, à ma soeur, Jessica, et à ma très chère Marilou.

Introduction

Dans la théorie de l'estimation, on peut s'intéresser à l'observation d'un évènement à valeurs extrêmes. Des exemples de telles études peuvent être : les pertes maximales d'un assureur, le risque d'une action financière, le nombre maximal de tremblements de terre au cours de l'année. Pour ce faire, on étudie la fonction de survie $\mathbb{P}(X > x)$ du phénomène aléatoire à l'aide de la théorie des valeurs extrêmes.

D'autres parts, il est possible d'étudier et de modéliser la force du lien de dépendance qui régit deux ou plusieurs phénomènes aléatoires simultanément. Pour ce faire, une approche est d'estimer la copule des variables aléatoires liées aux évènements. Le terme copule, apporté par Sklar en 1959, peut être interprété comme la fonction de dépendance qui lie nos variables aléatoires ensemble. Cette approche est intéressante puisque selon Ding et Li (2015), la copule permet de capturer la totalité de la dépendance entre les variables, contrairement au coefficient de détermination, par exemple.

Sklar (1959) a montré qu'une fonction de répartition conjointe F , d'un vecteur aléatoire $(X,Y) \in \mathcal{X} \times \mathcal{Y}$, peut toujours se décomposer dans la forme :

$$F(x,y) = C(F_X(x), F_Y(y)), \quad (x,y) \in \mathcal{X} \times \mathcal{Y}$$

où F_X et F_Y sont les fonctions de répartition marginales respectives de X et Y , C est la fonction de dépendance - ou copule.

Le contexte dans lequel baigne les contributions de ce mémoire est celui où l'on s'intéresse à l'observation de deux événements à valeurs extrêmes simultanément. Autrement dit, on s'intéresse à la distribution du vecteur aléatoire lié aux deux événements extrêmes, en tenant compte du lien de dépendance qui les régit. Ce contexte nous permet de réunir la théorie des valeurs extrêmes en prenant les marges associées à nos événements et de les joindre par une copule de valeurs extrêmes. Un exemple d'application, présenté dans Tawn (1988), est l'étude du niveau de la mer maximal de deux ports maritimes britanniques pour évaluer les risques d'inondation.

Une des particularités des copules de valeurs extrêmes a été montrée par Pickands (1981), où ce dernier a formulé le théorème portant son nom qui dit que pour la copule de valeurs extrêmes C , il existe une fonction A telle que

$$C(u,v) = \log(uv)A\left(\frac{\log(u)}{\log(uv)}\right), \quad u,v \in [0,1]$$

Avec

$$A(t) \geq \max\{1-t, t\}, \quad t \in [0,1]$$

$$A(0) = A(1) = 1,$$

A est convexe sur $[0,1]$.

La copule qui lie nos variables aléatoires à valeurs extrêmes peut alors s'exprimer à partir de la fonction A qui satisfait aux contraintes décrites précédemment. Depuis la publication de ce théorème en 1981, cette fonction A - de Pickands - est le sujet d'estimation dans le contexte d'estimation de phénomènes à valeurs extrêmes bivariés.

La difficulté de l'estimation de la fonction de Pickands repose naturellement sur les contraintes. Plusieurs estimateurs non paramétriques performent bien asymptotiquement, mais ne respectent pas les contraintes imposées des fonctions de Pickands sur des échantillons de tailles finies. Plus particulièrement, c'est le cas pour les estimateurs non paramétriques les plus populaires qui suivent. Le premier fût proposé par Pickands lui-même et porte également son nom. Ses propriétés asymptotiques, quant à elles, furent montrées par Deheuvels (1984). Deheuvels proposa aussi, en 1991, une variante de l'estimateur pour corriger le biais lorsque évalué en 0 et en 1. Quelques années plus tard, Capéraà *et al.* (1997) proposent un estimateur performant mieux sur les échantillons de petites tailles. Hall et Tajvidi (2000) proposent à leur tour un estimateur basé sur les estimateurs précédents, qui a l'avantage de satisfaire la contrainte de la borne inférieure exigée.

Suivant le non-respect des contraintes imposées, des corrections d'estimateurs ont été proposées. Selon Fils-Villetard *et al.* (2008), certaines vont prendre le plus grand minorant convexe de l'estimateur initial qui respecte les contraintes de A , tandis que d'autres vont approcher l'estimateur initial par une spline contrainte dans \mathcal{A} – l'espace des fonctions de Pickands.

Notre projet consiste à étudier quelques-unes de ces corrections et à en suggérer une nouvelle plus simple.

- Dans le premier chapitre, nous introduisons la théorie des valeurs extrêmes (TVE) qui est essentielle pour obtenir la forme des distributions de variables aléatoires à valeurs extrêmes univariées.
- Dans le second, nous introduisons la notion de copule qui complète la construction de la loi conjointe. Pour le contexte de copules de valeurs extrêmes qui nous intéresse, il est question d'avoir une copule qui lie deux distributions univariées de valeurs extrêmes, d'où l'importance du chapitre 1. Nous voyons que cette copule s'exprime à partir d'une fonction de Pickands, qui est le sujet de notre estimation. Nous y présentons ses caractéristiques.
- Dans le troisième chapitre, nous introduisons brièvement quelques estimateurs non paramétriques et type de corrections les plus populaires, en plus d'en justifier quelques-uns.
- Dans le quatrième chapitre, nous présentons une nouvelle correction d'estimateurs et un nouvel estimateur bayésien, accompagnés de simulations pour mesurer leur performance comparativement aux estimateurs traditionnels.

Chapitre 1

Théorie des valeurs extrêmes

Ce premier chapitre présente la TVE univariée et sa structure suit celle présentée dans le livre de de Haan et Ferreira (2006). Il est construit en vue de donner au lecteur les notions de base de la TVE univariée, en mettant l'emphasis sur les concepts analogues à ceux qui se trouvent dans la théorie des copules de valeurs extrêmes bivariées – la généralisation bivariée de la TVE.

Le théorème des valeurs extrêmes – qui est le résultat principal de la TVE univariée – peut être présenté en parallèle au théorème central limite, comme le font Kreinovich *et al.* (2017). Malgré le contraste qu'un s'intéresse à la moyenne partielle d'un ensemble de variables aléatoires tandis que l'autre s'intéresse au maximum partiel, les deux statistiques convergent en probabilité et peuvent avoir une distribution limite, sous certaines conditions et après normalisation. La différence majeure entre les théorèmes est lorsqu'une loi asymptotique existe pour le maximum normalisé, elle prend la forme de l'une des lois de probabilité suivantes : Weibull, Gumbell ou Fréchet.

Théorème des valeurs extrêmes

Soit X_1, X_2, \dots, X_n , des variables aléatoires i.i.d., on définit

$$M_n = \max(X_1, X_2, \dots, X_n).$$

La distribution de M_n est le point focal du chapitre.

Dans l'exemple de de Haan et Ferreira (2006) du niveau de la mer lors de tempêtes à Delfzijl, aux Pays-Bas, on s'intéresse à ériger une digue de manière à ce qu'il y ait une probabilité d'inondation inférieure à 10^{-4} au cours d'une année. Des données récoltées lors de 1877 tempêtes au cours des 100 dernières années forment un jeu de données d'observations approximativement indépendantes. On cherche à calculer la hauteur m que doit avoir la

digue pour que

$$\mathbb{P}(M_n > m) \leq 10^{-4},$$

où M_n est la variable aléatoire de la hauteur maximale du niveau de la mer atteinte lors d'une année.

Asymptotiquement, le problème d'identification de la distribution de M_n n'est pas très intéressant. En effet, on remarque que pour X_1, X_2, \dots, X_n , des variables aléatoires indépendantes et de fonction de répartition F , on a

$$M_n \xrightarrow[n \rightarrow \infty]{P} x^*,$$

où x^* est le supréum essentiel. En effet,

$$\mathbb{P}(M_n \leq x) = F^n(x)$$

qui converge à 1 si $x \geq x^*$ et à 0 si $x < x^*$, lorsque n tend vers l'infini.

On s'intéresse alors à la distribution asymptotique de M_n .

Définition 1.0.1. On dit que la fonction de répartition F appartient au domaine d'attraction de G , noté $F \in D(G)$, s'il existe deux suites $(a_n)_{n \in \mathbb{N}} > 0$ et $(b_n)_{n \in \mathbb{N}}$ telles que

$$\lim_{n \rightarrow \infty} F^n(xa_n + b_n) = G(x),$$

où G est une fonction de répartition non dégénérée. De plus, si G existe, sa loi est appelée une loi de valeurs extrêmes.

Nous cherchons alors, si elles existent, des suites de valeurs $(a_n)_{n \in \mathbb{N}} > 0$ et $(b_n)_{n \in \mathbb{N}}$, telles que

$$\frac{M_n - b_n}{a_n} \xrightarrow[n \rightarrow \infty]{\mathcal{L}} G,$$

avec G une fonction de répartition non dégénérée.

Autrement dit, on cherche $(a_n)_{n \in \mathbb{N}} > 0$ et $(b_n)_{n \in \mathbb{N}}$, telles que $F \in D(G)$.

Définition 1.0.2. Soit la fonction de répartition G de X_1, X_2, \dots, X_n i.i.d.. On définit G comme étant max-stable s'il existe des suites $(a_n)_{n \in \mathbb{N}} > 0$ et $(b_n)_{n \in \mathbb{N}}$ telles que

$$G^n(a_n x + b_n) = G(x),$$

pour tout $n > 0$. Une définition alternative et équivalente est de dire que

$$\mathbb{P}\left(\frac{M_n - b_n}{a_n} \leq x\right) = \mathbb{P}(X_1 \leq x),$$

pour tout $n > 0$.

Par définition, la loi de G qui est max-stable est une loi de valeurs extrêmes. Le théorème suivant énonce la réciproque de cette remarque.

Théorème 1.0.3. Soit F une fonction de répartition telle que $F \in D(G)$, alors G est max-stable si et seulement si la loi de G est une loi de valeurs extrêmes.

Démonstration. Voir Annexe 1. □

Les dernières définitions sont donc équivalentes. Elles mènent au théorème suivant qui est le résultat principal de cette section. Il agit comme analogue au théorème principal du cas bivarié.

Théorème 1.0.4 (Fisher et Tippet (1928), von Mises (1936), Gnedenko (1943)). La fonction de répartition d'une loi de valeurs extrêmes est donnée par

$$G_\gamma\left(\frac{x - \mu}{\sigma}\right),$$

avec μ la moyenne et σ l'écart-type de la loi de valeurs extrêmes, et

$$G_\gamma(x) = \begin{cases} \exp\left(- (1 + \gamma x)^{-1/\gamma}\right) & \text{si } \gamma \neq 0, \\ \exp(-e^{-x}) & \text{si } \gamma = 0. \end{cases} \quad 1 + \gamma x > 0$$

L'indice γ est appelé l'indice de valeurs extrêmes.

Démonstration. Le lecteur s'intéressant à la démonstration du théorème est référé à de Haan et Ferreira (2006). □

L'intérêt de ce théorème est de dire que la classe de toutes les distributions de valeurs extrêmes se caractérise en fonction d'un seul paramètre, en excluant les paramètres de centralisation et de normalisation. Le problème d'identification de la distribution asymptotique de M_n devient alors un problème d'estimation de γ . Pour faire l'analogie, le résultat bivarié indique qu'une copule de valeurs extrêmes s'écrit à l'aide d'une fonction, que nous chercherons à estimer.

Le paramètre γ de la distribution de valeurs extrêmes permet de contrôler les queues de la distribution. En effet,

- si $\gamma < 0$, la fonction de répartition G_γ appartient à la loi de Weibull. Cette distribution est à queue courte.

Densités de valeurs extrêmes

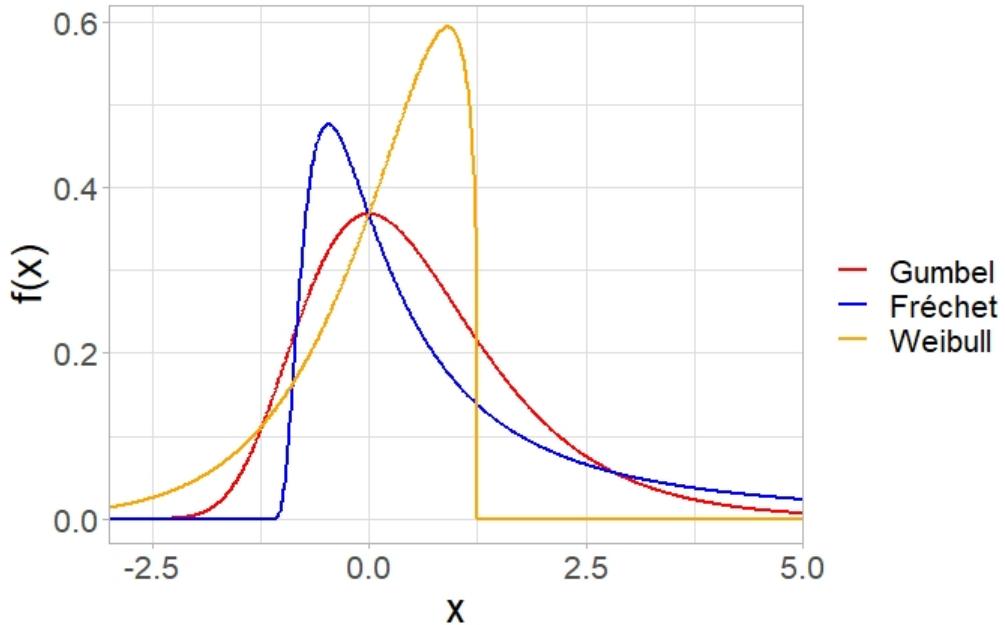


Figure 1.1. Densités de valeurs extrêmes.

- si $\gamma = 0$, la fonction de répartition G_γ appartient à la loi de Gumbel. Cette distribution est à queue légère et tous ses moments existent.
- si $\gamma > 0$, la fonction de répartition G_γ appartient à la loi de Fréchet. Cette distribution est à queue droite épaisse.

Un exemple visuel pour chacune de ces distributions est illustré dans la figure 1.1. Les densités des distributions sont paramétrisées par $\mu = 0$ et $\sigma = 1$. Les valeurs de γ pour les lois de Gumbel, de Fréchet et de Weibull sont respectivement de 0, de 0,8 et de -0,8.

Chapitre 2

Copules de valeurs extrêmes

"La pertinence des copules vient de leur rôle dans la modélisation de la structure de dépendance d'une distribution multivariée, sans tenir compte des marges" - Fils-Villetard et al. (2008)

Copules

La structure de cette section est basée sur le chapitre 2 du livre de Nelsen (2006) et le lecteur est encouragé à consulter l'ouvrage pour une introduction plus approfondie à la théorie des copules.

Le but de ce chapitre est de donner au lecteur un outil lui permettant d'étudier la structure de dépendance qui régit deux variables aléatoires. Un exemple pourrait être d'analyser la dépendance entre les notes d'un étudiant dans son cours de calcul intégral et ses notes dans son cours d'algèbre linéaire. Le résultat de l'analyse pourrait nous informer à savoir si l'étudiant sacrifie du temps d'étude d'un cours au détriment de l'autre, ou au contraire, s'il essaie d'uniformément répartir son temps d'étude entre les cours. Cet outil qui lie – ou accouple – deux variables aléatoires par un lien de dépendance est la copule.

Au cours du chapitre, il est question de représenter une fonction de répartition conjointe de deux événements aléatoires comme une fonction de ses marges. Cette fonction des marges est une copule bivariée associée aux deux événements aléatoires. La copule doit donc satisfaire aux conditions des fonctions de répartitions bivariées.

Ainsi, il pourrait intéresser le lecteur de rappeler certaines propriétés élémentaires des fonctions de répartition bivariées avant de présenter formellement les copules bivariées et leur usage en théorie des probabilités.

Pour le restant du chapitre, définissons F la fonction de répartition du vecteur aléatoire (X, Y) à valeurs dans $\mathcal{X} \times \mathcal{Y}$, F_X et F_Y les fonctions de répartition marginales respectives.

On a

- $F : \mathcal{X} \times \mathcal{Y} \rightarrow [0,1]$
- X et Y sont indépendantes $\Leftrightarrow F(x, y) = F_X(x)F_Y(y)$, $(x, y) \in \mathcal{X} \times \mathcal{Y}$
- $F_X(x) = \lim_{b \rightarrow \infty} F(x, b)$, $x \in \mathcal{X}$.

Définition 2.0.1. On définit la copule, C , comme une fonction

$$C : [0,1]^2 \rightarrow [0,1],$$

telle que :

(1) Pour $u \in [0,1]$,

$$C(u, 0) = C(0, u) = 0,$$

et

$$C(u, 1) = C(1, u) = u.$$

(2) $\forall u_1, u_2, v_1, v_2 \in [0,1]$ avec $u_1 \leq u_2$, $v_1 \leq v_2$,

$$C(u_2, v_2) - C(u_2, v_1) - C(u_1, v_2) + C(u_1, v_1) \geq 0.$$

Note : La deuxième propriété énumérée définit une fonction 2-croissante.

Théorème de Sklar

Le théorème suivant permet le lien entre la définition des copules présentée et l'usage probabiliste qu'il en est fait. Comme le précise Ding et Li (2015), le théorème de Sklar décompose la fonction de répartition conjointe en deux composantes : la copule et les marges uniformes des variables.

Théorème 2.0.2 (Sklar (1959)). Soit F une fonction de répartition conjointe avec F_X , F_Y comme fonctions de répartition marginales.

Il existe une copule C telle que

$$F(x, y) = C(F_X(x), F_Y(y)), (x, y) \in \mathcal{X} \times \mathcal{Y}$$

De plus, si F_X et F_Y sont continues, alors C est unique ; autrement l'unicité de C se réalise sur l'image de F_X et F_Y . Le résultat mentionne également que si F_X et F_Y sont des fonctions de répartition marginales, alors F telle que définie est une fonction de répartition conjointe de marges F_X et F_Y .

La force d'une copule est donc de prendre deux fonctions de répartition marginales et d'en créer une conjointe avec un lien de dépendance entre les variables. Le but de ce mémoire, on le rappelle, est d'estimer des copules de valeurs extrêmes.

Avant de voir les implications de ce théorème, il pourrait intéresser le lecteur de se familiariser davantage avec ce dernier à l'aide d'un exemple. Ci-dessous est présenté des données simulées selon différentes copules. La pertinence de cet exemple visuel se trouve dans le fait que les distributions marginales de X et de Y sont identiques entre elles, mais aussi au travers des graphiques. L'exemple illustre pleinement la force des copules à altérer la dépendance entre les variables aléatoires.

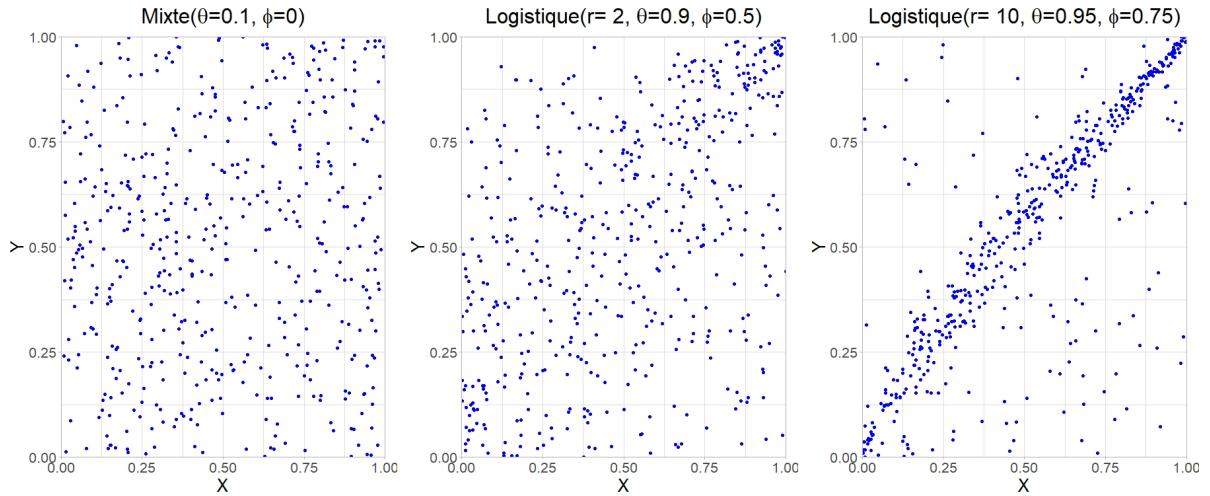


Figure 2.1. Données selon différentes forces de dépendance.

L'exemple ci-haut représente des données à valeurs extrêmes simulées de distributions marginales exponentielles. Le graphique est exprimé en fonction du vecteur de quantiles $(F_X(x_i), F_Y(y_i))$ des observations (x_i, y_i) et les copules utilisées sont respectivement de type et de paramètres : mixte ($\theta = 0.1, \phi = 0$), logistique ($\theta = 0.9, \phi = 0.5, r = 2$), logistique ($\theta = 0.95, \phi = 0.75, r = 10$). Ces copules de valeurs extrêmes sont explorées plus en détails dans la seconde section du chapitre, mais l'exemple illustre parfaitement la pertinence des copules.

Avant de présenter les copules de valeurs extrêmes, quelques propriétés, à être présentées, seront utiles dans la seconde section de ce chapitre.

D'abord, un corollaire direct du théorème de Sklar permettant de construire des copules. Il suffit de définir l'inverse généralisé, $F^{(-1)} : [0,1] \rightarrow [-\infty, \infty]$, d'une fonction de répartition F comme étant

$$F^{(-1)}(u) = \inf\{x : F(x) \geq u\}, \quad u \in [0,1].$$

On obtient alors :

Corollaire 2.0.3. *Soit F , F_X , F_Y , C comme dans le théorème de Sklar, $F_X^{(-1)}$, $F_Y^{(-1)}$ les inverses généralisés de F_X et de F_Y , alors*

$$C(u,v) = F\left(F_X^{(-1)}(u), F_Y^{(-1)}(v)\right), \quad (u,v) \in [0,1]^2$$

Si F_X et F_Y sont continues, avec la connaissance des fonctions de répartition marginales et conjointes, il est possible de construire des copules bivariées.

Sans avoir d'impact direct sur le restant du chapitre, ce corollaire peut servir d'outil de compréhension et d'analyse intéressant au lecteur ; il peut servir pour explorer et analyser théoriquement différents types et niveaux de dépendance. Toutefois, pour revenir à la définition 2.0.1 des copules, dans le cas où F_X et F_Y sont continues, la condition 1 reflète le principe qu'une fonction de répartition suit une distribution uniforme sur l'intervalle $[0,1]$;

$$\begin{aligned} C(u, 1) &= F(F_X^{(-1)}(u), F_Y^{(-1)}(1)) \\ &= F_{X|Y}(F_X^{(-1)}(u)|F_Y^{(-1)}(1))F_Y(F_Y^{(-1)}(1)) \\ &= F_X(F_X^{(-1)}(u)) = u. \end{aligned}$$

De même,

$$\begin{aligned} C(u, 0) &= F(F_X^{(-1)}(u), F_Y^{(-1)}(0)) \\ &= F_{Y|X}(F_Y^{(-1)}(0)|F_X^{(-1)}(u))F_X(F_X^{(-1)}(u)) = 0. \end{aligned}$$

La condition 2 découle du fait que la fonction de répartition doit être une fonction 2-croissante, pour ne pas permettre de probabilités négatives.

Un cas particulier des copules est celui où les variables aléatoires étudiées sont indépendantes. On rappellera que dans cette situation,

$$F(x,y) = F_Y(x)F_Y(y), \quad (x,y) \in (\mathcal{X}, \mathcal{Y}).$$

Dans le contexte des copules, ceci implique

$$C(u,v) = C(F_X(x), F_Y(y)) = F(x,y) = F_X(x)F_Y(y) = uv.$$

Cette remarque sera utilisée dans le contexte des copules de valeurs extrêmes.

Un aspect fort intéressant des copules est celui de leur invariance aux transformations strictement croissantes sur le domaine des variables X et Y . Autrement dit, pour des transformations qui n'affectent pas l'ordonnancement des valeurs de X et de Y , la copule des variables transformées est identique à celle originale.

Bornes de Fréchet-Hoeffding

Les dernières propriétés pertinentes pour notre problème sont les bornes de la copule. En 1951 pour Fréchet et simultanément pour Hoeffding dans des journaux allemands, les auteurs ont distinctement publié le résultat suivant.

Théorème 2.0.4 (Bornes de Fréchet-Hoeffding). Soit la copule C . On a

$$\max(u + v - 1, 0) \leq C(u,v) \leq \min(u,v), \quad u,v \in [0,1]$$

En plus de borner explicitement la fonction copule, les bornes de Fréchet-Hoeffding sont d'intérêt pour leur analogue à valeurs extrêmes qui seront présentés au cours des pages suivantes. Elles serviront de points de repère pour l'interprétation de la fonction de Pickands estimée.

Copules de valeurs extrêmes

Suivant les sujets des premiers chapitres, les copules de valeurs extrêmes se présentent naturellement. D'abord, lorsqu'il est question de valeurs extrêmes, un sens d'ordonnancement est requis. Dans le contexte actuel, nous définissons l'ordonnancement des données Z_i par composante. C'est-à-dire :

$$\max_{1 \leq i \leq n} Z_i = (\max_{1 \leq i \leq n} X_i, \max_{1 \leq i \leq n} Y_i).$$

Comme le précise Tawn (1988), la difficulté de cette statistique est qu'elle peut être impossible à observer sous la forme d'une donnée.

Définition 2.0.5. Une copule C est définie max-stable si

$$C(u,v) = C^r(u^{1/r}, v^{1/r}), \quad u,v \in [0,1],$$

pour tout $r > 0$.

La pertinence de cette définition se trouve dans la définition suivante :

Définition 2.0.6. Une copule C est dite copule de valeurs extrêmes, s'il existe une copule C^* telle que

$$C(u,v) = \lim_{n \rightarrow \infty} C^{*n}(u^{1/n}, v^{1/n}), \quad u,v \in [0,1].$$

En effet, les définitions précédentes se rejoignent par le théorème suivant :

Théorème 2.0.7. Une copule est max-stable si et seulement si elle est une copule de valeurs extrêmes.

DÉMONSTRATION. Par définition, une copule max-stable est une copule de valeurs extrêmes. Si C est une copule de valeurs extrêmes, alors il existe C^* telle que

$$C^r(u^{1/r}, v^{1/r}) = \lim_{n \rightarrow \infty} C^{*rn}(u^{1/rn}, v^{1/rn}) = C(u,v), \quad u,v \in [0,1]$$

pour tout $r > 0$. □

Suite à quoi, le résultat principal de cette section est présenté : le théorème de Pickands.

Théorème 2.0.8 (Pickands (1981)). Soit C , une copule max-stable. Soit X et Y des variables aléatoires et distribuées selon la copule C , alors on a

$$C(u,v) = \exp \left\{ \log(uv) A \left(\frac{\log(u)}{\log(uv)} \right) \right\}, \quad u,v \in [0,1],$$

où A est la fonction de dépendance de la copule de valeurs extrêmes C - aussi appelée fonction de Pickands. Pour que le résultat soit valide, A doit respecter les conditions suivantes :

$$\max\{1-t, t\} \leq A(t) \leq 1, \quad t \in [0,1],$$

A est convexe.

Démonstration. La démonstration vient de Nelsen (2006).

Comme C est max-stable et X, Y sont des variables exponentielles de copule de survie C , les fonctions de survies marginales sont $\bar{F}_X(x) = e^{-x}$, $\bar{F}_Y(y) = e^{-y}$, $x, y \geq 0$.

La distribution de survie conjointe est alors

$$\bar{F}(x, y) = \mathbb{P}(X > x, Y > y) = C(e^{-x}, e^{-y}).$$

Comme C est max-stable, on a

$$\bar{F}(rx, ry) = C^r(e^{-x}, e^{-y}) = \bar{F}^r(x, y).$$

En choisissant $A : [0,1] \rightarrow [0,1]$ telle que

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

qui implique $C(e^{-(1-t)}, e^{-t}) = \exp\{-A(t)\}$.

En posant $(x,y) = (r(1-t), rt)$ avec $t \in [0,1], r > 0$, on a

$$\begin{aligned}\bar{F}(x, y) &= \bar{F}(r(1-t), rt) = \bar{F}^r((1-t), t) \\ &= C^r(e^{-(1-t)}, e^{-t}) = \exp\{-rA(t)\} \\ &= \exp\{-(x+y)A(y/(x+y))\}\end{aligned}$$

En remarquant $C(u,v) = \bar{F}(-\log(u), -\log(v))$, on obtient finalement

$$C(u,v) = \exp\{\log(uv)A(\log(u)/\log(uv))\}.$$

Pour que le côté droit soit une copule, il est impératif d'avoir

$$A(0) = A(1) = 1, \max\{1-t, t\} \leq A(t) \leq 1$$

et A est convexe.

□

Le problème d'identification de la copule C est alors équivalent au problème d'identification de la fonction A . Les estimateurs couverts dans le chapitre suivant sont conçus pour estimer A . Il est également pertinent de remarquer que la fonction A est définie sur le domaine unidimensionnel $[0,1]$. La variable aléatoire $\log(U)/\log(UV)$ est notée par Z .

Une remarque intéressante est que les bornes de Fréchet-Hoeffding pour C se traduisent, pour A , par $\max\{1-t, t\} \leq A(t) \leq 1$. Pour le cas d'indépendance entre U et V ,

$$A(z) = \log(C(u, v))/\log(uv) = \log(uv)/\log(uv) = 1, \quad (u, v) \in [0,1]^2.$$

Pour le cas de dépendance parfaite, $A(t) = \max[t, 1-t]$, $t \in [0,1]$. Les bornes des fonctions de Pickands servent alors de points de référence pour évaluer la dépendance entre les variables.

Plusieurs mesures de dépendance s'y basent telles que la norme infinie $\|1-A(t)\|_\infty$, la norme euclidienne $\|1-A(t)\|_2$, la norme absolue $\|1-A(t)\|_1$.

Par Tawn (1988), les exemples les plus courants de copules de valeurs extrêmes sont :

(1) Logistique : $A(t) = (\theta^r(1-t)^r + \theta^r t^r)^{1/r} + (\theta - \phi)t + 1 - \theta$,

avec $0 \leq \theta, \phi \leq 1, r \geq 1$.

(2) Mixte : $A(t) = \phi t^3 + \theta t^2 - (\theta + \phi)t + 1$,

avec $0 \leq \theta, \theta + \phi \leq 1, \theta + 2\phi \leq 1, \theta + 3\phi \geq 0$

Note : le modèle logistique est symétrique pour $\theta = \phi = 1$ et celui mixte l'est pour $\phi = 0$.

L'exemple présenté plus tôt dans le chapitre est repris pour illustrer les fonctions de Pickands. Dans la Figure 2.2 ci-dessous, le graphique de gauche présente les données simulées suivant la fonction de Pickands illustrée dans le graphique de droite. Telle que présentée plus haut, la borne supérieure de la fonction A représente l'indépendance des variables X et Y. La fonction de Pickands longeant cette borne indique un niveau de dépendance plus faible que celui présenté dans l'exemple de la Figure 2.3.

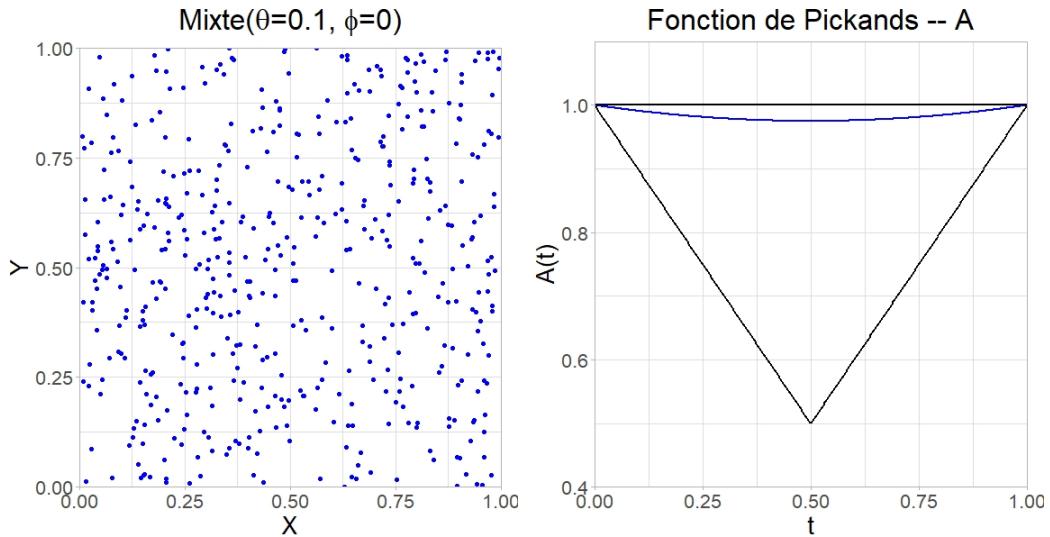


Figure 2.2. Données avec faible dépendance (Copule mixte ($\theta = 0.1, \phi = 0$)).

Dans ce second exemple de fonction de Pickands (Figure 2.3), le graphique de droite présente une fonction de Pickands longeant la borne inférieure suggérant une forte dépendance entre les variables X et Y. Cette dépendance est observée dans le graphique de gauche de la Figure 2.3.

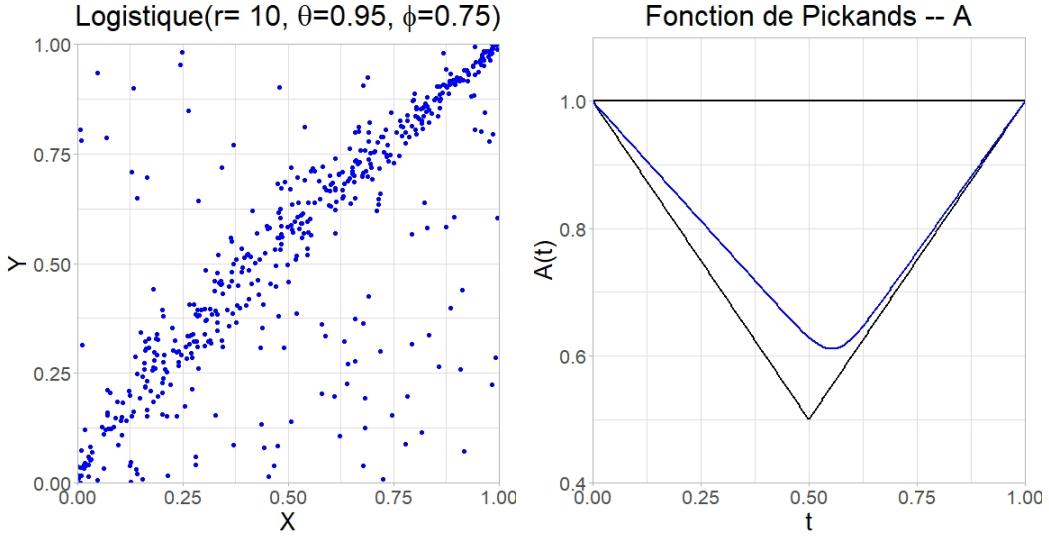


Figure 2.3. Données avec forte dépendance (Copule logistique ($\theta = 0.95, \phi = 0.75, r = 10$)).

Comme le fait remarquer Ahmadabadi et Ucer (2017), les copules de valeurs extrêmes permettent une modélisation de la dépendance entre des variables associées positivement. Cette remarque est causée par la nature de M_n , qui considère uniquement les maximums par composantes.

Une seconde remarque intéressante porte sur les résultats suivants présentés dans Capéraà et al. (1997) : deux des mesures d'association populaires, le coefficient de Kendall et de Spearman, peuvent respectivement être exprimées en fonction de la fonction de Pickands des variables :

$$\tau = \int_0^1 \frac{t(1-t)}{A(t)} dA'(t), \quad \rho = 12 \int_0^1 (A(t) + 1)^{-2} dt - 3.$$

Chapitre 3

Estimateurs et corrections non paramétriques

Dans ce chapitre, quelques méthodes d'estimation non paramétriques les plus courantes de la fonction de Pickands sont présentées et quelques-unes sont davantage expliquées. Comme le rapporte Tawn (1988), le choix non paramétrique est favorisé pour éviter les risques de modélisation apportés par les approches paramétriques.

Les estimateurs et corrections présentés sont évalués sur un jeu de données de 30 observations. Ces observations, présentées dans la Figure 3.1, ont été générées à partir d'une copule logistique($r = 1.5, \theta = 1, \phi = 1$) qui donne une fonction de Pickands symétrique.

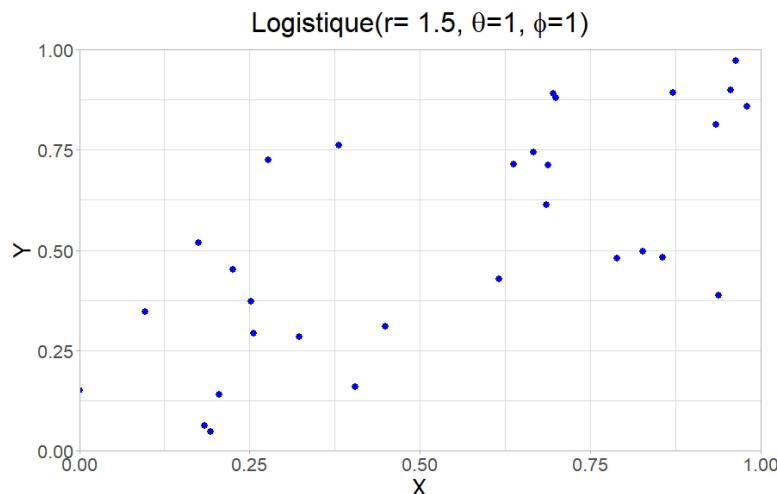


Figure 3.1. Jeu de données de 30 observations.

Estimateurs

Estimateur de Pickands (1981)

Le premier estimateur de la fonction de Pickands est celui portant le même nom. L'estimateur de Pickands, introduit en même temps que la fonction de Pickands, a vu ses propriétés asymptotiques être démontrées trois ans plus tard par Deheuvels (1984).

L'estimateur de Pickands A_n^P est donné par :

$$\frac{1}{A_n^P(t)} = \frac{1}{n} \sum_{i=1}^n \hat{\xi}_i(t), \quad t \in [0,1].$$

Les réalisations sont

$$\hat{\xi}_i(t) = \min\left\{\frac{-\log(u_i)}{t}, \frac{-\log(v_i)}{1-t}\right\},$$

où $u_i = F_X(x_i)$ et $v_i = F_Y(y_i)$, $i \in \{1, \dots, n\}$, avec (x_i, y_i) les observations associées au vecteur aléatoire (X, Y) de marges F_X et F_Y .

Comme le présente Ahmadabadi et Ucer (2017), l'obtention de cet estimateur est justifiée comme suit. On pose

$$\begin{aligned} \mathbb{P}(U \leq u, V \leq v) &= C(u, v) = \exp\left(\log(uv)A\left(\frac{\log(u)}{\log(uv)}\right)\right) \\ S &= -\log(U) \\ T &= -\log(V), \\ \xi(t) &= \min\left(\frac{S}{t}, \frac{T}{1-t}\right), \quad t \in [0,1], \end{aligned}$$

où (U, V) est distribué selon la copule C , S et T suivent une distribution exponentielle de paramètre 1.

On obtient alors

$$\mathbb{P}(\xi(t) > x) = \mathbb{P}(U < e^{-tx}, V < e^{-(1-t)x}) = C(e^{-tx}, e^{-(1-t)x}) = e^{-xA(t)}.$$

Ainsi la distribution de $\xi(t)$ est une Exponentielle($A(t)$) et que $\mathbb{E}(\xi(t)) = 1/A(t)$, d'où l'estimation par moyenne empirique des $\hat{\xi}_i(t)$.

Gudendorf et Segers (2009) font remarquer que malgré l'inconvénient majeur de ne pas respecter les critères de bornes et de convexité des fonctions de Pickands, cet estimateur est encore considéré lorsqu'il est question de comparer de nouveaux estimateurs de la fonction de Pickands. Une des principales raisons vient de ses propriétés asymptotiques.

La Figure 3.2 présente l'estimateur de Pickands évalué sur les données précédemment générées. Comme le graphique l'illustre, l'estimateur de Pickands ne respecte pas les contraintes de bornes et de convexité imposées par la fonction de Pickands.

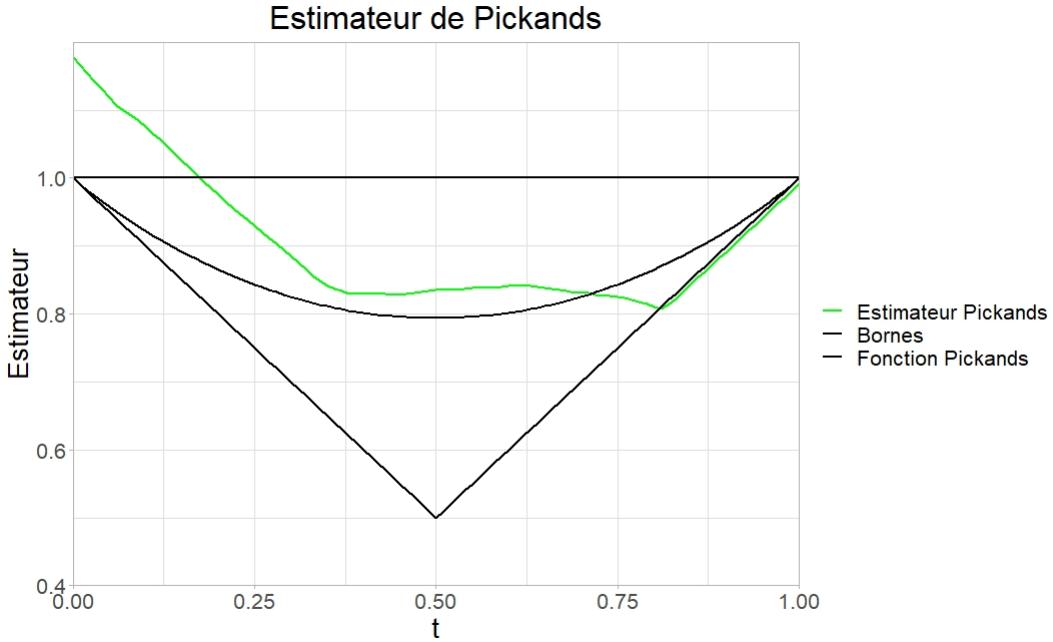


Figure 3.2. Estimation par l'estimateur de Pickands.

Estimateur de Deheuvels (1991)

Pour corriger le biais de la fonction évaluée aux extrémités de l'intervalle $[0,1]$, Deheuvels (1991) propose l'estimateur alternatif A_n^D suivant :

$$\frac{1}{A_n^D(t)} = \frac{1}{n} \sum_{i=1}^n \hat{\xi}_i(t) - \frac{t}{n} \sum_{i=1}^n \hat{\xi}_i(1) - \frac{1-t}{n} \sum_{i=1}^n \hat{\xi}_i(0) + 1, \quad t \in [0,1],$$

avec $\hat{\xi}_i(t)$ défini comme pour l'estimateur précédent. Comme le précise Hall et Tajvidi (2000), malgré cet ajustement apporté à l'estimateur de Pickands, l'estimateur de Deheuvels ne garantit pas la convexité et peut, de même que pour son enveloppe convexe, se trouver hors des bornes exigées des fonctions de Pickands.

La Figure 3.3 présente l'estimateur de Deheuvels évalué sur les données générées plus tôt. L'estimateur de Pickands est illustré à titre comparatif. Sans respecter les contraintes de bornes et de convexité, l'estimateur corrige le biais de l'estimateur de Pickands évalué aux extrémités.

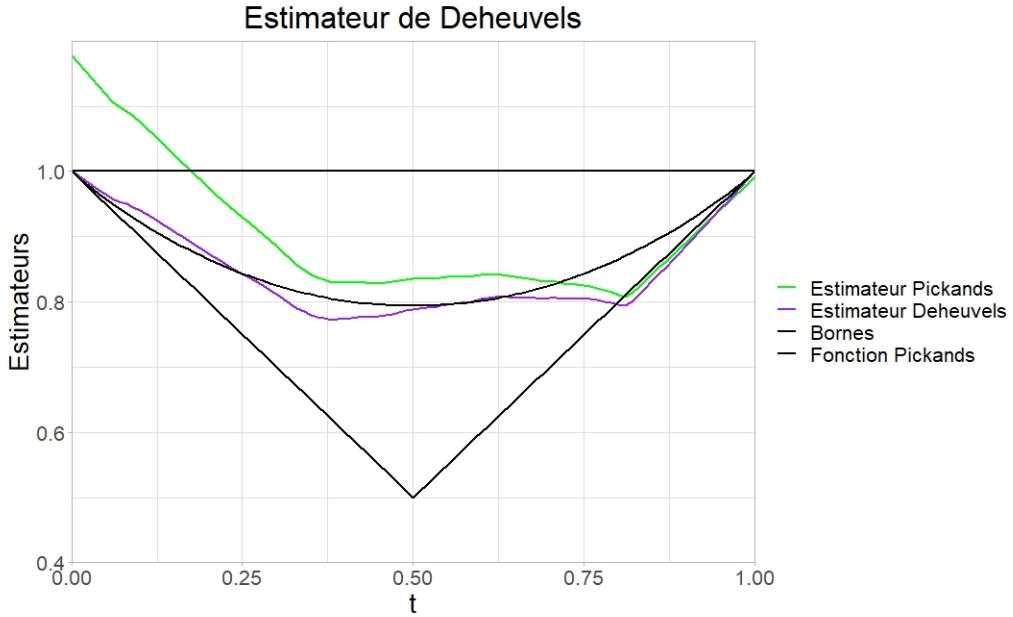


Figure 3.3. Estimation par l'estimateur de Deheuvels.

Estimateur de Capéraà, Fougères et Genest (1997)

L'estimateur présenté ici a été publié en promouvant sa capacité d'estimation pour des échantillons de petites tailles, comparativement à celles des estimateurs présentés jusqu'à maintenant. Dans l'article de Capéraà et al. (1997), les auteurs comparent la performance de leur estimateur avec celle des estimateurs précédents en se basant sur des simulations numériques de tailles d'échantillon $n = 100$.

L'estimateur de CFG, A_n^{CFG} , est défini comme suit. Soit

$$Q_i = \left\{ \prod_{k=1}^i z_{(k)} / (1 - z_{(k)}) \right\}^{1/n},$$

$z_i = \log(u_i) / \log(u_i v_i)$ et $z_{(k)}$ = k^e statistique d'ordre des z_i .

On définit alors

$$A_n^{CFG}(t) = \begin{cases} (1-t)Q_n^{1-p(t)} & \text{Si } 0 \leq t \leq z_{(1)}, \\ t^{i/n}(1-t)^{1-i/n}Q_n^{1-p(t)}Q_i^{-1} & \text{si } z_{(i)} \leq t \leq z_{(i+1)}, \\ tQ_n^{-p(t)} & \text{si } z_{(n)} \leq t \leq 1. \end{cases}$$

Les z_i sont distincts et $0 \leq p(t) \leq 1$. Les auteurs suggèrent $p(t) = 1-t$.

La justification de l'estimateur est la suivante. Pour $Z = \log(U)/\log(UV)$ avec (U,V) distribué selon la copule C , de fonction de Pickands A , on cherche à calculer

$$H(z) = \mathbb{P}(Z \leq z).$$

Ceci est obtenu en calculant $\mathbb{P}(U \leq u, Z \leq z) = \mathbb{P}(U \leq u, V \leq U^{\frac{1-z}{z}})$.

Pour ce faire, il faut d'abord savoir :

$$\begin{aligned} \frac{\partial}{\partial u} F(u, v) &= \frac{\partial}{\partial u} \exp \left\{ \log(uv) A \left(\frac{\log(u)}{\log(uv)} \right) \right\} \\ &= \exp \left\{ \log(uv) A \left(\frac{\log(u)}{\log(uv)} \right) \right\} \left[\frac{v}{uv} A \left(\frac{\log(u)}{\log(uv)} \right) + \right. \\ &\quad \left. \log(uv) A' \left(\frac{\log(u)}{\log(uv)} \right) \frac{\log(uv) - \log(u)}{u \log^2(uv)} \right] \\ &= \exp \left\{ \log(uv) A \left(\frac{\log(u)}{\log(uv)} \right) \right\} \frac{1}{u} \left[A \left(\frac{\log(u)}{\log(uv)} \right) + A' \left(\frac{\log(u)}{\log(uv)} \right) \frac{\log(uv) - \log(u)}{\log(uv)} \right] \\ &= \exp \left\{ \log(uv) A \left(\frac{\log(u)}{\log(uv)} \right) \right\} \frac{1}{u} \left[A \left(\frac{\log(u)}{\log(uv)} \right) + A' \left(\frac{\log(u)}{\log(uv)} \right) \left(1 - \frac{\log(u)}{\log(uv)} \right) \right]. \end{aligned}$$

On peut ainsi calculer

$$\begin{aligned} \mathbb{P}(U \leq u, Z \leq z) &= \mathbb{P}(U \leq u, V \leq U^{\frac{1-z}{z}}) = \int_0^u \int_0^{t^{1/z-1}} \frac{\partial}{\partial u} \frac{\partial}{\partial v} F(t, s) ds dt \\ &= \int_0^u \frac{\partial}{\partial u} F(t, t^{1/z-1}) dt \\ &= \int_0^u \frac{1}{t} \exp \left\{ \log(t^{1/z}) A \left(\frac{\log(t)}{\log(t^{1/z})} \right) \right\} \left[A \left(\frac{\log(t)}{\log(t^{1/z})} \right) + \right. \\ &\quad \left. A' \left(\frac{\log(t)}{\log(t^{1/z})} \right) \left(1 - \frac{\log(t)}{\log(t^{1/z})} \right) \right] dt \\ &= \int_0^u \frac{1}{t} \exp \left\{ \log(t^{1/z}) A(z) \right\} \left[A(z) + A'(z)(1-z) \right] dt \\ &= \int_0^u t^{A(z)/z-1} dt \left[A(z) + A'(z)(1-z) \right] \\ &= u^{A(z)/z} \frac{z}{A(z)} \left[A(z) + A'(z)(1-z) \right] \\ &= u^{A(z)/z} \left[z + \frac{A'(z)}{A(z)} (1-z) z \right] \end{aligned}$$

Ce qui donne $H(z) = \mathbb{P}(U \leq 1, Z \leq z) = z + \frac{A'(z)}{A(z)} (1-z) z$.

En résolvant le système d'équations différentielles pour A , on obtient

$$\begin{aligned}
\frac{H(z) - z}{z(1-z)} &= \frac{A'(z)}{A(z)} \Leftrightarrow A(z) \frac{H(z) - z}{z(1-z)} = A'(z) \\
&\Rightarrow A(z) = \exp \left\{ \int_0^z \frac{H(u) - u}{u(1-u)} du + C \right\} \\
&\Rightarrow \frac{A(t)}{A(s)} = \exp \left\{ \int_s^t \frac{H(u) - u}{u(1-u)} du \right\} \\
&\Rightarrow A(t) = \exp \left\{ \int_0^t \frac{H(u) - u}{u(1-u)} du \right\},
\end{aligned}$$

puisque $A(0) = 1$, par définition de la fonction de Pickands. De même, on peut obtenir

$$A(t) = \exp \left\{ - \int_t^1 \frac{H(u) - u}{u(1-u)} du \right\}.$$

L'estimateur présenté se base sur ces formes de A . En effet, les auteurs proposent alors deux estimateurs non paramétriques :

$$A_n^0(t) = \exp \left\{ \int_0^t \frac{H_n(u) - u}{u(1-u)} du \right\} \quad \text{et} \quad A_n^1(t) = \exp \left\{ - \int_t^1 \frac{H_n(u) - u}{u(1-u)} du \right\},$$

où H_n est la fonction de répartition empirique des z_i .

L'estimateur de CFG combine ces deux estimateurs de sorte que

$$\log A_n^{CFG}(t) = p(t) \log A_n^0(t) + (1 - p(t)) \log A_n^1(t).$$

Des calculs, présentés dans l'annexe 2, prouvent l'équivalence de cette forme de l'estimateur avec celle présentée plus tôt.

Il a été mentionné que rien ne garantit le respect des bornes pour A_n^{CFG} . En fait, l'estimateur ne les respecte pas presque sûrement.

Soit n la taille de l'échantillon, $t \in [0, z_{(1)}]$ et $t' \in [z_{(n)}, 1]$ où $z_{(i)}$ est la i^e statistique d'ordre des z_j , $j \in \{1, 2, \dots, n\}$, on a

$$\begin{aligned}
A_n^{CFG}(t) &= (1 - t) Q_n^{1-p(t)} < 1 - t \\
\Leftrightarrow Q_n^{1-p(t)} &< 1 \Leftrightarrow \prod_{k=1}^n \frac{z_{(k)}}{1 - z_{(k)}} < 1
\end{aligned}$$

et

$$\begin{aligned} A_n^{CFG}(t') &= t' Q_n^{-p(t)} < t' \\ \Leftrightarrow Q_n > 1 &\Leftrightarrow \prod_{k=1}^n \frac{z(k)}{1-z(k)} > 1. \end{aligned}$$

Comme $\mathbb{P}\left(\prod_{k=1}^n \frac{z(k)}{1-z(k)} = 1\right) = 0$ pour des z_j indépendants, l'estimateur de CFG ne respecte pas les bornes, presque sûrement.

L'estimateur de CFG a la propriété d'être consistant et asymptotiquement normal. Par simulations, Gudendorf et Segers (2009) indiquent que cet estimateur est plus performant que ceux de Pickands, de Deheuvels et de Hall-Tajvidi.

De plus, par Segers (2004), l'estimateur de CFG s'écrit dans une expression simple qui fait intervenir la forme des estimateurs précédents. On a

$$\log(A_n^{CFG}(t)) = -\frac{1}{n} \sum_{i=1}^n \log(\hat{\xi}_i(t)) + \frac{p(t)}{n} \sum_{i=1}^n \hat{\xi}_i(0) + \frac{1-p(t)}{n} \sum_{i=1}^n \hat{\xi}_i(1).$$

Ou encore, puisque la constante d'Euler $\gamma = 0.577.. = -\mathbb{E}(\log(X))$, avec X une variable aléatoire exponentielle de paramètre 1 et $\xi(t)$ est de loi exponentielle de paramètre $A(t)$, une approximation de l'estimateur de CFG devient

$$\log(A_n^{CFG}(t)) = \frac{-1}{n} \sum_{i=1}^n \log(\hat{\xi}_i(t)) - \gamma, \quad t \in [0,1].$$

La Figure 3.4 présente l'estimateur de CFG évalué sur les données générées plus tôt. Les estimateurs précédents sont illustrés à titre comparatif. L'estimateur ne respecte pas les contraintes de bornes et de convexité imposées, mais performe mieux pour des échantillons de petites tailles que les précédents, selon Capéraà et al. (1997) et Gudendorf et Segers (2009).

Estimateur de Hall-Tajvidi (2000)

L'estimateur de Hall et Tajvidi (2000) est le seul, parmi ceux présentés jusqu'à maintenant, à respecter la contrainte de borne inférieure de la fonction de Pickands.

Il est défini comme suit. Soit

$$\bar{\xi}_i(t) = \min\left\{\frac{r_i}{t}, \frac{s_i}{1-t}\right\},$$

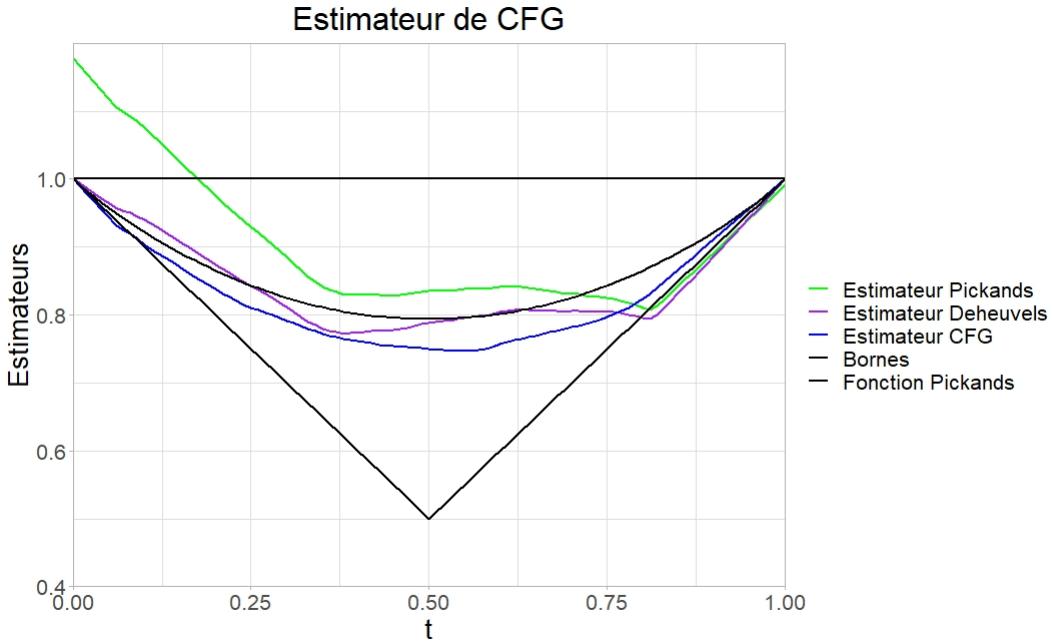


Figure 3.4. Estimation par l'estimateur de CFG.

où

$$r_i = \frac{\log(u_i)}{\frac{1}{n} \sum_{i=1}^n \log(u_i)} \quad \text{et} \quad s_i = \frac{\log(v_i)}{\frac{1}{n} \sum_{i=1}^n \log(v_i)},$$

l'estimateur de Hall-Tajvidi A_n^{HT} est donné par

$$\frac{1}{A_n^{HT}(t)} = \frac{1}{n} \sum_{i=1}^n \bar{\xi}_i(t), \quad t \in [0,1].$$

Une remarque intéressante par Hall et Tajvidi (2000) est la distinction entre leur estimateur et celui de Deheuvels : «Les deux étant des modifications de l'estimateur de Pickands, l'estimateur de Deheuvels propose un ajustement additif pour le biais tandis que celui présenté à l'instant propose un ajustement multiplicatif qui a pour effet de normaliser les moments marginaux à la forme adéquate».

Gudendorf et Segers (2009) précisent que parmi les estimateurs de Pickands, de Deheuvels et de Hall-Tajvidi, celui de Hall-Tajvidi a la plus petite variance asymptotique.

La Figure 3.5 présente l'estimateur de Hall-Tajvidi évalué sur les données générées plus tôt. Les estimateurs précédents sont illustrés à titre comparatif. L'estimateur de Hall-Tajvidi est le seul à respecter les contraintes de bornes imposées, sans toutefois respecter celle de convexité.

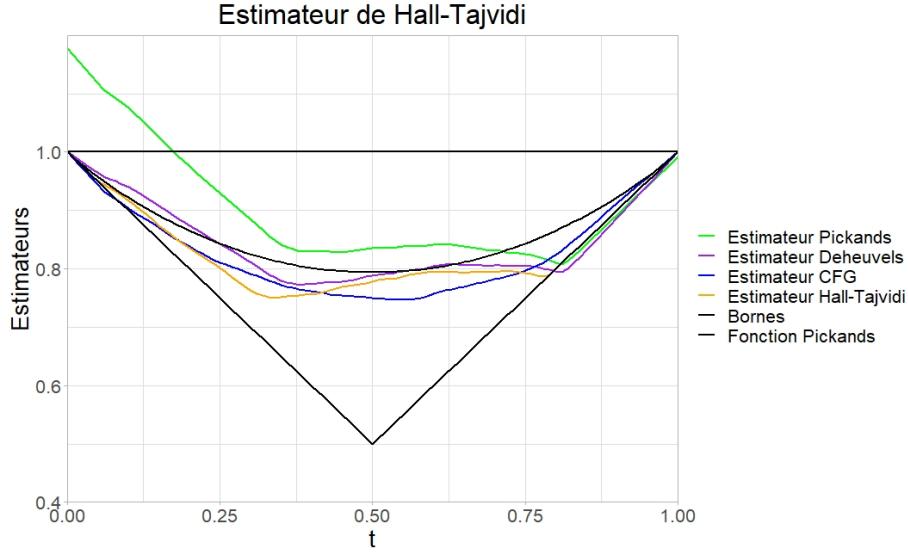


Figure 3.5. Estimation par l'estimateur de Hall-Tajvidi.

Correction d'estimateurs

Dans la seconde partie de ce chapitre, on présente des corrections d'estimateurs de la fonction de Pickands. C'est-à-dire des techniques à être appliquées sur un estimateur pour lui permettre de respecter les contraintes de la fonction de Pickands.

Selon Fils-Villetard et al. (2008), deux principales approches sont généralement empruntées. La plus évidente est de forcer l'estimateur à demeurer à l'intérieur des bornes et de prendre son plus grand minorant convexe, tandis que la seconde est d'approcher l'estimateur obtenu à l'aide d'une spline définie dans l'espace des fonctions de Pickands \mathcal{A} . Ces auteurs mentionnent le résultat pertinent que pour chacune de ces approches, l'estimateur corrigé est consistant si l'estimateur initial l'est.

Voici un exemple pour chacune de ces approches.

Plus grand minorant convexe

L'idée de ce type de correction est de prendre le plus grand minorant de l'estimateur initial qui est élément de \mathcal{A} . Un exemple de cette approche qui est similaire à celle qui sera présentée dans la correction proposée est de définir

$$\hat{A}^*(t) = \min(1, \max(\hat{A}(t), 1-t, t)), \quad t \in [0,1]$$

où $\hat{A}(t)$ est l'estimateur initial de la fonction de Pickands. La correction du plus grand minorant convexe peut être approchée en prenant l'enveloppe convexe de $\hat{A}^*(t)$ sur un ensemble fini de points.

La définition de $\hat{A}^*(t)$ assure le respect de la contrainte de bornes, tandis que l'enveloppe convexe garantit la convexité du nouvel estimateur.

La Figure 3.6 présente la correction du plus grand minorant convexe (PGMC) appliquée aux estimateurs de Pickands et de CFG. Pour obtenir les plus grands minorants convexes, les enveloppes convexes des estimateurs ont été calculées en $t = \{0; 0,01; 0,02; \dots; 0,99; 1\}$. Les nouveaux estimateurs obtenus respectent les contraintes imposées par les fonctions de Pickands.

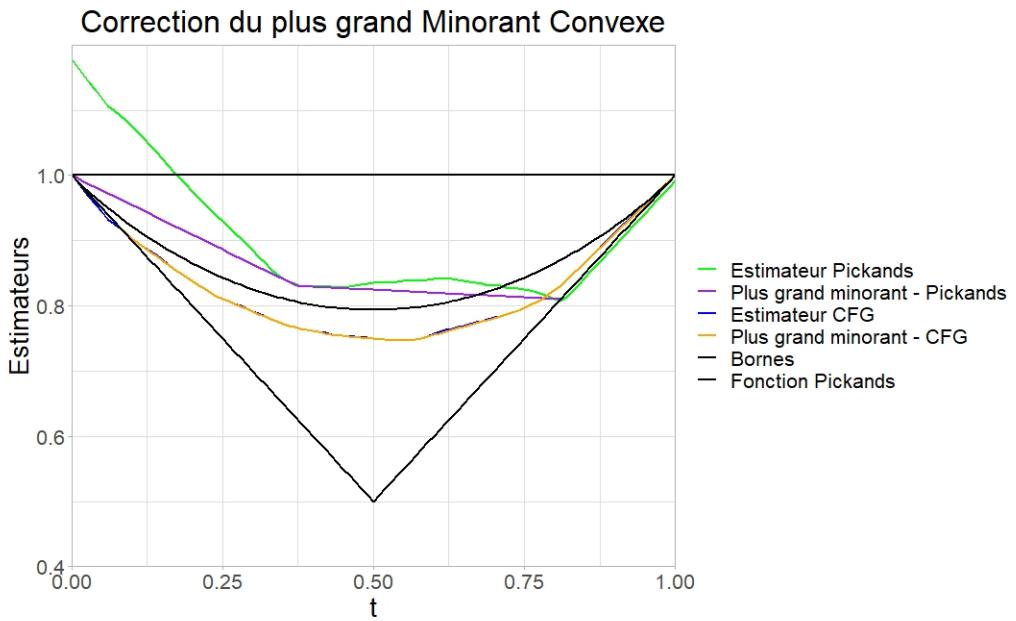


Figure 3.6. Exemple de correction par le PGMC.

Spline définie dans \mathcal{A}

L'idée de ce type de correction est d'approcher l'estimateur initial \hat{A} de la fonction de Pickands par une spline dans \mathcal{A} . Plusieurs méthodes existent pour ce faire, une des plus populaires est celle présentée dans Hall et Tajvidi (2000) à l'aide d'une spline de degrés 3 ou plus.

L'exemple utilisé est la correction de Fils-Villetard et al. (2008). L'idée de la correction FGS est de projeter l'estimateur initial dans l'espace des fonctions de Pickands, \mathcal{A} , de sorte à minimiser la distance quadratique entre l'estimateur et sa projection dans \mathcal{A} . N'ayant pas une forme explicite pour la projection optimale, les auteurs proposent l'algorithme suivant permettant d'obtenir une spline linéaire approximative. Les auteurs précisent que les avantages d'autres types de splines sont négligeables.

La spline linéaire est composée de $m + 1$ segments, où m est un hyperparamètre que doit fournir l'usager. Les auteurs mentionnent que $m = 20$ est généralement suffisant.

Le lecteur est encouragé à consulter l'article de Fils-Villetard et al. (2008) pour la justification théorique. Seul l'algorithme est présenté ici.

Algorithme de la correction FGS :

(1) Estimer la fonction de Pickands A par un estimateur initial \hat{A} .

(2) Choisir le nombre de segments, $m+1$, de la spline linéaire pour approcher \hat{A} dans \mathcal{A} .

(3) Définir les constantes du problème

$$\mathbf{A}_{m+1} = \frac{1}{6m} \begin{bmatrix} 2 & 1 & 0 & 0 & \cdots & 0 \\ 1 & 4 & 1 & 0 & \cdots & 0 \\ 0 & 1 & 4 & 1 & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \ddots & \\ 0 & \cdots & 1 & 4 & 1 & \\ 0 & \cdots & 0 & 1 & 2 & \end{bmatrix}$$

et

$$\mathbf{b}_i = \begin{cases} \frac{1}{3m} \left(\frac{1}{2} \hat{A}(0) + \hat{A}\left(\frac{1}{2m}\right) \right) & \text{si } i = 0 \\ \frac{1}{3m} \left(\hat{A}\left(\frac{2i-1}{2m}\right) + \hat{A}\left(\frac{i}{m}\right) + \hat{A}\left(\frac{2i+1}{2m}\right) \right) & \text{si } i = 1, \dots, m-1 \\ \frac{1}{3m} \left(\hat{A}\left(\frac{2m-1}{2m}\right) + \frac{1}{2} \hat{A}(1) \right) & \text{si } i = m \end{cases}$$

(4) Trouver λ qui minimise $\lambda^T \mathbf{A} \lambda - 2\lambda^T \mathbf{b}$, $\lambda \in \Lambda_m \subset \mathbb{R}^{m+1}$ fermé et convexe,
s.à.

$$\begin{cases} \lambda_0 = \lambda_m = 1 \\ 1 - \lambda_1 \leq \frac{1}{m} \\ \lambda_{i-1} - 2\lambda_i + \lambda_{i+1} \geq 0 \\ 1 - \lambda_{m-1} \leq \frac{1}{m} \end{cases}$$

De sorte que $A^{FGS} = \sum_{i=0}^m \lambda_i h_{im}$, avec h_{im} qui est l'unique fonction linéaire par morceaux à m nœuds telle que $h_{im}(j/m) = \delta_i(j)$. Autrement dit, A^{FGS} est la fonction qui relie linéairement les points $(\frac{i}{m}, \lambda_i)_{i=0,1,\dots,m}$.

Un inconvénient de l'algorithme présenté est la nécessité de fournir l'hyperparamètre m .

La Figure 3.7 illustre la correction FGS appliquée à l'estimateur de Pickands et de CFG pour les données présentées plus tôt. L'estimateur de Pickands initial est considérablement corrigé. L'estimateur de CFG initial a une forme près de celle souhaitée, la correction apportée est légère. Les nouveaux estimateurs obtenus respectent les bornes et la convexité demandées. Contrairement à la correction du PGMC, la correction par spline permet l'existence de t_0 tel que pour l'estimateur corrigé \hat{A}^* , $\hat{A}^*(t_0) \geq \hat{A}(t_0)$ (pour t_0 tel que $\hat{A}(t_0)$ respecte les bornes de la fonction de Pickands).

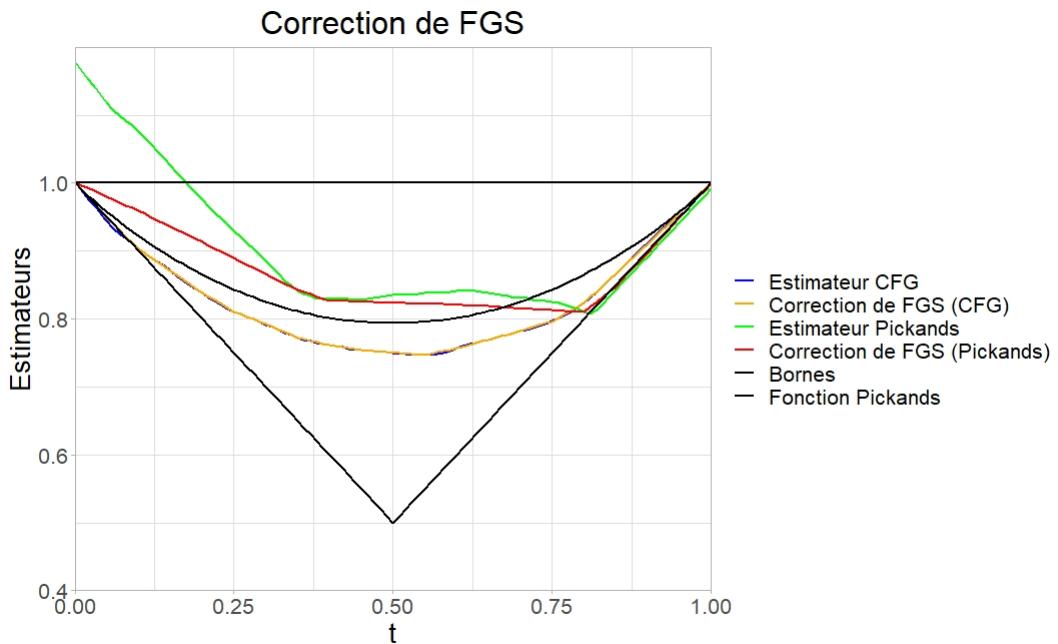


Figure 3.7. Correction d'estimateurs par FGS.

Conclusion

Les estimateurs présentés ci-haut ainsi que la correction FGS appliquée sur l'estimateur de CFG seront repris lors de simulations dans le prochain chapitre. L'estimateur de CFG sert également de point de départ pour le développement de la correction convexe et de l'estimateur bayésien qui forment le sujet de ce mémoire.

Chapitre 4

Correction convexe et estimateur bayésien

Dans ce dernier chapitre, la correction d'estimateurs et l'estimateur bayésien proposés sont développés. Pour chacun, les prérequis théoriques sont présentés, suivi de la correction ou de l'estimateur, puis quelques résultats et une application sur les données générées au chapitre 3. Suivant la présentation des deux contributions de ce mémoire, elles sont évaluées et comparées à quelques estimateurs et corrections du chapitre 3 sur un ensemble de données générées à partir de 18 distributions de copules de valeurs extrêmes. L'évaluation est basée sur l'erreur quadratique moyenne entre l'estimateur obtenu et la fonction de Pickands qui a servi à générer les données.

Correction par centralisation et enveloppe convexe

Comme son nom le suggère, la correction proposée est divisée en deux étapes. Une première de centralisation, puis une seconde d'obtention de l'enveloppe convexe de l'estimateur. La centralisation des données est une idée originale. Elle est basée sur la propriété développée ci-bas. La correction est conçue pour les estimateurs qui sont une fonction des observations z_i . L'estimateur de CFG en est un exemple.

Propriété de la variable aléatoire Z

On définit la variable aléatoire $Z = \log(U) / \log(UV)$, avec (U, V) distribué selon la copule C . Le lemme suivant permet d'obtenir la propriété de Z qui justifie la correction proposée.

Lemme 4.0.1. Soit X une variable aléatoire dont l'espérance existe et H sa fonction de répartition, on obtient

$$\mathbb{E}(X) = \int_0^\infty (1 - H(t))dt - \int_{-\infty}^0 H(t)dt.$$

Démonstration. En effet,

$$\begin{aligned}
\mathbb{E}(X) &= \int_{-\infty}^{\infty} x \mathbb{P}(dx) = \int_0^{\infty} x \mathbb{P}(dx) + \int_{-\infty}^0 x \mathbb{P}(dx) \\
&= \int_0^{\infty} \int_0^x dt \mathbb{P}(dx) - \int_{-\infty}^0 \int_x^0 dt \mathbb{P}(dx) \\
&= \int_0^{\infty} \int_t^{\infty} \mathbb{P}(dx) dt - \int_{-\infty}^0 \int_{-\infty}^t \mathbb{P}(dx) dt \\
&= \int_0^{\infty} (1 - H(t)) dt - \int_{-\infty}^0 H(t) dt
\end{aligned}$$

□

Posons $Y = \log\left(\frac{Z}{1-Z}\right)$, $dy = \frac{dz}{z(1-z)}$, on a

$$\mathbb{P}(Y \leq y) = \mathbb{P}\left(Z \leq \frac{e^y}{1+e^y}\right) = H\left(\frac{e^y}{1+e^y}\right) := G(y).$$

Soit A la fonction de Pickands qui permet de générer Z et H la fonction de répartition de Z , on rappelle que Capéraà et al. (1997) ont montré

$$A(t) = \exp\left\{ \int_0^t \frac{H(z) - z}{z(1-z)} dz \right\} = \exp\left\{ - \int_t^1 \frac{H(z) - z}{z(1-z)} dz \right\}.$$

Ceci donne,

$$\log A(1/2) = \int_0^{1/2} \frac{H(z) - z}{z(1-z)} dz = \int_0^{1/2} \frac{H(z)}{z(1-z)} dz + \log(1/2)$$

et

$$\begin{aligned}
\log A(1/2) &= - \int_{1/2}^1 \frac{(H(z) - 1) - (z - 1)}{z(1-z)} dz = \int_{1/2}^1 \frac{(1 - H(z))}{z(1-z)} dz + \log(1/2). \\
&\Rightarrow \int_{1/2}^1 \frac{(1 - H(z))}{z(1-z)} dz - \int_0^{1/2} \frac{H(z)}{z(1-z)} dz = 0 \\
&\Rightarrow \int_0^{\infty} (1 - G(y)) dy - \int_{-\infty}^0 G(y) dy = \mathbb{E}(Y) = 0, \text{ par le lemme.}
\end{aligned}$$

Ainsi, $\mathbb{E}(Y) = 0 = \mathbb{E}\left(\log\left(\frac{Z}{1-Z}\right)\right)$. Cette propriété de Y – et de Z – est centrale dans notre proposition de correction.

Enveloppe convexe

La deuxième partie de notre correction se base sur le concept de l'enveloppe convexe d'une fonction – ou d'un estimateur dans notre cas. On doit alors proprement définir l'enveloppe convexe et donner des méthodes pour l'approcher numériquement.

Définition 4.0.2. Soit E un espace et A un sous-ensemble de E , on définit l'enveloppe convexe de A comme l'intersection de tous les sous-ensembles convexes de E qui contiennent A .

Une propriété de l'enveloppe convexe est d'être le plus petit ensemble convexe de E qui contient A . Cette propriété est fort désirable puisqu'elle nous permet d'obtenir le plus grand minorant convexe d'un estimateur.

Pour un estimateur quelconque de la fonction de Pickands, il faut employer une approximation de son enveloppe convexe parce qu'on ne peut pas obtenir la fonction explicite. Plusieurs algorithmes existent pour obtenir cette approximation : *Gift Wrapping* (1973), *Quick-hull* (1977), *Chan's algorithm* (1996).

Correction par centralisation et enveloppe convexe

Soit un estimateur initial de la fonction de Pickands \hat{A} , on définit

$$\hat{A}^*(t) = \min(1, \max(\hat{A}(t), 1 - t, t)), \quad t \in [0,1].$$

La correction de l'estimateur de la fonction de Pickands consiste à obtenir l'estimateur \hat{A}^* basé les observations z_i transformées -notées z_i^* - et à prendre l'enveloppe convexe de \hat{A}^* pour garantir la convexité.

Soit les observations z_i , nous centrons les observations $y_i = \log\left(\frac{z_i}{1-z_i}\right)$,

de telle sorte que

$$y_i^* = y_i - \bar{y} \quad \text{et} \quad z_i^* = \frac{e^{y_i^*}}{1 + e^{y_i^*}},$$

où \bar{y} est la moyenne empirique des y_i , et $i = \{1, 2, \dots, n\}$. Nous rappelons que la raison de la centralisation des y_i est le résultat du lemme 4.1 qui mentionne $\mathbb{E}(Y) = 0$.

Pour l'estimateur de CFG, l'action de centraliser les observations z_i vient garantir le respect des bornes inférieures de la fonction de Pickands. En effet, on peut montrer que $A_n^0(t)$ et $A_n^1(t)$ respectent les bornes inférieures sur $[0, 1/2]$ et $[1/2, 1]$ respectivement et que $A_n^0 = A_n^1$.

Effectivement, on a sous cette égalité

$$\log(A_n^{CFG}(t)) = p(t) \log(A_n^0(t)) + (1 - p(t)) \log(A_n^1(t)) = \log(A_n^0(t)) = \log(A_n^1(t))$$

où A_n^0 et A_n^1 respectent les bornes sur $[0, 1/2]$ et $[1/2, 1]$ respectivement, ce qui implique que A_n^{CFG} respecte les bornes sur tout son domaine.

Montrons d'abord que $A_n^0(t)$ et $A_n^1(t)$ respectent les bornes sur $[0, 1/2]$ et $[1/2, 1]$ respectivement.

Pour $0 \leq t \leq 1$, H une fonction de répartition quelconque à domaine $[0, 1]$, $A^0(t)$ et $A^1(t)$ étant définies telles que

$$A^0(t) = \exp \left\{ \int_0^t \frac{H(u) - u}{u(1-u)} du \right\} \quad \text{et} \quad A^1(t) = \exp \left\{ - \int_t^1 \frac{H(u) - u}{u(1-u)} du \right\}.$$

Ceci implique

$$\begin{aligned} \log A^0(t) &= \int_0^t \frac{H(u) - u}{u(1-u)} du = \int_0^t \frac{H(u)}{u(1-u)} du - \int_0^t \frac{1}{(1-u)} du \\ &= \int_0^t \frac{H(u)}{u(1-u)} du + \log(1-t) \\ \Rightarrow A^0(t) &\geq 1 - t, \text{ car } H(u) \geq 0 \text{ et } 0 \leq u \leq t \leq 1. \end{aligned}$$

De même, on a

$$\begin{aligned} \log A^1(t) &= - \int_t^1 \frac{H(u) - u}{u(1-u)} du = \int_t^1 \frac{u - H(u)}{u(1-u)} du = \int_t^1 \frac{(1 - H(u)) - (1 - u)}{u(1-u)} du \\ &= \int_t^1 \frac{(1 - H(u))}{u(1-u)} du - \int_t^1 \frac{1}{u} du \\ &= \int_t^1 \frac{(1 - H(u))}{u(1-u)} du + \log(t) \\ \Rightarrow A^1(t) &\geq t. \end{aligned}$$

Montrons maintenant que $A_n^0 = A_n^1$.

D'abord, on a

$$\begin{aligned} \prod_{i=1}^n \left(\frac{z_i^*}{1 - z_i^*} \right)^{1/n} &= \exp \left(\log \left(\prod_{i=1}^n \left(\frac{z_i^*}{1 - z_i^*} \right)^{1/n} \right) \right) \\ &= \exp \left(1/n \sum_{i=1}^n \log \left(\frac{z_i^*}{1 - z_i^*} \right) \right) \\ &= \exp(\bar{y}^*) = \exp(0) = 1. \end{aligned}$$

Estimateur CFG corrigé

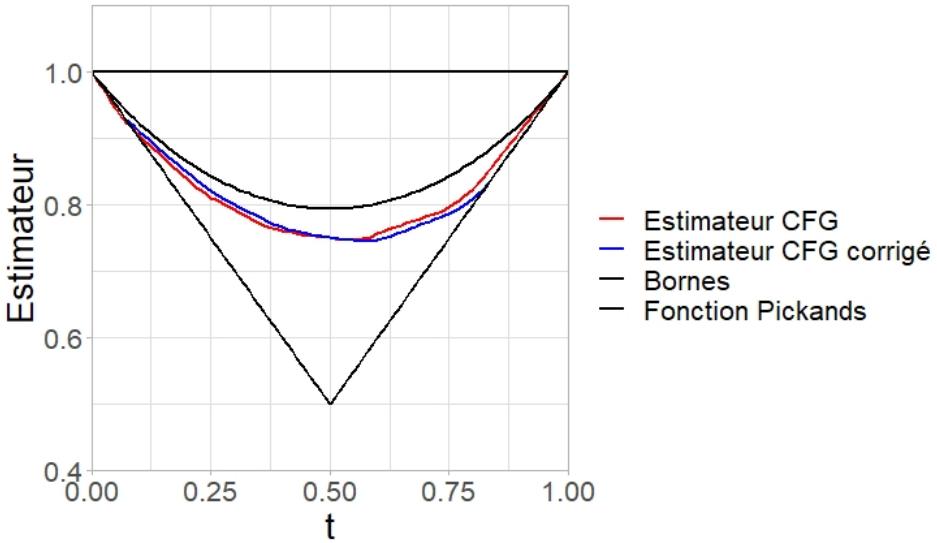


Figure 4.1. Estimateur de CFG corrigé.

Ceci implique

$$\begin{aligned}
 1 &= \prod_{i=1}^n \frac{z_i^*}{1 - z_i^*} \\
 &= \prod_{i=1}^j \frac{z_i^*}{1 - z_i^*} \prod_{i=j+1}^n \frac{z_i^*}{1 - z_i^*} \\
 \Leftrightarrow \prod_{i=1}^j \frac{1 - z_i^*}{z_i^*} &= \prod_{i=j+1}^n \frac{z_i^*}{1 - z_i^*}.
 \end{aligned}$$

Donc, par leur forme développée dans l'annexe 2,

$$A_n^0(t) = t^{j/n}(1-t)^{1-j/n} \prod_{i=1}^j \frac{1 - z_i^*}{z_i^*} = t^{j/n}(1-t)^{1-j/n} \prod_{i=j+1}^n \frac{z_i^*}{1 - z_i^*} = A_n^1(t),$$

pour $t \in [z_{(j)}, z_{(j+1)}]$ et $n \in \{1, 2, \dots, n\}$. Pour $t \in [0, z_{(1)}]$ et $t \in [z_{(n)}, 1]$, l'égalité est directe.

L'enveloppe convexe sur l'estimateur de CFG évalué sur les z_i^* permet alors de respecter les contraintes de borne supérieure et de convexité.

La correction de l'estimateur de CFG sur les données simulées au chapitre 3 est illustrée dans la Figure 4.1. L'estimateur initial est illustré à titre comparatif.

Estimateur bayésien

On présente maintenant notre estimateur bayésien de la fonction de Pickands qui est aussi basé sur une idée originale. Pour comprendre son fonctionnement, le lecteur doit être familier avec le contexte bayésien et avec les processus de Dirichlet. Les pages suivantes contiennent le nécessaire à la compréhension de l'estimateur proposé.

Contexte bayésien

La statistique est majoritairement présentée à l'aide de l'approche classique. C'est-à-dire que nous présentons un modèle générateur d'observations qui requiert des paramètres fixes. Ces paramètres sont généralement notés $\theta \in \mathbb{R}^p$. À l'aide d'observations recueillies et du modèle, on peut inférer sur la valeur de θ . L'estimation de θ est alors entièrement basée sur l'échantillon, noté \underline{x} .

La différence de l'approche bayésienne se trouve dans l'interprétation du paramètre θ . A contrario du cas classique, nous supposons que θ est la réalisation d'une variable aléatoire Θ . La variable Θ suit une distribution que doit définir le(la) statisticien(ne) ou l'expérimentateur(-trice) avant d'observer \underline{x} . Cette distribution est appelée *a priori* et est notée π . Elle doit capturer, du mieux possible, les croyances de l'expérimentateur(-trice) sur θ . L'approche bayésienne laisse place à un peu de subjectivité qui peut engendrer des controverses.

La méthode bayésienne va alors comme suit. Puisqu'on suppose θ inconnu et aléatoire, on définit

$\pi(\theta)$: *La distribution a priori de θ , qui capture les croyances sur le paramètre avant l'observation de \underline{x} .*

$f(\underline{x}|\theta)$: *Le modèle qui permet de générer les observations et qui dépend du paramètre θ . C'est à l'aide de ce modèle qu'on peut tirer de l'information sur θ à partir de \underline{x} .*

$\pi(\theta|\underline{x})$: *La distribution a posteriori de θ suivant l'observation de \underline{x} . Elle représente la mise-à-jour de nos croyances sur θ suivant l'expérience.*

La distribution a posteriori est obtenue par une application du théorème de Bayes.

$$\begin{aligned}\pi(\theta|\underline{x}) &= \frac{f(\theta, \underline{x})}{f(\underline{x})} = \frac{f(\underline{x}|\theta)\pi(\theta)}{f(\underline{x})} \\ &\propto f(\underline{x}|\theta)\pi(\theta)\end{aligned}$$

$f(\underline{x}) = \int_{\Theta} f(\underline{x}|\theta)\pi(\theta)d\theta$ est la densité marginale de la variable aléatoire \underline{X} . Elle peut être ignorée lors du calcul de la distribution a posteriori puisqu'elle ne dépend pas de θ et peut facilement être retrouvée.

L'inférence bayésienne de θ est basée sur la distribution a posteriori. Les estimateurs les plus courants sont :

- $\mathbb{E}(\theta|\underline{x})$: La moyenne a posteriori.
- $\underset{\theta \in \Theta}{\operatorname{argmax}} \pi(\theta|\underline{x})$: Le mode a posteriori.
- θ_0 tel que $\mathbb{P}(\theta < \theta_0|\underline{x}) = \mathbb{P}(\theta > \theta_0|\underline{x}) = 1/2$: La médiane a posteriori.

Dans certains cas, la distribution a posteriori obtenue n'est pas familière ou peut être impossible à manipuler. Si l'on souhaite faire numériquement l'estimation d'une fonction de θ , on peut recourir à des simulations numériques pour estimer les valeurs qui nous intéressent par la loi forte des grands nombres, par exemple. Quelques méthodes pour générer des observations de telles distributions sont les algorithmes de Monte-Carlo par chaînes de Markov (MCMC).

Nous venons de présenter l'approche bayésienne paramétrique, c'est-à-dire pour une famille de modèles connue. Dans le contexte de l'estimation de la fonction de Pickands, nous désirons faire de l'estimation non paramétrique. L'idée de l'estimation bayésienne non paramétrique est la suivante.

On pose une mesure de probabilité *a priori* π sur \mathcal{F} – l’ensemble des distributions possibles de \underline{X} . Sans nécessairement avoir de forme explicite pour π , savoir comment générer aléatoirement de π est suffisant. Cette distribution *a priori* permet de modéliser les croyances *a priori* de l’expérimentateur(-trice) sur la forme du modèle qu’on souhaite estimer. Depuis son introduction par Ferguson (1973), les processus de Dirichlet sont le choix de π le plus populaire.

Processus de Dirichlet

Seule l’intuition et les propriétés pertinentes des processus de Dirichlet sont présentées ici. Le lecteur s’intéressant à leur développement théorique peut se référer à Ghosh et Ramamoorthi. (2003).

On définit d’abord \mathcal{X} , l’ensemble des valeurs que peut prendre la variable aléatoire X et $M(\mathcal{X})$, l’ensemble des mesures de probabilités sur \mathcal{X} . On cherche ensuite à définir une mesure de probabilité π sur $M(\mathcal{X})$.

Voici une série d’exemples tirés de Ghosh et Ramamoorthi. (2003) qui visent à faciliter la compréhension du concept de mesure de probabilité sur $M(\mathcal{X})$.

Exemple : Soit $\mathcal{X} = \{1, 2\}$.

On a

$$M(\mathcal{X}) = \{(p_1, p_2) : p_1, p_2 \geq 0 \text{ et } p_1 + p_2 = 1\},$$

où $M(\mathcal{X})$ devient alors l’ensemble des valeurs que peut prendre p_1 , puisque $p_2 = 1 - p_1$. Une mesure de probabilité sur $M(\mathcal{X})$ est alors une distribution pour p_1 . Un exemple est la distribution bêta.

Exemple : On peut généraliser l’exemple précédent à $\mathcal{X} = \{1, 2, \dots, n\}$.

On a

$$M(\mathcal{X}) = \{(p_1, p_2, \dots, p_{n-1}) : p_i \geq 0 \text{ pour } 1 \leq i \leq n-1 \text{ et } \sum_{i=1}^{n-1} p_i \leq 1\}$$

Une mesure de probabilité π sur $M(\mathcal{X})$ est alors une distribution pour le vecteur $(p_1, p_2, \dots, p_{n-1})$. Un exemple est la généralisation multivariée des lois bêta : la loi de Dirichlet.

Un processus de Dirichlet, ce qui nous intéresse, est une généralisation des exemples précédents pour $\mathcal{X} = \mathbb{R}$. Les processus de Dirichlet sont utiles pour modéliser nos croyances sur la distribution de X , une variable aléatoire à valeur réelle. Ils requierent deux paramètres :

- (1) F_0 : La meilleure estimation subjective a priori de la fonction de répartition de X .
- (2) α : Le niveau d'importance qu'on attribue à F_0 . Plus α est grand, plus le processus de Dirichlet a priori est concentré autour de F_0 .

Comme pour le cas paramétrique, les observations de \underline{x} permettent de mettre à jour les croyances sur la distribution de X . Dans le contexte où l'on s'intéresse à l'estimation d'une fonction de répartition, on a le résultat suivant pour la distribution a posteriori :

Théorème 4.0.3. Soit X_1, X_2, \dots, X_n des variables aléatoires i.i.d. de fonction de répartition F et π un processus de Dirichlet pour la distribution a priori de paramètres F_0 et α – notée $Dir(\alpha, F_0)$, on a que la distribution a posteriori de F est

$$F|\underline{x} \sim Dir(\alpha + n, F_n^*)$$

où $F_n^* = \frac{n}{n+\alpha}F_n + \frac{\alpha}{n+\alpha}F_0$ et F_n est la fonction de répartition empirique des x_i . La moyenne a posteriori est F_n^* .

Dans le contexte des copules de valeurs extrêmes, nous utilisons ce résultat pour estimer la fonction de répartition des z_i^* qui permet d'obtenir une estimation de la fonction de Pickands.

Estimateur bayésien

L'estimateur que nous proposons est le suivant

$$A_n(t) = \exp\left\{ \int_0^t \frac{H_n^*(u) - u}{u(1-u)} du \right\}, \quad t \in [0,1]$$

où H_n^* est l'espérance a posteriori d'un processus de Dirichlet de paramètres :

$$F_0 = F, \quad \alpha = b,$$

où F est une fonction de répartition bêta($\alpha = a, \beta = b$) a priori pour les observations z_i et avec $1 < a \in \mathbb{R}$, $b \in \mathbb{R}^+$. Nous appliquons ensuite la correction proposée dans la première section de ce chapitre pour obtenir la convexité.

Pour le paramètre F_0 du processus de Dirichlet, on peut choisir n'importe quelle fonction de répartition H telle que

$$A(t) = \exp \left\{ \int_0^t \frac{H(u) - u}{u(1-u)} du \right\}, \quad t \in [0,1]$$

où A est une fonction de Pickands. Le choix de la fonction de répartition a priori d'une loi bêta(a,a) est arbitraire, mais permet de générer une infinité de fonctions de Pickands symétriques. Le choix a priori de H est équivalent à choisir une fonction de Pickands a priori.

La Figure 4.2 illustre des fonctions de Pickands définies respectivement par des fonctions de répartition de lois bêta(a,a) avec $a = \{1,1; 1,2; 1,4; 1,6; 2; 3\}$. Ces fonctions sont utiles lorsque l'expérimentateur(-trice) possède des croyances a priori sur le niveau de dépendance entre les données.

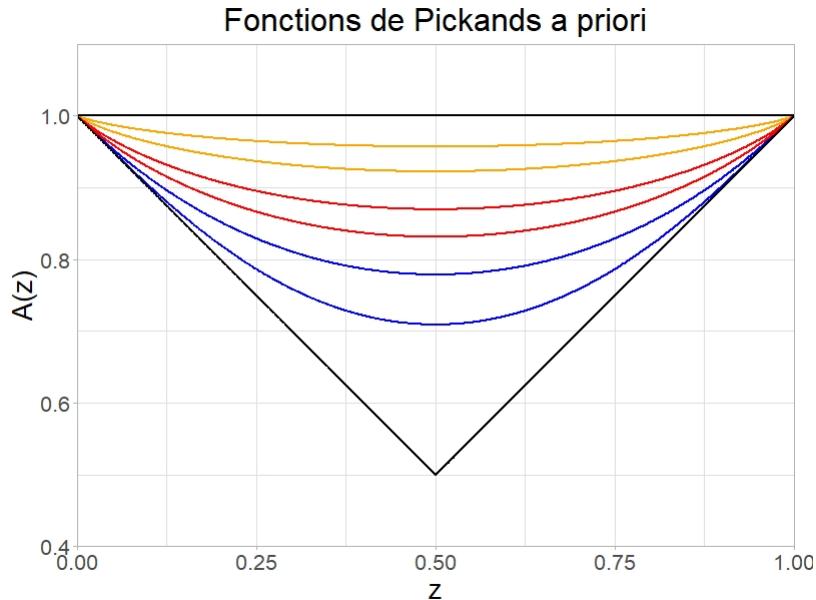


Figure 4.2. Fonctions de Pickands a priori.

La pertinence du contexte bayésien dans l'estimation d'une fonction de Pickands est due à la connaissance a priori que la fonction de répartition H de Z doit être convexe sur une première partie de l'intervalle $[0,1]$ et concave sur la seconde. Autrement dit, il existe un z_0

tel que

$$\begin{cases} H(z) - z \leq 0 & \text{si } z \leq z_0 \\ H(z) - z \geq 0 & \text{si } z > z_0. \end{cases}$$

En effet, on a

$$\frac{\partial}{\partial t} \log(A(t)) = \frac{H(t) - t}{t(1-t)}.$$

Par la convexité de A , il existe un t_0 tel que

$$\begin{cases} A'(t) \leq 0 & \text{si } 0 \leq t \leq t_0 \\ A'(t) \geq 0 & \text{si } t_0 \leq t \leq 1. \end{cases}$$

Ainsi,

$$\begin{cases} H(t) - t \leq 0 & \text{si } t \leq t_0 \\ H(t) - t \geq 0 & \text{si } t > t_0. \end{cases}$$

Nous détenons ainsi des croyances *a priori* sur H que nous pouvons fournir à l'aide du contexte bayésien.

De plus, le résultat suivant garantit que l'estimateur bayésien respecte la borne inférieure des fonctions de Pickands. En effet, elle montre que

$$A_n^0(t) = \exp \left\{ \int_0^t \frac{H_n^*(u) - u}{u(1-u)} du \right\} = \exp \left\{ - \int_t^1 \frac{H_n^*(u) - u}{u(1-u)} du \right\} = A_n^1(t)$$

lorsque l'estimateur est évalué sur les z_i^* . On le rappelle, A_n^0 et A_n^1 respectent la borne inférieure sur $[0, 1/2]$ et $[1/2, 1]$ respectivement. L'égalité de A_n^0 et A_n^1 implique le respect de la borne inférieure sur tout le domaine.

Proposition 4.0.4. Comme la fonction répartition d'une loi bêta($\alpha = a, \beta = a$) permet de créer une fonction de Pickands et la fonction de répartition échantillonale des z_i^* permet $A_n^0 = A_n^1$ pour l'estimateur de CFG, on a pour

$$H_n^*(z) = \alpha/nF(z) + (n-\alpha)/nF_n(z),$$

avec F la fonction de répartition de loi bêta(a, a) et F_n la fonction de répartition empirique des z_i :

$$\begin{aligned}
\exp \left\{ \int_0^t \frac{H_n^*(u) - u}{u(1-u)} du \right\} &= \exp \left\{ \int_0^t \frac{\alpha/nF(u) + (n-\alpha)/nF_n(u) - u}{u(1-u)} du \right\} \\
&= \exp \left\{ -\frac{\alpha}{n} \int_t^1 \frac{F(u) - u}{u(1-u)} du - \frac{(n-\alpha)}{n} \int_t^1 \frac{F_n(u) - u}{u(1-u)} du \right\} \\
&= \exp \left\{ - \int_t^1 \frac{H_n^*(u) - u}{u(1-u)} du \right\}
\end{aligned}$$

*Un premier avantage de l'estimateur bayésien est qu'il respecte les bornes de la fonction de Pickands. Les autres principaux avantages sont ceux que l'on trouve généralement dans l'approche bayésienne. C'est-à-dire l'usage de connaissances *a priori* sur le phénomène observé et la consistante de l'estimateur lorsqu'on détient un grand nombre d'observations.*

En effet,

$$\begin{aligned}
\int_0^t \frac{H_n^*(u) - u}{u(1-u)} du &= \int_0^t \frac{\alpha/nF(u) + (n-\alpha)/nF_n(u) - u}{u(1-u)} du \\
&= \frac{\alpha}{n} \int_0^t \frac{F(u) - u}{u(1-u)} du + \frac{(n-\alpha)}{n} \int_0^t \frac{F_n(u) - u}{u(1-u)} du \\
\Rightarrow \lim_{n \rightarrow \infty} \log(A_n(t)) &= \lim_{n \rightarrow \infty} \left(\frac{\alpha}{n} \int_0^t \frac{F(u) - u}{u(1-u)} du + \frac{(n-\alpha)}{n} \int_0^t \frac{F_n(u) - u}{u(1-u)} du \right) \\
&= \int_0^t \frac{H(u) - u}{u(1-u)} du = \log(A(t)),
\end{aligned}$$

Avec H , la vraie fonction de répartition de Z .

L'avant-dernière égalité est due au fait que l'estimateur de CFG est consistant. Cela rend l'estimateur proposé consistant.

*La Figure 4.3 illustre des estimateurs bayésiens de différentes distributions *a priori* évaluées sur les 30 données générées au chapitre 3. Les processus de Dirichlet *a priori* ont comme paramètres des fonctions de répartition bêta($\alpha = 1, \beta = 1$), bêta($\alpha = 2, \beta = 2$), bêta($\alpha = 3, \beta = 3$) et ont tous le paramètre $\alpha = 5$.*

Simulations

Pour faire l'évaluation de la correction et de l'estimateur proposés, leur erreur quadratique moyenne est comparée à celle des estimateurs de CFG, de Hall-Tajvidi et de la correction

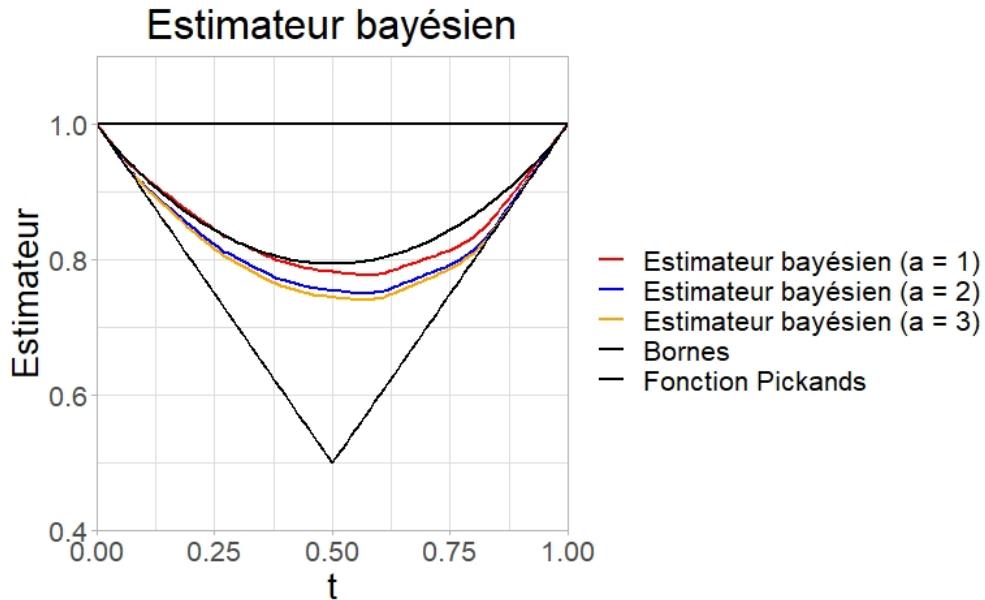


Figure 4.3. Estimation par l'estimateur bayésien.

FGS sur l'estimateur de CFG ($m = 20$), tous présentés au chapitre 3. La comparaison s'effectue à l'aide d'échantillons de tailles $n = 30$ et $n = 100$ générés à partir de 18 distributions de copules de valeurs extrêmes. Pour chacune des distributions et tailles d'échantillons, 1000 échantillons sont générés pour mesurer les erreurs d'estimation.

Le Tableau 4.1 présente les distributions utilisées dans la simulation.

Les Figure 4.4 et Figure 4.5 illustrent les 18 fonctions de Pickands des copules de valeurs extrêmes utilisées pour générer les données.

Le Tableau 4.2 présente les erreurs quadratiques moyennes des estimateurs et des corrections appliqués sur les échantillons générés, selon la taille de l'échantillon et la distribution. Pour utiliser la flexibilité de la correction proposée, la correction de FGS appliquée sur les données centrées s'y trouve également sous le nom de correction FGSc. Sur cet ensemble de données simulées, la centralisation donne un estimateur équivalent.

Comme le montre le Tableau 4.2, l'estimateur bayésien proposé a les erreurs quadratiques moyennes les plus faibles. La correction proposée appliquée sur l'estimateur de CFG réduit l'erreur moyenne de l'estimateur. L'avantage de la correction proposée comparativement à celle de FGS est de ne pas nécessiter d'hyperparamètre et a un temps de calcul moyen environ 15 fois plus rapide (0,006 seconde en moyenne par évaluation comparativement à 0,091).

Tableau 4.1. Description des distributions à partir desquelles les échantillons de la simulation ont étées générées.

<i>Distribution</i>			
<i>Logistique</i>	<i>r</i>	θ	ϕ
	1.5	1	1
	1.5	0.9	0.5
	2	1	1
	2	0.9	0.5
	2	0.75	0.95
	3	1	1
	3	0.9	0.5
	3.25	0.75	0.95
	10	0.75	0.95
<i>Mixte</i>		θ	ϕ
–	–	0.9	0
–	–	0.1	0
–	–	0.5	0
–	–	0.1	0.25
–	–	0.5	0.2
–	–	0.1	0.4
–	–	1	-0.25
–	–	0.5	-0.1
–	–	1.25	-0.3

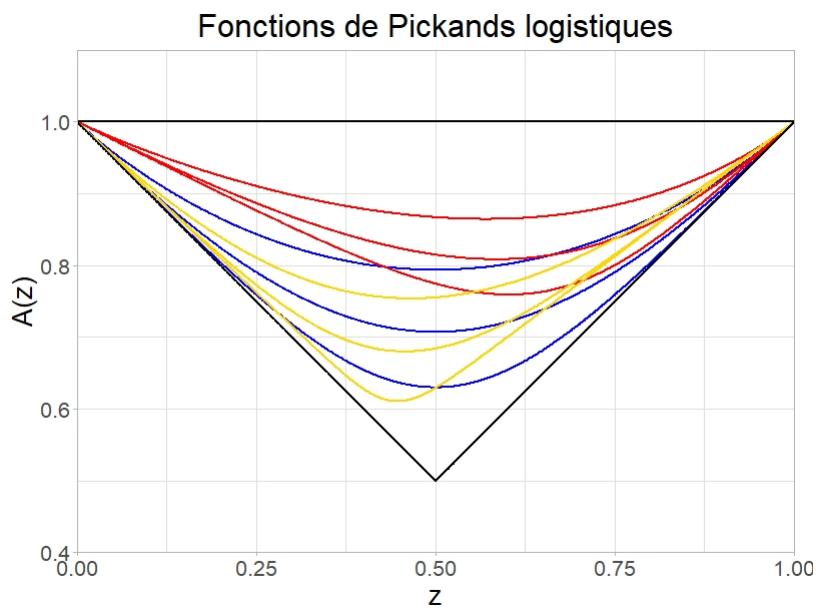


Figure 4.4. Copules logistiques utilisées.

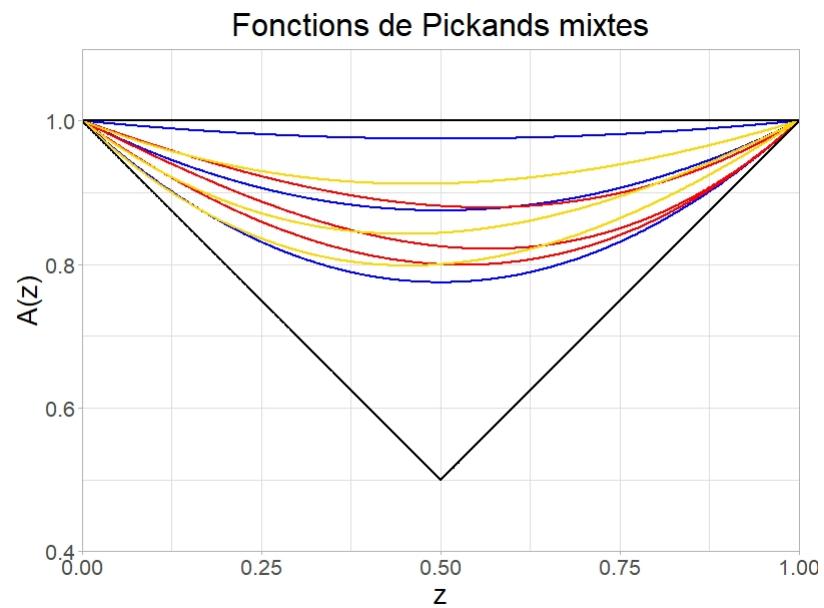


Figure 4.5. Copules mixtes utilisées.

Tableau 4.2. Description des erreurs quadratiques moyennes des estimateurs et des corrections appliqués sur les échantillons générés, selon la taille de l'échantillon.

Estimateurs	$n = 30$	$n = 100$
Logistique		
<i>CFG</i>	$2,28 \pm 4,37$	$0,72 \pm 1,12$
<i>Hall-Tajvidi</i>	$2,73 \pm 4,92$	$0,83 \pm 1,22$
<i>CFG avec correction FGS</i>	$1,55 \pm 2,23$	$0,60 \pm 0,87$
<i>CFG avec correction proposée</i>	$2,04 \pm 2,84$	$0,72 \pm 1,04$
<i>CFG avec correction FGSc</i>	$1,61 \pm 2,17$	$0,62 \pm 0,87$
<i>Estimateur bayésien ($a = 3$)</i>	$1,36 \pm 1,78$	$0,57 \pm 0,76$
<i>Estimateur bayésien ($a = 1.6$)</i>	$1,40 \pm 1,86$	$0,61 \pm 0,81$
<i>Estimateur bayésien ($a = 1.2$)</i>	$1,57 \pm 2,06$	$0,65 \pm 0,86$
Mixte		
<i>CFG</i>	$4,51 \pm 7,68$	$1,32 \pm 1,94$
<i>Hall-Tajvidi</i>	$5,70 \pm 9,28$	$1,74 \pm 2,39$
<i>CFG avec correction FGS</i>	$2,52 \pm 3,04$	$0,93 \pm 1,24$
<i>CFG avec correction proposée</i>	$3,02 \pm 3,32$	$1,04 \pm 1,34$
<i>CFG avec correction FGSc</i>	$2,61 \pm 3,07$	$0,96 \pm 1,24$
<i>Estimateur bayésien ($a = 3$)</i>	$2,62 \pm 3,18$	$0,94 \pm 1,24$
<i>Estimateur bayésien ($a = 1.6$)</i>	$2,22 \pm 2,75$	$0,90 \pm 1,20$
<i>Estimateur bayésien ($a = 1.2$)</i>	$2,07 \pm 2,58$	$0,89 \pm 1,19$
Total		
<i>CFG</i>	$3,39 \pm 6,02$	$1,02 \pm 1,53$
<i>Hall-Tajvidi</i>	$4,21 \pm 7,10$	$1,28 \pm 1,80$
<i>CFG avec correction FGS</i>	$2,03 \pm 2,63$	$0,78 \pm 1,05$
<i>CFG avec correction proposée</i>	$2,53 \pm 3,08$	$0,88 \pm 1,19$
<i>CFG avec correction FGSc</i>	$2,11 \pm 2,62$	$0,79 \pm 1,05$
<i>Estimateur bayésien ($a = 3$)</i>	$1,99 \pm 2,48$	$0,75 \pm 1,00$
<i>Estimateur bayésien ($a = 1.6$)</i>	$1,81 \pm 2,30$	$0,75 \pm 1,01$
<i>Estimateur bayésien ($a = 1.2$)</i>	$1,82 \pm 2,32$	$0,77 \pm 1,02$

Les mesures sont présentées sous forme de moyenne \pm écart-type.

Les valeurs de la table sont de l'ordre 10^{-3} .

FGSc : Correction de FGS sur les données centrées

Conclusion

Nous avons été confrontés au problème d'estimation de la fonction de Pickands associée à une copule de valeurs extrêmes. Plusieurs estimateurs existent tels que ceux de CFG, de Hall-Tajvidi et la correction de FGS.

*Les estimateurs de CFG et de Hall-Tajvidi sont les estimateurs classiques qui performent le mieux en pratique. Toutefois, ils ne respectent pas les contraintes imposées de la fonction de Pickands. Nous proposons alors deux alternatives pour faire l'estimation. La première est une correction d'estimateurs basée sur une nouvelle propriété de la variable aléatoire des observations qui vient améliorer la performance d'estimateurs en les forçant à respecter les contraintes exigées de la fonction de Pickands. La seconde est un estimateur bayésien basé sur les connaissances *a priori* que la fonction de répartition de Z, une transformation de la variable aléatoire des observations, doit prendre une certaine forme. Cette idée originale, combinée avec la correction proposée, donne un estimateur performant de la fonction de Pickands qui respectent ses contraintes.*

Corriger un estimateur de la fonction de Pickands pour qu'il respecte les contraintes imposées est une idée qui date de Pickands (1981), dans la publication du premier estimateur. Fils-Villetard et al. (2008) ont proposé une correction d'estimateurs performante. Par contre, elle nécessite l'ajustement d'un hyperparamètre et elle fonctionne à l'aide d'un algorithme complexe et couteux en temps de calculs (il requiert une optimisation à contraintes linéaires). La nouvelle propriété découverte nous permet de proposer une correction d'estimateurs simple et rapide qui rivalise celle de FGS sur les cas particuliers obtenus dans notre simulation. De plus, notre correction est flexible puisque l'ajustement basé sur la nouvelle propriété peut être combiné avec d'autres corrections de la convexité pour tenter de la rendre plus performante.

*Parmi tous les estimateurs considérés dans ce mémoire et sur les bases de simulations, l'estimateur bayésien proposé a obtenu, pour trois distributions *a priori* différentes et des paramètres non ajustés, les erreurs quadratiques moyennes les plus basses. De plus, notre estimateur ne considérait que des fonctions de répartition *a priori* de loi bêta. L'estimateur peut être ajusté au problème dans lequel il sera appliqué en utilisant d'autres distributions*

a priori et des paramètres reflétant mieux les connaissances a priori de l'expérimentateur(-trice) que ceux utilisés dans cet ouvrage.

Ce mémoire propose alors deux méthodes d'estimation de la fonction de Pickands表演antes, basées sur des idées originales et qui ont la flexibilité d'être ajustées et modifiées de manière à mieux répondre aux besoins pour lesquels elles sont requises.

Bibliographie

- A. AHMADABADI et B. Hudaverdi UCER : *Bivariate nonparametric estimation of the pickands dependence function using bernstein copula with kernel regression approach.* Computational Statistics, 32(4):1515–1532, 2017.
- P. CAPÉRAÀ, A.-L. FOUGÈRES et C. GENEST : *A nonparametric estimation procedure for bivariate extreme value copulas.* Biometrika, 84(3):567–577, 1997.
- Laurens de HAAN et Ana FERREIRA : Extreme Value Theory An Introduction. Springer Series in Operations Research and Financial Engineering, New York, NY, 2006.
- P. DEHEUVELS : *Probabilistic aspects of multivariate extremes.* Springer, pages 117–130, 1984.
- P. DEHEUVELS : *On the limiting behavior of the pickands estimator for bivariate extreme-value distributions.* Statistics & Probability Letters, 12(5):429–439, 1991.
- A.A. DING et Y. LI : *Copula correlation : An equitable dependence measure and extension of pearson’s correlation.* Proceedings of the National Academy of Sciences, 2015.
- T.S. FERGUSON : *A bayesian analysis of some nonparametric problems.* Annals of Statistics, 1:209–230, 1973.
- A. FILS-VILLETARD, A. GUILLOU et J. SEGERS : *Projection estimators of pickands dependence functions.* The Canadian Journal of Statistics, 36(3):369–382, 2008.
- J.K. GHOSH et R.V. RAMAMOORTHI. : Bayesian Nonparametrics. Springer Series in Statistics, New York, NY, 2003.
- G. GUDENDORF et J. SEGERS : *Extrem-value copulas.* Springer, pages 127–145, 2009.
- P. HALL et N. TAJVIDI : *Distribution and dependence-function estimation for bivariate extreme-value distributions.* Bernoulli, 6(5):835–844, 2000.
- V. KREINOVICH, H.T. NGUYEN, S. SRIBOONCHITTA et O. KOSHELEVA : *Modeling extremal events is not easy : Why the extreme value theorem cannot be as general as the central limit theorem. in uncertainty modeling.* Springer, Cham., pages 123–133, 2017.
- Roger B. NELSEN : An Introduction to Copulas. Springer Series in Statistics, New York, NY, 2006.
- J. PICKANDS : *Multivariate extreme value distribution.* Proceedings 43th, Session of International Statistical Institution, 1981.

- J. SEGERS : Non-parametric inference for bivariate extreme-value copulas. 2004.*
- A. SKLAR : Fonction de répartition à n dimensions et leurs marges. Publications de L'Institute de Statistique, pages 229–231, 1959.*
- J. TAWN : Bivariate extreme value theory : Models and estimation. Biometrika, 75(3):397–415, 1988.*

Annexe 1 : Preuve univariée d'équivalence du max-stable

Par définition, si G est max-stable, elle est une distribution de valeurs extrêmes.

Si G est une distribution de valeurs extrêmes, on a pour $a_n = n^\gamma$ et $b_n = (n^\gamma - 1)/\gamma$,

$$\begin{aligned} G_\gamma^n(a_n x + b_n) &= \exp^n(-(1 + \gamma[n^\gamma x + (n^\gamma - 1)/\gamma]^{-1/\gamma})) \\ &= \exp^n(-[n^\gamma(1 + \gamma x)]^{-1/\gamma}) \\ &= \exp(-(1 + \gamma x)^{-1/\gamma}) \\ &= G(x) \end{aligned}$$

Annexe 2 : Calculs de Capéraà, Fougères et Genest (1997)

Sans perte de généralité, on suppose ici que les variables aléatoires Z_i sont ordonnées. C'est-à-dire

$$0 < Z_1 < Z_2 < \dots < Z_n < 1.$$

Si $t \in [0, Z_1]$, on a :

$$\begin{aligned} \log A_n^0(t) &= \int_0^t \frac{H_n(z) - z}{z(1-z)} dz = \int_0^t \frac{-1}{(1-z)} dz = \log(1-z)|_{z=0}^{z=t} = \log(1-t) \\ \log A_n^1(t) &= - \int_t^1 \frac{H_n(z) - z}{z(1-z)} dz = - \left(\sum_{k=1}^{n-1} \int_{Z_k}^{Z_{k+1}} \frac{k/n - z}{z(1-z)} dz + \int_t^{Z_1} \frac{-1}{(1-z)} + \int_{Z_n}^1 \frac{1}{z} dz \right) \\ &= - \left(\sum_{k=1}^{n-1} \int_{Z_k}^{Z_{k+1}} \frac{k/n}{z(1-z)} dz - \sum_{k=1}^{n-1} \int_{Z_k}^{Z_{k+1}} \frac{1}{(1-z)} dz + \log(1-z)|_{z=t}^{z=Z_1} + \log(z)|_{z=Z_n}^{z=1} \right) \\ &= - \left(\sum_{k=1}^{n-1} k/n (\log z - \log(1-z))|_{z=k}^{z=k+1} + \sum_{k=1}^{n-1} \log(1-z)|_{Z_k}^{Z_{k+1}} + \log(1-Z_1) - \right. \\ &\quad \left. \log(1-t) - \log(Z_n) \right) \\ &= - \left(\sum_{k=1}^{n-1} k/n \left(\log \left(\frac{Z_{k+1}(1-Z_k)}{Z_k(1-Z_{k+1})} \right) \right) + \log(1-Z_n) - \log(1-t) - \log(Z_n) \right) \\ &= - \log \left(\prod_{k=1}^{n-1} \left(\frac{Z_{k+1}}{(1-Z_{k+1})} \frac{(1-Z_k)^{k/n}}{Z_k} \right) \right) + \log(Z_n/(1-Z_n)) + \log(1-t) \\ &= - \log \left(\prod_{k=1}^{n-1} \left(\frac{(1-Z_k)}{Z_k} \right)^{1/n} \right) - \log \left(\frac{1-Z_n}{Z_n}^{n-1/n} \right) + \log(Z_n/(1-Z_n)) + \log(1-t) \\ &= - \log \left(\prod_{k=1}^n \left(\frac{(1-Z_k)}{Z_k} \right)^{1/n} \right) + \log(1-t) \end{aligned}$$

Ainsi, on a

$$\begin{aligned}
A_n(t) &= A_n^0(t)^{p(t)} A_n^1(t)^{1-p(t)} \\
&= (1-t)^{p(t)} * \exp \left(-\log \left(\prod_{k=1}^n \left(\frac{(1-Z_k)}{Z_k} \right)^{1/n} \right) + \log(1-t))^{1-p(t)} \right) \\
&= (1-t)^{p(t)} * \exp \left(\log \left(\prod_{k=1}^n \left(\frac{Z_k}{(1-Z_k)} \right)^{1/n} \right) + \log(1-t))^{1-p(t)} \right) \\
&= (1-t)^{p(t)} * \exp \left(\log \left(\prod_{k=1}^n \left(\frac{Z_k}{(1-Z_k)} \right)^{1/n} \right) \right) (1-t))^{1-p(t)} \\
&= (1-t) * \left(\prod_{k=1}^n \left(\frac{Z_k}{(1-Z_k)} \right)^{1/n} \right) = (1-t)Q_n
\end{aligned}$$

Si $t \in [Z_n, 1]$, on a :

$$\begin{aligned}
\log A_n^0(t) &= \int_0^t \frac{H_n(z) - z}{z(1-z)} dz = \int_0^{Z_1} \frac{-dz}{1-z} + \sum_{i=1}^{n-1} \int_{Z_i}^{Z_{i+1}} \frac{i/n - z}{z(1-z)} dz + \int_{Z_n}^1 dz/z \\
&= \log(1-z)|_{z=0}^{z=Z_1} + \log(z)|_{Z_n}^{z=t} + \sum_{i=1}^{n-1} i/n \int_{Z_i}^{Z_{i+1}} \frac{1}{z(1-z)} dz - \\
&\quad \sum_{i=1}^{n-1} \int_{Z_i}^{Z_{i+1}} \frac{-1}{(1-z)} dz \\
&= \log(1-Z_1) + \log(t) - \log Z_n + \sum_{i=1}^{n-1} i/n \left[\log(z) - \right. \\
&\quad \left. \log(1-z) \right]_{z=Z_i}^{z=Z_{i+1}} - \sum_{i=1}^{n-1} \log(1-z)|_{z=Z_i}^{z=Z_{i+1}} \\
&= \log(t) - \log Z_n + \sum_{i=1}^{n-1} \left[\log(z/(1-z))^{i/n} \right]_{z=Z_i}^{z=Z_{i+1}} + \log(1-Z_n) \\
&= \log(t) + \log((1-Z_n)/Z_n) + \sum_{i=1}^{n-1} \left[\log \left(\frac{Z_{i+1}}{1-Z_i+1} \frac{1-Z_i}{Z_i} \right)^{i/n} \right] \\
&= \log(t) + \log((1-Z_n)/Z_n) + \left[\log \prod_{i=1}^{n-1} \left(\frac{Z_{i+1}}{1-Z_i+1} \frac{1-Z_i}{Z_i} \right)^{i/n} \right] \\
&= \log(t) + \log \prod_{i=1}^n \left(\frac{1-Z_i}{Z_i} \right)^{1/n} \\
\log A_n^1(t) &= - \int_t^1 \frac{H_n(z) - z}{z(1-z)} dz = - \int_t^1 dz/z = - \log(z)|_{z=t}^{z=1} = \log(t)
\end{aligned}$$

Ainsi, on a

$$\begin{aligned}
A_n(t) &= A_n^0(t)^{p(t)} A_n^1(t)^{1-p(t)} = t^{p(t)} \left[\prod_{i=1}^n \left(\frac{1-Z_i}{Z_i} \right)^{1/n} \right]^{p(t)} t^{1-p(t)} \\
&= t \left[\prod_{i=1}^n \left(\frac{Z_i}{1-Z_i} \right)^{1/n} \right]^{-p(t)}
\end{aligned}$$

Si $t \in [Z_j, Z_{j+1}]$, on a :

$$\begin{aligned}
\log A_n^0(t) &= \int_0^t \frac{H_n(z) - z}{z(1-z)} dz = \int_0^{Z_1} \frac{-dz}{1-z} + \sum_{i=1}^{j-1} \int_{Z_i}^{Z_{i+1}} \frac{i/n - z}{z(1-z)} dz + \int_{Z_j}^t \frac{j/n - z}{z(1-z)} dz \\
&= \log(1 - Z_1) + \sum_{i=1}^{j-1} i/n \left[\log(z) - \log(1-z) \right]_{z=Z_i}^{z=Z_{i+1}} + \sum_{i=1}^{j-1} \log(1-z)|_{z=Z_i}^{z=Z_{i+1}} \\
&\quad + j/n \int_{Z_j}^t \frac{dz}{z(1-z)} + \log(1-z)|_{z=Z_j}^{z=t} \\
&= \log(1 - Z_1) + \sum_{i=1}^{j-1} i/n \left[\log(z) - \log(1-z) \right]_{z=Z_i}^{z=Z_{i+1}} + \log(1 - Z_j) - \log(1 - Z_1) \\
&\quad + j/n \log(t) - j/n \log(1-t) - j/n \log Z_j + j/n \log(1 - Z_j) + \log(1-t) - \log(1 - Z_j) \\
&= \log(t^{j/n}) + \log \left(\frac{1 - Z_j^{j/n}}{Z_j} \right) + \log(1-t)^{(n-j)/n} + \sum_{i=1}^{j-1} \log \left[\frac{Z_{i+1}}{1 - Z_{i+1}} \frac{1 - Z_i^{i/n}}{Z_i} \right] \\
&= \log(t^{j/n}) + \log(1-t)^{(n-j)/n} + \sum_{i=1}^j \log \left[\frac{1 - Z_i^{1/n}}{Z_i} \right] \\
-\log A_n^1(t) &= \int_t^1 \frac{H_n(z) - z}{z(1-z)} dz = \int_t^{Z_{j+1}} \frac{j/n - z}{z(1-z)} dz + \sum_{i=j+1}^{n-1} \int_{Z_i}^{Z_{i+1}} \frac{i/n - z}{z(1-z)} dz + \int_{Z_n}^1 \frac{1}{z} dz \\
&= j/n(\log Z_{j+1} - \log(1 - Z_{j+1}) - \log(t) + \log(1-t)) + \log(1 - Z_{j+1}) - \log(1 - t) \\
&\quad + \sum_{i=j+1}^{n-1} i/n(\log \left(\frac{Z_{j+1}}{1 - Z_{j+1}} \frac{1 - Z_i}{Z_i} \right)) + \sum_{i=j+1}^{n-1} \log(1-z)|_{z=Z_i}^{z=Z_{i+1}} - \log(Z_n) \\
&= \log \left(\frac{1 - t^{j/n}}{t} \right) - \log(1-t) + \sum_{i=j+1}^{n-1} \log \left(\frac{1 - Z_i^{1/n}}{Z_i} \right) + \log \left(\frac{Z_n}{1 - Z_n} \right)^{(n-1)/n} \\
&\quad + \log \left(\frac{1 - Z_n}{Z_n} \right) \\
&= \log \left(\frac{1 - t^{j/n}}{t} \right) - \log(1-t) + \sum_{i=j+1}^n \log \left(\frac{1 - Z_i^{1/n}}{Z_i} \right) \\
&= \log((1-t)^{(j-n)/n}) - \log(t^{j/n}) + \sum_{i=j+1}^n \log \left(\frac{1 - Z_i^{1/n}}{Z_i} \right)
\end{aligned}$$

Ainsi,

$$\log A_n^1(t) = \log((1-t)^{1-j/n}) + \log(t^{j/n}) + \sum_{i=j+1}^n \log \left(\frac{Z_i}{1-Z_i} \right)^{1/n}$$

On obtient alors

$$\begin{aligned} A_n(t) &= \left[t^{j/n} (1-t)^{1-j/n} \prod_{i=1}^j \frac{1-Z_i}{Z_i} \right]^{p(t)} \left[t^{j/n} (1-t)^{1-j/n} \prod_{i=J+1}^n \frac{Z_i}{1-Z_i} \right]^{1-p(t)} \\ &= t^{j/n} (1-t)^{1-j/n} Q_j^{-p(t)} (Q_n/Q_j)^{1-p(t)} \\ &= t^{j/n} (1-t)^{1-j/n} Q_j^{-1} Q_n^{1-p(t)} \end{aligned}$$

Annexe 3 : Algorithme de l'estimateur bayésien

Algorithme de l'estimateur bayésien :

- (1) Choisir une fonction de Pickands ou une fonction de répartition de Z a priori basée sur nos croyances.
- (2) Choisir le paramètre α du processus de Dirichlet (le niveau d'importance qu'on attribue à notre fonction de répartition a priori).
- (3) Centré les données $\log\left(\frac{z_i}{1-z_i}\right)$.
- (4) Calculer numériquement l'intégrale

$$A_n(t) = \exp\left\{\int_0^t \frac{H_n^*(u) - u}{u(1-u)} du\right\}, \quad t \in [0,1]$$

où H_n^* est l'espérance a posteriori d'un processus de Dirichlet basée sur les données centrées.

- (5) Appliquer une correction pour la convexité.

Implémentation de l'algorithme en R :

```

centralize_Z <- function(z) {
  # To the Y space
  y <- log(z / (1-z))

  # centralization
  y_bar <- mean(y)
  y_ <- y - y_bar

  # To the Z space
  return(exp(y_) / (1 + exp(y_)))
}

cfg_bayes <- function(t, alpha = 5, data, centralize = F, a = 3, b = 3)

#Set up data
n <- nrow(data)
z <- log(data[, 1]) / log(data[, 1] * data[, 2])
z <- sort(z)
z <- centralize_Z(z)

# Empirical function
Fn <- stats::ecdf(z)
z_ <- c(0, z, 1)

# Term to be integrate for the estimator
integrande <- function(x) {

  (n / (n+alpha) * Fn(x) + alpha / (n + alpha) * pbeta(x, a, b) - x) /
  (x*(1-x))
}

}

```

```

# Integration to obtain the log estimator (in the form of CFG)
cfg_estimates <- sapply(2:(length(z_)-1), function(i) {
  integrate(integrand, lower = z_[i-1], upper = z_[i]) $value
})
Az <- pmin(exp(c(cumsum(c(0, cfg_estimates)), 0)), 1)

# Convexity obtention
convhull_chull <- sort(chull(cbind(z_, Az)))
t_ <- unique(sort(c(t, z_[convhull_chull])))

estimates <- rep(NA, length(t_))
estimates[which(t_ %in% z_[convhull_chull])] <- Az[convhull_chull]

data_ <- cbind(t = t_, A = estimates)
cfg_estimates <- zoo::na.approx(data_[, "A"], x = t_)

cfg_estimates <- cfg_estimates[which(t_ %in% t)]
ordered_est <- rep(0, length(cfg_estimates))
ordered_est[order(t)] <- cfg_estimates

ordered_est
}

```