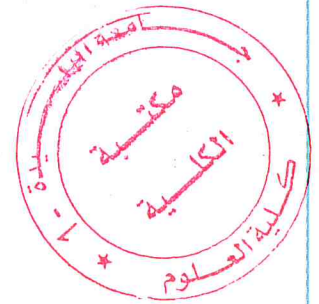


RÉPUBLIQUE ALGÉRIENNE DÉMOCRATIQUE ET POPULAIRE
MINISTÈRE DE L'ENSEIGNEMENT SUPÉRIEURE ET DE LA RECHERCHE SCIENTIFIQUE
 Université Saad Dahlab Blida
 Faculté des Sciences
 Département de Mathématiques



**MEMOIRE POUR L'OBTENTION DU DIPLOME DE MASTER
 EN MATHÉMATIQUES**

Thème

**Estimation semi paramétrique des moments
 de queue conditionnelle des distributions**

Présenté par :

Benharkat
 Fatiha

Ouadah
 Afifa

Encadré par :

Dr. A.Rassoul

Soutenue le /10/2018 devant le jury composé de :

Nom et Prénom	Grade	Etablissement	En Qualité de
Tami Omar	M. A.A	U.S.D.Blida	Président
Laidi Mohamed	M. A.A	E.N.S.T.Alger	Examineur

0.1 Dédicace

Je dédie ce travail, comme preuve de respect, de gratitude, et de reconnaissance :

A mes chers parents, Je vous remercie pour tout le soutien et l'amour que vous me portez depuis mon enfance et j'espère que votre bénédiction m'accompagne toujours. Que ce modeste travail soit l'exaucement de vos vœux tant formulés, le fruit de vos innombrables sacrifices, bien que je ne m'en acquitterai jamais.

à mes chers frères, Ahmed, Sedik, Yakoub, et ma chère sœur, Fadhila.

a ma grande mère chérie, Qui m'a accompagné par ses prières, sa douceur, puisse Dieu lui prêter longue vie et beaucoup de santé et de bonheur dans les deux vies. A toute ma famille ainsi qu'à mes amis.

Merci pour m'avoir toujours supporté dans mes décisions. Merci pour tout votre amour et votre confiance.

B.Fatiha

0.2 Dédicace

A mes chers parents, pour tous leurs sacrifices, leur amour, leur tendresse, leur soutien et leurs prières tout au long de mes études,

A mes chères sœurs, Hayet, Ahlem, Douaa pour leurs encouragements permanents, et leur soutien moral,

A toute ma famille pour leur soutien tout au long de mon parcours universitaire,

Que ce travail soit l'accomplissement de vos vœux tant allégués, et le fruit de votre soutien infailible,

Merci d'être toujours là pour moi.

Ou.Afifa

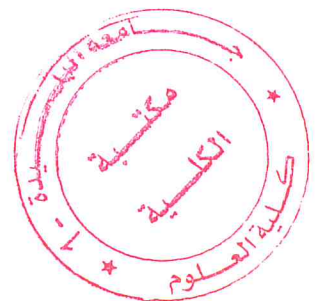
0.3 Remerciement

Nous tenons tout d'abord à remercier Dieu le tout puissant et miséricordieux, qui nous a donné la force et la patience d'accomplir ce Modeste travail.

En second lieu, nous tenons à remercier notre encadreur Mr : (A.RASSOUL), son précieux conseil et son aide durant toute la période du travail.

Nos vifs remerciements vont également aux membres du jury pour l'intérêt qu'ils ont porté à notre recherche en acceptant d'examiner notre travail Et de l'enrichir par leurs propositions.

Enfin, nous tenons également à remercier toutes les personnes qui ont participé de près ou de loin à la réalisation de ce travail.



ملخص

المعروف من هذا العمل هو اقتراح تقدير العزم الشرطي لتوزيعات ذات الذيل الثقيل، فنحن نقدم أولاً مقاييس المخاطر المختلفة الموجودة ويمكن استخدام العزم الشرطي لتوزيعات ذات الذيل الثقيل من أجل تحديد أي قياس المخاطر على أساس اللحظات المشروطة. ثم نقوم بعمل التقدير المعلمي ونجد أن النتائج التجريبية لا تقرب بشكل كافٍ النتيجة النظرية لتوزيعات ذات الذيل الثقيل. يقترح إجراء التقدير شبه المعنوي من خلال النهج الكمي Weismann الذي يعطي نتائج أقرب من القيمة النظرية. تنتهي هذه المذكرة مع أمثلة من المحاكاة مع R.

0.4 Résumé

L'objet de ce travail est de proposer de faire l'estimation de Moment conditionnel des queues (CTM), nous présentons dans un premier temps les différentes mesures de risque qui existe, et le CTM peut être utilisé dans le but de définir n'importe quelle mesure de risque basée sur les moments conditionnels. Ensuite nous avons fait l'estimation paramétrique et on trouve que, les résultats empiriques n'approchent pas suffisamment le résultat théorique pour des distributions à queues lourdes. On propose de faire l'estimation semi paramétrique par l'approche quantile de Weisman qui donne des résultats plus proches que la valeur théorique. On finit ce mémoire par des exemples de simulation avec R.

0.5 Abstract

The purpose of this work is to propose to estimate Conditional Moment Queues (CTM), we first present the different risk measures that exist, and the CTM can be used to define any risk measure based on the conditional moments. Then we make the parametric estimation and we find that the empirical results do not sufficiently approximate the theoretical result for heavy-tailed distributions. It is proposed to make the semi-parametric estimation by the Weisman quantile approach which gives results closer than the theoretical value. we finish this memory with examples of simulation with R.

Table des matières

0.1	Dédicace	5
0.2	Dédicace	5
0.3	Remerciement	6
0.4	Résumé	7
0.5	Abstract	7
0.6	Introduction Générale	8
1	Rappels sur la théorie des valeurs extrêmes	10
1.1	Introduction	10
1.2	Approche des maxima par blocs : lois GEV	12
1.2.1	Valeurs extrêmes et lois α -stables	12
1.3	Théorème fondamental des valeurs extrêmes	15
1.4	Domaines d'attraction	19
1.5	Estimation des modèles GEV	23
1.5.1	Estimation pour la loi de Gumbel	23
1.5.2	Méthode du maximum de vraisemblance	24
1.5.3	Méthode des moments	24
1.5.4	Estimation pour les lois de Fréchet et de Weibull	25
1.5.5	Méthode du maximum de vraisemblance	25
1.5.6	Méthode des moments de probabilités pondérés PWM :	25
1.5.7	Méthodes non paramétriques pour l'estimation de ξ	26
1.5.8	Estimateur de Pickands	26
1.5.9	Estimateur de Hill	27
1.5.10	Estimateur des moments de Dekkers-Einmahl-De Haan	28
1.6	Approche du loi GPD	28
1.6.1	Modélisation des excès	28
1.6.2	Sélection du seuil	31
1.7	Estimation du modèle GPD	34
1.7.1	Méthode des moments	34
1.7.2	Méthode du maximum de vraisemblance(EMV)	34
1.7.3	Méthode des moments pondérés(EMP)	35

1.8	Estimation de la queue de la distribution	36
1.9	Estimation des quantiles	36
1.9.1	Quantile d'ordre p	36
1.9.2	Fonction des quantiles	37
1.9.3	Estimation empirique	37
1.9.4	Intervalles de confiance des quantiles empirique	38
1.9.5	Approche POT	39
1.9.6	Estimation de Weismann	39
1.9.7	Approche basée sur un estimateur d'indices positifs ($\gamma > 0$)	40
1.9.8	Approche basée sur un estimateur d'indices quelconques	40
2	Mesures des risques	41
2.1	Introduction	41
2.2	Risque	41
2.3	Mesure de risque	42
2.4	Définition Formelle d'une Mesure de Risque	42
2.4.1	VaR(Value-at-Risque)	43
2.5	Mesure de risque Cohérente	46
2.5.1	La Tail Value-at-Risk	49
2.5.2	Conditionnelle VaR	50
2.5.3	Mesure de Wang	53
2.6	Mesures de risque par distorsion	58
2.7	Moment conditionnel des queues (CTM)	60
3	Estimation de CTM	62
3.1	Estimation de CTM	62
3.1.1	Estimation empirique de CTM	63
3.2	Simulations	69
3.3	Estimateur semi parametrique	70
3.3.1	Théorie des limites pour les moments extrêmes de la queue conditionnelle	70
3.3.2	Preuve de l'expression (3.46).	74
3.3.3	Preuve de l'expression (3.47).	78
3.4	Simulations	81
3.5	Conclusion générale	82

TABLE DES FIGURES

1.1	Explication graphique de quantiles extrêmes et queue de distribution. . . .	17
1.2	Classification des lois selon les queues, (El Adlouni et al.(2007))	19
1.3	Méthode des maxima (minima) par bloc.	21
1.4	Représentation de la fonction de répartition : Gumbel($\xi = 0$), Fréchet ($\xi = 1$) et Weibull ($\xi = 1$)	22
1.5	Lois des valeurs extrêmes	22
1.6	Loi du domaine d'attraction $\mathcal{D}(\Lambda), \mathcal{D}(\oplus_\xi)$ et $\mathcal{D}(\ominus_\xi)$	26
1.7	Représentation des excès Y issue des dépassements X au-delà d'un seuil u , (Raggad,2009).	33
2.1	Graphique d'une fonction de survie modélisant la distribution des pertes. On a représenté une Value-at-Risk d'un certain niveau α	48
2.2	Illustration graphique de deux fonctions de survies distinctes ayant la même $VaR(\alpha)$	49

Liste des tableaux

1.1	Lois du domaine d'attraction de Gumbel.	20
1.2	Lois du domaine d'attraction de Weibull	21
1.3	Lois du domaine d'attraction de Fréchet	21
2.1	Principales mesures de risque par distorsion (où p est une constante comprise entre 0 et 1)	56
2.2	Liens entre le CTM et les mesures de risque classiques.	60
3.1	Simulation du CTM_1 pour distribution de Pareto avec $\gamma = 3/4$ et $t = 0.95$ sachant que la valeur théorique est 29.822979	70
3.2	Simulation du CTM_1 pour distribution de Pareto avec $\gamma = 2/3$ et $t = 0.95$ sachant que la valeur théorique est 19.31929	70
3.3	Simulation du CTM_1 pour distribution de Pareto avec $\gamma = 3/4$ et $t = 0.95$ sachant que la valeur théorique est 29.822979	81
3.4	Simulation du CTM_1 pour distribution de Pareto avec $\gamma = 2/3$ et $t = 0.95$ sachant que la valeur théorique est 19.31929	81

0.6 Introduction Générale

Ce mémoire présente l'une des applications de la théorie des valeurs extrêmes sur un problème concret issu du monde de la finance ou de l'assurance, nous avons choisi de développer des outils capables de mesurer des caractéristiques de phénomènes réels. Nous avons présenté l'essentiel de la théorie qui nous permet de prendre une réflexion théorique pour comprendre et analyser des interactions complexes observées dans les séries financières.

La théorie des valeurs extrêmes trouve des applications dans nombreux domaines. un exemple parmi bien d'autres est celui des entreprises qui ont recours à celle ci pour estimer leur production optimale. On utilise aussi en hydrologie pour calculer la probabilité que la hauteur d'eau d'un fleuve dépasse un certain seuil. Cette théorie est basée sur la connaissance de la loi asymptotique du maximum d'un échantillon. L'expression de cette loi asymptotique que l'on nomme loi des valeurs extrêmes, ainsi que les conditions d'existence de celle-ci, sont données dans de nombreux ouvrages (voir Embrecht, Reiss Thomas, de Haan,...). La loi des valeurs extrêmes, lorsque'elle existe, est indexées par un paramètre appeler indice des valeurs extrêmes, ainsi qu'éventuellement par des paramètres d'échelle et de position. La connaissance de l'indice des valeurs extrêmes est importante car elle controle la lourdeur de la queue de distribution. Ainsi de nombreux estimateurs de l'indice de valeurs extrêmes ont été proposées dans la littérature, les deux estimateurs les plus connues étant les estimateurs de Hill (1975) et Pickand (1978).

La principale difficulté dans l'estimation des quantiles extrêmes réside dans le fait que le quantile que l'on cherche à estimer est situé au-delà de l'observation maximale. On ne peut donc pas, comme pour l'estimation de quantiles "classique", inverser tout simplement la fonction de répartition empirique. Il faut donc, à l'aide des plus grandes observations, estimer la fonction de répartition au-delà de l'observation maximale.

Au long de ce mémoire, je vais résumer mon travail en 3 chapitres principaux.

- Le premier chapitre présente un état de l'art sur la théorie des valeurs extrêmes, nécessaire pour estimer le quantile extrêmes qui est la problématique principale de ce mémoire. Tout d'abord, on présentera les principaux résultats et définitions de la théorie des valeurs extrêmes utiles dans nos travaux. Après avoir introduit le comportement du maximum d'un échantillon, on présentera les deux principaux outils servant à modéliser le comportement des valeurs extrêmes : la loi des valeurs extrêmes et la loi des excès. On s'intéressera ensuite à la caractérisation des domaines d'attraction. Puis on rappellera les différentes méthodes d'estimation des paramètres de la GEV et la loi des excès. Enfin on

- présentera une estimation de la queue de la distribution, estimation empirique, l'approche pot et estimation de quantile du Weisman.
- Nous avons consacré le deuxième chapitre pour exposer la théorie de mesure de risque telles que, la variance, Value at Risk (VaR), Tail Value-at-Risk (TVaR) cette dernière est connue par Conditional Tail Expectation (CTE) lorsque les risques sont continus, l'expected shortfall (ES) et les mesures basées sur la distorsion, le moment conditionnel des queues (CTM). Le but de ce chapitre est de présenté au lecteur les différentes mesures de risque qui existent.
 - Au dernier chapitre, on a traité l'estimation empirique et semi paramétrique du moment conditionnel des queues noté par CTM, le comportement asymptotique des ces estimateurs et on a fait quelques résultats de simulations pour avoir le comportement des estimateurs proposés avec le logiciel R.

Chapitre 1

Rappels sur la théorie des valeurs extrêmes

1.1 Introduction

La théorie des valeurs extrêmes communément appelée « **Extreme Value Theory** » (EVT) en anglais, est une vaste théorie dont le but est d'étudier les événements rares c'est-à-dire les événements dont la probabilité d'apparition est faible. Autrement dit elle essaie d'amener des éléments de réponses aux intempéries, aux inondations, aux catastrophes naturelles, aux problèmes financiers, etc. en prédisant leurs occurrences dans les années à venir. En d'autres termes on veut estimer des petites probabilités ou des quantités dont la probabilité d'observation est très faible c'est-à-dire proche de zéro. Ces quantités sont appelées quantiles extrêmes car l'ordre de ces quantiles tend vers zéro lorsque la taille de l'échantillon, n , tend vers l'infini.

On considère, pour cela, n variables aléatoires réelles $(Y_i)_{1 \leq i \leq n}$ indépendantes et identiquement distribuées de fonction de répartition F non nécessairement continue. Et soit $Y_{(1)} \leq Y_{(2)} \leq \dots \leq Y_{(n)}$ l'échantillon ordonné associé. Nous voulons déterminer, à partir de l'échantillon, le quantile extrême d'ordre $\alpha_n \rightarrow 0$ quand $n \rightarrow \infty$ défini par :

$$q(\alpha_n) = \bar{F}^{\leftarrow}(\alpha_n) = \inf\{y : \bar{F}(y) \leq \alpha_n\}, \quad (1.1)$$

avec la convention $\inf\{\emptyset\} = \infty$, où \bar{F}^{\leftarrow} est l'inverse généralisée de $\bar{F} = 1 - F$. Si la taille de l'échantillon tend vers l'infini, nous avons

$$\begin{aligned}
P(Y_{(n)} < q(\alpha_n)) &= P(Y_i < q(\alpha_n), \forall i = 1, \dots, n) \\
&= \mathbb{P} \left(\bigcap_{i=1}^n \{Y_i < q(\alpha_n)\} \right) \\
&= \prod_{i=1}^n P(Y_i < q(\alpha_n)) \\
&= F^n(q(\alpha_n)) \\
&= (1 - \alpha_n)^n \\
&= \exp(n \log(1 - \alpha_n)) \\
&= \exp(-n\alpha_n(1 + o(1))) \quad \text{quand } \alpha_n \rightarrow 0.
\end{aligned}$$

Cette probabilité dépend donc du comportement asymptotique de $n\alpha_n$. Estimer les quantiles extrêmes revient à étudier la limite de $n\alpha_n$ lorsque n tend vers l'infini. Ainsi, on distingue trois situations selon la vitesse de convergence de α_n vers 0 :

– **Première situation** : Si $n\alpha_n \rightarrow \infty$ alors $\mathbb{P}(Y_{(n)} < q(\alpha_n)) \rightarrow 0$.

Dans cette situation, le quantile à estimer se trouve avec une grande probabilité dans l'échantillon disponible. Elle correspond au cas où α_n converge lentement vers 0 autrement dit le quantile à estimer converge lentement vers l'infini lorsque n tend vers l'infini. Dans ce cas estimer le quantile revient à interpoler à l'intérieur de l'échantillon et donc on obtient $Y_{(n - \lfloor n\alpha_n \rfloor + 1)}$ comme estimateur.

– **Deuxième situation** : Si $n\alpha_n \rightarrow c \in [1, \infty[$ alors $\mathbb{P}(Y_{(n)} < q(\alpha_n)) \rightarrow e^{-c}$. Dans ce deuxième cas, nous avons une faible probabilité que le maximum de l'échantillon soit supérieur au quantile. Cela signifie que l'estimation du quantile extrême repose sur les grandes observations situées au voisinage de la frontière de l'échantillon et toujours dans l'ensemble des données. De ce fait, l'estimateur proposé dans la première situation est celui envisagé.

– **Troisième situation** : Si $n\alpha_n \rightarrow c \in [0, 1[$ alors $\mathbb{P}(Y_{(n)} < q(\alpha_n)) \rightarrow e^{-c}$. Dans ce cas, le quantile à estimer est supérieur au maximum des observations disponibles. Proposer un estimateur du quantile est impossible avec **la fonction de répartition empirique** définie :

$$F_n(y) = \frac{1}{n} \sum_{i=1}^n \mathbf{1}_{\{Y_i \leq y\}} = \begin{cases} 0 & \text{si } y < Y_{(1)}, \\ \frac{i-1}{n} & \text{si } Y_{(i-1)} \leq y < Y_{(i)}, 2 \leq i \leq n \\ 1 & \text{si } y \geq Y_{(n)}. \end{cases} \quad (1.2)$$

Cependant pour estimer le quantile extrême il faut extrapoler au delà du maximum des observations dont on dispose car $F_n(y) = 1$ si $y \geq Y_{(n)}$. Ce phénomène peut être expliqué en utilisant le schéma suivant :

de ces résultats provient du fait qu'il n'est pas nécessaire de connaître la loi F du processus Y que nous souhaitons prédire. Cependant, l'**EVT** est analogue au Théorème Central Limite (**TCL**) mais pour les phénomènes extrêmes. Ainsi, là où le **TCL** montre que la moyenne empirique de la variable Y converge vers une loi normale (indépendamment de la loi de la variable d'intérêt et lorsque les moments d'ordre 1 et 2 existent); l'**EVT** établit des résultats analogues mais pour les valeurs extrêmes de Y .

Définition 1.1 On dit qu'une variable aléatoire X suit une loi **stable** si pour tout $n \geq 2$, il existe un réel strictement positif C_n et réel D_n tel que :

$$X_1 + X_2 + \dots + X_n \stackrel{\text{loi}}{=} C_n X + D_n \quad (1.4)$$

Où X_1, X_2, \dots, X_n sont n copies indépendantes de X .

Le réel strictement positif C_n est nécessairement de la forme $n^{1/\alpha}$, pour un certain $\alpha \in]0, 2]$, d'où l'appellation loi α -stable (Nikias et Shao (1995)).

Définition 1.2 On dit qu'une variable aléatoire X suit une loi α -**stable** si et seulement si sa fonction caractéristique est de la forme :

$$\phi(t) = \exp \{iat - \gamma |t|^\alpha [1 + i\beta \text{sign}(t) \omega(t, \alpha)]\} \quad (1.5)$$

Avec

$$\omega(t, \alpha) \begin{cases} -\tan \frac{\alpha\pi}{2} & \text{si } \alpha \neq 1 \\ \frac{2}{\pi} \log |t| & \text{si } \alpha = 1 \end{cases} \quad (1.6)$$

Où $\alpha \in]0, 2]$, $\beta \in]-1, 1[$, $\gamma > 0$ et $a \in \mathbb{R}$. Les paramètres α , a , γ et β représentent respectivement l'exposant caractéristique de la loi stable, le paramètre de position, le paramètre d'échelle et le paramètre d'asymétrie.

Le théorème suivant montre que les lois stables peuvent être utilisées particulièrement, pour approximer la loi d'une somme de variables aléatoire i.i.d.

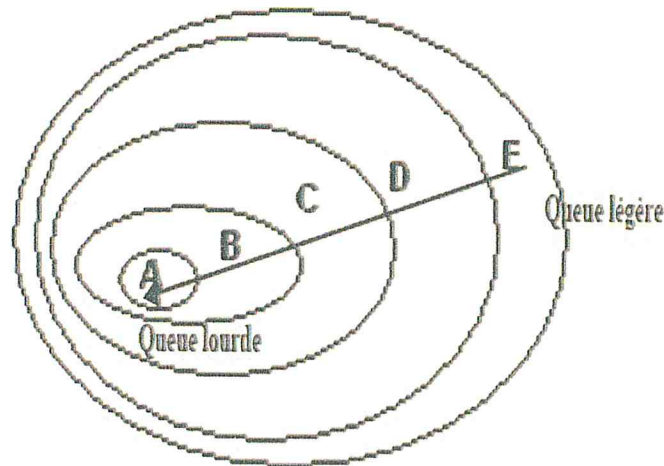
Théorème 1.1 Soit X_1, \dots, X_n des variables aléatoires indépendantes, de même loi. Une variable aléatoire X est limite en loi de la suite :

$$\frac{X_1 + \dots + X_n}{a_n} - b_n$$

pour une suite de réels strictement positifs (a_n) et une suite de réels (b_n) , si et seulement si, X suit une loi stable (pour la démonstration de ce théorème, (voir Nolan 1996)).

les lois α -stable présentent un grand intérêt dans la modélisation de nombreux problèmes physiques. La caractéristique de ces lois est leur index de stabilité α qui indique la vitesse de décroissance des queues. Ainsi, les lois peuvent être regroupées selon le comportement de leurs queues. La classification suivante a été proposée récemment par El Adlouni et al. (2007) :

- (E) : Les lois exponentielles dont les moments n'existent pas.
- (D) : Les lois subexponentielles .
- (C) : Les lois à variation régulières.
- (B) : Les lois avec un comportement de Pareto .
- (A) : Les lois α -stables avec $\alpha < 2$.



Classification des lois selon les queues, (El Adlouni et al.(2007))

La classe (E) contient toutes les lois tel que $\mathbb{E} [e^X] = \infty$. La probabilité au dépassement, $\bar{F} = \Pr(X \geq x) = 1 - F(x)$, pour les extrêmes de cette classe, décroît moins rapidement que celle des lois ayant une queue plus lourde que la loi normale.

La classe (D) quant à elle, contient les lois tel que \bar{F} décroît plus lentement que n'importe quelle loi exponentielle.

Pour la classe (C) dite classe des lois à variations régulières, la probabilité au dépassement des extrêmes décroît suivant une fonction puissance (appelée aussi décroissance géométrique).

La classe (B) est celle des lois de Pareto. Enfin, la classe (A) est celle qui regroupe des lois à différentes asymétries, avec des queues très lourdes.

1.3 Théorème fondamental des valeurs extrêmes

Le résultat de base de la *TVE* consiste à décrire la loi asymptotique du maximum de n variables indépendantes et identiquement distribuées (i.i.d). La forme de cette loi (notée *GEV*) dépend d'un seul paramètre qui permet de spécifier le comportement de la queue de la loi considérée. Ce paramètre est l'indice des valeurs extrêmes (noté ξ). En d'autres termes, le résultat de base de la *TVE* assure que la loi du maximum de n variables (i.i.d) est toujours dans le domaine d'attraction d'une loi GEV_ξ .

Définition 1.3 Soit G une fonction de répartition définie sur \mathbb{R} Une fonction de répartition F sera dite appartenir au domaine d'attraction $\mathcal{D}(G)$ de G s'il existe deux suites (a_n) ($a_n > 0$) et (b_n) telles que :

$$\lim_{n \rightarrow \infty} F^n(a_n x + b_n) = G(x) \quad (1.7)$$

pour tout point de continuité x , de G .

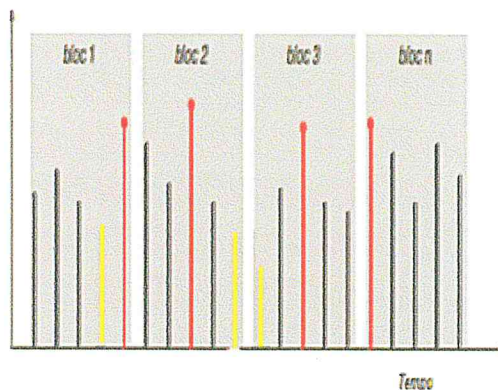
Dans le cas où F est connue, G peut être identifiée à partir de la loi des statistiques d'ordre. En effet, si nous supposons que les données observées sont des réalisations d'une suite des variables aléatoires X_1, X_2, \dots, X_n (i.i.d) de loi F , alors la fonction de répartition de la statistique d'ordre $X_{i:n}$ est donnée par :

$$F_{i:n}(x) = p(X_{i:n} \leq x) = \sum_{k=1}^n \binom{n}{k} [F(x)]^k [1 - (F(x))]^{n-k} \quad (1.8)$$

En conséquence, les lois de $M_n = X_{n,n} = \max(X_1, X_2, \dots, X_n)$ et de $W_n = X_{1,n} = \min(X_1, X_2, \dots, X_n)$, notées respectivement F_M et F_W , sont données par :

$$\forall x \in \mathbb{R} : F_M(x) = [F(x)]^n ; F_W(x) = 1 - [1 - F(x)]^n \quad (1.9)$$

Dans la plupart des cas F n'est pas connue et parfois, l'utilisation de la loi des statistiques d'ordre aboutissent à des calculs compliqués. D'où l'utilité d'établir un résultat asymptotique (résultat de base de la *TVE*), connu sous le nom du théorème de Fisher-Tippett (1928). Par analogie avec le théorème central limite (*TCL*), le théorème de Fisher-Tippett établit la convergence en loi de M_n (le théorème est aussi applicable à W_n) :



Méthode des maxima (minima) par bloc.

Théorème 1.2 (Fisher – Tippet) . Soient $X_1, \dots, X_i, \dots, X_n$ n variables aléatoires indépendantes et de même loi de probabilité F . S'il existe des constantes a_n et b_n et une loi G non dégénérée (c'est à dire différente d'une masse de Dirac) tel que :

$$\lim_{n \rightarrow \infty} \Pr \left(\frac{M_n - b_n}{a_n} \leq x \right) = G(x), \forall x \in \mathbb{R} \quad (1.10)$$

Alors G appartient à l'un des trois types de loi :

- Loi de Gumbel ($\xi = 0$) :

$$\Lambda_{\mu, \sigma}(x) = \exp \left(- \exp - \frac{x - \mu}{\sigma} \right), x \in \mathbb{R} \quad (1.11)$$

- Loi de Fréchet ($\xi > 0$) :

$$\Phi_{\mu, \sigma, \xi}(x) = \begin{cases} \exp - \left(\frac{x - \mu}{\sigma} \right)^{-\frac{1}{\xi}} & \text{si } x > \mu \\ 0 & \text{si } x \leq \mu \end{cases} \quad (1.12)$$

- Loi de Weibull ($\xi < 0$) :

$$\Psi_{\mu, \sigma, \xi}(x) = \begin{cases} \exp - \left(- \frac{x - \mu}{\sigma} \right)^{-\frac{1}{\xi}} & \text{si } x < \mu \\ 1 & \text{si } x \geq \mu \end{cases} \quad (1.13)$$

Les trois lois de probabilité ci-dessus sont appelées lois des valeurs extrêmes.

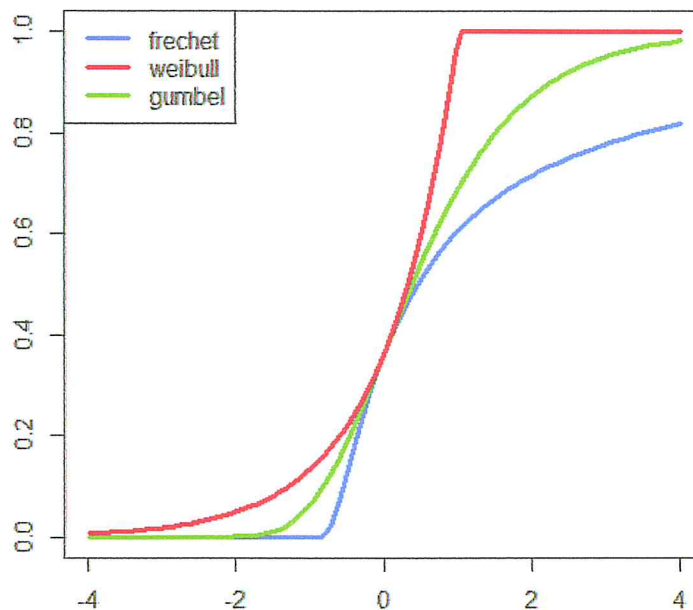
Pour la preuve, nous renvoyons à Embrechts et al. (1997), chapitre 3, p.122. Il faut bien signaler que chacune des trois lois des valeurs extrêmes

peut s'obtenir par une transformation fonctionnelle de l'autre. D'une façon similaire, on définit les lois des valeurs extrêmes associées au minimum.

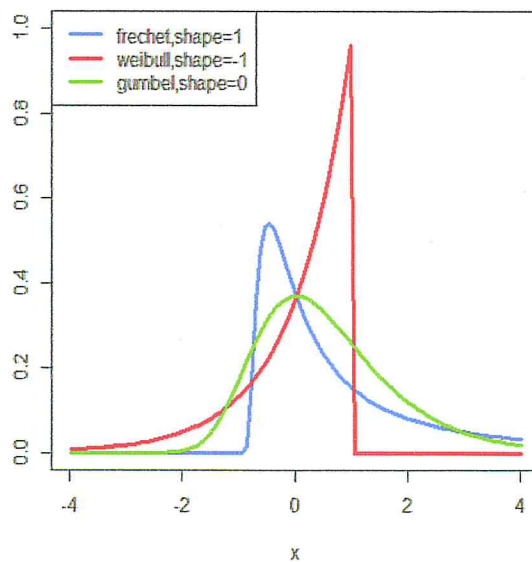
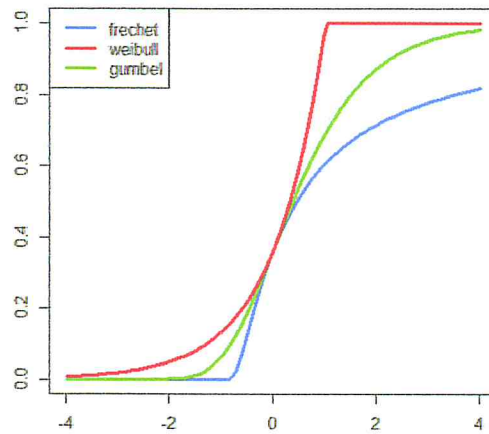
S'il existe des Suites de normalisation $a_n > 0$ et $b_n \in \mathbb{R}$ et une loi non dégénérée H^* tel que :

$$\lim_{n \rightarrow \infty} P \left[\frac{W_n - b_n}{a_n} \leq x \right] = \lim_{n \rightarrow \infty} (1 - F(a_n x + b_n))^n = H^*(x) \quad (1.14)$$

alors H^* est une loi des valeurs extrêmes associée au minimum.



Représentation de la fonction de répartition :
Gumbel ($\xi = 0$), Fréchet ($\xi = 1$) et Weibull ($\xi = -1$) :



lois des valeurs extremes.

En effet : $\min(X_i) = -\max(-X_i)$. Ainsi,

$$\forall x \in \mathbb{R}, \forall 1 \leq i \leq n : P[\min(X_i) > x] = P[\max(-X_i) \leq -x] \implies 1 - F_W(x) = F_M(-x) \quad (1.15)$$

Von Mises (1954,) et Jenkinson (1955) ont établi un résultat. théorique qui permet d'identifier une famille de lois GEV. qui regroupe les trois types

de lois associés aux valeurs extrêmes. Le résultat théorique est formulé par le biais du théorème suivant :

Théorème 1.3 *Sil existe des suites de normalisations $a_n > 0$ et $b_n \in \mathbb{R}$ et une loi G^* tel que :*

$$\lim_{n \rightarrow \infty} P \left[\frac{M_n - b_n}{a_n} \leq x \right] = G^*(x); \forall x \in \mathbb{R} \quad (1.16)$$

alors G^* est donnée par :

$$G_{\mu, \sigma, \xi}(x) = \exp \left\{ - \left[1 + \xi \left(\frac{x - \mu}{\sigma} \right) \right]^{-\frac{1}{\xi}} \right\} \quad (1.17)$$

Où x est tel que $1 + \xi \frac{(x - \mu)}{\sigma} > 0$, $-\infty < \mu < +\infty$, $\sigma > 0$, et $-\infty < \xi < +\infty$. Les paramètres μ , σ et ξ sont respectivement, les paramètres de position, d'échelle et de forme.

Une démonstration détaillée de ce théorème est donnée dans l'ouvrage de Resnick (1987), avec des développements dans le livre d'Embrechts et al. (1997, p. 152) et des illustrations dans le livre de Galambos (1987, pp.53-54).

1.4 Domaines d'attraction

Il faut bien noter que le paramètre ξ conditionne le type de la loi des valeurs extrêmes. Nous présentons dans ce qui suit, les domaines d'attractions dans les trois cas correspondant au paramètre ξ .

On commence par définir les fonctions à variations régulières.

Définition 1.4 *Une fonction L est dite à variation lente si $L(t) > 0$ pour t assez grand et si pour tout $x > 0$, on a :*

$$\lim_{t \rightarrow \infty} \frac{L(tx)}{L(t)} = 1 \quad (1.18)$$

1. Cas où $\xi = 0$: La loi présente dans la queue une décroissance de type exponentielle, ce qui permet de caractériser dans ce cas, le domaine d'attraction de Gumbel $\mathcal{D}(\Lambda)$. Ce dernier est délicatement traitable, car il n'y a pas de lien direct entre la queue de la loi et les fonctions à variations lentes définies par (Delmas et Jourdain (2006)) :

Von Mises (1936,) a donné une caractérisation simple pour le domaine d'attraction de Gumbel, formulée par le biais du théorème suivant :

Théorème 1.4 *S'il existe une fonction mesurable R , appelée fonction auxiliaire telle que :*

$$\lim_{x \rightarrow \omega(F)} \frac{1 - F(t + xR(t))}{1 - F(t)} = \exp(-x) \quad (1.19)$$

avec

$$\omega(F) = \sup \{x \in \mathbb{R} : F(x) < 1\} \quad (1.20)$$

est le point terminal de F , alors $F \in \mathcal{D}(\Lambda)$

Le tableau suivant fournit quelques exemples de lois qui appartiennent au domaine d'attraction de Gumbel.

TAB. 1.1 – Lois du domaine d'attraction de Gumbel.

Loi	$1 - F(x)$
Benktender $I I$, $\alpha, \beta > 0$	$x^{-(1-\beta)} \exp(-\frac{\alpha}{\beta} x^\beta)$
Logistic	$\frac{1}{1+\exp(x)}$

2. Cas où $\xi < 0$: Le domaine d'attraction dans ce cas est celui de Weibull $\mathcal{D}(\Psi_\xi)$. Les lois de ce domaine sont bornées à droite, et par conséquent, le point terminal $\omega(F)$ est fini. Une caractérisation d'appartenance à ce domaine d'attraction est donnée par le théorème suivant (pour la démonstration, voir Gnedenko (1943,)) :

Théorème 1.5 *Une fonction de répartition F appartient au $\mathcal{D}(\Psi_\xi)$ si et seulement si $\omega(F) < +\infty$ et*

$$\bar{F}(\omega(F) - \frac{1}{x}) = x^{-\frac{1}{\xi}} l(x) \quad (1.21)$$

avec: \bar{F} est la fonction de survie donnée par $\bar{F}(x) = 1 - F(x)$, l est une fonction à variation lente.

Le tableau ci-dessous présente quelques lois qui appartiennent au domaine d'attraction de Weibull.

TAB. 1.2 – Lois du domaine d'attraction de Weibull

Loi	$1 - F(\omega(F) - \frac{1}{x})$	ξ	$l(x)$
Uniforme	$\frac{1}{x}, x > 1$	-1	1
Weibull	$1 - \exp(-x^{-\alpha})$	$-\frac{1}{\alpha}$	$1 - \frac{x^{-\alpha}}{2} + o(x^{-\alpha})$
Reverse Burr (λ, β, τ)	$(\frac{\beta}{\beta+x^\tau})^\lambda$ avec $x > 0$	$-\frac{1}{\lambda\tau}$	$\beta^\lambda(1 - \lambda\beta x^{-\tau} + o(x^{-\tau}))$

3. Cas où $\xi > 0$:

Ce cas correspond au domaine d'attraction de Fréchet. Les lois appartenant à ce domaine d'attraction sont caractérisées par une queue à décroissance lente (polynomiale) à l'infini, et un point terminal $\omega(F) = +\infty$. Elles sont dites aussi lois à queues lourdes. Une caractérisation de ce domaine d'attraction noté $\mathcal{D}(\Phi_\xi)$ est donnée par le théorème suivant (pour la démonstration, voir Gnedenko (1943)) :

Théorème 1.6 Une fonction de répartition F appartient au $\mathcal{D}(\Phi_\xi)$ si et seulement si sa fonction de survie est donnée par :

$$\bar{F}(x) = x^{-\frac{1}{\xi}} l(x) \quad (1.22)$$

où l est une fonction à variation lente.

Le tableau ci-dessous présente quelques lois qui appartiennent au domaine d'attraction de Fréchet .

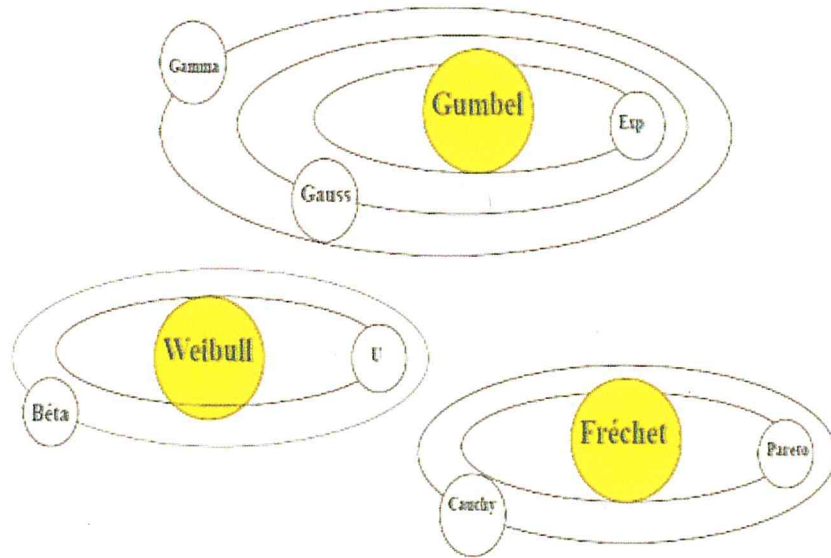
TAB. 1.3 – Lois du domaine d'attraction de Fréchet

Loi	$1 - F(x)$	ξ	$l(x)$
Pareto(α) avec $\alpha > 0$	$x^{-\alpha}$	$\frac{1}{\alpha}$	1
Burr(η, τ, λ) avec $\eta, \tau, \lambda > 0$	$(\frac{\eta}{\eta+x^\tau})^\lambda, x > 0$	$\frac{1}{\lambda\tau}$	$(\frac{\eta}{1+\frac{\eta}{x^\tau}})^\lambda$
Pareto Généralisé(σ, ξ) avec $\sigma, \xi > 0$	$(1 + \frac{\xi x}{\sigma})^{-\frac{1}{\xi}}, x > 0$	ξ	$(\frac{\sigma}{\xi})^{\frac{1}{\xi}} (1 + \frac{\sigma}{\xi x})^{-\frac{1}{\xi}}$

1. Une fonction de répartition F peut appartenir à deux domaines d'attraction différents \mathcal{D}_1 et \mathcal{D}_2 (pour plus de détails, voir Resnick (1987)).

2. Les variables aléatoires dans les trois domaines d'attraction de Gumbel, de Fréchet et de Weibull sont liées par la relation suivante (voir Embrecht et al. (1997)) :

$$X \in \mathcal{D}(\Phi_\alpha) \iff \log(X^{\frac{1}{\xi}}) \in \mathcal{D}(\Lambda) \iff -X^{-1} \in \mathcal{D}(\Psi_\xi) \quad (1.23)$$



Loi du domaine d'attraction $\mathcal{D}(\Lambda), \mathcal{D}(\Phi_\xi)$ et $\mathcal{D}(\Psi_\xi)$.

3. Les trois domaines d'attraction $\mathcal{D}(\Lambda), \mathcal{D}(\Phi_\xi)$ et $\mathcal{D}(\Psi_\xi)$ ne sont pas restrictifs. Il se peut, qu'une loi n'appartienne à aucun de ces trois domaines d'attraction ; c'est le cas de la loi définie par :

$$F(x) = 1 - \frac{1}{\log(x)}, x \geq e. \quad (1.24)$$

Outre les domaines d'attraction qui caractérisent, chacune des lois *GEV*, les champs d'application conditionnent leurs choix. En effet, la loi de Gumbel par exemple, sert de référence car elle permet de mesurer l'écart entre les lois empiriques et la loi normale. Elle est généralement utilisée pour prévoir le niveau des crues d'un fleuve. Elle peut aussi servir à prédire la probabilité d'un événement, critique comme un tremblement de terre. En ce qui concerne la loi de Fréchet, elle est couramment évoquée en finance et en macroéconomie, où les variations sont généralement non bornées. Par ailleurs, cette loi s'applique en hydrologie pour modéliser des événements extrêmes tels que les débits maximaux des rivières. La loi de Weibull quant à elle, est largement utilisée grâce à sa grande flexibilité. Elle est préconisée en particulier, dans les analyses de durée de vie ou pour l'estimation des potentiels éoliens.

1.5 Estimation des modèles GEV

Les lois GEV se distinguent, essentiellement selon le paramètre ξ . Ainsi, l'estimation sera établie selon deux cas : $\xi = 0$ (loi de Gumbel) et $\xi \neq 0$ (lois de Fréchet et de Weibull).

La littérature propose une multitude de méthodes, consacrées principalement à l'estimation des paramètres des lois GEV. Nous pouvons citer par exemple, les méthodes d'estimation empirique (Gumbel et Mustafi (1967) ou Tiago de Oliveira (1975)), la méthode du maximum de vraisemblance (Smith (1987 et 1985), Prescott et Walden (1980 et 1983), Tiago de Oliveira (1982), Hougaard (1986)), la méthode des moments (Christopeit, (1994)), la méthode des moments de probabilité pondérés (Greenwood et al., (1979)) ou encore, les méthodes bayésiennes (Lye et al., (1993)).

La comparaison de ces méthodes d'estimation s'effectue dans la plupart du temps dans un cadre empirique (voir Benkhaled (2007) et Lang et al. (2012)).

Par ailleurs, d'autres approches principalement non paramétriques, ont été dédiées à l'estimation de l'indice de queue. Nous pouvons citer à titre d'exemple l'estimateur de Pickands (1975), l'estimateur de Hill (1975) (pour la loi de Fréchet uniquement) et l'estimateur de Dekkers-Einmahl-de Hann (Dekkers et al. (1989)).

1.5.1 Estimation pour la loi de Gumbel

Dans cette partie, nous notons la loi de Gumbel par $G_0(\mu_0, \sigma_0)$, où μ_0 et σ_0 désignent respectivement, les paramètres de position et d'échelle ($\xi = 0$ dans le cas de la loi de Gumbel). La fonction de densité de Gumbel est notée g_0 .

Méthode de la régression linéaire :

Soient Y_1, \dots, Y_n les maxima par bloc issus d'une loi de Gumbel et soient $Y_{1,m} \leq \dots \leq Y_{m,m}$ les statistiques d'ordre associées à Y_1, \dots, Y_m . La méthode d'estimation par régression linéaire consiste à ajuster le nuage de points :

$$(-\log(-\log(G_0(Y_{i,m}))); Y_{i,m}) \quad (1.25)$$

à une droite d'équation $Y_{i,m} = aX + b$, avec $X = -\log(-\log(G_0(Y_{i,m})))$, en calculant les coefficients a et b par la méthode des moindres carrés.

Le coefficient directeur (a) et l'ordonnée à l'origine (b) de la droite ajustée, permettent de calculer les estimateurs $\hat{\mu}_0^{reg}$ et $\hat{\sigma}_0^{reg}$:

$$\begin{cases} \hat{\mu}_0^{reg} = b \\ \hat{\sigma}_0^{reg} = a \end{cases} \quad (1.26)$$

L'estimation des quantiles extrêmes dépassés avec une faible probabilité p , pour la loi de Gumbel G_0 et pour une période de retour $T = \frac{1}{p} = \frac{1}{1-G_0(y)}$, est obtenue en remplaçant μ et σ par leurs estimations dans la formule suivante :

$$q_0 = \mu_0 + \sigma_0 \left(-\log \left(-\log \left(1 - \frac{1}{T} \right) \right) \right) \quad (1.27)$$

Ainsi, l'estimateur du quantile extrêmes \hat{q}_0^{reg} est :

$$\hat{q}_0^{reg} = \hat{\mu}_0^{reg} + \hat{\sigma}_0^{reg} \left(-\log \left(-\log \left(1 - \frac{1}{T} \right) \right) \right) \quad (1.28)$$

1.5.2 Méthode du maximum de vraisemblance

Cette méthode consiste à chercher les paramètres $\Theta_0 = (\mu_0, \sigma_0)$ qui maximisent la fonction de vraisemblance $l(\mu_0, \sigma_0) = \prod_{i=1}^m g_0(y_i, \mu_0, \sigma_0)$ ou log-vraisemblance $L(\mu_0, \sigma_0) = \log l(\mu_0, \sigma_0)$, en résolvant le système suivant :

$$\begin{cases} \frac{\partial \log L(\Theta_0)}{\partial \mu_0} = 0 \\ \frac{\partial \log L(\Theta_0)}{\partial \sigma_0} = 0 \end{cases} \quad (1.29)$$

Cependant, la solution n'est pas explicite et il faut passer par une résolution numérique en utilisant par exemple, la méthode de *Newton – Raphson*.

1.5.3 Méthode des moments

La méthode consiste à égaliser respectivement, les deux premiers moments théoriques m_1 et m_2 aux deux premiers moments observés : moyenne (\bar{y}_m) et variance S_Y^2 empiriques.

Dans le cas d'une variable aléatoire issue d'une loi de Gumbel, la méthode du maximum de vraisemblance abouti au système suivant :

$$\begin{cases} \mathbb{E}[Y] = \mu_0 + \gamma_* \sigma_0 = \bar{y}_m \\ \mathbb{V}[Y] = \frac{1}{6} \pi^2 \sigma_0^2 = S_Y^2 \end{cases} \quad (1.30)$$

où $\gamma_* \approx 0.57721$ représente la constante d'Euler. La résolution de ce système d'équation fournit les valeurs des estimateurs $\hat{\mu}_0^{mom}$ et $\hat{\sigma}_0^{mom}$:

$$\begin{cases} \hat{\sigma}_0^{mom} = \frac{\sqrt{6} S_Y}{\pi} \\ \hat{\mu}_0^{mom} = \bar{y}_m - \gamma_* \hat{\sigma}_0^{mom} \end{cases} \quad (1.31)$$

1.5.4 Estimation pour les lois de Fréchet et de Weibull

Pour $\xi \neq 0$, les domaines d'attractions sont ceux de Fréchet et de Weibull. Ces lois seront notées $G_\xi(\mu, \sigma, \xi)$ ($\xi > 0$ pour une loi de Fréchet et $\xi < 0$ pour une loi de Weibull). La fonction de densité sera notée g_ξ .

Nous présentons pour les deux lois de Fréchet et de Weibull, les méthodes d'estimation du maximum de vraisemblance et des moments de probabilités pondérés.

1.5.5 Méthode du maximum de vraisemblance

La méthode consiste à chercher les paramètres $\Theta = (\mu, \sigma, \xi)$ qui maximise la fonction de log-vraisemblance. Pour l'obtention de $\hat{\mu}^{mv}$, $\hat{\sigma}^{mv}$ et $\hat{\xi}^{mv}$, la maximisation conduit à résoudre le système suivant :

$$\begin{cases} \frac{\partial \log L(\Theta)}{\partial \sigma} = 0 \\ \frac{\partial \log L(\Theta)}{\partial \xi} = 0 \\ \frac{\partial \log L(\Theta)}{\partial \mu} = 0 \end{cases} \quad (1.32)$$

Smith(1985) a étudié le problème de l'estimation par la méthode du maximum de vraisemblance. Il a obtenue les résultats suivant :

- Si $\xi > -0.5$, les estimateurs du maximum de vraisemblance possèdent des propriétés asymptotiques telle que la convergence vers la vraie valeur du paramètre inconnu, l'invariance par rapport à une transformation paramétrique et l'efficacité asymptotique.
- Si $-1 < \xi < -0.5$, les estimateurs du maximum de vraisemblance ne possèdent pas des propriétés asymptotiques standards.
- Si $\xi < -1$, l'obtention des estimateurs du maximum de vraisemblance n'est pas garantie.

1.5.6 Méthode des moments de probabilités pondérés PWM :

Cette méthode a été introduite par Greenwood et al. (1979) dans le but de surmonter les problèmes de convergence, surtout pour les petits échantillons ($n < 25$) (Lubes et al, 1991).

Les moments des probabilités pondérés sont définis par :

$$M_{p,r,s} = \mathbb{E}[Y^p(F(Y))^r(1 - F(Y))^s] \quad (1.33)$$

où p, r, s sont des réels et Y est une variable aléatoire issue de la loi F .

$$\hat{\xi}_{k,n}^P = \frac{1}{\ln 2} \ln \left(\frac{X_{(n-k+1,n)} - X_{(n-2k+1,n)}}{X_{(n-2k+1,n)} - X_{(n-4k+1,n)}} \right) \quad (1.39)$$

Où X_1, \dots, X_n un échantillon de v.a (i.i.d) et n est la taille de l'échantillon observé. L'estimateur de Pickands est un estimateur convergent. Plus précisément :

$$\sqrt{k}(\hat{\xi}_{k,n}^P - \xi) \rightarrow \mathcal{N}(0, \sigma^2(\xi)) \quad (1.40)$$

où

$$\sigma^2(\xi) = \frac{\xi^2(2^{2\xi+1} + 1)}{(2(2\xi - 1) \ln 2)^2} \quad (1.41)$$

1.5.9 Estimateur de Hill

L'estimateur de Hill (1975) est l'un des estimateurs les plus répandus dans les applications de la théorie des valeurs extrêmes (TVE). Introduit par Hill, il est défini de la façon suivante (pour tout $\xi > 0$) :

$$\hat{\xi}_{k,n}^H = \frac{1}{k} \sum_{i=n-k+1}^n \ln X_{(i,n)} - \ln X_{(n-k+1,n)} \quad (1.42)$$

Le théorème suivant présente les propriétés de l'estimateur de Hill :

Théorème 1.7 Soit X_1, \dots, X_n est un échantillon de loi F tel que $\bar{F}(X) = x^{-\frac{1}{\xi}} L(x)$, $x > 0$ et L une fonction à variations régulières d'indice 0, alors:

- Si X_1, \dots, X_n un échantillon d'observation i.i.d, $k \rightarrow \infty$ et $k/n \rightarrow 0$ pour $n \rightarrow \infty$, alors l'estimateur de Hill $\hat{\xi}_{k,n}^H$ converge en probabilité vers ξ .
- Si $k/n \rightarrow 0$, $k/\ln(\ln(n)) \rightarrow \infty$ et les X_1, \dots, X_n sont i.i.d, alors l'estimateur de Hill converge presque sûrement vers ξ .
- Si les X_1, \dots, X_n sont i.i.d et si \bar{F} est à variations régulières, alors :

$$\sqrt{k}(\hat{\xi}_{k,n}^H - \xi) \rightarrow \mathcal{N}(0, \xi^2) \quad (1.43)$$

Pour la preuve de ce théorème, voir Hill (1975).

1.5.10 Estimateur des moments de Dekkers-Einmahl-De Haan

L'estimateur de Dekkers-Einmahl-De Haan présente une extension de l'estimateur de Hill pour $\xi \in \mathbb{R}$. Il est défini comme suit :

$$\hat{\xi}_{k,n}^D = 1 + H_n^{(1)} + \frac{1}{2} \left(\frac{(H_n^{(1)})^2}{H_n^{(2)}} - 1 \right)^{-1} \quad (1.44)$$

où :

$$H_n^{(1)} = \frac{1}{k} \sum_{i=n-k+1}^n \ln X_{(i,n)} - \ln X_{(n-k,n)} \quad (1.45)$$

et

$$H_n^{(2)} = \frac{1}{k} \sum_{i=n-k+1}^n (\ln X_{(i,n)} - \ln X_{(n-k,n)})^2 \quad (1.46)$$

Sous certaines conditions sur k , l'estimateur de Dekkers-Einmahl-De Haan converge asymptotiquement vers la loi normale.

1.6 Approche du loi GPD

1.6.1 Modélisation des excès

Alternativement à l'approche des maxima par bloc, le modèle de renouvellement pour des v.a (i.i.d), a donné naissance à un autre volet de la théorie des valeurs extrêmes, illustré formellement par l'approche des dépassements de seuil (Peaks over Threshold (*POT*)). Contrairement à l'approche du maxima par bloc, la méthode *POT* consiste à utiliser toutes les observations, appelés excès, qui dépassent un certain seuil suffisamment élevé. L'objectif est d'analyser leur comportement asymptotique.

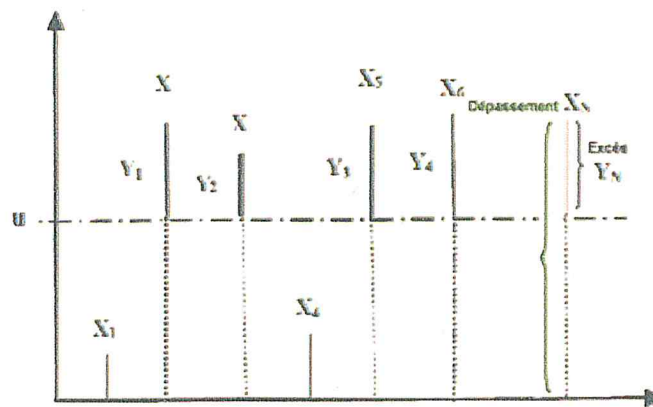
L'approche *POT* a été introduite pour remédier aux défauts de l'approche classique de la TVE. Les critiques sont essentiellement dues au fait que la méthode d'estimation de la loi *GEV* se base sur les "block component -wise", ce qui implique systématiquement une perte d'information. En plus, certains blocs peuvent contenir plusieurs valeurs extrêmes issues de la loi initiale, alors que d'autres peuvent ne pas en contenir (Mc Neil et Frey (2002), Katz (2002)).

Historiquement, la méthode *POT* introduite par pickands (1975) et reprise par de Haan et Rootzen (1993), a été initialement appliqué pour le traitement

des données hydrologiques, notamment le volume et la durée des déficits en eau. Quant aux volets théoriques de la méthode, ils ont été abondamment développés par divers auteurs tels que Todorovic et Zelenhasie (1970), Todorovic et Rousselle (1971), Smith (1987), Davison et Smith (1990), Reiss et Thomas (2001). Le point de départ de l'approche POT consiste à choisir un seuil $u \in \mathbb{R}$ et de considérer les variables Y_j définies, à partir des variables initiales X_1, \dots, X_n , par :

$$\begin{cases} N_u = \text{card} \{i : i = 1, \dots, n / X_i > u\} \\ Y_j = X_i - u > 0, \end{cases} \quad \text{pour } 1 \leq j \leq N_u \quad (1.47)$$

où : X_1, \dots, X_n un échantillon de v.a (i.i.d), N_u est le nombre des dépassements du seuil u et Y_1, \dots, Y_{N_u} les excès correspondants.



Représentation des excs Y issue des dpassements X au-del d'un seuil u , (Raggad,2009).

Le nombre de dépassements d'un seuil u suit une loi de Poisson de paramètre τ , notée $P(\tau)$. Ce résultat est justifié sur la base de l'approximation de la loi binomiale par la loi de Poisson. En effet, si on désigne par $\{X_i > u_n\}$ l'événement "Dépassement d'un seuil u_n " et sa fonction indicatrice $I_{\{X_i > u_n\}}$ définie par :

$$I_{\{X_i > u_n\}} = \begin{cases} 1, & X_i > u_n \\ 0, & \text{sin on} \end{cases} \quad (1.48)$$

alors le nombre de dépassements de u_n , égal à $N_{u_n} = \sum_{i=1}^n I_{\{X_i > u_n\}}$, suit une loi binomiale $Bin(n, 1 - F(u_n))$ qui peut être approchée par une loi de Poisson

lorsque $n \rightarrow +\infty$, $1 - F(u_n) \rightarrow 0$ et $n(1 - F(u_n)) \rightarrow \tau$. Une fois les excès identifiés, l'approche POT consiste, en partant de la loi F de X , à déterminer une loi conditionnelle F_u , pour les variables aléatoires Y qui dépassent le seuil u .

La loi conditionnelle F_u est définie par :

$$F_u(y) = \mathbb{P}(X - u \leq y \mid X > u) = \begin{cases} \frac{F(u+y) - F(u)}{1 - F(u)}, & \text{si } y \geq 0 \\ 0, & \text{si } y < 0 \end{cases} \quad (1.49)$$

La loi asymptotique associée à $F_u(y)$ est donnée par le théorème suivant :

Théorème 1.8 (Pickands, Balkema et de Haan) . Si F (la loi de X) appartient à l'un des trois domaines d'attraction : $\mathcal{D}(\text{Fréchet})$, $\mathcal{D}(\text{Gumbel})$ ou $\mathcal{D}(\text{Weibull})$, alors il existe une fonction $\sigma(u)$ positive, tel que :

$$\lim_{u \rightarrow \omega(F)} \sup_{0 < y < \omega(F) - u} |F_u(y) - H_{\xi, \sigma(u)}(y)| = 0 \quad (1.50)$$

où

$$\omega(F) = \sup \{x \in \mathbb{R} : F(x) < 1\} \quad (1.51)$$

est le point terminal de F et $H_{\xi, \sigma(u)}(y)$ est la fonction de répartition de la loi de Pareto Généralisée (GPD, Generalized Pareto Distribution), définie par :

$$H_{\xi, \sigma(u)}(y) = \begin{cases} 1 - (1 + \frac{\xi y}{\sigma(u)})^{-\frac{1}{\xi}}, & \text{si } \xi \neq 0 \\ 1 - \exp(-\frac{y}{\sigma(u)}), & \xi = 0. \end{cases} \quad (1.52)$$

Pour la démonstration de ce théorème, voir Pickands (1975).

Nous pouvons estimer la période de retour par la relation suivante (Rosbjerg (1987)) :

$$T_u = \frac{1}{\tau p} = \frac{1}{\tau(1 - H_{\xi, \sigma(u)}(y))} \quad (1.53)$$

Ainsi, si on s'intéresse particulièrement à un niveau de dépassement, une fois chaque N années, nous pouvons utiliser l'équation 1.53 pour déduire que ce niveau de dépassement est donné par :

$$Z_n = \begin{cases} \mu + \frac{\sigma}{\xi} [(Nn_y \zeta_u)^\xi - 1], & \text{si } \xi \neq 0; \\ \mu + \sigma \log(Nn_y \zeta_u), & \xi = 0. \end{cases} \quad (1.54)$$

où $\zeta_u = \mathbb{P}\{X > u\}$ et n_y est le nombre de dépassements.

Il faut bien signaler que les deux approches ; blocs des maxima et dépassements d'un seuil, sont équivalentes. Ainsi, pour un seuil adéquat, la loi des excès peut être approchée par une loi GPD d'indice extrême ξ , identique à celui de la loi GEV. En effet :

$$\mathbb{P}\{X > x|X > u\} = [1 + \xi(\frac{x - \mu}{\sigma})]^{-\frac{1}{\xi}} \mathbb{P}\{X > x\} = \zeta_u [1 + \xi(\frac{x - \mu}{\sigma})]^{-\frac{1}{\xi}} \quad (1.55)$$

La fonction de répartition ($H_{\xi, \sigma(u)}$) met en exergue trois paramètres : μ , σ et ξ . Le paramètre ξ , appelé indice des valeurs extrêmes ou indice de queue, renseigne sur le type de la loi asymptotique (Fréchet pour $\xi > 0$, Gumbel pour $\xi = 0$ et Weibull pour $\xi < 0$). En outre, plus ξ est élevé en valeur absolue, plus le poids des extrêmes est important. Nous parlons dans ce cas d'une loi à queue épaisse.

L'identification du seuil est la phase la plus déterminante dans l'implémentation de l'approche *POT*, vu que la qualité du modèle en dépend. C'est à partir de la définition d'un bon seuil u que nous pouvons garantir la convergence des excès vers une *GPD*, estimer les autres paramètres et évaluer le temps de retour.

D'une manière similaire à la détermination des blocs dans la première approche de la *TVE*, le choix du seuil doit établir un compromis entre biais et variance. Concrètement, le seuil doit être suffisamment grand pour pouvoir utiliser les résultats asymptotiques, mais pas trop élevé afin d'obtenir des estimations précises. Par contre, le choix d'un seuil faible risque de déclarer abusivement des observations extrêmes, introduire un biais dans l'estimation et par conséquent, mal approximer la loi asymptotique. Dans ce sens, plusieurs méthodes de détection du seuil ont été proposées. Certaines ont le défaut d'être subjectives, et donc elles ne doivent être utilisées que pour proposer un intervalle de seuils candidats. D'autres par contre, ont l'avantage d'être objectives.

1.6.2 Sélection du seuil

L'estimation des paramètres de la distribution *GPD* pose le problème de la détermination du seuil u , le seuil ne doit pas être trop grand car il faut suffisamment de données pour avoir une bonne précision des estimateurs. La technique graphique est la plus utilisée pour estimer u .

Le Mean Excess Plot

Etant donnée un échantillon aléatoire X_1, X_2, \dots, X_n d'une distribution F et soit u un seuil donné. Le Mean Excess Plot est le graphe des points $(u, e(u))$ où $e(u)$ est la moyenne des excès au delà du seuil u , appelée aussi la Mean Excess Function définie par :

$$e(u) = E[X - u | X > u] \quad 0 < u < \infty \quad (1.56)$$

Elle correspond à une fonction de u que l'on peut exprimer à l'aide de $1 - F$. On suppose que pour le modèle proposé, $E(X) < \infty$. En pratique, la fonction des excès moyens $e(u)$ est estimée par $\hat{e}_n(u)$:

$$\hat{e}_n(u) = \frac{\sum_{i=1}^n (x_i - u) 1_{(x_i > u)}(x_i)}{\sum_{i=1}^n 1_{(x_i > u)}(x_i)} \quad (1.57)$$

Un des principaux outils du choix du seuil est le graphe de la sample mean excess function $\hat{e}_n(u)$ ou le mean excess plot.

Définition 1.5 *Le mean excess plot ou ME-plot est défini par le nuage de points suivant :*

$$\{(u, \hat{e}_n(u)), x_{1,n} < u < x_{n,n}\}$$

où $x_{n,n}$ et $x_{1,n}$ sont respectivement le maximum et le minimum de l'échantillon.

Comment interprète-t-on ce graphe ? Il faut remarquer tout d'abord que la mean excess function d'une $GPD(\sigma, \xi)$, pour $\xi < 1$, est

$$E[X - u | X > u] = \frac{\xi}{1 - \xi} u + \frac{\sigma}{1 - \xi} \quad (1.58)$$

Dans ce cas, le seuil à retenir u , est celui pour lequel la moyenne des excès est approximativement linéaire. Plus les queues des distributions sont épaisses, plus la fonction $e(u)$ tend rapidement vers l'infini. En effet, la fonction moyenne des excès relative à une distribution de loi Pareto de paramètre $\alpha > 0$, appartenant au domaine d'attraction de Fréchet d'indice des valeurs extrêmes $\xi = \frac{1}{\alpha}$, s'écrit comme suit :

$$e(u) = \frac{k + u}{\alpha - 1} \quad (1.59)$$

Dans la pratique, si la ME-plot semble avoir un comportement linéaire au-dessus d'une certaine valeur de u , cela signifie que les excès au-dessus de

ce seuil suivent une *GPD*. Cette valeur correspond à un arbitrage acceptable entre le souhait d'avoir un seuil élevé et celui d'obtenir un échantillon d'excès de taille suffisante.

Seuil aléatoire

McNeil et Frey choisissent un seuil aléatoire : $N_u = k$ (Il y a k observations excédentaires) où $k \ll N$ et le seuil est donc la $(k+1)^{i\grave{e}me}$ statistique d'ordre.

Soient $z_{1,n} \leq z_{2,n} \leq \dots \leq z_{n,n}$, les excès ordonnés. La loi Pareto généralisée de paramètres ξ et σ est ajustée aux données :

$$(z_{1,n} - z_{k+1,n}, \dots, z_{k,n} - z_{k+1,n}).$$

Ces auteurs font une étude simulatoire à partir d'une distribution *t* de Student pour déterminer une valeur de k appropriée.

Bootstrap

Caers, Beirlant et Maes utilisent une méthode basée sur le bootstrap semi-paramétrique afin de déterminer l'erreur moyenne carrée (**MSE**). Ceci permet de choisir le seuil qui donne la plus petite *MSE*. La méthodologie est la suivante :

1. Un seuil u est fixé pour lequel les paramètres de la Pareto généralisée, sont estimés par $\hat{\xi}_u, \hat{\sigma}_u$ avec la méthode de notre choix (par exemple, avec le maximum de vraisemblance). Ceci nous permet d'utiliser la densité semi-paramétrique suivante :

$$\hat{F}_s(x \setminus u) = \begin{cases} (1 - \hat{F}(u)) \left(1 - \left(1 + \frac{\hat{\xi}_u}{\hat{\sigma}_u}(x - u)\right)\right)^{-\frac{1}{\hat{\xi}_u}} + \hat{F}(u), & x > u, \\ \hat{F}(x), & x \leq u. \end{cases} \quad (1.60)$$

2. Cette densité estimée permet de faire du bootstrap semi-paramétrique. Donc B échantillons sont tirés à partir $\hat{F}_s(x \setminus u)$:

$$X^{(b)} = \{X_1^{(b)}, \dots, X_n^{(b)}\}, \quad (1.61)$$

où $b = 1, \dots, B$.

3. À partir de ces échantillons, un paramètre (par exemple l'indice de queue ξ) est estimé. De ces B estimés, le biais et la variance de l'estimateur sont calculés et combinés pour former un estimé de la **MSE**.
4. Le seuil ayant donné la plus petite **MSE** pour l'estimation du paramètre qui nous intéresse est retenu.

Il n'est pas clair que cette méthode produise un estimé raisonnable du biais du choix de seuil u , le biais correspondrait plutôt à la méthode d'estimation du paramètre d'intérêt (par exemple ξ) pour un seuil fixé.

L'estimation de ξ et σ pose le problème de la détermination du seuil u . Il doit être suffisamment grand pour que l'on puisse appliquer le résultat précédent, mais ne doit pas être trop grand afin d'avoir suffisamment de données pour obtenir des estimateurs de bonne qualité.

1.7 Estimation du modèle GPD

1.7.1 Méthode des moments

la Méthode consiste à égaliser les moments théoriques et les moments empiriques de façon à obtenir :

$$\begin{cases} \mathbb{E}(Y) = \frac{\sigma_u}{1-\xi} = \bar{Y} \\ \mathbb{V}(Y) = \frac{\sigma_u^2}{(1-\xi)^2(1-2\xi)} = S_Y^{*2} \end{cases} \quad (1.62)$$

Où \bar{Y} et S_Y^{*2} représentent respectivement, la moyenne et la variance empirique des excès.

Deux méthodes d'estimation sont ici encore réalisables : l'Estimation par Maximum de Vraisemblance (*EMV*) et celle par les Moments Pondérés (*EMP*).

1.7.2 Méthode du maximum de vraisemblance(EMV)

Supposons que notre échantillon des excès $Y = (Y_1, \dots, Y_{N_u})$ est i.i.d avec comme fonction de distribution la *GPD*. La fonction de densité $g_{\xi, \sigma}$ de *GPD* $G_{\xi, \sigma}$ est

$$g_{\xi, \sigma}(y) = \begin{cases} \frac{1}{\sigma} (1 + \frac{\xi y}{\sigma})^{-\frac{1}{\xi}-1} & \text{si } \xi \neq 0 \\ \exp(-\frac{y}{\sigma}) & \text{si } \xi = 0 \end{cases}, \quad \sigma > 0. \quad (1.63)$$

La fonction log-vraisemblance est donc égale à

$$l((\xi, \sigma); X) = -N_u \ln \sigma - \left(\frac{1}{\sigma} + 1\right) \sum_{i=1}^{N_u} \ln\left(1 + \frac{\xi y_i}{\sigma}\right). \quad (1.64)$$

prendre des dérivées partielles de $l((\xi, \sigma); X)$ par rapport à ξ et σ , on obtient :

$$\begin{cases} \frac{1}{\sigma^2} \sum_{j=1}^{N_u} \ln\left(1 + \frac{\xi y_j}{\sigma}\right) - \sum_{i=1}^{N_u} \frac{y_i}{\sigma + \xi x_i} = 0 \\ -N_u + (1 + \xi) \sum_{i=1}^{N_u} \frac{y_i}{\sigma + \xi y_i} = 0 \end{cases} \quad (1.65)$$

où y_1, \dots, y_{N_u} est une réalisation de Y_1, \dots, Y_{N_u} . EMV $(\hat{\xi}_{N_u}, \hat{\sigma}_{N_u})$ de (ξ, σ) comme une solution de se système d'équations. Notez que ce système n'a pas de solution explicite et que des méthodes numériques sont donc nécessaires pour calculer les valeurs estimées.

Smith montre la normalité asymptotique de $(\hat{\xi}_{N_u}, \hat{\sigma}_{N_u})$ fournie $\xi > -\frac{1}{2}$ plus précisément, nous avons

$$\sqrt{N_u} \begin{pmatrix} \hat{\xi}_{N_u} - \xi \\ \hat{\sigma}_{N_u} - \sigma \end{pmatrix} \xrightarrow{d} \mathcal{N}_2(0, \mathbb{Q}^{-1}) \text{ comme } N_u \rightarrow \infty, \quad (1.66)$$

où

$$\mathbb{Q}^{-1} = (1 + \xi) \begin{pmatrix} 1 + \xi & -1 \\ -1 & 2 \end{pmatrix}, \quad (1.67)$$

et $\mathcal{N}_2(\varepsilon, \omega)$ représente la distribution normale bivariée avec le vecteur moyen ε et la matrice de covariance ω . avec ce résultat, les intervalles de confiance pour les estimations de paramètres sont facilement construits.

1.7.3 Méthode des moments pondérés(EMP)

La définition des moments de probabilité pondérés pour une GPD, dans le cas $p = 1, r = 0, s = 0$ ou $s = 1$, est donnée par :

$$M_{1,0,s} = \frac{\sigma_u}{(s+1)(s+1-\xi)}, \quad \xi < 1 \quad (1.68)$$

Avec :

$$\hat{M}_{1,0,s} = \frac{1}{k} \sum_{i=1}^k \left(1 - \frac{i}{k+1}\right) Y_{i,k} \quad (1.69)$$

La résolution de l'équation définissant $M_{1,0,s}$ en fonction de σ_u et ξ , conduit pour $s = 0$ et $s = 1$ aux estimateurs *PWM* suivants :

$$\hat{\xi}_u^{pwm} = 2 - \frac{\hat{M}_{1,0,0}}{\hat{M}_{1,0,0} - 2\hat{M}_{1,0,1}} \quad (1.70)$$

$$\hat{\sigma}_u^{pwm} = 2 \frac{\hat{M}_{1,0,0} \hat{M}_{1,0,1}}{\hat{M}_{1,0,0} - 2\hat{M}_{1,0,1}} \quad (1.71)$$

1.8 Estimation de la queue de la distribution

La distribution de Pareto généralisée est utilisée pour modéliser la queue (supérieure ou inférieure) d'une distribution présentant des valeurs extrêmes. Pour pouvoir faire de l'estimation des quantiles, il faut avoir une formulation qui combine la distribution estimée dans la queue et la distribution centrale. Une telle formulation est donnée par l'égalité suivante, $\forall u < x < x_F$

$$\bar{F}(x) = \bar{F}_u(x - u) \bar{F}(u), u < x < x^*. \quad (1.72)$$

où

$$F_u(y) = P(X - u \leq y / X > u). \quad (1.73)$$

Autrement dit, $\forall x > u$:

$$p(X > x) = P(X > u) P(X > x / X > u). \quad (1.74)$$

L'estimation se fait de la façon suivante :

$\bar{F}(x)$ est estimée avec la probabilité de dépassement empirique N_u/n .

$$\hat{\bar{F}}(u) = \bar{F}_n(u) = \frac{1}{n} \sum_{i=1}^n I_{\{X > u\}} = N_u/n, u < x_F. \quad (1.75)$$

La queue conditionnelle F_u de F est estimée par

$$\hat{F}(x - u) = 1 - G_{\hat{\xi}_u, \hat{\sigma}_u}(x - u) = \left(1 + \hat{\xi}_u \frac{x - u}{\hat{\sigma}_u}\right)^{-\frac{1}{\hat{\xi}_u}}, u < x < x_F. \quad (1.76)$$

L'estimateur de la queue de la distribution est donc :

$$\hat{F}(u) = \frac{N_u}{n} \left(1 + \hat{\xi}_u \frac{x - u}{\hat{\sigma}_u}\right)^{-\frac{1}{\hat{\xi}_u}}, u < x < x_F. \quad (1.77)$$

1.9 Estimation des quantiles

1.9.1 Quantile d'ordre p

Définition 1.6 Soit X une variable aléatoire d'une distribution F . On suppose que F est continue, Le quantile d'ordre p vérifie :

$$F(x_p) = p(X < x_p) = p. \quad \text{où } p \in [0, 1]. \quad (1.78)$$

x_p : est un quantile d'ordre p . Par définition le quantile d'ordre p c'est la fonction inverse de F c'est-à-dire $x_p = F^{-1}(p)$.

1.9.2 Fonction des quantiles

Si on a F est une fonction continue et monotone alors F est bijective donc F^{-1} existe.

Définition 1.7 Soit X une variable aléatoire et X_1, X_2, \dots, X_n un échantillon de X .

$$Q(p) = F^{-1}(p) = \inf \{x; F(x) \geq p\} \quad \text{avec } p \in [0, 1]. \quad (1.79)$$

"la fonction des quantiles s'appelle la fonction inverse générale".

1.9.3 Estimation empirique

L'estimation quantile joue un rôle important dans le contexte de la gestion des risques, où il est crucial d'évaluer correctement le risque d'une grosse perte qui survient très rarement. La principale difficulté de cette estimation est due au fait que lorsque p est très petit, le point x_p est au-delà de la plage de l'échantillon (X_1, \dots, X_n) tiré d'un cdf F inconnu.

Comme nous utilisons la théorie asymptotique, p doit dépendre de la taille de l'échantillon n , c'est-à-dire, $p = p_n$.

Deux cas sont possibles pour x , à l'intérieur et à l'extérieur de l'échantillon.

- si $p_n \rightarrow 0$ avec $np_n \rightarrow c \in [1, \infty]$ comme $n \rightarrow \infty$, le $(1-p)$ -quantile est dans l'échantillon, et
- si $p_n \rightarrow 0$ avec $np_n \rightarrow c \in [0, 1]$ comme $n \rightarrow \infty$, le $(1-p)$ -quantile est en dehors de l'échantillon.

En d'autres termes, l'estimation intra-écha

ntillon est possible jusqu'au quantième $(1/n)$ alors que pour $p < 1/n$, les estimations quantiles sont au-delà de la plage des données. Ce dernier cas est le plus pertinent pour les applications réelles. Pour la première situation, nous avons $Q_n(s) = X_{n-i+1,n}$, puis avec $s = 1 - p = 1 - (i-1)/n$ pour $i = 2, \dots, n$, on a

$$Q_n \left(1 - \frac{(i-1)}{n} \right) = X_{n-i+1,n}, i = 2, \dots, n. \quad (1.80)$$

Ainsi, $X_{n-i+1,n}$ semble être un estimateur naturel pour le $(1 - \frac{(i-1)}{n})$ -quantile.

Dans le second cas, nous devons déduire au-delà des limites de l'échantillon en extrapolant à partir des quantiles intermédiaires. Évidemment, cela ne peut pas être fait sans une sorte d'information sur les queues est alors nécessaire.

1.9.4 Intervalles de confiance des quantiles empirique

Proposition 1.1 (*Intervalles de confiance 1*) : soit $p \in]0, 1[$ supposons que la loi de X_1 possède une densité f continue en x_p et telle que $f(x_p) > 0$.

On suppose de plus que $k(n) = np + o(\sqrt{n})$ c-à-d

$$\lim_{n \rightarrow \infty} \frac{k(n)}{n} = p. \quad (1.81)$$

Alors, on a la convergence en loi suivante :

$$\frac{\sqrt{n} (X_{(k(n),n)} - x_p) f(x_p)}{\sqrt{p(1-p)}} \xrightarrow{n \rightarrow \infty} \mathcal{N}(0, 1). \quad (1.82)$$

Alors soit $\alpha > 0$, pour calculer l'intervalle de confiance $1 - \alpha$ on procède de la façon suivante, on veut :

$$1 - \alpha = p(|N(0, 1)|). \quad (1.83)$$

Donc l'intervalle de confiance de x_p , de niveau asymptotique $1 - \alpha$ est l'intervalle aléatoire :

$$\left[X_{(k(n),n)} \pm \delta \frac{\sqrt{p(1-p)}}{f(X_{(k(n),n)}\sqrt{n})} \right]. \quad (1.84)$$

ou δ est le quantile d'ordre $1 - \frac{\alpha}{2}$ de la loi $\mathcal{N}(0, 1)$.

Proposition 1.2 (*Intervalles de confiance 2*) : Soit $p \in]0, 1[$, soit δ est le quantile d'ordre $1 - \frac{\alpha}{2}$ de la loi $\mathcal{N}(0, 1)$. On considère par la suite que, les entiers $i_n = \left\lfloor np - \sqrt{n\delta}\sqrt{p(1-p)} \right\rfloor$ et $j_n = \left\lfloor np + \sqrt{n\delta}\sqrt{p(1-p)} \right\rfloor$. Pour n assez grand les entiers i_n et j_n sont compris entre 1 et n . De plus l'intervalle aléatoire :

$$[X_{(i(n),n)}, X_{(j(n),n)}].$$

est un intervalle de confiance pour x_p de niveau asymptotique $1 - \alpha$.

En d'autres termes, une bonne estimation de l'indice de queue est essentielle au processus d'estimation quantile extrême.

Puisque l'estimation des quantiles élevés est directement liée à l'estimation de l'IVE, on s'attendrait à trouver, dans la littérature, autant d'estimateurs quantiles que d'estimateurs de l'indice de queue.

1.9.5 Approche POT

La méthode POT s'appuie sur le théorème de Balkema-de Haan-Pickands pour estimer x_p et cet estimateur est obtenu en inversant $\hat{F}(x)$ dans l'équation 1.77.

L'estimateur obtenu par cette méthode s'écrit sous la forme :

$$\hat{x}_p = u + \frac{\hat{\sigma}_u}{\hat{\xi}_u} \left[\left(\frac{N_u}{np} \right)^{\hat{\xi}} - 1 \right]. \quad (1.85)$$

où N_u désigne le nombre d'excès au-delà du seuil u , $\hat{\sigma}_u$ et $\hat{\xi}_u$ sont des estimateurs des paramètres de la loi GPD.

Quand $\xi > 0$, le point terminal est fini et il est estimé par :

$$\hat{x}_F = u - \frac{\hat{\sigma}_u}{\hat{\xi}_u}. \quad (1.86)$$

En pratique, u est choisi égal à l'une des statistiques d'ordre et en Prenant $u = X_{n-k,n}$ la $(k+1)^{ieme}$ observation, donne $N_u = k$.

Les estimateurs résultants des paramètres ξ et σ sont respectivement dénotés par $\hat{\xi}^{(POT)}$ et $\hat{\sigma}^{(POT)}$. Dans ce cas, l'estimateur des quantiles est de la forme suivante :

$$\hat{x}_p^{(POT)} = X_{n-k,n} + \frac{\hat{\sigma}^{(POT)}}{\hat{\xi}^{(POT)}} \left(\left(\frac{k}{np} \right)^{\hat{\xi}^{(POT)}} - 1 \right) \quad \text{pour } p < \frac{k}{n}, \quad (1.87)$$

le point terminal est estimé par :

$$\hat{x}_F = X_{n-k,n} - \frac{\hat{\sigma}^{(POT)}}{\hat{\xi}^{(POT)}}. \quad (1.88)$$

1.9.6 Estimation de Weismann

Dans l'analyse des extrêmes, une exigence typique consiste à trouver les grandes valeurs de sorte que les probabilités de les dépasser soient très faibles (c-à-d proches de zéro). Ces quantités sont appelées quantiles extrêmes car l'ordre de ces quantiles tend vers zéro lorsque la taille n de l'échantillon tend vers l'infini. Nous nous sommes donc intéressés à l'estimation des quantiles extrêmes d'ordre p .

1.9.7 Approche basée sur un estimateur d'indices positifs ($\gamma > 0$)

On estime le quantile extrême, à partir de l'estimateur de Hill de γ en extrapolant le long de la droite du « Paréto quantile plot » d'équation :

$$y = \log X_{n-kn,n} + \hat{\gamma}_{k_n}^H \left(x + \log \frac{k_n + 1}{n + 1} \right). \quad (1.89)$$

de point d'origine $(\log((n+1)/(kn+1)), \log X_{n-kn,n})$, Ceci conduit en prenant $x = -\log \alpha_n$ à l'estimateur le plus connu d'un quantile extrême :

$$q(\alpha_n) = F^{*-}(1 - \alpha_n) = Q(1 - \alpha_n) = U(1/\alpha_n). \quad (1.90)$$

introduit par *Weissman* (1978) :

$$\hat{q}_{k_n}^H(\alpha) = X_{n-k_n,n} \left(\frac{k_n + 1}{(n + 1)\alpha} \right)^{\hat{\gamma}_{k_n}^+}. \quad (1.91)$$

et $\hat{\gamma}_{k_n}^+ = \hat{\gamma}_{k_n}^H$.

1.9.8 Approche basée sur un estimateur d'indices quelconques

En utilisant l'estimateur des moments on a proposé d'estimer $q(\alpha_n) = U(1/\alpha_n)$ pour $\alpha_n \rightarrow 0$ comme suit :

$$\hat{q}_{k_n}^M(\alpha_n) = X_{n-K_n,n} + \hat{a} \left(\frac{k_n}{n} \right) \frac{\left(\frac{n}{k_n} \alpha_n \right)^{-\hat{q}_{k_n}^M} - 1}{\hat{q}_{k_n}^M}. \quad (1.92)$$

avec $\hat{a}(k_n/n) = X_{n-K_n,n} \hat{\gamma}_{k_n}^H \max(1 - \hat{q}_{k_n}^M, 1)$.

Chapitre 2

Mesures des risques

2.1 Introduction

La maîtrise des risques est un sujet de préoccupation aussi bien en hydrologie qu'en finance et en actuariat. Se prémunir contre les risques en finance et en actuariat est primordial afin d'anticiper respectivement des crises financières ou d'importants sinistres coûteux pour les assurances. En hydrologie, l'étude des pluies et débits extrêmes est d'un grand intérêt afin de se prémunir contre les catastrophes telles que des inondations et ainsi permettre l'aménagement du territoire.

Pour cela, il existe des outils permettant de quantifier et de prédire le risque : les mesures de risque. Elles permettent d'évaluer un niveau de dangerosité d'un risque, mais également de comparer différents risques entre eux. La quantification, la prévention, la comparaison et la prédiction du risque sont des éléments essentiels de notre société.

2.2 Risque

Le risque est un mot avec des implications diverses. Certaines personnes définissent le risque différemment des neutres. Ces désaccords cause de graves confusions dans le domaine de l'évaluation des risques et de leur gestion.

Définition 2.1 *Fixons un espace mesurable (Ω, F) . Un risque est une v.a. définie sur (Ω, F) désigné par X . Il représente la perte nette finale d'une position (éventualité) actuellement détenue. Lorsque $X > 0$, on dit qu'il y a une perte, la classe de ces v.a. sur (Ω, F) est noté par ϕ .*

2.3 Mesure de risque

Le principal outil théorique pour calculer le besoin en fonds propres est défini sous le vocable « mesures de risque ». Certaines de ces mesures de risque sont manipulées depuis fort longtemps par les actuaires, plus particulièrement dans le domaine de la tarification. Les mesures de risque utilisées pour le besoin en fonds propres ont donné lieu à de nombreux travaux actuariels ces dernières années. De manière générale, elles visent à fixer un niveau de capital pour un portefeuille de risque donné, et mesurent le risque en un ou plusieurs nombres. Dans ce qui suit-on donne la définition d'une mesure de risque et les propriétés associées qui peuvent être recherchées pour évaluer le besoin en fonds propres, puis on présentera les principales mesures de risque permettant de fixer un niveau optimal de fonds propres.

2.4 Définition Formelle d'une Mesure de Risque

Une mesure de risque constitue un outil important, n'importe quel établissement financière peut faire appel à diverses techniques pour mesurer et contrôler le risque qu'elle assume dans ses diverses activités. Pour chaque type d'activité, elle peut déterminer les mesures du risque les plus convenables.

Dans la suite, la variable aléatoire d'intérêt dont on souhaite quantifier le risque sera notée X . Elle désignera également la variable aléatoire du montant de perte. Donnons à présent la définition précise d'une mesure de risque.

Définition 2.2 *On appelle mesure de risque une fonction \mathcal{R} associant à X une valeur positive ou nulle telle que :*

$$\begin{aligned} \mathcal{R} & : X \rightarrow \mathbb{R}^+ \\ X & \rightarrow \mathcal{R}(X) \end{aligned} \tag{2.1}$$

De grandes valeurs de $\mathcal{R}(X)$ indiqueront que X est "dangereux". $\mathcal{R}(X)$ peut être vu comme le capital à détenir pour faire face aux pertes X .

Censée à quantifier le risque par la v.a. X définie sur (Ω, \mathcal{F}, P) .

Une mesure de risque définie en tant qu'une trace d'une classe des v.a. définies sur un espace mesurable (représentant les risques actuels) aux nombre réelles. Un bon sens de la valeur numérique d'une mesure de risque est donné par la certitude équivalente pour une mesure de risque donnée et pour une v.a. donnée, l'équivalent de certitude est le nombre réel qui est tout aussi risque.

En termes économiques, nous interprétons $\mathcal{R}(x)$ comme le montant de capital qui devrait être ajouté en tant qu'amortisseur pour un portefeuille

avec une perte donnée par X , de sorte que le portefeuille devient acceptable à un contrôleur externe ou interne de risque (elle doit saisir le risque sur les préférences du décideur).

Définition 2.3 *On dit que est une mesure de risque monétaire si :*

- \mathcal{R} est croissant : $X \geq Y \implies \mathcal{R}(X) \geq \mathcal{R}(Y)$.
- invariante par translation : $\mathcal{R}(X + m) = \mathcal{R}(X) + m$ pour tout $m \in \mathbb{R}$.

Il existe de nombreuses façons de mesurer le risque, la mesure la plus répandue est la Value-at-Risk (VaR) établie par JP Morgan (1994). VaR comme mesure de risque est profondément dénigrée et souffre des contradictions en raison de ne pas être une mesure cohérente de risque selon Artzner et al. (1999).

2.4.1 VaR(Value-at-Risque)

Depuis le milieu des années 90, le pivot du processus de mesure du risque de marché, est sans aucun doute, la Value-at-Risque (**VaR**). Le **VaR** constitue un repère de mesure de risque notamment pour la mesure du capital économique et réglementaire. Certaines directives s'appuient par exemple sur cette mesure de risque, qui est adoptée par le Comité de Bâle comme mesure de référence.

Il y a de vaste littérature consacrée aux méthodes d'évaluation et d'utilité de cette quantité dans les applications; la littérature peut être également vaste également consacrée aux critiques de **VaR**. Pour une vue d'ensemble générale et récente au sujet de **VaR**.

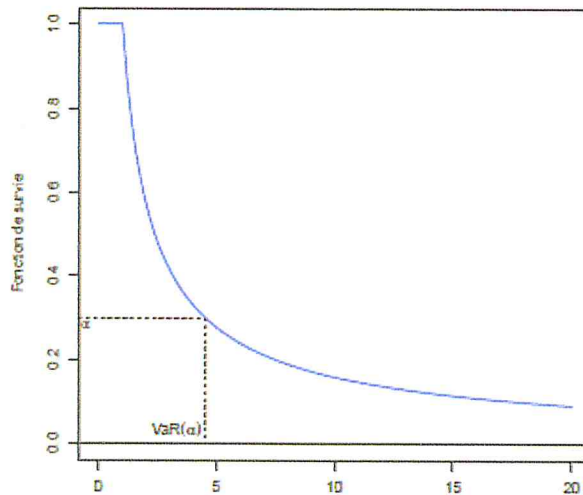
Définition 2.4 (*La VaR*)

*En termes mathématiques la mesure de **VaR** n'est rien d'autre que $(1 - \alpha)$ niveau de quantile de la fonction de répartition F , En effet pour un risque X sur une période donnée $[0, T]$ et $0 < \alpha < 1$, la **VaR** est :*

$$\begin{aligned} VaR(X) &= \alpha \times 100\% \text{ quantile} & (2.2) \\ &= F^{\leftarrow}(1 - \alpha) \\ &= \sup \{x \in \mathbb{R} \mid \Pr(X \geq x) > 1 - \alpha\} \end{aligned}$$

L'idée derrière le concept de la **VaR** est la suivante : on fixe un seuil probabiliste et on calcule une valeur **VaR**(α) qui sera telle que la probabilité que la catastrophe survienne soit plus petite que α . On peut voir la **VaR**(α) comme le montant d'extra capital dont une entreprise a besoin afin de réduire à α la probabilité de faire faillite.

En termes statistiques, la **VaR** de niveau $\alpha \in]0, 1[$ correspond au quantile de niveau $1 - \alpha$ de la fonction de répartition de la loi des pertes.



Graphique d'une fonction de survie modélisant la distribution des pertes. On a représenté une Value-at-Risk d'un certain niveau α .

Cette mesure de risque présente cependant plusieurs défauts.

En premier lieu, elle ne donne qu'une information ponctuelle au quantile $q(\alpha)$ et aucune information au-delà de ce point ; elle ne prend ainsi pas en compte l'importance du sinistre lorsqu'il survient mais seulement sa fréquence. De ce fait, comme elle ne prend pas en considération ce que serait la perte au-delà de ce quantile, elle ne fait donc pas la différence entre une catastrophe qui coûtera 1 million et celle qui coûtera 1 milliard d'euros. La **VaR** sous-estime les pertes.

Un des principaux reproches fait à la **VaR** est que des variables aléatoires à queues légères et à queues lourdes peuvent avoir la même $VaR(\alpha)$.

“Value at Risk is to finance what body temperature is to a patient, an indication of bad health but not an instrument telling us what is wrong and for less a clue on how to get the patient system healthy again”.

Paul Embrechts

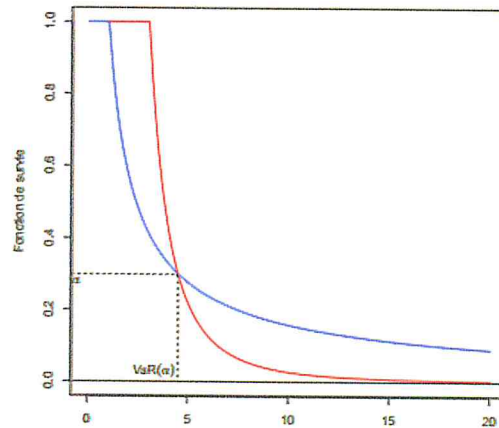


Illustration graphique de deux fonctions de survies distinctes ayant la mme $VaR(\alpha)$.

Limites de la VaR

VaR en dépit de son universalité, plusieurs auteurs ont précisé ses déficiences. **VaR** comme mesure de risque est fortement critiquée :

- **VaR** est difficile à optimiser. En outre, il est inadéquat d'employer la VaR dans la pratique en raison de sa non convexité elle possède beaucoup de minimum locaux. Ce qui peut avoir beaucoup d'extrémités locales, qui mènent au rang instable de risque.
- VaR est un modèle de mesure de risque dépendante parce que, par définition, elle dépend de la référence de probabilité initiale.
- VaR n'est pas une mesure de risque cohérente, car elle n'obéit pas l'axiome de sous additivité proposée par Artzner et al. (1999) et Acerbi et Tasche (2002) ce qui cause des contradictions, par conséquent il peut avoir un plus gros risque surgissant de la diversification. Ce résultat implique que l'agrégation des portefeuilles peut mener à une augmentation de risque.
- La VaR ne mesure pas les pertes excédant le seuil de confiance, elle ignore toutes les informations concernant la queue de la distribution, et les propriétés statistiques de la perte significative au delà du seuil, par exemple elle ne s'inquiète de pas au risque de queue, ne nous indique rien au sujet de la taille potentielle de la perte qui l'excède, ne tient pas compte de la sévérité d'un événement encouru de dommages.
- Pour les portefeuilles des titres, les modèles de VaR utilisés sont habi-

tuellement basés sur l'hypothèse (souvent implicite) que les rentabilités suivent une distribution normale. Or, la normalité ne peut pas capturer l'épaisseur de la queue de distribution. Sous des fluctuations extrêmes des prix des actifs ou une structure extrême de la dépendance des actifs, la VaR peut sous-estimer de manière significative le risque. Yamai et Yoshida (2002) prouvent que la VaR n'a pas de risque de queue quand les distributions sont de type elliptique. La VaR est sous additive seulement si la distribution de la variable financière est normale (ou de manière générale elliptique), ce qui n'est pas souvent le cas, même pour les rentabilités des actifs sur le marché.

2.5 Mesure de risque Cohérente

Le risque d'un portefeuille mesure par VaR peut être supérieur à la somme des VaR de ses composantes. En basons à ce type de critique Artzner et al. (1999) ont introduit le concept de mesure de risque cohérente.

Proposition 2.1 (*Caractérisations des mesures du risque cohérentes*).

Si \mathcal{R} est une mesure du risque cohérente, alors il existe un ensemble de mesures de probabilité \mathcal{Q} tel que :

$$\mathcal{R}(X) = \sup_{\mathbb{Q} \in \mathcal{Q}} \{\mathbb{E}_{\mathbb{Q}}(X)\} \quad (2.3)$$

Selon Artzner et al. (1999) une mesure de risque est cohérente si elle adhère aux axiomes que nous énumérons maintenant :

Définition 2.5 (*Mesure Cohérente*) :

Une mesure du risque $\mathcal{R} : \mathcal{F} \rightarrow \mathbb{R}$ est cohérente si elle vérifiée ($X, Y \in \mathcal{F}$) :

– *La sous additivité :*

$$\mathcal{R}(X + Y) \leq \mathcal{R}(X) + \mathcal{R}(Y), \quad (2.4)$$

plus généralement

$$\mathcal{R}(nX) = \mathcal{R}(X + \dots + X) \leq n\mathcal{R}(X), n = 1, 2, \dots \quad (2.5)$$

– *L'homogénéité positive :*

$$\mathcal{R}(\lambda X) \leq \lambda \mathcal{R}(X) \text{ pour tout } \lambda \in \mathbb{R}_+, \quad (2.6a)$$

plus généralement

$$\mathcal{R}(nX) \leq n\mathcal{R}(X), n = 1, 2, \dots \quad (2.7a)$$

De même qu'une fusion ne crée pas risque supplémentaire ($\mathcal{R}(\lambda X) \leq \lambda \mathcal{R}(X)$), une fusion sans diversification ne réduit pas le besoin global en capital.

- La monotonie :

$$X \geq Y \Rightarrow \mathcal{R}(X) \geq \mathcal{R}(Y). \quad (2.8)$$

Si les pertes encourues avec le risque X sont toujours supérieures à celles obtenues avec Y , le besoin en capital pour X doit être supérieur à celui pour Y .

- L'invariance par translation :

$$\mathcal{R}(X + c) \leq \mathcal{R}(X) + c \text{ pour tout } c \in \mathbb{R}. \quad (2.9)$$

Spécialement, nous avons $\mathcal{R}(X + \mathcal{R}(X)) = \mathcal{R}(X) - \mathcal{R}(X) = 0$ c'est-à-dire, en ajoutant $\mathcal{R}(X)$ à la position initiale X , nous obtenons une position "neutre". On notera en particulier que $\mathcal{R}(X) = -\alpha$ avec la convention que $\mathcal{R}(0) = 0$.

- Convexe :

La convexité implique que la diversification n'augmente pas le risque, car la valeur de risque du portefeuille diversifié $\lambda X + (1 - \lambda)Y$ est inférieure ou égale à la moyenne re-pondérer des différentes valeurs de risque.

\mathcal{R} est dite convexe si pour tout $\lambda \in [0, 1]$

$$\mathcal{R}(\lambda X + (1 - \lambda)Y) \leq \lambda \mathcal{R}(X) + (1 - \lambda) \mathcal{R}(Y). \quad (2.10)$$

Une mesure de risque \mathcal{R} convexe et positivement homogène est dite cohérente.

Définition 2.6 Une mesure de risque est dite :

1. monétaire si elle est monotone et invariante par translation,
2. convexe si elle est monétaire et convexe,
3. cohérente si elle est monétaire, homogène et sous-additive.

Corollaire 2.1 Si \mathcal{R} est une mesure de risque monétaire, homogène et normalisée à 0, alors la convexité et la sous-additivité sont des notions équivalentes.

Définition 2.7 Si \mathcal{R} est une mesure de risque, on définit la région de risque acceptable pour la mesure \mathcal{R} comme

$$\mathcal{A}\{X, \mathcal{R}(X) \leq 0\} \quad (2.11)$$

Réciproquement, si \mathcal{A} est une région de risque acceptable, la mesure de risque induite \mathcal{R} est

$$\mathcal{R}(X) = \inf \{m, X - m \in \mathcal{A}\} \quad (2.12)$$

Proposition 2.2 Si \mathcal{R} est une mesure de risque monétaire alors

- \mathcal{R} est convexe si et seulement si \mathcal{A} est convexe,
- \mathcal{R} est positivement homogène si et seulement si \mathcal{A} est un cône.

Preuve: Pour le premier point, si \mathcal{R} est convexe, alors \mathcal{A} est convexe. Réciproquement, si \mathcal{A} est convexe, soient X_1, X_2, m_1 et m_2 tels que $X_i - m_i \in \mathcal{A}$, pour $i = 1, 2$. Par hypothèse, pour tout λ , $\lambda[X_1 - m_1] + (1 - \lambda)[X_2 - m_2] \in \mathcal{A}$, c'est-à-dire que $\mathcal{R}(\lambda[X_1 - m_1] + (1 - \lambda)[X_2 - m_2]) \leq 0$, soit, par la propriété d'invariance par translation, et par monotonie,

$$\mathcal{R}(X_1 + (1 - \lambda)X_2) \leq \lambda m_1 + (1 - \lambda)m_2 \quad (2.13)$$

et ceci pour tout m_1 et m_2 . Il suffit de le faire pour $m_i = \mathcal{R}(X_i)$. Pour le second point, si \mathcal{A} est un cône, alors pour tout $X - m \in \mathcal{A}$, $\lambda(X - m) \in \mathcal{A}$, pour tout $\lambda > 0$. Donc $\mathcal{R}(\lambda X - \lambda m) \leq 0$, d'où $\mathcal{R}(\lambda X) \leq \lambda m$. Si $m = \mathcal{R}(X)$, on en déduit que $\mathcal{R}(\lambda X) \leq \lambda \mathcal{R}(X)$. Et si $X - m \notin \mathcal{A}$; alors $\lambda(X - m) \notin \mathcal{A}$, et $\mathcal{R}(\lambda X) > \lambda m$. On fait alors tendre m vers $\mathcal{R}(X)$ pour avoir le résultat souhaité. On dispose du théorème de représentation suivant ■

Théorème 2.1 \mathcal{R} est une mesure de risque monétaire convexe si et seulement si pour tout X bornée ($X \in L^\infty$)

$$\mathcal{R}(X) = \max_{\mathbb{Q} \in \mathcal{M}} \{\mathbb{E}_{\mathbb{Q}} - \alpha(\mathbb{Q})\}, \quad (2.14)$$

où \mathcal{M} est l'ensemble des mesures additives et normalisées à 1, et

$$\alpha(\mathbb{Q}) = \sup_{X \in \mathcal{A}} \{\mathbb{E}_{\mathbb{Q}}(X)\} \quad (2.15)$$

où \mathcal{A} est l'ensemble d'acceptation associé à \mathcal{R} . On notera que \mathcal{M} contient plus que des mesures de probabilité.

Proposition 2.3 Si \mathcal{R} est une mesure de risque convexe, les trois propriétés suivantes sont équivalentes

1. \mathcal{R} est fortement cohérente,
2. \mathcal{R} est additive pour des risques comonotones,
3. \mathcal{R} est une mesure de corrélation maximale.

Proposition 2.4 Une mesure de risque cohérente \mathcal{R} est additive pour des risques comonotones si et seulement s'il existe une fonction décroissante positive g sur $[0; 1]$ telle que

$$\mathcal{R}(X) = \int_0^1 g(t) F_X^{-1}(1-t) dt \quad (2.16)$$

où $F_X(x) = P(X \leq x)$.

2.5.1 La Tail Value-at-Risk

Si la Value-at-Risk s'intéressait aux probabilités d'événements rares et extrêmes, avec la Tail-Value-at-Risk, nous nous intéresserons à ce qui se passe (en moyenne) lorsque ces événements extrêmes surviennent.

Définition 2.8 La Tail Value-at-Risk au niveau α , notée $TVaR(X; \alpha)$ est définie par :

$$TVaR(X; \alpha) = \frac{1}{1-\alpha} \int_\alpha^1 VaR(X; t) dt. \quad (2.17)$$

Autrement dit, la $TVaR$ apparaît comme la moyenne des VaR de niveau supérieur à α .

Remarque 2.1 Il est intéressant de noter qu'il existe une fonction de répartition \tilde{F}_X , appelée transformée de Hardy-Littlewood de F_X , telle que pour tout α :

$$\tilde{F}_X^{-1}(\alpha) = TVaR(X; \alpha). \quad (2.18)$$

Si on note \tilde{X} une variable aléatoire de fonction de répartition \tilde{F}_X , on a :

$$TVaR(X; \alpha) = VaR(\tilde{X}; \alpha). \quad (2.19)$$

La $TVaR$ d'un risque X est donc la VaR de la transformée de Hardy-Littlewood de X .

Notons que, $TVaR[X; 0] = \mathbb{E}[X]$. Et comme

$$TVaR[X; \alpha] = \frac{1}{1-\alpha} \left\{ \mathbb{E}[X] - \int_0^\alpha VaR[X; \xi] d\xi \right\}. \quad (2.20)$$

on en déduit que la Tail Value-at-Risk est une fonction croissante du niveau α .

En effet :

$$\frac{d}{d\alpha} TVaR(X; \alpha) = \frac{TVaR(X; \alpha)}{1 - \alpha} - \frac{VaR(X; \alpha)}{1 - \alpha}. \quad (2.21)$$

Et comme $\alpha \mapsto VaR[X; \alpha]$ est une fonction croissante :

$$TVaR[X; \alpha] = \frac{1}{1 - \alpha} \int_{\alpha}^1 \underbrace{VaR[X; t]}_{\geq VaR[X; \alpha]} dt \geq VaR[X; \alpha], \quad (2.22)$$

on en déduit que $\frac{d}{d\alpha} TVaR[X; \alpha] \geq 0$, et :

$$TVaR[X; \alpha] \geq TVaR[X; 0] = \mathbb{E}[X]. \quad (2.23)$$

Définition 2.9 La *Conditional Tail Expectation* au niveau de probabilité α , notée $CTE[X; \alpha]$, représente la perte attendue sachant que la VaR au niveau α est dépassée :

$$CTE[X; \alpha] = \mathbb{E}[X \mid X > VaR[X; \alpha]]. \quad (2.24)$$

Il s'agit donc de la version mathématique du concept intuitif de « perte moyenne dans les pires $1 - \alpha\%$ des cas ».

2.5.2 Conditionnelle VaR

Définition 2.10 La *Conditional-VaR* au niveau α de probabilité, notée $CVaR[X; \alpha]$, est la valeur moyenne des pertes qui excèdent la VaR , correspondant à l'excédent moyen au-delà de la VaR ou encore :

$$\begin{aligned} CVaR[X; \alpha] &= \mathbb{E}[X - VaR[X; \alpha] \mid X > [X; \alpha]] \\ &= e_X(VaR[X; \alpha]) \\ &= CTE[X; \alpha] - VaR[X; \alpha]. \end{aligned} \quad (2.25)$$

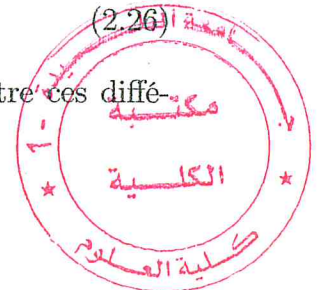
où

$$e_X(s) = \mathbb{E}[X \mid X > s].$$

Définition 2.11 L'*Expected shortfall* au niveau de probabilité α , notée $ES[X; \alpha]$, n'est autre que la prime stop-loss dont la rétention est fixée à $VaR[X; \alpha]$:

$$ES[X; \alpha] = \mathbb{E}[(X - VaR[X; \alpha])_+]. \quad (2.26)$$

Le résultat suivant détaille les liens qui peuvent exister entre ces différentes mesures de risque.



Proposition 2.5 *Quel que soit le niveau de probabilité $\alpha \in (0, 1)$, les identités suivantes sont vérifiées :*

$$TVaR[X; \alpha] = VaR[X; \alpha] + \frac{1}{1 - \alpha} ES[X; \alpha]. \quad (2.27)$$

$$CTE[X; \alpha] = VaR[X; \alpha] + \frac{1}{\bar{F}_X(VaR[X; \alpha])} ES[X; \alpha]. \quad (2.28)$$

Preuve: L'expression (2.27) provient de

$$\begin{aligned} ES[X; \alpha] &= \int_0^1 (VaR[X; \xi] - VaR[X; \alpha])_+ d\xi \\ &= \int_\alpha^1 VaR[X; \xi] d\xi - VaR[X; \alpha] (1 - \alpha). \end{aligned} \quad (2.29)$$

L'expression (2.28) se déduit de

$$ES[X; \alpha] = \mathbb{E}[X - VaR[X; \alpha] \mid X > VaR[X; \alpha]] \bar{F}_X(VaR[X; \alpha]). \quad (2.30)$$

■

La Propriété (2.20) permet d'établir le résultat suivant (qui fournit l'interprétation usuelle de la TVaR).

Proposition 2.6 *La CTE et la TVaR coïncident pour des risques dont la fonction de répartition est continue :*

$$CTE[X; \alpha] = TVaR[X; \alpha], \text{ pour } \alpha \in (0, 1). \quad (2.31)$$

Preuve: Il suffit de noter que si F_X est continue alors

$$F_X(VaR[X; \alpha]) = \alpha \text{ et } \bar{F}_X(VaR[X; \alpha]) = 1 - \alpha \quad (2.32)$$

■

de sorte que 2.27-2.28 donnent le résultat annoncé.

La TVaR est invariante par translation et homogène. En effet, comme la VaR est invariante par translation,

$$\begin{aligned} TVaR[X + c; \alpha] &= \frac{1}{1 - \alpha} \int_\alpha^1 VaR[X + c, \xi] d\xi \\ &= \frac{1}{1 - \alpha} \int_\alpha^1 (VaR[X; \xi] + c) d\xi \\ &= TVaR[X; \alpha] + c. \end{aligned} \quad (2.33)$$

De la même manière, l'homogénéité de la VaR garantit l'homogénéité de la TVaR.

Proposition 2.7 Soient le risque X et le niveau de perte x tels que $\bar{F}_X(x) > 0$. Quel que soit l'événement aléatoire E tel que $\mathbb{P}[E] = \bar{F}_X(x)$, on a $\mathbb{E}[X | E] < \mathbb{E}[X | X > x]$.

Preuve: Notons que l'on peut écrire

$$\mathbb{P}[E | X > x] = \frac{\mathbb{P}[X > x | E] \mathbb{P}[E]}{\mathbb{P}[X > x]} = \mathbb{P}[X > x | E]$$

Aussi

$$\begin{aligned} \mathbb{E}[X | X > x] &= x + \mathbb{E}[X - x | X > x, E] \mathbb{P}[E | X > x] \\ &\quad + \mathbb{E}[X - x | X > x, \bar{E}] \mathbb{P}[\bar{E} | X > x] \\ &\geq x + \mathbb{E}[X - x | X > x, E] \mathbb{P}(E | X > x) \\ &= x + \mathbb{E}[X - x | X > x; E] \mathbb{P}[X > x | E] \\ &\geq x + \mathbb{E}[X - x | X > x, E] \mathbb{P}[X > x | E] \\ &\quad + \mathbb{E}[X - x | X \leq x, E] \mathbb{P}[X \leq x | E] \\ &= \mathbb{E}[X | E]. \end{aligned}$$

ce qui achève la preuve. ■

Ce résultat est intuitif : $\mathbb{E}[X | E]$ représente le montant moyen de perte lorsqu'on sait que l'événement E est réalisé. La Propriété 2.28 nous apprend que la pire information qui soit est que X soit grande, correspondant à $X > x$. Cette proposition 2.7 permet d'affirmer que la **TVaR** est sous-additive lorsque les risques sont continus. En effet, dans ce cas, la **TVaR** et la **CTE** coïncident

et

$$\begin{aligned} TVaR[X + Y; \alpha] &= \mathbb{E}[X | X + Y > VaR[X + Y; \alpha]] + \\ &\quad \mathbb{E}[Y | X + Y > VaR[X + Y; \alpha]] \\ &\leq \mathbb{E}[X | X > VaR[X; \alpha]] + \\ &\quad \mathbb{E}[Y | Y > VaR[Y; \alpha]] \\ &= TVaR[X; \alpha] + TVaR[Y; \alpha]. \end{aligned}$$

De la même manière, la **TVaR** est monotone, puisque lorsque $\mathbb{P}[X \leq Y] = 1$

$$\begin{aligned} TVaR[Y; \alpha] &= \mathbb{E}[Y | Y > VaR[Y; \alpha]] \\ &\geq \mathbb{E}[Y | X > VaR[X; \alpha]] \\ &\geq \mathbb{E}[X | X > VaR[X; \alpha]] \\ &= TVaR[X; \alpha]. \end{aligned}$$

Proposition 2.8 *La TVaR est cohérente pour les risques continus, et coïncide alors avec la CTE.*

On peut montrer que la TVaR est la plus petite mesure de risque majorant la VaR qui soit cohérente .

2.5.3 Mesure de Wang

Nous appellerons désormais fonction de distorsion toute fonction croissante $g : [0, 1] \rightarrow [0, 1]$ telle que $g(0) = 0$ et $g(1) = 1$. Nous allons présenter dans cette section les mesures par distorsion (qui seront étudiées plus en détails dans la section suivante) dans le cas particulier proposé par Wang, à savoir le cas des variables aléatoires (strictement) positives, soit $F(0) = \mathbb{P}(X \leq 0) = 0$.

Définition 2.12 *La mesure de risque de Wang associée à la fonction de distorsion g , notée \mathcal{R}_g , est définie par :*

$$\mathcal{R}_g(X) = \int_0^\infty g(1 - F_X(x)) dx = \int_0^\infty g(\bar{F}_X(x)) dx. \quad (2.34)$$

Remarque 2.2 *On notera que la fonction de distorsion $g(q) = q$ (fonction identité) correspond à l'espérance mathématique $\mathbb{E}[X]$. De plus, si $g(q) \geq q$ quel que soit $q \in [0; 1]$, on a alors $\mathcal{R}_g[X] \geq \mathbb{E}[X]$, de sorte que les mesures de risque de Wang associées à de telles fonctions de distorsion contiennent un chargement de sécurité. Enfin, il est intéressant de remarquer que lorsque $g_1(q) \leq g_2(q)$ quel que soit $q \in [0; 1]$ nous avons alors $\mathcal{R}_{g_1}[X] \leq \mathcal{R}_{g_2}[X]$.*

En substituant $\int_0^{\bar{F}_X(x)} dg(\alpha)$ à $g(\bar{F}_X(x))$ dans ?? et en permutant les intégrales, on obtient le résultat suivant.

Proposition 2.9 *Quel que soit le risque X , la mesure de risque de Wang associée à la fonction de distorsion g peut s'écrire :*

$$\mathcal{R}_g[X] = \int_0^1 VaR[X; 1 - \alpha] dg(\alpha). \quad (2.35)$$

Ainsi, les mesures de risque de Wang sont des moyennes pondérées de VaR.

Exemple 2.1 En particulier, si nous considérons la fonction de distorsion $g : [0, 1] \rightarrow [0, 1]$ définie par :

$$g_\alpha(x) = 1[x \geq 1 - \alpha].$$

pour une valeur $\alpha \in (0, 1)$ fixée, il vient alors par 2.35 :

$$\mathcal{R}_{g_\alpha}[X] = \text{VaR}[X; \alpha].$$

qui montre que le **VaR** au niveau de probabilité α est une mesure de Wang particulière correspondant à une fonction de distorsion passant de 0 à 1 en $1 - \alpha$. Dans ce cas, g_α est une fonction de répartition, associée à une masse de Dirac en $1 - \alpha$.

Exemple 2.2 De la même manière, en repartant de 2.35 avec la fonction de distorsion :

$$g_\alpha(x) = \min \left\{ \frac{x}{1 - \alpha}, 1 \right\},$$

pour $\alpha \in (0; 1)$ fixé on obtient :

$$\mathcal{R}_{g_\alpha}[X] = \frac{1}{1 - \alpha} \int_0^{1 - \alpha} \text{VaR}[X; 1 - \xi] d\xi = \text{TVaR}[X; \alpha].$$

Dans ce cas également, g_α est une fonction de répartition, correspondant à une loi uniforme sur $[0, 1 - \alpha]$.

En revanche, ce n'est généralement pas le cas de l'expected shortfall (*ES*). Supposons par l'absurde que $ES[X; \alpha]$ puisse effectivement être représentée sous la forme d'une mesure de risque de Wang, associée à une fonction de distorsion g_α , $ES[X; \alpha] = \mathcal{R}_{g_\alpha}[X]$ quel que soit le risque X . Prenons $X \sim \mathcal{U}(0; 1)$. Dans ce cas

$$\begin{aligned} ES[X; \alpha] &= \int_\alpha^1 (1 - x) dx = \frac{1}{2} (1 - \alpha)^2 \\ &= \mathcal{R}_{g_\alpha}[X] = \int_0^1 g_\alpha(1 - x) dx. \end{aligned} \quad (2.36)$$

Considérons à présent $Y \sim \mathcal{B}(q)$ pour $0 < q \leq 1 - \alpha$. On obtient facilement

$$ES[Y; \alpha] = q = \mathcal{R}_{g_\alpha}[Y] = g_\alpha(q).$$

qui implique $g(q) = q$ pour $0 < q \leq 1 - \alpha$. En insérant cela dans 2.36 on obtient

$$\begin{aligned} \frac{1}{2} (1 - \alpha)^2 &= \int_0^\alpha g_\alpha(1 - x) dx + \int_\alpha^1 (1 - x) dx \\ &\geq \frac{1}{2} (1 - \alpha)^2 + \alpha (1 - \alpha), \end{aligned}$$

ce qui nous mène à une contradiction puisque $0 < \alpha < 1$.

Nous pouvons formellement établir par l'absurde que la *CTE* n'est pas une mesure de risque de Wang. Il suffit en fait de procéder comme pour l'*ES*. Par l'absurde, supposons qu'il existe une fonction de distorsion g_α telle que $CTE[X; \alpha] = \mathcal{R}_{g_\alpha}[X]$ quel que soit le risque X . Pour $X \sim \mathcal{U}(0, 1)$ il vient à partir de 2.28

$$CTE[X; \alpha] = \alpha + \frac{1}{2}(1 - \alpha) = \int_0^1 g_\alpha(1 - x) dx.$$

qui donne après simplification

$$\int_0^1 g_\alpha(x) dx = \frac{1}{2}(1 + \alpha). \quad (2.37)$$

En passant à $Y \sim \mathcal{B}(q)$ et invoquant à nouveau 2.28, on obtient $CTE[Y; \alpha] = 1$ quels que soient q et α . Dès lors,

$$g_\alpha(q) = \mathcal{R}_{g_\alpha}[Y] = CTE[Y; \alpha] = 1.$$

d'où l'on tire $g_\alpha(\cdot) \equiv 1$ sur $(0, 1]$, ce qui contredit l'équation 2.37 et achève de prouver que la *CTE* n'est pas une mesure de risque de Wang.

Exemple 2.3 *En prenant*

$$g(x) = 1 - (1 - x)^\xi, \xi \geq 1;$$

nous obtenons

$$\mathcal{R}_g[X] = \int_{x>0} \left(1 - \{F_X(x)\}^\xi\right) dx.$$

Si ξ est entier, $\mathcal{R}_g[X]$ peut être interprétée comme la valeur attendue du maximum $M_\xi = \max\{X_1, \dots, X_\xi\}$ d'un ensemble de ξ variables aléatoires indépendantes et de même loi que X . En effet, la fonction de queue de M_ξ est donnée par $\mathbb{P}[M_\xi > x] = 1 - \mathbb{P}[X_1 \leq x, \dots, X_\xi \leq x] = 1 - \{F_X(x)\}^\xi$, de sorte que $\mathcal{R}_g[X] = \mathbb{E}[M_\xi]$.

Exemple 2.4 *Considérons la fonction de distorsion*

$$g(x) = x^{1/\rho}, \rho \geq 1.$$

La mesure de risque PH , introduite dans [65] et [64] est donnée par

$$PH[X] = \mathcal{R}_g[X] = \int_{x \geq 0} \{\bar{F}_X(x)\}^{1/\rho} dx.$$

Notons que pour $\xi = 1$, $PH_1[X] = \mathbb{E}[X]$. Nous reviendrons par la suite sur cette mesure de risques, liée à des propriétés de hasard proportionnel dans les modèles de durées.

Les principales propriétés des mesures de risque de Wang sont les suivantes :

Proposition 2.10 *Les mesures de risque de Wang sont*

- homogènes,
- invariantes par translation,
- monotones.

Preuve: En travaillant à partir de la Propriété 2.28, et en tenant compte du fait que la **VaR** est invariante par translation, on obtient :

$$\begin{aligned} \mathcal{R}_g[X + c] &= \int_0^1 VaR[X + c; 1 - \alpha] dg(\alpha) \\ &= \int_0^1 (VaR[X; 1 - \alpha] + c) dg(\alpha) \\ &= \mathcal{R}_g[X] + c(g(1) - g(0)) = \mathcal{R}_g[X]. \end{aligned}$$

■

TAB. 2.1 – Principales mesures de risque par distorsion (où p est une constante comprise entre 0 et 1)

mesure de risque R	fonction de distorsion g
VaR	$g(x) = 1 [x \geq p]$
TVaR	$g(x) = \min \{x/p, 1\}$
PH	$g(x) = x^p$
Dual Power	$g(x) = 1 - (1 - x)^{1/p}$
Gini	$g(x) = (1 + p)x - px^2$
Transformation exponentielle	$g(x) = (1 - p^x) / (1 - p)$

Ceci établit l'homogénéité. De la même manière, on déduit l'homogénéité et la monotonie des mesures de Wang des propriétés correspondantes de la **VaR**.

Si la fonction de distorsion g est concave, la fonction $x \mapsto g(\overline{F}_X(x))$ est continue à droite et est donc la fonction de queue d'une certaine variable aléatoire. En pareille situation $\mathcal{R}_g[X]$ est effectivement une espérance mathématique (non pas de X mais d'une variable Y dont la fonction de queue est $y \mapsto g(\overline{F}_X(y))$).

Proposition 2.11 *Lorsque la fonction de distorsion est concave, la mesure de risque correspondante est sous-additive.*

Corollaire 2.2 *En combinant les Propositions 2.7 et 2.8, on constate que les mesures de risque de Wang correspondant à des fonctions de distorsion concaves sont cohérentes.*

Le résultat suivant montre que la **TVaR** est la plus petite mesure de risque de Wang associée à une fonction de distorsion concave et excédant la **VaR** de même niveau de probabilité. En d'autres termes, la **TVaR** est la plus petite mesure de risque cohérente excédant la **VaR**.

Proposition 2.12 *Quel que soit le niveau de probabilité $0 < \alpha < 1$ et le risque X , nous avons*

$$TVaR[X; \alpha] = \min \{ \mathcal{R}_g[X] \mid g \text{ est concave et } \mathcal{R}_g[X] \geq VaR[X; \alpha] \}. \quad (2.38)$$

Preuve: La **TVaR** est la mesure de risque de Wang associée à la fonction de distorsion concave $\min(\frac{x}{1-\alpha}, 1)$. De plus, $TVaR[X; \alpha] \geq VaR[X; \alpha]$. Ceci implique que

$$TVaR[X; \alpha] \geq \inf \{ \mathcal{R}_g[X] \mid g \text{ est concave et } \mathcal{R}_g[X] \geq VaR[X; \alpha] \}.$$

■

Afin d'établir l'inégalité contraire, considérons une fonction de distorsion concave g telle que

$$\mathcal{R}_g[Y] \geq VaR[Y; \alpha],$$

soit valable quelle que soit le risque Y . En prenant alors q tel que $1 - \alpha < q < 1$, et $Y_q \sim \mathcal{B}(q)$, il est facile de vérifier que $VaR[Y_q; \alpha] = 1$, et aussi $\mathcal{R}_g[Y_q] = g(q)$. Comme $g(x) \leq 1$, nous voyons que la condition $\mathcal{R}_g[Y_q] \geq VaR[Y_q; \alpha]$ peut se réécrire $g(q) = 1$. Cette condition doit être satisfaite

pour tout q tel que $1 - \alpha < q < 1$. Ceci signifie que g vaut 1 sur l'intervalle $(1 - \alpha, 1]$. Comme g est concave, ceci entraîne

$$g(x) \geq \min \left(\frac{x}{1 - \alpha}, 1 \right) \text{ pour } 0 < x < 1;$$

d'où l'on déduit

$$\mathcal{R}_g[X] \geq TVaR[X; \alpha].$$

Cette dernière inégalité est satisfaite quelle que soit la fonction de distorsion concave g telle que $\mathcal{R}_g[X] \geq VaR[X; \alpha]$. Ceci implique

$$TVaR[X; \alpha] \leq \inf \{ \mathcal{R}_g[X] \mid g \text{ est concave et } \mathcal{R}_g[X] \geq VaR[X; \alpha] \}. \quad (2.39)$$

et achève la démonstration.

2.6 Mesures de risque par distorsion

Les mesures de Wang sont en fait un cas très particulier des mesures de risque par distorsion, définies sur les variables aléatoires qui ne sont pas nécessairement positives. L'écriture étant toutefois plus simple sur la classe des mesures de Wang (et donc les démonstrations plus courtes) nous admettrons ici la plupart des propriétés énoncées.

Définition 2.13 On appelle mesure de risque par distorsion la quantité :

$$\mathcal{R}(X) = \int_0^1 F_X^{-1}(1 - u) dg(u).$$

où g est une fonction de répartition sur $[0, 1]$, appelée fonction de distorsion.

De même que pour les mesures de Wang, il est possible de réécrire cette expression sous la forme

Proposition 2.13 $\mathcal{R}(X)$ peut se réécrire :

$$\mathcal{R}(X) = \int_0^{+\infty} g(1 - F_X(x)) dx - \int_{-\infty}^0 [1 - g(1 - F_X(x))] dx. \quad (2.40)$$

Preuve:

$$\begin{aligned}\mathcal{R}(X) &= \int_0^1 F_X^{-1}(1-u) dg(u) \\ &= \int_0^{F_X(0)} F_X^{-1}(1-u) dg(u) + \int_{F_X(0)}^1 F_X^{-1}(1-u) dg(u).\end{aligned}$$

ce qui donne, en faisant une intégration par parties, et un changement de variable, $u = F_X(x)$, une réécrire la mesure de risque sous la forme

$$\mathcal{R}(X) = \int_0^{+\infty} g(1 - F_X(x)) dx - \int_{-\infty}^0 [1 - g(1 - F_X(x))] dx. \quad (2.41)$$

■

Remarque 2.3 Dans le cas particulier où g est la fonction de répartition de la loi uniforme sur $[0, 1]$, de telle sorte que g est la fonction identité sur $[0, 1]$:

$$\mathcal{R}(X) = \int_0^{+\infty} [1 - F_X(x)] dx - \int_{-\infty}^0 F_X(x) dx = \mathbb{E}(X). \quad (2.42)$$

Remarque 2.4 Soit \mathbb{Q} la mesure induite par la transformation g sur \mathbb{P} :

$$\mathbb{Q}([a, b]) = G \circ \mathbb{P}([a, b]).$$

La croissance de g sur $[0, 1]$ permet de construire une capacité.

Exemple 2.5 Considérons la fonction de répartition $g(x) = x^k$. On appellera mesure de risque à hasard proportionnel la mesure induite par cette transformation

$$\mathcal{R}_k(X) = \int_0^1 F_X^{-1}(1-u) k u^{k-1} du.$$

Lorsque $k < 1$, la fonction g est concave.

Remarque 2.5 On peut montrer que les mesures de distorsion de risque avec g concave vérifient la propriété de sous-additivité.

Remarque 2.6 Il existe dans la littérature un autre type de mesures de risque, appelées mesures spectrales. Une fonction spectrale (parfois aussi appelée fonction d'aversion pour le risque) est $\phi : [0, 1] \rightarrow \mathbb{R}_+$, décroissante, telle que $\int_0^1 \phi(t) dt = 1$. La mesure de risque spectrale induite est

$$\mathcal{R}(X) = \int_0^1 F_X^{-1}(t) \phi(t) dt.$$

Ces mesures de risque sont cohérentes. En fait, les mesures de distorsion de fonction de distorsion g concave sont des mesures spectrales, avec $\phi = g'$.

2.7 Moment conditionnel des queues (CTM)

Ce nouvel outil sera appelé le Conditional Tail Moment (noté **CTM**). Il est défini comme le moment d'ordre a , ($a \geq 1$) de la loi des pertes sachant que l'on se trouve au-dessus de la VaR de niveau .

On montrera que l'estimation du **CTM** permet d'estimer toutes les mesures de risque basées sur les moments conditionnels d'ordre arbitraire sachant que les pertes sont supérieures à la **VaR**. Il est par exemple clair que le Conditional Tail Moment d'ordre un est la Conditional Tail Expectation.

Notre seconde contribution sera d'estimer le **CTM** dans le cas de pertes extrêmes, en se servant pour cela de la théorie des valeurs extrêmes. Bien que de nombreux liens entre la théorie des valeurs extrêmes et les mesures de risque aient fait l'objet de recherches citons notamment Embrechts, Embrechts et al., McNeil et al. l'estimation de mesures de risque est habituellement réalisée dans la littérature pour des valeurs fixées de voir par exemple Deme et al. et Necir et al.

Le but premier de ce chapitre est d'unifier les définitions des mesures de risque énoncées juste auparavant. Pour cela, on introduit une nouvelle mesure de risque, le Conditional Tail Moment (noté **CTM**) définie par :

$$CTM_a(\alpha) := \mathbb{E}(Y^a \mid Y > VaR(\alpha));$$

où $a \geq 0$ est tel que le moment d'ordre a de Y existe. Il est facile de voir que toutes les mesures de risque précédentes de niveau peuvent être réécrites :

$$\Phi(VaR(\alpha), CTM_1(\alpha), CTM_2(\alpha), CTM_3(\alpha)),$$

où la fonction $\Phi : \mathbb{R}^4 \rightarrow \mathbb{R}$ est prise dans le Tableau 2.2 ci-dessous :

TAB. 2.2 – Liens entre le CTM et les mesures de risque classiques.

Mesure de risque	$\Phi(t_0, t_1, t_2, t_3)$
$CTE(\alpha)$	t_1
$CTV(\alpha)$	$t_2 - t_1^2$
$CTS(\alpha)$	$t_3 / (t_2 - t_1^2)^{3/2}$
$CVaR(\alpha)$	$\lambda t_0 + (1 - \lambda t_1), \lambda \in [0, 1]$
$SP(\alpha)$	$a(t_1 - t_0)$

De manière plus générale, le **CTM** peut être utilisé dans le but de définir n'importe quelle mesure de risque basée sur les moments conditionnels de la variable aléatoire des pertes sachant que les pertes sont supérieures à la **VaR** de niveau. En effet, on peut par exemple introduire le Conditional Tail Kurtosis à l'aide de la fonction :

$$\Phi(t_0, t_1, t_2, t_3, t_4) = \frac{t_4}{(t_2 - t_1^2)^2}.$$

Chapitre 3

Estimation de CTM

3.1 Estimation de CTM

On sait que le CTM est donné par :

$$CTM_a(\alpha) = \mathbb{E}(Y^a | Y > VaR(\alpha))$$

où

$$VaR(\alpha) = Q(1 - \alpha)$$

alors

$$CTM_a(\alpha) = \mathbb{E}(Y^a | Y > VaR(\alpha)) = \frac{1}{1 - F(\alpha)} \int_{VaR(\alpha)}^{+\infty} y^a dF(y). \quad (3.1)$$

On a

$$\mathbb{E}(Y^a) = \int_{-\infty}^{+\infty} y^a dF(y); \quad \text{tel que } 0 \leq F(X) \leq 1 \quad (3.2)$$

Posons

$$\begin{aligned} 1 - F(x_q) &= \mathbb{P}(X > x_q) = s \\ 1 - F(x_q) = s &\Rightarrow F(x_q) = 1 - s \\ &\Rightarrow x_q = F^{-1}(1 - s) \\ &\Rightarrow x_q = Q(1 - s) \end{aligned}$$

on remplace dans l'équation(3.2), on obtient :

$$\mathbb{E}(Y^a) = - \int_1^0 Q^a(1 - s) ds = \int_0^1 Q^a(t) dt. \quad (3.3)$$

On sait que $VaR(\alpha) = Q(1 - \alpha)$, on note $\mathbb{E}(Y^a | Y > VaR(\alpha)) := C^a(\alpha)$, alors

$$C^a(\alpha) = \frac{1}{1 - \alpha} \int_{\alpha}^1 Q^a(s) ds. \quad (3.4)$$

3.1.1 Estimation empirique de CTM

Soit X_1, X_2, \dots, X_n n copies de la v.a. X ; de distribution F . L'estimateur empirique de F est défini par :

$$F_n(x) = \frac{1}{n} \sum_{i=1}^n 1_{\{X_i < x\}}$$

où 1_A est la fonction indicatrice sur l'ensemble A .

La fonction des quantiles ou l'inverse généralisé de la fonction de distribution F notée par Q : pour chaque entier $n \geq 1$ est définie par

$$Q(s) = F^{\leftarrow}(s) = \inf \{t : F(t) \geq s\}, 0 \leq s \leq 1$$

La fonction des quantiles empiriques : notée Q_n est définie par :

$$Q_n(s) = F_n^{\leftarrow}(s) = \inf \{t : F_n(t) \geq s\}, 0 \leq s \leq 1$$

Soit $X_{1;n}, X_{2;n}, \dots, X_{n;n}$ est la statistique d'ordre basé sur l'échantillon X_1, X_2, \dots, X_n . L'estimateur empirique du quantile basé sur l'échantion X_1, X_2, \dots, X_n est donné par :

$$Q_n(x) = \begin{cases} X_{1;n} & \text{si } x = 0 \\ X_{k;n} & \text{si } \frac{i-1}{n} < k \leq \frac{i}{n} \\ X_{n;n} & \text{si } x = 1 \end{cases} \quad (3.5)$$

Par le remplacement de l'estimation empirique du quantile donné par l'équation (3.5) dans l'équation (3.4), on obtient un estimateur empirique de $C^a(\alpha)$, défini par :

$$\begin{aligned} \hat{C}_n^a(\alpha) &= \frac{1}{1 - \alpha} \int_{\alpha}^1 (Q_n(s))^a ds \\ &= \frac{1}{1 - \alpha} \sum_{i=k+1}^n \int_{1-i/n}^{i/n} X_{i,n}^a ds \\ &= \frac{1}{1 - \alpha} \frac{1}{n - k} \sum_{i=k+1}^n X_{i,n}^a \end{aligned} \quad (3.6)$$

Alors

$$\hat{C}_n^a(\alpha) = \frac{1}{1-\alpha} \frac{1}{n-k} \sum_{i=k+1}^n X_{i,n}^a. \quad (3.7)$$

est un estimation empirique de CTM.

La normalié asymptôtique de l'estimateur (3.7) est présentée dans le théorème suivant :

Théorème 3.1 *On suppose $E(X^{2+\varepsilon}) < \infty, \varepsilon > 0$ et pour $\alpha \in]0, 1[$ fixé. Alors*

$$\sqrt{n} \left(\hat{C}_n^a(\alpha) - C^a(\alpha) \right) \rightarrow N(0; \sigma^2) \quad (3.8)$$

où

$$\sigma^2 := \frac{1}{(1-\alpha)^2} \int_{F^{-1}(\alpha)}^{\infty} \int_{F^{-1}(\alpha)}^{\infty} (F_{X,Y}^a(x,y) - F_X^a(x) F_Y^a(y)) \mathbf{I}_{\{F_{X,n}(x) \geq \alpha\}} \mathbf{I}_{\{F_{Y,n}(y) \geq \alpha\}} dx dy \quad (3.9)$$

Théorème 3.2 *Laissez la fonction de distribution F_X être continue. Supposons que le moment $E(X^{2+\varepsilon})$ est fini pour certains $\varepsilon > 0$, et laisse le moment $E[X_+^r]$ être aussi fini pour certains $r > 2$, où $X_+ = \max(X, 0)$. Si la fonction de pondération $q : (0, 1) \rightarrow [0, \infty)$ est telle que, pour certains $v > 0$,*

$$\sup_{0 < \alpha < 1} \frac{q(\alpha)}{(1-\alpha)^{1/r+1/2+v}} < \infty, \quad (3.10)$$

alors nous avons cela

$$\sqrt{n} \sup_{0 < \alpha < 1} q(\alpha) |\hat{C}_n^a(\alpha) - C^a(\alpha)| \xrightarrow{d} \sup_{0 < \alpha < 1} \frac{q(\alpha)}{1-\alpha} \left| \int_{F^{-1}(\alpha)}^{\infty} \beta(F(x))^a dx \right| \quad (3.11)$$

où β est le pont brownien sur l'intervalle $[0, 1]$.

à partir de 3.11 nous obtenons la bande de confiance asymptotique suivante $(1-\omega)100\%$ pour la fonction C^a :

$$C_n^a(\alpha) \pm \frac{s_\omega}{q(\alpha)\sqrt{n}} \text{ pour tout } \alpha \in [0, 1] \text{ tel que } q(\alpha) > 0, \quad (3.12)$$

où s_ω est le $(1-\omega)$ ième quantile de la fonction de distribution de la variable aléatoire limitante du côté droit de 3.11.

Evidemment, s_ω dépend à la fois de la fonction de distribution F^a et de la fonction de poids q .

Since s_ω dépend de la fonction de distribution (inconnue) F^a , il doit être estimé de manière empirique. Cela peut être fait en utilisant une approximation bootstrap analogue à celle décrite dans les deux paragraphes ci-dessous 3.8 mais maintenant avec

$$\sqrt{n} \sup_{0 < \alpha < 1} q(\alpha) |\hat{C}_n^{*a}(\alpha) - C^a(\alpha)| \quad (3.13)$$

au lieu de la quantité 3.9. Nous obtenons la bande de confiance de niveau asymptotique $(1 - \omega)100\%$ pour la fonction C^a :

$$C_n^a(\alpha) \pm \frac{x^*}{q(\alpha)\sqrt{n}} \text{ pour tout } \alpha \in [0, 1], \text{ tel que } q(\alpha) > 0. \quad (3.14)$$

Notez que si une bande de confiance pour C^a est souhaitée, par exemple sur un intervalle $[\alpha_1, \alpha_2] \subseteq [0, 1]$, alors il est naturel de choisir $q(\alpha) = 0$ pour tout α, s en dehors l'intervalle $[\alpha_1, \alpha_2]$.

Preuve: de théorème 3.1 On commence par la différence suivante

$$\begin{aligned} \hat{C}_n^a(\alpha) - C^a(\alpha) &= \frac{1}{1 - \alpha} \left(\int_\alpha^1 (Q_n(s))^a ds - \int_\alpha^1 Q^a(s) ds \right) \quad (3.15) \\ &= \frac{1}{1 - \alpha} \int_\alpha^1 [(Q_n(s))^a - Q^a(s)] ds \end{aligned}$$

Notre prochaine étape consiste à extraire une somme de variables aléatoires du côté droit de 3.15. Pour comprendre comment faire cela, nous devons maintenant regarder l'intégrale

$$\int_\alpha^1 [(Q_n(s))^a - Q^a(s)] ds \quad (3.16)$$

Lorsque $\alpha = 0$, l'intégrale est la différence entre la moyenne empirique $n^{-1} \sum_{i=1}^n X_i$ et le théorique $\mathbf{E}[X]$. Cette différence est certainement une somme de variables aléatoires i.i.d centrées, qui est une représentation désirée. Notez également que la différence peut être écrite

$$- \int_{-\infty}^{\infty} [(Q_n(x))^a - Q^a(x)] dx \quad (3.17)$$

ce qui est encore une somme de variables aléatoires i.i.d centrées. Cette dernière représentation donne un indice très important sur la façon d'extraire une somme de i.i.d variables aléatoires de l'intégrale dans 3.16. C'est ce que

nous accomplissons rigoureusement ensuite. Nous commençons avec l'équation

$$\begin{aligned} \int_{\alpha}^1 [(Q_n(s))^a - Q^a(s)] ds &= \int_{\alpha}^1 [(F_n^{-1}(s))^a - (F^{-1}(s))^a] ds \\ &= - \int_{F^{-1}(\alpha)}^{\infty} (F_n^a(x) - F^a(x)) dx + R_n^a \end{aligned} \quad (3.18)$$

où le reste du terme $R_n(\alpha)$ est défini par equation 3.18 lui-même. Notez que $R_n(\alpha)$ est égal à 0 lorsque $\alpha = 0$, qui résulte de la déjà noté l'égalité des intégrales dans 3.16 et 3.17. Cette égalité, à son tour, implique la représentation suivante :

$$R_n(\alpha) = -V_n \quad (3.19)$$

où

$$V_n^a(\alpha) = \int_{\alpha}^1 [(F_n^{-1}(s))^a - (F^{-1}(s))^a] ds + \int_0^{F^{-1}(\alpha)} (F_n^a(x) - F^a(x)) dx$$

A première vue, il est difficile de voir pourquoi on devrait préférer la représentation 3.19 à la définition originale de $R_n(\alpha)$ donnée par equation. 3.18. La raison en est la suivante. Le processus V_n a été étudié d'une manière approfondie dans la littérature et est connu comme le (général) Processus de Vervaat (cf., par exemple, Zitikis, 1998, et Davydov et Zitikis, 2004). Parmi de nombreux faits à propos de ce processus, nous savons, par exemple, que $V_n(\alpha)$ est non négatif pour tout $\alpha \in]0, 1[$ et vérifie la borne suivante :

$$V_n^a(\alpha) \leq - (F_n^a((F^{-1}(\alpha)))^a - \alpha) ((F_n^{-1}(\alpha))^a - (F^{-1}(\alpha))^a) \quad (3.20)$$

pour chaque fonction de répartition F . Si, cependant, on sait que la fonction de distribution F est continue au point $F^{-1}(\alpha)$, alors nous avons l'égalité $\alpha = F(F^{-1}(\alpha))$ et donc, à son tour, les limites suivantes :

$$\begin{aligned} |V_n^a(\alpha)| &\leq |(F_n^a((F^{-1}(\alpha)))^a - \alpha)| |((F_n^{-1}(\alpha))^a - (F^{-1}(\alpha))^a)| \\ |V_n^a(\alpha)| &\leq \sup_{x \in \mathbb{R}} |F_n(x) - F(x)| |F_n^{-1}(\alpha) - F^{-1}(\alpha)| \end{aligned} \quad (3.21)$$

Par le théorème classique de Kolmogorov - Smirnov (cf., par exemple, Shorack et Wellner, 1986).), le supremum sur le côté droit de 3.21 est de l'ordre $\mathcal{O}_{\mathbf{P}}(n^{-1/2})$. Ainsi, $\sqrt{n}|V_n^a(t)| = o_{\mathbf{P}}(1)$ chaque fois que $F_n^{-1}(\alpha) \rightarrow \mathbf{P}F^{-1}(\alpha)$. La dernière convergence tient si la fonction de distribution F est continue au

point $F^{-1}(\alpha)$. Par conséquent, pour tout α fixe et $\alpha \in]0, 1[$, nous avons

$$\begin{aligned}\sqrt{n} \left(\hat{C}_n^a(\alpha) - C^a(\alpha) \right) &= -\frac{1}{1-\alpha} \left(\int_{F^{-1}(\alpha)}^{\infty} \sqrt{n} \left((F_n(x))^a - (F(x))^a \right) dx \right) - \frac{1}{1-\alpha} \sqrt{n} (V_n(\alpha)) \\ &= -\frac{1}{1-\alpha} \left(\int_{F^{-1}(\alpha)}^{\infty} \sqrt{n} \left((F_n(x))^a - (F(x))^a \right) dx \right) + o_{\mathbf{P}}(1) \\ &= \frac{1}{\sqrt{n}} \sum_{i=1}^n H(X_i, \alpha) + o_{\mathbf{P}}(1)\end{aligned}\tag{3.22}$$

où

$$H(X_i, \alpha) = -\frac{1}{1-\alpha} \int_{F^{-1}(\alpha)}^{\infty} (1_{\{X_i \leq x\}} + (F(x))^a) dx$$

Pour chaque $\alpha \in [0, 1]$, les variables aléatoires $H(X_i; \alpha)$, $1 \leq i \leq n$, sont centrées, i.i.d et ont des variances $\sigma_X^2(\alpha)$. La variance σ_X^2 est fini pour tout $\alpha \in [0, 1]$ si le second moment de X est fini. Ce dernier tient par hypothèse. Ceci complète la preuve de Théorème 3.1. ■

Preuve: de théorème 3.2

Avec $V_{X,n}$ dénotant le processus de Vervaat, nous avons l'égalité

$$q(\alpha) \left(\hat{C}_n^a(\alpha) - C^a(\alpha) \right) = -\frac{q(\alpha)}{1-\alpha} \int_{F^{-1}(\alpha)}^{\infty} \left((F_n(x))^a - (F(x))^a \right) dx - \frac{q(\alpha)}{1-\alpha} (V_n(\alpha))^a\tag{3.23}$$

Nous allons maintenant prouver que

$$\sqrt{n} \sup_{0 < \alpha < 1} \frac{q(\alpha)}{1-\alpha} (V_n(\alpha))^a \rightarrow \mathbf{p}0.\tag{3.24}$$

En utilisant la première borne de 3.21, nous avons cela $|V_n(\alpha)|$ ne dépasse pas le produit de $\alpha^{1/2-v} (1-\alpha)^{1/2-v} |(F_n^{-1}(\alpha))^a - (F^{-1}(\alpha))^a|$ et

$$\sup_{0 < u < 1} \frac{|F_n(F^{-1}(u)) - F(F^{-1}(u))|}{u^{1/2-v} (1-u)^{1/2-v}}.\tag{3.25}$$

Par le théorème pondéré de Kolmogorov - Smirnov (cf. Shorack et Wellner, 1986.), la quantité 3.25 est de l'ordre $\mathcal{O}_{\mathbf{P}}(n^{-1/2})$ pour tout $v > 0$. Par conséquent, la déclaration 3.24 Maintient à condition que

$$\sup_{0 < \alpha < 1} \frac{q(\alpha) \alpha^{1/2-v}}{(1-\alpha)^{1/2-v}} |F_n^{-1}(\alpha) - F^{-1}(\alpha)| \rightarrow \mathbf{p}0.\tag{3.26}$$

Au vu de l'hypothèse 3.10, la déclaration 3.26 découle de la déclaration

$$\sup_{0 < \alpha < 1} \alpha^{1/r_1} (1-\alpha)^{1/r_1} |(F_n^{-1}(\alpha))^a - (F^{-1}(\alpha))^a| \rightarrow \mathbf{p}0.\tag{3.27}$$

Puisque, par hypothèse, $\mathbf{E}[X_-^{r_1}] < \infty$ avec quelques $r_1 > 2$ et $\mathbf{E}[X_+^{r_1}] < \infty$, nous avons (cf., par exemple, Shorack et Wellner, 1986) cette déclaration 3.27 tient. Cela prouve l'affirmation 3.24. De l'égalité 3.23 et de la déclaration 3.24, nous avons cela

$$\sqrt{n} \sup_{0 < \alpha < 1} q(\alpha) \left| \hat{C}_n^a(\alpha) - C^a(\alpha) \right| = \sqrt{n} \sup_{0 < \alpha < 1} \frac{q(\alpha)}{1 - \alpha} \left| \int_{F^{-1}(\alpha)}^{\infty} ((F_n(x))^a - (F(x))^a) dx \right| + o_{\mathbf{P}}(1). \quad (3.28)$$

Nous devons établir une convergence dans la distribution du terme principal sur le côté droit de 3.28. Pour cela, nous réécrivons le terme comme suit :

$$\sup_{0 < \alpha < 1} \frac{q(\alpha)}{1 - \alpha} \left| \int_{F^{-1}(\alpha)}^{\infty} \left[\frac{\sqrt{n} ((F_n(x))^a - (F(x))^a)}{(1 - (F(x))^a)^{1/2-\delta}} \right] (1 - (F(x))^a)^{1/2-\delta} dx \right| \quad (3.29)$$

où $\delta > 0$ sera spécifié ci-dessous. Le processus par rapport à $-\infty < x < \infty$ entre crochets $[\cdot]$ ci-dessus converge dans les espaces fonctionnel au processus $\mathcal{B}((F(x))^a) / (1 - (F(x))^a)^{1/2-\delta}$ (cf., par exemple, Shorack et Wellner, 1986). Par conséquent, si la quantité

$$\sup_{0 < \alpha < 1} \frac{q(\alpha)}{1 - \alpha} \left| \int_{F^{-1}(\alpha)}^{\infty} (1 - (F(x))^a)^{1/2-\delta} dx \right| \quad (3.30)$$

est fini pour quelques $\delta > 0$, puis par le théorème de cartographie continue (cf., par exemple, Shorack et Wellner, 1986) nous avons cette quantité 3.29 converge dans la distribution du côté droit de (3.7). Par conséquent, afin de terminer la preuve du théorème 3.2, nous avons seulement besoin vérifier que la quantité de 3.30 est finie pour quelques $\delta > 0$. Si nous remplaçons le supremum en quantité 3.30 par le supremum sur $0 < \alpha \leq 1/2$, alors la quantité résultante sera finie pour suffisamment petit $\delta > 0$, en raison de l'hypothèse $\mathbf{E}[X_+^{r_1}] < \infty$ que pour certains $\varepsilon > 0$. Ainsi, il nous reste à vérifier cette quantité 3.30 est également fini si nous remplaçons le supremum par le supremum sur $1/2 \leq \alpha < 1$. Nous faisons ceci comme suit. Tout d'abord, en utilisant 3.10, nous avons cela

$$\begin{aligned} & \sup_{1/2 \leq \alpha < 1} \frac{q(\alpha)}{1 - \alpha} \left| \int_{F^{-1}(\alpha)}^{\infty} (1 - (F(x))^a)^{1/2-\delta} dx \right| \quad (3.31) \\ & \leq c \sup_{1/2 \leq \alpha < 1} \frac{1}{(1 - \alpha)^{1-(1/r+1/2+v)}} \int_{F^{-1}(\alpha)}^{\infty} (1 - (F(x))^a)^{1/2-\delta} dx \quad (3.32) \end{aligned}$$

Maintenant, nous choisissons $v > 0$ suffisamment petit pour que $(1/r + 1/2 + v) < 1$ détiendrait (un tel choix est possible puisque $r > 2$). Continuer avec 3.32,

nous avons

$$\begin{aligned}
& \sup_{1/2 \leq \alpha < 1} \frac{1}{(1-\alpha)^{1-(1/r+1/2+v)}} \int_{F^{-1}(\alpha)}^{\infty} (1 - (F(x))^a)^{1/2-\delta} dx \quad (3.33) \\
& \leq \sup_{1/2 \leq \alpha < 1} \int_{F^{-1}(\alpha)}^{\infty} \frac{(1 - (F(x))^a)^{1/2-\delta}}{(1 - (F(x))^a)^{1-(1/r+1/2+v)}} dx \\
& \leq \int_{F^{-1}(1/2)}^{\infty} (1 - (F(x))^a)^{1/r+v-\delta} dx \quad (3.34)
\end{aligned}$$

Choisissez maintenant $\delta = v/2$. Depuis $\mathbf{E}[X_+^{r+1}] < \infty$ par hypothèse, la droite la plus intégrale dans 3.34 est finie. Cela termine la preuve de Théorème 3.2.

■

Preuve de la formule 3.9. Les égalités suivantes sont simples :

$$\begin{aligned}
\sigma_{X,Y,n}^2(\alpha) &= \frac{1}{(1-\alpha)^2} \int_{F^{-1}(\alpha)}^{\infty} \int_{F^{-1}(\alpha)}^{\infty} (F_{X,Y,n}^a(x,y) - F_{X,n}^a(x) F_{Y,n}^a(y)) \mathbf{I}_{\{F_{X,n}(x) \geq \alpha\}} \mathbf{I}_{\{F_{Y,n}(y) \geq \alpha\}} \\
&= \frac{1}{(1-\alpha)^2} \sum_{n\alpha \leq j \leq n-1} \sum_{n\alpha \leq k \leq n-1} \int_{X_{j,n}}^{X_{j+1,n}} \int_{Y_{k,n}}^{Y_{k+1,n}} \left(\frac{1}{n} \sum_{i=1}^n \mathbf{I}_{\{X_i \leq x, Y_i \leq y\}} - \frac{j}{n} \frac{k}{n} \right) dx dy \\
&= \frac{1}{(1-\alpha)^2} \sum_{n\alpha \leq j \leq n-1} \sum_{n\alpha \leq k \leq n-1} \left(\frac{k_n(j,k)}{n} \frac{j}{n} \frac{k}{n} \right) (X_{j+1;n} - X_{j;n}) (Y_{k-1;n} - Y_{k;n}). \quad (
\end{aligned}$$

3.2 Simulations

Pour illustrer la performance de cet mesure nous avons effectué une petite étude de simulation basée sur la distribution de Pareto $F(x) = 1 - x^{-1/\gamma}$, $x > 1$, avec l'indice de queue $\gamma = 2/3$ et $\gamma = 3/4$ et on prend $\alpha = 0.95$ comme valeur fixé. Nous avons généré 500 indépendants répliques de trois échantillons de tailles $n = 500, 1000, 2000$.

Pour chaque échantillon simulé, nous avons obtenu des estimations $\hat{C}_n^a(t)$, nous avons calculé les moyennes arithmétiques sur les valeurs des 500 répliques, avec l'erreur absolue (*error*), et le Biais du nouvel estimateur $\hat{C}_n^a(t)$ rapporté dans le tableau 1 ($\gamma = 3/4$) et le tableau 2 ($\gamma = 2/3$) avec $a = 1$ pour la distribution Pareto après les calculs on obtient une valeur théorique fixe pour tout les échantillons, et des valeurs empirique pour chaque échantillon.

TAB. 3.1 – Simulation du CTM_1 pour distribution de Pareto avec $\gamma = 3/4$ et $t = 0.95$ sachant que la valeur théorique est 29.822979

n	$\hat{C}_n^a(\alpha)$	<i>erreur</i>
500	31.8196	0.0667
1000	29.341	0.05222
2000	28.8796	0.0078

TAB. 3.2 – Simulation du CTM_1 pour distribution de pareto avec $\gamma = 2/3$ et $t = 0.95$ sachant que la valeur théorique est 19.31929

n	$\hat{C}_n^a(\alpha)$	<i>erreur</i>
500	21.3675	0.09555
1000	20.3524	0.0666
2000	19.6726	0.01476

3.3 Estimateur semi parametrique

3.3.1 Théorie des limites pour les moments extrêmes de la queue conditionnelle

Soit $\{X_i\}$ une suite strictement stationnaire de r.v.s non négatifs, dont la queue droite sera étudiée comme c'est la queue dans la théorie des valeurs extrêmes. En pratique, la non-négativité peut être réalisée via un simple transformation, par exemple $X_i I_{\{X_i \geq 0\}}$ où $-X_i I_{\{-X_i \geq 0\}}$ si l'intérêt se concentre sur la queue droite ou gauche respectivement. Définir la fonction de survie $\bar{F}(\cdot) = 1 - F(\cdot)$, où F désigne la fonction de distribution de X_1 , nous prendrons des queues des variantes régulière $\bar{F}(\cdot) \in RV_{-1/\gamma}$, c.à.d.

$$\lim_{x \rightarrow \infty} \frac{\bar{F}(\lambda x)}{\bar{F}(x)} = \lambda^{-1/\gamma} \quad \forall \lambda > 0, \quad (3.36)$$

où $\gamma > 0$ est appelé l'indice des valeurs extrême et $\alpha = 1/\gamma$ l'indice de la queue. Notons que la condition (3.36) est équivalent à la condition suivante

$$F(x) = x^{-1/\gamma} L(x) \text{ où } L(\cdot) \text{ varie lentement, c.-à-d } \lim_{x \rightarrow \infty} \frac{L(\lambda x)}{L(x)} = 1. \quad (3.37)$$

Ceci est équivalent à (de Haan et Ferreira, 2006, p.25)

$$U(x) = x^\gamma L_U(x), \text{ où } U(x) = F^{\leftarrow}(1 - 1/x) \text{ et } L_U(\cdot) \text{ varie lentement.} \quad (3.38)$$

Puisque (3.36) est une relation asymptotique, nous avons besoin d'une suite intermédiaire $k_n \rightarrow \infty$ avec $k_n = o(n)$ ($k/n \rightarrow 0$) et $k_n < n$. Cette séquence k_n est limitée par l'hypothèse suivante.

Hypothèse 1. Il existe une fonction $A(\cdot)$ avec $\lim_{x \rightarrow \infty} A(x) = 0$ telle que pour certains $\rho < 0$

$$\lim_{x \rightarrow \infty} \frac{\frac{\bar{F}(\lambda x)}{\bar{F}(x)} - \lambda^{-1/\gamma}}{A(x)} = \lambda^{-1/\gamma} \frac{\lambda^{\rho/\gamma}}{\gamma \rho}, \quad \forall \lambda > 0. \quad (3.39)$$

Avec la condition asymptotique $\sqrt{k_n} A(U(n/k_n)) \rightarrow 0$, comme $n \rightarrow \infty$.

Remarque 3.1 Cette hypothèse contrôle la vitesse de convergence de 3.36 et est par conséquent référée à comme une condition de second ordre dans la théorie de la valeur extrême (**EVT**). De manière équivalente, il peut également être écrit dans termes de la fonction quantile $U(\cdot)$ de 3.38 (voir de Haan et Ferreira, 2006, Théorème 2.3.9). Dans ce forme, il est largement utilisé dans l'index de queue (par exemple, Einmahl et al., 2016; Hoga, 2017 + a) et quantile extrême estimation (par exemple Chan et al., 2007, Hoga, 2017 + b). Des exemples de d.f.s satisfaisant à l'hypothèse 1 sont

$$\bar{F}(x) = c_1 x^{-1/\gamma} + c_2 x^{-1/\gamma + \rho/\gamma} (1 + o(1)), \quad x \rightarrow \infty, \quad (c_1 > 0, c_2 \neq 0, \gamma > 0, \rho < 0) \quad (3.40)$$

remplir l'hypothèse 1 avec γ et ρ , et $k_n = o(n^{-2\rho/(1-2\rho)})$ (de Haan et Ferreira, 2006, pp. 76-77). Plus ρ est négatif, plus la queue est proche de la décroissance réelle de Pareto ($\rho = -\infty$).

Définir $x_p = F^{\leftarrow}(1 - \alpha)$ comme le $(1 - \alpha)$ quantile. La plupart de la littérature, se concentrent sur le cas où $\alpha \in (0, 1)$ est fixé. **EVT** permet cependant $\alpha = \alpha_n \rightarrow 0$ comme $n \rightarrow \infty$. Les approximations dérivées d'**EVT** fournissent souvent de meilleures approximations lorsque α est petit - le cas d'un intérêt particulier dans la gestion des risques -, car ils prennent la queue semi-paramétrique 3.36 en compte. Les deux motivations suivantes montrent comment la variation régulière de la queue est prise en compte.

Premièrement, nous utilisons l'hypothèse de variation régulière 3.36 pour estimer x_{p_n} dans $CTMa(\alpha_n) = E[X^a | X > x_{\alpha_n}]$ comme suit. Notez que α_n peut être très petit, de sorte que x_{α_n} peut se trouver en dehors de la gamme des observations X_1, \dots, X_n . Ensuite, l'idée est de baser l'estimation de x_{α_n} sur un quantile moins extrême (dans l'échantillon) $x_{k_n/n}$ et utiliser 3.36 pour extrapoler à partir de cette estimation. Concrètement, mettre $x = x_{k_n/n}$, $\lambda = x_{\alpha_n}/x_{k_n/n}$ et utiliser 3.36 comme une approximation pour obtenir

$$\left(\frac{x_{\alpha_n}}{x_{k_n/n}} \right)^{-1/\gamma} \approx \frac{1 - F(x_{\alpha_n})}{1 - F(x_{k_n/n})} \approx \frac{n\alpha_n}{k_n}. \quad (3.41)$$

Remplaçant la population par des quantités empiriques, cette approximation motive le soi-disant Weissman (1978) estimateur $\hat{x}_{\alpha_n} = d_n^\gamma X_{(kn+1)}$, où $d_n = k_n/(n\alpha_n)$. Il a été utilisé dans, par exemple, Drees (2003), Chan et al. (2007), ou Hoga et Wied (2017). Bien sûr, il existe un large éventail d'estimateurs $\hat{\gamma}$. Nous utiliserons l'estimateur de Hill (1975)

$$\hat{\gamma} = \frac{1}{k_n} \sum_{i=1}^{k_n} \log (X_{(i)}/X_{(k_n+1)})$$

dans ce qui suit, qui est sans doute le plus populaire (voir, par exemple, Hsing, 1991, Hill, 2010, et références ici). Pour la deuxième approximation, nous exploitons de nouveau 3.36

on a

$$\begin{aligned} \hat{C}_n^a(\alpha) &= \mathbb{E}(Y^a \mid Y > VaR(\alpha)) \\ &= \frac{1}{1-\alpha} \int_{\alpha}^1 (Q_n(s))^a ds \\ &= \frac{1}{1-\alpha} \left[\int_{\alpha}^{1-k/n} Q_n^a(s) ds + \int_{1-k/n}^1 Q_w^a(s) ds \right] \end{aligned} \quad (3.42)$$

pour $s \in]1 - k/n, 1[$ alors

$$Q_w(1-s) = X_{n-k,n} \left(\frac{k}{n} \right) s^{-\hat{\gamma}_n}$$

est introduit par Weismann 1978, alors on remplace dans 3.42

$$\begin{aligned} \int_{\alpha}^{1-k/n} Q_n^a(s) ds + \int_{1-k/n}^1 Q_w^a(s) ds &= \int_{\alpha}^{1-k/n} Q_n^a(s) ds + \int_{1-k/n}^1 \left(X_{n-k,n} \left(\frac{k}{n} \right) s^{-\hat{\gamma}_n} \right)^a ds \\ &= \int_{\alpha}^{1-k/n} Q_n^a(s) ds + \int_{1-k/n}^1 X_{n-k,n}^a \left(\frac{k}{n} \right)^a s^{-a\hat{\gamma}_n} ds \\ &= \int_{\alpha}^{1-k/n} Q_n^a(s) ds + \left[X_{n-k,n}^a \left(\frac{k}{n} \right)^a \frac{s^{-a\hat{\gamma}_n+1}}{1-a\hat{\gamma}_n} \right]_{1-k/n}^1 \end{aligned}$$

avec après le changement de variables on obtient :

$$\begin{aligned} \hat{C}_n^a(\alpha) &= \frac{1}{1-\alpha} \left(\int_{\alpha}^{1-k/n} Q_n^a(s) ds + \left[X_{n-k,n}^a \left(\frac{k}{n} \right)^a \frac{t^{-a\hat{\gamma}_n+1}}{1-a\hat{\gamma}_n} \right]_0^{k/n} \right), 0 < \alpha < k/n \\ &= \frac{1}{1-\alpha} \left(\int_{\alpha}^{1-k/n} Q_n^a(s) ds + \left[X_{n-k,n}^a \left(\frac{k}{n} \right)^a \frac{1}{1-a\hat{\gamma}_n} \right] \right). \end{aligned}$$

Finalement, on a le semi parametrique estimateur de CTM, comme suivant :

$$\hat{C}_n^a(\alpha) = \frac{1}{1-\alpha} \left(\int_{\alpha}^{1-k/n} Q_n^a(s) ds + \left[\left(\frac{k}{n} \right)^a \frac{X_{n-k,n}^a}{1-a\hat{\gamma}_n} \right] \right). \quad (3.43)$$

Théorème 3.3 Soit F une distribution de X_i vérifiant la condition $F(x) = x^{-1/\gamma}L(x)$ avec γ l'indice de valeurs extrême telle que $\gamma \in]0.5, 1[$; avec k_n est une suite $k_n \rightarrow \infty$ et $k_n = o(n)$ ($k/n \rightarrow 0$) avec $\sqrt{k_n}A(U(n/k_n)) \rightarrow 0$, lorsque $n \rightarrow \infty$. Alors :

$$\sqrt{n} \frac{\hat{C}_n^a(\alpha) - C_n^a(\alpha)}{\left(\frac{k}{n} \right)^a X_{n-k,n}^a} \xrightarrow{\mathcal{D}} \mathcal{N}(0, \sigma^2(\gamma, a))$$

avec

$$\sigma^2(\gamma, a) = \frac{a^2 \gamma^4}{(1-a\gamma)^4 (2a\gamma - 1)}. \quad (3.44)$$

Preuve: Nous commençons la preuve du théorème 3.3 avec la décomposition

$$(\tilde{C}_n^a(t) - C^a(t))(1-t) = A_{n,1}^a(t) + A_{n,2}^a, \quad (3.45)$$

où

$$A_{n,1}^a(t) = \int_t^{1-k/n} (Q_n^a(s) - Q^a(s)) ds,$$

$$A_{n,2}^a = \frac{k/n}{1-a\hat{\gamma}_n} X_{n-k,n} - \int_{1-k/n}^1 Q^a(s) ds.$$

Nous montrerons plus loin qu'il existe des ponts browniens B_n tels que

$$\frac{\sqrt{n}A_{n,1}^a(t)}{(k/n)^{1/2}Q^a(1-k/n)} = - \frac{\int_0^{1-k/n} B_n^a(s) dQ(s)}{(k/n)^{1/2}Q^a(1-k/n)} + o_{\mathbf{P}}(1) \quad (3.46)$$

et

$$\frac{\sqrt{n}A_{n,2}^a}{(k/n)^{1/2}Q^a(1-k/n)} = \frac{a\gamma^2}{(1-a\gamma)^2} \sqrt{\frac{n}{k}} B_n^a(1-k/n) - \frac{a\gamma}{(1-a\gamma)^2} \sqrt{\frac{n}{k}} \int_{1-k/n}^1 \frac{B_n^a(s)}{1-s} ds + o_{\mathbf{P}}(1). \quad (3.47)$$

En supposant pour le moment que les déclarations (3.46) et (3.47) sont valides, nous complétons la preuve du théorème 3.3. Pour simplifier la présentation, nous utilisons la notation :

$$\begin{aligned} W_{1,n} &= -\frac{\int_0^{1-k/n} B_n^a(s) dQ^a(s)}{(k/n)^{1/2} Q^a(1-k/n)}, \\ W_{2,n} &= \frac{a\gamma^2}{(1-a\gamma)^2} \sqrt{\frac{n}{k}} B_n^a(1-k/n), \\ W_{3,n} &= -\frac{a\gamma}{(1-a\gamma)^2} \sqrt{\frac{n}{k}} \int_{1-k/n}^1 \frac{B_n^a(s)}{1-s} ds. \end{aligned}$$

Par conséquent, nous avons la représentation asymptotique

$$\frac{\sqrt{n}(\tilde{\mathbb{C}}_n^a(t) - \mathbb{C}^a(t))(1-t)}{(k/n)^{1/2} Q^a(1-k/n)} = W_{1,n} + W_{2,n} + W_{3,n} + o_{\mathbf{P}}(1).$$

La somme $W_{1,n} + W_{2,n} + W_{3,n}$ est une variable aléatoire gaussienne centrée. Pour calculer sa variance asymptotique, nous établissons les limites suivantes :

$$\begin{aligned} \mathbf{E}[W_{1,n}^2] &\rightarrow \frac{2a\gamma}{2a\gamma-1}, & \mathbf{E}[W_{2,n}^2] &\rightarrow \frac{a^2\gamma^4}{(1-a\gamma)^4}, & \mathbf{E}[W_{3,n}^2] &\rightarrow \frac{2(a\gamma)^2}{(1-a\gamma)^4}, \\ 2\mathbf{E}[W_{1,n}W_{2,n}] &\rightarrow \frac{-2(a\gamma)^2}{(1-a\gamma)^2}, & 2\mathbf{E}[W_{1,n}W_{3,n}] &\rightarrow \frac{2a\gamma}{(1-a\gamma)^2}, \\ 2\mathbf{E}[W_{2,n}W_{3,n}] &\rightarrow \frac{-2a^2\gamma^3}{(1-a\gamma)^4}. \end{aligned}$$

Résumant les côtés droits des six limites ci-dessus, nous obtenons σ_γ^2 , dont l'expression en termes du paramètre γ est donnée dans le théorème(3.3). Enfin, depuis $X_{n-k:n}/Q(1-k/n)$ converge en probabilité de 1 (voir, par exemple, la preuve du corollaire dans Necir et Meraghni (2009)), le lemme classique de Sultsky complète la preuve du théorème(3.3). Bien sûr, il nous reste à vérifier les instructions(3.46) et (3.47), qui font le contenu des deux sous-sections suivantes. ■

3.3.2 Preuve de l'expression (3.46).

Nous ne supposons pas la différentiabilité de Q et donc une nouvelle preuve est nécessaire, qui est fondamentalement basée sur le processus de

Vervaat (voir Zitikis (1998), Davydov et Zitikis (2003), Davydov et Zitikis (2004), et les références)

$$V_n^a(t) = \int_0^t (Q_n^a(s) - Q^a(s)) ds + \int_{-\infty}^{Q(t)} (F_n^a(x) - F^a(x)) dx.$$

Par conséquent, pour chaque t tel que $0 < t < 1 - k/n$, ce qui est satisfait pour tout n suffisamment grand puisque t est fixé, nous avons cela

$$\begin{aligned} A_{n,1}^a(t) &= \int_0^{1-k/n} (Q_n^a(s) - Q^a(s)) ds - \int_0^t (Q_n^a(s) - Q^a(s)) ds \\ &= - \int_{Q(t)}^{Q(1-k/n)} (F_n^a(x) - F^a(x)) dx + V_n^a(1 - k/n) - V_n^a(t). \end{aligned} \quad (3.48)$$

C'est bien connu (voir Zitikis (1998), Davydov et Zitikis (2003), Davydov et Zitikis (2004)) $V_n(t)$ est non négatif et ne dépasse pas $-(F_n(Q(t)) - t)(Q_n^a(t) - Q^a(t))$. Puisque le cdf F est continu par hypothèse, nous avons donc

$$\sqrt{n} V_n^a(t) \leq |e_n(t)| |Q_n^a(t) - Q^a(t)|$$

où $e_n(t)$ est le processus empirique uniforme $\sqrt{n} (F_n^a(Q(t)) - F^a(Q(t)))$, qui pour un grand n ressemble au pont brownien $B_n(t)$. Notez également qu'avec la notation nouvellement introduite e_n , l'intégrale sur le côté droit de l'équation(3.48) est égale à $\int_{Q(t)}^{Q(1-k/n)} e_n(F(x)) dx$. Par conséquent,

$$\begin{aligned} \frac{\sqrt{n} A_{n,1}^a(t)}{(k/n)^{1/2} Q^a(1 - k/n)} &= \frac{\int_{Q(t)}^{Q(1-k/n)} e_n^a(F^a(x)) dx}{(k/n)^{1/2} Q^a(1 - k/n)} \\ &\quad + O_{\mathbf{P}}(1) \frac{|e_n^a(1 - k/n)| |Q_n^a(1 - k/n) - Q^a(1 - k/n)|}{(k/n)^{1/2} Q^a(1 - k/n)} \\ &\quad + O_{\mathbf{P}}(1) \frac{|e_n^a(t)| |Q_n^a(t) - Q^a(t)|}{(k/n)^{1/2} Q^a(1 - k/n)}. \end{aligned} \quad (3.49)$$

Nous remplacerons ensuite le processus empirique e_n par un pont brownien approprié B_n dans la première intégrale du côté droit de l'équation (3.49) avec un terme d'erreur de magnitude $o_{\mathbf{P}}(1)$, et nous montrerons aussi que les deuxième et troisième sommands du côté droit de l'équation(3.49) sont de l'ordre $o_{\mathbf{P}}(1)$. Le remplacement de e_n par B_n peut être accompli en utilisant, par exemple, Corollary 2.1 sur p. 48 de Csörgő et al. (1986), qui stipule que sur un espace de probabilité construit de manière appropriée et pour tout $0 \leq v < 1/4$, nous avons

$$\sup_{1/n \leq s \leq 1-1/n} \frac{|e_n(s) - B_n(s)|}{s^{1/2-\nu} (1-s)^{1/2-\nu}} = O_{\mathbf{P}}(n^{-\nu}). \quad (3.50)$$

Ce résultat est applicable dans la situation actuelle puisque nous pouvons toujours placer notre problème original dans l'espace de probabilité requis, car nos résultats principaux sont 'en probabilité'. De plus, depuis $Q(t) \leq x \leq Q(1 - k/n)$, nous avons cela $t \leq F(x) \leq 1 - k/n$. Par conséquent, statement (3.50) implique que

$$\frac{\int_{Q(t)}^{Q(1-k/n)} e_n(F^a(x)) dx}{(k/n)^{1/2} Q^a(1 - k/n)} = -\frac{\int_{Q(t)}^{Q(1-k/n)} B_n^a(F^a(x)) dx}{(k/n)^{1/2} Q^a(1 - k/n)} + O_{\mathbf{P}}(1) \frac{\int_{Q(t)}^{Q(1-k/n)} (1 - F^a(x))^{1/2-\nu} dx}{n^\nu (k/n)^{1/2} Q^a(1 - k/n)}. \quad (3.51)$$

Changer les variables d'intégration et utiliser la propriété $(k/n)^{1/2} Q^a(1 - k/n) \rightarrow \infty$ quand $n \rightarrow \infty$, nous obtenons cela

$$-\frac{\int_{Q(t)}^{Q(1-k/n)} B_n^a(F^a(x)) dx}{(k/n)^{1/2} Q^a(1 - k/n)} = -\frac{\int_0^{1-k/n} B_n^a(s) dQ^a(s)}{(k/n)^{1/2} Q^a(1 - k/n)} + o_{\mathbf{P}}(1). \quad (3.52)$$

Le terme principal sur le côté droit de l'équation (3.52) est $W_{1,n}$. Nous montrerons ensuite que le summand le plus à droite de l'équation (3.51) converge vers 0 quand $n \rightarrow \infty$.

En changeant la variable d'intégration puis en l'intégrant par parties, on obtient la borne

$$\begin{aligned} & \frac{\int_{Q(t)}^{Q(1-k/n)} (1 - F^a(x))^{1/2-\nu} dx}{n^\nu (k/n)^{1/2} Q^a(1 - k/n)} \\ & \leq \frac{(1 - s)^{1/2-\nu} Q^a(s) \Big|_t^{1-k/n}}{n^\nu (k/n)^{1/2} Q^a(1 - k/n)} + O(1) \frac{\int_t^{1-k/n} (1 - s)^{-1/2-\nu} Q^a(s) ds}{n^\nu (k/n)^{1/2} Q^a(1 - k/n)}. \end{aligned} \quad (3.53)$$

Nous voulons montrer que le côté droit de bound (3.53) converge vers 0 quand $n \rightarrow \infty$. Pour cela, nous notons d'abord que

$$\frac{(1 - s)^{1/2-\nu} Q^a(s) \Big|_t^{1-k/n}}{n^\nu (k/n)^{1/2} Q^a(1 - k/n)} = \frac{1}{k^\nu} - \frac{(1 - t)^{1/2-\nu} Q^a(t)}{n^\nu (k/n)^{1/2} Q^a(1 - k/n)} \rightarrow 0. \quad (3.55)$$

Ensuite, avec la notation $\phi(u) = Q^a(1 - u)/u^{1/2+\nu}$ nous avons cela

$$\frac{\int_t^{1-k/n} (1 - s)^{-1/2-\nu} Q^a(s) ds}{n^\nu (k/n)^{1/2} Q^a(1 - k/n)} = \frac{1}{k^\nu} \frac{\int_{k/n}^{1-t} \phi(s) ds}{\phi(k/n)} \rightarrow 0 \quad (3.56)$$

lorsque $n \rightarrow \infty$, où la convergence à 0 découle du résultat 1 de l'annexe de Necir et Meraghi (2009). En prenant des déclarations (3.53) – (3.56) ensemble, nous avons cela la sommation la plus à droite de l'équation (3.51)

converge vers 0 lorsque $n \rightarrow \infty$. Par conséquent, afin de compléter la preuve de déclaration (3.46), il nous reste à montrer que les deuxième et troisième sommants du côté droit de l'équation (3.49) sont de l'ordre $o_{\mathbf{P}}(1)$.

Le troisième est de l'ordre $o_{\mathbf{P}}(1)$ car $|e_n^a(t)| |Q_n^a(t) - Q^a(t)| = O_{\mathbf{P}}(1)$ et $(k/n)^{1/2} Q^a(1 - k/n) \rightarrow \infty$. Par conséquent, il ne nous reste plus qu'à montrer que le deuxième sommet sur le côté droit de l'équation (3.49) est de l'ordre $o_{\mathbf{P}}(1)$, pour lequel nous montrerons que

$$\frac{|e_n^a(1 - k/n)|}{(k/n)^{1/2}} \left| \frac{Q_n^a(1 - k/n)}{Q^a(1 - k/n)} - 1 \right| = o_{\mathbf{P}}(1). \quad (3.57)$$

Pour prouver la déclaration (3.57), nous notons d'abord que

$$\frac{|e_n^a(1 - k/n)|}{(k/n)^{1/2}} \leq \frac{|e_n^a(1 - k/n) - B_n^a(1 - k/n)|}{(k/n)^{1/2}} + \frac{|B_n^a(1 - k/n)|}{(k/n)^{1/2}}. \quad (3.58)$$

La première sommelle sur le côté droit de bound (3.58) est de l'ordre $O_{\mathbf{P}}(1)$ en raison de la déclaration (3.50) avec $v = 0$. Le deuxième sommet sur le côté droit de bound (3.58) est de l'ordre $O_{\mathbf{P}}(1)$ en raison d'une déclaration sur p. 49 de Csörgö et al. (1986) (voir la déclaration ci-dessous liée ci-dessous (2.39)). Par conséquent, pour compléter la preuve de déclaration (3.57), nous devons vérifier que

$$\frac{Q_n^a(1 - k/n)}{Q^a(1 - k/n)} = 1 + o_{\mathbf{P}}(1). \quad (3.59)$$

Observez que, pour chaque n , la distribution de $Q_n^a(1 - k/n)$ est la même que celle de $Q^a(E_n^{-1}(1 - k/n))$, où E_n^{-1} est la fonction quantile empirique uniforme. De plus, les processus $\{1 - E_n^{-1}(1 - s), 0 \leq s \leq 1\}$ et $\{E_n^{-1}(s), 0 \leq s \leq 1\}$ sont égaux dans la distribution. Par conséquent, l'instruction (3.59) équivaut à

$$\frac{Q^a(1 - E_n^{-1}(k/n))}{Q^a(1 - k/n)} = 1 + o_{\mathbf{P}}(1). \quad (3.60)$$

Du théorème de Glivenko-Cantelli nous avons cela $E_n^{-1}(k/n) - k/n \rightarrow 0$ presque sûrement, ce qui implique aussi que $E_n^{-1}(k/n) \rightarrow 0$ depuis $k/n \rightarrow 0$ par notre choix de k . De plus, nous savons d'après Théorème 0 et Remarque 1 de Wellner (1978) que

$$\sup_{1/n \leq s \leq 1} s^{-1} |E_n^{-1}(s) - s| = o_{\mathbf{P}}(1), \quad (3.61)$$

à partir de laquelle nous concluons que

$$nE_n^{-1}(k/n)/k = 1 + o_{\mathbf{P}}(1). \quad (3.62)$$

Depuis la fonctions $s \mapsto Q(1-s)$ varie lentement à zéro, en utilisant l'inégalité de Potter (voir la 5^{eme} affirmation de la Proposition B.1.9 à la page 367 de de Haan et Ferreira (2006)), nous obtenons que

$$\frac{Q^a(1 - E_n^{-1}(k/n))}{Q^a(1 - k/n)} = (1 + o_{\mathbf{P}}(1)) (nE_n^{-1}(k/n)/k)^{-\gamma \pm \theta} \quad (3.63)$$

pour tout $\theta \in (0, \gamma)$. Au vu de (3.62), le membre de droite de l'équation (3.63) est égal à $1 + o_{\mathbf{P}}(1)$, ce qui implique statement (3.60) et termine ainsi la preuve de l'instruction (3.46). En vue de (3.62), le côté droit de l'équation (3.63) est égal à $1 + o_{\mathbf{P}}(1)$, ce qui implique une déclaration (3.60) et termine ainsi la preuve de déclaration (3.46).

3.3.3 Preuve de l'expression (3.47).

La preuve de déclaration (3.47) est similaire à celle du théorème 2 de Necir et al. (2007), bien que certains ajustements soient nécessaires étant donné que nous nous intéressons maintenant à la mesure du risque de CTM. Nous présentons donc les blocs principaux de la preuve avec des références précises à Necir et al. (2007) pour des détails techniques spécifiques.

Nous commençons la preuve avec la fonction $\mathbb{U}(z) = Q^a(1 - 1/z)$ qui était déjà utilisé dans la formulation du théorème (3.3). Par conséquent, si Y est une variable aléatoire avec la fonction de distribution $G(z) = 1 - 1/z$, $z \geq 1$, puis $\mathbb{U}(Y) = Q(G(Y)) =_d X$ depuis $G(Y)$ est un uniforme sur l'intervalle $[0, 1]$ Variable aléatoire. Par conséquent,

$$A_{n,2}^a = \frac{k/n}{1 - a\hat{\gamma}_n} \mathbb{U}(Y_{n-k:n}) - \int_0^{k/n} \mathbb{U}(1/s) ds,$$

et nous avons donc

$$\begin{aligned} \frac{\sqrt{n}A_{n,2}^a}{(k/n)^{1/2}Q^a(1 - k/n)} &= \sqrt{k} \left(\frac{1}{1 - a\hat{\gamma}_n} \frac{\mathbb{U}(Y_{n-k:n})}{\mathbb{U}(n/k)} - \frac{(n/k) \int_0^{k/n} \mathbb{U}(1/s) ds}{\mathbb{U}(n/k)} \right) \\ &= \sqrt{k} \left(\frac{1}{1 - a\hat{\gamma}_n} \frac{\mathbb{U}(Y_{n-k:n})}{\mathbb{U}(n/k)} - \frac{1}{1 - a\gamma} \right) \\ &\quad + \sqrt{k} \left(\frac{1}{1 - a\gamma} + \frac{\int_1^\infty s^{-2} \mathbb{U}(ns/k) ds}{\mathbb{U}(n/k)} \right). \end{aligned} \quad (3.64)$$

Nous montrons ensuite que le terme le plus à droite dans l'équation (3.64) converge vers 0 quand $n \rightarrow \infty$. Pour cette raison, nous réécrivons d'abord le terme comme suit :

$$\sqrt{k} \left(\frac{1}{1 - a\gamma} + \frac{\int_1^\infty s^{-2} \mathbb{U}(ns/k) ds}{\mathbb{U}(n/k)} \right) = \sqrt{k} \int_1^\infty \frac{1}{s^2} \left(\frac{\mathbb{U}(ns/k)}{\mathbb{U}(n/k)} - s^\gamma \right) ds. \quad (3.65)$$

Le côté droit de l'équation (3.65) converge vers 0 (voir les notes de la page 149 de Necir et al., 2007) en raison de la condition de second ordre (3.39), qui peut être réécrit de manière équivalente

$$\lim_{z \rightarrow \infty} \frac{1}{A(z)} \left(\frac{\mathbb{U}(zs)}{\mathbb{U}(z)} - s^\gamma \right) = s^\gamma \frac{s^\rho - 1}{\rho}$$

pour chaque $s > 0$, pour chaque $A(z) = \gamma^2 a(\mathbb{U}(z))$. Notez que $\sqrt{k} A(n/k) \rightarrow 0$ quand $n \rightarrow \infty$. Par conséquent, afin de compléter la preuve de déclaration (3.47), nous devons vérifier cela

$$\begin{aligned} \sqrt{k} \left(\frac{1}{1 - a\hat{\gamma}_n} \frac{\mathbb{U}(Y_{n-k:n})}{\mathbb{U}(n/k)} - \frac{1}{1 - a\gamma} \right) &= \frac{\gamma^2}{(1 - a\gamma)^2} \sqrt{\frac{n}{k}} B_n^a(1 - k/n) \\ &\quad - \frac{\gamma}{(1 - a\gamma)^2} \sqrt{\frac{n}{k}} \int_{1-k/n}^1 \frac{B_n^a(s)}{1 - s} ds + o_{\mathbf{P}}(1). \end{aligned} \quad (3.66)$$

Avec l'estimateur de Hill écrit dans le formulaire

$$\hat{\gamma}_n = \frac{1}{k} \sum_{i=1}^k \log \left(\frac{\mathbb{U}(Y_{n-i+1:n})}{\mathbb{U}(Y_{n-k:n})} \right),$$

nous procédons à la preuve de déclaration (3.66) comme suit :

$$\begin{aligned} \sqrt{k} \left(\frac{1 - \gamma}{1 - a\hat{\gamma}_n} \frac{\mathbb{U}(Y_{n-k:n})}{\mathbb{U}(n/k)} - 1 \right) &= \sqrt{k} \frac{1 - \gamma}{1 - a\hat{\gamma}_n} \left(\frac{\mathbb{U}(Y_{n-k:n})}{\mathbb{U}(n/k)} - \left(\frac{Y_{n-k:n}}{(n/k)^a} \right)^\gamma \right) \\ &\quad + \sqrt{k} \frac{1 - \gamma}{1 - a\hat{\gamma}_n} \left(\left(\frac{Y_{n-k:n}}{(n/k)^a} \right)^\gamma - 1 \right) + \sqrt{k} \frac{a\hat{\gamma}_n - \gamma}{1 - a\hat{\gamma}_n}. \end{aligned} \quad (3.67)$$

De plus, nous avons cela

$$\begin{aligned} \sqrt{k} \frac{a\hat{\gamma}_n - \gamma}{1 - a\hat{\gamma}_n} &= \frac{1}{1 - a\hat{\gamma}_n} \frac{1}{\sqrt{k}} \sum_{i=1}^k \left(\log \left(\frac{\mathbb{U}(Y_{n-i+1:n})}{\mathbb{U}(Y_{n-k:n})} \right) - \gamma \log \left(\frac{Y_{n-i+1:n}}{Y_{n-k:n}} \right) \right) \\ &\quad + \frac{\gamma}{1 - a\hat{\gamma}_n} \frac{1}{\sqrt{k}} \sum_{i=1}^k \left(\log \left(\frac{Y_{n-i+1:n}}{Y_{n-k:n}} \right) - 1 \right). \end{aligned} \quad (3.68)$$

Arguments sur p. 156 de Necir et al. (2007) impliquent que le premier terme du côté droit de l'équation (3.68) est de l'ordre $O_{\mathbf{P}}(\sqrt{k} A(Y_{n-k:n}))$, et une note sur p. 157 de Necir et al. (2007) dit que $\sqrt{k} A(Y_{n-k:n}) = o_{\mathbf{P}}(1)$. Par conséquent, le premier terme sur le côté droit de l'équation (3.68) est de l'ordre $o_{\mathbf{P}}(1)$. Considérations analogues mais maintenant en utilisant bound

(2.5) au lieu de (2.4) sur p. 156 de Necir et al. (2007) impliquent que le premier terme du côté droit de l'équation (3.67) est de l'ordre $o_{\mathbf{P}}(1)$. Par conséquent, en résumé, nous avons cela

$$\begin{aligned} \sqrt{k} \left(\frac{1 - \gamma}{1 - a\hat{\gamma}_n} \frac{\mathbb{U}(Y_{n-k:n})}{\mathbb{U}(n/k)} - 1 \right) &= \frac{1 - \gamma}{1 - a\hat{\gamma}_n} \sqrt{k} \left(\left(\frac{Y_{n-k:n}}{(n/k)^a} \right)^\gamma - 1 \right) \\ &+ \frac{\gamma}{1 - a\hat{\gamma}_n} \frac{1}{\sqrt{k}} \sum_{i=1}^k \left(\log \left(\frac{Y_{n-i+1:n}}{Y_{n-k:n}} \right) - 1 \right) + o_{\mathbf{P}}(1). \end{aligned} \quad (3.69)$$

Nous devons maintenant connecter le côté droit de l'équation (3.69) avec des ponts browniens B_n . À cette fin, nous convertissons d'abord le Y -les statistiques d'ordre basé sur U -basée (c'est-à-dire uniforme sur $[0, 1]$) statistiques de commande. Pour cela, nous rappelons que la cdf de Y est G , Et ainsi Y est égal à la distribution à $G^{-1}(U)$, lequel est $1/(1 - U)$. par conséquent,

$$\begin{aligned} \sqrt{k} \left(\frac{1 - \gamma}{1 - a\hat{\gamma}_n} \frac{\mathbb{U}(Y_{n-k:n})}{\mathbb{U}(n/k)} - 1 \right) &= \frac{1 - \gamma}{1 - a\hat{\gamma}_n} \sqrt{k} \left(\left(\frac{1}{(n/k)^a (1 - U_{n-k:n})} \right)^\gamma - 1 \right) \\ &+ \frac{\gamma}{1 - a\hat{\gamma}_n} \frac{1}{\sqrt{k}} \sum_{i=1}^k \left(\log \left(\frac{(1 - U_{n-k:n})}{(1 - U_{n-i+1:n})} \right) - 1 \right) + o_{\mathbf{P}}(1). \end{aligned} \quad (3.70)$$

Ensuite, nous choisissons une séquence de ponts browniens B_n (voir pages 158-159 dans Necir et al. (2007) et les références qui y sont citées) de telle sorte que les deux représentations asymptotiques suivantes tiennent :

$$\sqrt{k} \left(\left(\frac{1}{(n/k)^a (1 - U_{n-k:n})} \right)^\gamma - 1 \right) = -\gamma \sqrt{\frac{n}{k}} B_n^a(1 - k/n) + o_{\mathbf{P}}(1)$$

et

$$\begin{aligned} \frac{1}{\sqrt{k}} \sum_{i=1}^k \left(\log \left(\frac{(1 - U_{n-k:n})}{(1 - U_{n-i+1:n})} \right) - 1 \right) &= \sqrt{\frac{n}{k}} B_n(1 - k/n) \\ &- \sqrt{\frac{n}{k}} \int_{1-k/n}^1 \frac{B_n(s)}{1-s} ds + o_{\mathbf{P}}(1). \end{aligned}$$

En utilisant ces deux déclarations sur le côté droit de l'équation (3.70) et aussi en gardant à l'esprit que $\hat{\gamma}_n$ est un estimateur cohérent de γ (voir Mason (1982)), nous avons cela

$$\begin{aligned} \sqrt{k} \left(\frac{1 - \gamma}{1 - a\hat{\gamma}_n} \frac{\mathbb{U}(Y_{n-k:n})}{\mathbb{U}(n/k)} - 1 \right) &= \frac{\gamma^2}{1 - a\gamma} \sqrt{\frac{n}{k}} B_n^a(1 - k/n) \\ &- \frac{\gamma}{1 - \gamma} \sqrt{\frac{n}{k}} \int_{1-k/n}^1 \frac{B_n(s)}{1-s} ds + o_{\mathbf{P}}(1). \end{aligned} \quad (3.71)$$

Diviser les deux côtés de l'équation (3.71) par $1 - \gamma$, nous arrivons à l'équation (3.66). Ceci complète la preuve de déclaration(3.47) et de Théorème 3.3 ainsi que.

3.4 Simulations

Pour illustrer la performance de cet mesure nous avons effectué une petite étude de simulation basée sur la distribution de Pareto $F(x) = 1 - x^{-1/\gamma}$, $x > 1$, avec l'indice de queue $\gamma = 2/3$ et $\gamma = 3/4$ et on prend $\alpha = 0.95$ comme valeur fixé. Nous avons généré 500 indépendants répliques de trois échantillons de tailles $n = 500, 1000, 2000$.

Dans notre étude actuelle, nous employons la méthode de Cheng et Peng 2001, pour avoir la valeur appropriée k_{opt} de paramètre k , ayant calculé l'estimateur de Hill et déterminé par conséquent l'observation $X_{n-k_{opt}:n}$, nous calculons ensuite les valeurs correspondantes de $\hat{C}_n^a(t)$.

Pour chaque échantillon simulé, nous avons obtenu des estimations $\hat{C}_n^a(t)$, nous avons calculé les moyennes arithmétiques sur les valeurs des 500 répliques, avec l'erreur absolue (*error*), et le Biais du nouvel estimateur $\hat{C}_n^a(t)$ rapporté dans le tableau 1 ($\gamma = 3/4$) et le tableau 2 ($\gamma = 2/3$) avec $a = 1$ pour la distribution Pareto après les calculs on obtient une valeur théorique fixe pour tout les échantillons, et des valeurs semi paramétrique avec le biais pour chaque échantillon .

TAB. 3.3 – Simulation du CTM_1 pour distribution de Pareto avec $\gamma = 3/4$ et $t = 0.95$ sachant que la valeur théorique est 29.822979

n	k_{opt}	CTM	erreur	LCB	UCB	cov	length
500	25	30.47157	0.641779	21.3363	39.6066	0.788	18.1
1000	46	28.96359	0.582416	19.84526	34.70406	0.897435	11.4809
2000	98	28.90842	0.03088773	18.92141	33.09218	0.907489	8.367534

TAB. 3.4 – Simulation du CTM_1 pour distribution de Pareto avec $\gamma = 2/3$ et $t = 0.95$ sachant que la valeur théorique est 19.31929

n	k_{opt}	CTM	erreur	LCB	UCB	cov	length
500	26	20.74979	0.074045	14.74166	27.02793	0.920380	12.5562
1000	51	19.74367	0.062207	12.00948	26.03278	0.975221	11.0.23375
2000	99	19.53764	0.0113025	13.219509	22.8602	0.982062	6.6451081

3.5 Conclusion générale

Dans ce mémoire, on a étudié les deux approches de la théorie des valeurs extrêmes, maximum et POT méthode, l'estimation des paramètres des distribution généralisées des valeurs extrêmes (GEV) et les distributions de Paréto généralisées, l'estimation des indices des distributions et les quantiles extrêmes. Les mesures de risques et ses importance dans les domaines de finance et assurance, telle que VaR, CVaR, ES,.... l'étude des propriétés des mesures de risque notamment la propriété de cohérence qui est importante pour les praticiens. Le mesure de risque Moment de la queue conditionnel CTM qui comporte un classe importante des mesures comme CTE, CTVaR, CS, CK. On a étudier les estimateurs existants et proposé des estimateurs empiriques et semi paramétrique pour le CTM pour des distributions à queues lourdes. Le comportement asymptotique de ces estimateurs sont montré et illustrés par quelques outils de simulations pour des différents modèles dangereux, à savoir, Pareto.

Bibliographie

- [1] Acerbi, Carlo, and Dirk Tasche. "Expected shortfall : a natural coherent alternative to Value at Risk." *Economic notes* 31.2 (2002) : 379-388.
- [2] ARTZNER, Philippe, DELBAEN, Freddy, EBER, Jean-Marc, et al. Coherent measures of risk. *Mathematical finance*, 1999, vol. 9, no 3, p. 203-228.
- [3] Beirlant J., Goegebeur Y., Teugels J. and Segers J., *Statistics of extremes*. Wiley Series in Probability and Statistics. John Wiley and Sons Ltd, Chichester, 2004.
- [4] Cheng, S. et Peng, L. (2001). Confidence Intervals for the Tail Index. *Bernoulli* 7, 751-760.
- [5] Csorgo, Miklos, et al. "Weighted empirical and quantile processes." *The Annals of Probability* (1986) : 31-85.
- [6] Davison, A. et Smith, R. (1990). Models for exceedances over high thresholds. *Journal of the Royal Statistical Society Series B*, 52(3) :393-442.
- [7] Davydov, Youri, and Ričardas Zitikis. "Convex rearrangements of random elements." *Asymptotic Methods in Stochastics* 44 (2004) : 141-171.
- [8] Davydov, Youri, and Ričardas Zitikis. "Generalized Lorenz curves and convexifications of stochastic processes." *Journal of Applied Probability* 40.4 (2003) : 906-925.
- [9] de Haan, L. et Ferreira, A. (2006). *Extreme Value Theory : An Introduction*. Springer Series in Operations Research and Financial Engineering, New York Inc.
- [10] de Haan L. and Rootzen H., On the estimation of high quartiles J. of *Statistical Planning and Inference*, 35, n.1, 1-13, 1993.
- [11] Dekkers A., Einmahl J. and De Haan L., A moment estimator for the index of an extreme-value loi, *Ann. Statist.* 17, p. 1833-1855, 1989.
- [12] Drees, Holger. "Extreme quantile estimation for dependent data, with applications to finance." *Bernoulli* 9.4 (2003) : 617-657.

- [13] Embrechts P., Kluppelberg C. and Mikosch T., Modelling extremal events, Springer Verlag, Berlin, 1997.
- [14] Galambos, Janos. The asymptotic theory of extreme order statistics. No. 04 ; QA274, G3.. 1978.
- [15] Greenwood J., Landweher J., Matalas N. and Wallis J., Probability weighted moments : Definition and relation to parameters of several lois expressible in inverse form, Water Resources Research 15, p. 1049-1054, 1979.
- [16] Gnedenko B., Sur la distribution limite du terme maximum d'une série aléatoire, Ann. Math., 44(3), 423-453, 1943.
- [17] Gumbel E.J. and Mustafi C.K., Some analytical properties of bivariate extremal distributions, Am. Stat. Assoc. J., 62, p.569-589, 1967.
- [18] Hill B. M., A simple general approach to inference about the tail of a Loi, Annals of Statistics 5, p. 1163-1174, 1975.
- [19] Hosking, J., Wallis, J. et Wood, E. (1985). Estimation of the generalized extreme-value distribution by the method of probability-weighted moments. Technometrics, 27(3) :251–261.
- [20] Hsing, Tailen. "On tail index estimation using dependent data." The Annals of Statistics (1991) : 1547-1569.
- [21] Jenkinson A.F., The frequency distribution of the annual maximum (or minimum) values of meteorological events, Journal of the Royal Meteorological Society , 81, p. 158-272, 1955.
- [22] Katz R.W, Techniques for estimating Uncertainty in climate change Scenarios and impact studies Climate research, 20, 167-185, 2002.
- [23] Lubes H. and Masson J., Méthode des moments de probabilité pondérés. Application à la loi de Jenkinson, Hydrol. continent., vol. 6, n 1, p. 67-84, 1991.
- [24] Mc Neil A.J. and Frey R, Estimation of tail-related risk measures for heteroscedastic financial time series : An extreme value approach Journal of Empirical Finance, 7, p. 271-300, 2002.
- [25] Necir, A., Rassoul, A. et Zitakis, R. (2010). Estimating the conditional tail expectation in the case of heavy-tailed losses. Journal of Probability and Statistics, ID 596839 :17 pages.
- [26] Pickands J., Statistical inference using extreme order statistics, Annals of Statistics 3, p. 119-131, 1975.
- [27] Prescott P. and Walden A.T., Maximum likelihood estimation of the parameters of the three-parameter generalized extreme-value distribution

- for censored samples, *Journal of Statistical Computation and Simulation*, 16(3-4) : 241 – 250, 1983.
- [28] Prescott P. and Walden A.T., Maximum likelihood estimation of the parameters of the generalized extreme-value distribution, *Biometrika* ,67, 723-724, 1980.
- [29] Raggad B., *Fondements de la théorie des valeurs extrêmes, ses principales applications et son apport à la gestion des risques du marché pétrolier*, Mathematics and Social Sciences. p. 29-63, 2009.
- [30] Reiss R.D. and Thomas M., *Statistical Analysis of Extreme Values with Applications to Insurance, Finance, Hydrology and Other Fields*, Birkhauser Basel. 2nd edition, 2001.
- [31] Resnick Si, *Extreme Values, Regular variation and Point Processes*, Springer-Verlag, New-York, 1987.
- [32] Shorack, G. R., and J. A. Wellner. "Empirical Processes with Applications to Statistics Wiley Series in Probability and Mathematical Statistics : Probability and Mathematical Statistics." New York (1986).
- [33] Smith R.L., Estimating tails of probability lois. *The Annals of Statistics* 3, p. 1174-1207, 1987.
- [34] Smith R.L., Maximum likelihood estimation in a class of nonregular cases. *Biometrika* 72, p. 67-90, 1985.
- [35] Von Mises R., La distribution de la plus grande de n valeurs, *Revue de Mathématique, Union Interbalcanique*, p. 141-160, 1936.
- [36] Von Mises R., La distribution de la plus grande de n valeurs, *American Mathematical Society, RI, USA*, vol. II, pp. 271-294, 1954.
- [37] S.S. Wang. Insurance pricing and increased limits ratemaking by proportional hazards transforms. *Insurance Math. Econom.*, 17(1) :43-54, 1995.
- [38] S.S. Wang. Ordering of risks under PH-transforms. *Insurance Math. Econom.*, 18(2) :109-114, 1996.
- [39] S.S. Wang and V.R. Young. Risk-adjusted credibility premiums using distorted probabilities. *Scand. Actuar. J.*, (2) :143-165, 1998.
- [40] Weissman I., Estimation of Parameters and Large Quantiles Based on the k-Largest Observations, *J. Amer. Statist. Assoc*, 73, 812-815, 1978.
- [41] Wellner, Jon A. "Limit theorems for the ratio of the empirical distribution function to the true distribution function." *Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete* 45.1 (1978) : 73-88.
- [42] Zitikis, Ričardas. "-The Vervaat process." *Asymptotic methods in probability and statistics*. 1998. 667-694.

