

Estimation en théorie des valeurs extrêmes

Mémoire de mes recherches autodidactes sur la théorie des valeurs extrêmes

Par

KANE HAMIDOU

Réalisé en 2024

* Student at IMT Atlantique, hamidou.kane@imt-atlantique.net

Remerciements

Ce document rassemble des développements et démonstrations en **théorie des valeurs extrêmes** (EVT). Bien qu'il s'agisse d'un travail à dominante mathématique, ma motivation profonde est ma **passion pour la finance quantitative et les statistiques** : comprendre et maîtriser les outils permettant d'appréhender les **risques extrêmes** (queues lourdes, quantiles extrêmes) afin d'améliorer l'évaluation et la **gestion du risque** en finance, notamment à travers la *Value-at-Risk* (*VaR*) et l'*Expected Shortfall* (*ES*).

Je remercie chaleureusement mon ami **Polneau Neba Ibrahim**, étudiant à l'École polytechnique de paris, **médaille d'or aux Olympiades panafricaines de mathématiques (champion d'Afrique, 2017)** et **mention honorable aux Olympiades internationales de mathématiques (IMO 2018)**. Ses conseils, sa rigueur et la **passion de la recherche** qu'il m'a transmise au fil de nos séances consacrées aux démonstrations des grands résultats et théorèmes m'ont **donné le goût des mathématiques**, appris la **résilience** et offert des **méthodes de travail** solides.

Je remercie également mon ami **Akpa Simon Pierre**, étudiant à l'ESTP Paris : **nous préparions ensemble, de manière rigoureuse et complète, les stages de préparation aux Olympiades de mathématiques en Côte d'Ivoire**, ce qui a renforcé mon exigence, mon goût de l'effort et nourri mon attrait pour l'abstraction.

*Sans l'environnement exigeant des olympiades et sans ces amitiés, je n'aurais sans doute pas développé la même passion qu'aujourd'hui. La fréquentation régulière de problèmes souvent abstraits, parfois peu balisés par la théorie et les méthodes classiques, a allumé en moi la flamme d'explorer l'inconnu, renforcé ma curiosité et profondément développé ma capacité de **résolution de problèmes**. C'est cet état d'esprit — fait de persévérance, de précision et d'ouverture — que je mobilise désormais dans mes travaux.*

Table des matières

Remerciements	ii
Table des matières	iii
Table des figures	vi
Liste des tables	vii
Introduction	1
1 Loi de statistique d'ordre	3
1.1 Notions générales de probabilité et statistique	3
1.1.1 Espace probabilisé	3
1.1.2 Variable aléatoire et lois de probabilité	4
1.2 Statistique d'ordre	5
1.2.1 Définition de la statistique d'ordre	5
1.2.2 Loi de la statistique d'ordre	7
1.2.3 Densité conjointe de deux statistique d'ordre	7
1.3 Distribution empirique et quantile	8
1.3.1 Inverse généralisé	8
1.3.2 La fonction de répartition empirique	9

1.3.3	Point terminal	9
1.3.4	Convergence de la fonction de répartition empirique	10
1.3.5	Quantile d'ordre p	10
1.3.6	Quantile extrême	10
1.3.7	Quantile empirique	11
1.3.8	Lois des grands nombres	12
1.3.9	Théorème central limite	12
2	Distribution des valeurs extrêmes généralisées	13
2.1	Distribution des valeurs extrêmes usuelles	13
2.1.1	Distribution de probabilité de Fréchet	13
2.1.2	Distribution de probabilité de Weibull	16
2.1.3	Distribution de probabilité de Gumbel	18
2.2	Comportement asymptotique du maximum	19
2.3	Caractérisation des domaines d'attraction et Conditions nécessaire et suffisantes de	
	convergence	22
2.3.1	Notion de fonction à variation régulière	22
2.3.2	Domaines d'attraction	23
2.4	Constantes de normalisations	25
2.5	Distribution des excès	26
2.6	Distribution de Pareto Généralisée (GPD)	26
3	L'estimation de l'indice de queue	29
3.1	Introduction	29
3.2	L'estimation de l'indice de queue par des méthodes non paramétrique	30
3.2.1	L'estimateur de pickands	30

3.2.2	Estimateur de moment	32
3.2.3	L'estimateur de Hill	33
3.2.4	Comparaison des différents estimateurs	35
3.3	L'estimation de l'indice de queue par des méthodes paramétrique	36
3.3.1	Méthode de maximum de vraisemblance	36
3.3.2	Méthode des moments	39
Conclusion		43
Bibliographie		44
Annexe : Abréviations et Notations		46

Table des figures

2.1	Fonction de densité de Fréchet : (a) pour différentes valeur de μ , (b) pour différentes	
	valeur de σ	14
2.2	Fonction de densité de Weibull : (a) pour différentes valeurs de μ , (b) pour différentes	
	valeurs de σ	16
2.3	Fonction de densité de Gumbel : (a) pour différentes valeurs de μ , (b) pour différentes	
	valeurs de σ	18
2.4	Densité de probabilité des valeurs extrêmes avec différentes valeurs de γ	22
2.5	Dépassement du seuil (POT)	26
2.6	Distribution GPD	28
3.1	Estimateur de Pickands , en fonction du nombres des extrêmes avec l'intervalle de	
	confiance 95%, pour la distribution uniforme ($\gamma = 1$) basé sur 100 échantillons de	
	3000 observations.	31
3.2	Estimateur du Moment, en fonction du nombres des extrêmes avec l'intervalle de	
	confiance 95%, pour la distribution de Burr(1,1,1) basée sur 100 échantillons de	
	3000 observations.	33
3.3	Estimateur de Hill , en fonction du nombres des extrêmes avec l'intervalle de confiance	
	95% , pour la distribution uniforme ($\gamma = 1$) basée sur 100 échantillons de 3000 ob-	
	servations.	35
3.4	Comparaison de l'erreur d'estimation en fonction de la méthode.	36

Liste des tableaux

2.1 Domaines d'attraction et quelques lois associées	25
--	----

Introduction

La modélisations des évènements extrêmes (tels les ouragan, les tremblements de terre, les inondations, les crues, les crises financières, Krachs, chocs pétrolier, les tempêtes, ...) est aujourd'hui un champ de recherches particulièrement actif , notamment par l'importance de leurs impacts économiques et sociaux.

Ces évènements extrêmes peuvent causer des dégâts humains et matériels considérables. De telles catastrophes ne peuvent pas toujours être évitées. Cependant, la société peut prendre des actions préventives pour minimiser leurs effets. Pour cela, les spécialistes ont à leur disposition la théorie statistique des valeurs extrêmes. Elle donne un résultat très intéressant .

La Théorie des Valeurs Extrêmes est une branche des statistiques apparue entre 1920 et 1943 grâce à Fréchet (1927) ; Fisher et Tippett (1928) et Gnedenko (1943). Cette branche a connue un grand essor depuis les années cinquante bien que le théorème principal ait été énoncé par Fisher et Tippett dès 1928 . Elle a pour but de modéliser et de décrire la survenue et l'intensité des événements extrêmes, c'est à dire qui présentent des variations de très grande amplitude (ayant une faible probabilité d'apparition).

Au cours des dernières décennies, nous avons pu observer dans le monde scientifique un intérêt renouvelé pour la modélisation et l'analyse statistique des événements rares. L'étude de ces événements revient à l'analyse des queues de distribution des fonctions ou de manière équivalente, à l'analyse de la plus grande ou de la plus petite observation d'un échantillon. En ce sens, nous pouvons considérer la théorie des valeurs extrêmes comme la contrepartie de la théorie statistique, qui est principalement basée sur l'étude des tendances centrales d'un échantillon plutôt que sur

des observations extrêmes.

La théorie des valeurs extrêmes est largement utilisée, elle a trouvé des champs d'application en :

- Hydrologie, domaine dans lequel la prévision des crues et la connaissance de leurs débits est particulièrement importante pour la conception des aménagements des cours d'eau et la protection des zones urbaines inondables, souvent théâtre d'épisodes de crues violentes causant d'importants dommages économiques.
- Climatologie : étude des événements climatiques extrêmes (précipitations, températures, chutes de neige), modélisation des grands feux de forêt.
- Assurance : survenue des sinistres d'intensité exceptionnelle (ouragan Katrina en 2005, importants incendies en risque industriels, sinistres grave en responsabilité civile automobile) qui peuvent avoir des conséquences négatives sur les résultats et la solvabilité des organismes d'assurance.
- Finance : fortes variations du cours d'actifs financiers, gestion risque opérationnel des banques (la crise fin des années 2000)
- L'industrie, où l'on s'intéresse à la fiabilité des structures.

Ce travail est organisé en trois chapitres, le premier chapitre rappelle les notions générales de probabilité et statistique et la définition des statistiques d'ordre d'un échantillon de taille n . suivis du chapitre deux qui est consacré aux concepts fondamentaux de la théorie des valeurs extrêmes dans le cas univarié et essentiellement dans le cas i.i.d. On expose les résultats asymptotique de base, notamment le théorème de convergence de Fisher-Tippett , ceci conduira à donner des critères d'appartenance aux différents domaines d'attraction. Enfin le dernier chapitre, on s'intéresse à l'estimation de l'indice des valeurs extrêmes par des méthodes paramétriques et semi paramétriques.

Nous souhaitons que toutes les personnes désireuses étudier ou mettre à jour leurs connaissances en "L'estimation en Théorie Des Valeurs Extrêmes " trouveront la lecture de ce mémoire très utile et agréable.

Chapitre 1

Loi de statistique d'ordre

L'étude des statistiques d'ordre et des valeurs extrêmes nécessitent un rappel sur les généralités et notions élémentaires sur les probabilités et les statistiques, et bien sûr les premières bases sont : l'espace probabilisé, les variables aléatoires et leur distributions .

1.1 Notions générales de probabilité et statistique

1.1.1 Espace probabilisé

Soit Ω l'ensemble de tous les résultats possibles lié à une expérience considérée si on le muni d'une tribu \mathcal{A} de sous ensembles de Ω , on dit alors que (Ω, \mathcal{A}) est un espace probabilisable .

On appelle probabilité toute application P de \mathcal{A} dans \mathbb{R} , vérifiant les propriétés :

1. $P(\Omega) = 1$.
2. Si $(B_n)_{n \in \mathbb{N}}$ est une suite (éventuellement finie) d'éléments de \mathcal{A} , deux à deux incompatibles (i.e. $B_i \cap B_j = \emptyset$ si $i \neq j$), alors :

$$P\left(\bigcup_{n \in \mathbb{N}} B_n\right) = \sum_{n \in \mathbb{N}} P(B_n).$$

Alors le triplet (Ω, \mathcal{A}, P) s'appelle espace de probabilité ou espace probabilisé .

1.1.2 Variable aléatoire et lois de probabilité

Variable aléatoire

Définition 1.1.1 *On appelle variable aléatoire, l'application mesurable de (Ω, A) dans U noté X , et selon l'ensemble d'arrivé U , on dit que X est discrète si U est fini ou au plus dénombrable ($U \subseteq \mathbb{N}^*$), elle est dite continue si U est une partie de \mathbb{R} ou \mathbb{R} ($U \subseteq \mathbb{R}$).*

Distribution d'une variable aléatoire

L'application définie de U vers $[0, 1]$ est la loi de probabilité de X .

Soit X une variable aléatoire, on note par $P(X \leq x)$ la fonction cumulative ou fonction de répartition de X , cette fonction est noter par $F_X(x)$. Dans le cas d'une variable aléatoire discrète, la fonction de densité évaluée en x représente la probabilité que la variable X prenne la valeur x donc $P(X = x)$. Dans le cas d'une variable aléatoire continue sa fonction de densité notée $f_X(x)$ doit être multipliée par une distance infiniment petite dx pour obtenir une probabilité $\left(f_X(x) = \frac{dF_X(x)}{dx}\right)$.

Caractéristiques d'une variable aléatoire

L'esperance d'une variable aléatoire notée $E(X)$ est définie comme suit :

$$E(X) = \int_{\mathbb{R}} x f_X(x) dx$$

et $Var(X)$ représente la variance de X telle que :

$$\begin{aligned} Var[X] &= E[(X - E[X])^2] \\ &= E[X^2] - E[X]^2 \end{aligned}$$

L'écart type σ soit la racine carrée de la variance est une mesure de la variabilité d'une v.a et peut

être vu comme une mesure de risque. En finance, des rendements avec un plus grand écart type sont souvent vus comme plus risqués.

1.2 Statistique d'ordre

pour commencer notre étude et les explications de la théorie des valeurs extrêmes, il faut avoir un grand bagage. Celle-ci est un outil essentiel de modélisation des risques.

Soit (X_1, \dots, X_n) une suite de variables aléatoires indépendantes et identiquement distribuées (i.i.d) de fonction de répartition F ($F(x) = P(X \leq x)$, pour $x \in \mathbb{R}$).

Soit A_n l'ensemble des permutations de $\{1, \dots, n\}$.

1.2.1 Définition de la statistique d'ordre

La statistique d'ordre de l'échantillon (X_1, \dots, X_n) est le réarrangement croissant de (X_1, \dots, X_n) .

On le note par $(X_{1,n}, X_{2,n}, \dots, X_{n,n})$ et est tel que :

$X_{1,n} \leq \dots \leq X_{n,n}$, il existe une permutation aléatoire $\sigma_n \in A_n$ telle que

$$(X_{1,n}, \dots, X_{n,n}) = (X_{\sigma_n(1)}, \dots, X_{\sigma_n(n)})$$

Le vecteur $(X_{1,n}, \dots, X_{n,n})$ est appelé l'échantillon ordonné associé à l'échantillon (X_1, \dots, X_n) , et $X_{k,n}$ étant la $k^{\text{ème}}$ statistique d'ordre.

Dans un échantillon de taille n deux statistiques d'ordre sont particulièrement intéressantes pour l'étude des événements extrêmes et qui sont :

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

En utilisant la propriété d'indépendance des variables aléatoires X_1, \dots, X_n , nous en déduisons que les lois du maximum $X_{n,n}$ et du minimum $X_{1,n}$ de la statistique d'ordre associés à l'échantillon

X_1, \dots, X_n sont :

$$\begin{aligned} F_{X_{n:n}}(x) &= P(X_{n:n} \leq x) = P(\max(X_1, \dots, X_n) \leq x) \\ &= P\left[\bigcap_{i=1}^n (X_i \leq x)\right] \\ &= \prod_{i=1}^n P(X_i \leq x) \\ &= \prod_{i=1}^n F_X(x) \end{aligned}$$

donc

$$F_{X_{n:n}}(x) = [F_X(x)]^n$$

En utilisant les même propriétés on a :

$$\begin{aligned} F_{X_{1:n}}(x) &= P(X_{1:n} \leq x) = 1 - P(X_{1:n} > x) \\ &= 1 - P(\min(X_1, \dots, X_n) > x) \\ &= 1 - P\left[\bigcap_{i=1}^n X_i > x\right] \\ &= 1 - \prod_{i=1}^n P(X_i > x) \\ &= 1 - \prod_{i=1}^n [1 - F_X(x)] \end{aligned}$$

D'où on déduit

$$F_{X_{1:n}}(x) = 1 - [1 - F_X(x)]^n$$

De ces résultats on conclut que le maximum $X_{n:n}$ est une v.a dont la fonction de répartition correspond à F^n .

1.2.2 Loi de la statistique d'ordre

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

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

Preuve. On rappelle que, puisque $f(x)$ est la dérivée de $F(x)$

$$\begin{aligned} f(x) &= \lim_{\varepsilon \rightarrow 0} \frac{F(x + \varepsilon) - F(x)}{\varepsilon} \\ &= \lim_{\varepsilon \rightarrow 0} \frac{P(x \leq X < x + \varepsilon)}{\varepsilon} \end{aligned}$$

alors :

$$\begin{aligned} F_{X_{1,n}, \dots, X_{n,n}}(x_1, \dots, x_n) &= P(x_1 \leq X_{1,n} < x_1 + \varepsilon, \dots, x_n \leq X_{n,n} < x_n + \varepsilon) \\ &= n! P(x_1 \leq X_{1,n} < x_1 + \varepsilon) \dots P(x_n \leq X_{n,n} < x_n + \varepsilon) \end{aligned}$$

d'où

$$\frac{P(x_1 \leq X_{1,n} < x_1 + \varepsilon_1, \dots, x_n \leq X_{n,n} < x_n + \varepsilon_n)}{\varepsilon_1 \dots \varepsilon_n} = n! \frac{P(x_1 \leq X_{1,n} < x_1 + \varepsilon_1)}{\varepsilon_1} \dots \frac{P(x_n \leq X_{n,n} < x_n + \varepsilon_n)}{\varepsilon_n}$$

alors

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

■

1.2.3 Densité conjointe de deux statistique d'ordre

La densité conjointe de $(X_{i,n}, X_{j,n})$ pour $i < j$ est donnée par :

$$f_{X_{i,n}, X_{j,n}}(x) = \frac{n!}{(i-1)!(j-i-1)!(n-j)!} [F(x)]^{i-1} [F(y) - F(x)]^{j-i-1} [1 - F(y)]^{n-j} f(x)f(y), \quad x < y < \infty.$$

La fonction de répartition $F_{X_{i,n}, X_{j,n}}(x)$ est définie par :

a) Pour $x \geq y$:

$$F_{X_{i,n}, X_{j,n}}(x) = F_{X_{j,n}}(y)$$

b) Pour $x < y$:

$$\begin{aligned} F_{X_{i,n}, X_{j,n}}(x) &= P(X_{i,n} \leq x, X_{j,n} \leq y) \\ &= \sum_{i=s}^n \sum_{j=r}^i \frac{n!}{(i-1)!(j-i-1)!(n-j)!} [F(x)]^i [F(y) - F(x)]^{j-1} [1 - F(y)]^{n-j} \end{aligned}$$

1.3 Distribution empirique et quantile

1.3.1 Inverse généralisé

On appelle inverse généralisée de F , l'application notée F^{\leftarrow} définie par :

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

L'inverse généralisée F^{\leftarrow} coïncide avec l'inverse F^{-1} lorsque F est strictement croissante et continue

.

1.3.2 La fonction de répartition empirique

On appelle fonction de répartition empirique associée à un échantillon (X_1, \dots, X_n) , la fonction F_n définie par :

$$\begin{aligned} F_n(x) &= \frac{1}{n} \sum_{k=1}^n I_{\{X_k \leq x\}} \\ &= \frac{1}{n} \text{card} \{k \in \{1, \dots, n\} : X_k \leq x\}. \end{aligned}$$

$$F_n(x) = \begin{cases} 0 & \text{si } x < X_{1.n} \\ \frac{1}{n} & \text{si } X_{1.n} \leq x < X_{2.n} \\ \cdot & \\ \cdot & \\ \frac{k}{n} & \text{si } X_{k.n} \leq x < X_{k+1.n} \\ \cdot & \\ \cdot & \\ 1 & \text{si } x \geq X_{n.n} \end{cases}$$

où I désigne la fonction indicatrice :

$$I_A(x) = \begin{cases} 0 & \text{si } x \notin A \\ 1 & \text{si } x \in A \end{cases}$$

1.3.3 Point terminal

Le point terminal noté x_F d'une distribution F est définie par :

$$x_F = \sup\{x \in \mathbb{R} : F(x) < 1\}.$$

Proposition 1.3.1 *La fonction de répartition F_X n'étant pas souvent connue, il n'est généralement pas possible de déterminer la distribution du maximum. On s'intéresse alors à la distribution*

asymptotique du maximum en faisant tendre n vers l'infini . On a :

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

On constate que la distribution asymptotique du maximum déterminée en faisant tendre n vers l'infini, donne une loi dégénérée (ils prennent les valeurs 0 ou 1 seulement) .

1.3.4 Convergence de la fonction de répartition empirique

Théorème 1.3.1 (Glivenco-Cantelli) : Soit $(X_n, n \in \mathbb{N})$ une suite de variables aléatoires réelles indépendantes et de même lois F , alors :

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

1.3.5 Quantile d'ordre p

On appelle quantile ou fractile d'ordre p , le nombre définie par :

$$x_p = \inf\{x \in \mathbb{R} : F(x) \geq p\}.$$

Si F est strictement croissante et continue, alors x_p est l'unique nombre réel tel que :

$$F(x_p) = p.$$

1.3.6 Quantile extrême

On appelle quantile extrême , le quantile d'ordre $(1 - p)$, définie par :

$$\begin{aligned} x_p &= \inf\{x \in \mathbb{R} : F(x) \geq p\} \\ &= F^{\leftarrow}(1 - p). \end{aligned}$$

où p proche de zéro .

1.3.7 Quantile empirique

La fonction quantile empirique de l'échantillon (X_1, \dots, X_n) est donnée par :

$$\begin{aligned} Q_n(p) &= F_n^{-1}(p) \\ &= \inf\{x \in \mathbb{R} : F_n(x) \geq p\} \text{ avec } 0 < p < 1 \end{aligned}$$

la fonction quantile empirique de la queue correspondante est définie par :

$$U(t) = F^{\leftarrow}\left(1 - \frac{1}{t}\right) \text{ avec } 1 < t < \infty$$

Définition 1.3.1 Soit (X_1, \dots, X_n) un échantillon issu d'une loi F et $(X_{1.n}, \dots, X_{n.n})$ l'échantillon ordonné . Soit $p \in]0, 1[$, on appelle quantile empirique d'ordre p , la variable aléatoire notée $\hat{X}_{[np].n}$ définie par :

$$\hat{X}_{[np].n} = \begin{cases} \frac{X_{[np].n} + X_{[np]+1.n}}{2} & \text{si } [np] \in \mathbb{N} \\ X_{[np]+1.n} & \text{sinon} \end{cases}$$

où $[np]$ désigne la partie entière de np .

En particulier $X_{[\frac{n}{2}]+1.n}$ est la médiane empirique.

Proposition 1.3.2 (Propriété de Markov) : Quand F est continue , l'échantillon ordonné $(X_{1.n}, \dots, X_{n.n})$ forme une chaîne de Markov . En d'autres termes, nous avons pour $i = 2, \dots, n$

$$P(X_{i.n} \leq x / X_{1.n} = x_1, \dots, X_{i-1.n} = x_{i-1}) = P(X_{i.n} \leq x / X_{i-1.n} = x_{i-1})$$

1.3.8 Lois des grands nombres

Ces loi décrivent le comportement asymptotique de la moyenne de l'échantillon :

$$\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i$$

Elles sont de deux types : loi faible mettant en jeu la convergence en probabilité et loi forte relative à la convergence presque sûre.

Théorème 1.3.2 (*Lois des grands nombres*) : *Si (X_1, \dots, X_n) est une suite d'une variable aléatoire X tel que $E|X| < \infty$, Alors :*

- *La loi faible :*

$$\bar{X}_n \xrightarrow{P} \mu \text{ quand } n \rightarrow \infty$$

- *La loi forte :*

$$\bar{X}_n \xrightarrow{P.s} \mu \text{ quand } n \rightarrow \infty$$

où $\mu := E(X)$ est la moyenne de X .

1.3.9 Théorème central limite

L'étude de somme de variables indépendantes et de même loi joue un rôle capitale en statistique . Le théorème suivant connu sous le nom de théorème central limite(TCL) établit la convergence vers la loi gaussienne .

Théorème 1.3.3 (*TCL*) : *Si (X_1, \dots, X_n) est une suite de variable aléatoire définie sur le même espace de probabilité de variance σ^2 finie et de moyenne μ alors :*

$$\frac{1}{\sqrt{n}} \left(\frac{S_n - n\mu}{\sigma} \right) \rightarrow^D N(0, 1) \text{ quand } n \rightarrow \infty \text{ tel que } S_n = \sum_{i=1}^n X_i$$

Dans les applications pratiques , ce théorème permet en particulier de remplacer une somme de variables aléatoires en nombre assez grand mais fini par une approximation normale .

Chapitre 2

Distribution des valeurs extrêmes généralisées

Le théorème des valeurs extrêmes donne la distribution asymptotique du maximum d'un échantillon. Les distributions limites possibles obtenues sont paramétrées par l'indice des valeurs extrêmes γ . Cet indice contrôle la lourdeur de la queue de distribution. La valeur de l'indice γ permet de répartir les distributions limites en trois domaines d'attraction, **Fréchet**, **Weibull** et **Gumbel**.

2.1 Distribution des valeurs extrêmes usuelles

Les relations donnant les principaux moments centrés et non centrés ainsi que les coefficients d'asymétrie et d'aplatissement des distributions de Fréchet, de Weibull et de Gumbel sont présentés dans cette section.

2.1.1 Distribution de probabilité de Fréchet

La distribution dite de Fréchet a été introduite par le mathématicien français **Maurice René Fréchet**. Il a développé cette distribution comme loi de valeurs extrêmes. Les travaux de Fréchet ont inspiré les recherches de **Fisher-Tippett**(1928) et ceux de **Gumbel**.

La distribution de Fréchet est l'une des distributions les plus populaires dans de nombreux domaines, par exemple en hydrologie, la distribution de Fréchet est utilisée pour des événements extrêmes tels que le maximum annuel des précipitations journalières ou le débit des rivières. En finance et macroéconomie, la distribution de Fréchet est utilisée pour calculer la valeur à risque .

Définition 2.1.1 *On dit qu'une variable aléatoire X est de distribution de **Fréchet** de paramètres $\mu, \sigma > 0, \gamma > 0$, si sa densité de probabilité s'écrit :*

$$h_{\gamma,\mu,\sigma}(x) = \begin{cases} 0 & \text{si } x < \frac{1}{\gamma} \\ \frac{\gamma}{\sigma} \left(\frac{x-\mu}{\sigma}\right)^{-1-\gamma} \exp\left(-\left(\frac{x-\mu}{\sigma}\right)^{-\gamma}\right) & \text{si } x \geq \frac{1}{\gamma} \end{cases} \quad (2.1)$$

où γ est le paramètre de forme, σ est le paramètre d'échelle et μ est le paramètre de position.

Dans la figure 2.1 nous allons visualiser l'effet de σ et de μ sur la densité de la distribution de Fréchet

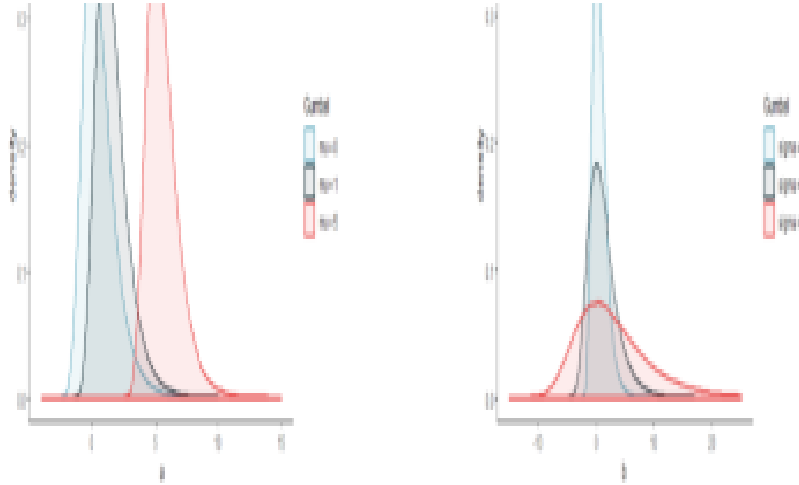


FIG. 2.1 – Fonction de densité de Fréchet : (a) pour différentes valeur de μ , (b) pour différentes valeur de σ .

Il est clair sur la figure 2.1(a) que la moyenne μ joue un rôle de translation. Par contre, sur la figure 2.1 (b) le paramètre d'échelle σ contrôle l'aplatissement de la distribution.

Remarque 2.1.1 *On constate sur la figure 2.1 que la fonction de densité de Fréchet est bornée à gauche. De plus, la distribution de Fréchet admet un Kurtosis positif, donc sa médiane est supérieure à sa moyenne.*

Définition 2.1.2 *La fonction de répartition de Fréchet est donnée la formule suivante :*

$$H_{\gamma,\mu,\sigma}(x) = P(X \leq x) = \begin{cases} 0 & \text{si } x < \frac{1}{\gamma} \\ \exp(-(\frac{x-\mu}{\sigma})^{-\frac{1}{\gamma}}) & \text{si } x \geq \frac{1}{\gamma} \end{cases} \quad (2.2)$$

Définition 2.1.3 *La fonction de survie de Fréchet s'écrit donc :*

$$\bar{H}_{\gamma,\mu,\sigma}(x) = 1 - H_{\gamma,\mu,\sigma}(x) = \begin{cases} 1 & \text{si } x < \frac{1}{\gamma} \\ 1 - \exp(-(\frac{x-\mu}{\sigma})^{-\frac{1}{\gamma}}) & \text{si } x \geq \frac{1}{\gamma} \end{cases} \quad (2.3)$$

Moments d'ordre k

Soit X une variable aléatoire distribuée selon une distribution Fréchet (2.1). Il est facile de montrer que la moyenne $E(X)$ et la variance $Var(X)$ de cette variable sont données respectivement par

$$E(X) = \mu + \frac{\gamma}{\sigma}[1 - \Gamma(1 + \gamma)]$$

$$Var(X) = (\frac{\sigma}{\gamma})^2[\Gamma(1 + 2\gamma) - \Gamma^2(1 + \gamma)].$$

où $\Gamma(x) = \int_{\infty}^0 t^{x-1} e^{-t} dt$ et $\Gamma(n + 1) = n!$, pour $n \in \mathbb{N}^*$.

Les moments d'ordre 3 et 4 sont déterminés par :

$$\mu_3(X) = -(\frac{\sigma}{\gamma})^3[\Gamma(1 + 3\gamma) - 3\Gamma(1 + 2\gamma)\Gamma(1 + \gamma) + 2\Gamma^3(1 + \gamma)]$$

$$\mu_4(X) = (\frac{\sigma}{\gamma})^4[\Gamma(1 + 4\gamma) - 4\Gamma(1 + \gamma)\Gamma(1 + 3\gamma) + 6\Gamma^2(1 + \gamma)\Gamma(1 + 2\gamma) - 3\Gamma^4(1 + \gamma)].$$

Coefficients du Skewness et du Kurtosis

En utilisant les résultats précédent, on peut déduire le coefficient C_s de Skewness ou d'asymétrie

$$C_s = \frac{\mu_3}{\mu_2^{\frac{3}{2}}} = -\frac{|\gamma|}{\gamma} \frac{[\Gamma(1 + 3\gamma) - 3\Gamma(1 + 2\gamma)\Gamma(1 + \gamma) + 2\Gamma^3(1 + \gamma)]}{[\Gamma(1 + 2\gamma) - \Gamma^2(1 + \gamma)]^{\frac{3}{2}}}$$

ainsi que le coefficient K du Kurtosis ou d'applatissage

$$K = \frac{\mu_4}{\mu_2^2} = \frac{[\Gamma(1+4\gamma) - 4\Gamma(1+\gamma)\Gamma(1+3\gamma) + 6\Gamma^2(1+\gamma)\Gamma(1+2\gamma) - 3\Gamma^4(1+\gamma)]}{[\Gamma(1+2\gamma) - \Gamma^2(1+\gamma)]^2}.$$

Il est clair que les coefficients C_s et K ne dépendent que du paramètre de forme γ (indice des queues) .

2.1.2 Distribution de probabilité de Weibull

La distribution de Weibull a été introduite par le mathématicien **Suédois Waloddi Weibull**. De nos jours, cette distribution est couramment utilisée dans l'analyse de survie et en fiabilité des systèmes . La distribution de Weibull peut également être utilisée dans de nombreux autres domaines, notamment la biologie, l'économie, l'industrie, l'hydrologie, etc...

Définition 2.1.4 *On dit qu'une variable aléatoire X est de distribution de Weibull de paramètres $\mu, \sigma > 0, \gamma < 0$, si sa densité de probabilité s'écrit :*

$$h_{\mu,\sigma,\gamma}(x) = \begin{cases} 0 & \text{si } x > \frac{1}{\gamma} \\ \frac{\gamma}{\sigma} \left(\frac{x-\mu}{\sigma}\right)^{\gamma-1} \exp\left(-\left(\frac{x-\mu}{\sigma}\right)^\gamma\right) & \text{si } x \leq \frac{1}{\gamma} \end{cases} \quad (2.4)$$

où γ est le paramètre de forme, σ le paramètre d'échelle et μ le paramètre de position.

Dans la figure 2.2 nous allons visualiser l'effet de μ et de σ sur la densité de la distribution de Weibull .

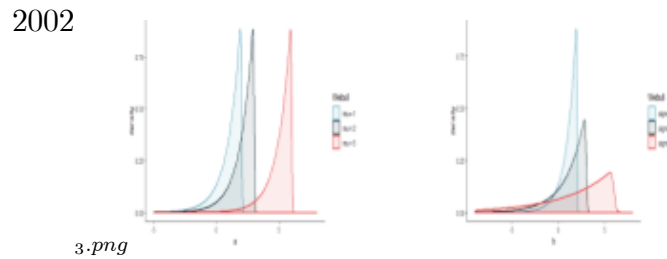


FIG. 2.2 – Fonction de densité de Weibull : (a) pour différentes valeurs de μ , (b) pour différentes valeurs de σ .

Il est clair sur la figure 2.2(a) que la moyenne μ joue un rôle de translation . Par contre, sur la figure 2.2(b) le paramètre d'échelle σ contrôle l'aplatissement de la distribution .

Remarque 2.1.2 *On constate sur la figure 2.2 que la fonction de densité de Weibull est bornée à droite . De plus, la distribution de weibull admet un Kurtosis négatif, donc sa médiane inférieure à sa moyenne .*

Définition 2.1.5 *La fonction de distribution de Weibull est donnée par la formule suivante :*

$$H_{\gamma,\mu,\sigma}(x) = \begin{cases} 0 & \text{si } x > \frac{1}{\gamma} \\ \exp(-(-\frac{x-\mu}{\sigma})^{-\frac{1}{\gamma}}) & \text{si } x \leq \frac{1}{\gamma} \end{cases}$$

Définition 2.1.6 *La fonction de survie de Weibull est définie par :*

$$\bar{H}_{\gamma,\mu,\sigma}(x) = 1 - H_{\gamma,\mu,\sigma}(x) = \begin{cases} 1 & \text{si } x > \frac{1}{\gamma} \\ 1 - \exp(-(-\frac{x-\mu}{\sigma})^{-\frac{1}{\gamma}}) & \text{si } x \leq \frac{1}{\gamma} \end{cases}$$

Moments d'ordre k

Soit X une variable aléatoire distribuée selon une loi de Weibull (2.2) . On peut montrer que la moyenne $E(X)$ et la variance $Var(X)$ de cette variable sont données respectivement par :

$$E(X) = \mu + \frac{\gamma}{\sigma}[1 - \Gamma(1 + \gamma)]$$

$$Var(X) = (\frac{\sigma}{\gamma})^2[\Gamma(1 + 2\gamma) - \Gamma^2(1 + \gamma)].$$

Les moments d'ordre 3 et 4 sont alors :

$$\mu_3(X) = -(\frac{\sigma}{\gamma})^3[\Gamma(1 + 3\gamma) - 3\Gamma(1 + 2\gamma)\Gamma(1 + \gamma) + 2\Gamma^3(1 + \gamma)]$$

$$\mu_4(X) = (\frac{\sigma}{\gamma})^4[\Gamma(1 + 4\gamma) - 4\Gamma(1 + \gamma)\Gamma(1 + 3\gamma) + 6\Gamma^2(1 + \gamma)\Gamma(1 + 2\gamma) - 3\Gamma^4(1 + \gamma)].$$

Coefficients du Skewness et du Kurtosis

Utilisant les résultats précédant, on peut déduire le coefficient C_s du Skewness :

$$C_s = \frac{\mu_3}{\mu_2^{\frac{3}{2}}} = -\frac{|\gamma|}{\gamma} \frac{[\Gamma(1+3\gamma) - 3\Gamma(1+2\gamma)\Gamma(1+\gamma) + 2\Gamma^3(1+\gamma)]}{[\Gamma(1+2\gamma) - \Gamma^2(1+\gamma)]^{\frac{3}{2}}}$$

ainsi que le coefficient K du Kurtosis :

$$K = \frac{\mu_4}{\mu_2^2} = \frac{[\Gamma(1+4\gamma) - 4\Gamma(1+\gamma)\Gamma(1+3\gamma) + 6\Gamma^2(1+\gamma)\Gamma(1+2\gamma) - 3\Gamma^4(1+\gamma)]}{[\Gamma(1+2\gamma) - \Gamma^2(1+\gamma)]^2}$$

2.1.3 Distribution de probabilité de Gumbel

La distribution de Gumbel a été introduit par le mathématicien Américain **Émil Julius Gumbel**. Il a développé cette distribution comme loi de valeurs extrêmes. La distribution de Gumbel est l'une des distributions les plus populaires dans de nombreux domaines grâce à sa flexibilité.

Définition 2.1.7 On dit qu'une variable aléatoire X est de distribution de gumbel de paramètres μ et $\sigma > 0$, si sa densité de probabilité s'écrit telle que :

$$h_{\mu,\sigma}(x) = \frac{1}{\sigma} \exp\left\{-\frac{(x-\mu)}{\sigma} - \exp\left(-\frac{(x-\mu)}{\sigma}\right)\right\}$$

où σ est le paramètre d'échelle et μ est le paramètre de position.

Dans la figure 2.3 nous allons visualiser l'effet de σ et de μ sur la densité de la distribution de Gumbel

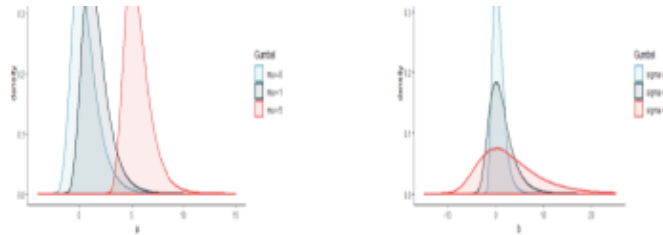


FIG. 2.3 – Fonction de densité de Gumbel : (a) pour différentes valeurs de μ , (b) pour différentes valeurs de σ .

Il est clair sur la figure 2.3(a) que la moyenne μ joue un rôle de translation. Par contre, sur la figure 2.3(b) le paramètre d'échelle σ contrôle l'applatissage de la distribution.

Définition 2.1.8 *La fonction de répartition de la variable de Gumbel s'écrit :*

$$H_{\mu,\sigma}(x) = \exp\{-\exp(-\frac{x-\mu}{\sigma})\} \quad (2.5)$$

Moments d'ordre k

Soit X une variable aléatoire distribuée selon la loi de Gumbel 2.3. On peut montrer que la moyenne et la variance de cette variable sont données respectivement par :

$$E(X) = \mu + \sigma C$$
$$Var(X) = \frac{\sigma^2 \pi^2}{6}$$

où $C = 0.5772...$ est la constante d'Euler.

Coefficients du Skewness et du Kurtosis

Les coefficients du Skewness ou d'asymétrie C_s et du Kurtosis ou d'applatissage K sont donnés respectivement par :

$$C_s = 1.139$$

$$K = 2.40$$

2.2 Comportement asymptotique du maximum

Distribution du maximum d'un échantillon

On s'intéresse principalement au problème de la distribution du maximum d'un échantillon .

Considérons un échantillon aléatoire X_1, \dots, X_n d'une variable aléatoire de fonction de répartition

F . Notons le maximum de cet échantillon par $X_{n,n}$ défini par :

$$X_{n,n} = \max\{X_1, \dots, X_n\}$$

Soit X_1, \dots, X_n une suite de variable aléatoire indépendantes et identiquement distribuées (i.i.d) , de fonction de répartition F . Il est naturel de s'intéresser à la distribution asymptotique de $X_{n,n}$. On vérifie aisément que :

$$F_{X_{n,n}}(x) = P[X_{n,n} \leq x] = F^n(x)$$

$$\text{et donc } \lim F_{X_{n,n}}(x) = \begin{cases} 0 & \text{si } x < x_F \\ 1 & \text{si } x = x_F \end{cases} \quad \text{et donc la loi limite de } X_{n,n} \text{ est dégénérée.}$$

Par analogie avec le théorème central limite, on s'autorise alors une normalisation linéaire de $X_{n,n}$ et on étudie le comportement limite de la quantité $\frac{X_{n,n} - b_n}{a_n}$, avec a_n et b_n des suites réelles bien choisies pour éviter la dégénérescence de la loi limite. On pose deux suites $(a_n)_{n \geq 1}$, $a_n > 0$ et $(b_n)_{n \geq 1}$, $b_n \in \mathbb{R}$ telle que :

$$P\left(\frac{X_{n,n} - b_n}{a_n} \leq x\right) \longrightarrow H(x) \text{ quand } n \longrightarrow \infty$$

en tout point x de continuité de H ,

Le problème de la limite a été résolu par **Fisher-Tippett**(1928) , **Gnedenko** (1943) et a ensuite été relancé et rationalisé **De Haan** (1970).

Théorème de Fisher-Tippet

Théorème 2.2.1 *S'il existe deux suites $(a_n)_{n \geq 1}$, $a_n > 0$ et $(b_n)_{n \geq 1}$, $b_n \in \mathbb{R}$ tels que :*

$$\lim_{n \rightarrow \infty} F_{X_{n,n}}(a_n x + b_n) = H(x)$$

Alors H est non dégénérée et elle est l'une des trois loi suivantes :

Type I : Fréchet

$$\Phi_{\alpha}(x) = \begin{cases} 0 & \text{si } x \leq 0 \\ e^{-x^{-\alpha}} & \text{si } x > 0 \end{cases}$$

Type II : Weibull

$$\Psi_{\alpha}(x) = \begin{cases} e^{-(-x)^{-\alpha}} & \text{si } x < 0 \\ 1 & \text{si } x \geq 0 \end{cases}$$

Type III : Gumbel

$$\Lambda_{\alpha}(x) = e^{-e^{-x}}, \quad x \in \mathbb{R}.$$

Avec Φ_{α} , Ψ_{α} et Λ_{α} sont des lois limites possibles pour le maximum .

Représentation de Jenkinson-Von mises

Jenkinson (1955) a proposé ce qu'il a appelé la distribution généralisée des valeurs extrêmes (GEV) . cette distribution a l'avantage de combiner les trois types de lois des valeurs extrêmes développées par Fisher-Tippet . La forme générale de la fonction de répartition d'une GEV (Generalized extreme value) s'exprime par :

$$H_{\mu,\sigma,\gamma}(x) = \begin{cases} \exp[-(1 + \frac{\gamma}{\sigma}(x - \mu))]^{-\frac{1}{\gamma}} & \text{si } \gamma \neq 0 \\ \exp[-\exp(-(\frac{x-\mu}{\sigma}))] & \text{si } \gamma = 0 \end{cases} \quad (2.6)$$

où μ est un paramètre de position , σ est celui d'échelle, γ le paramètre de queue et H_{γ} est définie sur l'ensemble D .

$$D = \{x \in \mathbb{R} : (1 + \frac{\gamma}{\sigma}(x - \mu)) > 0\}.$$

Pour $\mu = 0$ et $\sigma = 1$, on obtient la forme standard des trois types de loi des valeurs extrêmes . Cette loi dépend du seul paramètre de forme γ appelé l'indice des valeurs extrêmes , ainsi la relation (2.3) devient :

$$H_{\mu,\sigma,\gamma}(x) = \begin{cases} \exp[-(1 + \gamma x)]^{-\frac{1}{\gamma}} & \text{si } \gamma \neq 0 \text{ avec } 1 + \gamma x > 0 \\ \exp[-\exp(-x)] & \text{si } \gamma = 0 \end{cases}$$

La fonction de densité de probabilité prend deux formes obtenues en dérivant (2.3) :

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

de GEV

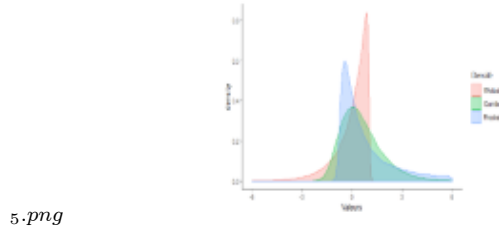


FIG. 2.4 – Densité de probabilité des valeurs extrêmes avec différentes valeurs de γ .

2.3 Caractérisation des domaines d'attraction et Conditions nécessaire et suffisantes de convergence

Il faut noter que le paramètre de forme γ conditionne le type de la loi des valeurs extrêmes . Nous présentons dans ce qui suit les domaines d'attraction et les concepts théoriques qui leur sont associées dans les trois cas correspondant aux signes du paramètre de forme γ .

On dit qu'une distribution F appartient au domaine d'attraction de H_γ , et on note $F \in D(H_\gamma)$ s'il existe des suites réelles $(a_n)_{n \geq 1}$, $a_n > 0$ et $(b_n)_{n \geq 1}$, $b_n \in \mathbb{R}$ telles que

$$\lim_{n \rightarrow \infty} F^n(a_n x + b_n) = H_\gamma(x).$$

2.3.1 Notion de fonction à variation régulière

Nous présentons ici le résultat principal de la variation régulière ;

Définition 2.3.1 Une fonction $f : \mathbb{R}_+ \longrightarrow \mathbb{R}_+$ mesurable au sens de Lebesgue, est à variation régulière à l'infini, si et seulement s'il existe un réel α tel que pour tout $x > 0$

$$\lim_{t \rightarrow \infty} \frac{f(tx)}{f(t)} = x^\alpha$$

et on écrit $f \in \mathfrak{R}v_\alpha$, α est appelé indice (ou exposant) de la fonction à variation régulière f .

Remarque 2.3.1 Pour $\alpha = 0$, on retrouve le cas de la fonction à variation lente dans la définition suivante.

Définition 2.3.2 Une fonction L est dite à variation lente si $L(t) > 0$ pour t assez grand et si pour tout $x > 0$, on a

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

Exemple 2.3.1 $f(x) = \log(x)$

$$\lim_{t \rightarrow \infty} \frac{\log(tx)}{\log(t)} = \lim_{t \rightarrow \infty} \left(1 + \frac{\log(x)}{\log(t)}\right) = 1$$

donc la fonction \log est à variation lente.

2.3.2 Domaines d'attraction

Domaine d'attraction de la loi 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 $x_F = +\infty$. Elles sont dites aussi lois à queue lourde. Le domaine d'attraction de Fréchet est noté $D(\Phi_\gamma)$.

F appartient au domaine d'attraction de Fréchet ssi $F^{-1}(1) = +\infty$ et $\exists \gamma > 0$ tel que :

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

on peut écrire,

$$x^{-\frac{1}{\gamma}} = -\log \Phi_\gamma(x)$$

Théorème 2.3.1 VonMises(1936) : Une fonction de répartition $F \in D(\Phi_\gamma)$ ssi sa fonction de survie est donnée par :

$$\bar{F}(x) = x^{-\frac{1}{\gamma}} L(x)$$

où L est une fonction à variation lente et \bar{F} est la fonction de survie.

Domaine d'attraction de la loi de Weibull

Les lois de ce domaine sont bornées à droite et par conséquent, le point terminal x_F est fini ; le domaine d'attraction de weibull est noté $D(\Psi_\gamma)$.

$F \in D(\Psi_\gamma)$ ssi $F^{-1}(x) < +\infty$ et $\exists \gamma < 0$ tel que :

$$\lim_{t \rightarrow \infty} \frac{1 - F(F^{-1}(1) - tx)}{1 - F(F^{-1}(1) - t)} = (-x)^{-\frac{1}{\gamma}}, \forall x \leq 0$$

On peut écrire

$$(-x)^{-\frac{1}{\gamma}} = -\log \Psi_\gamma(x)$$

Théorème 2.3.2 Gnedenko(1943) : Une fonction de répartition $F \in D(\Psi_\gamma)$ ssi $x_F < +\infty$ et

$$\bar{F}(x_F - \frac{1}{x}) = x^{-\frac{1}{\gamma}} L(x)$$

où L est une fonction à variation lente et \bar{F} est la fonction de survie .

Domaine d'attraction de la loi de Gumbel

La loi présente dans la queue une décroissance de type exponentiel, ce qui permet de caractériser dans ce cas, le domaine d'attraction de Gumbel noté $D(\Lambda)$.

F appartient au domaine d'attraction de Gumbel ssi $E[X/X > c] < +\infty$ et $\forall c < F^{-1}(1)$ tel que :

$$\lim_{t \rightarrow \infty} \frac{1 - F(t + xE[X - t/X > t])}{1 - F(t)} = e^{-x}, \forall x \in \mathbb{R}$$

Théorème 2.3.3 VonMises (1936) : S'il existe une fonction mesurable R , appelée fonction

auxiliaire telle que :

$$\lim_{x \rightarrow x_F} \frac{1 - F(t + xR(t))}{1 - F(t)} = e^{-x}$$

où $x_F = \sup\{x \in \mathbb{R}, F(x) < 1\}$, est le point terminal de F , alors $F \in D(\Lambda)$.

Le tableau suivant présente quelques lois qui appartiennent aux domaines d'attraction de Fréchet, Weibull et Gumbel.

Domaines d'attraction	Gumbel $\gamma = 0$	Fréchet $\gamma > 0$	Weibull $\gamma < 0$
Lois	Normale	Cauchy	Uniforme
	Exponentielle	Paréto généralisée	Beta
	Log-Normal	Student	
	Weibull	Log-gamma	
	Gumbel	Log-logistique	
	Gamma	Fréchet	
	Logistique	Chi-deux	
		Burr	

TAB. 2.1 – Domaines d'attraction et quelques lois associées

2.4 Constantes de normalisations

Nous allons présentés ici le choix des suites a_n, b_n (constantes de normalisations) afin que $P(\frac{X_{n,n}-b_n}{a_n} \leq x)$ converge vers $H(x)$ quand n tends vers l'infini selon le domaine d'attraction .

Théorème 2.4.1 *On peut choisir des constantes $a_n > 0$ et $b_n \in \mathbb{R}$ telles que :*

$$\lim_{n \rightarrow \infty} F_{X_{n,n}}(a_n x + b_n) = H_\gamma(x)$$

soit vérifiée de la manière suivante :

$$1. \ b_n = 0, \ a_n = F^{-1}(1 - \frac{1}{n}) \quad \text{si } H = \Phi.$$

2. $b_n = F^{-1}(1), \quad a_n = F^{-1}(1) - F^{-1}(1 - \frac{1}{n}) \quad \text{si } H = \Psi.$
3. $b_n = F^{-1}(1 - \frac{1}{n}), \quad a_n = F^{-1}(1 - \frac{1}{ne}) - F^{-1}(1 - \frac{1}{n}) \quad \text{si } H = \Lambda.$

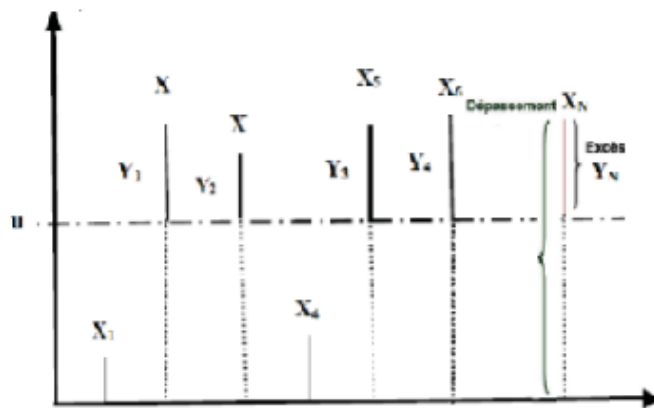
2.5 Distribution des excès

Soit maintenant une suite d'observations X_1, \dots, X_n indépendantes et identiquement distribuées, de fonction de répartition F et x_F un point terminal. Alors, pour un seuil $u < x_F$ fixé, on définit la variable $Y_j = X_j - u, j = 1, \dots, n$, l'excès au dessus du seuil u .

La fonction de distribution des excès de X au-dessus du seuil u est :

$$F_u(x) = P(X - u \leq x / X > u) = \frac{F(x + u) - F(u)}{1 - F(u)}$$

POT



6.png

FIG. 2.5 – Dépassement du seuil (POT)

2.6 Distribution de Pareto Généralisée (GPD)

La distribution de Pareto généralisée, joue un rôle essentiel dans la modélisation des excès.

Définition 2.6.1 Une distribution $G_{\gamma, \beta}$ est dite de Pareto généralisée de paramètre $\gamma \in \mathbb{R}$ et $\beta > 0$

si elle s'écrit sous la forme suivante :

$$G_{\gamma,\beta}(x) = \begin{cases} 1 - (1 + \frac{\gamma}{\beta}x)^{-\frac{1}{\gamma}} & \text{si } \gamma \neq 0 \\ 1 - \exp(-\frac{x}{\beta}) & \text{si } \gamma = 0 \end{cases}$$

cette distribution est définie pour

$$\begin{cases} x \geq 0 & \text{si } \gamma \geq 0 \\ 0 \leq x < -\frac{\beta}{\gamma} & \text{si } \gamma < 0 \end{cases}$$

telle que γ est le paramètre de queue, et β est le paramètre d'échelle .

Théorème de Balkema-De Haan-Pickands

Le théorème suivant fait le lien entre le comportement asymptotique de la distribution des excès et la loi de Pareto généralisée.

Théorème 2.6.1 Soit F_u la distribution des excès, lorsque le seuil u tend vers le point terminal x_F , on a :

$$\lim_{u \rightarrow x_F} \sup_{0 \leq x \leq x_F - u} |F_u(x) - G_{\gamma,\beta(u)}(x)| = 0.$$

où $\beta(u)$ est une fonction positive mesurable et $G_{\gamma,\beta(u)}$ est la fonction de répartition de la loi de Pareto Généralisée définie par :

$$G_{\gamma,\beta_u}(x) = \begin{cases} 1 - (1 + \frac{\gamma}{\beta_u}x)^{-\frac{1}{\gamma}}, & \text{si } \gamma \neq 0 \\ 1 - \exp(-\frac{x}{\beta_u}), & \text{si } \gamma = 0 \end{cases}$$

Remarque 2.6.1 Selon le signe de γ , nous avons les cas suivants :

1. $\gamma > 0$: distribution de type Pareto à queue lourde .
2. $\gamma < 0$: distribution de type Beta bornée au dessus de $u - \frac{\beta_u}{\gamma}$. (loi de Pareto de type II)
3. $\gamma = 0$: distribution de type exponentielle à queue légère.

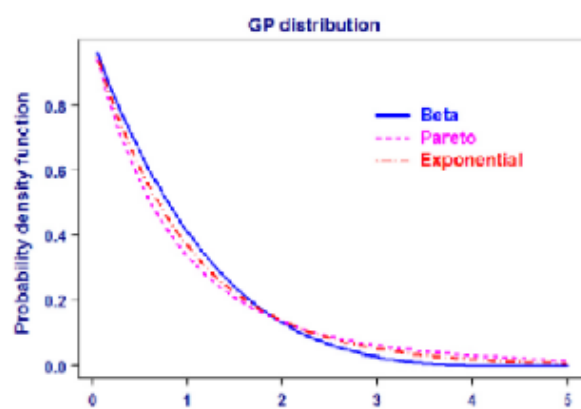


FIG. 2.6 – Distribution GPD

Chapitre 3

L'estimation de l'indice de queue

3.1 Introduction

Dans ce chapitre nous nous intéressons aux différentes méthodes d'estimation du paramètre γ intervenant dans la distribution des valeurs extrêmes généralisée . En se référant à la littérature, diverses méthodes paramétriques, non_paramétriques et graphique ont été proposées pour l'estimation de cet indice.

Pour les méthodes paramétriques, nous pouvons mentionner par exemple, les méthodes d'estimation empirique, la méthode de maximum de vraisemblance, la méthode des moments ou encore les méthodes bayésiennes. D'autres part, des approches non_paramétriques ont été consacrées à l'estimation de l'indice de queue ; les estimateurs les plus utilisés sont les estimateurs de Hill et Pickands . Le premier est appliqué seulement dans le domaine d'attraction de Fréchet ($\gamma > 0$) et le deuxième est valable pour les différentes lois limites des extrêmes.

Hosking et Wallis ont utilisé la méthode des moments et proposent un estimateur des moments pondérés défini seulement pour $\gamma < 1$.

Dans ce chapitre, on considère une suite $(X_n)_n$ de variables aléatoires i.i.d de fonction de répartition

F . On note $X_{1,n}, \dots, X_{n,n}$ la statistique d'ordre associée à l'échantillon X_1, \dots, X_n telle que :

$$X_{1,n} \leq X_{2,n} \leq \dots \leq X_{n,n}.$$

On dit qu'une suite $(k_n)_n$ d'entiers est intermédiaire si :

$$\lim_{n \rightarrow \infty} k_n = \infty \quad \text{et} \quad \lim_{n \rightarrow \infty} \frac{k_n}{n} = 0.$$

Pour simplifier l'écriture, on notera $k = k_n$.

3.2 L'estimation de l'indice de queue par des méthodes non paramétrique

3.2.1 L'estimateur de pickands

Il est défini par la statistique suivante pour $\gamma \in \mathbb{R}$ et $1 \leq k \leq \frac{n}{4}$:

$$\hat{\gamma}_n^{(p)} = \frac{1}{\ln 2} \ln \left[\frac{X_{n-k,n} - X_{n-2k,n}}{X_{n-2k,n} - X_{n-4k,n}} \right]$$

il a prouvé la consistance faible de l'estimation.

Il présente l'intérêt d'être valable quelle que soit la distribution des extrêmes (Gumbel, Weibull ou Fréchet). La représentation graphique de cet estimateur en fonction du nombre k d'observations considérées montre un comportement en général très volatil au départ, ce qui nuit à la lisibilité du graphique. De plus, cet estimateur est très sensible à la taille de l'échantillon sélectionné, ce qui le rend peu robuste. Il est donc d'un maniement délicat.

Propriétés asymptotiques de $\hat{\gamma}_n^{(p)}$

supposons que $F \in D(H_\gamma)$, $\gamma \in \mathbb{R}$, $k \rightarrow \infty$ et $k/n \rightarrow 0$ quand $n \rightarrow \infty$:

a) Consistance faible :

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

b) Consistance forte : si $k/\ln \ln n \rightarrow \infty$ quand $n \rightarrow \infty$, alors :

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

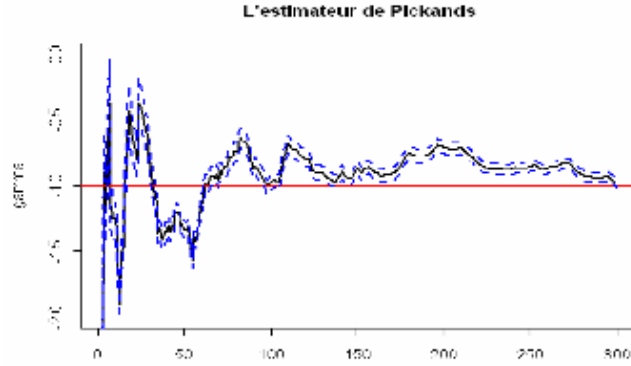
c) Normalité asymptotique : sous certaines conditions pour k et F on a :

$$\sqrt{k} \frac{\hat{\gamma}_n - \gamma}{\sigma(\gamma)} \longrightarrow N(0, 1) \quad \text{quand } n \rightarrow \infty$$

lorsque $k \rightarrow \infty$ la variance asymptotique étant donnée par :

$$\sigma(\gamma) = \frac{\gamma \sqrt{2^{2\gamma+1} + 1}}{2(2^\gamma - 1) \ln(2)}.$$

pickands



8.png

FIG. 3.1 – Estimateur de Pickands , en fonction du nombres des extrêmes avec l'intervalle de confiance 95%, pour la distribution uniforme ($\gamma = 1$) basé sur 100 échantillons de 3000 observations.

Une généralisation de l'estimateur de Pickands a été introduite comme suit :

$$\hat{\gamma}_n^{(Y)} = \hat{\gamma}_n^{(Y)}(k ; u, v) = \frac{1}{\ln v} \ln \frac{X_{n-k+1.n} - X_{n-[uk]+1.n}}{X_{n-[vk]+1.n} - X_{n-[uvk]+1.n}}$$

où u, v sont des nombres réels positifs différent de 1 tels que $[vk]$, $[uk]$ et $[uvk]$ ne dépassent pas n . Pour $u = v = 2$, nous avons $\hat{\gamma}_n^{(P)}$.

3.2.2 Estimateur de moment

Cet estimateur est proposé dans **Dekkers et al.** (1989). Il est défini par la statistique :

$$\hat{\gamma}_n^{(M)} = 1 + \hat{\gamma}_n^{(1)} - \frac{1}{2} \left[1 - \frac{(\hat{\gamma}_n^{(1)})^2}{\hat{\gamma}_n^{(2)}} \right]^{-1}$$

avec $\hat{\gamma}_n^{(i)} = \left[\frac{1}{k} \sum_{j=1}^k \log \left[\frac{X_{j,n}}{X_{k-1,n}} \right] \right]^i$, $i = 1, 2$

Propriétés asymptotique de $\hat{\gamma}_n^{(M)}$

supposons que $F \in D(H_\gamma)$, $\gamma \in \mathbb{R}$, $k \rightarrow \infty$ et $k/n \rightarrow 0$ quand $n \rightarrow \infty$:

a) Consistance faible :

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

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

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

c) Normalité asymptotique :

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

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

des moments

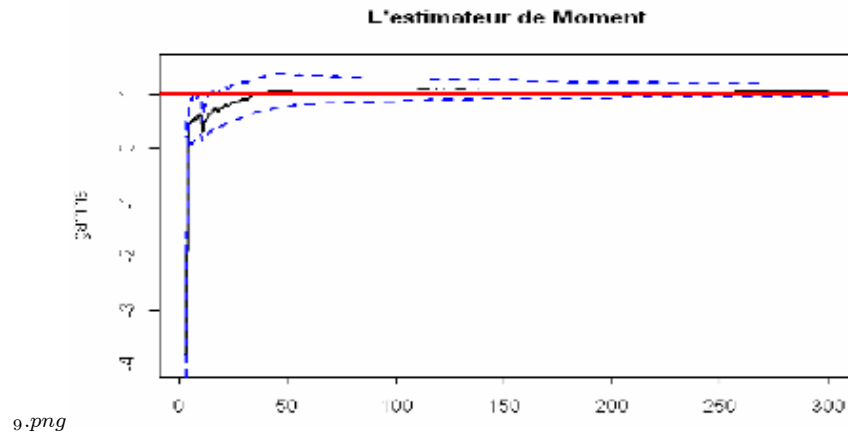


FIG. 3.2 – Estimateur du Moment, en fonction du nombre des extrêmes avec l'intervalle de confiance 95%, pour la distribution de Burr(1,1,1) basée sur 100 échantillons de 3000 observations.

3.2.3 L'estimateur de Hill

L'estimateur de **Hill** n'est utilisable que pour les distributions de Fréchet ($\gamma > 0$) pour lesquelles il fournit un estimateur de l'indice de queue plus efficace que l'estimateur de Pickands. Il est défini par la statistique suivante :

$$\hat{\gamma}_n^{(H)} = \frac{1}{k} \sum_{i=1}^k \ln \left[\frac{X_{n-i+1:n}}{X_{n-k:n}} \right], \text{ pour } k \in \{1, \dots, n-1\}$$

Construction de l'estimateur de Hill

Il y a plusieurs approches pour construire l'estimateur de Hill, par exemple :

Approche par EMV On considère une suite de variables aléatoires i.i.d de paramètre de Pareto de paramètre $\lambda > 0$, de fonction de répartition est donnée par :

$$F(x) = 1 - x^{-\lambda} \text{ pour } x \geq 1$$

alors

$$f(x) = \lambda x^{-\lambda-1}$$

On veut trouver l'estimateur du maximum de vraisemblance (E.M.V) de λ pour cela, on donne $L(\cdot)$ la fonction de vraisemblance et les dérivées du logarithme de la fonction vraisemblance par rapport à λ .

$$\begin{aligned} L(x_1, \dots, x_n, \lambda) &= \prod_{i=1}^n f(x_i) \\ &= \lambda^n \prod_{i=1}^n x_i^{-\lambda-1} \end{aligned}$$

$$\begin{cases} \frac{d \log L(x_1, \dots, x_n, \lambda)}{d\lambda} = \frac{n}{\lambda} - \sum_{i=1}^n \log x_i \\ \frac{d^2 \log L(x_1, \dots, x_n, \lambda)}{d\lambda^2} = -\frac{n}{\lambda^2} < 0 \end{cases}$$

l'estimateur du maximum de vraisemblance de $\frac{1}{\lambda}$ et donc la statistique :

$$\hat{T} = \frac{1}{n} \sum_{i=1}^n \log X_i = \frac{1}{n} \sum_{i=1}^n \log X_{i,n}$$

Une généralisation concerne

$$F(x) = Cx^{-\lambda} = Cx^{-\frac{1}{\gamma}}, \quad x > 0$$

Si on pose $C = u^\lambda$ avec $0 < u \leq x$, on obtient de l'E.M.V de λ :

$$\hat{\lambda} = \left(\frac{1}{n} \sum_{i=1}^n \log \left(\frac{X_{i,n}}{u} \right) \right)^{-1} = \left(\frac{1}{n} \sum_{i=1}^n \log X_{i,n} - \log u \right)^{-1}$$

Propriétés asymptotique de $\gamma_n^{(H)}$

supposons que $F \in D(H_\gamma)$, $\gamma > 0$, $k \rightarrow \infty$ et $k/n \rightarrow 0$ quand $n \rightarrow \infty$:

a) Consistance faible :

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

b) Consistance forte : Si $k/\log(\log(n)) \rightarrow \infty$ alors :

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

c) Normalité asymptotique :

$$\sqrt{k} \frac{\hat{\gamma}_n^{(H)} - \gamma}{\gamma} \xrightarrow{L} N(0, 1)$$

On peut associer à l'estimateur de Hill un intervalle de confiance asymptotique $I_N(\alpha)$ de niveau α

$$I_N(\alpha) = \left[\hat{\gamma} - z_{\alpha/2} \hat{\gamma} \frac{1}{\sqrt{k}}, \hat{\gamma} + z_{\alpha/2} \hat{\gamma} \frac{1}{\sqrt{k}} \right],$$

où $z_{\alpha/2}$ est le quantile d'ordre $1 - \frac{\alpha}{2}$ de la loi normale centrée réduite.

de hill

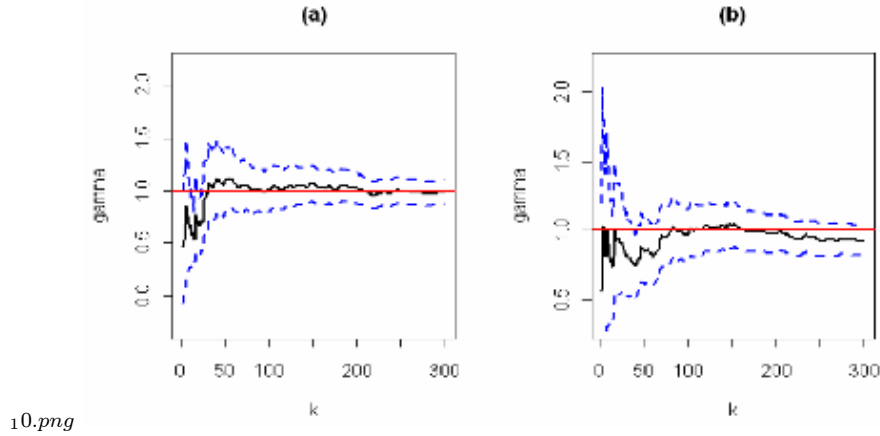


FIG. 3.3 – Estimateur de Hill , en fonction du nombres des extrêmes avec l'intervalle de confiance 95% , pour la distribution uniforme ($\gamma = 1$) basée sur 100 échantillons de 3000 observations.

Le principal inconvénient de l'estimateur de Hill est qu'il n'est valable que dans le cas d'un indice positif.

3.2.4 Comparaison des différents estimateurs

Le comportement en termes de précision des différents estimateurs présentés est illustré ci_ après dans le cas d'une loi de Pareto ; on obtient l'allure topique suivante :

des estimat

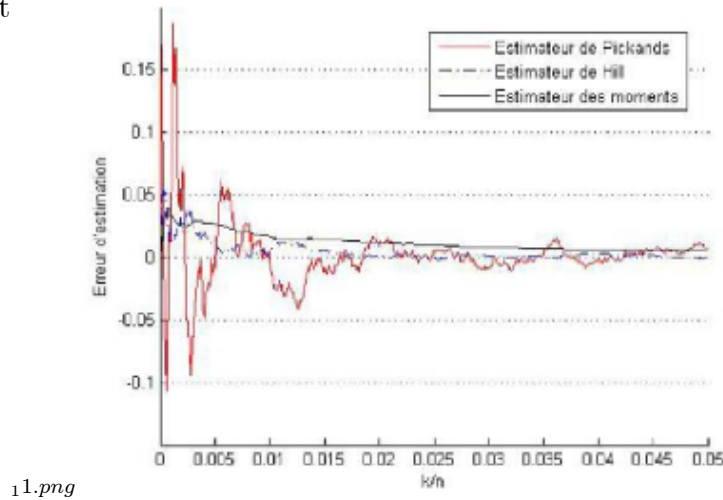


FIG. 3.4 – Comparaison de l'erreur d'estimation en fonction de la méthode.

On remarque que l'estimateur de pickands est moins efficace dans l'estimation de l'indice de queue . Par contre, on observe une efficacité et une suprématie de l'estimateur de Hill sur ceux de Pickands et des moments. Par ailleurs, pour $k < 0.02 \times n$ l'estimateur de Hill est relativement volatile. On serait donc amené à utiliser de l'ordre de 2.5% des données les plus extrêmes pour estimer l'épaisseur de la queue de distribution.

3.3 L'estimation de l'indice de queue par des méthodes paramétrique

3.3.1 Méthode de maximum de vraisemblance

L'estimation par le maximum de vraisemblance donne les résultats asymptotiques efficaces . Soit X_1, \dots, X_n un échantillon de n variables aléatoires supposées i.i.d , de densité h_θ , où $\theta = (\gamma, \mu, \sigma)$.

1. Dans le cas où $\gamma = 0$ (loi de Gumbel) : Soit $h_{\mu,\sigma}(x)$ la densité de probabilité de gumbel telle que :

$$h_{\mu,\sigma}(x) = \frac{1}{\sigma} \exp \left(-\frac{x - \mu}{\sigma} \right) \exp \left\{ -\exp \left(-\frac{x - \mu}{\sigma} \right) \right\}$$

La log-vraisemblance de l'échantillon aléatoire X_1, \dots, X_n est donnée par :

$$\log L(\mu, \sigma) = -n \log(\sigma) - \sum_{i=1}^n \frac{x_i - \mu}{\sigma} - \sum_{i=1}^n \exp\left(\frac{x_i - \mu}{\sigma}\right) \quad (3.1)$$

En prenant la dérivée de $\log L(\mu, \sigma)$ par rapport à μ et σ , on obtient le système d'équations qui permet de maximiser cette fonction :

$$n - \sum_{i=1}^n \frac{x_i - \mu}{\sigma} + \sum_{i=1}^n \frac{x_i - \mu}{\sigma} \exp\left(\frac{x_i - \mu}{\sigma}\right) = 0 \quad (3.2)$$

$$-n + \sum_{i=1}^n \exp\left(\frac{x_i - \mu}{\sigma}\right) = 0 \quad (3.3)$$

On peut toute fois simplifier ce système en écrivant l'équation (3.3) sous la forme

$$\mu = -\sigma \log\left(\frac{1}{n} \sum_{i=1}^n \exp\left(-\frac{x_i}{\sigma}\right)\right) \quad (3.4)$$

et en remplaçant μ dans (3.2) par (3.4) on obtient :

$$\sigma = \bar{x} - \left\{ \sum_{i=1}^n x_i \exp\left(-\frac{x_i}{\sigma}\right) \right\} \left\{ \sum_{i=1}^n \exp\left(-\frac{x_i}{\sigma}\right) \right\}^{-1} \quad (3.5)$$

Les équations (3.4) et (3.5) sont équivalentes au système (3.2) et (3.3). On peut ainsi, pour un échantillon donné x_1, \dots, x_n , déterminer $\hat{\sigma}$ par (3.5) et ensuite connaissant $\hat{\sigma}$ en déduire $\hat{\mu}$ par (3.4).

Pour résoudre (3.5), on doit employer une méthode itérative comme celle de Newton-Raphson.

Les estimateurs du maximum de vraisemblance sont asymptotiquement non biaisés. Par contre, pour une taille d'échantillon finie, ils sont généralement biaisés. En particulier, les estimateurs $\hat{\mu}$ et $\hat{\sigma}$ obtenus dans le cas de la loi de Gumbel possèdent un biais.

2. Dans le cas où $\gamma \neq 0$: Soit X_1, \dots, X_n n variable aléatoire indépendantes de même distribution

GEV , la fonction de densité de probabilité d'une GEV est donné par :

$$h_{\mu,\sigma,\gamma}(x) = \frac{1}{\sigma} \left(1 + \frac{\gamma}{\sigma}(x - \mu)\right)^{-\frac{1}{\gamma}-1} \exp \left\{ - \left(1 + \frac{\gamma}{\sigma}(x - \mu)\right)^{-\frac{1}{\gamma}} \right\} \quad (3.6)$$

Considérons la variable auxiliaire

$$y(x) = -\frac{1}{\gamma} \log \left(1 + \frac{\gamma}{\sigma}(x - \mu)\right)$$

ce qui est équivalent à

$$1 + \frac{\gamma}{\sigma}(x - \mu) = e^{-\gamma y(x)}$$

La relation (3.6) devient :

$$h_{\mu,\sigma,\gamma}(x) = \frac{1}{\sigma} (e^{-\gamma y(x)})^{-\frac{1}{\gamma}-1} \exp \left[- (e^{-\gamma y(x)})^{-\frac{1}{\gamma}} \right]$$

ce qui s'écrit encore

$$h_{\mu,\sigma,\gamma}(x) = \frac{1}{\sigma} e^{(\gamma+1)y(x)} \exp \left[-e^{y(x)} \right]$$

La log-vraisemblance est alors :

$$\log L(\mu, \sigma, \gamma) = -n \log \sigma + (1 + \gamma) \sum_{i=1}^n y(x_i) - \sum_{i=1}^n e^{y(x_i)} \quad (3.7)$$

Les estimateurs du maximum de vraisemblance sont obtenus en maximisant la fonction (3.7)

. On peut montrer que le système d'équations à résoudre est :

$$\frac{d \log L(\mu, \sigma, \gamma)}{d\mu} = \frac{Q}{\sigma} \quad (3.8)$$

$$\frac{d \log L(\mu, \sigma, \gamma)}{d\sigma} = \frac{1}{\sigma} \frac{P - Q}{\gamma} \quad (3.9)$$

$$\frac{d \log L(\mu, \sigma, \gamma)}{d\gamma} = \frac{1}{\gamma} \left(R - \frac{P + Q}{\gamma} \right) \quad (3.10)$$

où

$$P = n - \sum_{i=1}^n e^{y(x_i)}$$

$$Q = (1 + \gamma) \sum_{i=1}^n e^{\gamma y(x_i)} - \sum_{i=1}^n e^{(1+\gamma)y(x_i)}$$

$$R = n - \sum_{i=1}^n y(x_i) + \sum_{i=1}^n y(x_i) e^{-y(x_i)}$$

Ce système d'équations ne peut être résolu explicitement. C'est pourquoi nous devons utiliser une méthode itérative pour trouver la solution .

3.3.2 Méthode des moments

1. Dans le cas où $\gamma = 0$: Le moment d'ordre 1 d'une variable aléatoire X de distribution de Gumbel est tel que :

$$\mu'_1(X) = E(X) = \int_{-\infty}^{+\infty} \frac{x}{\sigma} \exp\left(-\frac{x-\mu}{\sigma}\right) \exp\left\{-\exp\left(-\frac{x-\mu}{\sigma}\right)\right\} dx$$

Procédons au changement de variable $\frac{x-\mu}{\sigma} = y$. D'où $x = \mu + \sigma y$ et $dx = \sigma dy$.

Alors

$$E(X) = \int_{\mathbb{R}} (\mu + \sigma y) \exp(-y) \exp\{-\exp(-y)\} dy$$

Procédons alors à un deuxième changement de variable $\exp(-y) = z$. D'où $y = -\log z$ et $dy = -\frac{dz}{z}$. Donc :

$$E(X) = \mu \int_0^{+\infty} \exp(-z) dz - \sigma \int_0^{+\infty} \log(z) \exp(-z) dz$$

De plus l'intégrale $\int_0^{+\infty} \log(z) \exp(-z) dz = C = 0.5772\dots$, constante d'Euler . Finalement :

$$E(X) = \mu - \sigma C \tag{3.11}$$

Le moment centré d'ordre 2 d'une distribution de Gumbel est tel que :

$$Var(X) = \int_{-\infty}^{+\infty} \frac{x^2}{\sigma} \exp\left(-\frac{x-\mu}{\sigma}\right) \exp\left\{-\exp\left(-\frac{x-\mu}{\sigma}\right)\right\} dx - [E(X)]^2$$

On procède au même changement de variable en y que précédemment. On aura :

$$Var(X) = \int_{\mathbb{R}} (\mu + \sigma y)^2 \exp(-y) \exp\{-\exp(-y)\} dy - [E(X)]^2$$

Idem pour le changement de variable de y en z :

$$\begin{aligned} Var(X) + [E(X)]^2 &= \int_0^{\infty} (\mu - \sigma \log z)^2 \exp(-z) dz \\ &= \mu^2 \int_0^{\infty} \exp(-z) dz - 2\mu\sigma \int_0^{\infty} \log z \exp(-z) dz + \sigma^2 \int_0^{\infty} (\log z)^2 \exp(-z) dz \end{aligned}$$

où $\int_0^{\infty} \exp(-z) dz = 1$, $\int_0^{\infty} \log z \exp(-z) dz = C$ et $\int_0^{\infty} (\log z)^2 \exp(-z) dz = \frac{\Pi^2}{6} + C^2$

Finalement

$$Var(X) = \frac{\sigma^2 \Pi^2}{6} \quad (3.12)$$

En égalant les équations (3.11) et (3.12) à la moyenne \bar{X} et l'écart type S , on obtient le système suivant :

$$\bar{X} = \mu + C\sigma$$

$$S = \frac{\Pi}{\sqrt{6}}\sigma$$

d'où l'on déduit aisément les estimateurs des moments :

$$\hat{\sigma} = 0.77969S \quad (3.13)$$

$$\hat{\mu} = \bar{X} - 0.45005S \quad (3.14)$$

Les variances et les covariances asymptotiques des estimateurs des moments sont obtenus à partir des équations (3.13) et (3.14)

2. Dans le cas où $\gamma \neq 0$: On utilise la méthode des moments pour l'ajustement de la loi GEV lorsque $\gamma \neq 0$. Toutefois, elle peut fournir des estimations pour les paramètres de cette loi lorsque la méthode de maximum de vraisemblance ne converge pas.

Rappelons d'abord les résultats du calcul de la moyenne μ'_1 , de la variance μ_2 et du skewness C_s d'une loi GEV :

$$\mu'_1(X) = \mu + \frac{\gamma}{\sigma} [1 - \Gamma(1 + \gamma)] \quad (3.15)$$

$$\mu_2(X) = \left(\frac{\sigma}{\gamma}\right)^2 [\Gamma(1 + 2\gamma) - \Gamma^2(1 + \gamma)] \quad (3.16)$$

$$\begin{aligned} C_s(X) &= \frac{\mu_3(X)}{\mu_2^{\frac{3}{2}}} \\ &= -\frac{|\gamma|}{\gamma} \frac{[\Gamma(1 + 3\gamma) - 3\Gamma(1 + 2\gamma)\Gamma(1 + \gamma) + 2\Gamma^3(1 + \gamma)]}{[\Gamma(1 + 2\gamma) - \Gamma^2(1 + \gamma)]^{\frac{3}{2}}} \end{aligned} \quad (3.17)$$

La méthode des moments appliquées à la loi GEV consiste à évaluer la moyenne, l'écart type et le coefficient d'asymétrie de l'échantillon aux valeurs empiriques correspondantes de la population :

$$\bar{X} = \mu'_1(X) \quad (3.18)$$

$$S = \sqrt{\mu_2(X)} \quad (3.19)$$

$$\hat{C}_s = \frac{\mu_3(X)}{\mu_2^{\frac{3}{2}}} \quad (3.20)$$

où

$$\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i$$

$$S^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2$$

et

$$\hat{C}_s = \frac{n^2}{(n-1)(n-2)} \frac{m^3}{S^3}$$

avec

$$m^3 = \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^3$$

A partir des équations (3.18), (3.19), (3.20) et de (3.15), (3.16) et (3.17), on déduit le système d'équations suivant :

$$\bar{X} = \mu + \frac{\gamma}{\sigma} [1 - \Gamma(1 + \gamma)] \quad (3.21)$$

$$S = \frac{\sigma}{|\gamma|} [\Gamma(1 + 2\gamma) - \Gamma^2(1 + \gamma)]^{\frac{1}{2}} \quad (3.22)$$

$$\hat{C}_s = -\frac{|\gamma|}{\gamma} \frac{[\Gamma(1 + 3\gamma) - 3\Gamma(1 + 2\gamma)\Gamma(1 + \gamma) + 2\Gamma^3(1 + \gamma)]}{[\Gamma(1 + 2\gamma) - \Gamma^2(1 + \gamma)]^{\frac{3}{2}}} \quad (3.23)$$

Evidemment, on ne peut résoudre ce système de manière explicite. On doit utiliser une procédure itérative . Remarquons toutefois que l'équation (3.23) ne dépend que du paramètre γ . Il suffit alors de résoudre cette équation par la méthode de Newton-Raphson pour obtenir l'estimateur des moments $\hat{\gamma}$, et ensuite de déterminer $\hat{\sigma}$ et $\hat{\mu}$ à l'aide des équations suivantes déduites de (3.21) et (3.22) :

$$\hat{\sigma} = |\hat{\gamma}| [\Gamma(1 + 2\hat{\gamma}) - \Gamma^2(1 + \hat{\gamma})]^{-\frac{1}{2}} S$$

$$\hat{\mu} = \bar{X} - \frac{\hat{\sigma}}{\hat{\gamma}} [1 - \Gamma(1 + \hat{\gamma})]$$

Conclusion

Nous avons, dans ce mémoire essayé de présenter d'une façon très simple et accessible l'estimation paramétrique surtout dans le cas des distributions des valeurs extrêmes, vu leurs importance et leurs utilisation dans beaucoup de domaines sensibles telsque les finances, l'actuariat et l'assurance, l'hydrologie...etc.

Après avoir étudié des valeurs extrêmes et leurs distributions possibles, nous avons présentés les trois forme de distributions des valeurs extrêmes citées notamment dans *Jenkison – Von Mises* pour la *GEV* et dans *Balkema – De Haan – Pickands* pour la distribution de *Pareto Généralisé (GPD)* .

L'estimation paramétrique des paramètres de position, échelle et de l'indice de queue ou l'indice des valeurs extrêmes sont aussi détaillés dans notre travail .

Bibliographie

- [1] Stuart Coles. An Introduction to Statistical Modeling of Extreme Values. Springer Series in Statistics, Springer-Verlag, London, 2001.
- [2] A. L. M. Dekkers et L. de Haan. On the estimation of the extreme value index and large quantile estimation. Ann. statist., 17(4) : 1795-1832, 1989.
- [3] DEKKERS A.L.M., EINMAHL J.H.J., DE HAAN. (1989), A moment estimator for the index of an extrem value distribution, The Annals of Statistics, Vol. 17 n° 4, 1833-1855.
- [4] De Haan L., Ferreira A. (2006). Extreme Value Theory : An Introduction Springer-Verlag.
- [5] Frédéric Planchet. Utilisation de la théorie des valeurs extrêmes dans le cadre de solvabilité 2, Support de cours 2017-2018, d'Institut de Science Financière et d'Assurance.
- [6] J.Beirlant, Y. Geogebur, J. Teugels (2005), Statistics of extremes theory and application, copyright 2004 John Weley and Sons.
- [7] Kots S et Nadarajah S 2000. Extreme Value Distributions. Imprial College Press.
- [8] PICKANDS J. (1975), Statistical inference using extreme order statistics, Ann. Statis, 3, 119-131.
- [9] S. Resnick, Extreme Values, Regular Variation, and Point Processes. Springer Verlag. 1987.
- [10] Toubia Sonia. L'approche des Valeurs Extrêmes en Statistique des séries financières. Thèse de Doctorat de l'Université de Mohamed Khider, Biskra, 2005-2006.

Annexe : Abréviations et Notations

Les différentes abréviations et notations utilisées tout au long de ce mémoire sont expliquées ci-dessous :

v : a : variable aléatoire.
 i : i : d : Indépendant et identiquement distribué.
 $F(\cdot)$: Fonction de distribution.
 $F_n(\cdot)$: Fonction de distribution empirique.
TCL : Théorème centrale limite.
TVE : Théorie des valeurs extrêmes.
 $(X_1:n; \dots; X_n:n)$: Statistique d'ordre de $(X_1; \dots; X_n)$.
 $E[X]$: Espérance de la v.a X .
 $\text{Var}(X)$: Variance de la v.a X .
GEV : Distribution des Valeurs Extrêmes Généralisée.
EMV : Méthode de Maximum de Vraisemblance.
GPD : Distribution de Pareto Généralisée.
POT : L'approche par dépassement de seuil.
 H : Distribution des valeurs extrêmes.
 $D(H)$: Domaine d'attraction de H .
 Q : Fonction inverse généralisées de F .
 Q_n : Fonction de quantile empirique.
 $S : O$: Statistique d'ordre.
 IA : Fonction indicatrice de l'événement A .