الجمهورية الجزائرية الديمقراطيـة الشعبيــة
République algérienne démocratique et populaire
وزارة التـعليـم العالـي والبحـث العلمـــي
Ministère de l'enseignement supérieur et de la recherche scientifique
جامعة عين تموشنت بلحاج بوشعيب
Université –Ain Temouchent- Belhadj Bouchaib
Faculté des Sciences et de Technologie
Département de Mathématiques et Informatique

Projet de Fin d'Etudes
Pour l'obtention du diplôme de Master en Mathématiques
Domaine : Mathématiques et Informatique
Filière : Mathématiques
Spécialité : Probabilités et Statistique Appliquées

Thème :

Méthodes d'estimation d'une distribution alpha-stable

Présenté Par :

1) M^{elle} Sekka Aicha

Devant le jury composé de :

BALASKA Lamia	M.A.A	UAT.B.B (Ain Temouchent)	Présidente
Dr. BENNAFLA Djamila	M.C.B	UAT.B.B (Ain Temouchent)	Examinatrice
Dr. MAMI Tawfiq Fawzi	M.C.A	UAT.B.B (Ain Temouchent)	Encadrant

Année Universitaire 2023/2024

Dédicace

Avec mes plus profondes expressions

d'amour, de gratitude

et d'appréciation, Je dédie ce travail à :

A mes chers parents pour leurs affections

et leur amour

Que Dieu les garde et les protège

A mes chers frères

A ma chère sœur

A toutes personnes qui m'ont aidé de

près ou de loin

À mes amies.

Je dédicace ce modeste travail.

Remerciements

Avant toute chose, on tient à remercier Dieu le tout puissant, pour nous avoir donné la force et la patience.

Nous tenons tout d'abord à remercier notre encadreur de ce mémoire monsieur MAMI Tawfiq Fawzi professeur à l'université Belhadj BOUCHAIB AIN TEMOUCHENT.

Recevez ici nos sincères remerciements pour la confiance, les conseils que vous nous avez accordés tout le long de ce travail. Merci également pour votre encadrement, votre disponibilité et votre gentillesse.

Nous vous adressons notre profonde reconnaissance pour vos remarques et conseils en vue d'améliorer ce manuscrit.

Nous tenons à exprimer nos vifs remerciements à tous les membres de jury pour le temps qu'ils nous ont accordé, ainsi que pour les remarques que nous accepterons avec humilité.

Enfin on remercie tous les enseignants du parcours Master et à l'ensemble des membres du Département de Mathématique et informatique - Université de Belhadj BOUCHAIB Ain t'émouchent pour leur aide pédagogique durant nos années.

Liste des figures

Liste des tableaux

Index des notations

v.a.r.	Variable aléatoire réelle
i.e.	c'est à dire
i.i.d.'s	Indépendamment et identiquement distribué(e)s
f.c.	Fonction caractéristique
$S(\alpha, \sigma, \beta, m)$	Loi stable de paramètres α, σ, β, m
f(x)	Fonction de densité de la v.a.r. X
$sign(\theta)$	c'est la fonction $\theta/\|\theta\|$
$I_{[a,b[}$	Fonction indicatrice : I(x)=1 si $x \in [a, b[$ et I(x)=0 sinon
exp()	Fonction exponentielle
$\mathcal{R}$	Ensemble des nombres réels
$\mathcal{R}^+$	Ensemble des nombres réels positifs
$\stackrel{d}{=}$	"suit la même loi que" ou "égalité en distribution"
$S\alpha S(\sigma)$	Loi α-stable symétrique de paramètre σ

Table des matières

Introduction

Les distributions stables de Lévy dites α-stables sont des lois de probabilité qui présentent certaines caractéristiques que d'autres lois ne possèdent pas. D'abord, leurs densités lisses et unimodales jouissent d'une certaine flexibilité dans leurs allures globales ; Elles peuvent être symétriques par rapport à une valeur centrale comme elles peuvent avoir une dissymétrie par rapport à cette valeur ; Tout celà, contrôlé par le paramètre d'asymétrie de la loi ; Elles peuvent avoir des queues épaisses gardant assez d'informations aux extrémités et ceci grâce au paramètre de stabilité ; On peut agrandir ou rétrécir la densité grâce à son paramètre d'échelle et enfin, on peut déplacer le graphe de cette densité sur l'axe des abscisses selon les valeurs de son paramètre de localisation.

Apparemment, avec ces quatre paramètres on peut ajuster n'importe quelle suite de données statistiques seulement qu'il faut satifaire certaines conditions requises lors de l'adaptation du modèle à cette suite.

Les lois stables ont été étudiées par Paul Lévy et Yakovlevich Khintchine entre les années 1920 et 1930. Initialement, l'idée était de généraliser la propriété constatée sur les lois gaussiennes (et même sur celles de Cauchy) dont l'objet était qu'une somme finie de variables gaussiennes indépendantes suit une loi de même type c'est à dire une loi gaussienne. Autrement dit, il y a conservation du type de loi par la somme ; C'est la stabilité par addition. Cette stabilité est encore conservée, à une constante d'échelle et une constante de position près, si on considérait des combinaisons linéaires finies de variables aléatoires réelles (v.a.r.'s) gaussiennes indépendantes, donc il y a conservation de la stabilité même par linéarité.

En passant par les fonctions caractéristiques, Paul Lévy est arrivé par son instinct de mathématicien à trouver toute une classe de lois ayant cette propriété de conservation du type de loi par linéarité en partant d'une suite finie de v.a.r.'s indépendantes et de même loi de distribution convenablement renormalisée. Ce sont les distributions α-stables de Lévy. D'ailleurs, il a été établi que ce sont les seules lois limites de sommes de variables indépendantes et identiquement distribuées (i.i.d.'s) normalisées ; Seulement que ces lois ne sont pas explicitées à l'aide des fonctions classiques à l'exception de la loi de Gauss ou de Cauchy ou la gaussienne inverse dite "loi de Lévy". Pour le reste, elle sont calculées via des procédés numériques de simulation.

Les lois stables jouissent encore d'une autre propriété dite "d'infinie divisibilité". Cette propriété consiste à ce que pour une v.a.r. X et pour tout indice naturel $n > 1$, il existe une suite finie de n v.a.r.'s i.i.d.'s ayant pour somme X. Cette dernière est du même type de loi que celles qui forment la somme. Intuitivement, on peut toujours "fragmenter" la variable X en somme finie, à n'importe quel ordre, de v.a.r's indépendantes ayant même loi commune. Cette propriété est d'ailleurs partagée par pas mal de variables continues ou discrètes comme la loi de Poisson, la loi de Gauss ou celle de Cauchy etc. Signalant qu'en général, une loi infiniment divisible n'est pas nécessairement une loi stable.

Pour ce qui est de l'estimation des paramètres d'une loi stable, il existe une multitude d'approches qui mènent à trouver les bonnes estimations pour ces paramètres. La méthode du maximum de vraissemblance donne de meilleurs résultats seulement qu'elle est coûteuse en terme de temps d'exécution de la procédure de calcul. La méthode de régression de Koutrouvilis est une autre approche qui donne des résultats assez bons au fur et à mesure que le pas de l'algorithme avance dans son exécution. Pour ce qui de la méthode de MacCulloch basée sur les quantiles, elle reste la méthode qui donne de bons résultats pour un temps de calcul assez rapide seulement qu'elle est valable pour des valeurs d'indice de stabilité supérieur à 0.6.

Dans le premier chapitre, on commencera dans un premier temps à rappeler les lois infiniment (on dit aussi indéfiniment) divisibles, leur définition, l'expression de leur fonction caractéristique avec quelques exemples de ces lois. Dans un deuxième temps et comme cas particulièrement remarquable de ces lois et ayant des propriétés intéressantes, on trouve les lois stables. On donnera ainsi la définition théorique d'une loi stable en exhibant sa fonction caractéristique et ses quatre paramètres dont elle dépend avec leurs significations respectives. Puis, on étalera les propriétés que possède cette loi surtout celles relatives aux queues qui leur fait perdre le deuxième moment donc la perte de la variance (sauf le cas particulier de la loi de Gauss). A la fin du chapitre, un algorithme de simulation sera décrit montrant comment on arrive à développer un procédé permettant de simuler une loi stable et à calculer sa densité.

Dans le deuxième chapitre, on examinera en détails quelques approches d'estimation des paramètres d'une loi stable les plus usitées dans la pratique tout en évoquant leurs performances du point de vue temps calcul et rapidité d'exécution. Bien sûr, il existe une multitude d'autre méthodes dont l'objectf principal est d'une part, d'avoir la meilleure précision pour les estimations et d'autres part, l'aptitude pour l'estimateur à aller vite vers la valeur estimée.

Dans la dernière section du dernier chapitre, une étude comparative fera l'objet d'une simulation lancée sur le logiciel R pour mettre en exergue l'importance du temps calcul et de la précision des estimateurs (consistance et efficacité). On montrera également les étapes à suivre pour ajuster la séquence de données à un modèle de distribution α-stable.

Chapitre I

Généralités sur les lois stables

I.1 Lois indéfiniment divisibles

Avant de définir les lois $\alpha-$stables, nous allons introduire une famille de lois plus générales : Les lois indéfiniment divisibles. C'est à partir de ces lois que sera précisée par la suite, la forme d'une fonction caractéristique d'une loi stable.

L'intérêt principal de telles lois réside dans la solution du problème suivant : comment détermine-t-on toutes les distributions qui s'expriment comme limite d'une somme de n variables aléatoires réelles (v.a.r's) indépendantes et identiquement distribuées (i.i.d.'s) ?

Intuitivement, si une suite de variables aléatoires réelles S_n converge en loi vers une v.a.r. S et que ces S_n s'expriment comme une somme de n v.a.r's indépendantes de même loi, alors S va aussi s'exprimer de la même manière.

Définition I.1.1

Une variable aléatoire réelle X est dite indéfiniment divisible si et seulement si pour tout $n \geq 1,$ il existe n variables aléatoire $X_1, ..., X_n$ indépendantes ayant même loi de distribution telles que :

$$X \stackrel{d}{=} X_1 + X_2.... + X_n.$$

N.B : Autrement exprimé, une variable aléatoire réelle X est indéfiniment divisible si et seulement si X est décomposable en une somme arbitrairement finie de v.a.r's i.i.d.'s.

Le symbole $\stackrel{d}{=}$ désigne l'égalité en distribution.

Remarque I.1.1 *Notez que les variables aléatoires réelles X_i n'ont pas nécesairement la même* *distribution que X mais, appartiennent à la même classe de distributions (lois de même type). D'autre part, on peut référer la propriété d'indéfinie divisibilité à la loi de distribution de la variable aléatoire en disant "loi indéfiniment divisible".*

Le résultat suivant due à lévy-Khinchin répond à la question de problème posé ci-dessus :

Théorème I.1.1 (Formule de lévy-Khintchin)

Si X est une v.a.r de loi indéfiniment divisible, alors sa fonction caractéristique se présente comme suit :

$$\varphi_X(\theta) = \exp\left\{ im\theta + \int_{-\infty}^{+\infty} \left(e^{i\theta x} - 1 - \frac{i\theta x}{1+x^2} \right) \frac{1+x^2}{x^2}\, dM(x) \right\}$$

où m est un réel et M est une fonction bornée, non décroissante, continue à droite, vérifiant $M(-\infty) = 0$ et la fonction sous le signe intégrale est prolongeable par continuité en $x = 0$ en prenant la valeur $-\theta^2/2$.

Preuve : Voir démonstration dans [4] et dans [12].

Une autre version de Lévy est donnée par le théorème suivant :

Théorème I.1.2

Si X est une v.a.r. de loi indéfiniment divisible, alors sa fonction caractéristique admet la représentation suivante :

$$\varphi_X(\theta) = \exp\left\{ im\theta - \frac{\sigma^2\theta^2}{2} + \int_{|x|>0} \left(e^{i\theta x} - 1 - \frac{i\theta x}{1+x^2} \right) dL(x) \right\}$$

où m est un réel, σ^2 une constante non négative et $L(x)$ est une fonction non décroissante sur les intervalles $]-\infty, 0[$ et $]0, \infty[$ satisfaisant les conditions :

a) $L(-\infty) = 0$, $L(+\infty) = 0$

b) $\displaystyle\int_{0<|x|<\delta} x^2 dL(x) < \infty$ pour tout $\delta > 0$ fini.

La fonction $L(x)$ est appelée "Fonction spectrale de Lévy".

I.1.1 Exemples de lois indéfiniment divisibles :

Il existe pas mal de lois connues qui sont indéfiniment divisibles. Dans ce qui suit, on en citera quelques unes.

Exemple I.1.1 ***Loi de Gauss :***

La fonction caractéristique d'une loi normale $N(m, \sigma^2)$ s'écrit :

$$\begin{aligned}
\varphi_X(\theta) &= \exp\left\{im\theta - \frac{\theta^2\sigma^2}{2}\right\} \\
&= \left[\exp\left\{i\frac{m\theta}{n} - \frac{\frac{\theta^2\sigma^2}{2}}{n}\right\}\right]^n \\
&= \left[\exp\left\{i\left(\frac{m}{n}\right)\theta - \frac{\theta^2\left(\frac{\sigma^2}{n}\right)}{2}\right\}\right]^n
\end{aligned}$$

C'est une puissance $n^{ième}$ de la fonction caractéristique d'une loi normale $N(\frac{m}{n}, \frac{\sigma^2}{n})$. Autrement dit, une v.a.r. normale est la somme de n v.a.r.'s indépendantes de même loi gaussienne et ceci pour tout ordre n de N.

Exemple I.1.2 ***Loi de Poisson :***

La loi de poisson $P(\lambda)$ a pour fonction caractéristique :

$$\begin{aligned}
\varphi_X(\theta) &= \exp\left\{\lambda(e^{i\theta} - 1)\right\} \\
&= \left[\exp\left\{\frac{\lambda}{n}(e^{i\theta} - 1)\right\}\right]^n
\end{aligned}$$

Elle s'écrit comme puissance $n^{ième}$ de la fonction caractéristique d'une loi de poisson $P(\frac{\lambda}{n})$. Une v.a.r. de Poisson peut être considérée comme une somme arbitraire de n v.a.r.'s indépendantes et ceci pour tout ordre de fractionnement n.

Exemple I.1.3 ***Loi Gamma :***

Pour une loi Gamma $\Gamma(r, \lambda)$, la fonction caractéristique s'écrit :

$$\begin{aligned} \varphi_X(\theta) &= \left(1 - \frac{i\theta}{\lambda}\right)^{-r} \\ &= \left[\left(1 - \frac{i\theta}{\lambda}\right)^{-\frac{r}{n}}\right]^n \end{aligned}$$

comme puissance $n^{\text{ième}}$ de la fonction caractéristique d'une loi Gamma $\Gamma\left(\frac{r}{n}, \lambda\right)$.

Une v.a.r. de loi Gamma est donc somme de n v.a.r.'s indépendantes de même loi Gamma $\Gamma\left(\frac{r}{n}, \lambda\right)$ et ceci pour tout entier positif n.

N.B. Une loi de $\chi_2(m)$ (m étant son degré de liberté) ou une la loi exponentielle de paramètre λ sont des cas particuliers de la loi Gamma. En effet, $\chi_2(m) = \Gamma(\frac{m}{2}, \frac{1}{2})$ et $\exp(\lambda) = \Gamma(1, \lambda)$.

Exemple I.1.4 ***Loi de Cauchy :***

La fonction caractéristique d'une variable aléatoire de Cauchy $C(m, \mu)$ a pour expression :

$$\begin{aligned} \varphi_X(\theta) &= \exp(im\theta - \mu|\theta|) \\ &= \exp\left(i\frac{m}{n}\theta - \frac{\mu}{n}|\theta|\right)^n \end{aligned}$$

comme puissance $n^{\text{ième}}$ de la fonction caractéristique d'une loi de Cauchy $C(m/n, \mu/n)$. Une v.a.r. de Cauchy est donc somme de n v.a.r.'s indépendantes de même loi de Cauchy $C(\frac{m}{n}, \frac{\mu}{n})$ et ceci pour tout ordre de fractionnement n.

N.B : Il existe des v.a.r.'s de lois qui ne sont pas indéfiniment divisibles comme par exemple une loi uniforme, une loi binomiale, une loi Bêta etc. Ce sont des lois de distribution à supports bornés.

I.2 Variables aléatoires α-stables

Parmi les exemples de lois indéfiniment divisibles il existe encore un notre type de loi ayant des propriétés très intéressantes. Ce sont les lois de Lévy dites α-stables.

Définition I.2.1

Une variable aléatoire X est dite possédant une distribution stable si pour n'importe quelle suite finie $X_1, \ldots, X_n$ de v.a.r.'s indépendantes de même loi que X, il existe un nombre positif C_n et un nombre réel D_n, tels que :

$$X_1 + X_2 + ...X_n \stackrel{d}{=} C_n X + D_n, \tag{I.1}$$

N.B. L'égalité considérée est une égalité entre lois de distribution. C'est à dire la somme du premier membre possède la même loi que celle de la combinaison du second membre.

Remarque I.2.1

— *Lorsque* $D_n = 0$, *on parle de distribution strictement stable.*

— *On montre dans [4] que* C_n *s'écrit sous la forme :* $C_n = n^{\frac{1}{\alpha}}$ *appelée "constante de normalisation" avec un certain* $0 < \alpha < 2$.

— *Cette constante joue un rôle très important, c'est pourquoi on appelle une distribution stable :* ***"loi α-stable"***.

I.3 Fonction caractéristique d'une loi stable

Une loi stable admet une fonction caractéristique donnée par le théorème suivant :

Théorème I.3.1

Pour que la variable aélatoire X soit stable, il est nécéssaire et suffisant que sa fonction caractéristique soit de la forme :

$$\varphi_X(\theta) = \begin{cases} \exp\left\{-\sigma^\alpha|\theta|^\alpha(1 - i\beta sign(\theta)\tan\frac{\pi\alpha}{2}) + im\theta\right\} \text{ si } \alpha \neq 1, \\ \\ \exp\left\{-\sigma|\theta|(1 + i\beta\frac{2}{\pi}sign(\theta)\log|\theta|) + im\theta\right\} \text{si } \alpha = 1. \end{cases} \tag{I.2}$$

où $\alpha, \beta, \sigma,$ et m sont des constantes telles que : $0 < \alpha \leq 2$, $\sigma \geq 0$, $-1 \leq \beta \leq 1$ et $m \in R$ avec :

$$sign(\theta) = \{1 \text{ si } \theta > 0, \quad 0 \text{ si } \theta = 0, \ \ -1 \text{ si } \theta < 0\}.$$

Preuve : Voir démonstration dans [5].

Cette représentation de la fonction caractéristique est appelée paramétrisation standard. Il existe d'autres paramétrisations de la fonction caractéristique comme celle décrite dans zotolarev [17] :

$$\varphi_X(\theta) = \begin{cases} \exp\left\{-\sigma^\alpha|\theta|^\alpha\left(1 + i\beta sign(\theta)\tan\frac{\pi\alpha}{2}[(\sigma|\theta|)^{1-\alpha} - 1] + im\theta\right)\right\} \text{ si } \alpha \neq 1, \\ \\ \exp\left\{-\sigma|\theta|\left(1 + i\beta\frac{2}{\pi}sign(\theta)\log(\sigma|\theta|)\right) + im\theta\right\} \text{si } \alpha = 1. \end{cases} \tag{I.3}$$

de sorte que les paramètres $\alpha, \sigma,$ β sont ceux de la paramétrisation standard, mais m et m_0 sont reliés par les relations :

$$m = \begin{cases} m_0 - \sigma\beta\tan(\pi\alpha/2) \text{ quand } \alpha \neq 1, \\ \\ m_0 - \sigma\beta\tan(2/\pi)\log(\sigma) \text{ quand } \alpha = 1. \end{cases} \tag{I.4}$$

Dans ce cas de paramétrisation, l'avantage est que les fonctions de densité et de distribution cumulative sont continues comme fonctions des quatre paramètres. Cette paramétrisation est bien conditionnée pour le calcul numérique chose que ne possède pas la paramétrisation standard car elle n'est pas continue en ses paramètres.

Proposition I.3.1 *Une variable stable est indéfiniment divisible.*

Preuve :

Pour les v.a.r.'s $X_1, X_2, \ldots, X_n$ indépendantes de même loi que X , soient $Y_1, Y_2, \ldots, Y_n$ les v.a.r.'s pareillement i.i.d.'s avec $Y_k = \frac{1}{C_n}(X_k - \frac{D_n}{n})$ pour $k = 1, 2, \ldots n$ et si $C_n \neq 0$,
on constate que :

$$Y_1 + Y_2 + \ldots + Y_n = \frac{1}{C_n}(X_1 + X_2 + \ldots + X_n) - \frac{D_n}{C_n}$$

Or, par définition de la stabilité

$$X_1 + X_2 + \ldots + X_n \stackrel{d}{=} C_n X + D_n$$

et donc,

$$Y_1 + Y_2 + \ldots + Y_n \stackrel{d}{=} X$$

N.B. La réciproque de ce résultat est généralement fausse.

I.4 Paramètres et Notations

Le théorème nous indique que la fonction caractéristique d'une distribution α-stable dépend de quatre paramètres. Ces paramètres donnent la possibilité d'avoir une grande flexibilité à représenter graphiquement les densités et les répartitions de telles lois. Cet avantage n'existe pas dans la plupart des distributions de probabilité habituelles.

— **Le paramètre α est appelé "indice de stabilité"** : D'après le théorème (I.3), il appartient à l'intervalle]0,2]. C'est le paramètre principal car il caractérise les queues de distribution ; Plus α diminue, plus les queues sont " épaisses" on dit qu'elles sont "lourdes" ; En augmentant, les queues deviennent moins épaisses et par suite moins lourdes. On l'appelle aussi **"exposant caractéristique"**.

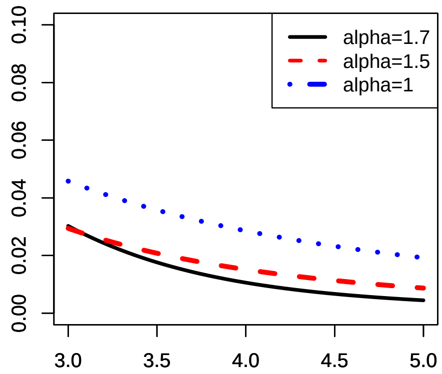

FIGURE I.1 – Différentes queues de lois stables.

— Le paramètre β est appelé **"paramètre d'asymétrie"** ; Il prend ses valeurs dans l'intervalle [-1,1] et mesure le degré de dissymétrie de la densité relativement à un point central.

Si $\beta = 0$, la distribution est parfaitement symétrique.

Si $\beta = -1$ (resp $\beta = 1$), la distribution est totalement asymétrique à gauche (resp à droite).

Lorsque β est positif, la queue de distribution est plus épaisse à droite qu'à gauche.

Lorsque β est négatif, la queue de distribution est plus épaisse à gauche qu'à droite.

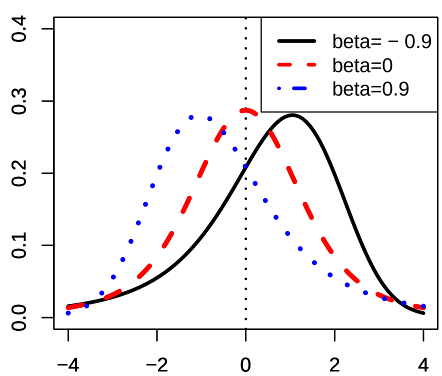

FIGURE I.2 – Différentes valeurs de β.

— **Le paramètre σ est appelé paramètre d'échelle** : Il prend ses valeurs dans l'intervalle $[0, +\infty[$ et joue un rôle similaire à celui de l'écart-type dans le cas d'une variable gaussienne. Il mesure le degré de dispersion des valeurs de la distribution autour du point central.

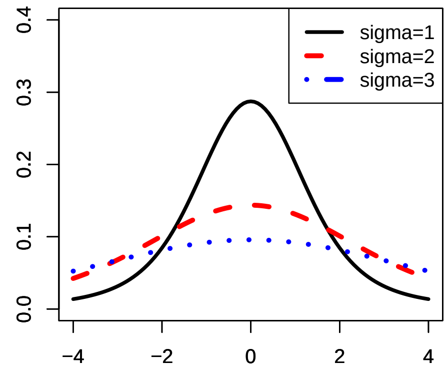

FIGURE I.3 – Différentes valeurs de σ.

— **le paramètre m est dit "paramètre de position" ou "paramètre de localisation"** : il indique la tendance centrale de la distribution qui peut être soit une moyenne soit une médiane soit une valeur centrale selon les valeurs des paramètres α et β.

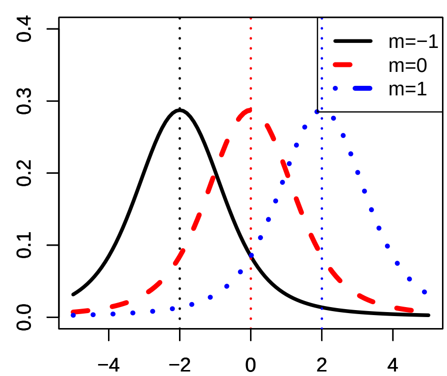

FIGURE I.4 – Différentes valeurs de m.

Notations

— Une loi stable de paramètres $\alpha, \beta, \sigma,\ m$ est notée : $S(\alpha,\ \sigma,\ \beta,\ m)$ et parfois $S_\alpha(\sigma,\ \beta,\ m)$.

— Lorsque la loi stable est symétrique par rapport à l'origine (i.e $\beta = 0$ et $m = 0$) elle est notée $S\alpha S(\sigma)$. Sa fonction caractéristique prend la forme simple : $\varphi_X(\theta) = \exp -\sigma^\alpha |\theta|^\alpha$. On dit dans ce cas qu'elle est $\alpha-$stable symétrique de paramètre σ.

I.5 Densité d'une loi stable

La densité d'une variable aléatoire stable ne s'exprime pas à l'aide des fonctions classiques sauf dans trois cas particuliers qui sont la densité d'une loi gaussienne, ou celle de Cauchy ou celle de la gaussienne inverse dite loi de Lévy. En effet :

— La distribution gaussienne $N(m, \sigma\sqrt{2}),$ dont la densité est

$$f(x) = (2\sigma\sqrt{\pi})^{-1} \exp\left(-\frac{(x-m)^2}{4\sigma^2}\right)$$

est une loi 2-stable avec : $N(m, \sigma\sqrt{2}) = S(2, \sigma, 0, m)$.

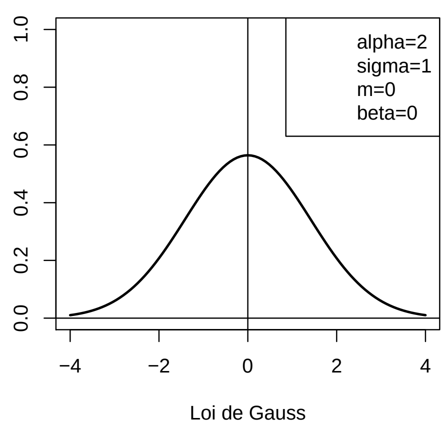

FIGURE I.5 – Densité de Gauss.

— La distribution de Cauchy $C(m, \sigma)$ dont la densité est :

$$f(x) = \frac{\sigma}{\pi((x-m)^2 + \sigma^2)}$$

est une loi 1-stable avec : $C(m, \sigma) = S(1, \sigma, 0, m)$.

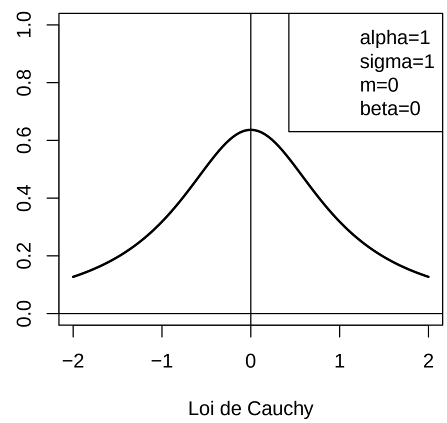

FIGURE I.6 – Densité de Cauchy.

— La distribution de Lévy dont la densité est :

$$f(x) = \left(\frac{\sigma}{2\pi}\right)^{1/2} \left(\frac{1}{x-m}\right)^{3/2} \exp\left\{-\frac{\sigma}{2(x-m)}\right\} \times I_{]m,+\infty[}$$

est une loi $\frac{1}{2}$-stable $S(1/2, \sigma, 1, m)$ totalement asymétrique à droite.

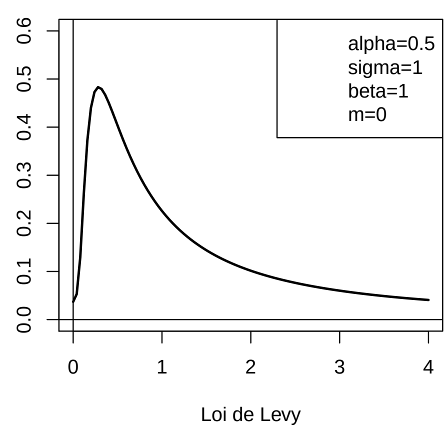

FIGURE I.7 – Densité de la gaussienne inverse

Ici, I_A indique la fonction indicatrice de la partie A :

$$I_A(x) = \begin{cases} 1 \text{ si } x \in A, \\ \\ 0 \text{ si } x \notin A. \end{cases} \tag{I.5}$$

Pour le reste des cas de densités ou répartitions d'une loi stable, on a pas d'expressions fermes seulement qu'on peut les exprimer à l'aide de la transformée de Fourier inverse comme suit :

$$f(x) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} e^{-i\theta x} \varphi_X(\theta) \; d\theta$$

.

qui donne l'expression intégrale suivante :

$$f(x) = \frac{1}{\pi} \int_0^{+\infty} \exp(-\sigma^\alpha \theta^\alpha) \cos[(x-m)\theta + \beta\theta^\alpha w(\theta,\alpha)] \; d\theta. \tag{I.6}$$

où :

$$w(\theta,\alpha) \quad = \quad \begin{cases} tg(\frac{\alpha\pi}{2}) \quad si \quad \alpha \neq 1, \\ \\ \frac{2}{\pi} \ln|\theta| \quad si \quad \alpha = 1. \end{cases}$$

Sinon, on passe par l'intégration numérique directe utilisée par (Nolan 1997,1999) (voir [9]) et qui permet de donner des représentations graphiques pour toute loi α-stable quel que soient ses paramètres.

I.6 Propriétés des lois stables

I.6.1 Propriété de(s) queue(s)

Proposition I.6.1 *Soit X une v.a.r. $S(\alpha,\sigma,\beta,m)$, avec $0 < \alpha < 2$ on a les deux relations suivantes :*

$$\begin{cases} \lim\limits_{x\to+\infty} P(X > x) \quad = \quad C_\alpha \dfrac{1+\beta}{2} \sigma^\alpha \; x^{-\alpha}, \\ \\ \lim\limits_{x\to+\infty} P(X < -x) \quad = \quad C_\alpha \dfrac{1-\beta}{2} \sigma^\alpha \; x^{-\alpha}, \end{cases} \tag{I.7}$$

Où C_α est une constante qui ne dépend que de α. Elle est évaluée par la relation :

$$C_\alpha = \left(\int_0^{+\infty} t^{-\alpha} \sin t \; dt\right)^{-1} = \begin{cases} 2/\pi & si \quad \alpha = 1 \\ \dfrac{1-\alpha}{\Gamma(2-\alpha)\cos(\pi\alpha/2)} & si \quad \alpha \neq 1 \end{cases}$$

N.B. $\Gamma(x)$ est la fonction Gamma définie , pour x>0, par :

$$\Gamma(x) = \int_0^{+\infty} t^{x-1} e^{-t} \; dt.$$

Le terme en puissance $x^{-\alpha}$ dans les deux relations (I.7) rappelle la loi dite "de Pareto" qui a un comportement asymptotique similaire aux lois stables et dont voici la définition :

Définition I.6.1.1

Une variable aléatoire X suit une **Loi de Pareto** de paramètres $x_0 > 0$ et $\alpha > 0$ si elle admet pour fonction de répartition :

$$F(x) = 1 - Cx^{-\alpha} \quad \text{où} \quad C = x_0^\alpha > 0$$

Toute loi de probabilité ayant un comportement au voisinage de l'infini en $x^{-\alpha}$ avec $\alpha > 0$ est dite de "type Pareto". Précisément on a encore la définition suivante :

Définition I.6.1.2

Une variable aléatoire X suit une loi de **type Pareto** si sa fonction de répartition vérifie :

$$1 - F(x) = x^{-\alpha} L(x) \tag{I.8}$$

avec $x > 0$, $\alpha > 0$ et $L(x)$ est une fonction à variation lente i.e. :

$$\lim_{x\to\infty} \frac{L(tx)}{L(x)} = 1.$$

Remarque I.6.1 *Une variable aléatoire de type Pareto avec $0 < \alpha < 2$ décroît lentement vers zéro que toute queue d'une variable aléatoire suivant une loi exponentielle. Sa queue droite est plus épaisse que celle d'une exponentielle, on dit qu'elle est "lourde" et sa variance est infinie.*

Par conséquent, les distributions α-stables qui ont un comportement asymptotique de type Pareto avec un α vérifiant $0 < \alpha < 2$, sont donc à queue(s) lourde(s) et leurs variances sont infinies.

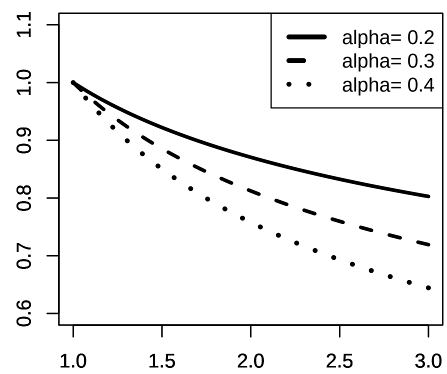

FIGURE I.8 – Queues de Pareto pour différentes valeurs de α

I.6.2 Moments fractionnaires absolus d'une loi stable

Comme on l'a déjà évoqué en haut, une loi α-stable ne possède pas de variance finie ni même l'espérance quand $0 < \alpha \leq 1$. Toutefois, en cas d'absence de ces deux caractéristiques, on peut définir des moments d'ordres fractionnaires comme suit :

Définition I.6.2.1

On appelle moment "fractionnaire absolu" d'une v.a.r X la quantité :

$$E(|X|^p) = \int_{-\infty}^{+\infty} |x|^p f(x)dx$$

où p est un nombre **réel** quelconque.

Pour une loi α-stable, on a la résultat suivant :

Proposition I.6.2 *Si la v.a.r. X suit une loi $S(\alpha, \sigma, \beta, m)$, alors :*

a) $E|X|^p < +\infty, \quad \forall p \in [0, \alpha[$,

b) $E|X|^p = +\infty, \quad \forall p \in [\alpha, 2[$.

En particulier :

1. *Si* $\alpha < 2$ *on a :* $Var(X) = \infty$.
2. *Si* $\alpha > 1$ *on a :* $E|X| < \infty$.
3. *Si* $\alpha \leq 1$ *on a :* $E|X| = \infty$.

Preuve : voir démonstration [4] et [7].

N.B. Pour le cas très particulier de loi stable $\alpha = 2$, on a : $E|X|^p < +\infty, \quad \forall p \in R$ et par conséquent, l'espérance et la variance existent à la fois. C'est le cas d'une loi gaussienne.

I.6.3 Propriétés de la densité

Une loi α-stable est de classe $\mathcal{C}^\infty$:

En effet, dans Zolotarev [17] il est montré le résultat suivant :

Théorème I.6.3.1

Toutes les lois stables non dégénérées sont continues et leurs densités sont unimodales (donc bornées) et infiniment différentiables.

Une loi α-stable vérifie la propriété de réflexion :

Si on suppose : $m = 0$ et $\sigma = 1$ et si on pose $f(x;\alpha,\beta) := f(x;\alpha,1,\beta,0)$ désignant la densité d'une loi stable de paramètres $(\alpha,1,\beta,0)$, on a le résultat suivant :

Proposition I.6.3 *Pour n'importe quels α et β la densité vérifie :*

$$f(x;\alpha,\beta) = f(-x;\alpha,-\beta) \tag{I.9}$$

et la fonction de répartition vérifie :

$$F(x;\alpha,\beta) = 1 - F(-x;\alpha,-\beta) \tag{I.10}$$

Preuve :

- La relation $(I.9)$ vient en remplaçant m par 0 dans $(I.6)$, x et β par leurs symétriques $(-x)$ et $(-\beta)$ puis utiliser la parité du cos.

- Pour $(I.10)$, en remarquant que :

$$\begin{aligned} F(-x;\alpha,-\beta) &= \int_{-\infty}^{-x} f(t;\alpha,-\beta)\,dt = \int_{-\infty}^{-x} f(-t;\alpha,\beta)\,dt \\ &= \int_{x}^{\infty} f(t;\alpha,\beta)dt \qquad = 1-\int_{-\infty}^{x} f(t;\alpha,\beta)dt \\ &= 1-F(x,\alpha,\beta) \end{aligned}$$

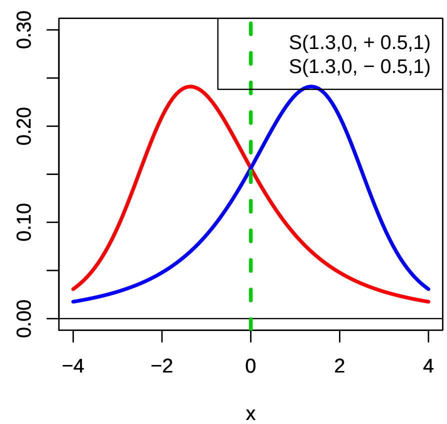

FIGURE I.9 – Propriété de réflexion.

Une loi α-stable est développable en série entière :

Toujours, si on considère $f(x;\alpha,\beta)$ définie ci-dessus, on montre que les lois stables dont les densités possédant des dérivées uniformément bornées de tout ordre, admettent aussi des représentations par des séries entières convergentes (voir [2] et [16]) :

1. Si $0<\alpha<1$ et $x>0$, alors :

$$f(x;\alpha,\beta)=\frac{1}{\pi}\sum_{n=1}^{\infty}(-1)^{n+1}\sin\left(\frac{n\pi}{2}\alpha(1+\beta)\right)\frac{\Gamma(n\alpha+1)}{n!}x^{-n\alpha-1} \qquad \text{(I.11)}$$

2. Si $1 < \alpha < 2$ et $x > 0$, alors :

$$f(x;\alpha,\beta) = \frac{1}{\pi}\sum_{n=0}^{\infty}(-1)^{n+1}\sin\left(\frac{n\pi}{2}\left(\alpha+(2-\alpha)\beta\right)\right)\frac{\Gamma(n\alpha+1)}{n!}x^{-n\alpha-1} \tag{I.12}$$

3. Si $\alpha = 1$ et $x > 0$, alors :

$$f\left(x+\frac{2\beta}{\pi}\ln x, 1, \beta\right) = \frac{1}{\pi}\sum_{n=0}^{\infty}\frac{c_n}{n!}x^{-n-1},$$

où

$$c_n = Im\int_0^{\infty} e^{-t}t^n\left[(1+\beta)i - \frac{2\beta}{\pi}\ln t\right]^n dt$$

Commentaires sur ces résultats :

— Il faut mentionner ici que la série (I.11) diverge pour les valeurs de $\alpha > 1$, de même pour la série (I.12) pour les valeur de $\alpha < 1$.

— La démonstration de ce résultat passe par la théorie des variables complexes et des fonctions analytiques developpables en séries entières, appliqués à certaines expressions intégrales qui sont fonctions de la fonction de répartition de la loi stable (voir Zolotarev dans [17])

I.6.4 Propriétés arithmétiques

1. **Standardisation d'une loi stable** :

Proposition I.6.4 *Pour $\alpha \neq 1$ nous avons l'équivalence qui suit :*
La v.a.r X suit une loi $S(\alpha,\sigma,\beta,m) \Leftrightarrow Y = \dfrac{X-m}{\sigma}$ suit une loi $S(\alpha,1,\beta,0)$.

Preuve :

<u>Condition nécéssaire</u> :

Si on pose : $Y = aX + b$, on sait que :

$$\varphi_{aX+b} = e^{i\theta b}\varphi_X(a\theta)$$

et en prenant : $a = \dfrac{1}{\sigma}$ et $b = -\dfrac{m}{\sigma}$, on obtient :

$$\varphi_Y(t) = \exp\left(-\frac{im\theta}{\sigma}\right).\varphi_X\left(\frac{\theta}{\sigma}\right)$$

il s'en suit que :

$$\varphi_Y(\theta) = \exp\left(-\frac{im\theta}{\sigma}\right)\exp\left\{\frac{im\theta}{\sigma} - |\theta|^\alpha\left[1 - i\beta.sign\left(\frac{\theta}{\sigma}\right).w(\theta,\alpha)\right]\right\}$$

mais, $\text{sign}\left(\frac{\theta}{\sigma}\right) = \text{sign}(\theta)$ puisque $\sigma > 0$, soit alors,

$$\varphi_Y(\theta) = \exp\{-|\theta|^\alpha[1 - i\beta.sign(\theta).w(\theta,\alpha)]\}$$

qui est bien la forme de la fonction caractéristique de la loi $S(\alpha, 1, \beta, 0)$.

Pour la condition suffisante, même raisonnement en prenant $a = \sigma$ et $b = m$.

N.B. Pour la simulation, il suffit de générer des lois $S(\alpha, 1, \beta, 0)$ et par changement de variables, on obtient les lois $S(\alpha, \sigma, \beta, m)$.

2. **L'ajout d'une constante à une loi stable :**

Proposition I.6.5 *Soit une loi stable $X \sim S(\alpha, \sigma, \beta, m)$ et C une constante réelle, alors : $X + C \sim S(\alpha, \sigma, \beta, m + C)$.*

Preuve :

En effet :

$$\begin{aligned}
\varphi_{(X+C)}(\theta) &= \exp\left(iC\theta\right)\ \varphi_X(\theta) \\
&= \exp\left(iC\theta\right)\ \exp\left\{im\theta - \sigma^\alpha|\theta|^\alpha\left(1 + i\beta\frac{\theta}{|\theta|}w(\theta,\alpha)\right)\right\} \\
&= \exp\left\{i(m + C)\theta - \sigma^\alpha|\theta|^\alpha\left(1 + i\beta\frac{\theta}{|\theta|}w(\theta,\alpha)\right)\right\}
\end{aligned}$$

qui est bien une v.a.r. $S(\alpha, \sigma, \beta, m + C)$.

3. **Somme de deux lois stables :**

Proposition I.6.6 *Considérons deux variables aléatoires α-stables indépendantes $X_1 \sim S(\alpha, \sigma_1, \beta_1, m_1)$ et $X_2 \sim S(\alpha, \sigma_2, \beta_2, m_2)$, alors :*

$$X_1 + X_2 \sim S\left(\alpha, (\sigma_1^\alpha + \sigma_2^\alpha)^{1/\alpha}, \frac{\beta_1\sigma_1^\alpha + \beta_2\sigma_2^\alpha}{\sigma_1^\alpha + \sigma_2^\alpha}, m_1 + m_2\right).$$

Preuve :

Nous avons :

$$\begin{aligned}
\varphi_{X_1+X_2}(\theta) &= \varphi_{X_1}(\theta)\varphi_{X_2}(\theta) \\
&= \exp\left\{im_1\theta - \sigma_1^\alpha|\theta|^\alpha\left(1 + i\beta_1\frac{\theta}{|\theta|}w(\theta,\alpha)\right)\right\}\exp\left\{im_2\theta - \sigma_2^\alpha|\theta|^\alpha\left(1 + i\beta_2\frac{\theta}{|\theta|}w(\theta,\alpha)\right)\right\} \\
&= \exp\left\{i(m_1+m_2)\theta - (\sigma_1^\alpha + \sigma_2^\alpha)|\theta|^\alpha - i(\beta_1\sigma_1^\alpha + \beta_2\sigma_2^\alpha)|\theta|^\alpha\frac{\theta}{|\theta|}w(\theta,\alpha)\right\} \\
&= \exp\left\{i(m_1+m_2)\theta - (\sigma_1^\alpha + \sigma_2^\alpha)|\theta|^\alpha\left[1 + i\left(\frac{\beta_1\sigma_1^\alpha + \beta_2\sigma_2^\alpha}{\sigma_1^\alpha + \sigma_2^\alpha}\right)\frac{\theta}{|\theta|}w(\theta,\alpha)\right]\right\}.
\end{aligned}$$

N.B. $w(\theta, \alpha)$ étant la fonction définie dans la formule (I.6).

I.6.5 Simulation des variables α-stables

Le procédé de Chambers et al. dans [1] permet de générer des lois stables standardisées. Il est basé sur le résultat suivant :

Proposition I.6.7 *Soit deux variables aléatoires indépendantes, Ψ suivant une loi uniforme sur l'intervalle $[-\frac{\pi}{2}, \frac{\pi}{2}]$ et W suivant une loi exponentielle de moyenne égale à 1 et si on pose $\psi_0 = \arctan(\beta\tan(\pi\alpha/2))/\alpha$. Alors, pour tous α et β tels que $0 < \alpha \leq 2$, $-1 < \beta < +1$, la variable Z définie par :*

$$Z = \begin{cases} \dfrac{\sin\alpha(\psi_0 + \Psi)}{(\cos\alpha\psi_0\cos\Psi)^{1/\alpha}}\left[\dfrac{\cos(\alpha\psi_0 + (\alpha-1)\Psi)}{W}\right]^{(1-\alpha)/\alpha} & , \quad \alpha \neq 1 \\[2ex] \dfrac{2}{\pi}\left[(\pi/2 + \beta\Psi)\tan\Psi - \beta\ln\left(\dfrac{\pi/2\ W\cos\Psi}{\pi/2 + \beta\Psi}\right)\right] & , \quad \alpha = 1 \end{cases} \tag{I.13}$$

possède une distribution stable $S(\alpha, 1, \beta, 0)$

Preuve : voir preuve dans Noland [10].

A partir de deux lois U_1 et U_2 uniformes et indépendantes sur l'intervalle $[0,1]$, on obtient facilement les variables Ψ et W en posant :

$$\left\{ \begin{array}{rcl} \Psi & = & \pi(U_1 - \frac{1}{2}) \\ \\ W & = & -\ln(1 - U_2) \end{array} \right. \tag{I.14}$$

Il suffit alors de générer aléatoirement deux échantillons de ces deux variables pour obtenir un échantillon d'une distribution $S(\alpha, 1, \beta, 0)$.

Pour ce qui est d'une distribution $S(\alpha, \sigma, \beta, m)$ de position m et d'échelle σ, on utilisera tout simplement le résultat de la proposition I.6.4.

Cas particuliers :

1. Lorsque $\alpha = 2$: $\psi_0 = 0$ et $Z = 2\sqrt{W} \sin \Psi$ qui est la représentation bien connue de Box-Müler permettant de générer une loi de Gauss $N(0, 2)$ (voir [15]).

2. Lorsque $\alpha = 1$, $\beta = 0$: on retrouve la formule bien connue $Z = \tan \psi$ qui permet de simuler une loi de Cauchy standard.

3. Lorsque $\alpha = \frac{1}{2}$, $\beta = 1$, on trouve : $Z = \dfrac{1}{2W \cos^2 \Psi}$.

4. Lorsque $\alpha \neq 1$, $\beta = 0$: $Z = \dfrac{\sin \alpha \Psi}{(\cos \Psi)^{1/\alpha}} \left[\dfrac{\cos(\alpha - 1)\Psi}{W} \right]^{(1-\alpha)/\alpha}$ qui est le cas d'une loi $S\alpha S(\sigma)$ pour $\sigma = 1$.

Chapitre II

Estimation des paramètres d'une loi stable

Il existe toute une panoplie de méthodes pour estimer les paramètres d'une loi stable $S_\alpha(\sigma, \beta, m)$ à partir d'un échantillon i.i.d. de v.a.r.'s $X_1, X_2, \ldots, X_n$. Dans ce qui suit on présentera quatre méthodes avec leurs algorithmes si possible ; Elles donnent des estimateurs ayant de bonnes propriétés telles que la consistance et la normalité asymptotique.

II.1 Méthode des moments

C'est une méthode qui a été proposé par Press(1972) dans [13] et [14]. Elle repose sur la fonction caractéristique (f.c.). Dans un premier temps, on prend le logarithme du module de la f.c. ; on obtient alors une relation qui va permettre d'avoir des estimateurs pour les paramètres α et σ de la loi cherchée et ceci en remplaçant la f.c. théorique par la f.c. empirique. Dans un deuxième temps, en passant cette fois-ci à l'argument de la f.c., on obtient des estimateurs pour β et m, soit :

$$|\varphi(\theta)| = \exp\left(-\sigma^\alpha|\theta|^\alpha\right). \tag{II.1}$$

Alors

$$-\log|\varphi(\theta)| = \sigma^\alpha|\theta|^\alpha.$$

Dans le cas $\alpha \neq 1,$ on prend les valeurs non nulles de $\theta,\ \theta_1 \neq \theta_2$

$$-\log|\varphi(\theta_k)| = \sigma^\alpha|\theta_k|^\alpha, \qquad Pour \quad k = 1, 2, \tag{II.2}$$

la résolution de ces deux équations simultanément pour α et σ et en remplaçant $\varphi(\theta)$ par $\widehat{\varphi}(\theta)$

donne :

$$\widehat{\alpha} = \frac{\log\left[\dfrac{\log|\widehat{\varphi}(\theta_1)|}{\log|\widehat{\varphi}(\theta_2)|}\right]}{\log\left|\dfrac{\theta_1}{\theta_2}\right|}, \tag{II.3}$$

et

$$\log\widehat{\sigma} = \frac{\log|\theta_1|\log\Big(-\log|\widehat{\varphi}(\theta_2)|\Big) - \log|\theta_2|\log\Big(-\log|\widehat{\varphi}(\theta_1)|\Big)}{\log\left[\dfrac{\log|\widehat{\varphi}(\theta_1)|}{\log|\widehat{\varphi}(\theta_2)|}\right]}. \tag{II.4}$$

Pour l'estimation de β et m, on va définir $v(\theta) = Im\left(\log(\varphi(\theta))\right)$, alors :

$$v(\theta) = m\theta + \sigma^{\alpha}|\theta|^{\alpha}\beta sign(\theta)\tan\frac{\alpha\pi}{2}. \tag{II.5}$$

On prend deux valeurs non nulles de θ, $\theta_3 \neq \theta_4$ il vient que :

$$\frac{v(\theta_k)}{\theta_k} = m + \beta\left[\sigma^{\alpha}|\theta_k|^{\alpha-1}\tan(\alpha\pi/2)\right], \qquad pour \quad k = 3, 4. \tag{II.6}$$

comme :

$$\widehat{\varphi}(\theta) = \left(\frac{1}{n}\sum_{i=1}^{n}\cos(\theta x_i)\right) + i\left(\frac{1}{n}\sum_{i=1}^{n}\sin(\theta x_i)\right), \tag{II.7}$$

alors,

$$\tan\widehat{v}(\theta) = \frac{\displaystyle\sum_{i=1}^{n}\cos(\theta x_i)}{\displaystyle\sum_{i=1}^{n}\sin(\theta x_i)}. \tag{II.8}$$

En remplaçant $v(\theta)$ dans (II.6) par la valeur $\widehat{v}(\theta)$ à partir de (II.8) et les paramètres par leurs estimateurs puis, en résolvant les deux équations linéaires simultanément pour β et m ce qui donne les estimateurs :

$$\widehat{\beta} = \frac{\dfrac{\widehat{v}(\theta_4)}{\theta_4} - \dfrac{\widehat{v}(\theta_3)}{\theta_3}}{\left[|\theta_4|^{\widehat{\alpha}-1} - |\theta_3|^{\widehat{\alpha}-1}\right] \widehat{\sigma}^{\widehat{\alpha}} \ \tan(\widehat{\alpha}\pi/2)}, \tag{II.9}$$

et

$$\widehat{m} = \frac{|\theta_4|^{\widehat{\alpha}-1}\left[\dfrac{\widehat{v}(\theta_3)}{\theta_3}\right] - |\theta_3|^{\widehat{\alpha}-1}\left[\dfrac{\widehat{v}(\theta_4)}{\theta_4}\right]}{|\theta_4|^{\widehat{\alpha}-1} - |\theta_3|^{\widehat{\alpha}-1}}. \tag{II.10}$$

Le cas $\alpha = 1$ et plus simple, en suivant le même raisonnement, on arrive aux résultats suivants :

$$\widehat{\sigma} = -\frac{\log|\widehat{\varphi}(\theta_1)|}{\theta_1} \tag{II.11}$$

$$\widehat{\beta} = \frac{\dfrac{\widehat{v}(\theta_3)}{\theta_3} - \dfrac{\widehat{v}(\theta_4)}{\theta_4}}{(2/\pi)\widehat{\sigma}\log\left|\dfrac{\theta_4}{\theta_3}\right|} \tag{II.12}$$

et

$$\widehat{m} = \frac{\log|\theta_4|\left[\dfrac{\widehat{v}(\theta_3)}{\theta_3}\right] - \log|\theta_3|\left[\dfrac{\widehat{v}(\theta_4)}{\theta_4}\right]}{\log|\theta_4| - \log|\theta_3|}. \tag{II.13}$$

Mise en oeuvre de l'algorithme :

1. Choisir deux points θ_1 et θ_2 et calculer $\widehat{\varphi}(\theta_1)$ et $\widehat{\varphi}(\theta_2)$.
2. Calculer $\widehat{\alpha}$ et $\widehat{\sigma}$ d'après les expressions (II.3) et (II.4).
3. Choisir deux points θ_3 et θ_4 et calculer $\widehat{\beta}$ et $\widehat{m}$ des expressions (II.9) et (II.10).

Cette méthode est très facile à implémenter et elle est très efficace en temps calcul.

II.2 Estimation du maximum de vraisemblance

Etant donné un échantillon $X_1, X_2, \ldots, X_n$ de variables aléatoires i.i.d.'s avec $X_i \sim S(\alpha, \sigma, \beta, m)$. L'estimation par maximum de vraisemblance pour les distributions α-stables est la même que de celle des autres lois. Il s'agit de maximiser la fonction de vraisemblance :

$$l(\alpha, \sigma, \beta, m) = \sum_{i=1}^{n} \log f(x_i, \alpha, \sigma, \beta, m).$$

dans l'espace des paramètres est $\Omega =]0,2]\times]0,+\infty[\times[-1,1]\times R$.

Elle permet d'obtenir des estimateurs consistants et, s'il sont uniques, il sont asymptotiquement sans biais, efficaces et normaux. Seulement qu'elle reste difficile à mettre en oeuvre car les lois stables n'admettent pas de formules simples et fermes pour leurs densités afin d'écrire la vraissemblance .

Il y a eu des tentatives de la part de plusieurs statisticiens qui ont pu développer des algorithmes efficaces pour approcher la densité et la fonction de répartition pour des valeurs $\alpha > 0.85$. Noland (1996) dans [9] obtînt des formules basées comme celles de Zolotarev (1986, [17]), sur la représentation intégrale d'une loi stable qui permettent de calculer de façon précise et les densités et les répartitions ainsi que les quantiles dans tout l'espace paramétrique. Ceci implique que le coût calcul par la méthode de vraisemblance est significatif. Mais, ça n'empêche pas d'après Ojeda [11] d'affirmer que les estimateurs obtenus soient les plus précis.

II.3 Méthode basée sur la régression

Koutrouvelis dans [6] propose une autre méthode de type régression qui permet d'estimer les paramètres d'une loi stable à partir de l'expression de la fonction caractéristique et d'en déduire un ensemble d'équations comme suit :

$$\log(-\log|\varphi(\theta)|^2) = \log(2\sigma^\alpha) + \alpha\log|\theta|. \tag{II.14}$$

Pour $\alpha \neq 1$ les parties réelle et imaginaire de $\varphi(\theta)$ sont données par :

$$Re[\varphi(\theta)] = \exp\Big(-|\sigma\theta|^\alpha\Big)\cos\left[m\theta + |\sigma\theta|^\alpha\beta sign(\theta)\tan(\pi\alpha/2)\right],$$

et

$$Im[\varphi(\theta)] = \exp\Big(-|\sigma\theta|^\alpha\Big)\sin\left[m\theta + |\sigma\theta|^\alpha\beta sign(\theta)\tan(\pi\alpha/2)\right].$$

D'après les deux dernières équations, on peut déduire sans considération sur la valeur principale de la fonction arctan :

$$\arctan\left(\frac{Im[\varphi(\theta)]}{Re[\varphi(\theta)]}\right) = m\theta + \beta\sigma^\alpha\tan(\pi\alpha/2)sign(\theta)|\theta|^\alpha. \tag{II.15}$$

L'équation (II.14) ne dépend que de α et σ et suggère que l'on estime ces paramètres par régression de $y = \log(-\log|\varphi(\theta)|^2)$ sur $\omega = \log|\theta|$ dans le modèle :

$$y_k = \mu + \alpha\omega_k + \varepsilon_k \ , \ \text{k=1,2,....,n} \tag{II.16}$$

où $\{\theta_k; k = 1, 2, ..., n\}$ est un ensemble approprié de nombres réels, et ε_k désigne un terme d'erreur. Koutrouvelis (1980) propose d'utiliser $\theta_k = \pi k/25, \ k = 1, 2, \ldots, n$ sans perte significative d'efficacité, nous avons simplifié sa méthode de choisir K (Voir tableau (II.1)).

l'indice α	K
1.5 - 2.0	10
0.6 - 1.5	20
0.4 - 0.6	60
inférieur à 0.4	120

TABLE II.1 – Valeurs optimales de K pour différentes valeurs de α.

Une fois $\widehat{\alpha}$ et $\widehat{\sigma}$ obtenus, alors les estimations de β et m peuvent s'obtenir à partir de (II.15). Soit :

$$g_n(u) = \arctan\left(Im[\widehat{\varphi}(t)]/Re[\widehat{\varphi}(t)]\right)$$

$\widehat{\varphi}$ étant la f.c. empirique définie dans (II.7). On peut alors estimer β et m en régressant $z = g_n(u) + \pi k_n(u)$ à u et $sign(u)|u|^\alpha$ dans le modèle suivant :

$$z_r = mu_r + \beta\sigma^\alpha \tan(\pi\alpha/2) sign(u_r)|u_r|^\alpha + \eta_r, \ \text{r=1,2,...,L} \tag{II.17}$$

où $\{u_r, \ r = 1, 2, ..., L\}$ est un ensemble approprié de nombres réels et η_r désigne un terme d'erreur et le nombre $k_n(u)$ représente les branches possibles de la fonction arctan. Koutrouvelis [6] propose d'utiliser $u_r = \pi r/50, \ r = 1, 2, ..., L$ avec L entre 9 et 70 pour les différentes estimations du paramètre α et les tailles d'échantillons.

Mise en oeuvre de l'algorithme :

1. Fixer un K approprié et pour les points $\theta_k = \frac{\pi k}{25}, \theta = 1, 2, \ldots, K$ calculer la fonction caractéristique empirique $\widehat{\varphi}(\theta)$.
2. Calculer $y_k = \log(-\log|\varphi(\theta_k)|^2)$.
3. Ajuster la régression linéaire $y_k = \mu + \alpha\omega_k + \epsilon_k$, pour obtenir des estimateurs $\widehat{\alpha}$ et $\widehat{\mu}$.

4. Obtenir l'estimateur $\widehat{\sigma}$ à partir de $\mu = \log(2\widehat{\sigma})^{\widehat{\alpha}}$.
5. Calculer, pour un L approprié , $z_r(u)$ pour $u_r = \frac{\pi r}{50}$, $r = 1, 2, \ldots, L$.
6. Ajuster la régression : $z_r = mu_r + \beta\widehat{\sigma}^{\alpha} \tan\frac{\pi\widehat{\alpha}}{2} sign(u_r)|u_r|^{\widehat{\alpha}} + \eta_r$, $r = 1, 2, \ldots, L$, pour obtenir les estimateurs de m et β.

Dans le cas symétrique, il y a seulement deux paramètres à fixer α et σ, pour le paramètre α l'estimateur est obtenu par la première régression linéaire :

$$\widehat{\alpha} = \frac{K\sum_{k=1}^{K}\omega_k\widehat{y}_k - \sum_{k=1}^{K}\omega_k\sum_{k=1}^{K}\widehat{y}_k}{k\sum_{k=1}^{K}\omega_k^2 - \left(\sum_{k=1}^{K}\omega_k\right)^2}.$$

Si on prend $\sum_{k=1}^{K}\omega_k = 0$ l'estimation précédente est égale à :

$$\widehat{\alpha} = \frac{k\sum_{k=1}^{K}\omega_k\widehat{y}_k}{k\sum_{k=1}^{K}\omega_k^2}.$$

et pour l'estimateur de σ, il est obtenu par la deuxième régression linéaire :

$$\widehat{\sigma} = \exp\left\{\frac{1}{\widehat{\alpha}}\log(e^{\widehat{\mu}}/2)\right\},$$

où

$$\widehat{\mu} = \frac{1}{K}\sum_{k=1}^{k}\widehat{y}_k.$$

II.4 Méthode de McCulloch

McCulloch dans [8] a fourni des estimateurs consistants pour les quatre paramètres avec la restriction $\alpha > 0.6$. C'est une généralisation de l'approche basée sur les quantiles empiriques, initiée par Fama et Roll [3]. Ces derniers ont fourni des estimateurs consistants d'une distribution stable symétrique SαS(σ) mais pour des valeurs allant de 1 jusqu'à 2 du paramètre α et pour $\sigma > 0$; Alors que McCulloch a obtenu des estimateurs convergents et asymptotiquement de lois normales pour l'ensemble des quatre paramètres à la fois d'une distribution stable $S(\alpha, \sigma, \beta, m)$ et pour un indice de stabilité α allant cette fois-ci de 0.6 jusqu'à 2.

L'idée commence par l'évaluation de deux paramètres seulement α et β à l'aide de simples indices, fonctions de cinq quantiles prédéterminés, calculés à partir de l'échantillon de données considéré $X_1, X_2, \ldots, X_n$, soit :

$$v_\alpha = \frac{x_{0.95} - x_{0.05}}{x_{0.75} - x_{0.25}} \quad , \quad v_\beta = \frac{x_{0.95} + x_{0.05} - 2x_{0.50}}{x_{0.95} - x_{0.05}}$$

Notons qu'ici x_p désigne le quantile d'ordre p de la distribution.

Ces indices ne dépendent ni de m ni de σ ; De plus, ils sont respectivement des fonctions décroissante et croissante de α et de β. Ce qui permet d'inverser ses dernières pour obtenir α et β comme fonctions de ces deux indices. Soit :

$$\alpha = \varphi_1(v_\alpha, v_\beta) \text{ et } \beta = \varphi_2(v_\alpha, v_\beta)$$

Considérons maintenant leurs estimateurs consistants (voir [8]) :

$$\widehat{v}_\alpha = \frac{\widehat{x}_{0.95} - \widehat{x}_{0.05}}{\widehat{x}_{0.75} - \widehat{x}_{0.25}} \quad , \quad \widehat{v}_\beta = \frac{\widehat{x}_{0.95} + \widehat{x}_{0.05} - 2\widehat{x}_{0.50}}{\widehat{x}_{0.95} - \widehat{x}_{0.05}}$$

lesquels vont permettre d'estimer les paramètres α et β par :

$$\widehat{\alpha} = \varphi_1(\widehat{v}_\alpha, \widehat{v}_\beta) \quad \text{and} \quad \widehat{\beta} = \varphi_2(\widehat{v}_\alpha, \widehat{v}_\beta)$$

Dans un deuxième temps, on estimera le reste des paramètres en utilisant les indices suivants :

$$v_\sigma = \frac{x_{0.75} - x_{0.25}}{\sigma} := \varphi_3(\alpha, \beta) \quad , \quad v_m = \frac{m - x_{0.5}}{\sigma} := \varphi_4(\alpha, \beta)$$

Lesquels donnent à leur tour les estimateurs consistants (voir [8]) :

$$\widehat{\sigma} = \frac{\widehat{x}_{0.75} - \widehat{x}_{0.25}}{\varphi_3(\widehat{\alpha}, \widehat{\beta})} \quad , \quad \widehat{m} = \widehat{\sigma}\ \varphi_4(\widehat{\alpha}, \widehat{\beta}) + \widehat{x}_{0.5}$$

Comme l'estimateur $\widehat{x}_p$ est consistant et asymptotiquement normal pour x_p et que les fonctions φ_i sont continues alors, les estimateurs des paramètres de la loi stable sont aussi consistants et asymptotiquement de lois normales.

II.5 Simulation et comparaison entre ces méthodes

Dans cette section nous avons généré des échantillons à partir de lois stables dont les paramètres sont connus au départ ; Puis, à partir de ces échantillons, on a estimé les paramètres de la loi stable en question par les différentes méthodes présentées dans ce chapitre et ceci dans le but de montrer le degré d'efficacité de chaque méthode et en même temps estimer le temps d'exécution nécessaire de chacune d'elles pour arriver au résultat espéré. Nous avons fait varier deux paramètres seulement à savoir l'indice de stabilité et le paramètre d'asymétrie car ce sont les seuls indicateurs de forme pour la densité, les deux autres agissent uniquement sur la position et l'échelle.

Dans une première étape, nous avons simulé des lois stables pour deux valeurs différentes de l'indice de stabilité α tout en préservant les valeurs des autres paramètres.

Dans une deuxième étape, nous avons fait varier la valeur du paramètre d'asymétrie sans changer le reste des paramètre et ceci en vue de voir s'il y a influence de β sur les calculs en termes de temps calcul et d'efficacité des estimations obtenues.

Les tableaux qui vont suivre résument les résultats obtenus et le temps écoulé de chaque opération.

	$\alpha = 1.2$	$\sigma = 1$	$\beta = 0$	$m = 2$
Méthode	$\hat{\alpha}$	$\hat{\sigma}$	$\hat{\beta}$	$\hat{m}$
McCulloch	1.24227518	0.98447200	0.04422733	1.98693397
Erreur absolue	0.04	0.015	0.04	0.014
Temps d'exécution : 0 s				
Koutrouvilis	1.21136190	0.98457752	-0.03142572	1.84461124
Erreur absolue	0.01	0.016	0.03	0.156
Temps d'exécution : 12 s				
EMV	1.19959231	0.98050373	-0.00472542	2.02794997
Erreur absolue	0.001	0.02	0.004	0.027
Temps d'exécution : 5 min, 21 s s				
Moments	1.209102864	0.982466621	0.007542575	2.014959320
Erreur absolue	0.009	0.018	0.007	0.014
Temps d'exécution : 8 s				

TABLE II.2 – $S(1.2, 1, 0, 2)$

	$\alpha = 0.8$	$\sigma = 1$	$\beta = 0$	$m = 2$
Méthode	$\hat{\alpha}$	$\hat{\sigma}$	$\hat{\beta}$	$\hat{m}$
McCulloch	0.76356528	0.91039457	0.07248489	1.98666301
Erreur absolue	0.037	0.09	0.07	0.02
Temps d'exécution : 0 s				
Koutrouvilis	0.84453260	0.91687471	-0.07643853	1.94356877
Erreur absolue	0.04	0.09	0.07	0.06
Temps d'exécution : 13 s				
EMV	0.86675791	1.04441918	-0.06763236	1.94880380
Erreur absolue	0.06	0.04	0.06	0.06
Temps d'exécution : 5 min, 3 s				
Moments	0.7809351	1.0597598	-0.0500750	2.0598158
Erreur absolue	0.02	0.05	0.05	0.05
Temps d'exécution : 3 s				

TABLE II.3 – $S(0.6, 1, 0, 2)$

	$\alpha = 1.2$	$\sigma = 1$	$\beta = 1/3$	$m = 2$
Méthode	$\hat{\alpha}$	$\hat{\sigma}$	$\hat{\beta}$	$\hat{m}$
McCulloch	1.2391667	0.9526848	0.3264515	1.9884297
Erreur absolue	0.03	0.05	0.01	0.02
Temps d'exécution : 0 s				
Koutrouvilis	1.2081532	1.0215572	0.3304358	1.9261666
Erreur absolue	0.008	0.02	0.03	0.08
Temps d'exécution : 0 s				
EMV	1.1503275	0.9985308	0.3731558	1.9722416
Erreur absolue	0.05	0.01	0.03	0.03
Temps d'exécution : 5 min ,50 s				
Moments	1.177666	1.042920	0.341717	1.990277
Erreur absolue	0.03	0.04	0.03	0.01
Temps d'exécution : 3 s				

TABLE II.4 – $S(1.2, 1, 1/3, 2)$

	$\alpha = 1.2$	$\sigma = 1$	$\beta = -1/2$	$m = 2$
Méthode	$\hat{\alpha}$	$\hat{\sigma}$	$\hat{\beta}$	$\hat{m}$
McCulloch	1.1963392	1.0425808	-0.4149694	2.0639317
Erreur absolue	0.01	0.04	0.09	0.06
Temps d'exécution : 0 s				
Koutrouvilis	1.1266435	0.9897922	-0.4529803	1.9494465
Erreur absolue	0.08	0.02	0.05	0.06
Temps d'exécution : 0 s				
EMV	1.1998022	1.0672422	-0.4839019	1.9255033
Erreur absolue	0.01	0.06	0.02	0.08
Temps d'exécution : 4 min , 44 s				
Moments	1.2985885	1.0845143	-0.3720513	1.9893530
Erreur absolue	0.09	0.08	0.13	0.02
Temps d'exécution : 3 s				

TABLE II.5 – $S(1.2, 1, -1/2, 2)$

II.6 Commentaires

Les calculs ont été effectués sur le logiciel libre R version 4.4.0 avec un AMD A4 9125 RADEON R 3.4 CORES 2.30 GHz. Pour chacune des méthodes nous avons généré un échantillon aléatoire de taille 1000 à partir d'une loi stable $S(\alpha, \sigma, \beta, m)$ qui représente la loi théorique.

Dans chaque tableau, nous avons indiqué en haut les valeurs des paramètres de la loi théorique. Pour chacune des méthodes indiquées, nous avons mis les valeurs estimées des paramètres au-dessous de chaque valeur théorique de manière correspondante ; Puis, nous avons indiqué les erreurs absolues des estimations relatives à chaque paramètre et enfin, juste en dessous, nous avons enregistré la durée qu'aurait pris le RADEON pour donner les valeurs $(\widehat{\alpha}, \widehat{\sigma}, \widehat{\beta}, \widehat{m})$ du vecteur des paramètres théoriques $(\alpha, \sigma, \beta, m)$.

Dans le premier tableau (II.2), les estimateurs du MV sont les meilleurs pour les paramètres de forme α et β avec des erreurs les plus petites parmi les autres méthodes mais avec un temps calcul le plus lent (ici plus de 5 mn). Cependant, pour l'échelle σ et pour la position m (ici c'est une moyenne), c'est celle de McCulloch qui est la bonne pour un temps presque instantané.

Dans le deuxième tableau (II.3), c'est la méthode des moments qui donne les meilleurs résultats pour presque tous les paramètres avec un temps d'exécution très court (ici 3 s).

Dans le troisième tableau (II.4), l'estimateur de régression de Koutrouvelis donne la meilleure estimation pour α avec une erreur de l'ordre de 10^{-3} et un temps d'exécution presque instantané ; Pour le reste des paramètres, les quatre méthodes donnent des résultats du même ordre de grandeurs de l'ordre de 10^{-2} avec des écarts minimes.

Dans le quatrième tableau (II.5), les méthodes de McCulloch et du MV sont les meilleures avec une même erreur absolue ici 0.01 mais, c'est celle des moments qui donne la bonne estimation pour l'asymétrie β avec une erreur de 0.13 et pour la position m, elle donne la plus petite erreur absolue de 0.02 pour un temps d'exécution de $3s$.

Conclusions à tirer :

Ce qu'on peut tirer à partir de ces résultats et pour une taille d'échantillon moyenne de l'ordre de millier d'unités ce qui suit :

— La méthode de McCulloch est très efficace en temps calcul, c'est la plus rapide à tous les coups, seulement que ses estimations ne sont pas toujours les meilleures mais, elles ne sont pas mauvaises dans leur ensemble. Elle sert aussi à faire démarrer, comme valeur initiale, d'autres méthodes telle que la méthode itérative de Koutrouvilis.

— La méthode de Koutrouvelis donne des résultats satisfaisants avec des erreurs absolues de l'ordre de 10^{-2} à quelques exceptions près. Dans [6], Koutrouvelis a proposé une variante de sa propre méthode dont les résultats sont bien meilleurs pour une plus grande région de l'espace paramétrique. C'est la méthode de régression itérative qui n'a pas été décrite dans ce travail.

— La méthode des moments donne généralement de bonnes estimations pour l'ensemble des paramètres avec un temps calcul assez court ne dépassant pas les dix secondes. Elle est très facile à implémenter. Cependant, elle est imprécise si l'échantillon n'est pas normalisé.

— La méthode du maximum de vraissemblance donne des estimations très satisfaisantes et elle est généralement la plus précise surtout lorsque l'échantillon est de taille importante. Cependant, elle reste la méthode la plus lente en temps d'exécution.

Conclusion

Les lois stables constituent une classe très riche de distributions de probabilité, très intéressantes pour la modélisation de phénomènes financiers qui présentent de grandes variabilités.
Ce mémoire comporte deux chapitres. Dans le premier chapitre, nous avons présenté la définition d'une loi stable et sa paramétrisation standard via sa fonction caractéristique. Nous avons donné l'interprétation pratique de chaque paramètre de cette distribution et ses diverses propriétés pour la densité, les propriétés arithmétiques, la stabilité, le calcul des moments etc.

Dans un deuxième chapitre nous avons présenté quelques approches pour estimer les paramètres caractérisant les lois stables ; En réalité, il existe plusieures méthodes pour les estimer ; Nous avons traîté, parmi elles, quatre méthodes qui donnent de bonnes estimations. Nous avons fait une comparaison entres ces différentes méthodes à l'aide de simulations tout en montrant l'efficacité en temps de calcul de chacune d'elles.

Bibliographie

[1] Chambers J.M., Mallows C.L. et Stuck B.W. (1976). A method for simulating stable random variables. Journal of the American Statistical Association, vol.71, n° 354 pages 340-344.

[2] Deheuvels, P., Haeusler, E., Mason, D. (1988). Almost sure convergence of the Hill estimator. Math. Proc. Camb. Philos. Soc. 104, 371-381.

[3] Fama. E, and Roll. R, (1973) Paramèters Estimates for Symetric Stable Distributions. Journal of the American Statistical Association, **66**, 331-339

[4] Feller. W, (1966). An Introduction To Probability Theory and its Applications, Vol 2. 1st edn, Wisley, New York.

[5] Gnedenko & Kolmogorov (1954). Limit Distributions for Sums of Independent Random Variables.

[6] Koutrouvelis I.A. (1980). Regression-type Estimation of the Parameters of Stable Laws. Journal of the American Statistical Association, **75**, 918-928.

[7] Loève, M. (1963). Probability Theory (third ed. vol 1). Princeton : D. Van Nostrand Company. bitemc16McCulloch, J.H. (1986) Simple Consistent Estimators of Stable Distribution Parameters. Commun. Statist.-Simula., **15**(4) ,1109-1136.

[8] *McCulloch J.H. (1986). Simple consistent estimators of stable distribution parameters. Communications in Statistics. Simulation and Computation, 15, 1109-1136.*

[9] Nolan, J. P.(1997). Numerical Calculation of Stable Densities and Distribution Functions,Communications in Statistics-Stochastic Models 13 : 759-774.

[10] Nolan, J. P.(2005). Stable Distribution Models For Heavy Tailed Data.Math/Stat Department American University.

[11] Ojeda, D. (2001). Comparison of Stable Estimators. Ph.D. Thesis, Department of Mathematics and Statistics, American University.

[12] Petrov.V.(1987)première version. Limit Theorems of Probability Theory. Sequences Of Independent Random Variables.

[13] Press S.J. (1972). Estimation in Univariate and Multivariate Stable Distributions. J. Amer. Stat. Assoc., **67** , 842-846.

[14] Press S.J. (1972). Applied Multivariate Analysis. Holt, Rinehart and Winston, Inc.,New York.

[15] Samorodnitsky, G. and Taqqu, M. S.(1994). Stable Non-Gaussian Random Processes, Chapman & Hall.

[16] Shorokhod, A. V.(1954). An asymptotic Formula for Stable Distribution Laws, Dokl. Akad. Nayk USSR, 98, 731-734.

[17] Zolotarev, V. M. (1986). One-dimentional stable distribution , volume 65 In Translations of Mathematical Monographs. American Mathematical Society, Providence.

Résumé

Les lois stables de lévy sont une classe assez utilisée dans le domaine de l'Actuariat et la Finance. Elles sont décrites par quatre paramètres. Il existe un tas de méthodes pour les estimer. Ce mémoire développera quelques unes de ces approches.

Mots clés : Distribution stable ; Queue lourde ; Variance infinie ; Méthode d'estimation ; Simulation.

Abstract

Levy's stable laws are a class of laws widely used in actuarial and financial science. They are defined by four parameters. There are a number of methods for estimating them. This paper will develop some of these approaches.

Key words : Stable distribution, Heavy tail, Infinite variance, Estimation method, Simulation.

الملخص

قوانين ليفي المستقرة هي فئة من القوانين المستخدمة على نطاق واسع في البحوث الاكتوارية والمالية. وهي موصوفة بأربعة بارامترات. وهناك عدد من الطرق لتقديرها. ستعمل هذه الأطروحة على تطوير بعض هذه الأساليب.

الكلمات المفتاحية : التوزيع المستقر، الذيل الثقيل، التباين اللانهائي، طريقة التقدير، المحاكاة