

Approche des Valeurs Extrêmes dans la Modélisation des Séries Financières

Djamel Meraghni, Abdelhakim Necir

Laboratoire de Mathématiques Appliquées
Université Mohamed Khider, B.P. 145 R.P. 07000. Biksra, Algérie.
dmeraghni@yahoo.fr, necirabdelhakim@yahoo.fr

Résumé. Dans les années 60, les travaux de Mandelbrot sur les fluctuations boursières montrèrent que le modèle gaussien ne convenait pas pour décrire les rendements d'actifs. Mandelbrot (1963) puis Fama (1965) proposèrent la distribution Lévy-stable, dont les propriétés sont très proches de celles des distributions empiriques à queues lourdes, comme alternative pour modéliser les séries financières. Ce choix est justifié par au moins deux bonnes raisons ; la première est le théorème central limite généralisé selon lequel les lois stables sont les seules distributions limites possibles pour des sommes, convenablement normalisées et centrées, de variables aléatoires (v.a.) indépendantes et identiquement distribuées (i.i.d.) et la deuxième est le fait que les distributions stables peuvent être dissymétriques et permettent des queues épaisses de telle sorte qu'elles ajustent les distributions empiriques beaucoup mieux que ne le font les distributions gaussiennes. C'est ce que confirme l'exemple des rendements quotidiens de l'indice boursier CAC 40 que nous traitons à la fin de cet article dans lequel nous nous intéressons à l'estimation des paramètres caractérisant les lois Lévy-stables, ce qui constitue une étape essentielle dans le processus de modélisation des séries financières

1 Introduction

Les lois stables (dites aussi α -stables, Paréto-stables ou Lévy-stables) ont été introduites par Paul Lévy en 1924 (Lévy, 1925) lors de ses investigations sur le comportement des sommes de variables aléatoires indépendantes. Les propriétés de la distribution Lévy-stable sont une des raisons qui font d'elle le modèle que beaucoup de gens préfèrent pour décrire les observations dans les domaines où les données présentent une très grande variabilité tels que l'économie, la finance (voir Bassi *et al.*, 1998), les télécommunications (voir Bestavros *et al.*, 1998),... Les propriétés de ces lois ont été étudiées de manière approfondie par Samorodnitsky et Taqqu (1994) aussi bien dans le cas univarié que multivarié. Nous consacrons la deuxième section de cet article à un bref rappel de ces propriétés. Dans la troisième section, nous traitons l'estimation des paramètres stables par l'approche des valeurs extrêmes. La dernière section est réservée à l'étude des rendements journaliers de l'indice boursier CAC 40.

Enfin, notons que les simulations et applications numériques sont faites au moyen du logiciel de traitement statistique **R** (Ihaka et Gentleman, 1996).

2 Loi Lévy-stable

2.1 Caractérisation

Une v.a. X est dite de distribution stable si et seulement si, pour tout entier $n \geq 1$ et pour toute famille X_1, X_2, \dots, X_n de v.a. i.i.d. de même loi que X , il existe deux nombres réels $a_n > 0$ et b_n tels que

$$\frac{(X_1 + X_2 + \dots + X_n) - b_n}{a_n} \stackrel{d}{=} X,$$

où $\stackrel{d}{=}$ désigne l'égalité en distribution. Il est montré dans Feller (1971) qu'il existe une constante réelle $0 < \alpha \leq 2$ telle que $a_n = n^{1/\alpha}$.

Les variables de distribution Lévy-stable ont l'inconvénient de ne pas posséder (sauf dans trois cas) de formes explicites pour la densité de probabilité et la fonction de répartition. Une v.a. X de loi stable est typiquement décrite par sa fonction caractéristique φ (transformée inverse de Fourier de la densité de probabilité f) définie sur \mathbb{R} par

$$\varphi(t) := E(e^{itX}) = \int_{-\infty}^{\infty} e^{itx} f(x) dx,$$

et ayant plusieurs représentations selon les différentes paramétrisations des lois stables. La plus célèbre de ces représentations est donnée dans Samorodnitsky et Taquq (1994) par

$$\log(\varphi(t)) = \begin{cases} i\mu t - \sigma^\alpha |t|^\alpha \left\{ 1 - i\beta \operatorname{sign}(t) \tan \frac{\alpha\pi}{2} \right\}, & \alpha \neq 1, \\ i\mu t - \sigma |t| \left\{ 1 + i\beta \operatorname{sign}(t) \frac{2}{\pi} \log |t| \right\}, & \alpha = 1. \end{cases}$$

La distribution α -stable est donc caractérisée par quatre paramètres réels α , β , μ et σ .

Le paramètre α , appelé exposant caractéristique ou indice de stabilité, est un indicateur du degré d'épaisseur des queues de la distribution : plus il est petit, plus les queues sont épaisses ce qui correspond à de très grandes fluctuations. C'est le paramètre le plus important, il est compris entre 0 et 2 ($0 < \alpha \leq 2$). Sa valeur maximale $\alpha = 2$ correspond à une loi stable particulière : la loi de Gauss ou loi normale. β est le paramètre de dissymétrie, il varie entre -1 et 1 ($-1 \leq \beta \leq 1$) et quand il est nul, la distribution est symétrique par rapport à μ . Lorsque α s'approche de 2, β perd de son effet entraînant une tendance vers la distribution normale. Les paramètres $\mu \in \mathbb{R}$ et $\sigma > 0$ représentent les usuelles caractéristiques de position et d'échelle respectivement avec la remarque que pour la loi gaussienne, l'écart type vaut $\sigma\sqrt{2}$. Enfin, une v.a. X de distribution stable sera notée, selon Samorodnitsky et Taquq (1994), par $X \sim S_\alpha(\sigma, \beta, \mu)$.

Les trois exceptions mentionnées plus haut sont la très célèbre loi de Gauss $S_2(\sigma, 0, \mu)$ et les lois moins connues celle de Cauchy $S_1(\sigma, 0, \mu)$ et celle de Lévy $S_{1/2}(\sigma, 1, \mu)$.

	$0 < \alpha \leq 1$	$1 < \alpha < 2$	$\alpha = 2$
EX	∞	μ	μ
$VarX$	∞	∞	$2\sigma^2$

TAB. 1 – Moyenne et variance théoriques d’une loi α -stable

α	2	1.75	1.5	1.25	1	0.75
moyenne	-0.022	0.013	0.044	-0.799	-0.902	-44.3615
variance	2.051	3.245	19.174	902.617	1913.208	1499629

TAB. 2 – Moyenne et variance empiriques d’une loi stable symétrique

La loi stable jouit d’une propriété d’additivité selon laquelle la somme de deux v.a. stables indépendantes de même indice de stabilité α est encore stable avec le même exposant caractéristique α . Cette propriété très intéressante, est utilisée en finance pour étudier les portefeuilles où deux actifs ayant la même valeur pour α peuvent être considérés ensemble.

Une des particularités de la distribution stable est qu’elle est à variance infinie dès que α est strictement inférieur à 2. En fait, les moments d’ordre $p \in \mathbb{N}^*$ d’une v.a. $X \sim S_\alpha(\sigma, \beta, \mu)$ sont tels que pour $\alpha = 2$,

$$E|X|^p < \infty, \quad \forall p,$$

pour $0 < \alpha < 2$,

$$E|X|^p \begin{cases} < \infty, & p < \alpha \\ = \infty, & p \geq \alpha \end{cases}.$$

En particulier, pour la moyenne $E(X)$ et la variance $Var(X)$, on résume les résultats dans le tableau suivant :

On constate que, dès que $\alpha < 2$, la variance d’une loi stable est infinie et dès que $\alpha \leq 1$, c’est la moyenne qui devient infinie. Lorsque $\alpha > 1$, la moyenne existe et vaut μ . A titre d’exemple, on simule (voir algorithme ci-dessous) 2500 réalisations d’une v.a. $X \sim S_\alpha(1, 0, 0)$ et on calcule la moyenne et la variance empiriques pour différentes valeurs de l’indice de stabilité α . Les résultats obtenus sont présentés dans le tableau 2 où l’explosion de la variance lorsque α décroît est très remarquable.

2.2 Queues lourdes

Lorsque l’exposant caractéristique d’une v.a. $X \sim S_\alpha(\sigma, \beta, \mu)$ est strictement inférieur à 2, la variance est infinie et les queues sont asymptotiquement équivalentes à celle d’une distribution de Pareto. De façon plus précise, il est montré que (voir par exemple Samorodnitsky et Taqqu, 1994)

$$\lim_{x \rightarrow \infty} x^\alpha P(X > x) = C_\alpha \frac{1 + \beta}{2} \sigma^\alpha, \tag{1}$$

et

$$\lim_{x \rightarrow \infty} x^\alpha P(X < -x) = C_\alpha \frac{1 - \beta}{2} \sigma^\alpha, \tag{2}$$

où

$$C_\alpha = \left(\int_0^\infty x^{-\alpha} \sin x dx \right)^{-1} = \frac{2}{\pi} \Gamma(\alpha) \sin \frac{\alpha\pi}{2}. \tag{3}$$

Les deux relations (1) et (2), entraînent que, lorsque $n \rightarrow \infty$

$$P(|X| > x) \sim C_\alpha \sigma^\alpha x^{-\alpha}. \tag{4}$$

On voit donc que la loi stable prend en considération les queues de distribution qui sont souvent porteuses d'informations essentielles, alors que la loi de Gauss néglige ces queues entraînant ainsi une erreur qui peut être fatale pour l'investisseur.

2.3 Algorithme de génération

Il existe plusieurs algorithmes permettant de générer des v.a. stables. Parmi ces algorithmes, celui de Chambers, Mallows et Stuck (1976) occupe une bonne place. Nous présentons ci-dessous une version légèrement modifiée de cet algorithme (voir Janicki et Weron, 1994 et Weron et Weron, 1995), dont le temps de calcul est quasiment nul (il faut moins d'une minute pour simuler un échantillon de taille 60 000).

1. Posons

$$C = \frac{\arctan\left(\beta \tan \frac{\alpha\pi}{2}\right)}{\alpha}, \quad \text{et} \quad D = \left(1 + \beta^2 \tan^2 \frac{\alpha\pi}{2}\right)^{1/2\alpha}.$$

2. Calculons

$$Y = \begin{cases} D \times \sin(\alpha(V + C))(\cos V)^{1/\alpha} \times (\cos(V - \alpha(V + C)))W^{(1-\alpha)/\alpha}, & \alpha \neq 1, \\ 2\pi \left((\pi^2 + \beta V) \tan V - \beta \log(W \cos V \frac{\pi}{2} + \beta V) \right), & \alpha = 1, \end{cases}$$

où V est une v.a. uniforme sur $[-\frac{\pi}{2}, \frac{\pi}{2}]$ et W une v.a. exponentielle de paramètre 1. La v.a. Y ainsi obtenue est de distribution $S_\alpha(1, \beta, 0)$.

3. Enfin, le changement de variable

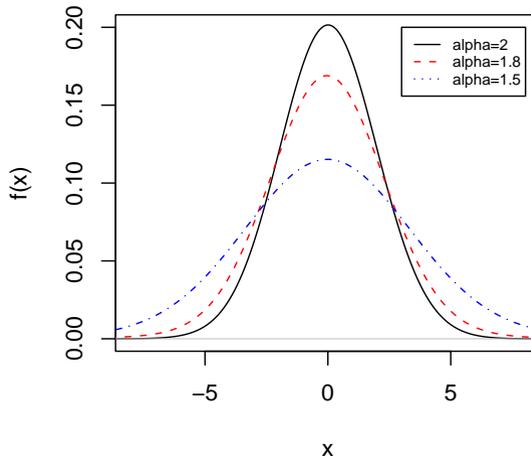
$$X = \begin{cases} \sigma Y + \mu, & \alpha \neq 1 \\ \sigma Y + \frac{2}{\pi} \beta \sigma \log \sigma + \mu, & \alpha = 1 \end{cases},$$

nous permet d'obtenir une v.a. $X \sim S_\alpha(\sigma, \beta, \mu)$.

On remarque que pour $\alpha = 2$, $\sigma = 1$ et $\mu = \beta = 0$, on retrouve l'algorithme de simulation d'une variable normale standard $\mathcal{N}(0, 1)$.

$$Y = 2\sqrt{W} \sin V \sim \mathcal{N}(0, 1).$$

On applique cet algorithme pour générer trois variables stables symétriques ($\beta = \mu = 0$ et $\sigma = 1$) puis on trace les graphes des densités correspondantes avec $\alpha = 2, 1.8$ et 1.5 ; on constate sur la figure 1) que plus le paramètre α est petit, plus les queues des distributions sont épaisses.

FIG. 1 – Densité de $S_\alpha(1, 0, 0)$.

2.4 Estimation des paramètres

L'estimation des paramètres de la loi stable $S_\alpha(\sigma, \beta, \mu)$ est sévèrement affectée par l'absence de forme explicite pour la fonction de densité de probabilité, d'où le recours à des méthodes numériques pouvant être complexes et lentes dans leur exécution. Un grand nombre d'algorithmes, plus ou moins efficaces et plus ou moins rapides, sont proposés selon des approches différentes. Nous pouvons citer l'approche des quantiles empiriques (voir par exemple McCulloch, 1986), l'approche de la fonction caractéristique empirique (voir par exemple Koutrouvelis, 1980 ; Kogon et Williams, 1998) et l'approche du maximum de vraisemblance (voir par exemple DuMouchel, 1973 ; Nolan, 2001).

En général, les performances de toutes ces méthodes sont bonnes, mais Ojeda (2001) observe, dans une étude comparative, que les méthodes du maximum de vraisemblance donnent les estimations les plus précises avec cependant un temps d'exécution plus long. C'est ce que confirment des études de simulation faites par Weron (2004).

3 Approche des valeurs extrêmes

En finance, comme dans d'autres secteurs de la vie socio-économique, une des préoccupations majeures des gestionnaires est d'être capables de se prémunir contre les crises (crashes financiers, catastrophes naturelles, ...). Il est donc fondamental pour l'investisseur d'être en mesure d'évaluer la probabilité, malheureusement non nulle, d'occurrence d'événements rares. Une forte variation du prix d'un certain actif financier est un événement rare qui se traduit par l'apparition de valeurs extrêmes dans la série des rendements de cet actif, d'où une certaine

épaisseur dans les queues de distribution. C'est justement à cause de la nature de ces queues que l'on peut faire intervenir la théorie des valeurs extrêmes qui peut jouer (à travers l'estimation des paramètres stables) un rôle important dans la connaissance du comportement des marchés financiers. L'avantage de l'approche des valeurs extrêmes, par rapport aux méthodes citées ci-dessus, dans l'estimation des paramètres stables est qu'elle donne des estimateurs de formes explicites permettant ainsi un calcul plus direct des valeurs estimées. De plus, la propriété de normalité asymptotique de ces estimateurs rend possible la construction d'intervalles de confiance. Pour une revue de la théorie des valeurs extrêmes et de ses principaux résultats se référer par exemple à Härdle, Klinkle et Müller (2000).

La construction des estimateurs est basée sur la relation (4) qui traduit le comportement des queues de la v.a. $X \sim S_\alpha(\sigma, \beta, \mu)$. Si on désigne par G la fonction de répartition de $Z = |X|$ alors on peut écrire quand $x \rightarrow \infty$

$$1 - G(x) := P(Z > x) \sim C_\alpha \sigma^\alpha x^{-\alpha},$$

ce qui entraîne que la queue de G vérifie la condition de variation régulière, d'indice α , suivante (voir par exemple Geluk et de Haan, 1987) :

$$\lim_{t \rightarrow \infty} \frac{1 - G(tx)}{1 - G(t)} = x^{-\alpha} \text{ pour } x > 0. \tag{5}$$

D'après Gnedenko (1943), la relation (5) équivaut à dire que G appartient au domaine d'attraction de Fréchet. Plus précisément, pour un échantillon Z_1, \dots, Z_n ($n \geq 1$) de la variable Z , on a

$$\frac{\max(Z_1, \dots, Z_n)}{G^{-1}(1 - n^{-1})} \xrightarrow{d} \Phi_\alpha,$$

où

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

et

$$G^{-1}(1 - n^{-1}) := \inf \{x \in \mathbb{R}, G(x) \geq 1 - n^{-1}\},$$

avec \xrightarrow{d} désignant la convergence en distribution.

Afin de prouver la normalité asymptotique des estimateurs, on impose une condition plus stricte sur G . On suppose que G appartient à la classe des modèles de Hall, (Hall et Welsh, 1985), i.e.

$$1 - G(x) = Ax^{-\alpha} \{1 + Bx^{\alpha\rho} (1 + o(1))\} \text{ lorsque } x \rightarrow \infty,$$

où $0 < \alpha < 2$ (indice de stabilité), $A > 0$, $B \in \mathbb{R}^*$ et $\rho < 0$ (paramètre du second ordre). C'est un cas particulier de la condition plus générale de variation régulière du second ordre très utilisée en statistiques des extrêmes (voir par exemple de Haan et Stadtmüller, 1996).

Les ouvrages de Feller (1971) et Bingham et al. (1987) sont de très bonnes références pour une présentation générale de la notion de variation régulière.

Les estimateurs ainsi construits dépendent essentiellement du nombre k de statistiques d'ordre à prendre en compte dans le calcul des valeurs estimées des différents paramètres. Ce nombre

influe sur le comportement des estimateurs, dans le sens que d'une part utiliser trop de données entraîne un biais et d'autre part utiliser trop peu d'observations résulte en une variance substantielle, d'où la nécessité de bien le choisir afin de garantir les meilleures estimations possibles. La question (fondamentale) qui se pose est alors la suivante : comment déterminer le nombre k optimal ? Il existe plusieurs approches pour répondre à cette question. Une première approche graphique qui consiste à considérer l'estimateur comme fonction du nombre k et puis chercher sur le graphe correspondant une région de stabilité autour de la ligne horizontale représentant la vraie valeur du paramètre, cette région correspond au nombre optimal de statistiques extrêmes nécessaires pour estimer le paramètre. Une deuxième méthode est basée sur l'idée de minimiser l'erreur quadratique moyenne asymptotique (AMSE : Asymptotic Mean Squared Error) de l'estimateur comme fonction de k , le point qui réalise le minimum représente un choix optimal pour le nombre de valeurs extrêmes devant être utilisées dans l'estimation. Une troisième procédure consiste à minimiser la précision de couverture (en valeur absolue) pour des intervalles de confiance. Pour plus de détails sur ce problème de sélection du nombre k de valeurs extrêmes, se référer par exemple à Hall et Welsh (1985), Hall (1990), Dekkers et de Haan (1993), Drees (1998), Cheng et Peng (2001), Danielson et al. (2001) et Ferreira et de Vries (2004). Plus récemment, Neves et Fraga Alves (2004) ont fait une évaluation de la performance d'une autre méthode, celle proposée par Reiss et Thomas (1997). Enfin, notons que le processus de choix du nombre optimal (de statistiques extrêmes devant intervenir dans le calcul des estimations) est malheureusement rendu difficile par le fait que ce nombre dépend aussi d'un nouveau paramètre inconnu (paramètre du second ordre) et qui doit être estimé de manière consistante afin d'avoir une estimation du nombre k optimal.

3.1 Estimation du paramètre α

L'exposant caractéristique α d'une variable Lévy-stable représente l'inverse de l'indice de queue $\gamma = 1/\alpha$ d'une distribution à queues lourdes. Ce dernier est largement étudié en théorie des valeurs extrêmes où plusieurs bons estimateurs sont proposés, (voir par exemple Pickands, 1975 ; Hill, 1975 ; Csörgő, Deheuvels et Mason, 1985 ; Dekkers, Einmahl et de Haan, 1989 ; Groeneboom, Lopushaä et de Wolf, 2003 ;...), dont le plus célèbre (mais pas nécessairement le meilleur) est celui de Hill (Hill, 1975). C'est un estimateur consistant et asymptotiquement normal (voir par exemple Peng, 1998), il est défini comme suit :

$$\hat{\alpha}_n := \left(\frac{1}{k} \sum_{i=1}^k \log Z_{n-i+1,n} - \log Z_{n-k,n} \right)^{-1},$$

où $Z_{1,n} \leq \dots \leq Z_{n,n}$ sont les statistiques d'ordre associées à un échantillon Z_1, \dots, Z_n ($n \geq 1$) de la v.a. Z et $k = k(n)$ est une suite entière vérifiant $k \rightarrow \infty$ et $k/n \rightarrow 0$ lorsque $n \rightarrow \infty$.

Discutant la performance de cet estimateur, Weron (2001) note que pour $\alpha \leq 1.5$, l'estimation est assez raisonnable mais lorsque α s'approche de 2 il y a une surestimation significative quand on travaille avec des échantillons de tailles typiques. Pour de telles valeurs de α , il est nécessaire de disposer d'un très grand nombre d'observations (un million ou plus) pour obtenir des estimations acceptables, car le vrai comportement des queues des distributions Lévy-stables est visible seulement dans le cas d'ensembles de données extrêmement grands.

De nos jours et grâce à une technologie très sophistiquée, le stockage et le traitement de quantités considérables de données sont devenus chose aisée aussi bien en finance qu'en d'autres secteurs de la vie socio-économique.

3.2 Estimation du paramètre μ

Dans le cas $1 < \alpha < 2$, le paramètre de position μ n'est autre que la moyenne EX (voir TAB.2). Un estimateur naturel pour μ est donc la moyenne empirique \bar{X} dont la normalité asymptotique ne peut être établie dans le cas des lois α -stables en raison de leur variance infinie. Pour résoudre ce problème, Peng (2001) propose un estimateur $\hat{\mu}_n$ basé sur les valeurs extrêmes et prouve sa consistance et sa normalité asymptotique. Cet estimateur est défini comme suit :

$$\hat{\mu}_n := \hat{\mu}_n^{(1)} + \hat{\mu}_n^{(2)} + \hat{\mu}_n^{(3)},$$

où

$$\hat{\mu}_n^{(2)} = \frac{1}{n} \sum_{i=1}^{n-k} X_{i,n},$$

est la moyenne empirique tronquée, et

$$\hat{\mu}_n^{(1)} = \frac{k}{n} X_{k,n} \frac{\hat{\alpha}_n^{(1)}}{\hat{\alpha}_n^{(1)} - 1}, \quad \hat{\mu}_n^{(3)} = \frac{k}{n} X_{n-k+1,n} \frac{\hat{\alpha}_n^{(3)}}{\hat{\alpha}_n^{(3)} - 1},$$

avec

$$\hat{\alpha}_n^{(1)} = \left(\frac{1}{k} \sum_{i=1}^k \log(-X_{i,n}) - \log(-X_{k+1,n}) \right)^{-1},$$

et

$$\hat{\alpha}_n^{(3)} = \left(\frac{1}{k} \sum_{i=1}^k \log(X_{n-i+1,n}) - \log(X_{n-k,n}) \right)^{-1}.$$

Récemment, Necir (2006b) a étudié la convergence forte de l'estimateur $\hat{\mu}_n$ dans le but de construire un test séquentiel de puissance 1 pour la moyenne d'une distribution Lévy-stable. Les statistiques $\hat{\alpha}_n^{(1)}$ et $\hat{\alpha}_n^{(3)}$ sont deux versions consistantes de l'estimateur de Hill pour le paramètre α (voir Mason, 1982). Leur convergence presque sûre est prouvée, dans un cadre général, par Necir (2006a).

3.3 Estimation du paramètre σ

Pour le paramètre d'échelle σ , la combinaison des relations (3) et(4) et les résultats de la théorie des valeurs extrêmes nous permettent de construire un estimateur $\hat{\sigma}_n$ dont la forme est la suivante :

$$\hat{\sigma}_n := Z_{n-k,n} \left(\frac{k\pi}{2n\Gamma(\hat{\alpha}_n) \sin \pi\hat{\alpha}_n 2} \right)^{1/\hat{\alpha}_n},$$

(voir Meraghni et Necir, 2005).

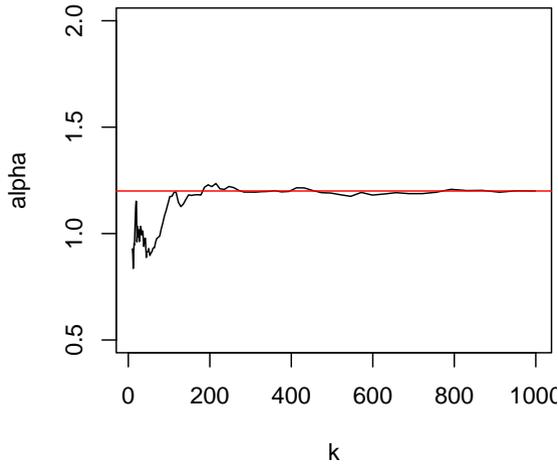


FIG. 2 – Estimateur de l'indice de stabilité de $S_{1.2}(0.5, 0, 0)$.

Comme notre estimateur $\hat{\sigma}_n$ dépend de manière étroite de $\hat{\alpha}_n$, il semble que la remarque de Weron (2001) sur la performance de $\hat{\alpha}_n$ reste valable pour $\hat{\sigma}_n$. Le comportement asymptotique de $\hat{\sigma}_n$ est discuté dans Meraghni et Necir (2005) où la consistance et la normalité asymptotique sont établies.

Une procédure analogue nous permettra d'aboutir à la construction (en cours de réalisation) d'un estimateur pour le paramètre de dissymétrie β .

3.4 Simulation

A titre d'exemples, on s'intéresse au comportement des estimateurs $\hat{\alpha}_N$, $\hat{\mu}_N$ et $\hat{\sigma}_N$ en simulants, par l'algorithme de Chambers, Mallows et Stuck (1976), des échantillons de taille $N = 10000$ de distributions α -stables symétriques $S_\alpha(\sigma, 0, 0)$ avec $\alpha = 1.2$ et $\sigma = 0.5$. On trace (séparément) les graphes des estimateurs $\hat{\alpha}_N$, $\hat{\mu}_N$ et $\hat{\sigma}_N$ comme fonctions du nombre k pour avoir une idée sur le nombre optimal de statistiques d'ordre nécessaires à l'estimation des paramètres α , μ et σ .

On voit sur les figures 2 et 3 que si on utilisait entre 300 et 400 statistiques extrêmes dans le calcul de $\hat{\alpha}_N$ et $\hat{\mu}_N$ on obtiendrait de très bonnes estimations pour les paramètres α et μ respectivement, alors que la figure 4 suggère un choix optimal autour de 600 pour le nombre de valeurs extrêmes à prendre en considération lors de l'estimation du paramètre d'échelle σ .

4 Modélisation des rendements du CAC 40

Le rendement journalier (hebdomadaire, mensuel,...) d'un actif financier représente les accroissements quotidiens (hebdomadaires, mensuels,...) du logarithme de son prix, il est défini

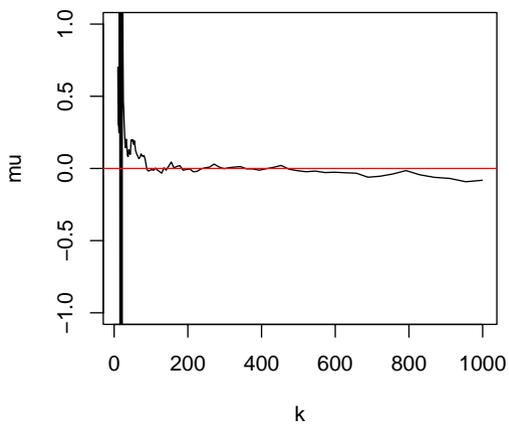


FIG. 3 – Estimateur du paramètre de position de $S_{1.2}(0.5, 0, 0)$

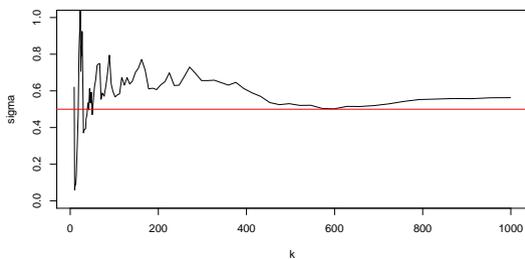


FIG. 4 – Estimateur du paramètre d'échelle de $S_{1.2}(0.5, 0, 0)$.

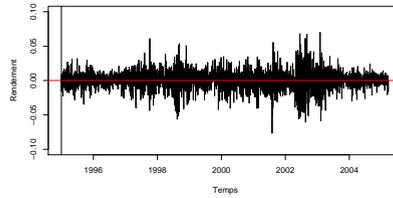


FIG. 5 – *Serie des rendements quotidiens du CAC40 (03/01/95-29/04/05).*

comme suit :

$$R_t := \log \frac{P_t}{P_{t-1}}, \quad t \geq 1,$$

où P_t désigne le prix de l'actif (le cours du titre financier) à l'instant t . Cela correspond au gain (perte) relatif. Il est donc plus intéressant pour l'investisseur que le prix lui-même, car il lui permet de déterminer les profits qu'il peut réaliser. C'est pour cela que les analyses financières sont basées sur les rendements des actifs au lieu de leurs prix.

Les rendements d'actifs financiers ont été pour longtemps décrits, en vertu du théorème central limite, par la distribution normale, car étant souvent considérés comme résultant du cumul d'un grand nombre d'informations et de décisions individuelles. Mais cette modélisation est contredite, depuis les travaux de Mandelbrot dans les années 60, par beaucoup d'indications empiriques parmi lesquelles on peut citer :

- Une possibilité de dissymétrie dans la distribution empirique des rendements.
- Des queues épaisses dans l'histogramme, révélant un comportement asymptotique plus lent que celui de la distribution normale.
- Une variance empirique infinie ou plutôt instable (elle continue à varier considérablement quelle que soit la taille de l'échantillon).
- Un coefficient d'aplatissement (kurtosis) empirique très supérieur à 3 (caractéristique de la loi de Gauss).

Donc pour la plupart des actifs financiers, en particulier lorsque la valeur du kurtosis dépasse nettement 3, la loi normale ajuste mal les données. Mandelbrot (1963) puis Fama (1965) lui préfèrent la loi Lévy-stable qui apporte une nette amélioration surtout aux queues de distribution.

Dans ce qui suit, nous allons traiter, comme exemple, les rendements journaliers de l'indice boursier CAC 40. Les données, prises dans le site internet "www.finance.yahoo.com", concernent les prix (de clôture) quotidiens durant la période allant du 03/01/1995 au 29/04/2005, soit 2609 observations. Les 2608 rendements correspondants sont représentés dans la figure 5, sur laquelle on constate la non régularité des accroissements et l'apparition d'un certain nombre de variations extrêmes tant positives que négatives.

La série des rendements, dont les caractéristiques statistiques sont résumées dans le tableau 3, est distribuée comme l'indique la figure 6.

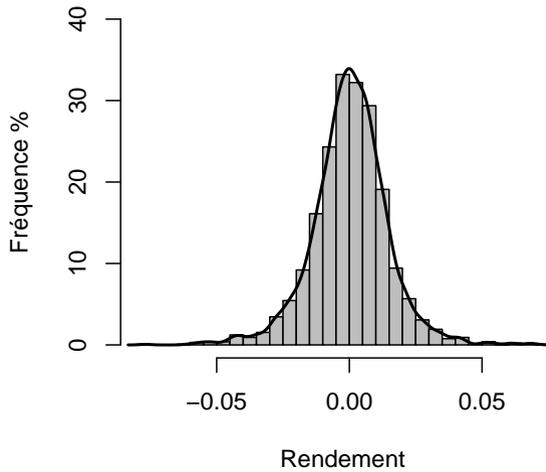


FIG. 6 – Distribution des rendements du CAC 40 (03/01/95629/04/95).

moyenne	0.0003
écart type	0.0144
minimum	-0.0768
maximum	0.0700
1° quartile	-0.0073
médiane	0.0005
3° quartile	0.0084

Tableau 3 : Caractéristiques empiriques des rendements du CAC 40

La valeur du kurtosis de cette série financière est $5.41 > 3$, ce qui entraînerait la non validité de l’ajustement Gaussien. Le qq-plot de la figure 7 montre une nette déviation par rapport à la direction droite surtout aux extrémités, confirmant ainsi le rejet de l’hypothèse selon laquelle les rendements seraient normalement distribués. D’autre part, la valeur de la statistique du test de Shapiro-Wilk est de 0.9727 avec une $p - value < 2.2 \cdot 10^{-16}$. Ceci nous amène aussi à écarter la normalité des rendements de l’indice boursier du CAC 40.

Pour ajuster une loi stable à nos données, on estime les quatre paramètres par la méthode de McCulloch (1986), qui fournit des estimateurs consistants avec la restriction $\alpha \geq 0.6$ améliorant ainsi la méthode de Fama et Roll (1971) applicable seulement dans le cas $\beta = 0$ et $\alpha \geq 1$. Les résultats obtenus sont résumés dans le tableau 4. Il est à noter que les titres financiers ont, en général, des exposants caractéristiques strictement compris entre 1 et 2 (voir par exemple Adam, 2001).

La loi Lévy-stable apporte une nette amélioration à la description des rendements quotidiens (extrêmes) de l’indice boursier CAC 40 par rapport à la loi de Gauss. Les grandes fluctuations

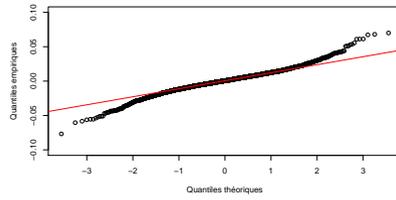


FIG. 7 – *Q-Q plot des rendements du CAC 40*

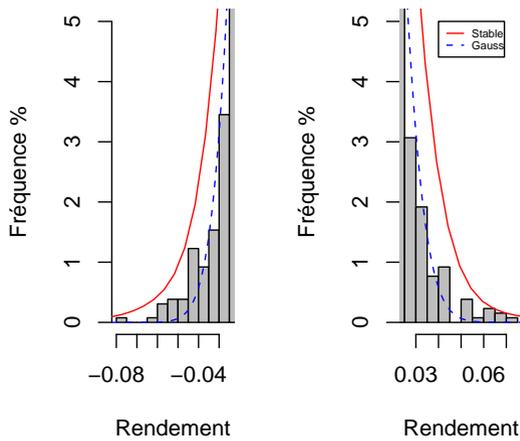


FIG. 8 – *Rendements extrêmes du CAC 40*

α	β	σ	μ
1.618	-0.166	0.008	0.000

TAB. 3 – *Paramètres stables des rendements du CAC 40*

jusqu' alors négligées par la distribution normale sont ainsi mises en valeur par la distribution α -stable.

En conclusion, on peut dire que l'ajustement normal des rendements financiers peut être dangereux pour l'investisseur dans la mesure où il lui fait sous-estimer le risque d'avoir des pertes très importantes.

Remerciements

Nous tenons à remercier les referees pour leurs précieuses critiques, lesquelles ont permis une amélioration substantielle de l'article.

Références

Adam, E. (2001). L'analyse fractale des marchés financiers. *Stage effectué à Finama Asset Management*.

Bassi, F., Embrechts, P. and Kafetzaki, M. (1998). Risk Management and Quantile Estimation. *In Adler, Feldman et Taqqu (eds). A practical guide to heavy tails. Birkhäuser, Boston*, 111-130.

Bestavros, A., Crovella, M. and Taqqu, M. (1998). Heavy-Tailed Probability Distributions in the World of Wide Web. *In Adler, Feldman et Taqqu (eds). A practical guide to heavy tails. Birkhäuser, Boston*, 3-25.

Bingham, N. H., Goldie, C. M., Teugels, J. L. (1987). Regular Variation. Cambridge University Press, Cambridge.

Chambers, J.M., Mallows, C.L. and Stuck, B.W. (1976). A Method for Simulating Stable Random Variables. *Journal of the American Statistical Association* **71**, 340-344.

Cheng, S. and Peng, L. (2001). Confidence Intervals for the Tail Index. *Bernoulli* **7**, 751-760.

Csörgő, S., Deheuvels, P. and Mason, D. (1985). Kernel Estimates of the Tail Index of a Distribution. *Annals of Statistics* **13**, 1050-1077.

Danielson, J., de Haan, L., Peng, L. and de Vries, C.G. (2001). Using a Bootstrap Method to Choose the Sample Fraction in Tail Index Estimation. *Journal of Multivariate Analysis* **76**, 226-248.

Dekkers, A.L.M., Einmahl, J.H.J. and de Haan, L. (1989) A Moment Estimator for the Index of an Extreme Value Index. *Annals of Statistics* **17**, 1833-1855.

Dekkers, A.L.M. and de Haan, L. (1993). Optimal Choice of Sample Fraction in Extreme-Value Estimation. *Journal of Multivariate Analysis* **47**, 173-195.

Drees, H. (1998). Optimal Rates of Convergence for Estimates of the Extreme Value Index. *Annals of Statistics* **26**, 434-448.

DuMouchel, W.H. (1973). On the Asymptotic Normality of the Maximum-Likelihood Estimate when Sampling from a Stable Distribution. *Annals of Statistics* **1** (5), 948-957.

Fama, E.F. (1965). The Behavior of Stock Market Prices. *Journal of Business* **38**, 34-105.

Fama, E.F. and Roll, R. (1971). Parameter Estimates for Symmetric Stable Distributions. *Journal of the American Statistical Association* **66**, 331-338.

Feller W. (1971). An introduction to probability theory and its applications. Vol. II. *JohnWiley & Sons Inc.*, 2 ed.

Ferreira, A. and de Vries, C. G., (2004). Optimal Confidence Intervals for the Tail Index and High Quantiles. *Tinbergen Institute Discussion Paper*, 090/2.

Geluk, J. and de Haan, L. (1987). Regular Variation, Extensions and Tauberian Theorems. *Mathematical Centre Tracts* **40**, Centre for Mathematics and Computer Science, Amsterdam.

Gnedenko, B.V. (1943). Sur la distribution limite du terme maximum d'une série aléatoire. *Annales de Mathématiques* **44**, 423-453.

Groeneboom, P., Lopuhaäa, H.P. and de Wolf, P.P. (2003). Kernel Estimators for the Extreme Value Index. *Ann. Statist.*, **31**, 1956-1995.

de Haan, L. and Stadtmüller, U. (1996). Generalized Regular Variation of Second Order. *J. Australian Math. Soc. (Series A)* **61**, 381-395.

Hall, P. (1990). Using the Bootstrap to Estimate Mean Squared Error and Select Smoothing Parameter in Nonparametric Problems. *Journal of Multivariate Analysis* **32**, 177-203.

Hall, P. and Welsh, A.H. (1985). Adaptive Estimates of Parameters of Regular Variation. *Annals of Statistics* **13**, 331-341.

Härdle, W., Klinke, S., and Müller, M. (2000). XploRe Learning Guide, *Springer*.

Hill, B. (1975). A Simple Approach to Inference About the Tail of a Distribution, *Annals of Statistics* **3**, 1163-1174.

Ihaka, R. and Gentleman, R. (1996). R : A Language for Data Analysis and Graphics. *Journal of Computational and Graphical Statistics* **5**, 299-314.

Janicki, A. and Weron, A. (1994). Simulation and Chaotic Behavior of α -Stable Stochastic Processes. *New York, Marcel Dekker*.

Kogon, S.M. and Williams, D.B. (1998). Characteristic Function Based Estimation of Stable Parameters. In R. Adler, R. Feldman, M. Taqqu (eds.), *A Practical guide to heavy tails. Birkhäuser*, 311-335.

Koutrouvelis, I.A. (1980). Regression-Type Estimation of the Parameters of Stable Law. *Journal of the American Statistical Association* **75**, 918-928.

Lévy, P. (1925). Calcul des Probabilités. *Paris, Gauthier-Villars*.

Mandelbrot, B. (1963). The Variation of Certain Speculative Prices. *Journal of Business* **36**, 394-419.

Mason, D. M. (1982). Laws of Large Numbers for Sums of Extreme Values. *Annals of Probability* **10**, 754-764.

McCulloch, J. H. (1986). Simple Consistent Estimators of Stable Distribution Parameters. *Communications in Statistics – Computation and Simulation* **15**, 1109-1136.

Meraghni, D. and Necir, A. (2005). Estimating the Scale Parameter of a Lévy-Stable Distribution via the Extreme Value Approach. (*soumis*).

Necir, A. (2006a). A Functional Law of the Iterated Logarithm for Kernel-type estimators of the Tail Index. *Journal of Statistical Planning and Inference* **136**, 780-802.

Necir, A. (2006b). A Nonparametric Sequential Test with Power 1 for the Mean of Lévy-Stable Distributions with Infinite Variance. *Methodology and Computing in Applied Probability*, (à paraître).

Neves, C. and Fraga Alves, M.I. (2004). Reiss and Thomas' Automatic Selection of the Number of Extremes. *Computational Statistics and Data Analysis* **47**, 689-704.

Nolan, J.P. (2001). Maximum Likelihood Estimation and Diagnostics for Stable Distributions. In O.E. Barndorff-Nielsen, T. Mikosch, S. Resnick (eds.), *Lévy Processes, Birkhäuser, Boston*.

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

Peng, L. (1998). Asymptotically Unbiased Estimators for the Extreme-Value Index. *Statistics and Probability Letters* **38**, 107-115.

Peng, L. (2001). Estimating the Mean of a Heavy Tailed Distribution. *Statistics and Probability Letters* **52**, 255-264.

Pickands III, J. (1975). Statistical Inference Using Extreme Order Statistics. *Annals of Statistics* **3**, 119-131.

Reiss, R.D. and Thomas, M. (1997). Statistical Analysis of Extreme Values with Applications to Insurance, Finance, Hydrology and Other Fields. *Birkhäuser, Basel*.

Samorodnitsky, G. and Taqqu, M.S. (1994). Stable Non-Gaussian Random Processes : Stochastic models with infinite variance. *Chapman & Hall, New York*.

Weron A. and Weron R. (1995). Computer Simulation of Lévy α -Stable Variables and Processes. *Lecture Notes in Physics*, **457**, 379-392.

Weron, R. (2001). Levy-Stable Distributions Revisited : Tail Index 2 Does Not Exclude the Levy-Stable Regime. *International Journal of Modern Physics C*. **12** (2), 209-223.

Weron, R. (2004). Computationally Intensive Value at Risk Calculations. In J. E. Gentle, W. Härdle, Y. Mori (eds). *Handbook of Computational Statistics*. Springer, Berlin, 911-950.

Summary

In the 1960's, Mandelbrot's works on stock market fluctuations showed that the Gaussian model was not appropriate to describe asset returns. Mandelbrot (1963) then fama (1965) proposed the Lévy-stable distribution, whose properties are very close to those of the empirical distributions with heavy tails, as an alternative to model financial series. This choice is justified by at least two reasons. The first one is the Generalized Central Limit Theorem which states that the stable law is the only possible limit distribution for adequately centered and normalized sums of independent identically distributed random variables. The second reason is that the stable distribution allows skewness and fat tails in such a way that it fits the empirical distributions much better than the Gaussian distribution does. This seems to be confirmed by the example of the daily returns of the French stock index CAC 40. In this paper, we are interested in the crucial issue of estimating the four parameters that characterize the Lévy-stable law.

