

MODÈLES FINANCIERS EN ASSURANCE ET ANALYSES DYNAMIQUES

Support de cours 2023-2024

Utilisation de la théorie des valeurs extrêmes dans le cadre de Solvabilité 2

Frédéric PLANCHET

Version 5.3

Décembre 2023

SOMMAIRE

1. Introduction 3
1.1. Contexte 3
1.2. Qu'est-ce qu'un événement rare ? 5
2. Rappels sur la théorie des valeurs extrêmes 5
2.1. Distribution des valeurs extrêmes 6
2.1.1. Résultats principaux 6
2.1.2. La distribution GEV 8
2.1.3. L'identification du type de valeurs extrêmes 9
2.2. Les estimateurs non paramétriques de l'indice de queue 10
2.2.1. L'estimateur de Pickands 11
2.2.2. L'estimateur de Hill 11
2.2.3. L'estimateur de Hill Alternatif 12
2.2.4. L'estimateur des moments 12
2.2.5. Comparaison des différents estimateurs 13
2.3. Un estimateur paramétrique de l'indice de queue 13
2.4. Séparation des données standards et des données de queue 14
3. L'estimation des distributions au-dessus d'un seuil 15
3.1. Le nombre d'excès au-delà d'un seuil 15
3.2. La distribution de Pareto généralisée (GPD) 16
3.2.1. Propriétés 17
3.2.2. Estimation des paramètres 18
3.3. L'espérance au-delà d'un seuil 20
3.3.1. Estimation 21
3.3.2. Exemple : la loi de Pareto 22
3.3.3. Détermination du seuil 22
3.4. Estimation des quantiles élevés 23
3.4.1. Approximation GPD 23
3.4.2. Estimateur de Hill 24
3.4.3. Illustration : cas de la loi de Pareto 25
4. Application : calcul d'une VaR en assurance 26
4.1. L'approche paramétrique directe 27
4.1.1. Estimation ponctuelle 27
4.1.2. Rappels sur les méthodes bootstrap 27
4.1.3. Intervalles de confiance par la méthode du boostrap 29
4.1.4. Application numérique 31
4.2. Exemples 32
4.2.1. Loi log-normale (DA de Gumbel) 32
4.2.2. Loi de Pareto (DA de Fréchet) 33
4.2.3. Loi de Weibull 34
5. Références 36

1. Introduction

L'objectif de ce document est de présenter les principaux résultats de la théorie des valeurs extrêmes pouvant être utilisés dans les problématiques assurantielles, notamment dans un contexte Solvabilité 2. Il ne s'agit donc que d'une brève synthèse qui ne prétend pas à l'exhaustivité. Le lecteur intéressé par le sujet pourra utilement se reporter à EMBRECHTS et al. [1997] ainsi qu'à GARRIDO [2002] pour les aspects théoriques et ZAJDENWEBER D. [2000] pour les aspects pratiques.

1.1. Contexte

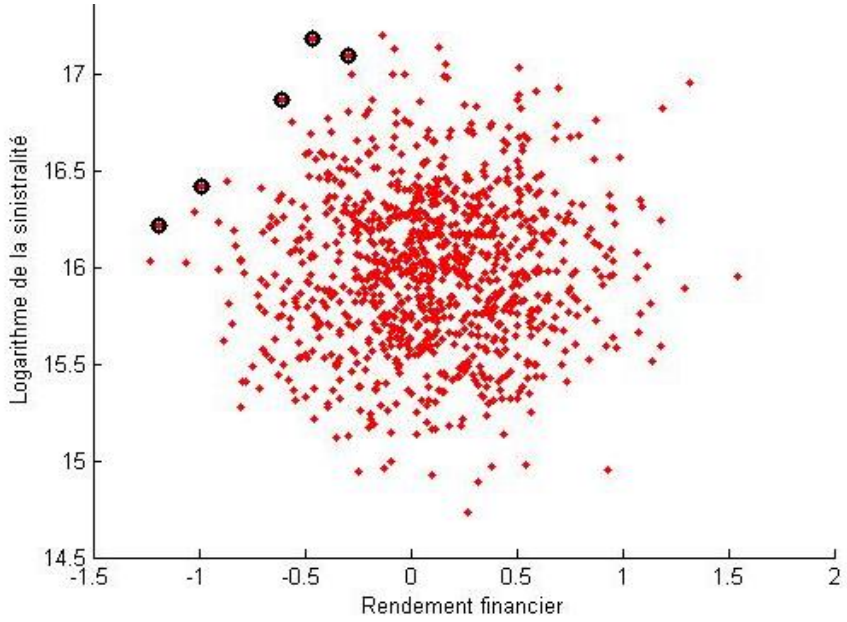
L'une des difficultés de Solvabilité 2 est que l'on doit évaluer des grandeurs qui dépendent de la queue de distribution, pour laquelle, quasiment par définition, on ne dispose pas d'observations (ou du moins on ne dispose que de très peu d'observations).

L'analyse des queues de distribution est cependant indispensable pour une appréciation pertinente et une gestion efficace du risque puisque l'estimation de la *VaR* dépend de prédictions correctes sur des événements rares. En particulier le quantile à 99,5 %, pivot du calcul du capital économique dans le dispositif Solvabilité 2[1] ne doit pas être sous-estimé.

La figure ci-dessous présente un échantillon de 1 000 réalisations du montant d'actif dont doit disposer la société en 0 pour ne pas être en ruine en 1 dans le cadre d'un modèle interne très simplifié (illustration reprise de THÉROND et PLANCHET [2007]). Les cinq points soulignés correspondent aux cinq scénarios qui ont conduits aux valeurs maximales.

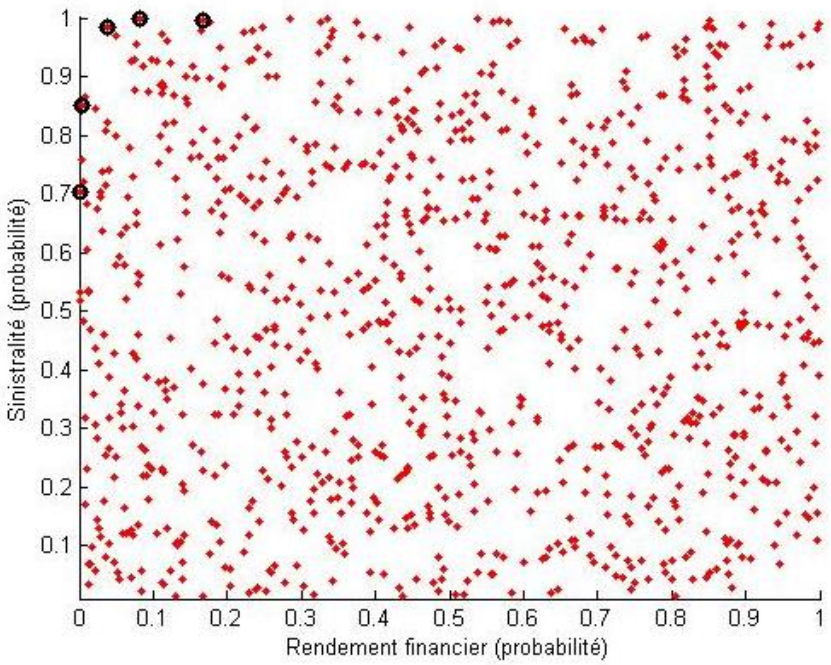
[1] Et pour le monde bancaire dans les accords de Bâle.

Fig. 1 - **Modèle interne simplifié : identification des valeurs extrêmes**



On remarque sur la figure ci-après qui représente les probabilités associées à chaque grandeur, que ces points se situent dans la queue de distribution d'au moins une des deux variables de base.

Fig. 2 - **Modèle interne simplifié : identification des valeurs extrêmes**



Or justement aux queues de distribution, l'adéquation des données au modèle retenu est souvent imparfaite.

La théorie des valeurs extrêmes fournit des outils pour estimer la loi des observations au-delà d'un seuil donné et pour calculer les quantiles extrêmes à l'aide de cette loi estimée. En effet, la loi des valeurs « au-delà d'un seuil » pour un échantillon iid, est, sous des conditions assez générales, une loi universelle (la distribution de Pareto généralisée), et il existe des méthodes statistiques pour estimer les paramètres de cette « loi de queue ».

1.2. Qu'est-ce qu'un événement rare ?

On considère un événement dont la probabilité d'apparition est p ; dans un modèle à temps discret, la distribution du premier instant de réalisation de l'événement est une loi géométrique de paramètre p, *i.e.* $P(T=k) = p(1-p)^{k-1}$. En particulier, le temps moyen d'attente de la réalisation de l'événement est $E(T) = \frac{1}{p}$.

On peut donc apprécier le caractère exceptionnel d'un événement au travers de sa fréquence : un événement qui se produit une fois par siècle est ainsi associé à une probabilité de survenance dans l'année de 1 %.

Le projet Solvabilité 2 retient un seuil de 0,5 % ; Pour fixer les idées[2], d'après Standard and Poor's la notation AAA correspond à une probabilité de défaut annuelle de 0,04 %, BBB à 0,6 %.

Cette approche de la notion d'événement rare est heuristique et on peut souhaiter disposer d'une caractérisation plus formelle : un événement rare est alors un événement associé à la queue de la distribution. La question est alors de définir la frontière entre les observations issues d'un échantillon que l'on peut considérer comme « normales » et celles associées à des « valeurs extrêmes ». Ce point sera développé ci-après (*cf.* 2.4 ci-dessous).

2. Rappels sur la théorie des valeurs extrêmes

Soit $X_1,\ldots,X_n$ un échantillon iid. Si F_X est la fonction de répartition commune des X_t alors :

$$P(M_n \le x) = P(\mathbf{max}(X_1,\ldots,X_n) \le x) = P(X_1 \le x,\ldots,X_n \le x) = (F_X(x))^n$$

avec $M_n = \mathbf{max}(X_1,\ldots,X_n) = -\mathbf{min}(-X_1,\ldots,-X_n)$. Ce résultat a une double conséquence : d'une part, dans le cas d'un échantillon iid, la loi des maxima s'obtient facilement dès lors que $F_X(.)$ est connue, et d'autre part la loi des minima se déduit très simplement des résultats obtenus sur les maxima.

[2] Standard and Poor's : Annual Global Corporate Default Study: Corporate Defaults Poised to Rise in 2005", published January 2005

On considère dans la suite l'échantillon ordonné $X_{(1)} \le \ldots \le X_{(n)}$; afin d'alléger les notations on notera parfois $X_{k,n} = X_{(n-k+1)}$, de sorte que le sous-échantillon des *k* valeurs les plus élevées s'écrive : $X_{k,n} \le \ldots \le X_{1,n}$. On notera dans la suite $S(x) = 1 - F(x)$ la fonction de survie associée à *F*.

2.1. Distribution des valeurs extrêmes

Les principaux résultats de la théorie des valeurs extrêmes sont rappelés ici sans démonstration. Le lecteur intéressé par de plus amples détails et par les démonstrations pourra se référer à EMBRECHTS et al. [1997].

2.1.1. Résultats principaux

Le résultat fondamental de la théorie des valeurs extrêmes est que, sous des conditions de régularité très générales, la loi des extrêmes appartient à une famille composée de trois lois de probabilités. Le théorème de Gnedenko (1943) affirme en effet que si $(X_1, \ldots, X_n)$ est une suite de variables aléatoires iid, et s'il existe 2 suites normalisantes de réels $\mu_n \in \Re, \sigma_n > 0$ et une fonction de distribution H non dégénérée tels que $\frac{M_n - \mu_n}{\sigma_n}$ converge en loi vers H alors H doit être une distribution standard des valeurs extrêmes, qui est nécessairement d'une des trois formes suivantes :

Tab. 1. Distributions des extrêmes

Fréchet	*Weibull*	*Gumbel*
$\Phi_\alpha(x) = \begin{cases} 0, & x \le 0 \\ \exp(-x^{-\alpha}), & x > 0 \end{cases}$	$\Psi_\alpha(x) = \begin{cases} \exp\{-(-x)^\alpha\}, & x \le 0 \\ 1, & x > 0 \end{cases}$	$\Lambda(x) = \exp(-e^{-x}),\ x \in R$

On dit alors que *F* est dans le domaine d'attraction maximum de *H* (que l'on note $DA(H)$ ou $DAM(H)$). Ces trois lois caractérisent chacune un comportement spécifique des valeurs extrêmes. On les distingue au travers de deux caractéristiques :

- ✓ la vitesse de décroissance de la probabilité des grandes valeurs. Si la décroissance est rapide, de type exponentielle, alors la loi des extrêmes est la loi de Gumbel. Si au contraire, elle est plus lente, de type puissance, on se réfère au second critère :
- ✓ l'étendue de variation des données. Si celle-ci est bornée, alors la loi des extrêmes est la loi de Weibull, si ce n'est pas le cas, c'est la loi de Fréchet.

En suivant les notations employées par EMBRECHTS et al. [1997], on a :

Tab. 2. Paramètres des distributions des extrêmes

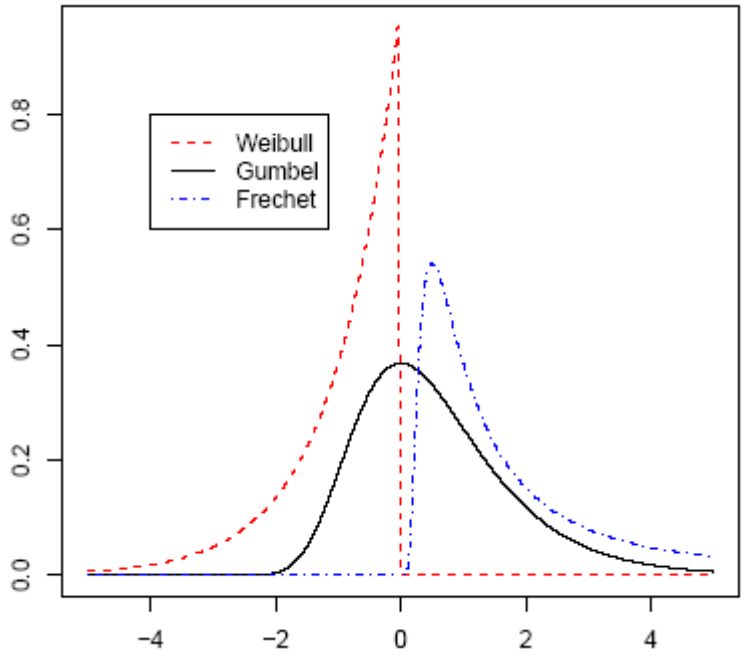
Fréchet	*Weibull*	*Gumbel*
$\xi = \alpha^{-1} > 0$ $\mu_n = 0$ $\sigma_n = F^{-1}\left(1-\frac{1}{n}\right)$	$\xi = -\alpha^{-1} < 0$ $\mu_n = \omega(F)$ $\sigma_n = \omega(F) - F^{-1}\left(1-\frac{1}{n}\right)$	$\xi = 0$ $\mu_n = F^{-1}\left(1-\frac{1}{n}\right)$ $\sigma_n = \frac{1}{S(\mu_n)} \int_{\mu_n}^{\infty} S(t)\,dt$

où le paramètre ξ est appelé l'indice de queue. Plus cet indice est élevé en valeur absolue, plus le poids des extrêmes dans la distribution initiale est important. On parle alors de distribution à « queue épaisse ». Plus généralement on peut montrer que la loi de Gumbel est la loi des extrêmes pour les lois normales et log-normales, que la loi de Fréchet est associée aux distributions de Cauchy, Pareto et Student ; une loi uniforme a pour loi des extrêmes la loi de Weibull. JANSSEN et de VRIES [1991] ont montré que dans le cas de données modélisées par un processus ARCH, la loi limite pour les extrêmes était la loi de Fréchet. Cette dernière loi apparaît donc comme étant *a priori* assez naturelle dans un contexte financier ou assurantiel.

En introduisant le point terminal de la distribution défini par $\omega(F) = \mathbf{sup}\{x \,/\, F(x) < 1\}$, on peut remarquer que pour $F \in DA(Fréchet)$ on a $\omega(F) = +\infty$, qu'à l'opposé pour toute loi de $DA(Weibull)$ $\omega(F) < \infty$ et que pour le domaine d'attraction de Gumbel, on peut rencontrer les deux situations. La forme générale des densités de ces lois est présentée sur le graphique ci-dessous[3] :

[3] Voir BENGIO et CARREAU [2004].

Fig. 3 - **Distributions des extrêmes**



2.1.2. La distribution GEV

La représentation de Jenkinson-Von Mises fournit une caractérisation synthétique des lois extrêmes : la distribution généralisée des valeurs extrêmes (*Generalized Extreme Value*, ou GEV). La fonction de répartition GEV a la forme suivante :

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

μ est un paramètre de localisation, il est directement lié à la valeur la plus probable de la loi, il indique donc approximativement où se trouve le cœur de la distribution. σ est un paramètre de dispersion, il indique l'étalement des extrêmes. ξ est l'indice de queue, déjà introduit.

La loi initiale *F* influe sur la distribution limite *H* au travers de la forme de ses queues de distribution, qui détermine les paramètres μ et σ. Le résultat de convergence présenté en 2.1.1 ci-dessus peut alors se formuler de la manière suivante : il existe 2 suites normalisantes de réels μ_n et $\sigma_n > 0$ telles que :

$$\lim_{n\to+\infty} F^n(\mu_n + \sigma_n x) = H_\xi(x)$$

avec $H_\xi = H_{\xi,0,0}$:

$$H_\xi(x) = \begin{cases} \exp\left\{-(1+\xi x)^{-\frac{1}{\xi}}\right\} & \text{si } 1+\xi x > 0, \xi \neq 0 \\ \exp\left\{-\exp(-x)\right\} & \text{si } \xi = 0 \end{cases}$$

Une fois la forme de la loi des extrêmes établie, il existe plusieurs méthodes pour estimer les paramètres de cette loi, et plus particulièrement le paramètre de queue ξ. Les approches paramétrique et non paramétrique sont possibles. Cette dernière utilise l'ensemble des données initiales, et non seulement les maxima ou les minima, pour estimer ξ alors que la méthode paramétrique consiste à estimer les trois paramètres de la fonction *H* par le maximum de vraisemblance sur les séries des maxima et des minima. Ces différentes approches sont décrites ci-après.

Exemple : Loi exponentielle

Dans le cas d'une loi exponentielle de paramètre $\lambda > 0$ on a $F(x) = 1 - e^{-\lambda x}$ d'où l'on déduit que $P\left(\lambda X_{(n)} - \ln(n) \leq x\right) = P\left(X_1 \leq \frac{x + \ln(n)}{\lambda}\right)^n = \left(1 - \frac{e^{-x}}{n}\right)^n$ et donc :

$$\lim_{n\to\infty} P\left(\lambda X_{(n)} - \ln(n) \leq x\right) = \exp\left(-\exp(-x)\right)$$

ce qui montre que le domaine d'attraction de la loi exponentielle est le domaine de Gumbel.

2.1.3. L'identification du type de valeurs extrêmes

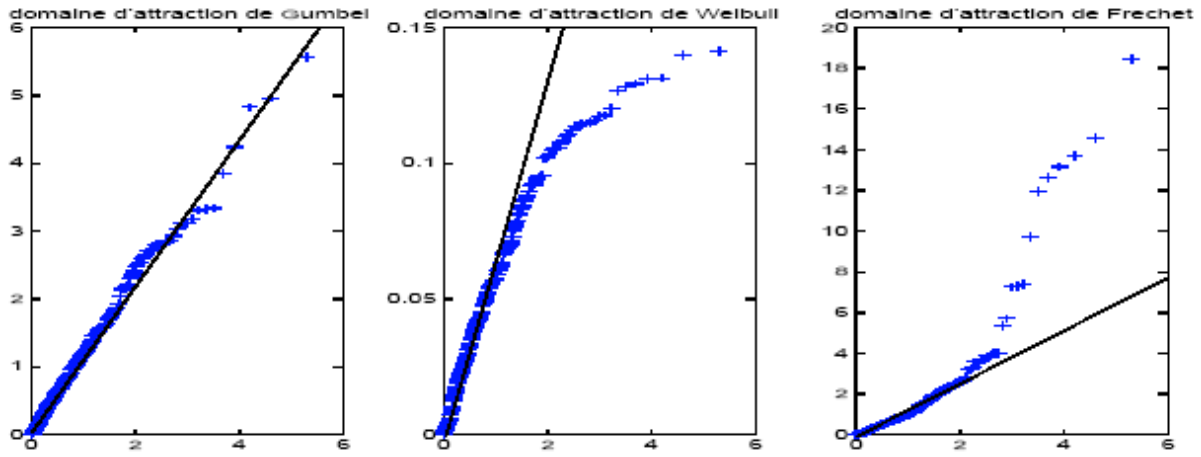
En pratique, on souhaite déterminer, pour un échantillon donné, la forme de la queue de distribution, et, à tout le moins, avoir une idée du domaine d'attraction de la distribution. La méthode des graphiques quantile-quantile permet de fournir une indication de l'appartenance probable à un des domaines d'attraction possibles.

Son principe est de représenter graphiquement les quantiles de la distribution exponentielle[4] $\left(-\ln\left(\frac{i}{k}\right) / 1 \leq i \leq k\right)$ contre les excès ordonnés $\left(X_{(n-k+i)} - X_{(n-k)} / 1 \leq i \leq k\right)$; on obtient une représentation ayant l'allure suivante[5] :

[4] Associée au domaine d'attraction de Gumbel.

[5] Les graphes sont construits avec des échantillons issus respectivement des lois normale, Béta et Student ; l'échantillon utilisé est de taille 1000 avec 200 excès ; ils sont repris de GARRIDO [2002].

Fig. 4 - **Utilisation des QQ-Plots**



Si la loi dont sont issues les données appartient à $DA(Gumbel)$ alors les points du graphique sont approximativement alignés et on peut déterminer les paramètres d'échelle et de position de la loi par régression linéaire. Si en revanche les points du graphique ont tendance à se disperser vers le haut, les données sont vraisemblablement issues d'une loi de $DA(Fréchet)$, et inversement de $DA(Weibull)$ si la dispersion est orientée vers le bas.

2.2. Les estimateurs non paramétriques de l'indice de queue

On présente ici trois estimateurs différents, tous basés sur la statistique d'ordre $X_{k,n} \leq ... \leq X_{1,n}$, obtenue à partir de la série initiale en considérant les *k* valeurs les plus grandes (ou les plus petites). *k* dépend *a priori* de *n*, même si on ne le mentionnera pas dans la notation : l'idée est d'avoir $k \to \infty$ lorsque $n \to \infty$, mais sans prendre « trop » de valeurs de l'échantillon, ce qui conduit à imposer $\frac{k}{n} \to 0$.

Incidemment, cela implique que se posera la question du choix optimal de *k*. En effet, il est indispensable de calculer ces estimateurs sur les queues de distribution. Choisir un *k* trop élevé engendre le risque de prendre en compte des valeurs qui ne sont pas extrêmes, à l'inverse, un sous-échantillon trop petit ne permet pas aux estimateurs d'atteindre leur niveau de stabilité. Ce point délicat est abordé en 2.4 ci-dessous.

Enfin, on retiendra que l'approche non paramétrique n'est envisageable que si l'on dispose d'un nombre important d'observations : dans le cas où les échantillons sont de petite taille, on se tournera vers l'approche paramétrique (y compris l'approche directe décrite en 4.1 ci-dessous).

2.2.1. L'estimateur de Pickands[6]

Il est défini par la statistique :

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

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

$$\sqrt{k}\,\frac{\hat{\xi}_{k,n}^{P} - \xi}{\sigma(\xi)} \to N(0,1)$$

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

$$\sigma(\xi) = \frac{\xi\sqrt{2^{2\xi+1}+1}}{2\left(2^{\xi}-1\right)\ln(2)}$$

2.2.2. L'estimateur de Hill

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

$$\hat{\xi}_{k,n}^{H} = \frac{1}{k-1} \sum_{j=1}^{k-1} \ln\left(\frac{X_{j,n}}{X_{k,n}} \right)$$

Si on choisit $k, n \to +\infty$ de sorte que $\frac{k}{n} \to 0$ alors on peut montrer que $\lim_{k \to +\infty} \hat{\xi}_{k,n}^{H} = \xi$ et l'estimateur de Hill est de plus asymptotiquement normal :

$$\sqrt{k}\,\frac{\hat{\xi}_{k,n}^{H} - \xi}{\xi} \to N(0,1)$$

la convergence étant en loi. Cet estimateur est l'estimateur du maximum de vraisemblance dans le cas particulier du modèle $S(x) = 1 - F(x) = Cx^{-1/\xi}$; on reconnaît ici une distribution de Pareto d'indice $\alpha = \frac{1}{\xi}$. Dans le cas général du domaine de Fréchet, la

[6] Voir PICKANDS [1975].

fonction de survie est de la forme $S(x) = 1 - F(x) = x^{-1/\xi} L(x)$ avec *L* une fonction à variation lente. Cela induit un biais important sur l'estimateur de Hill, qui est donc en pratique d'un maniement délicat. Dans le cas général, la fonction *L* apparaît comme un paramètre de nuisance de dimension infinie, qui complique l'estimation (*cf.* BERTAIL [2002]).

2.2.3. L'estimateur de Hill Alternatif[7]

Même si l'estimateur de Hill est bien moins volatil que celui de Pickand (quand on le représente en fonction du nombre d'observations considérées), il manque de précision et sa phase de stabilisation, qui devrait correspondre à la valeur de l'indice de queue, est parfois difficile à identifier. RESNICK et STARICA [1997] ont mis au point une méthode très simple pour effectuer un « agrandissement » du démarrage du graphe de l'estimateur de Hill. Pour ce faire, l'idée est de changer d'échelle sur l'axe des abscisses pour passer en échelle logarithmique. Ainsi, la partie de courbe correspondant aux premières valeurs extrémales est agrandie et plus précise, par opposition à la partie de courbe correspondant aux valeurs qui ne font probablement pas partie de la queue de la distribution. En pratique, au lieu de représenter graphiquement $\left\{ k, \hat{\xi}_{k,n}^{H} \right\}$, on trace $\left\{ k, \hat{\xi}_{n^{\theta},n}^{H} \right\}$ avec $0 \le \theta \le 1$. Le calcul du *k* optimal qui détermine la séparation de la queue de distribution et du reste des données est détaillé en 2.4 ci-dessous.

2.2.4. L'estimateur des moments

Cet estimateur est proposé dans DEKKERS et al. [1989]. Il est défini par la statistique :

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

avec $\hat{\xi}_{k,n}^{(i)} = \left(\frac{1}{k}\sum_{j=1}^{k} \ln\left(\frac{X_{j,n}}{X_{k-1,n}}\right)\right)^{i}$. Cet estimateur est convergent et asymptotiquement gaussien :

$$\sqrt{k}\,\frac{\hat{\xi}_{k,n}^{H} - \xi}{\sqrt{1+\xi^2}} \to N(0,1)$$

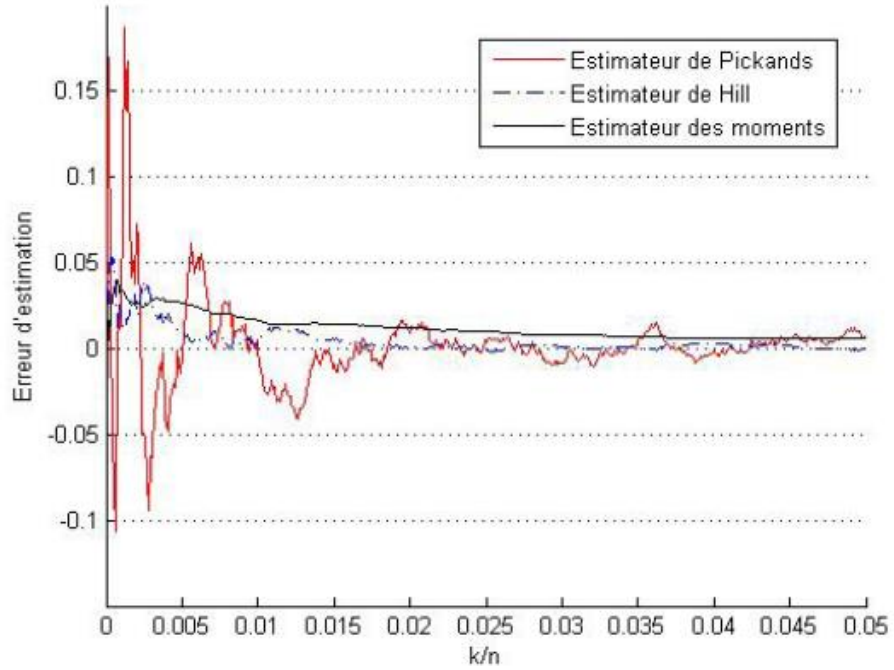
[7] *Cf.* RESNICK et STARICA [1997].

Comme pour l'estimateur de Hill, il n'est utilisable que pour le domaine d'attraction de Fréchet $(\xi > 0)$.

2.2.5. Comparaison des différents estimateurs

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

Fig. 5 - **Comparaison de l'erreur d'estimation en fonction de la méthode**



On note en particulier que l'estimateur de Pickands « paye » sa généralité par une moindre efficacité dans l'estimation de l'indice de queue. On observe la suprématie de l'estimateur de Hill sur ceux de Pickands et de Dekkers-Einmahl-de Haan. Par ailleurs, pour $k < 0{,}02 \times n$, l'estimateur de Hill est relativement volatile. On serait donc amené à utiliser de l'ordre de 2,5 % des données les plus extrêmes pour estimer l'épaisseur de la queue de distribution.

2.3. Un estimateur paramétrique de l'indice de queue

L'approche paramétrique consiste alors, à partir de la représentation de Jenkinson-von Mises introduite en 2.1 ci-dessus, à exprimer la densité de la distribution des extrêmes :

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

Cela conduit à écrire la log-vraisemblance de $h_{\xi,\mu,\psi}(x)$ que l'on peut ensuite maximiser :

$$\mathbf{ln}\, L(\xi,\mu,\psi) = \sum_{t=1}^{n} \mathbf{ln}\left(h_{\xi,\mu,\psi}(x_t) 1_{\left\{1+\xi\frac{x-\mu}{\psi}>0\right\}}(x_t) \right)$$

sur la série des maxima et des minima. On obtient ainsi les trois paramètres de la distribution des extrêmes. L'estimation des paramètres, et notamment du paramètre de queue ξ peut s'effectuer par une approche graphique. Elle consiste à utiliser une représentation quantile-quantile[8], sur l'inverse de la distribution GEV. Cette représentation n'utilise pas les paramètres de localisation et de dispersion, mais utilise en revanche le paramètre de queue. Ce graphique quantile-quantile se construit à partir de l'inverse de la distribution GEV centrée réduite $H_{\xi,0,1}(x)$, et il est défini par :

$$\left\{ \left(X_{k,n}, \frac{1}{\xi}\left(\mathbf{exp}\left(-\xi\, \mathbf{ln}\left(-\mathbf{ln}\left(\frac{n-k+1}{n+1} \right) \right) \right) - 1 \right) \right) : k = 1, \ldots, n \right\}$$

Ce graphique permet, après l'estimation des paramètres de la loi GEV par le maximum de vraisemblance, de vérifier le pouvoir explicatif de cette loi centrée réduite, mais avec le paramètre de queue estimé, sur la distribution empirique des maxima ou des minima. La non linéarité du QQ Plot ainsi tracé ne réside plus que dans les paramètres de localisation et de dispersion qui peuvent ensuite être approchés aisément par le calcul de l'ordonnée à l'origine et de la pente de la courbe. Ce recoupement permet de valider éventuellement les résultats des estimations obtenues.

2.4. Séparation des données standards et des données de queue[9]

On a $\hat{\xi}^H_{k,n} = \frac{1}{k-1}\sum_{j=1}^{k-1}\left(\ln X_{j,n} - \ln X_{k,n}\right)$ avec $X_{k,n} \leq \ldots \leq X_{1,n}$, l'échantillon des valeurs ordonnées. Le *k* optimal recherché doit être construit de façon qu'il minimise la moyenne du carré des erreurs :

$$k^* = \mathbf{arg\,min}_k\, E\left[\left(\xi^H_{k,n} - \xi \right)^2 \right]$$

Le premier point important est le choix de ξ dans le critère à optimiser, il représente un indice de référence à partir duquel les écarts vont être minimisés. Pour s'assurer que cette valeur ne décale le *k* optimal vers des valeurs non extrêmes, il faut calculer le ξ de référence avec des données qui sont dans la queue de la distribution. DANIELSSON et de VRIES [1997] proposent par exemple de calculer ξ avec 1 % des données les plus grandes (ou les plus petites).

[8] « QQ plot »
[9] La méthode présentée ici a été proposée par HALL [1990] et DANIELSSON et de VRIES [1997].

On doit aussi déterminer à quel moment stopper les calculs. On peut par exemple fixer un seuil consistant à conserver *x* % des données comme échantillon minimal (avec par exemple $x = 10\%$). La taille de l'échantillon ainsi obtenu est notée *N*.

Il reste ensuite à calculer l'espérance des écarts quadratiques. Pour cela, on peut utiliser une approche de type « bootstrap » : on choisit un sous-échantillon de taille inférieure à *N*, noté N_1, et on tire au hasard (avec remise) *Z* sous-échantillons de taille N_1 dans *N*. Soit $i = 1,..,Z$ l'indice du *i*ième sous-échantillon tiré dans *N* ; l'espérance des écarts quadratiques est approximativement égale à : $\frac{1}{Z}\sum_{i=1}^{Z}\left(\xi_{k_1,N_1}^{H,i} - \xi\right)^2$. La minimisation de cette dernière expression en k_1 donne le k_1 optimal. On doit ensuite réajuster ce k_1 à la taille *N* choisie (puisqu'il est calibré sur N_1) ce qui conduit après quelques manipulations à :

$$k^* = k_1^*\left(\frac{N}{N_1}\right)^{\frac{2\beta}{2\beta+\frac{1}{\xi}}} \qquad \text{avec } \hat{\beta} = \frac{\sqrt{\Delta}-1}{\xi},\ \Delta = \frac{\tilde{m}^{(1)} - \frac{\tilde{m}^{(2)}}{2\tilde{m}^{(1)}}}{\frac{\tilde{m}^{(3)}}{3\tilde{m}^{(2)}} - \frac{\tilde{m}^{(4)}}{4\tilde{m}^{(3)}}}, \text{ et } \tilde{m}^{(i)} = \frac{1}{k-1}\sum_{j=1}^{k-1}\left(\ln\left(\frac{x_{j,n}}{x_{k,n}}\right)\right)^i$$

L'estimateur de Hill optimal est alors $\hat{\xi} = \xi_{k^*-1,N}^H$.

3. L'estimation des distributions au-dessus d'un seuil

La mesure et la gestion des risques exigent qu'on s'intéresse aux queues de distribution et notamment aux pertes que l'on peut subir avec des probabilités faibles mais non négligeables.

Cette problématique est traitée par l'introduction du concept de valeur à risque (« Value at Risk »). La *VaR* au seuil α est le montant de la perte maximale avec une probabilité α. La *VaR* calculée pour répondre aux exigences de Solvabilité 2 traite d'événements « rares », au-delà du seuil 99,5 %. Techniquement, la notion de *VaR* au seuil α est équivalente à la notion probabiliste de quantile d'ordre α ou fractile d'ordre $1-\alpha$ de la distribution.

De plus, une fois le seuil atteint, on est conduit à s'intéresser à l'espérance de la perte, qui conduit à introduire la notion de *TVaR*.

La théorie des valeurs extrêmes fournit des outils pour estimer la loi des observations au-dessus d'un seuil bien choisi et calculer les quantiles extrêmes à l'aide de cette loi estimée. Dans la suite on s'intéresse au comportement (statistique) des observations au-delà d'un seuil *u* fixé (*a priori u* sera choisi « grand »).

3.1. Le nombre d'excès au-delà d'un seuil

Avant de considérer la loi du phénomène observé au-delà du seuil *u*, on peut s'intéresser au nombre de fois où, dans un échantillon de taille *n*, le seuil est dépassé. On montre (*cf.* DENUIT et CHARPENTIER [2005]) que si le seuil u_n est choisi de sorte que

$\lim_{n\to\infty} n\left(1-F\left(u_n\right)\right) = \lim_{n\to\infty} n \times P\left(X > u_n\right) = \lambda$ alors le nombre de dépassement du seuil, $N\left(u_n\right)$ a pour loi asymptotique une loi de Poisson de paramètre λ.

Cette propriété est la conséquence directe du fait que $N\left(u_n\right)$ est distribué selon une loi binomiale de paramètres *n* et $p_n = 1-F\left(u_n\right)$ et que la distribution binomiale $\mathrm{B}\left(n, \frac{\lambda}{n}\right)$ converge en loi vers $\mathrm{P}\left(\lambda\right)$ lorsque $n \to \infty$. Incidemment, ce résultat justifie l'apparition naturelle de la loi de Poisson pour modéliser des événements « rares ».

3.2. La distribution de Pareto généralisée (GPD)

Lorsqu'on s'intéresse aux valeurs extrêmes d'un phénomène, il est naturel de considérer la loi des observations « au-delà d'un seuil » ; on est donc conduit à considérer la distribution conditionnelle :

$$F_u\left(x\right) = P\left(X - u \le x \,\middle|\, X > u\right) = 1 - \frac{S\left(x+u\right)}{S\left(u\right)},\ x \ge 0$$

où $S\left(x\right) = 1 - F\left(x\right)$ est la fonction de survie. Cette approche s'appelle la « méthode POT[10] » et a été introduite par DE HAAN et ROOTZEN [1993] ; on peut noter l'analogie entre $F_u\left(x\right)$ et le taux de décès à l'âge $u+x$ sachant que l'individu est vivant à l'âge *u*, ${}_xq_u$. Pour une classe assez large de lois de probabilité[11], le théorème de Pickands[12] affirme qu'il existe une constante de normalisation positive $a\left(u\right)$ telle que, pour $x \ge 0$:

$$\lim_{u\to\omega\left(F\right)} F_u\left(xa\left(u\right)\right) = G_{\xi,\beta}\left(x\right)$$

avec :

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

et $\beta > 0$. Cette distribution limite est la distribution de Pareto généralisée (GPD). La loi des valeurs au-delà d'un seuil admet ainsi une distribution limite. L'égalité ci-dessus peut

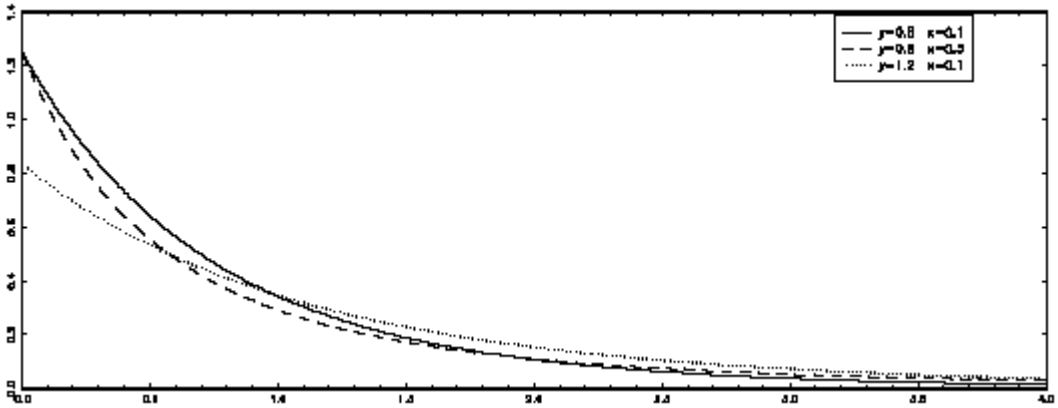
[10] Pour « Peak Over the Threshold »
[11] En pratique, les distributions appartenant à l'un des trois domaines d'attraction présentés en 2.1.
[12] PICKANDS [1975].

être reformulée de la manière suivante : lorsque *u* est grand (c'est à dire proche du point terminal $\omega(F)$), alors on peut trouver $\beta(u)$ tel que :

$$F_u(x) \approx G_{\xi,\beta(u)}(x).$$

De plus, la fonction $\beta(u)$ est unique à une équivalence près. Dans le cas où $F \in DA(Gumbel)$, on peut montrer que l'on peut prendre $\beta(u) = \frac{1}{S(u)} \int_u^{\omega(F)} S(t)\,dt$. Ce résultat peut s'exprimer en disant que la loi des excès au-delà d'un seuil converge vers une loi GPD lorsque la taille de l'échantillon, et donc le seuil, tendent vers l'infini. L'indice de queue ξ de cette distribution est le même que pour la distribution GEV. Le paramètre β est un indicateur de la taille de la queue à une distance finie. La distribution GPD a l'allure suivante[13] :

Fig. 6 - **Distribution GPD**



3.2.1. Propriétés

On montre (cf. DENUIT et CHARPENTIER [2005]) que la loi de Pareto généralisée peut être vue comme une log-gamma ou encore, dans le cas où $\xi > 0$, comme un mélange de lois exponentielles dont le paramètre suit une loi gamma. Les résultats suivants[14] sont énoncés pour une variable aléatoire *X* distribuée selon une loi GPD de paramètre (ξ, β).

Propriété 1. Si $\xi < 1$, on a

$$\mathbf{E}\left(1+\frac{\xi}{\beta}X\right)^{-r} = \frac{1}{1+\xi}, \text{ pour } r > -1/\xi,$$

[13] Ce graphique et les suivants sont repris de JONDEAU et ROCKINGER [1999].
[14] Voir EMBRECHTS et al. [1997].

$$\mathbf{E}\left[\ln\left(1+\frac{\xi}{\beta}X\right)\right]^k = \xi^k k!, \text{ pour } k \in \mathbb{N},$$

$$\mathbf{E}\left(X\left(1-G_{\xi,\beta}(X)\right)^r\right) = \frac{\beta}{(r+1-\xi)(r+1)}, \text{ pour } (r+1)/\xi > 0.$$

Dès lors que l'on se place dans le contexte de distributions à queues épaisses, l'existence des moments aux différents ordres n'est plus assurée. Précisément, on a :

Propriété 2. La variable aléatoire *Y* admet des moments jusqu'à l'ordre $\left[\xi^{-1}\right]$ et l'on a

$$\mathrm{E}\left[X^r\right] = \frac{\beta^r \Gamma\left(\xi^{-1}-r\right)}{\xi^{r+1}\Gamma\left(\xi^{-1}+1\right)}, \text{ pour } r \le \left[\xi^{-1}\right].$$

En particulier $\mathbf{E}(X) = \frac{\beta}{1-\xi}$.

Propriété 3. (Stabilité) La variable aléatoire $X_u = [X-u \mid X>u]$ est distribuée selon une GPD de paramètre $(\xi, \beta+\xi u)$. On en déduit que si $\xi<1$, alors pour tout $u<\omega_F$,

$$\mathbf{E}[X-u \mid X>u] = \frac{\beta+\xi u}{1-\xi}, \text{ pour } \beta+u\xi>0.$$

On rappelle que ω_F est la borne supérieure du support de *X*, soit $\omega_F = \mathbf{sup}\{y \in \mathbb{R}, F(y)<1\}$.

3.2.2. Estimation des paramètres

L'estimation des paramètres peut être effectuée par le maximum de vraisemblance. La densité de la distribution GPD s'écrit en effet :

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

et on en déduit la log-vraisemblance :

$$\ln L(\xi, \beta; X_1, \ldots, X_n) = -n\ln\beta - \left(1+\frac{1}{\xi}\right)\sum_{i=1}^{n}\ln\left(1+\frac{\xi}{\beta}X_i\right).$$

En utilisant la reparamétrisation $\tau = \xi/\beta$, l'annulation des dérivées partielles de la log-vraisemblance conduit au système

$$\begin{cases} \xi = \frac{1}{n}\sum_{i=1}^{n} \ln(1+\tau X_i) =: \hat{\xi}(\tau), \\ \frac{1}{\tau} = \frac{1}{n}\left(\frac{1}{\xi}+1\right)\sum_{i=1}^{n} \frac{X_i}{1+\tau X_i}. \end{cases}$$

L'estimateur du maximum de vraisemblance de (ξ,τ) est $\left(\hat{\xi} = \hat{\xi}(\hat{\tau}), \hat{\tau}\right)$ où $\hat{\tau}$ est solution de

$$\frac{1}{\tau} = \frac{1}{n}\left(\frac{1}{\hat{\xi}(\tau)}+1\right)\sum_{i=1}^{n} \frac{X_i}{1+\tau X_i}.$$

Cette dernière équation se résout numériquement de manière itérative pour autant que l'on dispose d'une valeur initiale τ_0 pas trop éloigné de τ. En pratique, cette valeur initiale pourra être obtenue par la méthode des moments (pour autant que ceux-ci existent jusqu'à l'ordre 2) ou par la méthode des quantiles. Lorsque $\xi > -1/2$, HOSKING et WALLIS [1987] ont montré la normalité asymptotique des estimateurs du maximum de vraisemblance :

$$n^{1/2}\left(\hat{\xi}_n - \xi, \frac{\hat{\beta}_n}{\beta} - 1\right) \xrightarrow[n\to\infty]{\mathcal{L}} \mathcal{N}\left[0, (1+\xi)\begin{pmatrix} 1+\xi & -1 \\ -1 & 2 \end{pmatrix}\right].$$

Ce résultat permet en particulier de calculer les erreurs approximatives d'estimation commises par les estimateurs du maximum de vraisemblance. On peut aussi remarquer que la fonction quantile a une expression explicite, lorsque $\xi > 0$:

$$x_p = G^{-1}_{\xi,\beta}(p) = \frac{\beta}{\xi}\left((1-p)^{-\xi} - 1\right).$$

Cela permet de proposer un estimateur de ξ par la méthode des quantiles en remarquant que :

$$\frac{x_{p_1}}{x_{p_2}} = \frac{(1-p_1)^{-\xi} - 1}{(1-p_2)^{-\xi} - 1}.$$

Une fois l'estimation terminée, on peut vérifier graphiquement la pertinence des estimations en comparant la distribution GPD estimée avec la distribution empiriquement observée des observations au-delà du seuil.

D'un point de vue pratique l'estimation des paramètres (ξ,β) pour un seuil *u* fixé conduit à l'estimation suivante de la distribution au-delà du seuil : $\hat{F}_u(x) \approx G_{\hat{\xi},\hat{\beta}}(x)$. On notera que dans l'approximation proposée par la formule $F_u(x) \approx G_{\xi,\beta(u)}(x)$, le paramètre β dépend du seuil *u* choisi.

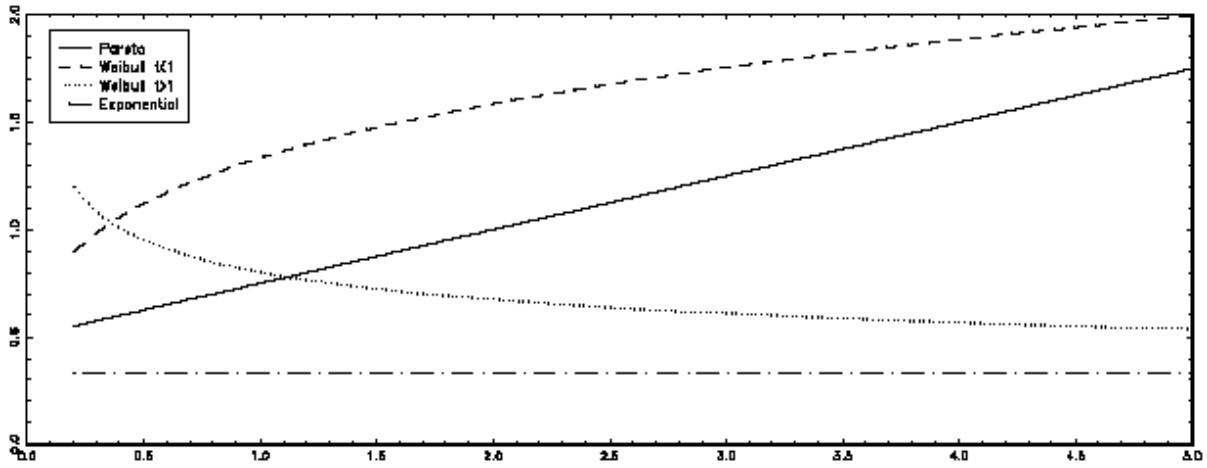
3.3. L'espérance au-delà d'un seuil

L'espérance au-delà du seuil *u* (ou « moyenne en excédent ») est définie pour une distribution *F* par :

$$e(u) = E\left(X - u \middle| X > u\right)$$

Il s'agit donc de l'espérance de la distribution conditionnelle introduite en 3.2 ci-dessus. Cette fonction de *u* n'est autre que l'espérance de vie résiduelle classique dans les modèles de durée ; elle s'exprime à l'aide de la fonction de survie de *F*. Plus les queues de distribution sont épaisses, plus cette fonction a tendance à tendre vite vers l'infini. Le graphique ci-dessous présente quelques cas de distributions usuelles :

Fig. 7 - **Espérance de vie résiduelle**



Cette notion est directement liée à la *Conditional Tail Expectation* (*CTE*), qui est la valeur moyenne des pertes au-delà de la *VaR*[15]:

$$CTE(X,\alpha) = \mathbf{E}\left[X \mid X > VaR(X,\alpha)\right]$$

Intuitivement, la *CTE* représente le « risque moyen » dans les pires 1-α cas. On a donc :

$$CTE(X,\alpha) = VaR(X,\alpha) + e\left(VaR(X,\alpha)\right)$$

On rappelle que lorsque la distribution est continue, la *CTE* coïncide avec la *TVaR* :

$$TVaR(X,\alpha) = \frac{1}{1-\alpha}\int_{\alpha}^{1} VaR(X,s)\,ds$$

[15] Pour une présentation détaillée de ces notions, voir par exemple DENUIT et CHARPENTIER [2004].

3.3.1. Estimation

En pratique, si N_u est le nombre d'observations au-dessus du seuil *u*, on a l'estimateur empirique suivant de $e(u)$:

$$\hat{e}_1(u) = \frac{1}{N_u} \sum_{j=1}^{n} (x_j - u) 1_{\{x>u\}}(x_j).$$

On peut également proposer comme estimateur :

$$\hat{e}_2(u) = \frac{\hat{\beta}(u)}{1-\hat{\xi}},$$

avec les paramètres (ξ, β) estimés par le maximum de vraisemblance. Enfin, une troisième méthode permet de se ramener à l'estimation du seul paramètre ξ, dans le cas particulier où $\xi > 0$. Pour cela on observe que la fonction de survie de la distribution peut se mettre sous la forme $S(x) = 1 - F(x) = x^{-1/\xi} L(x)$ où *L* est une fonction « à variation lente » à l'infini au sens où $\lim_{t\to\infty} \frac{L(tx)}{L(t)} = 1$.

Avec le seuil *u*, on peut donc écrire que $S_u(x) = \frac{S(x)}{S(u)} = \frac{L(x)}{L(u)} \left(\frac{x}{u}\right)^{-1/\xi}$ et en considérant que le rapport des fonctions à variation lente est proche de 1 on trouve :

$$S_u(x) \approx \left(\frac{x}{u}\right)^{-1/\xi}.$$

Dès lors que $\xi < 1$ on trouve donc que $e_x(u) = \int_0^{+\infty} S_u(u+t)\,dt \approx \frac{\xi u}{1-\xi}$. Il suffit donc de déterminer un estimateur de ξ, par exemple l'estimateur de Hill, pour proposer comme estimateur de l'espérance au-delà de *u* :

$$\hat{e}_3(u) = \frac{\hat{\xi} u}{1-\hat{\xi}}.$$

$\hat{\xi}$ peut être choisi estimé pour un seuil $u_0 \le u$ fixé, ou réestimé pour chaque valeur de *u*, ce qui conduit alors à $\hat{\xi} = \hat{\xi}(u)$.

3.3.2. Exemple : la loi de Pareto

On illustre la mise en œuvre de ces estimateurs dans le cas d'une loi de Pareto de fonction de répartition $F(x) = 1 - \left(\frac{x_m}{x}\right)^{\alpha}$ pour $x \geq x_m$. Dans cette configuration l'approximation GPD est en fait exacte :

$$S_u(x-u) = P(X > x \,|\, X > u) = \left(\frac{x}{u}\right)^{-\alpha} = \left(1 + \frac{\xi}{\beta}(x-u)\right)^{-1/\xi}$$

en posant $\xi = \frac{1}{\alpha}$ et $\beta(u) = \xi u$. Dans ce cas particulier, du fait de l'invariance par changement de seuil de la loi de Pareto, l'estimation est en fait ramenée à l'estimation du seul paramètre $\xi = \frac{1}{\alpha}$.

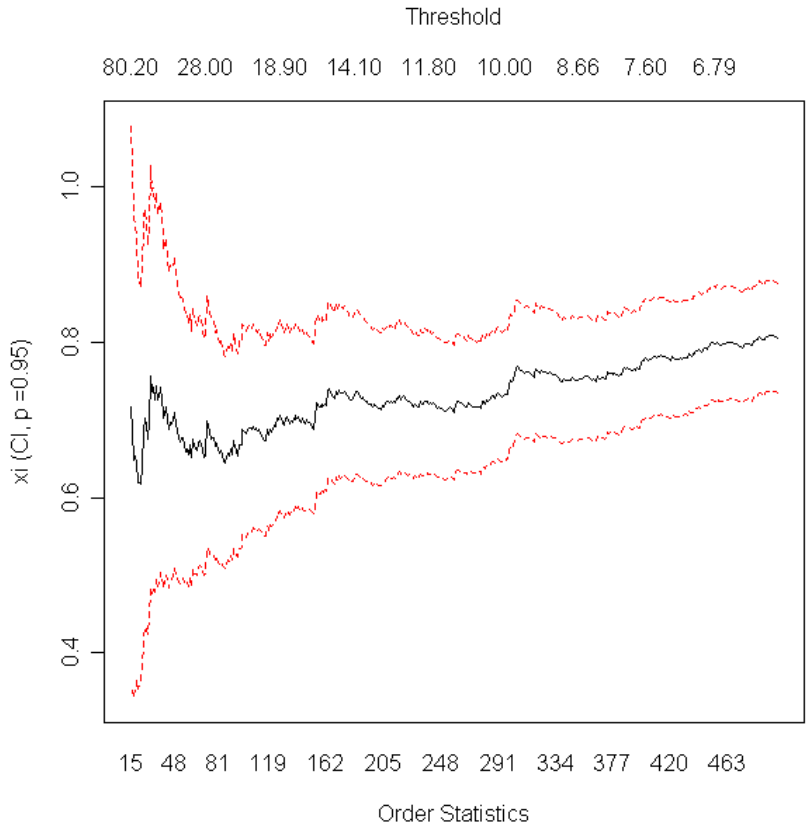
3.3.3. Détermination du seuil

On déduit de la propriété 3 de 3.2.1 que pour un seuil *u* fixé tel que $S_u(x) \approx S_{\xi,\beta(u)}(x)$ et si on considère un $v \geq u$ alors en notant $Y = [X - u \,|\, X > u]$ on a l'égalité en loi $[X - v \,|\, X > v] = [Y - (v-u) \,|\, Y > v - u]$ et comme *Y* suit une loi de Pareto généralisée $(\alpha, \beta(u))$:

$$\mathbf{E}[X - v \,|\, X > v] = \frac{\beta(u) + \xi \times (v-u)}{1-\xi}.$$

Ainsi, l'espérance conditionnelle des excès est une fonction affine de *v*, lorsque *v* est supérieur au seuil de référence *u*. Cela donne un moyen empirique de tester si le seuil *u* choisi pour les calculs est suffisamment élevé : les espérances des excès au-delà des seuils plus élevés doivent être alignée sur une droite de pente $\frac{\xi}{1-\xi}$. Une approche alternative consiste à calculer l'estimateur de Hill de l'indice de queue pour différents seuils et à chercher à partir de quel seuil cet estimateur est approximativement constant :

Fig. 8 - **Identification du seuil**



3.4. Estimation des quantiles élevés

L'estimation des quantiles d'ordre élevé constitue une étape déterminante du mécanisme de détermination du capital économique dans le dispositif « Solvabilité 2 », qui fait explicitement référence à un quantile à 99,5 % de la distribution du résultat. Cette estimation peut être effectuée dans le cadre de la théorie des valeurs extrêmes. L'approche directe à partir de l'ensemble d'un échantillon paramétrique est évoquée à la section 4 ci-dessous.

3.4.1. Approximation GPD

Les historiques de données dont on dispose en général sont beaucoup trop courts pour mettre en évidence les phénomènes exceptionnels qui ne se reproduisent que tous les 30 ou 50 ans. La distribution de Pareto généralisée est un outil utile dans ce contexte puisque, en plus de reproduire fidèlement la forme de la distribution empirique au-dessus d'un seuil, la GPD permet d'estimer les quantiles élevés de la distribution. Cela revient de fait à pouvoir extrapoler les données au-delà de l'échantillon dont on dispose. ROOTZEN et TAJVIDI [1996] ont en effet montré que le quantile à l'ordre p peut être approché, pour un seuil u fixé arbitrairement tel que $u < x_p$, par :

$$\hat{x}_p = u + G^{-1}_{\hat{\xi},\hat{\beta}}\left(\frac{p - F_n(u)}{1 - F_n(u)}\right) = \begin{cases} u + \frac{\hat{\beta}}{\hat{\xi}}\left[\left(\frac{1-p}{1-F_n(u)}\right)^{-\hat{\xi}} - 1\right] \text{ si } \hat{\xi} \neq 0 \\ u - \hat{\beta}\ln\left(\frac{1-p}{1-F_n(u)}\right) \text{ si } \hat{\xi} = 0 \end{cases}$$

On peut noter qu'on a très simplement (avec les notations de 3.1 ci-dessus) $1 - F_n(u) = \frac{N(u)}{n}$; Cette égalité s'obtient en remarquant que la séparation entre le cœur de la distribution et les données extrêmes peut être traduite par la décomposition suivante de la fonction de répartition à l'aide d'un seuil :

$$F(x) = P(X \leq x) = (1 - P(X \leq u))F_u(x-u) + P(X \leq u)$$

Cette égalité exprime simplement la relation entre la fonction de survie et la fonction de survie conditionnelle $S(x) = S(u)S_u(x)$ avec $S(x) = 1 - F(x)$. Le premier terme est associé à la queue de la distribution, le second aux valeurs standards. Cette quantité peut être estimée par :

$$\hat{F}(x) = (1 - F_n(u))G_{\hat{\xi},\hat{\beta}}(x) + F_n(u),$$

$F_n(u)$ étant la fonction de répartition empirique de l'échantillon d'origine au point *u* :

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

En prenant comme cas particulier un seuil $u = X_{(n-k)}$ et en remarquant que $F_n\left(X_{(n-k)}\right) = \frac{k}{n}$, on trouve que pour $p > 1 - \frac{k}{n}$

$$\hat{x}_p = \begin{cases} X_{(n-k)} + \frac{\hat{\beta}}{\hat{\xi}}\left[\left(\frac{n}{k}(1-p)\right)^{-\hat{\xi}} - 1\right] \text{ si } \hat{\xi} \neq 0 \\ X_{(n-k)} - \hat{\beta}\ln\left(\frac{n}{k}(1-p)\right) \text{ si } \hat{\xi} = 0 \end{cases}$$

3.4.2. Estimateur de Hill

On se restreint aux fonctions de répartition du domaine de Fréchet ($\xi > 0$) ; alors la fonction de survie de la distribution peut se mettre sous la forme

$S(x) = 1 - F(x) = x^{-1/\xi} L(x)$ où L est une fonction « à variation lente » à l'infini au sens où $\lim_{t\to\infty} \frac{L(tx)}{L(t)} = 1$.

On peut donc écrire que $\frac{S(x)}{S\left(X_{(n-k)}\right)} = \frac{L(x)}{L\left(X_{(n-k)}\right)} \left(\frac{x}{X_{(n-k)}} \right)^{-1/\xi}$ et en considérant que le rapport des fonctions à variation lente est proche de 1 on trouve :

$$S(x) = S\left(X_{(n-k)}\right) \left(\frac{x}{X_{(n-k)}} \right)^{-1/\xi}.$$

On déduit de cette expression un estimateur de $F(x)$ pour $x > X_{(n-k)}$:

$$\hat{F}(x) = 1 - \frac{k}{n} \left(\frac{x}{X_{(n-k)}} \right)^{-1/\hat{\xi}^H_{k,n}},$$

avec $\hat{\xi}^H_{k,n}$ l'estimateur de Hill de l'indice de queue. Par inversion de cette expression on obtient l'estimateur de Hill de la fonction quantile :

$$\hat{x}_p = X_{(n-k)} \left(\frac{n}{k}(1-p) \right)^{-\hat{\xi}^H_{k,n}}.$$

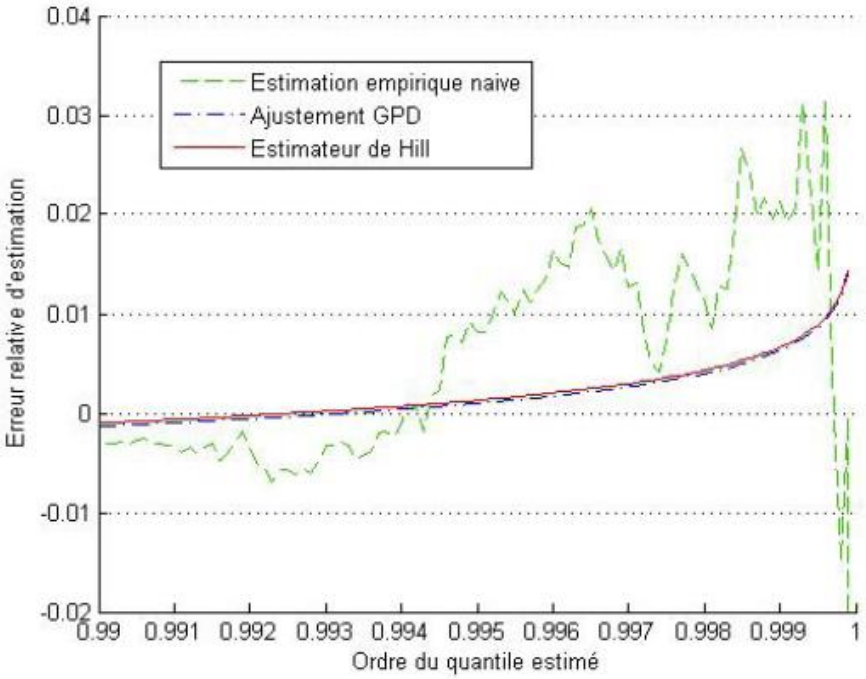
3.4.3. Illustration : cas de la loi de Pareto

Dans le cas d'une loi de Pareto de fonction de répartition $F(x) = 1 - \left(\frac{x_m}{x} \right)^{\alpha}$ pour $x \geq x_m$, dont on a vue que l'approximation GPD est en fait exacte :

$$1 - F_u(x-u) = P(X > x \,|\, X > u) = \left(\frac{x}{u} \right)^{-\alpha} = \left(1 + \frac{\xi}{\beta}(x-u) \right)^{-1/\xi}$$

en posant $\xi = \frac{1}{\alpha}$ et $\beta(u) = \xi u$. On obtient la représentation graphique suivante :

Fig. 9 - **Estimation d'un quantile d'ordre élevé**



NB : l'estimation empirique naïve consiste à utiliser l'estimateur direct de $F^{-1}(p)$:

$$\hat{F}^{-1}(p) = \left([pn] - pn + 1\right) X_{([pn])} + \left(pn - [pn]\right) X_{([pn]+1)}.$$

4. Application : calcul d'une VaR en assurance

On s'intéresse ici au calcul pratique d'une valeur à risque (*VaR*) ; il s'agit donc d'estimer un quantile d'une distribution à partir d'un échantillon de taille *n* donné. On se trouve en pratique confronté à deux situations nécessitant potentiellement des traitements distincts :

- ✓ le calcul d'un quantile de niveau « petit » : VaR à 75 % pour le niveau des provisions en solvabilité 2 ;
- ✓ le calcul d'un quantile de niveau élevé : contrôler la probabilité de ruine à un an pour qu'elle ne dépasse pas 0,05 %.

A priori l'estimation d'un quantile d'ordre peu élevé sera plus aisée que celle d'un quantile d'ordre élevé, du fait de données plus nombreuses : les événements rares ne se produisent pas souvent, ce qui complique leur mesure…

On a examiné à la section 3.4 ci-dessus une méthode ne nécessitant pas d'hypothèse forte sur la loi sous-jacente, mais impliquant d'estimer les paramètres de la loi des extrêmes associée. On pourrait également imaginer d'utiliser l'approche « naïve » consistant à déterminer le quantile empirique au niveau souhaité à partir de la distribution empirique :

$$F_n(u) = \frac{1}{n}\sum_{i=1}^{n} 1_{\{X_i \le u\}}.$$

en posant par exemple :

$$\hat{F}^{-1}(p) = \left(\left[pn\right] - pn + 1\right) X_{(\left[pn\right])} + \left(pn - \left[pn\right]\right) X_{(\left[pn\right]+1)}$$

Cette approche peut raisonnablement être envisagée dans le cas d'un seuil peu élevé, par exemple pour la médiane ou le dernier quartile (conduisant à la provision « à 75 % » prévue par Solvabilité 2). Elle ne peut plus l'être pour un quantile d'ordre élevé, pour lequel elle conduit à un estimateur beaucoup trop instable. On examine ici l'approche directe, envisageable lorsque la loi sous-jacente est connue.

4.1. L'approche paramétrique directe

4.1.1. Estimation ponctuelle

On fait l'hypothèse que la loi du phénomène sous-jacent appartient à une famille paramétrique (normale, log-normale, Pareto, Benktander, etc.). L'estimation d'un quantile d'ordre quelconque peut alors être abordée de la manière suivante :

- ✓ estimation des paramètres (par maximum de vraisemblance ou autre méthode) ;
- ✓ inversion de la fonction de répartition (détermination de la fonction quantile) ;
- ✓ estimation de la *VaR* par $\hat{VaR}(\alpha) = F_{\hat{\theta}}^{-1}(\alpha)$.

Cette approche implique que l'on soit capable d'inverser la fonction de répartition, que ce soit de manière exacte (Pareto par exemple) ou numérique (dans le cas d'une loi log-normale). On rappelle que la fonction quantile existe systématiquement (même lorsque la fonction de répartition n'est pas inversible) en posant, par définition :

$$F^{-1}(y) = \mathbf{inf}\left\{x : F(x) \ge y\right\}$$

qui définit une fonction continue à gauche. L'estimateur ainsi obtenu de la *VaR* est en général biaisé ; l'obtention d'intervalles de confiance passe en général par l'utilisation du « bootstrap », dont on rappelle brièvement le principe ci-après.

4.1.2. Rappels sur les méthodes bootstrap

La méthode du bootstrap proposée initialement par EFRON [1979] est devenue d'application courante en assurance (cf. PARTRAT et JAL [2002] et VERALL et ENGLAND [1999]). Plus généralement, le lecteur intéressé trouvera une revue des applications de la méthode bootstrap en économétrie dans HOROWITZ [1997].

Le principe de la méthode consiste à remarquer que pour un échantillon de taille suffisante, la fonction de répartition de la loi sous-jacente peut être approchée par la

fonction de répartition empirique $F_n(x) = \frac{1}{n}\sum_{i=1}^{n} 1_{\{X_n \le x\}}$. Évaluer des statistiques par simulation se ramène alors à générer des échantillons à l'aide de la distribution empirique. Or un tirage dans la distribution empirique s'obtient simplement par un tirage avec remise des *n* valeurs dans l'échantillon initial. On obtient ainsi au plus n^n échantillons « boostrapés » à partir desquels on va calculer les estimateurs empiriques des statistiques d'intérêt.

En pratique ce principe très général conduit à décliner des méthodes bootstrap adaptées à différents contextes : séries chronologiques, modèles linéaires, modèles de régression non linéaires, etc. On en présente ci-après le principe dans le cas simple de la construction d'un intervalle de confiance pour un estimateur de tendance centrale.

Pour évaluer une fonctionnelle de *F* de la forme $I(g) = \int g(x)dF(x)$, l'approche par simulation classique suggère de calculer

$$I_n(g) = \int g(x)dF_n(x) = \frac{1}{n}\sum_{i=1}^{n} g(X_i),$$

qui constitue une approximation de $I(g)$. La théorie asymptotique fournie des informations sur la loi de la statistique $I_n(g)$: le théorème central limite permet en effet de prouver que la loi asymptotique est normale et d'en déduire, par exemple, des intervalles de confiance de la forme

$$J_\alpha = \left[I_n(g) - \frac{\sigma_g}{\sqrt{n}}\varphi^{-1}\left(1-\frac{\alpha}{2}\right), I_n(g) + \frac{\sigma_g}{\sqrt{n}}\varphi^{-1}\left(1-\frac{\alpha}{2}\right)\right],$$

avec $\sigma_g^2 = \mathrm{Var}(g(X))$ qu'il s'agit d'estimer. Toutefois dans certaines situations l'approximation asymptotique n'est pas suffisamment précise, et la méthode bootstrap fournit une approche alternative pour obtenir des informations sur la loi de $I_n(g)$. Cette technique consiste à considérer des réalisations $I_n^b(g)$, pour $b = 1,\ldots,B$ où $B \le n^n$ obtenues en remplaçant l'échantillon initial $(X_1,\ldots,X_n)$ par les échantillons « bootstrapés » $(X_1^b,\ldots,X_n^b)$ obtenus en effectuant des tirages avec remise dans $(X_1,\ldots,X_n)$. On dit que *B* est la taille de l'échantillon bootstrap. En effet, évaluer des statistiques par simulation se ramène alors à générer des échantillons à l'aide de la distribution empirique. Or un tirage dans la distribution empirique s'obtient simplement par un tirage avec remise des *n* valeurs dans l'échantillon initial. On obtient ainsi au plus n^n échantillons « boostrapés » à partir desquels on va calculer les estimateurs empiriques des statistiques d'intérêt. Le bootstrap permet ainsi, à partir d'un tirage aléatoire au sein

de l'échantillon d'origine, de le perturber afin d'obtenir de nouveaux estimateurs des paramètres. Une fois que l'on dispose de *B* estimateurs $I_n(g)$ de $I(g)$, en posant

$$\hat{\mu}\left[I_n^1(g),\ldots,I_n^B(g)\right] = \frac{1}{B}\sum_{b=1}^{B} I_n^b(g),$$

et

$$\hat{\sigma}^2\left[I_n^1(g),\ldots,I_n^B(g)\right] = \frac{1}{B-1}\sum_{b=1}^{B}\left[I_n^b(g)-\hat{\mu}\right]^2,$$

on obtient un intervalle de confiance de la forme :

$$J_\alpha = \left[\hat{\mu} - \frac{\hat{\sigma}}{\sqrt{B}}\Phi^{-1}\left(1-\frac{\alpha}{2}\right), \hat{\mu} + \frac{\hat{\sigma}}{\sqrt{B}}\Phi^{-1}\left(1-\frac{\alpha}{2}\right)\right].$$

Cette technique est complétée par deux techniques annexes qui permettent d'affiner les résultats : le *boostrap* percentile et la méthode BCa. Ces techniques sont décrites *infra* dans le cas particulier de l'estimation de quantiles.

La méthode bootstrap n'est toutefois pas pertinente dans tous les cas de figures : elle ne fonctionne par exemple pas lorsque l'on s'intéresse à la valeur maximum de l'échantillon. En effet, par construction, il est impossible de constituer des échantillons contenant des valeurs supérieures à la valeur maximale de l'échantillon de départ. Pour ce type de problème, d'autres méthodes existent telle que celle proposée par ZELTERMAN [1993] qui consiste à rééchantillonner les intervalles entre deux valeurs adjacentes des données de départ plutôt que les données elles-mêmes. Cette méthode semi-paramétrique est illustrée dans le cas de l'estimation par intervalle de l'intensité maximale des séismes dans différentes régions par BOTTARD [1996].

Tel quel, le bootstrap est en revanche très efficace dans toutes les statistiques basées de près ou de loin sur la moyenne. La technique du bootstrap est communément utilisée en assurance pour analyser la variabilité des montants de sinistres et obtenir des erreurs de prédiction pour différentes méthodes de provisionnement, et notamment pour les méthodes basées sur Chain Ladder et sur les modèles linéaires généralisés (cf. PARTRAT et JAL [2002]).

4.1.3. Intervalles de confiance par la méthode du boostrap

On se place dans le cas particulier où le paramètre θ est estimé par l'estimateur du maximum de vraisemblance $\hat{\theta}$; on déduit des propriétés générales de l'estimateur du maximum de vraisemblance que $\hat{VaR}(\alpha) = F_{\hat{\theta}}^{-1}(\alpha)$ est alors l'estimateur du maximum de vraisemblance de $VaR(\alpha)$. $\hat{VaR}(\alpha)$ est donc asymptotiquement sans biais et gaussien[16].

[16] Pour le rappel des principales propriétés de l'estimateur du maximum de vraisemblance, on pourra consulter SAPORTA [1990].

En obtenant une estimation de sa variance asymptotique on pourra donc construire un intervalle de confiance. La loi de la statistique $\hat{VaR}(\alpha)$ est toutefois difficile à déterminer, et on est conduit à se tourner vers des techniques de simulation, et tout particulièrement la méthode « Bootstrap[17] ».

Dans le cas de l'estimation d'une *VaR*, l'échantillon initial est constitué par les *n* observations de *X* utilisées pour estimer les paramètres du modèle. La statistique d'intérêt est $\hat{VaR}(\alpha) = F_{\hat{\theta}}^{-1}(\alpha)$. On peut procéder de trois manières pour construire l'intervalle de confiance recherché :

- ✓ estimation de la variance boostrapée de $\hat{VaR}(\alpha)$, que l'on utilise ensuite avec l'hypothèse de normalité asymptotique (« bootstrap classique ») ;
- ✓ classement des estimations ponctuelles obtenues pour chaque échantillon bootstrap et constitution des bornes de l'intervalle de confiance par les estimations correspondantes (« bootstrap percentile ») ;
- ✓ estimation directe d'un intervalle de confiance boostrapé via la méthode BCa (« Bias corrected and accelerated »).

On rappelle ci-après le principe de la méthode BCa. On notera B la taille de l'échantillon bootstrap. La méthode consiste à déterminer les bornes b_1 et b_2 de l'intervalle de confiance de sorte que :

$$P\left[b_1 \le VaR_\alpha(X) \le b_2\right] = 1 - a$$

en prenant pour b_i la valeur d'indice $B \times \beta_i$ de l'échantillon des $\hat{VaR}(\alpha)$ bootstrapé, les valeurs β_1 et β_2 étant définies par :

$$\beta_1 = \phi\left(z_0 + \frac{z_0 + u_{\frac{a}{2}}}{1 - \gamma\left(z_0 + u_{\frac{a}{2}}\right)}\right) \text{ et } \beta_2 = \phi\left(z_0 + \frac{z_0 + u_{1-\frac{a}{2}}}{1 - \gamma\left(z_0 + u_{1-\frac{a}{2}}\right)}\right)$$

où on a posé :

- ✓ u_c le quantile d'ordre c de la normale centrée réduite ;
- ✓ $z_0 = \phi^{-1}(\kappa)$ avec κ la proportion des échantillons bootstrapés pour lesquels la *VaR* obtenue est inférieure à $\hat{VaR}(\alpha)$;

[17] Les éléments présentés ici sont tirés de EFRON et TIBSHIRANI [1993].

✓ $$\gamma = \frac{\sum_{i=1}^{n} \left(V\tilde{a}R(\alpha) - V\tilde{a}R_i(\alpha) \right)^3}{6\left[\sum_{i=1}^{n} \left(V\tilde{a}R(\alpha) - V\tilde{a}R_i(\alpha) \right)^2 \right]^{3/2}}$$

en désignant par $V\tilde{a}R_\alpha^i(X)$ la *VaR* obtenue sur le *i*ième échantillon « jackknife » (c'est à dire l'estimation de la *VaR* obtenue à partir de l'échantillon initial duquel on a retiré la *i*ième valeur) et $V\tilde{a}R_\alpha(X) = \frac{1}{n}\sum_{i=1}^{n} V\tilde{a}R_\alpha^i(X)$.

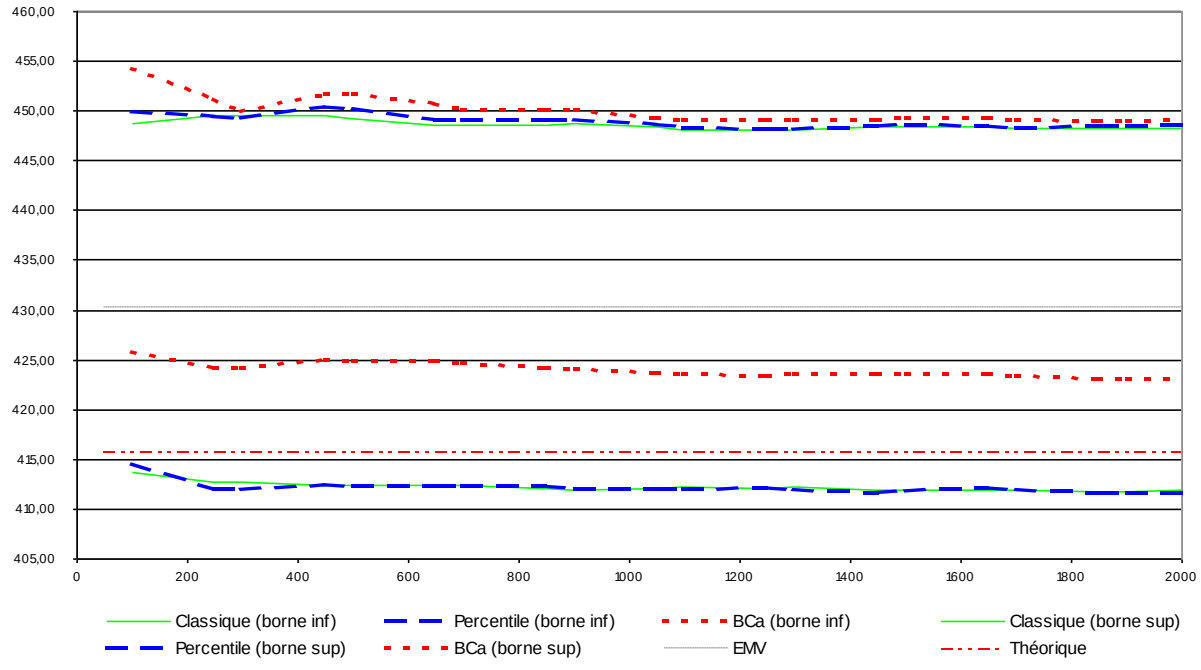
4.1.4. Application numérique

On applique les méthodes bootstrap décrite ci-dessus dans un modèle log-normal pour lequel (cf. *infra*) :

$$V\hat{a}R_p(X) = \exp\left(\hat{m} + \hat{\sigma}\phi^{-1}(p)\right)$$

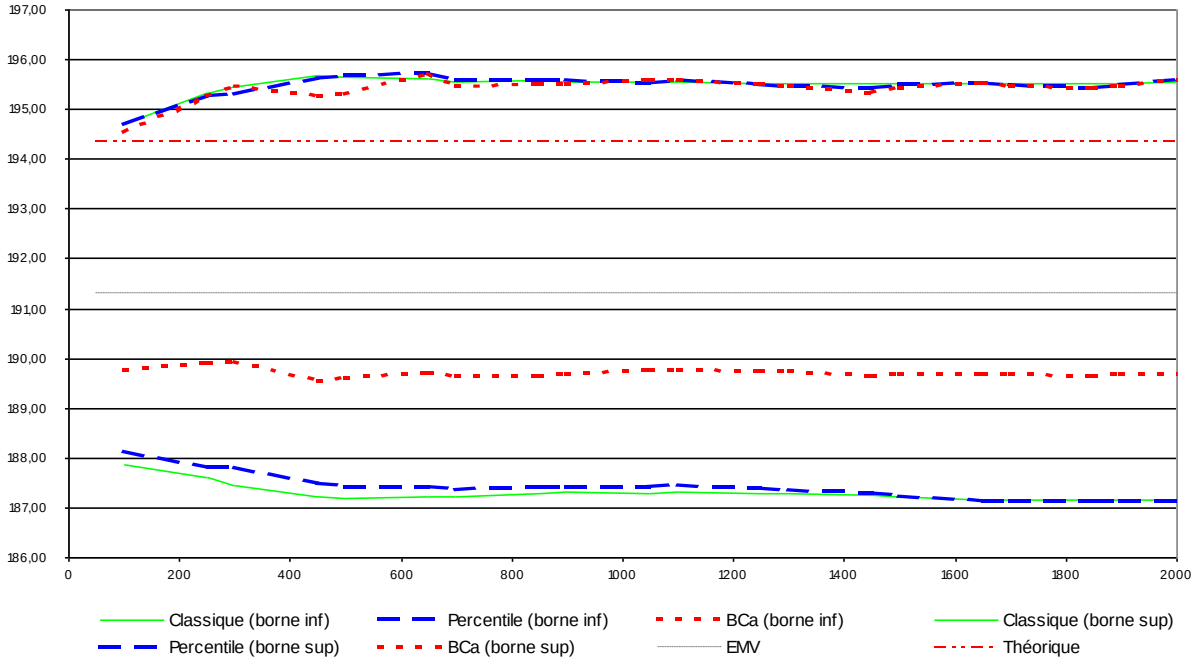
On retient les valeurs numériques suivantes : $m = 5$, $\sigma = 0{,}4$, ce qui conduit à une *VaR* théorique à 99,5 % de $VaR_{99,5\,\%}(X) = \exp\left(m + \sigma\phi^{-1}(p)\right) = 415{,}85$. On estime ce quantile sur la base d'un échantillon de taille 1 000 et on détermine les intervalles de confiance à 90 % avec les méthodes classique, percentile et BCA. On obtient en fonction de la taille de l'échantillon bootstrap les résultats typiques suivants :

Fig. 10 - **Comparaison des intervalles de confiance par méthode (VaR à 99,5 %)**



On note que les méthodes classique et percentile sont très proches. La méthode BCa fournit un intervalle dont la borne supérieure est proche des 2 autres méthodes, mais avec une borne inférieure plus élevée. Dans l'exemple ci-dessus on constate par exemple que si la vraie valeur est bien dans les intervalles de confiance « classique » et « percentile », elle est inférieure à la borne inférieure de l'intervalle fournit par la méthode BCa. Les résultats obtenus par les méthodes *bootstrap* restent peu robustes pour un quantile élevé. Pour un quantile d'ordre inférieur, comme par exemple le quantile à 75 %, les estimations deviennent bien entendu beaucoup plus fiables :

Fig. 11 - **Comparaison des intervalles de confiance par méthode (VaR à 75 %)**



4.2. Exemples

4.2.1. Loi log-normale (DA de Gumbel)

On considère ici le cas d'un risque *X* distribué selon une loi log-normale de paramètres $\left(m,\sigma^2\right)$, autrement dit $\ln(X)$ est distribué selon une loi normale de paramètres $\left(m,\sigma^2\right)$.
L'équation définissant la valeur à risque au niveau *p* se met sous la forme :

$$P\left[\frac{\ln(X)-m}{\sigma} \geq \frac{VaR_p(X)-m}{\sigma}\right] = 1-p$$

ce qui conduit à l'expression explicite de la fonction quantile suivante :

$$VaR_p(X) = F^{-1}(p) = \exp\left(m+\sigma\phi^{-1}(p)\right)$$

avec $\phi(x)$ la fonction de répartition de la loi normale centrée réduite. A partir d'un estimateur $(\hat{m}, \hat{\sigma}^2)$ des paramètres (m, σ^2) obtenu sur l'échantillon complet on obtient un estimateur simple de la valeur à risque :

$$\hat{VaR}_p(X) = \exp\left(\hat{m} + \hat{\sigma}\phi^{-1}(p)\right).$$

On peut par exemple utiliser les estimateurs du maximum de vraisemblance des paramètres (moyenne et écart-type empiriques). L'estimation des paramètres de la loi normale sous-jacente est entachée d'une erreur liée aux fluctuations d'échantillonnage. Il est important de mesurer l'incidence de cette erreur sur l'estimation de la valeur à risque qui s'en déduit. On trouve ici que les élasticités par rapport aux paramètres s'expriment selon :

$$\frac{1}{VaR_p(X)} \frac{\partial}{\partial m} VaR_p(X) = 1 \text{ et } \frac{1}{VaR_p(X)} \frac{\partial}{\partial \sigma} VaR_p(X) = \phi^{-1}(p)$$

On en déduit qu'une erreur de mesure de 1 % sur l'espérance conduit à une erreur de mesure de 1 % sur la valeur à risque ; mais parallèlement une erreur de mesure de 1 % sur la volatilité induit une erreur de $\phi^{-1}(p)$ sur la valeur à risque. Au seuil $p = 99,5\%$, on a $\phi^{-1}(99,5\%) = 2,58$, et donc une erreur de 1 % sur la volatilité conduit à une erreur 2,6 fois plus grande sur la *VaR*. On peut également noter en utilisant l'inégalité de Jensen pour une fonction convexe, $f(EX) \leq Ef(X)$, que, si on retient des estimateurs sans biais de (m, σ^2) on a :

$$\exp\left(E\left(\hat{m} + \hat{\sigma}\phi^{-1}(p)\right)\right) = VaR_p(X) \leq E\hat{VaR}_p(X).$$

En moyenne cet estimateur surestime la *VaR*.

4.2.2. Loi de Pareto (DA de Fréchet)

On s'intéresse ici au cas particulier d'un échantillon issu d'une loi de Pareto de première espèce de fonction de répartition $F(x) = 1 - \left(\frac{x_m}{x}\right)^{\alpha}$ pour $x \geq x_m$. On vérifie aisément que

l'estimateur du maximum de vraisemblance[18] de (x_m, α) est $\hat{x}_m = \mathbf{min}\{x_i, 1 \le i \le n\}$ et $\hat{\alpha} = \dfrac{n}{\sum_{i=1}^{n} \mathbf{ln}\left(1 + \dfrac{x_i}{\hat{x}_m}\right)}$. La fonction quantile a ici une expression simple :

$$VaR_p(X) = F^{-1}(p) = x_m (1-p)^{-1/\alpha}$$

La démarche paramétrique usuelle utilisée *supra* dans le cas de la loi log-normale fournit alors un estimateur naturel d'un quantile d'ordre quelconque :

$$\hat{F}^{-1}(p) = \hat{x}_m (1-p)^{-1/\hat{\alpha}},$$

et les élasticités par rapport aux paramètres s'en déduisent aisément :

$$\frac{1}{VaR_p(X)} \frac{\partial}{\partial x_m} VaR_p(X) = \frac{1}{x_m} \text{ et } \frac{1}{VaR_p(X)} \frac{\partial}{\partial \alpha} VaR_p(X) = \frac{\mathbf{ln}(1-p)}{\alpha^2}$$

Pour des valeurs du paramètre de forme α proches de 1, au seuil $p = 99{,}5\%$ on trouve que la sensibilité par rapport à α est de l'ordre de –5. L'estimation de la *VaR* par cette méthode est donc très sensible aux erreurs d'estimation de ce paramètre.

4.2.3. Loi de Weibull

On considère ici le cas d'une loi de Weibull de paramètres α et τ, dont l'expression de la densité est :

$$f(x) = \frac{\tau}{\alpha^{\tau}} x^{\tau-1} \mathbf{exp}\left\{-\left(\frac{x}{\alpha}\right)^{\tau}\right\}$$

et la fonction de répartition :

$$F(x) = 1 - \mathbf{exp}\left\{-\left(\frac{x}{\alpha}\right)^{\tau}\right\}.$$

L'inversion de la fonction de répartition est dans ce cas immédiatement et conduit à :

$$VaR_p(X) = F^{-1}(p) = \alpha \left[\mathbf{ln}\frac{1}{1-p}\right]^{1/\tau}.$$

Les élasticités par rapport aux paramètres s'en déduisent aisément :

[18] On fera attention dans ce cas au fait que l'estimateur EMV n'annule pas le vecteur des scores à l'optimum.

$$\frac{1}{VaR_p(X)} \frac{\partial}{\partial \alpha} VaR_p(X) = \frac{1}{\alpha} \quad \frac{1}{VaR_p(X)} \frac{\partial}{\partial \tau} VaR_p(X) = -\frac{\ln\left(\ln\left(\frac{1}{1-p}\right)\right)}{\tau^2}$$

Contrairement aux cas de la loi log-normale et de la loi de Pareto pour lesquels les équations de vraisemblance ont une solution explicite, on doit ici résoudre numériquement le système suivant :

$$\begin{cases} \alpha = \left(\frac{1}{n}\sum_{i=1}^{n} x_i^{\tau}\right)^{1/\tau} \\ \frac{1}{\tau} = \frac{\sum_{i=1}^{n} x_i^{\tau} \ln x_i}{\sum_{i=1}^{n} x_i^{\tau}} - \frac{1}{n}\sum_{i=1}^{n} \ln x_i \end{cases}$$

La deuxième équation définit un algorithme qui converge vers $\hat{\tau}$ pour autant qu'on lui fournisse une valeur initiale pas trop éloignée. En pratique, cette valeur initiale pourra être l'estimateur obtenu par la méthode des quantiles sur l'ensemble des observations complètes :

$$\breve{\tau} = \frac{\ln\left[\frac{-\ln(1-p_2)}{-\ln(1-p_1)}\right]}{\ln(x_{p_2}) - \ln(x_{p_1})}$$

Une fois $\hat{\tau}$ obtenu, $\hat{\alpha}$ s'en déduit grâce à la première équation.

Remarque : La loi de Weibull apparaît naturellement dans l'étude de la distribution limite du minimum d'un échantillon iid. En effet, si $(X_1,..,X_n)$ est un échantillon d'une loi *G* sur $]0,+\infty[$ dont le comportement à l'origine vérifie :

$$\lim_{x \to 0^+} \frac{G(x)}{\lambda x^{\alpha}} = 1$$

alors $n^{1/\alpha} X_{(1)}$ converge en loi lorsque *n* tend vers l'infini vers une distribution $W(\alpha, \lambda^{-1/\alpha})$.

Démonstration : On a $P\left(n^{1/\alpha} X_{(1)} > x\right) = \left[1 - G\left(\frac{x}{n^{1/\alpha}}\right)\right]^n$ et donc :

$$\ln\left(P\left(n^{1/\alpha} X_{(1)} > x\right)\right) = n \ln\left[1 - G\left(\frac{x}{n^{1/\alpha}}\right)\right] = n\left[-\lambda\left(\frac{x}{n^{1/\alpha}}\right)\alpha + o\left(\frac{1}{n}\right)\right]$$

D'où l'on tire que $\lim_{n\to\infty} P\left(n^{1/\alpha} X_{(1)} > x\right) = e^{-\lambda x^{\alpha}}$, ce qui achève la démonstration.

En fait, cette propriété est à l'origine de la forme de la loi proposée par W. WEIBULL dans son article de 1951. Il se propose en effet de résoudre des problèmes de rigidité des matériaux. L'exemple qu'il utilise pour illustrer sa distribution est celui d'une chaîne. Comment peut-on établir la probabilité qu'une chaîne se brise ? Son raisonnement est que la chaîne va se briser si le plus faible des maillons se brise. Ceci revient donc à trouver la distribution du minimum d'un grand nombre d'objets. En théorie des valeurs extrêmes, on établit que la distribution du minimum ne dépend pas de la fonction de probabilité de chaque objet si le nombre d'objets est suffisamment grand (GALAMBOS [1978], GUMBEL [1958]). LOGAN [1992] a utilisé cette distribution dans l'optique d'une course (race model). Imaginons par exemple un grand nombre de neurones en compétition pour émettre un signal. Le signal émis sera produit par le neurone le plus rapide.

5. Références

BENGIO Y., CARREAU J. [2004], « *Estimation de densité conditionnelle lorsque l'hypothèse de normalité est insatisfaisante* », CIRANO.

BERTAIL P. [2002], « *Évaluation des risques d'exposition à un contaminant alimentaire : quelques outils statistiques.* », CREST, working paper.

BOTTARD S. [1996] « Application de la méthode du Bootstrap pour l'estimation des valeurs extrêmes dans les distributions de l'intensité des séismes », *Revue de statistique appliquée* 44 (4), 5-17.

DANIELSON J., DE VRIES C.G. [1997], « *Extreme returns, tail estimator and Value-at-Risk* » , Working paper, London School of Economics.

DE HAAN L., ROTZEN H. [1993], « *On the estimation of high quantiles* », Journal of Statistical Planning and Inference, Vol. 35 n°1, 1-13.

DEKKERS A.L.M., EINMAHL J.H.J., DE HAAN L. [1989], « *A moment estimator for the index of an extrem value disribution* », The Annals of Statistics, Vol. 17 n°4, 1833-1855.

DENUIT M., CHARPENTIER A. [2004] « *Mathématiques de l'assurance non-vie. Tome 1 : principes fondamentaux de théorie du risque* ». Paris : Economica.

DENUIT M., CHARPENTIER A. [2005] *Mathématiques de l'assurance non-vie. Tome 2 : tarification et provisionnement*, Paris : Economica.

DIEBOLT J., EL ARAOUI M.A. [1997] « *Modélisation des queues de distribution et des quantiles extrêmes (1)* », Rapport de contrat INRIA-EDF.

EFRON B. [1979] « Bootstrap methods: Another look at the Jackknife », *Ann. Statist.*, 71-26.

EFRON B. TIBSHIRANI R.J. [1993], « *An introduction to the boostrap* », Monographs on statistics and applied probability, Chapman & Hall.

EMBRECHTS P., KLUPPELBERG C., MIKOSCH T. [1997], « *Modelling extremal events* », Berlin : Springer Verlag.

ENGLAND P.D., VERALL R.J. [1999] « Analytic and bootstrap estimates of prediction errors in claim reserving », *Insurance: mathematics and economics* 25, 281-93.

JANSSEN D.V., DE VRIES C.G. [1991], « *On the frequency of large stock returns: putting booms and busts into perspective* », Review of Economics and Statistics, 73, 18-24

JONDEAU E.,ROCKINGER M. [1999], « *The tail behavior of stock returns : emerging versus mature markets* », Banque de France, Note d'étude et de recherche n°66

GALAMBOS J. [1978] *The Asymptotic Theory of Extreme Order Statistics*. New York : John Wiley and Sons.

GARRIDO M. [2002], « *Modélisation des évènements rares et des quantiles extrêmes, méthodes de sélection de modèle pour les queues de distribution* », Thèse de doctorat, Université J. Fourier.

GAUTHIER C., PISTRE N. [2000], « *Evénements extrêmes sur les spreads de crédit* », ENSAE, Working Paper

GUMBEL E. J. [1958] *The Statistics of Extremes*. New York : Columbia University Press.

HOROWITZ J.L. [1997] « Bootstrap methods in econometrics: theory and numerical performance », in *Advances in Economics and Econometrics: Theory and Application*, volume 3, 188-222, D. M. Kreps, K. F. Wallis (eds), Cambridge: Cambridge University Press.

HOSKING J. R., WALLIS J. R. [1987] « Parameter and quantile estimation for the generalized pareto distribution », *Technometrics* 29, 339-49.

JAL P., PARTRAT Ch. [2004] « Evaluation stochastique de la provision pour sinistres », Conférence scientifique de l'Institut des Actuaires, Paris, 20 janvier 2004.

LOGAN G. D. [1992] « Attention and preattention in theories of automaticity ». *American Journal of Psychology*, 105, 317-339.

PICKANDS J. [1975], « *Statistical inference using extreme order statistics* », Ann. Statist, 3, 119-131.

PLANCHET F., THÉROND P.E. [2005] « L'impact de la prise en compte des sauts boursiers dans les problématiques d'assurance », *Proceedings of the 15th AFIR Colloquium*.

RESNICK S., STARICA C. [1997], « *Smoothing the Hill estimator* » , Advanced Applied Probabilities, 29, pp. 271-293

ROBERT C. [1998], « *Mouvements extrêmes des séries financières haute fréquence* », Finance, Vol. 19, pp. 221-271

SAPORTA G. [1990], « *Probabilités, analyse des données et statistique* », Paris : Editions Technip

THÉROND P.E., PLANCHET F. [2007] « Provisions techniques et capital de solvabilité d'une compagnie d'assurance : méthodologie d'utilisation de Value-at-Risk », *Assurances et gestion des risques*, Vol. 74 (4).

ZAJDENWEBER D. [2000] *Économie des extrêmes*, Paris : Flammarion

ZELTERMAN D. [1993] « A semiparametric Bootstrap technique for simulating extreme order statistics », *Journal of American Statistical Association* 88 (422), 477-85.