

**Mémoire présenté devant l'ENSAE ParisTech
pour l'obtention du diplôme de la filière Actuariat
et l'admission à l'Institut des Actuaires**

le 27 juin 2012

Par : Xavier SERVEL et Sylvie TRAN

Titre: Application de la Théorie des Valeurs Extrêmes dans le Cadre de
l'Implémentation d'un Modèle Interne

Confidentialité : NON OUI (Durée : 1 an 2 ans)

Les signataires s'engagent à respecter la confidentialité indiquée ci-dessus

Membre présents du jury de l'Institut
des Actuaires

signature

Entreprise :

Nom : ZGHAL Thameur

Signature :

Membres présents du jury de la filière

Directeur de mémoire en entreprise :

Nom : ZGHAL Thameur

Signature :

Invité :

Nom :

Signature :

**Autorisation de publication et de
mise en ligne sur un site de
diffusion de documents actuariels
(après expiration de l'éventuel délai de
confidentialité)**

Signature du responsable entreprise

Secrétariat

Signature du candidat

Bibliothèque :

Application de la théorie des valeurs extrêmes dans le cadre de l'implémentation d'un modèle interne

Xavier Servel

Sylvie Tran

29 mai 2012

Résumé

Le présent article s'intéresse aux méthodologies de calcul de quantiles extrêmes dans le cadre de Solvabilité II. Nous développons des méthodes basées sur la théorie des valeurs extrêmes afin d'évaluer ces quantités et montrons comment la théorie des valeurs extrêmes peut être utilisée dans le cadre d'un modèle interne lié du risque de marché. Nous travaillons pour cela sur un historique de l'indice IPD UK, CAC 40, taux d'intérêt et spreads de crédit. Nous modélisons ces facteurs de risque à l'aide de processus *ARMA – GARCH* et nous proposons une alternative aux pratiques standards en approchant les résidus avec une loi de Pareto Hybride. Ceci nous permet d'en déduire les distributions complètes des facteurs de risque dans un an qui pourront être utilisées dans un générateur de scénarios économiques, après avoir été corrélées entre elles.

Mots-clés : Solvabilité II, théorie des valeurs extrêmes, GEV, GPD, GARCH, GPD Test, Pareto Hybride

1 Introduction

Solvabilité II exige des assureurs qu'ils détiennent un montant de capital suffisant pour faire face à des événements extrêmes : ceux-ci doivent leur permettre d'éviter la ruine avec une probabilité à 99,5 % à horizon un an. Pour calculer ce montant capital - appelée Capital de Solvabilité Requis (SCR en anglais) - les assureurs ont le choix entre la possibilité d'appliquer la formule standard fournie par le régulateur ou de mettre en place un modèle interne mieux adapté à leur profil de risque. Dans ce dernier cas, ils doivent donc estimer avec des méthodes statistiques adéquates les quantiles extrêmes de leur distribution de résultat (perte bi-centenaire dans Solvabilité II), qui devront être validées par l'Autorité de Contrôle Prudenciel. Ces compagnies sont donc tenues de justifier la pertinence de leurs méthodes. Cependant, il reste difficile de modéliser les événements extrêmes, car les historiques disponibles sont dans la plupart des cas restreints (généralement 10 à 20 ans maximum). L'objectif ici est de construire un ensemble de modèles et d'outils basés sur la théorie des valeurs extrêmes permettant l'estimation des quantiles extrêmes dans le cadre d'un modèle interne pour le risque de marché.

Nous commençons par une description des principaux résultats de la théorie des valeurs extrêmes, avant de présenter l'application de ces méthodes à des séries de données financières. Enfin, nous procédons à une analyse des résultats des estimations dans une troisième partie.

2 Etude de la théorie des valeurs extrêmes

Afin de déterminer un quantile extrême, nous choisissons d'étudier les queues de distribution à l'aide de la théorie des valeurs extrêmes dont nous rappelons dans cette section les principaux résultats.

2.1 Distribution des valeurs extrêmes

La théorie des valeurs extrêmes repose sur le théorème de Fisher-Tippett-Gnedenko :

Soit $(X_i)_{i=1,\dots,n}$ une suite de variables aléatoires indépendantes et identiquement distribuées (i.i.d.), de fonction de répartition F , définissons la suite $(M_i)_{i=1,\dots,n}$ telle que $M_n = \max(X_1, \dots, X_n)$. S'il existe deux suites de constantes $(a_i)_{i=1,\dots,n}$ et $(b_i)_{i=1,\dots,n}$, et une fonction de répartition H non-dégénérée telles que $\frac{M_n - b_n}{a_n}$ converge en loi vers H , alors H appartient à la classe de fonctions suivante :

$$H_\xi(x) = \begin{cases} e^{-(1+\xi x)^{-\frac{1}{\xi}}} \mathbb{1}_{1+\xi x > 0} & \text{si } \xi \neq 0 \\ e^{-e^{-x}} & \text{si } \xi = 0 \end{cases}$$

H_ξ est la loi des valeurs extrêmes (ou GEV) et ξ est appelé "indice de valeurs extrêmes". Ce paramètre mesure l'épaisseur de la queue de distribution. Si $\xi < 0$, F appartient au domaine d'attraction de Weibull (distribution à queue finie), si $\xi = 0$, il s'agit du domaine d'attraction de Gumbel (distribution à queue fine), et enfin $\xi > 0$ correspond au domaine d'attraction de Fréchet (distribution à queue lourde).

En pratique, nous pouvons nous appuyer sur ce théorème pour estimer notre quantile d'intérêt q en présence d'un échantillon indépendant et identiquement distribué en étudiant la distribution de M_n pour un n fixé. Dans ce cadre, a_n et b_n deviennent alors des constantes que nous appelons respectivement σ et μ .

Pour ce faire, nous séparons l'échantillon en k blocs de taille r et nous considérons l'échantillon constitué par les maxima de chaque bloc. Pour r assez grand, nous pouvons supposer que la distribution de M_r est une distribution GEV de paramètres μ , σ , et ξ . Il s'ensuit immédiatement que le quantile q d'ordre p peut s'approcher par la relation suivante :

$$\hat{q} \simeq \hat{\mu} + \frac{\hat{\sigma}}{\hat{\xi}} ((-r \ln(p))^{-\hat{\xi}} - 1)$$

Cela nécessite de toute évidence une estimation des paramètres de la GEV.

2.1.1 Estimation des paramètres de la loi des valeurs extrêmes

Plusieurs méthodes existent pour estimer les paramètres de la GEV. Les deux plus classiques sont la méthode du maximum de vraisemblance (ou EMV) et la méthode des moments pondérés (ou PWM). Nous testons également d'autres méthodes développées plus récemment telles que la méthode des moments pondérés généralisés (ou GPWM) qui vise à étendre le domaine d'application de la méthode des moments pondérés (voir [3]), et la méthode de "rao-blackwellisation" qui améliore la qualité des estimateurs en réduisant leur variance (voir [5]).

De plus, pour ces différents estimateurs, nous construisons des intervalles de confiance reposant sur leur normalité asymptotique (voir [8]) ou sur le profil de vraisemblance (voir [2]).

Cette approche par blocs basée sur la GEV souffre d'un inconvénient important : l'utilisation d'un seul maximum par bloc conduit à une perte d'information contenue dans les autres grandes valeurs du bloc. Pour pallier ce problème, la méthode POT (*Peaks-over-Threshold*) ou "méthode des excès" a été introduite.

2.2 Distribution de la loi de Pareto Généralisée

Cette approche repose sur le théorème de Balkema-de Haan-Pickands et consiste à analyser l'échantillon des excès au dessus d'un certain seuil u .

Considérons toujours un échantillon X i.i.d. de fonction de répartition F . Il s'agit d'étudier l'échantillon $X_i - u$ pour tout i tel que $X_i > u$. D'après Balkema-de Haan-Pickands, si F vérifie les hypothèses du premier théorème, la distributions des excès lorsque u converge vers l'infini converge vers une loi de Pareto Généralisée (ou GPD) dont la fonction de répartition s'écrit :

$$G_{\xi,\sigma}(x) = \begin{cases} (1 - (1 + \frac{\xi x}{\sigma})^{-\frac{1}{\xi}}) \mathbb{1}_{x>0} & \text{si } \xi > 0 \\ (1 - (1 + \frac{\xi x}{\sigma})^{-\frac{1}{\xi}}) \mathbb{1}_{x \in]0, -\frac{\sigma}{\xi}[} & \text{si } \xi < 0 \\ (1 - e^{-\frac{x}{\sigma}}) \mathbb{1}_{x>0} & \text{si } \xi = 0 \end{cases}$$

Ce théorème permet également d'estimer le quantile q avec la relation :

$$q \simeq u + \frac{\hat{\sigma}}{\hat{\xi}} \left(\left(\frac{n}{N_u} (1-p) \right)^{-\hat{\xi}} - 1 \right) \text{ où } N_u \text{ est le nombre d'excès.}$$

L'estimation du quantile requiert donc la détermination du seuil u et une estimation des paramètres de la GPD.

2.2.1 Recherche du Seuil u

Le choix du seuil est primordial dans cette approche : afin de pouvoir approcher la queue de distribution par une loi GPD, le seuil doit être suffisamment élevé. Cependant, un seuil trop grand peut réduire le nombre de données disponibles. Nous choisissons d'implémenter plusieurs méthodes afin de sélectionner le seuil adéquat, que nous présentons ici.

Fonction moyenne des excès Afin de déterminer ce seuil, nous pouvons faire appel à la fonction moyenne des excès (ou *mean excess function*) qui se caractérise par sa linéarité en u pour les lois GPD :

$$e(u) = \mathbb{E}(X - u | X > u)$$

Il s'agit donc de chercher le seuil u à partir duquel la fonction est linéaire.

Hill Plot Pour cette méthode il faut calculer un estimateur de $\frac{1}{\xi}$ (valable pour $\xi > 0$). Lorsque nous utilisons tout l'échantillon pour cela, sa valeur est biaisée car seuls les dépassements au-dessus d'un seuil élevé suivent une loi GPD. L'idée est de calculer cet estimateur avec un grand échantillon et d'enlever au fur et à mesure les observations les plus petites, jusqu'à ce que l'estimateur se stabilise. Le point où il se stabilise est notre seuil (voir [4]).

Cunnane Test Asymptotiquement, le nombre d'excès au-dessus du seuil u suit une loi de Poisson (voir [10]). Nous utilisons ce résultat pour déterminer u , en faisant aussi l'approximation qu'une loi de Poisson est proche d'une loi normale pour un paramètre suffisamment élevé et en testant la proximité entre la loi du nombre d'excès et une loi normale par une statistique du χ^2 .

Minimisation de distance de Kolmogorov-Smirnov (méthode de Pickands) Cette méthode est proche de la précédente, à la différence qu'elle ne considère pas le nombre d'excès mais les excès en eux-mêmes, censés suivre une loi de Pareto généralisée. De plus, c'est la statistique de Kolmogorov-Smirnov qui est appliquée (voir [11]).

2.2.2 Estimation des Paramètres de la GPD

Comme pour la GEV, nous utilisons plusieurs méthodes pour estimer les paramètres de la GPD : la méthode du maximum de vraisemblance (EMV) et celle des moments pondérés (PWM).

2.3 Simulations d'échantillons i.i.d. et comparaison des différentes méthodes

Après avoir implémenté toutes les méthodes décrites dans les deux paragraphes précédents, nous testons ces outils sur des échantillons théoriques de lois connues en estimant plusieurs quantiles. Le tableau ci-dessous présente le champ d'application théorique de chaque estimateur.

	GEV			GPD	
	EMV	PWM	GPWM	EMV	PWM
Convergence de l'estimateur	ξ quelconque	$\xi < 1$	$\xi < 2$	ξ quelconque	$\xi < 1$
Normalité asymptotique	$\xi > -\frac{1}{2}$	$\xi < \frac{1}{2}$	$\xi < \frac{3}{2}$	$\xi > -\frac{1}{2}$	$\xi < \frac{1}{2}$

Nous choisissons d'étudier trois lois : la loi normale centrée réduite ($\xi = 0$), la loi de Student à 2 degrés de liberté ($\xi = 0,5$) et la loi de Student à 1 degré de liberté ($\xi = 1$). Pour chacune de ces lois, nous calculons les estimateurs des quantiles à 95 % et à 99,5 % pour trois tailles d'échantillon différentes : $n = 100, 1000$ et 10000 et nous étudions les erreurs quadratiques associées ainsi que la qualité des intervalles de confiances.

Cette étude nous permet de mettre en évidence le comportement des différentes méthodes d'estimation du quantile et de calcul d'intervalles de confiance en fonction de la taille des données et de l'épaisseur de leur queue de distribution. Comme attendu, les erreurs quadratiques sont les plus faibles dans trois cas : lorsque la taille de l'échantillon est importante, lorsque ξ est proche de 0 et lorsque le quantile recherché n'est pas "trop" extrême. Pour les échantillons de petite taille, les estimateurs des moments sont plus fiables que ceux du maximum de vraisemblance, qui ont tendance à être beaucoup plus volatiles. Cela est d'autant plus vrai lorsque ξ croît. En revanche, pour de grands échantillons, les estimateurs du maximum de vraisemblance sont préférables. Dans le cas de ξ élevé, les estimateurs des moments pondérés généralisés fournissent également de bons résultats.

Concernant l'approche par blocs, la rao-blackwellisation permet dans tous les cas d'améliorer la précision des quantiles estimés. Pour l'approche par excès, le choix du seuil par procédure de Pickands nous semble plus stable que les autres. De plus, comme nous avons davantage de valeurs pour cette approche, elle donne généralement de meilleurs résultats que celle par blocs (erreurs quadratiques plus faibles).

Les mêmes conclusions s'appliquent pour les intervalles de confiance.

A l'issue de cette étude, nous pouvons définir un mode opératoire lorsque nous serons en présence de données réelles : premièrement, nous choisirons une taille de blocs en fonction de la structure des données, et sélectionnerons le seuil par comparaison des différentes méthodes tout en privilégiant la procédure de Pickands lorsque celles-ci divergent. Nous estimerons ensuite ξ par les différentes méthodes et en déduirons l'applicabilité ou non des différentes méthodes. Si l'échantillon est de taille faible, nous privilégierons les méthodes des moments pondérés dans le cas où $\xi < 1$ (ou généralisés pour l'approche par blocs si $\xi < 1,5$). Cela nous permettra d'avoir un "vrai" estimateur de ξ et des autres paramètres pour en déduire un estimateur du quantile.

2.4 GPD Test

Une autre approche permettant le calcul de quantile paramétrique repose sur le GPD Test (voir [6]) qui vérifie l'adéquation de la queue de distribution des données $(X_i)_{i=1,\dots,n}$ à la famille de fonctions de répartition F_θ . Il s'agit du test suivant :

$$H_0 : X \in (F_\theta)_{\theta \in \Theta}$$

$$H_1 : X \notin (F_\theta)_{\theta \in \Theta}$$

Il s'agit alors de calculer un estimateur \hat{q}_{GPD} du quantile q d'ordre p à l'aide de l'approche des excès puis de comparer cette valeur avec le quantile paramétrique obtenu en prenant l'estimateur du maximum de vraisemblance $\hat{\theta}$ de l'échantillon :

$$\hat{q}_{param} = F_{\hat{\theta}}^{-1}(1 - p)$$

Le principe de notre test est de dire que si X peut être approché par la loi $F_{\hat{\theta}}$, alors l'écart entre le quantile de X obtenu par théorie des valeurs extrêmes et celui obtenu paramétriquement doit ressembler aux écarts que nous obtenons en prenant à la place de X des échantillons ayant été tirés suivant la loi $F_{\hat{\theta}}$. S'il y a une différence significative, c'est-à-dire si nous sortons de l'intervalle de confiance à 95 %, alors X ne peut pas être approché par la loi $F_{\hat{\theta}}$.

Nous avons appliqué ce test à une loi mélange : c'est une loi normale centrée réduite dans 99 %

des cas et une loi normale de moyenne 3 et d'écart-type 2 dans 1 % des cas. Alors qu'un test de Kolmogorov-Smirnov accepte l'hypothèse de normalité de cette distribution, l'utilisation d'un GPD Test permet de la rejeter. Ceci souligne que les tests d'adéquation usuels ne prennent en compte les queues de distribution que de manière limitée. Dans le cadre de la détermination d'un quantile extrême, l'application de tels tests n'est pas suffisante et il est nécessaire de porter attention aux queues de distribution en effectuant un GPD Test par exemple.

3 Application de la théorie des valeurs extrêmes sur des données financières

Comme vu dans la partie précédente, la théorie des valeurs extrêmes est un outil privilégié qui permet de calculer des quantiles extrêmes par une modélisation des queues de distribution. Cette théorie va donc nous permettre de trouver les chocs permettant de calculer le SCR de marché. Pour cela, nous avons à notre disposition des données s'étalant sur une dizaine d'années (obtenus par *Bloomberg*), retraçant l'historique des indices IPD UK, CAC 40, des taux d'intérêt de maturité 1 an à 30 ans ainsi que des spreads de crédit des obligations notées de AAA à BBB. Ces séries ne présentant pas de tendance déterministe et le test de Phillips-Perron indiquant leur non-stationnarité au seuil de 5 %, nous les rendons stationnaires en considérant les log-rendements et nous notons la nouvelle série X_t .

Notre objectif est de calculer un choc à la baisse pour les actions et l'immobilier et à la hausse pour les spreads de crédit. En ce qui concerne les taux, nous cherchons à calculer un choc à la hausse et à la baisse.

3.1 Modélisation des lois marginales de chaque facteur de risque

L'analyse de ces données suggère l'existence d'hétéroscédasticité, ce qui nous pousse à les modéliser avec des modèles $ARMA(P, Q) - GARCH(p, q)$:

$$\begin{cases} X_t = \mu + \sum_{i=1}^P a_i X_{t-i} + \sum_{j=1}^Q b_j \epsilon_{t-j} + \epsilon_t \\ \epsilon_t = \eta_t \sqrt{h_t} \\ h_t = \omega + \sum_{i=1}^q \alpha_i \epsilon_{t-i}^2 + \sum_{j=1}^p \beta_j h_{t-j} \end{cases}$$

Pour chaque série, nous choisissons le modèle le plus simple pour lequel les résidus sont indépendants et les coefficients des ordres les plus élevés significatifs. Une hypothèse usuelle de ce type de modèle est de considérer la normalité des résidus (η_t). Cependant, cette hypothèse est ici rejetée (au seuil de 5 %) pour chacune des séries et nous les modélisons par conséquent à l'aide d'une loi de Pareto Hybride (voir [7]). Cette loi permet de prendre en compte les valeurs extrêmes prises par les résidus tout en conservant l'aspect gaussien du milieu de la distribution. Plus formellement, il s'agit d'une loi qui au dessus du seuil u correspond à une loi GPD et qui sinon est une loi normale.

La méthodologie consiste alors à prendre un intervalle $[u - \epsilon, u + \epsilon]$ au sein duquel la densité est un mélange des deux densités. Concrètement, en appelant f la fonction de densité de la gaussienne, g la fonction de densité de la loi de Pareto généralisée, et l la fonction de densité de la loi de Pareto Hybride, nous avons l'expression suivante :

$$l(x) = \begin{cases} \kappa f(x) & \text{si } x < u - \epsilon \\ \kappa(g(q(x))q'(x) + f(p(x))p'(x)) & \text{si } x \in [u - \epsilon, u + \epsilon[\\ \kappa g(x) & \text{si } x \geq u + \epsilon \end{cases}$$

où p et q sont des fonctions de pondération caractérisant le mélange dans la zone où les deux densités s'appliquent, et où κ sert à ce que l'intégrale de l soit bien égale à 1. Cette loi permet ainsi de corriger la faible épaisseur de queue de la loi gaussienne qui conduit à une faible représentation des valeurs extrêmes.

Dans le cas des taux d'intérêt, il est nécessaire de calculer à la fois un choc à la hausse et à la baisse. Il est donc important de tenir compte des valeurs extrêmes à droite et à gauche de la distribution.

Pour cela, nous utilisons une loi Pareto Hybride à deux seuils, qui est caractérisée par une GPD en dessous et au dessus des seuils, et au milieu de sa distribution par une loi normale.

3.2 Algorithme d'obtention de la distribution à un an

En estimant les paramètres de la loi de Pareto Hybride, nous pouvons ainsi obtenir la distribution de X dans un an. En effet, connaissant $(h_t, \epsilon_t, X_t)_{t \leq T}$ en T , nous pouvons en déduire $(h_{T+1}, \epsilon_{T+1}, X_{T+1})$ en utilisant un tirage aléatoire de la variable η qui suit une loi de Pareto Hybride dont les paramètres ont été estimés :

$$\begin{cases} h_{T+1} = \hat{\omega} + \sum_{i=1}^q \hat{\alpha}_i \epsilon_{T-i+1}^2 + \sum_{j=1}^p \hat{\beta}_j h_{T-j+1} \\ \epsilon_{T+1} = \eta \sqrt{h_{T+1}} \\ X_{T+1} = \hat{\mu} + \sum_{i=1}^P \hat{a}_i X_{T-i+1} + \sum_{j=1}^Q \hat{b}_j \epsilon_{T-j+1} \end{cases}$$

Puis par itérations successives, nous obtenons donc une valeur possible de X dans un an. En utilisant cet algorithme un grand nombre de fois (50 000 généralement), nous obtenons la distribution de notre variable d'intérêt dans un an dont nous pouvons déduire le quantile à 0,5 % ou 99,5 % selon que nous calculons un choc à la hausse ou à la baisse. Ce quantile est obtenu par méthode classique (nous ordonnons les observations et prenons la 250^e meilleure ou pire valeur) car les échantillons sont très grands, et dans ce cas le recours à la théorie des valeurs extrêmes ne se justifie pas.

3.3 Etude de la structure de dépendance des facteurs de risque

Modéliser l'évolution des différents sous-jacents (immobilier, action, taux et spread) à l'aide d'un modèle *ARMA – GARCH* à résidus Pareto Hybride nous permet donc d'estimer la loi marginale de chaque sous-jacent et de calculer des quantiles extrêmes à horizon un an. Cependant, dans le cadre d'un modèle interne, il est également nécessaire de capturer la structure de dépendance entre les différents facteurs de risque.

Nous commençons donc par tenir compte de la dépendance entre les différents taux et spreads. Nous notons en effet l'existence d'une corrélation interne aux taux de différentes maturités ou interne aux différents spreads. Pour capter cette structure de dépendance, nous avons utilisé une pratique usuelle de marché : corrélérer les résidus de nos modèles *via* une copule gaussienne.

Enfin, nous agrégeons les différents facteurs de risque à l'aide de la matrice de corrélation du QIS. Ceci nous permet de projeter les actifs dans l'avenir à partir de paramètres reflétant la situation à une date donnée, et donc de fournir des scénarios monde-réel.

4 Résultats

Voici les résultats que nous obtenons :

Maturités	1 an	2 ans	3 ans	4 ans	5 ans	7 ans	10 ans	15 ans	20 ans	30 ans
Choc 0,5 %	-77%	-80%	-79%	-80%	-77%	-74%	-68%	-54%	-52%	-50%
Choc 99,5 %	+39%	+56%	+37%	+34%	+29%	+18%	+18%	+5%	+11%	+15%
Choc 0,5 % QIS 5	-75%	-65%	-56%	-50%	-46%	-39%	-31%	-27%	-29%	-30%
Choc 99,5 % QIS 5	+70%	+70%	+64%	+59%	+55%	+49%	+42%	+33%	+26%	+25%

TABLE 1 – Chocs des taux à un an

	spread AAA	spread AA	spread A	spread BBB	Immobilier	Action
Choc estimé	1,41 %	2,99 %	5,28 %	19,28 %	-15,50 %	-43,50 %
Choc QIS 5	0,90 %	1,10 %	1,40 %	2,50 %	-25,00 %	-30,00 %

TABLE 2 – Chocs des spreads, de l'immobilier et de l'action à un an

Pour les taux d'intérêt, nous notons que les chocs à la baisse sont nettement plus importants que ceux proposés dans le QIS 5. Par contre, le phénomène inverse est observé pour les chocs à la hausse. Cela peut s'expliquer par la dépendance de notre modèle aux dernières valeurs des taux dans notre historique qui reflétaient une forte tendance à la baisse, d'où cette asymétrie dans nos estimateurs.

Pour les spreads et les actions, nos chocs sont également plus élevés que ceux dans le QIS 5. Ils tiennent compte de la forte volatilité observée sur les marchés pour les données les plus récentes de notre base, qui favorise les valeurs extrêmes.

Nos résultats doivent cependant être nuancés. En effet, les chocs calculés dépendent des modèles choisis et des hypothèses que nous faisons. Dans le cadre de ce mémoire nous avons traité de modèles assez simples pour focaliser notre attention sur la théorie des valeurs extrêmes, mais ceux-ci peuvent être améliorés. D'autres raisons peuvent expliquer la différence avec les chocs du QIS 5 : les séries utilisées ne sont pas forcément les mêmes, et ne couvrent pas non plus exactement la même période. Un choc calculé à une date ou à une autre ne seront pas nécessairement proches car les conditions de marché peuvent être différentes.

5 Conclusion

Dans cet article, nous avons montré que la théorie des valeurs extrêmes pouvait s'appliquer dans un modèle interne partiel. Nous avons en effet développé une méthodologie à l'aide de cette théorie, permettant le calcul de quantiles extrêmes mais aussi une modélisation de la distribution complète de chaque facteur de risque. Il suffit alors de corrélérer entre les différents facteurs pour obtenir un générateur de scénarios économiques.

Nous avons également mis en évidence un test permettant de vérifier l'adéquation d'un échantillon à une loi en nous focalisant sur la proximité des queues de distribution. Utilisé conjointement à des tests d'adéquation plus classiques - comme le test de Kolmogorov-Smirnov - cela permet de nous assurer que notre modélisation est licite tant sur la partie centrale que sur les queues de distribution.

Cependant, les méthodes que nous proposons peuvent encore être améliorées à plusieurs niveaux. Ainsi, il est entre autres possible de choisir des modèles plus élaborés pour décrire l'évolution des indices. D'autres structures de dépendance plus sophistiquées que la copule gaussienne peuvent également être utilisées.

Bibliographie

- [1] BEIRLANT J., GOEGEBEUR Y., TEUGELS J., SEGERS J. - *Statistics of Extreme* - Wiley, 2004
- [2] COLES S. - *An Introduction to Statistical Modeling of Extreme Values* - Springer, 2001
- [3] DIEBOLT J., GUILLOU A., NAVEAU P., RIBEREAU P. - "Improving Probability-Weighed Moment methods for the Generalized Extreme-Value distribution" - *Revstat Statistical Journal*, 2008
- [4] EMBRECHTS P., KLÜPPELBERG C., MIKOSCH T. - *Modelling Extremal Events for Insurance and Finance* - Springer, 1997
- [5] FERRO C.A.T., PEZZULLI S. - "Rao-Blackwellised Estimators for Parameters of Extreme-Value Models"
- [6] GARRIDO M. - "Modélisation des événements rares et estimation des quantiles extrêmes, méthodes de sélection de modèles pour les queues de distribution" - Université Joseph Fourier, Grenoble I, 2002
- [7] HOLDEN L., HAUG O. - "A mixture model for unsupervised tail estimation" - 2009
- [8] HOSKING J.R.M., WALLIS J.R., WOOD E.F. - "Estimation of the Generalized Extreme-Value distribution by the method of Probability-Weighed moment" - *Technometrics*, 1985
- [9] HOSKING J.R.M., WALLIS J.R. - "Parameter and quantile estimation for the Generalized Pareto Distribution" - *Technometrics*, 1987
- [10] LEADBETTER M.R. - "On high level exceedance modeling and tail inference" - *Journal of Statistical Planning and Inference*, 1993
- [11] PICKANDS J. - "Statistical inference using extreme order statistics" - *The Annals of Statistics*, 1975

Implementing an internal model using Extreme Value Theory

Xavier Servel Sylvie Tran

May 29, 2012

Abstract

This article describes several methods to compute extreme quantiles within the Solvency II framework. Methods derived from Extreme Value Theory are developed so as to value these quantiles and we show how they can be used in a partial internal model. We base our work on historical data of the IPD UK index, the CAC 40 index, interest rates and credit spreads. We model these series using *ARMA – GARCH* processes with residuals that follow a Hybrid Pareto law. It allows us to obtain the distributions of the risk factors over a one-year time horizon that, after applying a correlation structure to them, can be used in an Economic Scenario Generator.

KEYWORDS : Solvency II, Extreme Value Theory, GEV, GPD, GARCH, GPD Test, Hybrid Pareto

1 Introduction

The Solvency II Directive framework requires insurance companies to meet capital requirements so that they can face extreme events: they are required to hold a sufficient amount of equities to insure solvency at a 99.5% confidence level over a one-year time horizon. Within this strict regulatory context, insurers can either apply a standard formula given by regulators or develop an internal model which would take into account their own risk profile. In the latter case, insurers have to estimate the extreme quantiles of the distribution of their risk factors using adequate statistical methods.

These methods will have to be closely monitored and validated by the French Prudential Supervisory Authority (*Autorité de Contrôle Prudentiel*). Thus, insurance companies have to demonstrate the reliability and robustness of their estimation methods. This is peculiarly difficult, taking into account that losses stemming from extreme events are often far more important than maximal losses observed in historical data, which usually cover a 10 to 20 year period.

Our purpose is to build a framework that includes practical methods and tools derived from Extreme Value Theory, that would allow us to estimate the extreme quantiles within an internal model framework for market risk.

We will first focus on the main results of Extreme Value Theory. Following this analysis, we will study how to apply these methods to financial data. Eventually, the results will be analyzed.

2 Extreme Value Theory

In order to measure an extreme quantile, we choose to model the distribution tail using Extreme Value Theory. We will here describe the main results of this theory.

2.1 Generalized Extreme Value distribution

Extreme Value Theory is derived from the Fisher-Tippett-Gnedenko theorem.

Let $(X_i)_{i=1,\dots,n}$ be a sequence of independent and identically-distributed random variables (i.i.d), with the cumulative distribution function F . Let $(M_i)_{i=1,\dots,n}$ be a sequence such that $M_n = \max(X_1, \dots, X_n)$. If there exist two sequences of real numbers $(a_i)_{i=1,\dots,n}$ and $(b_i)_{i=1,\dots,n}$ and a non-degenerated cumulative distribution function H such that $\frac{M_n - b_n}{a_n}$ converges in law to H , then H belongs to the following family:

$$H_\xi(x) = \begin{cases} e^{-(1+\xi x)^{-\frac{1}{\xi}}} \mathbb{1}_{1+\xi x > 0} & \text{if } \xi \neq 0 \\ e^{-e^{-x}} & \text{if } \xi = 0 \end{cases}$$

H_ξ is the Generalized Extreme Value (GEV) and ξ is called the shape parameter. This parameter determines the heaviness of the tail of the distribution. The sub-families defined by $\xi = 0$, $\xi > 0$ and $\xi < 0$ correspond, respectively, to the domain of attraction of Gumbel (thin-tailed distribution), Fréchet (fat-tailed distribution) and Weibull (finite-tailed distribution).

We can use this theorem in practice to estimate our quantile q from an i.i.d sample, by studying the distribution of M_n for a fixed n . Within this framework, a_n and b_n are then constants that we will now call σ and μ

In order to do that, we divide the sample into k r -sized blocks and we consider the sub-sample built with the maxima of each block. For a sufficiently high r , we assume that the distribution of M_r is a GEV distribution, with parameters σ and μ . We can then make the following approximation of the quantile of order p :

$$\hat{q} \simeq \hat{\mu} + \frac{\hat{\sigma}}{\hat{\xi}} ((-r \ln(p))^{-\hat{\xi}} - 1)$$

Thus, having an estimate of the parameters of the GEV law allows us to obtain an estimate of the extreme quantile.

2.1.1 Estimation of the parameters of Generalized Extreme Value distribution

There are several methods to estimate the GEV parameters. The most classical methods are the Maximum Likelihood Estimation (or MLE) and the Probability-Weighted Moment Estimation (or PWM). We will use other recently developed methods as well, such as the Generalized Probability-Weighted Moment estimation (GPWM) which aims at extending the application field of the PWM (see [3]) and the "Rao-Blackwell" method which leads to better quality estimators *via* a lowered variance (see [5]).

Moreover, we can build confidence intervals for these estimators based on their asymptotical normality (see [8]) or on the profile likelihood (see [2]).

However, this block maxima approach is wasteful of information, when data relevant to extremes (other than the block maxima) are also available. In order to avoid this problem, the Peaks-Over-Threshold (POT) method using a high threshold u has been developed.

2.2 Generalized Pareto Distribution

This method is based on the Balkema-de Haan-Pickands theorem, which consists in analyzing the sample of excess values that are above an appropriate threshold u .

Let X be an i.i.d sample with the cumulative distribution function F . We want to study the sample of the excesses $X_i - u$ for all i such that $X_i > u$. Using the Balkema-de Haan-Pickands theorem, if F verifies the first theorem hypothesis, the limiting conditional distribution of the excesses over a sufficiently high threshold follows the Generalized Pareto Distribution (GPD), which cumulative distribution function is given as:

$$G_{\xi,\sigma}(x) = \begin{cases} (1 - (1 + \frac{\xi x}{\sigma})^{-\frac{1}{\xi}}) \mathbf{1}_{x>0} & \text{if } \xi > 0 \\ (1 - (1 + \frac{\xi x}{\sigma})^{-\frac{1}{\xi}}) \mathbf{1}_{x \in]0, \frac{-\sigma}{\xi}[} & \text{if } \xi < 0 \\ (1 - e^{-\frac{x}{\sigma}}) \mathbf{1}_{x>0} & \text{if } \xi = 0 \end{cases}$$

We can also estimate the quantile q using this approximation:

$$q \simeq u + \frac{\hat{\sigma}}{\hat{\xi}} \left(\left(\frac{n}{N_u} (1-p) \right)^{-\hat{\xi}} - 1 \right)$$

where N_u is the number of exceedances.

Finding an estimate of the quantile thus requires to find the threshold u and to estimate the GPD parameters.

2.2.1 The choice of the threshold u

The choice of the extreme threshold u , where the GPD model provides a suitable approximation of the excess distribution is a critical problem with the POT method. The choice of the threshold should be large enough to satisfy the conditions to allow its application (u tends towards infinity), while at the same time leaving sufficient observations for the estimate. Different methods are examined in this section.

Mean excess function The mean excess function is a tool popularly used to aid this choice of u :

$$e(u) = \mathbb{E}(X - u | X > u)$$

This method consists of finding u such that the function starts to be linear. Indeed, if X follows a GDP, the mean excess function is linear in u .

Hill Plot The Hill plot displays an estimate of the tail index $\frac{1}{\xi}$ (which is valid for $\xi > 0$) for different exceedance levels. If u is too small this estimator is biased, thus u is chosen as the lowest threshold such that the Hill estimator is stabilized (see [4]).

Cunnane Test The number of excess values above the threshold u asymptotically follows a Poisson law. We use this result to determine the threshold, knowing the fact that a Poisson law is close to a normal law for a sufficiently high parameter (see [10]).

Minimisation of the Kolmogorov-Smirnov distance - Pickands' method This method is similar to the previous one, except that here we use the fact that the distribution of the conditional excesses is close to a GPD to find the threshold. We thus want to minimize the Kolmogorov-Smirnov distance (see [11]).

2.2.2 Estimation of the GPD Parameters

Similarly to the GEV, the GPD parameters can be determined by the Maximum Likelihood (ML) method or by the PWM method.

2.3 Simulation of i.i.d samples and comparison of the different methods

After implementing all the methods above, we tested them on theoretical samples from well-known laws, estimating several quantiles. The following table displays the theoretical scope of each estimator:

	GEV			GPD	
	EMV	PWM	GPWM	EMV	PWM
Convergence of the estimator	any ξ	$\xi < 1$	$\xi < 2$	any ξ	$\xi < 1$
Asymptotic normality	$\xi > -\frac{1}{2}$	$\xi < \frac{1}{2}$	$\xi < \frac{3}{2}$	$\xi > -\frac{1}{2}$	$\xi < \frac{1}{2}$

We chose to study three laws : the standard normal distribution ($\xi = 0$), the Student law with 2 degrees of freedom ($\xi = 0.5$) and the Student law with 1 degree of freedom ($\xi = 1$). We compute the estimators of the 95% and 99.5% quantiles for each of these laws using three different sample sizes: $n = 100, 1000$ and 10000 . We then study the quadratic errors that are obtained and the quality of the confidence intervals.

This study allows us to emphasize the impact of the data sample size and of the heaviness of the distribution tail on the quantile estimation methods. As expected, the quadratic errors are weaker in three cases: when the sample size is important, when ξ is close to 0 and when the order of the computed quantile is not "too high". For small sample sizes, the PWM method reduces the bias and the quadratic error. It is especially true when ξ is increasing. However, for bigger samples, the MLE are better. When ξ is high, the GPWM gives good results as well.

Regarding the block maxima approach, the rao-blackwellised estimators reduce the quadratic errors. We thus believe it is best to use the rao-blackwellised estimators when estimating the parameters of the GEV distribution. When using the excess approach, using the Pickands' method to choose the threshold seems to be the most stable method. In addition, this approach gives better results than the block maxima approach since there is more values in the sample. Same conclusions can be deduced for confidence intervals.

Using these results, we can now define a methodology when dealing with real financial data. Firstly, we will choose a block size in accordance to the database structure and will select the threshold by comparing the different methods. If they give different results, we will use the threshold given by the Pickands' method. We will then estimate ξ using the different procedures and according to its values, will check which methods can be applied. If the sample size is small, we will first apply the PWM method when $\xi < 1$ (or the GPWM method for the block approach when $\xi > 1.5$). Thus we will deduce estimates of ξ and of the other parameters, which will allow us to compute the quantile estimator.

2.4 GPD Test

Another method to compute the parametric quantile can be developed using the GPD test (see [6]), which checks the goodness-of-fit of the distribution tail of $(X_i)_{i=1,\dots,n}$ to the family of cumulative distribution functions F_θ . The test is performed as below:

$$H_0 : X \in (F_\theta)_{\theta \in \Theta} \qquad H_1 : X \notin (F_\theta)_{\theta \in \Theta}$$

The goal is then to compute an estimator \hat{q}_{GPD} of the p -quantile q using the excess approach. We then have to compare this value to the computed parametric quantile using the MLE estimator $\hat{\theta}$:

$$\hat{q}_{param} = F_{\hat{\theta}}^{-1}(1 - p)$$

The test is based on the simple idea that if X can be approximated by $F_{\hat{\theta}}$, then the difference between the quantile of X estimated using Extreme Value Theory and the parametric quantile is close to the same difference obtained just by replacing X by a sample with cdf $F_{\hat{\theta}}$. If it is not the case (i.e. the approximation error is not in the confidence interval of 95%), we reject H_0 .

We applied this test to the following mixing law: it is a standard normal law in 99% of the cases and a normal law with a mean that is equal to 3 and a standard deviation that is equal to 2 in 1% of the cases. Although the Kolmogorov-Smirnov test accepted the hypothesis that the data came from the normal distribution, the GPD test allowed us to formally reject it. This emphasizes the fact that usual adequation tests poorly take into account the distribution tails. Using such tests is not sufficient when computing an extreme quantile. Indeed, it is necessary to pay attention to distribution tails by using tests such as the GPD test.

3 Applying Extreme Value Theory to financial data

We saw in the previous section that Extreme Value Theory is a powerful tool that allows a computation of extreme quantiles by modelling the distribution tails. It will be applied to find the shocks that are used to compute the market SCR. The database that will be used is spread over a ten year period and is based on the values of the IPD UK index, the CAC 40 index, the 1 to 30 year interest rates and the credit spread of bonds that are rated from AAA to BBB (data from Bloomberg).

These series do not have any deterministic trend. Moreover, the Phillips-Perron test indicates that they are non-stationary (at the 5% confidence level). We therefore transform them into stationary series by taking the log-returns and name each new series X_t .

Our objective is to compute a downward shock for equities and for the property index, and an upward shock for credit spreads. As for rates, we will compute upward as well as downward shocks.

3.1 Modelling of the margin law of each risk factor

The analysis of the data gives us strong evidence of heteroskedasticity. We thus use $ARMA(P, Q) - GARCH(p, q)$ models:

$$\begin{cases} X_t = \mu + \sum_{i=1}^P a_i X_{t-i} + \sum_{j=1}^Q b_j \epsilon_{t-j} + \epsilon_t \\ \epsilon_t = \eta_t \sqrt{h_t} \\ h_t = \omega + \sum_{i=1}^q \alpha_i \epsilon_{t-i}^2 + \sum_{j=1}^p \beta_j h_{t-j} \end{cases}$$

For each serie we apply the simplest model where residuals are independent and where the coefficients of the highest order lags are significant. It is commonly assumed that the innovation process is following a standard normal distribution. However we empirically find that the distribution of the innovations (η_t) is heavier-tailed. Indeed, the Shapiro-Wilk test rejects the hypothesis of normality at the 5% level of confidence. That is why we decided to model the distribution of the implied model residuals using a Hybrid Pareto distribution (see [7]), which is an extension of the GDP to the whole real axis. This new distribution is built by adding a GDP tail to a Gaussian. The method is then to define a threshold u and a mixing zone $[u - \epsilon, u + \epsilon]$, and let f (the gaussian density), g (the GDP density) and l (the Hybrid Pareto density) be as follows:

$$l(x) = \begin{cases} \kappa f(x) & \text{if } x < u - \epsilon \\ \kappa(g(q(x))q'(x) + f(p(x))p'(x)) & \text{if } x \in [u - \epsilon, u + \epsilon] \\ \kappa g(x) & \text{if } x \geq u + \epsilon \end{cases}$$

where p et q are mixing functions and κ is chosen so that the integral of l equals 1. The Hybrid Pareto density is used to correct the kurtosis of the gaussian density, the latter usually underestimating the number of extreme values.

As in the interest rates case a computation of an upward and downward shock is required, we have to take into account extreme values in both distribution tails. We therefore use a double threshold Hybrid Pareto distribution, which is defined as a GPD below and above the thresholds and by a normal distribution between them.

3.2 Algorithm to obtain a one year distribution

We can obtain the one year distribution of X by estimating the Hybrid Pareto distribution. Indeed, as we know $(h_t, \epsilon_t, X_t)_{t \leq T}$ at time T , we can build $(h_{T+1}, \epsilon_{T+1}, X_{T+1})$ using a random generation of η that follows a Hybrid Pareto distribution (which parameters have been previously estimated):

$$\begin{cases} h_{T+1} = \hat{\omega} + \sum_{i=1}^q \hat{\alpha}_i \epsilon_{T-i+1}^2 + \sum_{j=1}^p \hat{\beta}_j h_{T-j+1} \\ \epsilon_{T+1} = \eta \sqrt{h_{T+1}} \\ X_{T+1} = \hat{\mu} + \sum_{i=1}^P \hat{a}_i X_{T-i+1} + \sum_{j=1}^Q \hat{b}_j \epsilon_{T-j+1} \end{cases}$$

Then by repeating the process, we obtain a possible value of X in a year. Using this algorithm N times, we obtain the distribution of our variable in one year. The 0.5% and 99.5% quantiles can then be deduced. We take $N = 50\,000$, therefore the estimator of the quantile is the classic one, obtained by taking the 250th lowest or highest value. Indeed, when the sample size is important, Extreme Value Theory is not relevant.

3.3 Dependence structure of risk factors

We saw that we could estimate the marginal law of each risk factor by modelling the evolution of the different underlyings (property index, equities, rates and spreads) using *ARMA – GARCH* models with residuals that follow a Hybrid Pareto distribution. These models allow us to calculate an estimator of the extreme quantiles of the distribution of the underlying in a year. However, we also have to model the dependency between the different risk factors in order to build an internal model.

We then start to take into account the correlation between the interest rates of different maturity and the spreads. In order to assess this dependency structure, we apply a widely used method, the Gaussian copula which consists in correlating the residuals of the *ARMA – GARCH* models.

Finally, we aggregate the different risk factors using the correlation matrix given by the QIS 5. This allows us to model the underlyings in the future using parameters which are given at a specified date, thus giving us real-world scenarios.

4 Results

Below are the results that are obtained:

Maturities	1 year	2 years	3 years	4 years	5 years	7 years	10 years	15 years	20 years	30 years
Shock 0,5% (down)	-77%	-80%	-79%	-80%	-77%	-74%	-68%	-54%	-52%	-50%
Shock 99,5% (up)	+39%	+56%	+37%	+34%	+29%	+18%	+18%	+5%	+11%	+15%
Shock 0,5% (down) QIS 5	-75%	-65%	-56%	-50%	-46%	-39%	-31%	-27%	-29%	-30%
Shock 99,5% (up) QIS 5	+70%	+70%	+64%	+59%	+55%	+49%	+42%	+33%	+26%	+25%

Table 1: One year shocks applied to interest rates

	spread AAA	spread AA	spread A	spread BBB	Property index	Equities
Estimated shocks	1.41%	2.99%	5.28%	19.28%	-15.50%	-43.50%
QIS 5 shocks	0.90%	1.10%	1.40%	2.50%	-25%	-30%

Table 2: One year shocks applied to spreads, property index and equities

As for interest rates, we notice that downward shocks are far more important than the ones found in QIS 5. However, the contrary phenomenon is observed for upward shocks. This can be explained by the sensitivity of our model to the last values of rates in our database, which displayed a strong downward trend, thus explaining this asymmetry of our estimators.

The shocks applied to spreads and equities are far more important than the ones found in QIS 5 as well. These shocks take into account the strong volatility that is observed on the markets for more recent data.

However our results must be interpreted cautiously. Computed shocks are indeed dependent of the choice of models and hypothesis. In our paper, we dealt with rather simple models in order to focus solely on Extreme Value Theory, but more refined models can be used. Other reasons can explain the difference with the shocks in QIS 5: the series that are used can be different and do not necessarily cover the same time frame. Shocks that are computed on different dates can be very different as market conditions are changing.

5 Conclusion

We developed methods using Extreme Value Theory to compute extreme quantiles in order to assess the capital required for risk market. Working on the correlation allows us to build an Economic Scenario Generator.

We also used a goodness-of-fit test that focused on the distribution tail and we showed that standard test could lead to a bad estimate of extreme quantiles.

However, the proposed methods can be improved regarding several aspects. For example, other models that are more sophisticated can be chosen to describe the evolution of our indices. Other dependency structures can be used as well.

Bibliography

- [1] BEIRLANT J., GOEGBEUR Y., TEUGELS J., SEGERS J. - *Statistics of Extreme* - Wiley, 2004
- [2] COLES S. - *An Introduction to Statistical Modeling of Extreme Values* - Springer, 2001
- [3] DIEBOLT J., GUILLOU A., NAVEAU P., RIBEREAU P. - "Improving Probability-Weighed Moment methods for the Generalized Extreme-Value distribution" - *Revstat Statistical Journal*, 2008
- [4] EMBRECHTS P., KLÜPPELBERG C., MIKOSCH T. - *Modelling Extremal Events for Insurance and Finance* - Springer, 1997
- [5] FERRO C.A.T., PEZZULLI S. - "Rao-Blackwellised Estimators for Parameters of Extreme-Value Models"
- [6] GARRIDO M. - "Modélisation des évènements rares et estimation des quantiles extrêmes, méthodes de sélection de modèles pour les queues de distribution" - Université Joseph Fourier, Grenoble I, 2002
- [7] HOLDEN L., HAUG O. - "A mixture model for unsupervised tail estimation" - 2009
- [8] HOSKING J.R.M., WALLIS J.R., WOOD E.F. - "Estimation of the Generalized Extreme-Value distribution by the method of Probability-Weighed moment" - *Technometrics*, 1985
- [9] HOSKING J.R.M., WALLIS J.R. - "Parameter and quantile estimation for the Generalized Pareto Distribution" - *Technometrics*, 1987
- [10] LEADBETTER M.R. - "On high level exceedance modeling and tail inference" - *Journal of Statistical Planning and Inference*, 1993
- [11] PICKANDS J. - "Statistical inference using extreme order statistics" - *The Annals of Statistics*, 1975

Remerciements

Nous tenons à remercier Grégory BOUTIER et Thameur ZGHAL pour leur encadrement, leur disponibilité et leurs conseils.

Résumé

Dans le cadre de Solvabilité II, les acteurs du marché de l'assurance sont amenés à s'intéresser à des chocs se produisant une fois tous les deux cents ans sur leurs divers facteurs de risque (actions, taux, immobilier et spreads par exemple pour le risque de marché). Afin de quantifier leurs risques, ils doivent ainsi calculer des quantiles extrêmes. Une estimation précise de ces quantités peut se faire en ayant recours à la théorie des valeurs extrêmes, que nous mettons en œuvre dans ce mémoire.

Dans le cadre de cette théorie, plusieurs méthodes peuvent être appliquées. Nous les testons toutes sur divers échantillons de lois connues pour juger de leur pertinence selon la nature des données. Ensuite, nous proposons un test d'adéquation à une loi se basant sur les queues de distribution, permettant de compléter les tests d'adéquation plus classiques. Cela nous assure que nos modélisations conservent une proximité avec l'échantillon étudié en queue de distribution, partie qui nous intéresse particulièrement.

A l'aide de ces résultats, nous montrons comment la théorie des valeurs extrêmes peut être utilisée dans le cadre d'un modèle interne : nous modélisons les différents facteurs de risque à l'aide de processus $ARMA - GARCH$ avec des résidus suivant une loi de Pareto Hybride ; nous obtenons ainsi les distributions complètes des facteurs de risque dans un an qui pourront être utilisées dans un générateur de scénarios économiques, après avoir été corrélées entre elles.

Mots-clés : Solvabilité II, théorie des valeurs extrêmes, GEV, GPD, GARCH, GPD Test, Pareto Hybride

Abstract

Within the Solvency II Directive framework, insurance companies are required to assess the impact of the shocks that may happen once every 200 years on different risk factors (e.g equities, interest rates, property index and credit spreads). In order to value these risks, they thus have to compute extreme quantiles. A precise estimate of these shocks can be given using Extreme Value Theory, which is applied in this paper. Different methods are derived from this theory. We test them all on samples derived from usual probability laws in order to check their reliability according to the properties of the database.

We then propose a goodness-of-fit test that focuses on the distribution tails, which is introduced to complete more classical goodness-of-fit tests. This new testing method assures that our modeling does fit the distribution tails which is our area of main interest in a distribution.

Using these theoretical results, we show how Extreme Value Theory can be used within an internal model framework : we model the different risk factors using *ARMA-GARCH* processes with residuals that follow Hybrid Pareto distributions. It allows us to obtain the distributions of the risk factors over a one-year time horizon that, after applying a correlation structure to them, can be used in an Economic Scenario Generator.

Keywords : Solvency II, Extreme Value Theory, GEV, GPD, GARCH, GPD Test, Hybrid Pareto

Table des matières

Introduction	4
I Contexte de notre étude	5
1 Cadre général : Solvabilité II	6
1.1 Pilier I : les exigences quantitatives en capital	6
1.1.1 Calcul des provisions économiques	6
1.1.2 Le Minimum de Capital Requis et le Capital de Solvabilité Requis	7
1.1.3 Les différents types de risques	9
1.1.4 Exigence en capital lié au risque de marché : le SCR_{Mkt}	10
1.2 Pilier II : gestion des risques et principes de gouvernance	14
1.3 Pilier III : discipline de marché	15
II Présentation de la théorie des valeurs extrêmes	16
2 Loi des valeurs extrêmes	17
2.1 Estimation d'un quantile grâce à la loi des valeurs extrêmes	19
2.2 Estimation des paramètres de la loi des valeurs extrêmes	20
2.2.1 Méthode du maximum de vraisemblance	20
2.2.2 Méthode des moments pondérés	20
2.2.3 Méthode des moments pondérés généralisés	21
2.2.4 Amélioration de l'estimation par méthode de Rao-Blackwell	21
2.3 Intervalles de confiance pour les quantiles extrêmes	22
2.3.1 Intervalles obtenus pour les estimateurs du maximum de vraisemblance	22
2.3.2 Intervalles obtenus par méthode des moments pondérés	24
2.3.3 Intervalles obtenus pour les estimateurs des moments pondérés généralisée	25
3 Loi de Pareto généralisée	27
3.1 Estimation d'un quantile grâce à la loi de Pareto généralisée	28
3.2 Estimation des Paramètres de la GPD	29
3.2.1 Valeur de ξ pour certaines lois connues	29
3.2.2 Méthode du maximum de vraisemblance	32
3.2.3 Méthode des moments pondérés	32
3.3 Recherche du Seuil u	33
3.3.1 Méthode 1 : utilisation de la fonction moyenne des excès	33
3.3.2 Méthode 2 : utilisation du graphique de Hill	35
3.3.3 Méthode 3 : utilisation du test de Cunnane	37
3.3.4 Méthode 4 : minimisation de la distance de Kolmogorov-Smirnov	38
3.4 Intervalles de confiance pour les quantiles extrêmes	39
3.4.1 Intervalles obtenus pour les estimateurs du maximum de vraisemblance	39

3.4.2	Intervalles obtenus pour les estimateurs des moments pondérés	40
4	Application de la théorie sur des données simulées	41
4.1	Estimation des paramètres de la loi des valeurs extrêmes	41
4.2	Estimation des paramètres de la loi de Pareto généralisée	43
4.3	Estimation des quantiles de loi connues	43
4.3.1	Loi Normale centrée réduite	44
4.3.2	Loi Student à 2 degrés de liberté	46
4.3.3	Loi Student à 1 degré de liberté	48
4.3.4	Analyse des résultats	50
5	Tests d'adéquations à une loi	51
5.1	ET Test	51
5.1.1	Utilisation de la loi limite d'un estimateur de quantile	51
5.1.2	Utilisation d'un bootstrap paramétrique	52
5.2	GPD Test	53
6	Extension aux séries stationnaires	55
6.1	Extension des résultats sur la loi des valeurs extrêmes aux séries stationnaires . .	55
6.1.1	Amélioration des estimateurs avec la rao-blackwellisation	56
6.2	Extension des résultats sur la loi des excès aux séries stationnaires	56
III	Applications	59
7	Estimation du quantile sur des données réelles	60
7.1	Définition des rendements et étude de leur stationnarité	60
7.1.1	Rendements trimestriels annualisés	60
7.1.2	Rendements annuels glissants	61
7.2	Estimation du quantile à 0,5 % des rendements trimestriels annualisés	62
7.2.1	Utilisation d'un modèle log-normal	62
7.2.2	Utilisation d'une loi Gumbel	64
7.2.3	Utilisation de la loi des valeurs extrêmes	65
7.2.4	Utilisation de la loi des excès	66
7.2.5	Analyse des résultats	68
8	Estimation des lois marginales des facteurs de risque	71
8.1	Loi de Pareto Hybride	71
8.1.1	Densité de la loi	71
8.1.2	Estimation des paramètres	74
8.1.3	Simulation d'un échantillon suivant une loi de Pareto Hybride	74
8.1.4	Ajout d'une troisième densité	74
8.2	Modélisation du facteur de risque immobilier	76
8.2.1	Modélisation des log-rendements par un modèle $AR - GARCH$	76
8.2.2	Obtention de la distribution à un an de l'immobilier	80
8.3	Modélisation du facteur de risque action	84
8.3.1	Modélisation des log-rendements par un modèle $AR - GARCH$	84
8.3.2	Obtention de la distribution à un an	89
8.4	Modélisation du facteur de risque taux	90
8.4.1	Utilisation d'une Analyse en Composantes Principales	92
8.4.2	Modélisation des taux maturité par maturité	102
8.5	Modélisation du facteur de risque lié aux variations de spread	106
8.6	Synthèse des résultats obtenus	114

8.7	Etude de la structure de dépendance entre les facteurs de risque	114
Conclusion		115
Bibliographie		117
Annexe		119
A	Calcul des moments pondérés	119
A.1	Loi des valeurs extrêmes	119
A.2	Loi de Pareto généralisée	120
B	Estimation de quantiles de lois connues	121
B.1	Loi normale centrée réduite	121
B.2	Loi de Student à 2 degrés de liberté	125
B.3	Loi de Student à 1 degré de liberté	128
C	Présentation des différents tests	132
C.1	Test de Kolmogorov-Smirnov	132
C.2	Test de Wald-Wolfowitz	132
C.3	Test de Ljung-Box	132
C.4	Test de Phillips-Perron	132
C.5	Test de Shapiro-Wilk	132
D	Résultats sur les composantes principales C_2 , C_3 , et C_4	133
D.1	C_2	133
D.2	C_3	133
D.3	C_4	134
E	Résultats pour la modélisation des taux maturité par maturité	136
E.1	Valeur et significativité des coefficients $ARMA(1,1) - GARCH(1,1)$	136
E.2	Tests d'indépendance des résidus	137
E.3	Tests d'adéquation à la loi de Pareto Hybride	138

Introduction

Depuis quelques années, les normes prudentielles relatives aux assureurs évoluent. À la norme Solvabilité I est en train de se substituer la norme Solvabilité II, mieux adaptée à la gestion des risques d'une compagnie d'assurance. Solvabilité II se décline en trois piliers. Le premier pilier est qualifié de "pilier quantitatif", il contient en effet les principes d'évaluation des divers éléments constitutifs du bilan économique (valeur des actifs, des provisions économiques, définition du capital éligible) d'une compagnie d'assurance ainsi que les exigences en termes de capital requis. Le deuxième pilier définit le cadre réglementaire dans lequel les entreprises gèrent les risques et organisent leur contrôle interne. De plus, il donne aussi un pouvoir renforcé aux superviseurs qui peuvent vérifier d'eux-mêmes le respect des engagements pris par l'assureur. Enfin, le troisième pilier oblige les compagnies d'assurance à publier régulièrement les informations essentielles concernant leur situation financière ainsi que leur solvabilité.

Le pilier qui va être l'objet de notre attention dans ce mémoire est le premier pilier, car il contient l'exigence en terme de capital requis : le Capital de Solvabilité Requis, généralement appelé SCR (pour *Solvency Capital Requirement*) qui est calculé comme un quantile à 0,5 % de la distribution de résultats à horizon un an. Or ce quantile est qualifié d'extrême au sens où il est atteint une fois tous les deux cents ans (en moyenne). Les données de l'assureur ne remontant pas aussi loin, une estimation fiable de ce quantile nécessite donc un recours à des techniques appropriées, comme la théorie des valeurs extrêmes.

Notamment réputée pour ses applications en hydrologie pour estimer la hauteur maximale pouvant être atteinte par un cours d'eau, celle-ci est aussi utilisée dans d'autres domaines comme la météorologie, la biologie, ou encore l'industrie du coton. Introduite en 1928 par FISHER et TIPPETT, cette théorie propose des résultats asymptotiques sur la distribution des maxima d'un échantillon de variables, ou encore sur la distribution des "dépassements" par rapport à un seuil élevé. Elle permet ainsi d'estimer de manière fiable des quantiles extrêmes comme celui demandé pour le calcul du SCR. L'objet de notre mémoire va être l'application de cette théorie pour une compagnie d'assurances dans le cadre d'un modèle interne.

Dans un premier temps, nous rappellerons le cadre réglementaire qui définit l'orientation de notre mémoire : Solvabilité II. Puis nous introduirons les théorèmes fondateurs et définirons les deux lois centrales de la théorie des valeurs extrêmes : la loi des valeurs extrêmes et la loi de Pareto généralisée. Nous verrons comment elles peuvent servir à estimer des quantiles extrêmes sur des échantillons de lois inconnues, ainsi que d'avoir des intervalles de confiance autour de ces valeurs. Nous étudierons aussi un test d'adéquation d'un échantillon à une loi paramétrique basé sur cette théorie, et se focalisant sur la proximité en queue de distribution. Ces éléments théoriques peuvent être appliqués à n'importe quels risques d'une compagnie d'assurance, nous les illustrerons sur des sous-jacents issus du risque de marché : immobilier, action, taux, actions et spreads de crédit. Nous montrerons que nos travaux pourront non seulement permettre de calculer un quantile extrême en modélisant la distribution marginale des facteurs de risque, mais aussi de construire un générateur de scénarios économiques.

Première partie

Contexte de notre étude

Chapitre 1

Cadre général : Solvabilité II

Dans ce chapitre, nous allons présenter les trois piliers de la réglementation Solvabilité II. Nous mettons l'accent sur le pilier I qui contient les éléments au cœur de ce mémoire, puisque celui-ci introduit les éléments clés pour évaluer la solvabilité d'une compagnie d'assurance.

1.1 Pilier I : les exigences quantitatives en capital

Le premier pilier instaure de nouvelles règles d'évaluation des éléments du bilan prudentiel des compagnies d'assurance :

	Solvabilité I	Solvabilité II
Evaluation actifs et passifs couvrables	Valeur comptable	Valeur de marché (<i>fair value</i>)
Evaluation passifs non couvrables	Valeur comptable	<i>Best Estimate</i> et marge de risque
Exigence de marge	Pourcentage des primes, des sinistres ou des provisions	Quantile à 0,5 % du résultat de la société

TABLE 1.1 – Méthode de valorisation du bilan prudentiel et de l'exigence de marge

Il fixe également les normes quantitatives de calcul des provisions techniques et des fonds propres.

1.1.1 Calcul des provisions économiques

On distingue deux types de risques : les risques couvrables - c'est-à-dire les risques qui peuvent être couverts ou répliqués par des instruments financiers - et les risques non couvrables.

Les provisions économiques pour risques couvrables correspondent au prix de la couverture financière permettant de construire une stratégie d'investissement autofinancée dans laquelle les instruments financiers répliquent parfaitement les flux futurs dans tous les états du monde.

Dans le cas de provisions techniques non couvrables, le calcul se fait en deux étapes :

- La Meilleure Estimation (ou *Best Estimate*) correspond à la moyenne pondérée par leur probabilité des flux de trésorerie actualisés en utilisant une courbe de taux pertinente. Il s'agit donc de la valeur actuelle des flux futurs de trésorerie.
- La Marge de Risque est une marge de prudence dans l'estimation des provisions économiques. Elle est égale au montant que l'on devrait ajouter au *Best Estimate* afin qu'un autre assureur accepte de racheter les engagements d'assurance. Elle est calculée comme un coût du capital.

1.1.2 Le Minimum de Capital Requis et le Capital de Solvabilité Requis

Solvabilité II introduit également deux mesures de garantie de la solvabilité :

- Le Minimum de Capital Requis (MCR) correspond au niveau minimum de fonds propres dont doit disposer une compagnie d'assurance. En dessous de ce niveau, l'intervention de l'autorité de contrôle - en France, l'Autorité de Contrôle Prudentiel - est automatique (retrait de l'agrément, plan de redressement, etc.). Il est calculé comme un pourcentage des provisions et des primes, et doit être compris entre 25 % et 45 % du Capital de Solvabilité Requis (d'après le QIS 5).
- Le Capital de Solvabilité Requis (SCR) est défini comme le montant cible de fonds propres économiques dont doit disposer la compagnie d'assurance pour que la probabilité de ruine soit inférieure à 0,5 % à horizon un an. Autrement dit, il s'agit des fonds propres nécessaires à la compagnie pour qu'elle soit solvable dans un an, dans 99,5 % des cas.

Le calcul du SCR dépend des types de risques auxquels sont exposées les compagnies d'assurance : risque de souscription vie et non-vie, risque de souscription santé, risque de marché, risque de crédit et risque opérationnel. Il est basé sur une formule standard fournie par les autorités de régulation, ou sur un modèle interne que les compagnies peuvent mettre en place et qui devra être validé par les autorités de contrôle.

Dans tous les cas, le niveau de confiance retenu de 99,5 % pour le SCR implique la nécessité d'estimer correctement un quantile d'ordre élevé, quantile que nous déterminerons dans ce mémoire en utilisant les outils que fournit la théorie des valeurs extrêmes.

Selon les spécifications du QIS 5, la formule standard définit le SCR de la manière suivante :

$$SCR = BSCR + Adj + SCR_{Op}$$

où :

- BSCR (ou *Basic Solvency Capital Requirement*) est le SCR de base défini par :

$$\sqrt{\sum_{ij} corr_{ij} \times SCR_i \times SCR_j} + SCR_{Intang}$$

- *Adj* sont les ajustements liés à l'effet d'absorption des risques des futures participations aux bénéfices et des impôts différés.
- SCR_{Op} est l'exigence de capital au titre du risque opérationnel, déterminée de manière forfaitaire en fonction des provisions mathématiques et des primes.

Le BSCR est déterminé en utilisant une méthode modulaire dite *Bottom-Up* dans laquelle le calcul se divise en modules de risque qui se décomposent eux-mêmes en sous-modules (comme le montre la formule ci-dessus) : dans un premier temps, on calcule à l'intérieur de chaque module, le capital requis au titre de des différents risques élémentaires puis ces capitaux sont agrégés à l'aide de matrices de corrélations et on obtient un SCR relatif à chaque module. Enfin, on agrège le tout avec une autre matrice de corrélation afin d'obtenir le BSCR.

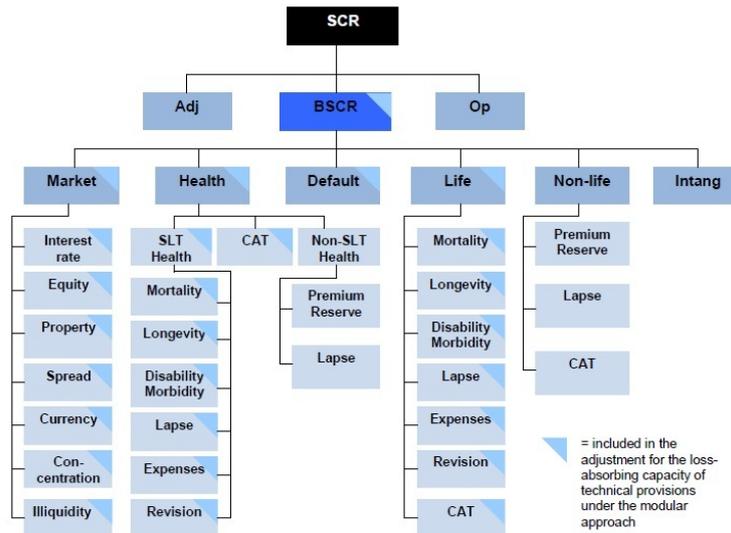


FIGURE 1.1 – Les modules de SCR d’après le QIS 5 - Formule Standard

Nous noterons :

- SCR_{Op} : exigence de capital au titre du risque opérationnel, déterminée de manière forfaitaire en fonction de différents éléments tels que le BSCR, les provisions techniques brutes de réassurance pour les engagements vie et non-vie et le montant des primes acquises.
- SCR_{Mkt} : exigence de capital au titre du risque de marché.
- SCR_{Health} : exigence de capital au titre du risque souscription santé.
- SCR_{Def} : exigence de capital au titre du risque de crédit.
- SCR_{Life} : exigence de capital au titre du risque souscription vie.
- SCR_{NL} : exigence de capital au titre du risque souscription non-vie.
- SCR_{Intang} : exigence de capital au titre du risque des actifs incorporels.

La corrélation entre les différents modules du SCR est prise en compte *via* la matrice de corrélation suivante fournie par le QIS 5 :

i \ j	Market	Default	Life	Health	Non-life
Market	1				
Default	0.25	1			
Life	0.25	0.25	1		
Health	0.25	0.25	0.25	1	
Non-life	0.25	0.5	0	0	1

FIGURE 1.2 – Matrice de corrélation des différents modules pour le calcul du BSCR - QIS 5

Il faut aussi prendre en compte des facteurs permettant de limiter le risque, comme la réassurance (mais qui en contrepartie ajoute un risque de défaut) ou les instruments financiers qui doivent respecter un certain nombre de principes (répertoriés par la Commission Européenne) pour pouvoir être pris en compte dans le calcul du SCR.

1.1.3 Les différents types de risques

a) Le risque de souscription

Il s'agit de l'ensemble des risques pris par l'assureur lors de la distribution de contrats d'assurance.

En assurance non-vie, on distingue plus précisément le risque de prime ou risque de sous-tarification, le risque de réserve lié à la nature aléatoire de l'évaluation des sinistres et à l'erreur d'estimation, et le risque catastrophe qui résulte d'événements extrêmes ou irréguliers (forte récurrence de sinistralité, catastrophes naturelles, etc.).

En assurance vie, c'est le risque de longévité, de mortalité, de morbidité et de rachat.

b) Le risque opérationnel

Ce risque est défini par l'EIOPA comme le "risque de pertes provenant de processus internes inadéquats ou défaillants, de personnes et systèmes ou d'événements externes". Il comprend donc les fraudes (interne ou externe), le dysfonctionnement de l'activité et des systèmes, les dommages aux actifs corporels, l'exécution et la gestion des processus et les pratiques en matières d'emploi et sécurité sur le lieu de travail.

c) Le risque de crédit

Il s'agit du risque de pertes lié au défaut ou à la dégradation de la qualité financière d'une contrepartie. Dans le cadre d'une compagnie d'assurance, il peut s'agir d'un réassureur ou d'un courtier.

d) Le risque de marché

Il s'agit du risque de pertes lors de mouvements adverses du marché. Dans les risques "quantifiables" identifiés par l'EIOPA, on distingue les risques liés à la baisse du marché des actions ou du marché immobilier, à la hausse ou baisse des taux, à une augmentation des spreads, aux taux de change ou aux concentrations. Le choix d'évaluer la solvabilité d'une entreprise dans

Solvabilité II à partir de la valeur économique de ses fonds propres rend la solvabilité apparente de l'entreprise encore plus exposée au risque de marché. Par la suite, nous nous intéresserons dans ce mémoire au module correspondant au risque de marché dont nous allons décrire les différents sous-modules dans le paragraphe suivant.

1.1.4 Exigence en capital lié au risque de marché : le SCR_{Mkt}

Le SCR_{Mkt} s'obtient en prenant en compte les chargements ci-dessous (toujours d'après le QIS 5) :

- $Mkt_{int,up}$: charge en capital relative au choc de hausse des taux.
- $Mkt_{int,down}$: charge en capital relative au choc de baisse des taux.
- Mkt_{eq} : charge en capital relative au choc sur les actions.
- Mkt_{prop} : charge en capital relative au choc sur les actifs immobiliers.
- Mkt_{sp} : charge en capital relative au choc de spread.
- Mkt_{conc} : charge en capital relative au choc de concentration.
- Mkt_{fx} : charge en capital relative au choc sur les taux de change.
- Mkt_{il} : charge en capital relative au choc sur la prime d'illiquidité.

Il se calcule en agrégeant les sous-risques avec la formule :

$$\max \left(\sqrt{\sum_{r,c} CorrMktUp_{r,c} \times Mkt_r^{up} \times Mkt_c^{up}}, \sqrt{\sum_{r,c} CorrMktDown_{r,c} \times Mkt_r^{down} \times Mkt_c^{down}} \right)$$

où les corrélations $CorrMktUp$ (corrélations pour les scénarios de hausse des niveaux de taux) et $CorrMktDown$ (corrélations pour les scénarios de baisse des niveaux de taux) sont définies par :

CorrMktUp	Interest	Equity	Property	Spread	Currency	Concentration	Illiquidity premium
Interest	1						
Equity	0	1					
Property	0	0.75	1				
Spread	0	0.75	0.5	1			
Currency	0.25	0.25	0.25	0.25	1		
Concentration	0	0	0	0	0	1	
Illiquidity premium	0	0	0	-0.5	0	0	1

FIGURE 1.3 – Matrice de corrélation pour le module risque de marché en cas de hausse des taux - QIS 5

CorrMktDown	Interest	Equity	Property	Spread	Currency	Concentration	Illiquidity premium
Interest	1						
Equity	0.5	1					
Property	0.5	0.75	1				
Spread	0.5	0.75	0.5	1			
Currency	0.25	0.25	0.25	0.25	1		
Concentration	0	0	0	0	0	1	
Illiquidity premium	0	0	0	-0.5	0	0	1

FIGURE 1.4 – Matrice de corrélation pour le module risque de marché en cas de baisse des taux - QIS 5

Pour chaque facteur de risque, la charge en capital est calculée en mesurant l'impact d'un choc associé au facteur de risque sur la NAV (ou *Net Asset Value*) qui correspond à la différence entre la valeur de l'actif et du passif.

a) Le risque de taux

Le risque de taux d'intérêt s'intéresse à l'impact résultant d'une modification de la structure par terme et/ou de la volatilité des taux d'intérêt sur l'actif et sur le passif. A l'actif, cela affecte la valeur des instruments financiers (par exemple les obligations ou les produits dérivés sur taux) et la rémunération des contrats en assurance-vie et modifie au passif le taux d'actualisation utilisé pour évaluer les engagements de l'assureur (provisions techniques).

Le détail des chocs à la baisse et à la hausse de la courbe des taux sans risque est donné par l'EIOPA.

Maturity t (years)	relative change $s^{up}(t)$	relative change $s^{down}(t)$
0.25	70%	-75%
0.5	70%	-75%
1	70%	-75%
2	70%	-65%
3	64%	-56%
4	59%	-50%
5	55%	-46%
6	52%	-42%
7	49%	-39%
8	47%	-36%
9	44%	-33%
10	42%	-31%
11	39%	-30%
12	37%	-29%
13	35%	-28%
14	34%	-28%
15	33%	-27%
16	31%	-28%
17	30%	-28%
18	29%	-28%
19	27%	-29%
20	26%	-29%
21	26%	-29%
22	26%	-30%
23	26%	-30%
24	26%	-30%
25	26%	-30%
30	25%	-30%

FIGURE 1.5 – Chocs relatifs à la hausse et à la baisse des taux - QIS 5

Le capital requis pour ce sous-module est le maximum entre la variation de la NAV avec choc à la hausse et choc à la baisse.

b) Le risque action

Le risque action est le risque induit par les variations du niveau des actions, ainsi que par celles de leur volatilité. On distingue deux types d'actions :

- les actions correspondant à l'indice dit "global", qui sont les actions cotées dans les pays de l'EEE et de l'OCDE.
- les actions "autres" qui regroupent les actions cotées dans les pays émergents, les actions non cotées et les investissements alternatifs.

Pour l'indice "global", un choc de 39 % à la baisse est à appliquer tandis que pour l'indice "autres", il s'agit d'un choc de 49 % à la baisse. Il y a de plus un ajustement (appelé *dampener*) de plus ou moins 10 % en fonction des conditions de marché. A titre d'exemple, les chocs étaient de 30 % et 40 % à fin 2009. Pour chaque indice, le chargement en capital correspond à la variation de NAV observée suite au choc puis les deux quantités sont agrégés avec un coefficient de corrélation de 75 %.

Il est à noter que ce sous-module ne tient compte que du risque systémique lié aux actions et non du risque idiosyncratique correspondant à une mauvaise diversification du portefeuille des actions. Ce dernier risque est traité dans le risque de concentration.

c) Le risque immobilier

Il s'agit du risque lié à la volatilité du prix de l'immobilier. Dans le cadre du QIS 5, on distingue l'immobilier coté qui est traité de la même manière que les actions, et l'immobilier non coté. Pour ce dernier, un choc de 25 % à la baisse est à appliquer.

d) Le risque de change

Le risque de change concerne l'ensemble des titres financiers exprimés dans une autre devise que l'euro. Le choc à appliquer est de 25 % à la hausse et à la baisse de la valeur de la devise par rapport à l'euro.

e) Le risque de spread

Le risque de spread est la part de risque due à la volatilité des spreads de crédit sur la structure par terme des taux d'intérêts. Le spread correspond à l'écart de taux actuariel entre une obligation et une obligation souveraine. Le calcul du Mkt_{sp} est la somme de la charge de capital relative à chacun des trois types d'actifs concernés : obligations à taux fixe, produits structurés ou obligations à taux variable et dérivés de crédit.

Pour les obligations à taux fixe (hors obligations souveraines), le besoin en capital est défini par :

$$Mkt_{sp}^{bond} = \sum_i VM_i \times F^{up}(Sign_i) \times Duration_i$$

où :

- VM_i est la valeur de marché de l'exposition au risque i
- $Duration_i$ est la duration de l'obligation i (en fonction de la notation de l'obligation, des bornes de duration sont définies)
- $Sign_i$ est la signature de l'exposition au risque i (notation)
- F^{up} est le choc lié à l'augmentation du spread et dépend de la signature :

	F^{up}	Duration Floor	Duration Cap
AAA	0,9%	1	36
AA	1,1%	1	29
A	1,4%	1	23
BBB	2,5%	1	13
BB	4,5%	1	10
B or lower	7,5%	1	8
Unrated	3,0%	1	12

FIGURE 1.6 – Valeurs de F^{up} en fonction de la notation - QIS 5

Pour les obligations des membres de l'EEE, émises dans leur propre devise, le risque de

spread est considéré comme nul. Par contre, pour les pays hors de l'EEE qui émettent des obligations dans leur propre devise ou pour les banques centrales, F^{up} est défini ainsi :

	F^{up}	Duration Floor	Duration Cap
AAA	0%	--	--
AA	0%	--	--
A	1,1%	1	29
BBB	1,4%	1	23
BB	2,5%	1	13
B or lower	4,5%	1	10
Unrated	3,0%	1	12

FIGURE 1.7 – Valeurs de F^{up} en fonction de la notation pour des Etats hors EEE - QIS 5

f) Le risque de concentration

Dans le QIS 5, le risque de concentration est lié à l'accumulation d'expositions à une même contrepartie. Il est constitué d'une part de la volatilité supplémentaire incluse dans le portefeuille d'actifs concentrés, et d'autre part de l'accumulation des risque de défaut (les obligations souveraines ne sont pas comprises dans ce module). Les chocs appliqués dépendent du rating du titre et du niveau de concentration. Les actifs considérés pour ce risque sont les produits de taux et les actions ainsi que les produits dérivés de ces actifs et les actifs immobiliers.

g) Le risque de prime de liquidité

La prime d'illiquidité modifie le calcul de la provision *Best Estimate* en augmentant le taux d'actualisation des flux et fait par conséquent diminuer la valeur de cette provision. Un besoin en capital est donc nécessaire pour ce risque et il se calcule en appliquant un choc de 65 % à la baisse sur la prime.

1.2 Pilier II : gestion des risques et principes de gouvernance

Le pilier 2 complète les exigences quantitatives du pilier 1 par des exigences qualitatives. Il traite à la fois du contrôle par les autorités et du contrôle interne des compagnies et vise à s'assurer que :

- Les compagnies sont correctement gérées (gestion saine, prudente et efficace de l'activité), qu'elles sont capable d'évaluer, calculer et maîtriser leurs risques avec une quantité suffisante de capital à tout moment.
- Elles possèdent le niveau de capital requis pour faire face à leurs risques.

Le régulateur aura le pouvoir de contrôler la qualité des données et des procédures d'estimation et les systèmes mis en place pour mesurer et maîtriser les risques au cas où ils se matérialiseraient. Dans le cadre de la mise en oeuvre du processus de contrôle, les autorités de supervision pourront imposer une exigence supplémentaire de capital, si le profil de risque de l'entreprise

s'écarte significativement des hypothèses utilisées dans le calcul du SCR (formule standard ou modèle interne) ou si le système de gouvernance s'écarte significativement des standards établis.

1.3 Pilier III : discipline de marché

Le troisième pilier fixe les règles de transparence financière et de la communication publique des compagnies d'assurances. Le but de ce pilier est de renforcer la discipline de marché et d'harmoniser les informations diffusées dans les Etats membres à différents niveaux (assurés, marché ou institutionnels, autorités de contrôle).

Deuxième partie

Présentation de la théorie des
valeurs extrêmes

Chapitre 2

Loi des valeurs extrêmes

Dans le cadre Solvabilité II, nous devons calculer des quantiles extrêmes à 0,5 % ou 99,5 % afin de déterminer le capital requis. Cependant, la recherche de ces quantiles peut s'avérer difficile lorsque les lois de départ sont inconnues ou lorsque l'échantillon est de taille trop faible, comme c'est le cas pour les compagnies d'assurance qui disposent souvent de données sur des périodes limitées. Pour remédier à ce problème, nous décidons donc d'utiliser la théorie des valeurs extrêmes, introduite par TIPPETT lorsqu'il étudiait la résistance des fils de coton. Dans ce chapitre, nous exposons le résultat fondateur de la théorie des valeurs extrêmes, et en expliquons l'intérêt pour notre étude.

Historiquement, le premier résultat en théorie des valeurs extrêmes est à attribuer à FISHER et TIPPETT qui énoncèrent en 1928 le théorème de Fisher-Tippett-Gnedenko :

Théorème 1 (Fisher-Tippett-Gnedenko).

Soient $(X_i)_{i=1,\dots,n}$ une suite de variables aléatoires indépendantes et identiquement distribuées, de fonction de répartition F . Soit la suite $(M_i)_{i=1,\dots,n}$ définie ainsi : $M_n = \max(X_1, \dots, X_n)$. S'il existe deux suites de constantes $(a_i)_{i=1,\dots,n}$ et $(b_i)_{i=1,\dots,n}$, et une fonction de répartition H non-dégénérée telles que :

$$\lim_{n \rightarrow +\infty} \mathbb{P} \left(\frac{M_n - b_n}{a_n} \leq x \right) = H(x)$$

Alors H appartient à la classe de fonctions suivante :

$$H_\xi(x) = \begin{cases} e^{-(1+\xi x)^{-\frac{1}{\xi}}} \mathbb{1}_{1+\xi x > 0} & \text{si } \xi \neq 0 \\ e^{-e^{-x}} & \text{si } \xi = 0 \end{cases}$$

H_ξ est la loi des valeurs extrêmes (ou GEV pour *Generalized Extreme Values*), et ξ est appelé indice de valeurs extrêmes et il mesure l'épaisseur de la queue. On distingue trois cas selon les valeurs de cet indice :

- Si $\xi < 0$, on dit que F appartient au domaine d'attraction de Weibull.
- Si $\xi = 0$, on dit que F appartient au domaine d'attraction de Gumbel.
- Si $\xi > 0$, on dit que F appartient au domaine d'attraction de Fréchet.

La densité h_ξ de la loi des valeurs extrêmes s'en déduit par simple dérivation. La figure 2.1 présente les différentes formes de h_ξ possibles en fonction de ξ .

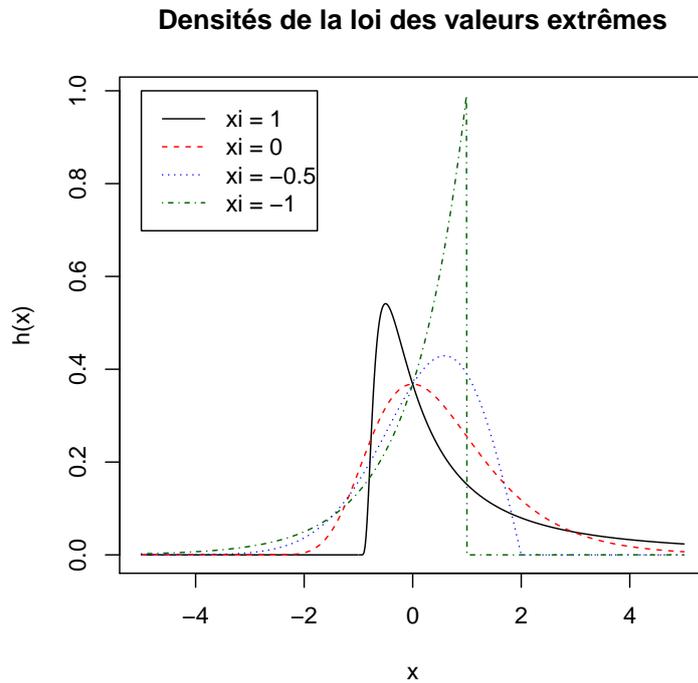


FIGURE 2.1 – Densités de la loi des valeurs extrêmes pour différentes valeurs de ξ

En pratique, un grand nombre de distributions vérifient les conditions du théorème. Les lois usuelles peuvent être classées selon leur domaine d'attraction :

WEIBULL	GUMBEL	FRÉCHET
$\xi < 0$	$\xi = 0$	$\xi > 0$
Uniforme, Bêta	Normale, Log-normale, Gamma, Exponentielle	Cauchy, Student, Pareto

Nous remarquons que le domaine d'attraction de Weibull contient les distributions à queue finie, c'est-à-dire pour lesquelles la fonction de survie est nulle à partir d'un certain point. Inversement, le domaine d'attraction de Fréchet accueille les fonctions à queue lourde, pour lesquelles la fonction de survie tend vers 0 comme une puissance de x .

En pratique, nous nous intéressons à la distribution de M_n pour un n fixé, a_n et b_n sont alors des constantes à estimer que nous appelons respectivement σ et μ . Nous faisons l'approximation :

$$P\left(\frac{M_n - \mu}{\sigma} \leq x\right) \simeq H_\xi(x)$$

$$P(M_n \leq \sigma x + \mu) \simeq H_\xi(x)$$

$$P(M_n \leq y) \simeq H_\xi\left(\frac{y - \mu}{\sigma}\right)$$

$$P(M_n \leq y) \simeq H_{\xi, \mu, \sigma}(y)$$

Nous définissons ainsi $H_{\xi,\mu,\sigma}(x) = H_{\xi}\left(\frac{x-\mu}{\sigma}\right)$, $H_{\xi,\mu,\sigma}$ est la fonction de répartition approchée de la variable M_n .

2.1 Estimation d'un quantile grâce à la loi des valeurs extrêmes

Il est possible d'estimer un quantile extrême en utilisant le résultat précédent sur la GEV. En effet, supposons que nous ayons un échantillon $(X_i)_{i=1,\dots,n}$ de variables i.i.d. de loi F et que nous séparions ces variables en k blocs de taille r . Si r est assez grand, nous pouvons alors faire l'hypothèse que la distribution de M_r est une distribution GEV de paramètres $\hat{\mu}$, $\hat{\sigma}$, et $\hat{\xi}$. Or la distribution de M_r n'est autre que F^r puisque les $(X_i)_{i=1,\dots,n}$ sont indépendants. Nous avons alors l'approximation suivante :

$$F^r(x) = H_{\hat{\xi},\hat{\mu},\hat{\sigma}}(x)$$

$$F^r(x) = \begin{cases} e^{-(1+\hat{\xi}\frac{x-\hat{\mu}}{\hat{\sigma}})^{-\frac{1}{\hat{\xi}}}} \mathbb{1}_{1+\hat{\xi}\frac{x-\hat{\mu}}{\hat{\sigma}} > 0} & \text{si } \hat{\xi} \neq 0 \\ e^{-e^{-\frac{x-\hat{\mu}}{\hat{\sigma}}}} & \text{si } \hat{\xi} = 0 \end{cases}$$

Nous cherchons q tel que $p = F(q)$, pour un quantile p élevé (par exemple 95 % ou 99,5 %). Dans la suite nous omettons la condition $1 + \hat{\xi}\frac{x-\hat{\mu}}{\hat{\sigma}} > 0$ car nous nous intéressons au cas $\hat{\xi} \geq 0$, cas dans lequel cette condition borne inférieurement la fonction de répartition (ce que nous voyons sur la figure 2.1). Cette condition est donc toujours vérifiée dès lors que nous cherchons un quantile élevé. Dans le cas $\hat{\xi} > 0$, il faut résoudre :

$$p^r = e^{-(1+\hat{\xi}\frac{\hat{q}-\hat{\mu}}{\hat{\sigma}})^{-\frac{1}{\hat{\xi}}}}$$

$$\Rightarrow -r \ln(p) = \left(1 + \hat{\xi}\frac{\hat{q}-\hat{\mu}}{\hat{\sigma}}\right)^{-\frac{1}{\hat{\xi}}}$$

$$\Rightarrow (-r \ln(p))^{-\hat{\xi}} = \left(1 + \hat{\xi}\frac{\hat{q}-\hat{\mu}}{\hat{\sigma}}\right)$$

Ce qui conduit finalement à :

$$\hat{q} = \hat{\mu} + \frac{\hat{\sigma}}{\hat{\xi}}((-r \ln(p))^{-\hat{\xi}} - 1) \quad (2.1)$$

Cela nous fournit donc une estimation d'un quantile extrême, pourvu que nous ayons estimé au préalable les paramètres de la loi GEV. Pour $\hat{\xi} = 0$ nous avons :

$$p^r = e^{-e^{-\frac{\hat{q}-\hat{\mu}}{\hat{\sigma}}}}$$

$$\Rightarrow -r \ln(p) = e^{-\frac{\hat{q}-\hat{\mu}}{\hat{\sigma}}}$$

$$\Rightarrow -\ln(-r \ln(p)) = \frac{\hat{q}-\hat{\mu}}{\hat{\sigma}}$$

Et donc :

$$\hat{q} = \hat{\mu} - \hat{\sigma} \ln(-r \ln(p)) \quad (2.2)$$

2.2 Estimation des paramètres de la loi des valeurs extrêmes

Les paramètres de la loi des valeurs extrêmes peuvent être estimés de plusieurs manières. Nous présentons ici différentes méthodes d'estimation que nous comparerons par la suite.

2.2.1 Méthode du maximum de vraisemblance

La méthode la plus classique est celle du maximum de vraisemblance consistant à maximiser la quantité :

$$L(\xi, \mu, \sigma, M_1, \dots, M_k) = \prod_{i=1}^k h_{\xi, \mu, \sigma}(M_i)$$

où $h_{\xi, \mu, \sigma}$ est la densité de la GEV. Nous pouvons écrire les dérivées partielles et les égaliser à 0, mais il n'existe pas de solution analytique à ce système. En pratique, il est plus aisé d'utiliser un algorithme cherchant à optimiser cette fonction.

2.2.2 Méthode des moments pondérés

Cette méthode consiste à obtenir les paramètres comme fonctions des moments $\omega_r = \mathbb{E}(MH_{\xi, \mu, \sigma}^r(M))$, où M suit une GEV de fonction de répartition $H_{\xi, \mu, \sigma}$ pour $r = 0, 1, 2$. En effet, nous pouvons montrer que pour $0 < \xi < 1$:

$$\omega_r = \frac{1}{r+1} \left(\mu - \frac{\sigma}{\xi} (1 - (r+1)^\xi \Gamma(1-\xi)) \right)$$

Et pour $\xi = 0$:

$$\omega_r = \frac{1}{r+1} (\mu - \sigma(\gamma - \ln(r+1)))$$

Γ est la fonction Gamma et γ est la constante d'Euler. Les démonstrations sont faites en annexe. Pour $\xi \geq 1$, les moments ne sont pas définis (la queue de distribution étant trop épaisse, ils ne sont pas finis).

HOSKING, WALLIS, et WOOD ont montré dans [HWW85] que ces quantités pouvaient être approchées par :

$$\hat{\omega}_r = \frac{1}{k} \sum_{j=1}^k \left(\frac{j - 0.35}{k} \right)^r M_{(j)}$$

où $M_{(1)} \leq M_{(j)} \leq M_{(k)}$ sont les maxima ordonnés.

Nous en déduisons alors les valeurs des paramètres :

$$\begin{cases} \frac{3\hat{\omega}_2 - \hat{\omega}_0}{2\hat{\omega}_1 - \hat{\omega}_0} &= \frac{3\hat{\xi} - 1}{2\hat{\xi} - 1} \\ \hat{\sigma} &= \frac{(2\hat{\omega}_1 - \hat{\omega}_0)\hat{\xi}}{\Gamma(1-\hat{\xi})(2\hat{\xi}-1)} \\ \hat{\mu} &= \hat{\omega}_0 + \frac{\hat{\sigma}}{\hat{\xi}}(1 - \Gamma(1-\hat{\xi})) \end{cases}$$

Ces estimateurs sont définis pour $\xi < 1$ mais ils ne sont asymptotiquement normaux que pour $\xi < \frac{1}{2}$. Nous utiliserons ce résultat dans la suite pour calculer des intervalles de confiance autour des quantiles estimés.

2.2.3 Méthode des moments pondérés généralisés

Cette méthode vise à généraliser la précédente en étendant son champ d'application. En effet, la méthode des moments pondérés perd de son intérêt lorsque ξ dépasse $\frac{1}{2}$ puisqu'il n'y a plus de normalité asymptotique de l'estimateur. Cette généralisation consiste à utiliser les moments $\nu_h = \mathbb{E}(Mh(H_{\xi,\mu,\sigma}(M)))$, où M suit une GEV de fonction de répartition $H_{\xi,\mu,\sigma}$. h est une fonction définie sur $[0, 1]$ et valant 0 sur les bornes. Ainsi, lorsque M prend une grande valeur, $H_{\xi,\mu,\sigma}(M)$ est proche de 1 et la fonction h ramène la valeur à 0. Intuitivement, nous voyons bien que l'impact des queues de distribution va être moins fort et donc que les moments existeront pour des valeurs de ξ plus élevées.

DIEBOLT a montré dans [DGNR08] que si nous choisissons $h(t) = h_{a,b}(t) = t^a(-\log(t))^b$ avec $a > 0$ et $b > 0$ alors ν_h vérifie :

$$\nu_h = \frac{\sigma}{\xi} \frac{1}{(a+1)^{b-\xi+1}} \Gamma(b-\xi+1) - \left(\frac{\sigma}{\xi} - \mu \right) \frac{1}{(a+1)^{b+1}} \Gamma(b+1)$$

En choisissant $(a, b) = (1, 1), (1, 2), (2, 1)$, nous pouvons ainsi obtenir des estimateurs des paramètres de la GEV avec les relations suivantes :

$$\begin{cases} \frac{\hat{\xi}}{1-(\frac{3}{2})^{\hat{\xi}}} &= \frac{2(\hat{h}_{1,1}-\hat{h}_{1,2})}{\hat{h}_{1,1}-2.25\hat{h}_{2,1}} \\ \hat{\sigma} &= (\hat{h}_{1,1}-\hat{h}_{1,2}) \frac{2^{3-\hat{\xi}}}{\Gamma(2-\hat{\xi})} \\ \hat{\mu} &= \frac{\hat{\sigma}}{\hat{\xi}} - \frac{\hat{\sigma}}{\hat{\xi}} 2^{\hat{\xi}} \Gamma(2-\hat{\xi}) + 4\hat{h}_{1,1} \end{cases}$$

Cette méthode est valable pour $\xi < \frac{3}{2}$, et la normalité asymptotique est vérifiée pour $\xi < 1$.

2.2.4 Amélioration de l'estimation par méthode de Rao-Blackwell

Soit $\hat{\phi} = (\hat{\xi}, \hat{\sigma}, \hat{\mu})$ estimateurs obtenus par maximum de vraisemblance. Il est possible d'améliorer ces estimateurs en effectuant N ($N \leq n!$) permutations sur notre échantillon i.i.d. de départ $(X_i)_{i=1,\dots,n}$. Nous obtenons alors les estimateurs dit "rao-blackwellisés" notés $\bar{\phi} = (\bar{\xi}, \bar{\sigma}, \bar{\mu})$ et définis par la formule suivante :

$$\bar{\phi} = \frac{1}{N} \sum_{i=1}^N \hat{\phi}_i$$

où $\hat{\phi}_i$ est obtenu par maximum de vraisemblance à partir de la i -ème permutation de l'échantillon.

Ces estimateurs sont intéressants car ils vérifient :

$$\begin{cases} \mathbb{E}(\bar{\phi}) &= \mathbb{E}(\hat{\phi}) \\ \mathbb{V}(\bar{\phi}) &= \frac{1}{N} \mathbb{V}(\hat{\phi}) + \frac{1}{N^2} \sum_{i \neq j} cov(\hat{\phi}_i, \hat{\phi}_j) \leq \mathbb{V}(\hat{\phi}) \end{cases}$$

La variance des estimateurs est donc réduite par rapport aux estimateurs classiques. De plus, nous pouvons définir deux estimateurs du quantile d'ordre p :

$$\begin{cases} \tilde{q} &= \bar{\mu} + \frac{\bar{\sigma}}{\bar{\xi}} ((-r \ln(p))^{-\bar{\xi}} - 1) \\ \bar{q} &= \frac{1}{N} \sum_{i=1}^N \hat{q}_i \end{cases}$$

où $\hat{q}_i = \hat{\mu}_i + \frac{\hat{\sigma}_i}{\hat{\xi}_i} ((r \ln(p))^{-\hat{\xi}_i} - 1)$ avec $(\hat{\xi}_i, \hat{\sigma}_i, \hat{\mu}_i)$ obtenus par maximum de vraisemblance à partir de la i^{eme} permutation de l'échantillon.

De la même manière, il est possible d'obtenir des estimateurs "rao-blackwellisés" basés sur la méthode des moments pondérés ou sur la méthode des moments pondérés généralisés.

A ce stade, nous sommes donc capables d'approcher la loi des maxima de notre échantillon par une loi des valeurs extrêmes, et d'en estimer les paramètres par diverses méthodes. Ainsi, en utilisant la formule 2.1 (ou 2.2) nous pouvons estimer la valeur d'un quantile extrême (p proche de 1) de notre échantillon de départ. Notre objectif est maintenant de mesurer la précision de notre estimateur, ce qui se fait par un calcul d'intervalle de confiance. Ceux-ci vont évidemment dépendre des paramètres que nous avons estimés, mais aussi de la méthode d'estimation utilisée.

2.3 Intervalles de confiance pour les quantiles extrêmes

2.3.1 Intervalles obtenus pour les estimateurs du maximum de vraisemblance

Intervalles obtenus à partir de la matrice de variance asymptotique

D'après SMITH dans [SMI84], lorsque $\xi > -\frac{1}{2}$, les estimateurs du maximum de vraisemblance sont asymptotiquement normaux :

$$\sqrt{k} \begin{pmatrix} \hat{\sigma} - \sigma \\ \hat{\xi} - \xi \\ \hat{\mu} - \mu \end{pmatrix} \xrightarrow{Loi} N \left(\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, \mathbb{V}_{as} \begin{pmatrix} \hat{\sigma} \\ \hat{\xi} \\ \hat{\mu} \end{pmatrix} \right)$$

où \mathbb{V}_{as} est l'inverse de l'information de Fisher I définie par :

$$\begin{aligned} I_{1,1} &= \frac{1}{\sigma^2 \xi^2} (1 - 2\Gamma(2 + \xi) + p_1) \\ I_{1,2} &= -\frac{1}{\sigma \xi^2} \left(1 - \gamma_* - p_2 + \frac{1 - \Gamma(2 + \xi)}{\xi} + \frac{p_1}{\xi} \right) \\ I_{1,3} &= -\frac{1}{\sigma^2 \xi} (p_1 - \Gamma(2 + \xi)) \\ I_{2,2} &= \frac{1}{\xi^2} \left(\frac{\pi^2}{6} + \left(1 - \gamma_* + \frac{1}{\xi} \right)^2 - \frac{2p_2}{\xi} + \frac{p_1}{\xi^2} \right) \\ I_{2,3} &= -\frac{1}{\sigma \xi} \left(p_2 - \frac{p_1}{\xi} \right) \\ I_{3,3} &= \frac{p_1}{\sigma^2} \end{aligned}$$

avec $\gamma_* = 0,577$ la constante d'Euler, $\Psi(x) = \frac{d \log \Gamma(x)}{dx}$ et :

$$\begin{aligned} p_1 &= (1 + \xi)^2 \Gamma(1 + 2\xi), \\ p_2 &= \Gamma(2 + \xi) \left(\Psi(1 + \xi) + \frac{1 + \xi}{\xi} \right) \end{aligned}$$

Or, l'estimateur du quantile d'ordre p s'écrit :

$$\begin{aligned} \hat{q} &= \hat{\mu} + \frac{\hat{\sigma}}{\hat{\xi}} ((-r \ln(p))^{-\hat{\xi}} - 1) \\ &= g(\hat{\sigma}, \hat{\xi}, \hat{\mu}) \end{aligned}$$

Nous pouvons donc en déduire avec la delta-méthode que q est asymptotiquement normal de variance asymptotique : $(\nabla g)' \nabla_{as}(\nabla g)$ en posant $(\nabla g)' = [\frac{\partial g}{\partial \sigma}, \frac{\partial g}{\partial \xi}, \frac{\partial g}{\partial \mu}]$ où :

$$\begin{cases} \frac{\partial g}{\partial \sigma} &= \frac{1}{\xi}((-r \ln(p))^{-\xi} - 1) \\ \frac{\partial g}{\partial \xi} &= -\frac{\sigma}{\xi^2}(c^{-\xi} - 1) - \frac{\sigma}{\xi}(\ln(c)c^{-\xi}) \\ \frac{\partial g}{\partial \mu} &= 1 \end{cases}$$

avec $c = -r \ln(p)$.

Il est alors aisé d'établir un intervalle de confiance asymptotique de niveau α pour q :

$$\left[\hat{q} \pm \frac{\sqrt{(\nabla g)' \nabla_{as}(\nabla g)}}{\sqrt{n}} z_{\frac{1+\alpha}{2}} \right]$$

où $z_{\frac{1+\alpha}{2}}$ est le quantile d'ordre $\frac{1+\alpha}{2}$ d'une loi normale centrée réduite.

Intervalles obtenus par méthode du profil de vraisemblance

Les intervalles de confiance classiques des estimateurs du maximum de vraisemblance reposent sur leur normalité asymptotique. Cependant, lorsque la taille de l'échantillon est faible, l'utilisation de cette normalité asymptotique est difficilement justifiable (à cause du biais). C'est pourquoi des auteurs comme COLES dans [COL01] recommandent l'utilisation d'intervalles de confiance basés sur le profil de vraisemblance. Ayant à notre disposition des échantillons de taille peu importante, cette méthode nous intéressera donc par la suite pour construire des intervalles de confiance pour le quantile q .

Reparamétrisons le modèle en écrivant μ sous la forme :

$$\mu = q - \frac{\sigma}{\xi}((-r \ln(p))^{-\xi} - 1),$$

et en injectant cette expression dans la fonction de logvraisemblance ℓ . La fonction de logvraisemblance dépend ainsi des paramètres (σ, ξ, q) , et nous pouvons alors définir le profil de vraisemblance associé à q :

$$\ell_p(q) = \max_{\sigma, \xi} \ell(\sigma, \xi, q)$$

De plus, nous avons le résultat asymptotique suivant :

$$2[\ell(\hat{\sigma}, \hat{\xi}, \hat{q}) - \ell_p(q)] \sim \chi_1^2$$

où $(\hat{\sigma}, \hat{\xi}, \hat{q})$ sont les estimateurs du maximum de vraisemblance obtenus en maximisant ℓ , et χ_1^2 est une loi du χ^2 à 1 degré de liberté.

Un intervalle de confiance asymptotique de niveau α pour q est alors donné par :

$$\left\{ q : \ell_p(q) \geq \ell_p(\hat{q}) - \frac{1}{2}c_\alpha \right\}$$

avec c_α quantile d'ordre α (par exemple, 95 %) d'une loi χ_1^2 .

En pratique, il suffit de tracer la fonction de profil de vraisemblance en fonction de q et de tracer la droite d'équation $y = \max_q \ell_p(q) - \frac{1}{2}c_\alpha$, l'intervalle de confiance correspondant aux

points d'intersection de la droite et de la courbe.

Par exemple, dans le cas d'une loi de Student à 2 degrés de liberté, nous trouvons un estimateur égal à 9,28 pour le quantile à 99,5 % (le quantile théorique étant égal à 9,925) et son intervalle de confiance donné par le profil de vraisemblance est : [8,45 ; 10,36]. Il contient donc la vraie valeur du quantile. Nous remarquons qu'à la différence des autres méthodes, cet intervalle n'est pas centré.

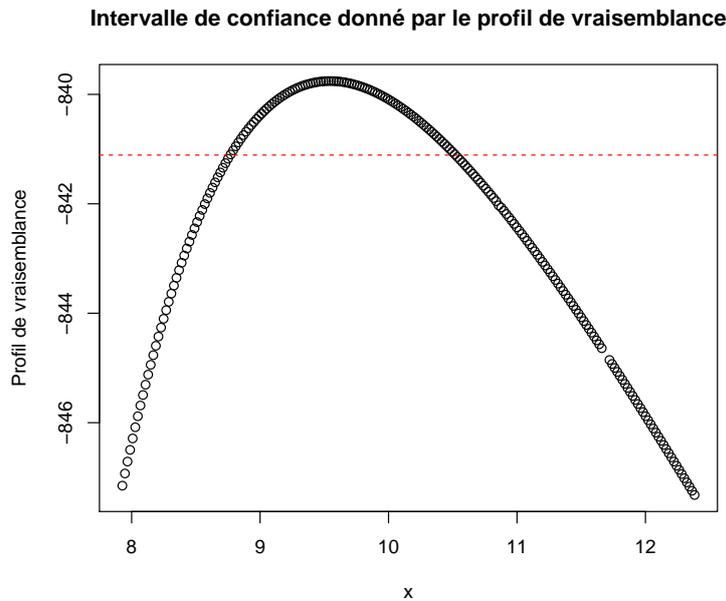


FIGURE 2.2 – Intervalle de confiance (de niveau 95 %) pour le quantile à 99,5 % d'une loi de Student à 2 degrés de liberté avec 1000 observations

2.3.2 Intervalles obtenus par méthode des moments pondérés

Intervalles obtenus à partir de la matrice de variance asymptotique

La distribution limite des estimateurs des moments pondérés se déduit de celle de $\hat{\omega} = (\hat{\omega}_0, \hat{\omega}_1, \hat{\omega}_2)$. HOSKING dans [HWW85] a montré la normalité asymptotique de ce vecteur sous la condition $\xi < \frac{1}{2}$ et sa matrice de variance asymptotique V est caractérisée par :

$$\begin{aligned}
 v_{r,r} &= \left(\frac{\sigma}{\xi} (r+1)^\xi \right)^2 \left(\Gamma(1-2\xi) K\left(\frac{r}{r+1}\right) - \Gamma(1-\xi)^2 \right) \\
 v_{r,r+1} &= \frac{1}{2} \left(\frac{\sigma}{\xi} \right)^2 \left[(r+2)^{2\xi} \Gamma(1-2\xi) K\left(\frac{r}{r+2}\right) + (r+1)^\xi ((r+1)^\xi - 2(r+2)^\xi) \Gamma(1-\xi)^2 \right] \\
 v_{r,r+2} &= \frac{1}{2} \left(\frac{\sigma}{\xi} \right)^2 \left[(r+3)^{2\xi} \Gamma(1-2\xi) K\left(\frac{r}{r+3}\right) - (r+2)^\xi \Gamma(1-2\xi) K\left(\frac{r+1}{r+2}\right) \right. \\
 &\quad \left. + 2(r+1)^\xi ((r+2)^\xi - (r+3)^\xi) \Gamma(1-\xi)^2 \right]
 \end{aligned}$$

avec $K(x) = F(-\xi, -2\xi, 1-\xi, -x)$ où F est la fonction hypergéométrique :

$$F(a, b, c, x) = \frac{1}{\int_0^1 t^{b-1} (1-t)^{c-b-1} dt} \int_0^1 t^{b-1} (1-t)^{c-b-1} (1-tx)^{-a} dt.$$

En appliquant la delta-méthode en version multidimensionnelle, nous pouvons en déduire la normalité asymptotique de $(\hat{\sigma}, \hat{\xi}, \hat{\mu})$ et calculer sa variance asymptotique. Cependant, comme il n'est pas possible d'obtenir directement une fonction exprimant $(\hat{\sigma}, \hat{\xi}, \hat{\mu})$ en fonction de $\hat{\omega}$, il est nécessaire de passer par des approximations numériques. Afin d'éviter ces approximations, il nous semble donc plus pratique de construire des intervalles de confiance utilisant le bootstrap non paramétrique que nous décrirons ici.

Intervalles obtenus à partir de bootstrap non paramétrique

La méthode du bootstrap non paramétrique est un autre procédé de construction des intervalles de confiance pour l'estimateur \hat{q} du quantile q . Il s'agit de générer un échantillon de taille n à partir de l'échantillon initial $(X_i)_{i=1, \dots, n}$ en procédant à des tirages aléatoires avec remise. Sur ce nouvel échantillon, il est alors possible de calculer un nouvel estimateur \hat{q}_1^* du quantile q . En répétant ces deux étapes B fois (en général, $B = 1\ 000, 5\ 000$ ou $10\ 000$), nous obtenons alors B estimateurs $(\hat{q}_j^*)_{j=1, \dots, B}$ de q que nous pouvons classer par ordre croissant. Nous pouvons donc en déduire un intervalle de confiance asymptotique de niveau α pour q dont les bornes sont définis par les percentiles $\frac{1-\alpha}{2} \%$ et $\frac{1+\alpha}{2} \%$ de l'échantillon des B estimateurs obtenus. Par exemple, pour $\alpha = 90\%$, les bornes de l'intervalle de confiance sont les percentiles 5% et 95% . Cette méthode peut être appliquée pour construire des intervalles de confiance pour tous les estimateurs. Nous l'utiliserons notamment pour les estimateurs provenant de la méthode PWM ainsi que des méthodes Rao-Blackwell pour lesquels nous n'avons pas d'expressions explicites de la variance.

2.3.3 Intervalles obtenus pour les estimateurs des moments pondérés généralisée

Les estimateurs des moments pondérés généralisés sont asymptotiquement normaux et DIEBOLT a montré dans [DGNR08] que la matrice de variance de $(\hat{\sigma}, \hat{\xi}, \hat{\mu})$ est de la forme : $M^{-1}\Omega(M^{-1})'$ où :

$$M = \begin{pmatrix} \frac{2^{\xi-2}}{\xi} \Gamma(2-\xi) - \frac{1}{4\xi} & \frac{\sigma}{\xi} 2^{\xi-2} [(\ln 2 - \frac{1}{\xi}) \Gamma(2-\xi) - \Gamma'(2-\xi)] + \frac{\sigma}{4\xi^2} & \frac{1}{4} \\ (\frac{2}{\xi} - 1) 2^{\xi-3} \Gamma(2-\xi) - \frac{1}{4\xi} & \frac{\sigma}{\xi} 2^{\xi-3} [(-\frac{2}{\xi} + \log 2(2-\xi)) \Gamma(2-\xi) - (2-\xi) \Gamma'(2-\xi)] + \frac{\sigma}{4\xi^2} & \frac{1}{4} \\ \frac{3^{\xi-2}}{\xi} \Gamma(2-\xi) - \frac{1}{9\xi} & \frac{\sigma}{\xi} 3^{\xi-2} [(\log 3 - \frac{1}{\xi}) \Gamma(2-\xi) - \Gamma'(2-\xi)] + \frac{\sigma}{9\xi^2} & \frac{1}{9} \end{pmatrix}$$

et Ω est une matrice définie pour $\xi < 1$, avec pour composantes :

$$\begin{aligned} \Omega_{ij} = & \sigma^2 \left[\int_0^1 \frac{1}{t} (-\log t)^{-\xi-1} h_i(t) \int_0^t (-\log s)^{-\xi-1} h_j(s) ds dt \right. \\ & + \int_0^1 (-\log t)^{-\xi-1} h_i(t) \int_t^1 \frac{1}{s} (-\log s)^{-\xi-1} h_j(s) ds dt \\ & \left. - \int_0^1 (-\log t)^{-\xi-1} h_i(t) dt \int_0^1 (-\log s)^{-\xi-1} h_j(s) ds \right], \quad i = 1, 2, 3, \quad j = 1, 2, 3 \end{aligned}$$

où $h_1 = h_{1,1}, h_2 = h_{1,2}, h_3 = h_{2,1}$.

Une approximation numérique de ces quantités nous permet de construire l'intervalle de confiance asymptotique du quantile lorsque $\xi < 1$.

Cette partie nous a permis de faire un tour d'horizon des différentes méthodes existantes pour calculer un intervalle de confiance. La question de savoir quelle méthode choisir dans quelle situation se pose alors. Nous essaierons d'y répondre dans une partie ultérieure.

En conclusion de ce chapitre, nous avons à notre disposition un outil permettant d'appréhender le comportement extrême de variables aléatoires, et qui peut servir à calculer des quantiles extrêmes sur des séries d'observations. Ceci est utile dans le cadre de la réglementation Solvabilité II, puisque si nous étions capable d'avoir un échantillon indépendant de réalisations du résultat à un an d'une compagnie d'assurance, la théorie des valeurs extrêmes nous permettrait de calculer de façon simple le SCR de cette compagnie. Nous disposerions aussi d'un intervalle de confiance autour de cette valeur. En réalité cette indépendance est difficile à obtenir comme nous le verrons plus loin.

Une limite de cette loi des valeurs extrêmes est qu'elle ne prend en compte que les maxima, et qu'elle peut donc ignorer une partie de l'information : par exemple, la situation n'est pas la même si le maximum est loin au-dessus des autres valeurs que s'il y a une dizaine d'autres valeurs très proches. Afin de remédier à ce défaut, des auteurs ont proposé de considérer non pas les maxima mais plutôt toutes les valeurs excédant un certain seuil. Cela permet aussi de prendre en compte davantage de valeurs qu'en ne prenant qu'un échantillon de maxima, et donc d'avoir des estimateurs plus robustes.

Chapitre 3

Loi de Pareto généralisée

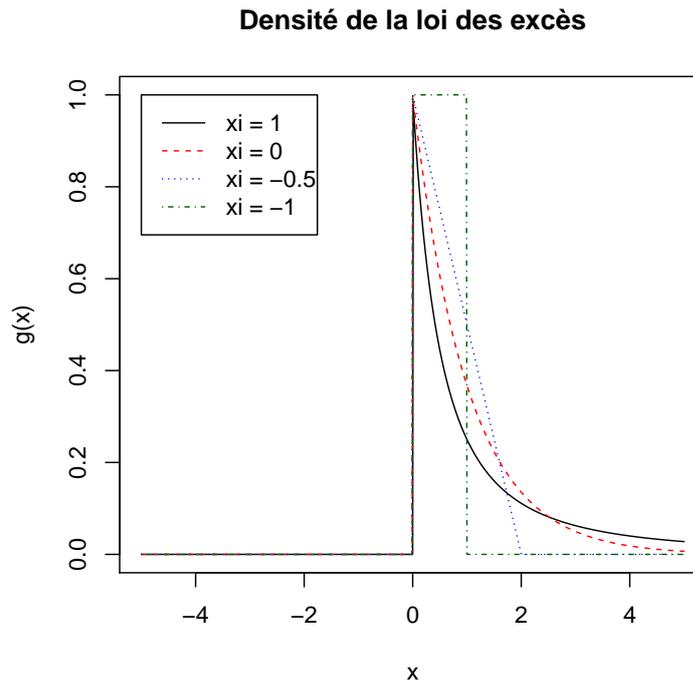
Il existe une autre approche de la théorie des valeurs extrêmes permettant d'estimer des quantiles. Pour cela, supposons que X (qui admet F comme fonction de répartition) puisse prendre au maximum la valeur x_F (cette valeur pouvant être infinie) et définissons la fonction de répartition des excès au-dessus d'un seuil $u < x_F$: $F_u(x) = P(X - u \leq x | X > u)$. Nous pouvons l'exprimer en fonction de F :

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

Enfin, définissons la loi de Pareto généralisée (ou GPD pour *Generalized Pareto Distribution*) $G_{\xi, \sigma}$:

$$G_{\xi, \sigma}(x) = \begin{cases} \left(1 - \left(1 + \frac{\xi x}{\sigma}\right)^{-\frac{1}{\xi}}\right) \mathbb{1}_{x > 0} & \text{si } \xi > 0 \\ \left(1 - \left(1 + \frac{\xi x}{\sigma}\right)^{-\frac{1}{\xi}}\right) \mathbb{1}_{x \in]0, \frac{-\sigma}{\xi}[} & \text{si } \xi < 0 \\ \left(1 - e^{-\frac{x}{\sigma}}\right) \mathbb{1}_{x > 0} & \text{si } \xi = 0 \end{cases}$$

De même que pour la GEV, nous pouvons obtenir la densité g_{ξ} de la GPD. La figure 3.1 présente les différentes formes de g_{ξ} possibles en fonction de ξ .


 FIGURE 3.1 – Densités de la loi des excès pour différentes valeurs de ξ

Nous pouvons alors énoncer le théorème de Balkema-de Haan-Pickands démontré en 1974-1975 :

Théorème 2 (Balkema-de Haan-Pickands).

Il existe une fonction σ positive mesurable telle que :

$$\lim_{u \rightarrow x_F} \sup_{0 \leq x < x_F - u} |F_u(x) - G_{\xi, \sigma(u)}(x)| = 0$$

si et seulement si F appartient au domaine d'attraction de la loi H_ξ .

3.1 Estimation d'un quantile grâce à la loi de Pareto généralisée

Supposons que nous ayons trouvé un seuil u convenable (nous définirons plus loin la notion de seuil "convenable"), nous pouvons approcher la quantité $\frac{F(x+u)-F(u)}{1-F(u)}$ par $G_{\hat{\xi}, \hat{\sigma}}(x)$, où $\hat{\xi}$ et $\hat{\sigma}(x)$ sont estimés sur ces données. En posant $y = x + u$, nous avons :

$$\frac{F(y) - F(u)}{1 - F(u)} \simeq G_{\hat{\xi}, \hat{\sigma}}(y - u)$$

Notre objectif est de trouver y tel que $F(y) = p$, avec p fixé. Lorsque $\xi > 0$ la relation précédente conduit à :

$$\begin{aligned}
 y &\simeq u + G_{\hat{\xi}, \hat{\sigma}}^{-1} \left(\frac{p - F(u)}{1 - F(u)} \right) \\
 \Rightarrow y &\simeq u + \frac{\hat{\sigma}}{\hat{\xi}} \left(\left(1 - \frac{p - F(u)}{1 - F(u)} \right)^{-\hat{\xi}} - 1 \right) \\
 \Rightarrow y &\simeq u + \frac{\hat{\sigma}}{\hat{\xi}} \left(\left(\frac{1 - p}{1 - F(u)} \right)^{-\hat{\xi}} - 1 \right)
 \end{aligned}$$

Nous pouvons donc estimer y via cette formule en estimant $1 - F(u)$ par $\frac{N_u}{n}$ qui est la proportion d'observations au-dessus du seuil u :

$$y \simeq u + \frac{\hat{\sigma}}{\hat{\xi}} \left(\left(\frac{n}{N_u} (1 - p) \right)^{-\hat{\xi}} - 1 \right) \quad (3.1)$$

De même lorsque $\xi = 0$:

$$\begin{aligned}
 y &\simeq u + G_{\hat{\xi}, \hat{\sigma}}^{-1} \left(\frac{p - F(u)}{1 - F(u)} \right) \\
 \Rightarrow y &\simeq u - \hat{\sigma} \ln \left(1 - \frac{p - F(u)}{1 - F(u)} \right) \\
 \Rightarrow y &\simeq u - \hat{\sigma} \ln \left(\frac{1 - p}{1 - F(u)} \right)
 \end{aligned}$$

Ce qui conduit à :

$$y \simeq u - \hat{\sigma} \ln \left(\frac{n}{N_u} (1 - p) \right) \quad (3.2)$$

3.2 Estimation des Paramètres de la GPD

3.2.1 Valeur de ξ pour certaines lois connues

Nous savons que si un échantillon a pour fonction de répartition F , alors quand u tend vers la borne supérieure du support de F , la fonction de répartition $F_u(x) = \frac{F(x+u) - F(u)}{1 - F(u)}$ définie pour $x > 0$ tend vers une limite $G_{\xi, \sigma(u)}(x)$. Il est possible de déduire de ce résultat la valeur de ξ lorsque F est connue. Une première approche est d'obtenir la convergence de façon théorique, par exemple pour une loi exponentielle de paramètre θ :

$$\begin{aligned}
 F_u(x) &= \frac{F(x+u) - F(u)}{1 - F(u)} \\
 &= 1 - \frac{\bar{F}(x+u)}{\bar{F}(u)} \\
 &= 1 - \frac{e^{-\theta(x+u)}}{e^{-\theta u}} \\
 &= 1 - e^{-\theta x}
 \end{aligned}$$

C'est encore une loi exponentielle, de même paramètre. Il s'agit de la propriété de perte de mémoire de la loi exponentielle. En réalité, il s'agit aussi d'une loi GPD de paramètres $\xi = 0$ et $\sigma = \frac{1}{\theta}$, $\forall u$.

Toutefois, cela s'avère rapidement compliqué dès que F n'a pas d'expression analytique simple. Une deuxième approche possible lorsque F est connue est d'utiliser le logiciel R pour représenter la fonction \bar{F}_u pour un u élevé. Dans ce cas, nous pouvons donc être aussi proches que nous le souhaitons de $\bar{G}_{\xi, \sigma(u)}$.

Nous remarquons alors que dans le cas $\xi = 0$, $\ln(\bar{G}_{\xi, \sigma(u)})$ est une droite de pente $-\frac{1}{\sigma}$. Ainsi, si nous obtenons une droite en prenant le logarithme de \bar{F}_u sous R, alors nous pouvons en déduire que $\xi = 0$ pour cette loi. Sinon, nous sommes dans le cas où $\bar{G}_{\xi, \sigma(u)}$ s'écrit sous la forme $(1 + \frac{\xi x}{\sigma})^{-\frac{1}{\xi}}$ et nous cherchons le réel k tel que \bar{F}_u^k soit une droite. Nous en déduisons alors que $\xi = -k$.

Nous pouvons utiliser cela pour trouver la valeur de ξ pour quelques lois usuelles. Considérons les quatre lois suivantes : une loi normale centrée réduite, une loi de Student à 1 degré de liberté, une loi de Student à 2 degrés de liberté, et une loi Gamma de paramètres $k = 2$ et $\theta = 3$ (rappelons que la densité de cette loi est $f(x, k, \theta) = x^{k-1} \frac{\theta^k e^{-\theta x}}{\Gamma(k)}$).

La figure 3.2 présente les différentes formes de F_u pour des u suffisamment grand pour chaque loi. $\ln(\bar{F}_u)$ est présenté en figure 3.3.

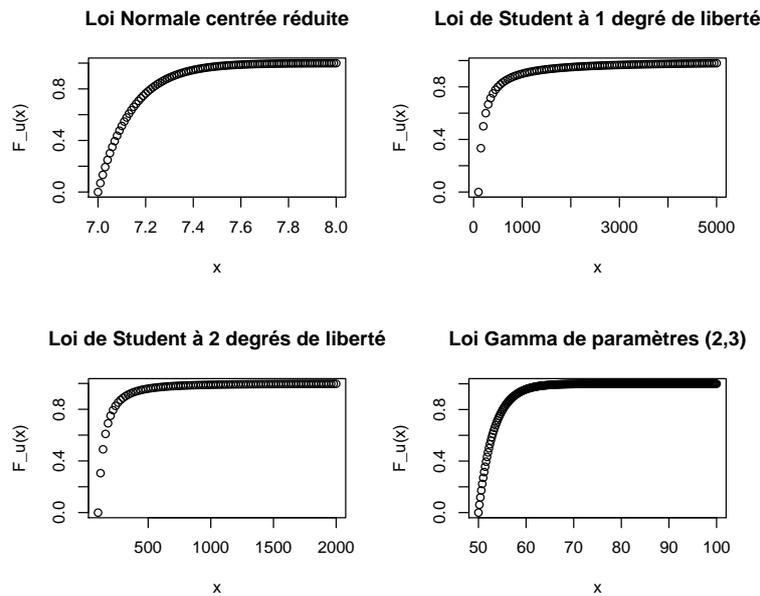


FIGURE 3.2 – Représentation des fonctions de répartition des excès F_u

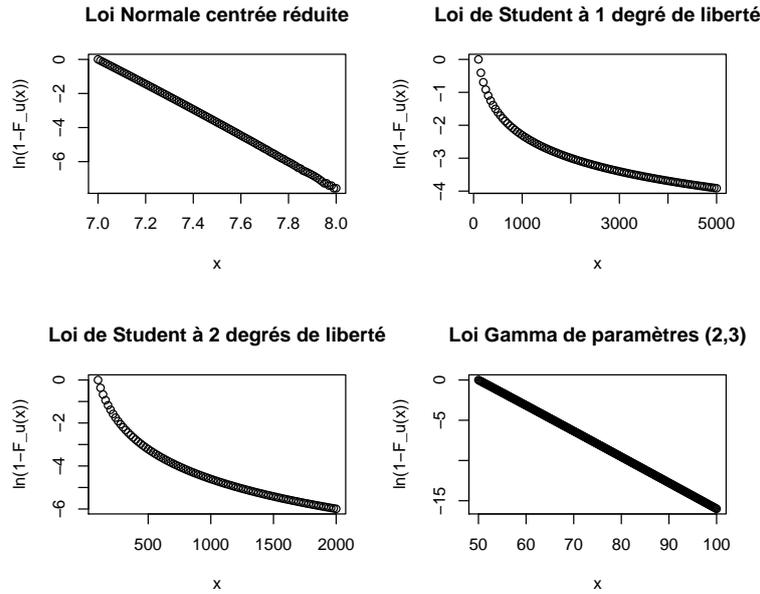


FIGURE 3.3 – Représentation des $\ln(\bar{F}_u)$

Nous remarquons que pour la loi normale et pour la loi Gamma, nous obtenons une droite, nous avons donc $\xi = 0$ pour ces lois. En revanche, ceci n'est pas vérifié pour les autres lois. En prenant respectivement $k = -1$ et $k = -0,5$, nous obtenons les graphiques présentées en figure 3.4. Nous avons ainsi trouvé les valeurs de ξ pour ces deux lois : 1 et 0,5.

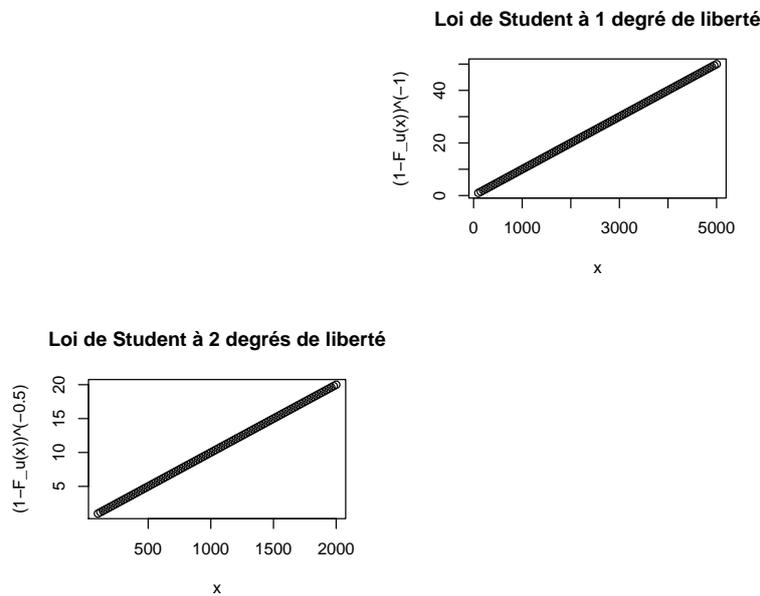


FIGURE 3.4 – Représentation des \bar{F}_u^k

Toutefois, lorsque la loi de départ de notre échantillon est inconnue, nous n'avons pas accès à l'expression exacte de F . Nous pourrions utiliser un estimateur de F pour en déduire ensuite un estimateur de F_u et appliquer la méthode décrite ci-dessus, mais en réalité, il existe des méthodes plus efficaces pour estimer ξ .

Dans la suite, nous nous plaçons à u fixé et nous cherchons à estimer ξ et de $\sigma(u) = \sigma$.

3.2.2 Méthode du maximum de vraisemblance

Définissons l'échantillon $(Z_i)_{i=1, \dots, N_u}$ des dépassements : si $X_i > u$, alors nous mettons $X_i - u$ dans l'échantillon, sinon nous regardons X_{i+1} et ainsi de suite. N_u est le nombre de dépassements. \hat{F}_u^r est la fonction de survie empirique de l'échantillon $(Z_i)_{i=1, \dots, N_u}$.

Par le théorème de Balkema-de Haan-Pickands, nous pouvons supposer que les $(Z_i)_{i=1, \dots, N_u}$ suivent une distribution GPD, dont nous pouvons estimer les paramètres par méthode de maximum de vraisemblance. La vraisemblance du modèle s'écrit dans le cas $\xi \neq 0$:

$$L(\xi, \sigma, Z_1, \dots, Z_{N_u}) = \frac{1}{\sigma^{N_u}} \prod_{i=1}^{N_u} \left(1 + \xi \frac{Z_i}{\sigma}\right)^{-1 - \frac{1}{\xi}}$$

Afin de trouver les paramètres optimaux, il est plus aisé de passer par la log-vraisemblance :

$$l(\xi, \sigma, Z_1, \dots, Z_{N_u}) = -N_u \ln(\sigma) - \left(1 + \frac{1}{\xi}\right) \sum_{i=1}^{N_u} \ln\left(1 + \xi \frac{Z_i}{\sigma}\right)$$

Nous pouvons alors calculer les dérivées partielles par rapport aux paramètres ξ et σ , et en les égalisant à 0 nous obtenons les paramètres maximisant la vraisemblance. En pratique, ces équations n'ont pas de solution explicite et nous effectuons donc une résolution numérique.

Dans le cas $\xi = 0$, nous avons une expression simplifiée de la log-vraisemblance :

$$l(0, \sigma, Z_1, \dots, Z_{N_u}) = -N_u \ln(\sigma) - \frac{1}{\sigma} \sum_{i=1}^{N_u} Z_i$$

L'estimateur du maximum de vraisemblance pour σ est alors $\frac{1}{N_u} \sum_{i=1}^{N_u} Z_i$.

3.2.3 Méthode des moments pondérés

Définissons les moments pondérés pour r un entier supérieur ou égal à 0 :

$$w_r(\xi, \sigma) = \mathbb{E}(X \bar{G}_{\xi, \sigma}^r(X))$$

Nous pouvons calculer cette quantité $\forall \xi < 1$:

$$w_r(\xi, \sigma) = \frac{\sigma}{(r+1)(r+1-\xi)}$$

La démonstration est faite dans l'annexe A.2. Les deux premiers moments s'écrivent donc :

$$\begin{cases} w_0 &= \frac{\sigma}{1-\xi} \\ w_1 &= \frac{\sigma}{2(2-\xi)} \end{cases}$$

Nous obtenons alors :

$$\begin{cases} \hat{\xi} &= \frac{w_0 - 4w_1}{w_0 - 2w_1} \\ \hat{\sigma} &= \frac{2w_0 w_1}{w_0 - 2w_1} \end{cases}$$

Il suffit d'estimer ces moments et d'appliquer les relations précédentes pour en déduire la valeur des paramètres de la GPD. Nous prenons comme estimateur :

$$\hat{w}_r(\xi, \sigma) = \frac{1}{N_u} \sum_{i=1}^{N_u} Z_i \hat{F}_u^r(Z_i)$$

Cette méthode n'est valable que pour $\xi < 1$. De plus, de même que pour la GEV la normalité asymptotique n'est vraie que pour $\xi < \frac{1}{2}$. Il existe là encore une généralisation de la méthode des moments pondérés étendant le domaine de validité des estimateurs, mais celle-ci comporte un biais qui la rend plus difficile à manier, notamment pour la construction d'intervalles de confiance. Nous décidons donc de ne pas l'utiliser.

3.3 Recherche du Seuil u

Une bonne détermination de u est primordiale si nous voulons appliquer la théorie des valeurs extrêmes. En effet, un seuil trop faible conduit à une mauvaise adéquation entre la fonction de survie et la loi GPD et les estimateurs sont biaisés. Inversement, un seuil trop élevé entraîne une faible taille d'échantillon de dépassements et donc des estimateurs volatiles. Nous allons voir dans la suite diverses méthodes que nous utiliserons ultérieurement pour estimer ce seuil.

3.3.1 Méthode 1 : utilisation de la fonction moyenne des excès

Afin de déterminer ce seuil, nous pouvons faire appel à la fonction moyenne des excès (plus connue sous le nom de *Mean excess function* dans la littérature, voir par exemple [EKM97]) :

$$e(u) = \mathbb{E}(X - u | X > u)$$

Nous pouvons réécrire :

$$\begin{aligned} e(u) &= \int_0^{x_F - u} x \frac{dF(x + u)}{1 - F(u)} \\ &= \frac{1}{\bar{F}(u)} \int_u^{x_F} (y - u) dF(y) \\ &= \frac{-1}{\bar{F}(u)} \int_u^{x_F} (y - u) d\bar{F}(y) \\ &= \frac{-1}{\bar{F}(u)} ([(y - u)\bar{F}(y)]_u^{x_F} - \int_u^{x_F} \bar{F}(y) dy) \\ &= \frac{1}{\bar{F}(u)} \int_u^{x_F} \bar{F}(x) dx \end{aligned}$$

Dans le cas de la loi GPD, nous avons $\bar{G}_{\xi, \sigma}(x) = (1 + \frac{\xi x}{\sigma})^{-\frac{1}{\xi}}$. Ainsi :

$$e(u) = \left(1 + \frac{\xi u}{\sigma}\right)^{\frac{1}{\xi}} \int_u^{x_F} \left(1 + \frac{\xi x}{\sigma}\right)^{-\frac{1}{\xi}} dx$$

Dans le cas $\xi < 1$ et $\xi \neq 0$:

$$\begin{aligned}
 e(u) &= \left(1 + \frac{\xi u}{\sigma}\right)^{\frac{1}{\xi}} \left[\frac{\sigma}{\xi - 1} \left(1 + \frac{\xi x}{\sigma}\right)^{\frac{\xi-1}{\xi}} \right]_u^{x_F} \\
 &= - \left(1 + \frac{\xi u}{\sigma}\right)^{\frac{1}{\xi}} \frac{\sigma}{\xi - 1} \left(1 + \frac{\xi u}{\sigma}\right)^{\frac{\xi-1}{\xi}} \\
 &= - \frac{\sigma}{\xi - 1} \left(1 + \frac{\xi u}{\sigma}\right) \\
 &= \frac{\sigma + \xi u}{1 - \xi}
 \end{aligned}$$

Si $\xi \geq 1$, l'intégrale ci-dessus est divergente et $e(u)$ n'existe pas : c'est le phénomène de "fuite vers l'infini", propre aux distributions à queues épaisses.

Enfin, si $\xi = 0$, nous avons :

$$\begin{aligned}
 e(u) &= e^{\frac{u}{\sigma}} \int_u^{+\infty} e^{-\frac{x}{\sigma}} \\
 &= e^{\frac{u}{\sigma}} [-\sigma e^{-\frac{x}{\sigma}}]_u^{+\infty} \\
 &= \sigma
 \end{aligned}$$

Dans le cas $\xi = 0$, la fonction moyenne des excès est donc une constante, sinon c'est une fonction linéaire en u . Nous pouvons ainsi en déduire un critère de sélection du seuil : nous cherchons u tel que la *mean excess function* soit linéaire à partir de ce seuil. En effet, cela signifie que la fonction F_u est proche de la GPD, et donc que notre hypothèse est raisonnable. Nous n'avons évidemment pas accès à la fonction e , nous l'estimons donc par la méthode habituelle :

$$\hat{e}_n(u) = \frac{1}{N_u} \sum_{i=1}^n (X_i - u)_+$$

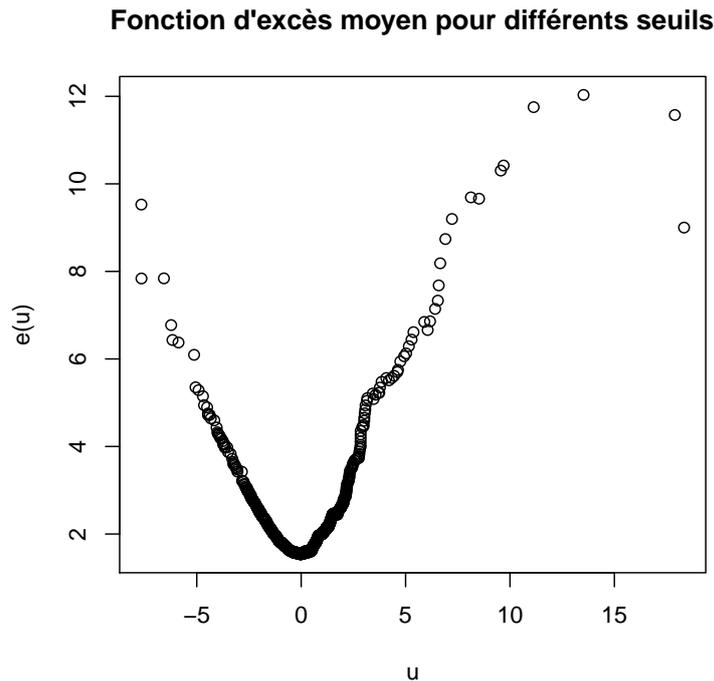


FIGURE 3.5 – Fonction moyenne des excès pour 1000 observations d’une Student à 2 degrés de liberté

Nous voyons que la fin de la courbe est erratique. Ceci est dû au fait que les dernières estimations sont faites sur un nombre très faible de données.

D’un point de vue pratique, l’utilisation de cette fonction pour le choix du seuil u nécessite un critère objectif déterminant à quel moment la fonction moyenne des excès devient une droite. Pour cela, nous décidons d’utiliser un algorithme de minimisation d’une fonction d’erreur, où l’erreur correspond à l’écart entre 1 et la valeur du R^2 calculé pour la régression de $\hat{e}_n(u)$ sur u pour la portion $[u_1, u_2]$. Les deux paramètres u_1 et u_2 sont susceptibles de bouger, mais la fonction est fortement pénalisée lorsque u_2 est plus bas que le quantile à 99 % de l’échantillon : l’algorithme est autorisé à supprimer les valeurs aberrantes, mais doit conserver la queue de distribution. De même, il y a pénalisation lorsque u_2 est plus élevé que le quantile à 97 % de l’échantillon, afin de conserver un nombre suffisant de valeurs pour l’estimation du quantile. Nous définissons alors le seuil comme étant le u_1 optimal.

Nous avons établi nous-mêmes ces règles, qui permettent de trouver u de manière automatique et donc de tester l’impact de cette procédure de choix de seuil sur un grand nombre d’échantillons. Si nous recherchons maintenant un seuil sur un échantillon de données réelles, nous pourrions apporter davantage de précision et tenir compte des caractéristiques de l’échantillon.

3.3.2 Méthode 2 : utilisation du graphique de Hill

Dans le cas où $\xi > 0$, posons $\alpha = \frac{1}{\xi}$. EMBRECHTS, KLÜPPELBERG et MIKOSCH ont démontré dans [EKM97] que nous avons alors $F(x) = x^{-\alpha}L(x)$ avec L fonction à variations lentes :

$$\forall t > 0, \lim_{x \rightarrow +\infty} \frac{L(tx)}{L(x)} = 1$$

Définissons $Y = \ln(X) | \ln(X) > a$ pour $a > 0$, et calculons la quantité suivante pour $t > a$:

$$\begin{aligned} \mathbb{P}(Y > t) &= \mathbb{P}(\ln(X) > t | \ln(X) > a) \\ \Rightarrow \mathbb{P}(Y > t) &= \mathbb{P}(X > e^t | X > e^a) \\ \Rightarrow \mathbb{P}(Y > t) &= \frac{\bar{F}(e^t)}{\bar{F}(e^a)} \\ \Rightarrow \mathbb{P}(Y > t) &= e^{-\alpha(t-a)} \frac{L(e^t)}{L(e^a)} \\ \Rightarrow \mathbb{P}(Y > t) &= e^{-\alpha(t-a)} \frac{L(e^a e^{t-a})}{L(e^a)} \end{aligned}$$

Nous voyons qu'en prenant a assez grand et $x = e^a$, la propriété des fonctions à variations lentes permet de faire l'approximation que la fraction est égale à 1. Dans ce cas, nous avons $\mathbb{P}(Y > t) = e^{-\alpha(t-a)}$, ce qui nous donne donc la loi suivie par Y : c'est une loi exponentielle de paramètre α , translatée de a . Autrement dit, $Y - a$ suit une loi exponentielle de paramètre α .

Un estimateur de α est donc l'estimateur classique du maximum de vraisemblance pour une loi exponentielle :

$$\hat{\alpha}_n = \frac{1}{\bar{Y} - a} = \frac{1}{\sum_{i=1}^{N_a} \ln(X_{(n-i+1)}) - a}$$

où N_a est le nombre de dépassements et $(X_{(i)})_i$ est l'échantillon trié par ordre croissant. Cet estimateur implique de connaître a , mais inversement, nous pouvons définir a comme la valeur à partir de laquelle cet estimateur converge. Concrètement, nous allons poser $a = \ln(X_{(n-k)})$, ce qui signifie que nous utilisons les k plus grandes valeurs de X pour estimer α . Nous définissons alors :

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

Notre objectif est de trouver k tel que cet estimateur reste stable, ce qui signifiera alors que nous avons utilisé un seuil suffisamment élevé pour que l'approximation ci-dessus soit licite. Toutefois, nous devons faire attention à ne pas choisir k trop faible, auquel cas l'estimation serait trop volatile. Pour prendre notre décision, nous traçons le graphique de Hill (*Hill Plot*), qui est la représentation graphique de $\hat{\alpha}_{n,k}$ en fonction de k .

Voici par exemple ce graphique lorsque X suit une loi de Student à 2 degrés de liberté :

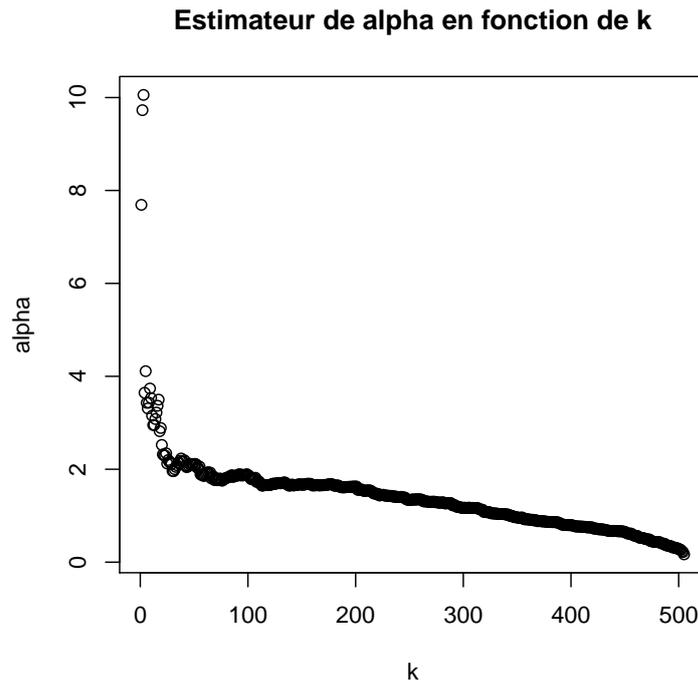


FIGURE 3.6 – Hill Plot pour 1000 observations d’une Student à 2 degrés de liberté

Nous lisons ce graphique en partant de la droite. Nous voyons sur la gauche une portion où le graphique n’est plus une fonction décroissante de k mais a un comportement plus erratique. Nous définissons le k optimal comme étant la valeur la plus à droite pour laquelle la partie du graphique située entre 0 et k est une constante ”bruitée”. Afin de définir de façon précise ce seuil optimal, nous utilisons un algorithme similaire à celui utilisé pour la fonction moyenne des excès : nous minimisons une erreur en fonction de deux valeurs k_1 et k_2 , où l’erreur est la moyenne des carrés des écarts entre les valeurs $\hat{\alpha}_{n,k_1}$ à $\hat{\alpha}_{n,k_2}$ et la droite horizontale minimisant ces écarts. Le seuil optimal est alors $X_{(n-k_2)}$, comme précédemment k_1 sert à éliminer les valeurs aberrantes.

3.3.3 Méthode 3 : utilisation du test de Cunnane

Des auteurs ont montré qu’asymptotiquement, le nombre d’excès au-dessus du seuil u suit une loi de Poisson (voir par exemple [LEA93b]). Nous utilisons ce résultat pour déterminer le seuil, en faisant aussi l’approximation qu’une loi de Poisson est proche d’une loi normale pour une valeur du paramètre suffisamment élevée. Soit $(m_i(u))_{i=1,\dots,k}$ un échantillon de nombre de dépassements, pour k échantillons indépendants. Nous allons tester l’hypothèse qu’ils soient tirés selon une loi $N(\lambda, \lambda)$. Dans ce cas :

$$\forall i, \frac{m_i(u) - \lambda}{\sqrt{\lambda}} \sim N(0, 1)$$

Nous en déduisons que :

$$\sum_{i=1}^k \frac{(m_i(u) - \lambda)^2}{\lambda} \sim \chi_k^2$$

Nous estimons λ par $S(m(u))$, la moyenne empirique des $(m_i(u))_{i=1,\dots,k}$:

$$S(m(u)) = \sum_{i=1}^k \frac{(m_i(u) - \bar{m})^2}{\bar{m}} \sim \chi_{k-1}^2$$

A l'aide de cette statistique de test, nous testons alors l'hypothèse H_0 : $m(u)$ suit une loi Normale au niveau α . Nous rejetons l'hypothèse nulle si $S(m(u)) > q_{k-1}(1 - \frac{\alpha}{2})$ ou $S(m(u)) < q_{k-1}(\frac{\alpha}{2})$ où $q_{k-1}(p)$ représente le p-quantile d'une χ_{k-1}^2 .

Nous choisissons le seuil u minimal pour lequel l'hypothèse nulle n'est pas rejetée pour tous les seuils supérieurs. D'un point de vue pratique, cette méthode nécessite l'approximation normale qui n'est licite que si le nombre de dépassements moyen est suffisamment élevé. Nous utilisons comme condition $\hat{\lambda} > 5$. Cela conduit à une impossibilité d'utiliser cette méthode lorsque nous avons un faible nombre d'observations (par exemple, 100) puisque dans ce cas la condition n'est pas vérifiée et l'approximation de la loi de Poisson par une loi normale est peu crédible.

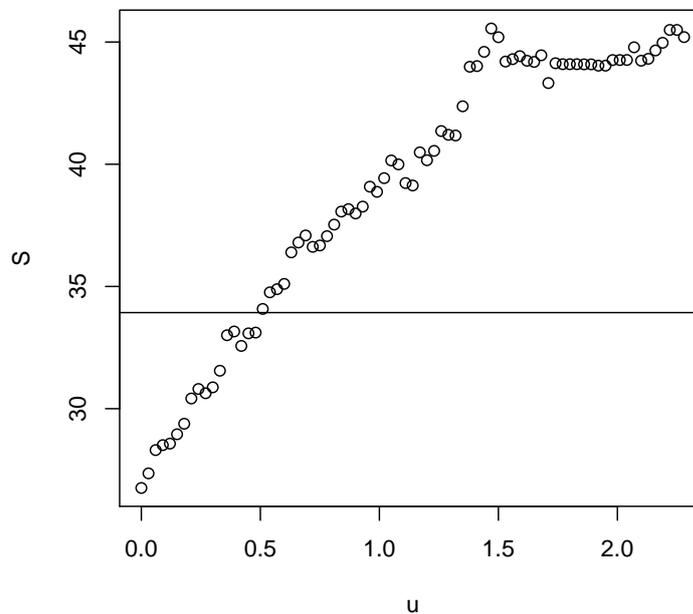


FIGURE 3.7 – Représentation du test de Cunnane pour 1000 observations d'une Student à 2 degrés de liberté. La barre horizontale correspond à la valeur de $q_{k-1}(\frac{\alpha}{2})$, le quantile $q_{k-1}(1 - \frac{\alpha}{2})$ est ici trop élevé pour figurer sur le graphique.

3.3.4 Méthode 4 : minimisation de la distance de Kolmogorov-Smirnov

Cette procédure de choix du seuil a été introduite par PICKANDS dans [PIC75]. L'idée est d'utiliser le fait que la distribution des excès doit être proche d'une loi de Pareto généralisée. Nous allons donc chercher le seuil u pour lequel la statistique de Kolmogorov-Smirnov est la plus petite. Cette statistique s'écrit ainsi :

$$\sup_x |\hat{G}_n(y_i) - G_{\hat{\xi}, \hat{\sigma}}(y_i)|$$

où \hat{G}_n est la fonction de répartition empirique des excès $(y_i)_{i=1, \dots, N_u}$, et $G_{\hat{\xi}, \hat{\sigma}}$ est la fonction de répartition d'une GPD de paramètres $\hat{\xi}$ et $\hat{\sigma}$ estimés sur les données. Pour un échantillon de taille infinie, cette distance est décroissante en fonction du seuil, du fait de la convergence des

excès vers une loi de Pareto généralisée. En revanche, pour un échantillon fini, la distance est décroissante puis croissante car lorsqu'il y a peu d'observations au-dessus du seuil, les estimateurs des paramètres sont volatiles. Nous choisissons alors comme seuil u le seuil qui rend minimal cette statistique. Pour un échantillon de taille n , nous prenons comme seuil la $(n - k)^{eme}$ plus grande valeur pour 100 valeurs de k comprises entre 5 et $\frac{n}{2}$. Les paramètres sont estimés selon deux méthodes : celle du maximum de vraisemblance et celle des moments pondérés, présentées dans la suite. Par soucis de cohérence, la manière dont les paramètres sont estimés lors du calcul du quantile est celle que nous avons utilisée pour trouver le seuil optimal. Par exemple, si nous avons trouvé k^* tel que la distribution des excès au-dessus de la $(n - k^*)^{eme}$ est la plus proche d'une GPD pour des paramètres estimés par maximum de vraisemblance, alors les paramètres utilisés dans (3.1) sont les paramètres estimés par maximum de vraisemblance.

3.4 Intervalles de confiance pour les quantiles extrêmes

Pour chaque méthode d'estimation, nous construisons des intervalles de confiance afin de mesurer la précision des estimateurs du quantile à un degré de confiance donné.

3.4.1 Intervalles obtenus pour les estimateurs du maximum de vraisemblance

SMITH a montré dans [SMI84] que pour $\xi > -\frac{1}{2}$, les estimateurs du maximum de vraisemblance étaient asymptotiquement normaux de variance asymptotique :

$$\mathbb{V}_{as} \begin{pmatrix} \hat{\sigma} \\ \hat{\xi} \end{pmatrix} = \begin{pmatrix} 2\sigma^2(1 + \xi) & -\sigma(1 + \xi) \\ -\sigma(1 + \xi) & (1 + \xi)^2 \end{pmatrix}$$

Autrement dit, nous avons la convergence en loi suivante :

$$\sqrt{n} \begin{pmatrix} \hat{\sigma} - \sigma \\ \hat{\xi} - \xi \end{pmatrix} \xrightarrow{Loi} N \left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \mathbb{V}_{as} \begin{pmatrix} \hat{\sigma} \\ \hat{\xi} \end{pmatrix} \right)$$

Rappelons l'expression de l'estimation du quantile :

$$\hat{q} \simeq u + \frac{\hat{\sigma}}{\hat{\xi}} \left(\left(\frac{n}{N_u} (1 - p) \right)^{-\hat{\xi}} - 1 \right)$$

Posons $a = \frac{n}{N_u} (1 - p)$ afin d'alléger les notations, et dérivons \hat{q} par rapport à $\hat{\xi}$ et $\hat{\sigma}$ pour appliquer la delta-méthode :

$$\begin{aligned} \frac{\partial \hat{q}}{\partial \hat{\xi}} &= -\frac{\hat{\sigma}}{\hat{\xi}} \left(\frac{a^{-\hat{\xi}} - 1}{\hat{\xi}} + \ln(a) a^{-\hat{\xi}} \right) \\ \frac{\partial \hat{q}}{\partial \hat{\sigma}} &= \frac{a^{-\hat{\xi}} - 1}{\hat{\xi}} \end{aligned}$$

Nous déduisons de la delta-méthode que \hat{q} est asymptotiquement normal, et nous pouvons estimer sa matrice de variance asymptotique :

$$\mathbb{V}_{as}(\hat{q}) = \left(\frac{\partial \hat{q}}{\partial \hat{\xi}} \right)^2 (1 + \hat{\xi})^2 + 2 \left(\frac{\partial \hat{q}}{\partial \hat{\sigma}} \right)^2 \hat{\sigma}^2 (1 + \hat{\xi}) - 2 \frac{\partial \hat{q}}{\partial \hat{\xi}} \frac{\partial \hat{q}}{\partial \hat{\sigma}} \hat{\sigma} (1 + \hat{\xi})$$

Un intervalle de confiance asymptotique de niveau α (par exemple $\alpha = 95\%$) pour q est donc $\left[\hat{q} \pm \frac{\sqrt{\mathbb{V}_{as}(\hat{q})}}{\sqrt{n}} z_{\frac{1+\alpha}{2}} \right]$, où $z_{\frac{1+\alpha}{2}}$ est le quantile d'ordre $\frac{1+\alpha}{2}$ d'une loi normale centrée réduite.

3.4.2 Intervalles obtenus pour les estimateurs des moments pondérés

HOSKING a montré dans [HWW85] que dans le cas des estimateurs des moments pondérés, nous avons aussi une normalité asymptotique, de variance asymptotique s'écrivant :

$$\mathbb{V}_{as} \begin{pmatrix} \hat{\sigma} \\ \hat{\xi} \end{pmatrix} = \frac{1}{(1-2\xi)(3-2\xi)} \begin{pmatrix} \sigma^2(7-18\xi+11\xi^2-2\xi^3) & -\sigma(2-\xi)(2-6\xi+7\xi^2-2\xi^3) \\ -\sigma(2-\xi)(2-6\xi+7\xi^2-2\xi^3) & (1-\xi)(2-\xi)^2(1-\xi+2\xi^2) \end{pmatrix}$$

Il nous suffit d'appliquer une nouvelle fois la delta-méthode pour en déduire l'intervalle de confiance pour cette méthode.

Comme pour la GEV, il est aussi possible d'utiliser le profil de vraisemblance ou le bootstrap non paramétrique.

Nous avons vu dans ce chapitre une autre méthode de calcul de quantiles extrêmes, similaire à la méthode utilisant la loi des valeurs extrêmes mais comportant des différences significatives. La question que nous nous posons alors est de savoir laquelle est la plus efficace, et sous quelles conditions : caractéristiques de l'échantillon, méthode d'estimation des paramètres, ou méthode de choix de seuil dans le cadre de la GPD. Pour y répondre nous menons dans la suite une application de ces théories sur des données simulées.

Chapitre 4

Application de la théorie sur des données simulées

Dans ce chapitre, nous allons utiliser les méthodes décrites dans les parties précédentes, sur des échantillons de lois connues que nous simulons afin de tester le comportement de nos estimateurs.

Les notations utilisées sont les suivantes :

- "EMV", "PWM", "GPWM" signifient respectivement que nous avons estimé les paramètres par méthode du maximum de vraisemblance, des moments pondérés, et des moments pondérés généralisés.
- "RB" signifie que nous avons amélioré les estimateurs par méthode de Rao-Blackwell.
- "m=10" signifie que nous avons pris des blocs de taille égale à 10
- "ME", "Hill", "Cunn1", "Cunn2", "Pick" signifient respectivement que la méthode de choix du seuil ont été celle de la fonction moyenne des excès, celle du graphique de Hill, celle du test de Cunnane avec $\alpha = 10\%$, celle du test de Cunnane avec $\alpha = 20\%$, et celle de Pickands.
- "(Fisher théo)" et "(Fisher obs)" signifient respectivement que nous avons utilisé l'information de Fisher théorique ou observée
- "(PF)" signifie que nous avons utilisé le profil de vraisemblance

Avant de commencer, rappelons les conditions d'application des diverses méthodes d'estimation des paramètres :

	GEV			GPD	
	EMV	PWM	GPWM	EMV	PWM
Convergence de l'estimateur	ξ quelconque	$\xi < 1$	$\xi < 2$	ξ quelconque	$\xi < 1$
Normalité asymptotique	$\xi > -\frac{1}{2}$	$\xi < \frac{1}{2}$	$\xi < \frac{3}{2}$	$\xi > -\frac{1}{2}$	$\xi < \frac{1}{2}$

TABLE 4.1 – Domaines de validité des différentes propriétés

Ce tableau sera utile pour l'interprétation des différents résultats dans la suite.

4.1 Estimation des paramètres de la loi des valeurs extrêmes

Nous souhaitons tester les méthodes d'estimation des paramètres (développées dans le chapitre 2) sur des lois GEV pour différentes tailles d'échantillon n . L'intérêt d'utiliser des GEV est que les paramètres de la loi sont connus et que nous pouvons donc mesurer l'écart entre la valeur réelle et celle estimée.

Nous simulons donc des échantillons suivant une loi GEV pour plusieurs tailles d'échantillon : 100, 1 000, et 10 000. Pour cela, nous faisons varier le paramètre ξ tout en gardant σ et μ constants : $\sigma = 1$ et $\mu = 0$. A chaque fois, nous cherchons à retrouver les valeurs des paramètres par méthode de maximum de vraisemblance, méthode des moments pondérés, et méthode des moments pondérés généralisée. Afin de mesurer l'écart entre les paramètres estimés et les paramètres réels, nous calculons l'écart-quadratique moyen. Celui-ci est défini pour N échantillons de taille n différents (pour le paramètre ξ) de la manière suivante :

$$EQM = \frac{1}{N} \sum_{j=1}^N (\xi - \hat{\xi}_j)^2$$

où $\hat{\xi}_j$ correspond à l'estimateur sur le j^{eme} échantillon. Nous prenons $N = 500$ pour avoir des valeurs robustes.

Nous choisissons une taille des blocs égale à 10 quand $n = 100$ et 50 sinon.

Echantillon		EMV						RB EMV					
n	ξ	ξ estimé	EQM ξ	σ estimé	EQM σ	μ estimé	EQM μ	ξ estimé	EQM ξ	σ estimé	EQM σ	μ estimé	EQM μ
1 000	0,3	0,315	0,0680	0,952	0,0482	0,017	0,0695	0,311	0,0439	0,944	0,0314	0,013	0,0467
1 000	0,8	0,873	0,1435	0,935	0,0953	0,003	0,0755	0,873	0,085	0,943	0,0820	-0,003	0,0434
1 000	1,5	1,796	0,8852	4,818	2482	0,617	64,3	1,807	0,5972	4,497	2247	0,591	60,71
10 000	0,3	0,301	0,0042	0,992	0,0041	0,001	0,0059	0,300	0,0029	0,994	0,0030	0,002	0,0041
10 000	0,8	0,803	0,0065	0,995	0,0098	0,001	0,0069	0,804	0,0048	0,996	0,0069	0,001	0,0043
10 000	1,5	1,507	0,0206	1,007	0,0274	0,006	0,0069	1,505	0,0131	1,018	0,0863	0,008	0,0072

TABLE 4.2 – Estimation des paramètres d'une GEV par méthode de maximum de vraisemblance

Echantillon		PWM						RB PWM					
n	ξ	ξ estimé	EQM ξ	σ estimé	EQM σ	μ estimé	EQM μ	ξ estimé	EQM ξ	σ estimé	EQM σ	μ estimé	EQM μ
100	0,3	0,215	0,0551	0,936	0,1134	0,092	0,1541	0,212	0,0492	0,9410	0,0974	0,099	0,1150
100	0,8	0,417	0,1944	1,442	3,7363	0,283	0,4365	0,415	0,1925	1,448	3,6407	0,302	0,4293
100	1,5	0,604	0,8425	35,114	164935	7,482	6523	0,607	0,8328	35,07	164897	7,442	6516
1 000	0,3	0,248	0,0346	0,977	0,0579	0,049	0,0679	0,250	0,0305	0,971	0,0474	0,046	0,0502
1 000	0,8	0,533	0,1080	1,306	1,0343	0,195	0,1645	0,535	0,1047	1,301	1,0621	0,185	0,11413
1 000	1,5	0,727	0,6243	15,831	146858	2,907	417,9	0,728	0,6204	15,880	14681	2,915	419,0
10 000	0,3	0,289	0,0058	0,996	0,0055	0,007	0,0063	0,289	0,0056	0,998	0,0046	0,008	0,0049
10 000	0,8	0,685	0,0238	1,101	0,0314	0,058	0,0126	0,685	0,0238	1,105	0,0307	0,059	0,0111
10 000	1,5	0,922	0,3371	70,793	2107478	10,091	42212	0,922	0,3369	70,785	2107508	10,089	42212

TABLE 4.3 – Estimation des paramètres d'une GEV par méthode des moments pondérés

Echantillon		GPWM						RB GPWM					
n	ξ	ξ estimé	EQM ξ	σ estimé	EQM σ	μ estimé	EQM μ	ξ estimé	EQM ξ	σ estimé	EQM σ	μ estimé	EQM μ
100	0,3	0,259	0,0931	0,789	0,1260	-0,212	0,1602	0,246	0,0391	0,789	0,0890	-0,204	0,1104
100	0,8	0,523	0,1876	0,772	0,1823	-0,153	0,1141	0,514	0,1460	0,772	0,1551	-0,133	0,0860
100	1,5	0,888	0,4918	1,122	3,1346	0,071	0,4352	0,894	0,4603	1,122	1,6436	0,048	0,2070
1 000	0,3	0,175	0,0697	0,941	0,0478	-0,107	0,0699	0,177	0,0469	0,941	0,0326	-0,110	0,0485
1 000	0,8	0,580	0,1189	0,913	0,1078	-0,059	0,0680	0,584	0,0966	0,913	0,0817	-0,069	0,0446
1 000	1,5	1,034	0,3107	1,054	0,5210	0,026	0,0965	1,033	0,3041	1,054	0,4465	0,023	0,0693
10 000	0,3	0,280	0,0061	1,003	0,0049	-0,001	0,0062	0,281	0,0041	1,003	0,0037	-0,001	0,0043
10 000	0,8	0,778	0,0095	1,004	0,0125	0,002	0,0074	0,778	0,0086	1,004	0,0095	0,003	0,0052
10 000	1,5	1,413	0,0336	1,072	0,0459	0,033	0,0104	1,412	0,0331	1,072	0,0414	0,032	0,0083

TABLE 4.4 – Estimation des paramètres d'une GEV par méthode des moments pondérés généralisés

Notons tout d'abord que la méthode de rao-blackwellisation permet dans tous les cas de réduire l'erreur quadratique des estimateurs classiques. Cependant le biais est conservé, mais il est le plus faible pour les estimateurs obtenus par maximum de vraisemblance. Les estimateurs calculés par la méthode des moments pondérés ou la méthode des moments pondérés généralisés sont en effet fortement biaisés lorsque $n = 100$ et $n = 1 000$. Néanmoins, pour des petits échantillons ($n = 100$), ils permettent de réduire l'erreur quadratique.

Pour $\xi = 0,3$, il y a convergence de tous les estimateurs vers les vraies valeurs de ξ , σ et μ et leur erreur quadratique est comparable.

Pour $\xi = 0,8$ et $\xi = 1,5$, la méthode des moments pondérés n'est plus applicable : cela se traduit dans les tableaux par des erreurs quadratiques très importantes qui augmentent en fonction de ξ (notamment pour σ et μ). En revanche, la GPWM reste applicable. Remarquons enfin que pour $\xi = 1,5$ et $n = 1\,000$, les estimateurs EMV sont beaucoup plus volatiles que les estimateurs GPWM.

4.2 Estimation des paramètres de la loi de Pareto généralisée

Nous testons maintenant les méthodes de maximum de vraisemblance et des moments pondérés sur des échantillons simulés selon une loi GPD :

Echantillon			EMV			
n	ξ	σ	ξ estimé	EQM ξ	σ estimé	EQM σ
100	0,3	1	0,272	0,0197	1,029	0,0290
100	0,8	1	0,772	0,0329	1,037	0,0402
100	1,5	1	1,485	0,0646	1,036	0,0585
1 000	0,3	1	0,296	0,00171	1,002	0,00290
1 000	0,8	1	0,796	0,00344	0,997	0,00362
1 000	1,5	1	1,498	0,00685	1,003	0,00523
10 000	0,3	1	0,30	0,000156	0,999	0,000271
10 000	0,8	1	0,801	0,000334	0,999	0,000377
10 000	1,5	1	1,500	0,000660	1,000	0,000449

TABLE 4.5 – Estimation des paramètres d'une GPD par méthode de maximum de vraisemblance

Echantillon			PWM			
n	ξ	σ	ξ estimé	EQM ξ	σ estimé	EQM σ
100	0,3	1	0,265	0,0165	1,030	0,0265
100	0,8	1	0,655	0,0410	1,148	0,0641
100	1,5	1	0,899	0,367	3,96	119
1 000	0,3	1	0,297	0,00195	1,002	0,00292
1 000	0,8	1	0,743	0,00951	1,054	0,00894
1 000	1,5	1	0,970	0,281	3,211	226
10 000	0,3	1	0,300	0,000188	0,999	0,000276
10 000	0,8	1	0,775	0,00316	1,021	0,00247
10 000	1,5	1	0,991	0,259	2,191	2,36

TABLE 4.6 – Estimation des paramètres d'une GPD par méthode de maximum de vraisemblance

Dans les cas $\xi = 0,3$ nous voyons une nette convergence des estimateurs EMV et PWM vers les vraies valeurs de ξ et σ , avec des valeurs d'écart-quadratique moyens comparables entre les deux estimateurs.

Pour $\xi = 0,8$ il y a là encore convergence des deux estimateurs, mais l'estimateur des moments pondérés a une convergence plus lente.

Enfin, pour $\xi = 1,5$, seul l'estimateur du maximum de vraisemblance converge, l'estimateur des moments étant très nettement biaisé. Cela correspond à ce que nous avons annoncé lorsque nous avons présenté cette méthode : celle-ci n'est définie que pour $\xi < 1$, au-dessus de cette valeur les queues de distribution sont trop épaisses et les moments ne sont pas définis.

4.3 Estimation des quantiles de loi connues

Dans cette section, nous allons appliquer les outils fournis par la théorie des valeurs extrêmes à l'estimation de quantiles de lois connues. Cela nous permettra de comparer les diverses méthodes d'estimation des paramètres en fonction de la loi, du quantile recherché, et du nombre d'observations.

Pour chaque loi, nous simulons 500 fois un échantillon de taille n , sur lequel nous appliquons toutes les méthodes exposées ci-dessus. Nous obtenons donc 500 estimations différentes d'un même quantile, du fait que les valeurs présentes dans l'échantillon ne sont pas les mêmes. Appelons \hat{q}_j un estimateur du quantile sur l'échantillon j . Nous nous intéressons à la valeur moyenne estimée, c'est-à-dire à $\frac{1}{500} \sum_{j=1}^{500} \hat{q}_j$. Plus cette valeur est proche de la valeur théorique de ce quantile q^{th} , mieux c'est.

De plus, nous regardons l'écart-quadratique moyen : $\frac{1}{500} \sum_{j=1}^{500} (\hat{q}_j - q^{th})^2$. Plus cette valeur est proche de 0, meilleur est l'estimateur. A titre de comparaison, nous comparons les résultats obtenus *via* les méthodes de théorie des valeurs extrêmes avec la méthode la plus simple d'estimation d'un quantile : la valeur empirique. C'est-à-dire que si nous recherchons le quantile à 95 % d'un échantillon de 100 valeurs, nous définissons le quantile empirique comme étant la cinquième plus grande valeur.

Enfin, nous calculons aussi les intervalles de confiance à 95 % et 90 % pour notre quantile, et les comparons selon deux critères : le taux de couverture qui doit être le plus proche possible du taux attendu (95 % ou 90 %), et la taille moyenne de l'intervalle qui doit être la plus faible possible - sous contrainte d'un bon taux de couverture.

Les trois tailles d'échantillon testées sont $n = 100$, 1 000 et 10 000. Les résultats concernant $n = 1 000$ et $n = 10 000$ figurent dans l'annexe B.

4.3.1 Loi Normale centrée réduite

Ici, X suit une loi normale centrée réduite. Nous testons trois tailles différentes d'échantillon : 100, 1 000 et 10 000. De plus, nous cherchons à estimer les quantiles à 95 % et à 99,5 %. Les valeurs théoriques de ces quantiles sont respectivement 1,645 et 2,576.

	Quantile moyen	Ecart quadratique	Seuil moyen	Ecart-type seuil
Empirique	1,59	0,043		
PWM m=10	1,62	0,041		
GPWM m=10	1,45	0,076		
RB PWM m=10	1,62	0,033		
RB GPWM m=10	1,45	0,067		
ME EMV	1,69	0,050	0,45	0,66
ME PWM	1,62	0,038	0,45	0,66
Hill EMV	1,70	0,063	1,21	0,25
Hill PWM	1,56	0,041	1,21	0,25
Pick EMV	1,66	0,035	0,22	0,32
Pick PWM	1,63	0,039	0,42	0,48

TABLE 4.7 – Quantile à 95 % pour un échantillon Normale centré réduit de taille 100

	Quantile moyen	Ecart quadratique	Seuil moyen	Ecart-type seuil
Empirique	2,15	0,29		
PWM m=10	2,65	0,17		
GPWM m=10	2,54	0,26		
RB PWM m=10	2,65	0,17		
RB GPWM m=10	2,55	0,19		
ME EMV	2,42	0,19	0,45	0,66
ME PWM	2,66	0,29	0,45	0,66
Hill EMV	2,42	0,19	1,21	0,25
Hill PWM	2,70	0,21	1,21	0,25
Pick EMV	2,41	0,19	0,22	0,32
Pick PWM	2,75	0,25	0,41	0,48

TABLE 4.8 – Quantile à 99,5 % pour un échantillon Normale centré réduit de taille 100

Premièrement, remarquons que nous n'avons pas les résultats pour les estimateurs EMV pour la méthode par blocs. En effet, les paramètres sont estimés sur 10 valeurs seulement et il arrive souvent que l'algorithme ne converge pas. Deuxièmement, nous pouvons noter que pour tous les estimateurs, une augmentation de la taille de l'échantillon réduit le biais et les erreurs quadratiques. Il y a donc bien convergence des quantiles estimés par méthode des blocs ou des excès vers la valeur théorique du quantile.

Plus précisément, dans l'approche par blocs, le choix de la taille des blocs s'avère fondamental pour l'estimation du quantile à 95 % : passer de $m = 50$ à $m = 30$ divise ainsi les erreurs quadratiques par deux pour $n = 1\,000$ et par quatre pour $n = 10\,000$. Cet effet est aussi visible pour le quantile à 99,5 %. Les estimateurs rao-blackwellisés semblent davantage robustes à la taille des blocs. De plus, nous observons que parmi tous les estimateurs utilisés, ceux obtenus avec la méthode GPWM présentent en général le biais et l'erreur quadratique les plus importants et cela pour toutes les tailles d'échantillon. Les estimateurs obtenus par EMV ou méthode PWM sont quant à eux relativement proches. Nous remarquons néanmoins que pour les échantillons de taille 1 000, les moments pondérés sont légèrement meilleurs si nous comparons les erreurs quadratiques, alors que pour $n = 10\,000$ c'est l'inverse.

L'approche des excès donne des résultats du même ordre de grandeur que l'approche par blocs (à condition d'avoir choisi la bonne taille des blocs). En ce qui concerne le choix du seuil, nous remarquons que la méthode de Hill donne des seuils plus élevés que les autres méthodes (entre 1,2 et 1,4 de seuil moyen alors qu'aucune autre méthode ne donne de seuil moyen supérieur à 1). Si nous regardons toutes les tailles d'échantillon, cette méthode est la plus stable, ce qui la rend globalement plus efficace que les autres, principalement pour $n = 10\,000$ où elle a un écart quadratique moyen bien inférieur aux autres procédures de choix de seuil.

Nous nous intéressons principalement aux échantillons de taille $n = 100$ puisque les données que nous analyserons dans la suite auront cet ordre de grandeur. Nous voyons que pour le quantile à 95 % les meilleures méthodes en terme d'écart-quadratique moyen sont celle des moments pondérés rao-blackwellisés (0,033) pour l'approche par blocs et celle par EMV avec choix de seuil par méthode de Pickands pour l'approche par excès (0,035). Ces deux méthodes restent supérieures ou égales aux autres pour le quantile à 99,5 %.

Maintenant que nous avons un estimateur du quantile, nous aimerions connaître notre degré de précision. Nous construisons donc des intervalles de confiance autour des quantiles que nous avons estimés, uniquement pour les méthodes s'étant révélés efficaces du point de vue de l'écart quadratique moyen, et ne présentons dans les tableaux que celles ayant un taux de couverture correct : il faut que ce taux soit supérieur à 50 % et ne soit pas largement inférieur au taux de couverture fourni par une autre méthode. Notons que nous n'avons pas de résultat théorique permettant de construire des intervalles de confiance dans le cadre rao-blackwellisé, et des intervalles de confiance obtenus numériquement par méthodes de Monte-Carlo ou bootstrap demandent beaucoup de temps de calcul.

	Taille moyenne	Taux de couverture	Borne inf moyenne	Borne sup moyenne
PWM	0,801	94,2 %	1,227	2,028
ME PWM	0,671	90,2 %	1,286	1,956
Pick EMV (Fisher théo)	0,585	87,1 %	1,364	1,949
Pick PWM	0,661	85,5 %	1,304	1,964

TABLE 4.9 – Propriétés des intervalles de confiance à 95 % pour le quantile à 95% d'un échantillon Normale centré réduit de taille 100

	Taille moyenne	Taux de couverture	Borne inf moyenne	Borne sup moyenne
PWM	0,671	89,8 %	1,286	1,957
ME PWM	0,563	84,5 %	1,339	1,902
Pick EMV (Fisher théo)	0,491	80,4 %	1,411	1,902
Pick PWM	0,554	81,1 %	1,357	1,911

TABLE 4.10 – Propriétés des intervalles de confiance à 90 % pour le quantile à 95 % d'un échantillon Normale centré réduit de taille 100

	Taille moyenne	Taux de couverture	Borne inf moyenne	Borne sup moyenne
PWM	1,047	76,2 %	1,922	2,969
Hill PWM	2,849	96,3 %	1,278	4,127
Pick PWM	2,452	96,6 %	1,529	3,981

TABLE 4.11 – Propriétés des intervalles de confiance à 95 % pour le quantile à 99,5 % d'un échantillon Normale centré réduit de taille 100

	Taille moyenne	Taux de couverture	Borne inf moyenne	Borne sup moyenne
PWM	0,915	72,8 %	2,017	2,932
ME PWM	1,896	85,1 %	1,709	3,605
Hill PWM	2,391	94,0 %	1,507	3,898
Pick PWM	2,058	94,8 %	1,726	3,784

TABLE 4.12 – Propriétés des intervalles de confiance à 90 % pour le quantile à 99,5 % d'un échantillon Normale centré réduit de taille 100

Pour les échantillons de taille 100, les meilleurs intervalles de confiance pour le quantile à 95 % sont fournis par la méthode par blocs avec estimateurs PWM. En effet, les taux de couverture sont très proches des taux ciblés : 94,2 % pour l'intervalle de confiance à 95 % et 89,8 % pour celui à 90 %. Pour le quantile à 99,5 %, nous observons en revanche l'inverse : la méthode des excès avec différents choix de seuil donne des résultats plus proches de ceux attendus. Remarquons que le seuil choisi par méthode de fonction moyenne des excès sous-estime l'intervalle à 90 % tandis que les méthodes Hill et Pickands le surestiment. De plus, les intervalles par méthode de Pickands sont plus petits que ceux obtenus par méthode de Hill mais ont paradoxalement un meilleur taux de couverture, la sélection de u semble donc plus pertinente.

4.3.2 Loi Student à 2 degrés de liberté

Nous faisons maintenant la même chose sur une loi de Student à 2 degrés de liberté, pour laquelle nous avons $\xi = \frac{1}{2}$: la queue de distribution est donc plus lourde que pour une loi normale. Le quantile à 95 % vaut 2,920 et celui à 99,5 % vaut 9,925.

	Quantile moyen	Ecart quadratique	Seuil moyen	Ecart-type seuil
Empirique	2,81	0,56		
PWM m=10	3,09	0,78		
GPWM m=10	2,43	0,62		
RB PWM m=10	3,09	0,70		
RB GPWM m=10	2,44	0,52		
ME EMV	3,05	0,61	0,87	0,66
ME PWM	2,90	0,47	0,87	0,66
Hill EMV	3,06	0,92	1,78	0,57
Hill PWM	2,74	0,46	1,78	0,57
Pick EMV	3,02	0,47	0,26	0,38
Pick PWM	2,90	0,43	0,57	0,73

TABLE 4.13 – Quantile à 95 % pour un échantillon Student à 2 degrés de liberté de taille 100

	Quantile moyen	Ecart quadratique	Seuil moyen	Ecart-type seuil
Empirique	6,07	24,22		
PWM m=10	9,19	25,42		
GPWM m=10	7,29	19,21		
RB PWM m=10	9,23	26,08		
RB GPWM m=10	7,33	18,09		
ME EMV	9,46	43,88	0,87	0,66
ME PWM	9,15	18,86	0,87	0,66
Hill EMV	10,08	62,10	1,78	0,57
Hill PWM	8,76	16,64	1,78	0,57
Pick EMV	9,97	83,94	0,26	0,38
Pick PWM	9,35	20,26	0,57	0,73

TABLE 4.14 – Quantile à 99,5 % pour un échantillon Student à 2 degrés de liberté de taille 100

Lorsque $n = 100$, pour la méthode par blocs, les estimateurs GPWM (normaux et rao-blackwellisés) fournissent des écarts quadratiques moyens plus faibles que les estimateurs PWM. Concernant la méthode des excès, les EMV sont moins bons au niveau de l'écart quadratique que les estimateurs PWM, et ce quelle que soit la procédure de choix du seuil. Paradoxalement, ils ont un biais plus faible, ce qui peut s'expliquer par le fait que les quantiles obtenus grâce aux EMV peuvent prendre des valeurs bien plus éloignées de la vraie valeur que pour les estimateurs PWM, ce qui se ressent davantage sur l'écart quadratique moyen que sur le biais. Pour les tailles d'échantillon supérieures, nous remarquons une convergence des biais et des écarts-quadratiques moyens vers 0. Lorsque $n = 1000$, les quantiles obtenus par méthode des excès ont des caractéristiques similaires avec une légère préférence pour la méthode de Cunnane avec $\alpha = 20\%$ et des estimateurs par maximum de vraisemblance. Concernant la méthode par blocs, $m = 30$ est préférable avec des meilleurs résultats pour les PWM rao-blackwellisés.

Nous pouvons maintenant étudier les propriétés des intervalles de confiance.

	Taille moyenne	Taux de couverture	Borne inf moyenne	Borne sup moyenne
PWM	3,973	95,8 %	1,849	5,822
GPWM	6,668	97,3 %	-0,887	5,781
ME EMV (Fisher théo)	2,120	85,4 %	1,968	4,088
ME PWM	2,071	84,8 %	1,860	3,932
ME EMV (Fisher obs)	2,076	83,4 %	2,010	4,087
Pick EMV (Fisher théo)	2,257	88,1 %	1,898	4,155
Pick PWM	2,148	83,2 %	1,809	3,957
Pick EMV (Fisher obs)	2,245	87,6 %	1,901	4,146

TABLE 4.15 – Propriétés des intervalles de confiance à 95 % pour le quantile à 95 % d'un échantillon Student à 2 degrés de liberté de taille 100

	Taille moyenne	Taux de couverture	Borne inf moyenne	Borne sup moyenne
PWM	3,342	93,2 %	1,993	5,334
GPWM	5,596	95,2 %	0,351	5,245
ME EMV (Fisher théo)	1,779	81,0 %	2,138	3,917
ME PWM	1,738	78,1 %	2,027	3,765
ME EMV (Fisher obs)	1,743	78,4 %	2,177	3,920
Pick EMV (Fisher théo)	1,894	84,0 %	2,079	3,974
Pick PWM	1,803	78,8 %	1,982	3,784
Pick EMV (Fisher obs)	1,884	83,5 %	2,082	3,966

TABLE 4.16 – Propriétés des intervalles de confiance à 90 % pour le quantile à 95 % d'un échantillon Student à 2 degrés de liberté de taille 100

	Taille moyenne	Taux de couverture	Borne inf moyenne	Borne sup moyenne
PWM	10,826	52,4 %	3,806	14,632
ME EMV (Fisher théo)	21,029	68,0 %	-0,870	20,159
ME PWM	17,349	76,4 %	-0,318	17,030
ME EMV (Fisher obs)	20,400	66,1 %	-0,725	19,675
Hill EMV (Fisher théo)	29,654	72,0 %	-3,944	25,710
Hill PWM	17,145	73,9 %	-0,060	17,084
Hill EMV (Fisher obs)	26,102	62,7 %	-2,939	23,163
Pick EMV (Fisher théo)	24,037	69,5 %	-1,982	22,055
Pick PWM	17,982	76,0 %	-0,475	17,508
Pick EMV (Fisher obs)	24,340	69,5 %	-2,174	22,167

TABLE 4.17 – Propriétés des intervalles de confiance à 95 % pour le quantile à 99,5 % d'un échantillon Student à 2 degrés de liberté de taille 100

	Taille moyenne	Taux de couverture	Borne inf moyenne	Borne sup moyenne
PWM	10,085	50,4 %	4,229	14,314
ME EMV (Fisher théo)	17,648	64,2 %	0,820	18,468
ME PWM	14,559	69,5 %	1,076	15,636
ME EMV (Fisher obs)	17,120	61,5 %	0,914	18,035
Hill EMV (Fisher théo)	24,886	68,7 %	-1,560	23,326
Hill PWM	14,388	70,4 %	1,318	15,706
Hill EMV (Fisher obs)	21,906	60,1 %	-0,841	21,065
Pick EMV (Fisher théo)	20,172	65,3 %	-0,049	20,123
Pick PWM	15,091	69,3 %	0,971	16,062
Pick EMV (Fisher obs)	20,427	64,7 %	-0,217	20,210

TABLE 4.18 – Propriétés des intervalles de confiance à 90 % pour le quantile à 99,5 % d'un échantillon Student à 2 degrés de liberté de taille 100

Il y a deux intervalles de confiance qui se détachent des autres : ceux obtenus par la méthode par blocs et estimateurs PWM; et ceux par la méthode des excès avec choix du seuil par procédure de Pickands et estimateurs du maximum de vraisemblance. En effet, ils conjuguent le mieux un bon taux de couverture et une taille d'intervalle raisonnable pour le quantile à 95 %. En revanche, nous n'avons pas de taux de couverture proche des taux ciblés lorsque nous nous intéressons au quantile à 99,5 %, il n'y a pas assez de données pour estimer précisément un quantile élevé de loi à queue épaisse. Sur des échantillons plus grands, nos taux de couverture redeviennent en accord avec ce que nous souhaitons.

4.3.3 Loi Student à 1 degré de liberté

Ici, X suit une loi de Student à 1 degré de liberté. ξ est donc égal à 1 : la queue de distribution est encore plus lourde que précédemment. Les quantiles à 95 % et 99,5 % valent respectivement 6,314 et 66,567.

	Quantile moyen	Ecart quadratique	Seuil moyen	Ecart-type seuil
Empirique	6,12	8,17		
EMV m=10	25,54	22034,25		
PWM m=10	9,62	67,99		
GPWM m=10	5,13	7,64		
RB EMV m=10	35,18	19574,67		
RB PWM m=10	9,55	62,42		
RB GPWM m=10	5,09	5,85		
ME EMV	6,42	10,70	1,62	1,91
ME PWM	6,09	6,25	1,62	1,91
Hill EMV	6,63	23,85	2,97	1,49
Hill PWM	6,11	7,59	2,97	1,49
Pick EMV	6,57	6,72	0,34	0,46
Pick PWM	6,03	5,89	0,99	1,82

TABLE 4.19 – Quantile à 95 % pour un échantillon Student à 1 degré de liberté de taille 100

	Quantile moyen	Ecart quadratique	Seuil moyen	Ecart-type seuil
Empirique	27,59	2235,66		
PWM m=10	56,89	6610,46		
GPWM m=10	32,17	2057,05		
RB PWM m=10	56,85	6010,43		
RB GPWM m=10	32,03	1901,26		
ME EMV	97,03	42298,35	1,62	1,91
ME PWM	38,21	1497,85	1,62	1,91
Hill EMV	98,27	34152,51	2,97	1,49
Hill PWM	37,22	1540,60	2,97	1,49
Pick EMV	92,17	26526,42	0,34	0,46
Pick PWM	36,78	1426,93	0,99	1,82

TABLE 4.20 – Quantile à 99,5 % pour un échantillon Student à 1 degré de liberté de taille 100

Tout d'abord, nous constatons que le niveau général des erreurs quadratiques est plus élevé que dans les cas précédents. C'est particulièrement vrai pour la méthode des moments pondérés dans l'approche par blocs : les estimateurs obtenus par cette méthode présentent en effet des erreurs quadratiques très importantes, ce qui les rend donc très peu fiables. Ce résultat est dû

au fait que nous sommes à la limite de validité théorique de ces estimateurs.

Dans l'approche par blocs, les meilleurs résultats proviennent de la méthode des moments généralisés (version rao-blackwellisée) alors que dans l'approche des excès, comme dans le premier exemple, la meilleure méthode dépend de la taille de l'échantillon : l'utilisation du maximum de vraisemblance est préférable pour les grands échantillons et celle de la méthode des moments pondérés pour les petits échantillons.

Pour $n = 100$, la méthode de Pickands est celle qui fournit le meilleur seuil au sens de l'écart quadratique moyen. Pour $n = 1\,000$, les méthodes de Pickands et Cunnane sont équivalentes. Notons que dans le cas du quantile à 99,5 %, les quantiles obtenus par la méthode PWM sont meilleurs en terme d'écart quadratique moyen mais sont beaucoup plus biaisés que ceux obtenus par la méthode EMV. Enfin, pour $n = 10\,000$, les méthodes EMV sont globalement meilleures avec notamment le choix du seuil par méthode de Hill qui donne un écart quadratique moyen de 45,8 contre 78,1 pour l'estimateur empirique.

	Taille moyenne	Taux de couverture	Borne inf moyenne	Borne sup moyenne
PWM	35,649	92,4 %	3,331	38,980
ME EMV (Fisher théo)	7,704	79,9 %	2,669	10,373
ME PWM	6,005	71,3 %	3,071	9,076
ME EMV (Fisher obs)	7,774	79,6 %	2,536	10,310
Hill EMV (Fisher théo)	7,245	77,0 %	3,053	10,297
Hill PWM	5,497	66,3 %	3,364	8,861
Hill EMV (Fisher obs)	7,030	73,4 %	3,114	10,143
Pick EMV (Fisher théo)	8,864	84,8 %	2,140	11,004
Pick PWM	5,759	73,9 %	3,198	8,957
Pick EMV (Fisher obs)	8,869	84,6 %	2,138	11,007

TABLE 4.21 – Propriétés des intervalles de confiance à 95 % pour le quantile à 95 % d'un échantillon Student à 1 degré de liberté de taille 100

	Taille moyenne	Taux de couverture	Borne inf moyenne	Borne sup moyenne
PWM	31,692	87,2 %	3,776	35,468
ME EMV (Fisher théo)	6,466	76,3 %	3,288	9,754
ME PWM	5,039	66,2 %	3,554	8,593
ME EMV (Fisher obs)	6,524	76,4 %	3,161	9,685
Hill EMV (Fisher théo)	6,080	71,1 %	3,635	9,715
Hill PWM	4,613	58,9 %	3,806	8,419
Hill EMV (Fisher obs)	5,900	68,0 %	3,679	9,578
Pick EMV (Fisher théo)	7,439	82,6 %	2,853	10,292
Pick PWM	4,833	67,3 %	3,661	8,494
Pick EMV (Fisher obs)	7,443	82,6 %	2,851	10,294

TABLE 4.22 – Propriétés des intervalles de confiance à 90 % pour le quantile à 95 % d'un échantillon Student à 1 degré de liberté de taille 100

	Taille moyenne	Taux de couverture	Borne inf moyenne	Borne sup moyenne
PWM	147,738	44,2 %	9,511	157,249
ME PWM	96,065	50,9 %	-6,212	89,853
Hill PWM	87,845	47,9 %	-5,342	82,502
Pick PWM	71,828	45,1 %	3,875	75,703

TABLE 4.23 – Propriétés des intervalles de confiance à 95 % pour le quantile à 99,5 % d'un échantillon Student à 1 degré de liberté de taille 100

	Taille moyenne	Taux de couverture	Borne inf moyenne	Borne sup moyenne
PWM	141,551	42,8 %	11,535	153,086
ME PWM	80,620	46,5 %	1,510	82,131
Hill PWM	73,722	43,2 %	1,719	75,441
Pick PWM	60,280	41,6 %	9,649	69,929

TABLE 4.24 – Propriétés des intervalles de confiance à 90 % pour le quantile à 99,5 % d'un échantillon Student à 1 degré de liberté de taille 100

Nous retrouvons l'effet vu pour la loi de Student à 2 degrés de liberté, mais amplifié : pour de petits échantillons, les intervalles de confiance sont très grands et peu précis. Pour le quantile à 95 %, la méthode des blocs fournit un taux de couverture proche du taux souhaité, mais le prix à payer est un intervalle de confiance à 95 % compris entre 3,3 et 39,0 ce qui est trop grand pour être utilisable en pratique. Lorsque les tailles d'échantillon augmentent, les intervalles sont évidemment plus fins et ont de meilleurs taux de couverture, mais nous retiendrons surtout que l'épaisseur de queue d'une distribution joue un grand rôle sur la qualité des intervalles de confiance, surtout sur des échantillons de petite taille.

4.3.4 Analyse des résultats

Ce chapitre nous a permis de mettre en évidence le comportement des différentes méthodes d'estimation du quantile et de calcul d'intervalles de confiance en fonction de la taille des données et de l'épaisseur de leur queue de distribution. Conformément à nos attentes, les erreurs quadratiques sont les plus faibles lorsque : les échantillons sont grands, ξ est proche de 0, et le quantile recherché n'est pas "trop" extrême. Par exemple pour le quantile à 95 % de la loi normale dans le cas $n = 10\,000$, nous obtenons des erreurs quadratiques de l'ordre de 10^{-3} , alors que dans la situation inverse (quantile à 99,5 % de la Student à 1 degré de liberté pour $n = 100$) elles sont de l'ordre de 10^3 dans le meilleur des cas.

Pour les échantillons de taille 100, les estimateurs des moments sont plus fiables que ceux du maximum de vraisemblance, qui ont tendance à être beaucoup plus volatiles. Cela est d'autant plus vrai lorsque ξ croît. En revanche, pour de grands échantillons, les estimateurs du maximum de vraisemblance sont préférables bien que les estimateurs des moments pondérés restent convenables pour $\xi < 0,5$. Dans le cas de ξ élevé, les estimateurs des moments pondérés généralisés fournissent également de bons résultats.

Concernant l'approche par blocs, la rao-blackwellisation permet toujours d'améliorer la précision des quantiles estimés. Le choix de la taille des blocs a aussi son importance, néanmoins sur des données réelles nous ne pouvons fixer celle-ci arbitrairement : il faut tenir compte de la structure des données (regroupement des données par mois, par an, etc.).

Pour l'approche par excès, le choix du seuil par procédure de Pickands nous semble plus stable que les autres : les erreurs quadratiques sont parfois plus petites que toutes les autres, et elles sont rarement les plus fortes. Comme nous avons davantage de valeurs pour cette approche, elle donne généralement de meilleurs résultats que celle par blocs (erreurs quadratiques plus faibles).

Les mêmes conclusions s'appliquent pour les intervalles de confiance.

Cette étude nous permet de définir un mode opératoire lorsque nous serons en présence de données réelles : premièrement nous choisirons une taille de blocs en fonction de la structure des données, et sélectionnerons le seuil par comparaison des différentes méthodes tout en privilégiant la procédure de Pickands lorsque celles-ci divergent. Nous estimerons ensuite ξ par les différentes méthodes et en déduirons l'applicabilité ou non des différentes méthodes. Si l'échantillon est de taille faible, nous privilégierons les méthodes des moments pondérés dans le cas où $\xi < 1$ (ou généralisés pour l'approche par blocs si $\xi < 1,5$). Cela nous permettra d'avoir un "vrai" estimateur de ξ et des autres paramètres pour en déduire un estimateur du quantile. Nous préfererons celui obtenu par méthode des excès, sauf dans des cas particuliers comme nous le verrons plus loin.

Une autre possibilité, plus simple, sera de tester l'adéquation de notre échantillon à une loi, en testant la proximité des queues de distribution. De cette manière, nous pourrions obtenir un quantile paramétrique. Le chapitre suivant présente cette approche.

Chapitre 5

Tests d'adéquations à une loi

Lorsque nous sommes en présence d'un échantillon dont nous ignorons la loi, une pratique usuelle est de choisir une famille de lois, d'estimer les paramètres qui correspondent le mieux à notre échantillon, et de tester la qualité de la modélisation *via* un test d'adéquation sur l'ensemble de la distribution (comme Kolmogorov-Smirnov). Toutefois, une bonne adéquation globale n'implique pas forcément une bonne adéquation en queue de distribution. Or, c'est cette zone qui nous intéresse le plus lorsque nous recherchons des quantiles extrêmes. Nous sommes donc intéressés par un test de la proximité des densités théorique et empirique en queue de distribution. C'est la raison pour laquelle nous traitons dans ce chapitre de l'ET Test et de son extension, le GPD Test.

5.1 ET Test

5.1.1 Utilisation de la loi limite d'un estimateur de quantile

L'idée de ce test est de s'interroger sur l'adéquation d'un échantillon de données $(X_i)_{i=1,\dots,n}$ de fonction de répartition F à la famille de fonctions de répartition F_θ , en prenant comme critère non pas une proximité en partie "centrale" des fonctions de répartition mais plutôt en queue de distribution. Il porte le nom d'*Exponential Tail Test* du fait qu'il s'applique uniquement aux échantillons pour lesquels F appartient au domaine d'attraction de Gumbel ($\xi = 0$).

Nous effectuons donc le test suivant :

$$H_0 : X \in (F_\theta)_{\theta \in \Theta} \qquad H_1 : X \notin (F_\theta)_{\theta \in \Theta}$$

Pour ce faire, nous estimons le quantile de l'échantillon en utilisant la théorie des valeurs extrêmes. Nous recherchons donc le seuil u optimal *via* les méthodes présentées précédemment puis nous estimons les paramètres ξ et σ qui nous permettent d'obtenir un estimateur du quantile d'ordre p de la loi :

$$\hat{q}_{ET} = u + \hat{\sigma} \ln \left(\frac{N_u}{n(1-p)} \right)$$

Nous comparons ce dernier au quantile paramétrique, obtenu en prenant l'estimateur du maximum de vraisemblance $\hat{\theta}$ de l'échantillon, et en définissant :

$$\hat{q}_{param} = F_{\hat{\theta}}^{-1}(1-p)$$

Nous n'allons évidemment pas obtenir la même valeur pour ces deux paramètres, la question est de quantifier l'écart "acceptable" pour que la loi $F_{\hat{\theta}}$ représente correctement la queue de distribution de l'échantillon X . Pour cela, nous construisons un intervalle de confiance autour de \hat{q}_{ET} et regardons si l'estimateur paramétrique appartient à cet intervalle.

En effet, nous pouvons décomposer l'écart entre le vrai quantile q et \hat{q}_{ET} :

$$q - \hat{q}_{ET} = (q - q_{ET}) + (q_{ET} - \hat{q}_{ET})$$

Le premier terme est une erreur d'approximation que nous pouvons noter δ , le deuxième une erreur d'estimation. GARRIDO montre dans [GAR02] la convergence en loi suivante :

$$q_{ET} - \hat{q}_{ET} \rightsquigarrow N \left(0, \frac{\left(\sigma \ln \left(\frac{N_u}{n(1-p)} \right) \right)^2}{N_u} \right)$$

Nous obtenons alors un intervalle de confiance pour q_{ET} :

$$IC_{ET}(\alpha) = \left[\hat{q}_{ET} \pm \frac{\sigma \ln \left(\frac{N_u}{n(1-p)} \right)}{\sqrt{N_u}} z_{1-\frac{\alpha}{2}} \right]$$

où $z_{1-\frac{\alpha}{2}}$ est le quantile d'ordre $1 - \frac{\alpha}{2}$ d'une loi normale centrée réduite. Ainsi, nous en déduisons un intervalle de confiance pour q :

$$IC_{Th}(\alpha) = \left[\hat{q}_{ET} + \delta \pm \frac{\sigma \ln \left(\frac{N_u}{n(1-p)} \right)}{\sqrt{N_u}} z_{1-\frac{\alpha}{2}} \right]$$

En pratique, la quantité δ n'est pas connue puisqu'elle nécessite la connaissance des "vraies" paramètres du modèle ET, ce qui est impossible puisque la loi de l'échantillon est inconnue. Il est possible d'en faire une approximation mais cette dernière dépend de beaucoup de paramètres. De plus, une approximation de cette quantité induirait une nouvelle erreur d'approximation qu'il faudrait mesurer. Pour ces raisons, une simplification du test en utilisant une méthode de bootstrap peut être effectuée.

5.1.2 Utilisation d'un bootstrap paramétrique

L'idée de cette partie est d'avoir un intervalle de confiance pour l'erreur d'approximation δ , et de s'en servir pour déterminer si l'écart entre q_{GPD} et q_{param} est acceptable ou non. Pour j allant de 1 à N , nous suivons la procédure suivante :

- Nous simulons un échantillon de taille n et de fonction de répartition $F_{\hat{\theta}}$. Sous H_0 , cet échantillon suit donc la même loi que X .
- Nous calculons $\hat{q}_{ET,j}^*$, défini comme au-dessus. Nous prenons comme seuil la même statistique d'ordre que pour l'échantillon X , les paramètres sont en revanche recalculés.
- Nous calculons l'estimateur du maximum de vraisemblance $\hat{\theta}_j^*$ du nouvel échantillon. Cet estimateur est évidemment proche de $\hat{\theta}$, à plus forte raison lorsque n est grand.
- Nous calculons $\hat{q}_{param,j}^*$ grâce à l'estimateur $\hat{\theta}_j^*$.
- Nous avons alors à disposition un échantillon d'écart $\delta_j^* = \hat{q}_{param,j}^* - \hat{q}_{ET,j}^*$ sous H_0 .

Nous ordonnons l'échantillon des δ_j^* par ordre croissant. Pour un test de niveau α , nous définissons $IC_{boot}(\alpha) = \left[\delta_{N \frac{1-\alpha}{2}}^*, \delta_{N \frac{1+\alpha}{2}}^* \right]$. Nous rejetons alors H_0 si $\hat{q}_{param} - \hat{q}_{ET} \notin IC(\alpha)$.

De manière moins formelle, le principe est simplement de se dire que si X peut être approché par la loi $F_{\hat{\theta}}$, alors l'écart entre le quantile de X obtenu par théorie des valeurs extrêmes et celui obtenu paramétriquement doit ressembler aux écarts que nous obtenons en prenant à la place de X des échantillons ayant été tirés suivant la loi $F_{\hat{\theta}}$. S'il y a une différence significative, c'est-à-dire si nous sortons de l'intervalle de confiance à 95 %, alors X ne peut pas être approché par la loi $F_{\hat{\theta}}$.

5.2 GPD Test

Le GPD Test est une généralisation de l'ET Test qui ne nécessite pas de faire l'hypothèse que F est dans le domaine d'attraction de GUMBEL. Nous estimons alors deux paramètres : ξ et σ . Ceux-ci conduisent à l'estimateur GPD que nous avons déjà vu :

$$\hat{q}_{GPD} = u + \frac{\hat{\sigma}}{\hat{\xi}} \left(\left(\frac{n}{N_u} (1-p) \right)^{-\hat{\xi}} - 1 \right)$$

La loi limite de l'erreur $q_{GPD} - \hat{q}_{GPD}$ n'est pas la même que pour $q_{ET} - \hat{q}_{ET}$, mais en réalité, nous n'utiliserons pas cette version du test. En effet, GARRIDO a montré qu'elle manquait de puissance pour les échantillons de petite taille, critère qui caractérise les échantillons sur lesquels nous travaillons. Nous utilisons donc la procédure bootstrap, qui est en tout point similaire à celle de la partie précédente en utilisant l'estimateur \hat{q}_{GPD} .

Dans la suite, nous présentons une application du GPD Test. Prenons un échantillon suivant une loi mélangée : normale centrée réduite avec probabilité 99 % et normale de moyenne 3 et d'écart-type 2 avec probabilité 1 %. La figure 5.1 présente un histogramme de cette loi, sur lequel nous remarquons une queue plus épaisse que pour une loi normale centrée réduite.

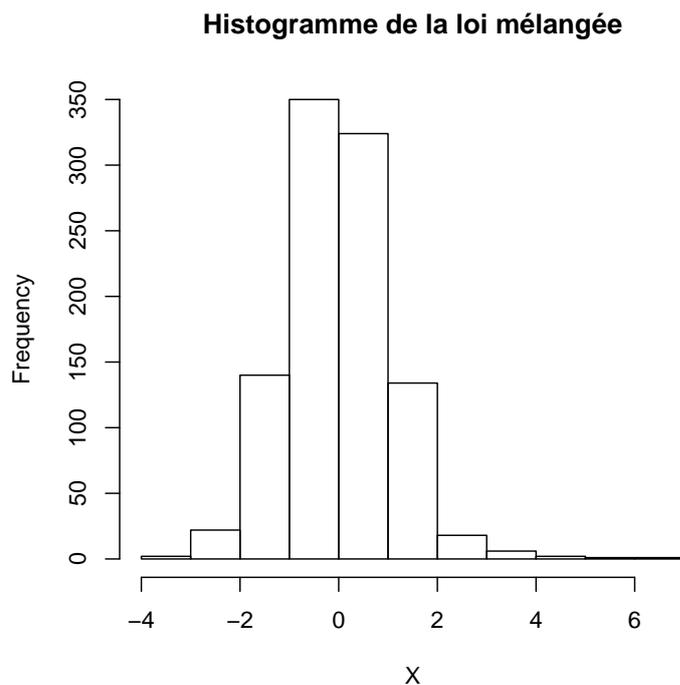


FIGURE 5.1 – Histogramme de la loi mélangée : $N(0, 1)$ avec probabilité 99 % et $N(3, 4)$ avec probabilité 1 %

Nous effectuons un test de Kolmogorov-Smirnov sous R pour tester l'adéquation à une loi normale : la p -value est de 0,622, nous acceptons donc l'hypothèse selon laquelle la loi mélangée suit une loi normale. Toutefois, nous savons que cela ne prend que marginalement en compte les queues de distribution.

Effectuons maintenant un GPD Test. Les estimateurs du maximum de vraisemblance pour la moyenne et l'écart-type d'une loi normale sont 0,02 et 1,08 : ils captent évidemment très peu

la queue de distribution. Si nous cherchons maintenant les paramètres GPD pour la queue de distribution (avec un seuil choisi par méthode de Hill), nous obtenons les valeurs suivantes :

	Echantillon de loi mélange	Paramètres moyen pour une $N(0,02; 1,08^2)$
$\hat{\xi}$	0,23	-0,20
$\hat{\sigma}$	0,50	0,60

TABLE 5.1 – Comparaison des paramètres estimés pour notre échantillon de loi mélange par rapport à une loi normale

Nous remarquons que la loi normale a un estimateur $\hat{\xi} < 0$ tandis que pour la loi mélange, nous avons $\hat{\xi} > 0$, ce qui est bien le résultat attendu puisque la queue de distribution de l'échantillon mélange est plus épaisse. Nous calculons maintenant le δ pour le quantile à 99,5 % de la loi mélange. Pour rappel, il s'agit de la différence entre le quantile estimé par méthode des excès, et le quantile paramétrique obtenu en faisant l'approximation de normalité. Nous obtenons $\delta = -0,83$.

Si nous calculons maintenant les δ_j^* pour 500 échantillons normaux de paramètres 0,02 et 1,08, ils sont compris entre -0,4 et 0,4. Nous obtenons comme intervalle de confiance à 95 % pour δ sous l'hypothèse que la loi mélange est une normale : $[-0,25; 0,22]$. δ n'appartient manifestement pas à cet intervalle de confiance, nous rejetons donc l'hypothèse de normalité. Nous voyons donc ici l'intérêt de ce test lorsque nous sommes intéressés par les queues de distribution.

Ce chapitre souligne les limites d'une pratique consistant à calculer des quantiles paramétriques en ayant au préalable effectué un test de Kolmogorov-Smirnov pour vérifier l'adéquation de la loi de l'échantillon avec une loi théorique. En effet, elle peut potentiellement induire une mauvaise estimation des quantiles extrêmes étant donné que ce test ne prend en compte les queues de distribution que de manière limitée. L'exemple précédent montre la nécessité d'effectuer également un test d'adéquation des queues de distribution lorsque nous cherchons à calculer des quantiles extrêmes.

Chapitre 6

Extension aux séries stationnaires

Les modèles décrits jusqu'à maintenant reposent sur une hypothèse mathématique contraignante : l'indépendance des données. Or, sur certains jeux de données, cette supposition peut sembler peu réaliste. Nous chercherons donc à étendre les résultats obtenus à des séries stationnaires.

6.1 Extension des résultats sur la loi des valeurs extrêmes aux séries stationnaires

Rappelons tout d'abord la définition de la stationnarité forte. La série $(X_i)_{i=1,\dots,n}$ est dite fortement stationnaire si : $\forall n > 0$ entier, $\forall h > 0$ réel, $\forall (t_1, t_2, \dots, t_n)$ réels, $(X_{t_1}, X_{t_2}, \dots, X_{t_n})$ a la même loi que $(X_{t_1+h}, X_{t_2+h}, \dots, X_{t_n+h})$.

Les résultats obtenus dans le cas i.i.d peuvent être généralisés aux séries stationnaires sous la condition de non-dépendance à long terme $D(u_n)$ qui s'écrit ainsi :
Soit $\chi = (X_i)_{i=1,\dots,n}$ une série stationnaire. On dit que la série vérifie la condition de non-dépendance à long terme $D(u_n)$ si :

$$\begin{aligned} & \forall i_1 < \dots < i_p < j_1 < \dots < j_q \text{ avec } j_1 - i_p > l_n, \\ & |P(X_{i_1} \leq u_n, \dots, X_{i_p} \leq u_n, X_{j_1} \leq u_n, \dots, X_{j_q} \leq u_n) - P(X_{i_1} \leq u_n, \dots, X_{i_p} \leq u_n)P(X_{j_1} \leq u_n, \dots, X_{j_q} \leq u_n)| \\ & \leq \alpha(n, l_n) \end{aligned}$$

avec $\alpha(n, l_n) \rightarrow 0$ pour l_n une suite telle que $l_n/n \xrightarrow{n \rightarrow +\infty} 0$, et u_n une suite réelle.

Intuitivement, la formulation mathématique de cette condition décrit les séries stationnaires pour lesquelles la corrélation temporelle entre deux observations décroît lorsqu'elles sont éloignées dans le temps. Il s'agit donc de séries avec une dépendance à long terme négligeable. Remarquons que les séries indépendantes vérifient la condition $D(u_n)$ pour toute suite (u_n) .

En 1983, LEADBETTER a montré le théorème suivant dans [LEA93a] :

Théorème 3 (Leadbetter).

Soit $\chi^* = (X_i^*)_{i=1,\dots,n}$ une suite de variable i.i.d avec la même distribution marginale que X série stationnaire. On définit les suites $M_n = \max(X_1, \dots, X_n)$ et $M_n^* = \max(X_1^*, \dots, X_n^*)$. S'il existe des constantes $a_n > 0$ et b_n et une distribution non-dégénérée telle que :

$$\lim_{n \rightarrow +\infty} P\left(\frac{M_n^* - b_n}{a_n} \leq x\right) = G(x)$$

et si le processus stationnaire X vérifie la condition $D(u_n)$ avec $u_n = a_n x + b_n$ pour tout x tel que $G_1(x) > 0$ alors : les maxima renormalisés de χ^* suivent asymptotiquement une GEV de paramètres (σ^*, ξ^*, μ^*) et la loi limite des maxima des blocs de la série χ est également une loi GEV de paramètres (σ, ξ, μ) et de fonction de répartition G^θ où $0 < \theta \leq 1$ est appelé *extremal index*.

Les paramètres (σ^*, ξ^*, μ^*) et (σ, ξ, μ) sont ainsi reliés par :

$$\begin{cases} \sigma &= \sigma^* \theta^{\xi^*} \\ \xi &= \xi^* \\ \mu &= \mu^* + \frac{\sigma^*}{\xi^*} (\theta^{\xi^*} - 1) \end{cases}$$

θ mesure la dépendance à court terme des valeurs extrêmes, le cas i.i.d correspondant à $\theta = 1$.

En supposant la condition $D(u_n)$ vérifiée, les paramètres de la GEV peuvent ainsi s'estimer de la même façon que dans le cas i.i.d. avec les méthodes EMV, PWM et GPWM.

6.1.1 Amélioration des estimateurs avec la rao-blackwellisation

Là encore, il est possible d'améliorer la qualité des estimateurs en effectuant des permutations et en calculant les estimateurs rao-blackwellisés. Cependant, la réduction de la variance dépend du choix des permutations. FERRO et PEZZULLI montrent dans [FP] que des permutations possibles sont les n permutations qui transforment l'échantillon χ en : $\chi_1 = \chi$ et $\chi_j = (X_j, \dots, X_n, X_1, \dots, X_{j-1})$ pour j entre 2 et n .

Elles permettent de conserver la structure de l'échantillon. En effectuant la moyenne des estimateurs calculés sur chacun de ces n échantillons, on obtient les estimateurs rao-blackwellisés.

6.2 Extension des résultats sur la loi des excès aux séries stationnaires

Le premier point à noter est que nous ne pouvons pas faire abstraction de la stationnarité des données, comme cela peut être fait pour la loi des valeurs extrêmes. Un des effets de la non-indépendance est qu'une valeur élevée de la série est susceptible d'être suivie d'autres valeurs élevées : il y a généralement une auto-corrélation positive dans les séries financières et actuarielles. Lorsque nous prenons un maximum, le fait qu'il y ait plus de valeurs élevées que d'habitude ne revêt pas une grande importance puisque nous n'en prenons qu'une seule, alors que si nous considérons tous les excès chaque valeur va compter. Ainsi, nous comprenons intuitivement pourquoi - si nous appliquons le même traitement que pour les séries indépendantes - nous risquons d'avoir des groupes (*clusters* dans la littérature) de valeurs supérieures au seuil qui vont avoir tendance à surestimer l'épaisseur de la queue de distribution.

Une solution est d'appliquer une procédure de *de-clustering*, consistant à extraire de chaque groupe une seule valeur d'excès. Les valeurs ainsi obtenues seront ensuite traitées comme dans le cas indépendant. Par exemple, considérons la série stationnaire suivante :

$$\forall t \geq 1, X_t = \rho X_{t-1} + \epsilon_t$$

avec $\rho = 0,8$, $\epsilon_t \sim N(0, 1) \forall t$, $X_0 \sim N\left(0, \frac{1}{1-\rho^2}\right)$, $(\epsilon_t)_t$ indépendants entre eux et de X_0 . Il s'agit d'un processus auto-régressif. Nous pouvons le représenter, en prenant par exemple un seuil égal à 1 :

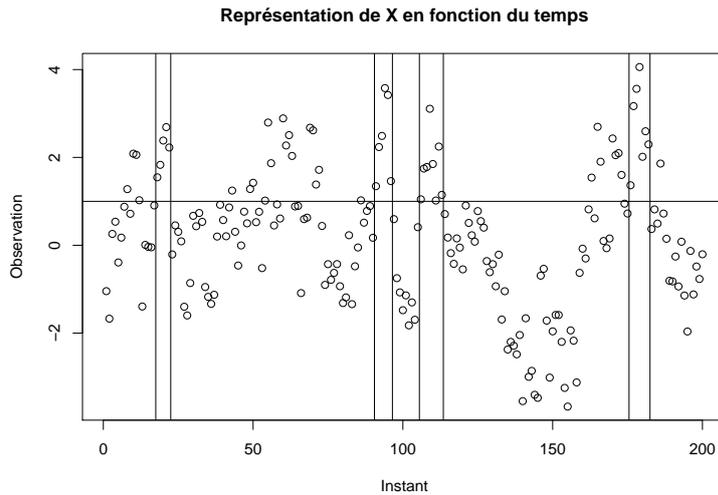


FIGURE 6.1 – Représentation d'une série auto-régressive de paramètre 0,8

Nous définissons de manière formelle un *cluster* comme un ensemble de valeurs consécutives toutes supérieures au seuil. Dès lors qu'une seule valeur est en-dessous du seuil, le *cluster* prend fin. Dans la figure 6.1, les barres verticales délimitent les *clusters* d'au moins 5 valeurs supérieures au seuil. Nous voyons qu'ils sont au nombre de 4. Au total, nous obtenons 16 *clusters*, et le plus grand compte 8 valeurs consécutives.

Nous pourrions nous demander si cela est réellement dû à la forte auto-corrélation de la série, ou si ces schémas sont monnaie courante même dans des séries indépendantes. Nous vérifions cela en regardant maintenant la série des $(\epsilon_t)_t$, qui elle est indépendante et qui a justement servi à générer les $(X_t)_t$. Cette série a une variance plus faible que la précédente, nous abaissons donc la valeur du seuil de manière à avoir exactement autant de dépassements (en l'occurrence, 53). Cela nous conduit à prendre un seuil de 0,57.

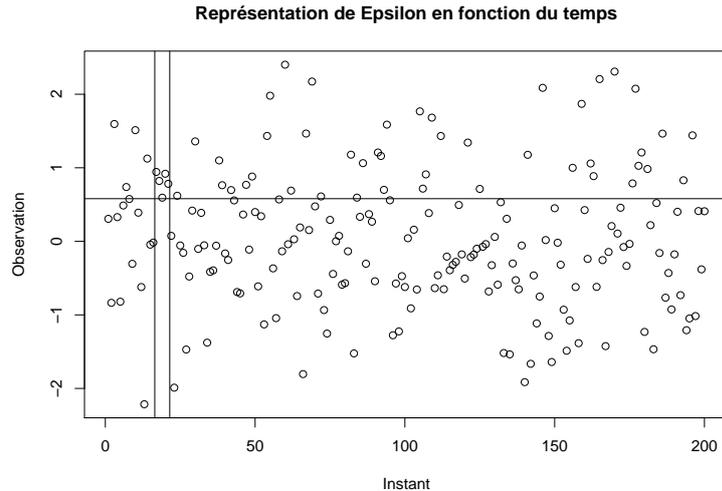


FIGURE 6.2 – Représentation d’une série normale centrée réduite indépendante

Il n’y a qu’un seul *cluster* de taille égale à 5, et aucun de taille supérieure. De plus, il y a un total de 38 *clusters* ce qui est bien plus élevé que les 16 trouvés précédemment, pour un nombre total de dépassements égal. Le phénomène de regroupement des valeurs extrêmes est donc bien présent.

COLES indique dans [COL01] que nous pouvons alors prendre le maximum de chaque *cluster* et considérer que ce sont des observations indépendantes de loi de Pareto généralisée. Cela nous permet d’estimer ξ et σ par les méthodes habituelles. De plus, COLES affirme que le quantile d’ordre p de la distribution marginale dont sont issues les données peut s’approcher par la formule suivante :

$$y \simeq u + \frac{\hat{\sigma}}{\hat{\xi}} \left(\left(\frac{n}{N_c} (1-p) \right)^{-\hat{\xi}} - 1 \right)$$

où N_c est le nombre de *clusters*.

Nous remarquons que l’expression est très proche de l’équation 3.1, dans laquelle N_u a été remplacée par N_c , ce qui semble logique puisque tout se passe comme si nous n’avions que N_c excès au-dessus du seuil. Cette procédure nous permet donc d’estimer le quantile de la distribution marginale dans le cas où l’échantillon n’est pas indépendant mais seulement stationnaire.

Au sujet des intervalles de confiance, rappelons que la delta-méthode faisait apparaître un paramètre a qui était égal à $\frac{n}{N_u} (1-p)$. De la même façon que pour le quantile, il suffit de poser $a = \frac{n}{N_c} (1-p)$ pour que les intervalles de confiance restent valables.

Troisième partie

Applications

Chapitre 7

Estimation du quantile sur des données réelles : indice immobilier

La partie précédente nous a permis d'introduire les concepts théoriques nécessaires à ce mémoire. Nous allons maintenant montrer comment ceux-ci s'appliquent à un modèle interne, et présenterons cela selon deux aspects : dans un premier temps nous verrons comment obtenir des valeurs de chocs extrêmes sur des séries financières avec un exemple sur l'indice immobilier IPD UK (toutes les données que nous utiliserons proviendront de *Bloomberg*), puis nous verrons comment modéliser l'intégralité de la distribution des facteurs de risque de marché.

Dans ce chapitre, nous utiliserons les méthodes décrites précédemment afin de les appliquer à un cas réel : les actifs immobiliers. Parmi les actifs possédés par les compagnies d'assurance, l'immobilier occupe généralement une part non-négligeable du portefeuille. Dans le cadre de la réglementation Solvabilité II, il convient donc de gérer le risque induit par la possession de biens immobiliers.

Le choc à 0,5 % sur l'immobilier est défini comme le quantile du même ordre de la distribution des rendements annuels de l'indice immobilier. Pour notre étude, nous choisissons de considérer l'IPD UK car il s'agit de l'un des indices avec le plus d'observations disponibles. De plus, cela nous permettra de comparer nos résultats avec le choc du QIS 5, qui a été calibré sur cet indice. Nous avons à notre disposition les valeurs trimestrielles de l'IPD UK du 31 décembre 1986 au 30 septembre 2010. Notre problématique est donc d'estimer la valeur de ce choc en utilisant la théorie développée dans les parties précédentes. Nous pourrions ainsi comparer le quantile obtenu par théorie des valeurs extrêmes avec celui obtenu par méthode standard. Toutefois, nous devons tout d'abord définir les rendements à utiliser puisque nous avons besoin de rendements annuels et que ceux à notre disposition sont trimestriels. Les possibilités à notre disposition sont d'annualiser les rendements trimestriels, ou alors de considérer des rendements annuels glissants.

7.1 Définition des rendements et étude de leur stationnarité

7.1.1 Rendements trimestriels annualisés

Appelons I_0 la valeur initiale de l'indice, et I_t sa valeur à l'instant t . Le rendement trimestriel est alors :

$$r_t = \frac{I_t - I_{t-1}}{I_{t-1}}, \forall t > 0$$

Nous définissons le rendement trimestriel annualisé R_t :

$$R_t = (1 + r_t)^4 - 1$$

Voici ce que nous obtenons :

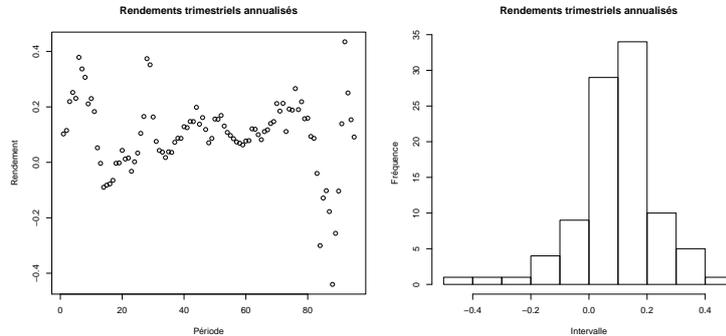


FIGURE 7.1 – Etude de la série des rendements trimestriels annualisés

Nous observons que la distribution est asymétrique avec une queue plus épaisse à gauche, qui est le côté qui nous intéresse. Nous le vérifions en regardant les moments de la distribution :

Minimum	Maximum	Moyenne	Ecart-type	Coefficient d'asymétrie	Excès de Kurtosis
-44,03 %	43,89 %	9,83 %	13,64 %	-83,71 %	280,70 %

TABLE 7.1 – Caractéristiques de la distribution des rendements trimestriels annualisés

Nous remarquons qu'au cours de certaines périodes, les valeurs prises sont très proches les unes des autres. Cela conduit à nous interroger sur la pertinence de l'hypothèse d'indépendance. Pour vérifier notre intuition, nous mettons en œuvre un test de Wald-Wolfowitz d'hypothèse nulle l'indépendance de la série, présenté succinctement en annexe. Nous obtenons une *p-value* de $2,2 \times 10^{-16}$, ce qui nous conduit donc à rejeter l'hypothèse d'indépendance. Un autre test d'indépendance dans le cadre des séries temporelles est le test de Ljung-Box de même hypothèse nulle que le test précédent, qui donne une *p-value* inférieure à $2,2 \times 10^{-16}$. Cela confirme le résultat du test précédent.

Afin de pouvoir estimer un quantile de cette distribution grâce aux techniques présentées ci-dessus, nous avons besoin que la série soit stationnaire. Un moyen de vérifier cela est d'effectuer un test de Phillips-Perron, dont l'hypothèse nulle est la non-stationnarité stochastique de la série. Nous obtenons une *p-value* de 0,023, ce qui nous conduit à rejeter l'hypothèse nulle au niveau 5 % : nous pouvons considérer qu'elle est stationnaire.

7.1.2 Rendements annuels glissants

Les rendements annuels glissants sont définis pour $t \geq 4$, ils s'écrivent ainsi :

$$\tilde{R}_t = \frac{I_t - I_{t-4}}{I_{t-4}}$$

Ceux-ci ont l'allure suivante :

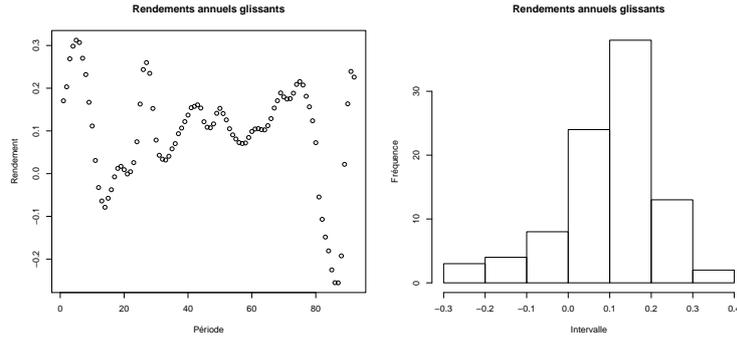


FIGURE 7.2 – Etude de la série des rendements annuels glissants

Nous voyons tout d’abord que cette série prend des valeurs moins extrêmes, à droite comme à gauche. Cela s’explique par le fait qu’il existe un phénomène de compensation des rendements (nous prenons en compte une année entière) qui n’existait pas dans la série précédente où chaque rendement trimestriel conduisait à une observation.

De plus, en nous intéressant au graphique représentant les valeurs de rendements dans le temps, nous remarquons une très forte auto-corrélation : des observations successives sont proches les unes des autres. En effet, deux rendements successifs ont en commun 9 mois : par exemple le rendement entre le 31 décembre 1994 et 31 décembre 1995 et celui entre le 31 mars 1995 et le 31 mars 1996 prennent tous deux en compte les variations de l’immobilier sur la période entre le 31 mars 1995 et le 31 décembre 1995. Ce n’est donc pas surprenant que d’un trimestre à l’autre la valeur du rendement annuel glissant soit proche.

Minimum	Maximum	Moyenne	Ecart-type	Coefficient d’asymétrie	Excès de Kurtosis
-25,56 %	31,23 %	9,46 %	11,98 %	-91,91 %	109,65 %

TABLE 7.2 – Caractéristiques de la distribution des rendements trimestriels annualisés

La série n’étant évidemment pas indépendante, nous souhaitons voir si elle est stationnaire. Un test de Phillips-Perron accepte la non-stationnarité. Le traitement des séries non-stationnaires par la théorie des valeurs extrêmes étant encore un point largement discuté aujourd’hui et ne disposant pas de méthodes aussi simples que pour les séries stationnaires, nous décidons d’étudier la série des rendements trimestriels annualisés. Une autre alternative aurait été d’étudier uniquement les rendements du 31 décembre de l’année N au 31 décembre de l’année $N + 1$, mais cela impliquerait de travailler sur 24 données, ce qui est trop peu.

7.2 Estimation du quantile à 0,5 % des rendements trimestriels annualisés

Dans l’étude précédente, nous recherchions un quantile à 99,5 %. Pour transposer facilement les méthodes, nous prenons dans la suite l’opposé des rendements. Nous allons commencer par étudier une pratique de marché d’estimation du choc en faisant l’hypothèse d’un modèle log-normal, avant de nous tourner vers des techniques utilisant la théorie des valeurs extrêmes.

7.2.1 Utilisation d’un modèle log-normal

Nous faisons l’hypothèse que l’indice IPD UK suit un brownien géométrique :

$$\frac{dI_t}{I_t} = \mu dt + \sigma dW_t$$

où W_t est un mouvement brownien. Nous retrouvons alors la forme habituelle pour I_t :

$$I_t = I_0 e^{\mu - \frac{\sigma^2}{2}t + \sigma W_t}$$

Nous souhaitons estimer les paramètres, pour cela nous pouvons remarquer la chose suivante :

$$\begin{aligned} \frac{I_t}{I_{t-1}} &= e^{\mu - \frac{\sigma^2}{2} + \sigma(W_t - W_{t-1})} \\ \Rightarrow \ln\left(\frac{I_t}{I_{t-1}}\right) &= \mu - \frac{\sigma^2}{2} + \sigma(W_t - W_{t-1}) \\ \Rightarrow \ln\left(\frac{I_t}{I_{t-1}}\right) &= \mu - \frac{\sigma^2}{2} + \sigma\epsilon \end{aligned}$$

où ϵ suit une loi normale centrée réduite.

Nous avons donc $\mathbb{E}\left(\ln\left(\frac{I_t}{I_{t-1}}\right)\right) = \mu - \frac{\sigma^2}{2}$ et $\mathbb{V}\left(\ln\left(\frac{I_t}{I_{t-1}}\right)\right) = \sigma^2$, ce qui nous permet d'estimer empiriquement les paramètres μ et σ en notant $X_t = \ln\left(\frac{I_t}{I_{t-1}}\right)$:

$$\begin{cases} \hat{\sigma} &= \frac{1}{T-1} \sum_{t=1}^T (X_t - \frac{1}{T} \sum_{t=1}^T X_t)^2 \\ \hat{\mu} &= \frac{1}{T} \sum_{t=1}^T X_t + \frac{\hat{\sigma}^2}{2} \end{cases}$$

Rappelons l'expression de R_t :

$$\begin{aligned} R_t &= \left(1 + \frac{I_t - I_{t-1}}{I_{t-1}}\right)^4 - 1 \\ \Rightarrow R_t &= \left(1 + \frac{I_t}{I_{t-1}} - 1\right)^4 - 1 \\ \Rightarrow R_t &= \left(\frac{I_t}{I_{t-1}}\right)^4 - 1 \\ \Rightarrow R_t &= e^{4[(\mu - \frac{\sigma^2}{2}) + \sigma\epsilon]} - 1 \end{aligned}$$

Nous pouvons ainsi calculer la valeur d'un choc à 0,5 % sur R_t en fonction de q_α le quantile d'ordre α d'une loi normale centrée réduite :

$$C = e^{4[(\mu - \frac{\sigma^2}{2}) + \sigma q_{0,5\%}]} - 1$$

Nous obtenons un choc à -23,5 % contre -25 % dans la formule standard. Cette modélisation a toutefois des limites. En effet, comparons la densité empirique des $\frac{I_t}{I_{t-1}}$ à une distribution log-normale de paramètres $\hat{\sigma}$ et $\hat{\mu}$:

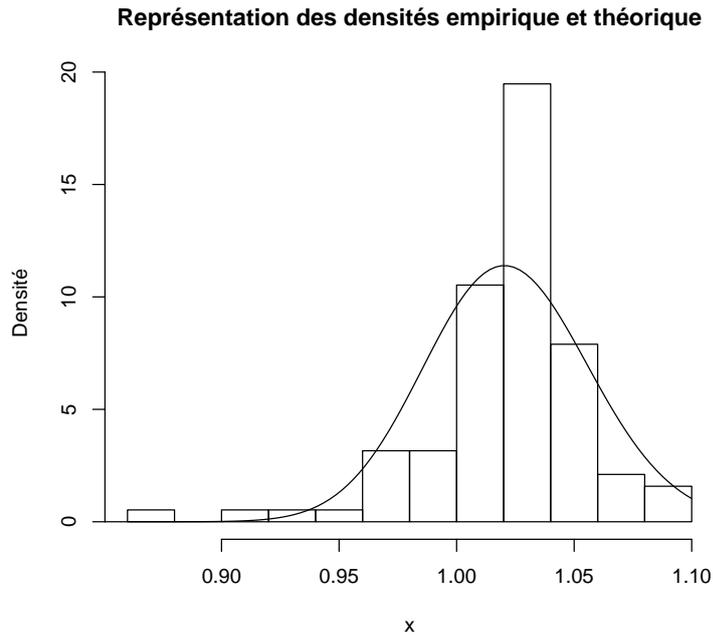


FIGURE 7.3 – Comparaison des densités théorique et empirique pour les $\frac{I_t}{I_{t-1}}$

Nous remarquons que les densités sont assez éloignées dans la partie centrale, et le sont davantage sur la queue de distribution à gauche qui est justement celle qui nous intéresse. Cela montre donc la faiblesse de l'hypothèse de log-normalité. Un test de Kolmogorov-Smirnov conduit à une *p-value* de 2,2 %, nous rejetons donc l'hypothèse de log-normalité au seuil 5 %.

7.2.2 Utilisation d'une loi Gumbel

En observant la série des rendements trimestriels annualisés, nous nous apercevons que leur densité est proche de celle d'une loi Gumbel dont les paramètres ont été estimés par méthode des moments. Pour rappel, la fonction de répartition d'une loi Gumbel est $F(x) = e^{-e^{-\frac{x-\mu}{\beta}}}$, son espérance vaut $\mu + \beta\gamma$ où γ est la constante d'Euler-Mascheroni, et sa variance vaut $\frac{\pi^2}{6}\beta^2$.

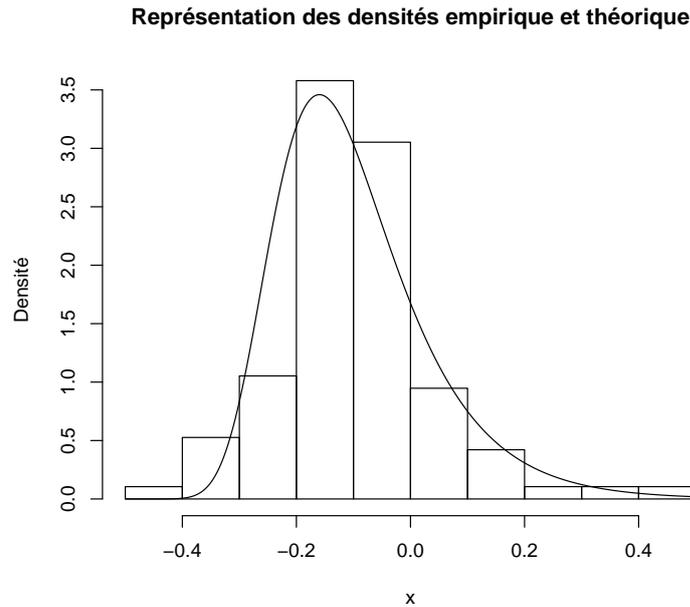


FIGURE 7.4 – Comparaison des densités théorique et empirique pour les rendements trimestriels annualisés

Un test d’adéquation de Kolmogorov-Smirnov conduit à une p -value de 22,0 %, nous acceptons donc l’hypothèse nulle. Afin de vérifier l’adéquation en queue de distribution, nous menons aussi un GPD Test : l’intervalle de confiance à 95 % obtenu pour δ est $[-0,18; 0,15]$, et nous obtenons $\delta = -0,08$. Nous acceptons donc également l’adéquation à une loi Gumbel en queue de distribution. Nous pouvons donc calculer le quantile paramétrique de la distribution des rendements :

$$q = \hat{\mu} - \hat{\beta} \ln(-\ln(p))$$

Comme nous travaillons sur l’opposé des rendements, nous prenons $p = 0,995$, ce qui nous renvoie $q = 40,4$ %. Le choc à 0,5 % est donc de -40,4 %. Nous remarquons qu’il est plus élevé (en valeur absolue) que celui obtenu avec la loi log-normale, ce qui est le résultat attendu puisque la queue de distribution n’était pas bien captée par la loi log-normale. Toutefois, ce résultat est à nuancer car l’estimation des paramètres suppose que les observations sont indépendantes, ce qui n’est pas le cas ici. Des méthodes tenant compte de cette non-indépendance sont donc préférables.

7.2.3 Utilisation de la loi des valeurs extrêmes

La méthode qui repose sur la loi log-normale décrite précédemment porte une attention limitée à la modélisation de la queue de distribution. Elle peut, par conséquent, donner un estimateur qui sous-estime la valeur du quantile.

Comparons maintenant le quantile estimé avec le modèle log-normal aux estimateurs obtenus avec les méthodes liées à la loi des valeurs extrêmes. La série des rendements trimestriels annualisés étant stationnaire, nous utilisons l’approche par blocs en supposant la condition $D(u_n)$ vérifiée. Toutes les méthodes s’appliquent ainsi sans traitement particulier excepté les méthodes Rao-Blackwell qui nécessitent des permutations spécifiques au cas stationnaire.

Disposant d’un échantillon regroupant 95 rendements trimestriels annualisés, il est logique en terme d’interprétation de choisir des blocs de taille proportionnelle à 4 (une taille de 4 cor-

respondant à des blocs annuels). De plus, la nécessité d'avoir à la fois un nombre de blocs et une taille des blocs importants nous a conduit à fixer des blocs de taille $m = 8$, ce qui correspond à des blocs de 2 ans.

Nous commençons par estimer le paramètre de queue ξ afin d'avoir un ordre de grandeur de l'épaisseur de la queue de distribution de l'opposé des rendements. Les méthodes de maximum de vraisemblance, des moments pondérés, des moments pondérés généralisés ainsi que leurs versions rao-blackwellisées donnent les estimateurs suivants :

	EMV	PWM	GPWM	RB EMV	RB PWM	RB GPWM
ξ	0.7526	0.5122	0.4939	0.6118	0.4027	0.3445

TABLE 7.3 – Estimateurs de ξ

Dans cet échantillon, le paramètre ξ semble être de l'ordre de 0,6, ce qui nous rapproche donc de notre étude sur la loi de Student de 2 degrés de liberté.

En estimant ensuite le quantile à 0,5% et en calculant les intervalles de confiance associés, nous obtenons :

	Quantile	Intervalle de confiance à 95 %		Intervalle de confiance à 90 %	
EMV (Fisher obs)	-50,8 %	-168,4 %	66,8 %	-149,5 %	47,9 %
EMV (PF)	-50,8 %	-53,8 %	-16,7 %	-53,8 %	-21,6 %
PWM (Boot)	-30,2 %	-58,7 %	-17,1 %	-56,6 %	-19,6 %
GPWM	-20,2 %	-319,5 %	279,1 %	-271,4 %	231,0 %
RB EMV (Boot)	-63,3 %	NA	NA	NA	NA
RB PWM (Boot)	-36,3 %	-57,3 %	-18,8 %	-55,6 %	-22,2 %
RB GPWM (Boot)	-28,2 %	-79,0 %	-16,4 %	-74,5 %	-18,4 %

TABLE 7.4 – Estimateurs du quantile à 0,5 % et leurs intervalles de confiance

Pour des petits échantillons ($n = 100$), nous avons vu que la méthode PWM et GPWM délivraient en général de meilleurs résultats en terme de biais et d'erreur quadratique. Or ici, au vu de la valeur des estimateurs de ξ , la méthode GPWM paraît plus appropriée que la méthode PWM, cette dernière n'étant pas valable (au sens de la normalité asymptotique) pour $\xi > 0,5$. Par conséquent, nous considérons que le meilleur estimateur du quantile est donné par la méthode des moments pondérés généralisés Rao-Blackwell : -28,2 %.

Comme pour l'étude sur les échantillons théoriques que nous avons menée dans le chapitre 4, la taille des intervalles de confiance est ici très importante, en particulier pour la méthode GPWM. Cependant, nous avons noté que même si les intervalles de confiance étaient peu fiables (les tailles étaient très importantes et les taux de couverture inférieurs au niveau souhaité), les estimateurs pouvaient rester satisfaisants, notamment pour la loi de Student à 2 degrés de liberté.

7.2.4 Utilisation de la loi des excès

Nous appliquons ici la méthode présentée précédemment pour l'utilisation de la loi des excès sur une série stationnaire. La première étape reste la même que celle pour les échantillon indépendants, il s'agit de trouver un seuil convenable. Voici les graphiques liés aux méthodes utilisées lors de l'étude théorique :

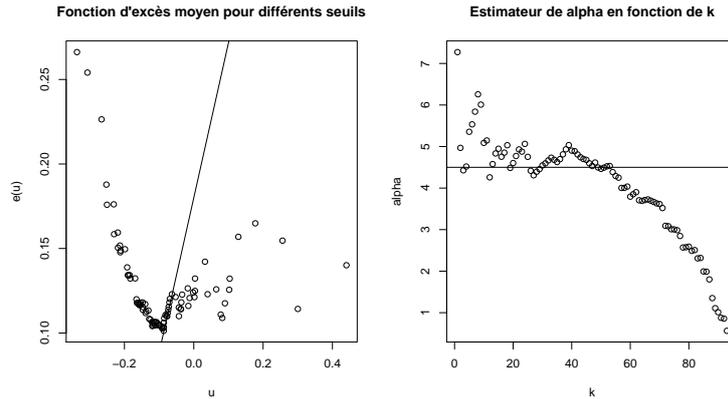


FIGURE 7.5 – Fonction d’excès moyens et graphique de Hill pour les rendements

A gauche, nous cherchons le point u tel que la fonction moyenne des excès soit une droite passé ce point. Ce n’est apparemment pas le cas ici. Toutefois, nous savons que lorsque u est trop élevé, la fonction moyenne des excès est mal estimée, il n’est donc pas pertinent de la regarder pour ces points. Nous remarquons une rupture de pente à partir du point $-0,0867$, qui conduit à une droite. Ce point est donc le seuil optimal pour la fonction moyenne des excès.

A droite, nous rappelons que le graphique de Hill estime la valeur de α en utilisant les valeurs de l’indice entre la plus grande et la $(n - k)^{eme}$ plus grande. Pour obtenir ce graphique, nous avons dû translater la série des rendements pour que toutes les valeurs soient positives, en effet le graphique de Hill nécessite de prendre le logarithme. L’inconvénient de cette modification est que la valeur de α que nous pouvons lire en ordonnées est erronée, toutefois cela nous permet toujours de rechercher le k le plus haut possible tel que pour tous les k plus petits, l’estimation de α reste stable. Ce principe a la même limitation que la fonction moyenne des excès, nous devons donc regarder les valeurs uniquement pour des k suffisamment éloignés de 0. Nous remarquons que la valeur de l’estimateur de α est croissante (en partant de k élevé, donc de la droite du graphique), puis devient décroissante pour $k = 40$ avant de redevenir croissante puis d’avoir un comportement plus erratique. Prendre $k = 40$ nous semble donc être un bon compromis, ce qui correspond à un seuil de $-0,0871$.

Nous remarquons que les deux seuils obtenus sont très proches l’un de l’autre, ce qui laisse supposer que la valeur obtenue est robuste. De plus, si nous appliquons la procédure de choix de seuil par méthode de Pickands, les deux méthodes d’estimation (maximum de vraisemblance et moments pondérés) fournissent un même seuil égal à $-0,108$.

Compte tenu de ces résultats, nous décidons donc de fixer comme seuil $u = -0,09$. Nous obtenons alors 6 *clusters* :

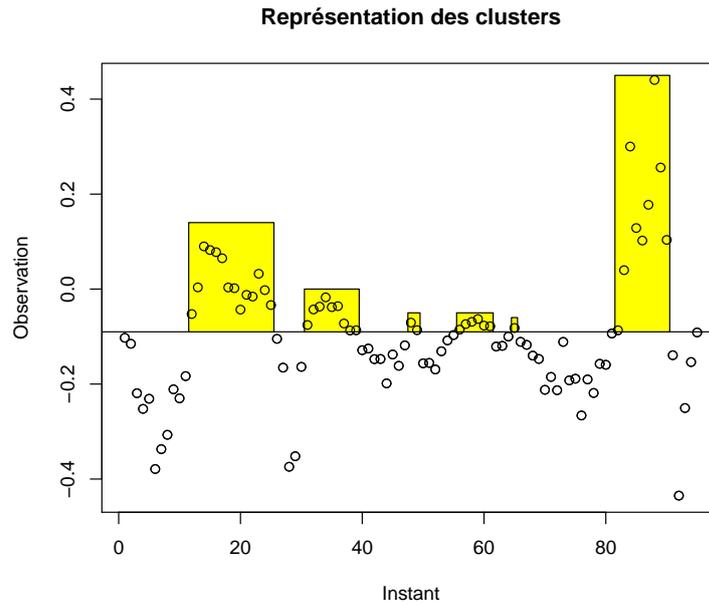


FIGURE 7.6 – Représentation de l’opposé des rendements, et identification des *clusters*

Nous estimons ensuite le quantile à 99,5 % selon la méthode adaptée aux séries stationnaires.

	Quantile	Intervalle de confiance à 95%		Intervalle de confiance à 90%	
EMV	-34,2 %	-114,6 %	46,1 %	-101,7 %	33,2 %
EMV obs	-34,2 %	-116,4 %	48,0 %	-103,2 %	34,8 %
PWM	-18,4 %	NA	NA	NA	NA

TABLE 7.5 – Estimateurs du quantile à 0,5 % et leurs intervalles de confiance

La valeur de $\hat{\xi}$ pour l’estimateur du maximum de vraisemblance est de 0,69, et elle est de 0,74 pour l’estimateur des moments pondérés. Cela explique que les intervalles de confiance ne soient pas définis pour cette méthode, puisque nous avons besoin de $\hat{\xi} < \frac{1}{2}$. Les intervalles de confiance obtenus sont plus larges que ceux donnés par l’approche par blocs car nous estimons ici sur moins de données.

7.2.5 Analyse des résultats

L’utilisation des méthodes basées sur la loi des valeurs extrêmes et la loi des excès nous permettent d’obtenir des estimateurs du quantile à 99,5 % des rendements trimestriels annu-
 lisés prenant en compte l’épaisseur des queues de distribution et la stationnarité de la série. Cependant, les résultats donnés sont très variables et nous obtenons des estimateurs allant de -18,4 % à -63,3 %. Cette variabilité peut s’expliquer par la taille de l’échantillon. En effet, pour la méthode reposant sur la loi des extrêmes, nous estimons les paramètres de la GEV à partir de 11 valeurs (après la mise en place des blocs) tandis que pour la GPD, seules 6 valeurs sont utilisées (suite au *de-clustering*). Cette faible quantité de données rend l’estimateur du maximum de vraisemblance peu pertinent. De plus, étant dans un cas où le paramètre de queue ξ est proche de 0,5, cela nous amène à considérer avec prudence les résultats fournis par la méthode PWM et à préférer ceux de la méthode GPWM. De plus, nous nous rappelons que les estimateurs rao-blackwellisés étaient plus précis que les estimateurs GPWM simples, c’est donc cette méthode que nous souhaitons utiliser.

Afin de confirmer ou d’infirmer notre démarche, nous traçons les histogrammes représentant d’une part les excès et d’autre part les maxima des blocs de la série des rendements trimestriels annualisés avec les densités des lois GPD et GEV :

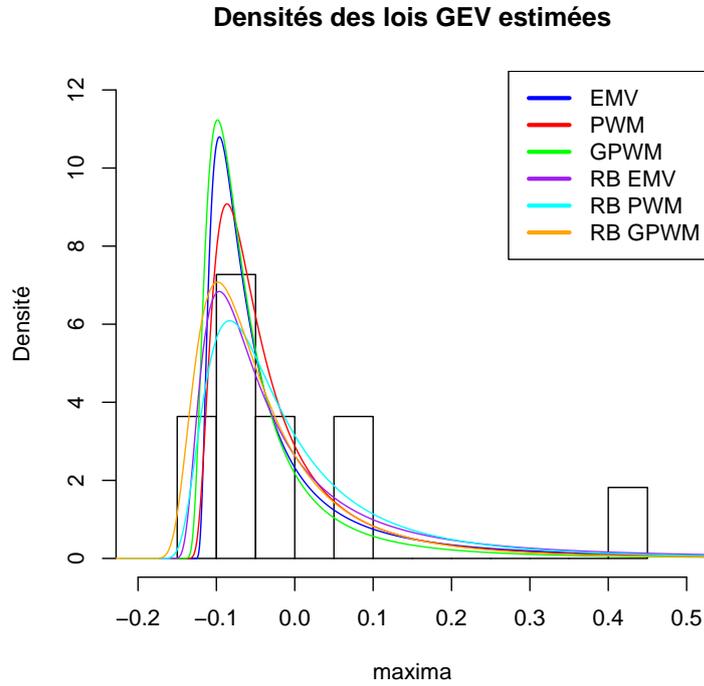


FIGURE 7.7 – Histogramme des maxima avec la densité des lois GEV estimées

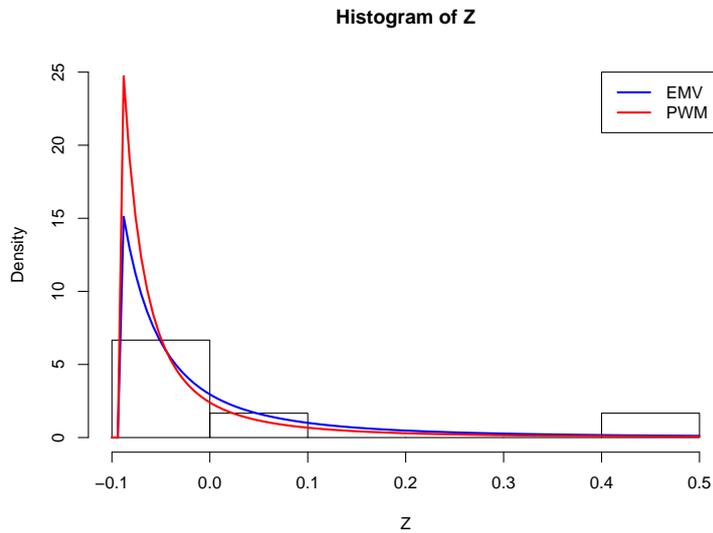


FIGURE 7.8 – Histogramme des excès avec la densité des lois GPD estimées

Nous observons qu’aucune des méthodes ne fournit une loi qui correspond parfaitement aux données, ce qui est normal compte tenu du très faible nombre de données. Au vu des graphiques nous comprenons les écarts entre les différentes méthodes : les densités sont visiblement différentes les unes des autres, il n’est donc pas surprenant que les quantiles estimés le soient aussi.

Ces graphiques ne nous apportant pas d'arguments supplémentaires, nous retenons donc le quantile fourni par l'estimateur GPWM rao-blackwellisé, qui est de $-28,2\%$. Toutefois, l'impact de ce résultat doit être nuancé car nous devons tenir compte de l'intervalle de confiance de l'estimateur RB GPWM qui est égal à $[-79,0\%; -16,4\%]$ pour un niveau de confiance à 95% et égal à $[-74,5\%; -18,4\%]$ pour un niveau de confiance à 90% .

Rappelons enfin plusieurs défauts de l'utilisation des rendements trimestriels annualisés. Tout d'abord, il n'y a pas de phénomène de compensation. En effet, les rendements trimestriels extrêmes sont supposés se répéter à l'identique au cours de tous les trimestres, ce qui conduit à une surestimation de la valeur du choc. De plus, dans ce chapitre, nous avons calculé un quantile de la distribution stationnaire des rendements trimestriels annualisés. Cependant, dans le cadre de Solvabilité II, ce qui nous intéresse est l'estimation d'un quantile d'horizon un an conditionnellement à la situation actuelle. Par la suite, nous allons essayer de nous affranchir de ces inconvénients en utilisant une modélisation dynamique des log-rendements trimestriels (dont la valeur est très proche des rendements trimestriels, nous utilisons cette quantité par soucis de cohérence avec le traitement des autres séries financières) qui nous permettra d'aboutir à des facteurs de risques indépendants et identiquement distribués. Ainsi, nous pourrions approcher de manière plus raisonnable la valeur du choc immobilier à horizon un an. Pour cela, nous aurons besoin d'une loi pour modéliser les facteurs de risques : il s'agira de la loi de Pareto Hybride.

Chapitre 8

Estimation des lois marginales des facteurs de risque

Par définition, une loi de Pareto Hybride est une loi dont une des queues - ou les deux - est (sont) modélisée(s) par des lois de Pareto généralisées, et où la partie centrale est modélisée par une autre loi : typiquement, une loi normale. L'utilisation de cette loi est motivée par plusieurs raisons. Tout d'abord, elle nous permet de poursuivre le travail que nous avons déjà effectué : la queue de distribution est modélisée par une loi que nous avons déjà étudiée. De plus, elle comporte un grand nombre de paramètres, ce qui donne davantage de degrés de liberté pour approcher la distribution empirique de l'échantillon.

Nous choisissons d'utiliser la loi normale en partie centrale car il s'agit d'une loi usuelle pour modéliser les séries financières.

8.1 Loi de Pareto Hybride

8.1.1 Densité de la loi

Une des problématiques principales pour ces lois est d'obtenir une régularité au point où l'on change de loi. En effet, une discontinuité de la fonction de densité de la loi de Pareto Hybride serait difficilement explicable en pratique. Un certain nombre d'articles ont traité du sujet, nous avons choisi d'utiliser l'article [HH09] écrit par HOLDEN et LAUG.

Supposons que nous souhaitons utiliser une loi de Pareto généralisée pour toutes les valeurs au-dessus d'un seuil u , et une gaussienne pour celles en-dessous. Leur méthodologie consiste à prendre un intervalle $[u - \epsilon, u + \epsilon]$ au sein duquel la densité est un mélange des deux densités. Concrètement, en appelant f la fonction de densité de la gaussienne, g la fonction de densité de la loi de Pareto généralisée, et l la fonction de densité de la loi de Pareto Hybride, nous avons l'expression suivante :

$$l(x) = \begin{cases} \kappa f(x) & \text{si } x < u - \epsilon \\ \kappa(g(q(x))q'(x) + f(p(x))p'(x)) & \text{si } x \in [u - \epsilon, u + \epsilon[\\ \kappa g(x) & \text{si } x \geq u + \epsilon \end{cases}$$

où p et q sont des fonctions de pondération caractérisant le mélange dans la zone où les deux densités s'appliquent, et où κ sert à ce que l'intégrale de l soit bien égale à 1. Par intégration nous obtenons la fonction de répartition :

$$L(x) = \begin{cases} \kappa F(x) & \text{si } x < u - \epsilon \\ \kappa(G(q(x)) + F(p(x))) & \text{si } x \in [u - \epsilon, u + \epsilon[\\ \kappa G(x) + 1 - \kappa & \text{si } x \geq u + \epsilon \end{cases}$$

Nous obtenons aisément que $\kappa = \frac{1}{1+F(u)}$ où F et G sont respectivement les fonctions de répartition de la gaussienne et de la Pareto généralisée.

La régularité de l impose certaines conditions sur p et q . En effet, la continuité en $u - \epsilon$ impose que $q'(u - \epsilon) = 0$, $p'(u - \epsilon) = 1$ et $p(u - \epsilon) = u - \epsilon$. De même, la continuité en $u + \epsilon$ impose que $q'(u + \epsilon^-) = 1$, $q(u + \epsilon^-) = u + \epsilon$ et $p'(u + \epsilon^-) = 0$. Les auteurs montrent qu'il est possible d'obtenir aussi la continuité de la dérivée, en utilisant les fonctions suivantes :

$$p(x) = \begin{cases} x & \text{si } x < u - \epsilon \\ \frac{1}{2}(x + u - \epsilon) + \frac{\epsilon}{\pi} \cos\left(\frac{\pi(x-u)}{2\epsilon}\right) & \text{si } x \in [u - \epsilon, u + \epsilon[\\ x - \epsilon & \text{si } x \geq u + \epsilon \end{cases}$$

$$q(x) = \begin{cases} x + \epsilon & \text{si } x < u - \epsilon \\ \frac{1}{2}(x + u + \epsilon) - \frac{\epsilon}{\pi} \cos\left(\frac{\pi(x-u)}{2\epsilon}\right) & \text{si } x \in [u - \epsilon, u + \epsilon[\\ x & \text{si } x \geq u + \epsilon \end{cases}$$

Nous pouvons représenter ces fonctions et leurs dérivées première et seconde :

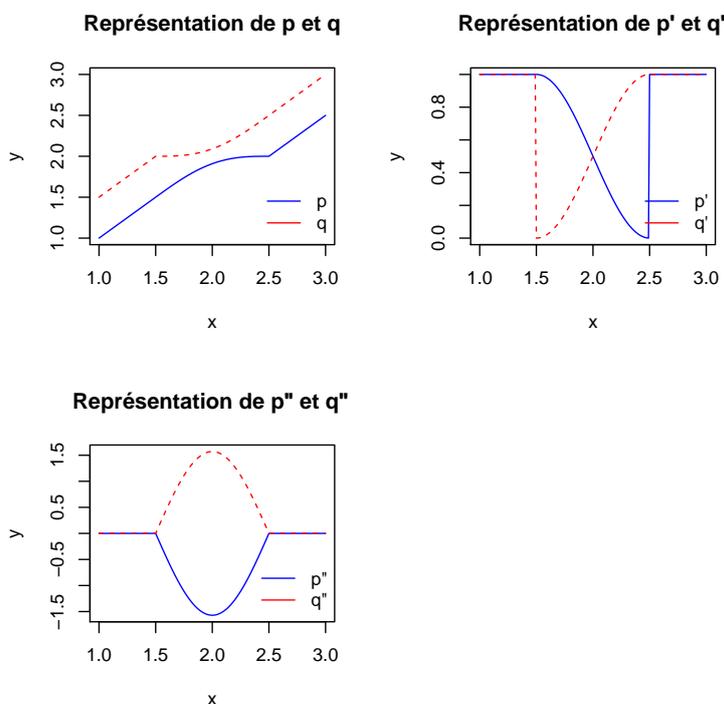


FIGURE 8.1 – Représentation des fonctions de pondération p et q , ainsi que leurs dérivées première et seconde

Le graphique suivant représente les fonctions de densité et de répartition d'une Pareto Hybride avec les caractéristiques suivantes :

- une gaussienne en partie centrale, de moyenne 0 et d'écart-type 0,5
- une GPD en queue de distribution à droite, de paramètre de queue $\xi = 0,8$, avec $\mu = 0,5$ et $\sigma = 1$
- $\epsilon = 1,5$

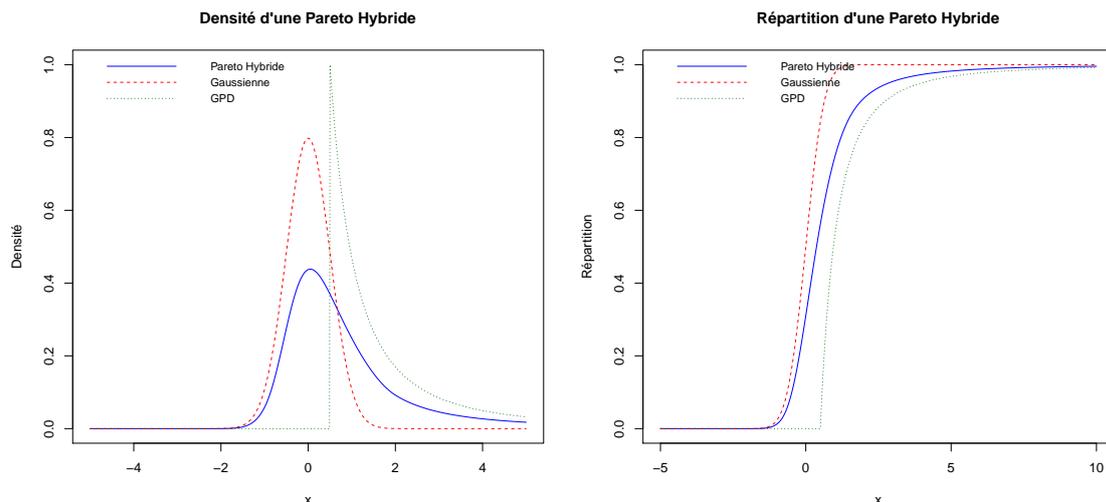


FIGURE 8.2 – Comparaison de la densité d'une Pareto Hybride avec celles d'une loi Gaussienne et d'une GPD

Nous voyons que la densité de la Pareto Hybride a les propriétés souhaitées : elle est proche de celle d'une gaussienne au centre de la distribution, mais elle a le comportement de queue d'une GPD. Elle est de plus régulière, sans point de discontinuité ni rupture de pente. Notons toutefois qu'ici ϵ est assez grand, pour une valeur plus faible de ϵ la décroissance de la fonction de densité est accentuée sur une portion de la courbe avant de baisser à nouveau. Par exemple, pour $\epsilon = 0,8$:

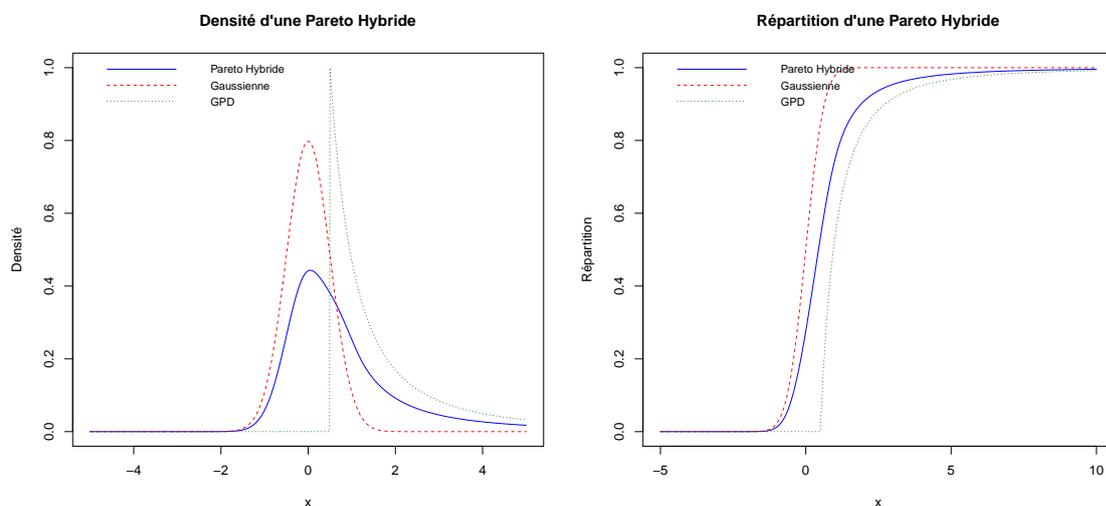


FIGURE 8.3 – Comparaison de la densité d'une Pareto Hybride avec celles d'une loi Gaussienne et d'une GPD

A la limite, nous obtenons une distribution bi-modale, par exemple avec $\epsilon = 0,2$:

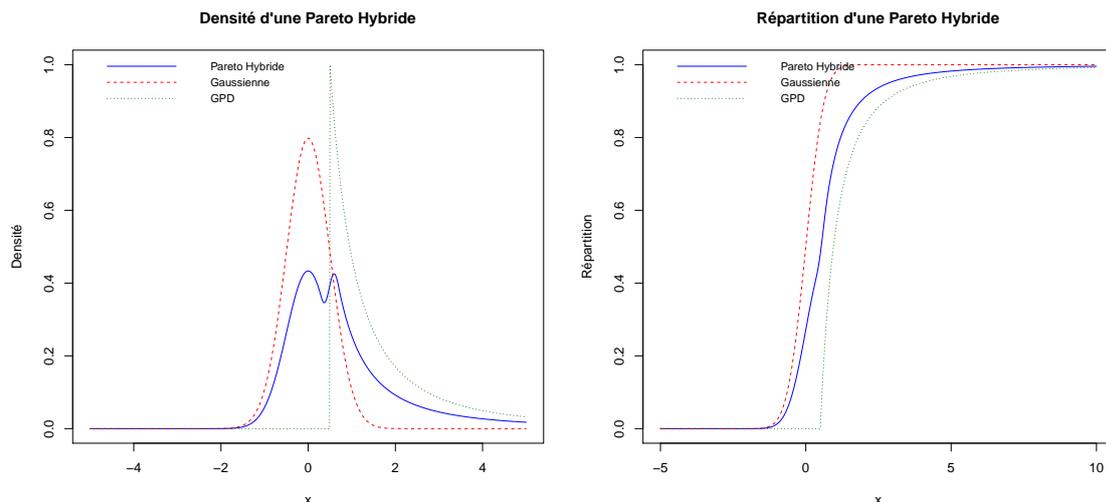


FIGURE 8.4 – Comparaison de la densité d’une Pareto Hybride avec celles d’une loi Gaussienne et d’une GPD

Le choix de ϵ est donc important si nous voulons que la densité soit la plus régulière possible.

8.1.2 Estimation des paramètres

Dans le cadre d’un modèle interne, nous sommes intéressés par une majeure partie de la distribution à un an - et pas seulement par la valeur d’un quantile extrême. Nous devons notamment tâcher d’être le plus précis possible dans la queue de distribution, car la valeur du SCR sera davantage influencée par les queues de distribution des facteurs de risque. Pour cette raison, nous décidons d’appliquer une estimation des paramètres en deux temps :

- estimer tout d’abord le seuil u puis les paramètres de queue ξ et σ par les méthodes habituelles
- puis estimer les paramètres de la gaussienne et ϵ conjointement par maximum de vraisemblance.

Ainsi, la queue de distribution est tout d’abord fixée sans tenir compte de la partie centrale, et celle-ci est ensuite ajustée sous contrainte de régularité de la distribution.

8.1.3 Simulation d’un échantillon suivant une loi de Pareto Hybride

Dans le cadre d’un modèle interne, nous avons besoin de projeter nos facteurs de risque, et donc de simuler des variables suivant des lois de Pareto Hybride. Pour cela, nous pouvons utiliser la fonction de répartition qui est entièrement déterminée. Il est donc aisé de l’inverser numériquement et de simuler selon cette loi par méthode d’inversion générique. Rappelons qu’elle consiste à simuler U selon une loi uniforme et à poser $X = L^{-1}(U)$, X suit alors une loi de Pareto Hybride. L^{-1} est classiquement évaluée comme suit :

$$L^{-1}(y) = \inf\{x | L(x) \geq y\}$$

8.1.4 Ajout d’une troisième densité

Pour certains risques, par exemple le risque de taux, nous souhaitons appréhender correctement les variations extrêmes à la hausse comme à la baisse. Nous ne pouvons donc pas nous contenter d’une approximation normale pour une des deux queues de distribution, il nous faut utiliser une loi de Pareto généralisée pour modéliser chaque queue de distribution. Nous rajoutons donc 4 paramètres à notre densité Pareto Hybride : les 3 paramètres ξ_2 , μ_2 (qui est aussi le

seuil u_2), et σ_2 de la deuxième loi de Pareto généralisée, et le paramètre de mélange ϵ_2 . Il faut aussi trois nouvelles fonctions de pondération définies ainsi (pour $u_1 < u_2$) :

$$\begin{aligned}
 q_1(x) &= \begin{cases} x & \text{si } x < u_1 - \epsilon_1 \\ \frac{1}{2}(x + u_1 - \epsilon_1) + \frac{\epsilon_1}{\pi} \cos\left(\frac{\pi(x-u_1)}{2\epsilon_1}\right) & \text{si } x \in [u_1 - \epsilon_1, u_1 + \epsilon_1[\\ x - \epsilon_1 & \text{si } x \geq u_1 + \epsilon_1 \end{cases} \\
 q_2(x) &= \begin{cases} x + \epsilon_1 & \text{si } x < u_1 - \epsilon_1 \\ \frac{1}{2}(x + u_1 + \epsilon_1) - \frac{\epsilon_1}{\pi} \cos\left(\frac{\pi(x-u_1)}{2\epsilon_1}\right) & \text{si } x \in [u_1 - \epsilon_1, u_1 + \epsilon_1[\\ x & \text{si } x \in [u_1 + \epsilon_1, u_2 - \epsilon_2[\\ \frac{1}{2}(x + u_2 - \epsilon_2) + \frac{\epsilon_2}{\pi} \cos\left(\frac{\pi(x-u_2)}{2\epsilon_2}\right) & \text{si } x \in [u_2 - \epsilon_2, u_2 + \epsilon_2[\\ x - \epsilon_2 & \text{si } x \geq u_2 + \epsilon_2 \end{cases} \\
 q_3(x) &= \begin{cases} x + \epsilon_2 & \text{si } x < u_2 - \epsilon_2 \\ \frac{1}{2}(x + u_2 + \epsilon_2) - \frac{\epsilon_2}{\pi} \cos\left(\frac{\pi(x-u_2)}{2\epsilon_2}\right) & \text{si } x \in [u_2 - \epsilon_2, u_2 + \epsilon_2[\\ x & \text{si } x \geq u_2 + \epsilon_2 \end{cases}
 \end{aligned}$$

La fonction q_2 est définie pour $u_1 + \epsilon_1 \leq u_2 - \epsilon_2$, ce qui en pratique n'est pas toujours vrai. En effet nous avons vu au-dessus ce qu'il pouvait se produire lorsque les paramètres de mélange étaient trop faibles, nous aimerions donc pouvoir définir q_2 dans le cas où $u_1 + \epsilon_1 > u_2 - \epsilon_2$. Celui-ci n'est pas évoqué dans [HH09], nous décidons donc d'adopter une solution simple qui est de définir la fonction suivante :

$$q_2^{bis}(x) = \begin{cases} x + \epsilon_1 & \text{si } x < u_1 - \epsilon_1 \\ \frac{1}{2}(x + u_1 + \epsilon_1) - \frac{\epsilon_1}{\pi} \cos\left(\frac{\pi(x-u_1)}{2\epsilon_1}\right) & \text{si } x \in [u_1 - \epsilon_1, u_2 - \epsilon_2[\\ ax + b & \text{si } x \in [u_2 - \epsilon_2, u_1 + \epsilon_1[\\ \frac{1}{2}(x + u_2 - \epsilon_2) + \frac{\epsilon_2}{\pi} \cos\left(\frac{\pi(x-u_2)}{2\epsilon_2}\right) & \text{si } x \in [u_1 + \epsilon_1, u_2 + \epsilon_2[\\ x - \epsilon_2 & \text{si } x \geq u_2 + \epsilon_2 \end{cases}$$

où a et b sont définis pour que la fonction soit continue.

Prenons $u_1 = -1$, $u_2 = 1$, et comparons les fonctions q_2 pour $\epsilon_1 = \epsilon_2 = \frac{1}{2}$ puis q_2^{bis} pour $\epsilon_1 = \epsilon_2 = 2$, ainsi que leurs dérivées :

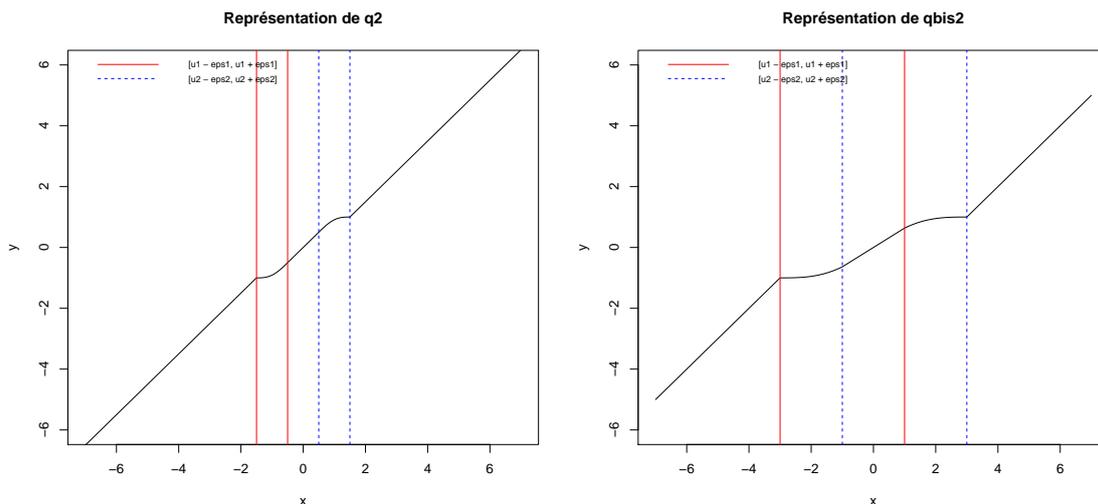


FIGURE 8.5 – Comparaison des fonctions q_2 et q_2^{bis} , avec en rouge les bornes de l'intervalle $[u_1 - \epsilon_1, u_1 + \epsilon_1]$ et en bleu celles de $[u_2 - \epsilon_2, u_2 + \epsilon_2]$

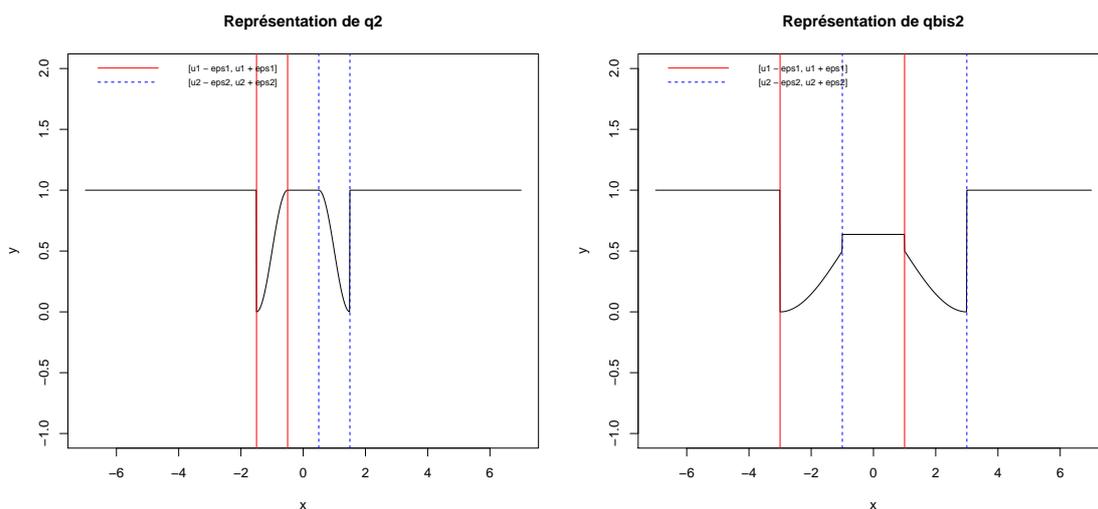


FIGURE 8.6 – Comparaison des dérivées de q_2 et q_2^{bis} , avec en rouge les bornes de l'intervalle $[u_1 - \epsilon_1, u_1 + \epsilon_1]$ et en bleu celles de $[u_2 - \epsilon_2, u_2 + \epsilon_2]$

Nous voyons que nous avons perdu la continuité de la dérivée dans le cas où les zones de mélange se chevauchent, ce qui entraîne la discontinuité de la densité de la loi de Pareto Hybride. Nous verrons plus tard l'influence que cela aura sur nos résultats. Nous pouvons aisément simuler un échantillon suivant une loi de Pareto Hybride à 3 densités, en utilisant là encore la méthode d'inversion générique, L étant connue analytiquement.

8.2 Modélisation du facteur de risque immobilier

8.2.1 Modélisation des log-rendements par un modèle $AR - GARCH$

Rappelons que les log-rendements sont définis de la façon suivante :

$$LR_t = \ln \left(\frac{I_t}{I_{t-1}} \right)$$

Voici leur évolution en fonction de t :

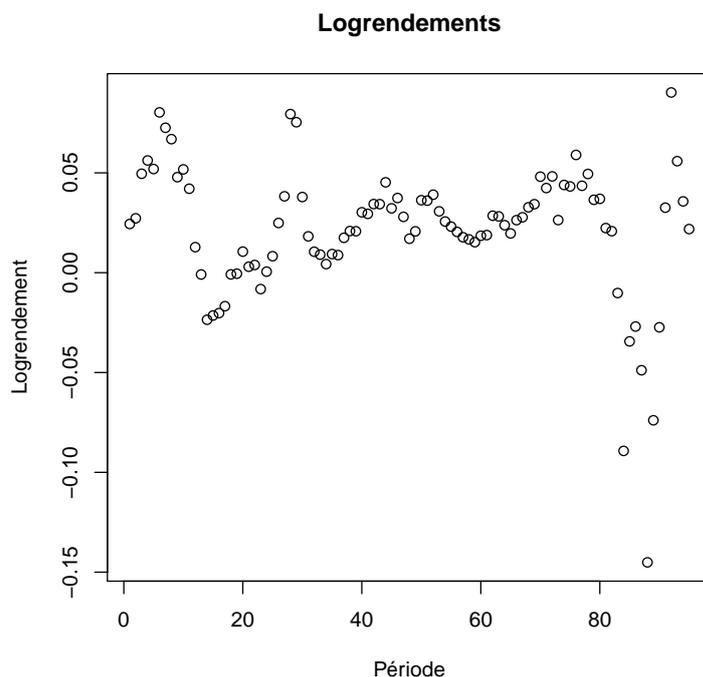


FIGURE 8.7 – Représentation de la série des log-rendements de l'immobilier

Il nous faut donc modéliser cette série. Un test de Phillips-Perron nous indique que cette série est stationnaire, nous pouvons donc choisir de la modéliser par un processus de type $ARMA(P, Q) - GARCH(p, q)$ qui s'écrit :

$$\begin{cases} X_t = \mu + \sum_{i=1}^P a_i X_{t-i} + \sum_{j=1}^Q b_j \epsilon_{t-j} + \epsilon_t \\ \epsilon_t = \eta_t \sqrt{h_t} \\ h_t = \omega + \sum_{i=1}^q \alpha_i \epsilon_{t-i}^2 + \sum_{j=1}^p \beta_j h_{t-j} \end{cases}$$

en notant X la série des log-rendements.

En effet, d'après la représentation des log-rendements dans la figure 8.7, la volatilité semble présenter une dépendance temporelle : à des périodes calmes présentant une faible volatilité se succèdent des périodes plus agitées avec des hausses de volatilité et des regroupements de volatilité. L'intérêt de cette modélisation par rapport à l'utilisation d'un simple processus auto-régressif est que nous tenons compte de ces fluctuations de la volatilité. Le fait de modéliser h_t (qui est la variance de ϵ_t dans le processus auto-régressif) par un modèle $GARCH$ permet de ne pas avoir des ϵ_t extrêmes au cours de certaines périodes et faibles le reste du temps. Ici, η peut donc être considéré comme un bruit blanc.

Nous cherchons les ordres (P, Q, p, q) du modèle tels que celui-ci soit le simple possible avec des résidus indépendants et des coefficients significatifs. De plus, nous nous limitons aux ordres p et q inférieurs à 2 afin de limiter le nombre de paramètres à estimer. Par ailleurs, des ordres supérieurs apportent généralement peu d'informations supplémentaires et il est donc courant de se limiter à ces ordres (voir [TSA05]). Nous choisissons finalement le modèle $AR(1) - GARCH(1, 1)$. Ce type de processus a l'avantage d'être relativement simple ce qui nous convient parfaitement car nous ne cherchons pas à obtenir la modélisation la plus exacte possible, mais une modélisation qui soit à la fois simple et cohérente avec nos hypothèses d'utilisation de la théorie des valeurs extrêmes qui est ce sur quoi nous nous focalisons. De plus, cela reste une modélisation

assez courante en finance.

X peut donc s'écrire sous la forme suivante :

$$\begin{cases} X_t = \mu + \rho X_{t-1} + \epsilon_t \\ \epsilon_t = \eta_t \sqrt{h_t} \\ h_t = \omega + \alpha \epsilon_{t-1}^2 + \beta h_{t-1} \end{cases}$$

Les paramètres du modèle peuvent être estimés par maximum de vraisemblance. Cette estimation repose sur l'hypothèse faite sur la loi des innovations η_t . On suppose en général que sachant les informations disponibles en $t - 1$, η_t suit une loi normale centrée réduite. Or, pour les séries financière, cette propriété de normalité est souvent remise en question ce qui devrait invalider les paramètres estimés. Cependant, GOURIÉROUX, MONFORT et TROGNON montrent dans [GMA84] que les estimateurs obtenus par pseudo-maximum de vraisemblance sont convergents et asymptotiquement normaux (sous des conditions de régularité). Les estimateurs restent donc corrects même si l'hypothèse de normalité est violée.

Le tableau suivant résume la valeur des paramètres ainsi que les p -values correspondant aux tests de significativité des coefficients :

$\hat{\mu}$	$3,420 \times 10^{-3}$	0,0223
$\hat{\rho}$	$8,760 \times 10^{-1}$	$< 2,2 \times 10^{-16}$
$\hat{\omega}$	$1,071 \times 10^{-5}$	0,2080
$\hat{\alpha}$	$6,022 \times 10^{-1}$	0,0359
$\hat{\beta}$	$5,662 \times 10^{-1}$	$5,84 \times 10^{-3}$

TABLE 8.1 – Valeurs des paramètres du modèle AR(1)-GARCH(1,1)

Nous remarquons que la série des log-rendements est fortement auto-corrélée, en effet ρ est égal à 0,876. De plus, tous les paramètres du modèle semblent significatifs mis à part $\hat{\omega}$ pour lequel le test de nullité est accepté. Celui-ci n'est toutefois pas déterminant du modèle puisqu'il s'agit du terme constant dans la volatilité. Nous pouvons vérifier la pertinence du modèle en effectuant des tests d'indépendance sur η , nous obtenons les p -values suivantes :

WALD-WOLFOWITZ	LJUNG-BOX ($m = 1$ an)	LJUNG-BOX ($m = 2$ ans)
0,080	0,077	0,210

TABLE 8.2 – Tests d'indépendance des résidus du modèle AR(1) – GARCH(1,1)

Nous ne pouvons donc pas rejeter l'indépendance des résidus au seuil 5 %, c'est cela qui va nous permettre de projeter la valeur du log-rendement immobilier dans un an comme nous le verrons par la suite.

Afin de vérifier que notre modèle reflète notre série de manière satisfaisante, nous représentons les valeurs estimées des log-rendements et les comparons à notre série initiale :

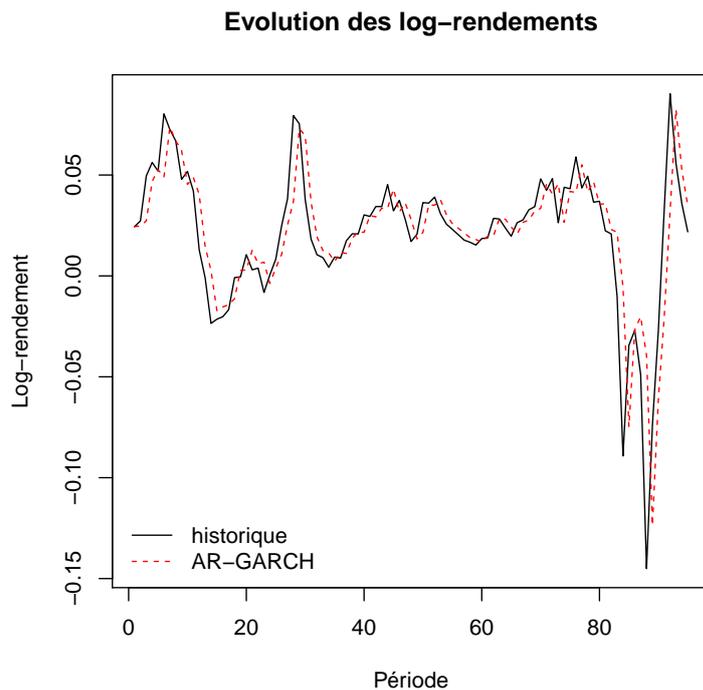


FIGURE 8.8 – Représentation de la série des log-rendements de l’immobilier et de la série des valeurs estimées par le modèle $AR(1) - GARCH(1, 1)$

Ce graphique est typique d’un modèle auto-regressif d’ordre 1 : la meilleure prédiction de la série à un instant donné est la valeur précédente de la série multipliée par un certain coefficient qui ici est proche de 1 (0,88), à laquelle s’ajoute une constante. Le fait que la volatilité soit stochastique n’est pas visible car elle s’applique aux résidus qui sont tous pris nuls. L’évolution de la volatilité conditionnellement aux observations passées est visible ci-dessous :

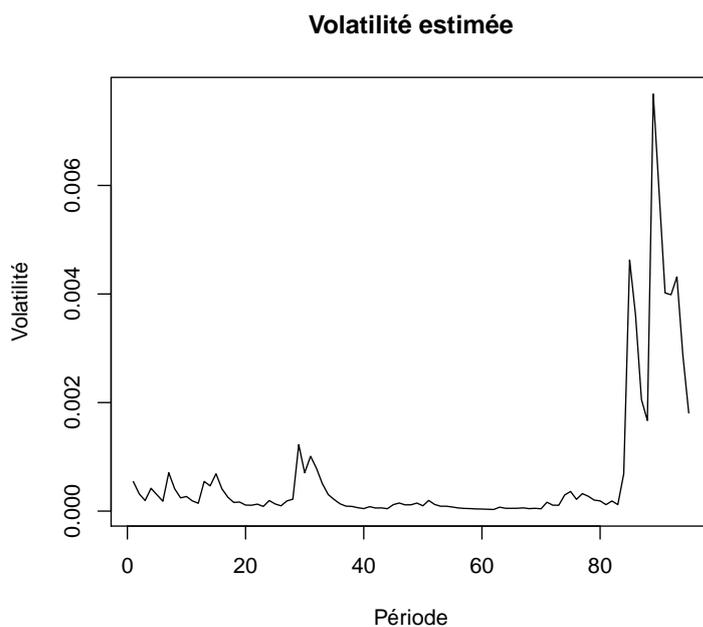


FIGURE 8.9 – Volatilité estimée par le modèle $AR(1) - GARCH(1, 1)$

Nous voyons que comme attendu, elle est forte en fin de période.

8.2.2 Obtention de la distribution à un an de l'immobilier

Intéressons-nous à la distribution des $(\eta_t)_t$:

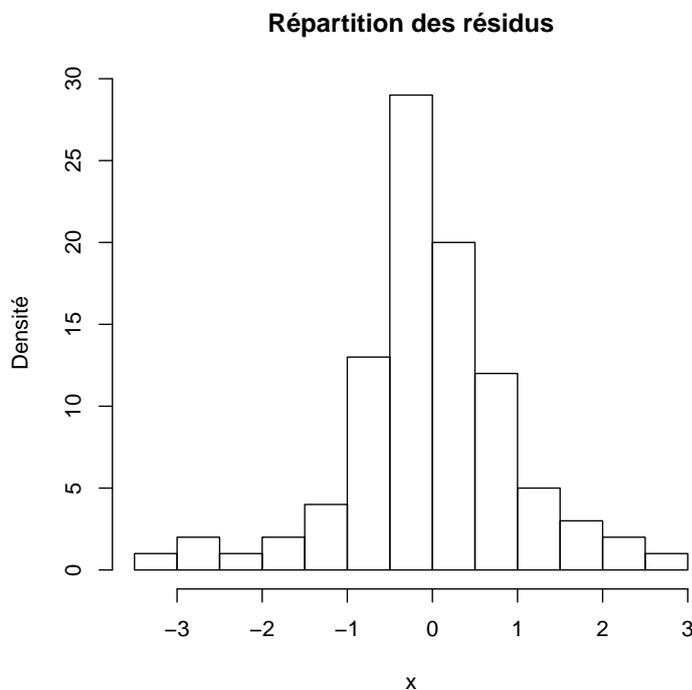


FIGURE 8.10 – Répartition des résidus $(\eta_t)_t$

Nous voyons que ceux-ci sont compris entre -3,2 et 2,9, avec des queues un peu plus lourdes sur la gauche de la distribution. Cela correspond à des valeurs négatives des log-rendements et donc à une baisse de l'immobilier. C'est ce phénomène qui nous intéresse particulièrement. Nous devons surtout tâcher de choisir une modélisation des résidus qui en tienne compte, ce qu'une loi gaussienne ne peut pas faire. C'est pour cette raison que nous nous tournons vers une loi de Pareto Hybride.

Puisque nous avons jusqu'ici appliqué nos théories à la queue droite des distributions, nous regardons $-\eta$. Nous pouvons alors représenter les fonctions moyenne des excès et de Hill :

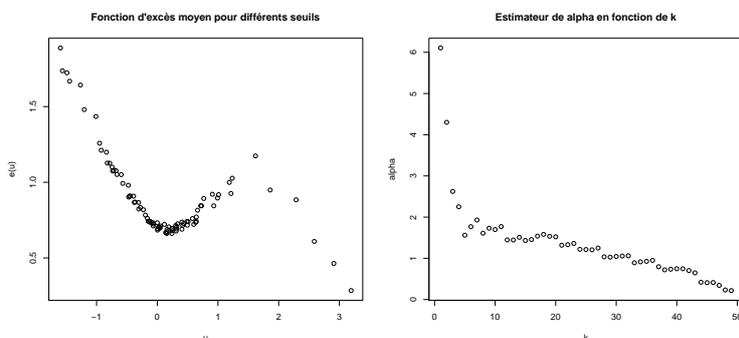


FIGURE 8.11 – Fonction d'excès moyen et graphique de Hill pour $-\eta$

Le graphique de Hill n'est pas très explicite, en revanche sur la fonction moyenne des excès, nous constatons un changement de pente peu après 0. La méthode de Pickands propose un seuil à 0,15 qui semble cohérent avec notre observation. Nous optons ainsi pour ce seuil et nous estimons alors les paramètres suivants :

$\hat{\xi}_{EMV}$	0,185	$\hat{\xi}_{PWM}$	0,261
$\hat{\sigma}_{EMV}$	0,535	$\hat{\sigma}_{PWM}$	0,481

TABLE 8.3 – Paramètres de queue de $-\eta$

L'étude théorique conduite précédemment nous avait montré que dans le cas d'une loi de Student à 2 degrés de liberté (dont la queue se rapproche davantage de celle observée ici), les estimateurs des moments pondérés étaient les plus efficaces. Nous avons de plus remarqué que pour de petits échantillons, cette méthode était la plus stable. Ce sont donc les estimateurs obtenus par cette méthode que nous utilisons par la suite pour modéliser la partie GPD de la loi de Pareto Hybride, et nous calibrons les trois paramètres restants par la méthode exposée précédemment. Voici ce que nous obtenons :

$\hat{\mu}_{norm}$	-0,144
$\hat{\sigma}_{norm}$	1,06
$\hat{\epsilon}_{mix}$	1,16

TABLE 8.4 – Paramètres de la partie centrale de $-\eta$

Nous pouvons confirmer la valeur de ces paramètres en comparant les densités empirique et théorique. Par soucis de comparaison, nous regardons aussi la densité d'une loi normale de moyenne et variance égales à celles empiriques.

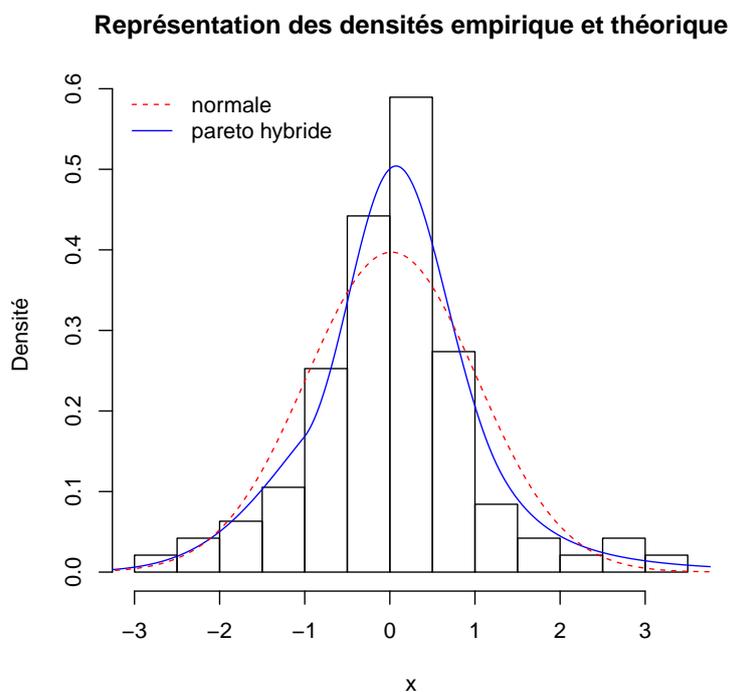


FIGURE 8.12 – Densités théoriques et empirique de $-\eta$

Nous voyons bien tout l'intérêt de la loi Pareto Hybride qui capte mieux la queue de distribution à droite de $-\eta$. De plus, elle semble aussi plus proche en partie centrale. Pour valider la représentation par une loi de Pareto Hybride, effectuons un test de Kolmogorov-Smirnov face à un échantillon de taille 1000 simulé suivant cette loi. Nous obtenons une *p-value* de 0,67, nous acceptons donc l'hypothèse d'adéquation des lois. En comparant à une loi normale nous obtenons là encore une *p-value* supérieure à 5 % (0,46), mais un GPD Test refuse l'égalité des lois : δ vaut $-1,49$ alors que l'intervalle bootstrap à 95 % est $[-0,98; 0,50]$. Bien sûr le même test par rapport à une loi de Pareto Hybride est accepté, par définition même du test.

Toutefois, ce qui nous intéresse est le choc à un an, or pour l'instant, nous avons traité des log-rendements trimestriels. Il nous faut donc commencer par exprimer la valeur du log-rendement dans un an en fonction des données d'aujourd'hui. Notons T le dernier instant pour lequel nous avons des observations, nous pouvons alors calculer les valeurs suivantes :

$$\begin{cases} h_{T+1} = \hat{\omega} + \hat{\alpha}\epsilon_T^2 + \hat{\beta}h_T \\ \epsilon_{T+1} = \eta\sqrt{h_{T+1}} \\ X_{T+1} = \hat{\mu} + \hat{\rho}X_T + \epsilon_{T+1} \end{cases}$$

Ces formules nous permettent de passer de (h_T, ϵ_T, X_T) à $(h_{T+1}, \epsilon_{T+1}, X_{T+1})$, en utilisant un tirage aléatoire de la variable η . Celle-ci est simulée à chaque itération selon la densité théorique obtenue ci-dessus. Nous pouvons faire cette opération grâce à l'indépendance mise en lumière plus tôt. Après 4 itérations, nous obtenons donc une valeur possible de X_{T+4} .

Afin d'obtenir une estimation fiable du quantile, nous appliquons cet algorithme 50 000 fois, ce qui nous conduit à avoir à notre disposition 50 000 valeurs possibles de X_{T+4} . La distribution du rendement à un an de l'immobilier est obtenue en considérant $e^{X_{T+4}} - 1$. Nous avons donc atteint l'objectif souhaité qui était de pouvoir projeter la valeur de l'indice immobilier à un an : la théorie des valeurs extrêmes est applicable dans le cadre d'un modèle interne, et apporte un degré de précision supplémentaire dans les queues de distribution de nos valeurs projetées. Nous pouvons obtenir le choc bicentenaire sur l'immobilier qui est défini comme le quantile à 0,5 % de cette distribution : il suffit de prendre la 250^{ème} pire valeur. Le résultat est -15,5 %. Le graphique suivant présente trois scénarios différents : un scénario central avec des résidus égaux à la moyenne et les deux scénarios ayant conduit aux quantiles à 0,5 % et 99,5 % de X_{T+4} .

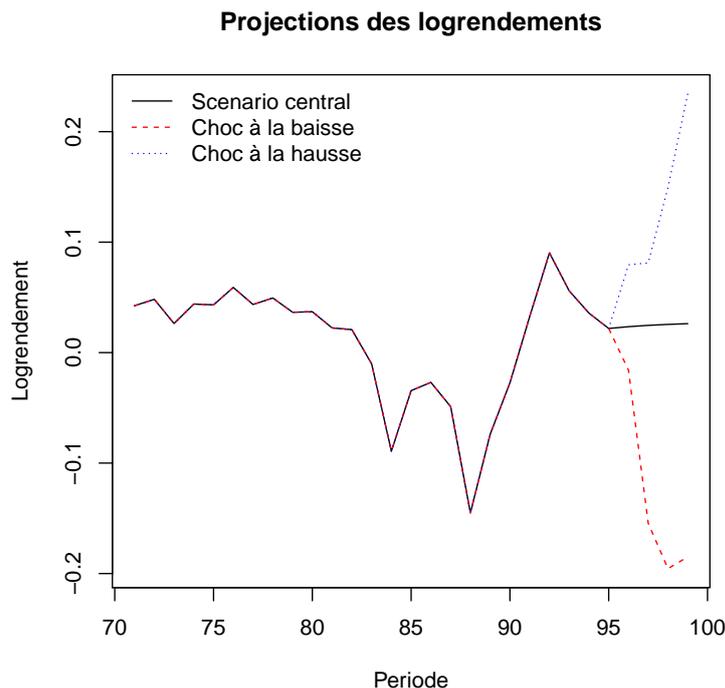


FIGURE 8.13 – Projections de la série des log-rendements

Nous prenons le quantile empirique - et non pas un quantile obtenu par application de la théorie des valeurs extrêmes sur cet échantillon - car l'échantillon est de très grande taille et nous avons vu précédemment que les quantiles estimés par TVE étaient meilleurs que le quantile empirique lorsque les échantillons étaient de petite taille, et que le rapport de force s'inversait pour de grandes tailles. Ici, il est donc plus pertinent d'utiliser le quantile empirique.

Tout d'abord, nous pouvons remarquer que la valeur du quantile estimée est plus faible que celle que nous avons trouvée dans le cas des rendements trimestriels annualisés, qui était de -28,2 %. La raison de cette différence notable est qu'ici nous avons permis un phénomène de compensation, qui n'était pas présent jusque-là puisque nous considérons qu'un choc trimestriel se répercutait avec la même intensité sur toute l'année. Le fait d'avoir pu mener l'étude sur des résidus indépendants nous a permis d'avoir une approche plus réaliste du comportement du rendement à un an.

Afin de vérifier la cohérence de cette valeur, regardons la figure 7.2 représentant les rendements annuels glissants. Nous avons vu qu'ils n'avaient pas de bonnes propriétés de stationnarité, toutefois ils ont le mérite d'être de vraies réalisations du rendement à un an. Nous voyons qu'il y a eu plusieurs rendements inférieurs à -15,5 %, mais ceux-ci ont eu lieu successivement donc il est préférable de ne les compter qu'une fois étant donnée la forte auto-corrélation de la série. De plus, pour la même raison il nous faut nous intéresser aux dernières réalisations du rendement, qui sont proches de 20 %. Il est donc raisonnable de penser que le choc bicentenaire sur l'immobilier ferait passer ce rendement de 20 % à -15,5 % en un an.

En conclusion de cette section, nous avons modélisé les log-rendements de l'immobilier à l'aide d'un modèle $AR - GARCH$ pour prendre en compte les variations de volatilité. Cela nous a permis d'obtenir une distribution à un an du rendement de l'immobilier, qui peut être utilisée dans un modèle interne. De plus, nous remarquons que notre choc bicentenaire sur l'immobilier correspond à un rendement de -15,5 % ce qui est moins important que celui donné par le QIS 5,

à savoir -25 %.

8.3 Modélisation du facteur de risque action

8.3.1 Modélisation des log-rendements par un modèle $AR - GARCH$

Les actifs obligataires sont privilégiés par les compagnies d'assurance car ils permettent de libérer le capital investi à une date connue tout en limitant le risque de marché. Néanmoins, les assureurs sont poussés à investir dans des actifs plus risqués, comme les actions afin d'améliorer leurs rendements et ainsi rester compétitifs. En effet, à long terme, les actions ont en général un rendement nettement plus élevé que les autres types de placement moins risqués grâce à leur prime de risque ; leur rôle dans les investissements des compagnies d'assurance est donc important. Afin d'évaluer la quantité de capital à détenir pour gérer le risque lié à la détention des actions, Solvabilité II inclut dans la formule standard un calcul basé sur un choc à 0,5 % sur les rendements des actions. Nous allons estimer ce choc à l'aide de la théorie des valeurs extrêmes et le comparer à la valeur du choc donné par le QIS 5.

Pour cela, nous disposons de données journalières (3391 données) retraçant l'historique du CAC 40 depuis le 1er janvier 1999 jusqu'au 30 décembre 2011.

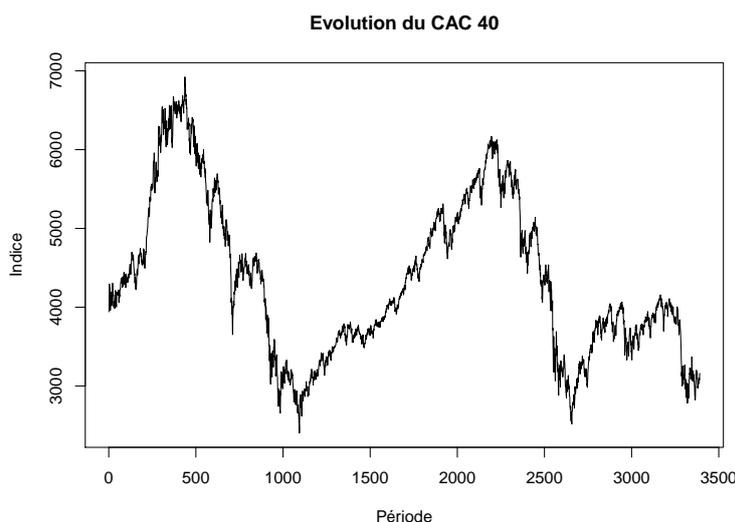


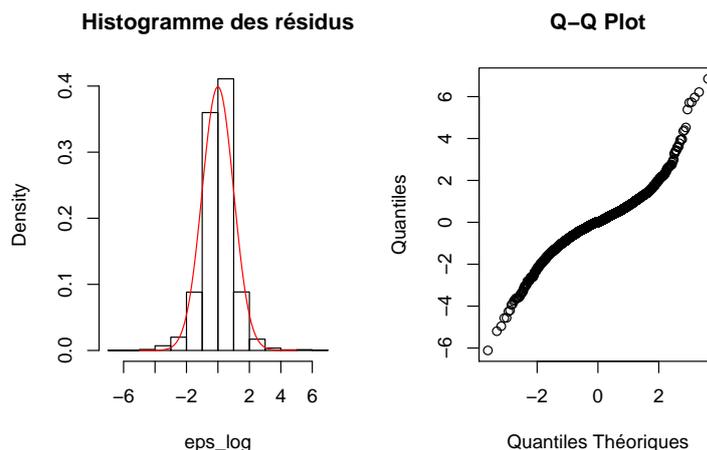
FIGURE 8.14 – Evolution de l'indice du CAC 40 entre 1999 et 2011

Le graphique nous laisse penser que l'indice CAC 40 est un processus non stationnaire. Nous le vérifions en effectuant un test de Phillips-Perron qui nous renvoie une p -value de 0,7502, ce qui nous confirme la non-stationnarité de la série.

Comme pour l'indice de l'immobilier IPD UK, nous préférons donc travailler avec la série des log-rendements journaliers définie par :

$$LR_t = \ln \left(\frac{I_t}{I_{t-1}} \right)$$

Dans ce cas encore, l'utilisation du modèle de Black-Scholes est inadéquate. En effet, elle suppose la normalité des log-rendements : $LR_t = m + \sigma \epsilon$ où ϵ suit une loi normale centrée réduite.


 FIGURE 8.15 – Comparaison de la distribution de ϵ avec la loi normale centrée réduite

L'histogramme et le graphique montrent que les queues de distribution sont sous-estimées si nous optons pour un modèle gaussien, ce qui aura pour conséquence une sous-estimation du risque de perte extrême et donc une sous-estimation du besoin en capital pour faire face au risque action. De plus, le test de Shapiro-Wilk qui a pour hypothèse nulle la normalité de la série a une p -value inférieure à 2×10^{-16} . Nous sommes donc en présence d'une distribution leptokurtique.

En outre, en observant la série des log-rendements, nous observons un regroupement des extrêmes (ou *volatility clustering*) : de fortes variations ou fortes valeurs tendent à être suivies par d'autres grandes variations.

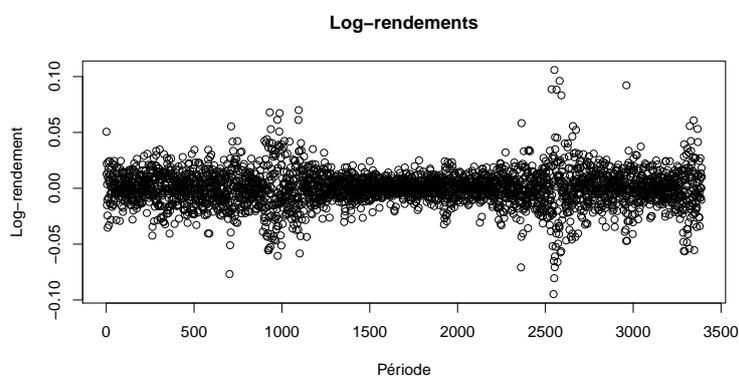


FIGURE 8.16 – Représentation de la série des log-rendements du CAC 40

Ces propriétés expliquent la difficulté de modéliser ces rendements avec des séries de type ARMA qui supposent une variance conditionnelle constante ou avec des modèles de type Black-Scholes qui supposent moyenne et variance constantes. Il nous semble donc judicieux de choisir une modélisation dynamique de la moyenne et de la volatilité des log-rendements.

Nous choisissons comme pour l'indice immobilier une modélisation $AR(1) - GARCH(1,1)$ qui permet de prendre en compte des faits stylisés tels que la variation de la volatilité dans le temps et le phénomène de *volatility clustering*. Elle permet de plus de simuler des rendements qui intègrent le niveau actuel de la volatilité constatée sur les marchés, ce qui est particulièrement intéressant dans le cadre de la mise en place d'un modèle interne. En gardant les mêmes notations que pour l'immobilier, nous avons :

$$\begin{cases} LR_t = \mu + \rho LR_{t-1} + \epsilon_t \\ \epsilon_t = \eta_t \sqrt{h_t} \\ h_t = \omega + \alpha \epsilon_{t-1}^2 + \beta h_{t-1} \end{cases}$$

Voici les estimateurs des paramètres obtenus ainsi que les *p-values* associés au test de significativité (qui a pour hypothèse nulle la nullité du paramètre) :

$\hat{\mu}$	$5,144 \times 10^{-4}$	0,00708
$\hat{\rho}$	$-3,029 \times 10^{-2}$	0,08367
$\hat{\omega}$	$1,892 \times 10^{-6}$	0,00118
$\hat{\alpha}$	$9,100 \times 10^{-2}$	$5,87 \times 10^{-13}$
$\hat{\beta}$	$9,028 \times 10^{-1}$	$< 2 \times 10^{-16}$

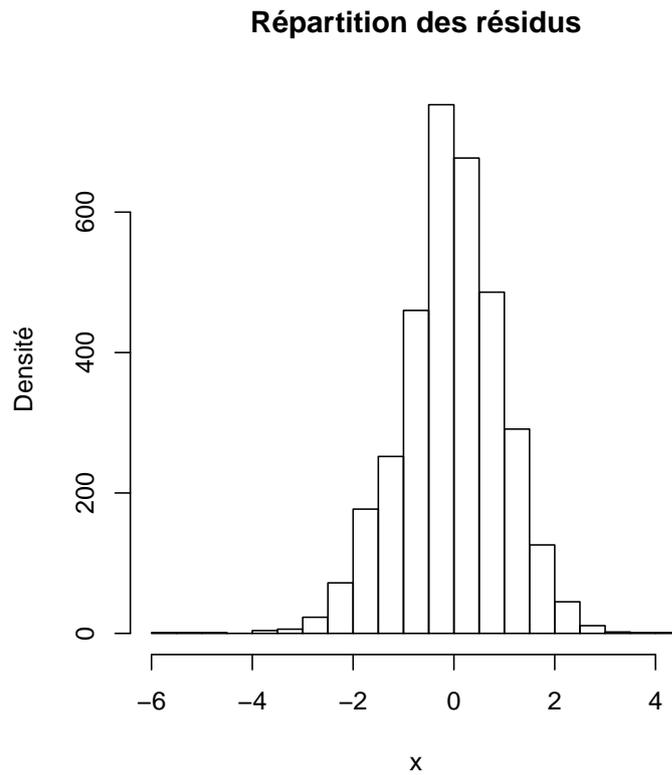
TABLE 8.5 – Valeurs des paramètres du modèle $AR(1) - GARCH(1,1)$

Tous les paramètres du modèle sont donc significatifs au seuil de 10 %. Afin de valider le modèle, nous vérifions l'indépendance des résidus $(\eta_t)_t$ avec le test de Wald-Wolfowitz et le test de Ljung-Box. Voici les *p-values* que nous obtenons :

WALD-WOLFOWITZ	0,246
LJUNG-BOX ($m = 1$ mois)	0,635
LJUNG-BOX ($m = 1$ an)	0,231
LJUNG-BOX ($m = 2$ an)	0,816

TABLE 8.6 – Tests d'indépendance des résidus du modèle $AR(1) - GARCH(1,1)$

Les résidus peuvent donc être considérés comme indépendants. Cependant, le test de Shapiro-Wilk montre qu'ils ne sont pas gaussiens (*p-value* = $6,903 \times 10^{-12}$). Ils sont en effet compris entre -5,7 à 4,3 et nous comptons plus de 4 valeurs en dessous de -4. Cela semble indiquer une queue épaisse du côté gauche de la distribution, ce qui est peu compatible avec l'hypothèse d'une loi normale centrée réduite.

FIGURE 8.17 – Répartition des résidus $(\eta_t)_t$

Afin de calculer le choc à 0,5 % le plus précisément possible, nous allons modéliser la distribution à l'aide d'une loi de Pareto Hybride, puis nous calculerons le quantile à 0,5 % d'horizon un an.

Pour cela, nous considérerons la série $(-\eta_t)_t$ afin de ramener le problème au calcul d'un quantile à 99,5 %.

Avant de déterminer les paramètres de la loi GPD, nous commençons par déterminer le seuil u . La méthode de Pickands fournit un seuil de 1,53. De plus, le graphique de la fonction moyenne des excès montre une rupture de pente vers 1,50. Nous décidons donc de retenir un seuil u égal à 1,50.

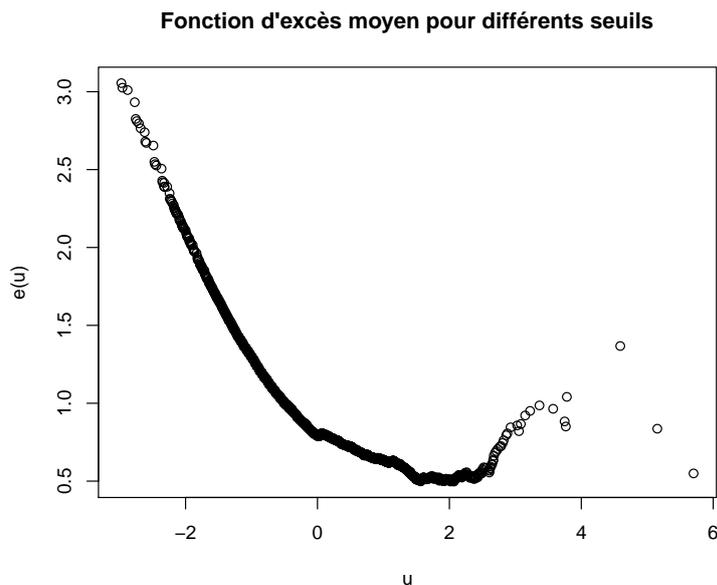


FIGURE 8.18 – Fonction d'excès moyen pour $-\eta$

Le seuil fixé, nous estimons ensuite les paramètres de la GPD :

$\hat{\xi}_{EMV}$	0,052	$\hat{\xi}_{PWM}$	0,032
$\hat{\sigma}_{EMV}$	0,484	$\hat{\sigma}_{PWM}$	0,495

TABLE 8.7 – Paramètres de queue de $-\eta$

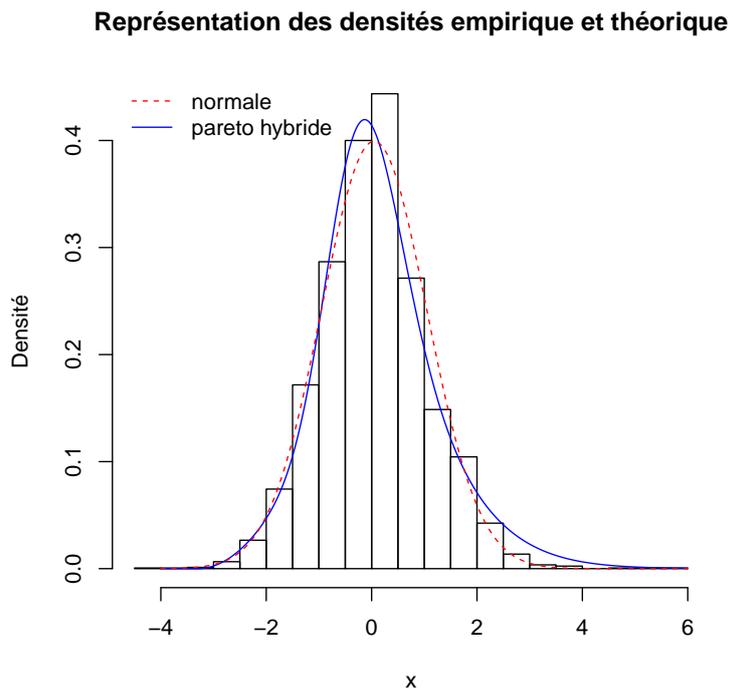
La méthode du maximum de vraisemblance et celle des moments pondérés donnant des résultats similaires, nous pouvons donc retenir les estimateurs EMV. L'estimateur de ξ qui est positif et très proche de 0, nous indique une queue de distribution à droite légèrement plus épaisse que la gaussienne. Un calcul de l'excès de kurtosis nous confirme cette idée car nous obtenons 0,82.

La calibration de la loi gaussienne correspondant à la partie centrale de la distribution donne les résultats suivants :

$\hat{\mu}_{norm}$	-0,345
$\hat{\sigma}_{norm}$	0,541
$\hat{\epsilon}_{mix}$	4,810

TABLE 8.8 – Paramètres de la partie centrale de $-\eta$

Comparons la densité de la loi de Pareto Hybride estimée avec la densité empirique de $-\eta$ afin de voir si elle donne un meilleur résultat que la loi gaussienne.


 FIGURE 8.19 – Densités théoriques et empirique de $-\eta$

Dans la partie centrale, nous remarquons que la loi normale et la Pareto Hybride donnent un résultat correct, avec un léger avantage pour la loi normale. En ce qui concerne la queue droite (région de la distribution qui nous intéresse le plus), la loi hybride semble, sur la graphique, surestimer la queue de distribution. En réalité, cette épaisseur de queue permet de capter les valeurs extrêmes de $-\eta$ qui sont comprises entre 4 et 6 et qui ne sont pas prises en compte par la gaussienne. La présence de ces valeurs extrêmes et le souci d'évaluer au mieux possible les risques extrêmes liés aux actions justifient le recours à une loi Pareto Hybride.

Un test de Kolmogorov-Smirnov peut nous confirmer cette idée : pour la loi normale, nous obtenons une p -value de 0,01 contre 0,24 pour la loi de Pareto Hybride. Nous acceptons donc la loi de Pareto Hybride à 5 %.

8.3.2 Obtention de la distribution à un an

A partir du modèle, nous pouvons facilement calculer en t la VaR en $t + 1$ de niveau p à l'aide du quantile α de u_t noté u_p :

$$VaR_{t+1}(p) = \hat{\mu} + \hat{\rho}LR_t + u_p\sqrt{\hat{\omega} + \hat{\alpha}(LR_t - \hat{\mu} - \hat{\rho}LR_{t-1})^2 + \hat{\beta}\hat{h}_t^2}$$

Cependant, dans Solvabilité II, ce qui nous intéresse n'est pas la VaR à un jour mais celle à un an. Comme nous ne pouvons pas directement expliciter une formule de cette quantité, nous procédons par simulation de la même manière que pour l'immobilier, mais en effectuant cette fois-ci 252 itérations puisque nous sommes ici en présence de données journalières.

Nous pouvons donc en déduire le choc à 0,5 % sur les log-rendements en prenant le quantile empirique de cette distribution. Nous obtenons -57,2 %, ce qui donne sur les rendements un choc à un an de -43,5 %.

En analysant la série des rendements annuels glissants, nous pouvons constater que ce choc a été dépassé une vingtaine de fois depuis 2 000 et que la valeur minimale du rendement à un an observé avoisine -49 %. La valeur du choc que nous obtenons semble donc raisonnable.

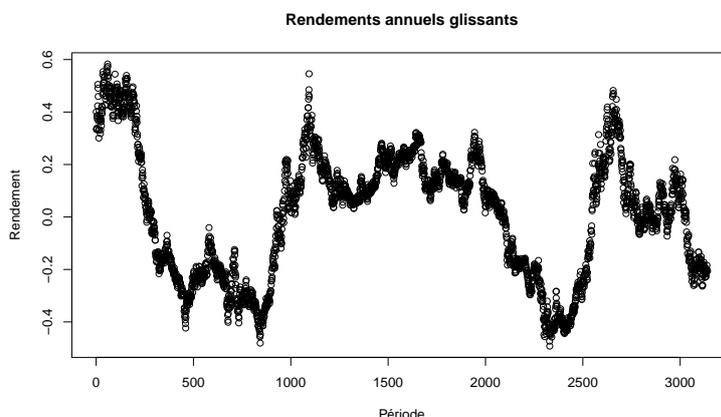


FIGURE 8.20 – Rendements annuels glissants entre 2000 et 2011

Nous notons que ce choc est supérieur à 39 % qui est celui recommandé dans le QIS 5 pour les actions. Avec la modélisation que nous avons adoptée, un choc égal à 30 % correspondrait à un quantile à 1,14 %.

8.4 Modélisation du facteur de risque taux

Le risque de taux est un risque qui affecte l'actif et le passif des assureurs. De plus, les assureurs-vie sont sensibles à une baisse des taux comme à une hausse. En effet, en cas de baisse des taux, ils peuvent être amenés à avoir des difficultés à honorer leurs contrats à taux garanti ; mais en cas de hausse de taux, les souscripteurs sont incités à racheter leurs contrats pour les placer à un taux plus avantageux pour eux. Pour ces raisons, il est capital pour les assureurs de bien gérer leur risque de taux.

Notre base de donnée est constituée de la valeur des taux d'intérêt à 1 an, 2 ans, 3 ans, 4 ans, 5 ans, 7 ans, 10 ans, 15 ans, 20 ans et 30 ans. Les observations sont faites quotidiennement (hors *week-end*) du 3 janvier 2000 au 30 décembre 2011. Voici par exemple la forme de la courbe des taux à différentes dates :

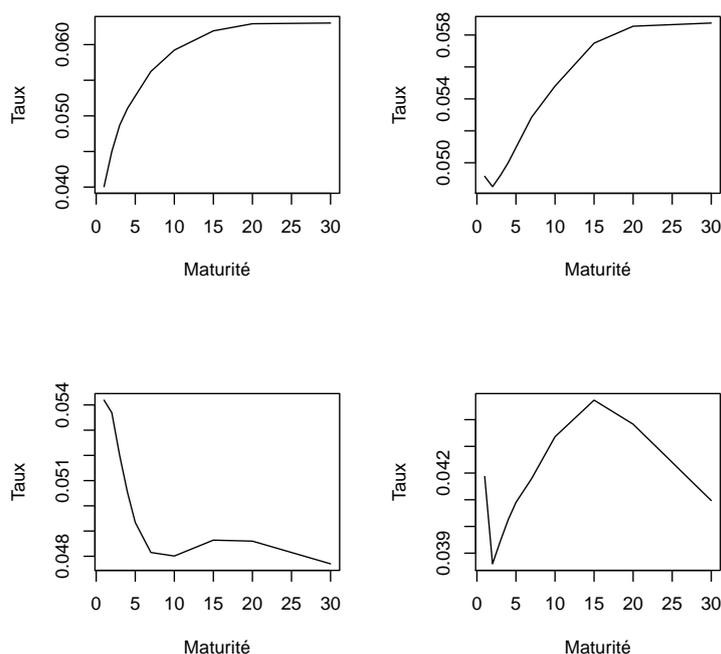


FIGURE 8.21 – Représentation des taux pour quatre dates distinctes

Les taux pour les maturités manquantes sont estimés par interpolation linéaire, ils ne seront de toute façon pas étudiés ici donc cette simplification n'aura pas d'impact sur la suite. Nous voyons que les courbes n'ont pas toutes la même forme, par exemple du 10 août 2007 au 19 décembre 2008, le taux d'intérêt 2 ans avait une valeur inférieure au taux d'intérêt à 1 an (ce qui se voit sur plusieurs courbes), ce qui est un phénomène particulier. Notons que la forme de courbe la plus courante dans notre historique est celle en haut à droite. Pour bien gérer le risque de taux, il faut un modèle capable de prendre en compte ces différentes formes de courbe.

Intéressons-nous maintenant aux caractéristiques des différentes maturités :

Moments / Maturités	1 an	2 ans	3 ans	4 ans	5 ans	7 ans	10 ans	15 ans	20 ans	30 ans
Moyenne	0,0304	0,0322	0,0341	0,0358	0,0372	0,0397	0,0423	0,0448	0,0458	0,0457
Ecart-type	0,0128	0,0119	0,0111	0,0105	0,0099	0,0092	0,0086	0,0082	0,0082	0,0087
Skewness	0,1703	0,1153	0,0788	0,0691	0,0850	0,1289	0,1504	0,1586	0,1029	-0,0183
Excès de Kurtosis	-1,258	-1,203	-1,090	-0,980	-0,894	-0,768	-0,668	-0,603	-0,589	-0,648

TABLE 8.9 – Moments des taux d'intérêt pour différentes maturités

Nous voyons que les taux ont tendance à monter en moyenne, tout en étant moins volatiles, ce qui est conforme à ce que nous savons des courbes de taux. De plus, nous remarquons que tous les taux ont des distributions plus aplaties que la loi normale.

Pour terminer cette étude préliminaire, regardons l'évolution des taux d'intérêt à 1 an, 3 ans, 10 ans, et 30 ans au cours du temps :

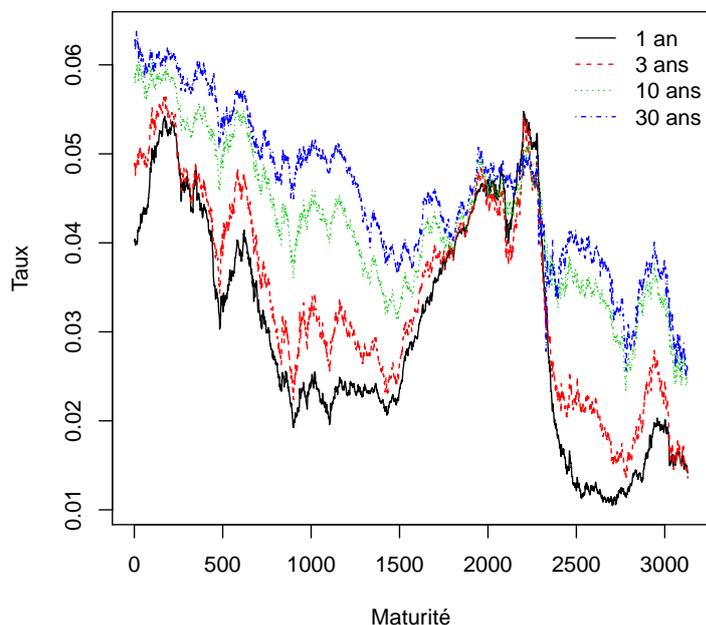


FIGURE 8.22 – Evolution du taux au cours du temps pour quatre maturités

Il apparaît clairement une dépendance des différents taux entre eux, les fluctuations sont très proches les unes des autres. Nous remarquons une période déjà mise en évidence précédemment durant laquelle les taux sont très proches les uns des autres et où le taux à un an a tendance à être supérieur, signe d'un manque de confiance dans le marché. Il nous semble donc indispensable de tenir compte de cette forte dépendance entre les taux de différentes maturités dans notre modèle.

8.4.1 Utilisation d'une Analyse en Composantes Principales

Une façon d'approcher les taux d'intérêt a été popularisée par LITTERMAN et SCHEINKMAN en 1991. Elle consiste à pratiquer une Analyse en Composantes Principales (ou ACP) sur la valeur des taux à différentes maturités pour identifier des facteurs de risque. Dans la suite, chaque date sera considérée comme un individu et chaque maturité comme une variable. Notre matrice de taux est donc une matrice à 3130 individus et 10 variables. Dans ce qui suit, nous considérons les logarithmes de taux afin d'obtenir des taux positifs à la fin.

Commençons par centrer et réduire chacune des variables et appelons Y la matrice obtenue. Elle a donc exactement les mêmes dimensions que la matrice des taux, la différence étant que chaque variable a une moyenne de 0 et un écart-type de 1. Cela permet de ne pas privilégier une maturité par rapport à une autre, toutes vont donc avoir le même impact. Le principe de l'ACP est de diagonaliser la matrice $\frac{YY'}{3130}$. Rappelons que cela revient à trouver une matrice D diagonale et une matrice V telles que :

$$\frac{YY'}{3130} = VDV'$$

Ceci est possible car YY' est symétrique. Les matrices D et V sont toutes deux des matrices carrées de dimension 10. Posons alors $C = YV$, les colonnes de C sont appelées "composantes principales". En nommant V_1, V_2, \dots, V_{10} les colonnes de V , nous avons par exemple $C_1 = YV_1$,

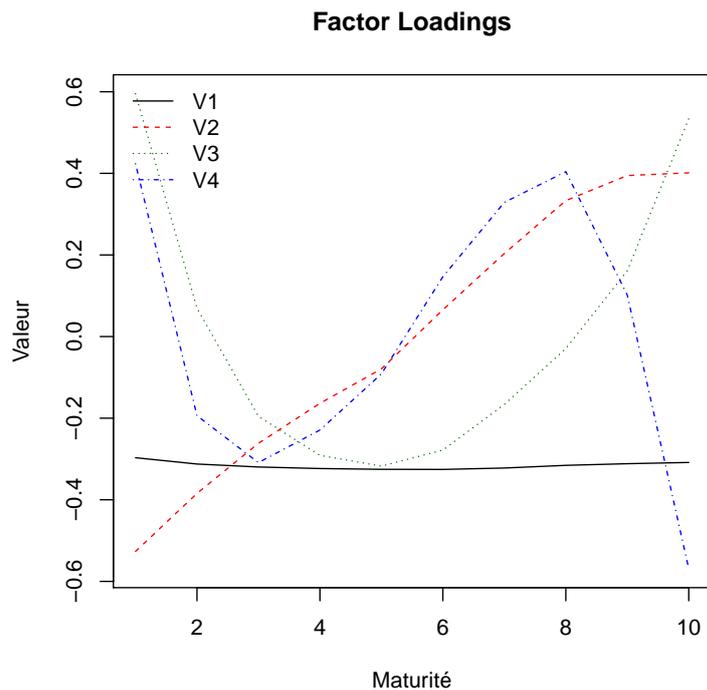
avec C_1 qui est la "première composante principale".

Nous avons la relation $Y = CV'$, qui n'est que la réécriture de la relation précédente. Mais si nous n'utilisons que C_1 au lieu d'utiliser l'intégralité de la matrice C , nous pouvons approcher Y par C_1V_1' (pour rappel, C_1 est un vecteur-colonne de taille 3130 et V_1' un vecteur-ligne de taille 10, le résultat est donc une matrice de 3130 lignes et 11 colonnes, soit la même taille que Y). Un résultat d'analyse de données nous dit alors que la variance de C_1 est égale à la plus grande valeur propre de D . Appelons $\lambda_1, \lambda_2, \dots, \lambda_{10}$ ces valeurs propres classées par ordre décroissant, la part de variance totale de Y expliquée par C_1V_1' est alors : $\frac{\lambda_1}{\sum_{i=1}^{10} \lambda_i}$, c'est-à-dire 93,48 % dans notre cas.

Afin de gagner en précision, nous pouvons intégrer une deuxième composante principale : Y est alors approché par $C_1V_1' + C_2V_2'$, ce qui permet d'augmenter la part de variance expliquée à 99,56 %. Nous allons finalement utiliser 4 composantes principales, nous approchons donc Y par $C_1V_1' + C_2V_2' + C_3V_3' + C_4V_4'$, ce qui permet d'expliquer 99,98 % de la variance.

L'intérêt de cette méthode est double. Tout d'abord, nous avons restreint les 10 maturités à 4 composantes principales : le nombre de paramètres à estimer est donc réduit. Mais surtout, un résultat d'analyse de données indique l'indépendance des composantes principales entre elles. Pour simuler une courbe de taux, il nous suffit donc de créer un nouvel individu qui aurait une valeur pour chacune des 4 premières composantes principales, ces valeurs étant tirées indépendamment les unes des autres selon leur loi marginale. Appelons-les c_1, c_2, c_3 , et c_4 ; ce sont 4 réels. $e^{c_1V_1'+c_2V_2'+c_3V_3'+c_4V_4'}$ est alors une courbe des taux à 10 maturités intégrant la structure de dépendance entre les différentes maturités. C'est là le grand avantage de cette approche : il n'est pas nécessaire de supposer une corrélation linéaire entre les différentes maturités, la structure de dépendance est directement captée par la matrice V , et l'indépendance des composantes principales entre elles facilite la génération de courbes et le calcul de quantiles.

Les colonnes de V sont appelées *factor loadings*, voici les valeurs de V_1 à V_4 pour les 10 maturités :

FIGURE 8.23 – Représentation des quatre premiers *Factor Loadings*

Le premier *factor loading* est presque constant, il s'agit du facteur de translation : il fixe le niveau de la courbe. Le deuxième est croissant, c'est le facteur de pente. Le troisième fixe la courbure, et enfin le quatrième est le facteur de rotation. Selon les valeurs prises par les composantes principales, ces facteurs vont donc influencer la forme de la courbe.

A titre d'exemple, étudions la reconstruction de deux courbes de taux par ces *factor loadings*. La première d'entre elles est une courbe relativement classique, qui est en fait la première courbe de notre historique :

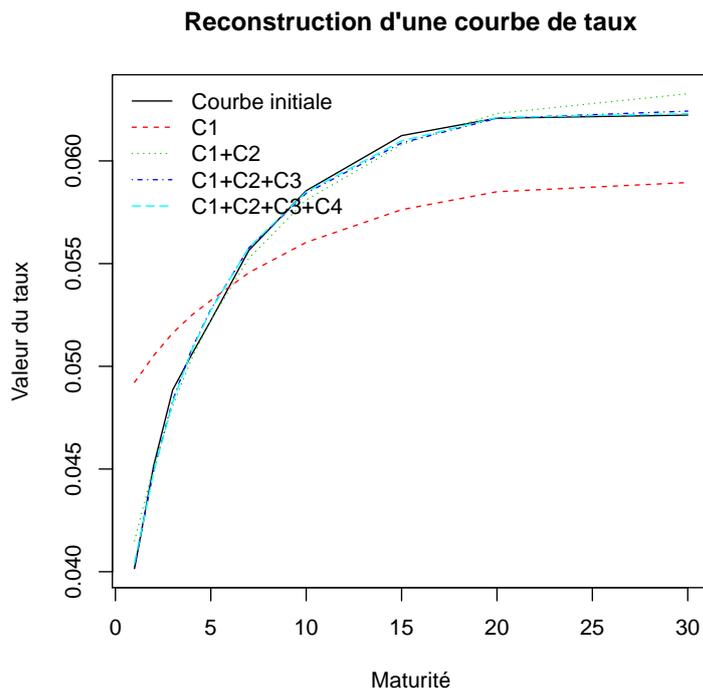


FIGURE 8.24 – Comparaison entre une vraie courbe des taux et ses approximations successives

La notation "C1" signifie que seule la première composante principale a été utilisée, "C1+C2" que les deux premières ont été utilisées, etc.

La première chose à remarquer est que la courbe "C1" n'est pas constante alors que l'on s'attendait à ce qu'elle le soit : en fait, la première étape de notre ACP a été de centrer et réduire les variables. Pour observer les courbes de taux, nous devons faire la même chose en sens inverse (ainsi que prendre l'exponentielle), c'est la raison pour laquelle elle n'apparaît pas constante. Si nous avions fait la même reconstruction sur les variables centrées réduites, le résultat aurait été similaire mais les courbes n'auraient pas eu les mêmes formes : la courbe "C1" aurait eu l'allure attendue.

Nous voyons qu'à partir de la 3^{ème} composante principale les courbes sont extrêmement proches, la 4^{ème} ne semble pas apporter grand-chose. Confirmons cela en regardant les valeurs des 4 premières composantes principales pour cette courbe :

$\frac{c_1}{\sqrt{\lambda_1}}$	$\frac{c_2}{\sqrt{\lambda_2}}$	$\frac{c_3}{\sqrt{\lambda_3}}$	$\frac{c_4}{\sqrt{\lambda_4}}$
-1,61	-1,51	-1,01	0,24

TABLE 8.10 – Valeurs des composantes principales

Nous divisons à chaque fois par la racine carrée de la valeur propre associée pour avoir des valeurs comparables car les composantes principales sont des variables centrées de variance égale à la valeur propre associée.

La valeur la plus importante relativement aux autres est celle qui fixe le niveau. En effet, si nous regardons la figure 8.22, nous voyons que la première courbe avait un niveau plus élevé que les autres. Ensuite viennent la pente, la courbure, et la rotation qui est comme prévue moins importante que les autres.

Intéressons-nous maintenant à une autre courbe :

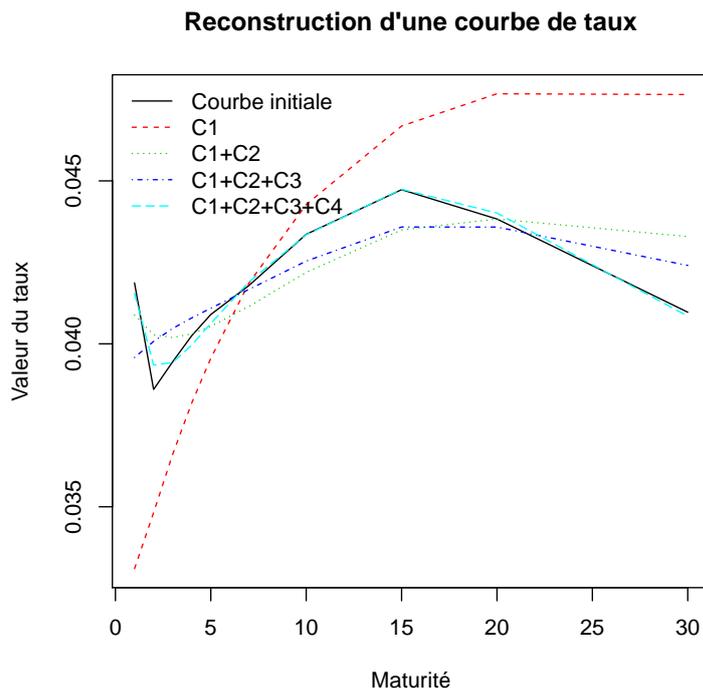


FIGURE 8.25 – Comparaison entre une vraie courbe des taux et ses approximations successives

Celle-ci a un effet de rotation qui n'était pas présent dans l'autre courbe, nous nous attendons donc à une plus grande importance de la 4^{eme} composante principale, ce qui se vérifie dans le tableau suivant :

$\frac{c_1}{\sqrt{\lambda_1}}$	$\frac{c_2}{\sqrt{\lambda_2}}$	$\frac{c_3}{\sqrt{\lambda_3}}$	$\frac{c_4}{\sqrt{\lambda_4}}$
-0,24	1,53	-1,06	2,93

TABLE 8.11 – Valeurs des composantes principales

Le facteur le plus important est effectivement le facteur de rotation, puis viennent la pente, la courbure, et enfin le niveau qui est faible. Cela signifie que la courbe est proche (en terme de valeur moyenne) de la moyenne des courbes.

Nous allons maintenant voir comment utiliser ces résultats pour calculer les valeurs de taux extrêmes pour chaque maturité dans le cadre de la formule standard, ainsi que simuler des courbes de taux futures dans le cadre du modèle interne. Nous devons pour cela modéliser les composantes principales. Observons tout d'abord leur comportement au cours du temps :

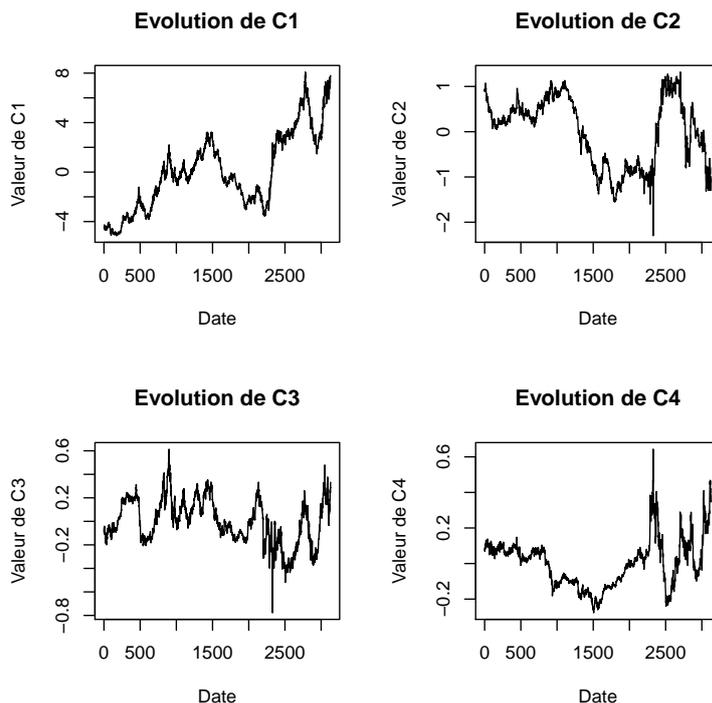


FIGURE 8.26 – Evolution des quatre premières composantes principales au cours du temps

Une dynamique temporelle apparaît clairement dans ces graphiques, nous ne pouvons pas considérer les valeurs successives de C_1 (par exemple) comme indépendantes. Si nous estimions un simple processus $AR(1)$, nous aurions un paramètre de valeur égale à 0,9993. Ce résultat laisse supposer que la série serait intégrée d'ordre 1. Nous le confirmons par un test de Phillips-Perron qui fournit une p -value de 0,73 : il y a bien une racine unitaire, C_1 est donc une marche aléatoire. Afin de la rendre stationnaire, nous nous intéressons à la série $C'_1 = \Delta C_1$, où Δ est l'opérateur de première différence :

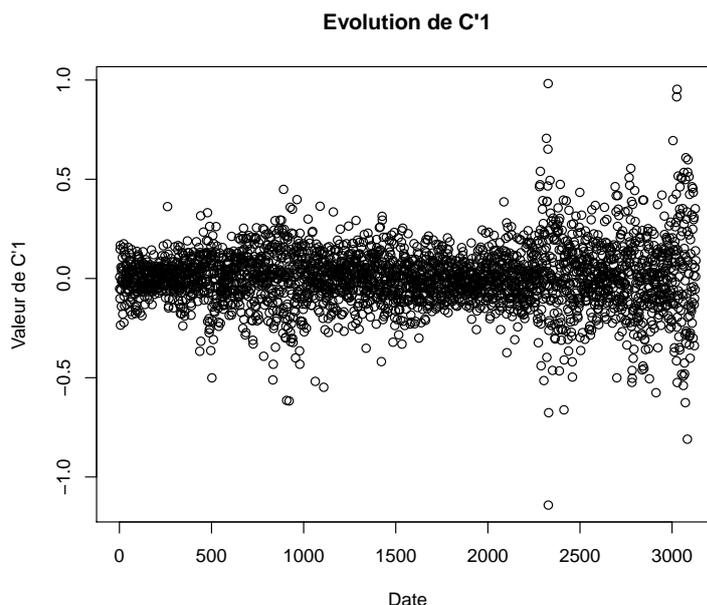


FIGURE 8.27 – Evolution de C'_1 au cours du temps

Nous voyons qu'à l'instar de l'indice immobilier ou de l'indice action, il existe des périodes où la série varie davantage que d'autres. Pour cette raison, nous choisissons une nouvelle fois une modélisation $GARCH(p, q)$, dont nous rappelons la forme générale :

$$\begin{cases} X_t = \mu + \epsilon_t \\ \epsilon_t = \eta_t \sqrt{h_t} \\ h_t = \omega + \sum_{i=1}^q \alpha_i \epsilon_{t-i}^2 + \sum_{j=1}^p \beta_j h_{t-j} \end{cases}$$

Nous commençons par traiter C'_1 en testant une modélisation simple : le modèle $GARCH(1, 1)$. Voici le tableau résumant la valeurs des paramètres ainsi que leur significativité :

$\hat{\mu}$	$1,039 \times 10^{-4}$	0,625
$\hat{\omega}$	$1,257 \times 10^{-4}$	0,031
$\hat{\alpha}$	$3,650 \times 10^{-2}$	$4,06 \times 10^{-6}$
$\hat{\beta}$	$9,562 \times 10^{-1}$	2×10^{-16}

TABLE 8.12 – Valeurs des paramètres du modèle $GARCH(1, 1)$

Tous les paramètres sont donc significatifs, à l'exception de la constante mais cela n'a pas d'importance. Nous vérifions aussi l'indépendance des résidus $(\eta_t)_t$ par des tests de Wald-Wolfowitz et de Ljung-Box avec des décalages allant d'un mois à deux ans. Nous considérons également les carrés des résidus :

Test	η	η^2
WALD-WOLFOWITZ	0,36	0,36
LJUNG-BOX (m = 1 mois)	0,48	0,52
LJUNG-BOX (m = 1 an)	0,63	0,06
LJUNG-BOX (m = 2 ans)	0,99	0,61

TABLE 8.13 – Tests d'indépendance des résidus

En conclusion, nous pouvons considérer les résidus comme indépendants, ce qui nous permet de les modéliser et ainsi de projeter la valeur de C'_1 (et donc C_1) dans le futur. Néanmoins, cette fois-ci, nous devons considérer les scénarios haussiers et les scénarios baissiers, il nous faut par conséquent intervenir sur les deux queues de distribution. En effet, voici ce que nous obtenons en utilisant une loi normale :

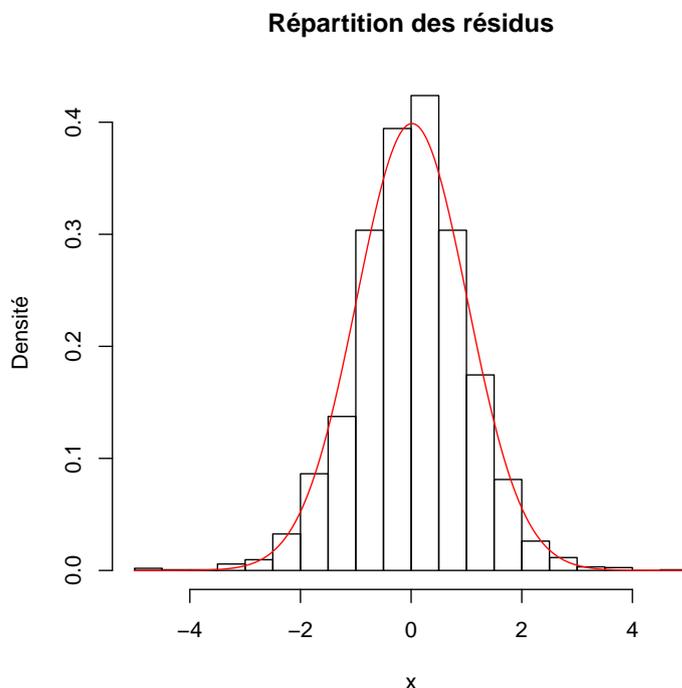


FIGURE 8.28 – Représentation des densités théorique et empirique de η

Le fait d'utiliser une GPD à droite et à gauche de la distribution pose davantage de problèmes d'estimation des paramètres qu'en ne considérant qu'un seul côté : empiriquement, nous constatons que si nous estimons chaque queue séparément, nous n'avons plus assez de degrés de liberté pour à la fois bien approcher l'histogramme en partie centrale et pour avoir une densité régulière. Nous décidons donc de fixer uniquement la valeur des seuils u_1 et u_2 et d'estimer tous les autres paramètres par maximum de vraisemblance. Voici les fonctions moyennes des excès pour η et $-\eta$:

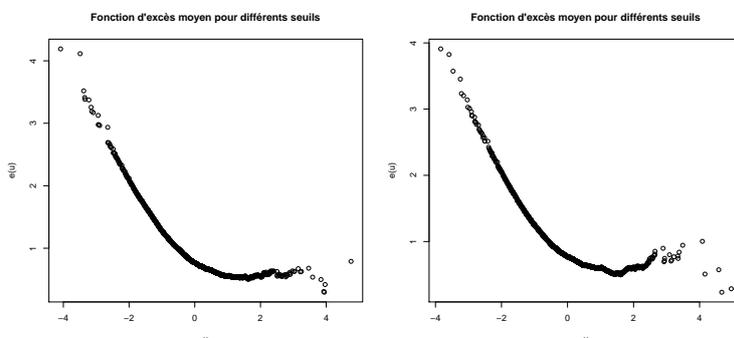


FIGURE 8.29 – Fonctions moyennes des excès pour η et $-\eta$

Celles-ci sont assez semblables, pour des raisons de simplicité nous prenons deux seuils op-

posés l'un de l'autre au niveau de la dernière rupture de pente : $u_1 = -1,5$ et $u_2 = 1,5$. Nous estimons alors les paramètres par maximum de vraisemblance, et voici la densité obtenue :

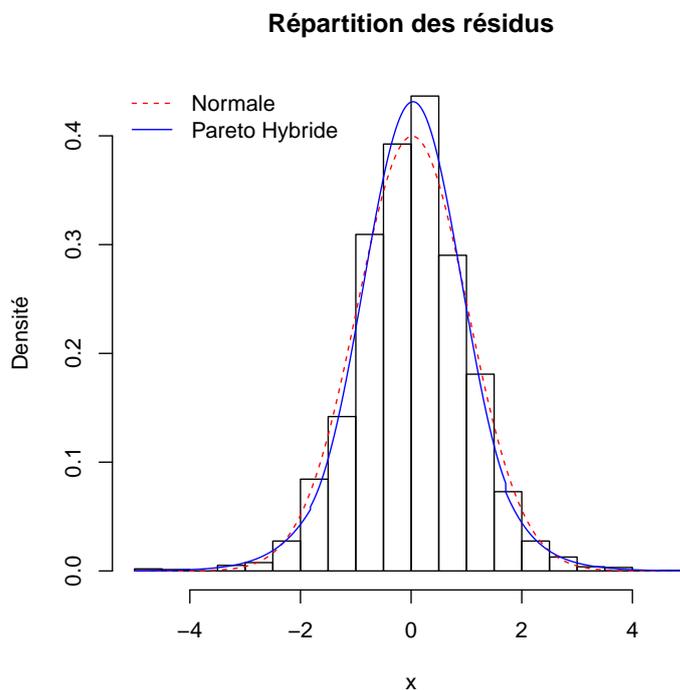


FIGURE 8.30 – Représentation des densités théorique et empirique de η

μ_{norm}	0,108
σ_{norm}	0,467
ξ_1	0,080
μ_1	-1,500
σ_1	0,307
ϵ_1	3,286
ξ_2	0,099
μ_2	1,500
σ_2	0,267
ϵ_2	3,361

TABLE 8.14 – Paramètres pour C_1

Cette densité semble plus proche de l'histogramme observé que la densité normale, cela justifie notre approche. Nous pouvons confirmer cela par des tests de KOLMOGOROV-SMIRNOV : la p -value pour notre loi est de 0,09 tandis qu'elle est de 0,03 pour une loi normale. Nous acceptons donc la Pareto Hybride à 5 % et rejetons la normale. Nous remarquons un point de discontinuité à gauche et à droite : il s'agit du prix à payer pour la loi de Pareto Hybride à 3 densités. Ce phénomène se produit car en partie centrale de la courbe, les trois densités se mélangent du fait des valeurs élevées de ϵ_1 et ϵ_2 , et la dérivée de la fonction de pondération q_2 n'est alors pas continue. Cependant, l'impact de cette discontinuité semble vraiment très limité, le "saut" est faible et cela se produit de plus en partie centrale. Si nous avions cherché à avoir une densité continue coûte que coûte, elle aurait été beaucoup moins proche de l'histogramme que nous cherchons à approcher. Nous sommes donc satisfaits de notre résultat.

Comme ce qui a été fait précédemment, nous projetons 10 000 scénarios possibles pour C'_1 à horizon un an en utilisant comme résidus des tirages aléatoires selon la densité que nous avons estimée. Nous obtenons alors 10 000 valeurs possibles pour C_1 dans un an en sommant toutes les valeurs C'_1 au cours de chaque scénario :

$$\begin{aligned}
 C_{1,T+252} &= C'_{1,T+252} + C_{1,T+251} \\
 C_{1,T+252} &= C'_{1,T+252} + C'_{1,T+251} + C_{1,T+250} \\
 C_{1,T+252} &= \sum_{k=1}^{252} C'_{1,T+k} + C_{1,T}
 \end{aligned}$$

où T est l'instant de notre dernière observation.

Nous appliquons la même procédure à l'ensemble des composantes principales afin d'avoir les distributions à un an de C_2 , C_3 , et C_4 . Il est important de noter que les hypothèses de modèle que nous avons faites sur C_1 semblent plus contestables : les tests d'indépendance des résidus (Wald-Wolfowitz notamment) sont négatifs quelles que soient les valeurs de p et q , ce qui constitue une limite de notre travail. Toutefois, étant donné que la non-indépendance n'apparaît pas de façon flagrante visuellement (par exemple, la figure ci-dessous présente l'évolution de η pour la 4^{ème} composante principale, ainsi que son autocorrélogramme d'ordre 1), nous appliquons quand même les méthodes utilisées ci-dessus, tout en sachant que pour des résultats plus justes, il faudrait revoir la modélisation des composantes principales. Les résultats détaillés concernant ces trois composantes principales sont disponibles dans l'annexe C.

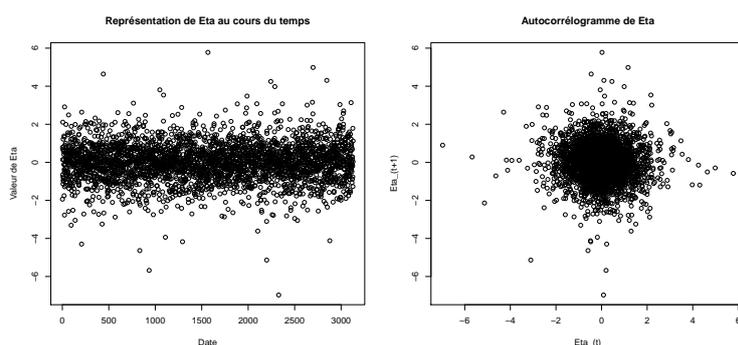


FIGURE 8.31 – Variations et autocorrélogramme de η

Nous avons donc réussi à générer des courbes de taux à un an. Pour rappel, si c_1 , c_2 , c_3 , et c_4 sont les observations des composantes principales pour un scénario donné, $e^{c_1V'_1+c_2V'_2+c_3V'_3+c_4V'_4}$ est alors la courbe de taux associée à ce scénario. En simulant 10 000 scénarios comme nous l'avons fait, une compagnie d'assurance est donc à même d'étudier la valeur de ses portefeuilles pour chaque scénario et donc de voir quelle serait sa perte bi-centenaire. L'utilisation de la théorie des valeurs extrêmes dans le cadre d'une modélisation des taux par ACP nous permet ainsi de faire un modèle interne pour les taux. Il est important de remarquer que dans ce cas, la courbe correspondant à ce choc n'est connue qu'*a posteriori* : c'est l'étude du portefeuille qui permet de remonter ensuite jusqu'à la courbe ayant produit cette valeur extrême.

Nous obtenons alors les quantiles suivants que nous pouvons comparer aux chocs donnés par la formule standard dans le QIS 5 :

Maturités	1 an	2 ans	3 ans	4 ans	5 ans	7 ans	10 ans	15 ans	20 ans	30 ans	50 ans
Taux 0.5% (baisse)	0,44%	0,44%	0,51%	0,62%	0,75%	1,00%	1,28%	1,55%	1,63%	1,55%	1,51%
Taux 99.5% (hausse)	4,44%	3,02%	2,76%	2,77%	2,87%	3,10%	3,40%	3,70%	3,83%	3,92%	4,16%

TABLE 8.15 – Valeurs extrêmes des taux dans un an

Afin de relativiser la valeur de ces taux, voyons quelles sont les valeurs de chocs associées en comparant avec la dernière courbe de taux de notre table :

Maturités	1 an	2 ans	3 ans	4 ans	5 ans	7 ans	10 ans	15 ans	20 ans	30 ans	50 ans
Taux en T	1,41%	1,31%	1,36%	1,54%	1,72%	2,07%	2,38%	2,66%	2,69%	2,56%	2,56%
Choc 0.5% (baisse)	-69%	-66%	-63%	-59%	-56%	-52%	-46%	-42%	-39%	-39%	-41%
Choc 99.5% (hausse)	+214%	+130%	+104%	+81%	+67%	+50%	+43%	+39%	+43%	+53%	+62%

TABLE 8.16 – Valeurs des chocs de taux à un an

Les chocs obtenus par notre méthode sont plus extrêmes que les chocs définis par le QIS 5, notamment ceux à la hausse. En effet, ils vont de 70 % à 25 % quand les nôtres vont de 214 % à 62 %, l'écart est donc très important. A la baisse, l'écart est moins visible, mais nous avons tout de même 10 % de différence pour chaque maturité en moyenne.

Il nous faut relativiser ces résultats en repensant à nos hypothèses de modèle, notamment le traitement des trois dernières composantes principales pour lesquelles nous avons mis en évidence la faiblesse de l'hypothèse d'indépendance, qui peut conduire à une mauvaise estimation des chocs. Il est intéressant de relever le fait que si nous avons utilisé des lois normales au lieu de lois de Pareto Hybride pour modéliser les résidus, nous aurions obtenu des chocs encore plus élevés. En effet, ces lois permettent d'atteindre des valeurs plus élevées mais en contrepartie, ont aussi davantage de valeurs autour de 0, et nous constatons par exemple que les quantiles à 25 % et 75 % sont plus élevés (en valeur absolue) pour la loi normale que pour la loi de Pareto Hybride, après les avoir calibrées sur les mêmes données. C'est la raison pour laquelle nous obtenons des chocs plus forts avec cette loi. Cependant, en comparant les vraisemblances calculées sur les données, la loi de Pareto Hybride approche mieux l'histogramme.

8.4.2 Modélisation des taux maturité par maturité

Notre deuxième approche consiste à regarder les taux maturité par maturité, et de corrélérer ensuite les résidus au moyen d'une matrice de variance-covariance calibrée sur les corrélations observées. Nous appliquons une nouvelle fois une modélisation $ARMA - GARCH$ sur les log-rendements des taux, ceux-ci étant stationnaires. Notre objectif est une nouvelle fois de prendre les ordres de l' $ARMA$ et du $GARCH$ tels que les résidus du modèle soient indépendants et en deuxième lieu que les paramètres soient significatifs. Nous nous apercevons que les η_i (si nous regardons la i^{eme} maturité) ne sont pas autocorrélés quelle que soit la maturité dès le modèle $ARMA(1, 1) - GARCH(1, 1)$, mais qu'à l'inverse η_i^2 l'est. C'est une limite de notre modélisation : nous ne captions pas encore assez bien les fluctuations de volatilité. Toutefois nous faisons le choix de conserver ce modèle en raison de sa relative simplicité, et modélisons donc les 10 séries avec au plus 4 paramètres - et parfois moins.

Après utilisation de la théorie des valeurs extrêmes sur chacune des séries η_i pour obtenir les paramètres de la série de Pareto Hybride correspondante, nous projetons 50 000 courbes de taux à un an et obtenons alors les taux haussiers et baissiers suivants :

Maturités	1 an	2 ans	3 ans	4 ans	5 ans	7 ans	10 ans	15 ans	20 ans	30 ans
Taux 0,5 % (baisse)	0,22%	0,38%	0,70%	0,35%	0,93%	0,69%	1,35%	1,90%	1,58%	1,81%
Taux 99,5 % (hausse)	9,43%	4,91%	1,66%	2,81%	4,13%	2,94%	4,81%	3,20%	3,61%	3,87%

TABLE 8.17 – Valeurs extrêmes des taux dans un an

Ce qui conduit aux chocs suivants :

Maturités	1 an	2 ans	3 ans	4 ans	5 ans	7 ans	10 ans	15 ans	20 ans	30 ans
Taux en T	1,41%	1,31%	1,36%	1,54%	1,72%	2,07%	2,38%	2,66%	2,69%	2,56%
Choc 0,5 % (baisse)	-84%	-71%	-95%	-77%	-46%	-67%	-43%	-29%	-41%	-29%
Choc 99,5 % (hausse)	+569%	+275%	+22%	+82%	+140%	+42%	+102%	+20%	+34%	+51%

TABLE 8.18 – Valeurs extrêmes des taux dans un an

Nous voyons qu'il n'y a pas vraiment de régularité dans les chocs, et qu'en fonction des paramètres choisis de l'*ARMA* – *GARCH* et de ceux de la loi de Pareto Hybride, nous pouvons obtenir des taux extrêmes bien différents. C'est un inconvénient majeur : rien dans les courbes de taux passées ne laisse à penser que le taux à 7 ans doit être choqué à la hausse à 140 % alors que celui à 4 ans doit l'être à 22 %. Il nous faut donc harmoniser nos modèles afin d'obtenir une courbe de chocs plus lisses. Une première étape est de prendre le même modèle pour toutes les maturités : *ARMA*(1, 1) – *GARCH*(1, 1). Nous avons dit plus haut que c'était le plus petit modèle tel que les séries de résidus étaient décorrélées pour toutes les maturités. Cela signifie aussi que pour certaines séries nous obtiendrons des paramètres non significatifs. L'utilisation d'un même modèle n'empêche pas des variations des paramètres, comme nous pouvons le constater sur cette figure :

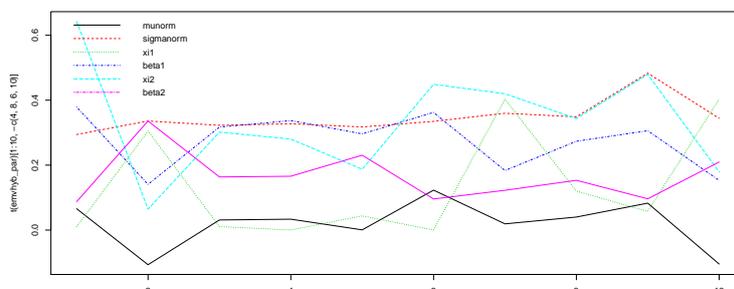


FIGURE 8.32 – Valeurs des différents paramètres pour chaque série

Nous voyons de façon assez nette que les paramètres évoluent librement en fonction de la maturité, même s'ils restent toujours du même ordre de grandeur. Notre intuition est que c'est ce phénomène qui produit les allures erratiques des courbes de chocs à la hausse et à la baisse. Pour vérifier cela, nous décidons de calibrer un seul jeu de paramètres pour l'ensemble des maturités. Cela se justifie par le fait que les estimateurs de l'*ARMA*(1, 1) – *GARCH*(1, 1) sont obtenus par quasi-maximum de vraisemblance, c'est-à-dire en supposant que les résidus suivent tous des lois normales centrées réduites et en maximisant la vraisemblance obtenue *via* cette hypothèse. Ici nous nous libérons de la normalité, mais cela nous semble raisonnable de considérer que tous les résidus sont issus d'une même loi. Nous obtenons alors un seul jeu de paramètres en calibrant sur tous les résidus à la fois :

μ_{norm}	0,028
σ_{norm}	0,344
ξ_1	0,042
μ_1	-2,000
σ_1	0,319
ϵ_1	4,969
ξ_2	0,304
μ_2	2,000
σ_2	0,158
ϵ_2	4,382

TABLE 8.19 – Paramètres communs à tous les η_i

Nous vérifions tout d'abord que ces paramètres conviennent à tous les résidus en effectuant des tests de Kolmogorov-Smirnov : ils sont tous acceptés au seuil 5 % (voir tableau 8.78 en

annexe). Nous pouvons alors projeter 50 000 taux à un an pour chaque maturité et comparer les courbes extrêmes avec la dernière courbe observée :

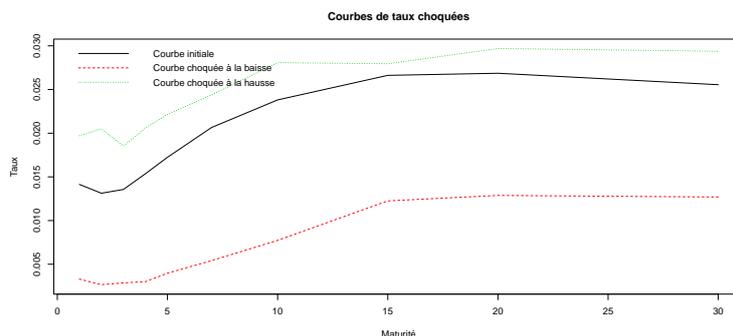


FIGURE 8.33 – Représentation des courbes de taux choquées à la hausse et à la baisse

Ce qui donne les chocs suivants :

Maturités	1 an	2 ans	3 ans	4 ans	5 ans	7 ans	10 ans	15 ans	20 ans	30 ans
Choc 0,5 % (baisse)	-77%	-80%	-79%	-80%	-77%	-74%	-68%	-54%	-52%	-50%
Choc 99,5 % (hausse)	+39%	+56%	+37%	+34%	+29%	+18%	+18%	+5%	+11%	+15%

TABLE 8.20 – Valeurs extrêmes des taux dans un an

Ceux-ci sont plus réguliers que les précédents. Nous remarquons que par rapport aux chocs obtenus par l'analyse en composantes principales, les chocs à la baisse sont du même ordre de grandeur au début mais sont plus extrêmes à la fin que ceux que nous avons trouvés. Toutefois, la différence majeure se situe au niveau des chocs à la hausse puisque nous avions auparavant des chocs bien plus élevés que dans le QIS alors que nous obtenons ici l'inverse. Cette différence est due à la très forte dépendance de notre modèle aux dernières variations des taux, qui ici étaient nettement à la baisse. La plupart des taux à un an projetés étaient donc plus faibles que la courbe initiale, ce qui a conduit à ces chocs à la hausse plutôt faibles. Cet écart met en lumière l'importance de la modélisation : selon les hypothèses que nous faisons nous pouvons arriver à des résultats bien différents. L'objet de notre mémoire n'est pas de statuer sur la meilleure façon de calculer les chocs de taux en formule standard, mais plutôt de montrer comment la théorie des valeurs extrêmes peut s'intégrer à ces calculs et ce qu'elle apporte. Il nous apparaît que dans le cas de séries non-indépendantes, la caractérisation de cette non-indépendance revêt une grande importance et la théorie des valeurs extrêmes ne solutionne pas cela. Cela dit, nous avons vu dans la partie précédente que l'utilisation d'une loi de Pareto Hybride - dont la densité était plus proche de l'histogramme empirique que celle d'une loi normale - fournissait des chocs moins extrêmes qu'une loi normale. Regardons si ce phénomène est là aussi présent :

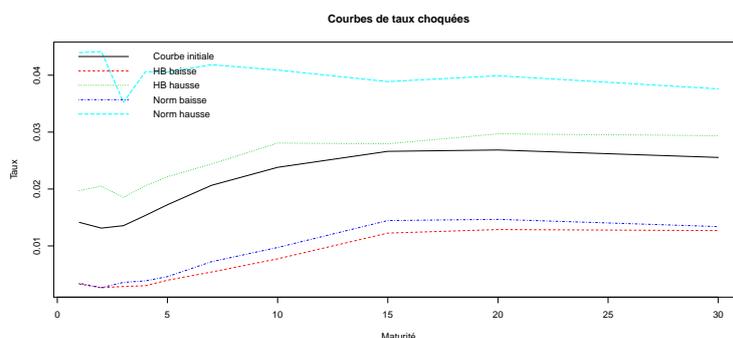


FIGURE 8.34 – Représentation des courbes de taux choquées à la hausse et à la baisse

Le résultat ici est légèrement différent : à la baisse nous sommes un peu plus extrêmes qu'une loi normale mais cela reste modéré, en revanche à la hausse l'hypothèse de normalité conduit à des chocs beaucoup plus élevés. Nous retrouvons ici le résultat précédent : le tableau 8.19 montre que le paramètre de queue à droite de la courbe ($\hat{\xi}_2$) est largement supérieur à 0. Cela implique que ce côté de la courbe a tendance à être leptokurtique, c'est-à-dire que les queues sont épaisses mais que du coup il y a aussi davantage de valeurs prises très proches de 0, d'où le fait d'être moins élevé après 252 itérations. Là encore, si nous comparons les log-vraisemblances des deux lois appliqués aux résidus des différentes maturités, celles pour la loi de Pareto Hybride sont plus élevées.

Jusqu'ici nous avons calculé des chocs maturité par maturité, maintenant nous allons voir comment faire pour étendre ceci au modèle interne, dans lequel il nous faut disposer de courbes de taux entières et non pas seulement de maturités indépendantes les unes des autres. Il nous est impossible de prendre uniquement des courbes pour lesquelles le taux en chaque maturité correspond au même quantile : nous n'aurions alors jamais de phénomène de renversement de courbe par exemple.

En réalité il est assez simple de simuler des courbes se rapprochant de vraies courbes de taux : il nous suffit de corrélérer entre eux les différents résidus. Rappelons que η_1 est la série des résidus observés pour la première maturité, η_2 la série des résidus observés pour la deuxième maturité, etc. Le graphique ci-dessous présente les valeurs prises au même instant avec η_1 en abscisse et η_2 en ordonnée, puis pour les couples (η_1, η_{10}) , (η_4, η_5) , (η_4, η_{10}) :

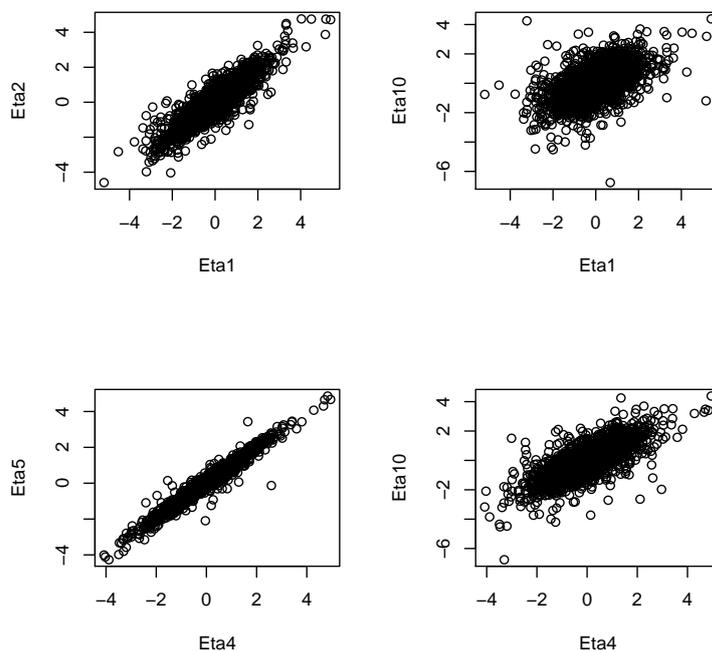


FIGURE 8.35 – Etude de la corrélation entre les résidus des différentes maturités

Il apparaît clairement une forte corrélation entre les résidus des différentes maturités, en effet plus les nuages de points sont proches de la première bissectrice et plus les séries sont corrélées. Cela se voit notamment lorsque les taux sont voisins (4 ans et 5 ans par exemple). A titre de comparaison, en cas d'indépendance les points se répartiraient sur l'ensemble du rectangle de façon égale, ce qui est loin d'être le cas ici. A ce stade, nous pouvons calculer les coefficients de corrélation entre les résidus, mais il n'est pas aisé de simuler des réalisations de

loi de Pareto Hybride de coefficients de corrélations fixés. De ce fait, il est difficile de reproduire cette corrélation lorsque nous projetons les valeurs futures des résidus. En revanche, il est très facile de simuler des lois normales de matrice de variance-covariance fixée, *via* la décomposition de CHOLESKY : supposons que nous voulions simuler un vecteur gaussien X de moyenne nulle et de matrice de variance-covariance Σ . Alors, si nous trouvons V telle que $\Sigma = VV'$, et que nous simulons Y un vecteur gaussien de même taille que X constitué de normales centrées réduites indépendantes, X et VY ont la même loi.

Nous faisons donc la chose suivante :

- Transformer les η_i en variables uniformes en posant $\eta'_i = \hat{F}_n^i(\eta_i)$, où $\hat{F}_n^i(x) = \sum_{j=1}^n \mathbb{1}_{\eta_{i,j} \leq x}$ est la fonction de répartition empirique de η_i
- Transformer les η'_i en variables gaussiennes centrées réduites en posant $\eta''_i = \Phi^{-1}(\eta'_i)$ où Φ est la fonction de répartition de la gaussienne centrée réduite
- Calculer les corrélations entre les η''_i pour obtenir la matrice de variance-covariance Σ

Nous disposons alors de la matrice Σ (symétrique), qui dans notre cas vaut :

$$\begin{pmatrix} 1,00 & 0,90 & 0,85 & 0,80 & 0,77 & 0,73 & 0,67 & 0,62 & 0,59 & 0,57 \\ & 1,00 & 0,96 & 0,91 & 0,89 & 0,86 & 0,79 & 0,75 & 0,71 & 0,69 \\ & & 1,00 & 0,93 & 0,92 & 0,91 & 0,85 & 0,80 & 0,77 & 0,75 \\ & & & 1,00 & 0,98 & 0,92 & 0,87 & 0,83 & 0,80 & 0,77 \\ & & & & 1,00 & 0,94 & 0,89 & 0,86 & 0,83 & 0,80 \\ & & & & & 1,00 & 0,97 & 0,94 & 0,92 & 0,90 \\ & & & & & & 1,00 & 0,97 & 0,95 & 0,92 \\ & & & & & & & 1,00 & 0,97 & 0,96 \\ & & & & & & & & 1,00 & 0,98 \\ & & & & & & & & & 1,00 \end{pmatrix}$$

Nous pouvons alors simuler les valeurs futures de nos résidus en appliquant le processus suivant :

- Simuler un vecteur gaussien X de taille 10 et de matrice de variance-covariance Σ
- Poser $U = \Phi(X)$, U est alors un vecteur dont les marginales sont uniformes non-indépendantes
- Poser $\eta = L^{-1}(U)$ où L est la fonction de répartition de la loi de Pareto Hybride

Cela nous permet d'obtenir les résidus nécessaires au passage d'un jour au suivant, en faisant cela 252 fois nous sommes donc à même d'obtenir une courbe des taux à un an.

Nous pouvons noter que cette procédure revient en réalité à utiliser une copule gaussienne.

8.5 Modélisation du facteur de risque lié aux variations de spread

Par convention, le spread de crédit est l'écart entre les rendements de titres obligataires émis par un emprunteur privé et les rendements de titres obligataires émis par un emprunteur public, supposé sans risque de défaut. Il correspond donc à la rémunération qu'exigent les investisseurs en compensation du risque global encouru par la détention du titre considéré. Il s'agit par conséquent d'une mesure du risque de crédit, assimilé à un risque de défaut, autrement dit de défaillance partielle ou totale de la contrepartie. Dès lors, l'analyse du spread de crédit se borne à une évaluation par le marché du risque de défaut de l'emprunteur. L'intérêt des firmes européennes pour les émissions obligataires est relativement récent du fait du rôle des banques dans le financement des entreprises. Toutefois, la création de la monnaie unique en 1999 a favorisé l'émergence d'un marché de la dette privée en euro. Majoritairement dominé à ses débuts par les obligations d'entreprises du secteur financier ayant de très bonnes notations (AAA et AA), ce marché s'est peu à peu ouvert à des émetteurs plus risqués, notamment ceux notés A et BBB. Prévoir et anticiper les mouvements des spreads de crédit sur les marchés obligataires est

donc important pour les détenteurs de titres obligataires dont les assureurs, pour la couverture de leurs positions face aux fortes variations de spreads.

Les spreads de crédit sont en général calculés sur le marché pour un ensemble d'obligations appartenant à une classe de notation donnée. La notation financière (ou *rating*) peut être considérée comme un indicateur de risque de défaut d'un émetteur. L'appréciation de la qualité de crédit d'un débiteur est généralement effectuée par les agences de notation qui évaluent la solvabilité d'une entreprise, d'une collectivité locale ou d'un Etat. La note "AAA" désigne la meilleure note, "BBB" renvoie à de mauvaises performances et constitue la limite inférieure des notes dites de catégorie d'investissement. En deçà de cette limite, les notes appartiennent à la catégorie dite spéculative.

Dans le cadre de Solvabilité II, nous nous intéresserons aux chocs de hausse des spreads que nous calculerons à l'aide de la loi de Pareto Hybride. Pour cela, nous avons à notre disposition des données journalières représentant l'historique des spreads des obligations d'entreprises notées AAA, AA, A et B, du 1er janvier 1999 au 6 février 2012 (3418 données).

Sur le graphique suivant, nous observons l'évolution de ces spreads en euro de 1999 à 2012 :

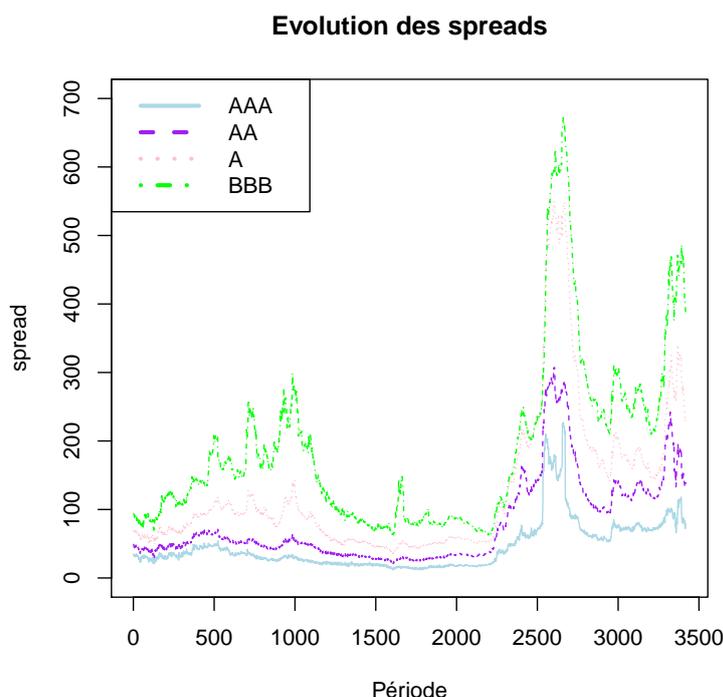


FIGURE 8.36 – Evolutions des spreads entre 1999 et 2012

Au vu du graphique, la hiérarchie des spreads semble respectée : plus la notation est basse, plus la prime de risque est importante. Les spreads correspondant aux obligations émises par les firmes AAA sont ainsi les plus bas car ces titres sont supposés être les moins risqués et les plus proches des obligations d'Etat. Le spread moyen pour les AAA est ainsi de 45,5 pb (points de base) sur la période considérée. A l'inverse, les obligations BBB ont un niveau de prime nettement supérieur avec un spread moyen de 188,1 pb, soit un spread 4 à 5 fois plus élevé en moyenne que les obligations AAA.

De plus, nous notons que la volatilité des spreads diminuent avec la qualité de la signature et augmentent en période de crise, en particulier en 2002-2003 (bulle internet) et en 2007-2008

(crise des subprimes). Cependant, cette augmentation du spread ne touche pas toutes les obligations de la même manière. En effet, au cours des périodes de tension sur les marchés financiers, les investisseurs ont tendance à rechercher de meilleures signatures et à délaisser les émetteurs de moins bonne qualité, ce qui a notamment abouti à des écarts de plus de 500 pb en 2007 entre les spreads AAA et BBB par exemple. C'est ce que l'on appelle le phénomène *flight to quality*.

Enfin, nous pouvons également remarquer l'asymétrie, le caractère leptokurtique et la non-stationnarité de la distribution des spreads de crédit sur la période étudiée ; des regroupements de volatilités sont également observés.

	AAA	AA	A	BBB
Min	12	21	35	62
Max	227	307	549	672
Moyenne	45,542	78,972	128,63	188,131
Ecart-type	36,066	62,588	107,404	127,732
Skewness	2,208	1,636	2,124	1,659
Kurtosis	8,979	5,201	7,546	5,664
Shapiro-Wilk (<i>p-value</i>)	$< 2,2 \times 10^{-16}$			
Phillips-Perron (<i>p-value</i>)	0,184	0,794	0,897	0,875

TABLE 8.21 – Propriétés des spreads (en pb)

Le test de Phillips-Perron montre l'existence d'une racine unitaire. Ceci est confirmé par l'observation des autocorrélogrammes qui présentent une décroissance très lente et linéaire des autocorrélations.

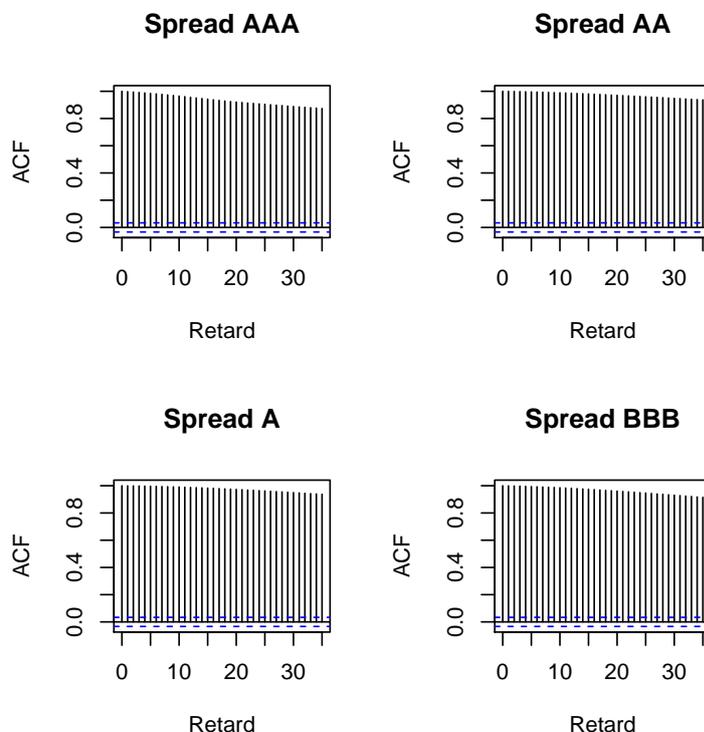


FIGURE 8.37 – Autocorrélogramme des séries de spreads

Nous décidons donc de considérer les log-rendements que nous notons : X_t^{AAA} , X_t^{AA} , X_t^A , X_t^{BBB} .

En plus de rendre les séries stationnaires, une modélisation des log-rendements permet de s'assurer que les valeurs des spreads obtenues restent positives.

Hormis la non-stationnarité, les séries des log-rendements conservent les propriétés que nous avons observées sur les spreads.

	AAA	AA	A	BBB
Excès de Kurtosis	8,915	5,059	4,624	31.977
Shapiro-Wilk (<i>p-value</i>)	$< 2,2 \times 10^{-16}$			
Phillips-Perron (<i>p-value</i>)	0,01	0,01	0,01	0,01

TABLE 8.22 – Propriétés des log-rendements

Les figures suivantes montrent l'évolution des log-rendements des spreads par notation. Elles mettent en évidence l'extrême volatilité de ces séries : la volatilité varie avec le temps et est plus importante pour les obligations les plus risquées. Nous pouvons aussi remarquer des regroupements de volatilité. Ce phénomène de *clustering* justifie le recours au modèle de type *GARCH*.

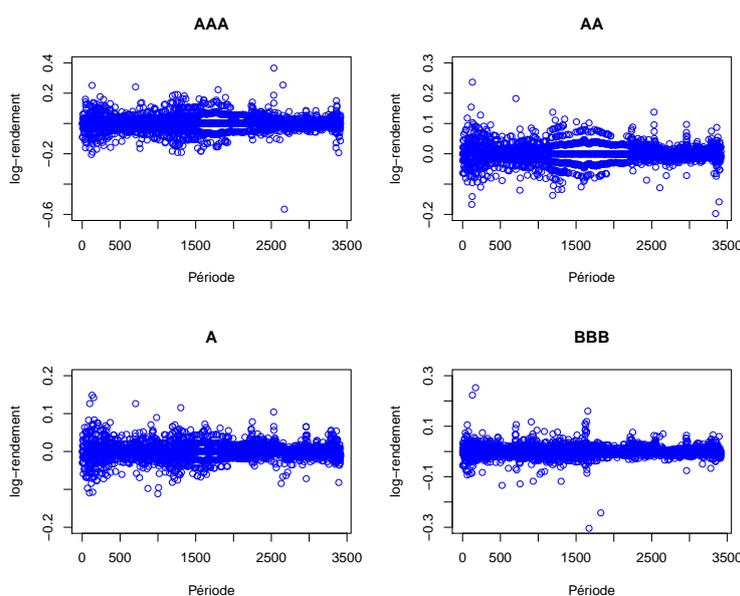


FIGURE 8.38 – Evolutions des log-rendements de spreads entre 1999 et 2012

Nous cherchons donc à modéliser ces séries avec un même modèle du type $ARMA(P, Q) - GARCH(p, q)$ qui pour rappel s'écrit sous la forme suivante :

$$\begin{cases} X_t = \mu + \sum_{i=1}^P a_i X_{t-i} + \sum_{j=1}^Q b_j \epsilon_{t-j} + \epsilon_t \\ \epsilon_t = \eta_t \sqrt{h_t} \\ h_t = \omega + \sum_{i=1}^q \alpha_i \epsilon_{t-i}^2 + \sum_{j=1}^p \beta_j h_{t-j} \end{cases}$$

Nous sélectionnons le modèle $ARMA(2, 1) - GARCH(1, 1)$ car il s'agit du modèle le plus simple permettant d'avoir un maximum de résidus indépendants et de paramètres significatifs.

Voici les paramètres auxquels nous aboutissons :

	AAA	AA	A	BBB
$\hat{\mu}$	$-2,413 \times 10^{-4}$	$-1,570 \times 10^{-4}$	$-4,452 \times 10^{-5}$	$-1,150 \times 10^{-4}$
\hat{a}_1	0,213	-0,226	0,772	0,828
\hat{a}_2	0,048	-0,026	0,173	0,057
\hat{b}_1	-0,589	-0,510	-0,890	-0,769
$\hat{\omega}$	$2,502 \times 10^{-5}$	$1,041 \times 10^{-5}$	$5,011 \times 10^{-6}$	$8,720 \times 10^{-7}$
$\hat{\alpha}$	0,093	0,094	0,067	0,039
$\hat{\beta}$	0,904	0,899	0,922	0,963

TABLE 8.23 – Valeurs des paramètres du modèle $ARMA(2,1) - GARCH(1,1)$ pour (X_t)

Mis à par les résidus au carré de la série X^{AAA} , l'application du test de Ljung-Box à η et à η^2 permet de ne pas rejeter l'hypothèse d'indépendance des résidus.

Test	η				η^2			
	AAA	AA	A	BBB	AAA	AA	A	BBB
LJUNG-BOX (m = 1 mois)	0,4972	0,2932	0,7950	0,2002	$< 2,2 \times 10^{-16}$	0,6565	0,1676	1
LJUNG-BOX (m = 1 an)	0,8870	0,0013	0,8781	0,6504	$8,826 \times 10^{-14}$	0,0013	0,9629	1
LJUNG-BOX (m = 2 ans)	0,8436	0,1820	0,9251	0,2381	0,148	1	0,1967	1

TABLE 8.24 – Tests d'indépendance des résidus des 4 modèles

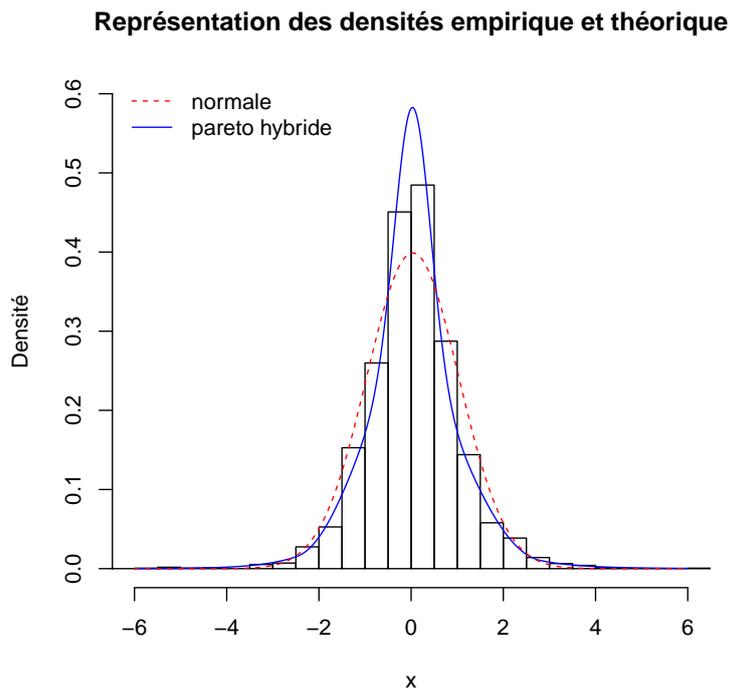
Les résidus n'étant pas gaussiens, nous choisissons de modéliser les résidus à l'aide d'une loi de Pareto Hybride à double seuil comme pour les taux. En effet, bien que nous soyons ici seulement intéressés par un choc à la hausse des spreads (et donc par la queue droite de distribution), nous remarquons de fortes valeurs négatives des résidus. Par exemple, pour les résidus correspondant au spread AA, nous comptons 7 valeurs en dessous de -5 avec des minimums allant jusqu'à -16 . Afin de calculer le quantile de la manière la plus précise possible, nous décidons de modéliser la queue gauche et la queue droite par des lois GPD.

Les paramètres estimés de la loi de Pareto Hybride sont les suivants :

	AAA	AA	A	BBB
μ_{norm}	0,018	0,274	0,107	-0,179
σ_{norm}	0,198	0,145	0,142	0,100
ξ_1	0,316	0,301	0,112	0,421
μ_1	-2,000	-1,000	-1,000	-1,000
σ_1	0,198	0,226	0,403	0,184
ϵ_1	4,949	3,768	3,810	2,955
ξ_2	0,560	0,292	0,269	0,129
μ_2	2,000	2,000	1,300	0,500
σ_2	0,111	0,173	0,279	0,532
ϵ_2	4,251	4,715	3,795	3,178

TABLE 8.25 – Paramètres de la Pareto Hybride

Nous pouvons évaluer l'adéquation de la loi gaussienne et de la loi de Pareto Hybride avec les résidus η en traçant des histogrammes. Par exemple, pour les spreads A, nous avons :


 FIGURE 8.39 – Densités théoriques et empiriques de η correspondant aux spreads A

Nous pouvons constater que la loi de Pareto Hybride a l'inconvénient de surestimer le nombre de valeurs en 0 contrairement à la loi normale où le phénomène inverse est observé. En revanche, elle permet de mieux modéliser les queues de distribution. Elle présente en effet des queues légèrement plus épaisses que la gaussienne. Ces remarques peuvent également s'appliquer pour les histogrammes des trois autres séries.

Nous pouvons confirmer nos observations avec des tests de Kolmogorov-Smirnov :

Loi testée	AAA	AA	A	BBB
loi normale	$2,7 \times 10^{-13}$	$4,8 \times 10^{-12}$	$2,8 \times 10^{-13}$	$< 2,2 \times 10^{-16}$
loi Pareto Hybride	0,29	0,36	0,31	0,24

 TABLE 8.26 – p -value des tests de Kolmogorov-Smirnov

Pour chaque série, la loi de Pareto Hybride est donc acceptée à 5 % alors que pour la loi normale, la p -value est très faible. L'introduction de la loi Pareto Hybride permet ainsi une meilleure modélisation des résidus.

Soit T la dernière valeur du spread connue. La valeur du spread dans un an S_{T+252} se déduit de :

$$\begin{aligned} S_{T+252} &= S_T e^{\sum_{i=1}^{52} \log\left(\frac{S_{T+i}}{S_{T+i-1}}\right)} \\ &= S_T e^{\sum_{i=1}^{52} X_{T+i}} \end{aligned}$$

Connaissant la S_T et en simulant pour tout i X_{T+i} , par itérations, nous pouvons ainsi en déduire une estimation de S_{T+252} . En recommençant cette manipulation 50 000 fois, nous

obtenons 50 000 valeurs de S_{T+252} , ce qui nous permet de calculer le choc à la hausse en prenant le quantile empirique à 99,5 %.

Pour chaque notation, nous calculons le choc obtenu avec la modélisation des résidus avec la loi de Pareto Hybride ainsi que celui obtenu en supposant la normalité des résidus :

Méthode	loi de Pareto Hybride	loi normale	QIS 5
AAA	1,41 %	3,62 %	0,90 %
AA	2,99 %	4,99 %	1,10 %
A	5,28 %	8,24 %	1,40 %
BBB	19,28 %	19,76 %	2,50 %

TABLE 8.27 – Chocs en fonction de la notation

Les chocs que nous obtenons en supposant une loi de Pareto Hybride pour les résidus sont plus faibles que ceux fournis par la loi normale. Cela n'est pas étonnant puisque sur la figure 8.39, nous observons que même si la loi de Pareto Hybride délivrait des queues plus épaisses, elle avait tendance à donner un nombre important de valeurs autour de zéro, contrairement à la loi normale.

Nous notons également que notre modélisation délivre des chocs supérieurs à ceux donnés dans le QIS 5. L'écart entre les chocs s'avère même considérable pour les spreads des obligations les plus risquées, c'est-à-dire celles notées A ou BBB. Pour les BBB, nous observons en effet un écart de l'ordre de 1 500 pb. Cependant, il est important de garder en tête que nos résultats dépendent fortement de la modélisation choisie et des hypothèses faites. Plusieurs facteurs montrent les limites de notre étude et nous incitent donc à considérer nos résultats avec prudence. En effet, comme dit précédemment, les modèles *ARMA – GARCH* que nous avons choisi ne respectent pas toujours les hypothèses de significativité des coefficients et d'indépendance des résidus, ce qui peut avoir des conséquences sur l'estimation des quantiles à un an. De plus, nous avons vu que la loi de Pareto Hybride utilisant une loi normale en partie centrale de la distribution bien qu'efficace pour modéliser la queue de distribution, était parfois moins performante que la loi gaussienne en milieu de distribution et pouvait conduire à une surestimation des valeurs proches de 0, ce qui peut aboutir à une sous-estimation du quantile à 99,5 %.

L'étude des résidus du modèle *ARMA – GARCH* nous a donc permis de modéliser chaque spread mais nous pouvons aller plus loin en analysant la structure de dépendance entre les différents spreads, mise en évidence dans la figure 8.36. En effet, les spreads semblent varier de façon similaire : en période agitée, ils sont très volatiles et ont tendance à augmenter tandis qu'en période calme, ils restent relativement bas. De plus, les courbes ne se croisent jamais.

Dans le cadre d'un modèle interne, il est crucial de tenir compte de ce phénomène et pour cela, nous décidons d'utiliser comme mesure de dépendance la corrélation linéaire. L'observation de la figure ci-dessous souligne en effet une forte corrélation entre les résidus η des différents spreads.

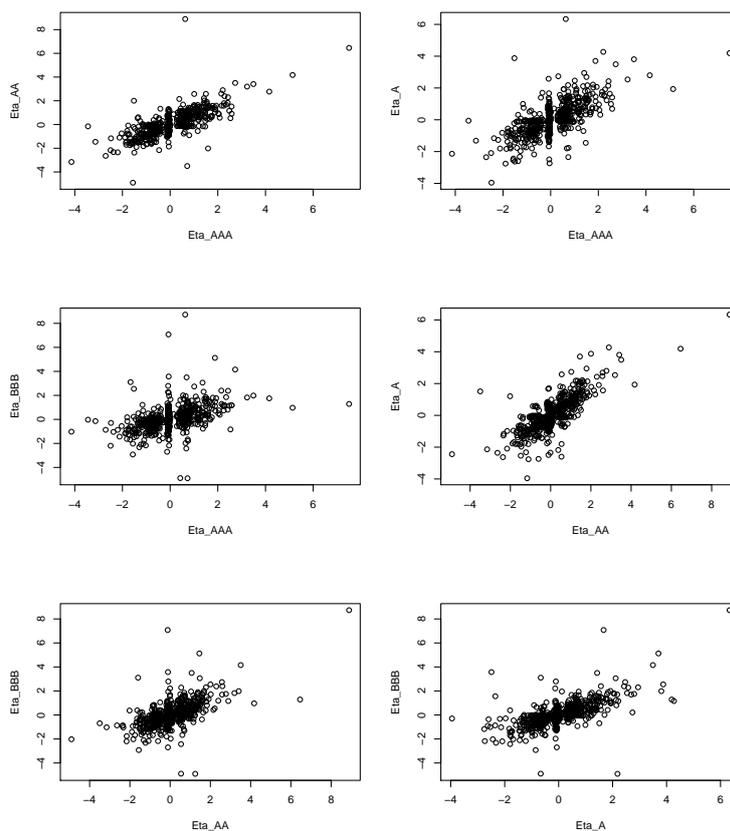


FIGURE 8.40 – Etude de la corrélation entre les résidus des différents spreads

Le choix de l'utilisation de la corrélation est également motivé par la simplicité du modèle à dépendance linéaire. De plus, il s'agit d'une pratique largement répandue en finance. Cependant, il est à noter que malgré son intérêt, elle souffre d'un certain nombre de défauts. Ainsi, par exemple :

- elle ne permet pas d'introduire de la dépendance dans les queues,
- elle peut s'annuler sans que les variables soient indépendantes,
- elle n'est pas invariante par transformations croissantes non linéaires.

Par conséquent, elle ne permet donc pas toujours de décrire de manière réaliste les structures de dépendance observées. Une étude plus approfondie sur ce sujet aurait ainsi nécessité l'étude de la théorie des copule qui propose des alternatives permettant de corriger ces problèmes.

En adoptant alors la même méthode que pour les taux, nous pouvons calculer la matrice de variance-covariance notée Σ du vecteur $(\eta''_{AAA}, \eta''_{AA}, \eta''_A, \eta''_{BBB})$:

$$\begin{pmatrix} 1,00 & 0,68 & 0,66 & 0,33 \\ & 1,00 & 0,65 & 0,32 \\ & & 1,00 & 0,49 \\ & & & 1,00 \end{pmatrix}$$

Il est alors aisé de générer à partir de cette matrice les résidus $(\eta^{AAA}, \eta^{AA}, \eta^A, \eta^{BBB})$ et d'obtenir des projections des spreads dans l'avenir.

8.6 Synthèse des résultats obtenus

Nous avons ainsi pu calculer des chocs sur les différents facteurs de risques que nous pourrions utiliser dans le cadre d'un modèle interne. Ceux-ci sont significativement distincts de ceux présents dans le QIS 5, les raisons sont les suivantes : d'une part les séries utilisées ne sont pas les mêmes, et ne couvrent pas non plus exactement la même période. Deux chocs calculés à deux dates différentes ne seront pas nécessairement proches car les conditions de marché peuvent être différentes. De plus, nous avons identifié un fort effet modèle. Une bonne calibration des chocs doit donc passer par un choix de modèle le plus pertinent possible. Dans le cadre de ce mémoire nous avons traité de modèles assez simples pour focaliser notre attention sur la théorie des valeurs extrêmes, mais ceux-ci peuvent être améliorés. Les critères de choix doivent inclure la proximité du modèle avec les observations, ainsi que l'indépendance des résidus afin de pouvoir appliquer correctement la théorie des valeurs extrêmes.

8.7 Etude de la structure de dépendance entre les facteurs de risque

Le chapitre précédent décrit la modélisation des lois marginales de chaque facteur de risque. Cependant, cette modélisation ne permet pas de refléter le risque réel auquel est exposée une compagnie d'assurance car nous omettons l'interaction entre les facteurs. Le montant du capital économique déterminé dans ces conditions pourrait alors être sous-estimé car on ne tiendrait pas compte par exemple de la corrélation entre les actions et les taux à la baisse. Il est donc nécessaire de modéliser les interactions entre les risques afin de pouvoir les agréger.

Le QIS 5 propose des matrices de corrélations (voir les figures 1.3 et 1.4) que nous pouvons utiliser pour corréler les facteurs de risque entre eux suivant une procédure similaire à celle que nous avons mis en œuvre pour les taux et les spreads. Comme nous l'avons déjà mentionné, cela revient à utiliser une copule gaussienne. Elle présente cependant l'inconvénient de ne pas tenir compte des dépendances extrêmes, ce qui peut nous amener à sous-évaluer les risques. Une solution pour remédier à ce problème pourrait être de considérer d'autres copules (comme la copule de Clayton, Gumbel, Frank etc.) comme cela a été fait par ARMEL dans [ARM10], mais cela sort du cadre de notre mémoire.

Conclusion

Notre mémoire propose des outils théoriques basés sur la théorie des valeurs extrêmes. Il met également en exergue des méthodes de mise en oeuvre pratique qui permettent une étude plus approfondie des queues de distribution et qui ont été ici appliquées à des données financières. La première étape de notre mémoire a nécessité l'examen du cadre réglementaire qui nous intéresse et l'étude d'éléments théoriques dans la littérature existante afin d'analyser l'avancement de la recherche sur le sujet. Cette analyse de l'existant a ensuite été suivie par une réflexion sur la manière de l'adapter à nos facteurs de risque et par une implémentation des différentes méthodes sous R. Nous avons ainsi pu les appliquer à nos données et analyser nos résultats.

Ce mémoire contribue à développer de nouvelles méthodologies, permettant de fournir une alternative concrète aux pratiques de marché. Dans un premier temps, nous avons montré que dans le cas où nous approchons une loi sur nos données, il existe un test (appelé GPD Test) permettant de se focaliser sur les queues de distribution. Celui-ci permet de compléter les tests d'adéquation classiques pour s'assurer que la modélisation est licite en partie centrale comme en queue de distribution. De plus, nous avons introduit une méthode de calcul de la Value-At-Risk (VaR) basée sur la théorie des valeurs extrêmes. Deux approches permettant le calcul des quantiles extrêmes ont été développées : l'approche par blocs basée sur la loi GEV et l'approche par excès basée sur la loi GPD. Nous avons également cherché à évaluer la précision du calcul des quantiles par la construction d'intervalles de confiance.

Nous avons pu appliquer nos méthodes à des données financières (valeurs immobilières, actions, taux, spreads). Deux schémas étaient a priori possibles :

- Une première solution aurait été de travailler nos données afin de les rendre stationnaires et y appliquer la théorie des valeurs extrêmes. En effet, le champ d'application de cette théorie, initialement restreint aux séries i.i.d., peut être étendu aux séries stationnaires.
- Nous pouvions également modéliser les séries de manière dynamique en incluant des paramètres de volatilité de type GARCH et appliquer nos méthodes sur les résidus i.i.d. issus de ce modèle, en les supposant non gaussiens.

Nous avons privilégié cette deuxième possibilité car elle permettait de nous rapprocher de la réalité dynamique des marchés financiers et nous avons alors choisi de modéliser les résidus en introduisant une loi de Pareto Hybride : elle permet d'avoir l'épaisseur de queue d'une loi GPD dans les queues de distribution tout en conservant les caractéristiques d'une loi gaussienne dans la partie centrale. Ceci nous a permis d'obtenir des projections dans l'avenir des indices étudiés, ce qui rend possible le calcul d'un quantile extrême. Une analyse détaillée de la structure de dépendance des facteurs de risque *via* l'utilisation de différentes copules aurait pu raffiner notre méthodologie mais faute de temps nous avons préféré utiliser une copule gaussienne afin de simplifier notre démarche.

Nous avons donc pu développer des implémentations intéressantes à partir de l'étude de la

théorie des valeurs extrêmes, dans le cadre de Solvabilité II : nous avons montré comment nos travaux pouvaient s'intégrer dans la construction d'un générateur de scénarios économiques et avons ainsi pu calculer les quantiles extrêmes des distributions à un an. Nous nous sommes essentiellement focalisés sur le risque de marché, mais les méthodes que nous avons étudiées peuvent également être appliquées afin d'analyser d'autres risques évoqués dans Solvabilité II.

Notre étude pourrait s'inscrire au sein d'une problématique plus large. En effet, nous nous sommes concentrés sur l'étude de la VaR qui est la mesure de risque préconisée dans le cadre réglementaire de Solvabilité II. Cette mesure souffre pourtant d'un défaut majeur : ce n'est pas une mesure cohérente. Une variation de la VaR, la Tail-VaR, permet de corriger ce problème, mais elle a l'inconvénient d'être délicate à calculer avec des méthodes conventionnelles car elle nécessite une connaissance précise des queues de distribution. Les méthodes évoquées dans notre mémoire pourraient cependant permettre un calcul aisé et pratique de la Tail-VaR. Nous pouvons donc voir en conclusion que la théorie des valeurs extrêmes est un outil riche qui mène à de multiples applications et qui peut répondre à de nombreuses problématiques du monde actuariel.

Bibliographie

- [ARM10] K. ARMEL. "Structure de dépendance des générateurs de scénarios économiques - Modélisation et simulation". Mémoire d'Actuariat, 2010.
- [BEN06] S. BENSEGHIR. "Calcul de la VaR selon l'approche historique et la théorie des valeurs extrêmes sur un fonds alternatif de Dexia Asset Management". Mémoire d'Actuariat, 2006.
- [BGTS04] J. BEIRLANT, Y. GOEGEBEUR, J. TEUGELS, and J. SEGERS. *Statistics of Extremes*. Wiley, 2004.
- [CH97] E. CASTILLO and A.S. HADI. "Fitting the Generalized Pareto Distribution to data". *Journal of the American Statistical Association*, 1997.
- [CLA08] P. CLAUSS. "Atelier de théorie des valeurs extrêmes - application en finance quantitative". Cours, 2008.
- [COL01] S. COLES. *An Introduction to Statistical Modeling of Extreme Values*. Springer, 2001.
- [DGNR08] J. DIEBOLT, A. GUILLOU, P. NAVEAU, and P. RIBEREAU. "Improving Probability-Weighted Moment methods for the Generalized Extreme Value distribution". *Revstat Statistical Journal*, 2008.
- [DV00] J. DANIELSSON and C.G. DE VRIES. "Value-at-Risk and extreme returns". *Annales d'Economie et de Statistique*, 2000.
- [EKM97] P. EMBRECHTS, C. KLÜPPELBERG, and T. MIKOSCH. *Modelling Extremal Events for Insurance and Finance*. Springer, 1997.
- [FP] C.A.T. FERRO and S. PEZZULLI. "Rao-Blackwellised estimators for parameters of extreme-value models".
- [GAR02] M. GARRIDO. *Modélisation des Evènements Rares et Estimation des Quantiles Extrêmes, Méthodes de Sélection de Modèles pour les Queues de Distribution*. PhD thesis, Université Joseph Fourier - Grenoble I, 2002.
- [GIR07] S. GIRARD. "Introduction à la statistique des valeurs extrêmes". 2007.
- [GMA84] C. GOURIEROUX, A. MONFORT, and A. MONFORT. "Pseudo Maximum Likelihood methods : Theory". *Econometrica*, 1984.
- [HH09] L. HOLDEN and O. HAUG. "A mixture model for unsupervised tail estimation". 2009.
- [HW87] J.R.M. HOSKING and J.R. WALLIS. "Parameter and quantile estimation for the Generalized Pareto Distribution". *Technometrics*, 1987.
- [HWW85] J.R.M. HOSKING, J.R. WALLIS, and E.F. WOOD. "Estimation of the Generalized Extreme Value distribution by the method of Probability-Weighted Moments". *Technometrics*, 1985.
- [LEA93a] M.R. LEADBETTER. "Extremes and local dependence in stationary sequences". *Z. Wahrscheinlichkeitstheorie verw. Gebiete*, 1993.

- [LEA93b] M.R. LEADBETTER. "On high level exceedance modeling and tail inference". *Journal of Statistical Planning and Inference*, 1993.
- [MF00] A.J. MCNEIL and R. FREY. "Estimation of tail-related risk measures for heteroscedastic financial time series : an extreme value approach". *Journal of Empirical Finance*, 2000.
- [PIC75] J. PICKANDS. "Statistical inference using extreme order statistics". *The Annals of Statistics*, 1975.
- [RON02] T. RONCALLI. "Théorie des valeurs extrêmes - modélisation des évènements rares pour la gestion des risques". Cours, 2002.
- [RSF08] C.Y. ROBERT, J. SEGERS, and C.A.T. FERRO. "A sliding blocks estimator for the extremal index". *Electronic Journal of Statistics*, 2008.
- [SMI84] R.L. SMITH. "Threshold methods for sample extremes". *Statistical Extremes and Applications*, 1984.
- [TSA05] R.S. TSAY. *Analysis of Financial Time Series*. Wiley, 2005.

Annexe

A Calcul des moments pondérés

A.1 Loi des valeurs extrêmes

Nous souhaitons calculer la valeur de $\omega_r = \mathbb{E}(MH_{\xi,\mu,\sigma}^r(M))$, où M suit une GEV de fonction de répartition $H_{\xi,\mu,\sigma}$, $\forall r \geq 0$, et $\xi \geq 0$.

Pour $0 < \xi < 1$:

$$\begin{aligned}\omega_r &= \int_{\mathbb{R}} m H_{\xi,\mu,\sigma}^r(m) dH_{\xi,\mu,\sigma}(m) \\ &= \int_0^1 H_{\xi,\mu,\sigma}^{-1}(y) y^r dy \\ &= \int_0^1 \left(\mu + \frac{\sigma}{\xi} ((-\ln(y))^{-\xi} - 1) \right) y^r dy \\ &= \int_0^1 \left(\mu - \frac{\sigma}{\xi} \right) y^r dy + \frac{\sigma}{\xi} \int_0^1 y^r (-\ln(y))^{-\xi} dy \\ &= \frac{1}{r+1} \left(\mu - \frac{\sigma}{\xi} \right) + \frac{\sigma}{\xi} \int_{-\infty}^0 e^{rz} (-z)^{-\xi} e^z dz \\ &= \frac{1}{r+1} \left(\mu - \frac{\sigma}{\xi} \right) + \frac{\sigma}{\xi} \int_{+\infty}^0 e^{-v} \left(\frac{v}{r+1} \right)^{-\xi} \frac{1}{r+1} dv \\ &= \frac{1}{r+1} \left(\mu - \frac{\sigma}{\xi} \right) + \frac{\sigma}{\xi} \frac{1}{r+1} (r+1)^\xi \Gamma(1-\xi)\end{aligned}$$

D'où : $\omega_r = \frac{1}{r+1} \left(\mu - \frac{\sigma}{\xi} (1 - (r+1)^\xi \Gamma(1-\xi)) \right)$. Et pour $\xi = 0$:

$$\begin{aligned}
\omega_r &= \int_{\mathbb{R}} m H_{0,\mu,\sigma}^r(m) dH_{0,\mu,\sigma}(m) \\
&= \int_0^1 H_{0,\mu,\sigma}^{-1}(y) y^r dy \\
&= \int_0^1 (\mu - \sigma \ln(-\ln(y))) y^r dy \\
&= \int_0^1 \mu y^r dy - \sigma \int_0^1 \ln(-\ln(y)) y^r dy \\
&= \frac{\mu}{r+1} + \sigma \int_{+\infty}^0 \ln(z) e^{-(r+1)z} dz \\
&= \frac{\mu}{r+1} - \frac{\sigma}{r+1} \int_0^{+\infty} \ln\left(\frac{u}{r+1}\right) e^{-u} du \\
&= \frac{\mu}{r+1} - \frac{\sigma}{r+1} \left(\int_0^{+\infty} \ln(u) e^{-u} du - \ln(r+1) \int_0^{+\infty} e^{-u} du \right) \\
&= \frac{\mu}{r+1} - \frac{\sigma}{r+1} (\gamma - \ln(r+1)) \\
&= \frac{1}{r+1} (\mu - \sigma(\gamma - \ln(r+1)))
\end{aligned}$$

Et donc : $\omega_r = \frac{1}{r+1} (\mu - \sigma(\gamma - \ln(r+1)))$ pour $\xi = 0$.

A.2 Loi de Pareto généralisée

$$w_r(\xi, \sigma) = \int_{-\frac{\sigma}{\xi}}^{+\infty} x \bar{G}_{\xi,\sigma}^r(x) dG_{\xi,\sigma}(x)$$

Le changement de variable $y = \bar{G}_{\xi,\sigma}^r(x)$ est licite car $\bar{G}_{\xi,\sigma}^r$ est bijective sur l'ensemble d'intégration. Nous obtenons alors :

$$\begin{aligned}
w_r(\xi, \sigma) &= - \int_1^0 y^r \bar{G}_{\xi,\sigma}^{-1}(y) dy \\
&= \int_0^1 \frac{\sigma}{\xi} y^r (y^{-\xi} - 1) dy \\
&= \frac{\sigma}{\xi} \int_0^1 (y^{r-\xi} - y^r) dy \\
&= \frac{\sigma}{\xi} \left[\frac{y^{r-\xi+1}}{r-\xi+1} - \frac{y^{r+1}}{r+1} \right]_0^1 \\
&= \frac{\sigma}{\xi} \left(\frac{1}{r-\xi+1} - \frac{1}{r+1} \right) \\
&= \frac{\sigma}{(r+1)(r+1-\xi)}
\end{aligned}$$

Remarquons que l'intégrale sur la troisième ligne n'est finie que pour $\xi - r < 1$, cela nous donne donc une condition sur ξ en fonction des moments que nous souhaitons utiliser. Nous pouvons procéder de même pour $\xi < 0$:

$$w_r(\xi, \sigma) = \int_{-\infty}^{-\frac{\sigma}{\xi}} x \bar{G}_{\xi, \sigma}^r(x) dG_{\xi, \sigma}(x)$$

Le changement de variable conduit à la même intégrale que ci-dessus. Le résultat est donc identique. Enfin, pour $\xi = 0$:

$$\begin{aligned} w_r(0, \sigma) &= \int_0^{+\infty} x \bar{G}_{0, \sigma}^r(x) dG_{0, \sigma}(x) \\ &= \int_0^{+\infty} x e^{-(r+1)\frac{x}{\sigma}} dx \\ &= \left[-\frac{\sigma x}{r+1} e^{-(r+1)\frac{x}{\sigma}} \right]_0^{+\infty} + \int_0^{+\infty} \frac{\sigma}{r+1} e^{-(r+1)\frac{x}{\sigma}} dx \\ &= \left[-\frac{\sigma}{(r+1)^2} e^{-(r+1)\frac{x}{\sigma}} \right]_0^{+\infty} \\ &= \frac{\sigma}{(r+1)^2} \end{aligned}$$

Puisque nous avons utilisé les moments d'ordre 0 et 1, cette méthode n'est valable que pour $\xi < 1$ à cause de la condition $\xi < 1 + r$.

B Estimation de quantiles de lois connues

B.1 Loi normale centrée réduite

	Quantile moyen	Ecart quadratique	Seuil moyen	Ecart-type seuil
Empirique	1,64	0,0047		
EMV m=30	1,66	0,0069		
PWM m=30	1,63	0,0066		
GPWM m=30	1,61	0,0080		
RB EMV m=30	1,63	0,0042		
RB PWM m=30	1,63	0,0042		
RB GPWM m=30	1,61	0,0054		
EMV m=50	1,67	0,0157		
PWM m=50	1,61	0,0143		
GPWM m=50	1,59	0,0204		
RB EMV m=50	1,66	0,0051		
RB PWM m=50	1,60	0,0062		
RB GPWM m=50	1,58	0,0079		
ME EMV	1,69	0,0107	0,10	0,54
ME PWM	1,65	0,0062	0,10	0,54
Hill EMV	1,65	0,0042	1,37	0,20
Hill PWM	1,64	0,0047	1,37	0,20
Cunn1 EMV	1,67	0,0044	0,53	0,14
Cunn1 PWM	1,67	0,0039	0,53	0,14
Cunn2 EMV	1,66	0,0041	0,70	0,20
Cunn2 PWM	1,66	0,0038	0,70	0,20
Pick EMV	1,67	0,0046	0,39	0,31
Pick PWM	1,66	0,0041	0,33	0,30

TABLE 8.28 – Quantile à 95 % pour un échantillon Normale centré réduit de taille 1000

	Quantile moyen	Ecart quadratique	Seuil moyen	Ecart-type seuil
Empirique	2,53	0,024		
EMV m=30	2,55	0,016		
PWM m=30	2,59	0,017		
GPWM m=30	2,50	0,026		
RB EMV m=30	2,55	0,015		
RB PWM m=30	2,59	0,016		
RB GPWM m=30	2,50	0,023		
EMV m=50	2,55	0,019		
PWM m=50	2,59	0,019		
GPWM m=50	2,49	0,025		
RB EMV m=50	2,56	0,016		
RB PWM m=50	2,59	0,016		
RB GPWM m=50	2,49	0,023		
ME EMV	2,53	0,071	0,10	0,54
ME PWM	2,34	0,125	0,10	0,54
Hill EMV	2,56	0,016	1,37	0,20
Hill PWM	2,59	0,017	1,37	0,20
Cunn1 EMV	2,58	0,019	0,53	0,14
Cunn1 PWM	2,51	0,026	0,53	0,14
Cunn2 EMV	2,58	0,017	0,70	0,20
Cunn2 PWM	2,53	0,024	0,70	0,20
Pick EMV	2,56	0,019	0,39	0,31
Pick PWM	2,51	0,028	0,33	0,30

TABLE 8.29 – Quantile à 99,5 % pour un échantillon Normale centré réduit de taille 1000

	Quantile moyen	Ecart quadratique	Seuil moyen	Ecart-type seuil
Empirique	1,643	0,00049		
EMV m=50	1,647	0,00123		
PWM m=50	1,643	0,00130		
GPWM m=50	1,642	0,00141		
RB EMV m=50	1,645	0,00043		
RB PWM m=50	1,641	0,00048		
RB GPWM m=50	1,640	0,00048		
EMV m=100	1,658	0,00613		
PWM m=100	1,641	0,00628		
GPWM m=100	1,617	0,00994		
RB EMV m=100	1,657	0,00146		
RB PWM m=100	1,639	0,00174		
RB GPWM m=100	1,616	0,00243		
ME EMV	1,657	0,00098	0,69	0,27
ME PWM	1,658	0,00063	0,69	0,27
Hill EMV	1,640	0,00043	1,39	0,11
Hill PWM	1,642	0,00045	1,39	0,11
Cunn1 EMV	1,666	0,00106	0,54	0,14
Cunn1 PWM	1,666	0,00088	0,54	0,14
Cunn2 EMV	1,655	0,00062	0,73	0,17
Cunn2 PWM	1,660	0,00066	0,73	0,17
Pick EMV	1,648	0,00064	0,90	0,33
Pick PWM	1,654	0,00065	0,71	0,31

TABLE 8.30 – Quantile à 95 % pour un échantillon Normale centré réduit de taille 10000

	Quantile moyen	Ecart quadratique	Seuil moyen	Ecart-type seuil
Empirique	2,570	0,00250		
EMV m=50	2,578	0,00200		
PWM m=50	2,580	0,00203		
GPWM m=50	2,571	0,00219		
RB EMV m=50	2,578	0,00168		
RB PWM m=50	2,580	0,00175		
RB GPWM m=50	2,571	0,00192		
EMV m=100	2,576	0,00252		
PWM m=100	2,575	0,00252		
GPWM m=100	2,565	0,00260		
RB EMV m=100	2,575	0,00197		
RB PWM m=100	2,575	0,00200		
RB GPWM m=100	2,564	0,00207		
ME EMV	2,611	0,00562	0,68	0,28
ME PWM	2,521	0,00841	0,68	0,28
Hill EMV	2,587	0,00177	1,39	0,10
Hill PWM	2,582	0,00181	1,39	0,10
Cunn1 EMV	2,624	0,00509	0,54	0,14
Cunn1 PWM	2,492	0,00980	0,54	0,14
Cunn2 EMV	2,614	0,00356	0,73	0,17
Cunn2 PWM	2,528	0,00492	0,73	0,17
Pick EMV	2,589	0,00322	0,90	0,33
Pick PWM	2,541	0,00569	0,71	0,31

TABLE 8.31 – Quantile à 99,5 % pour un échantillon Normale centré réduit de taille 10000

	Taille moyenne	Taux de couverture	Borne inf moyenne	Borne sup moyenne
EMV (PF)	0,328	93,0 %	1,489	1,817
EMV (Fisher obs)	0,320	91,8 %	1,501	1,820
EMV (Fisher théo)	0,318	91,6 %	1,501	1,819
PWM	0,329	98,2 %	1,469	1,798
GPWM	0,415	95,6 %	1,403	1,818
Cunn1 EMV (Fisher théo)	0,209	88,9 %	1,566	1,775
Cunn1 PWM	0,202	88,9 %	1,566	1,768
Cunn1 EMV (Fisher obs)	0,211	89,1 %	1,565	1,775
Cunn2 EMV (Fisher théo)	0,210	89,5 %	1,565	1,775
Cunn2 PWM	0,202	89,7 %	1,565	1,767
Cunn2 EMV (Fisher obs)	0,211	90,1 %	1,564	1,775
Pick EMV (Fisher théo)	0,205	87,3 %	1,565	1,770
Pick PWM	0,203	88,5 %	1,563	1,766
Pick EMV obs	0,205	86,9 %	1,563	1,768

TABLE 8.32 – Propriétés des intervalles de confiance à 95 % pour le quantile à 95 % d'un échantillon Normal centré réduit de taille 1000

	Taille moyenne	Taux de couverture	Borne inf moyenne	Borne sup moyenne
EMV (PF)	0,272	87,1 %	1,520	1,792
EMV (Fisher obs)	0,268	86,8 %	1,526	1,795
EMV (Fisher théo)	0,267	86,8 %	1,527	1,794
PWM	0,276	96,6 %	1,495	1,771
GPWM	0,348	93,2 %	1,436	1,784
Cunn1 EMV (Fisher théo)	0,175	80,1 %	1,581	1,757
Cunn1 PWM	0,170	81,2 %	1,582	1,751
Cunn1 EMV (Fisher obs)	0,177	80,5 %	1,582	1,759
Cunn2 EMV (Fisher théo)	0,176	82,5 %	1,582	1,758
Cunn2 PWM	0,170	82,7 %	1,581	1,751
Cunn2 EMV (Fisher obs)	0,177	83,1 %	1,581	1,758
Pick EMV (Fisher théo)	0,172	79,5 %	1,581	1,753
Pick PWM	0,170	82,5 %	1,580	1,750
Pick EMV (Fisher obs)	0,172	79,5 %	1,591	1,753

TABLE 8.33 – Propriétés des intervalles de confiance à 90 % pour le quantile à 95 % d'un échantillon Normal centré réduit de taille 1000

	Taille moyenne	Taux de couverture	Borne inf moyenne	Borne sup moyenne
EMV (Fisher obs)	0,468	89,9 %	2,319	2,787
EMV (Fisher théo)	0,464	89,8 %	2,321	2,785
EMV (PF)	0,543	91,8 %	2,348	2,891
PWM	0,471	92,1 %	2,340	2,811
Hill EMV (Fisher théo)	0,413	87,3 %	2,347	2,760
Hill PWM	0,514	94,6 %	2,330	2,843
Hill EMV (Fisher obs)	0,425	88,1 %	2,343	2,768
Cunn1 EMV (Fisher théo)	0,372	81,9 %	2,396	2,768
Cunn1 PWM	0,507	85,7 %	2,253	2,759
Cunn1 EMV (Fisher obs)	0,380	83,1 %	2,394	2,774
Cunn2 EMV (Fisher théo)	0,374	83,9 %	2,398	2,772
Cunn2 PWM	0,508	86,1 %	2,254	2,762
Cunn2 EMV (Fisher obs)	0,385	84,9 %	2,393	2,778
Pick EMV (Fisher théo)	0,370	83,7 %	2,386	2,757
Pick PWM	0,533	86,7 %	2,242	2,775
Pick EMV (Fisher obs)	0,368	82,1 %	2,388	2,756

TABLE 8.34 – Propriétés des intervalles de confiance à 95 % pour le quantile à 99,5 % d'un échantillon Normal centré réduit de taille 1000

	Taille moyenne	Taux de couverture	Borne inf moyenne	Borne sup moyenne
EMV (Fisher obs)	0,393	83,0 %	2,356	2,749
EMV (Fisher théo)	0,389	82,4 %	2,358	2,747
EMV (PF)	0,436	87,6 %	2,378	2,814
PWM	0,396	87,8 %	2,377	2,774
Hill EMV (Fisher théo)	0,347	80,9 %	2,380	2,727
Hill PWM	0,431	90,1 %	2,371	2,082
Hill EMV (Fisher obs)	0,357	82,0 %	2,377	2,733

TABLE 8.35 – Propriétés des intervalles de confiance à 90 % pour le quantile à 99,5 % d'un échantillon Normal centré réduit de taille 1000

	Taille moyenne	Taux de couverture	Borne inf moyenne	Borne sup moyenne
EMV (PF)	0,1409	95,8 %	1,5707	1,7116
EMV (Fisher obs)	0,1390	95,4 %	1,5777	1,7167
EMV (Fisher théo)	0,1389	95,2 %	1,5777	1,7166
PWM	0,1452	99,8 %	1,5674	1,7126
GPWM	0,1475	95,2 %	1,5682	1,7157
ME EMV (Fisher théo)	0,0657	76,2 %	1,624	1,690
ME PWM	0,0639	80,0 %	1,626	1,690
ME EMV (Fisher obs)	0,0652	76,4 %	1,625	1,690
Pick EMV (Fisher théo)	0,0589	78,6 %	1,617	1,676
Pick PWM	0,0634	77,6 %	1,622	1,685
Pick EMV (Fisher obs)	0,0586	78,0 %	1,615	1,674

TABLE 8.36 – Propriétés des intervalles de confiance à 95 % pour le quantile à 95 % d'un échantillon Normale centré réduit de taille 10000

	Taille moyenne	Taux de couverture	Borne inf moyenne	Borne sup moyenne
EMV (PF)	0,1176	90,8 %	1,584	1,702
EMV (Fisher obs)	0,1166	91,2 %	1,589	1,706
EMV (Fisher théo)	0,1165	90,6 %	1,589	1,705
PWM	0,1218	99,2 %	1,579	1,701
GPWM	0,1238	90,4 %	1,580	1,704
ME EMV (Fisher théo)	0,0551	65,4 %	1,631	1,686
ME PWM	0,0537	71,0 %	1,631	1,685
ME EMV (Fisher obs)	0,0546	65,5 %	1,629	1,684
Pick EMV (Fisher théo)	0,0494	69,6 %	1,620	1,670
Pick PWM	0,0532	69,8 %	1,627	1,680
Pick EMV (Fisher obs)	0,0492	69,7 %	1,622	1,671

TABLE 8.37 – Propriétés des intervalles de confiance à 90 % pour le quantile à 95 % d'un échantillon Normale centré réduit de taille 10000

	Taille moyenne	Taux de couverture	Borne inf moyenne	Borne sup moyenne
EMV (Fisher obs)	0,166	93,8 %	2,495	2,661
EMV (Fisher théo)	0,166	93,8 %	2,495	2,661
PWM	0,167	95,6 %	2,496	2,663
Hill EMV (Fisher théo)	0,140	91,2 %	2,517	2,657
Hill PWM	0,149	92,8 %	2,508	2,657
Hill EMV (Fisher obs)	0,140	91,2 %	2,517	2,656

TABLE 8.38 – Propriétés des intervalles de confiance à 95 % pour le quantile à 99,5 % d'un échantillon Normale centré réduit de taille 10000

	Taille moyenne	Taux de couverture	Borne inf moyenne	Borne sup moyenne
EMV (Fisher obs)	0,139	86,2 %	2,509	2,648
EMV (Fisher théo)	0,139	86,2 %	2,509	2,648
PWM	0,140	90,2 %	2,509	2,649
Hill EMV (Fisher théo)	0,117	84,8 %	2,528	2,645
Hill PWM	0,125	85,6 %	2,520	2,645
Hill EMV (Fisher obs)	0,117	84,8 %	2,528	2,645

TABLE 8.39 – Propriétés des intervalles de confiance à 90 % pour le quantile à 99,5 % d'un échantillon Normale centré réduit de taille 10000

B.2 Loi de Student à 2 degrés de liberté

	Quantile moyen	Ecart quadratique	Seuil moyen	Ecart-type seuil
Empirique	2,91	0,058		
EMV m=30	2,94	0,105		
PWM m=30	2,93	0,100		
GPWM m=30	2,81	0,102		
RB EMV m=30	2,88	0,063		
RB PWM m=30	2,94	0,064		
RB GPWM m=30	2,82	0,062		
EMV m=50	3,04	0,200		
PWM m=50	2,77	0,219		
GPWM m=50	2,68	0,256		
RB EMV m=50	3,06	0,085		
RB PWM m=50	2,78	0,111		
RB GPWM m=50	2,70	0,101		
ME EMV	3,01	0,061	0,78	0,63
ME PWM	3,01	0,064	0,78	0,63
Hill EMV	2,93	0,060	1,64	0,54
Hill PWM	2,91	0,054	1,64	0,54
Cunn1 EMV	3,01	0,054	0,64	0,19
Cunn1 PWM	3,03	0,059	0,64	0,19
Cunn2 EMV	2,98	0,052	0,86	0,26
Cunn2 PWM	2,98	0,058	0,86	0,26
Pick EMV	2,99	0,055	0,60	0,41
Pick PWM	2,98	0,059	0,67	0,48

TABLE 8.40 – Quantile à 95 % pour un échantillon Student à 2 degrés de liberté de taille 1000

	Quantile moyen	Ecart quadratique	Seuil moyen	Ecart-type seuil
Empirique	9,59	4,97		
EMV m=30	9,90	4,11		
PWM m=30	9,73	3,67		
GPWM m=30	8,83	4,05		
RB EMV m=30	9,81	3,62		
RB PWM m=30	9,73	3,42		
RB GPWM m=30	8,84	3,75		
EMV m=50	9,84	4,01		
PWM m=50	9,98	4,41		
GPWM m=50	8,81	4,11		
RB EMV m=50	9,86	5,46		
RB PWM m=50	9,95	4,06		
RB GPWM m=50	8,78	3,80		
ME EMV	9,31	3,10	0,78	0,63
ME PWM	9,52	3,72	0,78	0,63
Hill EMV	9,75	3,28	1,64	0,54
Hill PWM	9,69	3,58	1,64	0,54
Cunn1 EMV	9,37	2,89	0,64	0,19
Cunn1 PWM	9,74	4,40	0,64	0,19
Cunn2 EMV	9,53	2,93	0,86	0,26
Cunn2 PWM	9,77	4,19	0,86	0,26
Pick EMV	9,51	3,19	0,60	0,41
Pick PWM	9,67	3,61	0,67	0,48

TABLE 8.41 – Quantile à 99,5 % pour un échantillon Student à 2 degrés de liberté de taille 1000

	Quantile moyen	Ecart quadratique	Seuil moyen	Ecart-type seuil
Empirique	2,92	0,0052		
EMV m=50	2,92	0,0194		
PWM m=50	2,90	0,0274		
GPWM m=50	2,91	0,0221		
RB EMV m=50	2,92	0,0054		
RB PWM m=50	2,90	0,0125		
RB GPWM m=50	2,90	0,0065		
EMV m=100	3,01	0,1040		
PWM m=100	2,79	0,2900		
GPWM m=100	2,77	0,1942		
RB EMV m=100	2,99	0,0195		
RB PWM m=100	2,76	0,2061		
RB GPWM m=100	2,73	0,0984		
ME EMV	2,95	0,0071	1,36	0,66
ME PWM	2,95	0,0082	1,36	0,66
Hill EMV	2,93	0,0149	2,76	1,06
Hill PWM	2,94	0,0180	2,76	1,06
Cunn1 EMV	3,02	0,0156	0,64	0,19
Cunn1 PWM	3,06	0,0280	0,64	0,19
Cunn2 EMV	2,98	0,0088	0,91	0,25
Cunn2 PWM	3,00	0,0127	0,91	0,25
Pick EMV	2,94	0,0060	1,22	0,42
Pick PWM	2,94	0,0071	1,31	0,49

TABLE 8.42 – Quantile à 95 % pour un échantillon Student à 2 degrés de liberté de taille 10000

	Quantile moyen	Ecart quadratique	Seuil moyen	Ecart-type seuil
Empirique	9,85	0,50		
EMV m=50	9,90	0,38		
PWM m=50	9,89	0,39		
GPWM m=50	9,83	0,40		
RB EMV m=50	9,90	0,35		
RB PWM m=50	9,89	0,36		
RB GPWM m=50	9,83	0,37		
EMV m=100	9,87	0,52		
PWM m=100	9,97	0,57		
GPWM m=100	9,75	0,52		
RB EMV m=100	9,86	0,37		
RB PWM m=100	9,96	0,48		
RB GPWM m=100	9,75	0,41		
ME EMV	9,73	0,36	1,36	0,66
ME PWM	9,93	0,44	1,36	0,66
Hill EMV	9,88	0,35	2,76	1,06
Hill PWM	9,87	0,46	2,76	1,06
Cunn1 EMV	9,41	0,53	0,64	0,19
Cunn1 PWM	9,94	0,69	0,64	0,19
Cunn2 EMV	9,61	0,39	0,91	0,25
Cunn2 PWM	10,03	0,66	0,91	0,25
Pick EMV	9,80	0,35	1,22	0,42
Pick PWM	9,99	0,52	1,31	0,49

TABLE 8.43 – Quantile à 99,5 % pour un échantillon Student à 2 degrés de liberté de taille 10000

	Taille moyenne	Taux de couverture	Borne inf moyenne	Borne sup moyenne
EMV (PF)	1,203	94,4 %	2,384	3,587
EMV (Fisher obs)	1,173	93,2 %	2,355	3,528
EMV (Fisher théo)	1,155	92,2 %	2,364	3,519
PWM	1,262	99,2 %	2,396	3,658
GPWM	1,557	94,9 %	2,031	3,588
Cunn1 EMV (Fisher théo)	0,724	89,8 %	2,650	3,374
Cunn1 PWM	0,772	90,8 %	2,638	3,410
Cunn1 EMV (Fisher obs)	0,724	89,8 %	2,650	3,374
Cunn1 EMV (PF)	0,734	87,8 %	2,686	3,419
Cunn2 EMV (Fisher théo)	0,724	90,4 %	2,649	3,373
Cunn2 PWM	0,775	89,9 %	2,637	3,412
Cunn2 EMV (Fisher obs)	0,724	90,2 %	2,649	3,373
Cunn2 EMV (PF)	0,734	88,2 %	2,685	3,419
Pick EMV (Fisher théo)	0,715	87,8 %	2,631	3,346
Pick PWM	0,744	87,3 %	2,608	3,352
Pick EMV (Fisher obs)	0,715	87,6 %	2,631	3,346
Pick EMV (PF)	0,726	87,6 %	2,667	3,393

TABLE 8.44 – Propriétés des intervalles de confiance à 95 % pour le quantile à 95 % d'un échantillon Student à 2 degrés de liberté de taille 1000

	Taille moyenne	Taux de couverture	Borne inf moyenne	Borne sup moyenne
EMV (PF)	1,000	87,6 %	2,474	3,474
EMV (Fisher obs)	0,984	87,4 %	2,450	3,434
EMV (Fisher théo)	0,969	86,2 %	2,457	3,426
PWM	1,053	97,3 %	2,477	3,530
GPWM	1,307	91,3 %	2,156	3,463
Cunn1 EMV (Fisher théo)	0,607	83,0 %	2,708	3,316
Cunn1 PWM	0,648	83,5 %	2,700	3,348
Cunn1 EMV (Fisher obs)	0,607	83,0 %	2,708	3,316
Cunn1 EMV (PF)	0,613	80,4 %	2,734	3,347
Cunn2 EMV (Fisher théo)	0,608	82,4 %	2,708	3,315
Cunn2 PWM	0,651	82,9 %	2,700	3,350
Cunn2 EMV (Fisher obs)	0,607	82,4 %	2,708	3,315
Cunn2 EMV (PF)	0,613	80,4 %	2,733	3,346
Pick EMV (Fisher théo)	0,600	80,8 %	2,688	3,289
Pick PWM	0,624	78,8 %	2,668	3,292
Pick EMV (Fisher obs)	0,600	80,8 %	2,689	3,289
Pick EMV (PF)	0,606	77,4 %	2,715	3,320

TABLE 8.45 – Propriétés des intervalles de confiance à 90 % pour le quantile à 95 % d'un échantillon Student à 2 degrés de liberté de taille 1000

	Taille moyenne	Taux de couverture	Borne inf moyenne	Borne sup moyenne
EMV (Fisher obs)	7,818	89,2 %	5,993	13,811
EMV (Fisher théo)	7,711	88,8 %	6,047	13,758
EMV (PF)	4,411	78,0 %	7,500	11,911
PWM	6,118	81,6 %	7,166	13,284
ME EMV PF	5,979	86,0 %	7,315	13,295
Hill EMV (Fisher théo)	6,518	86,2 %	6,493	13,011
Hill EMV (Fisher obs)	6,507	86,6 %	6,498	13,006
Hill EMV (PF)	7,603	93,0 %	7,454	15,057
Cunn1 EMV (PF)	5,821	88,6 %	7,370	13,191
Pick EMV (PF)	6,270	87,6 %	7,402	13,672

TABLE 8.46 – Propriétés des intervalles de confiance à 95 % pour le quantile à 99,5 % d'un échantillon Student à 2 degrés de liberté de taille 1000

	Taille moyenne	Taux de couverture	Borne inf moyenne	Borne sup moyenne
EMV (Fisher obs)	6,561	84,4 %	6,621	13,183
EMV (Fisher théo)	6,471	84,5 %	6,666	13,138
EMV (PF)	4,148	75,6 %	7,699	11,847
PWM	7,364	88,4 %	6,825	14,189
Hill EMV (Fisher théo)	5,470	82,0 %	7,017	12,487
Hill EMV (Fisher obs)	5,461	82,4 %	7,022	12,482
Hill EMV (PF)	6,079	87,6 %	7,732	13,811
Cunn1 PWM	6,389	78,9 %	6,244	12,633
Cunn1 EMV (PF)	4,753	81,0 %	7,623	12,376
Cunn2 PWM	6,528	78,0 %	6,188	12,716
Cunn2 EMV (PF)	4,765	79,8 %	7,625	12,390
Pick EMV (PF)	5,100	79,0 %	7,666	12,766

TABLE 8.47 – Propriétés des intervalles de confiance à 90 % pour le quantile à 99,5 % d'un échantillon Student à 2 degrés de liberté de taille 1000

	Taille moyenne	Taux de couverture	Borne inf moyenne	Borne sup moyenne
EMV (PF)	0,499	92,2 %	2,664	3,163
EMV (Fisher obs)	0,491	92,0 %	2,677	3,169
EMV (Fisher théo)	0,491	92,0 %	2,678	3,168
PWM	0,592	99,6 %	2,586	3,178
GPWM	0,693	95,9 %	2,547	3,240
ME EMV (Fisher théo)	0,204	77,2 %	2,848	3,053
ME PWM	0,213	74,8 %	2,850	3,062
ME EMV (Fisher obs)	0,203	77,2 %	2,849	3,052
Pick EMV (Fisher théo)	0,213	84,9 %	2,832	3,045
Pick PWM	0,218	83,3 %	2,830	3,047
Pick EMV (Fisher obs)	0,213	84,9 %	2,832	3,045

TABLE 8.48 – Propriétés des intervalles de confiance à 95 % pour le quantile à 95 % d'un échantillon Student à 2 degrés de liberté de taille 10000

	Taille moyenne	Taux de couverture	Borne inf moyenne	Borne sup moyenne
EMV (PF)	0,416	85,7 %	2,708	3,125
EMV (Fisher obs)	0,412	84,8 %	2,717	3,129
EMV (Fisher théo)	0,412	84,8 %	2,717	3,129
PWM	0,4970	98,6 %	2,632	3,129
GPWM	0,582	93,2 %	2,603	3,184
ME EMV (Fisher théo)	0,171	68,5 %	2,865	3,036
ME PWM	0,179	67,2 %	2,867	3,045
ME EMV (Fisher obs)	0,171	68,5 %	2,865	3,036
Pick EMV (Fisher théo)	0,179	77,0 %	2,849	3,028
Pick PWM	0,183	74,4 %	2,847	3,030
Pick EMV (Fisher obs)	0,179	77,0 %	2,849	3,028

TABLE 8.49 – Propriétés des intervalles de confiance à 90 % pour le quantile à 95 % d'un échantillon Student à 2 degrés de liberté de taille 10000

	Taille moyenne	Taux de couverture	Borne inf moyenne	Borne sup moyenne
EMV (Fisher obs)	2,409	93,1 %	8,700	11,109
EMV (Fisher théo)	2,409	93,1 %	8,700	11,109
PWM	2,100	88,4 %	8,920	11,020
ME EMV (Fisher théo)	2,000	86,2 %	8,732	10,732
ME PWM	3,045	92,6 %	8,328	11,373
ME EMV (Fisher obs)	1,975	85,8 %	8,745	10,720
ME EMV (PF)	1,998	89,6 %	8,854	10,852
Hill EMV (Fisher théo)	2,108	89,9 %	8,831	10,939
Hill PWM	3,253	86,9 %	8,091	11,345
Hill EMV (Fisher obs)	2,106	89,9 %	8,832	10,938
Hill EMV (PF)	2,144	91,9 %	8,960	11,104
Cunn1 EMV (Fisher théo)	1,724	70,7 %	8,542	10,266
Cunn1 PWM	2,363	88,5 %	8,706	11,069
Cunn1 EMV (Fisher obs)	1,672	68,1 %	8,568	10,240
Cunn1 EMV (PF)	1,682	73,2 %	8,647	10,329
Cunn2 EMV (Fisher théo)	1,734	72,0 %	8,554	10,288
Cunn2 PWM	2,431	89,3 %	8,686	11,117
Cunn2 EMV (Fisher obs)	1,683	70,0 %	8,579	10,262
Cunn2 EMV (PF)	1,694	74,8 %	8,659	10,353
Pick EMV (Fisher théo)	2,034	87,5 %	8,783	10,817
Pick PWM	3,511	94,1 %	8,122	11,633
Pick EMV (Fisher obs)	2,019	87,3 %	8,791	10,810
Pick EMV (PF)	2,041	90,7 %	8,903	10,944

TABLE 8.50 – Propriétés des intervalles de confiance à 95 % pour le quantile à 99,5 % d'un échantillon Student à 2 degrés de liberté de taille 10000

	Taille moyenne	Taux de couverture	Borne inf moyenne	Borne sup moyenne
EMV (Fisher obs)	2,022	89,2 %	8,893	10,915
EMV (Fisher théo)	2,022	89,2 %	8,893	10,915
PWM	2,498	95,7 %	8,755	11,254
ME EMV (Fisher théo)	1,678	80,2 %	8,893	10,571
ME PWM	2,555	88,3 %	8,573	11,128
ME EMV (Fisher obs)	1,657	79,6 %	8,903	10,561
ME EMV (PF)	1,669	81,6 %	8,983	10,652
Hill EMV (Fisher théo)	1,769	84,7 %	9,000	10,770
Hill PWM	2,730	84,1 %	8,353	11,083
Hill EMV (Fisher obs)	1,767	84,5 %	9,001	10,769
Hill EMV (PF)	1,787	84,7 %	9,095	10,882
Cunn1 PWM	1,983	83,4 %	8,896	10,879
Cunn2 PWM	2,040	85,2 %	8,882	10,921
Pick EMV (Fisher théo)	1,707	83,7 %	8,947	10,654
Pick PWM	2,947	90,5 %	8,404	11,351
Pick EMV (Fisher obs)	1,694	83,1 %	8,953	10,647
Pick EMV (PF)	1,705	83,5 %	9,034	10,739

TABLE 8.51 – Propriétés des intervalles de confiance à 90 % pour le quantile à 99,5 % d'un échantillon Student à 2 degrés de liberté de taille 10000

B.3 Loi de Student à 1 degré de liberté

	Quantile moyen	Ecart quadratique	Seuil moyen	Ecart-type seuil
Empirique	6,31	0,85		
EMV m=50	6,48	1,63		
PWM m=50	6,23	9,83		
GPWM m=50	6,16	1,46		
RB EMV m=50	6,25	0,79		
RB PWM m=50	6,22	8,85		
RB GPWM m=50	6,15	0,80		
EMV m=100	7,09	4,31		
PWM m=100	0,65	384,07		
GPWM m=100	5,17	5,33		
RB EMV m=100	7,08	1,69		
RB PWM m=100	0,72	370,59		
RB GPWM m=100	5,14	2,61		
ME EMV	6,54	1,10	1,80	2,19
ME PWM	6,78	2,71	1,80	2,19
Hill EMV	6,42	1,54	2,00	1,74
Hill PWM	6,52	0,90	2,00	1,74
Cunn1 EMV	6,58	0,68	0,75	0,24
Cunn1 PWM	6,39	0,56	0,75	0,24
Cunn2 EMV	6,51	0,65	1,11	0,39
Cunn2 PWM	6,42	0,65	1,11	0,39
Pick EMV	6,52	0,68	0,78	0,66
Pick PWM	6,46	0,64	0,64	0,92

TABLE 8.52 – Quantile à 95 % pour un échantillon Student à 1 degré de liberté de taille 1000

	Quantile moyen	Ecart quadratique	Seuil moyen	Ecart-type seuil
Empirique	65,32	927,77		
EMV m=50	167,21	∞		
PWM m=50	75,47	7447,62		
GPWM m=50	53,70	489,07		
RB EMV m=50	8828,69	∞		
RB PWM m=50	75,70	7369,02		
RB GPWM m=50	53,98	463,57		
EMV m=100	58151,18	∞		
PWM m=100	89,04	16612,26		
GPWM m=100	54,23	558,22		
RB EMV m=100	62625,16	∞		
RB PWM m=100	88,88	16278,51		
RB GPWM m=100	54,32	498,76		
ME EMV	61,04	719,81	1,80	2,19
ME PWM	51,92	440,88	1,80	2,19
Hill EMV	66,25	585,05	2,00	1,74
Hill PWM	48,22	392,37	2,00	1,74
Cunn1 EMV	60,48	403,81	0,75	0,24
Cunn1 PWM	47,44	397,54	0,75	0,24
Cunn2 EMV	63,54	488,06	1,11	0,39
Cunn2 PWM	47,52	399,32	1,11	0,39
Pick EMV	63,32	551,36	0,78	0,66
Pick PWM	48,07	387,83	0,64	0,92

TABLE 8.53 – Quantile à 99,5 % pour un échantillon Student à 1 degré de liberté de taille 1000

	Quantile moyen	Ecart quadratique	Seuil moyen	Ecart-type seuil
Empirique	6,33	0,09		
EMV m=50	6,38	0,26		
PWM m=50	3,96	28,15		
GPWM m=50	6,21	0,47		
RB EMV m=50	6,39	0,11		
RB PWM m=50	3,98	27,37		
RB GPWM m=50	6,22	0,22		
EMV m=100	6,75	2,13		
PWM m=100	-5,75	867,03		
GPWM m=100	4,64	10,35		
RB EMV m=100	6,73	0,37		
RB PWM m=100	-5,81	858,79		
RB GPWM m=100	4,58	7,83		
ME EMV	6,37	0,13	3,60	3,08
ME PWM	6,39	0,47	3,60	3,08
Hill EMV	6,39	0,09	2,51	1,72
Hill PWM	6,50	0,15	2,51	1,72
Cunn1 EMV	6,53	0,11	0,77	0,23
Cunn1 PWM	6,48	0,09	0,77	0,23
Cunn2 EMV	6,46	0,09	1,16	0,41
Cunn2 PWM	6,45	0,09	1,16	0,41
Pick EMV	6,40	0,08	1,41	0,70
Pick PWM	6,46	0,10	0,79	0,68

TABLE 8.54 – Quantile à 95 % pour un échantillon Student à 1 degré de liberté de taille 10000

	Quantile moyen	Ecart quadratique	Seuil moyen	Ecart-type seuil
Empirique	63,9	78,1		
EMV m=50	64,4	64,6		
PWM m=50	71,6	1106,8		
GPWM m=50	63,6	61,0		
RB EMV m=50	64,4	60,0		
RB PWM m=50	71,5	1114,5		
RB GPWM m=50	63,6	56,2		
EMV m=100	64,5	134,9		
PWM m=100	78,9	2401,3		
GPWM m=100	63,1	76,3		
RB EMV m=100	64,4	65,2		
RB PWM m=100	78,8	2405,5		
RB GPWM m=100	63,0	61,0		
ME EMV	60,9	100,8	3,60	3,08
ME PWM	57,2	97,1	3,60	3,08
Hill EMV	62,9	45,3	2,51	1,72
Hill PWM	53,4	140,8	2,51	1,72
Cunn1 EMV	57,6	77,3	0,77	0,23
Cunn1 PWM	52,7	164,5	0,77	0,23
Cunn2 EMV	60,4	54,2	1,16	0,41
Cunn2 PWM	52,5	164,4	1,16	0,41
Pick EMV	62,4	57,0	1,41	0,70
Pick PWM	52,6	163,2	0,79	0,68

TABLE 8.55 – Quantile à 99,5 % pour un échantillon Student à 1 degré de liberté de taille 10000

	Taille moyenne	Taux de couverture	Borne inf moyenne	Borne sup moyenne
EMV (PF)	3,831	81,2 %	4,542	8,373
EMV (Fisher obs)	4,630	91,8 %	4,163	8,793
EMV (Fisher théo)	4,500	91,2 %	4,234	8,734
PWM	8,543	99,2 %	1,569	10,111
Cunn1 EMV (Fisher théo)	2,700	92,6 %	5,226	7,925
Cunn1 EMV (Fisher obs)	2,698	92,6 %	5,226	7,924
Cunn1 EMV (PF)	2,789	93,0 %	5,434	8,223
Cunn2 EMV (Fisher théo)	2,699	92,8 %	5,238	7,937
Cunn2 EMV (Fisher obs)	2,698	92,6 %	5,239	7,936
Cunn2 EMV (PF)	2,788	90,8 %	5,446	8,234
Pick EMV (Fisher théo)	2,689	92,4 %	5,173	7,862
Pick EMV (Fisher obs)	2,688	92,4 %	5,173	7,861
Pick EMV (PF)	2,777	91,2 %	5,382	8,159

TABLE 8.56 – Propriétés des intervalles de confiance à 95 % pour le quantile à 95 % d'un échantillon Student à 1 degré de liberté de taille 1000

	Taille moyenne	Taux de couverture	Borne inf moyenne	Borne sup moyenne
PWM	205,4	90,8 %	31,7	237,2
Hill EMV (Fisher théo)	90,5	89,2 %	21,0	111,5
Hill EMV (Fisher obs)	90,2	89,2 %	21,1	111,4
Hill EMV (PF)	113,3	93,4 %	37,8	151,1
Cunn1 EMV (PF)	83,6	90,8 %	36,2	119,8
Cunn2 EMV (PF)	82,1	91,6 %	36,0	118,1
Pick EMV (PF)	93,8	87,8 %	37,3	131,0

TABLE 8.57 – Propriétés des intervalles de confiance à 95 % pour le quantile à 99,5 % d'un échantillon Student à 1 degré de liberté de taille 1000

	Taille moyenne	Taux de couverture	Borne inf moyenne	Borne sup moyenne
EMV (PF)	3,446	80 %	4,805	8,251
EMV (Fisher obs)	3,886	87,2 %	4,535	8,421
EMV (Fisher théo)	3,776	86,4 %	4,596	8,372
PWM	7,298	96,6 %	2,090	9,388
Cunn1 EMV (Fisher théo)	2,265	86,2 %	5,442	7,708
Cunn1 PWM	1,905	80,8 %	5,442	7,347
Cunn1 EMV (Fisher obs)	2,264	86,2 %	5,443	7,707
Cunn1 EMV (PF)	2,314	83,6 %	5,594	7,908
Cunn2 EMV (Fisher théo)	2,265	85,0 %	5,455	7,720
Cunn2 PWM	1,910	81,0 %	5,449	7,359
Cunn2 EMV (Fisher obs)	2,264	85,0 %	5,456	7,720
Cunn2 EMV (PF)	2,314	82,4 %	5,606	7,920
Pick EMV (Fisher théo)	2,257	85,6 %	5,389	7,646
Pick EMV (Fisher obs)	2,256	85,6 %	5,390	7,645
Pick EMV (PF)	2,305	83,4 %	5,542	7,846

TABLE 8.58 – Propriétés des intervalles de confiance à 90 % pour le quantile à 95 % d'un échantillon Student à 1 degré de liberté de taille 1000

	Taille moyenne	Taux de couverture	Borne inf moyenne	Borne sup moyenne
PWM	181,0	86,2 %	35,4	216,4
Hill EMV (Fisher théo)	75,9	86,6 %	28,3	104,2
Hill EMV (Fisher obs)	75,7	86,2 %	28,4	104,1
Hill EMV (PF)	89,2	88,6 %	40,7	129,9
Cunn1 EMV (PF)	66,9	85,6 %	39,0	105,9
Cunn2 EMV (PF)	65,8	85,6 %	38,8	104,5
Pick EMV (PF)	74,9	82,6 %	40,1	114,9

TABLE 8.59 – Propriétés des intervalles de confiance à 90 % pour le quantile à 99,5 % d'un échantillon Student à 1 degré de liberté de taille 1000

	Taille moyenne	Taux de couverture	Borne inf moyenne	Borne sup moyenne
EMV (PF)	1,900	93,8 %	5,445	7,346
EMV (Fisher obs)	1,870	93,2 %	5,445	7,315
EMV (Fisher théo)	1,870	93,2 %	5,445	7,315
Hill EMV (Fisher théo)	0,791	82,9 %	5,993	6,783
Hill PWM	0,728	71,4 %	6,148	6,877
Hill EMV (Fisher obs)	0,790	82,9 %	5,993	6,783
Cunn1 EMV (Fisher théo)	0,840	81,5 %	6,108	6,948
Cunn1 PWM	0,762	81,1 %	6,096	6,858
Cunn1 EMV (Fisher obs)	0,840	81,5 %	6,108	6,948
Cunn1 EMV (PF)	0,841	78,1 %	6,132	6,972
Cunn2 EMV (Fisher théo)	0,840	79,9 %	6,108	6,947
Cunn2 PWM	0,763	79,1 %	6,098	6,862
Cunn2 EMV (Fisher obs)	0,839	79,9 %	6,108	6,947
Cunn2 EMV (PF)	0,840	76,7 %	6,131	6,971
Pick EMV (Fisher théo)	0,802	84,9 %	5,996	6,798
Pick PWM	0,742	77,7 %	6,095	6,838
Pick EMV (Fisher obs)	0,802	84,9 %	5,996	6,798
Pick EMV (PF)	0,803	83,7 %	6,017	6,820

TABLE 8.60 – Propriétés des intervalles de confiance à 95 % pour le quantile à 95 % d'un échantillon Student à 1 degré de liberté de taille 10000

	Taille moyenne	Taux de couverture	Borne inf moyenne	Borne sup moyenne
EMV (PF)	1,588	88,2 %	5,598	7,186
EMV (Fisher obs)	1,570	87,4 %	5,595	7,165
EMV (Fisher théo)	1,569	87,2 %	5,595	7,165
PWM	9,305	29,0 %	-3,445	5,860
Hill EMV (Fisher théo)	0,664	74,6 %	6,056	6,720
Hill EMV (Fisher obs)	0,663	74,6 %	6,056	6,720
Cunn1 EMV (Fisher théo)	0,705	72,0 %	6,176	6,881
Cunn1 PWM	0,640	70,6 %	6,157	6,797
Cunn1 EMV (Fisher obs)	0,705	72,0 %	6,176	6,881
Cunn1 EMV (PF)	0,704	69,6 %	6,193	6,897
Cunn2 EMV (Fisher théo)	0,705	71,2 %	6,175	6,880
Cunn2 PWM	0,640	70,6 %	6,160	6,800
Cunn2 EMV (Fisher obs)	0,704	71,2 %	6,175	6,880
Cunn2 EMV (PF)	0,704	69,2 %	6,192	6,896
Pick EMV (Fisher théo)	0,673	77,3 %	6,060	6,733
Pick PWM	0,623	67,8 %	6,155	6,778
Pick EMV (Fisher obs)	0,673	77,3 %	6,060	6,733
Pick EMV (PF)	0,673	75,1 %	6,076	6,749

TABLE 8.61 – Propriétés des intervalles de confiance à 90 % pour le quantile à 95 % d'un échantillon Student à 1 degré de liberté de taille 10000

	Taille moyenne	Taux de couverture	Borne inf moyenne	Borne sup moyenne
EMV (Fisher obs)	31,758	93,8 %	48,507	80,264
EMV (Fisher théo)	31,710	93,8 %	48,531	80,240
PWM	78,074	95,6 %	52,500	130,574
Hill EMV (Fisher théo)	26,528	93,6 %	49,670	76,198
Hill EMV (Fisher obs)	26,443	93,6 %	49,713	76,156
Hill EMV (PF)	27,028	97,2 %	51,696	78,725
Pick EMV (Fisher théo)	25,341	86,1 %	49,740	75,081
Pick EMV (Fisher obs)	25,279	85,7 %	49,771	75,050
Pick EMV (PF)	25,735	89,9 %	51,543	72,278

TABLE 8.62 – Propriétés des intervalles de confiance à 95 % pour le quantile à 99,5 % d'un échantillon Student à 1 degré de liberté de taille 10000

	Taille moyenne	Taux de couverture	Borne inf moyenne	Borne sup moyenne
EMV (Fisher obs)	26,652	89,2 %	51,059	77,711
EMV (Fisher théo)	26,612	89,2 %	51,080	77,691
PWM	68,601	89,0 %	54,842	123,443
Hill EMV (Fisher théo)	22,263	88,5 %	51,803	74,066
Hill EMV (Fisher obs)	22,192	88,3 %	51,838	74,030
Hill EMV (PF)	22,511	89,7 %	53,275	75,787
Pick EMV (Fisher théo)	21,266	80,3 %	51,777	73,044
Pick EMV (Fisher obs)	21,215	80,1 %	51,803	73,018
Pick EMV (PF)	21,458	80,5 %	53,086	74,544

TABLE 8.63 – Propriétés des intervalles de confiance à 90 % pour le quantile à 99,5 % d'un échantillon Student à 1 degré de liberté de taille 10000

C Présentation des différents tests

C.1 Test de Kolmogorov-Smirnov

L'hypothèse nulle de ce test est l'adéquation d'une série d'observations $(X_i)_{i=1,\dots,n}$ à une loi de fonction de répartition $F_{\hat{\theta}}$. Il consiste à calculer :

$$\max_i |\hat{F}_n(X_i) - F_{\hat{\theta}}(X_i)|$$

Où \hat{F}_n est la fonction de répartition empirique de l'échantillon. La loi de cette statistique sous l'hypothèse nulle est connue et tabulée, ce qui permet de rejeter l'hypothèse nulle lorsque celle-ci est trop élevée.

C.2 Test de Wald-Wolfowitz

L'hypothèse nulle de ce test est l'indépendance d'une série d'observations. Il consiste à récrire notre série d'observations en les remplaçant par un '+' si elles sont supérieures à la médiane, et par un '-' sinon. Le début de la série pourrait par exemple se réécrire '+ + + - - - - + + - + - - + - - -'. Nous appelons *run* le nombre de séquences d'un même signe (même s'il n'apparaît qu'une seule fois). Dans la série proposée, nous avons donc 8 *runs*, pour une taille totale de 20. Dans le cas d'une dépendance de la série, ce nombre pourrait être soit trop élevé (les signes alternent beaucoup, il y a une dépendance négative entre une valeur de la série et la suivante) soit trop faible (les signes alternent peu, il y a une dépendance positive). Le test de Wald-Wolfowitz utilise la loi limite du nombre de *runs* (celle-ci n'est pas connue explicitement, elle est tabulée) sous l'hypothèse d'indépendance pour calculer la *p-value* correspondant à notre série.

C.3 Test de Ljung-Box

L'hypothèse nulle de ce test est l'indépendance d'une série d'observations. Pour cela, il faut calculer les auto-corrélations r_k des observations pour k allant de 1 à m . Pour rappel, r_k est égal au coefficient de corrélation entre $(X_t)_t$ et $(X_{t-k})_t$. Nous prenons alors comme statistique de test :

$$Q = n(n+2) \sum_{k=1}^m \frac{r_k^2}{n-k}$$

Sous l'hypothèse nulle, celle-ci est distribuée selon une loi χ_m^2 .

C.4 Test de Phillips-Perron

L'hypothèse nulle de ce test est la non-stationnarité d'une série d'observations. Il s'appuie sur le modèle :

$$\Delta X_t = \rho X_{t-1} + \epsilon_t$$

Une statistique de Student pour la nullité du coefficient ρ est calculée, en utilisant une correction pour l'auto-corrélation et l'hétéroscédasticité. Si $\rho = 0$ est accepté, alors la série est intégrée d'ordre 1 et donc non-stationnaire.

C.5 Test de Shapiro-Wilk

L'hypothèse nulle de ce test est la normalité d'une série d'observations. Il compare les quantiles théoriques d'une loi normale avec les quantiles obtenus sur l'échantillon. Si ceux-ci sont suffisamment proches, alors la normalité est acceptée.

D Résultats sur les composantes principales C_2 , C_3 , et C_4

D.1 C_2

$\hat{\mu}$	$1,373 \times 10^{-3}$	0,034
$\hat{\omega}$	$4,497 \times 10^{-5}$	0,007
$\hat{\alpha}$	$1,142 \times 10^{-1}$	$4,81 \times 10^{-7}$
$\hat{\beta}_1$	$8,723 \times 10^{-1}$	$7,13 \times 10^{-5}$

TABLE 8.64 – Valeurs des paramètres du modèle $GARCH(1,1)$

Test	η	η^2
WALD-WOLFOWITZ	0,001	0,05
LJUNG-BOX (m = 1 mois)	0,52	0,80
LJUNG-BOX (m = 1 an)	0,73	0,29
LJUNG-BOX (m = 2 ans)	0,94	0,32

TABLE 8.65 – Tests d'indépendance des résidus

μ_{norm}	0,010
σ_{norm}	0,250
ξ_1	0,199
μ_1	-1,500
σ_1	0,290
ϵ_1	4,138
ξ_2	0,170
μ_2	1,500
σ_2	0,280
ϵ_2	4,110

TABLE 8.66 – Paramètres de la Pareto Hybride

0,31

TABLE 8.67 – p -value du test de Kolmogorov-Smirnov

D.2 C_3

$\hat{\mu}$	$2,895 \times 10^{-4}$	0,359
$\hat{\omega}$	$2,211 \times 10^{-6}$	0,068
$\hat{\alpha}$	$4,568 \times 10^{-2}$	$1,26 \times 10^{-4}$
$\hat{\beta}_1$	$9,503 \times 10^{-1}$	$< 2 \times 10^{-16}$

TABLE 8.68 – Valeurs des paramètres du modèle $GARCH(1,1)$

Test	η	η^2
WALD-WOLFOWITZ	0,30	$1,41 \times 10^{-8}$
LJUNG-BOX (m = 1 mois)	0,35	$1,73 \times 10^{-3}$
LJUNG-BOX (m = 1 an)	0,60	$4,80 \times 10^{-8}$
LJUNG-BOX (m = 2 ans)	0,91	$3,06 \times 10^{-3}$

TABLE 8.69 – Tests d'indépendance des résidus

	AAA	AA	A	BBB
μ_{norm}	0,018	0,274	0,107	-0,179
σ_{norm}	0,198	0,145	0,142	0,100
ξ_1	0,316	0,301	0,112	0,421
μ_1	-2,000	-1,000	-1,000	-1,000
σ_1	0,198	0,226	0,403	0,184
ϵ_1	4,949	3,768	3,810	2,955
ξ_2	0,560	0,292	0,269	0,129
μ_2	2,000	2,000	1,300	0,500
σ_2	0,111	0,173	0,279	0,532
ϵ_2	4,251	4,715	3,795	3,178

TABLE 8.70 – Paramètres de la Pareto Hybride

0,51

TABLE 8.71 – p -value du test de Kolmogorov-SmirnovD.3 C_4

$\hat{\mu}$	$1,500 \times 10^{-4}$	0,338
$\hat{\omega}$	$2,710 \times 10^{-7}$	0,108
$\hat{\alpha}$	$5,526 \times 10^{-2}$	$6,23 \times 10^{-5}$
$\hat{\beta}_1$	$9,468 \times 10^{-1}$	$< 2 \times 10^{-16}$

TABLE 8.72 – Valeurs des paramètres du modèle $GARCH(1, 1)$

Test	η	η^2
WALD-WOLFOWITZ	1,00	$1,44 \times 10^{-2}$
LJUNG-BOX (m = 1 mois)	0,99	
LJUNG-BOX (m = 1 an)	0,23	0,99
LJUNG-BOX (m = 2 ans)	0,71	1,00

TABLE 8.73 – Tests d'indépendance des résidus

μ_{norm}	0,028
σ_{norm}	0,283
ξ_1	0,241
μ_1	-1,500
σ_1	0,244
ϵ_1	3,967
ξ_2	0,240
μ_2	1,500
σ_2	0,251
ϵ_2	4,106

TABLE 8.74 – Paramètres de la Pareto Hybride

0,35

TABLE 8.75 – *p-value* du test de Kolmogorov-Smirnov

E Résultats pour la modélisation des taux maturité par maturité

E.1 Valeur et significativité des coefficients $ARMA(1,1) - GARCH(1,1)$

1 an	$\hat{\mu}$	$3,524 \times 10^{-4}$	0,0212
	$\hat{\rho}$	$2,541 \times 10^{-1}$	0,0806
	$\hat{\theta}$	$-1,033 \times 10^{-1}$	0,470
	$\hat{\omega}$	$1,536 \times 10^{-7}$	0,0347
	$\hat{\alpha}$	$6,045 \times 10^{-2}$	$1,27 \times 10^{-8}$
	$\hat{\beta}_1$	$9,425 \times 10^{-1}$	$< 2 \times 10^{-16}$
2 ans	$\hat{\mu}$	$2,488 \times 10^{-4}$	0,2912
	$\hat{\rho}$	$-4,610 \times 10^{-2}$	0,7945
	$\hat{\theta}$	$1,276 \times 10^{-1}$	0,4649
	$\hat{\omega}$	$3,326 \times 10^{-7}$	0,0671
	$\hat{\alpha}$	$5,423 \times 10^{-2}$	$1,14 \times 10^{-8}$
	$\hat{\beta}_1$	$9,474 \times 10^{-1}$	$< 2 \times 10^{-16}$
3 ans	$\hat{\mu}$	$1,074 \times 10^{-4}$	0,6733
	$\hat{\rho}$	$-1,475 \times 10^{-1}$	0,5220
	$\hat{\theta}$	$1,879 \times 10^{-1}$	0,4089
	$\hat{\omega}$	$4,645 \times 10^{-7}$	0,0744
	$\hat{\alpha}$	$4,698 \times 10^{-2}$	$2,14 \times 10^{-8}$
	$\hat{\beta}_1$	$9,530 \times 10^{-1}$	$< 2 \times 10^{-16}$
4 ans	$\hat{\mu}$	$-2,059 \times 10^{-5}$	0,9291
	$\hat{\rho}$	$-1,024 \times 10^{-1}$	0,5214
	$\hat{\theta}$	$1,221 \times 10^{-1}$	0,4599
	$\hat{\omega}$	$4,281 \times 10^{-7}$	0,0775
	$\hat{\alpha}$	$4,192 \times 10^{-2}$	$2,27 \times 10^{-8}$
	$\hat{\beta}_1$	$9,576 \times 10^{-1}$	$< 2 \times 10^{-16}$
5 ans	$\hat{\mu}$	$-1,183 \times 10^{-4}$	0,6659
	$\hat{\rho}$	$-4,109 \times 10^{-1}$	0,0515
	$\hat{\theta}$	$4,039 \times 10^{-1}$	0,0410
	$\hat{\omega}$	$4,464 \times 10^{-7}$	0,0647
	$\hat{\alpha}$	$4,408 \times 10^{-2}$	$8,33 \times 10^{-9}$
	$\hat{\beta}_1$	$9,552 \times 10^{-1}$	$< 2 \times 10^{-16}$
7 ans	$\hat{\mu}$	$-6,464 \times 10^{-5}$	0,4563
	$\hat{\rho}$	$5,344 \times 10^{-1}$	0,0171
	$\hat{\theta}$	$-5,527 \times 10^{-1}$	0,0124
	$\hat{\omega}$	$4,744 \times 10^{-7}$	0,0512
	$\hat{\alpha}$	$4,351 \times 10^{-2}$	$7,21 \times 10^{-10}$
	$\hat{\beta}_1$	$9,544 \times 10^{-1}$	$< 2 \times 10^{-16}$
10 ans	$\hat{\mu}$	$-7,213 \times 10^{-5}$	0,3482
	$\hat{\rho}$	$5,560 \times 10^{-1}$	0,0105
	$\hat{\theta}$	$-5,826 \times 10^{-1}$	0,0058
	$\hat{\omega}$	$5,479 \times 10^{-7}$	0,0337
	$\hat{\alpha}$	$4,504 \times 10^{-2}$	$2,04 \times 10^{-10}$
	$\hat{\beta}_1$	$9,511 \times 10^{-1}$	$< 2 \times 10^{-16}$
15 ans	$\hat{\mu}$	$-7,589 \times 10^{-5}$	0,2874
	$\hat{\rho}$	$5,480 \times 10^{-1}$	0,0028
	$\hat{\theta}$	$-5,879 \times 10^{-1}$	0,0009
	$\hat{\omega}$	$1,385 \times 10^{-6}$	0,0369
	$\hat{\alpha}$	$6,603 \times 10^{-2}$	0,0002
	$\hat{\beta}_1$	$9,207 \times 10^{-1}$	$< 2 \times 10^{-16}$
20 ans	$\hat{\mu}$	$-4,896 \times 10^{-5}$	0,3183
	$\hat{\rho}$	$6,611 \times 10^{-1}$	$5,58 \times 10^{-6}$
	$\hat{\theta}$	$-6,980 \times 10^{-1}$	$5,04 \times 10^{-7}$
	$\hat{\omega}$	$1,374 \times 10^{-6}$	0,0257
	$\hat{\alpha}$	$7,561 \times 10^{-2}$	$6,06 \times 10^{-5}$
	$\hat{\beta}_1$	$9,113 \times 10^{-1}$	$< 2 \times 10^{-16}$
30 ans	$\hat{\mu}$	$-4,064 \times 10^{-5}$	0,3179
	$\hat{\rho}$	$7,091 \times 10^{-1}$	$1,30 \times 10^{-8}$
	$\hat{\theta}$	$-7,444 \times 10^{-1}$	$3,10 \times 10^{-10}$
	$\hat{\omega}$	$1,614 \times 10^{-6}$	0,0181
	$\hat{\alpha}$	$9,042 \times 10^{-2}$	$2,97 \times 10^{-5}$
	$\hat{\beta}_1$	$8,957 \times 10^{-1}$	$< 2 \times 10^{-16}$

TABLE 8.76 – Valeurs des paramètres du modèle $ARMA(1,1) - GARCH(1,1)$ pour toutes les maturités

E.2 Tests d'indépendance des résidus

	Test	η	η^2
1 an	WALD-WOLFOWITZ	0,37	0,10
	LJUNG-BOX (m = 1 mois)	$7,3 \times 10^{-5}$	0,33
	LJUNG-BOX (m = 1 an)	0,11	$1,7 \times 10^{-2}$
	LJUNG-BOX (m = 2 ans)	0,91	0,66
2 ans	WALD-WOLFOWITZ	0,46	0,24
	LJUNG-BOX (m = 1 mois)	$2,0 \times 10^{-2}$	0,20
	LJUNG-BOX (m = 1 an)	0,48	$4,9 \times 10^{-7}$
	LJUNG-BOX (m = 2 ans)	0,99	$5,9 \times 10^{-4}$
3 ans	WALD-WOLFOWITZ	0,51	$4,0 \times 10^{-2}$
	LJUNG-BOX (m = 1 mois)	$4,3 \times 10^{-2}$	0,34
	LJUNG-BOX (m = 1 an)	0,39	$7,0 \times 10^{-5}$
	LJUNG-BOX (m = 2 ans)	0,98	$7,0 \times 10^{-2}$
4 ans	WALD-WOLFOWITZ	0,57	0,33
	LJUNG-BOX (m = 1 mois)	0,25	0,36
	LJUNG-BOX (m = 1 an)	0,52	$5,1 \times 10^{-3}$
	LJUNG-BOX (m = 2 ans)	0,99	0,13
5 ans	WALD-WOLFOWITZ	0,54	0,34
	LJUNG-BOX (m = 1 mois)	0,35	0,30
	LJUNG-BOX (m = 1 an)	0,48	$3,0 \times 10^{-2}$
	LJUNG-BOX (m = 2 ans)	0,98	0,32
7 ans	WALD-WOLFOWITZ	0,34	0,37
	LJUNG-BOX (m = 1 mois)	0,69	0,52
	LJUNG-BOX (m = 1 an)	0,73	0,13
	LJUNG-BOX (m = 2 ans)	0,99	0,60
10 ans	WALD-WOLFOWITZ	0,38	0,56
	LJUNG-BOX (m = 1 mois)	0,91	0,73
	LJUNG-BOX (m = 1 an)	0,81	0,69
	LJUNG-BOX (m = 2 ans)	0,99	0,93
15 ans	WALD-WOLFOWITZ	0,31	$1,7 \times 10^{-2}$
	LJUNG-BOX (m = 1 mois)	0,97	0,18
	LJUNG-BOX (m = 1 an)	0,91	0,54
	LJUNG-BOX (m = 2 ans)	1,00	0,77
20 ans	WALD-WOLFOWITZ	0,23	0,15
	LJUNG-BOX (m = 1 mois)	0,97	$4,8 \times 10^{-2}$
	LJUNG-BOX (m = 1 an)	0,71	0,43
	LJUNG-BOX (m = 2 ans)	0,99	0,47
30 ans	WALD-WOLFOWITZ	0,15	0,16
	LJUNG-BOX (m = 1 mois)	0,89	$2,2 \times 10^{-3}$
	LJUNG-BOX (m = 1 an)	0,71	$9,1 \times 10^{-2}$
	LJUNG-BOX (m = 2 ans)	0,99	$1,5 \times 10^{-2}$

TABLE 8.77 – Tests d'indépendance des résidus

E.3 Tests d'adéquation à la loi de Pareto Hybride

1 an	2 ans	3 ans	4 ans	5 ans	7 ans	10 ans	15 ans	20 ans	30 ans
0,85	0,50	0,56	0,22	0,30	0,12	0,12	0,11	0,17	0,18

TABLE 8.78 – *p-value* des tests de Kolmogorov-Smirnov