

Provisions techniques et capital de solvabilité d'une compagnie d'assurance : méthodologie d'utilisation de Value-at-Risk

Pierre-E. Thérond * Frédéric Planchet^{†§}

Résumé

The advent of the future European prudential framework (Solvency II) and, to a lesser extent, of the phase II of the IFRS dedicated to the insurance contracts, will systematize the use of the Value-at-Risk (VaR) risk measure in insurance. Especially used for financial purposes, the measure of an insurance risk by a VaR technique requires a specific adaptation. Schematically we may distinguish two different contexts, which impose distinct approaches :

- the measurement of the risk related to the sinistrality by the use of a VaR in the heart of the probability distribution : the technical provision will have to be enough to pay the claims with a 75 % probability ;
- the measurement of risk related to the ruin of the company by the mean of a very high order VaR : the solvency capital must to be calibrated to control the ruin if the insurer with a probability higher than 99.5 %.

In the first situation, the two standard approaches (historical VaR or modelling the sinistrality) can be based on a statistical material of relative but sufficient size to estimate a VaR in the heart of the probability distribution. In the second case, we are confronted to the absence of observations. Also we have to model the basic variables which influence the solvency of the company in order to be able to simulate the ruin of the company and finally to estimate the high order VaR. This last stage will require the use of Extreme Value Theory.

In this paper, we present the contexts of VaR computation in insurance, the related probabilistic results and the limits of these kinds of criterion for insurance purposes.

KEYWORDS : Solvency, extremes, Value-at-Risk.

*Actuaire conseil chez WINTER & Associés, doctorant et chargé de cours à l'ISFA.
Contact : +33 (0) 4 37 37 80 90, ptherond@winter-associes.fr.

†Actuaire associé chez WINTER & Associés, Professeur associé à l'ISFA. Contact :
fplanchet@winter-associes.fr.

‡Cabinet WINTER & Associés -18 avenue Félix Faure - F-69007 Lyon et 43-47 avenue
de la Grande Armée - 75116 Paris - France.

§ISFA, Université Lyon 1 - 50, avenue Tony Garnier - 69007 Lyon - France.

L'avènement du référentiel prudentiel Solvabilité II et, dans une moindre mesure, du nouveau cadre comptable induit par la phase II de la norme IFRS dédiée aux contrats d'assurance, va systématiser l'emploi de Value-at-Risk (VaR) en assurance. Particulièrement utilisées en finance de marché, les techniques de calcul de VaR exigent une adaptation spécifique dans un contexte assurantielle de par la nature des risques qui sont ainsi mesurés. Schématiquement on distingue deux contextes, qui imposent des approches différentes :

- la mesure du risque lié à la sinistralité au moyen d'une VaR dans le corps de la distribution : la provision devra suffire à payer les sinistres dans 75
- la mesure de risque lié à la ruine de la compagnie par le biais d'une VaR d'ordre très élevé : le capital de solvabilité devra permettre à la compagnie de ne pas être en ruine avec une probabilité supérieure à 99,5

Dans la première situation, que l'on adopte une approche VaR historique ou que l'on cherche à modéliser la sinistralité, on reste dans un cadre dans lequel on dispose d'un matériel statistique de taille relative mais suffisante pour estimer une VaR dans le corps de la distribution. Dans le second cas, on est en revanche confronté en l'absence d'observations et il convient alors, dans un premier temps, de modéliser les variables de base qui influent sur la solvabilité de la compagnie, dans un deuxième temps, de simuler la ruine de la compagnie et enfin d'estimer une VaR d'ordre très élevé. Cette dernière étape nécessitera le recours à la théorie des extrêmes.

Dans ce travail, on s'attache à présenter les contextes de calcul de VaR en assurance, les résultats probabilistes sous-jacents et les limites de ces critères.

KEYWORDS : Solvabilité, extrêmes, Value-at-Risk.

1 Introduction

A l'instar de ce qu'a connu le monde bancaire ces dernières années avec l'avènement des accords dits de Bâle II, le monde européen de l'assurance se prépare à se doter d'un nouveau référentiel prudentiel qui résultera du projet Solvabilité II. Construit également sur une structure à trois piliers, le futur système de solvabilité vise à ce que les entreprises d'assurance mesurent et gèrent mieux les risques auxquels elles sont soumises. Une bonne gestion des risques aura pour conséquence une moindre exigence en termes de capitaux propres. Le premier pilier, consacré aux exigences quantitatives, fait largement référence à des mesures de risques bien connues des financiers, les *Value-at-Risk* (VaR). Les travaux en cours sur Solvabilité II prévoient que les assureurs devront disposer d'un montant de provisions techniques¹ leur permettant d'honorer leurs engagements de payer les prestations avec une probabilité de 75 % (à ce stade, une approche *cost of capital* est également envisagée en assurance-vie). Ils devront de plus disposer d'un niveau de fonds propres leur permettant de ne pas être en ruine, à horizon un an, avec une très forte probabilité (*a priori* 99,5 %).

Pour mémoire, dans le système actuel, les montants de provisions sont, le plus souvent, déterminés par des méthodes déterministes prudentes (méthodes et hypothèses conservatrices d'estimation de la charge ultime de sinistres et absence d'actualisation ou actualisation à un taux résolument prudent). L'exigence minimale de fonds propres est quant à elle régie par le système de la marge de solvabilité : les assureurs doivent disposer d'un niveau de fonds propres et de plus-values latentes sur les placements financiers supérieur à une *exigence de marge de solvabilité* qui est exprimée en pourcentage des provisions mathématiques en assurance vie, et en pourcentage des primes perçues ou des sinistres payés en assurance non-vie. Ce système est donc entièrement déterministe et ne fait pas explicitement référence au risque réellement supporté par la compagnie. En particulier, deux entreprises qui ont des profils de risque très différents mais les mêmes éléments comptables se voient exiger un même niveau de capitaux propres². Il convient néanmoins de préciser que, d'un point de vue prudentiel, ce système s'avère relativement

¹Le passif d'une société d'assurance est essentiellement composé des provisions techniques d'une part, et des fonds propres d'autre part.

²Ce constat est néanmoins à modérer de par l'intervention de l'autorité de contrôle (l'ACAM) qui peut relever les exigences de capitaux propres des assureurs

efficace au vu du faible nombre de sociétés d'assurance européennes qui ont fait faillite au cours des dernières années.

Si elles font toutes les deux références à une VaR, les deux nouvelles exigences quantitatives prévues par Solvabilité II sont très différentes de par leur nature. L'exigence portant sur les provisions techniques ne va pas poser de problème pratique majeur. En effet, les assureurs disposent d'un nombre conséquent d'observations et de méthodes stochastiques de provisionnement robustes. De plus le quantile à estimer n'est pas très élevé, on se situe donc dans le coeur de la distribution pour lequel on dispose d'une information de qualité. Il en va différemment pour la détermination du niveau souhaitable de fonds propres ou *Solvency Capital Requirement* (SCR). En effet, l'assureur n'observe pas directement la variable d'intérêt (le résultat), il ne dispose donc pas de matériel statistique pour estimer directement la VaR. Une formule standard³ sera proposée et aura pour vocation d'approcher autant que possible le critère de probabilité de ruine. Le projet prévoit, qu'en parallèle, les sociétés pourront construire un *modèle interne* qui, sur la base de la modélisation de l'ensemble des variables influant la solvabilité de la compagnie, permettra de simuler la situation de la société à horizon un an et ainsi de mesurer le niveau de fonds propres dont a besoin aujourd'hui la compagnie pour ne pas être en ruine, un an plus tard, avec une probabilité de 99,5 %. La construction d'un tel modèle est une problématique à part entière : elle nécessite la modélisation fine des *variables de base* et de leurs interactions. Un tel modèle doit de plus intégrer des contraintes techniques liées à la puissance de calcul informatique. Par ailleurs, le montant à estimer est un quantile extrême de la distribution. Une estimation empirique naïve robuste par le quantile empirique à 99,5 % nécessiterait un nombre de simulations extrêmement important, chacune de ces réalisations résultant de la simulation de différents phénomènes tels que la sinistralité, l'écoulement des provisions techniques, l'évolution des actifs financiers, etc. L'estimation d'un tel quantile devra donc passer par l'utilisation des techniques issues de la théorie des extrêmes qui s'est particulièrement développée depuis le milieu des années 70 et les travaux de Pickands (1975) et Hill (1975), et plus

³Dans le cadre de l'étude d'impact quantitative QIS2 portant que les exigences de fonds propres auxquelles pourrait conduire le projet Solvabilité II, le Comité Européen des Contrôleurs d'Assurance et de Pensions Professionnelles (CEIOPS) a publié un modèle qui vise à la mesure de chaque risque et à l'agrégation des capitaux correspondants par une formule du type de celle du RBC américain.

récemment de Smith (1987), Dekkers et de Haan (1989) ou encore Dekkers et al. (1989) et a été rapidement appliquée aux problématiques financières et assurantielles (Embrechts et al. (1997)). Or ces résultats utilisent les données de queue ; cela pose un nouveau problème, celui de la robustesse de ces données qui, rappelons-le, ne sont pas observées mais simulées. En effet, les modèles internes passeront, le plus souvent, par une modélisation paramétrique des variables de base (évolution du cours des actions par exemple), or ces modèles ne collent pas parfaitement aux observations. En particulier les paramètres sont estimés sur l'ensemble de la distribution au risque de ne pas bien représenter les queues de distribution. Dans la problématique qui nous intéresse, ce sont justement ces queues de distribution qui vont engendrer les situations les plus catastrophiques en termes de solvabilité, i.e. les valeurs extrêmes sur lesquelles vont reposer l'estimation de la VaR à 99,5 %. Ceci pose le problème de la robustesse de ce critère de VaR à 99,5 %, avec en ligne de mire le souci de la sécurité des assurés : les sociétés d'assurance pourraient être tentées, en jouant sur les paramètres ou sur les modèles des variables de base, de sous-estimer le SCR.

Après avoir décrit les spécificités du calcul de VaR en assurance, nous présentons différentes méthodes d'estimation de quantiles extrêmes, puis nous explicitons et illustrons les limites de la mesure de risque proposée pour déterminer le SCR dans le projet Solvabilité II. Dans un souci de clarté, les principaux résultats de la théorie des extrêmes utilisés dans ce travail sont présentés en annexe.

2 Calcul de VaR en assurance

La transposition de la VaR (et de la TVaR) aux problématiques d'assurance est récente et impose une approche radicalement différente de l'approche bancaire. En effet, la situation de référence du monde bancaire consiste à estimer la VaR sur un échantillon important de gains / pertes sur un portefeuille ou une position. Les données sont disponibles en quantité, avec une fréquence importante. Les approches de type *VaR historique* sont ainsi le socle sur lequel sont construits de nombreux modèles bancaires (cf. Christoffersen et al. (2001)). L'adaptation des techniques bancaires au portefeuille d'actifs d'un assureur nécessite un certain nombre d'aménagements, essentiellement pour tenir compte de la durée de détention plus importante. On

pourra par exemple se reporter à Fedor et Morel (2006) qui présentent des analyses quantitatives sur ce sujet.

Dans le cadre de la détermination du capital de solvabilité, de nouvelles difficultés apparaissent ; la nature des données disponibles invalidant l'approche historique. Il convient ici de revenir sur les situations d'assurance dans lesquelles on est amené à évaluer des VaR ; en pratique on peut en distinguer principalement deux, dont on va voir qu'elle imposent des approches différentes :

- le calcul d'une provision via un quantile à 75 % de la distribution des sinistres ;
- la détermination du niveau du capital de solvabilité (SCR) pour contrôler une probabilité de ruine à un an en imposant que celle-ci soit inférieure à 0,5 % à horizon un an.

La première situation constitue une simple évolution par rapport à la situation qui prévaut actuellement ; en effet, les provisions sont aujourd'hui calculées sur la base d'une espérance (approche *best estimate*, cf. Blum et Otto (1998)). On conçoit que, du point de vue de la technique statistique, les deux calculs (calcul d'une espérance et calcul d'une VaR à 75 %) ne soient pas fondamentalement distincts. En particulier on dispose dans les deux cas de données permettant de mettre en oeuvre les techniques classiques de statistique inférentielle. En pratique, on propose souvent une modélisation de la charge des sinistres à l'aide d'un modèle paramétrique et tout se ramène alors à l'estimation des paramètres de la loi considérée (cf. Partrat et Besson (2005) pour une revue des lois les plus utilisées en pratique).

Pour ce qui concerne la détermination du capital de solvabilité (SCR), la situation est radicalement différente ; en effet, on ne dispose que de très peu de données directes (quelques années au plus d'observation du résultat par exemple) et d'aucune donnée dans la zone de la distribution concernée (si on en avait, l'assureur aurait fait faillite...). Alors que dans le cas précédent la variable d'intérêt était directement observable, la VaR à calculer est maintenant la résultante d'un modèle, souvent complexe, mettant en jeu les différents postes du bilan de l'assureur : prestations, provisions, actifs financiers, etc.

La démarche va donc consister à construire un modèle en décrivant chaque poste du bilan ; en pratique l'approche consiste à modéliser le passif, dont le poste le plus important est constitué des provisions techniques,

puis l'actif et enfin les éventuelles interactions actif / passif. Après avoir estimé les paramètres de ces modèles, on peut obtenir, en général par des techniques de simulation, une estimation de la loi du résultat ; de cette estimation on déduira enfin le niveau de la VaR cherchée. La mise en pratique d'une telle démarche comporte à chaque étape des risques qu'il importe *a minima* d'identifier :

- un risque de modèle : le modèle utilisé n'offre qu'une représentation imparfaite de la réalité ; au surplus, les modèles usuellement employés tant à l'actif qu'au passif conduisent à sous-estimer les situations extrêmes ; nous reviendrons sur ce point potentiellement très pénalisant dans l'approche Solvabilité II ;
- un risque d'estimation : les paramètres estimés qui alimenteront le modèle sont entachés d'une erreur. Les conséquences de cette erreur peuvent être lourdes dans le cas d'un modèle peu robuste (voir par exemple Windcliff et Boyle (2004) pour une illustration dans le cas du modèle de Markowitz) ;
- un risque de simulation : la distribution du résultat sera, en général, obtenue par simulation, et ne sera donc qu'approchée.

De plus, s'agissant de l'estimation d'un quantile d'ordre élevé (VaR à 99,5 %), et la forme de la loi du résultat n'étant en général pas aisée à ajuster globalement à un modèle paramétrique, il faudra se tourner vers les techniques de valeurs extrêmes pour calculer finalement la mesure de risque, avec l'apparition ici d'un nouveau risque d'estimation, sur la queue de la distribution simulée.

La figure 1 propose une synthèse de ces différents risques et des situations dans lesquelles on les rencontre. Typiquement les études de variations extrêmes de variables financières classiques s'inscrivent dans le cadre des branches A3 ou A1 : on dispose d'un nombre important d'observations et les variables sont relativement régulières (elles possèdent généralement des moments d'ordre 3 ou 4). En assurance, l'estimation d'une VaR à 75 % dans le cadre de la détermination d'une provision Solvabilité II par exemple, correspondra, dans le meilleur cas, à la branche A1 pour les risques pour lesquels on dispose de beaucoup d'observations (risques non-vie à forte fréquence) et le plus fréquemment aux branches B4 ou B5 (on se donne un modèle, paramétrique le plus souvent, puis selon le cas on estime les paramètres et l'on en déduit la VaR ou l'on procède par les méthodes de Monte Carlo).

L'estimation du SCR procède, quant à elle, de la branche B3 qui cumule les risques de modèle, de simulation et d'estimation (à deux reprises).

Les risques d'estimation et de simulation ne sont pas spécifiques à la problématique du présent travail, et les techniques propres à les contrôler sont robustes et accessibles.

En revanche, le risque de modèle prend ici une importance particulière ; de fait, et comme on l'illustre *infra*, la mesure correcte du capital de solvabilité impose de refondre l'ensemble de la modélisation des postes du bilan pour prendre en compte correctement les événements extrêmes et éviter ainsi une sous-estimation du capital de solvabilité.

3 Notations

Considérons un n -échantillon de variables aléatoires (v.a.) X_1, \dots, X_n indépendantes et identiquement distribuées (i.i.d.) de fonction de répartition F et de fonction de survie $\bar{F} : x \mapsto 1 - F(x)$. La statistique d'ordre associée sera notée $X_{n,n}, \dots, X_{1,n}$ et est définie par

$$\min\{X_1, \dots, X_n\} = X_{n,n} \leq X_{n-1,n} \leq \dots \leq X_{1,n} = \max\{X_1, \dots, X_n\}. \quad (1)$$

Par ailleurs, on notera F_u la fonction de répartition de la v.a. ${}_uX$ qui représente le surplus au-delà du seuil u de X , lorsque X dépasse le seuil u , soit

$$F_u(x) := \Pr[{}_uX \leq x] := \Pr[X - u \leq x | X > u]. \quad (2)$$

On notera de plus x_F la borne supérieure du support de X de fonction de répartition F , soit

$$x_F := \sup \{x \in \mathbf{R}, F(x) < 1\}. \quad (3)$$

Enfin, nous noterons N_u le nombre d'observations qui dépassent le seuil u , soit

$$N_u := \sum_{i=1}^n \mathbf{1}_{X_i > u}. \quad (4)$$

4 Estimation de quantiles extrêmes

Plaçons-nous dans la situation standard où l'on observe directement des réalisations du phénomène dont on souhaite déterminer la valeur qu'il ne dé-

passera qu'avec une faible probabilité. L'objet de ce paragraphe est de préciser les différentes méthodes d'estimation possibles et de mesurer les erreurs d'estimation commises relativement à la quantité de données disponibles. Les différentes méthodes sont illustrées à partir d'un échantillon simulé de réalisations d'une variable aléatoire de loi de Pareto de première espèce.

4.1 Estimation naturelle

Comme $X_{k,n}$ est un estimateur naturel du quantile d'ordre $1 - (k - 1)/n$, un estimateur naturel de $F^{-1}(p)$ ($p \in]0; 1[$) est

$$([\![pn] - pn + 1) X_{n-[\![pn]-1, n} + (pn - [\![pn]) X_{n-[\![pn], n}, \quad (5)$$

où $[\cdot]$ désigne l'opérateur partie entière.

Cette méthode nécessite pour être efficace en pratique un volume de données jamais disponible dans notre contexte.

4.2 Ajustement à une loi paramétrique

Une méthode naturelle consiste à ajuster l'ensemble des données à une loi paramétrique puis à estimer la fonction quantile au niveau de probabilité souhaité. En effet, on rappelle que si $\hat{\theta}$ est l'estimateur du maximum de vraisemblance (e.m.v.) du paramètre θ de la loi de X , alors $F_{\hat{\theta}}^{-1}(p)$ est l'e.m.v. de $F^{-1}(p)$. De plus, les méthodes quantiles bootstrap BC (BCa) (cf. Efron et Tibshirani (1993)) permettent d'obtenir des intervalles de confiance de l'estimation.

Cette méthode d'estimation se décompose en deux étapes :

- ajustement à une loi paramétrique : choix d'une ou plusieurs lois, estimation des paramètres, tests d'adéquation ;
- inversion de la fonction de répartition (analytiquement quand c'est possible, numériquement sinon).

La principale difficulté de cette méthode consiste en le choix des lois paramétriques qui vont être utilisées. Elles doivent répondre à la double contrainte de permettre une évaluation numérique de leur fonction quantile aux ordres souhaités et doivent être en adéquation avec les données. Ce dernier point est particulièrement problématique dans la mesure où l'estimation des paramètres est effectuée sur l'ensemble de la distribution observée et représente rarement bien les valeurs extrêmes. À moins d'être assuré que la loi choisie pour l'ajustement est la véritable loi d'où sont issues les données,

ce qui est exceptionnel en pratique, l'intérêt de cette méthode apparaît très limité dans un contexte d'estimation de quantiles extrêmes.

4.3 Approximation GPD

Connue sous le nom de méthode POT (*Peaks Over Threshold*), cette technique repose sur le théorème 6 qui établit que, pour u assez grand, F_u peut être approximée par une distribution Pareto généralisée (GPD). Une fois les paramètres de la GPD estimés (cf. le paragraphe A.3), comme $\bar{F}(x) = \bar{F}(u)\bar{F}_u(x-u)$, on dispose de l'approximation suivante :

$$\bar{F}(x) \approx \frac{N_u}{n} \left(1 + \frac{\hat{\xi}}{\hat{\beta}}(x-u) \right)^{-1/\hat{\xi}}, \quad (6)$$

pour $x > u$. L'inversion de cette fonction de répartition nous fournit un estimateur de la fonction quantile F^{-1} aux ordres élevés :

$$\hat{x}_p = u + \frac{\hat{\beta}}{\hat{\xi}} \left(\left(\frac{n}{N_u}(1-p) \right)^{-\hat{\xi}} - 1 \right). \quad (7)$$

N.B. Si l'on choisit comme seuil une des observations $X_{k+1,n}$ (par exemple le quantile empirique à 99 % pour estimer le quantile à 99,5 %) alors $N_u = k$ et l'estimateur se réécrit de la manière suivante

$$\hat{x}_p = X_{k+1,n} + \frac{\hat{\beta}}{\hat{\xi}} \left(\left(\frac{n}{k}(1-p) \right)^{-\hat{\xi}} - 1 \right), \quad (8)$$

pour $p > 1 - k/n$.

Cette méthode présente en pratique la difficulté du choix du seuil u . En effet, u doit être assez grand pour que l'approximation GPD soit bonne mais pas trop proche du quantile à estimer de manière à ce que l'on dispose de suffisamment de matériel statistique pour que l'estimation de la probabilité d'être entre u et ce quantile (estimation qui fait apparaître N_u) soit fiable.

La problématique posée par le choix de u est similaire à celle du choix du nombre d'observations à utiliser dans le cadre de l'estimation de Hill de l'épaisseur de la queue de distribution (cf. le paragraphe C.2.4).

Une approche pratique visant à choisir le seuil u consiste à tracer

$$e_n(u) := \frac{1}{N_u} \sum_{i=1}^n [X_i - u]^+, \quad (9)$$

l'estimateur empirique de l'espérance résiduelle de X et à choisir u de manière à ce que $e_n(x)$ soit approximativement linéaire pour $x \geq u$. En effet, comme la fonction espérance résiduelle d'une loi GPD de paramètre $\xi < 1$ est linéaire :

$$e(v) := \mathbb{E}[vX] = \frac{\beta + \xi v}{1 - \xi}, \quad (10)$$

on cherchera un u aussi petit possible sous la contrainte que l'estimateur empirique de l'espérance résiduelle des excès au delà de u soit approximativement linéaire.

Il ne faut pas attendre de cette technique la bonne valeur de u , elle fournit néanmoins une aide précieuse. En pratique, plusieurs valeurs de u doivent être testées.

4.4 Estimateur de Hill

Considérons les fonctions de répartition appartenant au domaine d'attraction maximum (DAM) de Fréchet ($\xi > 0$). Le théorème 5 nous indique que $\bar{F}(x) = x^{-1/\xi} \mathcal{L}_F(x)$, où \mathcal{L}_F est une fonction à variation lente. Ainsi, on a

$$\frac{\bar{F}(x)}{\bar{F}(X_{k+1,n})} = \frac{\mathcal{L}(x)}{\mathcal{L}(X_{k+1,n})} \left(\frac{x}{X_{k+1,n}} \right)^{-1/\xi}. \quad (11)$$

Pour k suffisamment petit et $x \geq X_{k+1,n}$, on peut négliger (au premier ordre) le rapport des fonctions à variation lente ; il vient alors

$$\bar{F}(x) \approx \bar{F}(X_{k+1,n}) \left(\frac{x}{X_{k+1,n}} \right)^{-1/\xi}. \quad (12)$$

D'où l'on déduit l'estimateur de F :

$$\hat{F}(x) = 1 - \frac{k}{n} \left(\frac{x}{X_{k+1,n}} \right)^{-1/\xi_k^H}, \quad (13)$$

pour $x > X_{k+1,n}$. L'inversion de cette fonction nous donne l'estimateur de Hill de la fonction quantile pour des ordres élevés ($p > 1 - k/n$) :

$$\hat{x}_p^H = X_{k+1,n} \left(\frac{n}{k} (1 - p) \right)^{-\xi_k^H}. \quad (14)$$

Le problème du choix de k est similaire à celui rencontré dans l'estimation du paramètre de queue ξ (cf. l'annexe C.2.4).

4.5 Illustration

La figure 2 illustre les différentes techniques dans le cadre de l'estimation d'un quantile extrême d'une loi de Pareto de première espèce, i.e. de fonction de répartition

$$F(x) = 1 - \left(\frac{x_0}{x}\right)^\alpha, \text{ pour } x > x_0. \quad (15)$$

Une telle loi appartient au DAM de Fréchet. La méthode POT (exacte dans le cas d'une loi de Pareto de première espèce) et l'estimateur de Hill donnent des résultats comparables et nettement plus précis que l'estimateur naturel.

5 Robustesse du SCR

L'objet de ce paragraphe est de préciser les limites d'un critère de type VaR extrême pour fixer le capital de solvabilité d'une société d'assurance. En effet la complexité du dispositif à mettre en place rend la pertinence des résultats issus du modèle interne toute relative. Deux points sont plus particulièrement développés et illustrés à l'aide d'un exemple simple : l'erreur d'estimation des paramètres des variables de base puis l'erreur de spécification de modèle de ces variables.

5.1 Estimation des paramètres des variables de base

Considérons le modèle interne simplifié inspiré des travaux de Deelstra et Janssen (1998) dans lequel les sinistres de l'année sont modélisés par une variable aléatoire de distribution log-normale et payés en fin d'année. Supposons de plus que le rendement des actifs financiers sur la période soit gaussien et indépendant de la sinistralité.

Formellement, notons $B \sim \mathcal{LN}(m, s)$ la v.a. modélisant le montant à payer en fin de période et $\rho \sim \mathcal{N}(\mu, \sigma)$ le rendement financier sur la période. Déterminons le montant minimal d'actif a_0 dont doit disposer l'assureur en début de période pour ne pas être en ruine, en fin de période, avec une probabilité supérieure à $1 - \alpha$. Formellement, a_0 est la solution d'un problème d'optimisation :

$$a_0 = \min \{a \geq 0 \mid \Pr [ae^\rho \geq B] > 1 - \alpha\}. \quad (16)$$

Comme $Be^{-\rho} \sim \mathcal{LN}(m - \mu, \sqrt{s^2 + \sigma^2})$, on dispose d'une expression analytique pour la valeur de a_0 :

$$a_0 = \exp \left\{ (m - \mu) + \sqrt{s^2 + \sigma^2} \Phi^{-1}(1 - \alpha) \right\}. \quad (17)$$

Étudions à présent le sensibilité de a_0 aux paramètres des modèles de base. Par exemple pour le passif, on a

$$\frac{1}{a_0} \frac{\partial a_0}{\partial m} = 1, \quad (18)$$

et

$$\frac{1}{a_0} \frac{\partial a_0}{\partial s} = \frac{\Phi^{-1}(1 - \alpha)}{\sqrt{1 + \sigma^2/s^2}}. \quad (19)$$

Ainsi une erreur relative d'estimation du paramètre m conduit à la même erreur relative sur a_0 . Par ailleurs, une erreur relative de 1 % sur s conduit à une erreur relative de $\Phi^{-1}(1 - \alpha)/\sqrt{1 + \sigma^2/s^2}$ sur a_0 , ce qui dans le cas où $s \simeq \sigma$, pour une VaR à 99,5 % représente une erreur 1,82 fois plus grande.

5.2 Simulation

La complexité inhérente à la modélisation du résultat d'une activité d'assurance rend incontournable le recours aux techniques de simulation pour obtenir des résultats numériques.

Si le principe de base de ces méthodes est simple et universel (utilisation de la convergence forte de la Loi des grand nombres), une mise en oeuvre efficace nécessite de prendre quelques précautions. En effet, l'utilisation de résultats approchés par simulation est source de différentes erreurs :

- fluctuations d'échantillonnage liées au nombre fini de tirages effectués ;
- biais de discrétisation lors de la transformation d'un modèle continu dans sa version discrète ;
- erreurs associées aux approximations utilisées par certaines techniques d'inversion ;
- biais induits par un choix mal approprié du générateur de nombres aléatoires.

Aussi, les algorithmes utilisés devront permettre un contrôle quantitatif de l'ensemble de ces sources d'erreurs afin de pouvoir calibrer le nombre de tirages nécessaire pour le degré de précision (un niveau de confiance étant fixé *a priori*) requis.

5.3 Spécification du modèle

5.3.1 Contexte et motivation

La construction d'un modèle interne nécessite la modélisation des différentes variables influant sur la solvabilité de l'assureur. Pour ce faire, la première étape consiste en la modélisation de ces différentes variables. Dans le cas d'ajustements paramétriques, il est naturel d'effectuer l'estimation des paramètres et les tests d'adéquation sur l'intégralité des données disponibles. Cependant, dans notre problématique, ce sont les queues de distribution qui vont influencer sur le niveau du SCR. Or celles-ci sont souvent mal décrites par l'approche globale utilisant une loi paramétrique simple unique. En particulier, la plupart des modèles usuels (log-normal, Benktander, etc.) conduisent à des lois dans le DAM de Gumbel alors que les observations tendent à opter pour des queues de distribution dans le DAM de Fréchet. En effet, les modèles stochastiques qui fonctionnent actuellement chez les assureurs ont initialement été créés pour effectuer des études de rentabilité ou calculer des provisions techniques ; les résultats étant, le plus souvent, appréciés par des critères de type espérance-variance (ou VaR dans le coeur de la distribution - 75 % par exemple), la modélisation des événements extrêmes a dans ce contexte une influence toute relative. Dans l'approche Solvabilité II, on ne s'intéresse quasiment plus qu'aux valeurs extrêmes de la variable d'intérêt : le capital dont on doit disposer aujourd'hui pour ne pas être en ruine dans un an. Cette remarque est à nuancer par le fait que la plupart des assureurs qui investissent dans le développement de modèles internes escomptent, qu'à quelques aménagements près, celui-ci leur permettra d'affiner leur plan stratégique et de répondre aux futures exigences comptables résultant du passage à la phase II de la norme dédiée aux contrats d'assurance.

Considérons le modèle présenté ci-dessus. La figure 3 présente un échantillon de 1000 réalisations du montant d'actif dont doit disposer la société en 0 pour ne pas être en ruine en 1. Les cinq points soulignés correspondent aux cinq scénarios qui ont conduits aux valeurs maximales. On remarque sur la figure 4, qui représente les probabilités associées à chaque grandeur, que ces points se situent dans la queue de distribution d'au moins une des deux variables de base. Or justement aux queues de distribution, l'adéquation des données au modèle retenu est souvent imparfaite.

5.3.2 Modélisations avancées

Considérons la figure 5 qui représente le QQ-plot normal du rendement journalier du titre TOTAL (de juillet 2005 à avril 2006). On peut d'ores et déjà observer graphiquement que l'ajustement semble globalement satisfaisant mais que les queues de distribution du modèle (en particulier celle des rendements négatifs pour ce qui nous intéresse) sont trop fines. Les tests statistiques de Jarque-Béra et de Lilliefors au seuil 5 % conduisent en effet à ne pas rejeter l'hypothèse de normalité du rendement. Cependant, dans le cadre de la détermination d'un SCR sur le fondement d'un critère VaR à 99,5 %, le modèle gaussien tendrait à minimiser le risque pris en investissant sur ce placement. Par exemple, sur les données du titre TOTAL, le quantile à 0,5 % du rendement quotidien observé vaut $-0,0309$ alors que le quantile du même ordre du modèle gaussien ajusté vaut $-0,0286$, ce qui représente une erreur de l'ordre de 7,5 %.

Pour remédier à cela, dans les situations où l'on dispose d'un nombre conséquent d'observations, on pourrait être tenté d'adopter une approche non-paramétrique en utilisant la distribution empirique. Cela n'est toutefois pas satisfaisant en pratique, car de telles approches sont techniquement plus difficiles à implémenter, nécessitent des temps de simulation plus longs et pénalisent une compréhension simplifiée du modèle. Dans ce contexte, il est naturel de conserver une approche paramétrique et de rechercher un modèle qui représente mieux les queues de distributions.

Le modèle suivant est une extension naturelle du rendement gaussien. Supposons que le rendement de l'actif sur la période soit régi par le processus suivant

$$\rho = \mu + \sigma_0 \epsilon_0 + \sigma_u \sum_{i=1}^N \epsilon_i, \quad (20)$$

où les $\epsilon_0, \epsilon_1, \dots$ sont des v.a.i.i.d. de loi $\mathcal{N}(0, 1)$ et indépendantes de $N \sim \mathcal{P}(\lambda)$ ($\lambda > 0$). Ce processus est une version mono-périodique du processus de Merton (1976). L'estimation des paramètres par maximum de vraisemblance nécessite l'utilisation de techniques numériques (cf. Planchet et Thérond (2005)).

Dans notre problématique, on peut toutefois souhaiter calibrer manuellement les paramètres de manière à alourdir les queues de distribution. Par exemple, en fixant l'espérance et la variance globale du rendement (identiques à celle du modèle gaussien), on peut pondérer la variabilité représentée par

la composante à sauts de manière à ce que la queue de distribution du modèle soit aussi lourde que celle des observations. Formellement, cela revient à choisir λ , σ_0 et σ_u sous la contrainte :

$$\sigma_0^2 + \lambda\sigma_u^2 = \sigma^2. \quad (21)$$

La figure 6 représente cet ajustement. En la comparant avec la figure 5, on observe que les queues de distributions ont été alourdies.

D'une manière générale, la démarche illustrée *supra* dans le cas de la modélisation du rendement financier devra être reconduite pour l'ensemble des *variables de base* à l'actif et au passif sous peine de sous-estimer le SCR.

6 Conclusion

Le projet Solvabilité II a pour objectif d'établir un référentiel prudentiel harmonisé et cohérent à l'ensemble des sociétés d'assurance européennes. Pour cela, il prévoit notamment d'inciter les sociétés à mieux modéliser leurs risques en les autorisant à construire des modèles internes qui aboutissent à un capital de solvabilité inférieur à celui fourni par la "formule standard".

Cependant, nous avons vu que la robustesse du critère VaR à 99,5 % sur des données non-observées mais générées par le modèle interne est toute relative. En particulier, les modèles stochastiques actuellement en place chez les assureurs (tant à l'actif qu'au passif) ne sont pas orientés vers l'estimation de valeurs extrêmes et sont, le plus souvent, construits à partir de modélisations des "variables de base" qui sous-estiment les queues de distribution. L'utilisation en l'état de tels modèles conduirait à sous-estimer le SCR.

Aussi les sociétés d'assurance qui souhaiteront développer un modèle interne devront modéliser plus finement les queues de distribution. Parallèlement, les autorités qui auditeront les modèles internes devront porter une attention particulière à celles-ci dans le processus de validation. En particulier, une harmonisation au niveau européen de ces procédures est indispensable sous peine d'introduire une distorsion de concurrence entre les différents intervenants.

Au global, la prise en compte du critère de contrôle d'une probabilité de ruine qui sous-tend le dispositif Solvabilité II implique donc une profonde refonte des modèles utilisés jusqu'alors dans les sociétés d'assurance.

A Loi de Pareto généralisée (GPD)

A.1 Définition

Définition A.1.1 Soit $H_{\xi,\beta}$ la fonction de répartition définie pour $\beta > 0$ par

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

Cette fonction de répartition correspond à la loi de Pareto généralisée (GPD) de paramètres ξ et β . Elle est définie pour $x > 0$ si $\xi > 0$ et pour $0 \leq x \leq -\beta/\xi$ si $\xi < 0$.

On notera dans la suite $D(\xi, \beta)$ le domaine de définition de $H_{\xi,\beta}$.

On montre (cf. Denuit et Charpentier (2005)) que la loi de Pareto généralisée peut être vue comme une log-gamma ou encore, dans le cas où $\xi > 0$, comme un mélange de lois exponentielles dont le paramètre suit une loi gamma.

A.2 Quelques propriétés

Les résultats suivants sont énoncés pour une variable aléatoire Y distribuée selon une GPD de paramètre (ξ, β) .

Propriété 1 Si $\xi < 1$, on a

$$\begin{aligned} \mathbb{E} \left(1 + \frac{\xi}{\beta} Y \right)^{-r} &= \frac{1}{1 + \xi r}, \text{ pour } r > -1/\xi, \\ \mathbb{E} \left[\ln \left(1 + \frac{\xi}{\beta} Y \right) \right]^k &= \xi^k k!, \text{ pour } k \in \mathbf{N}, \\ \mathbb{E} (Y \bar{H}_{\xi,\beta}^r) &= \frac{\beta}{(r+1-\xi)(r+1)}, \text{ pour } (r+1)/\xi > 0. \end{aligned}$$

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

$$\mathbb{E} [Y^r] = \frac{\beta^r \Gamma(\xi^{-1} - r)}{\xi^{r+1} \Gamma(\xi^{-1} + 1)} r!.$$

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

$$\mathbb{E} [Y - u | Y > u] = \frac{\beta + \xi u}{1 - \xi}, \text{ pour } \beta + u\xi > 0.$$

On rappelle que y_F est la borne supérieure du support de Y , soit

$$y_F = \sup \{y \in \mathbf{R}, F(y) < 1\}. \quad (22)$$

A.3 Estimation des paramètres

Considérons un n -échantillon (Y_1, \dots, Y_n) de la variable aléatoire Y de fonction de répartition $H_{\xi, \beta}$.

A.3.1 Méthode du maximum de vraisemblance

La densité f de Y vaut

$$f(y) = \frac{1}{\beta} \left(1 + \frac{\xi}{\beta} y\right)^{-1/\xi-1}, \text{ pour } y \in D(\xi, \beta). \quad (23)$$

On en déduit la log-vraisemblance

$$\ln L(\xi, \beta; Y_1, \dots, Y_n) = -n \ln \beta - \left(1 + \frac{1}{\xi}\right) \sum_{i=1}^n \ln \left(1 + \frac{\xi}{\beta} Y_i\right). \quad (24)$$

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

$$\begin{cases} \xi = \frac{1}{n} \sum_{i=1}^n \ln(1 + \tau Y_i) =: \hat{\xi}(\tau), \\ \frac{1}{\tau} = \frac{1}{n} \left(\frac{1}{\xi} + 1\right) \sum_{i=1}^n \frac{Y_i}{1 + \tau Y_i}. \end{cases} \quad (25)$$

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

$$\frac{1}{\tau} = \frac{1}{n} \left(\frac{1}{\hat{\xi}(\tau)} + 1\right) \sum_{i=1}^n \frac{Y_i}{1 + \tau Y_i}. \quad (26)$$

Cette dernière équation se résout numériquement de manière itérative pour autant que l'on dispose d'une valeur initiale τ_0 pas trop éloigné de τ . En pratique, cette valeur initiale pourra être obtenue par la méthode des moments (pour autant que ceux-ci existent jusqu'à l'ordre 2) ou par la méthode des quantiles.

Lorsque $\xi > -1/2$, Hosking et Wallis (1987) ont montré la normalité asymptotique des estimateurs du maximum de vraisemblance :

$$n^{1/2} \left(\hat{\xi}_n - \xi, \frac{\hat{\beta}_n}{\beta} - 1 \right) \rightarrow_d N \left[0, (1 + \xi) \begin{pmatrix} 1 + \xi & -1 \\ -1 & 2 \end{pmatrix} \right], \quad n \rightarrow \infty. \quad (27)$$

Ce résultat permet en particulier de calculer les erreurs approximatives d'estimation commises par les estimateurs du maximum de vraisemblance.

A.3.2 Méthode des moments

D'après les résultats du paragraphe A.2, si $\xi < 1/2$, les deux premiers moments de Y existent et l'on a

$$\mu_1 := E[Y] = \frac{\beta}{1 - \xi}, \quad (28)$$

et

$$\mu_2 := \text{Var}[Y] = \frac{\beta^2}{(1 - \xi)^2(1 - 2\xi)}. \quad (29)$$

On en déduit que

$$\xi = \frac{1}{2} - \frac{\mu_1^2}{2\mu_2} \quad \text{et} \quad \beta = \frac{\mu_1}{2} \left(1 + \frac{\mu_1^2}{\mu_2}\right). \quad (30)$$

En remplaçant μ_1 et μ_2 par leurs estimateurs empiriques, on obtient les estimateurs de la méthode des moments $\hat{\xi}_{MM}$ et $\hat{\beta}_{MM}$.

Ces estimateurs sont simples à mettre en oeuvre, mais ne fonctionnant que pour $\xi < 1/2$, ils nécessitent la connaissance *a priori* de cette information.

A.3.3 Méthode des moments pondérés par les probabilités

Une alternative intéressante à la méthode des moments a été proposée par Hosking et Wallis (1987). En remarquant que

$$\omega_r := E[Y \bar{H}_{\xi, \beta}^r(Y)] = \frac{\beta}{(r+1)(r+1-\xi)}, \quad \text{pour } r = 0; 1, \quad (31)$$

on obtient

$$\beta = \frac{2\omega_0\omega_1}{\omega_0 - 2\omega_1} \quad \text{et} \quad \xi = 2 - \frac{\omega_0}{\omega_0 - 2\omega_1}. \quad (32)$$

En remplaçant ω_0 et ω_1 par leurs estimateurs empiriques, on obtient les estimateurs des moments pondérés par les probabilités $\hat{\xi}_{PWM}$ et $\hat{\beta}_{PWM}$.

Dans le cas d'échantillons de taille réduite, Hosking et Wallis (1987) montrent que, lorsque $\xi < 1/2$, ces estimateurs sont plus efficaces que ceux du maximum de vraisemblance. Le domaine de validité de cette méthode est sa principale limite à une utilisation en assurance : si la plupart des risques admettent des moments d'ordre 2, ce n'est pas une généralité (tempêtes,

tremblements de terre, risques industriels, responsabilité civile, etc.). Des généralisations (Diebolt et al. (2005b)) permettent d'étendre cette méthode à $\xi < 3/2$.

A.3.4 Méthode des quantiles

La fonction quantile d'une loi GPD de paramètre (β, ξ) est donnée par

$$q_p := H_{\xi, \beta}^{-1}(p) = \frac{\beta}{\xi} \left[(1-p)^{-\xi} - 1 \right]. \quad (33)$$

En remarquant que

$$\frac{q_{p_2}}{q_{p_1}} = \frac{(1-p_2)^\xi - 1}{(1-p_1)^\xi - 1}, \text{ pour } p_1, p_2 \in]0; 1[, \quad (34)$$

la solution $\hat{\xi}_{MQ}$ de

$$\frac{Q_{p_2}}{Q_{p_1}} = \frac{(1-p_2)^\xi - 1}{(1-p_1)^\xi - 1}, \quad (35)$$

où Q_{p_1} et Q_{p_2} sont les quantiles empiriques d'ordres p_1 et p_2 , est l'estimateur des quantiles de ξ .

A.3.5 Méthodes bayésiennes

Des développements récents (cf. Coles et Powell (1996) ou Diebolt et al. (2005a)) proposent des approches bayésiennes pour estimer les paramètres de la GPD. Utilisant des algorithmes *Markov Chain Monte Carlo* (MCMC), ces méthodes permettent d'intégrer une information *a priori* (avis d'expert) dans des contextes où l'on dispose d'un nombre réduit d'observations.

B Résultats probabilistes

Sont successivement rappelés dans ce paragraphe des résultats de calcul des probabilités concernant la loi du maximum, l'épaisseur des queues de distribution et la loi des excès au-delà d'un seuil. Les différents résultats sont démontrés dans Embrechts et al. (1997).

B.1 Loi du maximum

Définition B.1.1 *Deux v.a. X et Y sont dites de même type s'il existe deux constantes $a \in \mathbf{R}$ et $b > 0$ telles que $Y \stackrel{d}{=} a + bX$.*

Théorème 4 (Fisher-Tippett-Gnedenko) *Considérons une suite X_1, X_2, \dots de v.a.i.i.d. S'il existe une suite de réels (a_n) , une suite positive (b_n) et une loi non-dégénérée H telles que*

$$\frac{X_{1,n} - a_n}{b_n} \rightarrow_d H,$$

alors H est du même type qu'une des lois suivantes :

$$G(x) = \exp \left\{ -(1 + \xi x)^{-1/\xi} \right\}, \quad \text{pour } 1 + \xi x \geq 0,$$

$$\Lambda(x) = \exp \{ -\exp(-x) \}, \quad \text{pour } x \in \mathbf{R}.$$

Les fonctions de répartition $G_+(\xi > 0)$, $\Lambda(\xi = 0)$ et $G_-(\xi < 0)$ correspondent respectivement aux lois de Fréchet, de Gumbel et de Weibull.

Définition B.1.2 *On dira d'une fonction de répartition F qui répond aux hypothèses du 4 qu'elle appartient au domaine d'attraction maximum (DAM) de H .*

B.2 Epaisseur des queues de distribution

Définition B.2.1 *Une fonction $h : \mathbf{R}_+^* \rightarrow \mathbf{R}_+$ est dite à variation régulière (en $+\infty$) d'indice α si h vérifie*

$$\lim_{x \rightarrow \infty} \frac{h(tx)}{h(x)} = t^\alpha \text{ pour tout } t > 0.$$

Si $\alpha = 0$, on parlera de variation lente ; si $\alpha = \infty$, de variation rapide.

Théorème 5 (Théorème Taubérien) *Une v.a. de fonction de répartition F et de transformée de Laplace L_F est à variation régulière d'indice $-\alpha$ ($\alpha \leq 0$) si les conditions équivalentes suivantes sont vérifiées ($\mathcal{L}_F, \mathcal{L}_{F^{-1}}, \mathcal{L}_L$ et \mathcal{L}_f désignent des fonctions à variation lente) :*

- (i) \bar{F} est à variation régulière d'indice $-\alpha$, i.e. $\bar{F}(x) = x^{-\alpha} \mathcal{L}_F(x)$.
- (ii) La fonction quantile est à variation régulière :

$$F^{-1}(1 - 1/x) = x^{1/\alpha} \mathcal{L}_{F^{-1}}(x).$$

- (iii) La transformée de Laplace de F vérifie $L_F(t) = t^\alpha \mathcal{L}_L(1/t)$.

- (iv) Si la densité existe et vérifie $xf(x)/\bar{F}(x) \rightarrow \alpha$ quand $x \rightarrow \infty$, alors la densité est à variation régulière d'indice $-(1 + \alpha)$: $f(x) = x^{-(1+\alpha)} \mathcal{L}_f(x)$.

La condition relative à la transformée de Laplace permet d'établir que la propriété de variation régulière à paramètre fixé est stable par convolution.

B.3 Loi des excès au-delà d'un seuil

Théorème 6 (Pickands-Balkema-de Haari) *Une fonction de répartition F appartient au domaine maximum d'attraction de G_ξ si, et seulement si, il existe une fonction positive $\beta(\cdot)$ telle que*

$$\lim_{u \rightarrow x_F} \sup_{x > 0} \{|_u F(x) - H_{\xi, \beta(u)}(x)|\} = 0.$$

Ce théorème établit le lien entre le paramètre de la loi du domaine d'attraction maximum et le comportement limite des excès au-delà d'un seuil grand. En particulier, l'indice de queue ξ est identique au paramètre de la loi GPD qui décrit le coût résiduel des sinistres dépassant un seuil suffisamment élevé.

Ceci permet notamment de distinguer les lois à queue épaisse qui appartiennent au DAM de Fréchet ($\xi > 0$) des lois à queue fine qui appartiennent au DAM de Gumbel ($\xi = 0$). Le tableau suivant indique le comportement limite de certaines lois usuelles.

Lois à queue épaisse $\xi > 0$	Lois à queue fine $\xi = 0$	Lois bornées à droite $\xi < 0$
Cauchy	Gamma	Uniforme
Pareto	normale	Beta
log-gamma	log-normale	
Student	Weibull	
α -stable ($\alpha < 2$)	Benktander	

La propriété suivante concerne le nombre N_u de dépassements d'un seuil u assez élevé.

Propriété 7 *Le nombre de dépassements du seuil u_n dans un échantillon de taille n est asymptotiquement distribué selon une loi de Poisson pour autant que la probabilité de dépasser u_n diminue proportionnellement à l'inverse de la taille de l'échantillon. Formellement, on a*

$$\lim_{n \rightarrow \infty} n\bar{F}(u_n) = \tau \Rightarrow N_{u_n} \rightarrow_d \mathcal{P}(\tau).$$

C Estimation du paramètre de queue

L'épaisseur de la queue de la fonction de répartition F est résumée par le paramètre ξ de la loi dont elle fait partie du DAM.

Ainsi les lois appartenant au DAM de la loi de Fréchet ($\xi > 0$) voient leur queue décroître en fonction puissance tandis que celles appartenant au DAM de Gumbel ($\xi = 0$) ont des queues qui décroissent de manière exponentielle.

Les estimateurs de quantiles extrêmes faisant appel à ce paramètre de queue ξ , son estimation doit faire l'objet d'une attention particulière.

L'objet de cette section est donc de présenter les principales méthodes d'estimation de l'épaisseur de la queue de distribution. Sont ainsi notamment présentés les estimateurs de Pickands (1975), de Hill (1975) et de Dekkers et al. (1989).

C.1 Méthodes paramétriques

C.1.1 Ajustement à la loi du maximum

Le théorème 4 nous donne la loi asymptotique de $X_{1,n}$. A supposer que l'on dispose de réalisations de cette v.a., i.e. de m échantillons permettant d'observer m réalisations $x_{1,n}^{(1)}, \dots, x_{1,n}^{(m)}$ de $X_{1,n}$, la méthode du maximum de vraisemblance permettrait d'estimer les paramètres de la loi limite et en particulier ξ .

C.1.2 Ajustement de la loi limite des excès

Considérons une distribution F appartenant au domaine d'attraction maximum de G_ξ . D'après le 6, il existe une fonction positive $\beta(\cdot)$ telle que

$$\lim_{u \rightarrow x_F} \sup_{x > 0} \{ |uF(x) - H_{\xi, \beta(u)}(x)| \} = 0. \quad (36)$$

En particulier, le paramètre ξ de la distribution Pareto généralisée $H_{\xi, \beta}$ est identique à celui de G_ξ . Ainsi pour un seuil u assez élevé, la distribution $H_{\xi, \beta(u)}$ est une bonne approximation de F_u dont on dispose de N_u observations $X_{N_u, n} - u, \dots, X_{1, n} - u$. Les techniques présentées au paragraphe A.3 permettent alors d'estimer ξ .

En particulier, si $\xi > -1/2$ est estimé par la méthode du maximum de vraisemblance, Smith (1987) montre que la variance asymptotique de cet estimateur vaut $(1 + \xi)^2 / N_u$.

C.2 Méthodes semi-paramétriques

C.2.1 Estimateur de Pickands

Propriété 8 Pour $k/n \rightarrow 0$, l'estimateur de Pickands, défini par

$$\xi_k^P = \frac{1}{\ln 2} \ln \frac{X_{k,n} - X_{2k,n}}{X_{2k,n} - X_{4k,n}},$$

est un estimateur convergent de ξ . De plus, sous certaines conditions supplémentaires portant sur k et F (cf. Dekkers et de Haan (1989)), il est asymptotiquement gaussien :

$$\sqrt{k} (\xi_k^P - \xi) \rightarrow_d \mathcal{N} \left(0, \frac{\xi^2 (2^{\xi+1} + 1)}{(2(2^\xi - 1) \ln 2)^2} \right).$$

C.2.2 Estimateur de Hill

Propriété 9 Pour $k/n \rightarrow 0$, l'estimateur de Hill défini par

$$\xi_k^H = \frac{1}{k} \sum_{j=1}^k \ln \frac{X_{j,n}}{X_{k,n}},$$

est un estimateur convergent de ξ . De plus, sous certaines conditions sur k et F (cf. de Haan et Peng (1998)), il est asymptotiquement gaussien :

$$\sqrt{k} (\xi_k^H - \xi) \rightarrow_d \mathcal{N} (0, \xi^2).$$

Bien que plus performant que l'estimateur de Pickands (cf. le rapport des variances asymptotiques), l'estimateur de Hill n'est utilisable que pour les distributions de Fréchet ($\xi > 0$).

C.2.3 Estimateur de Dekkers-Einmahl-de Haan

L'estimateur de Dekkers-Einmahl-de Haan est une généralisation de l'estimateur de Hill à l'ensemble des lois extrêmes ($\xi \in \mathbf{R}$).

Propriété 10 Pour $k/n \rightarrow 0$, l'estimateur de Dekkers-Einmahl-de Haan défini par

$$\xi_k^{DEdH} = 1 + \xi_k^{(1)} - \frac{1}{2} \left(1 - \frac{(\xi_k^{(1)})^2}{\xi_k^{(2)}} \right)^{-1},$$

avec $\xi_k^{(i)} = \frac{1}{k} \sum_{j=1}^k \left(\ln \frac{X_{j,n}}{X_{k+1,n}} \right)^i$, est un estimateur convergent de ξ . De plus, sous certaines conditions sur k et F (cf. Dekkers et al. (1989)), il est asymptotiquement gaussien :

$$\sqrt{k} (\xi_k^P - \xi) \rightarrow_d \mathcal{N} (0, 1 + \xi^2).$$

Cet estimateur est également connu sous l'appellation d'estimateur des moments, les $\xi_k^{(i)}$ pouvant s'interpréter comme des moments empiriques.

C.2.4 Nombre d'observations à utiliser

Les résultats concernant les estimateurs de l'indice de queue énoncés précédemment sont asymptotiques : ils sont obtenus lorsque $k \rightarrow \infty$ et $k/n \rightarrow 0$. Comme en pratique, on ne dispose que d'un nombre d'observations n fini, il s'agit de choisir k de manière à ce que l'on dispose de suffisamment de matériel statistique (les $X_{k,n}, \dots, X_{1,n}$) tout en restant dans la queue de distribution ($k \ll n$).

En particulier, l'estimateur de Hill satisfaisant la propriété asymptotique

$$\sqrt{k} (\xi_k^H - \xi) \rightarrow_d \mathcal{N} (0, \xi^2), \quad (37)$$

pour $k \rightarrow \infty$ avec un *certain taux de croissance* en fonction de n , on pourrait être tenté de choisir k aussi grand que possible de manière à minimiser l'erreur quadratique moyenne commise par ξ_k^H . Cependant, le comportement au deuxième ordre de la fonction à variation lente \mathcal{L}_F introduite dans le théorème 5 induit un biais lorsque k est trop grand. Des solutions pour fixer k de manière à disposer d'un estimateur asymptotiquement sans biais ont notamment été proposées par Goldie et Smith (1987) puis de Haan et Peng (1998).

Concernant l'estimateur de Hill pour des fonctions appartenant au DAM de Fréchet, de Haan et Peng (1998) ont proposé de retenir le nombre d'observation k^* qui réduit l'erreur quadratique moyenne de l'estimateur de Hill, i.e.

$$k^*(n) = \begin{cases} 1 + n^{2\xi/(2\xi+1)} \left(\frac{(1+\xi)^2}{2\xi} \right)^{1/(2\xi+1)}, & \text{si } \xi \in]0, 1[, \\ 2n^{2/3}, & \text{si } \xi > 1. \end{cases} \quad (38)$$

Toutefois, on remarque que k^* s'exprime en fonction de ξ qui n'est pas observé. De plus ce critère ne concerne que l'estimateur de Hill.

En pratique, il sera préférable de suivre la démarche proposée par Embrechts et al. (1997) qui suggèrent de tracer les estimateurs en fonction de k/n et de choisir k dans un intervalle où la droite des estimations est approximativement horizontale (cf. la figure 7 détaillée *infra*).

C.2.5 Illustration

La figure 7 illustre les différents estimateurs du paramètre de queue d'une loi de Pareto de première espèce (de fonction de répartition $F(x) = 1 - (x_0/x)^\alpha$, pour $x > x_0$). Cette famille de distributions faisant partie du DAM de Fréchet, les estimations de Pickands, de Hill et de Dekkers-Einmahl-de Haan ont pu être tracées en fonction de k . On observe la suprématie de l'estimateur de Hill sur ceux de Pickands et de Dekkers-Einmahl-de Haan. Par ailleurs, pour $k < 0,02n$, l'estimateur de Hill est relativement volatile. On serait donc amené à utiliser de l'ordre de 2,5 % des données les plus extrêmes pour estimer l'épaisseur de la queue de distribution.

Références

- Blum, K. A., Otto, D. J. (1998) Best estimate loss reserving : an actuarial perspective. CAS Forum Fall 1, 55–101.
- Christoffersen, P., Hahn, J., Inoue, A. (2001) Testing and comparing value-at-risk measures. *Journal of Empirical Finance* 8 (3), 325–342.
- Coles, S., Powell, E. (1996) Bayesian methods in extreme value modelling : a review and new developments. *Internat. Statist. Rev.* 64, 119–36.
- de Haan, L., Peng, L. (1998) Comparison of tail index estimators. *Statistica Neerlandica* 52 (1), 60–70.
- Deelstra, G., Janssen, J. (1998) Interaction between asset liability management and risk theory. *Applied Stochastic Models and Data Analysis* 14, 295–307.
- Dekkers, A., de Haan, L. (1989) On the estimation of the extreme-value index and large quantile estimation. *Annals of Statistics* 17, 1795–832.
- Dekkers, A., Einmahl, J., de Haan, L. (1989) A moment estimator for the index of an extreme-value distribution. *Annals of Statistics* 17, 1833–55.
- Denuit, M., Charpentier, A. (2005) *Mathématiques de l'assurance non-vie*. Vol. 2. Economica, Paris.
- Diebolt, J., El-Aroui, M., Garrido, S., Girard, S. (2005a) Quasi-conjugate bayes estimates for gpd parameters and application to heavy tails modelling. *Extremes* 8, 57–78.
- Diebolt, J., Guillou, A., Rached, I. (2005b) Approximation of the distribution of excesses using a generalized probability weighted moment method. *C. R. Acad. Sci. Paris, Ser. I* 340 (5), 383–8.
- Efron, B., Tibshirani, R. J. (1993) *An introduction to the bootstrap*. Chapman & Hall.
- Embrechts, P., Klüppelberg, C., Mikosch, T. (1997) *Modelling Extremal Events for Insurance and Finance*. Springer Verlag, Berlin.
- Fedor, M., Morel, J. (2006) Value-at-risk en assurance : recherche d'une méthodologie à long terme. Actes du 28e congrès international des actuaires, Paris.
- Goldie, C., Smith, R. (1987) Slow variation with remainder : a survey of the theory and its applications. *Quarterly Journal of Mathematics Oxford* 38 (2), 45–71.
- Hill, B. (1975) A simple general approach to inference about the tail of a distribution. *Annals of Statistics* 3, 1163–74.

- Hosking, J. R., Wallis, J. R. (1987) Parameter and quantile estimation for the generalized pareto distribution. *Technometrics* 29, 339–49.
- Merton, R. (1976) Option pricing when underlying stock returns are discontinuous. *Journal of Financial Economics* 3, 125–144.
- Partrat, C., Besson, J.-L. (2005) Assurance non-vie. Modélisation, simulation. Economica, Paris.
- Pickands, J. (1975) Statistical inference using extreme orders statistics. *Annals of Statistics* 3, 119–31.
- Planchet, F., Thérond, P.-E. (2005) L’impact de la prise en compte des sauts boursiers dans les problématiques d’assurance. Proceedings of the 15th AFIR Colloquium, Zürich.
- Smith, R. L. (1987) Estimating tails of probability distributions. *Annals of Statistics* 15, 1174–207.
- Windcliff, H., Boyle, P. P. (2004) The $1/n$ pension investment puzzle. *North American Actuarial Journal* 8 (3), 32–45.

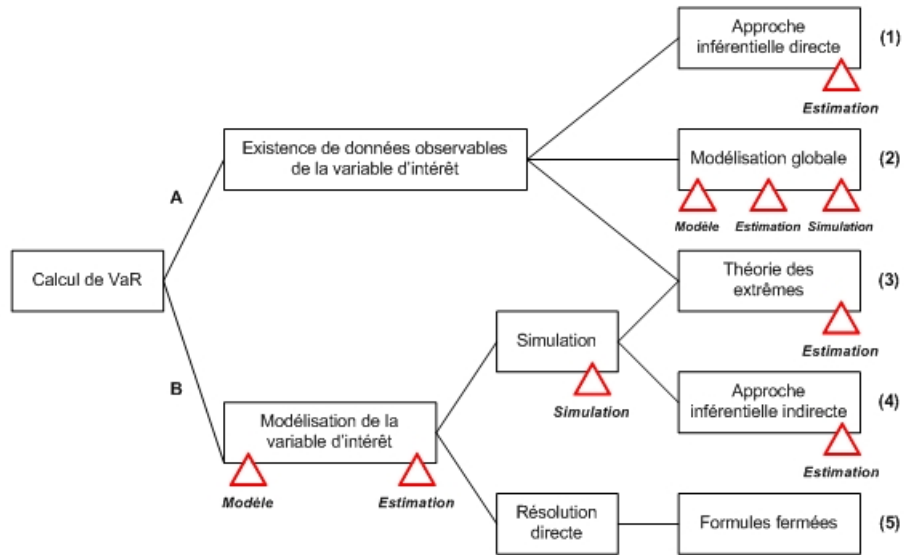


FIG. 1 – Typologie des différents risques rencontrés

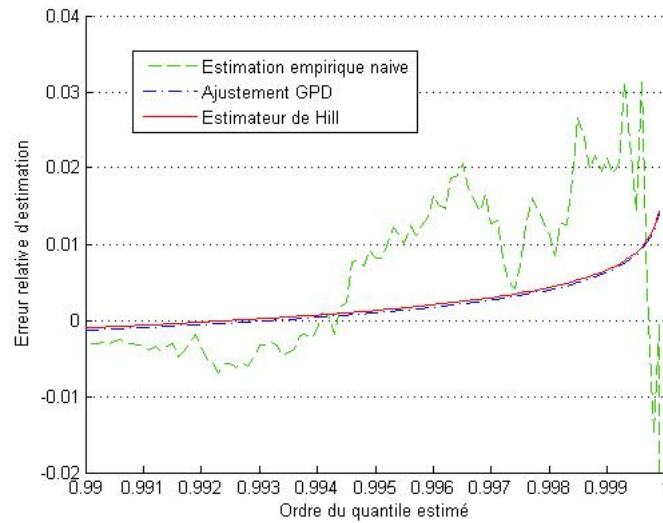


FIG. 2 – Estimation d'un quantile extrême : erreur relative d'estimation

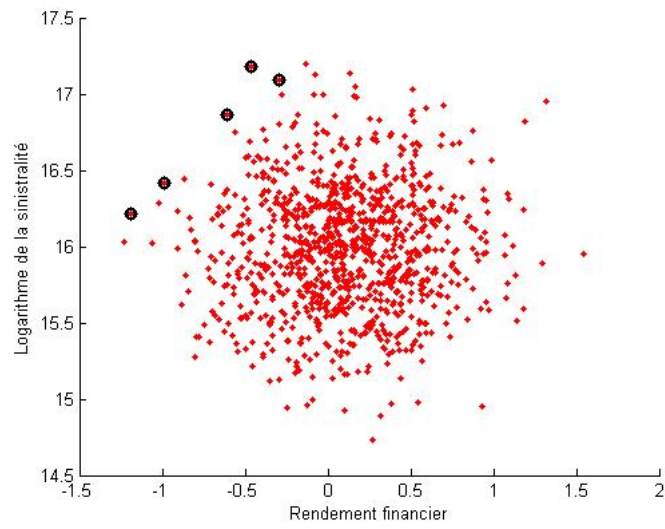


FIG. 3 – Modèle interne simplifié : identification des valeurs extrêmes

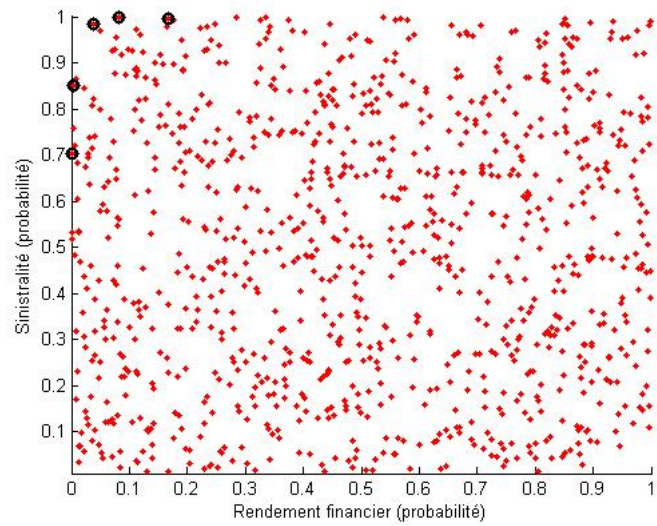


FIG. 4 – Modèle interne simplifié : identification des valeurs extrêmes

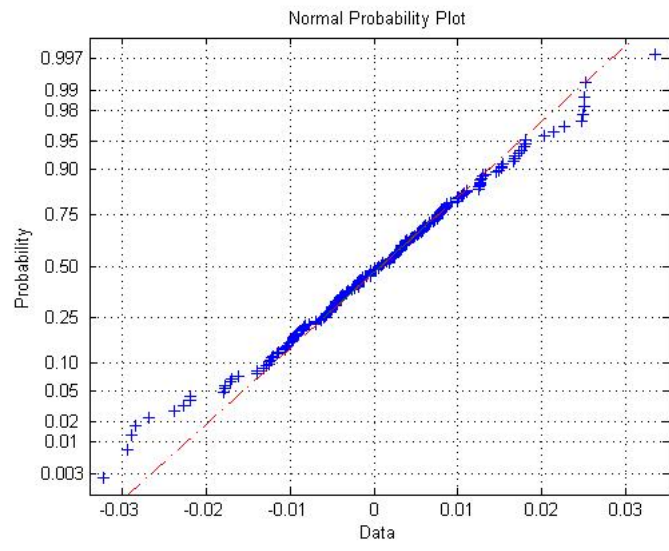


FIG. 5 – Rendement journalier du titre TOTAL : QQ-plot loi empirique vs loi normale

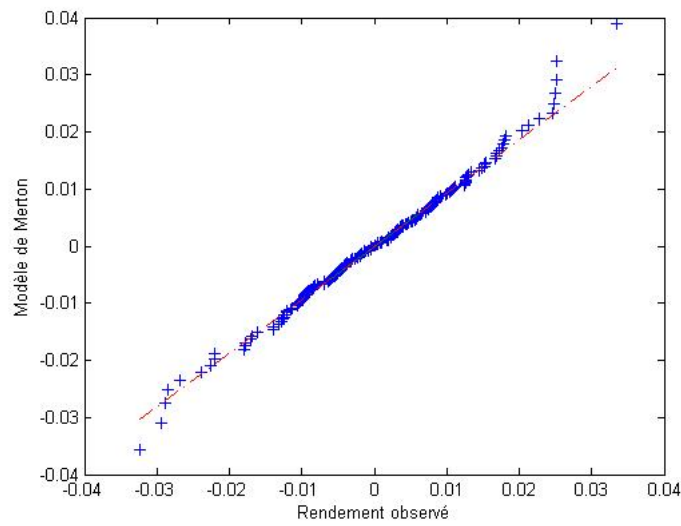


FIG. 6 – Rendement journalier du titre TOTAL : QQ-plot loi empirique vs modèle de Merton

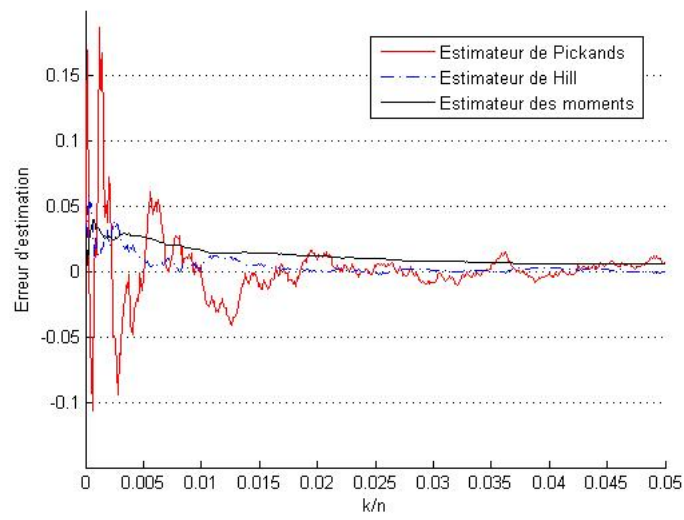


FIG. 7 – Estimation de l'épaisseur de la queue d'une distribution de Pareto