Théorie des valeurs extrêmes

De nouveaux types de risques bouleversent quotidiennement le travail des actuaires et des *risk managers*. La théorie des valeurs extrêmes peut toutefois leur permettre de mieux modéliser ces risques.

En quelques années, les assureurs ont subi des événements majeurs dont la probabilité d'occurrence est nulle ou presque si l'on s'en tient aux modèles probabilistes classiques. En effet, qu'il s'agisse d'événements climatiques, terroristes ou encore des fluctuations sur les marchés financiers (crise des *subprimes*, de la liquidité et de la dette souveraine), leur manifestation interpelle les assureurs sur la prise en compte de ces phénomènes dans la gestion des risques. À l'heure où Solvabilité II consacre une approche de *Value-at-Risk* à un niveau élevé (99,5%) pour déterminer le niveau de capital de solvabilité requis, il semble opportun de se tourner vers la théorie des valeurs extrêmes et les outils qu'elle met à disposition des *risk managers* pour mieux modéliser et donc appréhender ces risques.

Domaine d'attraction du maximum

Dans un premier temps, intéressons-nous au comportement extrême du maximum d'un n-échantillon i.i.d. Le théorème de Fisher-Tippett nous donne le résultat suivant.
Si le maximum normalisé[1] d'une suite de n variables aléatoires i.i.d. converge en loi vers une distribution non dégénérée H, alors H est une transformation d'une des lois suivantes :

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

Les fonctions de répartition

$$G_+(\xi > 0),\ \Lambda(\xi = 0) \text{ et } G_-(\xi < 0)$$

correspondent respectivement aux lois de Fréchet, de Gumbel et de Weibull.

Ce résultat fondamental conduit à définir les domaines d'attraction maximum (DAM) de Fréchet, Gumbel et Weibull, auxquels appartiennent nombre de distributions continues usuelles en assurance et en finance.

Comportement extrême

Le comportement extrême des lois appartenant aux différents DAM est significativement différent. Les lois appartenant au DAM de Weibull sont bornées à droite. Celles appartenant aux DAM de Gumbel et Fréchet ont un support infini à droite, mais les premières ont des queues fines alors que les deuxièmes ont des queues épaisses et exposent donc à des situations extrêmes beaucoup plus dangereuses.

Tableau 1 : Comportement extrême de lois usuelles en assurance

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

La représentation de Jenkinson-Von Mises donne une caractérisation synthétique des trois lois extrêmes au travers de la distribution généralisée des valeurs extrêmes (GEV), dont le signe du paramètre ξ identifie le domaine d'attraction maximum auquel la distribution appartient.

Excès au-delà d'un seuil élevé

Bien que les DAM soient définis par référence au comportement asymptotique de la loi du maximum, les résultats les plus utiles d'un point de vue opérationnel en assurance et en finance proviennent du comportement des excès au-delà d'un seuil élevé. En particulier, les résultats asymptotiques suivants sont essentiels.
Le nombre de dépassement N_{u_n} 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 n. Formellement, on a :

$$\lim_{n\to\infty} nF(u_n) = \tau \Rightarrow N_{u_n} \xrightarrow{d} \mathcal{P}(\tau).$$

On comprend aisément l'intérêt de ce résultat pour modéliser le nombre de survenances d'événements extrêmes (qu'il s'agisse de sinistres ou de fluctuations de variable financière), pour autant que le seuil à partir duquel cette propriété opère ait été déterminé.
Par ailleurs, considérant toujours un seuil suffisamment élevé, la distribution des coûts au-delà de ce seuil suit une loi de Pareto généralisée (GPD) dont le paramètre de forme est le paramètre ξ du DAM, auquel appartient la distribution considérée. La distribution de Pareto généralisée dispose des propriétés suivantes :

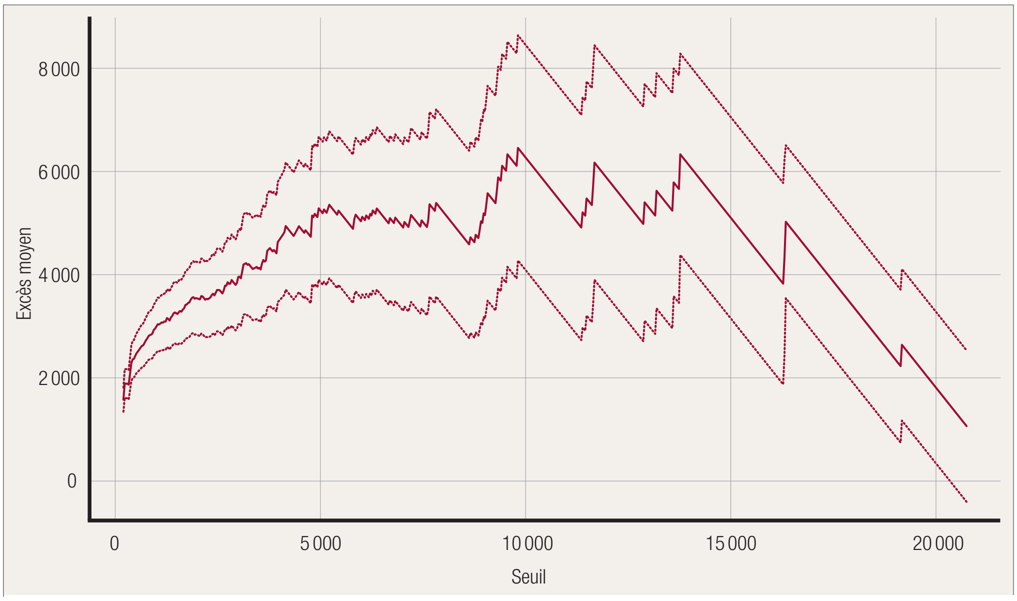
- les moments n'existent que jusqu'à l'ordre $[\xi^{-1}]$;
- le coût moyen (pour autant que celui-ci existe !), au-delà d'un seuil très élevé, est proportionnel au niveau du seuil en question.

Aussi, en plus d'être capable de déterminer la distribution du nombre d'événements dépassant un seuil élevé, la loi de probabilité du coût de ces événements est approchable au travers de l'approximation GPD.

Comment choisir le seuil ?

Sur le plan pratique, ces résultats fournissent les outils d'estimation du seuil à partir duquel les résultats de la théorie des valeurs extrêmes sont applicables.
La méthode la plus usitée est celle du *Picks-Over-Treshold* (POT). Cette dernière consiste à déterminer (au moyen de divers estimateurs et de méthodes graphiques) le seuil à partir duquel on commence à observer les phénomènes issus des résultats ci-dessus et en particulier la linéarité de l'excès moyen.

Figure 1 – Excès résiduel moyen au-delà du seuil en fonction du seuil



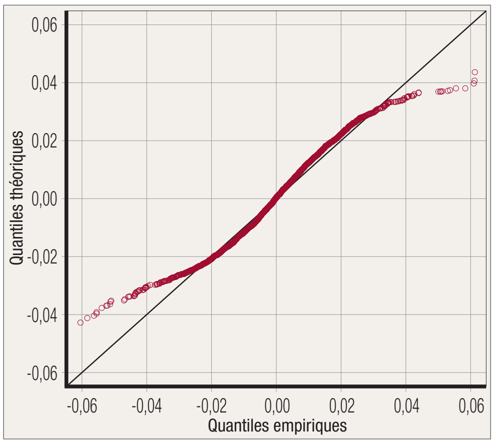
Lecture de la figure 1 : le seuil est choisi comme la plus petite valeur au-delà de laquelle l'excès résiduel moyen est linéaire. Cette sélection n'est pas toujours simple, comme on peut le constater sur cette illustration de sinistres automobiles.

…%… Ce point conduit néanmoins à une limite opérationnelle dans la mesure où, en pratique, on dispose d'un échantillon fini d'observations qui conduit à un arbitrage. En effet, plus le seuil est élevé, meilleure est l'approximation, mais moins on dispose d'observations au-delà du seuil pour calibrer la distribution Pareto généralisée.

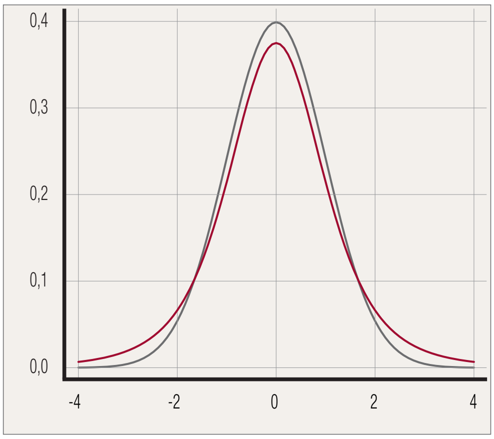
Est-ce vraiment utile ?

D'un point de vue opérationnel, les résultats énoncés *supra* ne font intervenir que les observations de la queue de distribution. Un écueil qui peut être rencontré en pratique est la bonne adéquation d'un modèle aux données sur l'ensemble de la distribution (avec des tests d'adéquation qui ne sont pas rejetés, par exemple) et l'utilisation de ce modèle, satisfaisant *a priori*, pour estimer des quantiles extrêmes. Cette approche peut conduire à des mésestimations sévères, comme le montrent les graphiques suivants.

QQ-plot des rendements journaliers du CAC 40 (modèle de Black & Scholes)



Densités gaussienne et Student



Le premier montre la proximité, dans le corps de la distribution, des densités d'une distribution gaussienne (à queue fine en gris) et d'une distribution de Student (à queue épaisse en rouge), qui « trompe » la plupart des tests d'adéquation effectués sur l'ensemble des observations alors que les comportements en queue de distribution sont singulièrement différents.

Le deuxième est un ajustement gaussien des rendements instantanés journaliers du CAC 40 (modèle de Black & Scholes). Si l'ajustement est satisfaisant dans le corps de la distribution, les rendements extrêmes (à la hausse, mais surtout à la baisse) et donc les variations de cours sont très significativement sous-estimés par un tel modèle.

Pierre Thérond (Galea & Associés, Isfa)
et Pierre Ribereau (Isfa)

1. *Via* une transformation affine au moyen de suites de réels c_n et d_n.

Quelques références bibliographiques

Modelling Extremal Events for Insurance and Finance, P. Embrechts, C. Klüppelberg, T. Mikosch, Springer, 1997.

The variation of certain speculative prices, Mandelbrot B., *The Journal of Business*, vol. 36, 1963.

Mesure et gestion des risques d'assurance, F. Planchet, P.E. Thérond, *Economica*, 2008.

Provisions techniques et capital de solvabilité d'une compagnie d'assurances : méthodologie d'utilisation de Value-at-Risk, P.E. Thérond, F. Planchet, *Assurances et gestion des risques* 74 (4), 533-63, 2007.

L'économie des extrêmes, D. Zajdenweber, Flammarion, 2000.