LES MODÈLES EN RÉASSURANCE

Arthur Charpentier

Université de Rennes I et École polytechnique

La modélisation des événements extrêmes est presque née avec la réassurance. En 1925, Karl Gustav Hagstrœm publiait dans la *Revue scandinave d'actuariat* un article sur « La loi de Pareto en réassurance ». Cette loi avait été proposée par Vilfredo Pareto sur les revenus fiscaux. Il avait noté que 20 % de la population possédait 80 % des richesses. Karl Gustav Hagstrœm a remarqué que cette répartition se retrouvait curieusement sur les grands risques, où 20 % des sinistres pouvaient représenter 80 % de la charge totale. C'est à partir de cette observation relativement anodine qu'ont été proposées les premières méthodes de tarification de traités en excédent, de manière empirique et heuristique... avant d'être légitimées presque cinquante ans plus tard par les statisticiens travaillant sur les valeurs extrêmes.

REINSURANCE MODELS

Modelling extreme events more or less came into being with reinsurance. In 1925, Karl Gustav Hangstrœm published an article in the Revue scandinave d'actuariat (Scandinavian Actuarial Review) entitled "Pareto's Law in Reinsurance". Vilfredo Pareto had proposed his law in respect of tax revenue. He had observed that 20 % of the population owned 80% of the wealth. Karl Gustav Hangstrœm noticed the curious fact that this distribution also occurred in respect of large risks, where 20% of losses could account for 80% of the total charge. It was from this relatively anodyne observation that the first methods of empirically and heuristically setting tariffs for excess of loss treaties were proposed ... before being confirmed almost fifty years later by statisticians working on extreme values.

Modèles statistiques sur les événements extrêmes

Historiquement, l'étude des valeurs extrêmes (dans une démarche de formalisation statistique) remonte aux travaux de Fisher et Tippett en 1920, qui tentèrent de dériver un théorème central limite pour le maximum d'un échantillon (convenablement normalisé, tout comme la moyenne dans le théorème central limite). Ils obtinrent – et cela fut formellement démontré en 1943 par Gnedenko – que les seules lois limites possibles pour le maximum sont les lois de Fréchet, de Gumbel et de Weibull. On peut réécrire la fonction de répartition du maximum sous la forme suivante :

$$\exp\left(1-\left[1-\frac{x-\mu}{\alpha\sigma}\right]^{\alpha}\right) \qquad (1)$$

où est appelé « indice de queue » (les deux autres paramètres étant des paramètres de normalisation, comme pour le théorème central limite).

Ces lois ont été très utiles en hydrologie par exemple, où Emil Gumbel les a appliquées sur les maximums annuels des niveaux des fleuves.

Mais ces lois n'ont pas vraiment eu d'application en assurance (sauf peut-être si l'on s'intéresse aux traités Ecomor introduits par Thépaut en 1950).

En fait, la loi de Pareto n'est réapparue que beaucoup plus tard. En effet, suite à l'inondation de 1953 qui causa la mort de plus de 1 800 personnes aux Pays-Bas, les autorités posèrent la question du niveau optimal auquel il conviendrait de construire des digues pour éviter les inondations côtières. La limite qui avait été retenue était la crue dix-millénaire.

Le problème était toujours lié aux valeurs extrêmes, mais cette fois la problématique n'était pas exprimée en termes de loi d'un maximum, mais consistait dans l'étude du franchissement d'un seuil.

Dans les années 1970, Balkema, Laurens de Haan et Pickands ont montré que la loi dite des excès (c'est-à-dire conditionnelle au franchissement d'un seuil) devait suivre une loi de type Pareto :

$$P\left(X > x + u \middle| X > u\right) = \left(1 + \frac{x}{\sigma}\right)^{-\alpha} \qquad (2)$$

Pour u suffisamment grand.

En outre, ils ont montré que les deux approches étaient très liées, le coefficient α apparaissant dans les équations (1) et (2) étant généralement identique [1]. Aussi, près de cinquante ans plus tard, les statisticiens néerlandais ont-ils pu légitimer l'intuition de Karl Hagstrœm.

Mais une question (posée dès 1930) reste en suspens : comment choisir le seuil à partir duquel les sinistres sont considérés comme importants ? Ou, plus formellement, à partir de l'équation (2), comment choisir le seuil u à partir duquel l'ajustement par une loi de Pareto est valide ?

De la difficulté de faire un ajustement statistique

Si l'on reprend les travaux de Karl Hagstrœm, on peut voir qu'une technique simple pour estimer l'indice de queue est de représenter le logarithme de laprobabilité de dépasser un montant en fonction du logarithme des coûts :

$$\log P\left(X > x + u \middle| X > u\right) \approx -\alpha \log(x) + \beta$$

L'indice de queue est alors tout simplement obtenu comme la pente de la droite passant par le nuage de points. D'un point de vue statistique, on ordonne les charges des sinistres

$$\min\{X_1, \ldots, X_n\} = X_{1:n} \le X_{2:n} \le \ldots$$

$$\le X_{n-1:n} \le X_{n:n} = \max\{X_1, \ldots, X_n\}$$

et on retient les κ plus grands sinistres (qui ont effectivement dépassé le seuil μ retenu).

L'estimateur naturel de la probabilité de dépassement est obtenu à l'aide de la fonction de répartition empirique

$$\hat{F}(x) = \frac{1}{k} \sum_{i=1}^{k} 1(X_{n-i+1:n} \le x)$$

qui vérifie

$$\hat{F}(X_{n-j+1:n}) = \frac{k - j + 1}{k}$$

Aussi, un estimateur « naturel » de l'indice de queue est obtenu en considérant la pente de la droite de régression du nuage de points

$$\left(\log(X_{n-j+1:n}),\frac{k-j+1}{k}\right)$$

pour $j = 1, ..., \kappa$

(on parlera de Pareto plot). C'est à peu de choses près l'estimateur proposé par Hill en 1975, à savoir

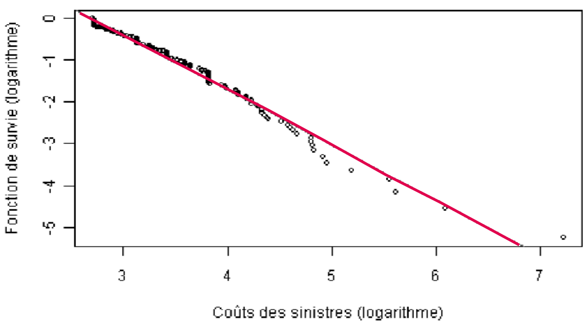
$$\hat{\alpha}_{(k)} = \frac{1}{k}\sum_{i=1}^{k}[\log(X_{n-i+1:n}) - \log(X_{n-k:n})]$$

À partir d'un échantillon $\{X_1, ..., X_n\}$ on obtient ainsi un ensemble d'estimateurs lié au choix du nombre de grands sinistres considérés (i.e. choix de μ ou de κ). Avec un arbitrage délicat d'un point de vue statistique,

• plus κ est petit (ou μ est grand), plus on s'intéresse aux événements extrêmes : l'estimateur aura un biais faible, mais comme il sera estimé avec peu de données, sa variance sera grande ;
• plus κ est grand (ou μ est petit), plus on s'éloigne des événements extrêmes : l'estimateur aura un biais fort (l'approximation par la loi de Pareto est valide pour μ très grand), sa variance sera faible car estimée sur un grand nombre de données.

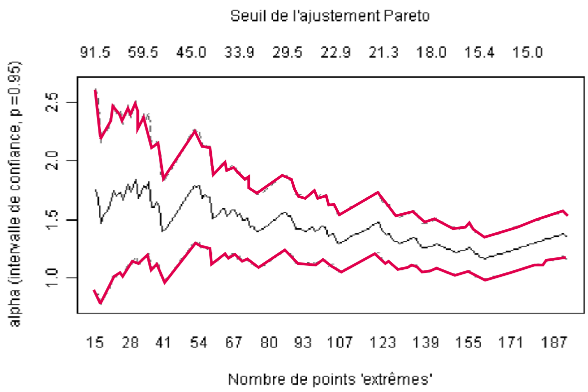
Considérons l'exemple des sinistres en perte d'exploitation en France [2]. La figure 1 est le Pareto plot sur les plus gros sinistres, montrant effectivement le très bon ajustement linéaire.

Figure 1. Pareto plot sur les sinistres en perte d'exploitation



On peut alors regarder (figure 2) l'impact de l'évolution du nombre de grands sinistres considérés (κ) ou du seuil à partir duquel on considère le sinistre comme grand (μ) sur l'estimation de l'indice de queue.

Figure 2. Estimateur de Hill de l'indice de queue $\alpha_{(\kappa)}$



Pour rappel, la loi de Pareto est une loi telle que si $\alpha \leq 1$, l'espérance mathématique n'est pas finie (la prime pure est infinie, ce qui signifierait que le risque n'est pas (ré)assurable), et si $\alpha \leq 2$, la variance est infinie.

Pourquoi chercher à définir une loi ?

Une méthode assez classique en réassurance consistait à utiliser le *burning cost*, c'est-à-dire tout simplement le coût moyen historique observé sur les données passées.

Mais cette approche trouve vite ses limites si l'on considère un programme de réassurance avec une multitude de tranches, dont quelques-unes très élevées où l'historique est très rare.

Une autre limite est que l'on peut parfois être intéressé par d'autres grandeurs que la prime pure, une prime de Wang par exemple. Et cette dernière suppose que l'on dispose de la distribution des coûts des sinistres (et pas seulement leur coût moyen).

Rappelons que la prime pure (par sinistre) pour un traité en excédent de priorité (et de portée infinie[3]) est de la forme

$$\pi = E(X-u|X>u) = \int_0^\infty S(x)dx$$

où

$$S(x) = P(X-u>x|X>u)$$

et la prime de Wang est donnée par

$$\pi(\lambda) = \int_0^\infty \Phi\left[\lambda + \Phi^{-1}\left(S(x)\right)\right]dx$$

où Φ désigne la fonction de répartition de la loi normale (centrée réduite, cf. Wang (2000)).

Plus généralement, toutes les mesures de risques par distorsion nécessitent de connaître la distribution des coûts de sinistre. Les figures 3 et 4 montrent ainsi l'évolution de la prime (par sinistre). Pour obtenir la prime effectivement payée par la cédante au réassureur, il convient de pouvoir estimer la probabilité qu'un sinistre atteigne la priorité. Mais, comme toujours, les modèles sur les nombres sont moins complexes que ceux sur les coûts (cf. Denuit et Charpentier, 2005).

Figure 3. Estimation de la prime pure (par sinistre) non paramétrique vs Pareto

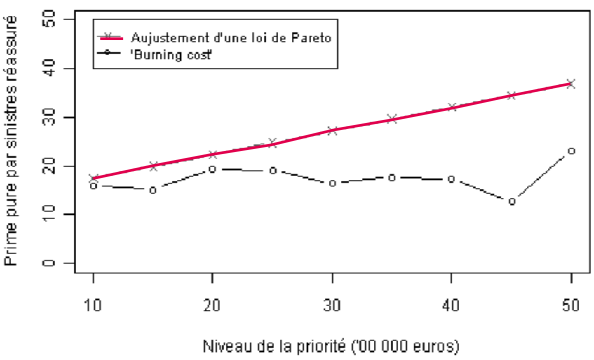
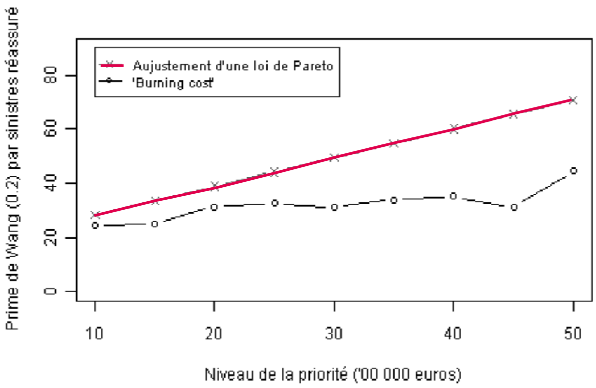


Figure 4. Estimation de la prime de Wang (par sinistre) non paramétrique vs Pareto



La robustesse et les limites de ces modèles

La crise financière récente a montré la limite des modèles utilisés en finance. Et, naturellement, les assureurs s'interrogent sur les modèles utilisés par les réassureurs. Si l'on y regarde de plus près, les modèles financiers sont arrivés du monde universitaire dans les établissements financiers. Dans les années 1950, Henry Markowitz avait proposé une méthode d'allocation d'actif en supposant les rendements gaussiens, de même, la formule de valorisation des options avait été proposée dans les années 1970 par Robert Merton, Myron Scholes et Fisher Black[4]. Les modèles de risques de crédit proposaient, dans les années 1990, d'utiliser des modèles toujours fondamentalement gaussiens pour modéliser la contagion dans les portefeuilles. Tous ces modèles se sont avérés faux après des tests empiriques.

Comme souvent, les modèles en actuariat ont découlé d'une toute autre logique. Ce sont les assureurs qui, dans les années 1920, ont proposé d'utiliser la loi de Pareto sur des constatations empiriques, et ce n'est qu'ensuite, cinquante ans plus tard, que les travaux académiques ont validé la pertinence de cette approche. De la même manière, les actuaires d'Amérique du Nord ont posé les bases de la crédibilité dans les années 1920, avant d'être confortés cinquante ans plus tard par les travaux d'Hans Bühlmann ou Bruno de Finetti. Tant que les actuaires se sont contentés de modéliser les risques, ils ont été brillants. C'est d'ailleurs ce que disait Emile Borel lorsqu'il écrivait, au sujet des sociétés d'assurances : « *nous partons ainsi d'un base pratique assez solide pour que nous ayons dans la théorie la confiance qui peut être nécessaire pour ne pas avoir à tenir compte du scepticisme qui peut toujours être opposé à toute tentative d'explication rationnelle* ».

Mais si ces modèles sont robustes, ils ont un certain nombre de limites. Tout d'abord des limites pratiques qu'ont rencontrées tous les praticiens : ces modèles sont valides à condition d'avoir des coûts de sinistre indépendants et identiquement distribués. En particulier, il ne doit pas y avoir d'inflation. Or déflater les montants des sinistres est délicat, comme l'avaient noté Pielke et Landsea (1998) dans leur étude des ouragans aux États-Unis. En effet, les événements extrêmes étant rares, on est souvent tenté d'utiliser un historique très long pour avoir une chance de trouver un événement centenaire. Mais les expositions ont radicalement changé sur une période aussi longue (par exemple, la population en Floride a triplé, ainsi que le coût potentiel d'un ouragan, même en dollars constants). L'autre problème est que des sinistres de toutes sortes sont souvent amalgamés afin de constituer des bases de données assez larges pour s'appliquer à des événements extrêmes, mais ces bases sont alors souvent très hétérogènes Formellement, il serait plus intéressant de réserver ces méthodes aux valeurs extrêmes, et d'utiliser ici plutôt les méthodes économétriques ou de crédibilité.

D'un point de vue théorique, de nombreuses questions restent ouvertes. Par exemple, comment traiter les risques dont l'espérance est théoriquement infinie ($\alpha \leq 1$) ? Compte tenu du faible nombre de réassureurs, de leurs capacités limitées et des participations croisées dans différentes tranches dans les montages de réassurance, un effet domino de faillites semblerait possible, en particulier si des réassureurs s'engagent sur des tranches très élevées (et potentiellement extrêmement rares). Comme le notait D. Zajdenweber (2000), cela pourrait se produire sur les catastrophes naturelles.

Un autre problème est que les couvertures de réassurance n'ont d'intérêt que si elles sont bien construites et effectivement adaptées à la sinistralité. L'exemple le plus classique est l'opposition souvent faite entre les couvertures proportionnelles et les couvertures non proportionnelles, où l'on nous explique que pour assurer la solvabilité d'une compagnie d'assurances (en particulier si elle traite des grands risques), c'est généralement une couverture en excédent qu'il convient de choisir (afin de faire face aux sinistres majeurs qui sont généralement la cause des faillites). Mais cela n'est aucunement justifié en théorie. Curieusement, une couverture proportionnelle permet, dans tous les cas, de réduire la probabilité de ruine ; alors qu'une mauvaise couverture non proportionnelle peut augmenter la probabilité de ruine si on se réassure mal. Il suffit qu'il existe une forte dépendance entre les survenances et les coûts (par exemple des périodes de l'année où surviennent une multitude de petits sinistres qui n'atteignent pas la priorité du traité). Mais il s'agit là d'un problème classique d'assurance, qui est l'adéquation entre la couverture et le risque : une mauvaise couverture fait payer aux sociétés d'assurances une cotisation sans qu'elles reçoivent (ou rarement) la contrepartie d'une indemnité, ce qui met en péril leur solvabilité.

Conclusion

Comme souvent, le problème ne vient pas des modèles utilisés, mais du fait que certains omettent d'y intégrer les hypothèses nécessaires à leur application. La formule standard de calcul de SCR proposée par Solvabilité II n'est valide que sous certaines hypothèses (essentiellement de lois gaussiennes, ou de type elliptique) ; mais on peut douter de sa validité sur les grands risques auxquels les réassureurs doivent faire face. Plus généralement, dès que l'on commence à voir apparaître en réassurance des modèles issus du monde financier et de son inoubliable loi normale, on peut alors s'inquiéter.

Notes

1. Pour davantage d'information sur les résultats théoriques associés à ces deux approches (approche par le maximum ou approche par les excès), il sera possible de se reporter à Embrechts et al. *(1997), ou à Beirlant et* al. *(1996).*

2. Ces données avaient été considérées dans Zajdenweber (2000), et sont téléchargeables en ligne (ainsi que tous les codes utilisés ici) à l'adresse
http://blogperso.univ-rennes1.fr/arthur.charpentier/

3. Mathématiquement, tous les contrats en excédent se déduisent de cette forme simple (priorité u et portée infinie). En effet, pour la prime pure par exemple, si on considère une portée finie

$$\pi(a,b) = \int_a^b (x-a)dF(x) = \pi(a,\infty) - \pi(b,\infty)$$

4. On notera que tous ces universitaires ont obtenu pour leurs travaux le prix de la Banque de Suède en sciences économiques en mémoire d'Alfred Nobel.

Bibliographie

BEIRLANT J. ; TEUGELS J. ; VYNCKIER P., *Practical Analysis of Extreme Values*, Leuven University Press, 1996.

CHARPENTIER A., « Value at risk et probabilité de ruine, entre vaccination et banque d'affaires », *Risques*, n° 76, 2009, pp. 103-106.

DENUIT M. ; CHARPENTIER A., *Mathématiques de l'assurance non vie*, tome 2, Economica, 2005.

EMBRECHTS P. ; KLÜPPELBERG C. ; MIKOSCH T., *Modelling Extremal Events for Insurance and Finance*, Springer-Verlag, 1997.

HAGSTRŒM K.G., « La loi de Pareto en réassurance », *Skandinavisk Aktuarietidskrift*, 1925.

PIELKE R.A. ; LANDSEA C.W., "Normalized US Hurricane Damage, 1925-1995", *Weather and Forecasting*, 1998.

THÉPAUT A., « Une nouvelle forme de réassurance. Le traité d'excédent du coût moyen relatif (Ecomor) », *Bulletin trimestriel de l'Institut des actuaires français*, 1950.

WANG S.W., "A Class of Distortion Operators for Pricing Financial and Insurance Risks", *Journal of Risk and Insurance*, 2000.

ZAJDENWEBER D., *Économie des extrêmes*, Flammarion, 2000.