Peut-on diversifier des risques extrêmes ?

Dans un contexte financier, diversifier les risques signifie investir dans une variété d'actifs, secteurs ou régions géographiques pour éviter que la mauvaise performance d'un investissement n'affecte trop l'ensemble du portefeuille. La diversification permet de réduire le risque, ou, dans sa formulation mathématique, de réduire la variance. Mais que se passe-t-il quand on est en présence de grands risques, de variance infinie ? Ou pire encore, d'espérance infinie ?

ARTHUR CHARPENTIER
Professeur,
Université du Québec
à Montréal (UQAM)

Risques extrêmes, et espérance infinie ?

Formaliser des grandeurs en lien avec des quantités aléatoires et incertaines est un exercice compliqué. Les probabilités, au sens où le mot est souvent entendu, sont définies comme des limites de fréquences observées par répétitions d'événements. La probabilité d'avoir 3 en lançant un dé est ⅙ car en lançant un dé un million de fois[1], un milliard de fois, la probabilité sera aussi proche qu'on veut de ⅙. C'est ce que dit la loi des grands nombres, dans sa version la plus faible. Dire que la probabilité qu'il pleuvra aujourd'hui est de ⅙ est totalement différent, car c'est un événement unique. Si je me fais tremper par une averse aujourd'hui, cela ne permettra aucunement de dire que la probabilité n'était pas, a priori, de ⅙, et que le modèle météorologique s'est trompé. Tout ça pour rappeler que lorsqu'on fait de la modélisation, on va essayer d'imaginer les valeurs petites d'événements rares, et qu'il est malheureusement très difficile de les valider.

Et quand on modélise les grands risques, les très grands risques, il n'est pas rare d'avancer l'idée que les risques sont de variance ou d'espérance infinie. Or la notion d'espérance infinie est à la fois étrange, et probablement contre-intuitive[2]. Si on considère une variable aléatoire X positive (pour simplifier), et si on note $S(x)=P(X>x)$ la fonction de survie, et $f(x)$ la densité (correspondant à l'opposée de la dérivée de S), on peut montrer que la moyenne empirique d'un million ou d'un milliard de tirages de cette variable va s'approcher autant qu'on veut d'une grandeur, appelée l'espérance mathématique.

$$E(X) = \int S(x)dx = \int x f(x)dx$$

Rien de bien étonnant ici : c'est encore la loi des grands nombres, énoncée dès 1713 par Jacob Bernoulli (le « golden theorem » de Raper (2018)) et surtout Pierre-Simon Laplace, en 1814. À condition toutefois que cette intégrale soit finie. Ce qui n'est pas garanti. Par exemple, la loi de Pareto d'indice α vérifie $S(x)=P(X>x)=x^{-\alpha}$. Dès 1925, Karl-Gustaf Hagstroem avait noté que cette loi semblait particulièrement adaptée à la modélisation des grands risques, et donc en réassurance[3]. Et pour une variable qui suit une loi de Pareto d'indice 1, son espérance est, mathématiquement, infinie.

Que signifie cette espérance infinie ? Il n'y aura aucun « sinistre de coût infini », et il sera toujours possible de calculer une moyenne empirique sur n observations. Mais cette moyenne va tendre vers l'infini quand n croît. Louis Bachelier, en évoquant le paradoxe de Saint-Pétersbourg (qui est un jeu dont l'espérance de gain est infinie), rappelle qu'« un résultat paradoxal, dans les sciences mathématiques, provient nécessairement d'un défaut de notre intelligence, inhabile à déchiffrer un ensemble trop complexe, inapte à se représenter l'infiniment grand [...] le bon sens ne peut être invoqué lorsqu'il s'agit de questions délicates ; il ne permet pas de reconnaître si l'aire comprise entre une courbe et son asymptote est finie ou non, si une série est convergente ou divergente. »

Cette moyenne va tendre vers l'infini quand n croît signifie qu'on peut être certain que la moyenne va toujours dépasser n'importe quelle valeur aussi grande qu'on puisse l'imaginer. On peut le visualiser en haut de la Figure 1 avec 10 simulations de 100,000 valeurs. À gauche, le cas où l'espérance finie (et la variance infinie) ; à droite, les deux sont infinies.

Une autre grandeur intéressante est le ratio du maximum sur n observations sur la somme (la part du plus gros sinistre dans la charge totale cumulée). Pour des variables d'espérance infinie, ce ratio ne tend pas vers 0. Il est alors possible, si les variables x désignent les coûts de sinistres, avec 100 000 sinistres d'espérance infinie, que le plus gros des sinistres représente plus de 90 % de la charge totale.

On le voit, cette propriété est importante, mais elle est difficile à identifier car il s'agit d'une propriété fondamentale du modèle[4] sous-jacent, en lien avec la distribution des observations, puisqu'il est toujours possible de calculer la moyenne. Par exemple, la suite suivante correspond à huit valeurs obtenues en tirant au hasard une loi de Pareto d'indice 1 (et donc théoriquement d'espérance infinie)

3.4773 1.2685 2.4451
1.1325 1.0633 21.9508
1.8936 1.1205 1.8135
2.1900 1.0451 2.2059

Comment savoir si un ensemble de coûts de sinistres suit une loi d'espérance finie, ou pas ? L'approche classique, présentée par exemple par Daniel Zajdenweber (1996, 2000), consiste à utiliser le graphique dit de Pareto, avec le logarithme des coûts sur l'axe des abscisses, et le logarithme de la probabilité de survie en ordonnées. Si les points sont alignés suivant une droite de pente $-a$, alors la loi de Pareto de paramètre a est parfaitement adaptée. En effet, si $P(X>x)=x^{-a}$ alors, en prenant le logarithme des deux grandeurs, et si on ordonne l'échantillon $(x_1 \leq x_2 \leq \ldots \leq x_n)$, alors :

$$log\left(\frac{n-i}{n}\right)=-a\times log(x_i)$$

Et si la pente est trop modérée, plus grande que –1, alors les coûts sont d'espérance infinie.

Cette hypothèse d'indice de Pareto proche de 1 n'est pas irréaliste quand on parle de catastrophes naturelles, ou industrielles :

- ouragans, Hsieh (1999), $a \approx 1.5$
- incendies entreprises, Biffis et al. (2014), $a \approx 1.25$
- perte d'exploitation, Zajdenweber (1996), $a \approx 1$

Fig. **1** Évolution de la moyenne $(x_1+\ldots+x_n)/n$ en fonction de n, avec une loi d'espérance finie et de variance infinie à gauche, et une loi d'espérance infinie à droite.

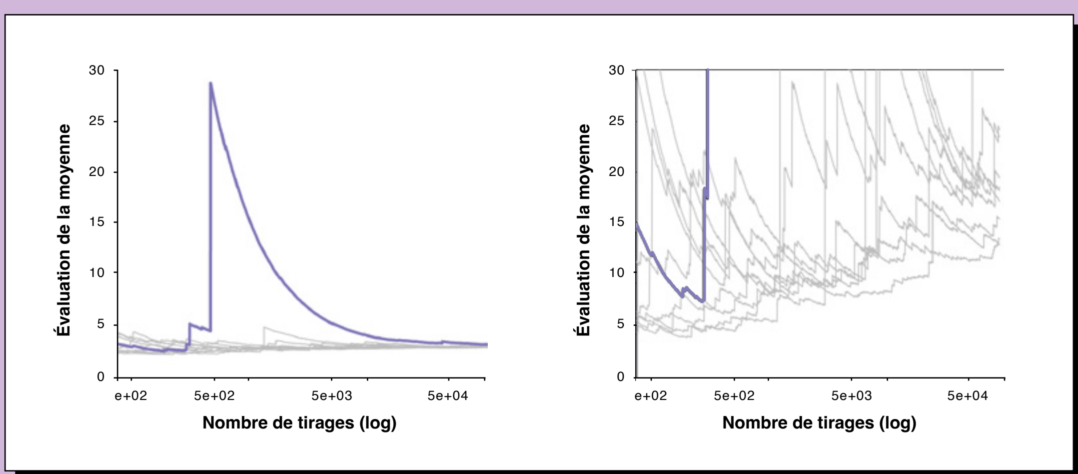
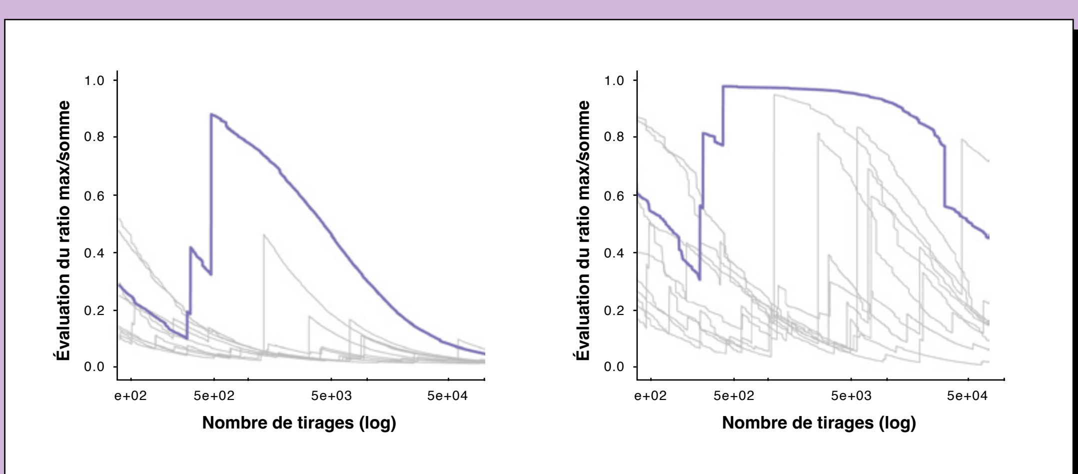


Fig. **2** Évolution du ratio $max\{x_1,\ldots,x_n\}/(x_1+\ldots+x_n)$ avec une loi d'espérance finie et de variance infinie à gauche, et une loi d'espérance infinie à droite.



Source : Auteur

- tremblement de terre, Sornette et al. (1996), α≈1
- tsunamis, Embrechts et al. (2024), α≈1
- risque opérationnel, Moscadelli (2004) et Chavez-Demoulin et al. (2006) α≈1
- risque cyber, Eling et al. (2019) α≈1
- risques nucléaires, Hofert et al. (2012), $\alpha \in (0.6\,;0.7)$

De la diversification des grands risques

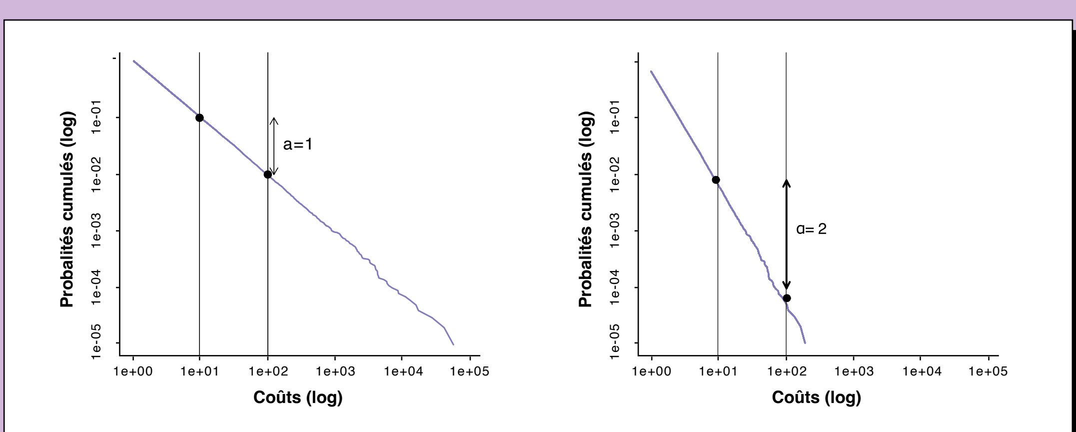
Au lieu de travailler par type de risque, on peut envisager l'agrégation de ces risques, tous ensemble. Heuristiquement, avoir des portefeuilles avec des risques d'inondation, de tremblement de terre, ou de sécheresse pourrait offrir un peu de « diversification ». Le concept de « diversification » peut être introduit avec la loi des grands nombres, dont on parlait auparavant, et il sera très proche de l'idée même d'assurance, de mutualisation des risques. Smith & Kane (1994), par exemple, rappellent que la contribution d'un n+1-ième risque, indépendant, dans un groupe de n risques, tarifés de manière actuariellement juste, permet généralement de faire marginalement diminuer le risque, ce qui renforce la mutualisation des risques par l'assureur. Cet effet de diversification fonctionne encore, même si les risques sont corrélés (mais pas parfaitement corrélés, et les gains de diversification diminuent avec la corrélation, comme le rappelait Charpentier (2011)).
Mais bien souvent, quand on parle de « diversification » on pense aux travaux de Harry Markowitz ou d'Arthur Roy en finance dans les années 1950, comme base de la théorie du portefeuille. Cette théorie montre comment des investisseurs rationnels peuvent utiliser la diversification, correspondant à la corrélation entre actifs, afin d'optimiser leur portefeuille financier. Dans cette approche, on suppose généralement que la préférence des investisseurs pour un couple risque / rendement peut être décrite par une fonction d'utilité quadratique. Autrement dit, seuls le rendement attendu (l'espérance de gain) et la volatilité (l'écart type) ou la variance, sont les paramètres considérés par l'investisseur. Ce que montre cette littérature, c'est qu'un investisseur peut réduire le risque de son portefeuille simplement en détenant des actifs qui ne sont pas (ou peu) corrélés, donc en diversifiant ses placements. Il peut alors obtenir la même espérance de rendement en diminuant la variabilité de son portefeuille.

Que se passe-t-il si la variance n'existe plus ?

Cette question revient à questionner l'utilisation de la loi normale pour modéliser les rendements financiers. La loi normale était intéressante en partie parce qu'elle vérifie une propriété de stabilité par sommation[5]. Garder cette propriété tout en considérant une loi ayant davantage d'extrêmes que la loi normale, revient à utiliser les lois « stables » étudiées par Paul Lévy, comme le proposait[6] Benoît Mandelbrot dans les années 1960.

Dans le cas où la variance est infinie, il convient d'utiliser une mesure de risque plus générale que l'écart-type. Heuristiquement, la « diversification » est liée à la sous-additivité de la mesure de risque : un portefeuille contenant la moyenne des avoirs de deux autres portefeuilles a un risque plus faible de la moyenne des risques des deux autres portefeuilles. Daníelsson *et al.* (2013) rappellent qu'en présence de grands risques (d'espérance infinie), la diversification ne fonctionne plus. Cette propriété avait été décrite et discutée par Paul Samuelson dès 1967, Stephen Ross en 1976, ou plus récemment Rustam Ibragimov, Dwight M. Jaffee, Johan Walden, Paul Embrechts ou Ruodu Wang, entre autres. L'introduction de Ibragimov *et al.* (2015) l'explique bien : « Il existe des limites à la diversification avec de telles distributions de risques [heavy-tailed distribution - distribution à queue lourde]. Plus précisément, alors que la diversification est préférée par les agents averses au risque lorsque les risques sont de faible envergure (le cas traditionnel a été largement étudié), il peut en réalité être préjudiciable pour les agents de se diversifier lorsque les risques sont de grande envergure. Les pièges de la non-diversification peuvent survenir lorsque la répartition des risques se présente à queue lourde et que les assureurs ont une responsabilité limitée. »

Fig. 3 Graphique de Pareto, avec $log\left(\frac{n-in}{n}\right)$ sur l'axe des ordonnées et $log(xi)$
Les points sont alignés suivant une droite, de pente $-\alpha$, correspondant à une loi de Pareto d'indice α. $\alpha \leq 1$ signifie que les risques sont d'espérance infinie.



Source : Auteur

Ces propriétés, largement discutées d'un point de vue mathématiques, sont compliquées à faire admettre car elles sont théoriques et contre-intuitives. De plus, il est souvent difficile de savoir pour qui la diversification devient dangereuse, puisqu'il y a plusieurs acteurs, les assurés, les assureurs, les réassureurs, l'État. Ibragimov *et al.* (2011) donnent des éléments de réponse, « Lorsque ces risques sont limités, le partage des risques est toujours optimal, tant pour les intermédiaires individuels que pour la société. Mais, en cas de risques modérément lourds, le partage des risques peut s'avérer sous-optimal pour la société, même si les intermédiaires individuels en bénéficient toujours [...]. Il est bien connu que la diversification peut s'avérer sous-optimale dans le cas d'une queue extrêmement lourde ».

Depuis une vingtaine d'années, les exemples où la diversification ne fonctionne pas sont nombreux et connus des praticiens. Fabozzi *et al.* (2014), évoquant la crise financière, le rappellent « La crise financière a clairement montré que lorsqu'on a besoin de diversification maintenant, cela peut ne pas fonctionner. » Quand on s'intéresse aux risques liés aux catastrophes climatiques, on voit que ces risques sont extrêmes, potentiellement inassurables car d'espérance potentiellement infinie. L'inassurabilité signifie surtout qu'un mécanisme de marché n'a pas de sens sans intervention de l'État. On pourrait aussi penser qu'il pourrait être intéressant de diversifier les risques, en offrant une couverture multi-périls (comme le propose le mécanisme CatNat actuel), ou bien en envisageant une diversification géographique, par exemple au niveau européen, comme le suggéraient récemment Carlo Cimbri, Thierry Derez et Philippe Lallemand. Mais la littérature scientifique nous rappelle que cette diversification est dangereuse, en tous cas inenvisageable sans une intervention forte et claire des États.

Références

- Bachelier, L. (1925). « Quelques curiosités paradoxales du calcul des probabilités ». *Revue de métaphysique et de morale, 32*(3), 311-320.
- Biffis, E., & Chavez, E. (2014). Tail risk in commercial property insurance. Risks, 2(4), 393–410.
- Charpentier, A. (2011). « La loi des grands nombres et le théorème central limite comme base de l'assurabilité ? » *Risques*, 86.
- Chavez-Demoulin, V., Embrechts, P., & Nešlehová, J. (2006). Quantitative models for operational risk: extremes, dependence and aggregation. *Journal of Banking & Finance*, *30*(10), 2635-2658.
- Chen, Y., Embrechts, P., & Wang, R. (2024). An unexpected stochastic dominance: Pareto distributions, dependence, and diversification. *Operations Research*.
- Cimbri, C., Derez, T. & Lallemand, P. (2024). « Mutualisons l'assurance pour offrir aux Européens une protection à la hauteur des risques actuels ! », *La Tribune*, 23 mai.
- Daníelsson, J., Jorgensen, B. N., Samorodnitsky, G., Sarma, M., & de Vries, C. G. (2013). Fat tails, VaR and subadditivity. *Journal of econometrics*, 172(2), 283-291
- Eling, M., & Wirfs, J. (2019). What are the actual costs of cyber risk events? *European Journal of Operational Research*, 272(3), 1109–1119.
- Embrechts, P., Hofert, M., & Chavez-Demoulin, V. (2024). *Risk Revealed: Cautionary Tales, Understanding and Communication*. Cambridge University Press.
- Fabozzi, F. J., Focardi, S. M., Jonas, C.: Investment Management: A Science to Teach or an Art to Learn?. *CFA Institute Research Foundation* (2014)
- Fama, E. F. (1965). Portfolio analysis in a stable Paretian market. *Management science*, 11(3), 404-419.
- Hagstroem, K.-G. (1925). Pareto and reinsurance. *Scandinavian Actuarial Journal*, 216–248
- Hofert, M., & Wüthrich, M. V. (2012). Statistical review of nuclear power accidents. *Asia-Pacific Journal of Risk and Insurance*, 7(1).
- Hsieh, P.-H. (1999). Robustness of tail index estimation. *Journal of Computational and Graphical Statistics*, 8(2), 318–332.
- Ibragimov, R., & Walden, J. (2007). The limits of diversification when losses may be large. *Journal of banking & finance, 31*(8), 2551-2569.
- Ibragimov, R., Jaffee, D., & Walden, J. (2011). Diversification disasters. *Journal of financial economics, 99*(2), 333-348.
- Ibragimov, M., Ibragimov, R., & Walden, J. (2015). *Heavy-tailed distributions and robustness in economics and finance* (Vol. 214). Springer.
- Lévy, Paul (1925). *Calcul des probabilités*. Paris: Gauthier-Villars.
- Mandelbrot, B. (1960). The Pareto–Lévy Law and the Distribution of Income. *International Economic Review.* 1 (2): 79–106.
- Markowitz, H. (1952). Portfolio Selection, *Journal of Finance*, 7 (1), 77-91.
- Markowitz, H. (1971). *Portfolio selection : efficient diversification of investments.* Yale University Press.
- Moscadelli, M. (2004). The modelling of operational risk: experience with the analysis of the data collected by the Basel committee. *Technical Report* 517, Banca d'Italia
- Raper, S. (2018). Turning points: Bernoulli's golden theorem. *Significance*, 15(4), 26-29.
- Ross, S. A. (1976). A note on a paradox in portfolio theory. *Unpublished Mimeo, University of Pennsylvania*.
- Roy, A. D. (1952). Safety first and the holding of assets. *Econometrica*, 431-449.
- Samuelson, P. A. (1967). Efficient portfolio selection for Pareto-Lévy investments. *Journal of financial and quantitative analysis*, 2(2), 107-122.
- Sornette, D., Knopoff, L., Kagan, Y. Y., & Vanneste, C. (1996). Rank-ordering statistics of extreme events: Application to the distribution of large earthquakes. *Journal of Geophysical Research: Solid Earth, 101*(B6), 13883-13893.
- Smith, M. L., & Kane, S. A. (1994). The law of large numbers and the strength of insurance. In *Insurance, Risk Management, and Public Policy: Essays in Memory of Robert I. Mehr* (pp. 1-27). Dordrecht: Springer Netherlands.
- Zajdenweber, D. (1996). Extreme values in business interruption insurance. *Journal of Risk and Insurance*, 95-110.
- Zajdenweber, D. (2000). *Économie des extrêmes.* Flammarion.

Notes

[1] Le cas des dés est un peu particulier car la géométrie du cube, en particularité sa régularité (on parle d'hexaèdre régulier, à 6 faces), permet d'inférer la probabilité sans faire la moindre expérience.

[2] La littérature théorique des probabilités s'est largement construite sur l'idée de variables d'espérances finies, et il est très difficile de s'en passer (tout raisonnement « en moyenne » devenant impossible).

[3] Il faudra attendre les années 1970 et les travaux de Guus Balkema ou Laurens de Haan, pour avoir une preuve mathématique de ce résultat. L'école néerlandaise de statistique a fait des avancées majeures sur l'analyse des évènements extrêmes suite au raz-de-marée de 1953 en mer du Nord, qui a eu des conséquences majeures et désastreuses aux Pays-Bas, comme le rappellent Embrechts *et al.* (2024)

[4] La loi des observations étant, par nature, inconnue, le choix du modèle est important et fait porter une lourde responsabilité sur les épaules du modélisateur ou de la modélisatrice.

[5] La somme (ou la moyenne) de variables normales indépendantes suit aussi une loi normale.

[6] Il appelle ces lois « Pareto-Lévy » pour souligner la forme des queues de distributions, correspondant à des lois de type Pareto, sur les pertes extrêmes (à gauche) et les gains extrêmes (à droite).