

La crédibilité : un pasteur et un philosophe pour soutenir les actuaires

Arthur Charpentier

Professeur en finance et science des actuaires, Ensaé

Pour les actuaires, la crédibilité n'est souvent qu'un vague concept enseigné lors de leurs études, qu'ils ont vite fait d'oublier une fois l'examen réussit¹. Quant aux non-actuaires, à part se demander si les réponses apportées par les actuaires sont crédibles...

Crédibilité : Lat. *credibilis*, de *credere* (voy. croire). Ce qui rend une chose digne de croyance. *Celui qui doute parce qu'il ne connaît pas les raisons de crédibilité n'est qu'un ignorant*, Diderot, Pens. phil. 24. (Le Littré).

Un point fondamental en tarification est la segmentation : pour des raisons d'aléa moral, il est important de pouvoir calculer une cotisation pour un contrat non pas sur l'ensemble du portefeuille (qui sera forcément hétérogène) mais en utilisant des informations plus fines, c'est-à-dire de valoriser sur un sous-portefeuille possédant des caractéristiques similaires. La loi des grands nombres permet d'utiliser la moyenne sur le portefeuille global comme étant une bonne approximation de la prime pure globale ; mais sur un sous-portefeuille beaucoup plus petit (dix ou vingt observations), ces théorèmes asymptotiques ne marchent plus. On ne peut généralement pas tarifier un contrat auto individuel en ne prenant en compte

que le passé sinistre d'un client et, en assurance incendie entreprise, on n'a souvent pas assez d'entreprises présentant des caractéristiques similaires pour les tarifier individuellement. L'idée des méthodes de crédibilité est de pondérer (en accordant un certain crédit) l'expérience de ce sous-portefeuille (dix années d'expérience dans le premier cas, ou vingt entreprises de caractéristiques similaires) et l'expérience sur l'ensemble du portefeuille.

■ Hétérogénéité d'un portefeuille et expérience *rating*

L'idée d'expérience *rating* suggère

d'utiliser l'expérience individuelle d'un assuré lors de la tarification. Hans Bühlman disait ainsi que « *l'expérience rating vise à assigner à chaque risque individuel une prime adéquate* ». On peut ainsi considérer l'exemple 1 (tiré de Norberg, 1979).

Exemple 1 - Considérons l'historique du tableau ci-après. Globalement, la fréquence d'accident est de 23 %, avec une forte hétérogénéité entre l'assuré I (70 %) ou A (60 %) et les assurés G, H ou J (0 %). Comment tarifier au sein de ce portefeuille ? Peut-on faire payer la même cotisation à tous ? Doit-on utiliser les dix années d'observation (A paiera alors trois fois plus cher que C, et G ne paiera rien) ?

	Assurés									
	A	B	C	D	E	F	G	H	I	J
1996									1	
1997	1	1	1						1	
1998	1								1	
1999			1						1	
2000									1	
2001		1								
2002	1	1		1	1					
2003	1			1		1			1	
2004	1				1					
2005	1								1	
moyenne	0,6	0,3	0,2	0,2	0,2	0,1	0,0	0,0	0,7	0

Si la prime collective est globalement juste, elle n'est cependant pas équitable. Depuis les travaux de George Arthur Akerloff, les assureurs savent qu'il est indispensable de prendre en compte l'hétérogénéité lors de la tarification. Parmi les méthodes permettant de prendre en compte le passé sinistre, on retiendra ainsi les systèmes bonus/malus (ou *no claim bonus*), les commissions de réassurance ou les méthodes dites de crédibilité.

■ Croyance, crédibilité et probabilité

Le terme « probabilité » remonte au Moyen-Âge en jurisprudence, issu du mot latin *probare* (signifiant prouver dans le sens que l'on retrouve dans probant) : il s'agissait, lors d'un jugement, d'accorder plus ou moins de poids, de crédibilité, à différents indices ou preuves, à partir de la notoriété (morale le plus souvent) de témoignages. C'est cette idée que l'on retrouvera à la base de *Arte Combinatoria*, la thèse présentée par Gottfried Wilhelm von Leibniz en 1666. En 1634, pour le moine Marin Mersenne dans ses *Questions inouïes*, la probabilité est alors le degré de crédibilité d'une opinion ou d'un jugement sur un fait. Il fallut attendre Antoine Augustin Cournot au XIX^e siècle pour qu'une distinction s'opère « [...] sur le double sens du mot de probabilité, qui tantôt se rapporte à une certaine mesure de nos connaissances, et tantôt à une mesure de la possibilité des choses, indépendamment de la connaissance que nous en avons ».

Les probabilités étaient alors des fréquences : si une urne contient cinq boules blanches et cinq noires, la pro-

babilité de tirer une boule noire est d'une chance sur deux. Le pasteur Thomas Bayes a été le premier à envisager d'induire les probabilités de résultats observés : si sur vingt tirages (avec remise) on a obtenu neuf fois une boule noire, c'est que la probabilité de tirer une boule noire est de 45 %. Cette probabilité est alors vue comme une variable aléatoire dont la loi va dépendre des réalisations observées. Cette théorie sera reprise puis développée par Siméon-Denis Poisson dans *Recherches sur la probabilité des jugements en matière criminelle et matière civile* (on notera que les probabilités restent alors appliquées aux domaines juridiques).

Cette idée avait été étudiée par différents mathématiciens avant d'être jugée « hérétique » par les statisticiens au début du XX^e siècle. La modélisation reposait sur le « Principe de la raison insuffisante » (pour ses partisans) ou l'« Hypothèse de distribution uniforme de l'ignorance » (pour ses détracteurs). Pierre-Simon Laplace avait ainsi essayé d'utiliser ce principe pour déterminer la probabilité d'avoir un garçon ou une fille à la naissance. Dans le chapitre consacré à la probabilité des causes et des événements futurs, tirée des événements observés, dans sa *Théorie analytique*, Pierre-Simon Laplace utilisa les notions de probabilité *a priori* et *a posteriori* avant même d'avoir lu les travaux de Thomas Bayes. La naissance d'un garçon suit une loi de Bernoulli de paramètre p . Ayant observé, entre 1745 et 1784, 393 386 naissances de garçons contre 377 555 naissances de filles, Pierre-Simon Laplace souhaitait quantifier la crédibilité de l'information « p excède $1/2$ » ; autrement dit, on a statistique-

ment plus de chances d'avoir un garçon qu'une fille. L'idée était de supposer *a priori* que p est uniforme sur $[0,1]$, puis d'en déduire la loi de p *a posteriori*, compte tenu des nombres respectifs de naissances observés, en notant que la loi du nombre de naissances de garçons conditionnellement à p suit une loi binomiale,

$$f(p|N=n) = \frac{f(p)}{P(N=n)} P(N=n|p) \propto \binom{n_0}{n} p^n (1-p)^{n_0-n}$$

où n_0 est le nombre total de naissances. Notons que cette réponse soulèvera de nombreux débats puisque, à l'époque, c'est l'opinion de John Arbuthnot qui prétendait que « *cette inégalité des mâles et des femelles n'est pas l'effet du hasard mais celui de la Providence divine* ».

Leonard Savage a été l'un des premiers, en 1954, à défendre la vision subjective des probabilités (*personalistic view of probability*) : « *le seul concept de probabilité essentiel à la science et aux autres activités faisant usage des probabilités* ». En statistique classique (ou fréquentiste), on refuse obstinément toute information subjective ou *a priori* alors qu'elle est la base de l'analyse bayésienne : Quelle crédibilité puis-je atteindre par l'ajout d'informations ? Pour Leonard Savage, *personal probability* est d'ailleurs synonyme de *subjective probability*, *degree of conviction*, ou *degree of credibility*.

Cette idée a été reformulée par un philosophe, prix Nobel de littérature en 1950, Bertrand Russell. Pour cela,

il distingue la probabilité mathématique, énoncée dans un cadre général (« la probabilité pour qu'un homme de 30 ans atteigne 90 ans est de 9,08 % »), de la probabilité particulière (« la probabilité pour que j'atteigne 90 ans est de 20 % »). Si la première peut être obtenue auprès de démographes, ou d'actuaire, il devra consulter un médecin, voire une voyante, pour déterminer la seconde. Et lorsque j'essaie de déterminer ma probabilité d'atteindre 90 ans, j'attribue des crédibilités – totales ou partielles – à ces différents jugements. Bertrand Russell fait alors la distinction entre la probabilité mathématique et ce qu'il appelle la crédibilité, correspondant à ce passage du général au particulier.

■ La prime pure et la loi des grands nombres

L'espérance mathématique de la sinistralité est appelée prime pure en assurance. En effet, depuis les travaux d'Abraham de Moivre ou Blaise Pascal au XVIII^e siècle, on sait que le prix juste d'un jeu (la règle des parties) est le produit scalaire entre les probabilités et les gains.

Quételet avait d'ailleurs justifié son utilisation en assurance incendie en notant que, en tarifant à ce prix, l'assureur serait à l'équilibre en moyenne. La loi des grands nombres nous dit que si un portefeuille est suffisamment grand, la sinistralité moyenne sera très proche de la prime pure, et le théorème central limite permet de quantifier la marge d'erreur. Mais ces théorèmes ne sont qu'asymptotiques : avec un million de contrats – homogènes, c'est une hypothèse de base de la loi des grands nombres – l'erreur

relative sera infime, mais pas avec vingt contrats. Or, faire une tarification fine prenant en compte l'hétérogénéité crée des classes tarifaires constituées de très peu de contrats. De même, tarifier en ne prenant en compte que l'expérience passée d'un assuré pousse à ne retenir que (trop) peu d'informations.

■ Les accidents du travail et la crédibilité américaine

Il y a un siècle, aux États-Unis, General Motors et Tucker (un petit constructeur indépendant) étaient assurés chez Allstate contre les accidents du travail (avec d'autres fabricants automobiles). Un taux moyen était calculé par Allstate, sur l'expérience de l'ensemble du portefeuille, et appliqué à chacun. General Motors calcula son taux de sinistralité sur son propre portefeuille et nota qu'il était plus faible que celui qui lui était imposé. La direction demanda alors à Allstate de lui faire payer son propre taux sous prétexte que son nombre d'employés était suffisamment important pour assurer la stabilité d'année en année. Si les actuaire d'Allstate ont été intuitivement d'accord avec ce principe plein de bon sens, où doit-on fixer la limite ? Peut-on appliquer la même règle à Tucker ?

Arthur Mowbray a été le premier à proposer une réponse en 1914 en utilisant le théorème central limite. Mais la réponse proposée était alors simplement à deux niveaux, 0 ou 1 : oui, la taille de l'entreprise suffit pour faire une tarification individuelle ou non.

C'est Albert Whitney en 1918 qui mentionna la « nécessité, par souci d'équité pour l'assuré, de pondérer d'un côté l'expérience collective, et de l'autre l'expérience individuelle ».

■ La crédibilité européenne ou les querelles entre statisticiens

Dans les travaux initiaux, menés par Albert Whitney, tous les risques étaient gaussiens et, par un développement laborieux, il proposa que la prime individuelle devait être de la forme

$$P = ZX + (1 - Z) C$$

où X est l'expérience individuelle, et C l'expérience collective. Ces deux grandeurs sont pondérées par un « facteur de crédibilité » Z dont Whitney propose la forme $n/(n + K)$ où n est la taille de l'expérience individuelle (le nombre d'employés chez le constructeur automobile, ou le nombre d'années d'expérience pour un assuré). Quant à K , Whitney propose une forme explicite dépendant des paramètres du modèle, mais suggère par souci de simplicité de le fixer de manière arbitraire.

Cette formule a été adoptée rapidement par les actuaire, bien que les fondements statistiques n'étaient pas encore posés. En 1950, Arthur Bailey entreprit d'approfondir les travaux de Whitney : « ... toutes les méthodes d'estimation présentées dans les ouvrages de statistiques, ou enseignées dans les universités américaines, sont basées sur l'hypothèse selon laquelle toute infor-

mation additionnelle, ou toute connaissance a priori est inutile [...] Des philosophes ont récemment étudié la crédibilité à accorder à divers éléments du savoir, remettant ainsi en doute la philosophie de la statistique. Par contre, il semble que ce ne soit que dans les domaines de l'actuariat que l'on ait assisté à une réelle révolte contre le refus de prendre en compte tout savoir a priori ... ». Arthur Bailey a ainsi montré que, en minimisant l'erreur quadratique dans un contexte bayésien, l'estimateur obtenu était une fonction linéaire des observations, et ce, pour des combinaisons de lois binomiale/beta, Poisson/gamma ou normale/normale (ce dernier cas ayant été obtenu par Whitney).

Arthur Bailey a ainsi introduit explicitement le principe de Bayes en tarification et démontré la linéarité de l'estimateur sous certaines conditions (même si l'on peut trouver quelques résultats éparpillés dans des travaux datant des années 1930). Mais il faut

attendre l'intervention de Bruno de Finetti au colloque ASTIN de Trieste en 1963, et celle d'Hans Bühlmann au colloque ASTIN de 1965 à Lucerne, pour que ces méthodes de crédibilité deviennent réellement populaires.

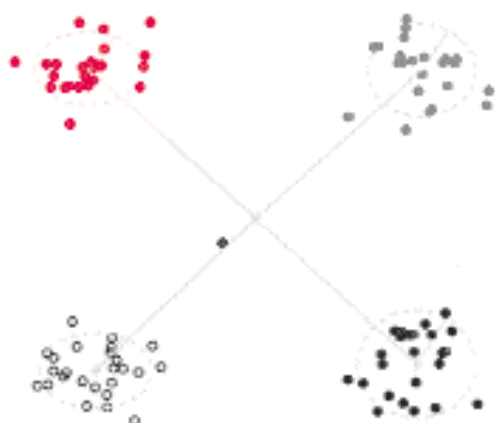
■ La réponse d'Hans Bühlmann

L'idée d'Hans Bühlmann fut de forcer la prime bayésienne à être linéaire et d'obtenir, dans un cadre non paramétrique – la forme de la loi de la variable d'intérêt et de la variable d'hétérogénéité n'a alors plus besoin d'être spécifiée –, un facteur de crédibilité de la forme $Z=n/(n+K)$ où K possède une forme simple, générale et basée sur l'idée très simple que l'on retrouve dans l'exemple 2.

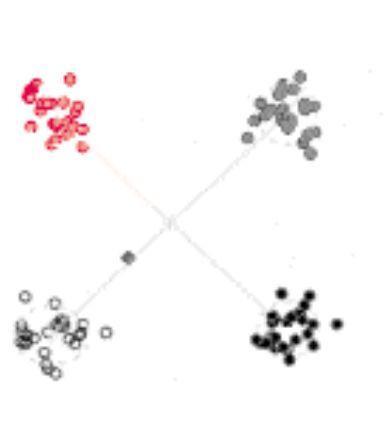
Exemple 2 : considérons un stand de tir où quatre tireurs visent une cible que l'on ne voit pas. On observe alors

les différents impacts que l'on peut rattacher à un tireur. Sur la figure 1, on note qu'un tireur vise plutôt en haut à gauche (les points rouges), un autre en bas à droite (les points noirs), etc. On observe soudain un nouvel impact, que l'on ne peut identifier, et on souhaite le rattacher à l'un des tireurs. Suivant la localisation de l'impact par rapport à l'impact moyen de chaque tireur, on peut le rattacher à tel ou tel tireur, avec plus ou moins de crédibilité. De manière assez naturelle, comme le montre la figure ci-après, si les tireurs visent bien (la dispersion parmi les impacts d'un même tireur est faible), on peut rattacher le nouvel impact à l'un des tireurs avec une crédibilité assez forte : le facteur de crédibilité croît si la *variance inter* diminue. De même, si les cibles sont très espacées, la crédibilité sera là aussi relativement forte : le facteur de crédibilité croît avec la *variance intra*.

Fléchettes : forte variance inter (cibles espacées)



Fléchettes : faible variance intra (bons tireurs)



Hans Bühlman propose ainsi une prime qui reflètera précisément ce comportement. Soit X la sinistralité et Θ le facteur d'hétérogénéité. Pour reprendre les notations de l'exemple 1, si n années ont été observées, alors la prime de crédibilité est :

$$Z\bar{X} + (1-Z)E(X) \text{ où } Z = \frac{n}{n+K} \text{ et } K = \frac{E(\text{Var}(X|\Theta))}{\text{Var}(E(X|\Theta))}.$$

■ Les raffinements de ces méthodes

Ce modèle proposé par Hans Bühlman a été amélioré tout d'abord avec Erwin Straub qui proposa d'ajouter des poids aux données, puis par William Jewell qui introduisit la crédibilité hiérarchique utilisée en France pour la tarification du risque incendie entreprise par exemple.

Plus récemment, Floriaen de Vylder, Hans Gerber, Marc Goovaerts ou Ragnar Norberg ont proposé des extensions dans des cadres plus généraux que le simple calcul d'une prime pure.

Note

1. Notons que pour l'actuariat nord-américain (CAS, SOA ou ICA), les exercices

sur la crédibilité représentent la moitié des exercices de la session trois (risk models), et qu'il y a dix sessions à valider pour devenir actuinaire qualifié (fellow). Les modèles de crédibilité sont alors plus importants que les modèles de taux par exemple, ou que le calcul des annuités.

Bibliographie

BAILEY A.L., « Credibility procedures, Laplace's generalization of Bayes rule and the combination of collateral knowledge with observed data », *Proceedings of the Casualty Actuarial Society*, 37, 1950, pp. 7-23.

BAYES T., « Essay Towards Solving a Problem in the Doctrine of Chances », *Philosophical Transactions of the Royal Society of London*, 1763.

BÜHLMAN H., « Experience rating and

credibility », *ASTIN Bulletin*, 5, 1969, pp. 157-165.

MOWBRAY A.H., « How extensive a payroll exposure is necessary to give a dependable pure premium ? », *Proceedings of the Casualty Actuarial Society*, 1, 1914, pp. 25-30.

NORBERG R., « The credibility approach to experience rating », *Scandinavian Actuarial Journal*, 4, 1979, pp. 181-221.

RUSSELL B., *Human knowledge, its scope and limits*, Routledge, 1948.

SAVAGE L.J., *The foundations of statistics*. John Wiley & Sons, 1954.

WHITNEY A.W., « The theory of experience rating », *Proceedings of the Casualty Actuarial Society*, 4, 1918, pp. 275-293.