

JEAN PINQUET

**Prise en compte de l'ancienneté des périodes
dans les modèles de risque en fréquence : aspects
théoriques et applications à la tarification des
risques en assurance automobile**

Journal de la société française de statistique, tome 144, n° 3 (2003),
p. 7-27

http://www.numdam.org/item?id=JSFS_2003__144_3_7_0

© Société française de statistique, 2003, tous droits réservés.

L'accès aux archives de la revue « Journal de la société française de statistique » (<http://publications-sfds.math.cnrs.fr/index.php/J-SFdS>) implique l'accord avec les conditions générales d'utilisation (<http://www.numdam.org/conditions>). Toute utilisation commerciale ou impression systématique est constitutive d'une infraction pénale. Toute copie ou impression de ce fichier doit contenir la présente mention de copyright.

NUMDAM

Article numérisé dans le cadre du programme
Numérisation de documents anciens mathématiques

<http://www.numdam.org/>

**PRISE EN COMPTE
DE L'ANCIENNETÉ DES PÉRIODES
DANS LES MODÈLES DE RISQUE
EN FRÉQUENCE :
ASPECTS THÉORIQUES ET APPLICATIONS
À LA TARIFICATION DES RISQUES
EN ASSURANCE AUTOMOBILE**

Jean PINQUET *

RÉSUMÉ

Ce papier rappelle les pratiques de tarification *a posteriori* en assurance automobile. Il décrit ensuite les disparités entre ces pratiques et les résultats obtenus par les modèles actuariels usuels, du point de vue de la dynamique des coefficients bonus-malus et de l'importance donnée aux historiques individuels. La principale conclusion des travaux présentés ici est que la prise en compte de l'ancienneté des périodes dans les modèles de description des risques permet de réduire ces disparités. L'analyse rétrospective modifie la structure des réductions de prime suite aux années sans sinistre. Celles-ci sont plus faibles que celles obtenus avec les modèles actuariels usuels si aucun sinistre n'a été observé dans le passé. Elles peuvent être beaucoup plus importantes si l'assuré a déclaré des sinistres dans un passé récent. Enfin, une analyse prospective conduit à une diminution des hausses de prime suite à sinistre, diminution d'autant plus forte que la fidélité anticipée de l'assuré à la compagnie est importante.

ABSTRACT

This paper presents real-world experience rating schemes used for motor insurance. Discrepancies between results drawn from usual actuarial models and real-world practices are described at two levels, namely the dynamic of bonus-malus coefficients and the credibility granted to individual histories. The main finding is that the allowance for the age of periods in a frequency risk model reduces these discrepancies. Retrospective analysis modifies the structure of the no-claim discounts. They are lower than those derived from usual actuarial models if the history is claim-free. On the other hand, they can be much greater if the policyholder reported a claim recently. Finally, a prospective analysis entails a decrease in the financial penalties after one claim. This decrease is all the more significant since the expected loyalty of the policyholder to the company is important.

* Université de Paris X, U.F.R. de sciences économiques, 200 avenue de la République, 92001 NANTERRE CEDEX

Adresses électroniques : pinquet@u-paris10.fr et jpinquet@free.fr

1. Introduction

La tarification dite *a posteriori* (c'est-à-dire fonction de l'historique du contrat) en assurance non-vie est basée le plus souvent sur des nombres d'événements. En assurance automobile, les accidents impliquant une responsabilité du conducteur sont retenus partout dans le monde dans la mise à jour des primes. D'autres types d'événements comme les infractions aux règles de sécurité routière sont utilisés dans la tarification *a posteriori* au Canada et aux États-Unis par exemple.

Les modèles statistiques permettant de calculer des primes de risque en fréquence (c'est-à-dire des estimations de la fréquence des accidents d'un certain type) sont basés sur des lois de Poisson avec composantes de régression et effets aléatoires. Ils sont exposés en section 2. Ils ont donné lieu à de nombreuses contributions dans la littérature actuarielle. Les résultats qu'ils produisent sont à bien des égards très éloignés des pratiques de marché. La section 3 rappelle le principe du système bonus-malus encore en vigueur en France, et passe en revue les différences entre les résultats produits par les modèles statistiques et les pratiques du monde réel.

Sauf quelques rares exceptions (Gerber, 1975, Sundt, 1988), les contributions actuarielles n'ont pas pris en compte l'ancienneté des périodes et des événements dans l'évaluation des risques. Or il est naturel de penser que le pouvoir de prédiction d'une période de l'historique décroît avec son ancienneté. En effet, un individu est plus proche en moyenne de ce qu'il était il y a un an que de ce qu'il était il y a dix ans, par exemple. Ceci vaut aussi bien pour les caractéristiques observables de l'individu – et susceptibles d'être intégrées dans le calcul des primes d'assurance – que pour des composantes inobservables et pesant sur le risque, comme le kilométrage annuel et le comportement au volant. Or la principale explication de la tarification *a posteriori* repose sur une interprétation des historiques comme révélateurs de ces composantes inobservables.

La section 4 présente un modèle qui permet d'intégrer l'ancienneté des périodes d'un historique dans l'évaluation des risques. Les effets aléatoires intégrés dans le modèle de Poisson varient avec le temps et sont supposés suivre un processus stationnaire. Ce faisant, le temps ne sera utilisé qu'au travers des décalages entre périodes d'exposition au risque et périodes d'estimation du risque. Par contre, on négligera les effets calendaires.

La section 5 applique le modèle à des données issues du portefeuille de la plus grande compagnie d'assurance espagnole. Elle reproduit des résultats publiés par Pinquet, Guillén et Bolancé (2001).

La principale conclusion obtenue est que la prise en compte de l'ancienneté des périodes dans le modèle d'évaluation du risque en fréquence permet de rapprocher les résultats des pratiques du monde de l'assurance.

2. Modèles de Poisson sans et avec effets aléatoires. Application à la prédiction des risques en fréquence

Dans cette section, nous rappelons les modèles statistiques utilisés couramment pour intégrer les historiques individuels dans l'évaluation du risque en fréquence. Les développements mathématiques sont concentrés dans les sections 2 et 4, tandis que les sections 3 et 5 sont dédiées aux applications et peuvent être lues presque indépendamment des deux autres.

Raisonnons d'abord sur une seule période. Soit un assuré i , observé pendant une durée d_i et décrit par un vecteur-ligne x_i à k composantes scalaires. On suppose donc que k variables sont retenues comme composantes de régression. En assurance automobile, ces variables sont le plus souvent des indicatrices de modalités de variables de tarification, qui se rapportent au type et à l'âge du véhicule, à la zone géographique, au type d'usage du véhicule, à la c.s.p. du conducteur principal, etc. Si les sinistres ne se produisent qu'un par un, si les accroissements du processus sont indépendants et si la probabilité instantanée d'apparition de ces sinistres est constante sur la période et égale à c_i , le nombre de sinistres observés suit une loi de Poisson de paramètre $\lambda_i = d_i c_i$. Comme c_i est positive, on retient en général la formulation : $c_i = \exp(x_i \beta)$ où β est un vecteur-colonne de \mathbb{R}^k .

La prise en compte de l'historique dans un tel modèle peut se faire de manière très simple. Une approche souvent retenue par les actuaires des compagnies d'assurance consiste à faire figurer des résumés de l'historique dans les composantes de régression. Par exemple, les nombres de sinistres observés durant les périodes précédentes peuvent être retenus comme composantes. Cette approche aisée à mettre en œuvre soulève toutefois des difficultés.

- On ne peut pas intégrer l'exposition au risque dans le passé : or une période du passé a un pouvoir prédictif d'autant plus grand que sa durée est importante ou que les variables observables situent le risque à un niveau élevé.
- Dans le même ordre d'idée, le panel doit être cylindré (chaque historique doit avoir le même nombre de périodes). En effet, supposons qu'on ait des historiques à deux et trois périodes, et qu'on explique le nombre de sinistres d'un certain type observés en dernière période. Le nombre de sinistres observés durant l'antépénultième période est nul pour les historiques à deux périodes parce que l'exposition au risque est inexistante. Cette variable n'est donc pas comparable à celle calculée pour les historiques à trois périodes. Cette limite est fâcheuse car retenir des panels cylindrés induit des biais de sélection.
- Enfin, la forme de la prime-fréquence (estimation de l'espérance du nombre des sinistres) sera contrainte par la régression. Si les nombres sont inclus dans un score additif, la prime-fréquence sera une fonction exponentielle des nombres. C'est peu ou prou le cas dans le système bonus-malus français, au moins dans la première phase de la vie de

l'assuré. Les modèles actuariels que nous développerons ci-après sont en fait sujets à la même critique.

- De tels modèles ne donnent pas directement une évaluation de l'historique. La prime calculée par la régression est le produit d'une prime fonction des caractéristiques présentes du contrat, et d'une autre résumant l'historique. La seconde prime est égale à un en l'absence de sinistres. La première représente donc le risque d'un contrat non sinistré dans le passé, et méritant d'habitude un « bonus ». Ce n'est pas une prime *a priori*, qui correspond à un contrat sans historique.

Les modèles actuariels reposent sur un principe de révélation par les historiques des composantes inobservées des lois de risque fréquence. On supposera dans un premier temps que ces composantes sont résumées par un effet fixe indépendant du temps. En reprenant le modèle à une période exposé en début de section et en notant N_i la variable nombre d'accidents d'un type donné, on écrira

$$N_i \text{ suit une loi } P(\lambda_i u_i^m),$$

$$\lambda_i = d_i \exp(x_i \beta) \Leftrightarrow \lambda_i u_i^m = d_i \exp(x_i \beta + u_i^a), u_i^a = \log(u_i^m) \quad (2.1)$$

Le paramètre individuel (additif u_i^a ou multiplicatif u_i^m) est un effet fixe, appelé aussi composante d'hétérogénéité. Il traduit l'effet des variables inobservables sur les lois de risques. Un coefficient bonus-malus actuariel sera un prédicteur de u_i^m basé sur l'historique. Il s'applique à la période suivante. La prime-fréquence pour cette nouvelle période est le produit d'une prime *a priori*, définie à partir des nouvelles caractéristiques du contrat, et du coefficient bonus-malus qui résume l'influence du passé sur l'évaluation du risque en fréquence.

Une définition précise de l'effet fixe conduit à la construction des modèles à effets aléatoires. Supposons que les variables inobservables soient représentées par le vecteur-ligne z_i pour le contrat i . En retenant une formulation analogue à celle du modèle de Poisson, on écrira

$$\lambda_i \exp(u_i^a) = d_i \exp(x_i \gamma + z_i \delta). \quad (2.2)$$

Un rapprochement avec l'équation (2.1) suggère la solution : $\gamma = \beta$, $u_i^a = z_i \delta$. Une telle formulation n'est pas satisfaisante car elle ne tient pas compte de la loi jointe entre observables et inobservables. Par exemple, le kilométrage annuel d'un véhicule décroît avec son ancienneté. La variable observable qu'est l'ancienneté du véhicule est ainsi un « proxy » de la variable inobservable : kilométrage annuel, laquelle pèse évidemment sur le risque en fréquence. On peut calculer un effet des variables inobservables sur le risque qui soit résiduel par rapport aux variables observables. L'effet fixe associé à un kilométrage annuel donné devrait croître avec l'ancienneté du véhicule, celui-ci étant d'autant plus risqué par rapport aux véhicules de même caractéristiques qu'il est âgé.

On obtient une définition résiduelle de l'effet fixe en définissant une loi sur des classes d'individus réels ne se différenciant que sur les caractéristiques inobservables. Soit par exemple la classe des individus de mêmes caractéristiques observables que i . Ils sont définis par les variables (x_i, z) . Avec une loi jointe sur l'ensemble des variables explicatives – observables ou non – on peut définir une loi sur cette classe d'individus de la manière suivante : si $l(n_i|x_i, z_i)$ est la vraisemblance de la loi définie en (2.2) dans le modèle à effets fixes, la vraisemblance $\tilde{l}(n_i|x_i)$ associée à la classe d'individus ayant les mêmes caractéristiques observables x_i est définie par

$$\tilde{l}(n_i|x_i) = E[l(n_i|X_i, Z)|X_i = x_i]. \quad (2.3)$$

Le couple (X_i, Z) suit une loi jointe qui est en général supposée indépendante de i . Notons $l^*(n_i|x_i, u_i^a) = l(n_i|x_i, z_i)$ la vraisemblance définie en (2.2) pour le modèle à effets fixes. Moyennant quelques hypothèses supplémentaires (cf. Pinquet, 2000), on obtient

$$\tilde{l}(n_i|x_i) = E[l^*(n_i|x_i, U_i^a)], \text{ avec} \quad (2.4)$$

$$U_i^a = Z_i\delta - E(Z_i\delta|X_i = x_i); \quad u_i^a = z_i\delta - E(Z_i\delta|X_i = x_i). \quad (2.5)$$

L'équation (2.4) définit la vraisemblance du modèle à effets aléatoires, obtenue en intégrant la vraisemblance du modèle à effets fixes sur l'effet aléatoire U_i^a . Le modèle à effets aléatoires est donc défini sur des classes d'individus réels ne se différenciant que sur les caractéristiques inobservables. L'effet fixe u_i^a est bien calculé de manière résiduelle dans l'équation (2.5).

Dans l'exemple précédent, on a :

$$E(U_i^a) = 0; \quad V(U_i^a) = V(Z_i\delta) - V(Z_i\delta|X_i = x_i). \quad (2.6)$$

Les lois sur les effets aléatoires sont en général caractérisées de manière paramétrique ou semi-paramétrique à partir de leurs moments du second ordre. L'exemple le plus courant est le modèle binomial négatif, où on retient des lois gamma d'espérance un pour les effets aléatoires multiplicatifs. Les paramètres d'échelle et de forme sont identiques, et égaux à l'inverse de la variance. On peut faire une estimation par le maximum de vraisemblance, car celle-ci est explicite. Dans ce cadre paramétrique, le prédicteur de l'effet fixe multiplicatif est défini par l'estimation de l'espérance de l'effet aléatoire calculée conditionnellement à l'historique (cf. Dionne et Vanasse, 1989). La prédiction se fait suivant une démarche de type « empirical Bayes ». L'*a priori* ne porte pas sur l'ensemble des données, comme dans l'approche Bayésienne classique, mais sur une classe d'individus dont la loi est définie dans le modèle à effets aléatoires. Dans le cas du modèle binomial négatif, les lois *a priori* et *a posteriori* sur l'effet aléatoire sont des lois gamma.

Cette approche basée sur une spécification paramétrique des effets aléatoires s'adapte difficilement à des situations plus complexes, comme par exemple au problème abordé dans ce papier. La vraisemblance ainsi que le prédicteur

par espérance conditionnelle cessent en général d'être explicites. On doit alors recourir à des approximations (intégration numérique ou simulations) coûteuses en temps de calcul et fournissant des résultats peu lisibles, selon l'auteur de cet article.

Une solution alternative très populaire dans la littérature actuarielle et appelée méthode de crédibilité linéaire (*cf.* Bühlmann, 1967) consiste à se placer dans un cadre semi-paramétrique au niveau des effets aléatoires. Le prédicteur de l'effet fixe multiplicatif est contraint à être une fonction affine du nombre des sinistres. Il est calculé par une régression probabiliste estimée dans le modèle à effets aléatoires. On reprend l'équation (2.1), en retenant une formulation multiplicative pour l'effet fixe. Si les données sont en panel, n_i et λ_i sont les sommes des variables associées à chaque période pour l'individu i , l'effet fixe étant supposé constant dans le temps. On fait les hypothèses suivantes sur les effets aléatoires :

$$E(U_i) = 1 ; V(U_i) = \sigma^2 \quad \forall i. \quad (2.7)$$

On en déduit les égalités

$$E(N_i) = \lambda_i ; V(N_i) = \lambda_i + \lambda_i^2 \sigma^2 ; \text{Cov}(U_i, N_i) = \lambda_i \sigma^2 \quad (2.8)$$

dans le modèle à effets aléatoires. La régression affine probabiliste de U_i sur N_i conduit au prédicteur suivant

$$\frac{1 + n_i \hat{\sigma}^2}{1 + \hat{\lambda}_i \hat{\sigma}^2} = (1 - cred_i) + cred_i \frac{n_i}{\hat{\lambda}_i}, \text{ avec } cred_i = \frac{\hat{\lambda}_i \hat{\sigma}^2}{1 + \hat{\lambda}_i \hat{\sigma}^2}. \quad (2.9)$$

Le coefficient $cred_i$ est appelé crédibilité donnée à l'historique dans la prédiction. Il correspond à un bonus en l'absence de sinistre, et il augmente avec l'exposition au risque mesurée par la prime-fréquence, ainsi qu'avec la variance estimée de l'effet aléatoire qui mesure l'hétérogénéité résiduelle. On obtient le même prédicteur par espérance conditionnelle dans le modèle binomial négatif. Les équations données en (2.8) conduisent à des estimateurs des paramètres β et σ^2 par la méthode des moments. L'estimateur de β , vecteur des paramètres associés aux composantes de régression, est celui du maximum de vraisemblance du modèle de Poisson sans effets fixes ni aléatoires. Les équations définissant $\hat{\beta}$ et $\hat{\sigma}^2$ dans le cas d'historiques à une période sont

$$\sum_i (n_i - \hat{\lambda}_i)^t x_i = 0 ; \hat{\sigma}^2 = \frac{\sum_i (n_i - \hat{\lambda}_i)^2 - n_i}{\sum_i \hat{\lambda}_i^2}, \quad (2.10)$$

avec $\hat{\lambda}_i = d_i \exp(x_i \hat{\beta})$. Les estimateurs peuvent être améliorés et atteindre une borne d'efficacité semi-paramétrique en intégrant la variance de l'effet aléatoire dans la régression (*cf.* Liang et Zeger, 1986). L'équation définissant

$\hat{\sigma}^2$ dans (2.10) est inchangée, et celle définissant $\hat{\beta}$ est :

$$\sum_i \frac{\hat{E}(N_i)}{\hat{V}(N_i)} (n_i - \hat{\lambda}_i)^t x_i = \sum_i \frac{n_i - \hat{\lambda}_i}{1 + \hat{\lambda}_i \hat{\sigma}^2} {}^t x_i = 0. \quad (2.11)$$

Cette équation en β est celle du maximum de vraisemblance dans le modèle binomial négatif. Un avantage certain à cette approche est qu'elle se généralise aisément à des situations plus complexes comme la prise en compte de l'ancienneté des périodes dans la tarification.

Les trois manières d'intégrer l'historique dans la prédiction du risque fréquence rappelées dans cette section (résumé de l'historique dans les composantes de régression, formulations paramétriques et semi-paramétriques sur des effets aléatoires) ont en commun le fait de contraindre la forme du prédicteur comme fonction des nombres de sinistres. Cette forme est explicite dans la première et la troisième approche. Dans le cas de la formulation paramétrique sur les effets aléatoires, la forme est induite par le choix de la famille de lois. On obtient des prédicteurs linéaires par rapport aux nombres avec des lois gamma pour les effets aléatoires, et ces prédicteurs deviennent convexes si les lois gamma sont remplacées par des lois log-normales. La seule manière d'obtenir des prédicteurs de forme non contrainte explicitement ou implicitement serait de faire une estimation non paramétrique de la loi des effets aléatoires (estimation de densité pour des lois continues par exemple). On en déduirait les prédicteurs par espérance conditionnelle. L'approche la plus populaire dans la littérature statistique consiste à représenter la loi des effets aléatoires par des moments d'ordre étagés. On va ainsi du semi-paramétrique vers le non-paramétrique, alors que l'approche actuarielle de crédibilité linéaire adopte un point de vue semi-paramétrique restreint aux moments du second ordre. Si la démarche est intéressante sur le plan théorique, elle se révèle d'application difficile dans un contexte où les fréquences sont faibles.

3. Les pratiques de tarification *a posteriori* en assurance automobile, et l'exemple du système bonus-malus français

La tarification *a posteriori* en assurance automobile dans le monde connaît une tendance à la déréglementation. En Europe, la troisième directive d'assurance non-vie édictée par la commission européenne interdit toute réglementation tarifaire. Les autorités de Bruxelles sont actuellement en conflit avec la France sur le maintien du système bonus-malus. On notera que les systèmes réglementés restent très présents sur le continent nord-américain.

Un système bonus-malus réglementé a le plus souvent une structure markovienne. Des coefficients multiplicatifs s'appliquant à une prime de base, fonction des caractéristiques du contrat, sont associés à des niveaux qu'on peut se représenter comme les barreaux d'une échelle. Le conducteur débutant se voit assigner un niveau en général assez élevé dans l'échelle. Ensuite, le changement de niveau se fait d'une année à la suivante sur la base du nombre des

événements d'un certain type déclarés durant l'année. Ces événements sont le plus souvent des accidents engageant la responsabilité du conducteur, comme en France. Mais d'autres événements associés à une notion de faute, comme les infractions aux règles de sécurité routière, sont également utilisés au Canada et aux États-Unis. On trouvera une étude exhaustive de ces systèmes dans le livre de Jean Lemaire (1995). Le plus souvent, l'assuré descend un barreau suite à une année sans sinistre (ou événement d'un autre type), pour autant qu'il n'ait pas atteint le niveau minimal. S'il a déclaré un ou plusieurs sinistres, il gravit plusieurs barreaux en un an. Les systèmes bonus-malus sont en général déséquilibrés à la baisse, et les assurés s'agglutinent vers les niveaux inférieurs de l'échelle au fur et à mesure que leur ancienneté augmente.

Dans le système français, le coefficient bonus-malus (appelé coefficient de réduction-majoration et noté CRM dans la suite de l'article) prend des valeurs quasiment continues. Il vaut un nombre entier de centièmes, ce nombre variant entre 50 et 350 et figurant obligatoirement sur le contrat d'assurance. La prime payée est le produit d'une prime de base – qui doit également être explicitée dans le contrat – et du CRM. Une remarque importante à faire est que la détermination de cette prime de base est totalement déréglementée. En particulier, le calcul peut intégrer des éléments de l'historique individuel de l'assuré. De ce fait, les griefs faits par la commission de Bruxelles au système français sont non fondés si on raisonne uniquement du point de vue de la déréglementation de la tarification du contrat, qui est à la base de la troisième directive européenne en assurance non-vie. Le CRM porte en fait mal son nom : il a plutôt pour fonction de fournir un résumé de l'historique de l'assuré accessible à tous les acteurs du marché, car il figure sur le contrat d'assurance. Les débats avec les autorités de régulation devraient porter sur l'évolution d'un système d'information commun aux compagnies d'assurance plutôt que sur la tarification qui est de fait complètement déréglementée.

Détaillons le calcul du CRM. On effectuera également des comparaisons avec les modèles statistiques exposés dans la section précédente. Un conducteur novice débute avec un CRM égal à 1. En l'absence de sinistre responsable dans l'année, le conducteur se voit attribuer un bonus de 5%. Le CRM est multiplié par 0,95 et arrondi au centième inférieur en restant borné inférieurement par 0,50. Si n sinistres responsables sont déclarés dans l'année, le CRM est multiplié par $1,25^n$ (malus de 25%) et arrondi au centième inférieur en restant borné supérieurement par 3,50. Comme $1,25 \times 0,95^4$ est voisin de 1 on peut considérer que si l'assuré descend un barreau suite à une année sans sinistres, il en gravit quatre suite à un sinistre tant que les bornes inférieures et supérieures ne sont pas atteintes. La fréquence moyenne annuelle des événements déclenchant un malus étant de l'ordre de un quinzième, le CRM moyen décroît avec l'ancienneté de l'assuré. Il vaut moins de 0,60 sur l'ensemble des assurés français, et plus de 60% d'entre eux ont le bonus maximum.

Mentionnons pour finir deux dispositions complémentaires. La clause dite de retour rapide au coefficient initial concerne la petite minorité des conducteurs ayant un CRM supérieur à un : deux années sans sinistre permettent de

retrouver une valeur de un, quelle que soit la valeur initiale du CRM. La motivation de cette clause est de créer des incitations à une conduite prudente pour les assurés avec un historique très médiocre. L'analyse économique des contrats d'assurance (*cf.* Henriet et Rochet, 1986) fait apparaître que les incitations créées par un système punition-récompense tel le CRM sont d'autant plus efficaces que les punitions et récompenses sont importantes et que le système a peu de mémoire. Cette clause de retour rapide satisfait à ces deux conditions d'efficacité.

L'autre disposition concerne la majorité des conducteurs ayant le CRM minimum de 50 %. Il est clair qu'à ce niveau on a une source importante d'inefficacité du système bonus-malus en terme de discrimination des risques. En effet, plus de 60 % des assurés ont le même coefficient, alors que leurs historiques permettraient de les différencier. Les assureurs français ont introduit depuis longtemps des rabais à la prime de base pour leurs meilleurs assurés, ainsi que des franchises de malus. Le possesseur d'une telle franchise voit son CRM inchangé s'il déclare un sinistre responsable. Il perd par contre sa franchise, qu'il récupère après quelques années sans sinistre. Ce système a été inclus au niveau national pour les assurés ayant le CRM maximum depuis trois ans, ce qui est aussi la durée pour récupérer la franchise.

Comparons maintenant le CRM français à des coefficients calculés à partir de données réelles, sur la base du modèle décrit dans la section précédente et résumé par la formule (2.9). La variance estimée $\hat{\sigma}^2$ de l'effet aléatoire mesure ce qu'apporte l'information cachée à l'information observable dans l'explication du risque en fréquence (*cf.* l'équation (2.6) et ce qui précède). On obtient en général une valeur comprise entre 0,5 et 1. Comme la prime-fréquence $\hat{\lambda}_i$ est très inférieure à un, cette variance estimée représente à peu de chose près un malus, c'est-à-dire une hausse relative de la prime-fréquence suite à un événement dont on explique la loi de nombre. Si cet événement est un sinistre responsable, on obtient donc une valeur nettement supérieure à 0,25, valeur du malus dans le système français pour les assurés n'ayant pas la franchise de malus. Comment expliquer cette disparité ? Il faut savoir que les tarifications du monde réel en assurance non-vie raisonnent implicitement à un horizon plus éloigné que l'année à venir, et que cet horizon est d'autant plus éloigné que la fidélité anticipée de l'assuré est élevée. Or la période de risque étudiée par les modèles actuariels est toujours l'année à venir, cette période se terminant ainsi à la date de renouvellement du contrat. Si le pouvoir prédictif d'un événement décroît avec son ancienneté, on conçoit qu'un malus de moyen ou long terme, comme celui d'un système du monde réel, soit plus faible que le malus de court terme calculé par un modèle actuariel. On reviendra sur ce problème à la fin de la section 5.

L'équation (2.9) donne le bonus cumulé pour un historique sans sinistre. Il est égal à la crédibilité donnée à l'historique. Après un an, on obtient à peu près 5 % si $\hat{\sigma}^2 = 0,75$; $\hat{\lambda}_i = 1/15$, qui représentent une valeur crédible de la variance et un risque moyen en fréquence. Si les résultats sont semblables à ce niveau, il ne faut pas oublier que le système de CRM français est déséquilibré. Un bonus de 5 % pour un conducteur novice après un an ne représente qu'une

baisse du coefficient de 2 ou 3% par rapport à la moyenne des CRM des conducteurs ayant un an d'ancienneté, par exemple. En raisonnant à plus long terme, les disparités entre les deux approches ne font que s'accroître. D'après l'équation (2.9), la crédibilité tend vers un quand le nombre de périodes croît vers l'infini. Ceci signifie par exemple que la prime d'un assuré n'ayant jamais déclaré de sinistre devrait tendre vers 0. Ce n'est pas du tout le cas avec le système de CRM français, pour lequel le coefficient minimal 0,50 n'est inférieur que d'à peine plus de 10% à la moyenne des CRM des assurés avec une ancienneté importante. La prise en compte de l'ancienneté des périodes dans le calcul actuariel permet d'expliquer cette différence importante et de la réduire notablement.

Par ailleurs, la dynamique des coefficients issus des deux systèmes est très différente. On a vu qu'il fallait quatre années sans sinistre pour effacer un malus dans le CRM français. Les coefficients issus de la formule (2.9) mettent en regard un nombre n_i et $\hat{\lambda}_i$, qui est l'estimation de son espérance. Suite à un sinistre, il faut une durée inverse de la prime fréquence annuelle pour annuler le malus. Avec une prime fréquence de un quinzième, quatorze années sans sinistre sont nécessaires pour retrouver le coefficient calculé au début de l'année de survenance du dernier sinistre. Le résultat est là aussi très différent de celui issu du CRM français. La prise en compte de l'ancienneté des périodes et des événements change radicalement les résultats, s'agissant des réductions de prime suite aux années sans sinistre. Les résultats, qui sont exposés dans la section 5, permettent de rapprocher les coefficients bonus-malus actuariels de ceux du monde réel.

4. Modèles de Poisson avec effets aléatoires dynamiques et stationnaires

Considérons un portefeuille de p contrats. Le contrat i , $i = 1, \dots, p$, est observé sur T_i périodes, avec $1 \leq T_i \leq T_{\max}$. Le panel est donc non cylindré *a priori*. Si $n_{i,t}$ est le nombre d'événements d'un certain type (par défaut les sinistres en responsabilité), on définit un modèle de Poisson avec des effets aléatoires dynamiques $U_{i,t}$ dans un contexte paramétrique à partir d'une loi jointe sur les $(U_{i,t})_{1 \leq t \leq T_i}$. La vraisemblance de la suite $(n_{i,t})_{1 \leq t \leq T_i}$ est définie par

$$L[(n_{i,t})_{1 \leq t \leq T_i}] = E \left[\prod_{t=1}^{T_i} P_{\lambda_{i,t} u_{i,t}}(n_{i,t}) \right],$$

$$\text{avec } \lambda_{i,t} = d_{i,t} \exp(x_{i,t} \beta) ; P_{\lambda}(n) = \exp(-\lambda) \frac{\lambda^n}{n!} \quad (4.1)$$

L'espérance mathématique est calculée sur les effets aléatoires, qui sont notés de manière multiplicative car on travaillera dans la suite avec des spécifications semi-paramétriques. Les moments des nombres se déduisent des moments des effets aléatoires multiplicatifs, d'où ce choix.

On peut interpréter le modèle à effets aléatoires comme on l'a fait en section 2. Si on suppose que le modèle à effets fixes s'applique aux individus réels, le produit de l'équation (4.1) signifie qu'on suppose l'indépendance des $(N_{i,t})_{t=1,\dots,T_i}$ pour le contrat i . La loi définie en (4.1) s'applique alors à une classe de contrats ayant les mêmes caractéristiques observables que le contrat i . L'approche actuarielle de prédiction suppose alors que l'historique a exclusivement un pouvoir de révélation des caractéristiques cachées des lois de risque, mais qu'il ne modifie pas les lois des individus réels. Cette hypothèse est certainement fautive : les incitations financières ainsi que les effets psychologiques (modifications de la perception du risque suite à un accident) induisent une baisse du niveau de risque suite à un sinistre. La mesure des effets incitatifs créés par un système bonus-malus est difficile à entreprendre, car elle se heurte à des problèmes d'identifiabilité de la nature de la dynamique sur les données, lesquels sont ignorés par l'approche actuarielle. On trouvera une synthèse et une application à des données françaises dans Abbring, Chiappori, Heckman et Pinquet (2003). Cela dit, un calcul de prime actuarielle ne nécessite pas d'avoir identifié les différentes sources de corrélation entre observations (effet révélation et effet modification). On observe la résultante de ces différentes causes, et c'est la seule chose qui importe dans une démarche de prédiction.

On peut également introduire une corrélation entre les $(N_{i,t})_{t=1,\dots,T_i}$ en définissant chaque nombre conditionnellement au précédent. Les modèles INAR(1) (Al-Osh et Alzaid, 1987, Mc Kenzie, 1988, Gouriéroux et Jasiak, 2002) définissent chaque nombre $N_{i,t}$ comme une somme de $N_{i,t-1}$ variables de Bernoulli indépendantes et de même loi, et d'une variable suivant une loi de Poisson et indépendante des précédentes. On obtient en quelque sorte un processus AR(1) adapté aux variables à valeurs entières, d'où le nom donné au modèle.

La vraisemblance définie en (4.1) n'a pas de forme explicite dans un contexte paramétrique pour les familles de lois usuelles. Gouriéroux et Jasiak (2001) proposent une analyse dans ce cadre. Le modèle de Poisson avec effets aléatoires dynamiques sera estimé suivant une approche semi-paramétrique, en faisant une hypothèse de stationnarité au second ordre pour les effets aléatoires. On supposera que des propriétés d'équidistribution et de stationnarité sont vérifiées au premier et au second ordre pour les effets aléatoires. Cela veut dire que les moments de $(U_{i,t})_{1 \leq t \leq T_i}$ sont ceux de $(U_t)_{1 \leq t \leq T_i}$ aux deux premiers ordres, et que la loi de $(U_t)_{1 \leq t \leq T_{\max}}$ est stationnaire au second ordre. On fait les hypothèses et notations suivantes :

$$E(U_t) = 1 \forall t = 1, \dots, T_{\max} ; \text{Cov}(U_t, U_{t-h}) = \sigma_U^2 \rho_U(h), \quad 0 \leq h < t \leq T_{\max} \quad (4.2)$$

avec : $-1 \leq \rho_U(h) \leq 1 ; \rho_U(0) = 1$

L'hypothèse sur l'espérance de l'effet aléatoire n'est pas restrictive. Pour une famille de lois sur $(\mathbb{R}^+)^{T_{\max}}$ telle les log-normales, l'espérance varie avec la variance. En divisant par l'espérance, on obtient une famille de lois paramétrée par les moments du second ordre. Les paramètres d'échelle sont

déjà représentés par la constante dans les composantes de régression du modèle de Poisson.

Avec les notations suivantes

$$\begin{aligned}\alpha &= \text{vec}_{0 \leq h < T_{\max}} (\sigma_U^2 \rho_u(h)); N_i = \text{vec}_{1 \leq t \leq T_i} (N_{i,t}); U_i = \text{vec}_{1 \leq t \leq T_i} (U_{i,t}); \\ \lambda_i &= \text{vec}_{1 \leq t \leq T_i} (\lambda_{i,t}); \Lambda_i = \text{diag}_{1 \leq t \leq T_i} (\lambda_{i,t})\end{aligned}\quad (4.3)$$

l'espérance et la variance dans le modèle à effets aléatoires de chaque série individuelle de nombres d'événements s'écrivent

$$E(N_i|\alpha, \beta) = \lambda_i; V(N_i|\alpha, \beta) = \Lambda_i + \Lambda_i V(U_i|\alpha) \Lambda_i. \quad (4.4)$$

Du fait de la constance de l'espérance des effets aléatoires, l'espérance de N_i ne dépend pas de α , vecteur des moments du second ordre des $U_{i,t}$. En conséquence l'estimateur de β du maximum de vraisemblance dans le modèle de Poisson sans effets aléatoires est convergent dans le modèle à effets aléatoires (cf. Gouriéroux, Monfort et Trognon, 1984a et 1984b). Si $\hat{\lambda}_{i,t}$ est la prime-fréquence déduite de cet estimateur, la deuxième partie de l'équation (4.4) fournit des estimateurs convergents de α par la méthode des moments, qui sont égaux à

$$\hat{\sigma}_U^2 = \frac{\sum_{i,t} (n_{i,t} - \hat{\lambda}_{i,t})^2 - n_{i,t}}{\sum_{i,t} \hat{\lambda}_{i,t}^2} \quad (4.5)$$

$$\hat{\sigma}_u^2 \hat{\rho}_U(h) = \frac{\sum_{i|T_i > h} \sum_{h < t \leq T_i} (n_{i,t} - \hat{\lambda}_{i,t})(n_{i,t-h} - \hat{\lambda}_{i,t-h})}{\sum_{i|T_i > h} \sum_{h < t \leq T_i} \hat{\lambda}_{i,t} \hat{\lambda}_{i,t-h}} \quad (1 \leq h < T_{\max}). \quad (4.6)$$

Les deux dernières équations expriment $\hat{\alpha}$ en fonction de $\hat{\beta}$. On écrira $\hat{\alpha} = \alpha_*(\hat{\beta})$. On peut améliorer la précision des estimateurs précédents en intégrant les moments estimés par $\hat{\alpha}$ dans la régression, comme on l'a fait dans l'équation (2.11). L'équation de vraisemblance en β du modèle de Poisson sans effets aléatoires s'écrit $\sum_{i,t} (n_{i,t} - \lambda_{i,t})^t x_{i,t} = 0$ ($0 \in \mathbb{R}^k$). Matriciellement, cette équation est équivalente à

$$\sum_i^t \left(\frac{\partial}{\partial \beta} E(N_i|\beta) \right) [V(N_i|0, \beta)]^{-1} (n_i - E(N_i|\beta)) = 0 \quad (4.7)$$

Les moments de N_i sont exprimés dans le modèle à effets aléatoires (cf. (4.4)). On a intégré le fait que $E(N_i|\alpha, \beta)$ est indépendant de α . L'estimateur amélioré de β s'obtient comme solution de l'équation obtenue en remplaçant dans (4.7) $V(N_i|0, \beta)$ par $V(N_i|\alpha_*(\hat{\beta}), \beta)$. Ses propriétés asymptotiques sont données dans Zeger (1988). Les équations (4.5) et (4.6) conduisent alors à un nouvel estimateur de α . Sur de grosses bases de données, les modifications

des estimations sont infimes, mais l'intérêt est d'avoir une appréciation plus réaliste de la précision de ces estimations.

Les deux estimateurs de α (fonction tronquée d'autocovariance des effets aléatoires) sont non contraints, alors que le domaine de définition de α est une partie propre de $\mathbb{R}^{T_{\max}}$ (c'est en fait un cône). De ce fait, on peut obtenir des estimateurs incohérents (variance estimée négative, coefficients d'autocorrélation estimés n'appartenant pas à $[-1,1]$, ou plus généralement matrice de variances-covariances non semi-définie positive). Dans ce cas, il faut simplifier (voire supprimer) la famille de lois sur les effets aléatoires pour retrouver une estimation utilisable. Le fait que les estimateurs soient non contraints ne peut être vu comme un inconvénient. Si les données conduisent à une estimation incohérente, un estimateur contraint (issu par exemple de la maximisation d'une vraisemblance approchée par intégration numérique ou par simulation) serait le plus souvent obtenu à la frontière du cône de paramétrisation. Comme avec l'approche non contrainte, il faudrait alors simplifier la famille de lois sur les effets aléatoires.

Une fois la fonction d'autocovariance estimée, on peut obtenir des prédicteurs pour les périodes futures avec l'approche de crédibilité linéaire exposée dans la section 2 pour le modèle de Poisson avec effets aléatoires constants. Soit (n_1, \dots, n_T) (avec $T < T_{\max}$) la suite des nombres d'événements observés sur un contrat. On a supprimé l'indice individuel pour alléger les notations. Un coefficient bonus-malus actuariel pour la période $T + 1$ sera obtenu par une régression affine probabiliste de U_{T+1} par rapport à N_1, \dots, N_T dans le modèle à effets aléatoires. Le prédicteur de l'effet fixe u_{T+1} basé sur cette régression est égal à

$$1 + \sum_{t=1}^T b_t (n_t - \hat{\lambda}_t)$$

$$\text{avec } \hat{V} \left(\text{vec} (N_t) \right) \times \text{vec} (b_t) = \text{vec} (\text{Cov}(N_t, U_{T+1})) \quad (4.8)$$

Les primes fréquences $\hat{\lambda}_t$ sont estimées dans le modèle de Poisson sans effets aléatoires car la spécification de l'espérance est inchangée en présence de ces effets. La prime-fréquence *a posteriori* pour la période $T + 1$ est le produit de la prime *a priori* $\hat{\lambda}_{T+1}$ et du prédicteur calculé en (4.8), qui est le coefficient bonus-malus. Les crédibilités associées à chaque période de l'historique sont égales à $\text{cred}_t = \hat{\lambda}_t b_t$, et le coefficient bonus-malus pour la période $T + 1$ est égal à

$$\left(1 - \sum_{t=1}^T \text{cred}_t \right) + \sum_{t=1}^T \text{cred}_t \frac{n_t}{\hat{\lambda}_t}. \quad (4.9)$$

Compte tenu des moments du premier et du second ordre de U_{T+1} et N_1, \dots, N_T (il suffit d'étendre les résultats donnés en (2.8) avec les hypothèses faites en (4.2)), les crédibilités sont solutions du système linéaire suivant :

$$cred_t + \left(\hat{\lambda}_t \sum_{t'=1}^T \hat{\sigma}_U^2 \hat{\rho}_U(|t'-t|) cred_{t'} \right) = \hat{\lambda}_t \hat{\sigma}_U^2 \hat{\rho}_U(T+1-t) \quad (t = 1, \dots, T) \quad (4.10)$$

D'après ces équations, la crédibilité $cred_t$ est équivalente à $\hat{\lambda}_t \hat{\sigma}_U^2 \hat{\rho}_U(T+1-t)$ ($t = 1, \dots, T$) si les primes-fréquences sont négligeables par rapport à un. De fait, les coefficients de crédibilité associés à un historique ont la forme du corrélogramme estimé sur les effets aléatoires, avec un indice de temps inversé (cf. la section 5).

Dans ce contexte, l'approche de crédibilité linéaire se révèle supérieure à des calculs d'espérance conditionnelle dans un cadre paramétrique, que ce soit au niveau du volume des calculs requis ou du point de vue de la lisibilité des résultats. La crédibilité linéaire a pourtant des limites importantes : par exemple, le prédicteur pourrait être négatif même si l'estimation d'un processus stationnaire était possible sur les données. Dans le cas fort improbable où le coefficient de corrélation d'ordre un serait négatif, un sinistre en première période entraînerait un bonus, et le prédicteur deviendrait négatif si le nombre des sinistres était assez grand. Mais en pratique, l'application des prédicteurs présentés ci-dessus ne pose pas de problème, y compris dans des analyses en longue période.

Dans la fin de cette section, nous abordons le problème de l'extension dans le temps de la structure de corrélation entre effets aléatoires estimée en (4.5) et (4.6). Dans l'étude empirique de la section 5, le nombre maximal des périodes des historiques individuels est $T_{\max} = 7$. On estimera donc les six premiers coefficients d'autocorrélation entre les effets aléatoires, mais on pourrait avoir à étudier des historiques plus longs ultérieurement, et surtout vouloir prévoir le risque au-delà de l'année qui vient, et intégrer une fidélité anticipée du client.

En estimant une famille paramétrée de processus stationnaires à partir des données, on obtiendra une estimation pour les décalages d'ordre supérieur ou égal à T_{\max} . Les spécifications usuelles sont de type moyenne mobile ou autorégressive. Les autocorrélations estimées en section 5 sont toutes significativement différentes de zéro, et les moyennes mobiles ne sont pas plausibles. Comme les corrélations décroissent avec le décalage, on étudiera des formulations autorégressives.

Ces classes de processus s'appliquent à des variables dont le support est non contraint *a priori*. De ce fait, l'estimation portera sur les effets aléatoires additifs $W_t = \log(U_t)$. Comme l'estimation initiale porte sur les moments du second ordre des effets aléatoires multiplicatifs, on créera un lien entre les moments des U_t et des W_t en se restreignant aux lois normales pour les effets additifs. Comme on a imposé $E(U_t) = 1$ les effets additifs et multiplicatifs sont liés par les relations suivantes :

$$U_t = \frac{\exp(W_t)}{E[\exp(W_t)]} ; W_t \text{ loi de } N(0, \sigma_W^2) \Rightarrow \sigma_U^2 = \exp(\sigma_W^2) - 1. \quad (4.11)$$

On a le même type de relation pour les covariances, et on peut écrire

$$\sigma_W^2 \rho_W(h) = \log(1 + \sigma_U^2 \rho_u(h)) \quad \forall h \geq 0. \quad (4.12)$$

À partir des autocorrélations estimées jusqu'à l'ordre $T_{\max} - 1$, on peut estimer des processus autorégressifs jusqu'au même ordre. Le choix d'un ordre adéquat s'obtient à partir des coefficients d'autocorrélation partielle. Le coefficient $r_W(h)$, $h \geq 1$ est celui de W_{t-h} dans la régression probabiliste de W_t sur W_{t-1}, \dots, W_{t-h} . Il s'interprète comme le coefficient de corrélation entre les résidus des régressions de W_t et W_{t-h} sur les variables intermédiaires (cf. Gouriéroux et Monfort, 1996). Le rejet par un test de la nullité de $r_W(h)$ revient à considérer qu'une spécification autorégressive d'ordre h sur les effets additifs surpasse significativement une spécification d'ordre $h - 1$. Si γ est la fonction d'autocovariance des effets additifs $\gamma_h = \text{Cov}(W_t, W_{t-h})$, les régressions qui définissent les coefficients d'autocorrélation partielle conduisent aux équivalences suivantes :

$$r_W(h) = 0 \Leftrightarrow g_h(\gamma) = 0, \text{ avec} \quad (4.13)$$

$$\begin{aligned} g_1(\gamma) &= \gamma_1 ; \quad g_2(\gamma) = \gamma_0 \gamma_2 - \gamma_1^2 ; \\ g_3(\gamma) &= \gamma_0^2 \gamma_3 + \gamma_2^2 \gamma_1 + \gamma_1^3 - \gamma_1^2 \gamma_3 - 2\gamma_0 \gamma_1 \gamma_2. \end{aligned} \quad (4.14)$$

Les tests associés reposent sur une estimation de la variance asymptotique des coefficients d'autocovariance des effets additifs (cf. Pinquet, Guillén, Bolancé, 2003).

5. Résultats empiriques

La base de données sur laquelle les modèles précédents ont été estimés est une extraction à 10 % du portefeuille des contrats automobile de la première compagnie d'assurances espagnole. L'intervalle d'observation va de début 1991 à fin 1997. Le portefeuille a connu une croissance importante durant ces sept années (plus de 60%), et le taux d'attrition annuel est de l'ordre de 10%. Le plan d'échantillonnage a consisté à retenir 10 % des contrats en 1991, à les garder ensuite aussi longtemps que possible, et à sélectionner chaque année 10 % des affaires nouvelles. Ainsi les taux d'entrée et de sortie sont du même ordre dans la base de travail que dans le portefeuille entier. On dispose d'un panel non cylindré de plus d'un million de contrats-année, pour lesquels on explique le nombre des sinistres causés en responsabilité par rapport à un tiers à partir de composantes de régression. Ces composantes sont issues des variables suivantes : sexe, zone de circulation, âge du conducteur, ancienneté du permis et de l'assuré dans la compagnie, niveau de couverture et puissance du véhicule. Les caractéristiques du contrat sont enregistrées à chaque anniversaire, ce qui fait que les durées des périodes sont toutes égales à un an. Dans l'idéal, chaque avenant au contrat (modification due à un changement de véhicule ou de niveau de couverture) devrait être enregistré aussitôt.

Reprenons les notations de la section précédente : on note $n_{i,t}$ et $\hat{\lambda}_{i,t}$ le nombre des sinistres et la prime-fréquence associés au contrat i et à la période t , et n_i , $\hat{\lambda}_i$ les mêmes variables sommées sur l'ensemble des périodes d'observation du contrat i . Deux estimateurs convergents de la variance d'un effet aléatoire multiplicatif supposé constant dans le temps sont obtenus par la méthode des moments. Ce sont respectivement

$$\frac{\sum_{i,t}(n_{i,t} - \hat{\lambda}_{i,t})^2 - n_{i,t}}{\sum_{i,t} \hat{\lambda}_{i,t}^2} = 1,269, \text{ et } \frac{\sum_i (n_{i,t} - \hat{\lambda}_i)^2 - n_i}{\sum_i \hat{\lambda}_i^2} = 0,779. \quad (5.1)$$

Le premier estimateur convient aussi pour la variance d'un processus stationnaire associé à des effets aléatoires dynamiques. D'après l'équation (2.9), cette variance exprime la hausse relative de la prime-fréquence suite à un sinistre pour un contrat avec une exposition au risque très faible. Retenir le premier estimateur reviendrait à multiplier par 2,269 la prime d'un conducteur débutant suite à un sinistre ! On est très loin du coefficient de 1,25 du système français. La prise en compte de l'ancienneté des sinistres permettra d'obtenir des résultats plus proches de ceux du monde réel. Sur les autres portefeuilles étudiés par l'auteur, la variance estimée à partir des nombres et primes désagrégés est le plus souvent comprise entre 0,4 et 0,8. La valeur élevée obtenue ici reflète l'importante hétérogénéité des risques en fréquence non expliquée par les composantes de régression. Elle traduit l'hétérogénéité intrinsèque de ces risques, mais aussi une certaine médiocrité des variables de tarification : par exemple, les typologies des unités géographiques et des véhicules sont peu performantes dans cette base de données.

L'estimateur sur données agrégées par individu est sensiblement inférieur au précédent. Comme les deux estimateurs devraient converger vers la même limite si les effets aléatoires étaient constants, on en déduit que cette hypothèse n'est pas plausible. En reprenant les développements de la section précédente, on peut dire que le premier estimateur traduit un effet prédictif de court terme des périodes et des sinistres sur le risque, alors que le second correspond à un horizon plus éloigné. Puisqu'il est nettement inférieur au premier, ce pouvoir prédictif doit décroître avec l'ancienneté des périodes, et on s'attend à trouver des coefficients d'autocorrélation globalement décroissants dans l'estimation d'un processus stationnaire. De fait, l'application des équations (4.5) et (4.6) conduit aux résultats suivants :

TABLEAU 1. – Coefficients d'autocorrélation estimés pour des effets aléatoires multiplicatifs

h (décalage)	1 an	2 ans	3 ans	4 ans	5 ans	6 ans
$\hat{\rho}_U(h)$	0,632	0,485	0,462	0,436	0,360	0,348

ANCIENNETÉ DES PÉRIODES DANS LES MODÈLES DE RISQUE EN FRÉQUENCE

On constate que la suite des coefficients est décroissante. C'est un résultat intéressant dans la mesure où il a été obtenu par une méthode d'estimation non contrainte. Par ailleurs les estimations obtenues sont plausibles, et on peut ajuster une loi log-normale multivariée sur les estimations du tableau. Le fait que $\hat{\rho}_U(1) = 0,632$ soit sensiblement inférieur à $\rho_U(0) = 1$ signifie que le pouvoir prédictif d'un sinistre sur le risque s'estompe fortement après un an.

Détaillons la contribution de chaque année à la crédibilité d'un historique de six ans. On retient un contrat ayant un risque moyen (soit une fréquence annuelle de 0,09). En résolvant le système de crédibilité linéaire donné en (4.10), on obtient :

TABLEAU 2. – Coefficients de crédibilité (en pourcentage) associés à un historique de six ans

$t = 1$	$t = 2$	$t = 3$	$t = 4$	$t = 5$	$t = 6$
2,66	2,68	3,46	3,71	3,94	5,65

La crédibilité totale associée à l'historique est de 22,10 %. Elle correspond par exemple à la réduction de la prime-fréquence *a priori* associée à six années sans sinistre. La crédibilité associée à la dernière année est supérieure à celle des deux premières.

Appliquons les résultats précédents à un calcul de coefficients bonus-malus calculés année après année sur le contrat suite à un sinistre en première période, et à cinq années sans sinistre ensuite. On comparera les résultats à ceux dérivés du modèle actuariel de base où les effets aléatoires sont constants, et utilisant la variance estimée sur les nombres et primes-fréquence sommés sur les historiques entiers.

TABLEAU 3. – Analyse « impulse-response » de coefficients bonus-malus après un sinistre déclaré durant la première année

Longueur de l'historique	1 an	2 ans	3 ans	4 ans	5 ans	6 ans
Effets aléatoires constants	166,2	156	147	139	131,7	125,2
Effets aléatoires dynamiques	165,5	140	131,7	123,8	111,4	107,5

Dans le tableau précédent, les coefficients sont donnés en pourcentage. Les coefficients calculés pour un historique d'un an sont appliqués pour la deuxième année (le coefficient initial est de 100 %). Si le malus suite à sinistre est du même ordre dans les deux approches, les réductions de prime suite à la deuxième année sans sinistres sont très différentes. Dans le modèle classique, le coefficient bonus-malus diminue de 6 % environ, ce qui traduit la hausse de l'exposition au risque sans déclaration de sinistre supplémentaire. Avec les effets aléatoires dynamiques, la baisse est beaucoup plus importante : elle est de 15 %, et elle s'explique par le fait que durant la deuxième année, le sinistre déclaré en première année vieillit, et perd de son pouvoir prédictif. En première approximation, on peut considérer que le vieillissement du sinistre induit une réduction de prime de 9 %, ce qui est supérieur à la baisse due à l'augmentation de l'exposition au risque.

On notera que la crédibilité totale associée à l'historique de six ans est sensiblement supérieure avec l'hypothèse de constance des effets aléatoires :

on obtient 29,6 % au lieu des 22,10 % issus du tableau 2. Les deux systèmes diffèrent donc essentiellement du point de vue des réductions de prime suite aux années sans sinistres. Celles-ci deviennent moins importantes quand on prend en compte l'ancienneté des périodes pour les historiques sans aucun sinistre, mais ont nettement plus d'ampleur quand il y a eu des sinistres avant la dernière période.

On peut également interpréter les résultats précédents du point de vue de la dynamique des coefficients bonus-malus. À la fin de la section 3, on a mentionné la « mémoire d'éléphant » des modèles actuariels classiques. Avec une prime-fréquence annuelle de 9 % (arrondie à un onzième), il faut dix années sans sinistre pour retrouver le coefficient bonus-malus d'avant l'année de survenance du sinistre. Avec les estimations précédentes, on aboutit presque au même résultat au bout de cinq ans quand on prend en compte l'ancienneté des sinistres. On se rapproche alors des pratiques du monde réel.

L'étude de plus long terme des systèmes bonus-malus nécessite le calcul d'un prolongement du corrélogramme estimé. On utilise pour cela les outils rudimentaires de l'analyse des séries chronologiques qui ont été rappelés dans la section précédente. On n'entrera pas dans le détail des calculs (*cf.* Pinquet, Guillén, Bolancé, 2003). Indiquons simplement les principales conclusions obtenues.

- La crédibilité totale d'un historique ne tend pas nécessairement vers un quand la prime-fréquence associée à celui-ci (c'est-à-dire l'exposition au risque) tend vers l'infini, alors que c'est le cas avec des effets aléatoires constants. Par exemple, le bonus associé à un historique sans sinistre ne tend pas vers 100 % dans ce contexte, ce qui est le cas dans le monde réel (*cf.* la section 3). Cela dit, les résultats asymptotiques dépendent fortement de la manière dont on a prolongé la structure de corrélation.
- En conséquence, la dispersion des coefficients bonus-malus obtenue à partir de modèles à effets aléatoires dynamiques est sensiblement inférieure à celle obtenue sous l'hypothèse usuelle. Là encore, la prise en compte de l'ancienneté rapproche les modèles actuariels de ceux du monde réel.
- En reprenant la problématique conduisant au tableau 2 (contribution des différentes périodes d'un historique à la crédibilité totale), mais sur un plus long terme, on constate que les dernières années sont de loin les plus informatives (*cf.* Pinquet, Guillén, Bolancé, 2003, pour des résultats précis). Dans les systèmes « punition-récompense » du monde réel, on observe une pratique d'oubli des événements déclenchant les sanctions dès lors que ceux-ci ont une certaine ancienneté, ce qui s'inscrit dans la même logique. Par exemple, une infraction déclenchant un retrait de points dans le système français de permis à points est amnistiée au bout de trois ans (la durée correspondante est de deux ans au Québec). Dans le système bonus-malus français, un sinistre en responsabilité est effacé au bout de quatre ans pour les conducteurs ayant le bonus maximum, pour autant qu'un autre sinistre ne soit pas déclaré dans l'intervalle. L'équité actuarielle conduit à des résultats relativement proches de ces pratiques quand on prend en compte l'ancienneté. On a vu dans la section 3 que ces pratiques d'amnistie trouvaient également une justification par une analyse de la structure incitative de ces systèmes « punition-récompense ».

Concluons cette section empirique avec une étude de risque qui considère l'horizon de prédiction comme une variable. Du moment qu'on peut calculer

des pouvoirs prédictifs pour des décalages quelconques, on peut augmenter ces décalages dans une logique prospective plutôt que rétrospective, c'est-à-dire considérer plusieurs périodes dans le futur. Les pratiques de tarification de l'industrie de l'assurance fournissent une motivation à des études de ce type. En effet, les primes en assurance non-vie prennent implicitement en compte une relation avec l'assuré qui va au-delà de l'année à venir, alors que les modèles actuariels bornent le futur à la prochaine date de renouvellement du contrat. Un exemple frappant est donné par l'influence de l'ancienneté du véhicule sur la prime. Toutes les études montrent que le risque décroît rapidement avec l'âge, et particulièrement pour les garanties dommages et vol. Or les primes d'assurance ne font jouer à l'ancienneté qu'un rôle assez secondaire, ce qui signifie que les véhicules neufs ou récents sont subventionnés par les véhicules plus anciens. On peut s'expliquer ces pratiques si l'assureur considère qu'un véhicule neuf assuré par lui vieillira dans son portefeuille avec une probabilité suffisante.

On calculera donc des coefficients bonus-malus intégrant une loi de durée sur la relation assureur-assuré. Supposons qu'un assuré ait été observé sur les périodes $1, \dots, T$. Un coefficient bonus-malus actuariel pour la période $T + 1$ est du type $\hat{E}(U_{T+1}|I_T)$, où U_{T+1} est l'effet aléatoire associé à la période $T + 1$, et I_T est l'information résumant l'historique de T périodes, soit les nombres et les primes-fréquence *a priori* dans le contexte du papier. Dans un calcul de crédibilité linéaire, l'espérance n'est pas une espérance conditionnelle, mais est obtenue par régression linéaire. Si $p_{T,h}$ est la probabilité estimée que l'assuré soit encore client à la période $T+h$, un prédicteur intégrant la loyauté anticipée de l'assuré est par exemple

$$\frac{\sum_{h=1}^H p_{T,h} \hat{E}(U_{T+1}|I_T)}{\sum_{h=1}^H p_{T,h}} \quad (5.2)$$

en bornant l'horizon de prédiction à H périodes. Le calcul est simpliste car les probabilités sont considérées ici comme exogènes alors qu'on devrait les faire dépendre entre autres des variations de prime. Si le taux d'attrition est constant et égal à r , on a $p_{T,h} = (1 - r)^h$.

Les taux d'attrition (taux de départ des assurés, que ce soit à leur initiative ou à celle de l'assureur) sont très variables suivant les compagnies. En France, on peut les situer entre 5 % et 20 % en moyenne par compagnie d'assurance. Ces taux diminuent avec l'ancienneté de l'assuré comme client de la compagnie, et sont plus faibles pour les mutuelles que pour les sociétés de capitaux. Ils peuvent atteindre 50 % dans des pays où il n'y a pas de tacite reconduction du contrat, comme au Royaume-Uni.

Le dernier tableau donne deux niveaux de malus calculés suite à un sinistre en première période pour un risque moyen. On a utilisé la formule (5.2) avec deux taux d'attrition $r = 5\%$ et $r = 20\%$, et une structure de corrélation obtenue par estimation d'un processus AR(6) sur des effets aléatoires additifs gaussiens, à partir du corrélogramme estimé dans le tableau 1.

TABLEAU 4. – Malus suite à sinistre, en fonction d'un taux d'attrition supposé constant

	$r = 5\%$	$r = 20\%$
Malus ($T = 1, n_1 = 1, H = 30$)	27,0%	43,6%

Le malus est d'autant plus faible que la fidélité anticipée de l'assuré est importante dans un contexte où le pouvoir prédictif des sinistres décroît avec l'ancienneté. On constate en effet en France que les mutualisations de risque entre les périodes sont d'autant plus fortes que les assurés sont fidèles. C'est flagrant au niveau de la tarification des jeunes conducteurs, où les surprimes imposées par les sociétés d'assurance sont d'autant faibles qu'on anticipe une fidélité importante.

La première estimation de l'équation (5.1) conduisait à un malus de plus de 120% suite à un sinistre en première période. La prise en compte de l'ancienneté des périodes, jointe à un point de vue prospectif sur le risque, conduit à des résultats très différents.

6. Références bibliographiques

- ABBRING J.H., CHIAPPORI P.A., HECKMAN J.J. et PINQUET J. (2003), Adverse selection and moral hazard in insurance : can dynamic data help to distinguish? *Journal of the European Economic Association, papers and proceedings*, **1**, 512-521 (une version plus complète cosignée par Abbring, Chiappori et Pinquet et intitulée « Testing for moral hazard on dynamic insurance data » est parue en 2003 dans *Journal of the European Economic Association* **1**, 767-820).
- AL-OSH M.A. et ALZAID A.A. (1987), First order integer valued autoregressive (INAR(1)) process. *Journal of Time Series Analysis*, **8**, 261-275.
- BÜHLMANN H. (1967), Experience rating and credibility. *Astin Bulletin*, **4**, 199-207.
- DIONNE G. et VANASSE C. (1989), A generalization of automobile insurance rating models : the negative binomial distribution with a regression component. *Astin Bulletin*, **19**, 199-212.
- HENRIET D. et ROCHET J.C. (1986), La logique des systèmes bonus-malus en assurance automobile. *Annales d'économie et de statistiques*, 133-152.
- GERBER H. et JONES D. (1975), Credibility formulas of the updating type. *Transactions of the society of actuaries*, **27**, 31-52.
- GOURIÉROUX C. et JASIAK J. (2001), *Non linear data models with dynamic heterogeneity*, Panel data econometrics : Future directions (éditeurs : J. Krishnakumar et E. Ronchetti), 127-148.
- GOURIÉROUX C. et JASIAK J. (2002), *Heterogenous INAR(1) model with application to car insurance*, mimeo.
- GOURIÉROUX C. et MONFORT A. (1996), *Time series and econometric models*. Cambridge University Press.
- GOURIÉROUX C., MONFORT A. et TROGNON A. (1984a), Pseudo likelihood methods : Theory. *Econometrica*, **52**, 681-700.
- GOURIÉROUX C., MONFORT A. et TROGNON A. (1984b), Pseudo likelihood methods : Applications to Poisson models. *Econometrica*, **52**, 701-720.
- LEMAIRE J. (1995), *Bonus-malus systems in automobile insurance*. Kluwer Academic Publishers.

ANCIENNETÉ DES PÉRIODES DANS LES MODÈLES DE RISQUE EN FRÉQUENCE

- LIANG K.Y. et ZEGER S.L. (1986), Longitudinal data analysis using generalized linear models. *Biometrika*, **73**, 13-22.
- Mc KENZIE E. (1988), Some ARMA models for dependent sequences of Poisson counts. *Advances in Applied Probability*, **20**, 822-835.
- PINQUET J. (2000), *Experience rating through heterogeneous models*. Handbook of insurance (éditeur : Georges Dionne), chapitre 14, 459-500. Kluwer Academic Publishers.
- PINQUET J., GUILLÉN M. et BOLANCÉ C. (2001), Allowance for the age of claims in bonus-malus systems. *Astin Bulletin*, **31**, 337-348.
- PINQUET J., GUILLÉN M. et BOLANCÉ C. (2003), Time-varying credibility for frequency risk models. Estimation and tests for autoregressive specifications for the random effects. *Insurance : Mathematics and Economics*, **33**, 273-282.
- SUNDT B. (1988), Credibility estimators with geometric weights. *Insurance : mathematics and economics*, **7**, 113-122.
- ZEGER L.S. (1988), A regression model for time series of counts. *Biometrika*, **74**, 721-729.