Available online at www.sciencedirect.com
SCIENCE DIRECT®
ELSEVIER
C. R. Acad. Sci. Paris, Ser. I 339 (2004) 283–286
COMPTES RENDUS MATHEMATIQUE

Probabilités/Statistique

Crédibilité paramétrique, corrélation et tarification bonus-malus

Daniel Pierre-Loti-Viaud

Université Pierre et Marie Curie, LSTA-ISUP, boîte 157, 4, place Jussieu, 75252 Paris cedex 05, France

Reçu le 10 novembre 2003 ; accepté après révision le 7 juin 2004

Disponible sur Internet le 28 juillet 2004

Présenté par Paul Deheuvels

Résumé

La tarification actuarielle bonus-malus du modèle de crédibilité paramétrique de mélange de lois de Poisson est-elle trop différenciée ? Pour répondre à cette question, nous élargissons le modèle de mélange de lois de Poisson en considérant le mélange de lois binomiales négatives. Le coefficient de corrélation entre les nombres annuels de sinistres pour un individu peut alors être ajusté, et nous calculons les différences de tarification engendrées par ce nouveau modèle de crédibilité paramétrique. Dans une application à un portefeuille automobile, l'ajustement de la corrélation peut conduire à des différences tarifaires réduites par rapport à celles du modèle classique. ***Pour citer cet article : D. Pierre-Loti-Viaud, C. R. Acad. Sci. Paris, Ser. I 339 (2004).***

Abstract

Parametric credibility, correlation and bonus-malus rating. Is an actuarial rating of bonus-malus type based on the parametric credibility model of mixed Poisson distributions too differentiated? To answer this question, we enlarge the model of mixed Poisson distributions by considering mixed negative binomial distributions. The correlation coefficient between the annual claim numbers for an individual can then be adjusted, and we calculate the differences between the ratings coming from this new model of parametric credibility. An application of these results to a car portfolio shows that adjustment of the correlation may yield significantly smaller rating differences than those coming from the classical model. ***To cite this article: D. Pierre-Loti-Viaud, C. R. Acad. Sci. Paris, Ser. I 339 (2004).***

1. Introduction

Nous nous intéressons aux modèles de crédibilité paramétrique sur le nombre de sinistres (voir, par exemple, [4,5] et [7]). La question que nous traitons est la suivante : le modèle de mélange de lois de Poisson utilisé pour in-

Adresse e-mail : pilovi@ccr.jussieu.fr (D. Pierre-Loti-Viaud).

1631-073X/$ – see front matter © 2004 Académie des sciences. Publié par Elsevier SAS. Tous droits réservés.
doi:10.1016/j.crma.2004.06.022

troduire une dépendance entre les nombres successifs de sinistres d'une police d'assurance ne permet pas d'ajuster la corrélation observée entre ces nombres de sinistres, et c'est le seul modèle de crédibilité paramétrique développé dans la littérature actuarielle. Nous proposons ici de considérer le modèle de mélange de lois binomiales négatives pour ajuster la corrélation observée entre les nombres successifs de sinistres, et nous donnons une évaluation de son influence dans le calcul de la tarification a posteriori en considérant le cas de valeurs des paramètres usuelles en assurance automobile.

Soit K fixé dans $\{1, 2, \ldots\}$ et un portefeuille constitué de K polices d'assurance. Pour l'exercice t, $t \in \mathbb{N}$, les observations $N_1(t), \ldots, N_K(t)$ donnent le nombre de sinistres de chacune des polices d'assurance de ce portefeuille. On suppose qu'à la $k^{\text{ième}}$ police est associée une variable aléatoire θ_k, prenant ses valeurs dans Θ sous-ensemble borélien de $\mathbb{R}$, et qui mesure la qualité du risque de cette police. On suppose aussi que les suites $(\theta_k, N_k(0), N_k(1), \ldots)$, $k \in \{1, \ldots, K\}$, sont des réalisations indépendantes et équidistribuées d'une même loi de probabilité, définie sur $\Theta \times \mathbb{N}^{\mathbb{N}}$, et vérifiant, pour tout $k \in \{1, \ldots, K\}$, tout borélien $B \subset \Theta$, tout $\ell \in \mathbb{N}$, et tout $(i_0, \ldots, i_\ell) \in \mathbb{N}^{\ell+1}$:

$$\mathbb{P}\big(\{\theta_k \in B\} \cap \{N_k(0) = i_0\} \cap \cdots \cap \{N_k(\ell) = i_\ell\}\big) = \int_B P_\theta(i_0) \cdots P_\theta(i_\ell)\, \mathrm{d}Q(\theta), \tag{1}$$

où Q est une loi de probabilité sur Θ et $(P_\theta, \theta \in \Theta)$ est une famille de lois de probabilité sur $\mathbb{N}$.

Le modèle (1) le plus étudié dans la littérature actuarielle est celui où $(P_\theta, \theta \in \Theta)$ est la famille des lois de Poisson (voir, par exemple, [1,4,5,7,10]). Deux autres modèles de crédibilité sont ceux où $(P_\theta, \theta \in \Theta)$ est la famille des lois binomiales ou bien celle des lois binomiales négatives (voir [4]). Ces deux derniers modèles sont cependant très peu étudiés, et une comparaison entre ces trois modèles de crédibilité paramétrique inexistante ; ces points sont l'objet principal du travail figurant dans [9].

Nous présentons dans cette Note des points issus de [9] et comparons sur un exemple les deux modèles de crédibilité correspondant aux lois de Poisson et aux lois binomiales négatives. La Section 2 utilise les résultats de [8] pour construire un modèle incluant ces deux types de lois et conduisant à une expression simple des premiers moments du nombre de sinistres et de la corrélation entre les nombres de sinistres. Dans la Section 3, nous présentons une méthode de calcul approché de l'indice de fréquence $I(t, \cdot)$, $t \in \{1, 2, \ldots\}$, qui est défini sur $\mathbb{N}^t$ et vérifie, pour $(n_0, \ldots, n_{t-1}) \in \mathbb{N}^t$:

$$I(t, n_0, \ldots, n_{t-1}) = \mathbb{E}(N_t / N_0 = n_0, \ldots, N_{t-1} = n_{t-1}) / \mathbb{E}N_t, \tag{2}$$

cet indice représentant, dans le cas du modèle collectif d'assurance, le rapport entre la prime pure du tarif basé sur l'expérience de sinistralité, tarif dit bonus-malus, et la prime pure classique [4]. Dans la dernière section, nous appliquons nos résultats au seul portefeuille automobile vu pendant deux ans de la littérature [1]. Nous montrons que la corrélation influe sur la différenciation tarifaire, et que celle-ci peut être significativement plus importante dans le modèle de lois de Poisson.

Indiquons qu'un test statistique basé sur la méthode du bootstrap de l'hypothèse « mélange de lois de Poisson » contre l'hypothèse « mélange de lois binomiales négatives » est proposé dans [2]. Et que, bien qu'ayant travaillé ici dans le cadre d'une tarification a posteriori, le cas de l'existence de variables tarifaires (tarification a priori) est considéré dans [3] sur un portefeuille automobile vu pendant deux ans où la corrélation observée est significativement plus petite que celle prédite par le modèle de lois de Poisson.

2. Mélanges de lois de Poisson et de lois binomiales négatives

En suivant [8], pour θ dans $\Theta = [0, \infty[$ et s dans $S = [0, \infty[$, soit $P_{\theta,s}$ la distribution de probabilité sur $\mathbb{N}$ qui vérifie, pour $i \in \mathbb{N}$:

$$P_{\theta,s}(i) = \frac{1}{i!}(1+s)(1+2s)\cdots\big(1+(i-1)s\big)\theta^i(1+s\theta)^{-(i+1/s)}, \tag{3}$$

où nous utilisons les conventions habituelles. Observons que, pour $s=0$ et $\theta \in \Theta$, on a $P_{\theta,0} = \mathcal{P}(\theta)$, loi de Poisson de paramètre θ, tandis que, pour $s \in]0,\infty[$ et $\theta \in \Theta$, on a $P_{\theta,s} = \mathcal{NB}(1/s, s\theta)$, loi binomiale négative de paramètres $1/s$ et $s\theta$ [6]. Notons $\mathcal{Q}$ l'ensemble des lois de probabilité définies sur Θ, et, pour $s \in S$ et $Q \in \mathcal{Q}$, $\mathcal{M}(s,Q)$ la loi mélange des lois $(P_{\theta,s}, \theta \in \Theta)$ par la loi Q. La loi $\mathcal{M}(s,Q)$ est la loi du nombre de sinistres dans le modèle de crédibilité (1) lorsque $P_\theta = P_{\theta,s}$, pour $\theta \in \Theta$. Dans ce qui suit, ce modèle général de crédibilité paramétrique est désigné par le modèle $(\mathcal{C})$. Pour P une loi de probabilité notons alors $M(P)$, $\sigma^2(P)$ et $\gamma_1(P)$, la moyenne, la variance et le coefficient d'asymétrie de P. Considérant $s \in S$ et $Q \in \mathcal{Q}$, notons aussi $\varrho = \varrho(s)$ le coefficient de corrélation commun aux couples de variables $(N_k(t), N_k(u))$, $t \neq u \in \mathbb{N}$, $k \in \{1,\dots,K\}$, dans $(\mathcal{C})$. Finalement notons $m = M(\mathcal{M}(s,Q))$, $\sigma^2 = \sigma^2(\mathcal{M}(s,Q))$ et $\gamma_1 = \gamma_1(\mathcal{M}(s,Q))$. La proposition suivante présente le calcul des premiers moments et de la corrélation dans le modèle $(\mathcal{C})$.

Proposition 2.1. *Soit $Q \in \mathcal{Q}$ et $s \in S$. Nous avons* :

$$m = M(Q), \quad \text{si } M(Q) < \infty, \quad \text{et} \quad \sigma^2 = (1+s)\sigma^2(Q) + m(1+sm), \quad \text{si } \sigma^2(Q) < \infty, \tag{4}$$

et, si en plus Q n'est pas la loi de Dirac en 0, *nous avons $\sigma^2 > 0$ avec* :

$$\varrho = \frac{\sigma^2(Q)}{m + \sigma^2(Q) + s(m^2 + \sigma^2(Q))} = \frac{1}{1+s}\left(1 - \frac{m(1+sm)}{\sigma^2}\right), \quad \text{si } \sigma^2(Q) < \infty, \tag{5}$$

$$\gamma_1 = \frac{1+2s}{\sqrt{1+s}}\left(1 - \frac{m(1+sm)}{\sigma^2}\right)^{3/2} \gamma_1(Q) + \frac{1+2sm}{\sigma}\left(3 - 2\frac{m(1+sm)}{\sigma^2}\right), \quad \text{si } \gamma_1(Q) < \infty. \tag{6}$$

Pour les mélanges de lois de Poisson soulignons que la corrélation vérifie toujours $\varrho = 1 - m/\sigma^2$. De plus, tant à loi Q fixée qu'à moyenne et variance fixées pour le nombre de sinistres, le mélange de lois binomiales négatives engendre une corrélation plus petite que celle du mélange de lois de Poisson, et le modèle $(\mathcal{C})$ permet d'ajuster la corrélation entre les observations de différents exercices.

3. Indice de fréquence

Les résultats que nous présentons maintenant sont vrais sous des hypothèses plus générales (voir [9]). Indiquons aussi que l'identité (7) est bien connue des actuaires.

Théorème 3.1. *Supposons $Q = \gamma(a,b)$, loi gamma de paramètres a et $b > O$. Pour $t \in \{1,2,\dots\}$ et $(n_0,\dots,n_{t-1}) \in \mathbb{N}^t$, posons $n = n_0 + \cdots + n_{t-1}$. Alors, l'indice de fréquence ne dépend que de n, i.e. $I(t,n_0,\dots,n_{t-1}) = I(t,n,s)$. De plus, pour le mélange de lois de Poisson, nous avons* :

$$I(t,n,0) = (1+n/a)/(1+t/b), \tag{7}$$

et, pour le mélange de lois binomiales négatives, il existe $s_0 = s_0(t,n) > 0$ vérifiant, pour $s \in [0,s_0]$:

$$\begin{aligned} I(t,n,s) = I(t,n,0)\biggl(1 &+ \frac{t(a+n+1) - n(b+t) - nt/2}{(b+t)^2}s \\ &+ \frac{t(a+n+1)[t(2(a+n)+3) - (b+t)((a+n) - 3n + 2)] + n(a+n+2)(b+t)^2}{(b+t)^4}s^2\biggr) \\ &+ J(t,n,s)s^3, \quad avec \sup_{s\in[0,s_0]} \bigl|J(t,n,s)\bigr| < \infty. \end{aligned} \tag{8}$$

Indiquons que conformément à la corrélation qui décroît quand s augmente, on voit sur (8) qu'au moins pour de petites valeurs de n et t les variations de tarif induites par le modèle de mélange de lois binomiales négatives sont plus faibles que celles dues au modèle de mélange de lois de Poisson. Cette propriété est vraie à loi Q fixée, mais aussi si on ajuste les paramètres de la loi Q pour que les deux premiers moments des deux modèles soient égaux. Cette propriété est aussi vraie pour des lois Q plus générales (voir [9]).

4. Importance pour le tarif a posteriori du choix du modèle de mélange

Considérons le portefeuille étudié par Besson et Partrat [1]. La différence entre la corrélation empirique et celle prédite par un mélange de lois de Poisson y est de l'ordre de -7% (en utilisant (5) pour les données de $t=0$), et elle dépasse largement les bornes de l'intervalle de confiance construit par la méthode du Z de Fisher autour de la corrélation empirique. Les résultats présentés maintenant impliquent que l'évolution entre les différences tarifaires induites par les mélanges de lois de Poisson et de lois binomiales négatives est du même ordre.

Proposition 4.1. *Fixons $0<m<\sigma^2$ et, pour $s\in S$, notons $a(s)$ et $b(s)>0$ les solutions de* (4) *avec $Q=\gamma(a(s),b(s))$. En considérant le modèle $\mathcal{M}(s,\gamma(a(s),b(s)))$, $s\in S$, nous avons alors pour $t\in\{1,2,\ldots\}$ et $n\in\mathbb{N}$, en notant $\alpha=a(0)$ et $\beta=b(0)$:*

$$\frac{I(t,n,s)-1}{I(t,n,0)-1}=1+(1+\alpha)\left(-\frac{\beta}{1+\beta}+\frac{(\alpha+n)\beta}{(1+\alpha)(\beta+t)^2}+\frac{\beta t(1-n/2)}{(1+\alpha)(n\beta-at)(\beta+t)^2}\right)s+\mathrm{o}(s)\;; \tag{9}$$

$$\frac{\varrho(s)-\varrho(0)}{\varrho(0)}=-(1+\alpha)\frac{s}{1+s}. \tag{10}$$

Alors, avec les données de [1] on observe $\alpha=1{,}67$, $\beta=9{,}39$, et le coefficient en facteur de $(1+\alpha)s$ dans (9) qui est de l'ordre de $-0{,}8$, au moins pour de petites valeurs de n et t (avec $(1+\alpha)s$ qui est de l'ordre de $0{,}07$). La différence de corrélation entre les deux modèles de mélange peut donc être réellement significative dans l'établissement d'une tarification a posteriori.

Remerciement

Nous remercions le rapporteur qui a permis d'améliorer sensiblement le traitement de l'exemple.

Références

[1] J.L. Besson, C. Partrat, Trend et systèmes de bonus-malus, ASTIN Bulletin 22 (1992) 11–31.
[2] E. Cohen, D. Pierre-Loti-Viaud, A test for assuming mixtures of Poisson distributions in a general parametric credibility system, prépublication LSTA, 2003, soumise pour publication.
[3] E. Cohen, D. Pierre-Loti-Viaud, Ajustement d'un modèle de crédibilité paramétrique sur un portefeuille automobile vu pendant deux ans, prépublication LSTA, 2003.
[4] F.E. De Vylder, Advanced Risk Theory, Éditions de l'Université de Bruxelles, Bruxelles, 1996.
[5] M. Goovaerts, R. Kaas, A. van Heerwaerden, T. Bauwelinckx, Effective Actuarial Methods, North-Holland, Amsterdam, 1990.
[6] N.L. Johnson, S. Kotz, A.W. Kemp, Univariate Discrete Distributions, Wiley, New York, 1992.
[7] J. Lemaire, Bonus-Malus Systems in Automobile Insurance, Kluwer Academic, Netherlands, 1995.
[8] D. Pierre-Loti-Viaud, Sur les ressemblances entre les lois de Poisson, binomiales et binomiales négatives mises en évidence par un modèle paramétré commun, Rev. Roumaine Math. Pures Appl. 48 (2) (2003) 205–210.
[9] D. Pierre-Loti-Viaud, Mélanges de lois sur ℕ et crédibilité paramétrique, prépublication LSTA, 2003, soumise pour publication.
[10] J.F. Walhin, J. Paris, Using mixed Poisson processes in connexion with bonus-malus systems, ASTIN Bulletin 29 (1999) 81–99.