ARTICLES

TREND ET SYSTÈMES DE BONUS-MALUS [1]

PAR JEAN-LUC BESSON

Assemblée Plénière des Sociétés d'Assurances Dommages, Paris

ET CHRISTIAN PARTRAT

Institut de Statistique, URA 1321, Université Pierre et Marie Curie, Paris

ABSTRACT

This paper deals with the Bonus-Malus system obtained when the claims frequency is submitted to trend. This system is specified in the two particular cases of Poisson-Gamma and Poisson-Inverse Gaussian distributions. The theoretical results are checked on data issued from automobile insurance policies observed during three years.

MOTS CLÉS

Loi de Poisson-mélange; Binomiale Négative; loi de Poisson-Inverse Gaussienne; systèmes de Bonus-Malus.

1. INTRODUCTION

L'étude de l'évolution dans le temps des systèmes de Bonus-Malus en assurance automobile nécessite la connaissance de la loi de probabilité de la fréquence annuelle des sinistres.

A cet effet la famille des lois de Poisson-mélange s'est avérée un outil efficace (KESTEMONT et PARIS, 1985).

Par leur bonne adéquation aux observations et les possibilités de calcul qu'elles permettent, la loi Binomiale Négative (LEMAIRE, 1985) et la loi de Poisson-Inverse Gaussienne (WILLMOT, 1986 et 1987) ont été les lois de cette famille les plus utilisées.

L'objet de cette étude est de présenter et comparer les systèmes de Bonus-Malus auxquels ces lois conduisent dans une modélisation tenant compte d'une variation tendancielle de la fréquence moyenne annuelle de sinistres du portefeuille. Plus précisément sont déterminés l'indice de fréquence et la loi de probabilité de la fréquence annuelle de l'exercice $(t+1)$ connaissant l'historique de sinistralité des t exercices précédents.

Ces résultats sont appliqués à un portefeuille de plus d'un million de polices d'assurance automobile en France dont la sinistralité a pu être observée pendant trois exercices consécutifs.

[1] Une première version de cet article a été présentée sous le titre «Loi de Poisson-Inverse Gaussienne et systèmes de Bonus-Malus» au XXII^{e} Astin Colloquium, Montreux 1990.

2. Loi Inverse Gaussienne Généralisée

La fonction de Bessel de 3^{e} espèce modifiée K_ν d'indice ν réel, définie sur l'ensemble $\mathbb{R}_+^*$ des réels > 0, a pour expression (Erdelyi-Magnus, 1953)

$$K_\nu(u) = \frac{1}{2}\int_0^{+\infty} x^{\nu-1} e^{-\frac{u}{2}\left(x+\frac{1}{x}\right)}\, dx.$$

Elle vérifie pour tout $u > 0$ les propriétés suivantes:

(i) $K_{-\nu}(u) = K_\nu(u)$ pour $\nu \in \mathbb{R}$

(ii) $K_{\nu+1}(u) = \dfrac{2\nu}{u} K_\nu(u) + K_{\nu-1}(u)$ pour $\nu \in \mathbb{R}$

(iii) $K_{1/2}(u) = \sqrt{\dfrac{\pi}{2u}}\, e^{-u}$.

On obtient (i) par le changement de variables $y = 1/x$ et (ii) par une intégration par parties dans $K_\nu(u)$. Le changement de variables $y = x^{1/2} - x^{-1/2}$ dans $K_{1/2}(u) + K_{-1/2}(u)$ donne (iii).

De (ii) et (iii) on déduit une expression de $K_{p+1/2}(u)$ pour $p \in \mathbb{N}$ (Willmot, 1987):

(iv) $K_{p+1/2}(u) = \sqrt{\dfrac{\pi}{2u}}\, e^{-u} \displaystyle\sum_{i=0}^{p} \frac{(p+i)!}{(p-i)!\, i!} (2u)^{-i}$.

La loi Inverse Gaussienne Généralisée de paramètres ν, μ, β ($\nu \in \mathbb{R}$, $\mu > 0$, $\beta > 0$) a pour densité sur $\mathbb{R}_+^*$

$$h_\nu(x) = \frac{x^{\nu-1} e^{-\frac{1}{2\beta}\left(x+\frac{\mu^2}{x}\right)}}{2\mu^\nu K_\nu(\mu/\beta)}.$$

Cette loi, notée dans la suite IGG (ν, μ, β) et étudiée dans (Jorgensen, 1982), a pour transformée de Laplace

$$\bar{h}_\nu(s) = \int_0^{+\infty} e^{-sx} h_\nu(x)\, dx = \frac{K_\nu\left[\frac{\mu}{\beta}(1+2\beta s)^{1/2}\right]}{K_\nu\left[\frac{\mu}{\beta}\right](1+2\beta s)^{\nu/2}} \quad (s > 0)$$

et pour moyenne $\dfrac{\mu K_{\nu+1}(\mu/\beta)}{K_\nu(\mu/\beta)}$.

On peut noter que si X suit une loi IGG (ν, μ, β) la loi de aX ($a > 0$) est une loi IGG (ν, $a\mu$, $a\beta$).

Le cas particulier $\nu = -1/2$ conduit à la loi Inverse Gaussienne IG (μ, β) (FOLKS-CHIKARRA, 1978) de densité $h_{-1/2}(x) = \dfrac{\mu}{\sqrt{2\pi\beta}\, x^{3/2}}\, e^{-\frac{1}{2\beta x}(x-\mu)^2}$ et transformée de Laplace $\bar{h}_{-1/2}(s) = e^{\frac{\mu}{\beta}[1-(1+2\beta s)^{1/2}]}$.

La moyenne et la variance de cette loi valent respectivement μ et $\mu\beta$.

3. LOIS DE POISSON-MÉLANGE

Cette famille de lois décrit la loi de la v.a.r. N fréquence de sinistres d'une police appartenant à une classe de risques; N est supposé distribué selon une loi de Poisson $\mathscr{P}(\lambda)$, λ étant la réalisation d'une v.a.r. $\wedge > 0$ dite variable de structure de la classe.

Si $\wedge$ a pour densité h_θ, dépendant d'un paramètre $\theta \in \Theta$ ouvert de $\mathbb{R}^q$, la loi de N est donnée par

$$p_n = P(N = n) = \int_0^{+\infty} e^{-\lambda} \frac{\lambda^n}{n!} h_\theta(\lambda)\, d\lambda \qquad (n \in \mathbb{N}).$$

La fonction génératrice de N s'exprime en fonction de la transformée de Laplace $\bar{h}_\theta$ de h_θ:

$$g_N(s) = E(s^N) = \sum_{n=0}^{+\infty} p_n s^n = \bar{h}_\theta(1-s) \text{ avec } g_N^{(n)}(0)/n! = p_n \text{ pour tout } n.$$

Quand ils existent, les moments factoriels de N valent

$$\mu_{[j]}(N) = E[N(N-1)\dots(N-j+1)] = E(\wedge^j) \qquad (j \in \mathbb{N}^* = \mathbb{N} - \{0\}).$$

D'où l'on déduit $m = E(N) = E(\wedge)$, $\sigma^2 = V(N) = V(\wedge) + E(\wedge) \geqslant E(N)$ et le coefficient d'asymétrie $\gamma_1(N)$ de Fisher

$$\gamma_1(N) = \frac{1}{\sigma^3}\left[3\sigma^2 - 2m + C\frac{(\sigma^2-m)^2}{m}\right]$$

où C est un coefficient, fonction de la loi de $\wedge$, qui, pour m et σ^2 fixés, est un indicateur du poids relatif de la queue de N (PANJER et WILLMOT, 1988).

Pour estimer θ on disposera de la réalisation d'un K-échantillon de N (par exemple l'observation sur un exercice de la fréquence de sinistres de K polices formant une classe de risques indépendants). Si l est la plus grande valeur observée de N et ν_j le nombre d'observations de N égales à j ($j = 0, \dots, l$) avec $\sum_{j=0}^{l} \nu_j = K$, la moyenne et la variance empiriques de N sont respectivement

$$\bar{n} = \frac{1}{K}\sum_{j=0}^{l} j\,\nu_j \qquad \hat{\sigma}_n^2 = \frac{1}{K}\sum_{j=0}^{l} \nu_j\,(j-\bar{n})^2.$$

a) Loi de Poisson-Gamma

Ce cas particulier détaillé dans (LEMAIRE, 1985) correspond à une variable de structure $\wedge$ distribuée selon une loi Gamma $\gamma(r, \alpha)$ de paramètre $\theta = (r, \alpha)$ $r, \alpha > 0$:

$$h_\theta(\lambda) = \frac{\alpha^r}{\Gamma(r)} \lambda^{r-1} e^{-\alpha\lambda} \quad \text{avec} \quad E(\wedge) = \frac{r}{\alpha}, \quad V(\wedge) = \frac{r}{\alpha^2}.$$

La loi de N est alors une loi Binomiale Négative BN $\left(r, \frac{\alpha}{\alpha+1}\right)$ telle que

$$p_n = \frac{\Gamma(r+n)}{\Gamma(r)\, n!} \left(\frac{\alpha}{\alpha+1}\right)^r \left(\frac{1}{\alpha+1}\right)^n \qquad (n \in \mathbb{N}),$$

$$m = \frac{r}{\alpha}, \ \sigma^2 = \frac{r(\alpha+1)}{\alpha^2} \quad \text{et} \quad C = 2.$$

Pour un K-échantillon et dans l'hypothèse où $\hat{\sigma}_n^2 > \bar{n}$ l'estimation $(\tilde{r}, \tilde{\alpha})$ de (r, α) donnée par la méthode des moments (e.m.m.) est

$$\tilde{r} = \frac{\bar{n}^2}{\hat{\sigma}_n^2 - \bar{n}} \qquad \tilde{\alpha} = \frac{\bar{n}}{\hat{\sigma}_n^2 - \bar{n}}.$$

Si on pose $f_j = \frac{v_j}{K}$ $(j = 0, \ldots, l)$ et si $\hat{r}$ estl'unique solution > 0 (SIMONSEN, 1976 et 1979; BRONS et TOLVER JENSEN, 1989) de l'équation en r:

$$\text{(E1)} \quad \sum_{j=1}^{l} f_j \left(\frac{1}{r+j-1} + \ldots + \frac{1}{r}\right) - \operatorname{Log}\left(1 + \frac{\bar{n}}{r}\right) = 0$$

l'estimation par le maximum de vraisemblance (e.m.v.) de (r, α) est $(\hat{r}, \hat{\alpha})$ avec $\hat{\alpha} = \frac{\hat{r}}{\bar{n}}$.

On peut noter que $\tilde{r}$ peut être prise comme valeur initiale de tout algorithme de résolution de l'équation ci-dessus.

b) Loi de Poisson-Inverse Gaussienne

Reprenant les résultats donnés dans (WILLMOT, 1987), si la variable de structure $\wedge$ suit la loi IG (μ, β), la v.a.r. N a pour fonction génératrice

$$g_N(s) = \exp\left\{\frac{\mu}{\beta}\left[1 - [1 + 2\beta(1-s)]^{1/2}\right]\right\}$$

et moments $m = \mu$, $\sigma^2 = \mu(1+\beta)$. Le coefficient C valant 3, cette loi mélange a une asymétrie plus marquée que la loi Binomiale Négative.

Par dérivation de la fonction génératrice, on obtient $p_0 = e^{\frac{\mu}{\beta}[1-(1+2\beta)^{1/2}]}$, $p_1 = \mu\, p_0 (1+2\beta)^{-1/2}$ et la relation de récurrence permettant d'obtenir de proche en proche toutes les probabilités p_n:

$$\text{(R1)} \quad (1+2\beta)\, n(n-1)\, p_n = \beta(n-1)(2n-3)\, p_{n-1} + \mu^2 p_{n-2} \qquad (n \geqslant 2).$$

Pour estimer $\theta = (\mu, \beta)$ à partir de la réalisation d'un K-échantillon de N et dans l'hypothèse où $\hat{\sigma}_n^2 > \bar{n}$, l'e.m.m. $(\tilde{\mu}, \tilde{\beta})$ est donnée par

$$\tilde{\mu} = \bar{n} \qquad \tilde{\beta} = \frac{\hat{\sigma}_n^2}{\bar{n}} - 1\,.$$

Si on note $p_n(\mu, \beta)$, la n^e probabilité d'une loi de Poisson-Inverse Gaussienne de paramètres μ et β et si on pose

$$T_j(\beta) = (j+1)\frac{p_{j+1}(\bar{n}, \beta)}{p_j(\bar{n}, \beta)} \qquad (j \in \mathbb{N})$$

on a $T_0(\beta) = \bar{n}(1+2\beta)^{-1/2}$ et, par (R1), la relation de récurrence:

$$T_j(\beta) = \frac{1}{(1+2\beta)}\left[(2j-1)\beta + \frac{\bar{n}^2}{T_{j-1}(\beta)}\right] \qquad (j \geqslant 1).$$

L'e.m.v. de (μ, β) est alors $(\hat{\mu}, \hat{\beta})$ avec $\hat{\mu} = \bar{n}$ et $\hat{\beta}$ solution > 0 de l'équation

$$\text{(E2)} \quad \sum_{j=0}^{l} f_j\, T_j(\beta) - \bar{n} = 0\,.$$

$\tilde{\beta}$ peut être prise comme valeur initiale de tout algorithme de résolution de l'équation (E2).

c) Loi de Poisson-Inverse Gaussienne Généralisée

Appelée aussi loi de Sichel (SICHEL, 1971), cette loi est obtenue quand $\wedge$ est distribuée selon une loi IGG (ν, μ, β).

La fonction génératrice et les moments factoriels de N sont alors (WILLMOT, 1986):

$$g_N(s) = \frac{K_\nu\left\{\dfrac{\mu}{\beta}[1-2\beta(s-1)]^{1/2}\right\}}{K_\nu(\mu/\beta)}\; [1-2\beta(s-1)]^{-\nu/2}$$

$$\mu_{[j]}(N) = \mu^j \frac{K_{\nu+j}(\mu/\beta)}{K_\nu(\mu/\beta)} \qquad (j \in \mathbb{N}^*).$$

Les probabilités $p_n = \dfrac{\mu^n}{n!} \dfrac{K_{\nu+n}\left[\dfrac{\mu}{\beta}(1+2\beta)^{1/2}\right]}{K_\nu(\mu/\beta)\,(1+2\beta)^{\frac{\nu+n}{2}}}$ $(n \in \mathbb{N})$ peuvent se déduire de

$$p_0 = \frac{K_\nu\left[\dfrac{\mu}{\beta}(1+2\beta)^{1/2}\right]}{K_\nu(\mu/\beta)\,(1+2\beta)^{\nu/2}} \quad \text{et} \quad p_1 = \mu \frac{K_{\nu+1}\left[\dfrac{\mu}{\beta}(1+2\beta)^{1/2}\right]}{K_\nu(\mu/\beta)\,(1+2\beta)^{\frac{\nu+1}{2}}}$$

par la relation de récurrence

(R2) $(1+2\beta)\,n(n-1)p_n = 2\beta(n-1)(\nu+n-1)p_{n-1} + \mu^2 p_{n-2} \qquad (n \geqslant 2).$

On peut noter que si $\nu = p+1/2$ avec $p \in \mathbb{N}$, les propriétés (ii) et (iii) de 2. conduisent à des expressions très simples de p_0 et p_1.

4. Systèmes de Bonus-Malus

Soit un risque dont la fréquence annuelle de sinistres est distribuée selon une loi de Poisson-mélange. Si pour tout i entier on désigne par N_i cette fréquence pour la i^e année, on dispose de la réalisation $(n_1, \ldots, n_t)$ du v.a. $(N_1, \ldots, N_t)$ correspondant à l'observation de ce risque pendant t années.

Notant $\wedge$ la variable de structure de la classe à laquelle appartient le risque considéré, on se propose de déterminer la loi prédictive de N_{t+1} connaissant $(N_1, \ldots, N_t)$ sous les hypothèses

(H1) les v.a.r. $N_1, \ldots, N_t, N_{t+1}$ sont conditionnellement en $\wedge$ indépendantes.

Le paramètre $\nu > 0$ représentant le taux de variation de la fréquence moyenne de sinistres d'un risque de la classe d'une année à l'autre

(H2) pour tout i la loi conditionnelle de N_i sachant $\wedge = \lambda$ est une loi de Poisson $\mathscr{P}(\lambda\nu^{i-1})$.

Remarque: Pour le modèle adopté le rapport

$$\frac{E(N_i)}{\sqrt{V(N_i)-E(N_i)}} \left(= \frac{E(\wedge)}{\sigma(\wedge)} \right) \text{ est constant } (i = 1, \ldots, t).$$

La loi conditionnelle de $(N_1, \ldots, N_t)$ sachant $\wedge = \lambda$ est alors, pour $(n_1, \ldots, n_t) \in \mathbb{N}^t$,

$$P^v(N_1 = n_1, \ldots, N_t = n_t / \wedge = \lambda) = \prod_{i=1}^{t} P(N_i = n_i / \wedge = \lambda)$$

$$= \frac{e^{-\lambda\, a_t(v)}\, v^{\sum_{i=1}^{t} n_i(i-1)}\, \lambda^{\sum_{i=1}^{t} n_i}}{\prod_{i=1}^{t} (n_i)!}$$

en posant $a_t(v) = \sum_{i=1}^{t} v^{i-1} = \begin{cases} t & \text{si } v = 1 \\ \dfrac{1-v^t}{1-v} & \phantom{\text{si }} v \neq 1 \end{cases}$

Si h_θ est la densité de $\wedge$ et par la formule de Bayes, la loi a posteriori de $\wedge$ sachant $N_1 = n_1, \ldots, N_t = n_t$ a pour densité

$$h_{v,\theta}^{n_1, \ldots, n_t}(\lambda) = \frac{e^{-\lambda\, a_t(v)}\, \lambda^{\sum_{i=1}^{t} n_i}\, h_\theta(\lambda)}{\int_0^{+\infty} e^{-\lambda\, a_t(v)} \cdot \lambda^{\sum_{i=1}^{t} n_i} \cdot h_\theta(\lambda)\, d\lambda} \propto e^{-\lambda\, a_t(v)}\, \lambda^{\sum_{i=1}^{t} n_i} \cdot h_\theta(\lambda)$$

où $\propto$ est un signe de proportionnalité (à une constante près).

Il résulte de (H1) que la loi prédictive de N_{t+1} sachant $N_1 = n_1, \ldots, N_t = n_t$ s'écrit, pour $m \in \mathbb{N}$,

$$P^{v,\theta}(N_{t+1} = m / N_1 = n_1, \ldots, N_t = n_t)$$
$$= \int_0^{+\infty} P(N_{t+1} = m / \wedge = \lambda)\, h_{v,\theta}^{n_1, \ldots, n_t}(\lambda)\, d\lambda$$

et par (H2)

$$P^{v,\theta}(N_{t+1} = m / N_1 = n_1, \ldots, N_t = n_t)$$
$$= \int_0^{+\infty} \frac{e^{-\lambda v^t}(\lambda v^t)^m}{m!}\, h_{v,\theta}^{n_1, \ldots, n_t}(\lambda)\, d\lambda$$

soit une loi de Poisson-mélange $\mathscr{P}(\wedge v^t)$ où la variable $\wedge$ a pour densité $h_{v,\theta}^{n_1, \ldots, n_t}(\lambda)$. De manière explicite on a

$$P^{v,\theta}(N_{t+1} = m/N_1 = n_1, \ldots, N_t = n_t)$$

$$= \frac{v^{mt}}{m!} \frac{\displaystyle\int_0^{+\infty} e^{-\lambda a_{t+1}(v)} \cdot \lambda^{\sum_{i=1}^t n_i + m} \cdot h_\theta(\lambda)\, d\lambda}{\displaystyle\int_0^{+\infty} e^{-\lambda a_t(v)} \cdot \lambda^{\sum_{i=1}^t n_i} \cdot h_\theta(\lambda)\, d\lambda}$$

avec $a_{t+1}(v) = a_t(v) + v^t$.

Remarque: Comme $P^{v,\theta}(N_{t+1} = m/N_1 = n_1, \ldots, N_t = n_t)$ ne dépend de $(n_1, \ldots, n_t)$ que par l'intermédiaire de $\sum_{i=1}^t n_i$, on a

$$P^{v,\theta}\left(N_{t+1} = m \Big/ \sum_{i=1}^t N_i = n\right) = P^{v,\theta}(N_{t+1} = m/N_1 = n_1, \ldots, N_t = n_t)$$

pour tout t-uple $(n_1, \ldots, n_t)$ tel que $n_1 + \ldots + n_t = n$.

Il en est de même de l'indice $I_{t+1}(n_1, \ldots, n_t; v, \theta)$ défini ci-dessous.

Une caractéristique essentielle d'un système de Bonus-Malus est la prime a posteriori à appliquer la $(t+1)^e$ année après avoir observé $(n_1, \ldots, n_t)$. Cette prime, exprimée en pourcentage de la prime a priori, devient l'indice de fréquence de la $(t+1)^e$ année

$$I_{t+1}(n_1, \ldots, n_t; v, \theta) = 100 \frac{E^{N_1 = n_1, \ldots, N_t = n_t}(N_{t+1})}{E(N_{t+1})}$$

qu'il est inutile de calculer à partir de la loi prédictive de N_{t+1} sachant $N_1 = n_1, \ldots, N_t = n_t$. En effet, en conditionnant par $\wedge$ et par (H1) en (H2):

$$E(N_{t+1}) = E[E(N_{t+1} | \wedge)] = v^t E(\wedge)$$

$$E^{N_1, \ldots, N_t}(N_{t+1}) = E^{N_1, \ldots, N_t}[E(N_{t+1} | N_1, \ldots, N_t, \wedge)]$$

$$E^{N_1, \ldots, N_t}(N_{t+1}) = E^{N_1, \ldots, N_t}[E(N_{t+1} | \wedge)] = v^t E^{N_1, \ldots, N_t}(\wedge)$$

d'où $I_{t+1}(n_1, \ldots, n_t; v, \theta) = 100 \dfrac{E^{N_1 = n_1, \ldots, N_t = n_t}(\wedge)}{E(\wedge)}$.

Dans le cas particulier où $\wedge$ suit une loi Gamma $\gamma(r, \alpha)$ [cf. LEMAIRE, 1985, pour $v = 1$], on a

$$h_{v,\theta}^{n_1, \ldots, n_t}(\lambda) \propto \lambda^{r + \sum_{i=1}^t n_i - 1} \cdot e^{-[\alpha + a_t(v)]\lambda}.$$

La loi a posteriori de $\wedge$ sachant $N_1 = n_1, \ldots, N_t = n_t$ est donc une loi

Gamma $\gamma\left(r+\sum_{i=1}^{t} n_i, \alpha+a_t(v)\right)$. De $E(\wedge)=\frac{r}{\alpha}$ et

$$E^{N_1=n_1,\ldots,N_t=n_t}(\wedge)=\frac{r+\sum_{i=1}^{t} n_i}{\alpha+a_t(v)}\text{, on déduit}$$

$$I_{t+1}(n_1,\ldots,n_t;v,\theta)=100\,\frac{r+\sum_{i=1}^{t} n_i}{r}\cdot\frac{\alpha}{\alpha+a_t(v)}.$$

Comme la loi a posteriori de $v^t\wedge$ sachant $N_1=n_1,\ldots,N_t=n_t$ est une loi Gamma $\gamma\left(r+\sum_{i=1}^{t} n_i, \frac{\alpha+a_t(v)}{v^t}\right)$, la loi prédictive de N_{t+1} sachant $N_1=n_1,\ldots,N_t=n_t$ est une loi de Poisson-Gamma de paramètres $r+\sum_{i=1}^{t} n_i$ et $\frac{\alpha+a_t(v)}{v^t}$ c'est-à-dire une loi Binomiale-Négative $\mathrm{BN}\left(r+\sum_{i=1}^{t} n_i, \frac{\alpha+a_t(v)}{\alpha+a_{t+1}(v)}\right)$.

5. Cas de la loi Inverse-Gaussienne

Dans le cas où $\wedge$ suit une loi IG (μ,β), on a

$$h_{v,\theta}^{n_1,\ldots,n_t}(\lambda)\propto e^{-\lambda a_t(v)}\,\lambda^{\sum_{i=1}^{t} n_i}\,\lambda^{-3/2}\,e^{-\frac{1}{2\beta}\left(\lambda+\frac{\mu^2}{\lambda}\right)}$$

$$\propto \lambda^{\sum_{i=1}^{t} n_i-3/2}\cdot e^{-\left[\lambda a_t(v)+\frac{1}{2\beta}\left(\lambda+\frac{\mu^2}{\lambda}\right)\right]}$$

soit

$$h_{v,\theta}^{n_1,\ldots,n_t}(\lambda)\propto\lambda^{\sum_{i=1}^{t} n_i-3/2}\times$$

$$\times\exp\left\{-\frac{1}{2\beta[1+2\beta a_t(v)]^{-1}}\left[\lambda+\left(\frac{\mu}{\sqrt{1+2\beta a_t(v)}}\right)^2\frac{1}{\lambda}\right]\right\}.$$

La loi de $\wedge$ sachant $N_1=n_1,\ldots,N_t=n_t$ est donc une loi Inverse Gaussienne Généralisée

$$\mathrm{IGG}\left(\sum_{i=1}^{t} n_i-1/2,\ \mu[1+2\beta a_t(v)]^{-1/2},\ \beta[1+2\beta a_t(v)]^{-1}\right).$$

De l'expression de la moyenne d'une loi Inverse Gaussienne Généralisée donnée en 2. on déduit

$$I_{t+1}(n_1, \ldots, n_t; v, \theta) = \frac{100}{\sqrt{1+2\beta a_t(v)}} \times \frac{K_{\sum_{i=1}^{t} n_i + \frac{1}{2}}\left[\frac{\mu}{\beta}(1+2\beta a_t(v))^{1/2}\right]}{K_{\sum_{i=1}^{t} n_i - \frac{1}{2}}\left[\frac{\mu}{\beta}(1+2\beta a_t(v))^{1/2}\right]}.$$

Pour un calcul effectif de I_{t+1} on peut noter qu'en posant

$Q_p(u) = \dfrac{K_{p+1/2}(u)}{K_{p-1/2}(u)}$, pour $p \in \mathbb{N}$ et $u > 0$, on a

$$I_{t+1}(n_1, \ldots, n_t; v, \theta) = \frac{100}{\sqrt{1+2\beta a_t(v)}} \times Q_{\sum_{i=1}^{t} n_i}\left[\frac{\mu}{\beta}(1+2\beta a_t(v))^{1/2}\right]$$

avec, par (i) et (ii) de 2., $Q_0(u) = 1$ et la relation de récurrence:

$$Q_p(u) = \frac{2p-1}{u} + \frac{1}{Q_{p-1}(u)} \quad (p \geqslant 1).$$

Il résulte de 2. que la loi de $v^t \wedge$ sachant $N_1 = n_1, \ldots, N_t = n_t$ est une loi Inverse Gaussienne Généralisée de paramètres

$$\sum_{i=1}^{t} n_i - 1/2,\ \mu\, v^t[1+2\beta a_t(v)]^{-1/2},\ \beta\, v^t[1+2\beta a_t(v)]^{-1}.$$

On en déduit que la loi prédictive de N_{t+1} sachant $N_1 = n_1, \ldots, N_t = n_t$ est une loi de Poisson-Inverse Gaussienne Généralisée de mêmes paramètres et donc, pour $m \in \mathbb{N}$, les probabilités

$$P_m = P_m^{v,\theta}\left(\sum_{i=1}^{t} n_i\right) = P^{v,\theta}(N_{t+1} = m / N_1 = n_1, \ldots, N_t = n_t)$$

se déduisent de P_0 et P_1 comme indiqué en 3. c.

6. ESTIMATION DES PARAMÈTRES

Soit à estimer (θ, v) à partir de la réalisation d'un K-échantillon $(N_1^{(k)}, \ldots, N_t^{(k)})_{k=1,\ldots,K}$ de $(N_1, \ldots, N_t)$ $(t \geqslant 2)$.

La loi de $(N_1, \ldots, N_t)$ a pour expression

$$P^{\nu,\theta}(N_1=n_1,\ldots,N_t=n_t)=\int_0^{+\infty} P^{\nu,\theta}(N_1=n_1,\ldots,N_t=n_t/\wedge=\lambda)h_\theta(\lambda)\,d\lambda$$

$$=\frac{\nu^{\sum_{i=1}^{t}(i-1)n_i}}{\prod_{i=1}^{t}(n_i!)}\int_0^{+\infty} e^{-\lambda a_t(\nu)}\lambda^{\sum_{i=1}^{t} n_i}h_\theta(\lambda)\,d\lambda$$

soit par le changement de variable $\lambda_1=\left(\sum_{i=1}^{t}\nu^{i-1}\right)\lambda$

$$=\frac{\left(\sum_{i=1}^{t} n_i\right)!}{\prod_{i=1}^{t}(n_i!)}\frac{\nu^{\sum_{i=1}^{t} n_i(i-1)}}{(a_t(\nu))^{\sum_{i=1}^{t} n_i}}\int_0^{+\infty}\frac{e^{-\lambda_1}}{\left(\sum_{i=1}^{t} n_i\right)!}\lambda_1^{\sum_{i=1}^{t} n_i}h_{\wedge_1}^{\nu,\theta}(\lambda_1)\,d\lambda_1$$

en notant $h_{\wedge_1}^{\nu,\theta}$ la densité de la variable $\wedge_1=\left(\sum_{i=1}^{t}\nu^{i-1}\right)\wedge$:

$$h_{\wedge_1}^{\nu,\theta}(\lambda_1)=h_\theta\left(\frac{\lambda_1}{a_t(\nu)}\right)\frac{1}{a_t(\nu)}.$$

On fera l'hypothèse de stabilité de la famille $\{h_\theta:\theta\in\Theta\}$ par homothétie:

(H3) Il existe une application $\eta=(\eta_1,\ldots,\eta_q)$ de $\mathbb{R}_+^*\times\Theta$ dans Θ de classe C^1 telle que $h_{\wedge_1}^{\nu,\theta}=h_{\eta(\nu,\theta)}\ \forall(\nu,\theta)\in\mathbb{R}_+^*\times\Theta$ et $\eta(\nu,.)$ bijective pour tout $\nu>0$.

Remarque: Dans le cas Poisson-Gamma où $\theta=(r,\alpha)$, on a $q=2$, $\Theta=\mathbb{R}_+^{*2}$, $\eta_1(\nu,r,\alpha)=r$ et $\eta_2(\nu,r,\alpha)=\dfrac{\alpha}{a_t(\nu)}$.

Dans le cas Poisson-Inverse Gaussienne où $\theta=(\mu,\beta)$, on a $q=2$, $\Theta=\mathbb{R}_+^{*2}$, $\eta_1(\nu,\mu,\beta)=a_t(\nu)\mu$ et $\eta_2(\nu,\mu,\beta)=a_t(\nu)\beta$.

De cette hypothèse il résulte que

$$P^{\nu,\theta}(N_1=n_1,\ldots,N_t=n_t)=\frac{\left(\sum_{i=1}^{t} n_i\right)!}{\prod_{i=1}^{t}(n_i!)}\frac{\nu^{\sum_{i=1}^{t} n_i(i-1)}}{(a_t(\nu))^{\Sigma n_i}}P^{\eta(\nu,\theta)}\left(M=\sum_{i=1}^{t} n_i\right)$$

où P^η est la probabilité d'une loi de Poisson-mélange avec densité de structure $h_{\eta(v,\theta)}$.

La fonction de vraisemblance associée à l'échantillon s'écrit alors en posant $\underline{n}^{(k)} = (n_1^{(k)}, \ldots, n_t^{(k)})$:

$$L[(\underline{n}^{(k)})_{k=1,\ldots K}; v, \theta] = \prod_{k=1}^{K} P^{v,\theta}[N_1 = n_1^{(k)}, \ldots, N_t = n_t^{(k)}]$$

$$\propto \frac{v^{\sum_{k=1}^{K} \sum_{i=1}^{t} n_i^{(k)}(i-1)}}{[a_t(v)]^{\sum_{k=1}^{K} \sum_{i=1}^{t} n_i^{(k)}}} L_1 \left[\left(\sum_{i=1}^{t} n_i^{(k)} \right)_{k=1,\ldots,K} ; \eta(v,\theta) \right]$$

où $L_1 = \prod_{k=1}^{K} P^{\eta(v,\theta)} \left(M^{(k)} = \sum_{i=1}^{t} n_i^{(k)} \right)$ est la vraisemblance associée à un K-échantillon de la loi de Poisson-mélange à densité de structure $h_{\eta(v,\theta)}$.

On en déduit

$$\text{Log } L = \text{constante} + K \left[\sum_{i=1}^{t} \bar{n}_i(i-1) \right] \text{Log } v - K \left(\sum_{i=1}^{t} \bar{n}_i \right) \text{Log} \left(\sum_{i=1}^{t} v^{i-1} \right) + $$
$$+ \text{ Log } L_1 \left[\left(\sum_{i=1}^{t} n_i^{(k)} \right)_{k=1,\ldots,K} ; \eta(v,\theta) \right]$$

et les équations de vraisemblance

$$\left| \begin{array}{l} \dfrac{\partial \text{ Log } L}{\partial \theta_l} = \displaystyle\sum_{j=1}^{q} \frac{\partial \text{ Log } L_1}{\partial \eta_j} \times \frac{\partial \eta_j(v,\theta)}{\partial \theta_l} = 0 \qquad l = 1, \ldots, q \\[2ex] \dfrac{\partial \text{ Log } L}{\partial v} = K \dfrac{\displaystyle\sum_{i=1}^{t} \bar{n}_i(i-1)}{v} - K \dfrac{\displaystyle\sum_{i=1}^{t} \bar{n}_i}{a_t(v)} \cdot \left(\displaystyle\sum_{i=1}^{t} (i-1) v^{i-2} \right) + \\[2ex] \qquad + \displaystyle\sum_{j=1}^{q} \frac{\partial \text{ Log } L_1}{\partial \eta_j} \cdot \frac{\partial \eta_j(v,\theta)}{\partial v} = 0 \end{array} \right.$$

Si on pose $a^* = \sum_{i=1}^{t} \bar{n}_i$ et $b^* = \sum_{i=1}^{t} i\bar{n}_i$, et en supposant qu'il existe $i > 1$ et $j < t$ tels que $\bar{n}_i > 0$ et $\bar{n}_j > 0$, on a $a^* < b^* < a^* t$.

Si $\hat{\eta} = (\hat{\eta}_1, \ldots, \hat{\eta}_q)$ est l'emv de η associé à la réalisation $\left(\sum_{i=1}^{t} n_i^{(k)}\right)_{k=1,\ldots,K}$ d'un K-échantillon de la loi de Poisson-mélange de densité de structure h_η et si $\hat{v}$ est solution de l'équation en v:

$$\text{(E3)} \quad b^* \sum_{i=1}^{t} v^{i-1} - a^* \sum_{i=1}^{t} iv^{i-1} = 0$$

l'emv de (v, θ) est alors $(\hat{v}, \hat{\theta})$ où $\hat{\theta}$ vérifie $\eta(\hat{v}, \hat{\theta}) = \hat{\eta}$.

Remarque: Comme $\Psi(v) = b^* \sum_{i=1}^{t} v^{i-1} - a^* \sum_{i=1}^{t} iv^{i-1}$ est un polynôme en v avec $\Psi_t(0) = b^* - a^* > 0$ et $\lim_{v \to +\infty} \Psi_t(v) = \lim_{v \to +\infty} (b^* - a^* t) v^{t-1} = -\infty$,

(E3) a au moins une racine > 0.

Cas particuliers:

1. Si $t = 2$, (E3) s'écrit $\bar{n}_2 - \bar{n}_1 v = 0$ et a pour solution unique $\hat{v} = \bar{n}_2/\bar{n}_1$.
2. Si $t = 3$ (E3) devient

$$v^2(\bar{n}_2 + 2\bar{n}_1) + v(\bar{n}_1 - \bar{n}_3) - (\bar{n}_2 + 2\bar{n}_3) = 0$$

dont l'unique solution > 0 est

$$\hat{v} = \frac{\sqrt{(\bar{n}_1 - \bar{n}_3)^2 + 4(2\bar{n}_1 + \bar{n}_2)(\bar{n}_2 + 2\bar{n}_3)} - (\bar{n}_1 - \bar{n}_3)}{2(2\bar{n}_1 + \bar{n}_2)}$$

7. APPLICATION

Les résultats précédents sont appliqués à un portefeuille d'assurance automobile de $K = 1\,044\,454$ polices observées durant une période de trois années consécutives (1979-1981, source: Statistique Commune Automobile A.P.S.A.D.).

La baisse tendancielle de la sinistralité moyenne du portefeuille au cours des trois exercices est sensible. Elle résulte d'une part de la baisse générale de la fréquence des sinistres observée en France au cours de cette période et d'une baisse spécifique due au vieillissement du portefeuille observé.

Les moyennes et variances des fréquences annuelles de sinistres au cours des trois exercices valent:

$$\bar{n}_1 = 0{,}17818 \qquad \bar{n}_2 = 0{,}16513 \qquad \bar{n}_3 = 0{,}15724$$
$$\hat{\sigma}_1^2 = 0{,}19739 \qquad \hat{\sigma}_2^2 = 0{,}18244 \qquad \hat{\sigma}_3^2 = 0{,}17485.$$

Les fréquences observées $(n_1^{(k)}, \ldots, n_t^{(k)})_{k=1,\ldots,K}$ sur les t premières années $(1 \leqslant t \leqslant 3)$ permettent d'obtenir, à l'aide des résultats établis dans les paragraphes précédents, pour chacune de ces trois valeurs de t, les emv de v (si $t \geqslant 2$), de (r, α) pour la loi Poisson-Gamma, de (μ, β) pour la loi Poisson-Inverse Gaussienne.

Afin de comparer l'adéquation de ces deux lois aux observations, on donne (tableaux 1, 4 et 5) les fréquences théoriques estimées pour ces deux lois, les fréquences observées correspondantes et, après regroupement des cellules pour respecter le critère de Cochran, la distance du Chi-deux: (obs-théor)^2/théor.

* $t = 1$: loi de Poisson: $\hat{\lambda} = \bar{n}_1 = 0{,}17818$.

Les équations (E1) et (E2) appliquées au K-échantillon $(n_1^{(k)})_{k=1,\ldots,K}$ donnent

$$\text{loi P-Gamma: } \hat{r} = 1{,}67305;\ \hat{\alpha} = \frac{\hat{r}}{\bar{n}_1} = 9{,}38950$$

$$\text{loi P-IG: } \hat{\beta} = 0{,}10812;\ \hat{\mu} = \bar{n}_1 = 0{,}17818.$$

Les fréquences théoriques $KP^{\hat{\theta}}(N_1 = n_1)$ sont données dans le tableau 1 ci-dessous.

TABLEAU 1

COMPARAISON DES FRÉQUENCES OBSERVÉES ET THÉORIQUES ($t = 1$)

Nombre de sinistres	Fréquence observée	Loi de Poisson		Loi P-Gamma		Loi P-I.G.	
		Fréquence théorique	Distance du Khi2	Fréquence théorique	Distance du Khi2	Fréquence théorique	Distance du Khi2
0	881.705	873.987,9	68,14	881.769,5	0,00	881.636,7	0,01
1	142.217	155.729,8	1.172,52	141.993,8	0,35	142.444,7	0,36
2	18.088	13.847,2	1.279,79	18.266,3	1,74	17.838,7	3,48
3	2.118	824,1	2.031,80	2.152,6	0,56	2.205,6	3,48
4	273	36,7	1.521,03	242,1	3,96	283,9	0,42
5	53	1,3	2.009,95	29,7	18,31	44,4	1,67
Total	1.044.454	1.044.454	8.083,23	1.044.454	24,92	1.044.454	9,42

* $t = 2$: les équations (E1) et (E2) appliquées à $(n_1^{(k)} + n_2^{(k)})_{k=1,\ldots,K}$ et (E3) conduisent à $\hat{v} = \bar{n}_2/\bar{n}_1 = 0{,}92676$ et

$$\text{loi P-Gamma: } \hat{r} = 1{,}69720;\ \hat{\alpha} = \frac{(1+\hat{v})\hat{r}}{\bar{n}_1 + \bar{n}_2} = \frac{\hat{r}}{\bar{n}_1} = 9{,}52520$$

$$\text{loi P-IG: } \hat{\beta} = 0{,}10760;\ \hat{\mu} = \frac{\bar{n}_1 + \bar{n}_2}{(1+\hat{v})} = \bar{n}_1 = 0{,}17818.$$

Les fréquences théoriques données dans le tableau 4 ci-dessous sont $KP^{\hat{v},\hat{\theta}}(N_1 = n_1, N_2 = n_2) = KP^{\hat{\theta}}(N_1 = n_1)\, P^{\hat{v},\hat{\theta}}(N_2 = n_2/N_1 = n_1)$, la probabilité $P^{\hat{v},\hat{\theta}}(N_2 = n_2/N_1 = n_1)$ étant calculée à l'aide des résultats donnés en 5 (P-Gamma) et 6. (P-IG).

* $t = 3$: les équations (E1) et (E2) appliquées à $(n_1^{(k)} + n_2^{(k)} + n_3^{(k)})_{k=1,\dots,K}$ et (E3) donnent $\hat{v} = 0{,}93914$ et

$$\text{loi P-Gamma: } \hat{r} = 1{,}65890;\ \hat{\alpha} = \frac{(1+\hat{v}+\hat{v}^2)\hat{r}}{\bar{n}_1+\bar{n}_2+\bar{n}_3} = 9{,}34950$$

$$\text{loi P-IG: } \hat{\beta} = 0{,}110917;\ \hat{\mu} = \frac{\bar{n}_1+\bar{n}_2+\bar{n}_3}{1+\hat{v}+\hat{v}^2} = 0{,}17743.$$

Les fréquences théoriques données dans le tableau 5 ci-dessous sont

$$KP^{\hat{v},\hat{\theta}}(N_1+N_2 = n, \quad N_3 = n_3) = KP^{\hat{v},\hat{\theta}}(N_1+N_2 = n) \quad P^{\hat{v},\hat{\theta}}(N_3 = n_3/N_1+N_2 = n)$$

avec

$$KP^{\hat{v},\hat{\theta}}(N_1+N_2 = n) = \sum_{\substack{n_1,\, n_2 \\ n_1+n_2 = n}} KP^{\hat{v},\hat{\theta}}(N_1 = n_1, N_2 = n_2)$$

et

$$P^{\hat{v},\hat{\theta}}(N_3 = n_3/N_1+N_2 = n) = P^{\hat{v},\hat{\theta}}(N_3 = n_3/N_1 = n_1, N_2 = n_2)$$

pour tout couple (n_1, n_2) tel que $n_1+n_2 = n$ [cf. remarque parag. 4].

Pour la loi Poisson-Gamma puis pous la Poisson-Inverse Gaussienne les distances $d = \Sigma$ (obs-théor)^2/théor du Chi-deux valent respectivement

24,92	et	9,42	($t = 1$; 5 classes)
138,6	et	108,1	($t = 2$; 30 classes)
321,1	et	245,2	($t = 3$; 39 classes).

L'adéquation est donc globalement meilleure pour la loi Poisson-Inverse Gaussienne que pour la Poisson-Gamma. L'écart entre ces 2 lois est cependant du aux classes à forte sinistralité. Pour les classes à sinistralité faible ou moyenne, la loi Poisson-Gamma fournit un meilleur ajustement.

Le nombre d'observations étant très important, le test du χ^2 rejette l'adéquation de l'une et l'autre loi. Dans cette situation on pourrait envisager (pour $t \geqslant 2$) un ajustement par une loi comportant plus de 2 paramètres [par exemple par une loi de Sichel décrite ci-dessus ou par une loi Poisson-Gamma à 3 paramètres (cf. RUOHONEN, 1988)].

Dans les systèmes de Bonus-Malus, la détermination de la prime est obtenue par référence aux indices de fréquence. Les indices estimes $\hat{I}_2(n_1)$ et $\hat{I}_3(n)$ pour la loi Poisson-Gamma et pour la loi Poisson-IG sont donnés dans les tableaux 2 et 3:

$\hat{I}_2(n_1) = I_2(n_1\,;\hat{\theta})$ où $\hat{\theta}$ est l'estimation de θ obtenue pour $t = 1$.
$\hat{I}_3(n) = I_3(n_1+n_2\,;\hat{v},\hat{\theta})$ pour tout couple (n_1, n_2) tel que $n_1+n_2 = n$

(cf. remarque parag. 4) où $(\hat{v},\hat{\theta})$ est l'estimation de (v,θ) obtenue pour $t = 2$.

Ces indices sont comparés aux indices de fréquence empiriques $I_2^*(n_1)$ et $I_3^*(n)$:

$$I_2^*(n_1) = 100\,\frac{\bar{n}_2(N_1 = n_1)}{\bar{n}_2}$$

où le numérateur désigne la fréquence moyenne de 2e année des polices ayant eu n_1 sinistres l'année précédente.

$$I_3^*(n) = 100\,\frac{\bar{n}_3(N_1+N_2 = n)}{\bar{n}_3}$$

où le numérateur représente la fréquence moyenne de 3e année des polices ayant eu n sinistres au cours des 2 premières années.

TABLEAU 2

INDICES DE FRÉQUENCES OBSERVÉS ET THÉORIQUES ($t = 1$)

Nombre de sinistres en 1re année n_1	0	1	2	3	4
$\bar{n}_2\ (N_1 = n_1)$	0,15004	0,23377	0,31811	0,44618	0,57876
$I_2^*(n_1)$ observé	90,86	141,56	192,64	270,19	350,48
$\hat{I}_2(n_1)$ Loi P-Gam	90,38	144,39	198,41	252,43	306,45
$\hat{I}_2(n_1)$ Loi P-I.G	90,68	140,57	208,17	288,96	377,70

TABLEAU 3

INDICES DE FRÉQUENCES OBSERVÉS ET THÉORIQUES ($t = 2$)

Nombre de sinistres cumulés n 1re et 2e années	0	1	2	3	4	5
$\bar{n}_3\ (N_1+N_2 = n)$	0,13055	0,20638	0,28696	0,37561	0,50509	0,62092
$I_3^*(n)$	83,03	131,25	182,50	238,88	321,23	394,89
$\hat{I}_3(n)$ Loi P-Gam	83,18	132,18	181,19	230,20	279,20	328,21
$\hat{I}_3(n)$ Loi P-I.G	84,08	126,77	183,83	251,89	326,88	405,82

TABLEAU 4

COMPARAISON DES FRÉQUENCES OBSERVÉES ET THÉORIQUES ($t = 2$)

n_1 \ n_2	0	1	2	3	4	5	Total
0	763.782 764.018,0 (0,1) 763.288,8 (0,3)	105.046 104.935,9 (0,1) 105.972,3 (8,1)	11.539 11.452,3 (0,7) 11.091,5 (18,1)	1.206 1.142,2 (3,6) 1.122,3 (6,2)	112 108,5 (0,1) 116,7 (0,2)	20 11,1 (7,1) 14,1 (2,5)	881.705 881.668,0 881.605,7
1	113.778 113.228,8 (2,7) 114.347,1 (2,8)	24.246 24.714,8 (8,9) 23.936,0 (4,0)	3.656 3.697,3 (0,5) 3.633,0 (0,1)	471 468,5 (0,0) 503,7 (2,1)	55 54,0 (0,0) 68,0 (2,5)	11 6,5 (3,1) 10,4 (0,0)	142.217 142.169,9 142.498,2
2	13.441 13.334,0 (0,9) 12.913,8 (21,5)	3.731 3.989,5 (16,7) 3.920,1 (9,1)	747 758,3 (0,2) 815,3 (5,7)	148 116,5 (8,5) 146,6 (0,0)	20 15,8 (1,1) 24,6 (0,9)	1 2,2 (0,7) 4,7 (2,9)	18.088 18.216,3 17.825,1
3	1.380 1.435,0 (2,1) 1.410,0 (0,6)	571 545,5 (1,2) 586,5 (0,4)	138 125,7 (1,2) 158,3 (2,6)	19 22,7 (0,6) 35,3 (7,5)	9 3,5 7,1	1 0,6 (8,5) 1,6 (0,2)	2.118 2.133,0 2.198,8
4	160 147,2 (1,1) 158,2 (0,0)	81 67,8 (2,6) 85,4 (0,2)	22 18,4 (0,7) 28,6 (1,5)	8 3,8 7,7	1 0,7 1,8	1 0,1 0,5	273 238,0 282,2
5	17 16,4 (0,0) 21,4 (0,9)	18 8,9 (9,3) 14,3 (1,0)	6 2,8 (3,7) 5,8 (0,0)	4 0,6 1,8	0 0,1 0,5	8 0,0 (52,6) 0,2 (7,2)	53 28,8 44,0

Chiffres supérieurs: fréquences observées
Chiffres intermédiaires: fréquences théoriques P-Gamma
Chiffres inférieurs: fréquences théoriques P-I.G.
} entre parenthèses: distance du χ^2 ie $(\text{obs-théor})^2/\text{théor}$.

TABLEAU 5

COMPARAISON DES FRÉQUENCES OBSERVÉES ET THÉORIQUES ($t = 3$)

n \ n_3	0	1	2	3	4	5	Total
0	673.792 674.380,0 (0,5) 672.645,5 (2,0)	81.180 81.072,4 (0,1) 82.553,9 (22,9)	7.996 7.810,7 (4,4) 7.549,6 (26,4)	720 690,4 (1,3) 661,5 (5,2)	88 58,3 (15,1) 59,2 (14,0)	6 5,3 (0,1) 6,2 (0,0)	763.782 764.017,0 763.475,9
1	179.722 178.246,6 (12,2) 181.504,0 (17,5)	33.802 34.345,6 (8,6) 33.197,3 (11,0)	4.613 4.553,4 (0,8) 4.363,2 (14.3)	605 512,5 (16,7) 520,9 (13,6)	68 52,5 (4,6) 60,3 (1,0)	8 5,7 (1,0) 7,9 (0,0)	218.824 217.716,3 219.653,6
2	37.530 37.756,3 (1,4) 36.494,0 (29,4)	9.667 10.011,3 (11,8) 9.593,1 (0,6)	1.683 1.690,0 (0,0) 1.717,9 (0,7)	299 231,0 (20,0) 265,3 (4,3)	39 27,9 (4.4) 38,1 (0,0)	8 3,4 (6,2) 6,1 (0,6)	49.226 49.719,9 48.114,3
3	7.022 7.337,0 (13,5) 7.030,5 (0,0)	2.297 2.477,1 (13,1) 2.518,0 (19,4)	528 507,9 (0,8) 583,2 (5,2)	111 81,7 (10,5) 111,6 (0,0)	15 11,3 19,2	0 1,6 (0,3) 19,2 (2,7)	9.973 10.416,6 10.266,1
4	1.294 1.361,6 (3,4) 1.384,0 (5,9)	561 558,4 (0,0) 641,2 (10,0)	150 134,7 (1,7) 184,0 (6,3)	45 24,9 (16,2) 42,2 (0,2)	10 3,9 8,5	1 0,6 (9,3) 1,9 (0,0)	2.061 2.084,1 2.261,8
5	264 247,2 (1,1) 284,6 (1,5)	130 119,2 (1,0) 163,4 (6,8)	47 33,1 (5,8) 56,2 (1,5)	12 6,9 (3,7) 15,1 (0,6)	5 1,2 3,5	1 0,2 (14,6) 0,9 (0,6)	459 407,9 523,6
6	34 43 (1,9) 59,3 (10,8)	30 23,9 (1,6) 40,8 (2,8)	17 7,5 (12,1) 16,4 (0,0)	6 1,8 5,0	3 0,3 1,3	0 0,1 (21,0) 0,4 (0,8)	90 76,5 123,1
7	8 6,9 (0,2) 12 (1,3)	9 4,3 (5,1) 9,6 (0,0)		15			24 13,3 28,1
8	4 1,2 2,8	3 0,8 (12,5) 2,6 (0,5)		2,5 (62,5) 8,6 (4,8)			15 2,5 7,4

Chiffres supérieurs : fréquences observées
Chiffres intermédiaires : fréquences théoriques P-Gamma } entre parenthèses : distance du χ^2 ie $(\text{obs-théor})^2/\text{théor}$.
Chiffres inférieurs : fréquences théoriques P-I.G.

TABLEAU 6

INDICES DE FRÉQUENCES THÉORIQUES ($t = 3$)

T \ $\sum_1^T n_i$	0	1	2	3	4	5	6	7	8	9	10
1	90,34 90,46	144,79 141,63	199,25 211,28	253,71 294,55	308,16 385,93	362,62 481,68	417,08 579,79	471,53 679,24	525,99 779,50	580,45 880,28	634,90 981,40
2	82,82 83,62	132,75 127,33	182,67 186,05	232,60 256,14	282,53 333,27	332,45 414,38	382,38 497,69	432,30 582,28	482,23 667,66	532,16 753,55	582,08 839,78
3	76,82 78,43	123,13 116,88	169,44 167,98	215,74 228,87	262,05 296,03	308,36 366,83	354,67 439,72	400,98 513,84	447,28 588,72	493,59 664,00	539,90 739,82
4	71,93 74,34	115,28 108,88	158,64 154,39	202,00 208,52	245,35 268,32	288,71 331,51	332,07 396,67	375,43 463,03	418,78 530,12	462,14 597,70	505,50 665,62
5	67,86 71,03	108,77 102,57	149,68 143,80	190,59 192,78	231,50 246,94	272,41 304,28	313,32 363,50	354,23 423,88	395,14 484,98	436,05 546,55	476,95 608,46
6	64,45 68,29	103,30 97,45	142,14 135,32	180,99 180,24	219,84 229,96	258,69 282,67	297,54 337,20	336,39 392,84	375,24 449,19	414,09 506,01	452,94 563,15
7	61,54 65,99	98,63 93,21	135,73 128,38	172,82 170,03	209,91 216,17	247,01 265,15	284,10 315,87	321,20 367,68	358,29 420,18	395,39 473,14	432,48 526,43

Chiffres supérieurs: loi B.N. (P - Gamma)
Chiffres inférieurs: loi P - I.G.

Les valeurs estimées sont proches des valeurs empiriques pour les historiques de sinistralité faible. Dans les autres cas, les indices déterminés pour la loi Poisson-Gamma sont nettement inférieurs aux indices empiriques alors que ceux calculés pour la loi de Poisson-Inverse Gaussienne sont supérieurs aux indices empiriques.

Dans le dernier tableau (tableau 6) sont donnés pour la loi Binomiale Négative et pour la loi de Poisson-Inverse Gaussienne les indices $I_{T+1}(n_1, \ldots, n_T; \hat{v}, \hat{\theta})$ de fréquence de l'année $T+1$ connaissant le nombre cumulé $\sum_1^T n_i$ des sinistres des T années précédentes, où $(\hat{v}, \hat{\theta})$ est l'estimation de (v, θ) obtenue pour $t = 3$.

On se limitera à $T \leqslant 7$ alors même que l'hypothèse d'une diminution annuelle de la fréquence des sinistres en progression géométrique sur une telle période puisse sembler un peu trop optimiste.

La loi de Poisson-Inverse Gaussienne conduit à un système de Bonus-Malus comportant des niveaux plus élevés que la loi Poisson-Gamma en cas de sinistralité lourde.

AVERTISSEMENT

We would like to thank one of the anonymous referees for his helpful comments on the submitted version of this paper.

RÉFÉRENCES

ACHER, J. (1985) Analyse de la survenance des sinistres en assurance automobile, systèmes de Bonus-Malus. *Journal de la Société Statistique de Paris*. Vol. 126, n° 2, 55–62.

ATHREYA, K. (1986) Another conjugate family for the normal distribution. *Stat. and Prob. Letters* **4**, 61–64.

BRØNS, H. et TOLVER JENSEN, S. (1989) Maximum likelihood estimation in the Negative Binomial distribution. *Comptes-rendus de la 47e session de l'Institut International de Statistique (Paris), Contributed papers*. Vol. 1, 168–169.

ERDELYI, A. et MAGNUS, W. (1953) *Higher transcendental functions*. Vol. 2, Mac Graw Hill.

FOLKS, J. et CHIKARA, R. (1978) The Inverse Gaussian distribution and its statistical application. A review — *Journal of the Royal Statistical Society* **B40**, 263–270.

GOSSIAUX, A. et LEMAIRE, J. (1981) Méthodes d'ajustement de distribution de sinistres. *Bulletin de l'Association des Actuaires Suisses MUSUM* **81**, 157–164.

JOHNSON, N. et KOTZ, S. (1969) *Distributions in Statistics: discrete distributions*. Wiley.

JØRGENSEN, B. (1982) *Statistical properties of the generalized Inverse Gaussian distribution*. Lecture Notes in Statistics. Vol. 9, Springer Verlag.

KESTEMONT, R. et PARIS, J. (1985) Sur l'ajustement du nombre de sinistres. *Bulletin de l'Association des Actuaires Suisses MUSUM* **85**, 157–164.

LEMAIRE, J. (1985) *Automobile Insurance: actuarial models*. Huebner International series on Risk, Insurance and Economic Security. Kluwer-Nijhoff Publishing.

PANJER, H. et WILLMOT, G. (1988) Motivating claim frequency models. *Comptes-Rendus du 23e Congrès International d'Actuaires (Helsinki)*. Vol. 3, 269–284.

PICARD, P. (1976) Généralisation de l'étude sur la survenance des sinistres en assurance automobile. *Bulletin trimestriel de l'Institut des Actuaires Français 1976*, 204–267.

RUOHONEN, M. (1988) A model for the claim number process. *ASTIN Bulletin* **18**, 57–68.

SICHEL, H. (1971) On a family of discrete distributions particularly suited to represent long tailed frequency data. *Proceedings of the third Symposium on Mathematical Statistics*, ed. N.F. Loubscher. Pretoria.
SIMONSEN, W. (1976) On the solution of a maximum likelihood equation of the Negative Binomial distribution. *Scandinavian Actuarial Journal 1976*, 220–231.
SIMONSEN, W. (1979) Correction note. *Scandinavian Actuarial Journal 1979*, 228–229.
STEIN, G. ZUCCHINI, W. et JURITZ, J. (1987) Parameter estimation for the Sichel distribution and its multivariate extension. *Journal of the American Statistical Association* **82**, 938–944.
TEUGELS, J.-L. et SUNDT, B. (1990) A Stop-Loss experience rating scheme for fleets of cars. XXII *Astin Colloquium* (Montreux).
TREMBLAY, L. (1992) Using the Poisson-Inverse Gaussian in Bonus-Malus Systems. *ASTIN Bulletin* **22**, 97–106.
WILLMOT, G. (1986) Mixed compound Poisson distributions. *ASTIN Bulletin* **16-S**, 59–79.
WILLMOT, G. (1987) The Poisson-Inverse Gaussian distribution as an alternative to the Negative Binomial. *Scandinavian Actuarial Journal 1987*, 113–127.
WILLMOT, G. et SUNDT, B. (1989) On posterior probabilities and moments in mixed Poisson processes. *Scandinavian Actuarial Journal 1989*, 139–146.

JEAN-LUC BESSON
Assemblée Plénière des Sociétés d'Assurances Dommages.
26 Boulevard Haussmann, 75009 Paris, France.

CHRISTIAN PARTRAT
Institut de Statistique, Université Pierre et Marie Curie,
4 Place Jussieu, 75252 Paris Cedex 05, France.