A NEW DISTRIBUTION OF POISSON-TYPE FOR THE NUMBER OF CLAIMS

By MICHEL DENUIT

Université Libre de Bruxelles, Bruxelles, Belgium

ABSTRACT

This paper is concerned with two methods to estimate the parameters of the Poisson-Goncharov distribution introduced recently by Lefèvre and Picard (1996). These methods are applied to fit, *inter alia*, the six observed claims distributions, from automobile insurance third party liability portfolios, studied by Gossiaux and Lemaire (1981) and analysed afterwards by several authors.

KEYWORDS

Abel-Goncharov polynomials; Poisson-Goncharov distribution; Generalized Poisson law; fitting; claims distribution.

1. INTRODUCTION

The importance of the third party liability automobile insurance has not to be demonstrated. In most of the industrialized countries, this branch represents a considerable share of the yearly premium collection (for instance, in Belgium, 29.5% during the year 1993). In view of this, many attempts have been made in the actuarial literature to find a probabilistic model for the distribution of the number of auto-accidents (see the review contained in Section 3). Our purpose here is to show that the Poisson-Goncharov distribution introduced recently by Lefèvre and Picard (1996) provides an appropriate probability model to describe the annual number of claims incurred by an insured motorist. We will then propose two methods to estimate the parameters and we will apply them to fit the six observed claims distributions in Gossiaux and Lemaire (1981), as well as recent data sets from Belgium.

Let us briefly survey the paper. In Section 2, we will shortly present the Poisson-Goncharov distribution, establishing two new properties of it. In Section 3, we will first introduce the problem under study and then bring up the Poisson-Goncharov model for the number of claims in automobile insurance. Section 4 will be devoted to the Maximum Likelihood method to estimate the parameters of the Poisson-Goncharov distribution. We will see that this method usually yields implicit Likelihood equations which have to be solved numerically. The starting values of the parameters will be obtained using the so-called "Ad Hoc" method developed in Panjer and Willmot (1992). In Section 5, we will propose a specific

Least Squares type estimation method, specially built for the Poisson-Goncharov distribution. We will see that this method provides explicit expressions for the estimators and, in most of the cases, accurate fits. Section 6 will be concerned with concluding remarks. Finally, Appendix will take all the numerical illustrations up.

2. THE POISSON-GONCHAROV DISTRIBUTION

Recently, Lefèvre and Picard (1996) introduced a new discrete probability distribution on the set of the non-negative integers (subsequently denoted by $\mathbb{N}$), called the Poisson-Goncharov law, which is constructed in terms of Abel-Goncharov polynomials and which extends the classical Poisson law as well as the Generalized Poisson distribution proposed by Consul and Jain (1973).

Very briefly, le $U = \{u_i,\ i \in \mathbb{N}\}$ be any given family of real numbers. To U is attached a unique family of Abel-Goncharov polynomials, $\{G_n(x|U),\ n \in \mathbb{N}\}$, of degree n in x, defined recursively, starting from $G_0(x|U) \equiv 1$, by

$$G_n(x|U) = \frac{x^n}{n!} - \sum_{i=0}^{n-1} \frac{u_i^{n-i}}{(n-i)!} G_i(x|U),\ n \geq 1. \tag{2.1}$$

The reader is referred to Oskolkov (1988) (and the references therein) for a presentation of these polynomials (AG polynomials, in short). Note that the only particular case in which an explicit expression is known for the $G_n(x|U)'s$ is the Abel one. Specifically, if $u_i = a + bi,\ i \in \mathbb{N}$, a and b being real constants, then

$$G_n(x|U) = (x-a)\frac{(x-a-nb)^{n-1}}{n!},\ n \in \mathbb{N}. \tag{2.2}$$

When $u_i = a,\ i \in \mathbb{N}$, (2.2) reduces to $G_n(x|U) = (x-a)^n/n!,\ n \geq 0$. In order to have all the $G_n(x|U)'s$ positive for $x \geq 0$, it suffices for U to be negative and non-increasing, i.e. $0 > u_0 \geq u_1 \geq \ldots \geq u_i \geq u_{i+1} \geq \ldots$ (this condition will be retained subsequently). Now, the Poisson-Goncharov distribution associated with U, negative and non-increasing, is the family $\{\mathcal{PG}_n(U),\ n \in \mathbb{N}\}$ defined by

$$\mathcal{PG}_n(U) = G_n(0|U)e^{u_n},\ n \in \mathbb{N}. \tag{2.3}$$

It is denoted by $\mathcal{PG}(U)$.

As announced earlier, the $\mathcal{PG}(U)$ law can be viewed as a distribution of Poisson-type. If the u_i's are linear in i, $u_i = -\theta - i\lambda$ say, with $\theta \in \mathbb{R}_0^+$ and $\lambda \in \mathbb{R}^+$, then, using (2.2),

$$\mathcal{PG}_n(U) = \frac{\theta(\theta+n\lambda)^{n-1}}{n!} e^{-\theta-n\lambda},\ n \in \mathbb{N}. \tag{2.4}$$

The distribution (2.4) is nothing else than the Lagrangian or Generalized Poisson law introduced by Consul and Jain (1973) (see the book by Consul (1989)). It is non-defective if and only if $\lambda \in [0,\ 1]$. In particular, if all the u_i's are equal to $-\theta$, say, with $\theta \in \mathbb{R}_0^+$, then $\mathcal{PG}(U)$ becomes the usual Poisson distribution with parameter θ.

It is worthwhile recalling that the Generalized Poisson law belongs to the wide class of discrete Lagrangian probability distributions, defined by Consul and Shenton (1972), by means of the Lagrange expansion formula. Moreover, it is also part of the Abel series distributions family introduced by Charalambides (1990). It has various fields of applications, in particular biostatistics (see, e.g., Janardan *et al.* (1979)) as well as actuarial sciences where it has been proposed initially by Consul (1990), and then by Ter Berg (1996), to model the annual number of accidents incurred by a motorist. See also Gerber (1990) for an application linking to the ruin model. We mention that recursive algorithms to evaluate compound Generalized Poisson probabilities have recently been developed, e.g. by Goovaerts and Kaas (1991) and Sharif and Panjer (1995).

Coming back to the $\mathcal{PG}(U)$ law, this corresponds typically to the distribution of the first crossing level L of a Poisson process $\mathcal{N} = \{N(t), t \in \mathbb{R}^+\}$ (with parameter 1, say) in a lower non-decreasing boundary $\mathcal{B}_U$ (such as represented in Figure 2.1). More precisely, we first observe that $\mathcal{B}_U$ may be reduced to the set of points that are eligible as levels of first-crossing, i.e. points with integer ordinate. Denoting this set of points by $\{(-u_i, i), i \in \mathbb{N}\}$, where $U = \{u_i, i \in \mathbb{N}\}$ is negative and non-increasing, it can then be proved that the law of L is provided by (2.3) (see Lefèvre and Picard (1996)). We notice that the Poisson law for L is obtained when $\mathcal{B}_U$ is vertical, and the Generalized Poisson law when $\mathcal{B}_U$ is linear.

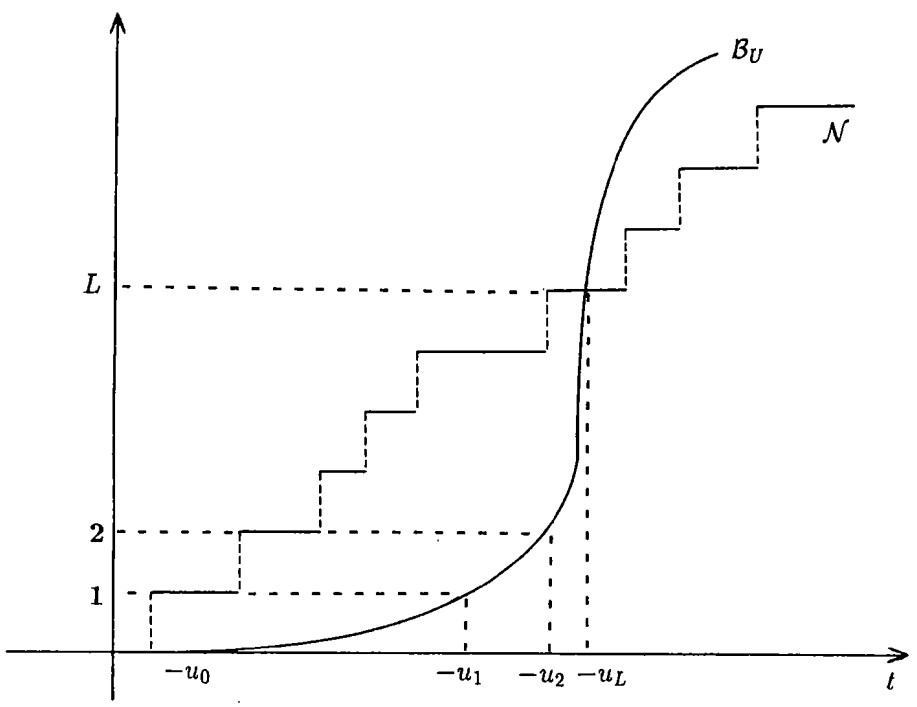

FIGURE 2.1. First-crossing level L of the Poisson process $\mathcal{N}$ in the lower boundary $\mathcal{B}_U$

Lefèvre and Picard (1996) have also pointed out the relationship between the $\mathcal{PG}(U)$ law and the number of customers served in the first busy period of a $M/D/1$ queueing system in which customers arrive according to a Poisson process with rate 1, a unique customer is present initially in the queue and the service times are deterministic but differ from a customer to another (let $-u_0$ be the service time of the initial customer and, for $i \geq 1$, let $u_{i-1} - u_i$ denote the service time of the i-th customer). In such a system, the number of new customers served during the first busy period is of law $\mathcal{PG}(U)$.

Now, the family U may represent any negative and non-increasing sequence of real numbers, so that the $\mathcal{PG}(U)$ law may depend on a large, even infinite, number of parameters. For statistical estimation, however, it is necessary to specify for the u_i's a simple analytical expression, depending on one, two or three parameters for example (like $u_i = \theta_1 + \theta_2 i + \theta_3 i^2$, $i \in I\!N, \theta_1 < 0, \theta_2 \leq 0, \theta_3 \leq 0$).

It is possible to show that the $\mathcal{PG}(U)$ law belongs to the Sundt's family, i.e. satisfies, for some fixed values of k and ω,

$$\mathcal{PG}_n(U) = \sum_{i=1}^{k} \left(\alpha_i + \frac{\beta_i}{n}\right) \mathcal{PG}_{n-i}(U), \; n = \omega + 1, \; \omega + 2, \; \ldots \tag{2.5}$$

Indeed, starting from the following identity,

$$\mathcal{PG}_n(U) = e^{-u_0} \sum_{i=1}^{n} \mathcal{PG}_i(U) \left(\frac{2i}{n} - 1\right) \mathcal{PG}_{n-i}(U), \; n \geq 1, \tag{2.6}$$

we obtain (2.5) with $k = n$, $\omega = 0$ and, for $i \geq 1$, $\alpha_i = -e^{u_i - u_0} G_i(0|U)$, $\beta_i = 2ie^{u_i - u_0} G_i(0|U)$. These α_i's and β_i's are in fact those suggested in Theorem 3 of Sundt (1992). Nevertheless, if we desire to obtain the distribution of the compound $\mathcal{PG}(U)$ sum, as pointed out by Panjer and Wang (1995), when $k = n$ in (2.5), the computing effort using Sundt's recursive formula is of the same order as that needed by a direct convolution approach (Sundt's recursion is interesting only when the claim frequency distribution satisfies (2.5) with small values for k and ω).

Let us recall that, given two random variables X and Y valued in $I\!N$, Y is said to stochastically dominate X, denoted by $X \preceq_{st} Y$, when $P[X \leq n] \geq P[Y \leq n]$ $\forall\, n \in I\!N$ (see, e.g., the recent books by Shaked and Shanthikumar (1994) and by Kaas *et al.* (1994)). It is well-known that $X \preceq_{st} Y$ if and only if there exist two random variables $\tilde{X}$ and $\tilde{Y}$, defined on the same probability space, such that X and $\tilde{X}$, as well as Y and $\tilde{Y}$, are identically distributed, and $P[\tilde{X} \leq \tilde{Y}] = 1$.

It is easy to prove that, if X (resp. Y) is distributed according to the $\mathcal{PG}(U)$ (resp. $\mathcal{PG}(V)$) law, with $u_i > v_i$ $\forall\, i \in I\!N$, then $X \preceq_{st} Y$. This follows immediately from a decomposition formula of the Raikov type, obtained by Lefèvre and Picard (1996), which states that it is always possible to decompose Y into the sum $Z_1 + Z_2$ such that Z_1 follows the $\mathcal{PG}(U)$ law and the law of Z_2, given $Z_1 = j$, $j \in I\!N$, is $\mathcal{PG}(V(j))$, where $V(j) = \{v_i(j), i \in I\!N\}$ with $v_i(j) = v_{i+j} - u_j$. Thus, with X and Y described above and using $\tilde{X} = Z_1$ and $\tilde{Y} = Z_1 + Z_2$, we have $X \preceq_{st} Y$.

If we come back to the representation of the $\mathcal{PG}(U)$ as the law of the first-crossing level L, the above result becomes obvious. Indeed, if X (resp. Y) is the first-crossing level of the Poisson process $\mathcal{N}$ in the lower boundary $\mathcal{B}_U$ (resp. $\mathcal{B}_V$) described by U (resp. by V), and if $v_i < u_i \forall i \in \mathbb{N}$ (that is $\mathcal{B}_V$ lies on the right of $\mathcal{B}_U$), it is clear that $X \preceq_{st} Y$ (since $X \leq Y$ almost surely). Choosing families like $\{-\theta, i \geq 0\}$ or $\{-\theta - i\lambda, i \geq 0\}$, with $\theta > 0$ and $\lambda \in [0, 1]$, yields straight corollaries for the Poisson or the Generalized Poisson laws.

3. THE POISSON-GONCHAROV MODEL FOR THE NUMBER OF CLAIMS IN AUTOMOBILE INSURANCE

Let us first introduce the problem under investigation in the present paper. In order to see if there exist some probability law applicable to claims distributions in automobile insurance third party liability portfolio, Gossiaux and Lemaire (1981) examined six observed claims distributions. Those came from five countries and were studied before by other researchers. Gossiaux and Lemaire (1981) fitted the Poisson distribution, the Generalized Geometric distribution, the Negative Binomial distribution and the mixed Poisson distribution to each of them by the Maximum Likelihood method and the method of Moments. They concluded that no single probability law seems to emerge as providing a good fit to all of them. Moreover, there was at least one example where each model gets rejected by a chi-square test (at the level 10%). Seal (1982) supplemented the paper by Gossiaux and Lemaire (1981) with an analysis of some automobile accidents data from California. He concluded that it supports the mixed Poisson hypothesis for the distribution of the number of claims. Kestemont and Paris (1985), using mixtures of Poisson processes, defined a large class of probability distributions and developed an efficient method of estimating its parameters. For the six data sets in Gossiaux and Lemaire (1981), they proposed a law depending on three parameters and they always obtained extremely good fits. Willmot (1987) showed that the Poisson-Inverse Gaussian law deserves consideration as a model for the claims distribution due to its good fit to the data. Furthermore, this law enjoys abundance of convenient mathematical properties. Willmot (1987) compared the Poisson-Inverse Gaussian distribution to the Negative Binomial one and concluded that the fits are superior with the Poisson-Inverse Gaussian in all the six cases studied by Gossiaux and Lemaire (1981). See also the note by Lemaire (1991) about the confrontation between Negative Binomial and Poisson-Inverse Gaussian on the basis of six data sets not related to insurance. Ruohonen (1987) considered a model for the claim number process. This model is a weighted Poisson process with a three-parameters Gamma distribution as the structure function and is compared with the two-parameters Gamma model giving the Negative Binomial distribution. He fitted his model to some data that can be found in the actuarial literature and the results were satisfying. Panjer (1987) proposed the Generalized Poisson-Pascal distribution, which includes three parameters, for the modelling of the number of automobile claims. The fits obtained were satisfactory, too. Note that the Pólya-Aeppli, the Poisson-Inverse

Gaussian and the Negative Binomial are special cases of this distribution. Willmot (1988) enumerated completely the class of claim frequency distributions discussed by Sundt and Jewell (1981). He demonstrated the good fit to automobile claim frequency data of one member (in fact, the Modified Extended Truncated Negative Binomial distribution), using the six data sets analyzed by Gossiaux and Lemaire (1981). Consul (1990) tried to fit the same six data sets by the Generalized Poisson distribution. Although the Generalized Poisson law is not rejected by a chi-square test, the fits obtained by Kestemont and Paris (1985), for instance, are always much better. Furthermore, Elvers (1991) reported that the Generalized Poisson distribution did not fit very well the data observed in an automobile third party liability insurance portfolio (the distribution hypothesis was, according to his note, in almost every case rejected by a chi-square test). More recently, Ter Berg (1996) considered a slightly different model, involving the Generalized Poisson, too. Moreover, he introduced a loglinear model, which is able to incorporate explanatory variables. The fits were found satisfactory. Islam and Consul (1992) suggested the Consul distribution as a probabilistic model for the distribution of the number of claims in automobile insurance. These authors approximated the chance mechanism which produces vehicle accidents by a branching process. They fit the model to the data sets used by Panjer (1987) and by Gossiaux and Lemaire (1981). Note that this model deals only with autos in accident. Consequently, the zero-class has to be excluded. The fitted values seem good. However, this has to be considered cautiously, due to the comments by Sharif and Panjer (1993). Indeed, these found serious flaws embedded in the fitting of the Consul model; in particular, the very restricted parameter space and some theoretical problems in the derivation of the Maximum Likelihood estimators. They refer to other simple probability models, as the Generalized Poisson-Pascal or the Poisson-Inverse Gaussian, whose fit were found quite satisfying. We end this brief review with two books. The first one is due to Panjer and Willmot (1992) in which Chapter 9 is devoted to the fitting risk model problem. In the second one, by Lemaire (1995), Chapter 3 focus on models for the claims number distribution. These authors give a remarkable account to the problem under investigation.

The probabilistic model for the number of claims incurred by a motorist introduced here extends both the classical Poisson and the Generalized Poisson models. We will use extensively the decomposition formula of the Raikov-type for the $\mathcal{PG}(U)$ recalled above. We split the total number of claims N_{tot} caused by an individual during a fixed period of time (say one year), which is distributed according to the $\mathcal{PG}(U)$ law, where $u_i = \theta_1 + \theta_2 i + \theta_3 i^2, i \in I\!N, \theta_1 < 0,\ \theta_2,\ \theta_3 \leq 0$ into $N_{Poisson}$ and N_{extra}, that is

$$N_{tot} = N_{Poisson} + N_{extra}, \tag{3.1}$$

where $N_{Poisson}$ follows a Poisson law with parameter $-\theta_1$ and N_{extra}, given $N_{Poisson} = j_1$, is distributed according to a $\mathcal{PG}(V(j_1))$ law, with $v_i(j_1) = \theta_2(i+j_1) + \theta_3(i+j_1)^2, i \in I\!N$. Note that $[N_{extra}|N_{Poisson} = j_1]$ increases in the stochastic dominance with j_1. By splitting up the extra claims, it is easily seen

that (3.1) consists in fact in breaking up N_{tot} into

$$N_{tot} = N_{Poisson} + N^{(1)}_{extra} + N^{(2)}_{extra} + N^{(3)}_{extra} + \dots \tag{3.2}$$

where $N^{(1)}_{extra}$, given $N_{Poisson} = j_1$, follows a Poisson law with parameter $-j_1\theta_2 - j_1^2\theta_3$, $N^{(2)}_{extra}$, given $N_{Poisson} = j_1$ and $N^{(1)}_{extra} = j_2$, follows a Poisson law with parameter $-j_2\theta_2 - j_2^2\theta_3 - 2j_1j_2\theta_3$, and so on. Considering N_{tot} as distributed according to the $\mathcal{PG}$ (U) law comes thus down to distinguish among the claims whether they are produced by one or another source, each source adduceing a number of accidents conditionnally distributed as a Poisson law, so that the model (3.2) seems intuitively acceptable.

4 MAXIMUM LIKELIHOOD ESTIMATORS

4.1 A general approach

Let us suppose that the u_i's, $i \in I\!N$, depend on m parameters, i.e. $u_i = u_i(\theta_1, \dots, \theta_m)$, where $(\theta_1, \dots, \theta_m) \in \Theta_1 \times \dots \times \Theta_m \subseteq I\!R^m$. We want to find the Maximum Likelihood estimators (MLE, in short) of the parameters $\theta_1, \dots, \theta_m$.

Let a random sample of size n, $(X_1, \dots, X_n)$, be taken from a population with the $\mathcal{PG}$ (U) law. The corresponding observations are $(x_1, \dots, x_n)$. Let $kmax$ be the largest observation; n_k, $0 \le k \le kmax$, the number of occurences for k; and $f_k = n_k/n$, $0 \le k \le kmax$, the observed frequencies for the different classes. The Likelihood function is

$$L_{\theta_1, \dots \theta_m}(n_0, \dots, n_{k\,max}) = \prod_{k=0}^{kmax} (e^{u_k} G_k(0|U))^{n_k}. \tag{4.1}$$

The MLE $\hat\theta_1, \dots, \hat\theta_m$ of the parameters $\theta_1, \dots, \theta_m$ are such that they maximize the Log-Likelihood function. This leads to the Likelihood equations

$$0 = \sum_{k=0}^{kmax} n_k \left[\frac{\partial u_k}{\partial \theta_j}\right]_{(\theta_1, \dots, \theta_m) = (\hat\theta_1, \dots, \hat\theta_m)} +$$

$$\sum_{k=0}^{kmax} n_k \left[\frac{\Gamma_{j,k}(0|U)}{G_k(0|U)}\right]_{(\theta_1, \dots, \theta_m) = (\hat\theta_1, \dots, \hat\theta_m)}, \; 1 \le j \le m, \tag{4.2}$$

where $\Gamma_{j,k}(x|U)$, $0 \le k \le kmax$, $1 \le j \le m$, denotes the first partial derivative of $G_k(x|U)$ with respect to the parameter θ_j. From (2.1), we see that the $\Gamma_{j,k}(0|U)$'s satisfy the following recurrence relations: starting from $\Gamma_{j,0}(0|U) = 0$, $1 \le j \le m$,

we have that

$$\Gamma_{j,k}(0|U) = -\sum_{i=0}^{k-1} \frac{u_i^{k-i-1}}{(k-i-1)!}\left(\frac{\partial u_i}{\partial \theta_j} G_i(0|U) + \frac{u_i}{k-i}\Gamma_{j,i}(0|U)\right), \tag{4.3}$$

which allow us to compute them recursively.

Unfortunately, the MLE $\hat{\theta}_1, \ldots, \hat{\theta}_m$, which are solutions of (4.2) cannot be obtained in a closed form (except in the Poisson case). Hence, they are computed *via* numerical maximization of the Log-Likelihood function. Let us quote in the next subsection some particular cases of special interest.

4.2 Particular cases

4.2.1. If $U = \{-\theta_1,\ i \in I\!N\}, \theta_1 > 0$, then $G_k(0|U) = \theta_1^k/k!, \Gamma_{1,k}(0|U) = \theta_1^{k-1}/(k-1)!$, and (4.2) gives $\hat{\theta}_1 = \bar{x}$, which is the classical result for the Poisson law.

4.2.2. If $U = \{\theta_1 + \theta_2 h_1(i), i \in I\!N\}, \theta_1,\ \theta_2$, and $h_1(.)$ such that U is negative and non-increasing, let us establish the two following results, which will give us expression for $\Gamma_{1,k}(x|U)$ and $\Gamma_{2,k}(x|U)$, $k \in I\!N$. First of all, we recall two interesting operational properties of the AG polynomials. For any integer n,

$$\frac{d^k}{dx^k} G_n(x|U) = \begin{cases} G_{n-k}(x|E^k U), & \text{if } n \geq k, \\ 0, & \text{otherwise}, \end{cases} \tag{4.4}$$

where $E^k U = \{u_{k+i}, i \in I\!N\}$ denotes the family U without its first k elements. We also have that, for $a,\ b \in I\!R$, and for $n \in I\!N$,

$$G_n(ax + b|aU + b) = a^n G_n(x|U), \tag{4.5}$$

where $aU + b = \{au_i + b,\ i \in I\!N\}$.

Lemma 4.1 *For $a,\ b \in I\!R$ and $n \in I\!N$,*

$$\frac{\partial}{\partial a} G_n(x|\{a + bf(i), i \in I\!N\}) = \begin{cases} -G_{n-1}(x|\{a + bf(i+1), i \in I\!N\}), & \text{if } n \geq 1, \\ 0, & \text{if } n = 0. \end{cases}$$

Proof. The result is obvious for $n = 0$. For $n \geq 1$, using (4.4) and (4.5) yields

$$\begin{aligned} \frac{\partial}{\partial a} G_n(x|\{a + bf(i), i \in I\!N\}) &= \frac{\partial}{\partial a} G_n(x - a|\{bf(i), i \in I\!N\}) \\ &= -G_{n-1}(x - a|\{bf(i+1), i \in I\!N\}) \\ &= -G_{n-1}(x|\{a + bf(i+1), i \in I\!N\}), \end{aligned}$$

hence the announced result. □

Lemma 4.2 *For a, $b \in I\!R$ and $n \in I\!N$,*

$$\frac{\partial}{\partial b} G_n(x|\{a+bf(i), i \in I\!N\}) = \begin{cases} \frac{n}{b} G_n(x|\{a+bf(i), i \in I\!N\}) \\ + \frac{a-x}{b} G_{n-1}(x|\{a+bf(i+1), i \in I\!N\}), & \text{if } n \geq 1, \\ 0, & \text{if } n = 0. \end{cases}$$

Proof. The result is obvious when $n = 0$. For $n \geq 1$, using (4.4) and (4.5) yields

$$\begin{aligned} \tfrac{\partial}{\partial b} G_n(x|\{a+bf(i), i \in I\!N\}) &= \tfrac{\partial}{\partial b}[b^n G_n((x-a)/b|\{f(i), i \in I\!N\})] \\ &= nb^{n-1} G_n((x-a)/b|\{f(i), i \in I\!N\}) \\ &\quad + b^{n-2}(a-x) G_{n-1}((x-a)/b|\{f(i+1), i \in I\!N\}), \end{aligned}$$

and this achieves the proof by (4.5). □

From (4.2) together with Lemma's 4.1 and 4.2, the *MLE* $\hat{\theta}_1$ and $\hat{\theta}_2$ of the parameters θ_1 and θ_2 are thus solutions of the following system:

$$\begin{cases} n = \sum_{k=1}^{kmax} n_k \frac{G_{k-1}\left(0 \middle| \left\{-\hat{\theta}_2 \overline{H}_1 - \bar{x} + \hat{\theta}_2 h_1(i+1), i \in N\right\}\right)}{G_k\left(0 \middle| \left\{-\hat{\theta}_2 \overline{H}_1 - \bar{x} + \hat{\theta}_2 h_1(i), i \in N\right\}\right)}, \\ \hat{\theta}_1 = -\hat{\theta}_2 \overline{H}_1 - \bar{x}, \end{cases} \tag{4.6}$$

where $\bar{x} = \frac{1}{n}\sum_{k=1}^{kmax} n_k k$ and $\overline{H}_1 = \frac{1}{n}\sum_{k=0}^{kmax} n_k h_1(k)$. The second equation of (4.6), with $h_1(i) = i$, gives $\hat{\theta}_1 = -\bar{x}\left(\hat{\theta}_2 + 1\right)$, which is the one obtained by Consul and Shoukri (1984) for the Generalized Poisson distribution. On the other hand, it is possible to show that, when $h_1(i) = i$, $\theta_1 < 0$ and $\theta_2 \in [-1, 0]$, the system (4.6) is equivalent to the one derived by these authors.

The first equation of (4.6), which provides the *MLE* for the parameter θ_2, is implicit. So, we have to use numerical methods to obtain the solution. To get the initial approximation of $\hat{\theta}_2$, we refer to the method described in paragraph 4.2.3.

4.2.3. If $U = \{\theta_1 + \theta_2 h_1(i) + \theta_3 h_2(i),\ i \in I\!N\}$, θ_1, θ_2, θ_3, $h_1(.)$, $h_2(.)$ such that U is negative and non-increasing, it is possible to obtain numerically the *MLE* of θ_1, θ_2 and θ_3, for instance using the method of Scoring, that can be found, e.g., in Panjer and Willmot (1992), pp. 326-328. We will utilize the starting values obtained by the "Ad Hoc" method (*ibidem*, pp. 303-305). The idea of "Ad Hoc" estimation is to equate sample statistics where "most of probability" is to corresponding theoretical quantities. Since there is a high proportion of zeros in the data sets concerning automobile claims, we propose estimators based upon lower classes frequencies, that is we equate the observed lower classes frequencies with the corresponding probabilities. The "Ad hoc" estimators $\ddot{\theta}_1$, $\ddot{\theta}_2$ and $\ddot{\theta}_3$ of the parameters θ_1, θ_2 and θ_3 are given by

$$\begin{cases} \ddot{\theta}_1 = ln(f_0) - \ddot{\theta}_2 h_1(0) - \ddot{\theta}_3 h_2(0)\,;\ \ddot{\theta}_2 = \frac{\alpha - \ln(f_0) - \ddot{\theta}_3(h_2(1)-h_2(0))}{h_1(1)-h_1(0)}\,; \\ \ddot{\theta}_3 = \frac{(\beta - ln(f_0))(h_1(1)-h_1(0)) + (h_1(0)-h_1(2))(\alpha - ln(f_0))}{(h_2(1)-h_2(0))(h_1(0)-h_1(2)) + (h_2(2)-h_2(0))(h_1(1)-h_1(0))}, \end{cases} \tag{4.7}$$

where $\alpha = ln(f_1) - ln(-ln(f_0))$ and $\beta = ln(f_2) - ln\left[-0.5(ln(f_0))^2 + \alpha\, ln(f_0)\right]$.

5. LEAST SQUARES TYPE ESTIMATORS

Let us first establish the following result. For any discrete observed distribution defined on a subset of $I\!N$, $\{f_k;\ 0 \le k \le kmax\}$, such that $f_k > 0$ for all k, there exists a unique family $\tilde{U}$ such that $f_k = e^{\tilde{u}_k} G_k(0|\tilde{U})$, for $k = 0, 1, ..., kmax$. The $\tilde{u}_i$'s are defined recursively as $\tilde{u}_i = ln(f_i) - ln(G_i(0|\tilde{U}))$ $i = 0, 1, ..., kmax$ (let us quote that, by definition, the AG polynomial $G_i(x|\tilde{U})$ depends only on $\tilde{u}_0, \tilde{u}_i, ..., \tilde{u}_{i-1}$). Note that $\tilde{U}$ built above is not always negative and non-increasing. In practice, it is often preferable to only consider those $\tilde{u}_i$'s which form a negative non-increasing family.

The idea is to minimize the function $S(\theta_1, ..., \theta_m)$ defined by

$$S(\theta_1, ..., \theta_m) = \frac{1}{n}\sum_{k=0}^{kmax} n_k(u_k(\theta_1, ..., \theta_m) - \tilde{u}_k)^2,$$

where $\tilde{U}$ is the optimal family constructed above. The Least Square type estimators ($LSTE$, in short) proposed here are thus those which minimize $S(\theta_1, ..., \theta_m)$, the weighted sum of the squared differences between the $\tilde{u}_i$'s and the u_i's having a specified parametric form. Let us mention that if we want to fit an observed distribution $\{(k, n_k),\ 0 \le k \le kmax\}$, we must at first group the classes in order to have all the n_k's positive.

The main advantage of this method is that it often provides explicit expressions for the estimators, as well as accurate fits. We give below the estimators in the case $U = \{\theta_1 + \theta_2 h_1(i) + \theta_3 h_2(i),\ i \in I\!N\}$, parameters θ_1, θ_2 and θ_3, and functions $h_1(.)$ and $h_2(.)$ such that U is negative and non-increasing. The estimators $\check{\theta}_1$, $\check{\theta}_2$ and $\check{\theta}_3$ of the parameters θ_1, θ_2 and θ_3 are those which minimize

$$S(\theta_1, \theta_2, \theta_3) = \frac{1}{n}\sum_{k=0}^{kmax} n_k(\theta_1 + \theta_2 h_1(k) + \theta_3 h_2(k) - \tilde{u}_k)^2.$$

They are given by

$$\check{\theta}_3 = \frac{H_{12}H_{1u} - H_{11}H_{2u}}{(H_{12})^2 - H_{11}H_{22}}\,;\ \check{\theta}_2 = -\theta_3\frac{H_{12}}{H_{11}} + \frac{H_{1u}}{H_{11}}\,;\ \check{\theta}_1 = \overline{U} - \check{\theta}_2\overline{H}_1 - \check{\theta}_3\overline{H}_2, \tag{5.1}$$

where $\overline{U} = \frac{1}{n}\sum_{i=0}^{kmax} n_i\tilde{u}_i$; $\overline{H}_j = \frac{1}{n}\sum_{i=0}^{kmax} n_i h_j(i)$, $j = 1, 2$; $H_{j\times k} = \frac{1}{n}\sum_{i=0}^{kmax} n_i h_j(i) h_k(i)$, $j, k = 1, 2$; $H_{j\times u} = \frac{1}{n}\sum_{i=0}^{kmax} n_i h_j(i)\tilde{u}_i$, $j = 1, 2$; $H_{jk} = H_{j\times k} - \overline{H}_j\overline{H}_k$, $j, k = 1, 2$; $H_{ju} = H_{j\times u} - \overline{H}_j\overline{U}$, $j = 1, 2$.

6. Concluding remarks

Looking at the numerical results presented in Appendix, we could say that the $\mathcal{PG}$ (U) law seems to be suitable to fit the discrete data sets met in automobile insurance. The fits are more accurate than most of the ones discussed before, and applied to recent data sets coming from Belgium, the methods proposed here provide good fits. Moreover, the underlying probabilistic model is intuitively acceptable. Nevertheless, other authors, like for instance Kestemont and Paris (1985), also provided accurate fits, but sometimes with more intricate models. On the other hand, the Least Squares type method is easy to understand and provides explicit expressions for the estimators of the parameters, while it yields satisfying results. We also mention that the simulation method proposed in Devroye (1992), which consists in the partial recreation of the queueing system described in Section 2, can easily be used to simulate the number of claims that affect some automobile insurance portfolio.

Acknowledgements

We greatefully thank Professor Claude Lefèvre for all the time he spent to guide us, and Professor Jean-Marie Reinhard for his precious advices. The author also acknowledges the insightful comments of the two anonymous referees. Their interesting suggestions really improved the paper. We also express our gratitude to the belgian insurance companies for providing us the two last data sets of Appendix.

Appendix: Numerical results

The reader will find here the fits of the six data sets in Gossiaux and Lemaire (1981), as well as of two recent belgian data sets, obtained with the $\mathcal{PG}$ (U), $U = \{\theta_1 + \theta_2 i + \theta_3 i^2,\ i \in \mathbb{N}\}$. To measure the goodness-of-fit, standard χ^2-statistics is used, with the following grouping procedure: the outside classes are gathered together in order to get theoretical class sizes greater or equal to 5 (that is, Rule B in Lemaire (1995)).

Belgium 1975-76				Zaïre 1974			
MLE: $\hat{\theta}_1 = -0.0981$, $\hat{\theta}_2 = -0.0250$; $\hat{\theta}_3 = -0.0037$				*MLE*. $\hat{\theta}_1 = -0.0728$; $\hat{\theta}_2 = -0.1546$; $\hat{\theta}_3 = -0.0005$			
LSTE: $\tilde{\theta}_1 = -0.0981$; $\tilde{\theta}_2 = -0.0212$. $\tilde{\theta}_3 = -0.0069$				*LSTE*: $\tilde{\theta}_1 = -0.0728$; $\tilde{\theta}_2 = -0.1429$; $\tilde{\theta}_3 = -0.0110$			
k	n_k	*ML*	*LST*	k	n_k	*ML*	*LST*
0	96 978	96 978.16	96 975.53	0	3 719	3 719.00	3 719.06
1	9 240	9 244.4	9 252.45	1	232	231.98	232.22
2	704	693.27	684.44	2	38	38.00	37.06
3	43	53.19	55.48	3	7	8.22	8.27
4	9	4.50	5.37	4	3	2.03	2.26
≥ 5	0	0.49	0.73	5	1	0.55	0.71
				≥ 6	0	0.22	0.43
χ^2_{obs}		0.82	4.76	χ^2_{obs}		0.00	0.06

Belgium 1958				Great-Britain 1968			
MLE: $\hat{\theta}_1 = -0.1879$; $\hat{\theta}_2 = -0.1045$; $\hat{\theta}_3 = -0.0078$				MLE: $\hat{\theta}_1 = -0.1285$; $\hat{\theta}_2 = -0.0182$; $\hat{\theta}_3 = -0.0048$			
$LSTE$: $\tilde{\theta}_1 = -0.1883$; $\tilde{\theta}_2 = -0.0699$; $\tilde{\theta}_3 = -0.0337$				$LSTE$: $\tilde{\theta}_1 = -0.1289$; $\tilde{\theta}_2 = -0.0183$; $\tilde{\theta}_3 = -0.0048$			
k	n_k	ML	LST	k	n_k	ML	LST
0	7 840	7 840.00	7 836.81	0	370 412	370 444.17	370 412.38
1	1 317	1 316.97	1 330.77	1	46 545	46 519.72	46 544.74
2	239	239.00	221.81	2	3 935	3 928.66	3 934.60
3	42	49.31	48.47	3	317	316.79	317.53
4	14	11.48	13.77	4	28	27.60	27.68
5	4	2.98	4.84	5	3	2.71	2.72
6	4	0.85	2.01	≥ 6	0	0.35	0.35
7	1	0.26	0.95				
≥ 8	0	0.14	1.57				
χ^2_{obs}		3.38	2.36	χ^2_{obs}		0.03	0.00

Switzerland 1961				Germany 1960			
MLE: $\hat{\theta}_1 = -0.1447$; $\hat{\theta}_2 = -0.0555$; $\hat{\theta}_3 = -0.0099$				MLE: $\hat{\theta}_1 = -0.1359$; $\hat{\theta}_2 = -0.0387$; $\hat{\theta}_3 = -0.0154$			
$LSTE$: $\tilde{\theta}_1 = -0.1447$; $\tilde{\theta}_2 = -0.0571$; $\tilde{\theta}_3 = -0.0078$				$LSTE$: $\tilde{\theta}_1 = -0.1358$; $\tilde{\theta}_2 = -0.0414$; $\tilde{\theta}_3 = -0.0130$			
k	n_k	ML	LST	k	n_k	ML	LST
0	103 704	103 706.62	103 708.19	0	20 592	20 592.00	20 593.44
1	14 075	14 056.34	14 061.18	1	2 651	2 650.60	2 648.58
2	1 766	1 778.12	1 781.10	2	297	297.19	298.96
3	255	256.80	251.79	3	41	40.30	39.76
4	45	43.68	40.95	4	7	6.90	6.52
5	6	8.69	7.64	5	0	1.46	1.30
6	2	1.99	1.62	6	1	0.37	0.31
≥ 7	0	0.72	0.53	≥ 7	0	0.18	0.13
χ^2_{obs}		1.17	0.91	χ^2_{obs}		0.11	0.06

Belgium 1993				Belgium 1994			
MLE: $\hat{\theta}_1 = -0.1017$; $\hat{\theta}_2 = -0.0165$; $\hat{\theta}_3 = -0.0185$				MLE: $\hat{\theta}_1 = -0.1000$; $\hat{\theta}_2 = -0.0253$; $\hat{\theta}_3 = -0.0095$			
$LSTE$: $\tilde{\theta}_1 = -0.1017$; $\tilde{\theta}_2 = -0.0183$; $\tilde{\theta}_3 = -0.0168$				$LSTE$: $\tilde{\theta}_1 = -0.1000$; $\tilde{\theta}_2 = -0.0243$; $\tilde{\theta}_3 = -0.0091$			
k	n_k	ML	LST	k	n_k	ML	LST
0	57 178	57 178.02	57 179.55	0	118 700	118 698.38	118 697.78
1	5 617	5 615.00	5 613.30	1	11 468	11 463.87	11 481.01
2	446	448.56	450.23	2	930	921.22	908.95
3	50	48.25	47.39	3	70	86.61	83.24
4	8	7.22	6.82	4	14	10.13	9.43
≥ 5	0	1.95	1.72	≥ 5	0	1.77	1.59
χ^2_{obs}		0.23	0.22	χ^2_{obs}		3.64	3.41

REFERENCES

CHARALAMBIDES, Ch.A. (1990). Abel series distributions with applications to fluctuations of sample functions of stochastic processes. *Commun. Statist. A, Theory Methods*, **19**, 255-263.

CONSUL, P.C. (1989). *Generalized Poisson Distributions, Properties and Applications*. Marcel Dekker. New York.

CONSUL, P.C. (1990). A model for distributions of injuries in auto-accidents. *MVSV*, 161-168.

CONSUL, P.C. and JAIN, G.C. (1973). A generalization of the Poisson distribution. *Technometrics*, **15**, 791-799.

CONSUL, P.C. and SHENTON, L.R. (1972). Use of Lagrange expansion for generating discrete generalized probability distributions. *SIAM J. Appl. Math.*, **23**, 239-248.

CONSUL, P.C. and SHOUKRI, M.M. (1984). Maximum Likelihood estimation for the Generalized Poisson distribution. *Comm. Statist. A, Theory Methods*, **13**, 1533-1547.

DEVROYE, L. (1992). The branching process method in Lagrange random variate generation. *Commun. Statist., Simula.*, **21**, 1-14.

ELVERS, E. (1991). A note on the Generalized Poisson distribution. *ASTIN Bulletin*, **21**, 167.

GERBER, H.U. (1990). When does the surplus reach a given target? *Insurance: Mathematics and Economics*, **9**, 115-119.

GOOVAERTS, M.J. and KAAS, R. (1991). Evaluating compound Generalized Poisson distributions recursively. *ASTIN Bulletin*, **21**, 193-198.

GOSSIAUX, A. M. and LEMAIRE, J. (1981). Méthodes d'ajustement de distributions de sinistres. *MVSV*, 87-95.

ISLAM, M.N. and CONSUL, P.C. (1992). A probabilistic model for automobile claims. *MVSV*, 85-93.

JANARDAN, K.G., KERSTER, H.W. and SCHAEFFER, D.J. (1979). Biological applications of the Lagrangian Poisson distribution. *Bioscience*, **29**, 599-602.

KAAS, R., VAN HEERWAARDEN, A.E. and GOOVAERTS, M.J. (1994). *Ordering of Actuarial Risks*. CAIRE (Education Series). Brussels.

KESTEMONT, R.-M. and PARIS, J. (1985). Sur l'ajustement du nombre de sinistres. *MVSV*, 157-164.

LEFÈVRE, Cl. and PICARD, Ph. (1996). On the first crossing of a Poisson process in a lower boundary. In *Athens Conference on Applied Probability and Time Series, Vol. I, Applied Probability*. C.C. Heyde, Yu V. Prohorov, R. Pyke and S.T. Rachev Editors. Lecture Notes in Statistics, **114**, pp. 159-175. Springer. Berlin.

LEMAIRE, J. (1991). Negative Binomial or Poisson-Inverse Gaussian? *ASTIN Bulletin*, **21**, 167-168.

LEMAIRE, J. (1995). *Bonus-Malus Systems in Automobile Insurance*. Klumer Academic Publishers. Boston.

OSKOLKOV, V.A. (1988). Abel-Goncharov problem. In *Encyclopaedia of Mathematics, Vol. I*, pp. 4-5. Kluwer Academic Publishers. Dordrecht.

PANJER, H.H. (1987). Models of claim frequency. In *Advances in the Statistical Sciences, Vol. VI, Actuarial Sciences*. I.B. Mac Neill and G.J. Umphrey Editors. The University of Western Ontario Series in Philosophy of Science, **39**, pp. 115-122. Reidel. Dordrecht.

PANJER, H.H. and WILLMOT, G.E. (1992). *Insurance Risk Models*. Society of Actuaries. Schaunburg III.

PANJER, H.H. and WANG, S. (1995). Computational aspects of Sundt's generalized class. *ASTIN Bulletin*, **25**, 5-17.

RUOHONEN, M. (1987). On a model for the claim number process. *ASTIN Bulletin*, **18**, 57-68.

SEAL, H.L. (1982). Mixed Poisson – an ideal distribution of claim number? *MVSV*, 293-295.

SHAKED, M. and SHANTHIKUMAR, J.G. (1994). *Stochastic Orders and their Applications*. Academic Press. New York.

SHARIF, A.H. and PANJER, H.H. (1993). A probabilistic model for automobile claims: a comment on the article by M.N. Islam and P.C. Consul. *MVSV*, 279-282.

SHARIF, A.H. and PANJER, H.H. (1995). An improved recursion for the compound Generalized Poisson distribution. *MVSV*, 93-98.

SUNDT, B. and JEWELL, W.S. (1981). Further results on recursive evaluation of compound distributions. *ASTIN Bulletin*, **12**, 27-39.

SUNDT, B. (1992). On some extension of Panjer's class of counting distributions. *ASTIN Bulletin*, **22**, 61-80.

TER BERG, P. (1996). A loglinear Lagrangian Poisson model. *ASTIN Bulletin*, **26**, 123-129.
WILLMOT, G.E. (1987). The Poisson-Inverse Gaussian distribution as an alternative to the Negative Binomial. *Scandinavian Actuarial Journal*, 113-127.
WILLMOT, G.E. (1988). Sundt and Jewell's family of discrete distributions. *ASTIN Bulletin*, **18**, 17-29.

MICHEL DENUIT
Institut de Statistique et de Recherche Opérationnelle
Universite libre de Bruxelles, CP 210
Boulevard du Triomphe
B-1050 Bruxelles
Belgium
e-mail: mdenuit@ulb.ac.be