SMOOTHED NPML ESTIMATION OF THE RISK DISTRIBUTION UNDERLYING BONUS-MALUS SYSTEMS

MICHEL DENUIT and PHILIPPE LAMBERT
Institut de Statistique
Université Catholique de Louvain
Voie du Roman Pays, 20
B-1348 Louvain-la-Neuve, Belgium
denuit@stat.ucl.ac.be   lambert@stat.ucl.ac.be

August 2000

Abstract

Mixed Poisson distributions are widely used for modeling claim counts when the portfolio is thought to be heterogeneous. The risk (or mixing) distribution then represents a measure of this heterogeneity. The aim of this paper is to use a variant of the Patilea and Rolin (1995) smoothed version of the Simar (1976) NonParametric Maximum Likelihood Estimator of the risk distribution in the mixed Poisson model. Empirical results based on two data sets from automobile third party liability insurance enhance the relevance of this approach. The design of merit-rating schemes is discussed in the second part of the paper.

Key words and phrases: Mixed Poisson model, nonparametric maximum likelihood estimation, credibility theory, Bonus-Malus systems

1 Introduction and Motivation

In most developed countries, the third party liability automobile insurance represents a considerable share of the yearly non-life premium collection (for instance, in Belgium, 26% during the year 1998). Therefore, many attempts have been made in the actuarial literature to find a probabilistic model for the distribution of the number of auto-accidents; for a review of the existing literature, we refer the interested reader e.g. to Lemaire (1995) or to Denuit (1997). Most of these models are parametric (i.e. an analytical expression is assumed for the probabilities that a policyholder reports k claims during an insurance period, depending on one or several parameters to be estimated on the basis of the observations). In order to see if there exist a universal model for claims distributions in automobile portfolios, Gossiaux and Lemaire (1981) examined six observed data sets. 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 a two-point 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. This author concluded that it supports the mixed Poisson hypothesis for the distribution of the number of claims.

In this paper, we will work in the mixed Poisson model, but no assumption will be made about the risk (or mixing) distribution. Following Walhin and Paris (1999), we first recall the basic features about the NonParametric Maximum Likelihood estimator (NPMLE, in short) of the risk distribution. As pointed out by these authors, the NPMLE suffers from some serious drawbacks in the design of Bonus-Malus systems. The problems are mainly due to its purely discrete nature. Therefore, we will propose a smoothed version of the NPMLE. In the second part of this paper, we focus on "Bonus-Malus Systems" (BMS, in short). A BMS is a particular form of experience rating. It penalizes insured responsible for one or more accidents by premium surcharges, or *maluses*, and rewards claim-free policyholders by awarding them discounts, or *bonuses*. An excellent account of these systems can be found in Lemaire (1995).

Let us consider a portfolio consisting of n policies, numbered 1 to n. Denote as K_{ij} the number of claims incurred by the ith policyholder during the jth year the policy is in force. In credibility theory, it is assumed that, for fixed i, the K_{ij}'s are conditionally independent and identically distributed given a random risk parameter Θ_i that represents unknown risk characteristics of the policy. After t years the available data are $(K_{i1}, K_{i2}, \dots, K_{it})$ and the insurance company wants to use these data to adjust the premium for year $t+1$; the premium for year $t+1$ is thus a function $\Psi(K_{i1}, K_{i2}, \dots, K_{it})$ of the past claims. Actuaries have traditionally applied minimization of the expected quadratic loss in order to determine Ψ, that is, Ψ minimizes $\mathbb{E}\left[\Psi(K_{i1}, K_{i2}, \dots, K_{it}) - \Theta_i\right]^2$, which is interpreted as the expected difference between the "true" premium Θ_i and the credibility premium $\Psi(K_{i1}, K_{i2}, \dots, K_{it})$. Henceforth, we assume that the sequences $\{\Theta_i, K_{i1}, K_{i2}, K_{i3}, \dots\}$ are independent and identically distributed; for the ease of explanation, we drop the policyholder's index i.

Considering the last paragraph, the very basic elements of a BMS are as follows:

1) an appropriate premium calculation principle;

2) a conditional distribution for the number of claims, that is, for the $[K_j|\Theta = \theta]$'s;

3) a distribution for the risk parameter Θ to describe how the conditional distributions vary accross the portfolio.

Let us give some details on these aspects. Considering the premium calculation principle, we use the expected value principle. The latter principle requires the insured to pay the pure premium plus a safety loading proportional to the pure premium. The pure premium will be the individual claim frequency per year multiplied by the average cost of a claim and can be scaled so that it will be equal to the claim frequency. The problem of the insurer is to predict, at the renewal of the policy, the claim frequency of the insured for this new year given the observations of the reported accidents in the preceding periods.

Let us now turn to the conditional distribution of the annual claim numbers. In automobile third party liability insurance portfolio, the Poisson distribution provides a good description of the number of claims incurred by an individual policyholder during a given reference period (one year, say). The assumptions underlying the Poisson counting model should indeed at least provide a good approximation to the accident generating mechanism; see e.g. Lemaire (1995). Therefore, we consider in the remainder of the paper that the number of claims incurred by a given policyholder during a reference period conforms to a Poisson distribution.

Now, individual driving abilities vary from individuals to individuals. Consequently the portfolio is heterogeneous and policyholders will have different Poisson parameters. This is indicated by the rejection of the homogeneous Poisson model when the latter is applied to fit data sets from automobile insurance portfolios; for empirical evidence supporting this assertion, see e.g. Gossiaux and Lemaire (1981). In order to reflect the different underlying risk profiles, each policyholder is characterized by the value of his mean claim frequency θ, and θ is considered to be a realization of a non-observable random variable Θ whose support is contained in the half-positive real line $\mathbb{R}^+ \equiv [0, +\infty)$. In other words, the conditional probability that a driver with annual mean claim frequency θ is involved in k accidents during the ith year is

$$\mathbb{P}[K_j = k|\Theta = \theta] \equiv p(k|\theta) = \exp(-\theta)\frac{\theta^k}{k!}, \;\; k \in \mathbb{N} \equiv \{0, 1, 2, \dots\}. \tag{1.1}$$

The annual number of accidents caused by a randomly selected policyholder of the portfolio during the jth year is then distributed according to a mixed Poisson law, that is,

$$\mathbb{P}[K_j = k] \equiv p(k|\Theta) = \int_{\theta \in \mathbb{R}^+} p(k|\theta) dF_\Theta(\theta), \;\; k \in \mathbb{N}, \tag{1.2}$$

where F_Θ denotes the cumulative distribution function (cdf, in short) of Θ, assumed to fulfill $F_\Theta(0) = 0$. The mixing distribution described by F_Θ represents the heterogeneity of the portfolio of interest; F_Θ is often called the structure function. It is worth mentioning that the mixed Poisson model (1.2) is an accident-proneness model: it assumes that a policyholder's mean claim frequency does not change over time but allows some insured persons to have higher mean claim frequencies than others.

Sometimes, (1.2) is taken to be a finite mixture model, that is, the mixing distribution is discrete and puts positive masses $\pi_1, \pi_2, \ldots, \pi_q$ on only a finite number q of positive real atoms $0 < \theta_1 < \theta_2 < \ldots < \theta_q$. Then,

$$p(k|\Theta) = \sum_{\ell=1}^{q} p(k|\theta_\ell)\pi_\ell, \;\; k \in \mathbb{N}. \tag{1.3}$$

The fact that Θ has a distribution with q support points means that the portfolio of interest consists of only q categories of policyholders. The special case $q = 2$ gives the classical "good risk/bad risk" model considered in Gossiaux and Lemaire (1981). Note that the actual reality of the insurance business is a finite mixture model (by taking q to be the number of policyholders in the portfolio). In risk theory, the finite mixture model (1.3) was first proposed by Grenander (1957a); see also Grenander (1957b).

Let us now consider the choice of F_Θ. Traditionally, actuaries have assumed that the distribution of θ values among all drivers is well approximated by a two-parameter Gamma distribution. This choice is particularly desirable because the class of the Gamma distributions is the natural conjugate family for the Poisson and facilitates a Bayesian approach towards updating mean frequency estimates. The resulting probability distribution for the number of claims is Negative Binomial. Other classical choices for F_Θ include the Inverse-Gaussian (which results in the Poisson-Inverse-Gaussian law for the number of claims; see e.g. Willmot (1987) and Tremblay (1992)) and Hoffman's distributions (see e.g. Kestemont and Paris (1985) and Walhin and Paris (1999)). However, there is no particular reason to believe that F_Θ belongs to some specified parametric family of distributions. Therefore, we would like to resort on a nonparametric estimator for F_Θ. This will thus lead to BMS relying on fewer assumptions than usual ones.

More precisely, after having recalled some key features of the model (1.2) in Section 2, we apply the Simar (1976) NPMLE of F_Θ in Section 3. The Maximum Likelihood approach results in a finite mixture model (1.3) with relatively few support points (see (3.1)). As pointed out by Walhin and Paris (1999), this model is undesirable for constructing BMS. Therefore, we propose in Section 4 to use a variant of the Patilea and Rolin (1995) Empirical Nonparametric Bayesian estimator for F_Θ; this estimator is a finite mixture of Gamma distributions and can be intuitively considered as a smoothed version of the NPMLE with the Gamma distribution playing the role of a kernel. In Section 5, we examine the BMS obtained with this model.

The present paper expands on several previous works. Albrecht (1982) gave a first account of statistical methods connected with model (1.2), mainly in a maximum likelihood approach. More recently, Walhin and Paris (1999) compared BMS obtained with Hofmann's parametric family and Simar's NPMLE for F_Θ. These authors showed that, although the NPMLE is powerful to evaluate functionals of claim counts, it is not suitable for building BMS (because it is purely discrete). Our approach consists in smoothing Simar's estimator with a Gamma kernel, and is thus comparable with Carrière's (1993b) study, who smoothed the Tucker-Lindsay moment estimator with a Log-Normal kernel.

Let us now detail some of the notations used throughout this paper. We denote as K_θ (resp. K_Θ) a random variable with probability distribution $\{p(k|\theta), \;\; k \in \mathbb{N}\}$ in (1.1) (resp. $\{p(k|\Theta), \;\; k \in \mathbb{N}\}$ in (1.2)). We denote by μ_k, $k = 1, 2, \ldots$, the moments $\mathbb{E}K_\Theta^k$ of K_Θ. Those of Θ are the ν_k's, $k = 1, 2, \dots$, that is, $\nu_k = \mathbb{E}\Theta^k$. By convention, $\mu_0 = \nu_0 \equiv 1$. Henceforth, we assume that we have observed an insurance collective consisting of n independent policies. The data that we have at our disposal are as follows: we know that n_k policies caused k claims during the reference period, $k = 0, 1, \dots, k_{\max}$; $k_{\max}$ is the maximal number of claims observed for a policy. The NonParametric Maximum Likelihood estimations of the claim frequencies are

$$\begin{cases} \widehat{p}(k) = \frac{n_k}{n}, & k = 0, 1, \dots, k_{\max}, \\ \widehat{p}(k) = 0, & k \geq k_{\max} + 1. \end{cases}$$

These unconstrained estimations reproduce exactly what is observed in the data. Thus, the NonParametric Maximum Likelihood estimations of the moments μ_k are

$$\widehat{\mu}_k = \frac{1}{n} \sum_{j=1}^{k_{\max}} j^k \widehat{p}(j), \quad k \in \mathbb{N}.$$

For the numerical illustrations, we used the two data sets presented in Appendix A. Portfolio 1 relates to Belgium and has been observed in 1958; it can be found e.g. in Gossiaux and Lemaire (1981). Portfolio 2 has been kindly provided to us by a large insurance company operating in the Benelux; it has been observed in 1995.

2 Basic Properties of the Mixed Poisson Model

2.1 Estimation of mixing functionals

According to Carrière (1993a), given a function $\phi : \mathbb{R}^+ \to \mathbb{R}$, the quantity $\mathbb{E}\phi(\Theta)$ is estimable if there exists a function $\psi : \mathbb{N} \to \mathbb{R}$ such that

$$\mathbb{E}\phi(\Theta) = \mathbb{E}\psi(K_\Theta). \tag{2.1}$$

Of course, such a function ψ theoretically always exists: it suffices to take $\psi(K_\Theta) = \mathbb{E}[\phi(\Theta)|K_\Theta]$ so that (2.1) holds provided ϕ is integrable. The actual meaning of (2.1) is that we desire an explicit expression for ψ. If ϕ possesses some desirable property, ψ can be obtained explicitly. This is for instance the case when ϕ is an absolutely monotone function, i.e. that all the derivatives $\phi^{(1)}, \phi^{(2)}, \phi^{(3)}, \dots$ of ϕ exist and are non-negative. Carrière (1993a) proved that the function ψ involved in (2.1) is then given by

$$\psi(\ell) = \sum_{k=0}^{\ell} \binom{\ell}{k} \phi^{(k)}(0), \quad \ell \in \mathbb{N}.$$

In practice, in order to estimate a quantity $\mathbb{E}\phi(\Theta)$, we use

$$\widehat{\mathbb{E}\phi(\Theta)} = \sum_{k=0}^{k_{\max}} \psi(k)\widehat{p}(k).$$

Carrière (1993a) proved the asymptotic normality for such estimators.

As simple examples, take $\phi(\theta) = \exp(t\theta)$ then

$$\psi(\ell) = \sum_{k=0}^{\ell} \left(\begin{array}{c} \ell \\ k \end{array} \right) t^k = (1+t)^\ell;$$

as a consequence, the moment generating function of Θ is estimable. The knowledge of $\{p(k|\Theta),\ k \in \mathbb{N}\}$ is thus equivalent to the knowledge of F_Θ. For $\phi(\theta) = \theta^k$, we get

$$\psi(\ell) = \ell(\ell-1)\ldots(\ell-k+1) \text{ for } \ell = k, k+1, \ldots;$$

the moments ν_k of Θ are thus estimable. More precisely, the ν_k's are estimated by

$$\begin{cases} \widehat{\nu}_k = \sum_{j=k}^{k_{\max}} j(j-1)\ldots(j-k+1)\widehat{p}(j), \quad k = 1, 2, \ldots, k_{\max}, \\ \widehat{\nu}_k = 0, \quad k \geq k_{\max}+1. \end{cases}$$

The estimator $\widehat{\nu}_k$ is unbiased and almost surely consistent for ν_k.

2.2 Testing the mixed Poisson hypothesis

The present work focuses on the model (1.2). Considering the possibility of misspecification, there is a need of a statistical test to decide whether the model (1.2) is reasonable to fit the data. To this end, let us present the non-parametric test proposed by Carrière (1993a). The reasoning behind this test is as follows. For any positive integer k, let $\mu_{[k]}$ be the kth descending factorial moment of K_Θ, i.e.

$$\mu_{[k]} = \mathbb{E}[K_\Theta(K_\Theta - 1)\ldots(K_\Theta - k + 1)].$$

If K_Θ has a mixed Poisson distribution then $\mu_{[k]} = \nu_k$. From Jensen inequality, we find that

$$\mu_{[k]} \geq (\mathbb{E}\Theta)^k = (\mu_{[1]})^k$$

must hold for $k = 2, 3, \ldots$ whenever K_Θ is mixed Poisson; therefore if $\mu_{[k]} < (\mu_{[1]})^k$ for some k then the underlying distribution cannot be of mixed Poisson type. Based on this fact, Carrière (1993a) suggested the test statistic $\sqrt{n}\left\{(\widehat{\mu}_{[1]}, \widehat{\mu}_{[k]}) - (\mu_{[1]}, \mu_{[k]})\right\}$ that weakly converges to a bivariate Normal distribution as $n \to +\infty$.

Carrière (1993a) constructed a Bonferroni multiple comparison test. In its simplest form, this statistical procedure is as follows. In order to decide whether the number of claims caused by a policyholder of the portfolio can conform to a mixed Poisson distribution (i.e. to test the null hypothesis H_0 that the underlying distribution is of the form (1.2)), it suffices to compute the value T_{obs} of the test statistic

$$T = \frac{\sqrt{n}(\widehat{\mu}_{[1]}^2 - \widehat{\mu}_{[2]})}{\sqrt{4(1-\widehat{\mu}_{[1]})(\widehat{\mu}_{[1]}^3 - 2\widehat{\mu}_{[2]}\widehat{\mu}_{[1]} + \widehat{\mu}_{[3]}) + \widehat{\mu}_{[4]} + 2\widehat{\mu}_{[2]} - \widehat{\mu}_{[2]}^2}}$$

and to reject H_0 if $T_{obs} > z_\alpha$ where z_α is such that

$$\frac{1}{\sqrt{2\pi}} \int_{t=-\infty}^{z_\alpha} \exp(-t^2/2)dt = 1 - \alpha,$$

and

$$\widehat{\mu}_{[k]} = \sum_{j=k}^{k_{\max}} j(j-1)\dots(j-k+1)\widehat{p}(j), \quad k = 1, 2, 3, 4.$$

Note that this test relies on the asymptotic properties of T so that n has to be large enough.

We performed Carrière's test on each of the two data sets presented in Appendix A. The model (1.2) was never rejected; for Portfolio 1, $T_{obs} = -7.53$ and for Portfolio 2, $T_{obs} = -12.93$.

2.3 Poisson vs. Poisson mixture

Let us now recall some basic facts about the model (1.2). First of all, it makes sense to study the mixed Poisson model through F_Θ. As noticed above, there is indeed a one-to-one correspondence between the mixing distribution and the resulting mixed distribution, that is, if K_{Θ_1} and K_{Θ_2} are identically distributed then Θ_1 and Θ_2 also are.

To each of the two data sets presented in Appendix A, we fitted an homogeneous Poisson distribution to the observations. These fits, given in column A, were clearly rejected (p-values smaller than 10^{-3}). This indicates that the two portfolios are heterogeneous. Another technique to check for the heterogeneity of the portfolio is described next. Therefore, let us recall that the model (1.2) enjoys the following nice property: for any Θ such that $\mathrm{Var}[\Theta] > 0$, the sequence $\{p(k|\Theta) - p(k|\nu_1),\ k \in \mathbb{N}\}$ has alternating signs "+,–,+". This result has been established by Shaked (1980). For data sets presented in Appendix A, we plotted in Figure 1 the sequence $\{\widehat{p}(k) - p(k|\widehat{\nu}_1),\ k = 0, 1, \dots, k_{\max}\}$. We expect to observe Shaked's sign sequence "+,–,+" if the data come from a Poisson mixture (1.2). We notice that the difference between the observed data and its Poisson fit has the "+,–,+" property as it is bound to have when the underlying distribution is a mixture of Poisson distributions. This indicates that the Poisson parameter varies from individual to individual.

3 NonParametric Estimation of the Risk Distribution

3.1 NPMLE

In a seminal paper, Simar (1976) gave a detailed description of the NPMLE of F_Θ, as well as an algorithm for its computation. The NPMLE is a discrete distribution, so that the resulting model is of the form (1.3). Simar (1976) obtained an upper bound for the size of the support of the NPMLE. This upper bound uses the quantity κ defined to be the number of observed distinct values, i.e.

$$\kappa = \#\{k \in \mathbb{N} \text{ such that } \widehat{p}(k) > 0\};$$

in most cases, $\kappa = k_{\max} + 1$. To be specific, Simar (1976) showed that the NPMLE $\widehat{F}_\Theta$ of F_Θ exists and is unique; the number of support points of the MLE is less than or equal to

$$\widehat{q} = \min\left\{\left[\frac{k_{\max}+1}{2}\right], \kappa\right\}, \tag{3.1}$$

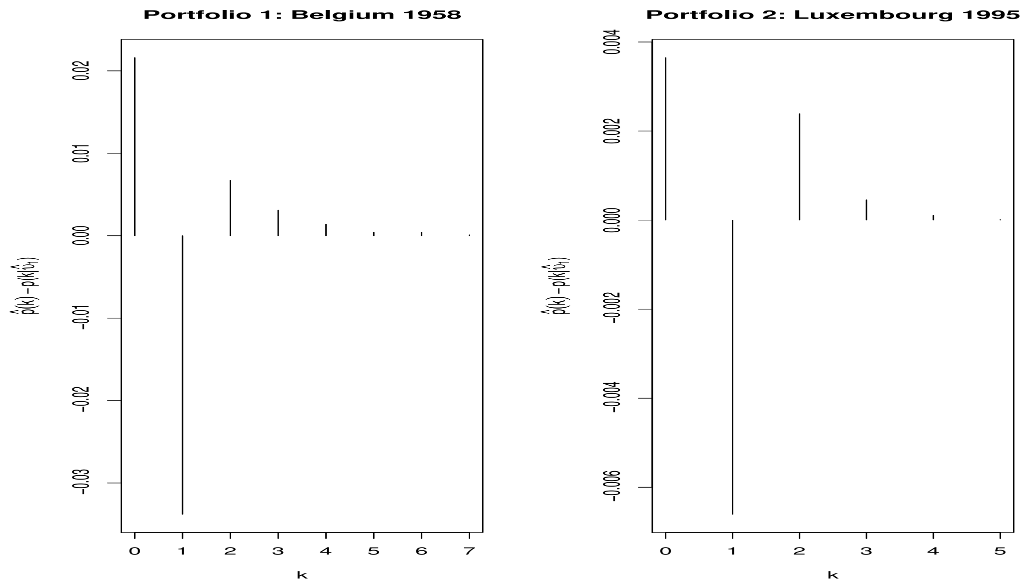

Figure 1: *Sequence* $\{\widehat{p}(k) - p(k|\widehat{\nu}_1),\ k = 0, 1, \ldots, k_{\max}\}$ *for the data sets presented in Appendix A.*

where $[x]$ denotes the integer part of the real x; for the data sets in Appendix A, $\widehat{q} = 4$ for Portfolio 1 and $\widehat{q} = 3$ for Portfolio 2. The solution $\widehat{F}_\Theta$ puts probability masses $\widehat{\pi}_1, \widehat{\pi}_1, \ldots, \widehat{\pi}_{\widehat{q}}$ at the atoms $\widehat{\theta}_1, \widehat{\theta}_1, \ldots, \widehat{\theta}_{\widehat{q}}$. In order to get a first approximation of $\widehat{F}_\Theta$, we resort on the moment estimator for F_Θ proposed by Tucker (1963) and suitably precised by Lindsay (1989a,b).

The NPMLE fits to each of the data sets can be found in Appendix A, together with the corresponding observed values of the χ^2-statistics. When $\widehat{q} \geq 3$, we fitted a model with 2 and 3 components. The results can be summarized as follows:

1. For Portfolio 1, the NPMLE of F_Θ has at most 4 support points. It appeared that a 3-point NPMLE gives a satisfactory fit (displayed in Column B), reflected in a p-value of 51%. The 3-point $\widehat{F}_\Theta$ is thus preferred in virtue of the statistical principle of parcimony. The NPMLE creates 3 categories of policyholders: the best ones (with a claim frequency of about 0) representing 41.8% of the portfolio, the standard ones (with a claim frequency of 33.6%) representing 57.3% of the portfolio, and the bad ones (with a claim frequency of 254.4%) representing 0.1% of the portfolio. The fit provided by a 2-point $\widehat{F}_\Theta$ (displayed in Column C) is rejected since the p-value is equal to 0.5%.

2. For Portfolio 2, we have $\widehat{q} = 3$ and we fitted the data with a 3-point (Column B) and a 2-point (column C) NPMLE. Since the quality of the two fits is similar (p-values of 26 and 29%, respectively), we prefer the 2-point $\widehat{F}_\Theta$. We thus have a good risk/bad risk model, with 93.3% of good drivers whose claim frequency is 6.8% and 6.7% of bad drivers with a claim frequency of 44.6%.

3.2 Smoothed NPMLE

In order to estimate F_Θ, Patilea and Rolin (1995) suggested to resort on a finite mixture of natural conjugate priors of the Poisson distribution; they call this estimator an Empirical NonParametric Bayesian estimator (ENBE, in short). These authors proved that the ENBE is an asymptotic Maximum Likelihood estimator. In other words, it is an estimator that almost maximizes the likelihood in the sense that the difference between the maximal value of the likelihood (as a function of F_Θ) and the value of the likelihood corresponding to the ENBE tends to zero as the sample size grows to $+\infty$. This ensures the consistency of the ENBE. We propose here a slightly modified version of the Patilea-Rolin estimator. In order to smooth the NPMLE of F_Θ, we let the family of natural conjugate priors play the role of a kernel. This technique is somewhat similar to the approach followed by Carrière (1993b); this author proposed to smooth the Tucker-Lindsay moment estimator with a Log-Normal kernel.

The natural way to smooth $\widehat{F}_\Theta$ consists in using

$$\sum_{k=1}^{\widehat{q}} \widehat{\pi}_k \Gamma(\theta | n\widehat{\pi}_k\widehat{\theta}_k, n\widehat{\pi}_k),\ \ \theta \in \mathbb{R}^+,$$

where $\Gamma(.|\alpha,\beta)$ denotes the cumulative distribution function corresponding to a two-parameter Gamma law with mean α/β and variance α/β^2, $\widehat{q}$ is Simar's upper bound (3.1) for the support size of the NPMLE and $\widehat{\pi}_k$'s and $\widehat{\theta}_k$'s are the corresponding masses and atoms. It is easily seen that the kth component of the mixture is centered at $\widehat{\theta}_k$; this corresponds to the intuitive idea that the NPMLE indicates the number and the locations of policyholders classes in the portfolio. Then, the distribution of the risk parameter in these classes is represented by a two-parameter Gamma distribution, resulting in a mixture of Gamma's. However, the variance of each component equals $\widehat{\theta}_k/n\widehat{\pi}_k$, which is virtually 0 since the number n of policies is usually very large. As a consequence, the smoothed estimator is more or less indistinguishable from the NPMLE. In order to avoid this phenomenon, we resort on an estimator of the form

$$\widetilde{F}(\theta) = \sum_{k=1}^{\widetilde{q}} \widetilde{\pi}_k \Gamma(\theta | n^{\widetilde{\lambda}}\widetilde{\pi}_k\widetilde{\theta}_k, n^{\widetilde{\lambda}}\widetilde{\pi}_k),\ \ \theta \in \mathbb{R}^+, \tag{3.2}$$

where $\widetilde{q}$ is taken as small as possible and in any case smaller than Simar's upper bound (3.1) for the support size of the NPMLE and the $\widetilde{\pi}_k$'s, $\widetilde{\theta}_k$'s and $\widetilde{\lambda}$ are maximum likelihood estimators. The main difference with Patilea and Rolin (1995)'s work is thus the introduction of the parameter λ in order to avoid the variance of each component of the mixture defining $\widetilde{F}_\Theta$ to be virtually zero.

With (3.2), (1.2) reduces to a mixture of Negative Binomial distributions, i.e.

$$\widetilde{p}(k|\Theta) = \sum_{j=1}^{\widetilde{q}} \widetilde{\pi}_j \begin{pmatrix} n^{\widetilde{\lambda}}\widetilde{\pi}_j\widetilde{\theta}_j + k - 1 \\ k \end{pmatrix} \left(\frac{n^{\widetilde{\lambda}}\widetilde{\pi}_j}{1+n^{\widetilde{\lambda}}\widetilde{\pi}_j}\right)^{n^{\widetilde{\lambda}}\widetilde{\pi}_j\widetilde{\theta}_j} \left(\frac{1}{1+n^{\widetilde{\lambda}}\widetilde{\pi}_j}\right)^k,\ \ k \in \mathbb{N}. \tag{3.3}$$

Let us now apply this method to the data sets of Appendix A. In both cases, we took $\widetilde{q} = 2$ in order to avoid overparameterization. In Figures 2-3, one can find the densities corresponding to the different components involved in the mixture $\widetilde{F}_\Theta$, as well as the resulting risk distribution (the continuous line represents $d\widetilde{F}_\Theta$ and the dotted line the classical two-parameter Gamma mixing with parameters estimated *via* maximum likelihood). The model proposed is a slight generalization of the good risk/bad risk model: the portfolio is splitted into two populations, each one having its own two-parameter Gamma structure function.

Let us now examine the fits obtained with the 2-component $\widetilde{F}_\Theta$:

1. For Portfolio 1, $\widetilde{\lambda} = 0.22$. The fit is given in Column E; it is very accurate and is regarded as satisfactory on the basis of the χ^2-criterion (p-value of 36%). It is worth mentioning that the Negative Binomial fit displayed in Column D is clearly rejected. Considering Figure 2, we see that $\widetilde{F}_\Theta$ puts more mass on large values than the classical two-parameter Gamma.

2. For Portfolio 2, we get $\widetilde{\lambda} = 0.28$. Again, the fit is satisfactory, and better than the Negative Binomial one. Figure 3 illustrates the difference between the Gamma mixing and $\widetilde{F}_\Theta$.

4 Resulting BMS

Let us now examine the merit-rating schemes obtained in the mixed Poisson model (1.2) using a quadratic loss function and the structure function $\widetilde{F}_\Theta$ defined in (3.2). The net premium for a new insured is given by $P_1 = \mathbb{E}[K_1] = \mathbb{E}[\Theta]$. After t years of coverage, the amount of premium for the $(t+1)$th period is $P_{t+1}(K_1, K_2, \ldots, K_t)$; it is determined so to minimize the expected squared difference between the true premium Θ and the premium P_{t+1} charged to the policyholder, i.e. to minimize $\mathbb{E}[P_{t+1}(K_1, K_2, \ldots, K_t) - \Theta]^2$. The solution of this optimization problem is the posterior mean $P_{t+1}(K_1, K_2, \ldots, K_t) = \mathbb{E}[\Theta | K_1, K_2, \ldots, K_t]$. Given $K_1 = k_1, K_2 = k_2, \ldots, K_t = k_t$, denote $k = \sum_{j=1}^t k_j$. We then get

$$\begin{aligned} P_{t+1}(k_1, k_2, \ldots, k_t) &= \int_{\theta \in \mathbb{R}^+} \theta d\mathbb{P}[\Theta \leq \theta | K_1 = k_1, K_2 = k_2, \ldots, K_t = k_t] \\ &= \frac{\int_{\theta \in \mathbb{R}^+} \theta \left\{ \prod_{i=1}^t \mathbb{P}[K_i = k_i | \Theta = \theta] \right\} dF_\Theta(\theta)}{\int_{\eta \in \mathbb{R}^+} \left\{ \prod_{i=1}^t \mathbb{P}[K_i = k_i | \Theta = \eta] \right\} dF_\Theta(\eta)} \\ &= \frac{\int_{\theta \in \mathbb{R}^+} \exp(-t\theta) \theta^{k+1} dF_\Theta(\theta)}{\int_{\eta \in \mathbb{R}^+} \exp(-t\eta) \eta^k dF_\Theta(\eta)} \equiv P_{t+1}(k); \end{aligned}$$

$P_{t+1}(k)$ appears as the ratio of two Mellin transforms, as expected from Albrecht (1984). It is interesting to note that the premium P_{t+1} only depends on the total number k of accidents caused in the past t years of insurance, and not on the history of these claims. This is a characteristic of the theoretical Bonus-Malus scales (with an infinite number of levels). In practice, since the Bonus-Malus scale is upper bounded, policyholders take always an advantage of concentrating all the claims during a single period.

Assuming that the first premium paid is 100, a Bonus-Malus scale depending on the number k of claims at fault during t years of coverage can be constructed with the aid of the

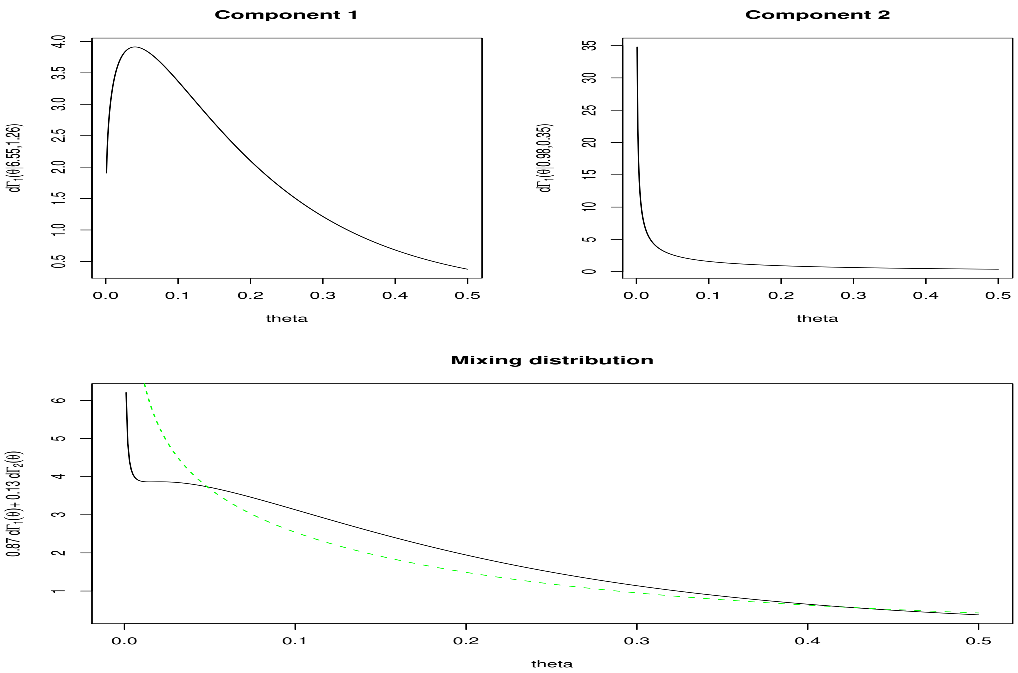

Figure 2: *Components of (3.2) and resulting $\widetilde{F}_\Theta$ for Portfolio 1 in Appendix A.*

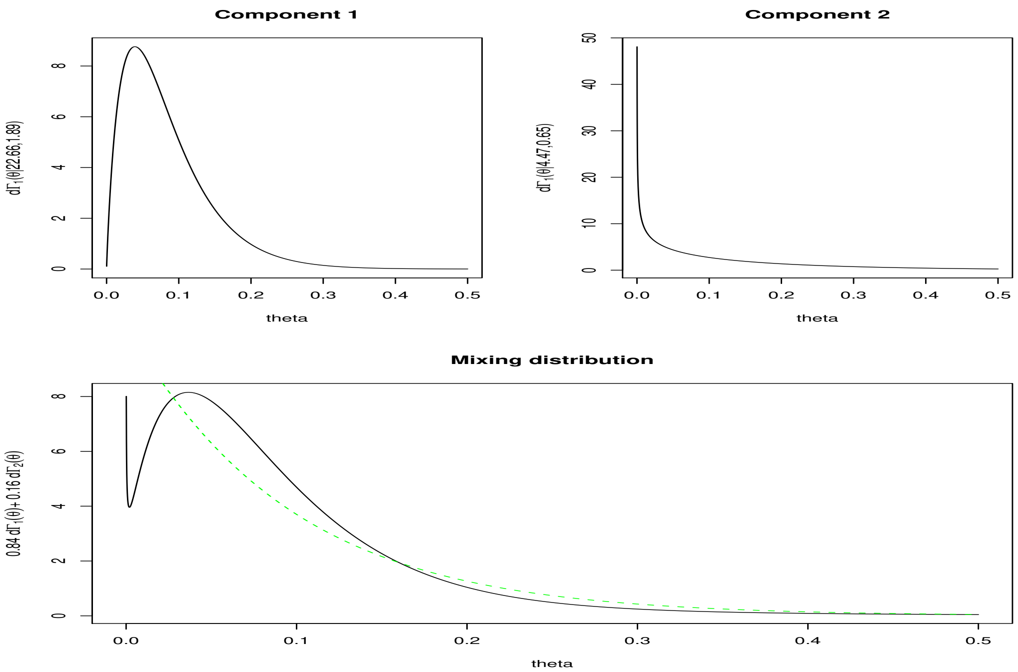

Figure 3: *Components of (3.2) and resulting $\widetilde{F}_\Theta$ for Portfolio 2 in Appendix A.*

formula

$$\beta(k,t) = 100 \times \frac{P_{t+1}(k)}{P_1}\%;$$

$\beta(k,t)$ is the relative level of premium for the $(t+1)$th year of coverage for an insured person who caused k accidents during the first t years.

In Appendix B, we considered Portfolio 2 (two support points for $\widehat{F}_\Theta$ and two components for $\widetilde{F}_\Theta$). We first built a BMS with the NPMLE $\widehat{F}_\Theta$. The $\beta(k,t)$'s so obtained are given in Tables B.1-B.2. A "block" structure is clearly apparent, each block with almost constant $\beta(k,t)$ corresponding to one support point of $\widehat{F}_\Theta$. In Figure 4, the evolution of the premium for a driver who caused 10 claims during $[0,t]$ is depicted as a function of $t \in \mathbb{N}$. A step behavior is clearly apparent. The policyholder is first put in the category $\widehat{\theta}_2 = 0.446$. Then, the BMS needs several claim-free years to decide that this individual belongs to the category $\widehat{\theta}_1 = 0.068$. *Grosso modo*, there is only one discount, the premium being constant before and after. At first, $\beta(10,1)$ equals 477.8946% (whereas it equals 477.8947% if we know that the driver is a bad risk), and after that, the premium decreases to $\beta(10,70) = 72.8767\%$ (it equals 72.8629% for good risks). Such a behavior, which is a byproduct of the purely discrete nature of the NPMLE, is undesirable. In order to avoid this, we need a smooth risk distribution, as (3.2). The $\beta(k,t)$'s derived from the estimator $\widetilde{F}_\Theta$ of the structure function F_Θ are given in Tables B.4-B.3, while Figure 5 is the counterpart of Figure 4. See Appendix B for the details of the computations. The BMS is now "smooth", with continuous variations of the $\beta(10,t)$'s; this can be regarded as commercially desirable.

To end with, let us mention that the $\beta(k,t)$'s of Tables B.4-B.3 can be transformed in a standard table following the method proposed by Coene and Doray (1996).

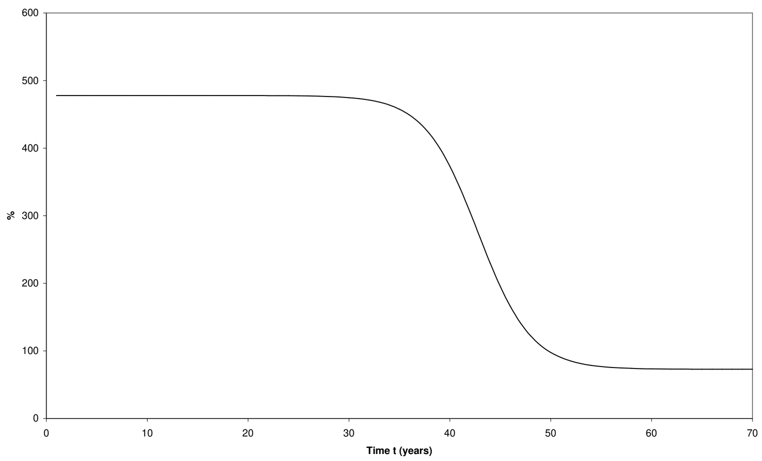

Figure 4: *Evolution of $\beta(10,t)$ as a function of $t = 1, 2, \ldots, 70$ with $\widehat{F}_\Theta$ for Portfolio 2.*

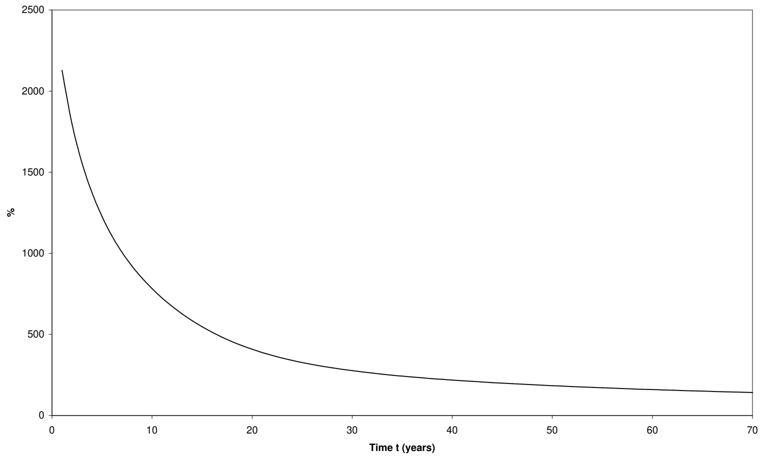

Figure 5: *Evolution of $\beta(10,t)$ as a function of $t = 1, 2, \ldots, 70$ with $\widetilde{F}_\Theta$ for Portfolio 2.*

5 Conclusions

In this paper, we demonstrated that an adequate smoothed version of the NPMLE is a good candidate for estimating the risk distribution in a mixed Poisson model for the claim count. This estimator is nonparametric; no assumption is thus made on the mixing distribution. Moreover, as a mixture of Gamma distributions, it is mathematically tractable, e.g. to elaborate BMS. In that respect, it performs better than the NPMLE which is purely discrete and results in "discontinuous" experience rating plans. Of course, the smoothed NPMLE does not provide accurate fits in all the cases: for instance, both NPMLE and smoothed NPMLE yielded poor fits for the data set relating to Belgium 1975-1976 provided in Gossiaux and Lemaire (1981)

In a forthcoming paper, the same problematic will be considered when *a priori* risk classification is enforced. Specifically, we will examine how to design merit rating plans in accordance with *a priori* ratemaking structure of the insurance company.

Acknowledgements

The support of the Belgian Government under "Projet d'Actions de Recherche Concertées" (No. 98/03-217) is gratefully acknowledged.

APPENDICES

A Data sets

The reader will find herefter two data sets from Benelux countries, together with all the fits considered in the present paper. To measure the goodness-of-fit, standard χ^2-statistics is used, with the following calculation procedure:

$$\chi^2_{obs} = -2 \sum_{k=0}^{k_{\max}} n_k \ln \left(\frac{\widehat{p(k|\Theta)}}{\widehat{p(k)}} \right).$$

Portfolio 1						
k	n_k	A	B	C	D	E
0	7,840	7,636	7,840	7,832	7,847	7,839
1	1,317	1,637	1,317	1,337	1,288	1,322
2	239	175	239	213	257	231
3	42	13	42	57	54	48
4	14	1	13	17	12	13
5	4	0	6	4	3	5
6	4	0	2	1	1	2
7	1	0	1	0	0	1
≥ 8	0		0	0	0	0
χ^2_{obs}		302.48	2.33	16.85	17.00	4.36
# d.f.		6	3	5	6	4
p-value		$< 10^{-3}$	0.51	0.005	0.009	0.36

Column A: expected frequency with homogeneous Poisson
Column B: expected frequency with 3-point NPMLE $\widehat{F}_\Theta$
$\widehat{\theta}_1 = 0.336$, $\widehat{\theta}_2 \approx 0.000$, $\widehat{\theta}_3 = 2.545$
$\widehat{\pi}_1 = 0.573$, $\widehat{\pi}_2 = 0.418$ and $\widehat{\pi}_3 = 0.001$
Column C: expected frequency with 2-point NPMLE $\widehat{F}_\Theta$
$\widehat{\theta}_1 = 0.147$, $\widehat{\theta}_2 = 1.231$, $\widehat{\pi}_1 = 0.938$ and $\widehat{\pi}_2 = 0.062$
Column D: expected frequency with Negative Binomial
Column E: expected frequency with 2-component $\widetilde{F}_\Theta$
$\widetilde{\lambda} = 0.22$, $\widetilde{\theta}_1 = 0.193$, $\widetilde{\theta}_2 = 0.355$, $\widetilde{\pi}_1 = 0.869$ and $\widetilde{\pi}_2 = 0.131$

Portfolio 2						
k	n_k	A	B	C	D	E
0	102,435	102,026	102,435	102,435	102,442	102,435
1	8,804	9,544	8,805	8,811	8,774	8,806
2	714	446	712	703	746	710
3	65	14	68	76	63	70
4	12	0	10	8	5	9
5	1	0	2	1	0	1
≥ 6	0	0	0	0	0	0
χ^2_{obs}		365.67	1.25	3.78	8.18	1.94
# d.f.		5	1	3	4	2
p-value		$< 10^{-3}$	0.26	0.29	0.09	0.38

<u>Column A:</u> expected frequency with homogeneous Poisson

<u>Column B:</u> expected frequency with 3-point NPMLE $\widehat{F}_\Theta$
$\widehat{\theta}_1 = 0.132$, $\widehat{\theta}_2 = 0.829$, $\widehat{\theta}_3 = 0.000$
$\widehat{\pi}_1 = 0.651$, $\widehat{\pi}_2 = 0.009$ and $\widehat{\pi}_3 = 0.340$

<u>Column C:</u> expected frequency with 2-point NPMLE $\widehat{F}_\Theta$
$\widehat{\theta}_1 = 0.068$, $\widehat{\theta}_2 = 0.446$, $\widehat{\pi}_1 = 0.933$ and $\widehat{\pi}_2 = 0.067$

<u>Column D:</u> expected frequency with Negative Binomial

<u>Column E:</u> expected frequency with 2-component $\widetilde{F}_\Theta$
$\widetilde{\lambda} = 0.28$, $\widetilde{\theta}_1 = 0.083$, $\widetilde{\theta}_2 = 0.145$, $\widetilde{\pi}_1 = 0.835$ and $\widetilde{\pi}_2 = 0.165$

B Theoretical BMS

Tables B.1-B.2 contain the Bonus-Malus coefficients $\beta(k,t)$ computed with the NPMLE $\widehat{F}_\Theta$ of F_Θ. Their counterparts B.4-B.3 are based on the smoothed NPMLE $\widetilde{F}_\Theta$. These quantities are computed on the basis of Portfolio 3, 2-point $\widehat{F}_\Theta$ and 2-component $\widetilde{F}_\Theta$. In Tables B.1-B.2,

$$\beta(k,t) = \frac{\sum_{j=1}^{\widehat{q}} \exp(-t\widehat{\theta}_j)\widehat{\theta}_j^{k+1}\widehat{\pi}_j}{\sum_{j=1}^{\widehat{q}} \exp(-t\widehat{\theta}_j)\widehat{\theta}_j^{k}\widehat{\pi}_j} \times \frac{100}{\sum_{j=1}^{\widehat{q}} \widehat{\theta}_j\widehat{\pi}_j}.$$

Let us briefly detail the computational aspects of Tables B.4-B.3. When the risk distribution is a Gamma mixture, i.e.

$$F_\Theta(\theta) = \sum_{j=1}^{q} \alpha_j \Gamma(\theta|a_j,\tau_j), \quad \theta \in \mathbb{R}^+, \tag{B.1}$$

we get

$$\begin{aligned} dF_\Theta(\theta|K_1=k_1,K_2=k_2,\ldots,K_t=k_t) &= \frac{\sum_{j=1}^{q}\alpha_j \exp(-t\theta)\theta^k d\Gamma(\theta|a_j,\tau_j)}{\sum_{i=1}^{q}\alpha_i \int_{\eta\in\mathbb{R}^+}\exp(-t\eta)\eta^k d\Gamma(\eta|a_i,\tau_i)} \\ &= \sum_{j=1}^{q} A(j,k)\, d\Gamma(\theta|a_j+k,\tau_j+t) \end{aligned}$$

where

$$A(j,k) = \alpha_j \frac{\int_{\eta\in\mathbb{R}^+}\exp(-t\eta)\eta^k d\Gamma(\eta|a_j,\tau_j)}{\sum_{i=1}^{q}\alpha_i \int_{\eta\in\mathbb{R}^+}\exp(-t\eta)\eta^k d\Gamma(\eta|a_i,\tau_i)}.$$

This yields

$$P_{t+1}(k_1,k_2,\ldots,k_t) = \sum_{j=1}^{q} A(j,k)\frac{a_j+k}{\tau_j+t}.$$

The coefficients $A(j,k)$'s are easy to compute. Indeed, they can be cast into

$$A(j,k) = \alpha_j \frac{\epsilon(j,k)}{\sum_{i=1}^{q}\alpha_i \epsilon(i,k)},$$

where

$$\begin{aligned} \epsilon(j,k) &= \int_{\eta\in\mathbb{R}^+} \frac{\exp(-t\eta)(t\eta)^k}{k!} d\Gamma(\eta|a_j,\tau_j) = \int_{\eta\in\mathbb{R}^+} \frac{\exp(-\eta)\eta^k}{k!} d\Gamma(\eta|a_j,\tau_j/t) \\ &= \binom{a_j+k-1}{k}\left(\frac{\tau_j}{\tau_j+t}\right)^{a_j}\left(\frac{t}{\tau_j+t}\right)^{k}. \end{aligned}$$

The $\epsilon(j,k)$'s satisfy the Panjer recurrence relations

$$\epsilon(j,k) = \frac{t}{\tau_j+t}\frac{a_j+k-1}{k}\epsilon(j,k-1), \quad k=1,2,\ldots,$$

starting from

$$\epsilon(j,0) = \int_{\eta\in\mathbb{R}^+} \exp(-\eta)d\Gamma(\eta|a_j,\tau_j/t) = \left(\frac{\tau_j}{\tau_j+t}\right)^{a_j}.$$

	k										
t	0	1	2	3	4	5	6	7	8	9	10
0	100										
1	92	172	348	451	473	477	478	478	478	478	478
2	86	146	313	439	472	477	478	478	478	478	478
3	82	126	275	424	469	476	478	478	478	478	478
4	79	111	237	404	465	476	478	478	478	478	478
5	77	100	202	378	459	475	477	478	478	478	478
6	76	92	171	347	450	473	477	478	478	478	478
7	75	86	146	312	439	471	477	478	478	478	478
8	74	82	126	274	424	469	476	478	478	478	478
9	74	79	111	236	403	464	476	478	478	478	478
10	74	77	100	201	377	459	475	477	478	478	478
11	73	76	92	170	346	450	473	477	478	478	478
12	73	75	86	145	311	439	471	477	478	478	478
13	73	74	82	125	273	423	468	476	478	478	478
14	73	74	79	110	235	403	464	476	478	478	478
15	73	74	77	99	200	377	458	475	477	478	478
16	73	73	76	91	170	345	450	473	477	478	478
17	73	73	75	86	145	310	438	471	477	478	478
18	73	73	74	82	125	272	423	468	476	478	478
19	73	73	74	79	110	234	402	464	476	478	478
20	73	73	74	77	99	199	376	458	475	477	478
21	73	73	73	76	91	169	345	450	473	477	478
22	73	73	73	75	86	144	309	438	471	477	478
23	73	73	73	74	82	124	271	422	468	476	478
24	73	73	73	74	79	110	233	402	464	476	478
25	73	73	73	74	77	99	199	375	458	475	477
26	73	73	73	73	76	91	168	344	449	473	477
27	73	73	73	73	75	85	143	308	438	471	477
28	73	73	73	73	74	82	124	270	422	468	476
29	73	73	73	73	74	79	109	232	401	464	476
30	73	73	73	73	74	77	99	198	375	458	475
31	73	73	73	73	73	76	91	168	343	449	473
32	73	73	73	73	73	75	85	143	307	437	471
33	73	73	73	73	73	74	82	124	269	421	468
34	73	73	73	73	73	74	79	109	232	400	464
35	73	73	73	73	73	73	77	98	197	374	458
36	73	73	73	73	73	73	76	91	167	342	449
37	73	73	73	73	73	73	75	85	142	306	437
38	73	73	73	73	73	73	74	81	123	268	421
39	73	73	73	73	73	73	74	79	109	231	400
40	73	73	73	73	73	73	73	77	98	196	373

Table B.1: $\beta(k,t)$ with $\widehat{F}_\Theta$ for Portfolio 2.

	k										
t	0	1	2	3	4	5	6	7	8	9	10
41	73	73	73	73	73	73	73	76	91	166	341
42	73	73	73	73	73	73	73	75	85	142	305
43	73	73	73	73	73	73	73	74	81	123	267
44	73	73	73	73	73	73	73	74	79	109	230
45	73	73	73	73	73	73	73	73	77	98	195
46	73	73	73	73	73	73	73	73	76	90	166
47	73	73	73	73	73	73	73	73	75	85	141
48	73	73	73	73	73	73	73	73	74	81	122
49	73	73	73	73	73	73	73	73	74	79	108
50	73	73	73	73	73	73	73	73	73	77	98
51	73	73	73	73	73	73	73	73	73	76	90
52	73	73	73	73	73	73	73	73	73	75	85
53	73	73	73	73	73	73	73	73	73	74	81
54	73	73	73	73	73	73	73	73	73	74	79
55	73	73	73	73	73	73	73	73	73	73	77
56	73	73	73	73	73	73	73	73	73	73	76
57	73	73	73	73	73	73	73	73	73	73	75
58	73	73	73	73	73	73	73	73	73	73	74
59	73	73	73	73	73	73	73	73	73	73	74
60	73	73	73	73	73	73	73	73	73	73	73
61	73	73	73	73	73	73	73	73	73	73	73
62	73	73	73	73	73	73	73	73	73	73	73
63	73	73	73	73	73	73	73	73	73	73	73
64	73	73	73	73	73	73	73	73	73	73	73
65	73	73	73	73	73	73	73	73	73	73	73
66	73	73	73	73	73	73	73	73	73	73	73
67	73	73	73	73	73	73	73	73	73	73	73
68	73	73	73	73	73	73	73	73	73	73	73
69	73	73	73	73	73	73	73	73	73	73	73
70	73	73	73	73	73	73	73	73	73	73	73

Table B.2: $\beta(k,t)$ with $\widehat{F}_\Theta$ for Portfolio 2 (continued).

						k					
t	0	1	2	3	4	5	6	7	8	9	10
0	100										
1	93	168	300	531	819	1084	1313	1524	1727	1928	2128
2	87	151	250	419	651	888	1096	1283	1458	1629	1799
3	82	139	219	346	531	740	932	1102	1259	1410	1558
4	78	130	197	298	445	625	802	961	1105	1241	1373
5	75	123	182	264	383	535	696	846	981	1106	1226
6	72	117	170	239	337	465	610	750	878	995	1106
7	69	112	160	220	302	411	539	669	791	902	1006
8	66	107	152	205	276	368	480	600	716	822	921
9	64	103	145	193	255	334	433	542	651	753	847
10	62	99	139	183	238	307	393	492	594	691	782
11	59	96	133	175	224	285	361	450	544	637	725
12	58	93	129	168	213	267	335	414	501	589	673
13	56	90	124	161	203	252	312	384	463	546	627
14	54	87	120	155	194	239	294	358	431	508	585
15	52	85	116	150	187	228	278	336	402	474	547
16	51	82	113	145	180	219	264	317	377	444	513
17	50	80	110	141	174	210	252	300	356	418	483
18	48	78	107	137	168	203	242	286	337	394	455
19	47	76	104	133	163	196	232	274	321	373	431
20	46	74	101	129	158	190	224	262	306	355	409
21	45	72	99	126	154	184	216	253	293	339	389
22	43	71	96	123	150	179	210	244	282	324	371
23	42	69	94	120	146	174	203	236	271	311	355
24	41	67	92	117	143	169	198	228	262	299	340
25	40	66	90	114	139	165	192	221	253	288	327
26	40	64	88	112	136	161	187	215	246	279	315
27	39	63	86	109	133	157	183	210	238	270	304
28	38	62	84	107	130	154	178	204	232	262	294
29	37	60	83	105	127	150	174	199	226	254	285
30	36	59	81	103	125	147	170	194	220	247	277
31	35	58	79	101	122	144	167	190	215	241	269
32	35	57	78	99	120	141	163	186	210	235	262
33	34	56	76	97	118	138	160	182	205	229	255
34	33	55	75	95	115	136	157	178	201	224	249
35	33	54	74	93	113	133	154	175	196	219	243
36	32	53	72	92	111	131	151	171	192	214	238
37	31	52	71	90	109	128	148	168	189	210	232
38	31	51	70	89	107	126	145	165	185	206	228
39	30	50	69	87	105	124	143	162	182	202	223
40	30	49	68	86	104	122	140	159	178	198	219

Table B.3: $\beta(k,t)$ with $\widetilde{F}_\Theta$ for Portfolio 2.

	k										
t	0	1	2	3	4	5	6	7	8	9	10
41	29	49	67	84	102	120	138	156	175	194	214
42	29	48	65	83	100	118	136	154	172	191	210
43	28	47	64	82	99	116	134	151	169	188	207
44	28	46	63	80	97	114	131	149	166	184	203
45	27	46	62	79	96	112	129	146	164	181	200
46	27	45	62	78	94	111	127	144	161	178	196
47	26	44	61	77	93	109	125	142	159	176	193
48	26	44	60	76	92	108	124	140	156	173	190
49	26	43	59	75	90	106	122	138	154	170	187
50	25	42	58	74	89	104	120	136	152	168	184
51	25	42	57	73	88	103	118	134	149	165	181
52	24	41	56	72	87	102	117	132	147	163	179
53	24	41	56	71	85	100	115	130	145	160	176
54	24	40	55	70	84	99	114	128	143	158	173
55	23	40	54	69	83	98	112	127	141	156	171
56	23	39	54	68	82	96	111	125	139	154	169
57	23	38	53	67	81	95	109	123	138	152	166
58	22	38	52	66	80	94	108	122	136	150	164
59	22	37	52	65	79	93	106	120	134	148	162
60	22	37	51	64	78	91	105	119	132	146	160
61	21	37	50	64	77	90	104	117	131	144	158
62	21	36	50	63	76	89	102	116	129	142	156
63	21	36	49	62	75	88	101	114	127	141	154
64	21	35	48	61	74	87	100	113	126	139	152
65	20	35	48	61	73	86	99	112	124	137	150
66	20	34	47	60	73	85	98	110	123	136	149
67	20	34	47	59	72	84	97	109	122	134	147
68	20	34	46	59	71	83	96	108	120	133	145
69	19	33	46	58	70	82	94	107	119	131	144
70	19	33	45	57	69	81	93	105	118	130	142
71	19	32	45	57	69	81	92	104	116	128	140

Table B.4: $\beta(k,t)$ with $\widetilde{F}_{\Theta}$ for Portfolio 2 (continued).

References

[1] Albrecht, P. (1982). On some statistical methods connected with the mixed Poisson process. *Scandinavian Actuarial Journal*, 1-14.

[2] Albrecht, P. (1984). Laplace transforms, Mellin tranforms and mixed Poisson processes. *Scandinavian Actuarial Journal*, 58-64.

[3] Carrière, J. (1993a). Nonparametric tests for mixed Poisson distributions. *Insurance: Mathematics & Economics* **12**, 3-8.

[4] Carrière, J. (1993b). A semiparametric estimator of a risk distribution. *Insurance: Mathematics & Economics* **13**, 75-81.

[5] Coene, G., and Doray, L.G. (1996). A financially balanced Bonus-Malus system. *ASTIN Bulletin* **26**, 107-116.

[6] Denuit, M. (1997). A new distribution of Poisson-type for the number of claims. *ASTIN Bulletin* **27**, 229-242.

[7] Grenander, U. (1957a). On the heterogeneity in non-life insurance. *Scandinavian Actuarial Journal*, 153-179.

[8] Grenander, U. (1957b). Some remarks on bonus systems in automobile insurance. *Scandinavian Actuarial Journal*, 180-197.

[9] Gossiaux, A.-M. and Lemaire, J. (1981). Méthodes d'ajustement de distributions de sinistres. *Bulletin of the Swiss Association of Actuaries*, 87-95.

[10] Kestemont, R.-M., and Paris, J. (1985). Sur l'ajustement du nombre de sinistres. *Bulletin of the Swiss Association of Actuaries*, 157-163.

[11] Lemaire, J. (1995). *Bonus-Malus Systems in Automobile Insurance*. Kluwer Academic Publisher, Boston.

[12] Lindsay, B. G. (1989a). On the determinants of moment matrices. *Annals of Statistics* **17**, 711–721.

[13] Lindsay, B. G. (1989b). Moment matrices: applications in mixtures. *Annals of Statistics* **17**, 722–740.

[14] Patilea, V., and Rolin, J.-M. (1995). Smooth vs. likelihood estimation for a class of discrete distributions. *Discussion Paper* **9516**, Insitut de Statistique, Université Catholique de Louvain, Louvain-la-Neuve, Belgium.

[15] Seal, H.L. (1982). Mixed Poisson - an ideal distribution of claim number ? *Bulletin of the Swiss Association of Actuaries*, 293-295.

[16] Shaked, M. (1980). On mixtures from exponential families. *Journal of the Royal Statistical Society – Series B* **42**, 192-198.

[17] Simar, L. (1976). Maximum likelihood estimation of a compound Poisson process. *The Annals of Statistics* **4**, 1200-1209.

[18] Tremblay, L. (1992). Using the Poisson-Inverse Gaussian distribution in Bonus-Malus Systems. *ASTIN Bulletin* **22**, 97-106.

[19] Tucker, H.G. (1963). An estimate of the compounding distribution of a compound Poisson distribution. *Theory of Probability and Its Applications* **8**, 195-200.

[20] Walhin, J.-F., and Paris, J. (1999). Using mixed Poisson distributions in connection with Bonus-Malus systems. *ASTIN Bulletin* **29**, 81-99.

[21] Willmot, G.E. (1987). The Poisson-Inverse Gaussian distribution as an alternative to the Negative Binomial. *Scandinavian Actuarial Journal*, 113-127.