Hacettepe Journal of Mathematics and Statistics
Volume 43 (6) (2014), 1047 – 1061

Computing credibility Bonus-Malus premiums using the total claim amount distribution

Emilio Gómez-Déniz* , Agustín Hernández-Bastida† and M. Pilar Fernández-Sánchez‡

Abstract

Assuming a bivariate prior distribution for the two risk parameters appearing in the distribution of the total claim amount when the primary distribution is geometric and the secondary one is exponential, we derive Bayesian premiums which can be written as credibility formulas. These expressions can be used to compute bonus-malus premiums based on the distribution of the total claim amount but not for the claims which produce the amounts. The methodology proposed is easy to perform, and the maximum likelihood method is used to compute the bonus-malus premiums for a real set of automobile insurance data, one that is well known in actuarial literature.

Keywords: Automobile Insurance, Bonus-Malus; Bivariate Distribution, Credibility; Premium

Received 01 : 08 : 2013 : Accepted 27 : 11 : 2013

1. Introduction

Bayesian methods have been successfully applied in actuarial statistics, and have proved to be a good tool for resolving problems related to credibility theory and the setting of insurance premiums. Given a risk group, it is usual to assume that the level of risk of each policy is represented by a risk parameter, or risk profile. It is also assumed that across the group there exists a random variable whose realizations are the values of the risk parameter for policies belonging to that group; its distribution or density function is called the prior distribution or structure function. Most automobile insurance schemes employ Bonus-Malus Systems (BMS). In this context, there is a finite number of classes and the premium applicable depends on the class to which the policyholder belongs. In each period (usually a year), a policyholder's class is determined on the basis

*Department of Quantitative Methods in Economics. University of Las Palmas de Gran Canaria, Spain.
Email: egomez@dmc.ulpgc.es
†Department of Quantitative Methods. University of Granada, Spain.
‡Department of Quantitative Methods. University of Granada, Spain.

of that assigned for the previous period and on the number of claims made during the period. The main purpose of a BMS is to decrease the premiums for good risks and to increase them for bad ones. With Bayesian methodology, this is achieved by dividing a posterior expectation by a prior expectation according to an estimate derived by means of an appropriate loss function (see Lemaire, 1979, 1985, 1995; Gómez-Déniz et al., 2002; Sarabia et al., 2004; Denuit et al., 2007, among others).

Nevertheless, it is obvious that not all accidents produce the same individual claim size and thus it does not seem fair to penalize all policyholders in the same way when they present a claim. In other words, when the bonus-malus premium is based only on the number of claims a policyholder who has an accident with an individual claim size of 100$ is penalized by the same amount as if the accident had produced an individual claim size of 500$. As different claims produce different claim amounts, it would seem that the best way to build a BMS would be based on both the number of claims and on the individual claim size. As Lemaire (2004) points out, when the claim amount is not incorporated into the bonus-malus premium this implies an assumption of independence between the variables "number of claims" and "claim amount", an assumption that is open to question.

In recent years, attempts have been made to include factors other than the number of claims in calculating bonus-malus premiums. Thus, Frangos and Vrontos (2001) and Mert and Saykan (2005) introduced a model where the number of claims and the individual claim size were used jointly to compute the bonus-malus premiums. Based on the independence assumption assumed in the collective risk model between these two random variables, they computed the premium by multiplying the bonus-malus premiums based only on the number of claims with the bonus-malus premiums based only on the individual claim size. Their empirical results show there is a positive correlation between these two random variables, and thus the assumption of some kind of dependence between them should be taken into account in calculating bonus-malus premiums.

If we wish to replace the distribution of the number of claims by the distribution of the total claim amount, this will depend on two parameters, one related to the random variable "number of claims" and the other related to the random variable "claim cost". It is possible to transfer a relation of dependence between these risk profiles, by assuming that both profiles fit a joint bivariate prior distribution.

Apart from looking for some kind of dependence, the bivariate prior distribution can be justified in the following manner. The aim of the actuary is to design a tariff system that will distribute the exact weight of each risk fairly within the portfolio when policyholders present different risks. For instance, in the automobile insurance market, the first approach to solving this problem, called tariff segmentation, consists in dividing policyholders into homogeneous classes according to certain variables believed to be influencing factors (a priori factors), such as the model and use of the car, the age and sex of the driver, the duration of the driving licence, etc. Once the actuary has classified policyholders, the premium can be established for each type of risk. However, some factors cannot be measured or introduced into the rates to calculate premiums according to tariff-segmentation methods. Consequently, heterogeneity continues to exist in every class defined with a priori factors. Some of these unmeasured or unknown characteristics probably have a significant effect on the number of claims and also on the individual claim size; for instance, in automobile insurance, swiftness of reflexes, knowledge of the Highway Code or the behaviour patterns of the driver. Given that many claims could be explained by these hidden features, they should be included in the tariff system. This is the goal of experience rating or credibility theory, the underlying idea of which is that past experience reveals information about hidden features.

In this paper, we assume a bivariate prior distribution for the two risk parameters appearing in the distribution of the total claim amount when the distribution of the random variable number of claims (primary distribution) is geometric and the distribution of the random variable individual claim size (secondary distribution) is exponential. This allows us to use the unconditional distribution of the total claim amount to compute the premiums, which can then be written as a credibility formula. These premiums are then used to obtain the bonus-malus premiums, based on the distribution of the total claim amount and not only on the claims which produced the amounts. The maximum likelihood method is used to estimate the parameters of the distribution in a real data set concerning automobile insurance and well known in actuarial literature.

The rest of this paper is structured as follows. Section **??** presents the basic collective risk model based on the geometric and the exponential distribution as the primary and the secondary distribution, respectively. The bivariate prior distribution is presented in Section **??**, where we also show the marginal and the unconditional distribution of the claim size. Credibility premiums are obtained in Section **??** and the parameters of the unconditional distribution of the claim amount are estimated in Section **??**. A numerical application with a real data set is presented in Section **??** and the main conclusions are drawn in the last Section.

2. The basic model

One of the main objectives of risk theory is to model the distribution of the aggregate claim amount for portfolios of policies, so that the insurance firm can take decisions taking into account just two aspects of the insurance business: the number of claims and the individual claim size. Therefore, the total claim amount over a fixed time period is modelled by considering the number of claims and the individual claim size separately. In this paper, we assume that the premiums in a bonus-malus system should be computed by taking into account both the number of claims and the individual claim size.

In the collective risk theory, the random variable of interest is the aggregate claim defined by $X = \sum_{i=1}^{N} X_i$, where N is the random variable denoting the number of claims and X_i, for $i = 1, 2, ...$ is the random variable denoting the individual claim size of the i-th claim. Assuming that $X_1, X_2, \dots,$ are independent and identically distributed random variables which are also independent of the random variable number of claims N, it is well-known (see Klugman et al. (2008) and Rolski et al. (1999), among others) that the probability density function of the aggregate claim (total claim amount) is given by $f_X(x) = \sum_{n=0}^{\infty} p_n f^{n^*}(x)$, where p_n denotes the probability of n claims (primary distribution) and $f^{n^*}(x)$ is the n-th fold convolution of $f(x)$, the probability density function of the claim amount (secondary distribution).

In automobile insurance, when the portfolio is considered to be heterogeneous, all policyholders have a constant but unequal underlying risk of having an accident. That is, the expected number of claims varies from policyholder to policyholder. As the mixed Poisson distributions have thicker tails than the Poisson distribution, the former provide a good fit to claim frequency data when the portfolio is heterogeneous. Frangos and Vrontos (2001), Gómez-Déniz (2002), Mert and Saykan (2005), among many others, consider the Poisson parameter, i.e. the expected number of claims, to follow a Gamma distribution. In this case, the unconditional distribution of the number of claims follows a negative binomial distribution. The advantage of this model is that the distribution of the total claim amount can be obtained in closed form expression when the secondary distribution is assumed to be exponential. Other models considered in the actuarial literature are the Poisson-inverse Gaussian distribution (Willmot (1987)) and the negative binomial-inverse Gaussian distribution (Gómez-Déniz et al. (2008)). Both models provide a good fit to the claim frequency data, and a recursive computation of the total claim amount can be obtained using Panjer's algorithm or a simple modification.

Assuming that the number of claims is represented by random variable N and that it follows a Poisson distribution with parameter $\lambda > 0$ denoting the differing underlying risk of each policyholder reporting a claim. Assume, moreover, that λ is distributed according to the exponential distribution with parameter $\theta_1/(1-\theta_1)$, with $0 < \theta_1 < 1$, i.e. $\pi(\lambda) \propto \exp\left(-\frac{\theta_1 \lambda}{1-\theta_1}\right)$, where $\pi(\lambda)$ represents the prior distribution of λ. It is a simple exercise to show that the unconditional distribution of the number of claims is given by

$$\Pr(N = n) = \theta_1(1-\theta_1)^n, \quad n = 0, 1, \ldots,$$

and therefore a geometric distribution with parameter θ_1.

Assuming that the individual claim size follows an exponential distribution (secondary distribution) with parameter $\theta_2 > 0$, the n-th fold convolution of exponential distribution has a closed form that is given as follows (see Klugman et al. (2008) and Rolski et al. (1999))

$$f^{*n}(x) = \frac{\theta_2^n}{(n-1)!} x^{n-1} e^{-\theta_2 x}, \; n = 1, 2, \ldots$$

i.e. it is a gamma distribution with shape parameter n and scale parameter θ_2. Now, it is easy to see that the probability density function of the random variable $X = \sum_{i=1}^{N} X_i$ is given by

$$f_X(x|\theta_1, \theta_2) = \begin{cases} \theta_1, & x = 0, \\ \theta_1(1-\theta_1)\theta_2 \exp(-\theta_1\theta_2 x), & x > 0. \end{cases} \qquad (2.1)$$

Observe that the probability density function of the claim amount has a jump of size θ_1 at the origin.

3. A suitable bivariate distribution

In this section we introduce a new continuous probability density function that will be used to derive, by mixing, the unconditional probability density function of the total claim amount in (**??**) and also to compute the bonus-malus premiums proposed in this paper.

We begin by introducing the new continuous bivariate probability density function, as follows. It can be shown straightforwardly that

$$f(x, y) = \frac{\sigma^\gamma}{B(\alpha-\gamma, \beta)\Gamma(\gamma)} x^{\alpha-1}(1-x)^{\beta-1} y^{\gamma-1} \exp(-\sigma x y), \qquad (3.1)$$

for $0 < x < 1,\; y > 0,\; \alpha > 0,\; \beta > 0, \gamma > 0, \sigma > 0$ and $\alpha > \gamma$ is a proper bivariate probability density function. In (**??**) we have that

$$\Gamma(z) = \int_0^\infty t^{z-1} e^{-t}\, dt$$

is the gamma function and $B(z_1, z_2)$ is the beta function given by

$$B(z_1, z_2) = \int_0^1 t^{z_1-1}(1-t)^{z_2-1}\, dt.$$

To the best of our knowledge, the bivariate distribution presented here has not been previously addressed in statistical literature.

Some computations provide that the distribution is unimodal with modal value at the point

$$\begin{aligned} x &= \frac{\alpha-\gamma}{\alpha+\beta-\gamma-1}, \\ y &= \frac{(\gamma-1)(\alpha+\beta-\gamma-1)}{\sigma(\alpha-\gamma)}. \end{aligned}$$

Now by a straightforward calculation we see that

$$E(XY) = \frac{\gamma}{\sigma}. \tag{3.2}$$

The marginal distribution of X and Y, which can be obtained by integrating (**??**) with respect to y and x, respectively, can be shown to be a known univariate distribution. Thus, the marginal distribution of X is a beta distribution with parameters $\alpha-\gamma$ and β, i.e.

$$f_X(x) = \frac{1}{B(\alpha-\gamma,\beta)} x^{\alpha-\gamma-1}(1-x)^{\beta-1}. \tag{3.3}$$

The marginal distribution of Y is given by

$$f_Y(y) = \frac{\sigma^\gamma \Gamma(\alpha+\beta-\gamma)}{\Gamma(\alpha+\beta)B(\alpha-\gamma,\gamma)} y^{\gamma-1}\, {}_1F_1(\alpha,\alpha+\beta,-\sigma y), \tag{3.4}$$

where ${}_1F_1(\cdot,\cdot,\cdot)$ is the confluent hypergeometric function, also called Kummer's function, given by

$${}_1F_1(m,n,z) = \sum_{k=0}^{\infty} \frac{(m)_k z^k}{(n)_k k!},$$

and $(m)_j = \Gamma(m+j)/\Gamma(m)$, $j \geq 1$, $(m)_0 = 1$ is the Pochhammer symbol.

Using Kummer's first theorem we have that (**??**) can be rewritten as

$$f_Y(y) = \frac{\sigma^\gamma \Gamma(\alpha+\beta-\gamma)}{\Gamma(\alpha+\beta)B(\alpha-\gamma,\gamma)} y^{\gamma-1} e^{-\sigma y}\, {}_1F_1(\beta,\alpha+\beta,\sigma y). \tag{3.5}$$

Expression (**??**) is reminiscent of the generalized exponential distribution given in Bhattacharya (1966), see expression (3.1) in this paper.

Since

$$E_Y(Y) = \frac{\gamma(\alpha+\beta-\gamma-1)}{\sigma(\alpha-\gamma-1)}, \tag{3.6}$$

by using (**??**) together with (**??**) and the mean of the distribution given in (**??**) we obtain the covariance of (**??**), which is given by

$$cov(X,Y) = \frac{\beta\gamma}{\sigma(\alpha+\beta-\gamma)(\gamma-\alpha+1)}, \tag{3.7}$$

which admits correlation of any sign. Thus, we have

$$cov(X,Y) \begin{cases} > 0 & \text{if } 0 < \alpha-\gamma < 1, \\ > 0 & \text{if } \alpha-\gamma > 1. \end{cases}$$

Let us now assume that the two parameters of the distribution of the total claim amount in (**??**) are random and that they follow a bivariate prior distribution as in (**??**), i.e. we have that

$$\pi(\theta_1,\theta_2)=\frac{\sigma^{\gamma}}{B(\alpha-\gamma,\beta)\Gamma(\gamma)}\theta_1^{\alpha-1}(1-\theta_1)^{\beta-1}\theta_2^{\gamma-1}\exp(-\sigma\theta_1\theta_2), \tag{3.8}$$

for $0<\theta_1<1$, $\theta_2>0$, $\alpha>0$, $\beta>0, \gamma>0, \sigma>0$ and $\alpha>\gamma$. The unconditional distribution of the total claim amount in (**??**) can be obtained by mixing, by computing the following integral

$$f_X(x|\alpha,\beta,\gamma,\sigma)=\int_0^{\infty}\int_0^1 f_X(x|\theta_1,\theta_2)\pi(\theta_1,\theta_2)\,d\theta_1\,d\theta_2.$$

Some algebra provides the following probability density function for the unconditional distribution of the total claim amount.

$$f(x|\alpha,\beta,\gamma,\sigma)=\begin{cases}\dfrac{\alpha-\gamma}{\alpha+\beta-\gamma}, & x=0,\\[2ex] \dfrac{\beta\gamma\sigma^{\gamma}}{\alpha+\beta-\gamma}\dfrac{1}{(x+\sigma)^{\gamma+1}}, & x>0,\end{cases} \tag{3.9}$$

which is obviously a two piece distribution with a jump of size $\frac{\alpha-\gamma}{\alpha+\beta-\gamma}$ at the origin.

Moments of order r of (**??**) are as follows:

$$E(X^r)=\frac{\beta\sigma^r r!\Gamma(\gamma-r)}{(\alpha+\beta-\gamma)\Gamma(\gamma)}, \quad \gamma>r. \tag{3.10}$$

In particular, we have that

$$\begin{aligned} E(X) &= \frac{\beta\sigma}{(\alpha+\beta-\gamma)(\gamma-1)}, \quad \gamma>1,\\ E(X^2) &= \frac{2\beta\sigma^2}{(\alpha+\beta-\gamma)(\gamma-1)(\gamma-2)}, \quad \gamma>2,\end{aligned}$$

from which we can obtain the variance of the distribution, given by

$$var(X)=\frac{\beta\sigma^2(2\alpha(\gamma-1)+\beta-2(\gamma-1)\gamma)}{(\alpha+\beta-\gamma)^2(\gamma-2)(\gamma-1)^2}, \quad \gamma>2.$$

4. Credibility premiums

When the premiums are based only on the number of claims, the distribution to be considered is, in this case, the geometric distribution with parameter $0<\theta_1<1$. Suppose now that the prior distribution on θ_1 is the beta distribution given in (**??**). Given a sample information $n_1,\dots,n_t$ the posterior distribution is a beta distribution with parameters $\alpha-\gamma+t$ and $\beta+t\bar{n}$ and it is simple to see that the unconditional distribution of the number of claims is given by

$$\Pr(N=n)=\frac{B(\alpha-\gamma+1,\beta+n)}{B(\alpha-\gamma,\beta)}, \tag{4.1}$$

which is the geometric–beta distribution.

In the actuarial context, the premium charged to a policyholder is computed on the basis of the past claims made and on that of the accumulated past claims of the corresponding portfolio of policyholders. To obtain an appropriate formula for this, various methods have been proposed, mostly in the field of Bayesian decision methodology. The procedure for premium calculation is modelled as follows. The number of claims made with respect to a given contract in a given period is specified by a random variable X following a probability density function $f(x|\theta)$ depending on an unknown risk parameter θ. A premium calculation principle (Gómez-Déniz et al. (2006) and Heilmann (1989)) assigns to each risk parameter θ a premium within the set $P \in \mathbb{R}$, the action space. Let $L : \Theta \times P \to \mathbb{R}$ be a loss function that assigns to any $(\theta, P) \in \Theta \times P$ the loss sustained by a decision-maker who takes the action P and is faced with the outcome θ of a random experience. The premium must be determined such that the expected loss is minimized.

From this parameter, the unknown premium $\mu(\theta)$, called the risk premium, can be obtained by minimizing the expected loss $E_f\left[L(\theta, P)\right]$. L is usually taken as the weighted squared-error loss function, i.e. $L(a, x) = h(x)(x - a)^2$. Using different functional forms for $h(x)$ different premium principles are obtained. For example, for $h(x) = 1$ we obtain the net premium principle (Heilmann (1989), Gerber (1979) and Klugman et al. (2008); among others). For a review of the net premium and the different premiums defined in the actuarial setting, see Bühlmann and Gisler (2005), Gerber (1979), Gómez et al. (2002, 2006), Heilmann (1989) and Rolski et al. (1999).

If experience is not available, the actuary computes the collective premium, μ, which is given by minimizing the risk function, i.e. minimizing $E_\pi\left[L(\mu(\theta), \theta)\right]$, where $\pi(\theta)$ is the prior distribution on the unknown parameter θ. On the other hand, if experience is available, the actuary takes a sample $\mathbf{x}$ from the random variables $X_i,\ i = 1, 2, \ldots, t$, assuming X_i i.i.d., and uses this information to estimate the unknown risk premium $\mu(\theta)$, through the Bayes premium μ^*, obtained by minimizing the Bayes risk, i.e. minimizing $E_{\pi_{\mathbf{x}}}\left[L(\mu(\theta), \theta)\right]$. Here, $\pi_{\mathbf{x}}$ is the posterior distribution of the risk parameter, θ, given the sample information $\mathbf{x}$.

Thus, in our case, if $L(x, a) = (x - a)^2$, the net risk, collective and Bayes premiums are given by

$$\begin{aligned}
\mu_C(\theta_1) &= \frac{1-\theta_1}{\theta_1}, \\
\mu_C &= \frac{\beta}{\alpha-\gamma-1}, \\
\mu_C^* &= \frac{\beta+k}{\alpha-\gamma+t-1} = Z(t)\bar{n} + (1 - Z(t))\mu_C, \qquad (4.2)
\end{aligned}$$

where the credibility factor is given by $Z(t) = t/(\alpha - \gamma + t - 1)$ and $\bar{n}$ is the sample mean based only on the claim frequency observed. The subscript C indicates that the premiums are based on the number of claims.

Now suppose that the practitioner chooses to compute the premium according to the individual claim size, assuming that this follows an exponential distribution with parameter $\theta_2 > 0$ and that the prior distribution on θ_2 is the distribution given in (**??**) with $\beta \to 0$ and $\alpha = 1$. In this case, (**??**) reduces to the gamma distribution with shape parameter γ and scale parameter σ. Given a sample information $x_1, \ldots, x_t$ the posterior distribution is a gamma distribution with parameters $\gamma + t$ and $\sigma + t\bar{x}$. Again, it is simple to see that the unconditional distribution of the individual claim size is given by

$$f(x) = \frac{\gamma\sigma^\gamma}{(\sigma + x)^{\gamma+1}},$$

and that the risk, collective and Bayes premiums are as follows:

$$\begin{aligned}
\mu_{CC}(\theta_2) &= \frac{1}{\theta_2}, \\
\mu_{CC} &= \frac{\sigma}{\gamma}, \\
\mu_{CC}^* &= \frac{\sigma + t\bar{x}}{\gamma + t} = Z(t)\bar{x} + (1 - Z(t))\mu, \qquad (4.3)
\end{aligned}$$

where the credibility factor is given by $Z(t)=t/(\gamma+t)$ and $\bar{x}$ is the sample mean based only on the size observed. The subscript CC denotes that the premiums are based on the individual claim size.

Frangos and Vrontos (2001) and Mert and Saykan (2005) introduced a model where the number of claims and the individual claim size were used jointly to compute the bonus-malus premiums. According to the independence assumption assumed in the collective risk model between the two random variables, they computed the premium by multiplying the bonus-malus premiums based only on the number of claims by the bonus-malus premiums based only on the individual claim size, i.e. multiplying (**??**) by (**??**).

The most reasonable model consists in working with both random variables but not in a separate way. To do so, let x_i, $i=1,2,\ldots,t$ be independent and identically distributed random variables following the probability density function (**??**), i.e.

$$f(x_1,\ldots,x_t)=\theta_1^t\theta_2^t(1-\theta_1)^t\exp(-t\bar{x}\theta_1\theta_2),$$

provided that $x_i>0,\ i=1,2,\ldots,t$. Let us suppose that (θ_1,θ_2) follows the prior distribution $\pi(\theta_1,\theta_2)$ given in (**??**), then the posterior distribution of (θ_1,θ_2) given the sample information $(x_1,\ldots,x_t)$ is of the same form as in (**??**) with the updated parameters $(\alpha^*,\beta^*,\gamma^*,\sigma^*)$ given by

$$\begin{aligned}\alpha^* &= \alpha+t,\\ \beta^* &= \beta+t,\\ \gamma^* &= \gamma+t,\\ \sigma^* &= \sigma+\kappa,\end{aligned}$$

where $\kappa=\sum_{i=1}^t x_i$.

When $x_i=0,\ i=1,2,\ldots,t$ then the posterior distribution has the following updated parameters

$$\begin{aligned}\alpha^* &= \alpha+t,\\ \beta^* &= \beta,\\ \gamma^* &= \gamma,\\ \sigma^* &= \sigma,\end{aligned}$$

Now, denoting the unknown risk premium by $\mu(\theta_1,\theta_2)=\mu(\Theta)$, and again using the square–error loss function, the net risk, collective and Bayes premiums are given by

$$\mu(\Theta) = \int xf(x|\Theta)dx, \tag{4.4}$$

$$\mu = \int \mu(\Theta)\pi(\Theta)d\Theta, \tag{4.5}$$

$$\mu^* = \int \mu(\Theta)\pi(\Theta|\underline{x})d\Theta, \tag{4.6}$$

respectively.

As mentioned above, it is clear that under the model Assumed, the net risk premium (**??**) is given by

$$\mu(\Theta)=E(X|\Theta)=\frac{1-\theta_1}{\theta_1\theta_2}$$

while the net collective premium in (**??**) is described by

$$\mu = \int_0^\infty\int_0^1 \mu(\Theta)\pi(\theta_1,\theta_2)\,d\theta_1\,d\theta_2=\frac{\beta\sigma}{(\alpha+\beta-\gamma)(\gamma-1)},\quad \gamma>1. \tag{4.7}$$

Finally, the net Bayes premium in (**??**) is given by

$$\mu^* = \frac{(\beta+t)(\sigma+t\bar{x})}{(\alpha+\beta-\gamma+t)(\gamma+t-1)}, \quad \gamma > 1, \tag{4.8}$$

for $x_i > 0,\ i = 1, 2, \ldots, t$. When $x_i = 0,\ i = 1, 2, \ldots, t$, the net Bayes premium is

$$\mu^* = \frac{\beta\sigma}{(\alpha+\beta-\gamma+t)(\gamma-1)}, \quad \gamma > 1, \tag{4.9}$$

Observe that (**??**) can be rewritten in the two following ways. Firstly, some simple computations provide that

$$\mu^* = \mathcal{H}(\alpha, \beta, \gamma, t)\mu^*_{CC},$$

where

$$\mathcal{H}(\alpha, \beta, \gamma, t) = \frac{(\beta+t)(\gamma+t)}{(\alpha+\beta-\gamma+t)(\gamma+t-1)}.$$

And secondly,

$$\mu^* = Z(t)h_1(\bar{x}) + (1 - Z(t))h_2(\mu),$$

where the credibility factor $Z(t)$ is given by

$$Z(t) = \frac{t}{t+\alpha+\beta-\gamma} \tag{4.10}$$

and the functions $h_1(\cdot)$ and $h_2(\cdot)$ are given by

$$\begin{aligned} h_1(x) &= \frac{(\beta+t)x+\sigma}{\gamma+t-1}, \\ h_2(x) &= \frac{\gamma-1}{\gamma+t-1}x. \end{aligned}$$

When $t \to \infty$ we have that, since $h_1(\bar{x}) \to \bar{x}$ and $Z(t) \to 1$, then $\mu^* \to \bar{x}$ and when $t \to 0$ it is easy to see that $\mu* \to \mu$. Thus, it is reasonable to assume that when the sample size tends to infinity the Bayes premium converges to the sample mean, and it converges to the collective premium when the sample size tends to zero.

The credibility factor for expression (**??**) is as in (**??**) where now $h_1(x) = 0$ and $h_2(x) = (\gamma - 1)x$.

5. Inference

Moment estimators can be obtained by equating the sample moments to the population moments in (**??**). Furthermore, the parameters of the unconditional distribution of the total claim amount can be estimated via maximum likelihood. To do so, consider a random sample $\{x_1, x_2, \ldots, x_t\}$. The likelihood function can be written as

$$f(x_1, \ldots, x_t|\alpha, \beta, \gamma, \sigma) = \left(\frac{\alpha-\gamma}{\alpha+\beta-\gamma}\right)^{t_0} \left(\frac{\beta\gamma\sigma^\gamma}{\alpha+\beta-\gamma}\right)^{t^*} \prod_{x_i>0} \frac{1}{(x_i+\sigma)^{\gamma+1}}, \tag{5.1}$$

where t_0 is the number of zero-observations and $t^* = t - t_0$ is the number of non-zero sample observations, where t is the sample size. Finally $\prod_{x_i>0}$ denotes the product over the t^* non-zero observations.

Taking logarithms of (**??**) we have the log-likelihood equation given by

$$\begin{aligned} \ell \equiv \ell(\alpha,\beta,\gamma,\sigma|x_1,\ldots,x_t) &= t_0 \log\left(\frac{\alpha-\gamma}{\alpha+\beta-\gamma}\right) + t^* \log\left(\frac{\beta\gamma\sigma^\gamma}{\alpha+\beta-\gamma}\right) \\ &- (\gamma+1)\sum_{x_i>0}\log(x_i+\sigma). \end{aligned} \tag{5.2}$$

Now, differentiating (**??**) with respect to the four parameters in turn, and equating to zero, we obtain the maximum likelihood estimating equations given by

$$\frac{\partial \ell}{\partial \alpha} = \frac{t_0}{\alpha-\gamma} - \frac{t}{\alpha+\beta-\gamma} = 0, \tag{5.3}$$

$$\frac{\partial \ell}{\partial \beta} = \frac{t^*}{\beta} - \frac{t}{\alpha+\beta-\gamma} = 0, \tag{5.4}$$

$$\frac{\partial \ell}{\partial \gamma} = \frac{t_0}{\alpha-\gamma} - \frac{t^*}{\gamma} + \frac{t}{\alpha+\beta-\gamma} - t^*\log\sigma + \sum_{x_i>0}\log(x_i+\sigma) = 0, \tag{5.5}$$

$$\frac{\partial \ell}{\partial \sigma} = \frac{t^*}{\sigma} - (\gamma+1)\sum_{x_i>0}\frac{1}{x_i+\sigma} = 0. \tag{5.6}$$

The second partial derivatives are as follows:

$$\begin{aligned} \frac{\partial^2 \ell}{\partial \alpha^2} &= \frac{t}{(\alpha+\beta-\gamma)^2} - \frac{t_0}{(\alpha-\gamma)^2}, \quad \frac{\partial^2 \ell}{\partial \alpha \beta} = \frac{t}{(\alpha+\beta-\gamma)^2}, \\ \frac{\partial^2 \ell}{\partial \alpha \partial \gamma} &= \frac{t_0}{(\alpha-\gamma)^2} - \frac{t}{(\alpha+\beta-\gamma)^2}, \quad \frac{\partial^2 \ell}{\partial \alpha \partial \sigma} = 0, \\ \frac{\partial^2 \ell}{\partial \beta^2} &= \frac{t}{(\alpha+\beta-\gamma)^2} - \frac{t^*}{\beta^2}, \quad \frac{\partial^2 \ell}{\partial \beta \gamma} = -\frac{t}{(\alpha+\beta-\gamma)^2}, \quad \frac{\partial^2 \ell}{\partial \beta \partial \sigma} = 0, \\ \frac{\partial^2 \ell}{\partial \gamma^2} &= \frac{t_0}{(\alpha-\gamma)^2} + \frac{t^*}{\gamma^2} + \frac{t}{(\alpha+\beta-\gamma)^2}, \quad \frac{\partial^2 \ell}{\partial \gamma \partial \sigma} = \sum_{x_i>0}\frac{1}{x_i+\sigma}, \\ \frac{\partial^2 \ell}{\partial \sigma^2} &= -\frac{t^*}{\sigma^2} + (\gamma+1)\sum_{x_i>0}\frac{1}{(x_i+\sigma)^2}. \end{aligned}$$

It is easy to see that

$$\begin{aligned} E\left(\frac{1}{X+\sigma}\right) &= \frac{\alpha+(-1+\alpha+\beta)\gamma-\gamma^2}{(\alpha+\beta-\gamma)(1+\gamma)\sigma}, \\ E\left[\frac{1}{(X+\sigma)^2}\right] &= \frac{(-2+\beta-\gamma)\gamma+\alpha(2+\gamma)}{(\alpha+\beta-\gamma)(2+\gamma)\sigma^2}. \end{aligned}$$

Therefore, Fisher's information matrix (not reproduced here) can be obtained easily, in closed form expression.

6. An application to a real data set

In order to compare the premiums based only on the number of claims with the premiums obtained when the total claim amount distribution is used, we examined a data set based on one-year vehicle insurance policies taken out in 2004 or 2005. This data set is available on the website of the Faculty of Business and Economics, Macquarie University (Sydney, Australia), see also Jong and Heller (2008). The first 100 observations of this data set are shown in Table **??**, with the following elements: from left to right, the policy number, the number of claims and the size of the claims. The total portfolio contains 67856 policies of which 4624 have at least one claim. Some descriptive statistics for this data set are shown in Table **??**. It can be seen that the standard deviation is very large for the size of the claims, which means that a premium based only on the mean size of the claims is not adequate for computing the bonus-malus premiums. The covariance between the claims and sizes is positive and takes the value 141.574.

Table 1. First 100 observations of the data set

1	0	0	21	0	0	41	2	1811.71	61	0	0	81	0	0
2	0	0	22	0	0	42	0	0	62	0	0	82	0	0
3	0	0	23	0	0	43	0	0	63	0	0	83	0	0
4	0	0	24	0	0	44	0	0	64	0	0	84	0	0
5	0	0	25	0	0	45	0	0	65	1	5434.44	85	0	0
6	0	0	26	0	0	46	0	0	66	1	865.79	86	0	0
7	0	0	27	0	0	47	0	0	67	0	0	87	0	0
8	0	0	28	0	0	48	0	0	68	0	0	88	0	0
9	0	0	29	0	0	49	0	0	69	0	0	89	0	0
10	0	0	30	0	0	50	0	0	70	0	0	90	0	0
11	0	0	31	0	0	51	0	0	71	0	0	91	0	0
12	0	0	32	0	0	52	0	0	72	0	0	92	0	0
13	0	0	33	0	0	53	0	0	73	0	0	93	0	0
14	0	0	34	0	0	54	0	0	74	0	0	94	0	0
15	1	669.51	35	0	0	55	0	0	75	0	0	95	0	0
16	0	0	36	0	0	56	0	0	76	0	0	96	1	1105.77
17	1	806.61	37	0	0	57	0	0	77	0	0	97	0	0
18	1	401.80	38	0	0	58	0	0	78	0	0	98	0	0
19	0	0	39	0	0	59	0	0	79	0	0	99	1	200
20	0	0	40	0	0	60	0	0	80	0	0	100	0	0

Figure **??** shows the complete number of claims and the total claim amount concerning these claims. It can be seen that the larger claim values appear in the case of single claims and that these values fall with larger numbers of claims. It is probable that a first severe accident encourages the driver to be more careful, which tends to reduce the size of the claims in future accidents. For this reason, we believe the bonus-malus premiums should not be based only on the number of claims but also on their size.

LISTPLOT-eps-converted-to.pdf

Figure 1. Number of claims and their costs

We used (**??**) to estimate the α and β parameters of this distribution when only the number of claims was used, and assumed that $\alpha \equiv \alpha - \gamma$. The maximum likelihood method does not provide a solution in this case, and so $\alpha = 1/\beta$ was assumed, which produced $\widehat{\beta} = 0.2528$ and -18684.10 for the value of the maximum of the log-likelihood function. The bonus-malus premiums (BMP) are computed according to the expression

$$\text{BMP} = \frac{\beta + k}{\alpha + t - 1} \frac{\alpha - 1}{\beta},$$

Table 2. Some descriptive data of claims and claim size for the data set

	Number of claims	Total claim amount
Mean	0.072	137.27
Standard deviation	0.278	1056.30
min	0	0
max	4	55922.10

where $k = t\bar{n}$. The resulting bonus-malus premiums are shown in Table **??**.

Now, using the expressions given in (**??**), (**??**), (**??**) and (**??**) we computed the maximum likelihood estimates of the parameters when the claim amount distribution is used to compute the bonus-malus premiums. In order to simplify the computations the values of the total claim amounts have been divided by 1000. These are given by $\widehat{\alpha} = 2.4282$, $\widehat{\beta} = 0.0299$, $\widehat{\gamma} = 2.0465$ and $\widehat{\sigma} = 2.2051$. Now, the value of the maximum of the log-likelihood function is -24111.80 and the estimated value of the covariance, using expression (**??**), is 0.1072. Observe that for these estimates the net collective premium based on both, the number of claims and the individual claim size, which is given in (**??**), is provided by $1000 \times 0.153068 = 153.068$, the latter value is close to the sample mean appearing in Table **??**.

Let us now compute the bonus-malus premiums using the expression

$$\text{BMP} = \frac{(\beta + t)(\sigma + \kappa)}{(\alpha + \beta - \gamma + t)(\gamma + t - 1)} \frac{(\alpha + \beta - \gamma)(\gamma - 1)}{\beta\sigma},$$

for $\kappa = t\bar{x} > 0$. When $\kappa = 0$ the bonus-malus premiums are given by

$$\text{BMP} = \frac{\beta\sigma}{(\alpha + \beta - \gamma + t)(\gamma - 1)} \frac{(\alpha + \beta - \gamma)(\gamma - 1)}{\beta\sigma}.$$

Table **??** shows the bonus-malus premiums obtained with the aggregate model, taking into account the number of claims and the individual claim size.

Observe that the premiums based only on the number of claims and on the total claim amount have several levels of premiums, but these levels have different meanings. Although the first is based on k and the second on κ we can consider both levels used by the insurance firm to move a policyholder from one column to another, i.e. to move the policyholder from one class to another. The first column is usually termed the bonus class and the other, the malus class.

Tables **??** and **??** show that the bonus-malus premiums under the model based on the total claim amount distribution are slightly lower for the bonus class and larger for the classes $k = 1, 2$ and 3, in comparison with the bonus-malus premiums based only on the number of claims. It seems reasonable that policyholders with no claims should pay less, taking into account that those reporting claims are now going to pay more. It could be said that the new bonus-malus system is very generous to drivers in the bonus class and very strict with those in the malus classes. The drivers in the bonus class, for the first claim free year, will receive 70.94% of the basic premium, while drivers who report one accident in the first year will have to pay a malus of 695.400% of the basic premium. It might be thought that this is dangerous for the insurance firm, because most policyholders would look for another company with more competitive prices, but it should be recalled that most of the policyholders in the portfolio do not make a claim.

Table 3. Bonus-malus premiums based only on the frequency component and the net premium principle

Year	Number of claims, k					
	0	1	2	3	4	5
0	1.00000					
1	0.74712	3.70161	6.65609	9.61058	12.5651	15.5196
2	0.59632	2.95449	5.31265	7.67081	10.0290	12.3871
3	0.49617	2.45831	4.42044	6.38257	8.34470	10.3068
4	0.42483	2.10482	3.78481	5.46480	7.14480	8.82479
5	0.37142	1.84021	3.30900	4.77780	6.24659	7.71538
6	0.32994	1.63471	2.93947	4.24423	5.54899	6.85376
7	0.29679	1.47049	2.64418	3.81787	4.99156	6.16525
8	0.26970	1.33625	2.40280	3.46935	4.53589	5.60244
9	0.24714	1.22447	2.20180	3.17913	4.15646	5.13379
10	0.22806	1.12995	2.03184	2.93373	3.83561	4.73750

Table 4. Bonus-malus premiums based on the severity component and the net premium principle

Year	Total claim amount, κ					
	0	1	2	3	4	5
0	1000.00					
1	290.57	7953.99	10435.70	12917.40	15399.10	17880.80
2	169.98	6166.55	8090.54	10014.50	11938.5	13862.50
3	120.12	4898.92	6427.41	7955.90	9484.40	11012.90
4	92.88	4040.46	5301.11	6561.76	7822.41	9083.06
5	75.71	3431.31	4501.90	5572.48	6643.07	7713.66
6	63.90	2979.23	3908.77	4838.31	5767.84	6697.38
7	55.27	2631.28	3452.25	4273.22	5094.20	5915.17
8	48.70	2355.53	3090.47	3825.41	4560.35	5295.29
9	43.52	2131.79	2796.92	3462.05	4127.18	4792.30
10	39.34	1946.68	2554.05	3161.43	3768.80	4376.18

In fact, for the policyholder studied here, 93.18% of the policyholders did not report any claim.

Conditional distributions form the theoretical basis of all regression analysis and therefore it is important to examine them. The conditional distribution of $X|Y = y$ is given by

$$f_{X|Y}(x|y) = \frac{x^{\alpha-1}(1-x)^{\beta-1}\exp(-\sigma xy)}{B(\alpha,\beta)\,{}_1F_1(\alpha,\alpha+\beta,-\sigma y)}, \quad 0 < x < 1, \tag{6.1}$$

which is the confluent hypergeometric distribution with parameters α, β and σy according to Gordy (1998).

The conditional distribution of $Y|X = x$ is given by

$$f_{Y|X}(y|x) = \frac{(\sigma x)^{\gamma}}{\Gamma(\gamma)} y^{\gamma-1} \exp(-\sigma x y), \quad y > 0, \tag{6.2}$$

which is a gamma distribution with shape parameter γ and scale parameter σx.

Furthermore, some algebra on (**??**) and (**??**) provides the conditional expectations (the regression of x on y and the regression of y on x), which are given by

$$E(X|Y = y) = \frac{\alpha}{\alpha+\beta} \frac{{}_1F_1(\alpha+1, \alpha+\beta+1, -\sigma y)}{{}_1F_1(\alpha, \alpha+\beta, -\sigma y)},$$

$$E(Y|X = x) = \frac{\gamma}{\alpha x}. \tag{6.3}$$

7. Conclusions

This paper presents an optimal BMS based on both random variables, i.e., the number of claims and the individual claim size. This model was constructed using a bivariate prior distribution for the two risk profiles on which the total claim amount depends. In consequence, we obtained premiums which can be written as credibility formula and are suitable for the computation of bonus-malus premiums. These premiums appear to be of most benefit to policyholders who report claims with a low individual claim size, while a larger premium is charged to those who produce a high individual claim size. It is concluded that it is fairer to charge policyholders premiums which not only take into account the number of claims, but also the total claim amount (which depends on both the number of claims and the individual claim size).

Acknowledgements. The authors are grateful to the referees for a very careful reading of the manuscript and suggestions which improved the paper. EGD was partially supported by ECO2009–14152 (MICINN, Spain).

References

[1] Bhattacharya, S. K., (1966). Confluent hypergeometric distributions of discrete and continuous type with applications to accident proneness. *Bull. Calcutta Statist. Assoc.*, 15, 20–31.

[2] Bühlmann, H. and Gisler, A. (2005). *A Course in Credibility Theory and its Applications.* Springer, New York.

[3] Denuit, M., Marechal, X., Pitrebois, S. and Walhin, J-F. (2007). *Actuarial Modelling of Claim Counts: Risk Classification, Credibility and Bonus-Malus Systems.* Wiley.

[4] Frangos, N. and Vrontos, S. (2001). Design of optimal Bonus-Malus systems with a frequency and a severity component on an individual basis in automobile insurance. *Astin Bulletin*, 31, 1, 1–22.

[5] Gómez-Déniz, E., Pérez, J.M., Hernández, A., Vázquez, F. (2002). Measuring sensitivity in a bonus-malus system. *Insurance: Mathematics and Economics*, 31, 105–113.

[6] Gómez-Déniz, E.; Pérez, J.; and Vázquez, F. (2006). On the use of posterior regret Γ-minimax actions to obtain credibility premiums. *Insurance: Mathematics and Economics*, 42, 39–49.

[7] Gómez-Déniz, E.; Sarabia, J.M. and Calderín, E. (2008). Univariate and multivariate versions of the negative binomial inverse Gaussian distributions with applications. *Insurance: Mathematics and Economics*, 39, 115–121.

[8] Heilmann, W. (1989) Decision theoretic foundations of credibility theory. *Insurance: Mathematics and Economics* 8, 1, 77–95.

[9] Gerber, H.U. (1979). *An Introduction to Mathematical Risk Theory.* Huebner Foundation Monograph 8.

[10] Gordy, M.B. (1998). Computationally convenient distributional assumptions for common-value auctions. *Computational Economics*, 12, 61–78.

[11] Jong, P. and Heller, G. (2008). *Generalized Linear Models for Insurance Data.* Cambridge University Press.

[12] Klugman, S.A., Panjer, H.H. and Willmot, G.E. (2008). *Loss Models: From Data to Decisions. Third Edition.* Wiley.

[13] Lemaire, J. (1979). How to define a bonus-malus system with an exponential utility function. *Astin Bulletin*, **10**, 274–282.

[14] Lemaire, J. (1985). *Automobile Insurance. Actuarial Models.* Kluwer-Nijhoff Publishing, Dordrecht.

[15] Lemaire, J. (1995). *Bonus-Malus Systems in Automobile Insurance.* Kluwer Academic Publishers, London.

[16] Lemaire, L. (2004). Bonus-Malus systems. In *Encyclopedia of Actuarial Science.* Wiley.

[17] Mert, M. and Saykan, Y. (2005). Bonus-Malus system where the claim frequency distribution is geometric and the claim severity distribution is Pareto. *Hacettepe Journal of Mathematics and Statistics*, 34, 75–81.

[18] Rolski, T., Schmidli, H. Schmidt, V. and Teugel, J. (1999). *Stochastic Processes for Insurance and Finance.* John Wiley & Sons.

[19] Sarabia, J.M., Gómez-Déniz, E. and Vázquez, F.J. (2004). On the use of conditional specification models in claim count distributions: An application to bonus-malus systems. *Astin Bulletin*, 34, 85–98.