Hacettepe Journal of Mathematics and Statistics
Volume 34 (2005), 75 – 81

ON A BONUS-MALUS SYSTEM WHERE THE CLAIM FREQUENCY DISTRIBUTION IS GEOMETRIC AND THE CLAIM SEVERITY DISTRIBUTION IS PARETO

Mehmet Mert* and Yasemin Saykan*

Received 21 : 03 : 2005 : Accepted 09 : 08 : 2005

Abstract

In this paper, the classical Bonus-Malus Systems (BMS) under which a premium is set by taking into account only the number of accidents each policyholder has, is compared with an optimal BMS under which the premium is set by taking into account both the frequency and the severity of the claims of each policyholder. The number and size of the claims an insured person has are assumed to follow a Geometric distribution and a Pareto distribution, respectively.

Keywords: Net premium principle, Quadratic error loss function, Mixed distributions.

1. Introduction

A BMS usually assigns each policyholder a premium based on the number of his/her accidents irrespective of their size. Under these systems, if an insured makes a claim he moves to a category where he is required to pay a higher premium (malus), and if he does not make a claim he either stays in the same category or moves to a category where he is required to pay a lower premium (bonus).

A BMS is called optimal if the total amount of bonuses is equal to the total amount of maluses, and if each policyholder pays a premium proportional to the risk he imposes to the pool.

As there is no difference between the policyholder having an accident with a small size of loss and a big size of loss, these systems can be said to be unfair. An optimal system which takes both the frequency and severity component into account must be used to set the premium an insured will pay [2].

As the parameters are estimated using the quadratic error loss function in the following sections, this method will be discussed in Section 2. In Section 3, the number of claims

*Hacettepe University, Faculty of Science, Department of Statistics, 06532 Beytepe, Ankara, Turkey.

for a given λ is considered to follow a Poisson distribution, and the expected number is exponentially distributed, so the number of claims has a Geometric distribution. In Section 4, the claim size with a given parameter θ is considered to have an Exponential distribution, and the parameter of the Exponential distribution is modelled using the Gamma distribution. So, the claim sizes have a Pareto distribution. In Section 5, as an application, the risk premium is calculated based only on the claim frequency, and then both on the claim frequency and the claim size. In Section 6, the results obtained are summarized.

2. The Estimation of a Parameter using the Quadratic Error Loss Function

The Bayesian approach to the parameter estimation problem is to use a loss function $L(\theta, \hat{\theta})$ to measure the loss incurred by estimating the value of the parameter θ as $\hat{\theta}$. Here $\hat{\theta}$ is chosen to minimize $\mathrm{E}\left\lfloor \mathrm{L}(\theta, \hat{\theta}) \right\rfloor$, where this expectation is taken over θ with respect to the posterior distribution of θ.

The quadratic loss function is defined as:

$$\text{(2.1)} \qquad \mathrm{L}(\theta, \hat{\theta}) = (\hat{\theta} - \theta)^2.$$

The expectation of the quadratic loss function given in (2.1) is:

$$\text{(2.2)} \qquad \mathrm{E}\left\lfloor \mathrm{L}(\theta, \hat{\theta}) \right\rfloor = \int (\theta - \hat{\theta})^2 \mathrm{P}(\theta \mid k_1, k_2, \ldots k_t) d\theta,$$

where k_i denotes the number of claims a policyholder had in year i, $i = 1, 2, \ldots, t$. So, by differentiating the expectation of the loss function given in (2.2) with respect to $\hat{\theta}$, the estimator $\hat{\theta}$ is found as follows:

$$\text{(2.3)} \qquad \hat{\theta} = \int \theta \mathrm{P}(\theta \mid k_1, k_2, \ldots, k_t) d\theta.$$

As seen from (2.3), the quadratic loss function is minimized by taking $\hat{\theta}$ to be the posterior mean.

3. Geometric Distribution (Poisson-Exponential Mixture) as a Claim Frequency Distribution

In automobile insurance, when the portfolio is considered to be heterogeneous, all policyholders will have a constant but unequal underlying risk of having an accident. That is, the expected number of claims differs from policyholder to policyholder. As the mixed Poisson distributions have thicker tails than the Poisson distribution, it is seen that the mixed Poisson distributions provide a good fit to claim frequency data when the portfolio is heterogeneous [5].

Frangos and Virontos [2] and also Dionne and Vanasse [1] considered the Poisson parameter λ, the expected number of claims, distributed according to Gamma. So, they used a Negative Binomial distribution to model the claim frequency data and Tremblay [3] used the Poisson-Inverse Gaussian distribution which provides a good fit to the claim frequency data. Willmot [5] used the Gamma, Generalized Inverse Gaussian, Beta and Uniform distributions to model the expected number of claims and obtained the probability distribution of the claim frequency. Walhin and Paris [4] used Hofmann's distribution to model the claim data.

Assume that the number of claims k is distributed according to Poisson with a given parameter λ,

$$\text{(3.1)} \qquad \mathrm{P}_\lambda(k \mid \lambda) = \frac{\mathrm{e}^{-\lambda}\lambda^k}{k!}, \ k = 0, 1, 2, \ldots \text{ and } \lambda > 0,$$

where λ denotes the differing underlying risk of each policyholder having an accident. Let us assume that λ is distributed according to the Exponential distribution with parameter θ (that is, the structure function of λ is assumed to be an Exponential distribution). The probability density function of λ is as follows:

$$\text{(3.2)} \qquad u(\lambda) = \theta \mathrm{e}^{-\lambda\theta}, \ \lambda > 0.$$

Then the unconditional distribution of k claims can be calculated by using (3.1) and (3.2) and taking the following integral:

$$\begin{aligned} \mathrm{P}(k) &= \int_0^\infty \mathrm{P}_\lambda(k \mid \lambda) u(\lambda) d\lambda \\ &= \int_0^\infty \frac{\mathrm{e}^{-\lambda}\lambda^k}{k!} \theta \mathrm{e}^{-\lambda\theta} d\lambda \\ &= \frac{\theta}{k!} \int_0^\infty \mathrm{e}^{-\lambda(1+\theta)} \lambda^k d\lambda \\ &= \frac{\theta}{k!} \frac{1}{(1+\theta)^{k+1}} \int_0^\infty \mathrm{e}^{-u} u^k du \\ &= \frac{\theta}{k!} \frac{1}{(1+\theta)^{k+1}} \Gamma(k+1) \\ &= \theta(1+\theta)^{k+1}, \ k = 0, 1, 2, \ldots. \end{aligned} \qquad \text{(3.3)}$$

As seen from (3.3), the unconditional distribution of k claims is a Geometric distribution.

If k_i denotes the number of claims a policyholder had in year i, $i = 1, 2, \ldots, t$, the total number of claims a policyholder had in t years is $K = \sum_{i=1}^t K_i$. Then, for a given λ, the conditional distribution of $K = \sum_{i=1}^t K_i$ claims in t years is:

$$\text{(3.4)} \qquad \mathrm{P}(k_1, k_2, \ldots, k_t \mid \lambda) = \frac{\mathrm{e}^{\lambda t}\lambda^K}{\prod_{i=1}^t k_i!}.$$

By applying the Bayesian theorem, the posterior structure function for a group of policyholders with a claim history $k_1, k_2, \ldots, k_t$ can be found as follows:

$$\begin{aligned} \mathrm{P}(k_1, k_2, \ldots, k_t \mid \lambda) &= \frac{\mathrm{e}^{-\lambda t}\lambda^K}{\prod k_i!}, \\ U(\lambda \mid k_1, k_2, \ldots, k_t) &\propto \mathrm{P}(k_1, k_2, \ldots, k_t \mid \lambda) U(\lambda) \\ &\propto \mathrm{e}^{-\lambda t}\lambda^K \mathrm{e}^{-\lambda\theta} \\ &\propto \mathrm{e}^{-\lambda(t+\theta)}\lambda^K. \end{aligned}$$

If

$$\int_0^\infty A\mathrm{e}^{-\lambda(t+\theta)}\lambda^K d\lambda = 1,$$

where A is a constant, then

$$A = \frac{(t+\theta)^{K+1}}{\Gamma(K+1)}.$$

Hence,

$$u(\lambda \mid k_1, k_2, \ldots, k_t) = \frac{(t+\theta)^{K+1}}{\Gamma(K+1)} \mathrm{e}^{-\lambda(t+\theta)} \lambda^K, \ \lambda > 0. \tag{3.5}$$

Using the quadratic loss function and following the procedure outlined in Section 2, the optimal choice for λ_{t+1}, the expected number of claims of a policyholder with a claim history $k_1, k_2, \ldots, k_t$, is:

$$\hat{\lambda}_{t+1} = \frac{K+1}{t+\theta} = \bar{\lambda}\frac{K+1}{t\bar{\lambda}+1}, \ \bar{\lambda} = \frac{1}{\theta}. \tag{3.6}$$

If the premium is determined only by taking the number of claims of a policyholder into account, assuming that the initial premium (premium at time $t = 0$) is 100, then at time $t + 1$ the policyholder is required to pay:

$$\text{Premium}_{t+1} = 100\frac{K+1}{t\bar{\lambda}+1} \tag{3.7}$$

As seen from (3.7), the risk premium payable at time $t + 1$ depends on the claim history of the policyholder and the parameter of the exponential distribution (θ).

4. Pareto Distribution (Exponential-Gamma mixture) as a Claim Severity Distribution

In an insurance portfolio, in addition to many small claim severities, high claim severities can also be observed. Therefore, long tail distributions such as Lognormal, Weibull, Pareto, Burr, etc. are widely used to model claim severity data.

Suppose that the amount x of the claim is distributed according to the Exponential distribution with a given parameter θ:

$$f(x \mid \theta) = \frac{1}{\theta}\mathrm{e}^{-x/\theta}. \tag{4.1}$$

As the mean claim size θ is not the same for all policyholders, an Inverse Gamma distribution with parameters s and m is taken as the prior distribution of the mean claim size:

$$g(\theta) = \frac{\frac{1}{m}\mathrm{e}^{-m/\theta}}{\left(\frac{\theta}{m}\right)^{s+1}\Gamma(s)}, \ \theta > 0. \tag{4.2}$$

So, the unconditional distribution of the claim size is a Pareto distribution whose probability density function is:

$$\begin{aligned} f(x) &= \int_0^\infty f(x \mid \theta) g(\theta) d\theta, \\ &= \int_0^\infty \frac{1}{\theta}\mathrm{e}^{-x/\theta} \frac{\frac{1}{m}\mathrm{e}^{-m/\theta}}{\left(\frac{\theta}{m}\right)^{s+1}\Gamma(s)} d\theta \\ &= \int_0^\infty \frac{m^s}{\Gamma(s)(x+m)^{s+1}} \int_0^\infty u^s \mathrm{e}^{-u} du \\ &= -m^s s(x+m)^{-(s+1)}, \ x > 0. \end{aligned} \tag{4.3}$$

If x_i, $i = 1, 2, \ldots, K$ denotes the amount of claim i, then the total amount claimed for a policyholder over the t years that he is in the portfolio will be equal to $\sum_{k=1}^{K} x_k$. So, the posterior distribution for θ can be obtained by applying the Bayesian theorem as follows:

$$\begin{aligned} g(\theta \mid x_1, x_2, \ldots, x_k) &\propto \frac{1}{\theta^{\sum k_i}} \mathrm{e}^{-\frac{\sum x_k}{\theta}} \frac{\mathrm{e}^{-m/\theta}}{\left(\frac{\theta}{m}\right)^{s+1}} \\ &\propto \mathrm{e}^{\frac{-(\sum x_k + m)}{\theta}} \frac{1}{\theta^{\sum k_i + s + 1}} \\ &= \mathrm{e}^{\frac{-(\sum x_k + m)}{\theta}} \theta^{-(\sum k_i + s + 1)} \end{aligned}$$

Since

$$\int_0^\infty A\mathrm{e}^{-\frac{\sum x_k + m}{\theta}} \theta^{-(\sum k_i + s + 1)} d\theta = 1,$$

where A is a constant, implies

$$A = \frac{(\sum x_k + m)^{K+s}}{\Gamma(s+K)}$$

we obtain:

$$(4.4) \qquad g(\theta \mid x_1, x_2, \ldots, x_K) = \frac{\frac{1}{\left(m + \sum_{k=1}^K x_k\right)} \mathrm{e}^{-\frac{m + \sum_{k=1}^K x_k}{\theta}}}{\left(\frac{\theta}{m + \sum_{k=1}^K x_k}\right)^{K+s+1} \Gamma(K+s)}, \ \theta > 0.$$

Using the quadratic loss function and following the procedure mentioned in section 2, the optimal choice for θ_{t+1} for a policyholder reporting claim amounts x_i, $i = 1, 2, \ldots, K$ over t years is estimated as:

$$(4.5) \qquad \hat{\theta}_{t+1} = \frac{m + \sum_{k=1}^K x_k}{s + K - 1}.$$

If the risk premium is determined not only by taking the number of claims into account but also the total amount of the claims, then the risk premium to be paid at time $t+1$ for a policyholder whose claim number history is $k_1, k_2, \ldots, k_t$, and whose claim amount history is $x_1, x_2, \ldots, x_K$, can be calculated according to the net premium principle as:

$$(4.6) \qquad \text{Premium}_{t+1} = \frac{K+1}{t+\theta} \cdot \frac{m + \sum_{k=1}^K x_k}{s + K - 1}.$$

As can be seen from (4.6), the risk premium that must be paid depends on the parameter of the Geometric distribution (θ), the parameters of the Pareto distribution (m and s), the number of years t that the policyholder is under observation, and his/her total number of claims and the amount of these claims.

5. Application

5.1. A BMS based on the claim frequency only. As the number of claims is assumed to follow a Geometric distribution, the parameter θ of this distribution is arbitrarily chosen as $1,25$. Using the net premium principle, an optimal bonus-malus system based only on the frequency component is considered. The risk premium is calculated by using (3.7). The risk premium at time $t = 0$ (base premium) is taken as 100 in order for us to be able to compare the risk premiums. If a policyholder had one claim ($K = 1$) in one year ($t = 1$), then the risk premium he/she has to pay is $100 \frac{1+1}{1/1,25+1} = 111$, and if he/she had two claims ($K = 2$) in 3 years ($t = 3$) then the risk premium is $100 \frac{1+2}{3/1,25+1} \equiv 88$. The risk premiums are calculated for years $t = 0, 1, \ldots, 7$ and for the number of claims $K = 0, 1, \ldots, 5$, and the results obtained are given in Table 1.

Table 1. Optimal BMS based on the A Posteriori Frequency Component

Year	Number of claims					
t	**0**	**1**	**2**	**3**	**4**	**5**
0	100					
1	56	111	167	222	278	333
2	38	77	115	154	192	231
3	29	59	88	118	147	176
4	24	48	71	95	119	143
5	20	40	60	80	100	120
6	17	34	52	69	86	103
7	15	30	45	61	76	91

As seen from Table 1, this BMS can be considered as being generous to good drivers and strict with bad drivers. For example, the bonuses given for the first claim free year are 44% of the basic premium. Drivers who have one accident over the first year will have to pay a malus of 11% of the basic premium.

5.2. A BMS based on both the claim frequency and the claim severity component. In order to see the effect of both the claim frequency and claim severity components, a BMS based on these components is considered. As the claim amounts are assumed to follow a Pareto distribution, the parameters m and s of this distribution are arbitrarily chosen as 495000 and 2,5 respectively. The aggregate claim amounts are taken as 250000 and 1000000 in order to compare the effect of the total claim amount on the risk premium to be paid, which is calculated using (4.6).

In Table 2 and Table 3, we can see the risk premiums that must be paid for various numbers of claims when the age of the policy is up to seven years. For example, a policyholder with one accident of claim size 250000 in the first year of observation will pay 264889, as seen from Table 2. On the other hand a policyholder with one accident of claim size 1000000 in the first year of observation will pay 531556, as seen from Table 3.

Table 2. Optimal BMS based on the Posteriori Frequency and Severity Component (Total Claim Size of 250000)

Year	Number of claims					
t	**0**	**1**	**2**	**3**	**4**	**5**
0	264000					
1	146667	264889	283810	294321	301010	305641
2	101538	183385	196484	203761	208392	211598
3	77647	140235	150252	155817	159358	161810
4	62857	113524	121633	126138	129004	130989
5	52800	95360	102171	105956	108364	110031
6	45517	82207	88079	91341	93417	94854
7	40000	72242	77403	80269	82094	83357

Table 3. Optimal BMS based on the Posteriori Frequency and Severity Component (Total Claim Size of 1000000)

Year	Number of claims					
t	**0**	**1**	**2**	**3**	**4**	**5**
0	264000					
1	146667	531556	569524	590617	604040	613333
2	101538	368000	394286	408889	418182	424615
3	77647	281412	301513	312680	319786	324706
4	62857	227810	244082	253122	258874	262857
5	52800	191360	205029	212622	217455	220800
6	45517	164966	176749	183295	187461	190345
7	40000	144970	155325	161077	164738	167273

If a policyholder with one accident of claim size 250000 in the first year of observation has one accident with claim size 750000 in the second year of observation, then a surcharge will be enforced and he/she will have to pay 394286, which is the premium for two accidents with an aggregate claim amount of 1000000 in two years of observation as seen from Table 3. If in the third year, he/she does not have an accident there will be a reduction in the premium because of a claim free year and he/she will pay 301513, which is the premium for two accidents of an aggregate claim amount of 1000000 in three years of observation as seen from Table 3.

6. Summary and Conclusion

In this paper the design of an optimal BMS based only on the claim frequency, and one based on both the claim frequency and the claim amount, are developed. As mixed distributions fit the claim number and claim amount data, the Geometric distribution - which is a mixture of a Poisson and an Exponential distributions - is used to model claim numbers, and the Pareto distribution - which is a mixture of an Exponential and a Gamma distribution - is used to model claim amounts. In an application, the risk premium is calculated using the net premium principle, and the results obtained by using the claim number only and by using both the claim number and claim amount are compared. It is concluded that it is fairer to charge policyholders premiums which not only take into account the number of claims, but also the aggregate amount of the claims.

References

[1] Dionne, G. and Vanesse, C. *A Generalization of automobile insurance rating models: the Negative Binomial distribution with a regression component*, ASTIN Bulletin **19**, 199–212, 1989.

[2] Frangos, N. E. and Vrontos, S. D. *Design of optimal bonus-malus systems with a frequency and a severity component on an individual basis in automobile insurance*, ASTIN Bulletin **31**, 1–22, 2001.

[3] Tremblay, L. *Using the Poisson Inverse Gaussian in bonus-malus systems*, ASTIN Bulletin **22**, 97–106, 1992.

[4] Walhin, J. F. and Paris, J. *The true claim amount and frequency distributions within a bonus-malus system*, ASTIN Bulletin **30**, 391–403, 2000.

[5] Willmot, G. *Mixed compound Poisson distributions*, ASTIN Bulletin **16**, 59–79, 1993.