# THEORETICAL AND PRACTICAL CONSIDERATIONS REGARDING BONUS-MALUS SYSTEM

Dorina Lazăr, Ioan Lazăr, Ilie Parpucea

Abstract. In this paper we add some remarks regarding the theoretical background and the applicability of bonus-malus system. Automobile insurance portfolios are heterogeneous because there are a series of characteristics that influence the damage frequency, respectively the severity. In the first part of the paper we present a direct justification of the property which sustains the building of the bonus-malus system using the negative binomial model. There will be also an empirical study for an important automobile insurance portfolio from our country. A private insurance company provided the database for which we intend to analyse the applicability of bonus-malus system. The numerical results suggest the necessity for the insurance providers to combine their a priori and a posteriori tariff systems.

*Keywords and phrases*: bonus-malus system, negative binomial distribution, a posteriori distribution, risk parameter.

## 1. Introduction

The bonus-malus system is used for including the information regarding the accident record of the policyholder in individual tariffs systems. The a posteriori rating system suits the non-homogeneous portfolios where the individual characteristics are difficult to be measured a priori. As a result of the automobile portfolios heterogeneity, the insurance companies very often combine the two tariff systems: a priori and a posteriori system.

The a priori tariff system consists of the portfolio segmentation in relatively homogene risk classes, after observable variables, with important influence on incidents history, and the assignment of the a priori tariff for each risk class. As a result of the ignorance of some risk factors, the risk classes will not be homogeneous and one should use the a posteriori bonus-malus tariff system.

The bonus-malus system stands on the accidents registered in the past. If we consider a heterogeneous portfolio, we have the information regarding the number of claims from the last $n$ years $X_t$, $t = 1, 2, \ldots, n$, where $X_t$ is the number of claims made by the policyholder for the year $t$. Let us consider an insurance policy (risk) randomly selected from a portfolio which contains several similar insurance policies.

For the incorporation of the portfolio heterogenity, we associate a random parameter $\Theta$ for each insurance policy (Bühlmann, 1967).

Given the risk parameter, $\Theta = \theta$, we suppose that the conditional variables $X_1, X_2, \ldots$ are independent random variables, which follow a Poisson distribution with parameter $\theta > 0$ (identical distributed) $Po(\theta)$:

$$f_\theta(x) = e^{-\theta}\frac{\theta^x}{x!}, \quad x = 0, 1, 2, \ldots \tag{1}$$

We also suppose that the distribution of $\Theta$ parameter for the portfolio is represented using the gamma distribution $G(k, \beta)$:

$$g(\theta) = \frac{1}{\beta^k \Gamma(k)}\theta^{k-1}\exp(-\theta/\beta), \quad x \geq 0 \tag{2}$$

Considering these hypothesis, the number of claims $X_t$ has a binomial negative probability distribution of parameters $BN\left(r = k, p = \frac{1}{1+\beta}\right)$.

Bayes' a posteriori estimator for expected number of claims (this peculiar form is because the conditional distribution is a Poisson distribution):

$$\mu(\Theta) = E(X_t|\Theta) = \Theta \tag{3}$$

results from the conditional mean:

$$\widehat{\mu}_C = E[\Theta|X_1, X_2, \ldots, X_n].$$

It is determined an estimator for the expected number of claims, for year $n+1$, considering the records in the past of number of claims during $n$ precedent years (Reinhart, 2005):

$$E(\Theta|X_1 = x_1, X_2 = x_2, \ldots, X_n = x_n)$$

$$= E(X_{n+1}|X_1 = x_1, X_2 = x_2, \ldots, X_n = x_n).$$

The calculus of this conditional mean requires the knowledge of the risk parameter conditional distribution $\Theta(X_1, X_2, \ldots, X_n)$.

## 2. Theoretical aspects regarding a bonus-malus system design

The design of a bonus-malus systems is realized starting from the hypothesis that the number of claims $X_t$ follows a negative binomial distribution.

Corollary 1 *Given the risk parameter $\Theta = \theta$, we suppose that the conditional variables $X_1, X_2, \ldots$ are independent and they follow a Poisson distribution $Po(\theta)$. The distribution of $\Theta$ inside the portfolio will be described using the gamma distribution $G(k, \beta)$. Under these hypothesis, the distribution of the risk parameter $\Theta$, conditioned by $X_1 = x_1$, $X_2 = x_2, \ldots$, $X_n = x_n$ (the a posteriori distribution) is a gamma distribution of parameters $G(k+s, 1/(n+\tau)$.*

One of the justifications of this corollary is realized in a larger context - that of structure conjugated functions (Reinhart, 2005). Other approaches are presented by Mikosch (2003) and Petauton (2000).

We shall give a direct justification of this corollary. Because the conditional variables $X_1, X_2, \ldots, X_n$ are independent and identically distributed, following a Poisson distribution of parameter $\theta$, results:

$$g(\theta|x_1, \ldots, x_n) = \frac{\displaystyle\prod_{i=1}^{n} f_\theta(x_i)g(\theta)}{\displaystyle\int_0^\infty f_z(x_1, \ldots, x_n)g(z)dz}$$

$$= \frac{\left(\prod_{i=1}^{n} e^{-\theta}\theta^{x_k}/x_i!\right) g(\theta)}{\int_0^\infty \left(\prod_{i=1}^{n} e^{-z} z^{x_k}/x_i!\right) g(z)dz}. \tag{4}$$

$$\begin{aligned} g(\theta|x_1,\ldots,x_n) &= \frac{e^{-\theta n}\theta^s g(\theta)}{\int_0^\infty e^{-zn}z^s g(z)dz} \\ &= \frac{1}{\int_0^\infty e^{-zn}z^s g(z)dz} e^{-\theta n}\theta^s \frac{\tau_k}{\Gamma(k)}\theta^{k-1}\exp(-\theta\tau) \\ &= \frac{\tau^k}{\Gamma(k)} \frac{1}{\int_0^\infty e^{-zn}z^s g(z)dz} e^{-\theta(n+\tau)}\theta^{k+s-1} \end{aligned}$$

where $s = x_1 + x_2 + \cdots + x_n$ is the total number of claims made on the insurance policy considered in $n$ years, and $\tau = 1/\beta$. Because:

$$\begin{aligned} \int_0^\infty e^{-zn}z^s g(z)dz &= \int_0^\infty e^{-zn}z^s \frac{\tau^k}{\Gamma(k)} e^{-z\tau}z^{k-1}dz \\ &= \frac{\tau^k}{\Gamma(k)} \int_0^\infty e^{-z(n+\tau)}z^{k+s-1}dz \end{aligned}$$

and

$$\begin{aligned} \int_0^\infty e^{-z(n+\tau)}z^{k+s-1}dz &= \frac{1}{(n+\tau)^{k+s}} \int_0^\infty e^{-z(n+\tau)}[z(n+\tau)]^{k+s-1}[(n+\tau)]dz \\ &= \frac{1}{(n+\tau)^{k+s}} \int_0^\infty e^{-x}x^{k+s-1}dx = \frac{\Gamma(k+s)}{(n+\tau)^{k+s}} \end{aligned}$$

results that:

$$g(\theta|x_1,\ldots,x_n) = \frac{(n+\tau)^{k+s}}{\Gamma(k+s)} e^{-\theta(n+\tau)}\theta^{k+s-1}. \tag{5}$$

This function represents the probability density function of the gamma distribution of parameters $G(k+s, 1/(n+\tau)$. The risk parameter conditional distribution is a gamma probability law, as in the unconditional case.

Considering the expected value of one variable which follows the gamma distribution, results:

$$E(\Theta|X_1 = x_1, X_2 = x_2, \ldots, X_n = x_n) = \frac{k+s}{n+1/\beta}. \tag{6}$$

This expression can be also obtained starting from the conditional distribution of the number of claims for the year $n+1$. As it follows, the expected number of claims for the year $n+1$ depends only by the total number of claims per contract $s$ for those $n$ precedents years (it does not depend by the repartition of the claims during this period).

Regarding the first insurance year, considering that the number of claims $X_1$ follows a negative binomial distribution of parameters

$$BN\left(r = k, p = \frac{1}{1+\beta}\right)$$

results the expected number of claims:

$$E(X_1) = k\left(1 - \frac{1}{\beta+1}\right)(\beta+1) = k\beta. \tag{7}$$

If during the calculus of the insurance premium we consider only the number of claims occurred in the past, then the insurance premium for the year $n+1$, for a policyholder who produced $s$ claims during $n$ precedent years, expressed as percents of the insurance premium for the first year, is (Reinhart, 2005):

$$\frac{(k+s)}{(n+1/\beta)k\beta} \cdot 100 \tag{8}$$

This expression can be used for the design of some bonus-malus specific tables.

## 3. Numerical results

We consider a portfolio in which the yearly number of claims follows the negative binomial law. The gamma distribution parameters $G(k, \beta)$, which model the risk parameter $\Theta$ distribution, where $k = 2$, $\beta = 0.1$.

*Table 1. The bonus-malus table, the average frequency in portfolio equals 20%*

| $n$ \ $s$ | 0 | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|---|
| 0 | $100,0$ | | | | | |
| 1 | $90,9$ | $136,4$ | $181,8$ | $227,3$ | $272,7$ | $318,2$ |
| 2 | $83,3$ | $125,0$ | $166,7$ | $208,3$ | $250,0$ | $291,7$ |
| 3 | $76,9$ | $115,4$ | $153,8$ | $192,3$ | $230,8$ | $269,2$ |
| 4 | $71,4$ | $107,1$ | $142,9$ | $178,6$ | $214,3$ | $250,0$ |
| 5 | $66,7$ | $199,0$ | $133,3$ | $166,7$ | $200,0$ | $233,3$ |
| 6 | $62,5$ | $93,7$ | $125,0$ | $156,2$ | $187,5$ | $218,7$ |
| 7 | $58,8$ | $88,2$ | $117,6$ | $147,1$ | $176,5$ | $205,9$ |
| 8 | $55,6$ | $83,3$ | $111,1$ | $138,9$ | $166,7$ | $194,4$ |

In the table 1 we have a bonus-malus table, computed starting from the expression (8). The risk parameter average for this portfolio, considering the gamma distribution average, is $k \cdot \beta = 0,2$. As it follows, the expected number of claims (a priori) is of 20%. We can observe that an accident occurred in the first year brings a malus of 36,4% (an increase of the insurance premium with 36,4%), a policyholder who had no claims during the last four years will benefit a bonus of 28,6%.

In the next section we will use a portfolio from our country insurance companies practice, which covers the automobile damages resulted from accidents caused by the policyholder. In table 2 we present the distribution of 16000 policies according with the number of claims occurred during one year. There are also the theoretical frequencies computed considering the hypothesis of fitting the negative binomial distribution.

The sample mean and the sample variance, for the number of claims $X$, are: $\overline{x} = 0,2747$ and $s^2 = 0,4917$ accordingly. We observe that the sample variance is greater than the sample mean, which suggests that the Poisson distribution is not appropriate, there should be introduced a distribution which permits a greater variation (heterogeneity). The negative binomial law is a serious candidate for fitting the distribution of this portfolio.

*Table 2. Fitting the number of claims distribution using the negative binomial law*

| Number of claims | Observed frequencies | Negative binomial frequencies |
|---|---|---|
| 0 | 13172 | 13067 |
| 1 | 1794 | 1806 |
| 2 | 674 | 597 |
| 3 | 238 | 206 |
| 4 | 84 | 76 |
| 5 | 28 | 29 |
| 6 | 7 | 11 |
| $\geq 7$ | 3 | 4 |

Using moments method, from the equations:

$$0,2747 = r(1-p)/p$$

$$0,4917 = r(1-p)/p^2$$

results the estimations for the negative binomial law parameters $\widehat{p} = 0,5587$ and $\widehat{r} = 0,3478$. The parameters of the gamma distribution are $\widehat{k} = \widehat{r} = 0,3478$ and $\widehat{\beta} = \dfrac{1-\widehat{p}}{\widehat{p}} = 0,7899$. The probability density function of a discrete variable $X$ which follows the negative binomial distribution $BN(r,p)$ of parameters $r > 0$ and $p \in (0,1)$ is:

$$P(X=k) = \frac{\Gamma(r+k)}{k!\Gamma(r)}p^r(1-p)^k, \quad k = 0,1,2,\ldots$$

As it follows:

$$P(X=0) = p_0 = p^r = 0,8167; \quad np_0 = 13067$$

$$P(X=1) = p_1 = \frac{\Gamma(r+1)}{1!\gamma(r)}p^r(1-p) \cong rp_0(1-p) = 0,1254; \quad np_1 = 1806.$$

In a similar way will be calculated the others theoretical frequencies from table 2. Considering the differences between the observed frequencies and theoretical frequencies, the value of the statistic $\lambda^2$:

$$\lambda^2_{calc} = \sum_{i=1}^{8} \frac{(n_i - np_i)^2}{np_i} = 9,23$$

do not conduct us to the rejection of the null hypothesis, for 5% level of significance. In this portfolio the yearly number of claims made by a policyholder is modeled using the binomial negative distribution $BN(\widehat{r} = 0,3478; \widehat{p} = 0,5587)$, where the parameters of gamma distribution $G(k, \beta)$ which models the risk parameter distribution $\Theta$ are $\widehat{k} = 0,3478$ and $\widehat{\beta} = 0,7899$. The table 3 contains the appropriate bonus-malus system for this portfolio. The average of the risk parameter is $k \cdot \beta = 0,2798$ while the average frequency (a priori) of claims is 27,98%.

*Table 3. The appropriate bonus-malus system for the portfolio, the average frequency in the portfolio 27,98%*

| $s$ / $n$ | 0 | 1 | 2 | 3 | 4 |
|---|---|---|---|---|---|
| 0 | 100,0 | | | | |
| 1 | 55,8 | 216,5 | 377,1 | 537,7 | 698,4 |
| 2 | 38,7 | 150,2 | 261,6 | 373,1 | 484,5 |
| 3 | 29,6 | 115,0 | 200,3 | 285,6 | 370,9 |
| 4 | 24,0 | 93,1 | 162,2 | 231,4 | 300,5 |
| 5 | 20,2 | 78,2 | 136,3 | 194,4 | 252,5 |
| 6 | 17,4 | 67,5 | 117,6 | 167,7 | 217,8 |
| 7 | 15,3 | 59,3 | 103,3 | 147,4 | 191,4 |
| 8 | 13,6 | 52,9 | 92,2 | 131,5 | 170,7 |

According to this table, the insurance premiums asked by the insurer company differs very much from one policyholder to another, a fact that the insurers cannot afford. It is difficult to implement in practice such a premium table. The heterogeneous portfolios are creating this situation. Making a comparison between the variances from table 1 and table 2, we can see that the insurer' portfolio is more heterogeneous. For the hypothesis which standed for the first table, the variance of number of claims is $r(1-p)/p^2 = 0,22024$ (because $p = 1/(1+\beta) = 0,909$).

We consider that these observations ask for the combination of the both tariff systems, a priori and posteriori, of the insurers. We should mention that in the same area, meaning pricing automobile insurance, Dionne & Vanasse (1989) have suggested a bonus-malus system using the negative binomial distribution, which used a priori and posteriori information simultaneously.

An other possible approach is the a priori construction of some risk homogeneous classes and then to build up a bonus-malus table for each class. Considering this purpose, Lazăr & all (2005) have estimated for the dependent variable - number of claims, and for this portfolio, the Poisson regression model. Between the significance explicative variables there could be remembered: the gender of the insured person, the cylinder capacity of the engine, the brand of the car, the age of the car. We also consider that the ignorance of the claim size in a posteriori tariff systems is not always justified.

## References

[1] Bühlmann, H., *Experience rating and credibility*, ASTIN Bulletin, 4 (1967), pp. 199-207.

[2] Dionne, G., Vanasse, C., A generalization of automobile rating models: the negative binomial distribution with a regression component, *Astin Bulletin*, 19 (1989), pp. 199-212.

[3] Goovaerts, M.J., Kaas, R., Van Heerwaarden, A.E., Bauwelinckx, T., *Effective Actuarial Methods*, North-Holland (1990) Amsterdam.

[4] Lazăr, D., Ciumas, C., Pop, F., On using econometrics in pricing automobile insurance. In Poloucek, S., Stavarek, D. (eds.) *Future of Banking after the Year 2000 in the World and in the Czech Republic*, Karvina: Silesian University (2005), pp. 1285-1298.

[5] Lemaire J., Hongmin, Z., Hight deductibles instead of Bonus-Malus. Can it work? *ASTIN Bulletin* 24 (1994), 75-88.

[6] Mikosch, T., *Non-Life Insurance Mathematics. A Primer*, Laboratory of Actuarial Mathematics, University of Copenhagen, www.math.ku.dk. (2003)

[7] Petauton P., *Theorie de l'Assurance Dommages*, Dunod, Paris (2000).

[8] Reinhart, J.M., *Assurance Non-Vie* (course notes), Université Libre de Bruxelles, Belgium (2005).

**Authors:**

Dorina Lazăr
Babeş-Bolyai University of Cluj-Napoca,
Romania.
E-mail: dlazar@econ.ubbcluj.ro

Ioan Lazăr
Babeş-Bolyai University of Cluj-Napoca,
Romania.
E-mail: ilazar@econ.ubbcluj.ro

Ilie Parpucea
Babeş-Bolyai University of Cluj-Napoca,
Romania.
E-mail: iparpucea@econ.ubbcluj.ro