Annales Univ. Sci. Budapest., Sect. Comp. 42 (2014) 19–35

ESTIMATION OF CLAIM NUMBERS IN AUTOMOBILE INSURANCE

Miklós Arató (Budapest, Hungary)
László Martinek (Budapest, Hungary)

Dedicated to András Benczúr on the occasion of his 70th birthday

Communicated by László Lakatos

(Received June 1, 2014; accepted July 1, 2014)

Abstract. The use of bonus-malus systems in compulsory liability automobile insurance is a worldwide applied method for premium pricing. Considering certain assumptions, an interesting and crucial task is to evaluate the so called claims frequency regarding the individuals. Here we introduce 3 techniques, two is based on the bonus-malus class, and the third based on claims history. The article is devoted to choose the method, which fits to the frequency parameters the best for specific input parameters. For measuring the goodness-of-fit we will use scores, similar to better known divergence measures. The detailed method is also suitable to compare bonus-malus systems in the sense that how much information they contain about drivers.

1. Introduction

The joint application of probabilistic, statistical and Bayesian methods related to specific issues in practice played an important role in several works of András Benczúr (see [1] for example). These methods are applied in the present article concerning an actuarial problem.

Key words and phrases: bonus-malus; claims frequency; Bayesian; scores; Markov chains.
2010 Mathematics Subject Classification: 97M30
The European Union and the European Social Fund have provided financial support to the project under the grant agreement no. TÁMOP 4.2.1/B-09/KMR-2010-0003.

1.1. Motivation

The concept of bonus-malus system is a worldwide, especially in Europe and Asia applied method in compulsory liability automobile insurance. Principally constructed in order to support the premium calculation of drivers. This might vary for different countries, but the idea is similar: policyholders with bad history should pay more than others without accidents in the past years. Schematically and mathematically described, there is an underlying graph with vertices called classes, among others an initial vertex, where every new driver begins. The graph contains a finite number of classes in practice, however, interesting articles are involved in the discussion of infinite systems, see book [10]. After a year without causing any accident he or she jumps up to another class, which has a cheaper premium. Otherwise, in case of causing an accident, the insured person goes downward, to a new class with higher premium, except he or she was already in the worst one.

Many other factors are taken into consideration when calculating one's premium, such as the engine type, purpose of use, habitat, age of the person etc., which result an a priori premium. Furthermore, the a posteriori premium arises as the product of a priori and the bonus-malus factor. A wide range of literature is available discussing these factors extensively, see [10], for example. Here we only concentrate on the bonus-malus system. Note that premium calculation is often realized involving credibility techniques, mixing the insurance institution's experiences with available national data, according to an appropriate proportion. The discussion of credibility is also omitted in the present article, see [7] for details, for instance. We also recommend book [3].

Our aim is to estimate the expected λ number of accidents triggered by insured drivers. This is usually called the claims frequency of the policyholder. First we review our necessary assumptions, among others about the distribution of claim numbers of a policyholder in one year, and the Markovian property of a random walk in such a system. Section 2 will discuss the basic problem illustrated with Belgian, Brazilian and Hungarian examples. Since also a Bayesian approach will be used for estimation of λ, a general a priori assumption is also needed. Based on this and a gappy information about the driver, the a posteriori expected value of λ will be calculated.

Fitting the possible best estimation for λ is a crucial task in the insurer's operation, since the expected value of claims, and which is the most important, the claims payments are forecasted using λ. Let us note that for evaluating the size of payments on the part of the insurer, the size of the property damages has to be approximated, not even the number of them. This part of actuarial calculations is not taken into account in the present article, for further discussions see [6, 12], for instance.

1.2. Bonus-malus systems

We assume the following constraints, simplifying real-world scenarios:

Assumptions 1.1. *1. λ is a constant value in time for each policyholder, and generally assumed to be the realization of a Λ random variable independent for each person. Remark that the case of time-dependent λ implies using double stochastic processes (Cox processes for instance), i.e., $\lambda(t)$ would also be a stochastic process.*

2. The random walk on the graph of classes is a homogeneous Markov chain, i.e., the next step depends only on the last state, and is independent of time.

3. The distribution of a policyholder's claim number for a year is Poisson(λ) distributed, conditionally on $\{\Lambda = \lambda\}$.

Notation 1.2. *For a bonus-malus system consisting of n classes, let C_1 denote the worst one with the highest premium, C_2 the second worst etc., and finally C_n the best premium class which can be achieved by a driver. Moreover, let Y_t be the class of the policyholder after t steps (years).*

Using these notations, the Markovian property can be written as $P(Y_t = C_i|Y_{t-1}, \ldots, Y_1) = P(Y_t = C_i|Y_{t-1})$. As the Y_t process is supposed to be homogeneous, it is correct to simply write p_{ij} instead of $P(Y_{t+1} = j|Y_t = i)$. These values specify an $n \times n$ stochastic matrix with non-negative elements, namely the transition probability matrix of the random walk on states. Let us denote it by $M(\lambda)$. Now to be more specific, we outline the example of three different systems, the Belgian, Brazilian and Hungarian.

Example 1.3 (Hungarian system)**.** *In the Hungarian bonus-malus system there are 15 premium classes, namely an initial (A_0), 4 malus ($M_4, \ldots, M_1$) and 10 bonus ($B_1, \ldots, B_{10}$) classes. Using the aforementioned notations, we can think of it as $C_1 = M_4, \ldots, C_4 = M_1, C_5 = A_0, C_6 = B_1, \ldots, C_{15} = B_{10}$. After every claim-free year the policyholder jumps one step up, unless he or she was in B_{10}, when there is no better class to go to. The consequence of every reported damage is 2 classes relapse, and at least 4 damages pulls the driver back to the worst M_4 state. Thus the transition probability matrix takes the form of Equation 7.1, see Appendix.*

Example 1.4 (Brazilian system)**.** *7 premium classes: $A_0, B_1, B_2, \ldots, B_6$. Sometimes written as classes $7, 6, 5, \ldots, 1$, e.g. in [11]. Transition rules can be found in the cited article.*

Example 1.5 (Belgian system)**.** *The new Belgian system was introduced in 1992. We address the transition rules regarding business-users, which can be found in article [11], among others. There are 23 premium classes:* M_8, M_7, ..., M_1, A_0, B_1, B_2, $\dots, B_{14}$ *(sometimes written as classes* $23, 22, 21, \dots, 2, 1$*).*

2. The Bayesian approach

Our realistic problem is the following. When an insured person changes insurance institution, the new company may not necessarily get his or her claim history, only the class where his or her life has to be continued. The new insurer also knows the number of years the policyholder has spent in the liability insurance system. Nevertheless, based on this two information we would like to provide the best possible estimation for the specific person's λ.

Recall that $Y_t = c$ denotes the event that the investigated policyholder has spent t years in the system (more precisely, from the initial class he or she has taken t steps), and arrived in class c.

Notation 2.1. π_0 *is the initial discrete distribution on the graph, which is a row vector of the form* $(0, \dots, 0, 1, 0, \dots, 0)$*. The ith element is 1, which means that each driver begins in the initial* C_i *state.*

As a Bayesian approach, suppose that claims frequency is also a random variable, and denote it by Λ. In practice, the gamma mixing distribution is a commonly used choice for that, therefore only this case will be discussed now. For other cases, the calculations can be similar, though not similarly nicely done.

Notation 2.2. $\Gamma(\alpha, \beta)$ *is the gamma distribution with* α *shape and* β *scale parameters, and with density function* $f(x) = \frac{x^{\alpha-1} \cdot \beta^{\alpha} \cdot e^{-\beta x}}{\Gamma(\alpha)}$.

Conditionally on $\{\Lambda = \lambda\}$, the number of property damages related to a driver is a random variable $X \sim \text{Poisson}(\lambda)$, i.e., the conditional probability is $P(X = k | \Lambda = \lambda) = \frac{\lambda^k}{k!} e^{-\lambda}$. It is well known (see [3] page 28 for details) that the unconditional distribution will be negative binomial with parameters $\left(\alpha, \frac{\beta}{1+\beta}\right)$ in our notations, i.e., if $\Lambda \sim \Gamma(\alpha, \beta)$. Accordingly, the a priori parameters can be estimated by a standard method of moments or maximum likelihood method.

Remark 2.1. *It is certainly necessary to make hypothesis testing after all, because our assumptions regarding the mixing distribution might be inaccurate.*

Remark 2.2. *Besides the gamma mixing distribution, other alternatives of claims frequency distributions might be worth considering in practice. For instance, if Λ is Inverse Gaussian, the unconditional distribution of X is Poisson Inverse Gaussian, see [5, 13]. Furthermore, the case of $\Lambda \sim$ Log-normal is also realistic, for description see [5]. For a more general parametric consideration see [14], which contains Poisson, Negative Binomial or Poisson Inverse Gaussian distributions as special cases, although requires 3 parameters. Besides the parametric assumptions, also the non-parametric estimation is worth considering, regarding the mixing distribution, see [2].*

2.1. Estimation of distribution parameters

In this subsection we describe an estimation method to compute the approximate values of α and β parameters. Assume that the insurance company has claim statistics from the past few years containing m policyholders. The ith insured person caused X_i accidents by his or her fault over a time period of t_i years, where t_i is not necessarily an integer. According to our assumption, the distribution of X_i is Poisson$(t_i \cdot \Theta)$, where t_i is a personal time factor and Θ is a Gamma(α, β)-distributed random variable. The unconditional distribution of X_i as mentioned above is Negative Binomial$\left(\alpha, \frac{\beta}{t_i+\beta}\right)$, thus its first two moments are

$$EX_i = t_i \frac{\alpha}{\beta}, \tag{2.1}$$

$$EX_i^2 = t_i^2 \frac{\alpha}{\beta^2} + t_i^2 \frac{\alpha^2}{\beta^2} + t_i \frac{\alpha}{\beta}. \tag{2.2}$$

Now we construct a method of moments estimation. Since this methodology implies several corresponding systems of equations, we have to choose one of them, described as follows. On the one hand, we can say that $\sum_{i=1}^{m} EX_i = \frac{\alpha}{\beta} \sum_{i=1}^{m} t_i$ and $\sum_{i=1}^{m} EX_i^2 = \left(\frac{\alpha}{\beta^2} + \frac{\alpha^2}{\beta^2}\right) \sum_{i=1}^{m} t_i^2 + \frac{\alpha}{\beta} \sum_{i=1}^{m} t_i$, thus the estimators of parameters are the result of the following equations

$$\frac{\hat{\alpha}}{\hat{\beta}} = \frac{\sum_{i=1}^{m} X_i}{\sum_{i=1}^{m} t_i}, \tag{2.3}$$

$$\frac{1}{\hat{\beta}} = \frac{\sum_{i=1}^{m} t_i}{\sum_{i=1}^{m} t_i^2} \left(\frac{\sum_{i=1}^{m} X_i^2}{\sum_{i=1}^{m} X_i} - 1 \right) - \frac{\sum_{i=1}^{m} X_i}{\sum_{i=1}^{m} t_i}. \tag{2.4}$$

On the other hand, $\sum_{i=1}^{m} \frac{EX_i}{t_i} = m\frac{\alpha}{\beta}$, thus the first equation for the estimators is the following $\frac{\hat{\alpha}}{\hat{\beta}} = \frac{\sum_{i=1}^{m} \frac{x_i}{t_i}}{m}$. Generally they do not provide the same results, except if $t_1 = t_2 = \ldots = t_m$. Numerous claims history scenarios were simulated for several portfolios, and according to our experiences, the solutions of system of equations 2.3 and 2.4 provided the best estimations for α and β. Finally we have to notice the maximum likelihood method as another obvious solution for this parameter estimation. Unfortunately, the likelihood function generally has no maximum, hence this method has been rejected.

2.2. Conditional probabilities

On the one hand, assume the a priori α, β parameters to be evaluated. On the other hand, consider the information $\{Y_t = c\}$ besides the initial class Y_0, which is fixed for each insured person. According to Bayes' theorem, the conditional density of Λ is

$$f_{\Lambda|Y_t=c}(\lambda|c) = \frac{P(Y_t = c|\Lambda = \lambda) \cdot f_\Lambda(\lambda)}{\int_0^\infty P(Y_t = c|\Lambda = \lambda) \cdot f_\Lambda(\lambda) d\lambda},$$

and the estimation for λ is the a posteriori expected value, denoted by

$$\hat{\lambda} = \int_0^\infty \lambda \cdot f_{\Lambda|Y_t=c}(\lambda|c) d\lambda.$$

Unfortunately, $P(Y_t = c|\Lambda = \lambda)$ introduces complications, because it only can be evaluated pointwise as a function of λ, calculating the tth power of transition matrix $M(\lambda)$. If I denotes the index of the initial class in the graph, this probability is exactly the $M(\lambda)^t_{(I,|c|)}$ element of the matrix, where $|c|$ is the index of class c. Using numerical integration it can be solved relatively fast. If we take a glance at Figure 1, we might see the claim frequency estimations for different countries and different α and β parameters. For example, the first figure was made based on the Brazilian bonus-malus system with parameters $\alpha = 1.2$ and $\beta = 19$, which indicates a relatively high risk portfolio. There is a line in the chart for each bonus class, which values show the estimated

λ frequencies as a function of time. Note that the mentioned functions do converge slowly to the stationary state, which might take even several decades. Thus the information regarding the elapsed time in the system of the individual is important.

We will refer to this method as `method.1` later.

3. Other methods for frequency estimation

In this section we shortly introduce 2 other (known) possible methods to evaluate policyholders' claim frequencies. Our assumptions for the distribution of claim numbers still definitely hold.

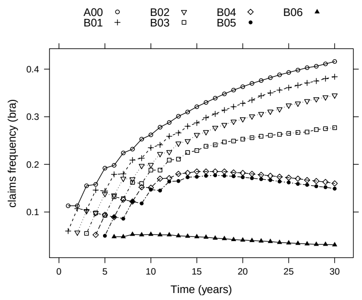
3.1. Average claim numbers of classes

The following method is probably the simplest, and we will refer to this as `method.2`. Suppose that statistics concerning the last years claim numbers are available. The estimator for a policyholders' claims frequency in class C is defined as the average number of claims related to the population in bonus class C in previous year, i.e., $\hat{\lambda}_C = \frac{\sum_{j=1}^{m} k_j \cdot \chi_{\{j\text{th policyholder is in class } C\}}}{\sum_{j=1}^{m} \cdot \chi_{\{j\text{th policyholder is in class } C\}}}$, where m is the total number of policies in previous year, k_j is the claim number regarding jth person, and χ is an indicator function, respectively.

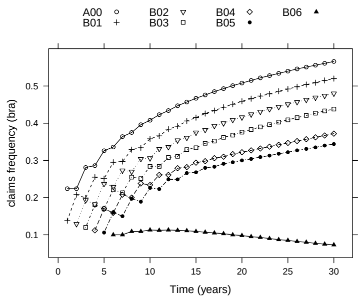
For instance, if last year our portfolio contained 5 policyholders each in class C with claims $0, 1, 0, 0, 2$, then the estimation for insured peoples' frequencies present year in class C is $\hat{\lambda}_C = 0.6$. Although it might sound as an oversimplification, in some cases it gives the best results.

3.2. Claim history of individuals

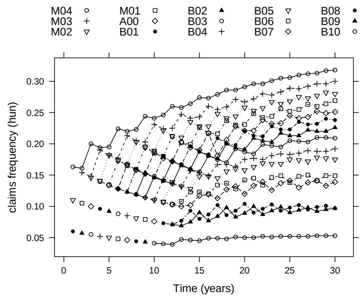
The third method is generally prevalent in the actuarial practice, and will be referred as `method.3`. Here we use the insured person's claim history, i.e., suppose that he or she was insured by our company for t years, and the distribution of claim numbers for each year is Poisson(λ). Let X denote the number of aggregate claims caused by this person in his t-year-long presence in the system, more precisely, in our field of vision. Then the conditional distribution of X is also Poisson with parameter λt. Let us remark that t does not



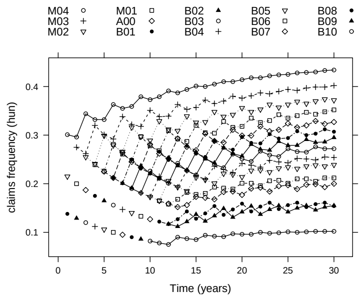
(a) Brazilian, $\alpha = 1.2, \beta = 19$



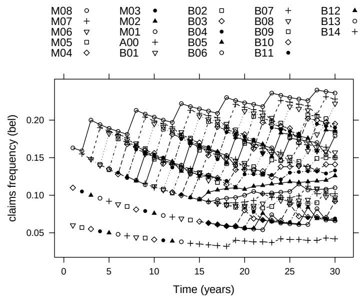
(b) Brazilian, $\alpha = 1.8, \beta = 12$



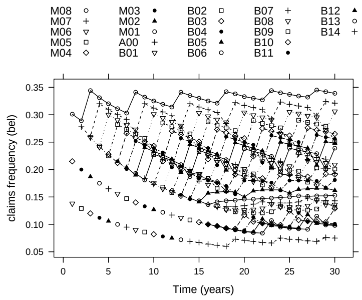
(c) Hungarian, $\alpha = 1.2, \beta = 19$



(d) Hungarian, $\alpha = 1.8, \beta = 12$



(e) Belgian, $\alpha = 1.2, \beta = 19$



(f) Belgian, $\alpha = 1.8, \beta = 12$

Figure 1: Estimated λ parameters for different countries and α, β parameters on a time horizon of 30 years. (Charts start at points for each class, from where the probability of being there is positive.)

have to be an integer (people often change insurer in the middle of the year in most countries). Based on this, our estimation on a special λ is the conditional expected value of the Gamma distributed Λ on condition $X = x$, i.e., as well known, $\hat{\lambda} = E(\Lambda|X = x) = \frac{x+\alpha}{t+\beta}$.

This is also a Bayesian approach as in the case of `method.1`, but the condition is different. Besides notice that first the α and β parameters have to be estimated exactly the way we have seen it before in Section 2.1.

4. Comparison using scores

The aim of this section is to make a decision which method gives the most accurate estimation of claims frequencies. In this context, we will use the theory of scores, a tool of probabilistic forecasting. For much more detailed information see article [9], as here we will discuss only the most important properties useful for our problem.

Scores are made for measuring the accuracy of probabilistic forecasts, i.e., measuring the goodness-of-fit of our evaluations. Let Ω be a sample space, $\mathcal{A}$ a σ-algebra of subsets of Ω and $\mathcal{P}$ is a family of probability measures on $(\Omega, \mathcal{A})$. Let a scoring rule be a function $S : \mathcal{P} \times \Omega \longrightarrow \overline{\mathbb{R}} = [-\infty, \infty]$. We work with the expected score $S(P, Q) = \int S(P, \omega)dQ(\omega)$, where measure P is our estimation and Q is the real one. Obviously one of the most important properties of this function is the inequality $S(Q, Q) \geq S(P, Q)$ for all $P, Q \in \mathcal{P}$. In this case S is proper relative to $\mathcal{P}$. Additionally, we call S regular relative to class $\mathcal{P}$, if it is real valued, except for the contingency of being $-\infty$ in case of $P \neq Q$. If these two properties hold, then the associated divergence function is $d(P, Q) = S(Q, Q) - S(P, Q)$.

Here we mention two main scoring rules, which will be used in sections below for comparing our evaluation methods. Remember that for an individual, the conditional distribution of the number of accidents is Poisson, so in our notations let p_i $(i = 0, 1, 2, \ldots)$ be $P(X = i|\Lambda = \hat{\lambda})$, i.e., the probabilities of certain claim numbers using the estimated $\hat{\lambda}$ as condition. Similarly, q_i $(i = 0, 1, 2, \ldots)$ is the same probability, but for the real λ frequency.

Two essential score definitions are introduced below. Although, other concepts can be found in the literature, as the spherical score, for instance. The application of the chosen scores is justified by the discrete probabilistic behavior of claim numbers. Note that in case of random variables with uncountable range, one has to apply other measures. It is inevitable if addressing claims severities, see [6, 12], for instance. For more details of density or distribution function forecasts in general see [9, 8].

4.1. Brier Score

Due to the associated Bregman divergence being $d(p,q) = \sum_i (p_i - q_i)^2$, Brier score is sometimes referred as quadratic score. For an analysis regarding precipitation forecasts using Brier scores see [4]. The score is defined as

$$S(P,Q) = 2\left(\sum_i p_i q_i\right) - \left(\sum_i p_i^2\right) - 1. \tag{4.1}$$

Here we use the unknown q_i probabilities deliberately. Although, in practice they are not known, in our simulations presented later in Section 5 we are able to make decisions using them. For the purpose of calculating the score of estimation subsequently, change q_i values to a Dirac delta depending on the number of claims caused. In other words, if the examined policyholder had i claims last year, then it implies the corresponding score to be

$$S(P,i) = 2\left(\sum_j p_j \delta_{ij}\right) - \left(\sum_j p_j^2\right) - 1 = 2p_i - \sum_j p_j^2 - 1.$$

4.2. Logarithmic Score

The logarithmic score is defined as

$$S(P,Q) = \sum_i q_i \log p_i. \tag{4.2}$$

(In case of i caused accidents $S(P,i) = \log p_i$.) We mention that the associated Bregman divergence is the Kullback-Leibler divergence $d(p,q) = \sum_i q_i \log \frac{q_i}{p_i}$.

5. Simulation and results

Example portfolios for testing have been simulated in R statistical program. The simulation technique can be used for frequency evaluation in practice, if we have the appropriate inputs. The main goal is to construct a ranking for given inputs among the estimation methods described in Section 2 and 3 according to scores measures, to support decision making related to the choice of evaluation technique. It is important to emphasize that the methodology might be applied for other actuarial models, as well. Note that the example simulations below are limited to Poisson distributed claim numbers with Gamma mixing distribution for frequencies.

First we need a portfolio containing N insured individuals, which is used to estimate the α and β parameters of the negative binomial distribution. We think of it as the policyholders' histories available in the insurer's database. Henceforth we refer this instituion as *our company.* Taking advantage of the entire claim and bonus-malus history, we have done the parameter evaluation exactly the way as described in Section 2.1. In the next step, we generated the history of a portfolio containing M policyholders, assuming that the distribution parameters are unchanged compared to the first portfolio. This might result some bias, but this is the best we can do based on our available data. The phenomenon occurs in the operation of insurance institutions. On the one hand, the details of the own portfolio are known. On the other hand, the recently acquired (or desired to acquire) portfolio involves only parts of relevant information.

After that, we estimated the claim frequency parameters of policyholders separately based on the three estimation methods described above, and compared them to the real λ parameters using scores. Our aim is to make a decision among the methods, and decide, which would give the best fit results in certain cases. In other words, for certain input parameters, which method yields good-fit estimations in expected value, where the method resulting higher score value means the better goodness-of-fit.

Considering that we are interested in the expected values of scores for given inputs, we will apply a Monte-Carlo-type technique. This means that we generate the above mentioned two portfolios r times independently, but in each first ones preserving the α and β distribution inputs. After that, based on the approximated $\hat{\alpha}$ and $\hat{\beta}$, we estimate the λ parameters of the second portfolios. Each method gives one score number for each simulation, which is the average of scores calculated for individuals. (Of course aggregate scores would be equally appropriate, since it differs only in an M multiplier from the average.) At last we take the mean of mean scores, and the method resulting higher score is the better.

For $i = 1, \dots, r$

- generate portfolio $P_i^{(1)}$ consisting of N individuals, using parameters α, β;
- calculate estimators $\hat{\alpha}^{(i)}, \hat{\beta}^{(i)}$ related to α, β;
- generate portfolio $P_i^{(2)}$ containing M policies, with parameters α, β; it results $\lambda_1^{(i)}, \dots, \lambda_M^{(i)}$ frequencies;

- according to $\hat{\alpha}^{(i)}, \hat{\beta}^{(i)}$ and an estimation method, calculate estimates $\hat{\lambda}_1^{(i)}, \ldots, \hat{\lambda}_M^{(i)}$;
- to each individual j assign score $S_j^{(i)}$, and calculate the mean value $\overline{S}^{(i)} = \frac{1}{M}(S_1^{(i)} + \ldots + S_M^{(i)})$.

In accordance with the above detailed notations, the ultimate score value regarding one estimation method is $S = \frac{\overline{S}^{(1)} + \ldots + \overline{S}^{(r)}}{r}$.

Remark 5.1. *For reference we will write and plot the score results also for comparing the real frequencies to the real frequencies, since these scores are not equal to 0, as in the divergence case. As a function of year steps, these scores should be constant, contrary to the charts below, where small differences can be observed. The reason is that in the Monte-Carlo simulations we generated also the λ parameters over and over.*

Remark 5.2. *In our simulations input parameters are the following:*

1. *α and β distribution parameters. These can be considered as the true underlying parameters, and unknown for the insurance company.*
2. *N the number of policyholders in the first portfolio, which is used to estimate the α and β parameters.*
3. *M the number of policyholders in the second portfolio. This contains individuals, whose λ parameters have to be evaluated. In practice, there might be an overlap between these two files.*
4. *Number of year steps. This means the time elapsed in years, since the certain individual is insured by our company. Note that it implies the knowledge of claims history and bonus classification, thus it is a very important feature, as it affects the goodness-of-fit of our estimation methods.*
5. *Number of years elapsed before entering our company. This affects `method.1` and `method.2`, because the Markov chain on the bonus classes converges slowly to the stationary distribution.*
6. *Transition rules of the examined country.*
7. *Number of simulated portfolios. As we approximate the scores via Monte-Carlo-type technique, it has to be large enough.*

Figure 2 shows an example. We simulated $r = 50$ times a portfolio containing 80 thousand people, estimated α and β parameters, then estimated the λ parameters of 20 thousand policyholders. After that we set the results of the three methods against the real frequencies using scores. For every simulation and country we got 2 scores, the Brier score and Log score. The points of the charts are the average scores for this 50 simulations for given year steps. Certain year steps mean that we generated the second portfolios (the current we are analysing) as we had information about policyholders $1, 2, 5, 10, 15$ and 20 years back, respectively. Intermediate points are approximated linearly. Note that standard deviation of the sample of Brier scores in the Hungarian example are under 0.0009, and under 0.0016 in case of Log scores.

Remark 5.3. *In practice, there are different lengths of claim histories available for different policyholders. In function of this length, we can decide that the parameters of a group of insured people will be evaluated using* `method.2`*, and the rest using* `method.3`*, for example.*

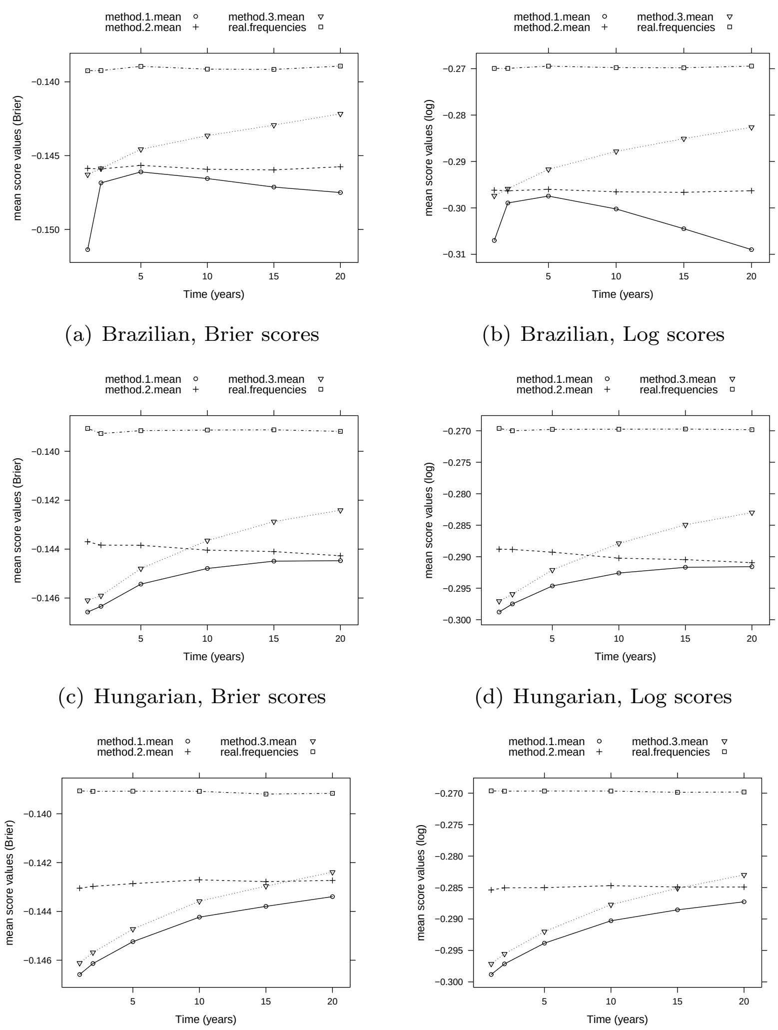
Remark 5.4. *Here (see Figure 2) we let the random walks of policyholders in the systems run for 15 years. Then they are assumed to be acquired by our company, which implies our observations to start at that point. In other words, year steps start at that time, when each driver has already spent some time randomly walking on the transition graph.*

The example clearly shows the differentiation capability of systems containing more bonus-malus classes. In other words, the more classes the system has, the more years needed for `method.3` to get the start of `method.2`. Here the Bayesian type `method.1` is the worst method in every case, but we shall not forget that in certain circumstances it can be useful. For example, if the claim history is largely deficient.

On the tested parameters the two types of scores gave almost the same results (difference is not significant), what we have been expecting. Meaning that if `method.x` is the best according to Brier scores, then it is the best according to Log scores, too.

6. Conclusions

In this article we presented the principle of bonus-malus systems, and the necessary assumptions about the distribution of claim numbers of policyholders, inter alia that the X number of claims caused in a year by an insured person is conditionally Poisson distributed. The goal was to evaluate these λ frequency parameters, which is the expected number of claims caused, and we did not deal with the size of them. Since the unconditional distribution is negative binomial, we can simply evaluate the shape and scale parameters based on the



(a) Brazilian, Brier scores

(b) Brazilian, Log scores

(c) Hungarian, Brier scores

(d) Hungarian, Log scores

(e) Belgian, Brier scores

(f) Belgian, Log scores

Figure 2: Mean Brier and Log scores in the Brazilian, Hungarian and Belgian system, when $N = 80000$, $M = 20000$, with real distribution parameters $\alpha = 1.2$ and $\beta = 14$, and 50 portfolio simulation.

insurer's claim history from past years. Though the second portfolio might have some different parameters, this file consisting of the insurers previously observed claims experiences is the best we can use.

We introduced 3 methods for frequency estimation, where one (`method.1`) was never used by actuaries to our knowledge, and the other two are known. Note that `method.1` is well applicable in case of the Hungarian MTPL insurance (motor third party liability insurance), the reason being that most frequently the data applied by this method are available regarding new policyholders. The main aim of this article was to decide which method is the most appropriate in certain circumstances, i.e., for given parameters. Our decision is made based on scores, which are devoted to measure the bias of two distributions. Assuming `method.x` gives $\hat{\lambda}_1, \ldots, \hat{\lambda}_M$ frequency parameters, and the real ones are $\lambda_1, \ldots, \lambda_M$, then `method.x` is the best choice among the other methods, if the average score is greater than in the other cases. The proposed Monte-Carlo-type algorithm can be used in order to make decisions in practice.

At last, but not least, our method presented in Section 2 includes a technique, which is suitable to compare bonus-malus systems in the following sense. In function of years, the longer the `method.2` is better than others, the more informative is the system, as we expect more accurate evaluation of claims frequencies using the past years' average claim numbers in different classes, than using other methods. For parameters chosen in example 2, in the Brazilian system, `method.3` based on claims history becomes the most appropriate in the second year, while in the Hungarian system it needs 7-8, and in the Belgian 16-17 years. The ranking methodology can be applied using other distributional and model constraints.

7. Appendix

7.1 shows the transition probability matrix of the Hungarian bonus-malus system.

References

[1] **Arató, M. and A. Benczúr,** Dynamic placement of records and the classical occupancy problem, *Computers and Mathematics with Applications,* **7** (2) (1981), 173-185.

$$M(\lambda)=\begin{pmatrix}
1-e^{-\lambda} & e^{-\lambda} & 0 & \dots & 0 & 0 & 0\\
1-e^{-\lambda} & 0 & e^{-\lambda} & \ddots & 0 & 0 & 0\\
1-e^{-\lambda} & 0 & 0 & \ddots & 0 & 0 & 0\\
1-e^{-\lambda} & \lambda e^{-\lambda} & 0 & \ddots & 0 & 0 & 0\\
1-(\lambda+1)e^{-\lambda} & 0 & \lambda e^{-\lambda} & \ddots & 0 & 0 & 0\\
1-(\lambda+1)e^{-\lambda} & \lambda^2\cdot e^{-\lambda}/2! & 0 & \ddots & 0 & 0 & 0\\
1-(\lambda^2/2!+\lambda+1)e^{-\lambda} & 0 & \lambda^2\cdot e^{-\lambda}/2! & \ddots & 0 & 0 & 0\\
1-(\lambda^2/2!+\lambda+1)e^{-\lambda} & \lambda^3\cdot e^{-\lambda}/3! & 0 & \ddots & 0 & 0 & 0\\
1-(\lambda^3/3!+\lambda^2/2!+\lambda+1)e^{-\lambda} & 0 & \lambda^3\cdot e^{-\lambda}/3! & \ddots & 0 & 0 & 0\\
\vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots\\
1-(\lambda^3/3!+\lambda^2/2!+\lambda+1)e^{-\lambda} & 0 & 0 & \ddots & 0 & 0 & e^{-\lambda}\\
1-(\lambda^3/3!+\lambda^2/2!+\lambda+1)e^{-\lambda} & 0 & 0 & \dots & \lambda e^{-\lambda} & 0 & e^{-\lambda}
\end{pmatrix} \tag{7.1}$$

[2] **Denuit, M. and P. Lambert,** Smoothed npml estimation of the risk distribution underlying bonus-malus systems, *Proc. of the Casualty Actuarial Soc. 2001* (to appear)

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

[4] **Ferro, C.A.T.,** Comparing probabilistic forecasting systems with the brier score, *Weather and Forecasting,* **22** (5) (2007), 1076-1088.

[5] **Fishman, G.,** *Monte Carlo: Concepts, Algorithms and Applications,* Springer Series in Operations Research and Financial Engineering, Springer, 1996.

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

[7] **Frees, E.W.,** *Regression Modelling with Actuarial and Financial Applications,* Cambridge University Press, 2010.

[8] **Gneiting, T., F. Balabdaoui and A.E. Raftery,** Probabilistic forecasts, calibration and sharpness, *J. of the Royal Stat. Soc.: Series B (Statistical Methodology),* **69** (2) (2007), 243-268.

[9] **Gneiting, T. and A.E. Raftery,** Strictly proper scoring rules, prediction and estimation, *J. of the American Statistical Association,* **102** (477) (2007), 359-378.

[10] **Lemaire, J.,** *Automobile Insurance,* Kluwer-Nijhoff, 1996.

[11] **Lemaire, J.,** Bonus-malus systems: The European and Asian approach to merit-rating, *North American Actuarial J.,* **2** (1) (1998), 26-38.

[12] **Mahmoudvand, R. and H. Hassani,** Generalized bonus-malus systems with a frequency and a severity component on an individual basis in automobile insurance, *ASTIN Bulletin,* **39** (5) (2009), 307-315.

[13] **Tremblay, l.,** Using the Poisson inverse Gaussian in bonus-malus systems, *ASTIN Bulletin,* **22** (1) (1992), 97-106.

[14] **Walhin, J.F. and J. Paris,** Using mixed Poisson processes in connection with bonus-malus systems, *ASTIN Bulletin,* **29** (1) (1999), 81-100.

Miklós Arató
Department of Probability Theory and Statistics
Eötvös Loránd University
Budapest
Hungary
`aratonm@ludens.elte.hu`

László Martinek
Department of Probability Theory and Statistics
Eötvös Loránd University
Budapest
Hungary
`martinek@cs.elte.hu`