ON THE CONVERGENCE RATE OF
BONUS-MALUS SYSTEMS

By Heikki Bonsdorff

Ministry of Social Affairs and Health
Helsinki, Finland

Abstract

Under certain conditions, a Bonus-Malus system can be interpreted as a Markov chain whose n-step transition probabilities converge to a limit probability distribution. In this paper, the rate of the convergence is studied by means of the eigenvalues of the transition probability matrix of the Markov chain.

Keywords

Bonus-Malus systems; rate of convergence; Markov chains; eigenvalues; automobile insurance.

1. Introduction

We consider Bonus-Malus systems (BMS) as Markov chains. For a general background the reader is referred e.g. to Loimaranta (1972), Norberg (1976), Prins and Roozenboom (1982) and Lemaire (1985). It is well known that, under certain conditions, the n-step transition probabilities between bonus classes converge to an invariant (stationary) probability distribution, when n tends to infinity. The aim of this paper is to study the convergence rate by means of the eigenvalues of the transition probability matrix.

The stationary distribution is often used as a basis for different characteristics associated to the BMS, e.g. the asymptotic expected premium of a driver and the efficiency (by Loimaranta) of the BMS, as well as for evaluations of the future occupation of bonus classes. Consequently, it is useful to study the rate of convergence to the stationary situation. The convergence rate also reflects the sensitivity to changes in the claims intensity. This sensitivity is important especially from the policyholders' point of view because the policyholder normally stays in the BMS his or her whole driving time. Further, when changing the Bonus-Malus system the convergence rate of the new system gives information on how fast the new stationary situation will be obtained.

In Chapter 2 we present a known result of convergence rate for finite Markov chains and apply this result to Bonus-Malus systems. In Chapter 3 we give some examples.

2. CONVERGENCE OF BONUS-MALUS SYSTEMS

We assume that the BMS has a finite set of bonus classes, $S = \{1, \ldots, N\}$. Further, it is assumed that the BMS has a "superbonus" class s, i.e. starting from any class, after a sufficient number of consecutive claim-free years the policy is in class s. In addition, we assume that the BMS is such that the yearly transitions are determined by the number of claims in the preceding year only. (If this is originally not the case, the above condition can often be satisfied by suitably splitting the bonus classes, cf. e.g. LOIMARANTA, 1972.) Finally, we make the standard assumption that for each policy the number of claims (per year) is Poisson distributed with intensity λ, characteristic of the policy in question. For each fixed policy the intensity is assumed to be independent of time.

Under the assumptions above, the BMS defines for each fixed policy a homogeneous Markov chain having transition probabilities

$$(2.1) \qquad p_{ij}(\lambda) = \sum_{k=0}^{\infty} \frac{e^{-\lambda} \lambda^k}{k!} \, t_{ij}(k) \quad \text{for all} \quad i, j \in S$$

where $\lambda > 0$ is the intensity of the policy in question and

$t_{ij}(k) = 0$ if k claims do not cause transition from bonus class i to class j,

$t_{ij}(k) = 1$ if k claims cause transition from class i to class j.

We denote by $p_{ij}^n(\lambda)$ the n-step transition probabilities of the Markov chain (2.1). It follows from the superbonus class assumption that for all $i \in S$, $p_{is}^n(\lambda) > 0$ for all sufficiently large n. This implies that the Markov chain defined by (2.1) is *uniformly ergodic*, i.e. it possesses a unique invariant probability distribution $(\pi_j(\lambda), j \in S) = \pi(\lambda)$ and the following convergence holds true: For some $a = a(\lambda) < \infty$ and $\rho = \rho(\lambda)$, $0 < \rho < 1$,

$$(2.2) \qquad \max_{i \in S} \sum_{j \in S} |p_{ij}^n(\lambda) - \pi_j(\lambda)| \le a\rho^n \quad \text{for all } n.$$

In what follows we often drop λ from our notation when no risk of ambiguity arises.

The rate of convergence of the n-step transition probabilities p_{ij}^n can be characterized by means of the eigenvalues of the transition probability matrix (2.1), cf. e.g. FELLER (1950) and KARLIN (1966). Under the assumptions above, 1 is a simple eigenvalue of the matrix (p_{ij}) and the other eigenvalues lie inside the unit circle of the complex plane. The rate of convergence in (2.2) can be expressed by means of the second biggest eigenvalue (with respect to the modulus). In fact, denote

$$(2.3) \qquad r(\lambda) = \max \{|\alpha_1(\lambda)|, \ldots, |\alpha_{N-1}(\lambda)|\},$$

where $\alpha_1(\lambda), \ldots, \alpha_{N-1}(\lambda)$ are the eigenvalues, *other than* 1, of the matrix (2.1). (Clearly, $r(\lambda) < 1$.) Then we have the result below.

The constant $r = r(\lambda)$ is the *convergence rate* in (2.2), in the following sense:

1° If $\rho > r$, there exists an $a < \infty$ such that (2.2) holds true,

2° If $\rho < r$, there does not exist an $a < \infty$ such that (2.2) holds true.

This result follows e.g. from the more general treatment by ISAACSON and LUECKE (1978) on denumerable Markov chains by noting that the spectrum of a finite matrix consists of eigenvalues only.

In the following we use the term convergence rate exclusively in the above sense.

Note, that the above result remains valid if the initial states i in (2.2) are replaced by initial probability distributions. In fact, let $\mu = (\mu_1, \ldots, \mu_N)$ be an arbitrary probability distribution on S. Denote μP^n the probability distribution defined by

$$(\mu P^n)_j = \sum_{i \in S} \mu_i p_{ij}^n, \ j \in S.$$

Then one can show (see the Appendix) that (2.2) is equivalent to condition

(2.4) $\displaystyle\sum_{j \in S} |(\mu P^n)_j - \pi_j| \leq a\rho^n$ for all probability distributions μ and for all n.

Let $b_j > 0$, $j \in S$ be the premiums associated with the BMS. Denote by $b_{i,n}$ the premium in year n under condition that the policy is in class i in year 0. Then the k:th moments of $b_{i,n}$ converge to the corresponding moments of the stationary distribution π with the convergence rate $\rho_{i,k} \leq r$. In fact, denote the k:th moment of $b_{i,n}$ by $Eb_{i,n}^k$ and that of the stationary distribution by πb^k. Then we have for an arbitrary $\rho > r$

$$|Eb_{i,n}^k - \pi b^k| = \left| \sum_j (p_{ij}^n - \pi_j) b_j^k \right| \leq \left(\max_j b_j^k \right) \sum_j |p_{ij}^n - \pi_j| \leq \left(a \max_j b_j^k \right) \rho^n.$$

A corresponding result holds true for variances, as well.

In the following, we consider matrix (2.1) as a function of the claims intensity λ, $\lambda > 0$, and define some related concepts. We call the function $\lambda \mapsto r(\lambda)$, defined by equation (2.3), the *convergence rate function* of the BMS. (See Figure 3.1 in Chapter 3.) Further, we call the family of all invariant probability distributions $\pi(\lambda)$, $\lambda > 0$, associated to (2.1), the *overall equilibrium* of the BMS. (See Figure A.1 in the Appendix.) Thus, for a fixed λ, $r(\lambda)$ can be called the *convergence rate of the* BMS *to equilibrium at intensity* λ.

Note. The considerations of this section can be carried out under more general assumptions on the claims process as in case (2.1). We illustrate this by an example associated to a situation where a hunger of bonus effect is present. Preserving the assumption that the basic number of claims process of a policyholder is Poisson with intensity λ, we assume, in addition, that the policyholder has for each bonus class j a retention limit $x_j \geq 0$ and that the

individual claims are i.i.d., cf. LEMAIRE (1985, Ch. 18). Then for each bonus class j the number of claims process is Poisson with intensity $\lambda_j \leq \lambda$. The superbonus assumption guarantees the uniform ergodicity of the Markov chain associated to the BMS in this case, as well. Accordingly, the considerations of this chapter remain valid. (Note that the assumption of fixed retention limits is, of course, simplified. In practice, a reasonable policyholder may change the limit x_j during the insurance year.)

3. EXAMPLES AND COMPARISONS

In this Chapter we give examples of the convergence rate function r (cf. 2.3) for some Bonus-Malus systems. We present the convergence rate functions for the following Bonus-Malus systems in automobile insurance: the *Dutch* BMS, see PRINS and ROOZENBOOM (1982), p. 95, the *Swiss* BMS, see GERBER (1990) (the new version) and the *Finnish* BMS in third party liability insurance (valid in 1990), see the Appendix at the end of this paper.
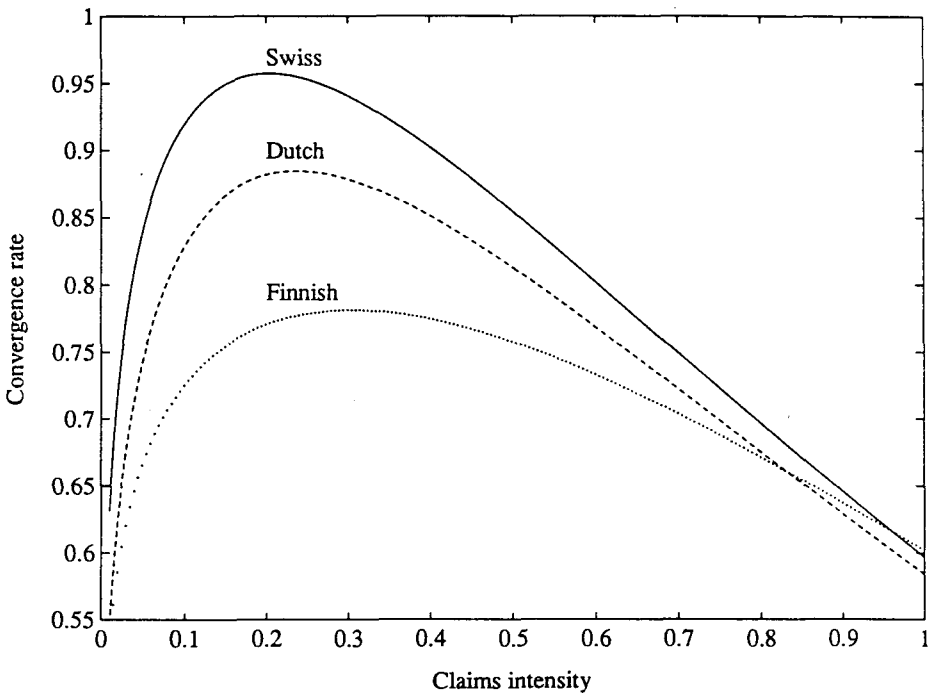

FIGURE 3.1. Convergence rate functions of certain Bonus-Malus systems; $0.01 \leq \lambda \leq 1$.

It can be seen in the figure that in each system the convergence rate for an ordinary driver is close to the slowest convergence rate of the system and that the convergence rate functions of the three systems differ considerably from each other.

The convergence rate function gives information about the maximal (with respect to the initial class i) distance of the distributions (p_{ij}^n) and (π_j), $j \in S$, in the sense of (2.2), for large values of n.

However, in practice, the magnitude of the rate is often reflected on the distance of these distributions also for "small" values of n and for fixed initial states. Let us illustrate this by an example. Let $\lambda = 0.12$. Consider the total variation

$$(3.1) \qquad\qquad \sum_j |p_{ij}^n - \pi_j|$$

in the following two cases:

1° when i is the initial state for new drivers in the BMS ($i = 3$ for the Dutch system),

2° when i is the class with the highest premium.

Taking $n = 25$ we get the following table.

TABLE 3.1
VALUES OF TOTAL VARIATION (3.1) FOR $n = 25$

	Case 1°	Case 2°
Finnish BMS	0.004	0.006
Dutch BMS	0.069	0.090
Swiss BMS	0.301	1.073

(The maximal possible value of (3.1) for any two probability distributions is, of course, equal to 2.)

In case 1° it takes 16 years for the Finnish system before (3.1) is less than 0.1, 23 years for the Dutch and 39 for the Swiss one, and in case 2°, correspondingly, 17 for the Finnish, 25 for the Dutch and 61 for the Swiss system.

Note that in the Swiss system the beginner starts from a relatively good class. As a consequence, the beginner will not be strongly charged in the Swiss system in spite of the slow convergence rate. The situation may be worse for a former bad driver who has changed into a good one. On the opposite, a high rate (small ρ) of convergence may be a sign of too mild handling of bad drivers.

APPENDIX

In this Appendix we present the Finnish BMS, illustrate the concept overall equilibrium and give a proof concerning convergence (2.4).

The table below should be read as follows: $a_{ij} = k$ means that k claims cause transition from class i to class j, and $a_{ij} = k+$ that a number of claims $K \geq k$ causes transition from class i to class j.

The Finnish BMS

	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1	1+			0										
2	1+			0										
3	2+	1			0									
4	2+	1			0									
5	2+			1		0								
6	3+	2			1		0							
7	3+	2			1			0						
8	3+			2		1			0					
9	3+			2		1				0				
10	4+	3			2		1				0			
11	4+	3			2		1					0		
12	4+	3			2		1						0	
13	4+	3			2		1							0
14	4+	3			2			1						0

$$= \begin{bmatrix} a_{1,1}, \ldots, a_{1,14} \\ \\ \\ \\ \\ \\ \\ \\ a_{14,1}, \ldots, a_{14,14} \end{bmatrix}$$

In Figure A.1 below the invariant probability distribution $\pi(\lambda)$ has been calculated, for each fixed λ, by norming the left eigenvector associated to the eigenvalue 1 of the transition probability matrix $(p_{ij}(\lambda))$, cf. (2.1), of the Dutch BMS.
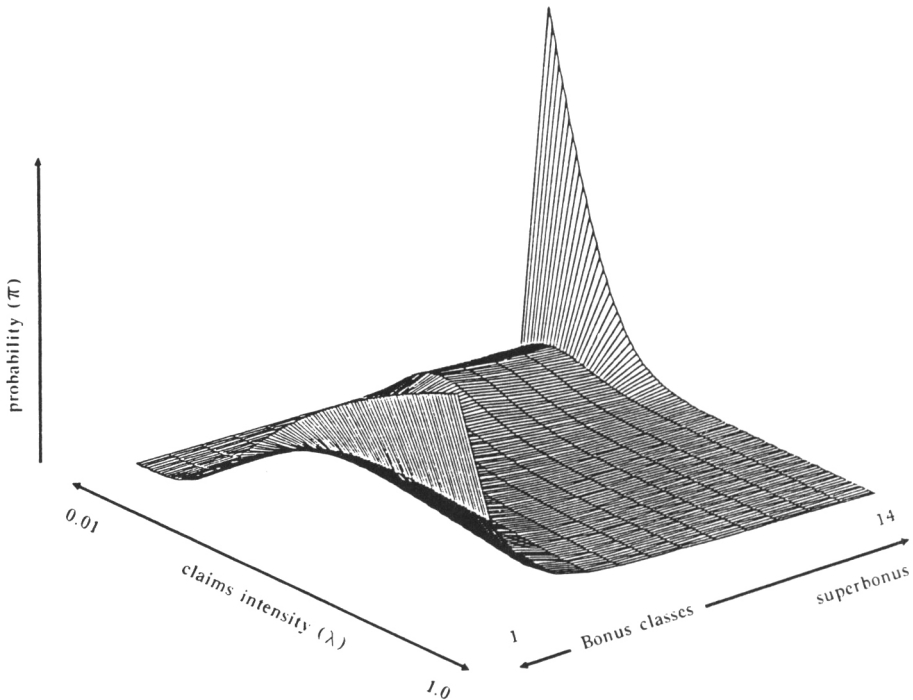

FIGURE A.1. Overall equilibrium of the Dutch BMS.

Finally, we prove that r defined by (2.3) is the convergence rate in (2.4). Recall that r is the convergence rate in (2.2). Clearly, (2.4) implies (2.2). Thus, it is sufficient to show that (2.2) implies (2.4).

Assume now (2.2). Let (μ_i) be an arbitrary probability distribution on S.

Since $\sum_i \mu_i = 1$, we get

$$
\begin{aligned}
\sum_j |(\mu P^n)_j - \pi_j| &= \sum_j \left| \sum_i \mu_i p_{ij}^n - \pi_j \right| \\
&= \sum_j \left| \sum_i \mu_i p_{ij}^n - \sum_i \mu_i \pi_j \right| \le \sum_j \sum_i \mu_i |p_{ij}^n - \pi_j| \\
&= \sum_i \left(\mu_i \sum_j |p_{ij}^n - \pi_j| \right) \le a\rho^n.
\end{aligned}
$$

ACKNOWLEDGEMENTS

REFERENCES

FELLER, W. (1950) *An Introduction to Probability Theory and its Applications*, Vol I. Wiley, New York.

GERBER, H. U. (1990) The new Bonus-Malus system of Switzerland. *Presented in the XXII ASTIN Colloquium*, in Montreaux, 2.27–2.29.

ISAACSON, D. and LUECKE, G. R. (1978) Strongly ergodic Markov chains and rates of convergence using spectral conditions. *Stochastic Processes Appl.* 7, 113–121.

KARLIN, S. (1966) *A First Course in Stochastic Processes*. Academic Press, New York.

LEMAIRE, J. (1985) *Automobile Insurance*. Kluwer-Nijhoff, Boston.

LOIMARANTA, K. (1972) Some asymptotic properties of bonus systems. *ASTIN Bulletin* 6, 233–245.

NORBERG, R. (1976) A credibility theory for automobile bonus systems. *Scand. Actuarial J.*, 92–107.

PRINS, H. J. and ROOZENBOOM, F. (1982) Bonus/Malus. In: *New Motor Rating Structure in the Netherlands.* (ed. DE WITT, G. W.), pp. 72–104. Publ. Astin-groep Nederland.

HEIKKI BONSDORFF
Ministry of Social Affairs and Health, Hallituskatu 5,
00170 Helsinki, Finland.