3. ARTÍCULOS DE APLICACIÓN

APPLICATION OF THE MARKOV PROCESSES THEORY IN AUTOMOBILE INSURANCE

Yongsheng Xing[1], **Shixia Ma**[2]
[1] College of Mathematics and information
Shandong Institute of Business and Technology, China
[2] School of Sciences
Hebei University of Technology, China

1. Introduction

The automobile insurance is an important branch of non-life insurance. Considering that an especial feature of automobile insurance is that the companies always adjust their premiums taking into account the past claims history, the bonus-malus system has been widely used in the European and Asian countries as a risk classification method. In this work, as an application of Markov process theory in insurance problems, we will show how an automobile insurance problem can be probabilistically modelled through a Markov chain. The work is structured as follows: in section 2, the probability model is formally and intuitively introduced; some illustrative examples are given in section 3 and, finally, a modified bonus-malus model is considered in section 4.

2. The probability model

Up to minor modifications, most automobile insurance companies employ the so called bonus-malus system. There is a finite number k of classes (tariff groups) and the premium depends of the class which the policy-holder belongs. Each year the class is determined on the basis of the class of the previous year and taking into account the number of reported claims during the current year. If no claim has been reported then the policy-holder gets a bonus expressed in the lowering to a class with a possibly lower premium. Depending on the number of reported claims, the policy-holder gets malus expressed by a shift to a higher class. Formally, we consider k classes numbered by $1, 2, \ldots, k$; we call class 1 superbonus and class k supermalus; the annual premium depends on the number of the current class and it is computed from a given premium scale $(b_1, b_2, \ldots, b_k)$, where $b_1 \leq b_2 \leq \ldots \leq b_k$. Transition rules, which say how to pass from one class to another, are determined once the number of claims is known. In fact, if l claims are reported then we have the matrix $(t_{ij}(l))_{i,j=1,2,\ldots,k}$ where, for each $i \in \{1, \ldots, k\}$, $t_{ij}(l) = 1$ if the policy-holder is transferred from the class i to the class j and $t_{ij}(l) = 0$ otherwise. Given a policy-holder, let us introduce the sequences of random variables $\{Y_n\}_{n\geq 1}$ and $\{X_n\}_{n\geq 0}$. The variable Y_n represents the cost associated to the reported claims by the policy-holder during the nth year, we will assume that $\{Y_n\}_{n\geq 1}$ is a sequence of independent and identically distributed variables, its common probability law will be denoted by $\{q_l\}_{l\geq 0}$, i.e. $q_l = P(Y_1 = l),\ l = 0, 1, \ldots$ The sequence $\{X_n\}_{n\geq 0}$ informs about the year-by-year classes corresponding to the policy-holder. Since we are assuming that the class for the next year is uniquely determined by the class of the preceding year and by the number of claims reported during the current year, one deduces that $\{X_n\}_{n\geq 0}$ is a homogeneous Markov chain with state space $\{1, \ldots, k\}$. If p_{ij} denotes the transition probability that the policy-holder passes from the state i to the state j then, one derives that $p_{ij} = \sum_{l=0}^{\infty} q_l t_{ij}(l)$. Usually, the number of claims reported by the policy-holder is assumed to be distributed according to a Poisson law, the mean λ possibly will depend of the policy-holder. In such a situation:

$$p_{ij} = p_{ij}(\lambda) = \sum_{l=0}^{\infty} \frac{\lambda^l}{l!} e^{-\lambda} t_{ij}(l), \quad i, j = 1, 2, \ldots, k.$$

For example, assuming $k = 5$, $q_l = \lambda^l (l!)^{-1} e^{-\lambda}$, $l = 0, 1, \ldots$, that for each claim reported the policy-holder goes one state up (if there are three or more claims, the policy-holder goes to the highest class direct) and fails to the lower state otherwise and considering that the occurrences of claims in different policy years are independent events then, one deduces the transition probabilities matrix:

$$\begin{pmatrix} q_0 & q_1 & q_2 & 0 & 1-q_0-q_1-q_2 \\ q_0 & 0 & q_1 & q_2 & 1-q_0-q_1-q_2 \\ 0 & q_0 & 0 & q_1 & 1-q_0-q_1 \\ 0 & 0 & q_0 & 0 & 1-q_0 \\ 0 & 0 & 0 & q_0 & 1-q_0 \end{pmatrix} \tag{2.1}$$

Clearly, $\{X_n\}_{n\geq 0}$ is an irreducible and aperiodic Markov chain and it has unique stationary distribution. Let S_n be the accumulated amount related to the claims reported in the first n years, namely

$$S_n = \sum_{i=1}^{n} Y_i$$

where recall that Y_i represents the aggregate amount associated to the claims for the book of insurance business in the ith policy year (note that it is possible that $q_0 = P(Y_i = 0) > 0$). Also, let us denote by C_i the premium that the driver paid in the ith policy year, which will depend of the last class of the driver in the bonus-malus system. The discrete-time surplus process or risk process for a book of insurance business at year n is defined (see [1], p. 83) in the form:

$$U(u,n) = u + \sum_{i=1}^{n} C_i - S_n, \quad n = 0, 1, \ldots$$

where $u \geq 0$ represents the initial surplus of the insurance company. Let

$$\begin{aligned} T(u) &= min\{n : \; U(u,n) \leq 0\}, \\ \Psi_j(u) &= P(T(u) < \infty \mid X_0 = j). \end{aligned} \tag{2.2}$$

Intuitively, $T(u)$ and $\Psi_j(u)$ represent, respectively, the ruin time and the probability of ruin in a finite time assumed that initially j is the class in the bonus-malus system. The probability of ruin enables one to compare portfolios with each other, but we can not attach any absolute meaning to such probability, as it doesn't actually represent the probability that the insurer will go bankrupt in the near future. First of all, it might take centuries for ruin to actually happen. Moreover, potential interventions in the process, for instance paying out dividends or raising the premium for risks with an unfavourable claims performance, are ruled out in the determination of the probability of ruin. Furthermore, the effects of inflation on the one hand and the return on the capital on the other hand are supposed to cancel each other out exactly. The ruin probability only accounts for the insurer risk, not the managerial blunders that might occur. Finally, the state of ruin is nothing but a mathematical abstraction (with a capital of -1 euro, the insurer is not broken in practice and with a capital of $+1$ euro, the insurer can hardly be called solvent). The calculation of the probability of ruin is one of the classical problems in actuarial science. A good deal has been written on ruin theory in the case where the premiums are received at a constant rate. Fewer papers in the literature consider the case of a varying premium rate. We will consider this problem in section 4 for an especial bonus-malus system. For the analysis of bonus-malus systems, the main problems are (see [2], p. 278): to determine the probability that in the nth year the policy-holder be in the state j and the expected accumulated premium paid by the policy-holder over the period of n years. In this paper, we will focus our interest about the following questions:

(a) To determine the initial class for a new policy-holder, that is, to decide how much premium should be paid for a new policy-holder in the first year.

(b) To study whether it is profitable for a policy-holder not to report small claims in order to avoid an increasing in the premium.

(c) To calculate the corresponding ruin probability for the insurance model.

The ultimate goal of a bonus-malus system is to make that the policy-holder pays a premium which be as near as possible the expected value of his (her) yearly claims. The expected value of the asymptotic premium to be paid is called the steady state premium, and it should be as the premium for new policy-holder in the first year. We will consider that the class which the policy-holder will be next year does not heavily depend on the initial class, it only depends on the claims filed during the last year.

3. Some illustrative examples

As illustration, next we provide two examples: **Example 1:** Let us consider a sequence $\{X_n\}_{n\geq 0}$

with the transition probabilities matrix given in (2.1) and let us denote by $\overrightarrow{\pi} = (\pi_1, \pi_2, \ldots, \pi_5)$ the associated stationary probability distribution, namely

$$\pi_j = \lim_{n\to\infty} p_{ij}^{(n)}, \ j = 1, \ldots, 5 \qquad (3.1)$$

where $p_{ij}^{(n)}$ is the n-step transition probability from state i to state j. We know that $\overrightarrow{\pi}$ is the unique solution of the system:

$$\pi_j = \sum_{i=1}^{5} \pi_i p_{ij}, \quad j = 1, \ldots, 5$$

Thus, the steady state premium will be:

$$b_0 = \sum_{i=1}^{5} b_i \pi_i.$$

For instance, if we consider the premium scale

$$b_1 = 100, \ b_2 = 120, \ b_3 = 130, \ b_4 = 150, \ b_5 = 160$$

then, it is easy to derive that: For $\lambda = 0{,}2$,

$$\overrightarrow{\pi} = (0{,}757; \ 0{,}168; \ 0{,}053; \ 0{,}016; \ 0{,}006)$$

and $b_0 \approx 106{,}14$ hence the initial class for a new policy-holder should be the class 1. For $\lambda = 0{,}5$,

$$\overrightarrow{\pi} = (0{,}318; \ 0{,}206; \ 0{,}181; \ 0{,}155; \ 0{,}140)$$

so $b_0 \approx 125{,}73$ hence the initial class for a new policy-holder should be the class 2 or 3. For $\lambda = 1$,

$$\overrightarrow{\pi} = (0{,}033; \ 0{,}056; \ 0{,}119; \ 0{,}253; \ 0{,}539)$$

and $b_0 \approx 149{,}69$ hence the initial class should be the class 4. Finally, for $\lambda = 1{,}5$,

$$\overrightarrow{\pi} = (0{,}003; \ 0{,}011; \ 0{,}046; \ 0{,}186; \ 0{,}753)$$

so $b_0 \approx 156{,}10$ hence the initial class should be the class 5.

Suppose that a driver, in order to earn more bonus in the near future, has the possibility of not filing small claims. In this case, the question is: When exactly is it profitable for him (her) to file his (her) claims? Of course, it is related with the driver's position (the scale class) and the probability of having one or more claims.

Example 2: Let us consider a driver with the bonus-malus system given in example 1. Suppose that, in state 1, he (she) causes a damage of size t in an accident. If the driver is not obliged to file this claim with the insurance company, when exactly is it profitable for him (her) to do so? We shall assume that, as some policies allow, the driver only has to decide on December 31st whether to file this claim, so it is certain that he (she) has not claims after this one concerning the same policy year. Since the effect of this particular claim on his (her) position on the bonus-malus scale will not vanished for a longtime (this is determined by convergence rate of (3.1)) in this paper, we only use a planning horizon of two years. His (her) costs in the coming two years (premiums plus claim) will depend on whether or not he (she) files the claim and whether he (she) is claim-free next year, so there are two possibilities:

(1) The claim is not filed. Then next year he (she) is also in the state 1 and the cost will be $c_0 = 200 + t$, $c_1 = 220 + t$, $c_2 = 230 + t$ or $c_3 = 260 + t$ if no claim, one claim, two claims or more than two claims, respectively, are reported during such a year.

(2) The claim is filed. Then next year he (she) is in state 2, and the cost will be $c_0^* = 220$, $c_1^* = 250$, $c_2^* = 270$ or $c_3^* = 280$ if no claim, one claim, two claims or more than two claims, respectively, are reported during such a year.

The driver should only file the claim if

$$\sum_{i=0}^{3} c_i q_i \geq \sum_{i=0}^{3} c_i^* q_i, \quad q_3 = 1 - q_0 - q_1 - q_2$$

or equivalently if $t \geq 20 + 10q_1 + 20q_2$.

For example: If $\lambda = 0{,}2$ then $q_1 = 0{,}164$, $q_2 = 0{,}016$ and one deduces that $t \geq 21{,}96$. If $\lambda = 0{,}5$ then $q_1 = 0{,}303$, $q_2 = 0{,}076$ so $t \geq 24{,}55$. If $\lambda = 1{,}0$ then $q_1 = 0{,}368$, $q_2 = 0{,}184$ hence $t \geq 27{,}36$. Finally, if $\lambda = 1{,}5$ then $q_1 = 0{,}335$, $q_2 = 0{,}251$ and one derives that $t \geq 28{,}37$.

We see that it is unwise to file very small claims, considering the loss of bonus in the near future. On the one hand, the insurer misses premiums that are his (her) due, because the insured in fact conceals that he (she) is a bad driver. But this is compensated by the fact that small claims also involve handling costs.

4. Some alternative models

The model used in the preceding sections could be much refined involving, for instance, a longer or infinite time-horizon, with discounting. Also the time in the year that a claim occurs is very important. Many articles have appeared in the literature, both on actuarial science and on stochastic operational research, considering these possibilities. For example, [1] introduces a modified bonus-malus system considering that a driver pays a high premium c if he (she) files claims in either of the two preceding years, otherwise pays a, where $a < c$. In consequence, we have a bonus-malus system with three possible states ($k = 3$):

1: Claim-free in the two latest policy years. Then the policy-holder has to pay a.

2: No claim in the previous policy year and claim in the year before. Then the policy-holder has to pay c.

3: Claim in the previous policy year and either claim or no claim in the year before. Then the policy-holder has to pay c.

In this section, we shall consider a risk model with this bonus-malus system. Also, we shall suppose that the driver has one or more claims in a policy year with probability p, ($q = 1 - p$ will be the probability of no claim) and that the occurrences of having claims in different policy years are independent events. Then, the transition probabilities matrix will be:

$$\begin{pmatrix} q & 0 & p \\ q & 0 & p \\ 0 & q & p \end{pmatrix} \tag{4.1}$$

It is easily verified that $\{X_n\}_{n\geq 0}$ is an irreducible and aperiodic Markov chain, its unique stationary distribution is given by: $\pi_1 = q^2$, $\pi_2 = pq$ and $\pi_3 = p$.

We shall consider a discrete-time insurance model with the above bonus-malus system. The surplus process of this model is given, for $n = 0, 1, \ldots$ by

$$U(k,n) = k + \sum_{i=1}^{n} C_i - \sum_{i=1}^{N(n)} Y_i \tag{4.2}$$

where k is a nonnegative integer which represents the initial surplus of the insurance company; C_i is the premium offered in the ith policy year, taking values a or c depending on the claim history of two preceding years; $N(n)$ is a binomial process which implies the number of policy years having one or more claims; Y_i denotes the cost associated to claims in the ith policy year which there are claims, these random variables are assumed to be positive, independent and identically distributed. By simplicity, we shall write $Y = Y_1$, $p_k = P(Y = k)$, $k = 1, 2, \ldots$, ($p_0 = 0$) and $\mu = E[Y]$. The ruin time $T(k)$ and the ruin probabilities $\Psi_j(k)$ (where recall that j is the initial state) corresponding to the model with transition matrix (4.1) are defined as in (2.2). Sometimes, it is more useful to consider the so called survival probability $\varphi_j(k) = 1 - \Psi_j(k)$. Without loss of generality, we shall consider that $a = 1$ and $c = 2$. Moreover, we shall assume that

$$\pi_1 + 2\pi_2 + 2\pi_3 - p\mu = 1 + (2 - \mu)p - p^2 > 0 \tag{4.3}$$

Since initially $k \geq 0$, note that, according to this bonus-malus system, in the case of no claims, a driver in state 1 or 2 will be next year in state 1 and a driver in state 3 will be next year in state 2. On the other hand, if there are claims then next year the driver will be in state 3 independently of his (her) current state. In consequence, one has for $k = 0, 1, \ldots$

$$\begin{aligned} \varphi_1(k) &= q\varphi_1(k+1) + pE[\varphi_3(k+1-Y)], \\ \varphi_2(k) &= q\varphi_1(k+2) + pE[\varphi_3(k+2-Y)], \\ \varphi_3(k) &= q\varphi_2(k+2) + pE[\varphi_3(k+2-Y)]. \end{aligned} \tag{4.4}$$

Hence, for $k = 0, 1, \ldots$

$$\begin{aligned} \varphi_2(k) &= \varphi_1(k+1), \\ \varphi_3(k) &= \varphi_2(k) + q(\varphi_2(k+2) - \varphi_1(k+2)) \\ &= \varphi_1(k+1) + q(\varphi_1(k+3) - \varphi_1(k+2)) \end{aligned} \tag{4.5}$$

and, consequently, it is sufficient to obtain $\varphi_1(k)$, $k = 1, 2, \ldots$ To this end, given a sequence of real number $\{a_k\}_{k\geq 0}$ we recall that its associated generating function is defined in the form:

$$A(s) = \sum_{k=0}^{\infty} a_k s^k.$$

Clearly, if $\{a_k\}_{k\geq 0}$ is bounded then $A(s)$ converges for $|s| < 1$. We say that $\{c_k\}_{k\geq 0}$ is the convolution of $\{a_k\}_{k\geq 0}$ and $\{b_k\}_{k\geq 0}$, and we denote $c_k = a_k * b_k$, if

$$c_k = \sum_{i=0}^{k} a_i b_{k-i}, \quad k = 0, 1, \ldots$$

If $A(s)$ and $B(s)$ (assumed to be convergent for $|s| < 1$) are the associated generating functions of $\{a_k\}_{k\geq 0}$ and $\{b_k\}_{k\geq 0}$, respectively, then it can be verified that the generating function associated to its convolution $\{c_k\}_{k\geq 0}$ is given by

$$C(s) = A(s)B(s), \quad |s| < 1.$$

In particular, let us consider the sequences $\{a_k\}_{k\geq 0}$ and $\{b_k\}_{k\geq 0}$ where, for $k = 0, 1, \ldots$

$$a_k = \sum_{j=k}^{\infty} p_{j+2} - qp_{k+3} \quad \text{and}$$

$$b_k = q - p^2 p_{k+3} + p\sum_{x=0}^{k} p_{x+3} + pq^{-1}\sum_{x=0}^{k-1} p_{x+2}$$

Taking into account (4.4) and (4.5) it can be deduced (see [3] for details) that

$$\varphi_1(0) = q(1 + (2-\mu)p - p^2),$$

$$\varphi_1(k) = \varphi_1(0)\sum_{n=0}^{\infty} p^n \{a_k\}^{*n} * b_k, \; k = 1, 2, \ldots$$

(by simplicity, g^{*n} denotes the n-fold convolution of g, $g^{*0} = 1$).

As illustration, we shall consider the degenerate case $p_3 = P(Y = 3) = 1$. In such a situation, condition given in (4.3) is

$$0 < p < \frac{\sqrt{5}-1}{2}$$

and it can be derived that $\varphi_1(0) = 1 - 2p + p^3$, $a_0 = p$, $b_0 = 1 - p^2$, $a_1 = b_1 = 1$, and $a_k = 0$, $b_k = (1-p)^{-1}$, $k = 2, 3, \ldots$ We deduce that the generating functions associated to the sequences $\{a_k\}_{k\geq 0}$ and $\{b_k\}_{k\geq 0}$ are, respectively, $A(s) = p + s$ and

$$B(s) = 1 - p^2 + s + \frac{1}{1-p}\cdot\frac{s^2}{1-s}.$$

Consequently, if $\Psi(s)$ denotes the generating function associated to $\{\varphi_1(k)\}_{k\geq 0}$ then,

$$\begin{aligned}\Psi(s) &= B(s)(1 - pA(s))^{-1}\varphi_1(0)\\ &= \varphi_1(0) + \sum_{k=1}^{\infty}(1 - \frac{p^k(1+p)}{(1-p^2)^{k-1}})s^k.\end{aligned}$$

hence,

$$\varphi_1(k) = 1 - \frac{p^k(1+p)}{(1-p^2)^{k-1}}, \quad k = 1, 2, \ldots$$

By (4.5), one derives

$$\varphi_2(0) = 1 - p - p^2, \quad \varphi_3(0) = \frac{1-p-p^2}{1-p^2} \quad \text{and}$$

$$\varphi_2(k) = 1 - \frac{p^{k+1}(1+p)}{(1-p^2)^k}, \; \varphi_3(k) = 1 - \frac{p^{k+1}}{(1-p^2)^{k+1}},$$

$k = 1, 2, \ldots$

Acknowledgements: This research has been supported by the Natural Science Foundation of China, grant number: 10771119.

References

[1] Kaas R., Goovaerts M., Dhaene J. and Denuit M. (2003). *Modern Actuarial Risk Theory.* Kluwer Academic Publishers.

[2] Rolski T., Schmidli H., Schmidt V. and Teugels J. (1999). *Stochastic Processes for Insurance and Finance.* John Wileyand Sons.

[3] Xing Y. (2007). The survival probability of an insurance model with bonus-malus system. *In referee process in Acta Mathematicae Applicatae Sinica, English Series.*