APPLICATION OF MARKOV CHAIN IN ACTUARIAL MODELLING

1. INTRODUCTION

In this research Markov processes are used for modelling a phenomenon in insurance, which changes over time of a random variable comprise a sequence of values in the future. Each of values depends only on the immediately preceding state, not on other past states. A Markov process is completely characterized by specifying the finite set S of possible states and the stationary probabilities of transition between these states. Random processes are of interest for describing the behaviour of a system evolving over period of time, hence they were greatly applied in the actuarial mathematics and enabled us to deal with very complicated actuarial problems.

Markov processes can be observed in dicreate-time and a continuous time. A discrete-time random process involves a system which is in a certain state at each step, with the state changing randomly between steps. The Markov property states that the conditional probability distribution for the system at the next step (and in fact at all future steps) depends only on the current state of the system, and not additionally on the state of the system at previous steps. Since the system changes randomly, it is generally impossible to predict with certainty the state of a Markov chain at a given point in the future.

A continuous-time Markov chain is a mathematical model which takes values in some finite or countable set. The time spent in each state takes non-negative real values and has an exponential distribution. It is a random process with the Markov property which means that future behavior of the model depends only on the current state of the model and not on historical behavior. The model is a continuous-time version of the Markov chain model, named because the output from such a process is a sequence (or chain) of states.

This research illustrates how the mathematics of Markov Processes can be used, in the actuarial modelling and calculation. First the possibility of application of the Markov chain was shown on the Belgrade Stock Exchange (BSE) in order to forecast stock prices and return of stocks, as an important part of the investment strategies. Markov chains are a simple non-parametric method with application in the stock market analysis, however, insufficiently researched in the area of return modelling. In a review of lot of articles published on this subject, it was found that no one provides the possibility to apply it on the BSE. Special attention in this research was brought to a bonus-malus system, as it can be considered as a special case of Markov processes. The nature of life assurance indicates to application of Markov chains. Examples show the modelling method for term life assurance as well as for the disability model.

2. MARKOV CHAINS

Important role in description of various events in the nature belongs to discrete Markov processes with a discrete parameter. They are called Markov chains (MC). These are discrete random processes $X(t), t \in T = \{0,1,2,...\}$ with finite number of states, i.e. possible states $S = \{1,2,...,N\}$. If $X(t) = i$ then the process in time t and state i.

Definition 1. (Markov property) Future development of the process at time t+1 depends only on state of the process at time t, and not on past development times. For all t=0,1,2,... and all states $i, j, i_{t-1},....,i_0 \in S$ is

$P(X_{t+1} = j \mid X_t = i, X_{t-1} = i_{t-1}, \ldots, X_0 = i_0) = P(X_{t+1} = j \mid X_t = i)$.

This process is called a Markov chain.

Markovian conditional probability of any future event, given any past event and the present state X event and depends only upon the present state.

The conditional probabilities $P(X_{t+1} = j / X_t = i\} = p_{ij}(t, t+1)$ are called transition probabilities from state i at time t to state j at time t+1.

The conditional probabilities $P(X_{t+s} = j / X_t = i\} = p_{ij}(t, t+s)$ are called transition probabilities from state i at time t to state j at time t+s.

Definition 2. If $\{X_t;\ t \in [0;\infty)\}$ is a discrete Markov process and if probabilities of transition $p_{ij}(t, t+s)$ do not depend on t and t+s, but only on the difference s, thus random process is called a homogenous Markov process.

We introduce the following notation:

$P(X(0) = i) = p_i(0)$, initial probability of state i,

$(p_1(0), p_2(0), ..., p_n(0)) = p(0)$, initial distribution of MC states,

$P(X(t) = i) = p_i(t)$, absolute probability of state i at time t,

$\left(p_1(t), p_2(t), ..., p_n(t)\right) = p(t),$ absolute distribution of MC states at time t.

Probability of transition after one step, $p_{ij}\left(t, t+1\right)$, can be arranged in the form of a *n* x *n* matrix known as the Transition Probability Matrix is given by

$$P = \begin{pmatrix} p_{11} & p_{12} & \cdots & p_{1n} \\ p_{21} & p_{22} & \cdots & p_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ p_{n1} & p_{n2} & \cdots & p_{nn} \end{pmatrix}$$

Transition matrix $P = (p_{ij} : i, j \in S)$ with $p_{ij} \geq 0$ for all i, j, is a stochastic matrix, meaning that $p_{ij} \geq 0$ for all i, j∈S and $\sum_{i=1}^{n} p_{ij} = 1$ (i.e. each row of P is a distribution over S). Transition Probability Matrix provides a precise description of the behaviour of a Markov chain. Each element in the matrix represents the probability of the transition from a particular state to the next state. The transition probabilities are usually determined empirically, that is based solely on experiment and observation.

Given the initial distribution p(0), let us treat it as a row vector, and using matrix P, we can describe dynamics of the process X(t)

$$P(X(0) = i_0, X(1) = i_1, ..., X(k) = i_k) = p_{i_0}(0) p_{i_0 i_1} \cdots p_{i_{k-1} i_k}$$

We denote transition probability from the state i to the state j after s steps

$$p_{ij}^{s} = p_{ij}\left(t, t+s\right) \text{ for } s = 1, 2, ..., \; p_{ij}^{0} = \begin{cases} 0, & i \neq j \\ 1, & i = j. \end{cases}$$

and transition probability matrix of the homogeneous Markov chain after s steps

$$P^{s} = \begin{pmatrix} p_{11}^{s} & p_{12}^{s} & \cdots & p_{1n}^{s} \\ p_{21}^{s} & p_{22}^{s} & \cdots & p_{2n}^{s} \\ \vdots & \vdots & \vdots & \vdots \\ p_{n1}^{s} & p_{n2}^{s} & \cdots & p_{nm}^{s} \end{pmatrix}$$

$$p_{ij}^{(2)} = P(X(t+2) = j \,|\, X(t) = i$$

$$= \sum_{k=1}^{n} P(X(t+2) = j \,|\, X(t+1) = k, X(t) = i) P(X(t+1) = k \,|\, X(t) = i)$$

$$= \sum_{k=1}^{n} P(X(t+2) = j \mid X(t+1) = k), P(X(t+1) = k \mid X(t) = i)$$

$$= \sum_{k=1}^{n} p_{kj} p_{ik} = P^{(2)} = P^2.$$

Similarly, continuing process we get the matrix $P^{(s)} = P^s$.

For initial vector

$$p_i(t) = P(X(t) = i) = \sum_{k=1}^{n} P(X(0) = k) \mid P(X(t) = i) \mid X(0) = k)$$

$$= \sum_{k=1}^{n} p_k(0) p_{ki}^t = p(0)P^{(t)} = p(0)P^t$$

In other words, the n-step transition probability of a Markov chain is the probability that it goes from state i to state j in n transitions.

In the literature the terms equilibrium, stationary, and steady state are used to mean the same thing.

In terms of long-term of the chain X(t) it is useful to determine absolute probability of states $p_i(t)$ for large t $(t \to \infty)$, then

$\lim\limits_{x\to\infty} p_{ik}^{(t)} = \pi_k$ and $\lim\limits_{x\to\infty} p_k(t) = \pi_k$, $i, k \in S$ and $\pi_1, \pi_2, ..., \pi_n$,

are unique solutions of $\pi_k = \sum_{j=1}^{n} \pi_j p_{jk} \wedge \sum_{j=1}^{n} \pi_j = 1.$

Now, the stationary probability distribution vector $\pi = (\pi_1, \dots \pi_n)$, is

$$\Pi = \begin{pmatrix} \pi \\ \pi \\ \vdots \\ \pi \end{pmatrix} = \begin{pmatrix} \pi_1, \pi_2, \dots \pi_n \\ \pi_1, \pi_2, \dots \pi_n \\ \cdots \\ \pi_1, \pi_2, \dots \pi_n \end{pmatrix}$$

limits are the matrix form as follows

$\lim\limits_{t\to\infty} p^{(t)} = \lim\limits_{t\to\infty} p^t = \Pi$ and $\lim\limits_{t\to\infty} p(t) = \lim\limits_{t\to\infty} p(0)P^t = \pi$,

where π is unique solution of

$$\pi = \pi P, \ \sum_j \pi_j = 1. \qquad (1)$$

If the initial probability distribution is stationary, ie. $p(0) = \pi$, then all absolute probability distributions p(t) are stationary, the chain is in the statistic equilibrium. Remarks that can be made are that a homogeneous Markov chain is characterized by the fact that the transition probabilities and therefore also the transition matrices only depend on the size of the time increment and that for homogeneous Markov chain one can simplify the Chapman – Kolmogorov equations to the semi group property

$$P(s+t) = P(s) \times P(t).$$

The forces of transition and the transition probability functions are related by Kolmogorov equations.

Forward differential equations are

$$\frac{\partial}{\partial t} p_{ij}(s,t) = \sum_{l=1}^{k} p_{il}(s,t)\mu_{lj}(s,t)$$

Backward differential equations

$$\frac{\partial}{\partial s} p_{ij}(s,t) = -\sum_{l=1}^{k} \mu_{lj}(s) p_{il}(s,t)$$

where μ_{lj} represents the force of transition from state l to state j.

Let's present now in the matrix the force of transition and transition probability. Let $Q_{k\times k}$ be the matrix with (i,j) entry μ_{ij} and $P(t)_{k\times k}$ matrix with (i,j) entry $p_{ij}(t),$ therefore, Kolmogorov equations can be writen as:

$$P'(t) = P(t)Q \quad \text{i} \quad P'(t) = QP(t).$$

However, application is limited because series may converge slowly. Cox and Miller (1965) reduced the problem of finding transition probability functions to the problem of determining the eigenvalues and eigenvectors of the force of transition matrix Q. If Q has distinct eigenvalues , $d_1,....,d_k$ then $Q = ADC,$ where C is inverse matrix of the matrix A, and D is a diagonal matrix $D = diag(d_1,....,d_k).$

Futhermore,

$$P(t) = Adiag(e^{d_{1l}},...,e^{d_{kl}})C. \qquad (2)$$

For an ergodic Markov chain with finite state-space, we have the mean first return times all finite. The next theorem shows us how to compute the mean first return times in terms of the stationary probability distribution for P.

Theorem 1. For an ergodic Markov chain, the expected first return time m_x for state x satisfies $m_x = 1/\pi_x$, where $\pi = [\pi_1, . . . , \pi_r]$ is the stationary probability vector for P.

3. APPLICATION OF MARKOV CHAIN TO STOCK MARKET ANALYSIS

The application of Markov chain to stock market analysis is the subject of numerous researches. Doubleday and Esunge (2011), in their research determine the relationship between a diverse portfolio of stocks and the market as a whole using Markov process. Svoboda and Lukáš (2012) were trying to predict the stock index trend of Prague stock exchange PX using Markov chain analysis. The overall objective of Otieno et al. (2015) study was to apply Markov Chain to model and forecast trend of one company shares trading in Nairobi Securities Exchange. In the Rajakumar and Shanthi work (2014), Markov process model is applied for forecasting the subsequent quarter EPS for companies in IT sector. They introduced two models with two states and extended state interval model with time independent transition probability matrices to predict the EPS for subsequent quarter for 4 securities.

The present research aims at trying to predict stock market asset prices on the Belgrade Stock Exchange, i.e. return of assets expounding Markov chain model. Specific features of the BSE were taken into account for modelling. Model includes five discrete state-spaces. The results of the short-term trend prediction using Markov Chain Analyses are shown for various stocks. The Markov prediction model consists of selecting input variable, data processing, classification of states, construction of state process, state probability, state transition probability matrix and forecasting the subsequent state probability of return.

Research was based on historical data on stock prices on the BSE. Stock prices at the closing of the BSE were observed in a period of one year, i.e. 252 trading days. We chose for our analysis stocks from various sectors of the economy. Stocks from all lists of the Stock Exchange were represented (Prime listing, Standard Listing, Open Market and MTP).

Based on stock prices, returns for every trading day were calculated by using a logarithmic approximation:

$$r_{i,t} = ln\left(\frac{P_{i,t}}{P_{i,t-1}}\right)$$

where $P_{i,t}$, $P_{i,t-1}$ are stock prices at the closing, respectively, on day t, t-1,… in period T.

On each trading day, return $\left(r_{i,t}\right)$ was compared to the return of the stock from previuos day $\left(r_{i,t-1}\right)$, and then all returns were classified in categories, i.e. intervals.

We shall mark states in the observed period with S_i and the state-space with $S=\left(S_1,\ldots.,S_n\right)$.

Probability of the state is a possibility of occurrence of various states of the system. Each state of the return was allocated with the initial probability vector by calculating relative frequencies of return in each of the stated states.

Then, the state vector is:

$$p\left(i\right)=(p_1,p_2,\ldots\ldots., p_n) \text{ for } i= 1,2,\ldots.n,$$

where p_j are probabilities of the state x_j, j=1,2,…,n.

For the observed period T, initial state vector will be

$$p\left(0\right)=\left(\ x_1/T,\ x_2/T\ ,\ \ldots,\ x_n/T\right)$$

Matrices of transition states are created on the basis of the previuos state of the system and probabilities of the next state are defined depending on the current state of the system.

Creation of vectors and matrices of transition states can enable further forecast of probabilities of the state of return in future, especially for each trading day (date).

Vectors of current states and transition matrices are created in MS-Excel, while calculation of transition matrices is done in the software package Matlab.

Model is characterised by 5 states of the system:

S1 – return is less than -0,5%

S2 – return is in interval [-0,5;0),

S3- return is 0 or unchanged,

S4- return is in interval (0;0,5%]

S5- return is higher than 0,5%

The initial analysis, which included 16 companies, eliminated stocks that were not traded for over 126 days in the previous period. Such stocks have the greatest number of returns in interval describing the state X3. This state is characterised with no significant changes in return, but also with no changes in the stock price compared to the previous trading day. Investors are not interested in such stocks. Further, stocks with less than 126 of trading days will not be considered as potential stocks in the portfoilo.

We will consider it through an example of trends of stock prices of the company Bambi. Figure 1 shows stock trading (marked as BMBI) in the observed period according to data from the BSE.

Figure 1. Historical data on trends of stocks of BMBI

Source: http://www.belex.rs

The graph with stock price trends of BMBI, in the observed period, shows that these stocks were not traded for several days, i.e. 214 days.
The initial vector is

$$p(0) = (0.051587 \quad 0.015873 \quad 0.892857 \quad 0.011905 \quad 0.027778)$$

that is interpreted as follows: probability that return from stocks of BMBI was less than -0.5% was 0.05159, that they were in interval [-0.5%, 0.5%] was 0.01587, probability with unchanged return or without return was 0.892857, then probability of return in interval [0, 0.5%] with the probability 0.01190 and finally probability of return higer than 0.5% was 0.02778. Initial distribution vectors were analogously calculated for other observed stocks. Calculation of a probability is analogous and interpreted in similar manner. Probabilities matrices were then calculated for each stock. Each element was calculated also as a relative frequency, where element p_{ij} is calculated as a ratio of number of transitions from i to j state and the number of returns in i state. The stated example shows that relative frequencies with the highest probability are in the state S3, and very low probability of state S4 and S5, so stocks of BMBI are not accepted into portfolio. Similarly we choose other stocks.

Stocks of SJPT fulfil the stated conditions so we will illustrate the process of calculating the transition matrix and the steady-state vector with these stocks. Figure 2 shows a graph of trends of stocks prices of SJPT in the observed period.

Figure 2. Historical data on on trends of stocks of SJPT

Source: http://www.belex.rs

Based on collected data on trends of return for each day individually in the observed period and on classification of returns in five states, the initial vector for stocks of SJPT was created

$$p(0)=[0.22619 \quad 0.076397 \quad 0.392857 \quad 0.0873 \quad 0.218254].$$

Then, transition probabilities were calculated by observing the number of transitions from the current state to some of the 5 states during 252 trading days, and transition states matrices were created

$$P_{SJPT}=\begin{pmatrix} 0.15789 & 0.087719 & 0.368421 & 0.10526 & 0.280702 \\ 0.42105 & 0.157895 & 0.411765 & 0 & 0.315789 \\ 0.21212 & 0.080808 & 0.444444 & 0.11111 & 0.151515 \\ 0.09091 & 0.045455 & 0.590909 & 0.04545 & 0.227273 \\ 0.30909 & 0.036364 & 0.345455 & 0.09091 & 0.218182 \end{pmatrix}$$

Probability distribution vectors were calculated for the next 4 days (Table 1) based on the initial vector and the matrix of PSJPT. Calculation was done by using Matlab, applying formula 1. Distribution equilibrium shows that steady state is achieved very fast, which points to stability of these stocks in the observed period. Upon observing distribution vectors, it can be expected that returns from these stocks will on the following day be in the state S3.

Table 1. Output results for distribution vectors of stocks of SJPT

π_1	π_2	π_3	π_4	Steady state	Mean return time
0.22619	0.22532	0.2254	0.22539	0.225390	4,44
0.07540	0.07543	0.07543	0.07544	0.075440	13,25
0.39286	0.39383	0.39385	0.39386	0.393860	2,54
0.09127	0.09109	0.0911	0.09110	0.091102	10,98
0.21429	0.21432	0.2142	0.21421	0.214208	4,67

Table 2. Output results for steady-state vectors

State/ stock		1	2	3	4	5
ALFA	π	0.123016	0.06746	0.531746	0.059524	0.218254
	m	8.13	1.48	1.88	16.8	4.58
FITO	π	0.159168	0.039695	0.522506	0.051818	0.226813
	m	6.28	25.19	1.91	19.30	4.41
AERO	π	0.321429	0.178571	0.039683	0.134921	0.325397
	m	3.11	5.60	25.20	7.41	3.07
MTLC	π	0.130952	0.071429	0.579365	0.083333	0.134921
	m	7.64	14.00	1.73	12.00	7.41
JESV	π	0.107143	0.055556	0.646826	0.071429	0.119048
	m	9.33	18.00	1.55	14.00	8.40
SJPT	π	0.22539	0.07544	0.39386	0.091102	0.214208
	m	4.44	13.26	2.54	10.98	4.67

Transition matrices of all selected stocks based on the stated criteria were calculated in the same manner. We observe stocks with increased return, i.e. greater probability that states S4 and S5 will occur. Based on this criterion we selected a portfolio with stocks shown in Table 2 and illustrated the steady-state vector for each stock.

Since methodology of Markov chains assumes that Markov matrix remains unchanged in time, which on capital markets is not a real assumption, it can be concluded that observing a long period by using this methodology is not recommended. Therefore, analysis is focused on the short-term forecast, i.e. next several days, and thus timing for investments or trading of stocks on exchange market is selected. In comparison with other methodologies, results are realised in simpler manner and the forecast is simpler.

4. NO CLAIM DISCOUNT MODELS

In a competitive environment it is important that insurers, as much as they can, find the real price of a risk for each individual insured by grouping them in narrow sub-groups. A priori variables are, for example: having in mind *a priori* variables (for example: age, car, or the township of residence, premium payment method, etc). *A posteriori* rating is a very efficient way of classifying policyholders into specific categories according to their risk and who pay premiums relative to their claims experience. However, the best forecast is based on past claims behaviour. Insureds mostly choose ranking according to merits, such as bonus and malus system. In practice, insurers have different bonus-malus systems. An insured without any claims is rewarded with a bonus, and an insured with claims is sanctioned with a malus.

An NCD can significantly reduce the cost of car insurance cover. This principle is meant to reward policy holders for not making claims during a year; that is, to grant a bonus to a careful driver. A bonus principle effects the policy holder's decision whether or not to claim in a particular instance. No claims discounts allow the driver to be more responsible about their vehicle and when driving. If no claims are made, each year the premium reduces. The discount is calculated as a straightforward percentage of the total cost of the insurance premium and will be discounted each year a claim is not made. An NCD system discourages small claims.

Bonus-malus systems using Markov chains

Bonus-malus system (BMS) presents a system of discounts granted to an insured and loadings paid by an insured depending on previous experience with claims. Bonus and malus are most often calculated at the beginning of insurance period by applying corresponding corrective factors to the basic premium. Application of bonus-malus system characterises those insurance lines with which risk occurrence greatly depends on behaviour and characteristics of an insured, as in motor third-party liability insurance. Bonus-malus systems can be considered as special cases of Markov processes. Actuarial literature studies bonus-malus systems since the early sixties, when it was introduced in Europe. Among the first papers was the paper written by Loimaranta (1972) who developed formulas for some asymptotic properties of bonus systems, where bonus systems are understood as Markov chains. There are a lot of works of

Lamaire where he studies the Markov chain theory for the design, evaluation, and comparison the BMSs of different countries.

Regarding an interruptible random variable, which in this case is the number of reported claims arising from the motor third-party liability insurance, we can say that it has the Poisson distribution, since the following conditions are fulfiled:

1. probability of number of occurred loss events is relatively small $\left(p \leq 5\%\right)$,
2. sample is big $\left(n \geq 20\right)$,
3. number of occurrence of loss events independent of time nuits, i.e. space,
4. probability of occurrence of loss events is proportional to the time or space uit, as well as
5. probability of simultaneous occurrence of two or more claims is insignificant

Probability that the average number of loss events per insured is x, under assumption that claims from motor third-party liability insurance have the Poisson distribution. Formula for calculation of a loading to the basic premium can be derived by using the Poisson distribution in order to ensure required funds for approval of the bonus to careful insureds. Modified bonus-malus system is used in motor third-party liability insurance, which classifies insureds into corresponding premium class. The following criteria serve as basis for classification of insureds into corresponding premium class:

- number of claims reported in previuos insurance period of one year for which the insured's liability is confirmed, and
- number of years without any claims,

and additional criteria, level of claims and technical result from previous period.

According to this bonus-malus system it is common that each premium class has a corresponding premium that presents a certain number of monetary units of the basic premium and can be calculated as a product of the basic premium and a corresponding corrective bonus-malus factor.

In Serbia, an insurance company is obliged to include and apply in its premium system, i.e. premium tariffs, the bonus-malus system. The basic criteria, data for application of the system and the maximum bonus is determined by the National Bank of Serbia. An insurance company can define additional criteria that are not contrary to these criteria.[1] The system consists of 12 premium classes with the bonus up to 15% and malus up to 250% of the basic premium from the 4^{th} premium class that is determined according to the premium tariffs.

[1] Decision on the Basic Criteria of the Bonus-Malus System, *Official Gazette of the Republic of Serbia*, No. 24/2010 and 60/2011.

Table 3. Serbian bonus-malus system for motor third-party liability

% premium	85	90	95	100	115	130	150	170	190	210	230	250
class	1	2	3	4	5	6	7	8	9	10	11	12
1	0			1			2			3		4+
2		0			1			2			3	4+
3			0			1			2			3+
4				x			1			2		3+
5					0			1			2	3+
6						0			1			2+
7							0			1		2+
8								0			1	2+
9									0			1+
10										0		1+
11											0	1+
12												1+

Source: Adapted from National Bank of Serbia. (2010). Decision on the Basic Criteria of the Bonus-Malus System. Belgrade: National Bank of Serbia, p. 7.

If an insurance lasted at least one year, and in that period no claim was reported, i.e. if insurance termination did not last longer than three years, and from the beginning until the expiry of the previous insurance policy there were no reported claims, the insured is granted for the next insurance period one lower premium class, and maximum up to the first premium class.
If claims were reported in the previous period, regardless of the insurance duration under which claims had been reported, according to each reported claim in that period the insured is moved by three premium classes higher, and maximum up to the 12th premium class.

Application of the bonus-malus system essentially has a multiple significance and an objective[2]:

1. to improve differentiation of insureds regarding the insurance premium level,
2. to stimulate and interest insureds in safer driving,
3. to stimulite insureds not to have claims so as not to lose bonus,
4. to discourage insureds from insurance frauds and

[2] Cerović, M. (2013). Bonus-malus motor liability insurance in comparative law and in Serbia. In: *Serbian insurance law in transition to EU insurance law*, Insurance Law Association of Serbia, p. 333.

5. to reduce claims through the self-retention that are at the expense of insurers.

In continuation we will demonstrate various schemas of the bonus-malus systems and average annual premium calculation using Markov chan.

Model 1. We will assume that an insurance company uses three categories of insurance:
I basic - 0,
II bonus 30%,
III bonus 50%.

Let X(t) be a random process which indicates the insurance category in the period t. The insured person is added to the appropriate category depending on the number of insurance events reported in the previous period. A new insured is in the first category, i.e. the state 0. If an insured did not have any claims in the first period, he is entitled to the bonus from category II. If an insured reports a claim in the next period, we classify him in the lower category. If an insured reports several claims, we classify him in the lowest category. Suppose that the number of claims in the insurance period t is a random variable
$Yt,\ t = 1,\ 2,\ \ldots,$ we can express previous as[3]

$$X(t+1) = \begin{cases} \min\{X(t)+1;2\}, & Y_t = 0 \\ \min\{X(t)-1;0\}, & Y_t = 1 \\ 0, & Y_t > 1. \end{cases}$$

We assume that Yt are independent random variables with the same Poisson distribution with parameter λ. Then X(t) is the homogeneous Markov chain, S = {0, 1, 2}. It has initial probability distribution p(0) = (1, 0, 0) and the transition probability matrix

$$P = \begin{pmatrix} 1-e^{-\lambda} & 1-e^{-\lambda} & 0 \\ 1-e^{-\lambda} & 0 & e^{-\lambda} \\ 1-e^{-\lambda}-\lambda e^{-\lambda} & \lambda e^{-\lambda} & e^{-\lambda} \end{pmatrix}$$

Transition matrix has in the first column non-negative elements, so there is a stationary distribution $\pi = \lim_{t\to\infty} p(t)$. Further, we deal with the equation system.

[3] Neubauer, J. (2014). Markov Models for Risk Assessment - Reliability and Risk Analysis, *ESF*, p.15.

For easier calculation we will introduce marks $a_0 = e^{-\lambda}$ and $a_1 = \lambda e^{-\lambda}$, then we have

$$\pi_0 = \pi_0(1-a_0) + \pi_1(1-a_0) + \pi_2(1-a_0-a_1),$$
$$\pi_1 = \pi_0 a_0 + \pi_2 a_1,$$
$$\pi_2 = \pi_1 a_0 + \pi_2 a_0,$$
$$\pi_0 + \pi_1 + \pi_2 = 1.$$
$$\pi_0 = \frac{1-a_0-a_0a_1}{1-\ a_0a_1} = \frac{1-e^{-\lambda}-\lambda e^{-2\lambda}}{1-\lambda e^{-2\lambda}} \quad ,$$
$$\pi_1 = \frac{a_0(1-a_0)}{1-a_0a_1} = \frac{e^{-\lambda}(1-e^{-\lambda})}{1-\lambda e^{-2\lambda}}$$
$$\pi_2 = \frac{a_0^2}{1-a_0a_1} = \frac{e^{-2\lambda}}{1-\lambda e^{-2\lambda}}.$$

Model 2. Let's consider now the following system Razmotrimo sada sledeći sistem: all its states are ergodic (it is possible to go from every state to every other state) and the chain is not cyclic. BMS has six levels of discount 0%, 20%, 25%, 35%, 45%, 50%. At the end of each policy year, policyholders change levels according to the following rules:

1. if no claim(s) during a policy year policyholder moves to the next higher discount level or remain at 50% if already at the highest level.
2. if at least one claim during a policy year, policyholder drops back to zero percent level.

We will mark with p_0 the probability of no claim and $(1-p_0)$ probability of at least one claim. Then we have,

$$P = (P_{ij})_{5\times 5},$$

$P_{i0} = 1 - p_0$, where $i = 0,1,\ldots,5$

$P_{i,i+1} = p_0$, where $i = 0,1,\ldots,4$

and

$$P_{55} = p_0.$$

General solution of Transition probability matrix is

$$\pi_0 = (1-p_0),\ \pi_1 = p_0(1-p_0),\ \pi_2 = p_0^2(1-p_0),$$
$$\pi_3 = p_0^3(1-p_0),\ \pi_4 = p_0^4(1-p_0),\ \pi_5 = p_0^5.$$

Now, average yearly premium is

$$\Pi\sum_{i=1}^{5}\pi_i d = \frac{\Pi}{100}(1-p_0)\left[100 + 80p_0 + 75p_0^2 + 65p_0^3 + 55p_0^4 + 50\frac{p_0^5}{1-p_0}\right]$$

where

Π - yearly amount of premium
d - levels different at discount of percentage

Model 3. Markov chains can be applied in classifying of drivers according to the risk, i.e. number of claims in one year. Category of good drivers with low risk are drivers without any claims in one year, medium risk drivers with 1 claim, and high risk drivers with 2 and more claims. For illustration purposes, we present the Markov system with 3 states with an initial vector

$$p(0) = \begin{bmatrix} 0.53125 & 0.260417 & 0.208333 \end{bmatrix}$$

Criterion for classification into a specific state is the number of claims per insured. In the observed period, majority of drivers are in the category of low-risk drivers who had no claims, with the probability of 0.5312, then there is the category of medium-risk drivers with the probability of 0.2604 that includes drivers with one claim. Category of high-risk drivers includes insureds with 2 and more claims and their probability is 0.2083. Further, transitions of drivers from one state to the other state were observed according to number of claims and the Transition Probability Matrix is created.

Transition Probability Matrix is given by

$$P = \begin{pmatrix} 0.60784 & 0.313725 & 0.078431 \\ 0.54902 & 0.352941 & 0.098039 \\ 0.56863 & 0.392157 & 0.039216 \end{pmatrix}$$

In this manner we can calculate the probability that a randomly selected driver will be next year, let's say, in the low-risk category. The Table 2 shows results of our system. Later states will change very little, if at all. Analysing the results after five states, we see that the percentage of low-risk drivers has increased to 58.5%. This means that in the long run, the number of low-risk drivers will be 0.5850, 0.3332 in mid and high-risk will be 0.081759.

Table 4. Output Results

State	Steady state	Mean return time
1	0.585037	1.71
2	0.333205	3
3	0.081759	12.23

Now, we will explain with a simple example the method for calculation of insurance premium for drivers in relation to whether they experience an accident or not in previous period.

Model 4. We know data on claims from previous period based on which we concluded that if a client had an accident in period t-1, the probability for an accident in period t would be 0,1. If in the period t-1 there were no accidents, the probability for an accident to happen in period t is 0,3. An insurance company charges customers annual premiums based on their accident history as follows:
- if no accident in last 2 years, annual premium is 200
- if accidents in each of last 2 years, annual premium is 700 and
- if accident in only 1 of last 2 years, annual premium is 360.

$$P = \begin{pmatrix} 0.97 & 0.03 & 0 & 0 \\ 0 & 0 & 0.9 & 0.1 \\ 0.97 & 0.03 & 0 & 0 \\ 0 & 0 & 0.9 & 0.1 \end{pmatrix}$$

and the steady state is

$$\pi = \begin{pmatrix} 0.93871 & 0.02903 & 0.02903 & 0.00323 \end{pmatrix}$$

Then, the long-run average annual premium is

$$\sum_{i=1}^{4} \pi_i \Pi_i = 0.93871 \cdot 200 + 0.029032 \cdot 360 + 0.029032 \cdot 360$$

$$+0.003226 \cdot 700 = 210.903.$$

IBNR reserving

Actuarial literature has proposed and developed numerous claims reserving models to estimate outstanding claims. The best known claims reserving models in actuarial literature are the Chain Ladder, Loss Cost, Cape Code and its variations. Chain Ladder model is probably the most widely used and its basic classification according to the exploitation data regarding the changes of the amount of claims in the past and projections about the frequency and intensity of claims. The method allows different data grouping, by the occurrence period, by the reporting period of claims and by the insurance contract year. Also, the data that we observe can be settled or settled claims and reported claims reserve. The projections may include incremental or cumulative amounts of claims.

Let $X_{i,j}$ represent the incremental claim amounts for accident year i and development year j; this is also known as the incremental payments of the change of reported claims. Let R represent the total outstanding claims liability. The claims that are incurred or the sum of all reported claims is termed as the cumulative claims amounts and denoted by $C_{i,j}$. Mathematically, the cumulative claim amounts for accident year i up to the development year j are given by

$$C_{i,j} = X_{i,1} + X_{i,2} + X_{i,3} + \cdots + X_{i,j} = \sum_{k=1}^{j} X_{i,k}$$

Assumptions from Wutrich, Buhlmann and Furrer (2007)

– $\mathcal{B}_k = \sigma\left\{ X_{i,j} : i + j \le I, j \le k \right\} = \sigma\left\{ C_{i,j} : i + j \le I, j \le k \right\}$

  $\mathcal{B}_k$ is σ-field, in upper triangle in time I (year for last received premium).

– The filtration T_t is generated with $\left\{ C_{i,j} : i + j \le t \right\}$

– payments $X_{i,j}$ in different accident years are independent

– $\left(C_{i,j}\right)_{j\ge 0}$ is a Markov chain with

  $f_0 < 0$ *and* $f_j > 0$ $(j \ge 1)$ *and* $\sigma_j^2 (j \ge 1)$ such that $\forall\ i \le I\ i\ j \ge 1$

$$E\left[C_{i,j} \middle| C_{i,j-1} \right] = f_{j-1} \cdot C_{i,j-1},$$

$$Var\left(C_{i,j} \middle| C_{i,j-1} \right) = \sigma_{j-1}^2 \cdot C_{i,j-1}.$$

Factor f_j is chain-ladder factor, development factor or age-to-age factor. Then we can define individual development factor:

$F_{i,j} = \dfrac{C_{i,j+1}}{C_{i,j}}$, with conditional $Var\left(F_{i,j} \middle| B_j \right) = \sigma_j^2 / C_{i,j}$

Under assumptions we have for all $i \ge 0, j < J, k \ge 1$

$$E\left[C_{i,J} \middle| C_{i,j} \right] = C_{i,j} \cdot f_j \cdots f_{J-1},$$

$$E\left[X_{i,j+k} \middle| C_{i,j} \right] = C_{i,j} \cdot f_j \cdots f_{j+k-2} \cdot \left(f_{j+k-1} - 1 \right)$$

and $T_t = \mathcal{B}_J$ for $t=I$

$$E_{t+k}^{(t)} = E\left[X_{t+k} \middle| \mathcal{T}_t \right] = \sum_{i+j=t+k} E\left[X_{i,j} \middle| \mathcal{B}_J \right]$$

$$= \sum_{i+j=t} C_{i,j} \cdot f_j \cdots f_{j+k-2} \cdot \left(f_{j+k-1} - 1 \right)$$

To estimate development factors we introduce folowing marks ($J=I=t$), $i^*(j) = I - j$ and $j^*(i) = J - i$, where $X_{i^*(j),j}$ and $X_{i,j^*(i)}$ are payments for last accounting year.

Then f_j, σ_j^2 are:

$$\hat{f}_j = \frac{\sum_{i=0}^{i^*(j+1)} C_{i,j+1}}{\sum_{i=0}^{i^*(j+1)} C_{i,j}} = \sum_{i=0}^{i^*(j+1)} \frac{C_{i,j}}{\sum_{i=0}^{i^*(j+1)} C_{i,j}} \cdot F_{i,j}$$

$$\hat{\sigma}_j^2 = \frac{1}{i^*(j+1)} \sum_{i=0}^{i^*(j+1)} C_{i,j} \cdot \left(\frac{C_{i,j+1}}{C_{i,j}} - \hat{f}_j \right)^2$$

If development factor is equal to 1 in the last year and the end of development is in year I-1, then $\hat{\sigma}_{I-1}^2 = 0$. Otherwise, extrapolating exponential descending serie $\hat{\sigma}_1, \hat{\sigma}_2, \ldots, \hat{\sigma}_{I-3}, \hat{\sigma}_{I-2}$, Mack (1993) proposed the following formula

$$\hat{\sigma}_{I-1}^2 = \min\left(\hat{\sigma}_{I-2}^4 / \hat{\sigma}_{I-3}^2, \min\left(\hat{\sigma}_{I-3}^2, \hat{\sigma}_{I-2}^2 \right) \right)$$

Lema 1. $\hat{f}_j$ is the $\mathcal{B}_{j+1}$ measurable unbiased estimator which has minimal conditional variance among all linear combination of unbiased estimators of $F_{i,j} = C_{i,j+1} / C_{i,j}$.

To solve numerical examples we also need next equality

$$E\left[\hat{f}_k^{\,2} | \mathcal{B}_k \right] = Var\left(\hat{f}_k | \mathcal{B}_k \right) + f_k^{\,2} = \frac{\sigma_k^2}{\sum_{i=0}^{i^*(k+1)} C_{i,k}} + f_k^{\,2}.$$

Further, we create known triangles.

5. MARKOV CHAINS IN LIFE INSURANCE

In life assurance, every transition of an insured from one state to the other can have a certain financial effect that needs to be quantified. Regarding term life

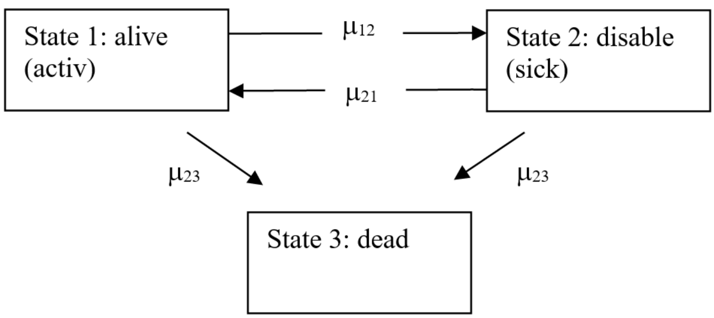
assurance we have a simple system with two states, alive or dead. Then we have $d_1 = -\mu_{12}$, $d_2 = 0$ and corresponding eigenvectors are $(1,0)'$ and $(1,1)'$.
By using a formula (2) we get

$$P(t) = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} e^{-\mu_{12}t} & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} e^{-\mu_{12}t} & 1-e^{-\mu_{12}t} \\ 0 & 1 \end{pmatrix}.$$

and still have probabilities
$p_{11}(t) = e^{-\mu_{12}t}$, $p_{12}(t) = 1-e^{-\mu_{12}t}$, $p_{22}(t) = 1$, $p_{21}(t) = 0$.

Regarding whole insurance, premium is paid while an insured is in the state 1, a death benefit is paid at the time of transition to state 3. Insured can make a transition from the state 1 to the state 2, but also vice versa. Also, it is possible to make a transition from the state of the disabled to dead. To illustrate the procedure described above, we use the three-state model shown in Figure 3 to describe select mortality.

Figure 3. Markov chain model for disabilities, recoveries, and death.



First, for attending age x, we calculated forces of mortality from the mortality rates. In the case of this model, we have

$$Q = \begin{pmatrix} -(\mu_{12}+\mu_{13}) & \mu_{12} & \mu_{13} \\ \mu_{21} & -(\mu_{21}+\mu_{23}) & \mu_{23} \\ 0 & 0 & 0 \end{pmatrix}$$

We can apply the Kolmogorov forward differential equations, for person who is active at time s:[4]

$$\frac{\partial}{\partial t} p_{00}(s,t) = p_{01}(s,t)\mu_{21}(t) - p_{00}(s,t)(\mu_{13}(t) + \mu_{12}(t)),$$

$$\frac{\partial}{\partial t} p_{01}(s,t) = p_{00}(s,t)\mu_{12}(t) - p_{01}(s,t)(\mu_{23}(t) + \mu_{21}(t)),$$

Initial conditions become

$$p_{00}\left(s,s\right) = 1 \text{ and } p_{01}\left(s,s\right) = 0.$$

For person who is disabled at time s, equations are:

$$\frac{\partial}{\partial t} p_{10}(s,t) = p_{11}(s,t)\mu_{21}(t) - p_{10}(s,t)(\mu_{13}(t) + \mu_{12}(t)),$$

$$\frac{\partial}{\partial t} p_{11}(s,t) = p_{10}(s,t)\mu_{12}(t) - p_{11}(s,t)(\mu_{23}(t) + \mu_{21}(t)),$$

Initial conditions become

$$p_{00}\left(s,s\right) = 0 \text{ and } p_{11}\left(s,s\right) = 1.$$

Many other problems can be described by the same schematic. For instance, in connection with a pension insurance with additional benefits to the spouse, states 0 and 1 would be single and married, and in connection with unemployment insurance they would be employed and unemployed.
Further, 4-stage model can be introduced for the health insurance. We will have 3 health states: healthy, sick and infected state. The fourth state, death, can occur as a result of the infection or death could also occur from causes not related with the disease infection. By variation of state space and intensities, the Markov model ie. its set-up is able to represent an extremely complex phenomena. For future research authors indicated the possibility of applying Markov chain model on actuarial equivalence principle, expected present values, reserves etc.

[4] Norberg, R. (2002). *Basic Life Insurance Mathematics*, London: London School of Economics, p. 76.