The IX International Academic Congress "Contemporary Science and Education in Americas, Africa and Eurasia"

(Brazil, Rio de Janeiro, 18-20 July 2015)

"UFRJ Press"
2015

ELSEVIER

Federal University
of Rio de Janeiro

“Contemporary Science and Education in Americas, Africa and Eurasia”

The IX International Academic Congress

(Brazil, Rio de Janeiro, 18-20 August 2015)

PAPERS AND COMMENTARIES

VOLUME III

“UFRJ Press”
2015

Gulyan Anahit, *Yerevan State University, the Faculty of Mathematics, Chair of Actiarial Mathematics and Risk management, Lecturer,*
Chitchyan Robert, *Yerevan State University, the Faculty of Mathematics, Chair of Actiarial Mathematics and Risk management, Docent, PhD of Physics and Mathematical Sciences*

Evaluation methods of the coefficients of transition Matrices in Extended Bonus-malus System

Abstract: In this article we presented the coefficients' estimations of the transition matrices of the extended bonus-malus system (BMS) by the EM algorithm and recursive methods. The extended BMS is constructed via hidden Markov models. The first method is based on the prior assumption of the coefficients' distribution and the second one can be applied only when initial distribution is known.

Keywords: Insurance, Markov chain, coefficient evaluation, likelihood function, filtration, Change of measure, Hidden Markov model, Bonus-Malus system.

1. **Introduction.** Let consider an extended model of BMS where the distribution of policyholders among the BMS classes at time n is described by Markov chain X_n which have a standard basis $\{e_1, \dots, e_{|S_X|}\}$[1].

The characteristic feature of this model is the specific rule of the transition between classes. In this model the movement between classes is modeled as a discrete time Markov chain, which is dependent on the processes of the claims number N_n and aggregate claims severity H_n which are also assumed to be Markovian and are based accordingly on the standard basis $\{f_1, \dots, f_{|S_H|}\}$ and $\{h_1, \dots, h_{|S_N|}\}$. The detailed description of the extended BMS is given in [1].

So, the model described with the following system of equations:

[1] Here $|S_.|$ is the cardinality or size of the state space and the vector e_i has 1 in the$i$-th position and zero's elsewhere.

$$\begin{cases} X_n = BX_{n-1} \otimes H_{n-1} \otimes N_{n-1} + V_n \\ H_n = Q_n H_{n-1} \otimes N_{n-1} + W_n \\ N_n = A_n N_{n-1} \otimes X_{n-1} + L_n \\ \breve{Q}_n = D\breve{Q}_{n-1} + R_n \\ \breve{A}_n = K\breve{A}_{n-1} + T_n \end{cases} \quad (1)$$

where $B = \left(p_{j,ikm}\right)_{j,i=1}^{|S_X|}\ _{k=1}^{|S_H|}\ _{m=1}^{|S_N|}$, $Q_n = (q_n(s,r,l))_{s=1}^{|S_H|}\ _{r=1}^{|S_H|}\ _{l=1}^{|S_N|}$,

$A_n = (a_n(u,v,w))_{u=1}^{|S_N|}\ _{v=1}^{|S_N|}\ _{w=1}^{|S_X|}$, $D = \left(d_{qp}\right)_{q,p=1}^{Q|S_H||S_N|}$, $K = \left(k_{qp}\right)_{q,p=1}^{A|S_X||S_N|}$ are the transition matrices with the following probability-members:

$$p_{j,ikm} \triangleq P\left(X_n = e_j \middle| X_{n-1} = e_i, H_{n-1} = f_k, N_{n-1} = h_m\right)$$
$$q_n(s,r,l) \triangleq P(H_n = f_s | H_{n-1} = f_r, N_{n-1} = h_l)$$
$$a_n(u,v,w) \triangleq P(N_n = h_u | N_{n-1} = h_v, X_{n-1} = e_w)$$
$$d_{щх} \triangleq P\left(\breve{Q}_n = b_{щ} \middle| \breve{Q}_{n-1} = b_{х}\right)$$
$$k_{\varpi\vartheta} \triangleq P\left(\breve{A}_n = m_{\varpi} \middle| \breve{A}_{n-1} = m_{\vartheta}\right)$$

which satisfy to the following normalization conditions:

$$\sum_{j=1}^{|S_X|} p_{j,ikm} = 1, \quad \sum_{s=1}^{|S_H|} q_n(s,r,l) = 1,$$
$$\sum_{u=1}^{|S_N|} a_n(u,v,w) = 1, \ \sum_{щ=1}^{Q|S_H||S_N|} d_{щх} = 1, \quad \sum_{\varpi=1}^{A|S_X||S_N|} k_{\varpi\vartheta} = 1 \quad (2)$$

All processes in this system are assumed as Markov chains. The concept of hidden Markov models (HMM) underlies the model construction.

The system is described with the following set of parameters:

$$и := \begin{Bmatrix} p_{j,ikm}, & i,j = \overline{1,|S_X|}, & k = \overline{1,|S_H|}, & m = \overline{1,|S_N|}, \\ q_n(s,r,l), & s,r = \overline{1,|S_H|}, & l = \overline{1,|S_N|}, & \\ a_n(u,v,w), & u,v = \overline{1,|S_N|}, & w = \overline{1,|S_X|}, & \\ d_{щх}, & щ,х = \overline{1,Q|S_H||S_N|}, & & \\ k_{\varpi\vartheta}, & \varpi,\vartheta = \overline{1,A|S_X||S_N|} & & \end{Bmatrix}$$

Our purpose is the estimation of the model parameters and it is presented here in two methods.

2. **Estimation with EM algorithm.** One of the best methods of HMM's coefficient estimation is the EM algorithm which is described detailed in [2].

We have to define a new set of parameters $\hat{и}$

$$\hat{и} := \begin{Bmatrix} \hat{p}_{j,ikm}(k), & i,j = \overline{1,|S_X|}, & k = \overline{1,|S_H|}, & m = \overline{1,|S_N|}, \\ \hat{q}_k(s,r,l), & s,r = \overline{1,|S_H|}, & l = \overline{1,|S_N|}, & \\ \hat{a}_k(u,v,w), & u,v = \overline{1,|S_N|}, & w = \overline{1,|S_X|}, & \\ \hat{d}_{щх}(k), & щ,х = \overline{1,Q|S_H||S_N|}, & & \\ \hat{k}_{\varpi\vartheta}(k), & \varpi,\vartheta = \overline{1,A|S_X||S_N|} & & \end{Bmatrix},$$

which maximizes the conditional pseudo log-likelihood functions and satisfies to conditions (2).

Define ${}^{X}\mathcal{T}_n^{j,ikm} \triangleq \sum_{t=1}^{n}\langle X_t,e_j\rangle\langle X_{t-1},e_i\rangle\langle H_{t-1},f_k\rangle\langle N_{t-1},h_m\rangle$

$$
\begin{aligned}
&{}^{H}\mathcal{T}_n^{srl} \triangleq \sum_{t=1}^{n}\langle H_t,f_s\rangle\langle H_{t-1},f_r\rangle\langle N_{t-1},h_l\rangle\\
&{}^{N}\mathcal{T}_n^{uvw} \triangleq \sum_{t=1}^{n}\langle N_t,h_u\rangle\langle N_{t-1},h_v\rangle\langle X_{t-1},e_w\rangle\\
&{}^{\breve{Q}}\mathcal{T}_n^{\text{щх}} \triangleq \sum_{t=1}^{n}\langle \breve{Q}_t,b_{\text{щ}}\rangle\langle \breve{Q}_{t-1},b_{\text{х}}\rangle\\
&{}^{\breve{A}}\mathcal{T}_n^{\varpi\vartheta} \triangleq \sum_{t=1}^{n}\langle \breve{A}_t,m_{\varpi}\rangle\langle \breve{A}_{t-1},m_{\vartheta}\rangle
\end{aligned}
\qquad (3)
$$

Each figure in (3) for the corresponding process mentioned on the left top angle represents the number of jumps of the process from one state to another up to time n.

Each figure in the next set of notations shows the number of occasions up to time n for which the corresponding Markov chain was in the mentioned state:

$$
\begin{aligned}
&{}^{X}\mathcal{O}_n^{ikm} \triangleq \sum_{t=1}^{n}\langle X_{t-1},e_i\rangle\langle H_{t-1},f_k\rangle\langle N_{t-1},h_m\rangle\\
&{}^{H}\mathcal{O}_n^{rl} \triangleq \sum_{t=1}^{n}\langle H_{t-1},f_r\rangle\langle N_{t-1},h_l\rangle\\
&{}^{N}\mathcal{O}_n^{vw} \triangleq \sum_{t=1}^{n}\langle N_{t-1},h_v\rangle\langle X_{t-1},e_w\rangle\\
&{}^{\breve{Q}}\mathcal{O}_n^{\text{х}} \triangleq \sum_{t=1}^{n}\langle \breve{Q}_{t-1},b_{\text{х}}\rangle\\
&{}^{\breve{A}}\mathcal{O}_n^{\vartheta} \triangleq \sum_{t=1}^{n}\langle \breve{A}_{t-1},m_{\vartheta}\rangle
\end{aligned}
\qquad (4)
$$

It is obvious that the figures in (3) and (4) are random variables.

Remark: For each process holds the relationship $\sum_{i=1}^{|S_*|} {}^{*}\mathcal{T}_n^{i*} = {}^{*}\mathcal{O}_n^{*}$.

To replace parameters и by $\hat{\text{й}}$ in (1) we define the following likelihood function:

$$
\begin{aligned}
\Gamma_n = {}&\prod_{t=1}^{n}\prod_{i,j=1}^{|S_X|}\prod_{k=1}^{|S_H|}\prod_{m=1}^{|S_N|}\left(\frac{\hat{p}_{j,ikm}(n)}{p_{j,ikm}}\right)^{\langle X_t,e_j\rangle\langle X_{t-1},e_i\rangle\langle H_{t-1},f_k\rangle\langle N_{t-1},h_m\rangle}\times\\
&\prod_{r,s=1}^{|S_H|}\prod_{l=1}^{|S_N|}\left(\frac{\hat{q}_n(s,r,l)}{q_n(s,r,l)}\right)^{\langle H_t,f_s\rangle\langle H_{t-1},f_r\rangle\langle N_{t-1},h_l\rangle}\times\\
&\prod_{u,v=1}^{|S_N|}\prod_{w=1}^{|S_X|}\left(\frac{\hat{a}_n(u,v,w)}{a_n(u,v,w)}\right)^{\langle N_t,h_u\rangle\langle N_{t-1},h_v\rangle\langle X_{t-1},e_w\rangle}\times\prod_{\text{щ},\text{х}=1}^{Q|S_H||S_N|}\left(\frac{\hat{d}_{\text{щх}}(n)}{d_{\text{щх}}}\right)^{\langle \breve{Q}_t,b_{\text{щ}}\rangle\langle \breve{Q}_{t-1},b_{\text{х}}\rangle}\times\\
&\prod_{\varpi,\vartheta=1}^{A|S_X||S_N|}\left(\frac{\hat{k}_{\varpi\vartheta}(n)}{k_{\varpi\vartheta}}\right)^{\langle \breve{A}_t,m_{\varpi}\rangle\langle \breve{A}_{t-1},m_{\vartheta}\rangle}
\end{aligned}
\qquad (5)
$$

where in the case of $\text{и} = 0$ we will take $\hat{\text{й}} = 0$ and $\frac{\hat{\text{й}}}{\text{и}} = 1$:

It is not difficult to show that Γ_n is a martingale-measure, so according to the Radon-Nycodim theorem there exists a measure $P_{\hat{\text{й}}}$ so that $\left.\frac{dP_{\hat{\text{й}}}}{dP_{\text{и}}}\right|_{\mathcal{G}_n} = \Gamma_n$ holds.

Lemma. Under the measure $dP_{\hat{\text{й}}}$the analogue of the system (1) holds for parameter set $\hat{\text{й}}$.

Proof. We will give the proof only for the process X and for the others' the statement of Lemma can be proved similarly.

Let $\mathcal{G}_n$ be the complete filtration generated by $\{X_k, H_{k-1}, N_{k-1}, \tilde{Q}_k, \tilde{A}_k, k \le n\}$. We have to show that under the measure $dP_{\hat{\text{й}}}$ the relationship $E_{\hat{\text{й}}}\left(\langle X_{n+1},e_j\rangle\middle|\mathcal{G}_n\right) = \hat{p}_{j,ikm}(n)$ holds. Note that in the case of information up to time n, (5) gets the form:

$$\Gamma_{n+1} = \left(\frac{\hat{p}_{j,ikm}(n)}{p_{j,ikm}}\right)^{\langle X_{n+1},e_j\rangle} \left(\frac{\hat{q}_n(s,r,l)}{q_n(s,r,l)}\right)^{\langle H_{n+1},f_s\rangle} \left(\frac{\hat{a}_n(u,v,w)}{a_n(u,v,w)}\right)^{\langle N_{n+1},h_u\rangle} \left(\frac{\hat{d}_{щх}(n)}{d_{щх}}\right)^{\langle \breve{Q}_{n+1},b_{щ}\rangle} \left(\frac{\hat{k}_{\varpi\vartheta}(n)}{k_{\varpi\vartheta}}\right)^{\langle \breve{A}_{n+1},m_{\varpi}\rangle}$$

According to the Bayes theorem

$$E_{\hat{й}}\left(\langle X_{n+1},e_j\rangle \middle| \mathcal{G}_n\right) = \frac{E\left(\langle X_{n+1},e_j\rangle \Gamma_{n+1} \middle| \mathcal{G}_n\right)}{E(\Gamma_{n+1}|\mathcal{G}_n)}.$$

Here the numerator is:

$$E\left(\langle X_{n+1},e_j\rangle \Gamma_{n+1} \middle| \mathcal{G}_n\right) = \sum_{s=1}^{|S_H|}\sum_{u=1}^{|S_N|}\sum_{щ=1}^{Q|S_H||S_N|}\sum_{\varpi=1}^{A|S_X||S_N|}\langle X_{n+1},e_j\rangle \left(\frac{\hat{p}_{j,ikm}(n)}{p_{j,ikm}}\right)\left(\frac{\hat{q}_n(s,r,l)}{q_n(s,r,l)}\right)\left(\frac{\hat{a}_n(u,v,w)}{a_n(u,v,w)}\right)\left(\frac{\hat{d}_{щх}(n)}{d_{щх}}\right)\left(\frac{\hat{k}_{\varpi\vartheta}(n)}{k_{\varpi\vartheta}}\right)\cdot$$
$$P\left(H_{n+1}=f_s, N_{n+1}=h_u, \breve{Q}_{n+1}=b_{щ}, \breve{A}_{n+1}=m_{\varpi} \middle| \mathcal{G}_n\right) = \sum_{s=1}^{|S_H|}\sum_{u=1}^{|S_N|}\sum_{щ=1}^{Q|S_H||S_N|}\sum_{\varpi=1}^{A|S_X||S_N|}\left(\frac{\hat{p}_{j,ikm}(n)}{p_{j,ikm}}\right)\left(\frac{\hat{q}_n(s,r,l)}{q_n(s,r,l)}\right)\left(\frac{\hat{a}_n(u,v,w)}{a_n(u,v,w)}\right)\left(\frac{\hat{d}_{щх}(n)}{d_{щх}}\right)\left(\frac{\hat{k}_{\varpi\vartheta}(n)}{k_{\varpi\vartheta}}\right)\cdot$$
$$q_n(s,r,l)a_n(u,v,w)d_{щх}\,k_{\varpi\vartheta}E\left(\langle X_{n+1},e_j\rangle \middle| \mathcal{G}_n\right) = \hat{p}_{j,ikm}(n):$$

And the denominator is:

$$E(\Gamma_{n+1}|\mathcal{G}_n) = \sum_{j=1}^{|S_X|}\sum_{s=1}^{|S_H|}\sum_{u=1}^{|S_N|}\sum_{щ=1}^{Q|S_H||S_N|}\sum_{\varpi=1}^{A|S_X||S_N|}\left(\frac{\hat{p}_{j,ikm}(n)}{p_{j,ikm}}\right)\left(\frac{\hat{q}_n(s,r,l)}{q_n(s,r,l)}\right)\left(\frac{\hat{a}_n(u,v,w)}{a_n(u,v,w)}\right)\left(\frac{\hat{d}_{щх}(n)}{d_{щх}}\right)\left(\frac{\hat{k}_{\varpi\vartheta}(n)}{k_{\varpi\vartheta}}\right)\cdot\cdot$$
$$P\left(H_{n+1}=f_s, N_{n+1}=h_u, \breve{Q}_{n+1}=b_{щ}, \breve{A}_{n+1}=m_{\varpi} \middle| \mathcal{G}_n\right) = 1.$$

the distributions of model processes and (2) conditions for the parameter set were used to get the results.

<u>*Theorem 1.*</u>The new estimates of parameter set $\hat{й}$ given the observations up to time n are given by

$$\hat{p}_{j,ikm} = \frac{{}^{X}\hat{\mathcal{T}}_n^{j,ikm}}{{}^{X}\hat{\mathcal{O}}_n^{ikm}},\ \hat{q}_n(s,r,l) = \frac{{}^{H}\hat{\mathcal{T}}_n^{srl}}{{}^{H}\hat{\mathcal{O}}_n^{rl}},\ \hat{a}_n(u,v,w) = \frac{{}^{N}\hat{\mathcal{T}}_n^{uvw}}{{}^{N}\hat{\mathcal{O}}_n^{vw}},$$
$$\hat{d}_{щх}(n) = \frac{{}^{\breve{Q}}\hat{\mathcal{T}}_n^{щх}}{{}^{\breve{Q}}\hat{\mathcal{O}}_n^{х}}, \hat{k}_{\varpi\vartheta}(n) = \frac{{}^{\breve{A}}\hat{\mathcal{T}}_n^{\varpi\vartheta}}{{}^{\breve{A}}\hat{\mathcal{O}}_n^{\vartheta}}$$

where ${}^{*}\hat{\mathcal{T}}_n^{*} = E\left({}^{*}\mathcal{T}_n^{*}\middle|.\right)$and ${}^{*}\hat{\mathcal{O}}_n^{*} = E\left({}^{*}\mathcal{O}_n^{*}\middle|.\right)$.

<u>*Proof.*</u> We give the proof for the first formula. Using (3) and (5) we get:

$$\ln\Gamma_n = \sum_{i,j=1}^{|S_X|}\sum_{k=1}^{|S_H|}\sum_{m=1}^{|S_N|}\sum_{t=1}^{n}\langle X_t,e_j\rangle\langle X_{t-1},e_i\rangle\langle H_{t-1},f_k\rangle\langle N_{t-1},h_m\rangle\left(\ln\hat{p}_{j,ikm} - \ln p_{j,ikm}\right)$$
$$+ R = \sum_{i,j=1}^{|S_X|}\sum_{k=1}^{|S_H|}\sum_{m=1}^{|S_N|}{}^{X}\hat{\mathcal{T}}_n^{j,ikm}\ln\hat{p}_{j,ikm} + R$$

where R is independent of $\hat{p}_{j,ikm}$. So,

$$E(\ln \Gamma_n \,|\mathcal{G}_n) = \sum_{i,j=1}^{|S_X|} \sum_{k=1}^{|S_H|} \sum_{m=1}^{|S_N|} \hat{\mathcal{T}}_n^{j,ikm} \ln \hat{p}_{j,ikm} + \hat{R} \quad (6)$$

where $\hat{p}_{j,ikm}$ must satisfy the analogue of (2) as well, that is

$$\sum_{j=1}^{|S_X|} \hat{p}_{j,ikm} = 1 \quad (7)$$

Note that the conditional expectations also satisfy the relationship $\sum_{j=1}^{|S_X|} {}^X\hat{\mathcal{T}}_n^{j,ikm} = {}^X\hat{\mathcal{O}}_n^{ikm}$.

Our purpose is to maximize (6) taking into account the constraints (2). Write л for the Lagrange multiplier and put

$$L(\hat{p}, л) = \sum_{i,j=1}^{|S_X|} \sum_{k=1}^{|S_H|} \sum_{m=1}^{|S_N|} {}^X\hat{\mathcal{T}}_n^{j,ikm} \ln \hat{p}_{j,ikm} + \hat{R} + л\left(\sum_{j=1}^{|S_X|} \hat{p}_{j,ikm} - 1\right).$$

Differentiating in л and $\hat{p}_{j,ikm}$ and equating the derivatives to 0 we get:

$$\frac{\partial L(\hat{p},л)}{\partial \hat{p}_{j,ikm}} = \frac{{}^X\hat{\mathcal{T}}_n^{j,ikm}}{\hat{p}_{j,ikm}} + л = 0, \quad \frac{\partial L(\hat{p},л)}{\partial л} = \sum_{j=1}^{|S_X|} \hat{p}_{j,ikm} - 1 = 0$$

Note that $л = -\,{}^X\hat{\mathcal{O}}_n^{ikm}$.So we get the statement of the Theorem1 for $\hat{p}_{j,ikm}$.

3. **Parameter Estimation with Recursion.** For getting the parameter estimation in the previous part we have to do some prior assumptions on the probability distribution of parameter set и, but if we have the initial distributions the recursive estimation of the parameters can be done. The new estimation in this case is presented as the previous estimation corrected with the new information.

We assume that и takes values in some set $\Theta \in \mathbb{R}^l$. Suppose we have the measure $\tilde{P}$ under which the processes of the system (1) are i.i.d.. Recall from [1] that

$$\Lambda_n = \prod_{t=1}^{n} л_t \quad (8)$$

where

$$л_t = \prod_{i,j=1}^{|S_X|} \prod_{k=1}^{|S_H|} \prod_{m=1}^{|S_N|} \left(\frac{p_{j,ikm}}{\tilde{p}_{j,ikm}}\right)^{\langle X_t,e_j\rangle\langle X_{t-1},e_i\rangle\langle H_{t-1},f_k\rangle\langle N_{t-1},h_m\rangle} \times$$
$$\prod_{r,s=1}^{|S_H|} \prod_{l=1}^{|S_N|} \left(\frac{q_n(s,r,l)}{\tilde{q}_n(s,r,l)}\right)^{\langle H_t,f_s\rangle\langle H_{t-1},f_r\rangle\langle N_{t-1},h_l\rangle} \times$$
$$\prod_{u,v=1}^{|S_N|} \prod_{w=1}^{|S_X|} \left(\frac{a_n(u,v,w)}{\tilde{a}_n(u,v,w)}\right)^{\langle N_t,h_u\rangle\langle N_{t-1},h_v\rangle\langle X_{t-1},e_w\rangle} \quad (9)$$

So, with the help of $\left.\frac{dP}{d\tilde{P}}\right|_{\mathcal{G}_n} \triangleq \Lambda_n$ we will define the “real world” measure P under which the system (1) holds. (see [1], Theorem1)

Define the unnormalized joint conditional density:

$$б_n(j, и) = \tilde{E}\left(\Lambda_n \langle X_n, e_j\rangle I(и \in dи)\middle| \mathcal{I}_n\right)$$

where $I(A)$ is the indicator function of A.

The normalized joint conditional density is:

$$f_n(j, и) = \frac{б_n(j, и)}{\sum_{j=1}^{|S_X|} \int_\Theta б_n(j, u) du}$$

We will assume given the initial density $f_0(\cdot)$.

Theorem 2. The unnormalized joint conditional density $б_n(j, и)$ satisfies the recursion:

$$б_n(j, и) = \sum_{i=1}^{|S_X|} \sum_{k=1}^{|S_H|} \sum_{m=1}^{|S_N|} \left(\frac{p_{j,ikm}}{\tilde{p}_{j,ikm}}\right) б_{n-1}(i, и)$$

Proof. Suppose $g(и)$ is any real valued Borel function on Θ. Then

$$\tilde{E}\left(\Lambda_n \langle X_n, e_j\rangle g(и) \middle| \mathcal{I}_n\right) = \int_\Theta б_n(j, u) g(u) du.$$

On the other hand using the distribution of X_n under the measure $\tilde{P}$ we get:

$$\tilde{E}\left(\Lambda_n \langle X_n, e_j\rangle g(и) \middle| \mathcal{I}_n\right) = \sum_{j=1}^{|S_X|} \Lambda_n g(и) \tilde{E}\left(\langle X_n, e_j\rangle \middle| \mathcal{I}_n\right)$$

Substitute the dynamics of X_n from (1):

$$б_n(j, и) = \sum_{i,j=1}^{|S_X|} \sum_{k=1}^{|S_H|} \sum_{m=1}^{|S_N|} \Lambda_n g(и) \tilde{E}\left(p_{j,ikm} \langle X_{n-1}, e_i\rangle \langle H_{n-1}, f_k\rangle \langle N_{n-1}, h_m\rangle \middle| \mathcal{I}_n\right)$$

Using the second relationship of (1), the distribution of H_n as well as the formulas (8) and (9) we get:

$$\begin{aligned} б_n(j, и) &= \sum_{i=1}^{|S_X|} \sum_{k=1}^{|S_H|} \sum_{m=1}^{|S_N|} \tilde{E}\left(\langle X_{n-1}, e_i\rangle \frac{\langle H_n, f_s\rangle}{q_n(s,k,m)} \Lambda_{n-1} g(и) \left(\frac{p_{j,ikm}}{\tilde{p}_{j,ikm}}\right) \prod_{s=1}^{|S_H|} \left(\frac{q_n(s,k,m)}{\tilde{q}_n(s,k,m)}\right)^{\langle H_n, f_s\rangle} \prod_{u=1}^{|S_N|} \left(\frac{a_n(u,m,i)}{\tilde{a}_n(u,m,i)}\right)^{\langle N_n, h_u\rangle} \middle| \mathcal{I}_n\right) \\ &= \sum_{i=1}^{|S_X|} \sum_{k=1}^{|S_H|} \sum_{m=1}^{|S_N|} \tilde{E}\left(\langle X_{n-1}, e_i\rangle \Lambda_{n-1} g(и) \left(\frac{p_{j,ikm}}{\tilde{p}_{j,ikm}}\right) \prod_{u=1}^{|S_N|} \left(\frac{a_n(u,m,i)}{\tilde{a}_n(u,m,i)}\right)^{\langle N_n, h_u\rangle} \middle| \mathcal{I}_n\right) \\ &= \sum_{i=1}^{|S_X|} \sum_{k=1}^{|S_H|} \sum_{m=1}^{|S_N|} \left(\frac{p_{j,ikm}}{\tilde{p}_{j,ikm}}\right) \tilde{E}(\langle X_{n-1}, e_i\rangle \Lambda_{n-1} g(и) | \mathcal{I}_{n-1}) = \sum_{i=1}^{|S_X|} \sum_{k=1}^{|S_H|} \sum_{m=1}^{|S_N|} \left(\frac{p_{j,ikm}}{\tilde{p}_{j,ikm}}\right) \int_\Theta б_{n-1}(i, u) g(u) du \end{aligned}$$

As g is arbitrary we see the result.

Write $g_n(щ, \varpi, и)$for the unnormalized joint conditional density of processes $\check{Q}_n$ and $\check{A}_n$:

$$g_n(щ, \varpi, и) = \tilde{E}\left(\Lambda_n \langle \check{Q}_n, b_щ \rangle \langle \check{A}_n, m_\varpi \rangle I(и \in dи) \middle| \mathcal{I}_n\right)$$

Theorem 3. Unnormalized joint conditional density $g_n(щ, \varpi, и)$ satisfies the recursion

$$g_n(щ, \varpi, и) = \langle \tfrac{B}{\tilde{B}} X_{n-1} \otimes H_{n-1} \otimes N_{n-1}, X_n \rangle \times \sum_{s=1}^{|S_H|} \langle \check{Q}_щ H_{n-1} \otimes N_{n-1}, f_s \rangle \times \sum_{u=1}^{|S_N|} \langle \check{A}_\varpi N_{n-1} \otimes X_{n-1}, h_u \rangle \times \sum_{х=1}^{Q|S_H||S_N|} \sum_{\vartheta=1}^{A|S_X||S_N|} d_{щх} k_{\varpi\vartheta} g_{n-1}(х, \vartheta, и)$$

The proof of this theorem can be done similarly as Theorem 5 in [3].

4. **Conclusion.** In this paper we considered two methods of coefficient estimation for the extended bonus-malus system constructed via hidden Markov models. EM algorithm estimations were expressed as fraction of number of jumps and occupation time. On the other hand the recursive estimations of parameters were expressed as previous estimation's correction with new information.

REFERENCES

1. Chitchyan R.N., Gulyan A.G. "A Bonus-Malus System with Aggregate Claim Amount Components", Izvestiya NAN Armenii (in English), v. 5, 2015 (in print).
2. Robert J. Elliott, Lakhdar Aggoun, John B. Moore, "Hidden Markov models: Estimation and Control", Springer-Verlag New York, 1995.
3. Aggoun, L., Benkgerouf, L. "An Insurance model with Bonus-Malus system", International Mathematical Forum 1, 2006, no. 26, pg. 1255-1272.