Some reward paths in semi Markov models with stochastic selection of the transition probabilities

Aleka A. Papadopoulou[1] and George M. Tsaklidis[2]

[1] General Department of Applied Sciences
Technological Institute of Thessaloniki,
P.O. Box 14561, 54101 Thessaloniki, Greece
(e-mail: alepapa@uth.gr)

[2] Department of Mathematics
Aristotle University of Thessaloniki,
Thessaloniki 54124, Greece
(e-mail: tsaklidi@math.auth.gr)

Abstract. In the present, the reward paths in non homogeneous semi Markov systems in discrete time are examined with stochastic selection of the transition probabilities. First, the basic probability equations of the reward paths are derived in terms of the main parameters of the system and a general formula is given. Then the expected rewards for the one unit time intervals are presented in relation to the entrance probabilities.
Keywords: Stochastic selection, semi-Markov process, reward.

1 Introduction

The definition of the non homogeneous semi Markov process was provided in Iosifescu-Manu (1972) for the continuous time case, in Janssen & De Dominicis (1984) for the discrete case and in De Dominicis & Manca (1985). A general definition of rewards can be found in Limnios & Oprisan (2001) and the study of the asymptotic behaviour of semi Markov reward process in Reza-Soltani & Khorshidian (1998). Later on the non homogeneous semi Markov system in discrete time was examined in Vassiliou and Papadopoulou (1992), and the asymptotic behavior of the same model was studied in Papadopoulou and Vassiliou (1994). Important theoretical results and applications for semi Markov models can be found in the work of Cinlar (1969,1975,1975), Teugels (1976), Pyke and Schaufele (1964), Keilson (1969,1971), Mclean and Neuts (1967), Howard (1971), McClean (1980,1986), Janssen (1986) and in Janssen and Limnios (1999). Continuing this effort in the present, we study the behaviour of the rewards paid during an interval of time along the reward paths. We consider rewards to be discrete random variables depending on the state occupancies, transition probabilities which are stochastically selected for every time unit, and the time spent at the state we examine before and after the time of reference. In order to examine the characteristics of the reward paths, we derive a general formula expressing the rewards per time unit for every state of the system and every time point. The expected reward for every time unit and every state can be easily evaluated by means of this general formula. Moreover the basic function used, helps us to look into the relative reward structure in the course of time. A specific type of the holding time probability functions leads to a characteristic result for the rewards evolution.

2 The semi Markov reward model with stochastic selection of the transition matrix

We consider a population which is stratified into a set of states according to various characteristics and we denote by $\mathcal{S} = \{1, 2, \ldots, N\}$ the set of states assumed to be exclusive and exhaustive, so that each member of the system may be in one and only one state at any given time. Time t is considered to be a discrete parameter and the state of the system at any given time could be described by the vector $\mathbf{N}(t) = [N_1(t), N_2(t), \ldots, N_N(t)]'$ where $N_i(t)$ is the expected number of members of the system in the i-th state at time t. The expected number of members of the system at time t is denoted by $T(t)$ and $N_{N+1}(t)$ is the expected number of leavers during the time interval $(t-1, t]$. We assume that $T(t) = T$, i.e. the total number of leavers equals to the total number of recruits for every t and that the individual transitions between the states occur according to a non homogeneous semi Markov chain *(embedded non homogeneous semi Markov chain)*. In this respect we denote by $\mathbf{F}(t)_{t=0}^{\infty}$ the sequence of matrices, the (i, j)th element of which is the probability of a member of the system to make its next transition to state j, given that it entered state i at time t. Let also $\mathbf{p}_{N+1}(t)$ be the Nx1 vector whose i-th element is the probability of leaving the system from i, given that the entrance in state i occured at time t and $\mathbf{p}_o(t)$ the Nx1 vector the j-th element of which is the probability of entering the system in state j as a replacement of a member who entered his last state at time t. A member entering the system holds a particular membership which moves within the states with the members (see also Bartholomew (1982), Vassiliou and Papadopoulou (1992), Vassiliou *et al.* (1990)). Since the size of the system is constant, when a member decides to leave the system, the empty membership is taken by a new recruit who behaves like the former one. Denote by $\mathbf{P}(t)$ the matrix described by the relation

$$\mathbf{P}(t) = \mathbf{F}(t) + \mathbf{p}_{N+1}(t)\mathbf{p}'_o(t)$$

Obviously $\mathbf{P}(t)$ is a stochastic matrix with the (i, j)th element equal to the probability that a membership of the system which entered state i at time t makes its next transition to state j. For the present, we consider that the transition probability matrix $\mathbf{P}(t)$ is selected from a pool of matrices $\mathbf{L} = [\mathbf{P}_1(t), \mathbf{P}_2(t), \ldots, \mathbf{P}_v(t)]$ with corresponding probabilities

$c_1(t), c_2(t), \ldots, c_v(t)$. Thus, whenever a membership enters state i at time t, it selects state j for its next transition according to the probabilities $p_{ij}(t)$. However before the entrance into j, the membership 'holds' for a time in state i. Holding times for the memberships are described by the holding time mass function $h_{ij}(m)$ which equals to the probability, a membership which entered state i at its last transition holds m time units in i before its next transition, given that state j has been selected.

Let also $y_{ij}(t)$ be the reward that a membership earns at time t after entering state i for occupying state i during the interval $[t, t+1)$ when its successor state is j, and $b_{ij}(m)$ be the bonus reward that the membership earns for making a transition from state i to j, after holding time m time units in state i. Thus if a membership enters state i at time s and decides to make a transition to j after m time units in i, then the total reward that it earns equals to

$$\sum_{t=s}^{s+m-1} y_{ij}(t) + b_{ij}(m).$$

Now, denote:

$A_{ij;k}(t)$ ={the reward that a membership earns during the time interval $[t, t+1)$ given that the membership entered state i at time 0, possesses state j at time t, and will undertake its next transition to state k}.
Moreover, entering some state j implies stay at j at least one time unit. Also, denote

$e_{ij}(n,t)$ =prob{that a membership which entered state i at time t will enter state j after n time units},
with corresponding probability matrix $\mathbf{E}(n,t) = \{e_{ij}(n,t)\}$. It is apparent that $e_{ij}(n,t)$ depend on the transition probabilities $p_{ij}(t)$ (see also Papadopoulou (1997)) or equivalently on the transition probability matrices $\mathbf{P}(t)$ which are selected from the pool $\mathbf{L} = [\mathbf{P}_1(t), \mathbf{P}_2(t), \ldots, \mathbf{P}_v(t)]$ with probabilities $c_1(t), c_2(t), \ldots, c_v(t)$. Thus $e_{ij}(n,t)$ become (define) random variables and we are interested in the expected value of matrix $\mathbf{E}(n,t)$. From Papadopoulou (1997) we have that

$$\mathbf{E}(n,t) = \mathbf{P}(t)\diamond\mathbf{H}(n) + \textstyle\sum_{j=2}^{n}\{\mathbf{P}(t)\diamond\mathbf{H}(j-1)\{\mathbf{P}(t+j-1)\diamond\mathbf{H}(n-j+1)\} + \sum_{j=2}^{n}\sum_{k=1}^{j-2}\mathbf{S}_j(k,s,m_k)\{\mathbf{P}(t+j-1)\diamond\mathbf{H}(n-j+1)\},$$

for every $n \geq 1$ and $\mathbf{E}(0,t) = \mathbf{I}$, where:
$\mathbf{P}(t)\diamond\mathbf{H}(n)$ is the Hadamard product of the matrices $\mathbf{P}(t)$, $\mathbf{H}(n)$,

$$\mathbf{S}_j(k,s,m_k) = \textstyle\sum_{m_k=2}^{j-k}\sum_{m_{k-1}=1+m_k}^{j-k+1}\cdots\sum_{m_1=1+m_2}^{j-1}\prod_{r=-1}^{k-1}\{\mathbf{P}(t+m_{k-r}-1)\diamond\mathbf{H}(m_{k-r-1}-m_{k-r})\}.$$

where the i, r element of $\mathbf{S}_j(k,s,m_k)$ is the probability that a membership which entered state i at time s makes a transition to state r after $j-1$ time units and k intermediate transitions during the interval $(s, s+j-1)$. Thus,

it is easily seen that
$E[\mathbf{E}(n,t)] = E[\mathbf{P}(t)]\diamond\mathbf{H}(n)+\sum_{j=2}^{n}\{E[\mathbf{P}(t)]\diamond\mathbf{H}(j-1)\}\{E[\mathbf{P}(t+j-1)]\diamond\mathbf{H}(n-j+1)\} + \sum_{j=2}^{n}\sum_{k=1}^{j-2}\tilde{\mathbf{S}}_j(k,s,m_k)\{E[\mathbf{P}(t+j-1)]\diamond\mathbf{H}(n-j+1)\}$,
where:
$E[\mathbf{P}(t)] = \sum_{x=1}^{v} c_x(t)\mathbf{P}_x(t)$,
$\tilde{\mathbf{S}}_j(k,s,m_k,\beta) = \sum_{m_k=2}^{j-k}\sum_{m_{k-1}=1+m_k}^{j-k+1}\cdots\sum_{m_1=1+m_2}^{j-1}\prod_{r=-1}^{k-1}E[\mathbf{P}(t+m_{k-r}-1)]\diamond\mathbf{H}(m_{k-r-1}-m_{k-r})$.
There are three different ways for a membership starting from state i (at time $t=0$) to occupy state j at time t. The three different ways are exclusive and exhaustive, and are illustrated below:

time t	time $t+1$
new entrance in j	new entrance in k
age in j equal to $m_p \geq 0$	residual life in j equal to $m_f > 0$
age in j equal to $m_p > 0$	new entrance in k

Thus $A_{ij;k}(t)$ is a random variable taking the values: $y_{jk}(t)+b_{jk}(1)$, $y_{jk}(t)$, $y_{jk}(t)+b_{jk}(m_p+1)$. The corresponding probabilities can be easily evaluated, for example:

$P\{A_{ij;k}(t) = y_{jk}(t)+b_{jk}(1)\}$
=prob{a membership which entered state i at time 0 will enter state j at time t}·prob{a membership which entered state j at time t will take its next transition to state k}·prob{a membership which entered state j at its last transition holds one time unit in j before its next transition given that state k has been selected}
$= e_{ij}(0,t)p_{jk}(t)h_{jk}(1)$.
Similarly we have:

$$P\{A_{ij;k}(t) = y_{jk}(t)\} = \sum_{m_p}\sum_{m_f} e_{ij}(0,t-m_p)p_{jk}(t-m_p)h_{jk}(m_p+m_f+1)$$
where $m_p+m_f \leq M-1$, $m_f \neq 0$, $M \in N$,
$$P\{A_{ij;k}(t) = y_{jk}(t)+b_{jk}(m_p+1)\} = e_{ij}(0,t-m_p)p_{jk}(t-m_p)h_{jk}(m_p+1).$$

The three cases given above can be summarized in the following general formula:

$$P\{A_{ij;k}(t) = y_{jk}(t)+\delta_{(m_f-1)}b_{jk}(m_p+m_f+1) = \sum_{m_p}\sum_{m_f} E[e_{ij}(0,t-m_p)]E[p_{jk}(t-m_p)]h_{jk}(m_p+m_f+1)$$
where $\delta_{(m_f-1)}$ stands for the unit impulse , i.e. $\delta_{(n)} = \begin{cases} 1, \text{ if n=0} \\ 0, \text{ if n}\neq 0 \end{cases}$.

Now, let us number by $1, 2, ..., N_i(0)$ the memberships having started their motion from state i, and denote by $A_i^{(r)}(t)$ the reward of the r-th membership paid in the time interval $[t, t+1)$ and by $A_i(t)$ the rewards paid to all the $N_i(0)$ memberships having started their motion from state i.

Let us assume that the r-th membership having started its motion from state i, possesses at time t state j, having hold for m_p time units in j and having attained the next state k after m_f time units. Then we correspond to the r-th membership a $N \times N \times M \times M$ vector (M stands for the bound of m_p, m_f) having the value

$$A_{ij;k}^{(r)}(t; m_p, m_f) = y_{jk}(t) + \delta_{(m_f-1)} b_{jk}(m_p + m_f)$$

in the position $(i-1)NM^2 + (j-1)M^2 + m_p M + m_f$ and 0 elsewhere.Then, the total reward paid in the interval $[t, t+1)$ for the memberships having started their motion from state i, is the r.v.

$$A_i(t) = \sum_{r=1}^{N_i(0)} A_i^{(r)}(t).$$

Symbolize by $f_i^{(r)}(t)$ the probability generating function (p.g.f.) of $A_i^{(r)}(t)$ and by $F_i(t)$ the p.g.f. of $A_i(t)$. Since the r.v's. $A_i^{(r)}(t)$ are independent with common p.g.f. $f_i^{(r)}(t) = f_i(t)$, $r = 1, 2, .., N_i(0)$, then

$$F_i(t) = \prod_{r=1}^{N_i(0)} f_i^{(r)}(t) = (f_i(t))^{N_i(0)}.$$

3 Conclusions

In the present paper we have derived, for the discrete time semi Markov system, formulas providing the probabilities of the rewards for one unit time interval by means of the entrance probabilities, the transition probabilities and the probabilities of the holding times. Then, relations for the evaluation of the total reward paid in one unit time interval to all the memberships of the system are given. The conclusions can be generalized for various reward paths of the memberships, and a reasonable perspective is treating the same questions for the continuous time case.

References

1.Bartholomew, D.J. (1982): *Stochastic models for social processes.* Wiley, Chichester.

2.Cinlar, E (1969): Markov renewal theory. *Adv. Appl. Prob.*, *1*, 123–187.

3.Cinlar, E (1975): Markov renewal theory: a survey. *Management Sci.*, *21*, 727–752.

4.Cinlar, E (1975): *Introduction to stochastic processes.*, Prentice-Hall, Englewood Cliffs, NJ.

5.De Dominicis, R. and R. Manca (1985): Some new results on the transient behaviour of semi Markov reward processes. *Methods of operations research and Computation*, *13*, 823–838.

6.Howard, R.A. (1971): *Dynamic Probabilistic systems*. Wiley, Chichester.

7.Iosifescu - Manu, A. (1972): Non homogeneous semi Markov processes. *Studiisi Cercetuari Matematice*, *24*, 529–533.

8.Janssen, J. (1986): *Semi-Markov models: Theory and Applications*. ed. J. Janssen, Plenum Press, New York.

9.Janssen, J. and R. De Dominics (1984): Finite non homogeneous semi Markov processes: Theoretical and computational aspects. *Insurance: Mathematics and Economics*, *3*, 157–165.

10.Janssen, J. and N. Limnios (1999): *Semi-Markov models and Applications*. J. Janssen and N. Limnios Eds, Kluwer Academic Publishers, Dordrecht.

11.Keilson, J (1969): On the matrix renewal function for Markov renewal processes. *Ann. Math. Statist.*, *40*, 1901–1907.

12.Keilson, J (1971): A process with chain dependent growth rate. Markov Part II: the ruin and ergodic problems. *Adv. Appl. Prob.*, *3*, 315–338.

13.Limnios, N. and G. Oprisan (2001): *Semi-Markov processes and reliability*. Birkhauser, Boston.

14.McClean, S.I. (1980): A semi Markovian model for a multigrade population. *J. Appl. Prob.*, *17*, 846–852.

15.McClean, S.I. (1986): Semi-Markov models for Manpower planning. In *Semi-Markov models: Theory and Applications*. Plenum Press, New York.

16.Mclean, R. A. and M. F. Neuts (1967): The integral of a step function defined on a semi Markov process. *Siam J. Appl. Math.*, *15*, 726–737.

17.Papadopoulou A.A. (1997): Counting transitions -Entrance probabilities in non homogeneous semi-Markov systems. *Applied Stoch. Models and Data Analysis*, *13*, 199–206.

18.Papadopoulou, A.A. & P.-C.G. Vassiliou (1994): Asymptotic behavior of non homogeneous semi-Markov systems. *Linear Algebra and Its Applications*, *210*, 153-198.

19.Pyke, R. and R. A. Schaufele (1964): Limit theorem for Markov renewal process. *Ann. Math. Statist.*, *55*, 1746–1764.

20.Reza Soltani, A. and K. Khorshidian (1998): Reward processes for semi-Markov processes: Asymptotic behaviour. *J. Appl. Prob.*, *35*, 833–842.

21.Teugels J.L. (1976): A bibliography on semi-Markov processes. *J. Comp. Appl. Math.*, *2*, 125–144.

22.Vassiliou, P.-C.G. and A.A. Papadopoulou (1992): Non homogeneous semi-Markov systems and maintainability of the state sizes. *J. Appl. Prob.*, *29*, 519–534.

23.Vassiliou, P.-C.G. , A. Georgiou and N. Tsantas (1990): Control of asymptotic variability in non homogeneous Markov systems. *J. Appl. Prob.*, *27*, 756–766.