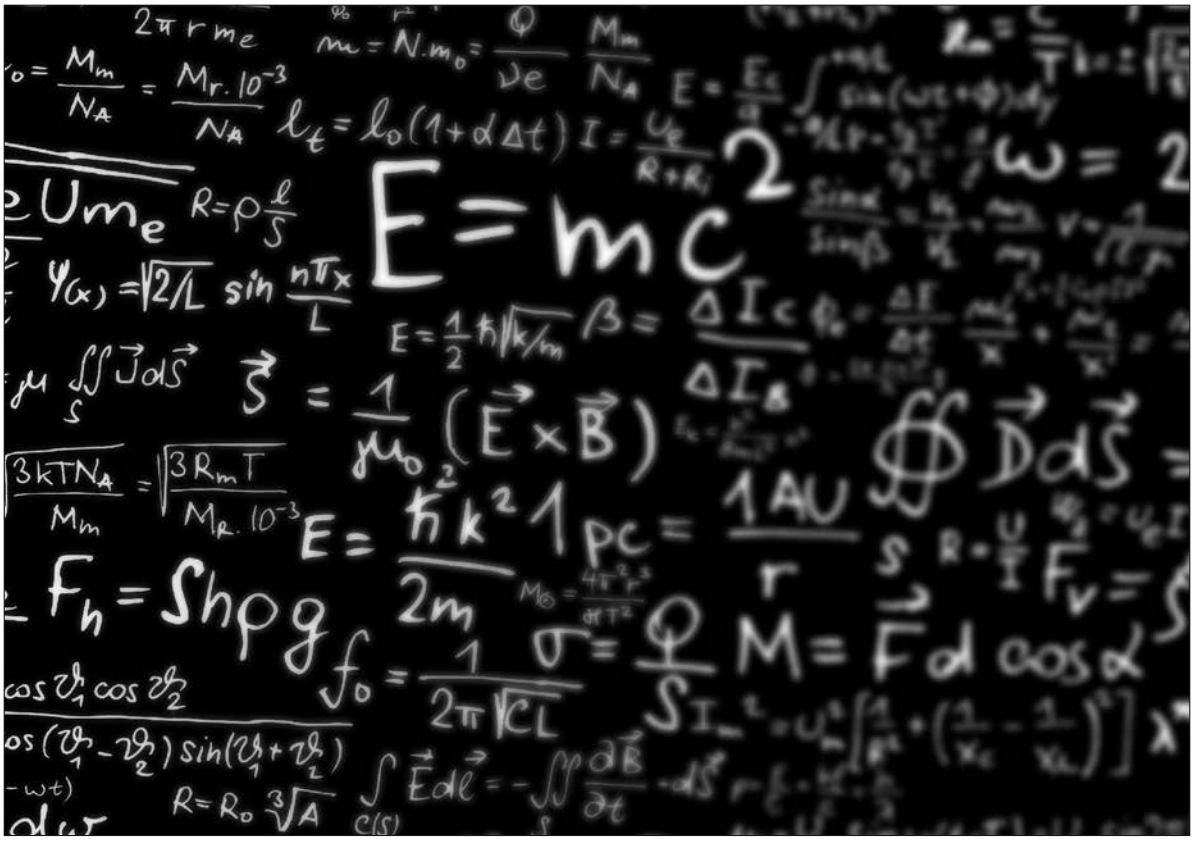

Asian Journal of Mathematics & Statistics

ISSN 1994-5418

science alert

ANSInet
an open access publisher
http://ansinet.com

Asian Journal of Mathematics and Statistics 3 (4): 310-315, 2010
ISSN 1994-5418 / DOI: 10.3923/ajms.2008.57.62

Strong Law of Large Numbers for Nonlinear Semi-Markov Reward Processes*

K. Khorshidian
Department of Statistics, Shiraz University, Shiraz, Iran

Abstract: We obtain a Strong Law of Large Numbers (SLLN) for the reward process {Z(t), $t \geq 0$}, the cumulative reward gained by operating a Semi-Markov system during the time interval [0, t]. The important and striking point in this study is leaving the usual assumption that rewards for each state are of constant rates. In most of applications this frequently used assumption is not realistic, therefore we deal with reward functions of general forms. The SLLN obtained is in the sense that $\text{Lim}_{t\to\infty} Z(t)/t = \alpha$, a.s. for some real value α. Mild conditions for this SLLN are existence of sojourn times means and integrability of reward functions with respect to sojourn time distributions. As it has been shown, the parameter α coincides with the shift parameter in asymptotic representation of $E[Z(t)]$, $t\to\infty$.

Key words: Semi-Markov processes, reward processes, law of large numbers

INTRODUCTION

In this study we deal with long run rate of reward processes defined on a Semi-Markov process, when the reward functions are not necessarily of linear form. The main difference between this work and other similar subjects, is inherent in the structure of functionals that we have mentioned. A Semi-Markov Process (SMP), or equivalently a Markov-Renewal Process (MRP) is an extension of a Markov Chain (MC) tied with ordinary renewal processes for each state. The successive states visited by this process forms a Markov chain, while the sojourn time in each state is not a constant and probabistically depends on the current state and may depend on the next state to be visited. It is well known that given the sequence of the states, the sequence of the sojourn times are independent. More details on Semi-Markov processes can be found in Cinlar (1969 and 1975a and b), Pyke (1961a and b) and Teugels (1976).

For processes with regenerative structure, a cost or reward function is often introduced. The total cost or reward gained corresponding to the performance of the system as a function of the time is called a reward process. In last decades many authors have considered different aspects of reward processes. Some have dealt with mean value of these processes and because of complicated forms of exact solutions obtained, most of these authors have tried for limiting behavior of mean value in long run of the process.

The definition of a reward process on a Semi-Markov process, backs to the first given by Pyke and Schaufele in (1964) as:

$$W_f(t)=\sum_{n=1}^{N(t)} f(J_{n-1}, J_n, \xi_n - \xi_{n-1}) \quad (1)$$

Where, J_n and ξ_n are the state and epoch of n-th transition and f is a real valued function. The mean and variance for these processes in univariate and multivariate case were obtained by Sumita and Masuda (1987) and Masuda and Sumita (1991), when they considered a reward process of the form:

Corresponding Author: K. Khorshidian, Department of Mathematics and Statistics, Persian Gulf University, Booshehr, Iran

Originally Published in Asian Journal of Mathematics and Statistics, 2008

$$Z(t)=\int_0^t \rho(J(s))ds=\sum_{n=0}^{N(t)-1}\rho(J_n)\Delta_n \tag{2}$$

Where, $\Delta_n = \xi_{n+1}-\xi_n$ is the sojourn time at state J_n. $\rho(k)$ is a constant at state k. Note that Eq. 2 is a special case of Eq. 1, when f is of constant rate and hence it is linear in Δ_n. They obtained exact formulas as well as asymptotic behavior for mean value of the process $E[Z(t)]$, $t\geq 0$ in univariate and multivariate case. Soltani (1996), introduced the nonlinear reward process as:

$$Z(t)=\sum_{n:\xi_{n+1}<t}\rho(J_n,\xi_{n+1}-\xi_n)+\rho(J(t),\gamma(t)) \tag{3}$$

Where, γ is the age process. The function ρ in Eq. 3 is called a reward function, where $\rho(j, \tau)$ measures the excess reward when time τ is spent in the state j. A closed form for mean of this process, when the reward functions assumed to be polynomials in time was obtained by Soltani (1996). Khorshidian and Soltani (2002 and 2003), obtained $E[Z(t)]$ and $Var[Z(t)]$ the expected value and variance of $Z(t)$, as well as $\Sigma(t)$, the covariance matrix for a multivariate reward process with components as Eq. 3. Using Markov-renewal theory and for general $\rho(.,.)$, Soltani and Khorshidian (1998), showed that

$$E[Z(t)]=\alpha_0+\alpha_1 t+\circ(1),\quad t\to\infty \tag{4}$$

Where, the constants α_0 and α_1 were fully specified. For reward functions of general form and in multivariate case, Khorshidian and Soltani 2002, used supplementary variables technique and showed that as $t\to\infty$,

$$E[\underline{Z}(t)]=A_0+A_1 t+\circ(1),\qquad \frac{\Sigma(t)}{t}=C_0+C_1 t+\circ(1) \tag{5}$$

Moreover, they calculated the vectors A_0 and A_1 and matrices C_0 and C_1.

In probability theory it is well known that for a sequence of random variables $\{X_n, n = 1, 2, ...\}$, with partial means (partial sums) $\{Y_n, n = 1, 2, ...\}$, if the limits exist and are constant values, $\lim_{n\to\infty} E[Y_n]$ does not necessarily coincides with $\lim_{n\to\infty} Y_n$. In what follows we will show that under some mild conditions $\lim_{t\to\infty} Z(t)/t = \lim_{t\to\infty} E[Z(t)]/t$, this is the SLLN we purposed.

To arrive at the mentioned strong law for the process Eq. 3, since the reward functions $\rho(.,.)$ assumed to have general forms, it is natural to impose some conditions on it. Most of the conditions we need are about integrability of ρ, just what has been done for calculating moments of $Z(t)$, by Soltani and Khorshidian (1998) and Khorshidian and Soltani (2002).

NOTATION AND PRELIMINARIES

Let $\{J(t), t\geq 0\}$ be a SMP with a MRP, $\{(J_n, \xi_n), n = 0, 1, 2, ...\}$. The state space of $\{J_n\}$ is assumed to be $E = \{s_0, s_1, ..., s_m\}$. J_n and ξ_n are the state and epoch of n-th transition $n\geq 0$ with $\xi_0 = 0$. Also let $\Delta_n = \xi_{n+1}-\xi_n$ denotes the sojourn time at J_n. The transition probabilities are purposed to be

$$\begin{aligned}P[J_n=j,\Delta_n\leq x\,|\,\xi_0,\ldots,\xi_n;J_0,\ldots,J_n=i] &= P[J_n=j,\ \Delta_n\leq x\,|\,J_n=i]\\ &= A_{ij}(x),\end{aligned}$$

$A_{ij}(x)=P_{ij}F_i(x)$, where $P_{ij}=A_{ij}(+\infty)$ are transition probabilities of the embeded Markov chain and $F_i(.)$ is the distribution of sojourn time at state i. The mean sojourn time at state i is:

$$m(i) = \int_0^\infty xF_i(dx)$$

The invariant stationary distribution of the underlying MC is $\underline{\pi} = (\pi_i),\ i \in E,\ \underline{\pi} = \underline{\pi}P$ and $\underline{p}_0$ is the initial probability row vector with $p_0(i) = P[J_0 = i],\ i \in E$.

The number of transitions in [0, t] is denoted by N(t). It is well known that if the SMP is irreducible, recurrent and the sojourn times have finite means, then

$$\lim_{t\to\infty}\frac{N(t)}{t} = \delta^{-1}, \qquad \delta = \sum_{i\in E}\pi_i m(i) \tag{6}$$

The elapsed life of the process after the last transition is the age process $\gamma(t) = t - \xi_{N(t)}$.

Throughout the study we assume that the reward functions ρ in Eq. 3 are nonnegative. For reward functions with possibly negative values, without loss of generality and by adding a constant finite value M, all of the undergoing results goes true, provided that $\inf_{u\in[0,\infty)} \min_{k\in E}\rho(k,u) > -M$.

SLLN FOR REWARD PROCESSES

In this section, using limit theorems for MC's and under integrability of reward functions $\rho(J_.,)$, with respect to the sojourn time distributions $F_j(.)$, we will arrive at a SLLN for Z given by Eq. 3.

Let $\{(J_n, \xi_n);\ n = 0,1,2,...\}$ ba an irreducible MRP, with finite state space E and corresponding SMP, $\{J(t),\ t \geq 0\}$ with $\underline{\pi}$ as the invariant stationary distribution of the underlying Markov chain $\{J_n,\ n \geq 0\}$. It is well known that when the state space is finite, the corresponding MC and MRP are positive recurrent and in this case,

$$\lim_{n\to\infty}(1/n)\sum_{k=0}^{n-1} f(J_n) = E_\pi[f(J_0)] = \sum_{j\in E}\pi_j f(j), \qquad \text{a.s.} \tag{7}$$

for any real valued function f. In the next theorem we will drive more general limiting results for summands similar to Eq. 7, with a bivariate reward function ρ: $E\times[0, \infty)\to[0, \infty)$ instead of f and a MRP instead of the MC. Thereafter we will show that a SLLN holds for $\{Z(t),\ t \geq 0\}$, in the sense that $\lim_{t\to\infty} Z(t)/t$ exists, under the assumptions that

$$\theta_0(j) = \int_0^\infty \rho(j,u)F_j(du) \quad \text{and} \quad m(j) = \int_0^\infty uF_j(du)$$

are finite for all $j \in E$.

Theorem 1

Suppose that $\{(J_n, \xi_n),\ n \geq 0\}$ is an irreducible recurrent MRP with finite state space E and stationary distribution $\underline{\pi}$. Also suppose that $\max_{j\in E} m_j$ and $\theta_0(j) < \infty$, for all $j \in E$, then

$$\lim_{t\to\infty}\frac{Z(t)}{t} = \delta^{-1}\sum_{j\in E}\pi_j\theta_0(j) = \alpha, \quad \text{a.s.}$$

where, Z(t) is given by Eq. 3 and δ by Eq. 6.

Proof: First note that,

$$\xi_{N(t)} \le t < \xi_{N(t)+1} \quad \text{and} \quad Z(\xi_{N(t)}) \le Z(t) \le Z(\xi_{N(t)+1})$$

and therefore,

$$\frac{N(t)}{\xi_{N(t)+1}} \frac{Z(\xi_{N(t)})}{N(t)+1} < \frac{Z(t)}{t} \le \frac{Z(\xi_{N(t)+1})}{N(t)} \frac{N(t)+1}{\xi_{N(t)})} \tag{8}$$

Also irreducibility assumption and finiteness of E implies that, $\lim_{t\to\infty}(N(t)/t) = \delta^{-1}$ a.s., which gives

$$\lim_{t\to\infty} \frac{N(t)}{\xi_{N(t)+1}} = \lim_{t\to\infty} \frac{N(t)+1}{\xi_{N(t)}} = \delta^{-1} \tag{9}$$

Moreover,

$$\begin{aligned} \lim_{t\to\infty} \frac{Z(\xi_{N(t)})}{N(t)+1} &= \lim_{t\to\infty} \frac{Z(\xi_{N(t)+1})}{N(t)} \\ &= \lim_{n\to\infty}(1/n)\sum_{k=0}^{n-1}\rho(J_k, \Delta_k) \\ &= \sum_{j\in E}\pi_j \int_0^\infty \rho(j,u)F_j(du) \\ &= \sum_{j\in E}\pi_j\theta_0(j) \end{aligned} \tag{10}$$

by the fact that $N(t)\to\infty$ as $t \to\infty$ and the argument will been given below for Eq. 10, So by Eq. 9 and 10

$$\begin{aligned} \lim_{t\to\infty}\frac{Z(t)}{t} &= \lim_{t\to\infty}\frac{N(t)}{\xi_{N(t)}}\frac{Z(\xi_{N(t)})}{N(t)} \\ &= \lim_{t\to\infty}\frac{N(t)}{\xi_{N(t)}}\lim_{t\to\infty}\frac{Z(\xi_{N(t)})}{N(t)} \\ &= \delta^{-1}\sum_{j\in E}\pi_j\theta_0(j). \end{aligned}$$

It remains to show that Eq. 10 holds in the sense that:

$$\lim_{n\to\infty}(1/n)\sum_{k=0}^{n-1}\rho(J_k, \Delta_k) = \sum_{j\in E}\pi_j\theta_0(j)$$

To do this, denote by $N_n(i)$ the number of visits to state i during first n-th transitions of the MRP and by subsequence k(1, i), k(2, i),... those indices at them the process visits i. Also, by $Y_{k(1,i)}$, $Y_{k(2,i)}$ denote the sojourn times in state i, during the above subsequence. Given J_0, J_1, J_2,..., the subsequence $Y_{k(1,i)}$, $Y_{k(2,i)}$ are IID with distributions F_i. Moreover, by Positive recurrence of the underlying MC, $N_n(i)\to\infty$ as $n\to\infty$ for all $i\in E$ and therefore by SLLN for IID random variables we have:

$$\begin{aligned} \frac{1}{N_n(i)}\sum_{j=1}^{N_n(i)}\rho(i, Y_{k(j,i)}) &\to E[\rho(i, Y_{k(j,i)})] \\ &= \int_0^\infty \rho(i,y)F_i(dy) \\ &= \theta_0(i) \end{aligned} \tag{11}$$

Thus,

$$\begin{aligned}(1/n)\sum_{k=0}^{n-1}\rho(J_k,\Delta_k) &= (1/n)\sum_{k=0}^{n-1}\sum_{i\in E}\rho(J_k,\Delta_k)I_{[J_k=i]}\\ &= (1/n)\sum_{i\in E}\sum_{j=1}^{N_n(i)}\rho(i,Y_{k(j,i)})\\ &= \sum_{i\in E}\frac{N_n(i)}{n}\sum_{j=1}^{N_n(i)}\frac{1}{N_n(i)}\rho(i,Y_{k(j,i)})\end{aligned} \tag{12}$$

Since the MRP is assumed to be irreducible, recurrent with finite state space, we conclude that as $n\to\infty$, $N_n(i)/n\to\pi_i$ and therefore Eq. 11 and 12 imply that:

$$\begin{aligned}\lim_{n\to\infty}(1/n)\sum_{k=0}^{n-1}\rho(J_k,\Delta_k) &= \lim_{n\to\infty}\sum_{i\in E}\frac{N_n(i)}{n}\sum_{j=1}^{N_n(i)}\frac{1}{N_n(i)}\rho(i,Y_{k(j,i)}(i))\\ &= \sum_{i\in E}\pi_i\theta_0(i)\end{aligned} \tag{13}$$

The proof is now complete.

Remark

As it has been expected, the limiting value α in Theorem 1, coincides with slope parameter α_1 in Eq. 4, which has been derived by Soltani and Khorshidian (1998), for asymptotic representation of E[Z(t), $t\to\infty$. Also it coincides with elements of the vector A_1, in Eq. 5 for multivariate {Z(t), $t\geq 0$, Moreover the assumptions on the reward functions and sojourn times for the above SLLN, are just as that which has been assumed for existence of E[Z(t)], $t\to\infty$ (Khorshidian and Soltani, 2002).

In the Theorem 1, finiteness of the state space is only used for embedded MC to have positive recurrent states. If in the infinite state space case, we impose conditions that by them, the states would be positive recurrent, then the result of this theorem hold. Therefore, we can arrive at the following corollary.

Corollary

Suppose that $\{(J_n, \xi_n), n\geq 0\}$ is an irreducible positive recurrent MRP with infinite state space E and stationary distribution $\underline{\pi}$. Also suppose that $m_j<\infty$ and $\theta_0(j)<\infty$, for all $j\in E$ then

$$\lim_{t\to\infty}\frac{Z(t)}{t}=\delta^{-1}\sum_{j\in E}\pi_j\theta_0(j)=\alpha,\quad \text{a.s.}$$

where Z(t) is given by Eq. 3 and δ by Eq. 6.

REFERENCES

Cinlar, E., 1969. Markov renewal theory. Adv. Applied Prob., 1: 123-187.

Cinlar, E., 1975a. Markov renewal theory: A survey. Manage. Sci., 21: 727-752.

Cinlar, E., 1975b. Introduction to Stochastic Processes. Prentice-Hall, Inc.

Khorshidian, K. and A.R. Soltani, 2002. Asymptotic behavior of multivariate reward processes with nonlinear reward functions. Bull. Iran. Math. Soc., 28: 1-17.

Khorshidian, K. and A.R. Soltani, 2003. Exact solutions for covariance matrix of multivariate reward processes with analytic reward functions. J. Sci. I.R. Iran, 14: 173-183.

Masuda, Y. and U. Sumita, 1991. A multivariate reward processes defined on a Semi-Markov process and its first passage time distributions. J. Applied Prob., 28: 360-373.

Pyke, R., 1961a. Markov renewal processes: Definitions and preliminary properties. Ann. Math. Statist., 32: 1231-1242.

Pyke, R., 1961b. Markov renewal processes finitely many states. Ann. Math. Stat., 32: 1243-1259.

Pyke, R. and R. Schaufele, 1964. Limit theorem for markov renewal processes. Ann. Math. Statist., 35: 1746-1764.

Soltani, A.R., 1996. Reward processes with nonlinear reward functions. J. Applied Prob., 33: 1011-1017.

Soltani, A.R. and K. Khorshidian, 1998. Reward processes for Semi-Markov processes: Asymptotic behavior. J. Applied Prob., 35: 833-842.

Sumita, U. and Y. Masuda, 1987. An alternative approach to the analysis of finite Semi-Markov and related processes. Commun. Stat. Stochastic Models, 3: 67-87.

Teugels, J.L., 1976. A bibliography on Semi-Markov processes. J. Comp. Applied Math., 2: 125-144.