Brazilian Journal of Probability and Statistics (2002), 16, pp. 69–85.

KERNEL ESTIMATORS FOR SEMI-MARKOV PROCESSES

Gregorio S. Atuncar[1] and **Krishna B. Athreya**[2]

[1] *Universidade Federal de Minas Gerais, Departamento de Estatística, Caixa Postal 702, 31270-901, Belo Horizonte, MG, Brazil. Email: gregorio@est.ufmg.br*

[2] *Iowa State University, Department of Statistics. Email: kba@iastate.edu*

Summary
In this article we define kernel estimator of sojourn time distribution for Semi - Markov processes and prove its consistency and asymptotic normality. We use techniques of regenerative processes to prove the results.

Key words: Kernel estimator; semi-Markov processes; sojourn time distribution.

1 Introduction and preliminaries

In this paper, sojourn time distributions for semi-Markov processes are studied. A semi - Markov process is a continuous - time stochastic process $\{Y_t : t \geq 0\}$, which makes transitions from state to state in accordance with a Markov chain, but the amount of time spent in each state before a transition occurs is a random variable depending on the transitions. So $\{Y_t : t \geq 0\}$ can be described as a marked process (T_1, Y_1), $(T_2, Y_2), \ldots,$ where $\{Y_n\}$ is the embedded Markov process and $\{T_1, T_2, \ldots\}$ are the sojourn times. As an example, consider a Shock model, a failure model where a system is studied in a random environment. The system is subject to a sequence of randomly occurring shocks and each shock causes a random amount of damage which accumulate over time. Let $\{T_n\}$ be the time points at which shocks occur and let Y_t be the accumulated damage level at time t ($Y_n = Y(T_n)$). An interesting problem is related to the distribution time for the next shock when the damage is at level y.

Let $\mathbf{X} = \{X_n : n = 0, 1, \ldots\}$ be a Markov chain with state space (S, Σ). In what follows we assume S the set $\Re$ of real numbers and $\Sigma = B(\Re)$, the Borel σ-algebra. Assume that there exists a point $\Delta \in \Re$ such that

$P_x(T_\Delta < \infty) = 1$ and $E_\Delta T_\Delta < \infty$. It can be seen in Atuncar (1994) that if $\{X_n : n = 0, 1, \ldots\}$ is a Harris chain then for some $n_0 \geq 1$, $\{X_{nn_0} : n = 0, 1, \ldots\}$ is a Markov chain for which a distinguished point Δ can be constructed by enlarging the state space. Thus the results of this study apply to Harris chains.

Let $\{G(x,.) : x \in \Re\}$ be a family of distribution functions such that $G(x, 0) = 0$ for all $x \in \Re$. Let $\{T_n : n = 0, 1, \ldots\}$ be a sequence of independent random variables such that, given $\{X_n : n = 0, 1, \ldots\}$,

$$P(T_n \leq t \mid \{X_n : n = 0, 1, \ldots\}) \quad = \quad G(X_n, t) \qquad \text{for all } t \geq 0.$$

These $\{T_n : n = 0, 1, \ldots\}$ are called sojourn times. Let $S_0 = 0$, $S_n = \sum_{i=0}^{n-1} T_i$ for $n \geq 1$ and let

$$Y_t \quad = \quad X_n \qquad\qquad S_n \leq t < S_{n+1},\ n = 0, 1, 2, \ldots$$

Since $P_x(X_n = \Delta \text{ for some } n \geq 1) = 1$, $\{X_n : n = 0, 1, \ldots\}$ hits Δ infinitely often, say at N_1, $N_2, \ldots$ and so by the strong law of large numbers, $\sum_i T_{N_i} = \infty$ with probability 1 and hence $S_n \to \infty$ with probability 1 and so Y_t is well defined for all t. The process $\{Y_t : t \geq 0\}$ has state space $\Re$ and is not Markovian unless $G(x,.)$ is exponential for all x. However, $\{Y_{S_n} = X_n : n = 0, 1, \ldots\}$ is still a Markov chain. Thus $\{Y_t : t \geq 0\}$ sampled at $t = S_n$, $n = 0, 1, 2, \ldots$ is Markov but not for all $t \geq 0$. For this reason it is called a **semi-Markov process.**

Results for kernel estimators for the stationary density f and the transition density t of the Markov chain $\{X_n : n = 0, 1, \ldots\}$ were established in Atuncar (1994).

The main goals in this paper will be to propose estimator for $G(x, t)$ and prove its properties. We observe the process up to time n. Besides the information $\{X_0, X_1, \ldots, X_n\}$, we have $\{T_0, T_1, \ldots, T_{n-1}\}$ where, for $i = 0, 1, \ldots, n-1$, T_i is the sojourn time in X_i.

Fix $x \in \Re \setminus \{\Delta\}$, the set of the real numbers different than Δ and take $A_n = (x - \delta_n, x + \delta_n)$. During $\{0, 1, \ldots, n\}$ let $N_1, N_2, \ldots, N_{L_n}$ be the times of visits by the process to A_n; i.e., for $i = 1, 2, \ldots, L_n$, $X_{N_i} \in A_n$; and let T_{N_i} be the corresponding sojourn times. Given $X_0, X_1, \ldots, X_n$; $T_{N_i} : i = 1, 2, \ldots, L_n$ are independent, but not identically distributed.

Define

$$G_n(x, t) \quad = \quad \frac{1}{L_n} \sum_{i=1}^{L_n} I(T_{N_i} \leq t).$$

We propose $G_n(x, t)$ as an estimator of $G(x, t)$. Using the techniques of regenerative processes, introduced by Athreya and Ney (1978), we will prove consistency of $G_n(x, t)$ in Section 2 and asymptotic normality in Section 3. We finish this section with the formal definition of Harris recurrence and establishing, without proofs, some results from the literature.

Definition 1.1 A Markov chain $\{X_n$:n=0,1,2,...$\}$ is called *Harris recurrent* if there exist a set $A \in \Sigma$, a probability measure ϕ on A, a real number $\epsilon > 0$, and an integer $n_0 > 0$ such that

$$P_x(\tau_A < \infty) = P_x(X_n \in A \text{ for some } n \geq 1) = 1 \ \forall x \in S, \tag{1.1}$$

$$P_x(X_{n_0} \in E) = P^{(n_0)}(x, E) \geq \epsilon\phi(E) \ \forall x \in A. \text{ and } \forall E \subset A. \tag{1.2}$$

In what follows, we assume $n_0 = 1$ and will say $(A, \epsilon, \phi, 1)$ recurrent instead of Harris recurrent.

Lemma 1.2 (Regeneration Lemma) *If* $\mathbf{X}$ *is* $(A, \epsilon, \phi, 1)$ *recurrent, then there exists a random time* N *such that* $P_x(N < \infty) = 1$ *and*

$$\begin{aligned} a(x, n, k) &\stackrel{\text{def}}{=} P_x(X_n \in A, X_{n+1} \in A_1, \ldots, X_{n+k} \in A_k, N = n) \\ &= P_x(N = n) \int_A P_y(X_1 \in A_1, \ldots, X_k \in A_k)\phi(dy). \end{aligned}$$

That is, the evolution of the process for $n \geq N$ *is independent of* X_1, X_2, ..., X_{N-1} *and* N *and has the same distribution as* $\mathbf{X}$ *where* X_0 *is distributed according to* ϕ. *Thus* N *is a random time such that the pre-N and post-N evolution are independent and the post-N process has a distribution independent of* $X_0, \ldots, X_{N-1}$.

Corollary 1.3 *If* $\mathbf{X}$ *is* $(A, \epsilon, \phi, 1)$ *recurrent, then there exists a sequence of random times* $\{N_i; i = 1, 2, \ldots\}$ *such that for any* x, *under* P_x, X_{N_i} *have distribution* ϕ *on* A, *and the random tours* $\{X_{N_i+j} : j = 0, 1, 2, \ldots, N_{i+1} - N_i - 1; N_{i+1} - N_i\}$ *are independent, identically distributed, and independent of* N_1.

The regeneration lemma can be used to show the existence of a stationary measure for Harris recurrent chains.

Theorem 1.4 *Let* N_1 *be the regeneration time as in Lemma 1.2. Define*

$$\nu(E) = E_\phi \Sigma_{i=0}^{N_1 - 1} I(X_i \in E). \tag{1.3}$$

Then ν *is a stationary measure for* $\mathbf{X}$, *and it is unique up to a multiplicative constant.*
Since $\nu(S) = E_\phi N_1$, ν *is finite if and only if* $E_\phi N_1 < \infty$.

Corollary 1.5 *A stationary probability distribution* $\pi(.)$ *for* $\mathbf{X}$ *exists if and only if* $E_\phi N_1 < \infty$, *and in this case,* $\pi(E) = \frac{\nu(E)}{\nu(S)}$.

Proofs are in Athreya and Ney (1978).

Another characterization of the stationary measure ν is given by the following theorem.

Theorem 1.6 *Let g be any bounded measurable function.*

(a) For a Harris recurrent chain, a measure ν satisfies (1.3) if and only if

$$\int_S g d\nu = E_\phi \Sigma_{i=0}^{N_1-1} g(X_i). \tag{1.4}$$

(b) Define $Tg(x) = E_x \Sigma_{j=1}^{N_1-1} g(X_j)$. Then

$$E_\phi(\Sigma_{j=0}^{N_1-1} g(X_j))^2 = \int_S g^2(x)\nu(dx) + 2\int_S g(x)(Tg)(x)\nu(dx) \tag{1.5}$$

Let Δ be a real number. For Theorems 1.7 and 1.8, define

$$\begin{aligned} T_\Delta^{(1)} &= inf\{k > 0 : X_k = \Delta\} \\ T_\Delta^{(2)} &= inf\{k > T_\Delta^{(1)} : X_k = \Delta\} \end{aligned}$$

Suppose $T_\Delta^{(i-1)}$ defined and define

$$T_\Delta^{(i)} = inf\{k > T_\Delta^{(i-1)} : X_k = \Delta\}$$

Theorem 1.7 *Let $\{X_n : n = 0, 1, \ldots\}$ be a real-valued Markov chain with stationary transition function. Assume that for any $x \in \Re$, $P_x(T_\Delta^{(1)} < \infty) = 1$ and $Var_\Delta(T_\Delta^{(1)}) < \infty$.*

Let $\pi(A) = \frac{E_\Delta \sum_{j=0}^{T_\Delta^{(1)}-1} I_A(X_j)}{E_\Delta T_\Delta^{(1)}}$ for any Borel set A in R. Assume that π is absolutely continuous w.r.t Lebesgue measure on $\Re \setminus \{\Delta\}$. Let f be the corresponding density. Let $\delta_n > 0$ and $p_n(x) = \frac{1}{2n\delta_n}\sum_{j=0}^{n} I_{A_n}(X_j)$ where $A_n = (x - \delta_n, x + \delta_n)$. If $\delta_n \to 0$ and $n\delta_n \to \infty$, then for almost all x and for every initial distribution,

$$p_n(x) \to f(x) \qquad \text{in probability.}$$

Proofs are in Atuncar (1994).

Theorem 1.8 *Let $\{X_n$:n=0,1,...$\}$ be a Harris chain with a recurrence point Δ. Let $\lambda = E_\Delta T_\Delta$ and $\sigma^2 = Var_\Delta(T_\Delta)$. Let K_n be the random number of visits to Δ by the chain during $\{0, 1, 2, \ldots, n\}$. That is $K_n =$*

$\sum_{j=0}^{n} I(X_j = \Delta)$. *Then the family* $\{n - T_\Delta^{(K_n)} : n = 1, 2, \ldots\}$ *is tight and*

$$\frac{K_n}{n} \to \frac{1}{\lambda} \text{ a.e. as } n \to \infty, \tag{1.6}$$

$$\sqrt{n}(\frac{K_n}{n} - \frac{1}{\lambda}) \xrightarrow{d} N(0, \frac{\sigma^2}{\lambda^3}). \tag{1.7}$$

Assertions (1.6) and (1.7) follow from the discrete version of Proposition 1.4, Chapter IV and Proposition 4.3, Chapter VI in Asmussen (1987). The first assertion is also derivable from the same material.

2 Consistency of $G_n(x,t)$

The main result of this section is Theorem 2.3 establishing weak consistency of $G_n(x,t)$.

Lemma 2.1 *Let* $\{k_n : n = 1, 2, \ldots\}$ *be a sequence of integers such that* $\frac{k_n}{n} \to \alpha$, $0 < \alpha < \infty$. *For* $i = 1, 2, \ldots, k_n$, *define*

$$\xi_{ni} = \sum_{j=T_\Delta^{(i)}}^{T_\Delta^{(i+1)}-1} \frac{1}{2\delta_n} |G(X_j, t) - G(x,t)| I(X_j \in A_n).$$

If $\lim_{\delta_n \to 0} \int_{|x-y|<\delta_n} \frac{1}{2\delta_n} |G(y,t) - G(x,t)| f(y) dy = 0$, *then*

$$\overline{\xi}_n \equiv \frac{1}{k_n} \sum_{i=1}^{k_n} \xi_{ni} \to 0 \text{ in probability.}$$

Proof: Since $\{\xi_{ni} : i = 1, 2, \ldots\}$ are i.i.d, see for example Asmussen (1987), and $E_\Delta \xi_{n1} \to 0$ by hypothesis, $E_\Delta \overline{\xi}_n \to 0$ and since $\overline{\xi}_n \geq 0$, the result follows. ■

Lemma 2.2 *Let* $\{K_n : n = 1, 2, \ldots\}$ *be a sequence of integer random variables such that* $\frac{K_n}{n} \to \alpha$ *with probability 1 for some* $0 < \alpha < \infty$. *Under the hypothesis of Lemma 2.1,*

$$\frac{1}{K_n} \sum_{i=1}^{K_n} \xi_{ni} \to 0 \quad \text{in probability.}$$

Proof: Let $\epsilon > 0$, $\theta > 0$ and let $A(n,\alpha,\epsilon) = \{n(\alpha-\epsilon) < K_n < n(\alpha+\epsilon)\}$

$$P_\Delta\left[|\frac{1}{K_n}\sum_{i=1}^{K_n}\xi_{ni}| > \theta\right] = P_\Delta\left[\frac{1}{K_n}\sum_{i=1}^{K_n}\xi_{ni} > \theta\right] + P_\Delta\left[\frac{1}{K_n}\sum_{i=1}^{K_n}\xi_{ni} < -\theta\right] = \alpha_{n1} + \alpha_{n2} \qquad \text{(say)}.$$

$$\alpha_{n1} = P_\Delta\left[\frac{1}{K_n}\sum_{i=1}^{K_n}\xi_{ni} > \theta, A(n,\alpha,\epsilon)\right] + P_\Delta\left[\frac{1}{K_n}\sum_{i=1}^{K_n}\xi_{ni} > \theta, A^c(n,\alpha,\epsilon)\right].$$

Since $\frac{K_n}{n} \to \alpha$ w.p. 1; the second term converges to zero. Now,

$$\begin{aligned} & P_\Delta\left[\frac{1}{K_n}\sum_{i=1}^{K_n}\xi_{ni} > \theta, A(n,\alpha,\epsilon)\right] \\ \leq\ & P_\Delta\left[\frac{1}{[n(\alpha-\epsilon)]}\sum_{i=1}^{[n(\alpha+\epsilon)]+1}\xi_{ni} > \theta, A(n,\alpha,\epsilon)\right] \\ =\ & P_\Delta\left[\frac{[n(\alpha+\epsilon)]+1}{[n(\alpha-\epsilon]}\frac{1}{[n(\alpha+\epsilon)]+1}\sum_{i=1}^{[n(\alpha+\epsilon)]+1}\xi_{ni} > \theta, A(n,\alpha,\epsilon)\right]. \end{aligned}$$

Since $\lim_{\epsilon\to 0}\lim_{n\to\infty}\frac{[n(\alpha+\epsilon)]+1}{[n(\alpha-\epsilon)]} = 1$ as $n \to \infty$, the last probability converges to zero by Lemma 2.1. Similarly it is proved that $\alpha_{n2} \to 0$. ■

Theorem 2.3 *Fix $x \in R \setminus \{\Delta\}$. Let $f(x) > 0$. Let $\delta_n \to 0$. If*

$$\lim_{\delta_n\to 0}\int_{|x-y|<\delta_n}\frac{1}{2\delta_n}|G(y,t)-G(x,t)|f(y)dy = 0,$$

$\frac{1}{2\delta_n}\int_{A_n} f(y)dy \to f(x)$, *and* $n\delta_n \to \infty$ *as* $n \to \infty$, *then*

$$G_n(x,t) \to G(x,t) \qquad \textit{in probability.}$$

Remark 2.4 The first condition means that the averages of $|G(y,t) - G(x,t)|$, weighted by f, are small on small intervals centered at x. x is a kind of Lebesgue point. In particular, if $G(.,t)$ is continuous at x this condition holds. If f is locally integrable in $(\Re, B(\Re))$ with respect to Lebesgue measure, the second condition holds by Lebesgue density theorem.

Proof: Recall that $N_1, N_2, \ldots, N_{L_n}$ are times of visits to A_n and

$$\begin{aligned} G_n(x,t) &= \frac{1}{L_n}\sum_{j=1}^{L_n} I(T_{N_j} \le t) \\ &= \frac{1}{L_n}\sum_{j=1}^{L_n}[I(T_{N_j} \le t) - G(X_{N_j}, t)] \\ &\quad + \frac{1}{L_n}\sum_{j=1}^{L_n}[G(X_{N_j}, t) - G(x,t)] + G(x,t) \\ &= \beta_{n1} + \beta_{n2} + G(x,t). \end{aligned}$$

It was proved in Atuncar (1994) that if $\frac{1}{2\delta_n}\int_{A_n} f(y)dy \to f(x)$ then $p_n(x) = \frac{L_n}{2n\delta_n} \to f(x)$ in probability. So, $L_n \to \infty$ in probability if $f(x) > 0$. Besides that, conditioned on $\{X_n : n = 0, 1, \ldots\}$, $(I(T_{N_j} \le t) - G(X_{N_j}, t))$ are independent random variables with mean zero. Then β_{n1} converges to zero in probability. Next we will prove that β_{n2} converges to zero in probability:

$$\begin{aligned} &\frac{1}{L_n}\sum_{j=1}^{L_n}[G(X_{N_j}, t) - G(x,t)] \le \frac{1}{L_n}\sum_{j=0}^{n}|G(X_j, t) - G(x,t)|I(X_j \in A_n) \\ &= \frac{1}{L_n}\sum_{j=0}^{T_\Delta^{(1)}-1}|G(X_j, t) - G(x,t)|I(X_j \in A_n) \\ &\quad + \frac{1}{L_n}\sum_{T_\Delta^{(1)}}^{T_\Delta^{(K_n)}-1}|G(X_j, t) - G(x,t)|I(X_j \in A_n) \\ &\quad + \frac{1}{L_n}\sum_{T_\Delta^{(K_n)}}^{n}|G(X_j, t) - G(x,t)|I(X_j \in A_n) \\ &= \beta_{n21} + \beta_{n22} + \beta_{n23} \qquad \text{(say)}. \end{aligned}$$

Since $L_n \to \infty$ in probability, β_{n21} converges to zero in probability. Also β_{n23} converges to zero in probability because $\frac{n - T_\Delta^{(K_n)}}{L_n}$ does by tightness of $(n - T_\Delta^{(K_n)})$.

Now,

$$\begin{aligned}\beta_{n22} &= \frac{K_n}{n}\frac{2n\delta_n}{L_n}\frac{1}{K_n}\sum_{i=1}^{K_n}\sum_{j=T_\Delta^{(i)}}^{T_\Delta^{(i+1)}-1}\frac{1}{2\delta_n}|G(X_j,t)-G(x,t)|I(X_j\in A_n)\\ &= \frac{K_n}{n}\frac{2n\delta_n}{L_n}\frac{1}{K_n}\sum_{i=1}^{K_n}\xi_{ni}.\end{aligned}$$

Therefore, by Theorem 1.8, Theorem 1.7, and Lemma 2.2, β_{n22} converges to zero in probability. ■

3 Asymptotic normality of $G_n(x,t)$

The main result of this section is Theorem 3.5.

Lemma 3.1 *For $i = 1, 2, \ldots$; define*

$$\Psi_{ni} = \frac{1}{2\delta_n}\sum_{j=T_\Delta^{(i)}}^{T_\Delta^{(i+1)}-1} G(X_j,t)(1-G(X_j,t))I(X_j\in A_n).$$

Let $\frac{1}{\delta_n}\int_{A_n} E_u T_\Delta f(u)du$ and $\frac{1}{\delta_n}\int_{A_n} f(u)du$ be bounded in n. Then

$$E_\Delta(\Psi_{n1}-E_\Delta\Psi_{n1})^2 = O(\frac{1}{\delta_n}).$$

Proof: Consider Ψ_{n_1} as defined above.

$$\begin{aligned}E_\Delta\Psi_{n1}^2 &= \frac{1}{4\delta_n^2}E_\Delta\sum_{j=0}^{T_\Delta^{(1)}-1} G^2(X_j,t)(1-G(X_j,t))^2 I(X_j\in A_n)\\ &\quad +\frac{1}{2\delta_n^2}E_\Delta\sum_{j=0}^{T_\Delta^{(1)}-1}\sum_{k=j+1}^{T_\Delta^{(1)}-1} G(X_j,t)(1-G(X_j,t))G(X_k,t)\\ &\qquad (1-G(X_k,t))I(X_j\in A_n)I(X_k\in A_n)\\ &= u_n+v_n \qquad \text{(say)}.\end{aligned}$$

$$\begin{aligned} u_n &= \frac{\lambda}{2\delta_n}\int_{A_n}\frac{1}{2\delta_n}G^2(u,t)(1-G(u,t))^2 f(u)du \\ &\leq \frac{\lambda}{64\delta_n}\int_{A_n}\frac{1}{\delta_n}f(u)du \\ &= O(\frac{1}{\delta_n}). \end{aligned}$$

Next,

$$\begin{aligned} v_n &= \frac{1}{\delta_n}E_\Delta\sum_{j=0}^{T_\Delta^{(1)}-1}G(X_j,t)(1-G(X_j,t))I(X_j\in A_n) \\ &\qquad E_{X_j}\sum_{K=1}^{T_\Delta^{(1)}-1}\frac{1}{2\delta_n}G(X_k,t)(1-G(X_k,t))I(X_k\in A_n) \\ &\leq \frac{\lambda}{\delta_n}\int_{A_n}\frac{1}{2\delta_n}G(u,t)(1-G(u,t))E_u(T_\Delta^{(1)}-1)f(u)du \\ &\leq \frac{\lambda}{8\delta_n}\int_{A_n}\frac{1}{\delta_n}E_uT_\Delta f(u)du \\ &= O(\frac{1}{\delta_n}). \end{aligned}$$

■

Lemma 3.2 *Assume the hypotheses of Theorem 2.3. Let* $\{k_n : n = 1, 2, \ldots\}$ *be a sequence of integers such that* $\frac{k_n}{n} \to \alpha$ *with* $0 < \alpha < \infty$. *Let* $\delta_n \to 0$ *and* $n\delta_n \to \infty$. *Then,*

$$\frac{1}{k_n}\sum_{i=1}^{k_n}\Psi_{ni} \quad\to\quad \lambda f(x)G(x,t)(1-G(x,t)) \qquad \textit{in probability.}$$

Proof:

$$\frac{1}{k_n}\sum_{i=1}^{k_n}\Psi_{ni} = \frac{1}{k_n}\sum_{i=1}^{k_n}(\Psi_{ni}-E_\Delta\Psi_{ni})+E_\Delta\Psi_{n1}.$$

By Theorem 1.6, $E_\Delta\Psi_{n1} = \lambda\int_{A_n}\frac{1}{2\delta_n}G(y,t)(1-G(y,t))f(y)dy$ and under the hypotheses of Theorem 2.3, this integral converges to $\lambda f(x)G(x,t)(1-$

$G(x,t))$. So, it is enough to prove that $\frac{1}{k_n}\sum_{i=1}^{k_n}(\Psi_{ni} - E_\Delta \Psi_{ni}) \to 0$ in probability:

$$\begin{aligned} P_\Delta(\frac{1}{k_n}\sum_{i=1}^{k_n}(\Psi - E_\Delta \Psi_{ni}) > \epsilon) &\leq \frac{1}{\epsilon^2 k_n} E_\Delta(\Psi_{n1} - E_\Delta \Psi_{n1})^2 \\ &= O(\frac{1}{k_n \delta_n}). \end{aligned}$$

Similarly, it is proved that $P_\Delta(\frac{1}{k_n}\sum_{i=1}^{k_n}(\Psi - E_\Delta \Psi_{ni}) < -\epsilon) = O(\frac{1}{k_n\delta_n})$. Since $k_n\delta_n \to \infty$, the proof is complete. ∎

Lemma 3.3 *Consider* $\{K_n : n = 1, 2, \ldots\}$ *a sequence of integer random variables such that* $\frac{K_n}{n} \to \alpha$ *with probability 1 for some* $0 < \alpha < \infty$ *and assume the hypotheses of Lemma 3.2. Then,*

$$\frac{1}{K_n}\sum_{i=1}^{K_n}\Psi_{ni} \quad\to\quad \lambda f(x)G(x,t)(1 - G(x,t)) \qquad \textit{in probability.}$$

The proof is similar to the proof of Lemma 2.2.

Lemma 3.4 *Under hypotheses of Lemma 3.1 and Lemma 3.2,*

$$\frac{1}{L_n}\sum_{j=0}^{L_n} G(X_{N_j}, t)(1 - G(X_{N_j}, t)) \to G(x,t)(1 - G(x,t))$$

in probability.

Proof: Notice that

$$\begin{aligned} &\frac{1}{L_n}\sum_{j=0}^{L_n} G(X_{N_j}, t)(1 - G(X_{N_j}, t)) \\ = &\frac{1}{L_n}\sum_{j=0}^{n} G(X_j, t)(1 - G(X_j, t))I(X_j \in A_n) \\ = &\frac{2n\delta_n}{L_n}\frac{K_n}{n}\frac{1}{K_n}\sum_{i=1}^{K_n}\frac{1}{2\delta_n}\sum_{j=T_\Delta^{(i)}}^{T_\Delta^{(1+1)}-1} G(X_j, t)(1 - G(X_j, t))I(X_j \in A_n) \\ &+\frac{1}{L_n}\sum_{j=0}^{T_\Delta^{(1)}-1} G(X_j, t)(1 - G(X_j, t))I(X_j \in A_n) \end{aligned}$$

$$
\begin{aligned}
&+\frac{1}{L_n}\sum_{j=T_\Delta^{(K_n)}}^{n} G(X_j,t)(1-G(X_j,t))I(X_j\in A_n)\\
&= \gamma_{n1}+\gamma_{n2}+\gamma_{n3.} \qquad .\text{(say)}
\end{aligned}
$$

By earlier arguments, γ_{n2} and γ_{n3} converge to zero in probability. Also, by Theorem 1.7, Theorem 1.8, and Lemma 3.3 it follows that γ_{n1} converges to $G(x,t)(1-G(x,t))$ in probability, concluding the proof. ■

Now we establish asymptotic normality of $G_n(x,t)$

Theorem 3.5 *Assume the following hypotheses:*

(i) $\frac{1}{\delta_n}\int_{A_n} f(y)dy$ *is bounded in* n,

(ii) $\frac{1}{2\delta_n}\int_{A_n} |G(y,t)-G(x,t)|f(y)dy \to 0$,

(iii) $\frac{1}{\delta_n}\int_{A_n} E_uT_\Delta f(u)du$ *bounded in* n,

(iv) $\delta_n\to 0$, $n\delta_n\to\infty$ *as* $n\to\infty$.
Assume also that:

(v) There exist $\alpha>0$ *and for each* t, *a* C_t *such that* $|G(y,t)-G(x,t)|\le C_t|x-y|^{2+\alpha}$ *for* y *near* x,

(vi) $n\delta_n^p\to 0$ *for some* $1<p\le 5+2\alpha$.

Let

$$
Z_n=\frac{\frac{1}{L_n}\sum_{j=1}^{L_n}(I(T_{X_{N_j}}\le t)-G(x,t))}{\sqrt{\frac{1}{L_n^2}\sum_{j=1}^{L_n}G(X_{N_j},t)(1-G(X_{N_j},t))}}
$$

Then, $Z_n\xrightarrow{d}N(0,1)$.

We write

$$
\begin{aligned}
Z_n &= \frac{\frac{1}{L_n}\sum_{j=1}^{L_n}(I(T_{X_{N_j}}\le t)-G(X_{N_j},t))}{\sqrt{\frac{1}{L_n^2}\sum_{j=1}^{L_n}G(X_{N_j},t)(1-G(X_{N_j},t))}}\\
&\quad+\frac{\frac{1}{L_n}\sum_{j=1}^{L_n}(G(X_{N_j},t)-G(x,t))}{\sqrt{\frac{1}{L_n^2}\sum_{j=1}^{L_n}G(X_{N_j},t)(1-G(X_{N_j},t))}}\\
&= Z_{n1}+Z_{n2} \qquad \textit{(say)}.
\end{aligned}
$$

To prove Theorem 3.5, first we will prove Lemma 3.7 and Lemma 3.8 below. For Lemma 3.7 we need the following result from Chung (1974), pp. 199.

Lemma 3.6 *Let* $\{\theta_{nj}, 1 \le j \le k_n, 1 \le n\}$ *be a double array of complex numbers satisfying the following conditions as* $n \to \infty$:

(i) $\max\{|\theta_{nj}| : 1 \le j \le k_n\} \to 0$*;*

(ii) $\sum_{j=1}^{k_n} |\theta_{nj}| \le M < \infty$*, where* M *does not depend on* n*;*

(iii) $\sum_{j=1}^{k_n} \theta_{nj} \to \theta$*, where* θ *is a (finite)complex number.*

Then we have

$$\prod_{j=1}^{k_n}(1+\theta_{nj}) \to \exp(\theta).$$

Lemma 3.7 *Let* $D = \sigma(X_0, X_1, \ldots)$ *be the* σ*-algebra generated by* $\{X_0, X_1, \ldots\}$*. Let* $\phi_n(\theta) = E_\Delta\left[\exp(i\theta Z_{n1}) \mid D\right]$*. Then for each* θ *in* $\Re$*,* $\phi_n(\theta) \to \exp(-\frac{1}{2}\theta^2)$ *in probability.*

Proof: Let $\delta_{nj} = I(T_{X_{N_j}} \le t)$, $p_{nj} = G(X_{N_j}, t)$, $w_n = \sqrt{\sum_{j=1}^{L_n} p_{nj}(1-p_{nj})}$.

For every θ in $\Re$,

$$\begin{aligned} \phi_n(\theta) &= E_\Delta\left[\exp(\frac{i\theta}{w_n}\sum_{j=1}^{L_n}(\delta_{nj}-p_{nj})) \mid D\right] \\ &= \prod_{j=1}^{L_n} E_\Delta\left[\exp(\frac{i\theta}{w_n}(\delta_{nj}-p_{nj})) \mid D\right] \end{aligned}$$

(by conditional independence of δ_{nj}). Let $a(\theta, n, j) = 1 + \theta_{nj}$ where $\theta_{nj} = E_\Delta\left[\exp(\frac{i\theta}{w_n}(\delta_{nj}-p_{nj})) - 1 \mid D\right]$.

Since $E_\Delta\left[(\delta_{nj}-p_{nj}) \mid D\right] = 0$,

$$\begin{aligned} \theta_{nj} &= E_\Delta\left[\{\exp(\frac{i\theta}{w_n}(\delta_{nj}-p_{nj})) - 1 - \frac{i\theta}{w_n}(\delta_{nj}-p_{nj})\} \mid D\right] \\ &= E_\Delta\left[\{\exp(\frac{i\theta}{w_n}(\delta_{nj}-p_{nj})) - 1 - \frac{i\theta}{w_n}(\delta_{nj}-p_{nj}) \right. \\ &\quad \left. -\frac{1}{2}(\frac{i\theta}{w_n})^2(\delta_{nj}-p_{nj})^2\} \mid D\right] - \frac{\theta^2}{2w_n^2}p_{nj}(1-p_{nj}). \end{aligned}$$

Since $|\exp(it) - 1 - it| \leq \frac{t^2}{2!}$ and $|exp(it) - 1 - it - \frac{(it)^2}{2!}| \leq \frac{|t|^3}{3!}$ for t real (See Feller (1971), pp 512), we have that

$$|\theta_{nj}| \leq \frac{\theta^2 p_{nj}(1-p_{nj})}{2w_n^2}.$$

Thus $\sum_{j=1}^{L_n} |\theta_{nj}| \leq \frac{\theta^2}{2}$ and $max\{|\theta_{nj}| : 1 \leq j \leq L_n\} \leq \frac{\theta^2}{8w_n^2}$.

From Lemma 3.4, $\frac{w_n^2}{L_n} \rightarrow G(x,t)(1 - G(x,t)$ in probability and since $L_n \rightarrow \infty$ in probability as well, $\max\{|\theta_{nj}| : 1 \leq j \leq L_n\} \rightarrow 0$ in probability under conditioning by D.

Also

$$\begin{aligned}\sum_{j=1}^{L_n}\theta_{nj} + \frac{\theta^2}{2} &= \sum_{j=1}^{L_n} E_\Delta\Big[\{\exp(\frac{i\theta}{w_n}(\delta_{nj}-p_{nj}) - 1 - \frac{i\theta}{w_n}(\delta_{nj}-p_{nj}) \\ &\quad -\frac{1}{2}(\frac{i\theta}{w_n})^2(\delta_{nj}-p_{nj})^2\} \mid D\Big].\end{aligned}$$

Then

$$\begin{aligned}|\sum_{j=1}^{L_n}\theta_{nj} + \frac{\theta^2}{2}| &\leq \frac{1}{3!}\sum_{j=1}^{L_n}(\frac{|\theta|}{w_n})^3 E_\Delta\Big[(\delta_{nj}-p_{nj})^3 \mid D\Big] \\ &\leq 2\frac{|\theta|^3}{3!w_n^3}\sum_{j=1}^{L_n} E_\Delta\Big[(\delta_{nj}-p_{nj})^2 \mid D\Big] \\ &\quad (\text{since } |\delta_{nj}-p_{nj}|^2 \leq 2) \\ &\leq \frac{1}{3}\frac{|\theta|^3}{w_n}.\end{aligned}$$

Since $w_n \rightarrow \infty$ in probability, $|\sum_{j=1}^{L_n}\theta_{nj} + \frac{\theta^2}{2}| \rightarrow 0$ in probability under conditioning by D.

For any subsequence n' there exists a further subsequence n'' such that along that, with probability one: $\max\{|\theta_{nj}| : 1 \leq j \leq L_n\} \rightarrow 0$, $\sum_{j=1}^{L_n}\theta_{nj} + \frac{\theta^2}{2} \rightarrow 0$, and $\sum_{j=1}^{L_n}|\theta_{nj}| \leq \frac{\theta^2}{2}$; and so by Lemma 3.6, $\phi_n(\theta) \rightarrow \exp(-\frac{\theta^2}{2})$ w.p. 1 along that subsequence n''. This being true for every subsequence n', the result follows. ■

Lemma 3.8 *Assume conditions (v) and (vi) in Theorem 3.5. Then* $Z_{n2} \rightarrow 0$ *in probability, where* Z_{n2}*, defined in Theorem 3.5, is*

$$Z_{n2} = \frac{\frac{1}{\sqrt{L_n}}\sum_{j=1}^{L_n}(G(X_{N_j},t) - G(x,t))}{\sqrt{\frac{1}{L_n}\sum_{j=1}^{L_n} G(X_{N_j},t)(1-G(X_{N_j},t))}}.$$

Proof: By Lemma 3.4, it is enough to prove that the numerator converges to zero in probability. To do this observe that

$$\begin{aligned}
&\frac{1}{\sqrt{L_n}}\sum_{j=1}^{L_n}(G(X_{N_j},t)-G(x,t))=\frac{1}{\sqrt{L_n}}\sum_{j=0}^{n}(G(X_j,t)-G(x,t))I(X_j\in A_n)\\
&=\frac{1}{\sqrt{L_n}}\sum_{j=0}^{T_\Delta^{(1)}-1}(G(X_j,t)-G(x,t))I(X_j\in A_n)\\
&\quad+\frac{1}{\sqrt{L_n}}\sum_{j=T_\Delta^{K_n}}^{n}(G(X_j,t)-G(x,t))I(X_j\in A_n)\\
&\quad+\frac{1}{\sqrt{L_n}}\sum_{j=T_\Delta^{(1)}}^{T_\Delta^{(K_n)}-1}(G(X_j,t)-G(x,t))I(X_j\in A_n)\\
&=b_{n1}+b_{n2}+b_{n3} \qquad \text{(say)}.
\end{aligned}$$

By earlier arguments, $b_{n1}+b_{n2}$ converges to zero in probability. Next,

$$\begin{aligned}
b_{n3} &= \frac{1}{\sqrt{L_n}}\sum_{j=T_\Delta^{(1)}}^{T_\Delta^{(K_n)}-1}(G(X_j,t)-G(x,t))I(X_j\in A_n)\\
&= \frac{\sqrt{2n\delta_n f(x)}}{\sqrt{L_n}}\frac{\sqrt{2n\delta_n f(x)}}{f(x)}\frac{K_n}{n}\frac{1}{K_n}\sum_{j=T_\Delta^{(1)}}^{T_\Delta^{(K_n)}-1}\frac{1}{2\delta_n}(G(X_j,t)\\
&\quad -G(x,t))I(X_j\in A_n).
\end{aligned}$$

By Theorem 1.7 and Theorem 1.8, it is enough to prove that

$$\sqrt{2n\delta_n}\frac{1}{K_n}\sum_{j=T_\Delta^{(1)}}^{T_\Delta^{(K_n)}-1}\frac{1}{2\delta_n}(G(X_j,t)-G(x,t))I(X_j\in A_n)\rightarrow 0$$

in probability.

Consider first K_n as non random. To go to the case in which K_n is random we follow the procedure used earlier. Let

$$\xi_{ni} = \sum_{j=T_\Delta^{(i)}}^{T_\Delta^{(i+1)}-1}\frac{1}{2\delta_n}(G(X_j,t)-G(x,t))I(X_j\in A_n).$$

Then,

$$
\begin{aligned}
& P_\Delta\left[\sqrt{2n\delta_n}|\frac{1}{K_n}\sum_{i=1}^{K_n}\xi_{ni}| > \epsilon\right] \\
& \leq \frac{2n\delta_n E_\Delta(\frac{1}{K_n}\sum_{i=1}^{K_n}\xi_{ni})^2}{\epsilon^2} \\
& = \frac{2n\delta_n}{\epsilon^2}\left[\frac{1}{K_n}E_\Delta\xi_{n1}^2 + \frac{K_n(K_n-1)}{2K_n^2}E_\Delta\xi_{n1}E_\Delta\xi_{n2}\right] \\
& = O(\delta_n E_\Delta\xi_{n1}^2) + O(n\delta_n(E_\Delta\xi_{n1})^2).
\end{aligned}
$$

where

$$
\begin{aligned}
E_\Delta\xi_{n1}^2 &= \frac{1}{4\delta_n^2}E_\Delta\sum_{j=0}^{T_\Delta^{(1)}-1}(G(X_j,t)-G(x,t))^2 I(X_j\in A_n) \\
& +\frac{1}{4\delta_n^2}E_\Delta\sum_{j=0}^{T_\Delta^{(1)}-1}\sum_{k=j+1}^{T_\Delta^{(1)}-1}(G(X_j,t)-G(x,t))(G(X_k,t)-G(x,t)) \\
& I(X_j\in A_n)I(X_k\in A_n) \\
&= b_{n31}+b_{n32} \qquad \text{(say)}.
\end{aligned}
$$

Notice that

$$
\begin{aligned}
b_{n31} &= \frac{1}{2\delta_n}\lambda\int_{|x-y|<\delta_n}\frac{1}{2\delta_n}(G(y,t)-G(x,t))^2 f(y)dy \\
&\leq \frac{\lambda C_t}{2\delta_n}\delta_n^{4+2\alpha}\int_{|x-y|<\delta_n}f(y)dy \\
&= O(\delta_n^{4+2\alpha}).
\end{aligned}
$$

By Theorem 1.6,

$$
\begin{aligned}
b_{n32} &= E_\Delta\sum_{j=0}^{T_\Delta^{(1)}-1}\frac{1}{2\delta_n}(G(X_j,t)-G(x,t))I(X_j\in A_n)E_{X_j} \\
& \sum_{k=1}^{T_\Delta^{(1)}-1}\frac{1}{2\delta_n}(G(X_k,t)-G(x,t))I(X_k\in A_n) \\
&= \lambda\int_{|x-y|<\delta_n}\frac{1}{2\delta_n}\left[(G(y,t)-G(x,t))E_y\sum_{k=1}^{T_\Delta^{(1)}-1}\frac{1}{2\delta_n}(G(X_k,t)\right.
\end{aligned}
$$

$$
\begin{aligned}
& \left. -G(x,t))I(X_k \in A_n)\right] f(y)dy \\
\leq \;& \frac{\lambda C_t}{2}\delta_n^{3+2\alpha}\int_{|x-y|<\delta_n}\frac{1}{2\delta_n}E_y(T_\Delta^{(1)}-1)f(y)dy \\
= \;& O(\delta_n^{3+2\alpha}).
\end{aligned}
$$

Hence, $E_\Delta \xi_{n1}^2 = O(\delta_n^{3+2\alpha})$. Also we have that

$$
\begin{aligned}
|E_\Delta \xi_{n1}| &= \lambda \int_{|x-y|<\delta_n}\frac{1}{2\delta_n}|G(y,t)-G(x,t)|f(y)dy \\
&\leq \frac{\lambda C_t}{2\delta_n}\int_{|x-y|<\delta_n}\delta_n^{2+\alpha}f(y)dy \\
&= O(\delta_n^{2+\alpha}) \qquad \text{(since } f \text{ is bounded)}.
\end{aligned}
$$

Therefore, $P_\Delta\left[\sqrt{2n\delta_n}|\frac{1}{K_n}\sum_{i=1}^{K_n}\xi_{ni}| > \epsilon\right] = O(n\delta_n^{5+2\alpha})$. ■

Proof of Theorem 3.5: For θ in R,

$$
\begin{aligned}
E(\exp(i\theta Z_{n1})) &= E(E_\Delta(\exp(i\theta Z_{n1} \mid D)) \\
&= E(\phi_n(\theta)).
\end{aligned}
$$

By Lemma 3.7 and the bounded convergence theorem, the last expectation converges to $exp(-\frac{\theta^2}{2})$. Thus $Z_{n1} \xrightarrow{d} N(0,1)$.

By Lemma 3.8, $Z_{n2} \to 0$ in probability. So by Slutsky's theorem, $Z_n = Z_{n1} + Z_{n2} \xrightarrow{d} N(0,1)$. ■

Remark 3.9 Notice that Z_n is pivotal for $G(x,t)$ since the limit law of Z_n is $N(0,1)$ and is independent of all parameters. Thus, Z_n could be used to obtain confidence intervals for $G(x,t)$.

Remark 3.10 In the hypotheses of Theorem 3.5 (i), (ii), and (iii) hold for almost all x by Lebesgue density theorem. $E_\Delta T_\Delta^2 < \infty$ is necessary for (iii).

4 Conclusions and further research

We have proved Consistency and Asymptotic Normality of $G_n(x,t)$, the naive estimator of the sojourn time distribution for Semi - Markov processes. As we know, the naive estimator for the density function in the i. i. d. case it can be improved defining a more general kernel estimator. We believe that the estimator defined in this paper can also be improved following the ideas from the i.i.d. case and we are working in that direction.

Acknowledgements

The authors are grateful to the referee for careful reading of the manuscript and helpful suggestions. We are also thankful to the Editor and the Executive Editor. Research supported by CAPES and FAPEMIG.

(Received March, 2002. Revised June, 2002.)

References

Asmussen, S. (1987). *Applied Probability and Queues.* New York: John Wiley and Sons Ltd.

Athreya. K. B. and Ney, P. (1978). A new approach to the limit theory of recurrent Markov chains. *Trans. Amer. Math. Soc.*, **245**, 493–501.

Atuncar, G.S. (1994). *Statistical Inference for real-valued Markov chains and some applications.* Ph. D. Dissertation. Iowa State University. Department of Statistics. Ames.

Chung, K. L. (1974). *A Course in Probability Theory.* 2.ed. New York: Academic Press, Inc.

Feller, W. (1971). *An Introduction to Probability Theory and its Applications.* Vol. II. New York: Wiley.