Bulletin of the Iranian Mathematical Society Vol. 34 *No. 1* (*2008*), *pp 23-35.*

CENTRAL AND LOCAL LIMIT THEOREMS IN MARKOV DEPENDENT RANDOM VARIABLES

A. RAHIMZADEH SANI

Communicated by Samad Hedayat

Abstract. We consider an irreducible and aperiodic Markov chain $\{k_n\}_{n=0}$ over the finite state space $E = \{1, \ \dots\ , \ p\}$ with positive regular transition matrix $P = \{p_{ij}\}$ and additive component $\{U_n\}$ such that $\{S_n\} = \{(k_n, U_n)\}$ is also a Markov chain over the state space $\mathcal{E}_1 = E \times \mathbb{R}$. We prove a central and a local limit theorem for this chain when the probability density functions of $\{S_n\}$, conditional on the first and the last states of $\{k_n\}_{n=0}$, exist.

1. Introduction

The purpose of this paper is to establish a central limit theorem and a local limit theorem for the sums of Markov dependent random variables when they are non-lattice and defined on the real line $\mathbb{R}$. The sequences of Markov dependent random variables have been frequently considered by different authors (see [5], [6], [7], [8], and [9]). Nagaev [7] considered sums of random variables defined on a lattice space, which are connected via a Markov chain, and proved a central limit theorem and a local limit theorem for these sums. Seva [9] considered sums of the form $S_n = \sum_{k=0}^{n} f(X_k)$, where $\{X_k\}$ is a Markov chain with a countable state space. He then gave the Nagaev like local limit theorem for lattice

MSC(2000): Primary: 60F05, 60J15

Keywords: Central limit theorem, Local limit theorem, Markov additive process, Maximum modulus eigenvalue

Received: 16 November 2006, Revised: 26 June 2007, Accepted: 16 October 2007

and non-lattice real valued functions. We consider a Markov chain $\{k_n\}$ on a finite state $E = \{1, \ldots, p\}$ with an additive component $\{U_n\}$ such that $\{S_n\} = \{(k_n,\ U_n)\}$ (as defined by (2.3)) is also a Markov chain on $\mathcal{E} = E \times \mathbb{R}$. Then we set the central limit theorem (Theorem 2.2)and the local limit theorem (Theorem 2.3) for S_n conditional on the first and the last positions of the Markov chain $\{k_n\}$. Theorem 2.3 can be interpreted as an extension of Nagaev theorem (Theorem 3.1, of [7]) from the lattice to the non-lattice random variables.

The original idea of this paper comes from our joint work [1] on multitype branching random walk, where the position of each individual in generation n on any sample path of the process is determined by a Markov dependent random walk. Similar to the classical proof of the central limit and the local limit theorems for $i.i.d.$ random variables, we use the Fourier inversion formula to change the discussion from densities to a study on characteristic functions. The new and the key point in this study is that the characteristic function of smoothed S_n asymptotically behaves as the $n-$th power of the spectral radius (or the maximum modulus eigenvalue given in Lemma 2.1) of the matrix of Laplace transforms with complex arguments. We set Lemma 3.1 to show that the characteristic functions of smoothed S_n conditional on the first and the last positions of $\{k_n\}$ have all the required properties to develop the theory. The proofs of Lemma 3.1 and the main results are contained in Section 3. Section 2 gives the notations, conventions and preliminary results being used through out the paper.

2. Notations and preliminary results

Similar to Jensen ([4], Chapter 9, Sec. 1), let $\{k_n\}_{n=0}$ be a homogeneous Markov chain over the finite space $E = \{1,\ \ldots\ ,\ p\}$ with a positive regular transition matrix $P = \{p_{ij}\}$ and an invariant probability distribution $\pi = (\pi_1,\ \ldots\ ,\ \pi_p)$. For any $i, j \in E$, let X_{ij} be a non-lattice random variable with distribution function $F_{ij}(x)$. We reserve the notations i and j to the members of E. Define $\nu(.) := \{\nu_{ij}(.)\} = \{p_{ij}F_{ij}(.)\}$. Then ν is an stochastic kernel on $\mathcal{E}_1 = E \times \mathbb{R}$. Therefore, for every fixed $i \in E$, $\sum_j p_{ij} \int_{\mathbb{R}} F_{ij}(dx) = 1$. Its n-fold convolution $\nu^{n*}(.) = \{\nu_{ij}^{n*}(.)\}$, by the semigroup property of convolution of measures, is

$$\nu_{ij}^{n*}(dx) \quad = \quad \sum_k \int_{\mathbb{R}} \nu_{ik}^{(n-1)*}(dy)\nu_{kj}(dx - y). \tag{2.1}$$

Then by induction it is easy to show that, for each $i,\ j$, and $n,\ \nu_{ij}^{n*}(\mathbb{R}) = p_{ij}^{n}$. We impose the following assumption:

$\mathcal{A}(1)$: *The matrix P is positive regular in the sense that for some positive integer $n,\ P^n$ has all positive entries.*

Although the assumption $\mathcal{A}(1)$ is equivalent to "irreduciblity and aperiodicity" in any finite Markov chain, we prefer to use $\mathcal{A}(1)$ instead, because it is more relevant with our work there. As a consequence of $\mathcal{A}(1)$, there is an n_0 such that for all $n \geq n_0,\ p_{ij}^n > 0$. Since we are dealing with large values of n, we will always assume that $n \geq n_0$, and hence we will have $p_{ij}^n > 0$, even if it is not emphasized.

Analogue to [1], for fixed $i,\ j$ and any complex $\lambda = \theta + \mathbf{i}\eta \in \mathbb{C}$, we define the complex valued Laplace transform related to $F_{ij}(.)$ by

$$\varphi_{ij}(\lambda) = \int_{\mathbb{R}} e^{-\lambda x} F_{ij}(dx), \quad (\lambda \in \mathbb{C})$$

wherever it exists. For real $\lambda = \theta \in \mathbb{R}$, $\varphi_{ij}(\theta)$ is a real valued continuous Laplace transform with $\varphi_{ij}(0) = 1$ and, for $\lambda = \mathbf{i}\eta$ on the imaginary axis, $\varphi_{ij}(\mathbf{i}\eta)$ becomes the characteristic function of F_{ij}. For each $\lambda = \theta + \mathbf{i}\eta$, we have $|\varphi_{ij}(\lambda)| \leq \varphi_{ij}(\theta)$. Thus each $\varphi_{ij}(\lambda)$ is a well-defined and analytic function on some open strip $L = (-h, h) \times \mathbb{R} \subset \mathbb{C}$ containing the imaginary axis (see [6], Introduction). Let $m_{ij}(\lambda) = p_{ij}\varphi_{ij}(\lambda)$ and $M(\lambda) = \{m_{ij}(\lambda)\}_{p\times p}$ with n-th power $M^n(\lambda) = \{m_{ij}^n(\lambda)\}$. Using induction, it is easy to show that (see [1]),

$$m_{ij}^n(\lambda) \ = \ \int_{\mathbb{R}} e^{-\lambda x} \nu_{ij}^{n*}(dx). \tag{2.2}$$

Let $\{S_n\} = \{(k_n,\ U_n)\}$ be a Markov chain over the state space $\mathcal{E} = E \times \mathbb{R}$ with transition densities $\nu(.) = \{\nu_{ij}(.)\}$, in the sense that, for any measurable set $K \times A \subset \mathcal{E}$,

$$P\,(\,S_n \in K \times A|\, S_0 = (i,\ u_0)) \ = \ \sum_{k\in K} \int_A \nu_{ik}^{n*}(dx - u_0). \tag{2.3}$$

We will always assume that $P(U_0 = 0) = 1$ and $k_0 = i$. To indicate the latter assumption, we will always put an index i in the probability and the expectation symbols. The n-th step distribution of $\{S_n\}$ is determined by $\nu_i^{n*}(.) = (\nu_{i1}^{n*}(.),\ \ldots\ ,\ \nu_{ip}^{n*}(.))$. Thus we will write (2.3) in the form

$$P_i\left(S_n \in K \times A\right) = \sum_{k\in K}\int_A \nu_{ik}^{n*}(dx).$$

The conditional probability of S_n, conditional on $k_n = j,\ (j \in K)$, is

$$\begin{aligned} P_i\left(\left. S_n \in K \times A \right| k_n = j\right) &= \frac{P_i\left(S_n \in \{j\} \times A,\ k_n = j\right)}{P_i\left(k_n = j\right)} \\ &= \frac{1}{p_{ij}^n}\int_A \nu_{ij}^{n*}(dx). \end{aligned}$$

The conditional Laplace transform of $\{S_n\}$, conditioned that the first and the last states of $\{k_n\}$ be i and j, respectively, is

$$\begin{aligned} \varphi_{ij}^{(n)}(\lambda) &= E_i\left[e^{-\lambda S_n}\middle| k_n = j\right] \\ &= \frac{1}{p_{ij}^n}\int_{\mathbb{R}} e^{-\lambda x}\nu_{ij}^{n*}(dx), \end{aligned} \tag{2.4}$$

which, by (2.1) and induction, is well-defined for all $\lambda \in L$. Comparing the equations (2.2) and (2.4), we get

$$m_{ij}^n(\lambda) = p_{ij}^n \varphi_{ij}^{(n)}(\lambda) \tag{2.5}$$

for all $i, j \in E, \lambda \in L$ and every integer n.

The entries of the matrix $M(\lambda)$ are complex-valued analytic functions in $\lambda \in L$. For those values of $\lambda = \theta \in L_0 = L \cap \mathbb{R}$, the entries are nonnegative and positive in a neighborhood of zero whenever $p_{ij} > 0$. Therefore, $M(\theta)$ has positive entries at least in the same positions as P does. This implies the positive regularity of $M(\theta)$ for all $\theta \in L_0$ and hence the conditions of Theorem 1 in [1] hold. We single out this fact here in the next lemma. Like the maximum eigenvalue in positive regular matrices, $\rho(\lambda)$ is *maximum modulus eigenvalue* of $M(\lambda)$ if it is a simple root of $\det(zI - M(\lambda)) = 0$ and for any other eigenvalue $\rho_1(\lambda)$ of $M(\lambda)$, $|\rho_1(\lambda)| < |\rho(\lambda)|$. Thus, $|\rho(\lambda)|$ is the spectral radius of $M(\lambda)$.

Lemma 2.1. *Let the matrix $M(\lambda) = \{m_{ij}(\lambda)\}$ of analytic functions defined on the open set $L \subset \mathbb{C}$ have the property that, for all $\theta \in L_0 = L \cap \mathbb{R}$, $\mathbb{M}(\theta)$ is positive regular. Then there is an open set $\Omega \subset L$ containing L_0 such that for any $\lambda \in \Omega, M(\lambda)$ has a simple maximum modulus eigenvalue $\rho(\lambda)$, with corresponding left and right eigenvectors $u(\lambda)$ and $v(\lambda)$ satisfying the following properties:*

(a) $\rho(\lambda),\ u(\lambda)$ *and* $v(\lambda)$ *are non-zero analytic functions in* $\lambda \in \Omega$; *and for the real argument,* $\rho(\lambda) > 0$ *and* $u,\ v$ *have positive components.*

(b) $u(\lambda)$ *and* $v(\lambda)$ *have nonzero analytic components and are normalized so that* $\sum_{i=1}^{p} u_i(\lambda)v_i(\lambda) = 1,$ *and* $\sum_{i=1}^{p} u_i(\lambda) = 1.$

(c) For any $\theta \in \Omega_0 = \Omega \cap \mathbb{R}$, *there is an open neighborhood* $B = B(\theta, \delta) \subset \Omega$ *and positive constants* $\gamma \in (0,\ 1)$ *and* K, *such that for all* $i,\ j$ *and* n,

$$\sup_{\lambda \in B} \left| \rho(\lambda)^{-n} m_{ij}^{n}(\lambda) - v_i(\lambda)u_j(\lambda) \right| \quad \leq \quad K\gamma^n.$$

□

Let $\theta \in \Omega_0$. The measure ν_{ij} is said to be degenerate if for some $\beta \in \mathbb{R}$, $m_{ij}(\theta) = e^{-\beta\theta}$. If for some diagonal matrix $D(\theta)$ of degenerate Laplace transforms, $M(\theta) = e^{-\beta\theta} D(\theta) M(0) D^{-1}(\theta)$, then $M(\theta)$ is said to be degenerate. When $M(\theta)$ is non-degenerate, $\rho(\theta)$ is a strictly log convex function in $\theta \in \Omega_0$ (see Section 5 in [5]), and $\mu = \rho'(\theta)$ is mean drift of the random walk S_n (see Section 2 in [6]). We now impose further assumptions:

$\mathcal{A}(2) : 0 \in \Omega_0$.

$\mathcal{A}(3)$: *The matrix* $M(\theta)$ *is non-degenerate in the sense that* $\rho(\theta)$ *is a strictly log-convex in* $\theta \in \Omega_0$.

For any fixed real number a and for any i and j, define $Y_{ij} = X_{ij} - a$. Then Y_{ij} is a random variable with distribution function $\bar{F}_{ij}(x) = F_{ij}(x+a)$. Define the stochastic kernel $\bar{\nu}(.) = \{\bar{\nu}_{ij}(.)\} = \{p_{ij}\bar{F}_{ij}(.)\}$ with its n-fold convolution,

$$\bar{\nu}_{ij}^{n*}(dx) \quad = \quad \sum_{k} \int_{\mathcal{R}} \bar{\nu}_{ik}^{(n-1)*}(dy)\bar{\nu}_{kj}(dx - y). \tag{2.6}$$

Similar to (2.3), let $\{\bar{S}_n\} = \{(k_n,\ \bar{U}_n)\}$ be a Markov chain over the state space $\mathcal{E} = E \times \mathbb{R}$ with transition densities $\bar{\nu}(.) = \{\bar{\nu}_{ij}(.)\}$. Given the initial distribution $P(\bar{S}_0 = (i, 0)) = P(k_0 = i,\ \bar{U}_0 = 0) = 1$, the distribution of the n-th step is determined by $\bar{\nu}_i^{n*}(.) = (\bar{\nu}_{i1}^{n*}(.),\ \ldots\ ,\ \bar{\nu}_{ip}^{n*}(.))$. The conditional complex valued Laplace transform of $\bar{S}^n$, conditional on the first and the last states of $\{k_n\}$, similar to (2.4), is

$$\bar{\varphi}_{ij}^{(n)}(\lambda) \quad = \quad E_i \left[e^{-\lambda \bar{S}_n} \middle| \, k_n = j \right].$$

Analogous to $M(\lambda)$, define $\bar{M}(\lambda) = \{\bar{m}_{ij}(\lambda)\} := \{p_{ij}\bar{\varphi}_{ij}^{(1)}(\lambda)\}$ with

$$\bar{M}^n(\lambda) \quad = \quad \{\bar{m}_{ij}^{n}(\lambda)\} = \{p_{ij}^{n}\bar{\varphi}_{ij}^{(n)}(\lambda)\}. \tag{2.7}$$

Then $\bar{M}(\lambda) = e^{a\lambda}M(\lambda)$ and both of these matrices are defined on the same set $L \subset \mathbb{C}$, and $\bar{M}(\lambda)$ satisfies all conditions of Lemma 2.1. We denote its simple maximum-modulus eigenvalue by $\bar{\rho}(\lambda)$ and the related left and right eigenvectors by $\bar{u}(\lambda)$ and $\bar{v}(\lambda)$, respectively. The equation $\bar{M}(\lambda) = e^{a\lambda}M(\lambda)$ implies that $\bar{\rho}(\lambda) = e^{a\lambda}\rho(\lambda)$, $\bar{u}(\lambda) = u(\lambda)$, and $\bar{v}(\lambda) = v(\lambda)$. We will prove these properties in proposition 2.7 at the end of this section. By this Proposition *we will be able to use* $u(\lambda)$ $(v(\lambda))$ *for the left (right) eigenvector of both matrices* $M(\lambda)$ *and* $\bar{M}(\lambda)$. We also assume that $\bar{\rho}'(\theta)|_{\theta=0} = 0$ or $a = -\rho'(0)/\rho(0)$ $(= -\rho'(0)$ by $\rho(0) = 1)$, and we let $\sigma^2 = (\bar{\rho}(\theta))''|_{\theta=0} > 0$. With this assumption on a, $\bar{M}(0) = P$, $\bar{\rho}(0) = 1$ and $v(0) = (1, \dots, 1)$. Note that, when $\theta \in \Omega_0$, $\bar{M}(\theta)$ is positive regular with equilibrium probabilities $\pi(\theta) = (\pi_1(\theta), \dots, \pi_p(\theta))$ with $\pi_i(\theta) = u_i(\theta)v_i(\theta)$ (see [6]). Our last assumption is:

$\mathcal{A}(4)$: *For any i and j, the characteristic function $\bar{\varphi}_{ij}^{(1)}(\mathbf{i}\eta)$ associated with $\bar{F}_{ij}(.)$ is absolutely integrable.*

Now we give The Central Limit Theorem 2.2 and The Local Limit Theorem 2.3 for $\bar{S}_n$ conditional on $k_0 = i$ and $k_n = j$, or the probability measure $(1/p_{ij}^n)\bar{\nu}_{ij}^{n*}(.)$. Between them the local limit Theorem 2.3 is new and as mentioned before, it can be interpreted as an extension of Negayev's local limit theorem to the non-lattice random variables. The central limit Theorems 2.2, 2.4 and 2.5 are well known (e.g., see Section 5 in [5]), but here we give a new proof based on the maximum modulus eigenvalue in Perron-Frobenius theory.

Theorem 2.2. *Suppose $\mathcal{A}(1)-\mathcal{A}(4)$ hold, $a = -\rho'(0)$, and $\sigma^2 = \bar{\rho}''(0) > 0$. Then for any fixed i, j, and $t \in \mathcal{R}$,*

$$\lim_n P\left(\bar{S}_n \le t/\sqrt{n} \,\middle|\, k_0 = i,\ k_n = j\right) \quad = \quad \int_{-\infty}^{t} \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{x^2}{2\sigma^2}} dx.$$

By $\mathcal{A}(4)$, for each i, j and n, the density function of $(1/p_{ij}^n)\bar{\nu}_{ij}^{n*}$ exists which will be denoted by g_{ij}^n.

Theorem 2.3. *Suppose $\mathcal{A}(1)-\mathcal{A}(4)$ hold, $a = -\rho'(0)$, and $\sigma^2 = \rho''(0) > 0$. Then,*

$$\lim_n \left\{ \sup_{x \in \mathbb{R}} \left| \sqrt{n} g_{ij}^n(x\sqrt{n}) - \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{x^2}{2\sigma^2}} \right| \right\} \quad = \quad 0.$$

The next theorem, which is an alternative version of the central limit Theorem 2.2, is a direct consequence of the integration of the uniform convergence sequences in Theorem 2.3.

Theorem 2.4. *Let the hypothesis of Theorem 2.2 hold and $b > 0$. Then for any bounded Borel measurable set $A \subset \mathbb{R}$, and any fixed $i,\ j \in E$,*

$$\lim_{n\to\infty} \left\{ \sup_{\substack{|A|\le b \\ c\in\mathbb{R}}} \left| \sqrt{n}\bar{\nu}_{ij}^{n*}(c+A) - \int_{c+A} \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{x^2}{2\sigma^2}} dx \right| \right\} \quad = \quad 0,$$

where $|A|$ is the Lebesgue measure of A. □

By the relation between the measures $\nu(.)$ and $\bar{\nu}(.)$, the analogous central limit theorem for the measure $\nu_{ij}^{n*}(.)$ may now be stated.

Theorem 2.5. *With the hypothesis of Theorem 2.4, for any bounded Borel measurable set $A \subset \mathbb{R}$, and any fixed $i,\ j \in E$,*

$$\lim_{n\to\infty} \left\{ \sup_{\substack{|A|\le b \\ c\in\mathbb{R}}} \left| \sqrt{n}\nu_{ij}^{n*}(na+c+A) - \int_{c+A} \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{x^2}{2\sigma^2}}(x) dx \right| \right\} \quad = \quad 0. \quad \square$$

Remark 2.6. In all of the theorems given above, conditioning on i and j means that the initial distribution of $\{k_n\}$ is concentrated at the state i and the n-th step distribution of $\{k_n\}$ is concentrated at the state j. As we have seen, the limiting results of these theorems are independent from the positions of the first and the last states of $\{k_n\}$ in our conditions. The author believes that, if instead of conditions on the first and the last states of $\{k_n\}$ one imposes a condition on any sample path of the process, the limiting results remain independent from that sample path of the process.

We conclude this section by the next proposition, which is a part of the proof of Theorem 1 in [6].

Proposition 2.7. *Let $\mathcal{A}(1) - \mathcal{A}(3)$ hold. Then we have:*
(i) $\bar{\rho}(\lambda) = e^{a\lambda}\rho(\lambda),\ \bar{u}(\lambda) = u(\lambda),$ *and* $\bar{v}(\lambda) = v(\lambda)$ *for all* $\lambda \in \Omega$.
(ii) $\bar{\rho}(\theta)$ *is strictly log-convex in* $\theta \in \Omega_0$, *and* $\bar{\rho}(\theta)'' > 0$.
(iii) If $a = -\rho'(\theta)/\rho(\theta)$ *then* $\bar{\rho}'(\theta) = 0$.

Proof. To prove (i), let $\lambda \in \Omega$. Then $\rho(\lambda)$ is a simple root of the characteristic equation $\det(zI - M(\lambda)) = 0$ with maximum modulus. Multiplying both sides of the characteristic equation by $e^{a\lambda}$ we get

$\det(e^{a\lambda}\rho(\lambda)I - \bar{M}(\lambda)) = 0$ implying that $\bar{\rho}(\lambda) = e^{a\lambda}\rho(\lambda)$ is a simple eigenvalue of $\bar{M}(\lambda)$. For any other eigenvalue γ of $\bar{M}(\lambda)$, from $\det(\gamma I - \bar{M}(\lambda)) = 0$ we get $\det(e^{-a\lambda}\gamma I - M(\lambda)) = 0$ and hence, $\gamma e^{-a\lambda}$ is an eigenvalue of $M(\lambda)$. This implies that $|e^{-a\lambda}\gamma| < \rho(\lambda)$ or $|\gamma| < |\bar{\rho}(\lambda|$. Thus, $\bar{\rho}(\lambda)$ has the maximum modulus. The equations $M(\lambda)v(\lambda) = \rho(\lambda)v(\lambda)$ and $e^{a\lambda}M(\lambda)v(\lambda) = e^{a\lambda}\rho(\lambda)v(\lambda)$ are equivalent, and so $v(\lambda)$ is the right eigenvector of $\bar{M}(\lambda)$. For the related left eigenvector of $\bar{\rho}(\lambda)$ we argue similarly.

To prove (ii), let $\theta \in \Omega_0$. Then $\log\bar{\rho}(\theta) = a\theta + \log\rho(\theta)$, implying that both $\log\bar{\rho}(\theta)$ and $\log\rho(\theta)$ have the same second derivatives. Then by $\mathcal{A}(3)$, both $\rho(\theta)$ and $\bar{\rho}(\theta)$ are strictly log-convex in $\theta \in \Omega_0$.

The proof for (iii) is obvious. □

3. **Proofs of Theorems 2.2 and 2.3**

This section contains the proofs for the central limit Theorem 2.2 and the local limit Theorem 2.3. The proofs are direct extensions of the well known central and local limit theorems for sums of *i.i.d.* random variables and is based on the properties of the characteristic functions (e.g., see the central and the local limit theorems in [3], Chapter 5, Sec. 10). The extended maximum modulus Perron-Frobenius theory, Lemma 2.1, leads to Lemma 3.1 on some properties of the characteristic functions $\bar{\varphi}_{ij}^{(n)}(\mathbf{i}\eta)$. This lemma is new and fundamental for our approach. Similar results for the real maximum Perron-Frobenius eigenvalue are previously considered (e.g., see Theorem 5.3 in [8]).

Lemma 3.1. *Let $\mathcal{A}(1) - \mathcal{A}(4)$ hold, $b > 0$, $a = -\rho'(0)$, and $\bar{\rho}''(0) = \sigma^2$. Then,*

(a) For any $i,\ j$, and $\eta \in \mathbb{R}$,

$$\lim_{n\to\infty} \bar{\varphi}_{ij}^{(n)}\left(\frac{\mathbf{i}\eta}{\sqrt{n}}\right) \quad = \quad e^{-\frac{\sigma^2\eta^2}{2}}.$$

(b) For any $i,\ j$ and $\eta \in \mathbb{R}$, let $\psi(\eta) = \max_{1\le i,j\le p}|\bar{\varphi}_{ij}(\mathbf{i}\eta)|$. Then for any n and $\eta \in \mathbb{R}$,

$$\left|\bar{\varphi}_{ij}^{(n)}(\mathbf{i}\eta)\right| \quad \le \quad \psi^n(\eta).$$

Moreover, all $\bar{\varphi}_{ij}^{(n)}(\mathbf{i}\eta)$ are absolutely integrable and their related bounded density functions g_{ij}^n exist.

(c) There are constants N_1 and $M > 0$ such that for all $n \geq N_1$ and $|\eta| \leq b$,

$$\left|\bar{\varphi}_{ij}^{(n)}\left(\frac{\mathbf{i}\eta}{\sqrt{n}}\right)\right| \quad \leq \quad Me^{-\frac{\sigma^2\eta^2}{4}}.$$

Proof. For each i, j, and $\eta \in \mathbb{R}$, from (2.7) we have $\bar{m}_{ij}^n(\mathbf{i}\eta/\sqrt{n}) = p_{ij}^n \bar{\varphi}_{ij}^{(n)}(\mathbf{i}\eta/\sqrt{n})$. Then the characteristic function $\bar{\varphi}_{ij}^{(n)}(\frac{\mathbf{i}\eta}{\sqrt{n}})$ can be written as:

$$\bar{\varphi}_{ij}^{(n)}\left(\frac{\mathbf{i}\eta}{\sqrt{n}}\right) \quad = \quad \left(\frac{1}{p_{ij}^n}\frac{\bar{m}_{ij}^n(\frac{\mathbf{i}\eta}{\sqrt{n}})}{(\bar{\rho}(\frac{\mathbf{i}\eta}{\sqrt{n}}))^n}\right)\left(\bar{\rho}\left(\frac{\mathbf{i}\eta}{\sqrt{n}}\right)\right)^n. \tag{3.1}$$

Since for each j, $v_j(0) = 1$, by Lemma 2.1, as $n \to \infty$, $\bar{m}_{ij}^n(\mathbf{i}\eta/\sqrt{n})\bar{\rho}(\mathbf{i}\eta/\sqrt{n})^{-n}$ converges to $u_j(0)$ for all $\eta \in \mathbb{R}$. Also as $n \to \infty$, p_{ij}^n converges to its equilibrium probability $u_j(0)$. Now we show that $\bar{\rho}(\mathbf{i}\eta/\sqrt{n})^n$ converges to the as required limit. In fact, $\bar{\rho}(\mathbf{i}\eta)$ has the same regularity conditions required for applying the well known methods in proving a classic central limit theorem (e.g., see the central limit theorem in [3], Chapter 5, Section 10). Observe that $\bar{\rho}(0) = 1$, $\bar{\rho}'(0) = 0$ and $\bar{\rho}''(0) = \sigma^2 > 0$. By the Taylor expansion of $\bar{\rho}(\lambda)$ in a neighborhood of zero (with a similar argument employed in the above mentioned theorem in [3]), for any $|\eta| \leq b$ as $n \to \infty$, we can write,

$$\bar{\rho}\left(\frac{\mathbf{i}\eta}{\sqrt{n}}\right) \quad = \quad \left(1 - \frac{\eta^2\sigma 2}{2n}\right) + O\left(\frac{1}{n\sqrt{n}}\right).$$

Then, for any $|\eta| \leq b$ as $n \to \infty$, we get,

$$\begin{aligned}\bar{\rho}\left(\frac{\mathbf{i}\eta}{\sqrt{n}}\right)^n \quad &= \quad \left(1 - \frac{\sigma^2\eta^2}{2n} + O(\frac{1}{n\sqrt{n}})\right)^n \\ &= \quad 1 - \frac{\sigma^2\eta^2}{2} + O(\frac{1}{\sqrt{n}}) \\ &= \quad e^{-\frac{\eta^2\sigma^2}{2}} + O(\frac{1}{\sqrt{n}}).\end{aligned}$$

This completes the proof of part (a).

To prove part (b), define $\psi(\eta) = \max_{i,j} |\bar{\varphi}_{ij}^{(1)}(\mathbf{i}\eta)|$ for $\eta \in \mathbb{R}$. Then, by $\mathcal{A}(4)$, ψ is absolutely integrable and $0 \leq \psi(\eta) < 1$ for all $|\eta| \neq 0$. We now proceed the proof by induction on n. Let i and j be fixed. Then, $|\bar{\varphi}^{(1)}_{ij}(\eta)| \leq \psi(\eta)$. For each n, by considering (2.7), using the convolution of measures in (2.6) and then anticipating a change of variable $x \to x+y$, it follows:

$$\begin{aligned}
\left|p_{ij}^{n+1}\bar{\varphi}_{ij}^{(n+1)}(\mathbf{i}\eta)\right| &= \left|\int_{\mathbb{R}} e^{-\mathbf{i}\eta x}\bar{\nu}_{ij}^{(n+1)*}(dx)\right| \\
&= \left|\sum_k \int_{\mathbb{R},x} e^{-\mathbf{i}\eta x}\int_{\mathbb{R},y}\bar{\nu}_{ik}^{n*}(dy)\bar{\nu}_{kj}(dx-y)\right| \\
&= \left|\sum_k p_{ik}^n p_{kj}\varphi_{ik}^{(n)}(\eta)\varphi_{kj}^{(1)}(\mathbf{i}\eta)\right| \\
&\leq p_{ij}^{n+1}\left(\psi(\mathbf{i}\eta)\right)^{n+1}.
\end{aligned}$$

Then, by induction we have $|\bar{\varphi}_{ij}^{(n)}(\mathbf{i}\eta)| \leq (\psi(\eta))^n$ for all $\eta \in \mathbb{R}$, i, j and n. Since ψ is integrable and $(\psi(\eta))^n \leq \psi(\eta)$, we get the absolute integrability of all $\bar{\varphi}_{ij}^{(n)}$ and hence the existence of their bounded density functions g_{ij}^n.

To prove part (c), as in the proof of part (a), we consider equation (3.1). The convergence of $\{\bar{m}_{ij}^n(\mathbf{i}\eta)\bar{\rho}(\mathbf{i}\eta)^{(-n)}\}$, by part (c) of Lemma 2.1, is uniform in a neighborhood of zero. So it is bounded by some M_1 for all $|\eta| \leq \delta_1$, for some δ_1. The convergent sequence $\{1/p_{ij}^n\}$ is also bounded by some M_2. If we let n be large enough, then for all $|\eta| \leq b$, we get $|\eta/\sqrt{n}| \leq \delta_1$, and hence,

$$\left|\bar{\varphi}_{ij}^{(n)}(\frac{\mathbf{i}\eta}{\sqrt{n}})\right| \leq M_1M_2\left|\left(\bar{\rho}\left(\frac{\mathbf{i}\eta}{\sqrt{n}}\right)\right)^n\right|.$$

Now it is enough to bound $\left(\bar{\rho}\left(\frac{\mathbf{i}\eta}{\sqrt{n}}\right)\right)^n$ by $e^{-\eta^2\sigma^2/4}$. Similar to part (a), the Taylor expansion of $\bar{\rho}(\lambda)$ in a neighborhood of zero, as $n \to \infty$, gives

$$\bar{\rho}\left(\frac{\mathbf{i}\eta}{\sqrt{n}}\right) = 1 - \frac{\sigma^2\eta^2}{2n} + O(\frac{1}{n\sqrt{n}}),$$

for all $|\eta| \leq b$. Let n be large enough so that $0 \leq 1 - b^2\sigma^2/2n < 1$ and $O(1/\sqrt{n}) \leq \sigma^2/4$. Then, we can write,

$$\left|\bar{\rho}\left(\frac{\mathbf{i}\eta}{\sqrt{n}}\right)\right| \leq 1 - \frac{\sigma^2\eta^2}{4n} - \frac{\eta^2}{n}\left(\frac{\sigma^2}{4} - O\left(\frac{1}{\sqrt{n}}\right)\right) \leq 1 - \frac{\sigma^2\eta^2}{4n}.$$

Let N_1 be such that for all $n \geq N_1$, the above inequalities hold. Since the sequence $\{(1-(\sigma^2\eta^2)/4n)^n\}$ is increasing and converges to $e^{-\eta^2\sigma^2/4}$, then for all $|\eta| \leq b$ and $n \geq N_1$, we get,

$$\left|\bar{\rho}\left(\frac{\mathbf{i}\eta}{\sqrt{n}}\right)^n\right| \leq e^{-\frac{\sigma^2\eta^2}{4}}.$$

This implies that for $M = M_1M_2$ and all $|\eta| \leq b,\ n \geq N_1$, and any $i,\ j$, we have $\left|\bar{\varphi}_{ij}^{(n)}\left(\frac{\mathbf{i}\eta}{\sqrt{n}}\right)\right| \leq Me^{-(\sigma^2\eta^2/4)}$, which completes the proof. ☐

Proof of Theorem 2.2 . From part (a) of Lemma 3.1, as $n \to \infty$, the characteristic functions $\bar{\varphi}_{ij}^{(n)}(\frac{\mathbf{i}\eta}{\sqrt{n}})$ converge to the characteristic function of normal distribution $e^{-\eta^2\sigma^2/2}$. Now, by the continuity theorem ([2], Chapter XV, Sec. 3), the proof is complete. ☐

Proof of Theorem 2.3 . Let g_{ij}^n be the density function of the probability measure $(1/p_{ij}^n)\bar{\nu}_{ij}^{n*}(.)$. Then its characteristic function $\bar{\varphi}_{ij}^{(n)}(\mathbf{i}\eta)$, by Fourier inversion formula ([2], Chapter XV, Sec. 3, Theorem 3), satisfies the equation,

$$g_{ij}^n(x\sqrt{n}) = \frac{1}{2\pi\sqrt{n}}\int_{\mathbb{R}} e^{\mathbf{i}\eta x}\bar{\varphi}_{ij}^{(n)}(\frac{\mathbf{i}\eta}{\sqrt{n}})d\eta. \tag{3.2}$$

The density of the normal distribution $N(0,\sigma^2)$ also satisfies the equation,

$$f_\sigma(x) := \frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{x^2}{2\sigma^2}} = \frac{1}{2\pi}\int_{\mathcal{R}} e^{\mathbf{i}\eta x}e^{-\frac{\eta^2\sigma^2}{2}}d\eta. \tag{3.3}$$

For any i, j, n, and any real x, by considering (3.2) and (3.3), we can write,

$$\begin{aligned}\left|\sqrt{n}g_{ij}^n(x\sqrt{n}) - f_\sigma(x)\right| &\leq \frac{1}{2\pi}\int_{\mathcal{R}}\left|\bar{\varphi}_{ij}^{(n)}(\frac{\mathbf{i}\eta}{\sqrt{n}}) - e^{-\frac{\eta^2\sigma^2}{2}}\right|d\eta \\ &= \frac{1}{2\pi}\int_{\mathcal{R}}\left|L_{ij}^n(\eta)\right|d\eta \quad (say).\end{aligned} \tag{3.4}$$

We show that, the last integral in (3.4) tends to zero, as $n \to \infty$, uniformly in x. Let $\delta > 0$ be as in Lemma 2.1, part (c), and $0 < b < \delta\sqrt{n}$. By dividing the integral in (3.4) into three parts as

$$\begin{aligned}\int_{\mathcal{R}}\left|L_{ij}^n(\eta)\right|d\eta &= \int_{|\eta|\leq b}\left|L_{ij}^n(\eta)\right|d\eta + \int_{b\leq|\eta|\leq\delta\sqrt{n}}\left|L_{ij}^n(\eta)\right|d\eta \\ &+ \int_{|\eta|\geq\delta\sqrt{n}}\left|L_{ij}^n(\eta)\right|d\eta,\end{aligned}$$

we show that, for suitable values of b, each part tends to zero, as $n \to \infty$.

Part one. $\int_{|\eta|\geq\delta\sqrt{n}} \left|L_{ij}^n(\eta)\right| d\eta$. Let $\beta_\delta = \sup_{|\eta|\geq\delta}\psi(\eta)$, Then, as in the proof of part (b) in Lemma 3.1 $0 < \beta_\delta < 1$. The function $\psi(\eta)$ is integrable. Let $\int_{\mathcal{R}}\psi(\eta)d\eta \leq M_1$. From the integrability of the function $e^{-(\eta^2\sigma^2/2)}$, as $n\to\infty$, $\delta_n = \int_{\delta\sqrt{n}}^\infty e^{-(\eta^2\sigma^2/2)}d\eta$ converges to zero, and so by part (b) of Lemma 3.1 we can write,

$$\begin{aligned}\int_{|\eta|\geq\delta\sqrt{n}} \left|L_{ij}^n(\eta)\right| d\eta &\leq \sqrt{n}\int_{|\eta|\geq\delta}\psi(\eta)^n d\eta + 2\delta_n \\ &\leq \sqrt{n}\beta_\delta^{n-1}\int_{|\eta|\geq\delta}\psi(\eta)d\eta + 2\delta_n \\ &\leq 2\sqrt{n}\beta_\delta^{n-1}M_1 + 2\delta_n.\end{aligned}$$

Since $0 \leq \beta_\delta < 1$, we can choose N_1 large enough to make $\int_{|\eta|\geq\delta\sqrt{n}} \left|L_{ij}^n(\eta)\right| d\eta$ small enough for all $n\geq N_1$.

Part two. $\int_{b\leq|\eta|\leq\delta\sqrt{n}} \left|L_{ij}^n(\eta)\right| d\eta$. Let $\delta > 0$ and N_1 be as in part one. Also, let $n\geq N_1$ and $b\leq\delta\sqrt{n}$. Then by part (c) of Lemma 3.1, we can write,

$$\int_{b\leq|\eta|\leq\delta\sqrt{n}} \left|L_{ij}^n(\eta)\right| d\eta \leq 2M\int_{b\leq\eta} e^{-\frac{\sigma^2\eta^2}{4}}d\eta + 2\int_{b\leq\eta} e^{-\frac{\sigma^2\eta^2}{2}}d\eta. \tag{3.5}$$

The integrals on the right hand side of the last inequality converge. Thus, by choosing N_1 and b large enough, we can make the right hand side of (3.5) small.

Part three. $\int_{|\eta|\leq b} \left|L_{ij}^n(\eta)\right| d\eta$. Let δ, b, and N_2 be such that the conditions of part two hold for all $n\geq N_2$. From part (a) of Lemma 3.1, the sequence of characteristic functions $\{\bar{\varphi}_{ij}^{(n)}(\mathbf{i}\eta)\}$ converges point-wise in $\mathbb{R}$. Since any point-wise convergent sequence of characteristic functions also converges uniformly in any neighborhood of zero (see Theorem 2 in [2], Chapter XV, Sec. 3). So $\lim_{n\to\infty}\left|L_{ij}^n(\eta)\right| = 0$ uniformly in $|\eta|\leq b$. The dominated convergence theorem implies that

$$\lim_{n\to\infty}\int_{|\eta|\leq b} \left|L_{ij}^n(\eta)\right| d\eta = 0.$$

Thus the right hand side of (3.4) converge to 0, as $n\to\infty$, and this completes the proof. □

Acknowledgment

The author thanks Professor J. D. Biggins for many valuable discussions. He also thanks the anonymous referee for his/her useful comments that helped to improve the presentation. This work has been partially supported by a grant No. S84-126 from Mosahab Institute of Mathematics (MIM).

References

[1] J. D. Biggins and A. Rahimzadeh Sani, Convergence results on multitype, multivariate branching random walks, *Adv. Appl. Prob.* **37** (2005), 681-705.

[2] W. Feller, *An Introduction to Probability Theory and Its Applications, Vol. II,* New York: Wiley, 1971.

[3] G. R. Grimmett and D. R. Stirzaker, *The Probability and Random Processes, Second Edition,* New York: Oxford University Press, 1993.

[4] J. L. Jensen, *Saddlepoint Approximations,* Oxford University Press, 1995.

[5] I. Kontoyiannis and S. P. Meyn, Spectral theory and limit theorems for geometrically ergodic Markov processes, *Ann. Appl. Prob.* **13** (2003), 304-362.

[6] H. D. Miller, A convexity property in the theory of random variables defined on a finite Markov chain, *Ann. Math. Statist*, **32** (1961), 1260-1270.

[7] S. V. Nagaev, Some limit theorems for stationary Markov chains, *Theory Prob. Appl*, **2** (1957), 378-406.

[8] P. Ney and E. Numellin, Markov Additive Processes, I and II, *Ann. Prob.* **15** (1987), 561-592 and 593-609.

[9] M. Seva, On the local limit theorem for non-uniformly ergodic Markov chain, *J. Appl. Prob.* **32** (1995), 32-62.

A. Rahimzadeh Sani
Department of Mathematics
Tarbiat Moallem University of Tehran
49 Mofatteh Avenue
Tehran, 15614, Iran
Email: rahimsan@saba.tmu.ac.ir