Rough Draft
1/29/99
On the Ergodicity of General State Space Markov Chains

James D. Lynch[1]
Department of Statistics
University of South Carolina
Columbia, SC 29208

Jayaram Sethuraman
Department of Statistics
Florida State University
Tallahassee, FL 32306-3033

January 1998

Abstract

Ergodicity questions about the chain are related to the almost sure (a.s.) and L^1 behavior of a reversed supermartingal referred to as the likelihood ratio trajectory and to the zero-oneness of the tail σ-field of the chain. Implications are that (i) convergence of the Markov simulation method is related to the a.s. convergence of the corresponding trajectory and that (ii) the variation norm between the distributions in the likelihood ratio regulates, through Doob's inequality, how far the trajectory of the simulation is from its limit.

AMS Subject Classification: Primary 60J05, 60B10. Secondary 60G42, 65U05.
Key words and phrases: general state Markov chains, Markov simulation, ergodic, zero-one σ-fields, reversed supermartingale.

[1] Supported by NSF Grant DMS 9503104

1. Introduction and Statement of Main Results

The ergodicity of general state Markov chains is considered in this paper. More specifically, variation norm ergodicity questions about the chain are related to the almost sure (a.s.) and L^1 behavior of a reversed supermartingal referred to as the likelihood ratio trajectory. Implications are that (i) convergence of the Markov simulation method is related to the a.s. convergence of the corresponding trajectory and that (ii) the variation norm between the distributions in the likelihood ratio regulates, through Doob's inequality, how far the trajectory of the simulation is from its limit. The following is needed to precisely state these relationships.

Let $\{\boldsymbol{X},\boldsymbol{B}\}$ be a measurable space and let P(x,B) be a ***transition function*** on $\{\boldsymbol{X},\boldsymbol{B}\}$, i.e., P(x,B) is

(i) a probability measure (p.m.) on $\{\boldsymbol{X},\boldsymbol{B}\}$ for each $x\in \boldsymbol{X}$ and
(ii) $\{\boldsymbol{X},\boldsymbol{B}\}$measurable for each $B\in B$.

From Tulcea's Theorem (see Neveu, 1965, Proposition V.1.1) it follows that there exists a unique p.m. P_x on $\{\Omega,\boldsymbol{A}\} = \prod_{t=0}^{\infty} \{\boldsymbol{X}, \boldsymbol{B}\}$ such that for every measurable rectangle $\prod_{t=0}^{n} F_t$,

$$P_x(\prod_{t=0}^{n} F_t) = I_{F_0}(x) \int_{x_1\in F_1} \cdots \int_{x_n\in F_n} \prod_{t=1}^{n} P(x_{t-1},dx_t).$$

For μ a p.m. on $\{\boldsymbol{X},\boldsymbol{B}\}$ let P_μ denote the p.m. on $\{\Omega,\boldsymbol{A}\}$ given by

$$P_\mu(A) = \int P_x(A)\mu(dx)$$

The mappings X_n: $\{\Omega,\boldsymbol{A}\} \to \{\boldsymbol{X},\boldsymbol{B}\}$,n=0,1,..., where $X_n(\omega)=x_n$, define a ***Markov chain*** $\{X_n\}$under P_μ since

$$\begin{aligned} P(X_{n+1}\in B|X_n,\ldots,X_0) &= P(X_{n+1}\in B|X_n) \qquad \text{a.s.} \quad P_\mu \\ &= P(X_n,B) \qquad \text{a.s.} \quad P_\mu. \end{aligned}$$

The ***m-step transition function*** of the chain is given by $P_m(x,B)$ where $P_1(x,B) \equiv P(x,B)$ and

$$P_m(x,B) = \int \ P(y,B)\ P_{m-1}(x,dy)$$

Expectations of the Markov chain $\{X_n\}$ under P_μ will be denoted by E_μ.

When there exists a ***stationary initial probability measure*** or ***equilibrium distribution***, π, on $\{\boldsymbol{X},\boldsymbol{B}\}$ for $P(x,B)$, i.e.,

$$\pi(B) = \int P(x,B)\pi(dx), \tag{1.1}$$

the chain $\{X_n\}$ is ***stationary*** under P_π, i.e.,

$$P_\pi((X_n,X_{n+1},...)\in A) = P_\pi((X_0,X_1,...)\in A) \text{ for every } A\in \mathcal{A}.$$

Let $S:\Omega\rightarrow\Omega$ denote the ***shift operator*** on Ω, i.e., $S((x_0,x_1,x_2,...)) = (x_1,x_2,x_3,...)$ and let $\boldsymbol{I} =\{A:S^{-1}A=A\}$ denote the σ-field of invariant subsets of $\mathcal{A}$. If $\boldsymbol{I}$ is a 0-1 σ-field under P_π, then the stationary process is said to be ***ergodic***. The classical ergodic theorem can be stated as follows for the context considered here. Note that, in the classical ergodic theorem, the sample averages remove any periodic behavior of the chain while reducibility information of the chain is contained in the invariant σ-field, $\boldsymbol{I}$.

The Ergodic Theorem Let $f: \boldsymbol{X}\rightarrow\boldsymbol{R}$ be a Borel measurable function for which $E_\pi(|f(X_0)|) < \infty$. Then,

$$\sum_{i=0}^{n} \frac{f(X_i)}{n+1} \rightarrow E(f(X_0)|\boldsymbol{I}) \qquad \text{a.s. } P_\pi.$$

If the chain is ergodic, then $E(f(X)|\boldsymbol{I}) = E_\pi(f(X))$ a.s. P_π and

$$\sum_{i=0}^{n} \frac{f(X_0)}{n+1} \rightarrow E_\pi(f(X)) \qquad \text{a.s. } P_\pi.$$

For the statement and proof of the ergodic theorem considered here, denote the distribution and expectations of the Markov chain $\{X_m: m=0,1,2,...\}$ initiated with distributions π_0 and μ_0 by P_{π_0}, P_{μ_0}, E_{π_0}, and E_{μ_0}, respectively. Let $\pi_m(B) = P_{\pi_0}(X_m\in B)$ and $\mu_m(B) = P_{\mu_0}(X_m\in B)$ for $B\in \boldsymbol{B}$. Let

$$\lambda(B) = \sum_{n=0}^{\infty} \frac{1}{2^n}(\pi_n(B)+\mu_n(B)).$$

Then, π_n and μ_n are absolutely continuous with respect to λ. Denote their densities (Radon-Nikodym derivatives) with respect to λ by $\pi_n(y)$ and $\mu_n(y)$, repectively. Let $L_m(y) = \frac{\mu_m(y)}{\pi_m(y)}$ where $L_m(y)$ is defined to be zero if $\pi_m(y)=0$. The sequence $\{L_m(X_m)$: $m=1,2,...\}$ will be referred to as the ***likelihood ratio trajectory*** of the Markov chain $\{X_m\}$.

Let

$$\Delta_m \equiv \|\mu_m(\bullet) - \pi_m(\bullet)\| = \sup_B|\mu_m(B) - \pi_m(B)| \tag{1.2a}$$

denote the ***variation norm*** between μ_m and π_m. Note that

$$\begin{aligned} \Delta_m &= \int(\pi_m(y) - \mu_m(y))^+\lambda(dy) = \int(1 - \frac{\mu_m(y)}{\pi_m(y)})^+\pi_m(y)\lambda(dy) \\ &= E_{\pi_0}(1-L_m(X_m))^+. \end{aligned} \tag{1.2b}$$

The main results are stated in the following theorem and lemma while their proofs are deferred to Section 4.

Ergodic Theorem In general, $L_m(X_m) \to L_\infty$ a.s. P_{π_0} and in L^1 and

(ia) $\Delta_m = E_{\pi_0}(1-L_m(X_m))^+ \to E_{\pi_0}(1-L_\infty)^+$ and

(ib) $\lim \mu_m(\pi_m(X_m)>0) = E_{\pi_0}(L_\infty)$

(ii) $\Delta_m \to 0$ if and only if $L_\infty=1$ a.s. P_{π_0}.

In addition if $\mu_m << \pi_m$, then

(iii) $\Delta_m = \frac{1}{2} E_{\pi_0}(|1-L_m(X_m)|)$

and if $\mu_m << \pi_m$ for every m, then

$$\Delta_m = \frac{1}{2} E_{\pi_0}(|1-L_m(X_m)|) \to \frac{1}{2} E_{\pi_0}(|1-L_\infty|).$$

Three examples are given that elucidate the behavior of Δ_m and of $L_m(X_m)$ when the chain is periodic or reducible or null recurrent.

Examples (a) Periodicity Let $\boldsymbol{X} = \{0,1\}$ and consider the two state Markov chain with transition funtion $P(x,\{y\}) = 0$ if $y=x$ and $= 1$ if $y\neq x$. Then, $\pi(\{x\})=\frac{1}{2}$, $x\in \boldsymbol{X}$, is the equilibrium distribution for the chain. Let $\mu_0=\pi$ and $\pi_0=\delta_0$ where δ_0 is the Dirac delta function and indicates that the process is initiated in state 0. Then, $\mu_n=\pi$ and $\pi_n=\delta_0$ if n is even and $\pi_n=\delta_1$ if n is odd. So,

$$L_n(x) = \frac{\mu_n(x)}{\pi_n(x)} = \frac{1}{2\delta_j(x)} \tag{1.3}$$

where j=0 if n is even and =1 if n is odd. Note that in (1.3) $L_n(x)=0$ if $\pi_n(x)=0$ since the observations from the chain are taken under π_0. Under π_0, $L_n(X_n)=\frac{1}{2}$, and so,

$$\Delta_n = E_{\pi_0}(1-L_n(X_n))^+ = \frac{1}{2}.$$

(b) Reducibility Here $\boldsymbol{X} = \{0,1\}$ but $P(x,\{y\}) = 1$ if $y=x$ and $= 0$ if $y \neq x$. Here every initial distribution is an equilibrium distribution. So, $L_n(x) = \frac{\mu(x)}{\pi(x)}$. If $\pi=\delta_0$, then $P_\pi(X_n=0)=1$ and $L_n(X_n) = \frac{\mu(0)}{\pi(0)} = \mu(0)$. So, $\Delta_n = 1-\mu(0) = \mu(1)$.

(c) Gaussian Random Walk (suggested by J. Berger) Let $X_n = X_0 + \sum_{i=1}^{n} Z_i$ where $Z_1, Z_2, \ldots$ are iid standard normal random variables. Fix $m \in \boldsymbol{X}$. Let $\pi_0=\delta_m$ and μ_0 be standard normal. Then, π_n is normal with mean m and variance n while μ_n is normal with mean 0 and variance n+1. Then,

$$L_n(X_n) \overset{d}{=} \left(\frac{n}{n+1}\right)^{1/2} \exp -\frac{1}{2}\left\{\frac{((n+1)^{1/2}Z)^2}{n+1} - \frac{((n+1)^{1/2}Z+m)^2}{n}\right\} \to 1$$

where Z is a standard normal random variable and $\overset{d}{=}$ denotes that the two random variables have the same distribution. Thus, since $(1-x)^+$ is bounded and continuous, $\Delta_n \to 0$. Note that in this example there is no stationary initial distribution but there is an invarient measure, Lebesgue measure on the real line. ■

The above ergodic theorem can be used for chains with a stationary distribution π by choosing $\mu_m = P_m(x,B)$ and $\pi_m = \pi$. In this case, a useful notion of ergodicity is

Variation Norm Ergodicity (VN-ergodicity):

$$\|P_m(x,\bullet) - \pi(\bullet)\| \equiv \sup_{B \varepsilon \boldsymbol{B}} \left| P_m(x,B)-\pi(B) \right| \to 0.$$

This is particularly useful for situations where one is interested in the dynamics of the chain under P_μ where (1.1) holds but μ is not the equilibrium distribution π. For example, in the Markov simulation method (see Athreya et al,1996, for a thorough discussion concerning this method) $\{\boldsymbol{X},\boldsymbol{B}\} = \{\boldsymbol{Y}\times\boldsymbol{Z},\boldsymbol{B}_1\times\boldsymbol{B}_2\}$, and one is interested in simulating from π. The p.m. π is the equilibrium distribution for the transition function, $P((y,z),(dy',dz')) = \pi(dy'|z')\, \pi(dz'|y)$ where the two terms on the right are the conditional p.m.'s y|z and z|y, respectively.

The classical ergodic theorem does not apply to the dynamics of the simulation since $X_0=(y,z)$ (i.e., under $P_{(y,z)}$), but variation norm ergodicity is particularly appropriate since, when $\{\boldsymbol{X},\boldsymbol{B}\}$ is a Borel space, there exists random variables, U and V, have joint distributions with marginals $P_n(x,)$ and π, respectively, such that $\|P_n(x,\bullet) - \pi(\bullet)\| =$

$P(U \neq V)$. Thus, the variation norm indicates just how indistiguishable the Markov simulation is from one for the desired distribution π. In particular, by reversing the roles of $P_m(x,B)$ and π in the likelihood ratio trajectory $\{L_m(X_m)\}$, $L_m(X_m) = \dfrac{\pi_m(X_m)}{p_m(X_m|x)}$, the variation norm regulates how far the Markov simulation is from the desired equilibrium distribution. Namely,

Lemma 1.1 $\varepsilon P_x(\sup_{m\geq n}(1-L_m(X_m))^+ \geq \varepsilon) \leq E_x(1-L_n(X_n))^+ = \Delta_n$.

Lemma 1.2 Let $\mu_m << \pi_m$ for every $m>n$. Then,

$$\varepsilon P_x(\sup_{m\geq n}|1-L_m(X_m)| \geq \varepsilon) \leq E_x(|1-L_n(X_n)|) = 2\Delta_n$$

and

$$\varepsilon P_x(\sup_{m\geq n}|L_\infty-L_m(X_m)| \geq \varepsilon) \leq E_x(|L_\infty-L_n(X_n)|)$$

where $L_m(X_m) \to L_\infty$ a.s.

The proofs of these results are given in the Section 4. They depend on showing that $\{L_m(X_m),\boldsymbol{F}_m\}$, $\boldsymbol{F}_m \equiv \sigma(X_m,X_{m+1},...)$, is a reversed supermartingale which is established in Section 2. The roles of the remote σ-fields in the ergodicity of the chain is discussed in Section 3. There it is shown that in the countable state space case the chain is ergodic iff and only if the tail σ-field is zero-one.

2. Martingale properties of $\{L_m(X_m),\boldsymbol{F}_m\}$

Lemma 2.1 Under P_{π_0}, $\{L_m(X_m):\ m=1,2,...\}$ is a nonegative reversed supermartingale adapted to the fields $\boldsymbol{F}_m \equiv \sigma(X_m,X_{m+1},...)$ and is a reversed martingal if $\mu_n << \pi_n$ for each n.

Proof Let $j < m$. Since $\{X_m:\ m=1,2,...\}$ is a Markov chain,

$$E(L_j(X_j)|\boldsymbol{F}_m) = E(L_j(X_j)|X_m)$$

Then, for $B\in \boldsymbol{B}$,

$$E_{\pi_0}(L_j(X_j)I(X_m\in B)) = E_{\pi_0}(L_j(X_j)I(X_m\in B,\pi_m(X_m)>0))$$

$$= E_{\pi_0}(E_{\pi_0}(L_j(X_j)I(X_m \in B, \pi_m(X_m)>0))|X_j)$$

$$= E_{\pi_0}(L_j(X_j)E_{\pi_0}(I(X_m \in B, \pi_m(X_m)>0))|X_j))$$

$$= E_{\pi_0}(L_j(X_j)P_{m-j}(X_j, X_m \in B, \pi_m(X_m)>0))$$

$$= \int L_j(y)P_{m-j}(y, X_m \in B, \pi_m(X_m)>0)\pi_j(y)\lambda(dy)$$

$$= \int P_{m-j}(y, X_m \in B, \pi_m(X_m)>0)\frac{\mu_j(y)}{\pi_j(y)}\pi_j(y)\lambda(dy)$$

$$= \int P_{m-j}(y, X_m \in B, \pi_m(X_m)>0)\mu_j(y)\lambda(dy)$$

$$- P_{\mu_o}(\pi_j(X_j)=0, \pi_m(X_m)>0, X_m \in B)$$

$$= P_m(x, X_m \in B, \pi_m(X_m)>0)$$

$$- P_{\mu_o}(\pi_j(X_j)=0, \pi_m(X_m)>0, X_m \in B)$$

$$= \int I(y \in B, \pi_m(y)>0)\mu_m(y)\lambda(dy)$$

$$- P_{\mu_o}(\pi_j(X_j)=0, \pi_m(X_m)>0, X_m \in B)$$

$$= \int I(y \in B)\frac{\mu_m(y)}{\pi_m(y)}\pi_m(y)\lambda(dy)$$

$$- P_{\mu_o}(\pi_j(X_j)=0, \pi_m(X_m)>0, X_m \in B)$$

$$= E_{\pi_o}(L_m(X_m)I(X_m \in B)) - P_{\mu_o}(\pi_j(X_j)=0, \pi_m(X_m)>0, X_m \in B)$$

$$\leq E_{\pi_o}(L_m(X_m)I(X_m \in B))$$

Thus,

(2.1) $$0 \leq E_{\pi_0}(L_1(X_1|\boldsymbol{F}_m) \leq E_{\pi_0}(L_j(X_j)|\boldsymbol{F}_m) \leq L_m(X_m)$$

where the inequalities in (2.1) are equalities if $\mu_j << \pi_j$ since

$P_{\mu_o}(\pi_j(X_j)=0, \pi_m(X_m)>0, X_m \in B) = 0$ in this case. ■

The following is an immediate consequence of convergence theorems from martingale theory.

Corollary 2.2 $L_m \to L_\infty$ a.s P_{π_0} and in L^1. If $\mu_n << \pi_n$ for every n, then $L_m = E_{\pi_0}(L_1(X_1|\boldsymbol{F}_m) \to E_{\pi_0}(L_1(X_1|\boldsymbol{F}_\infty) = L_\infty$. a.s P_{π_0} and in L^1.

Proof If $\{L_m, \boldsymbol{F}_m\}$ is a reversed nonegative supermartingale, then it is immedate from the upcrossing inequality that $L_m \to L_\infty$ a.s P_{π_0}. It follows from Fatou's lemma that $E_{\pi_0}(L_\infty) \leq \underline{\lim}\, E_{\pi_0}(L_m(X_m))$ and from (2.1) that $\overline{\lim}\, E_{\pi_0}(L_m(X_m)) \leq E_{\pi_0}(L_\infty)$. Thus, $E(L_m) \to E(L_\infty)$ and it follows that $L_m \to L_\infty$ in L^1 since $L_m \to L_\infty$ a.s.. The statement in the second sentence is immediate since $L_m = E_{\pi_0}(L_1(X_1|\boldsymbol{F}_m)$. ■

3. The Role of the Remote σ-fields

Here we investigate the relationship of the tail σ-field (the remote future σ-field) to the ergodicity of the Markov chain, $\{X_n\}$, when it has an equilibrium distribution π. To investigate this relationship it will be convenient to assume that the chain is stationary and, without loss of generality, doubly infinite. Let $\{\hat{X}_n: \ n=...,-1,0,1,...\}$, $\hat{X}_n = X_{-n}$, denote reversed chain. Under P_π, the reversed chain is stationary with equilibrium distribution π.

For n>0, let $F_{-n} = \sigma(X_{-n}, X_{-n-1},...)$ and $F_n = \sigma(X_n, X_{n+1},...)$. Let $F_{-\infty} = \bigcap_{n>0} \sigma(X_{-n}, X_{-n-1},...)$ and $F_\infty = \bigcap_{n>0} \sigma(X_n, X_{n+1},...)$. The σ-fields $F_{-\infty}$ and F_∞ are just the ***remote past** and **future σ-fields***. The next three theorems indicate the role of the remote σ-fields in the ergodicity of the chain. In particular, Theorem 3.3 shows that VN-ergodicity and F_∞ being zero-one are equivalent under an absolute continuity condition (which holds, for example, for countable state spaces) and indicates that reducibility and periodicity information about the chain is measurable with respect to these σ-fields. This result should be contrasted with Orey's (see Durrett, 1996, Theorem 5.8) which states that for a countable state irreducible recurrent Markov chain the tail σ-field is just the σ-field generated by the periodic classes.

To present these results, we need to define the weaker and more traditional notion of ergodicity, namely,

Markovian Ergodicity (M-ergodicity): $P_m(x,B) \to \pi(B)$ a.s. P_π.

Theorem 3.1 If $\{\hat{X}_n\}$ is M-ergodic, then $\mathcal{F}_\infty$ is zero-one.

Proof Since $\{X_n\}$ is Markovian,

$$(3.1)\quad P(X_o\ \varepsilon\ B|X_n) = P(X_o\ \varepsilon\ B|\mathcal{F}_n) \rightarrow P(X_o\ \varepsilon\ B|\mathcal{F}_\infty) \qquad \text{a.s.}$$

By M-ergodicity,

$$(3.2)\quad P(X_o\ \varepsilon\ B|X_n) \stackrel{d}{=} P(X_{-n}\ \varepsilon\ B|X_0) = P(\hat{X}_n\ \varepsilon\ B|\hat{X}_0) \equiv \hat{P}^n(X_0,B) \rightarrow \pi(B) \text{ a.s.}$$

Since $\pi(B)$ is a constant, it follows from (3.1) and (3.2) that

$$P(X_o\ \varepsilon\ B|\mathcal{F}_\infty) = \pi(B) \text{ a.s.}$$

Thus, X_o and $\mathcal{F}_\infty$ are independent. Since $\{X_n: n < 0\}$ and $\mathcal{F}_\infty$ are conditionally independent given X_o, it follows that $\{X_n: n \le 0\}$ and $\mathcal{F}_\infty$ are independent. A similar argument shows that $\{X_n: n \le k\}$ and $\mathcal{F}_\infty$ are independent for every k. Thus, $\{X_n\}$ and $\mathcal{F}_\infty$ are independent . By considering the reversed chain the same argument shows that $\{X_n\}$ and $\mathcal{F}_{-\infty}$ are independent. Since $\sigma(\ldots,X_{-1},X_0,X_1,\ldots) \supseteq \mathcal{F}_\infty$ and $\mathcal{F}_{-\infty}$, $\mathcal{F}_\infty$ and $\mathcal{F}_{-\infty}$ are zero-one. ■

Theorem 3.2 If $\mathcal{F}_\infty$ is zero-one, then $L_m \rightarrow E_\pi(L_\infty)$ a.s P_π and in L^1. Furthermore, if $E_\pi(L_\infty)=1$, then $\{X_n\}$ is VN-ergodic.

Proof The first result follows from Corollary 2.2, $L_\infty = E_\pi(L_\infty)$ a.s P_π if $\mathcal{F}_\infty$ is zero-one. If $E_\pi(L_\infty)=1$, then $\{X_n\}$ is VN-ergodic by (ia) of the ergodic theorem. ■

Theorem 3.3 Let $\mu_0=P(x,\ )$ and assume that $\mu_n \ll \pi$ for every n. Then, under P_π, $\{X_n\}$ is VN-ergodic if and only if $\mathcal{F}_\infty$ is zero-one.

Proof Note that, since $\mu_n \ll \pi$ for every n, by Corollary 2.2,

$$(3.3)\qquad E_\pi(L_\infty) = E_\pi(L_{-\infty}) = E_\pi(L_m)=1.$$

The proof of the "if part" follows from (3.3) and Theorem 3.2 since $\mathcal{F}_\infty$ is zero-one.

Since $\{X_n\}$ is VN-ergodic, it is M-ergodic. Thus, from Theorem 3.1, is $\mathcal{F}_{-\infty}$ is zero-one. By Theorem 3.2 and (3.3), $\{\hat{X}_n\}$is VN-ergodic, and hence, M-ergodic. This and another application of Theorem 3.1 shows that $\mathcal{F}_\infty$ is zero-one. ■

Application - Countable State Spaces Let the state space, $\boldsymbol{X}$, be countable. Let $D=\{y: \pi(y)>0\}$. Fix $x\in D$ and let $p^n(y|x)$ denote the n-step transition density. Suppose $p^n(y|x)>0$. Then, $\pi(y) = \sum_{z\in D} p^n(y|z)\pi(z) \geq p^n(y|x)\pi(x) > 0$. Thus, $\mu_n = P^n(x, \;)<<\pi$. Thus, by Theorem 3.3, $\lim p^n(y|x) = \pi(y)$ if and only if $\mathcal{F}_\infty$ is zero-one. ■

4. Proofs of Main Results

Proof of the Ergodic Theorem (i) Since $L_m(X_m)\rightarrow L_\infty$ a.s P_{π_0} by Corollary 3.2,

$$\Delta_m = E_{\pi_0}(1-L_m(X_m))^+ \rightarrow E_{\pi_0}(1-L_\infty)^+$$

by (1.2) and the bounded convergence theorem. This proves (ia).

Statement (ib) follow from the identity

$$(4.1) \quad E_{\pi_0}(L_m(X_m)) = \int L_m(y)\pi_m(y)\lambda(dy)) = \int I(\pi_m(y)>0)\mu_m(y)\lambda(dy)$$

$$= \mu_m(\pi_m(X_m)>0) \text{ ,}$$

since $L_m(X_m)\rightarrow L_\infty$ in L^1.

(ii) Statement (ii) follow from (ia) and (1.2) since $L_m(X_m)\rightarrow L_\infty$ a.s P_{π_0} and in L^1 by Corollary 2.2.

(iii) Note that since $\mu_m << \pi_m$, $\Delta_m = \frac{1}{2} E_{\pi_0}(|1-L_m(X_m)|)$. Since $L_m(X_m)\rightarrow L_\infty$ in L^1 by Corollary 3.2, it follows that $1-L_m(X_m)\rightarrow 1-L_\infty$ in L^1 and, from (4.1), that $\lim \mu_m(\pi_m(X_m)>0) = E_{\pi_0}(L_\infty)$. Thus,

$$\Delta_m = \frac{1}{2} E_{\pi_0}(|1-L_m(X_m)|) \rightarrow \frac{1}{2} E_{\pi_0}(|1-L_\infty|).$$

This proves (iii). ■

Proofs of Lemma 1.1 and 2 Lemma 1.1 follows directly from Doob's inequality since $\{1-L_m(X_m), \boldsymbol{F}_m\}$ is a reversed submartingale and $f(x)=(x)^+$ is an increasing convex function. Since $\{|1-L_m(X_m)|, \boldsymbol{F}_m\}$ and $\{|L_\infty-L_m(X_m)|, \boldsymbol{F}_m\}$ are reversed submartingales because $\{1-L_m(X_m), \boldsymbol{F}_m\}$ and $\{L_\infty-L_m(X_m), \boldsymbol{F}_m\}$ are reversed martingales and $f(x)=|x|$ is a convex function, Lemma 1.2 also follows from Doob's inequality. ■

References

1. Athreya, K. B., Doss, H and Sethuraman, J. (1996). A Proof of Convergence of the Markov Chain Simulation Method, ***Annals of Statistics***, 24, 69-100.

2. Durrett, R. (1996) ***Probability Theory and Examples***, Duxbury, Belmont, CA.

3. Neveu, Jacques (1965) ***Mathematical Foundations of the Calculus of Probability***, Holden-Day, Inc, San Francisco.