ELSEVIER

Statistics & Probability Letters 32 (1997) 125–131

STATISTICS & PROBABILITY LETTERS

A local limit theorem for hidden Markov chains

Michael Maxwell *, Michael Woodroofe

Department of Mathematics, University of Michigan, Ann Arbor, MI 48109-1003, USA

Received September 1995; revised February 1996

Abstract

A local limit theorem is proved for partial sums of a hidden Markov chain, assuming global asymptotic normality for a related sum, a fairly weak mixing condition, and a non-lattice condition. The proof proceeds by a study of the conditional characteristic functions, the analysis of which relies heavily on a theorem from Breiman (1968). The paper concludes with a Cesaro type limit theorem for the joint distributions of the Markov chain and the partial sums.

Keywords: Partial sum process; Stationary sequence; Ergodic theorem

1. Introduction

The global central limit theorem for normalized sums of dependent random variables has been investigated extensively, especially under the assumption of stationarity. See, e.g., the review by Peligrad (1986). Much less is known, however, about local limit theorems. See Wang (1990) for a recent contribution and further references. In this paper sufficient conditions are obtained for a local limit theorem to hold when the sequence forms a stationary hidden Markov chain.

Let $W_0, W_1, \ldots$ be a strictly stationary ergodic Markov chain. Let $X_1, X_2, \ldots$ denote a sequence of real random variables which are conditionally independent given $W = (W_0, W_1, \ldots)$ and for which the conditional distribution of X_k given the entire sequence depends only on W_k. In this case $X_1, X_2, \ldots$ is called a hidden Markov chain. The probability space on which the W_k and X_k are defined is denoted by $(\Omega, \mathscr{A}, P)$ and the initial stationary distribution by π, i.e., $P = P^{\pi}$.

Assume $EX_1 = 0$ and $0 < EX_1^2 < \infty$, and let $S_n = X_1 + \cdots + X_n$. The main result of this paper is that under quite modest conditions

$$\sqrt{n}\, P(a < S_n \leqslant b) \to c(b-a) \tag{1}$$

and

$$\frac{1}{n}\sum_{k=1}^{n} \sqrt{k}\, P(W_k \in A,\, a < S_k \leqslant b) \to c\,\pi(A)(b-a) \tag{2}$$

* Corresponding author.

PII S0167-7152(96)00064-8

as $n \to \infty$ whenever $-\infty < a < b < \infty$ and $A \in \mathscr{A}$. Here c is a positive constant which is identified in the proof.

2. Proofs

Let $F(w;\cdot)$ and $\phi(w;\cdot)$ denote the conditional distribution and characteristic function of X_k given $W_k = w$. Letting $\mu(w) = E(X_k \mid W_k = w)$, define $F_0(w;\cdot)$ and $\phi_0(w;\cdot)$ to be the conditional distribution and characteristic function of $X_k - \mu(W_k)$ given $W_k = w$.

The approach is to invoke Theorem 10.7 of Breiman (1968) by selecting an everywhere positive, symmetric, Lebesgue integrable function $h(x)$ with real Fourier transform $2\pi\hat{h}(t)$ which vanishes off a compact interval. Relation (1) is established by showing

$$\sqrt{n}\,E[g(S_n)] \to c\int g(x)\,\mathrm{d}x = 2\pi c\hat{g}(0) \tag{3}$$

for all functions $g(x)$ of the form $e^{ivx}h(x)$, v real. Using the conditional independence of $X_1, X_2, \ldots$ and Parseval's relation, (3) may be rewritten as

$$\sqrt{n}\,E[g(S_n)] = \sqrt{n}\int\!\!\int \hat{g}(t)\prod_{k=1}^{n}\phi(W_k;t)\,\mathrm{d}t\,\mathrm{d}P \to 2\pi c\hat{g}(0). \tag{4}$$

The proof is accomplished by a careful analysis of the left side of (4) using the inequality

$$|\sqrt{n}\,E[g(S_n)] - 2\pi c\hat{g}(0)| \leqslant |E[I_{n,1}(M) - 2\pi c\hat{g}(0)]| + EI_{n,2}(M,\varepsilon) + EI_{n,3}(\varepsilon), \tag{5}$$

where

$$I_{n,1}(M) = \sqrt{n}\int_{|t|\leqslant M/\sqrt{n}} \hat{g}(t)\prod_{k=1}^{n}\phi(W_k;t)\,\mathrm{d}t, \tag{6}$$

$$I_{n,2}(M,\varepsilon) = \sqrt{n}\int_{M/\sqrt{n}\leqslant|t|\leqslant\varepsilon} \left|\hat{g}(t)\prod_{k=1}^{n}\phi(W_k;t)\right|\mathrm{d}t \tag{7}$$

and

$$I_{n,3}(\varepsilon) = \sqrt{n}\int_{|t|\geqslant\varepsilon} \left|\hat{g}(t)\prod_{k=1}^{n}\phi(W_k;t)\right|\mathrm{d}t \tag{8}$$

for appropriate M and ε.

Observe that

$$\lim_{n\to\infty}\frac{1}{n}\sum_{k=1}^{n}\sigma^2(W_k) = \sigma^2 := \int\sigma^2(w)\pi(\mathrm{d}w)$$

w.p.1 by the assumed ergodicity and that the right side is finite since $EX_1^2 < \infty$. The first lemma is nearly as transparent.

Lemma 1. *For* $\varepsilon > 0$

$$\lim_{n\to\infty}\frac{1}{n}\sum_{k=1}^{n}\int_{|x|\geqslant\varepsilon\sqrt{n}} x^2 F_0(W_k;\mathrm{d}x) = 0 \quad w.p.1. \tag{9}$$

Proof. The ergodicity of $W_1, W_2, \ldots$ implies that for fixed M

$$\lim_{n\to\infty} \frac{1}{n}\sum_{k=1}^{n} \int_{|x|\geqslant \varepsilon\sqrt{M}} x^2 F_0(W_k; \mathrm{d}x) = \iint_{|x|\geqslant \varepsilon\sqrt{M}} x^2 F_0(w; \mathrm{d}x)\pi(\mathrm{d}w)$$

w.p.1, and the right side approaches zero as $M \to \infty$ by the Dominated Convergence Theorem. The lemma follows since the integrals in (9) are non-increasing in n. $\square$

Corollary 2. *For* $0<M<\infty$

$$\lim_{n\to\infty} \sup_{|t|\leqslant M} \left| \prod_{k=1}^{n} \phi_0\left(W_k; \frac{t}{\sqrt{n}}\right) - \mathrm{e}^{-(1/2)\sigma^2 t^2} \right| = 0 \quad w.p.1. \tag{10}$$

Proof. Relation (9) asserts that the Lindeberg–Feller condition holds w.p.1. Relation (10) follows. See, e.g., Feller (1966, Vol. 2, pp. 491–492). $\square$

Lemma 3. *Suppose that*

$$V_n := \frac{1}{\sqrt{n}} \sum_{k=1}^{n} \mu(W_k) \Longrightarrow V \sim \mathscr{N}(0, \sigma_0^2) \tag{11}$$

as $n \to \infty$ *where* $0 < \sigma_0^2 < \infty$. *Then*

$$\lim_{M\to\infty} \lim_{n\to\infty} EI_{n,1}(M) = \frac{2\pi\hat{g}(0)}{\sqrt{2\pi(\sigma_0^2+\sigma^2)}}.$$

Proof. First observe that for each fixed $M > 0$

$$\begin{aligned} I_{n,1}(M) &= \int_{-M}^{M} \mathrm{e}^{\mathrm{i}tV_n} \hat{g}\left(\frac{t}{\sqrt{n}}\right) \prod_{k=1}^{n} \phi_0\left(W_k; \frac{t}{\sqrt{n}}\right) \mathrm{d}t \\ &= \hat{g}(0) \int_{-M}^{M} \mathrm{e}^{\mathrm{i}tV_n} \mathrm{e}^{-(1/2)\sigma^2 t^2}\, \mathrm{d}t + \varepsilon_n \end{aligned} \tag{12}$$

where $\varepsilon_n \to 0$ boundedly w.p.1 by Corollary 2 and the Bounded Convergence Theorem. Thus

$$\begin{aligned} E[I_{n,1}(M)] &= \hat{g}(0) \int_{-M}^{M} E(\mathrm{e}^{\mathrm{i}tV_n})\, \mathrm{e}^{-(1/2)\sigma^2 t^2}\, \mathrm{d}t + \mathrm{o}(1) \\ &\to \hat{g}(0) \int_{-M}^{M} \mathrm{e}^{-(1/2)\sigma_0^2 t^2} \mathrm{e}^{-(1/2)\sigma^2 t^2}\, \mathrm{d}t \end{aligned}$$

as $n \to \infty$, and the last line approaches $2\pi\hat{g}(0)[2\pi(\sigma_0^2+\sigma^2)]^{-1/2}$ as $M \to \infty$. $\square$

An additional condition on the Markov chain is now imposed. It is assumed that the state space admits a topology $\mathscr{T}$ under which $\phi(w; t)$ is continuous in w for each fixed t, and that

$$\sum_{k=1}^{n} P(W_1 \in A, W_k \in A) - n\pi(A)^2 = \mathrm{o}(\sqrt{n}) \quad \forall\, A \in \mathscr{T} \cap \mathscr{A}. \tag{13}$$

If the left side of (13) holds for every measurable open set, then it holds for every measurable closed set as well since

$$P(W_1 \in A',\ W_k \in A') = P(W_1 \in A') - P(W_k \in A) + P(W_1 \in A,\ W_k \in A).$$

Of course, $\phi(w;t)$ is always continuous with respect to the discrete topology in which case (13) is required to hold for all measurable sets. If the continuity conditions are satisfied with respect to a smaller topology, then (13) requires less.

The analysis of $I_{n,2}(M,\varepsilon)$ and $I_{n,3}(\varepsilon)$ relies heavily on the next lemma, which is a consequence of (13).

Lemma 4. *For any measurable closed sets A, let*

$$N_n(A) = \#\{k \leqslant n :\ W_k \in A\} = \sum_{k=1}^{n} \mathbf{1}_A(W_k).$$

If (13) *holds then*

$$\operatorname{Var}(N_n(A)) = \mathrm{o}(n^{3/2})$$

and

$$P\left(N_n(A) \leqslant \tfrac{1}{2} n\pi(A)\right) = \mathrm{o}(n^{-1/2}). \tag{14}$$

Proof. The first assertion follows from

$$\operatorname{Var}(N_n(A)) = n\pi(A)(1-\pi(A)) + 2\sum_{k=2}^{n}(n-k+1)\left\{P(W_1 \in A,\ W_k \in A) - \pi(A)^2\right\}$$

and the assumption (13). Relation (14) then follows from Chebyshev's inequality. □

Lemma 5. *Under the assumption of* (13), *there is an $\varepsilon > 0$ so that*

$$\lim_{M\to\infty} \limsup_{n\to\infty} EI_{n,2}(M,\varepsilon) = 0.$$

Proof. For each characteristic function $\phi_0(w;\cdot)$ let $\psi_0(w;\cdot) = |\phi_0(w;\cdot)|^2$, and let $H(w;\cdot)$ denote the symmetrization of $F(w;\cdot)$. Since $\psi_0(w;\cdot)$ is real valued

$$\frac{1-\psi_0(w;t)}{t^2} = \int_{\mathscr{R}} \frac{1-\cos(xt)}{t^2} H(w;\mathrm{d}x) \to 2\sigma^2(w)$$

almost surely as $t \to 0$. Since $\sigma^2 > 0$, there is an $\varepsilon > 0$ so that the set

$$A = \left\{w : \frac{1-\psi_0(w;t)}{t^2} \geqslant \varepsilon\ \forall\ 0 < |t| < \varepsilon\right\}$$

has positive π-measure. With apparent notation, write

$$EI_{n,2}(M,\varepsilon) = E\left[I_{n,2}(M,\varepsilon), B_n\right] + E\left[I_{n,2}(M,\varepsilon), B_n'\right]$$

where $B_n = \{N_n(A) \leqslant n\pi(A)/2\}$. Since A is also closed, Lemma 4 implies

$$E\left[I_{n,2}(M,\varepsilon), B_n\right] \leqslant \mathrm{O}(\sqrt{n})P(B_n) \to 0$$

as $n \to \infty$. In addition, since $|\prod_{k=1}^{n} \phi(W_k;t)| \leqslant \exp(-t^2 N_n(A)\varepsilon/2)$ for each t,

$$E\left[I_{n,2}(M,\varepsilon), B_n'\right] \leqslant \iint_{M/\sqrt{n} \leqslant |t| < \infty} \sqrt{n}\, \mathrm{e}^{-(t^2/2)((1/2)n\pi(A))\varepsilon}\, \mathrm{d}t\, \mathrm{d}P$$

$$= \iint_{M\sqrt{\pi(A)\varepsilon/2} \leqslant |t| < \infty} \sqrt{\frac{2}{\pi(A)\varepsilon}}\, \mathrm{e}^{-(1/2)t^2}\, \mathrm{d}t\, \mathrm{d}P,$$

and the last line approaches 0 as $M \to \infty$. □

Lemma 6. *If*

$$P\left(|\phi(W_1;t)| < 1\right) > 0 \quad \forall t \neq 0 \tag{15}$$

then

$$\lim_{n\to\infty} EI_{n,3}(\varepsilon) = 0.$$

Proof. Since $\hat{g}$ is compactly supported, it suffices to show that for any $t_0 \neq 0$, there is a neighborhood U of t_0 such that

$$\sqrt{n} \iint_U \left| \hat{g}(t) \prod_{k=1}^{n} \phi(W_k;t) \right| \mathrm{d}t\, \mathrm{d}P \to 0$$

as $n \to \infty$. Since

$$\left| \frac{\mathrm{d}}{\mathrm{d}t} \phi(w;t) \right| \leqslant \int |x| F(w;\mathrm{d}x) \leqslant \sqrt{\sigma^2(w) + \mu^2(w)}$$

almost surely, the finiteness of σ^2 ensures that $|\phi(w;\cdot)|$ is Lipshitz continuous w.p.1. This observation, coupled with $P(|\phi(w;t_0)| < 1) > 0$, implies the existence of a neighborhhood U of t_0 and some $\delta > 0$ so that the set

$$B = \left\{ w \colon \sup_{t\in U} |\phi(w;t)| \leqslant 1 - \delta \right\}$$

has positive π-measure. Since B is also closed, (14) implies

$$\left| \sqrt{n} \iint_U \hat{g}(t) \prod_{k=1}^{n} \phi(W_k;t)\, \mathrm{d}t\, \mathrm{d}P \right| \leqslant C\sqrt{n} \int_{N_n(B) \geqslant (1/2)n\pi(B)} \int \left| \prod_{k=1}^{n} \phi(W_k;t) \right| \mathrm{d}t\, \mathrm{d}P + \mathrm{o}(1)$$

$$\leqslant C\sqrt{n}(1-\delta)^{(1/2)n\pi(B)} + \mathrm{o}(1) = \mathrm{o}(1). \quad \square$$

Theorem 7. *Under the assumptions* (11), (13) *and* (15), $\sqrt{n}P_0S_n^{-1}$ *converges weakly to* $c \times$ *Lebesgue measure, where* $c = [2\pi(\sigma_0^2 + \sigma^2)]^{-1/2}$. *In particular, for any* $-\infty < a < b < \infty$

$$\sqrt{n}P(a < S_n \leqslant b) \to \frac{b-a}{\sqrt{2\pi(\sigma_0^2 + \sigma^2)}}$$

as $n \to \infty$.

Proof. This follows directly from (5), Theorem 10.7 of Breiman (1968) and Lemmas 3, 5 and 6. □

The proof of (2) is presented next.

Lemma 8. *Suppose that* (11) *holds. Then for any $A \in \mathscr{A}$ and any open interval $B \subset \mathscr{R}$,*

$$\lim_{n\to\infty} \frac{1}{n}\sum_{k=1}^{n} P(W_k \in A,\ V_k \in B) = \pi(A)P(V \in B). \tag{16}$$

Proof. Fix an open interval B and a measurable set A. Given $\varepsilon > 0$, fix an integer m so large that

$$\int \left| \frac{1}{m}\sum_{l=0}^{m-1} p^l(A \mid w) - \pi(A) \right| \pi(\mathrm{d}w) < \varepsilon$$

by the mean ergodic theorem, where $p^l(A \mid w) = P(W_l \in A \mid W_0 = w)$ a.e. Next define

$$f_1(n) = \frac{1}{n}\sum_{k=1}^{n} P(W_k \in A,\ V_k \in B)$$

$$f_2(n) = \frac{1}{n}\sum_{k=m}^{n} \frac{1}{m}\sum_{l=0}^{m-1} P(W_k \in A,\ V_{k-l} \in B),$$

$$f_3(n) = \frac{1}{n}\sum_{k=1}^{n} \frac{1}{m}\sum_{l=0}^{m-1} P(W_{k+l} \in A,\ V_k \in B),$$

$$f_4(n) = \frac{\pi(A)}{n}\sum_{k=1}^{n} P(V_k \in B).$$

It suffices to show that $\limsup_n |f_1(n) - f_4(n)| < \varepsilon$, since $f_4(n)$ approaches the right side of (16) as $n \to \infty$ by (11).

Observe that for any fixed integer $s \geqslant 0$

$$|V_{n+s} - V_n| \leqslant \frac{\sqrt{n+s} - \sqrt{n}}{\sqrt{n(n+s)}} \left| \sum_{k=1}^{n} \mu(W_k) \right| + \frac{1}{\sqrt{n+s}} \left| \sum_{k=n}^{n+s} \mu(W_k) \right|,$$

and that the right side approaches zero almost surely as $n \to \infty$. Therefore $\max_{0 \leqslant s \leqslant m} P\left(V_n^{-1}(B) \,\Delta\, V_{n+s}^{-1}(B)\right) \to 0$, and consequently $|f_1(n) - f_2(n)| \to 0$, as $n \to \infty$. It is easy to check that $|f_2(n) - f_3(n)| \to 0$ as well.

Finally notice that for every n

$$\begin{aligned} |f_3(n) - f_4(n)| &= \left| \frac{1}{n}\sum_{k=1}^{n} \int \left\{ \frac{1}{m}\sum_{l=0}^{m-1} p^l(A \mid w) - \pi(A) \right\} P(V_k \in B | W_k = w)\, \pi(\mathrm{d}w) \right| \\ &\leqslant \int \left| \frac{1}{m}\sum_{l=0}^{m-1} p^l(A \mid w) - \pi(A) \right| \pi(\mathrm{d}w) < \varepsilon. \end{aligned}$$

That $\limsup_n |f_1(n) - f_4(n)| < \varepsilon$ now follows from the triangle inequality. □

Theorem 9. *Under the assumptions of Theorem 7,*

$$\frac{1}{n}\sum_{k=1}^{n}\sqrt{k}P(W_k \in A,\, a < S_k \leqslant b) \to \frac{\pi(A)(b-a)}{\sqrt{2\pi(\sigma^2+\sigma_0^2)}}, \tag{17}$$

for any $A \in \mathscr{A}$ and $a < b$.

Proof. Let $c = [2\pi(\sigma_0^2+\sigma^2)]^{-1/2}$ and fix $A \in \mathscr{A}$. Relation (17) is established by applying Theorem 10.7 of Breiman (1968) to the measures H_n defined by $H_n(f) = 1/n\sum_{k=1}^{n}\sqrt{k}E[f(S_k)\mathbf{1}_A(W_k)]$. For fixed g as in (3) observe that

$$\begin{aligned}
&\left| H_n(g) - E\left[\frac{1}{n}\sum_{k=1}^{n}\mathbf{1}_A(W_k)I_{k,1}(M)\right]\right| \\
&\quad = \left| E\left[\frac{1}{n}\sum_{k=1}^{n}\mathbf{1}_A(W_k)\{\sqrt{k}E[g(S_k)\mid W] - I_{k,1}(M)\}\right]\right| \\
&\quad \leqslant E\left[\frac{1}{n}\sum_{k=1}^{n}\mathbf{1}_A(W_k)\{I_{k,2}(M,\varepsilon)+I_{k,3}(\varepsilon)\}\right],
\end{aligned}$$

and so by Lemmas 5 and 6 it suffices to show that

$$E\left[\frac{1}{n}\sum_{k=1}^{n}\mathbf{1}_A(W_k)I_{k,1}(M)\right] \to 2\pi c\hat{g}(0)\pi(A) \tag{18}$$

as $n \to \infty$ and then $M \to \infty$. Using (12) the left side of (18) can be written as

$$\hat{g}(0)\int_{-M}^{M}\left\{\frac{1}{n}\sum_{k=1}^{n}E[\mathrm{e}^{\mathrm{i}tV_k}\mathbf{1}_A(W_k)]\right\}\mathrm{e}^{-(1/2)\sigma^2t^2}\mathrm{d}t + \mathrm{o}(1) \to \hat{g}(0)\int_{-M}^{M}\pi(A)\mathrm{e}^{-(1/2)(\sigma^2+\sigma_0^2)t^2}\mathrm{d}t,$$

where the last line follows from Lemma 8. The desired result is obtained by letting $M \to \infty$. □

References

Breiman, L. (1968), *Probability* (Addison-Wesley, Reading, MA).

Feller, W. (1966), An Introduction to Probability Theory and its Application, Vol. 2 (Wiley, New York).

Peligrad, M. (1986), Recent advances in the central limit theorem and its weak invariance principle for mixing sequences of random variables, in: E. Eberlein and M.S. Taqqu, eds., *Dependence in Probability and Statistics.* Progress Probab. Statist. No. 11 (Birkhäuser, Boston, MA) pp. 193–224.

Wang, M. and M. Woodroofe (1990), A local limit theorem for sums of dependent random variables, *Statist. Probab. Lett.* 9, 207–213.