The Annals of Probability
2000, Vol. 28, No. 2, 713–724

CENTRAL LIMIT THEOREMS FOR ADDITIVE FUNCTIONALS OF MARKOV CHAINS[1]

BY MICHAEL MAXWELL AND MICHAEL WOODROOFE

San Jose State University and University of Michigan

Central limit theorems and invariance principles are obtained for additive functionals of a stationary ergodic Markov chain, say $S_n = g(X_1) + \cdots + g(X_n)$, where $E[g(X_1)] = 0$ and $E[g(X_1)^2] < \infty$. The conditions imposed restrict the moments of g and the growth of the conditional means $E(S_n \mid X_1)$. No other restrictions on the dependence structure of the chain are required. When specialized to shift processes, the conditions are implied by simple integral tests involving g.

1. Introduction. Let $X_0, X_1, X_2, \ldots$ denote an ergodic stationary Markov chain with values in a measurable space $(\mathscr{X}, \mathscr{B})$, transition function Q, and stationary initial distribution π. Further, let $L^2(\pi)$ denote the space of (equivalence classes of) square integrable functions $g: \mathscr{X} \to \Re$ for which $\|g\|^2 := \int_{\mathscr{X}} g^2\, d\pi < \infty$, and let $L_0^2(\pi)$ denote the set of $g \in L^2(\pi)$ for which $\int_{\mathscr{X}} g\, d\pi = 0$. Given $g \in L_0^2(\pi)$, let

$$S_n = S_n(g) := g(X_1) + \cdots + g(X_n) \tag{1}$$

and

$$S_n^* = \frac{1}{\sqrt{n}} S_n.$$

for $n \geq 1$. The problem is to find conditions under which S_n^* is asymptotically normal. There are several approaches to this problem. If the chain has a recurrence point, then the problem may be reduced to the independent case, as in Meyn and Tweedie [(1993), pages 418–421]. If the chain exhibits suitable strong mixing, then blocking arguments may be useful, as in Dehling, Denker and Phillip (1986) or Peligrad (1986), for example. If there is a solution to Poisson's equation, $h = g + Qh$, then the problem may be reduced to the martingale case, as in Gordin and Lifsic (1978) and Bhattacharrya and Lee (1988). Here we explore some extensions of the latter approach, along the lines of Kipnis and Varadhan (1986), Toth (1986) and Woodroofe (1992), to cases where a solution to Poisson's equation is not required.

The condition imposed here for normality is a growth condition on the conditional mean $E(S_n \mid X_1 = x)$. To describe it, let Q denote both the conditional

---

Received March 1998; revised November 1999.
[1]Supported in part by NSF Grant DMS-96-26347.
AMS 1991 *subject classification*. 60F05.
Key words and phrases. Asymptotic normality, ergodic theorem, functional central limit theorem, Hilbert space, martingale, maximal inequality, one-sided shifts, Poisson's equation.

distribution of X_1 given X_0 and the operator defined by

$$Qh(x) = \int_{\mathscr{X}} h(y) Q(x; dy)$$

for a.e. $x \in \mathscr{X}$ and all $h \in L^2(\pi)$. Then Q is a contraction. Let

$$V_m h = \sum_{k=0}^{m-1} Q^k h$$

for $h \in L^2(\pi)$ and $m \geq 1$. Then each V_m is a bounded linear operator. Observe that $V_m g(x) = E[S_m(g) \mid X_1 = x]$ for a.e. $x(\pi)$. The condition required here for asymptotic normality is that

$$\sum_{n=1}^{\infty} n^{-3/2} \|V_n g\| < \infty, \tag{2}$$

where $\|\cdot\|$ denotes the norm in $L^2(\pi)$. A main result is that if (2) holds, then

$$\sigma^2 := \sigma^2(g) = \lim_{n \to \infty} E(S_n^{*2}) \tag{3}$$

exists and is finite, and

$$S_n^* \Rightarrow \text{Normal}[0, \sigma^2] \tag{4}$$

as $n \to \infty$, where $\Rightarrow$ denotes convergence is distribution. In fact, (4) holds conditionally given X_0, as described in Corollary 1 below. For a comparison with Gordin and Lifsic's (1978) result, observe that if there is a solution to Poisson's equation $h = g + Qh$, where $h \in L^2(\pi)$, then $V_n g = h - Q^n h$ and, therefore, $\|V_n g\| = O(1)$. Clearly, $\|V_n g\| = O(\sqrt{n})$, if (3) holds, since $\|V_n g\|^2 = E[E(S_n \mid X_1)^2] \leq E(S_n^2)$ for all $n \geq 1$. In Corollary 1, it is also shown that (3) and the conditional version of (4) imply $\|V_n g\| = o(\sqrt{n})$. In this rough sense, (2) is within a logarithmic term of being necessary.

Relations (3) and (4) are established in Section 2, and a functional central limit theorem is established in Section 3 under the stronger conditions that $g \in L^p(\pi)$ for some $p > 2$ and $\|V_n g\| = O(n^\alpha)$ for some $\alpha < 1/2$. In Section 4, the results are specialized to shift processes, and it is shown that the condition (2) is implied by simple integral tests.

2. Asymptotic normality.

For $\varepsilon > 0$, let h_ε be the solution to the equation $(1+\varepsilon)h = Qh + g$,

$$h_\varepsilon = \sum_{k=1}^{\infty} \frac{Q^{k-1} g}{(1+\varepsilon)^k} = \varepsilon \sum_{n=1}^{\infty} \frac{V_n g}{(1+\varepsilon)^{n+1}}. \tag{5}$$

Let π_1 be the joint distribution of X_0 and X_1, so that $\pi_1\{dx_0\, dx_1\} = Q(x_0; dx_1) \times \pi\{dx_0\}$; denote the norm in $L^2(\pi_1)$ by $\|\cdot\|_1$; and let

$$H_\varepsilon(x_0, x_1) = h_\varepsilon(x_1) - Q h_\varepsilon(x_0)$$

for $x_0, x_1 \in \mathscr{X}$. Then each H_ε is in $L^2(\pi_1)$; the norm of H_ε is $\|H_\varepsilon\|_1^2 = \|h_\varepsilon\|^2 - \|Qh_\varepsilon\|^2$; and $\int_{\mathscr{X}} H_\varepsilon(x_0, x_1)Q(x_0; dx_1) = 0$ for a.e. $x_0(\pi)$. With this notation,

$$S_n(g) = M_n(\varepsilon) + \varepsilon S_n(h_\varepsilon) + R_n(\varepsilon), \tag{6}$$

where

$$M_n(\varepsilon) = H_\varepsilon(X_0, X_1) + \cdots + H_\varepsilon(X_{n-1}, X_n)$$

and

$$R_n(\varepsilon) = Qh_\varepsilon(X_0) - Qh_\varepsilon(X_n).$$

For each fixed ε, $M_n(\varepsilon)$ is a martingale; $M_n(\varepsilon)/\sqrt{n}$ is asymptotically normal with mean 0 and variance $\|H_\varepsilon\|_1^2$; and $R_n(\varepsilon) = O_p(1)$.

Let $\delta_k = 2^{-k}$ for $k \geq 0$.

LEMMA 1. *If* (2) *holds, then* $\sqrt{\varepsilon}\|h_\varepsilon\| \to 0$ *as* $\varepsilon \to 0$ *and* $\sum_{k=1}^\infty \sqrt{\delta_k} \times \sup_{\delta_k \leq \varepsilon \leq \delta_{k-1}} \|h_\varepsilon\| < \infty$. *Further, if* $\|V_n g\| = O(n^\alpha)$ *for some* $\alpha > 0$, *then* $\|h_\varepsilon\| = O(\varepsilon^{-\alpha})$ *as* $\varepsilon \to 0$.

PROOF. Clearly, $\|h_\varepsilon\| \leq 2\delta_k \sum_{n=1}^\infty (1+\delta_k)^{-n}\|V_n g\|$ for $\delta_k \leq \varepsilon \leq \delta_{k-1}$. So,

$$\sum_{k=1}^\infty \sqrt{\delta_k} \sup_{\delta_k \leq \varepsilon \leq \delta_{k-1}} \|h_\varepsilon\| \leq 2 \sum_{n=1}^\infty \left[\sum_{k=1}^\infty \frac{\delta_k^{3/2}}{(1+\delta_k)^n} \right] \|V_n g\|.$$

The inner sum here is $O(n^{-3/2})$ by comparison with $\int_0^1 \sqrt{x}(1+x)^{-n}\,dx$, and the first two assertions follow easily. The third assertion follows similarly by writing $\|h_\varepsilon\| \leq \varepsilon \sum_{n=1}^\infty (1+\varepsilon)^{-n}\|V_n g\|$ and comparing the latter sum to $\int_o^\infty x^\alpha (1+\varepsilon)^{-x}\,dx$. □

LEMMA 2. *For* $0 < \varepsilon, \delta < \infty$, $\|H_\varepsilon - H_\delta\|_1^2 \leq (\varepsilon + \delta)[\|h_\varepsilon\|^2 + \|h_\delta\|^2]$.

PROOF. Using $\langle H_\varepsilon, H_\delta\rangle_1 = \langle h_\varepsilon, h_\delta\rangle - \langle Qh_\varepsilon, Qh_\delta\rangle$ and $Qh_\varepsilon = (1+\varepsilon)h_\varepsilon - g$,

$$\begin{aligned}\langle H_\varepsilon, H_\delta\rangle_1 &= \langle h_\varepsilon, h_\delta\rangle - [(1+\varepsilon)(1+\delta)\langle h_\varepsilon, h_\delta\rangle - (1+\varepsilon)\langle h_\varepsilon, g\rangle \\ &\qquad - (1+\delta)\langle h_\delta, g\rangle + \|g\|^2] \\ &= -(\varepsilon + \delta + \varepsilon\delta)\langle h_\varepsilon, h_\delta\rangle + [(1+\varepsilon)\langle h_\varepsilon, g\rangle + (1+\delta)\langle h_\delta, g\rangle - \|g\|^2]\end{aligned}$$

and

$$\begin{aligned}\|H_\varepsilon - H_\delta\|_1^2 &= -(2\varepsilon + \varepsilon^2)\|h_\varepsilon\|^2 + 2(\varepsilon + \delta + \varepsilon\delta)\langle h_\varepsilon, h_\delta\rangle - (2\delta + \delta^2)\|h_\delta\|^2 \\ &\leq 2(\varepsilon+\delta)\|h_\varepsilon\|\|h_\delta\| - (\varepsilon\|h_\varepsilon\| - \delta\|h_\delta\|)^2 \\ &\leq (\varepsilon+\delta)[\|h_\varepsilon\|^2 + \|h_\delta\|^2],\end{aligned}$$

as asserted. □

PROPOSITION 1. *If* (2) *holds, then* $H = \lim_{\varepsilon \downarrow 0} H_\varepsilon$ *exists in* $L^2(\pi_1)$.

PROOF. Let $\delta_k = 2^{-k}$, as above. Then $\|H_{\delta_k} - H_{\delta_{k-1}}\|_1^2 \le 3\delta_k\|h_{\delta_k}\|^2 + 2\delta_{k-1}\|h_{\delta_{k-1}}\|^2$ for all $k \ge 1$, by Lemma 2. So,

$$\sum_{k=1}^{\infty} \|H_{\delta_k} - H_{\delta_{k-1}}\|_1 \le 4 \sum_{k=0}^{\infty} \sqrt{\delta_k}\|h_{\delta_k}\| < \infty,$$

by Lemma 1 (using $\sqrt{a+b} \le \sqrt{a} + \sqrt{b}$ for nonnegative a and b). It follows that $H := \lim_{k\to\infty} H_{\delta_k}$ exists in $L^2(\pi_1)$. If $0 < \varepsilon < 1$, then there is a unique $k = k(\varepsilon)$ for which $\delta_k \le \varepsilon < \delta_{k-1}$. With this choice of k, $\|H_\varepsilon - H_{\delta_k}\|_1^2 \le 7\delta_k \sup_{\delta_k \le \varepsilon \le \delta_{k-1}} \|h_\varepsilon\|^2 \to 0$ as $\varepsilon \to 0$, so that $\lim_{\varepsilon\downarrow 0} H_\varepsilon = H$. □

THEOREM 1. *If* (2) *holds, then*

$$S_n = M_n + R_n, \tag{7}$$

where $M_1, M_2, \ldots$ *and* $R_1, R_2, \ldots$ *have strictly stationary increments,* $M_1, M_2, \ldots,$ *is a square integrable martingale, and* $E(R_n^2) = o(n)$ *as* $n \to \infty$.

PROOF. Let $M_n = H(X_0, X_1) + \cdots + H(X_{n-1}, X_n)$ for $n \ge 1$. Then $\lim_{\varepsilon\downarrow 0} M_n(\varepsilon) = M_n$ in $L^2(P)$ for each fixed n in (6), since $E\{[M_n(\varepsilon) - M_n]^2\} = n\|H_\varepsilon - H\|_1^2 \to 0$ by Proposition 1, and $\lim_{\varepsilon\downarrow 0} \varepsilon S_n(h_\varepsilon) = 0$ in $L^2(P)$ by Lemma 1. So, $R_n := \lim_{\varepsilon\downarrow 0} R_n(\varepsilon)$ exists in $L^2(P)$ for each fixed n, and (7) holds. It is clear that $M_1, M_2, \ldots$ and $R_1, R_2, \ldots$ have strictly stationary increments and that $M_1, M_2, \ldots,$ is a square integrable martingale. So, it remains to show that $E(R_n^2) = o(n)$ as $n \to \infty$. For each $n \ge 1$, let k_n be the unique integer k for which $2^{k-1} \le n < 2^k$ and let $\varepsilon_n = 2^{-k_n}$. Then $R_n = M_n(\varepsilon_n) - M_n + \varepsilon_n S_n(h_{\varepsilon_n}) + R_n(\varepsilon_n)$ and, therefore,

$$\begin{aligned} E(R_n^2) &\le 4E\{[M_n(\varepsilon_n) - M_n]^2\} + 4\varepsilon_n^2 E[S_n(h_{\varepsilon_n})^2] + 4E[R_n(\varepsilon_n)^2] \\ &\le 4n\|H_{\varepsilon_n} - H\|_1^2 + 4\|h_{\varepsilon_n}\|^2 + 8\|h_{\varepsilon_n}\|^2. \end{aligned} \tag{8}$$

The first term on the right side of (8) is $o(n)$ by Proposition 1, and the second two are $o(n)$ by Lemma 1. □

Asymptotic normality is a simple consequence of the theorem. Denote conditional probability given $X_0 = x$ by P^x and let

$$F_n(x; z) = P^x\{S_n^* \le z\}$$

for $z \in \Re$, $x \in \mathscr{X}$ and $n \ge 1$. Let Φ_σ denote the normal distribution with mean 0 and variance σ^2, and let Δ denote the Levy metric for distribution functions; that is, $\Delta(F, G) = \inf\{\varepsilon > 0\colon G(x - \varepsilon) - \varepsilon \le F(x) \le G(x + \varepsilon) + \varepsilon \text{ for all } x\}$.

COROLLARY 1. *If* (2) *holds, then* (3) *and* (4) *hold with* $\sigma^2 = \|H\|_1^2$ *and*

$$\lim_{n\to\infty} \int_{\mathscr{X}} \Delta[\Phi_\sigma, F_n(x; \cdot)]\, \pi\{dx\} = 0. \tag{9}$$

Conversely, if (3) *and* (9) *hold, then* $\|V_n g\| = o(\sqrt{n})$.

PROOF. Suppose first that (2) holds. Then (3) holds with $\sigma^2 = \|H\|_1^2$, because $E(M_n^2) = n\|H\|_1^2$ for all n, and $E(R_n^2) = o(n)$ as $n \to \infty$. For the normality, it suffices to establish (9). Observe first that $M_1, M_2, \dots$ is a martingale with respect to P^x for a.e. x (π), by the Markov property. Let $G_n(x; z) = P^x\{M_n^* \le z\}$ for $z \in \Re$, $x \in \mathscr{X}$, and $n \ge 1$. By the Ergodic theorem,

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{\infty} H(X_{k-1}, X_k)^2 = \|H\|_1^2$$

w.p.1 (P) and, therefore, w.p.1 (P^x) for a.e. x (π). It then follows from the Martingale central limit theorem that $G_n(x; \cdot) \Rightarrow \Phi_\sigma$ as $n \to \infty$ for a.e. x (π). See, for example, Durrett and Resnick (1978) or Hall and Heyde [(1981), Theorem 3.2]. Relation (9) then follows by integrating the inequality

$$\Delta[\Phi_\sigma, F_n(x; \cdot)] \le \Delta[\Phi_\sigma, G_n(x; \cdot)] + \varepsilon + P^x\left\{\left|\frac{R_n}{\sqrt{n}}\right| \ge \varepsilon\right\}$$

with respect to π and then letting $n \to \infty$ and $\varepsilon \to 0$.

For the converse, let $\sigma_n^2 = E[S_n^{*2}]$ and let $\sigma_n^2(x)$ denote the conditional variance of S_n^* given $X_0 = x$ [which is well defined for a.e. x (π)]. Then

$$\sigma_n^2 = \int_{\mathscr{X}} \sigma_n^2(x)\, \pi\{dx\} + \frac{1}{n}\|QV_n g\|^2$$

for all $n \ge 1$. It suffices to show that $\|V_{n_k} g\|/\sqrt{n_k} \to 0$ as $k \to \infty$ for every sequence n_k that increases sufficiently fast. If n_k is a subsequence for which $\Delta[\Phi_\sigma, F_{n_k}(x; \cdot)] \to 0$ for a.e. x (π) as $k \to \infty$, then $\liminf_{k\to\infty} \sigma_{n_k}^2(x) \ge \sigma^2$ for a.e. x (π) by Fatou's lemma and, therefore,

$$\limsup_{k\to\infty} \frac{1}{n_k}\|QV_{n_k} g\|^2 \le \lim_{n\to\infty} \sigma_n^2 - \liminf_{k\to\infty} \int_{\mathscr{X}} \sigma_{n_k}^2(x)\pi\{dx\} \le \sigma^2 - \sigma^2 = 0,$$

by another application of Fatou's lemma. The corollary follows since $0 \le \|V_n g\| - \|QV_n g\| \le 2\|g\|$. □

If the process is mixing, then it is possible to relate (2) to conditions on $\|Q^k g\|$, $k \ge 1$.

COROLLARY 2. *Relations* (3) *and* (4) *hold if either*

$$\sum_{k=1}^{\infty} \frac{1}{\sqrt{k}}\|Q^k g\| < \infty \tag{10}$$

or

$$\sum_{k=1}^{\infty} \log^{1+\delta}(k)\|Q^k g\|^2 < \infty \tag{11}$$

for some $\delta > 0$.

Both (10) and (11) imply (2). The easy proof is omitted. The condition (11) cannot be relaxed very much at this level of generality. An example in which (11) holds with $\delta = 0$, but $\sigma^2 = \infty$ in (3) may be constructed along the lines of Example 3 of Woodroofe (1992) by letting $a_r,\ r \geq 1$, be any square summable sequence and $b_k = 1/\{k \log(1+k) \log\log(2+k)\}$ for $k \geq 1$, in the notation of that example.

3. An invariance principle.

Let

(12) $$\mathbb{B}_n(t) = \frac{1}{\sqrt{n}} S_{\lceil nt \rceil}, \qquad 0 \leq t < 1,$$

and $\mathbb{B}_n(1) = \mathbb{B}_n(1-)$ for $n \geq 1$, where $\lceil x \rceil$ denotes the least integer that is greater than x. In Theorem 2 below, it is shown that $\mathbb{B}_n$ converges in distribution to a Brownian motion in the space $D[0,1]$, under slightly stronger conditions. Here $D[0,1]$ denotes the space of right continuous functions with left limits, as described by Billingsley [(1968), Chapter 3].

The proof uses the following simple maximal inequality.

PROPOSITION 2. *If $S_k,\ k \geq 1$, is any process with strictly stationary increments for which $E(S_n^2) \leq Cn$ for all $n = 1, 2, \ldots$, then*

(13) $$P\Big\{\max_{j \leq n} |S_j| > \lambda\Big\} \leq \frac{2^{6k} C n^{1+2^{-k}}}{\lambda^2}$$

for all $n \geq 1$, all $\lambda > 0$ and all $k \geq 0$.

PROOF. It is clear that (13) holds for $k = 0$. So, suppose inductively that it holds for a given $k \geq 1$. Let $m = \lceil \sqrt{n} \rceil$. Then

$$P\Big\{\max_{k \leq n} |S_k| > 2\lambda\Big\} \leq P\Big\{\max_{k \leq m} |S_{mk}| > \lambda\Big\} + mP\Big\{\max_{k \leq m} |S_k| > \lambda\Big\} \leq 2\frac{2^{6k} C m^{2+2^{-k}}}{\lambda^2},$$

where the second inequality follows by applying (13) twice, once to $S_k,\ k \leq m$, and once to $S_{mk},\ k \leq m$ (which satisfies the basic conditions with C replaced by Cm). Since $m/\sqrt{n} \leq 2$, it then follows that

$$P\Big\{\max_{k \leq n} |S_k| > 2\lambda\Big\} \leq 2^4 \frac{2^{6k} C n^{1+2^{-k-1}}}{\lambda^2}$$

and the proposition then follows by induction. □

The proposition is used in the following form.

COROLLARY 3. *For any $\beta > 1$ there is a constant Γ, depending only on β, for which*

$$P\Big\{\max_{k \leq n} |S_k| > \lambda\Big\} \leq \frac{\Gamma C n^{\beta}}{\lambda^2}$$

for all $\lambda > 0$ and all $n \geq 1$.

Now let $S_k = S_k(g),\ k \geq 1$, as in (1); define $\mathbb{B}_n$ by (12) and let $F_n(x;\cdot)$ denote the conditional distribution of $\mathbb{B}_n$ given $X_0 = x$, a probability measure on the Borel sets of $D[0,1]$, for each $x \in \mathscr{X}$. Further, let $\mathbb{B}$ be a standard Brownian motion, let Φ_σ denote the distribution of $\sigma\mathbb{B}$ and let Δ denote the Prokhorov metric on the space of probability measures on $D[0,1]$. See, for example, Billingsley [(1968), page 238].

THEOREM 2. *If there are* $p > 2$ *and* $\alpha < 1$ *for which* $\int_{\mathscr{X}} |g|^p \, d\pi < \infty$ *and* $E(R_n^2) = o(n^\alpha)$ *as* $n \to \infty$, *then*

$$\lim_{n\to\infty} \int_{\mathscr{X}} \Delta\big[\Phi_\sigma, F_n(x;\cdot)\big]\, \pi\{dx\} = 0. \tag{14}$$

PROOF. Let $M_n^*(t) = M_{\lceil nt \rceil}/\sqrt{n},\ 0 \leq t < 1$ and $M_n^*(1) = M_n^*(1-)$, and let $G_n(x;\cdot)$ denote the conditional distribution of M_n^* given $X_0 = x$. Then $G_n(x;\cdot)$ converges to Φ_σ for a.e. x (π), by Theorem 2.5 of Durrett and Resnick (1978), and

$$\int_{\mathscr{X}} \Delta\big[\Phi_\sigma, F_n(x;\cdot)\big]\, \pi\{dx\} \leq \int_{\mathscr{X}} \Delta\big[\Phi_\sigma, G_n(x;\cdot)\big]\, \pi\{dx\} + \varepsilon + P\Big\{\max_{k\leq n} |R_k| \geq \varepsilon\sqrt{n}\Big\}$$

for each $\varepsilon > 0$ and all $n \geq 1$. So, it suffices to show that $\max_{j\leq n} |R_j|/\sqrt{n} \to 0$ in probability as $n \to \infty$. Given α and p, as in the statement of the theorem, there are $0 < \gamma < 1/2 - 1/p$ and let $\beta > 1$ for which $\alpha\gamma + \beta(1-\gamma) < 1$. Next, let $l = \lceil n^\gamma \rceil$ and $m = \lceil n^{1-\gamma} \rceil$. Given $\varepsilon > 0$, define events

$$A_n = \Big\{\max_{0\leq k-j\leq l,\, k\leq n} |M_k - M_j| \leq \varepsilon\sqrt{n}\Big\},$$

$$B_n = \Big\{\max_{j\leq n} |g(X_j)| \leq \varepsilon\frac{\sqrt{n}}{l}\Big\}$$

and

$$C_n = \Big\{\max_{k\leq m} |R_{lk}| \leq \varepsilon\sqrt{n}\Big\}.$$

Then $A_n \cap B_n \cap C_n \subseteq \{\max_{j\leq n} |R_j| \leq 3\varepsilon\sqrt{n}\}$. For if A_n, B_n and C_n occur and $1 \leq j \leq n$, then $lk \leq j < l(k+1)$ for some $k < m$ and, therefore, $|R_j| \leq |R_{lk}| + |M_j - M_{lk}| + |S_j - S_{lk}| \leq 3\varepsilon\sqrt{n}$. Clearly, $P(A_n') \to 0$ as $n \to \infty$, by the functional martingale central limit theorem and

$$P(B_n') \leq nP\Big\{|g(X_1)| > \varepsilon\frac{\sqrt{n}}{l}\Big\} \leq n\left|\frac{l}{\varepsilon\sqrt{n}}\right|^p \int_{\mathscr{X}} |g|^p \, d\pi \to 0$$

as $n \to \infty$, by the choice of γ. By assumption, there is a constant c for which $E(R_j^2) \leq cj^\alpha$ for all $j = 1, 2, \ldots$. So, the sequence $R_{lk},\ k \geq 1$, satisfies the conditions of Proposition 2 with $C = cl^{\alpha}$. It follows that there is a constant Γ for which

$$P(C'_n) = P\Big\{\max_{k\le m}|R_{lk}| > \varepsilon\sqrt{n}\Big\} = \frac{c\Gamma}{\varepsilon^2}\frac{l^{\alpha}m^{\beta}}{n} \to 0$$

as $n \to \infty$, by the choice of β. Thus, $P\{\max_{j\le n}|R_j| > 3\varepsilon\sqrt{n}\} \le P(A'_n) + P(B'_n) + P(C'_n) \to 0$ as $n \to \infty$. □

COROLLARY 4. *If* $g \in L^p(\pi)$ *for some* $p > 2$ *and* $\|V_n g\| = O(n^{\alpha})$ *for some* $\alpha < 1/2$, *then* (14) *holds.*

PROOF. It suffices to show that $E(R_n^2) = O(n^{2\alpha})$, and this follows easily from (8). For $\|h_{\varepsilon_n}\| = O(n^{\alpha})$ by Lemma 1 and the definition of ε_n, and there is a constant C for which

$$\|H_{\varepsilon_n} - H\|_1 \le \sum_{j=k_n+1}^{\infty}\|H_{\delta_j} - H_{\delta_{j-1}}\|_1 \le C\sum_{j=k_n+1}^{\infty}\delta_j^{(1/2)-\alpha} = O(n^{\alpha-(1/2)}),$$

using Lemma 2. □

The following corollary is obvious.

COROLLARY 5. *If* $g \in L^p(\pi)$ *for some* $p > 2$ *and either*

$$\sum_{k=1}^{\infty} k^{\delta-1/2}\|Q^k g\| < \infty \tag{15}$$

or

$$\sum_{k=1}^{\infty} k^{\delta}\|Q^k g\|^2 < \infty, \tag{16}$$

for some $\delta > 0$, *then* (14) *holds.*

4. Shift processes. In this section, Theorems 1 and 2 are specialized to one-sided shift processes. These processes do not have recurrence points and are not strongly mixing. For such processes, it is possible to formulate an integral condition on g which is sufficient for asymptotic normality. The procedure is simplest for Bernoulli shifts.

Bernoulli shifts. Let ε_k, $k = 0, \pm1, \pm2, \dots$ be i.i.d. random variables that take the values 0 and 1 with probability 1/2 each and let

$$X_n = \sum_{k=1}^{\infty}(1/2)^{k+1}\varepsilon_{n-k}$$

for $n = 0, 1, 2, \dots$. Then X_n, $n \ge 0$, is an ergodic stationary Markov chain taking values in $I = [0, 1]$. The transition function is defined by $Q(x; \{x/2\}) = 1/2 = Q(x; \{(1 + x)/2\})$ for $x \in [0, 1]$ and the stationary initial distribution is the restriction, λ say, of Lebesgue measure to I. In this case, $L^2(\lambda)$ is a familiar space, and it is easy to relate the conditions in Theorems 1 and 2 to regularity properties of g.

PROPOSITION 3. *For the Bernoulli shift process, if* $g \in L_0^2(\lambda)$ *and*

$$\int_0^1 \int_0^1 [g(y) - g(x)]^2 \frac{1}{|x-y|} \log^{1+\delta}\left[\log\left(\frac{1}{|x-y|}\right)\right] dx\,dy < \infty \tag{17}$$

for some $\delta > 0$, *then* (11) *holds, and if*

$$\int_0^1 \int_0^1 [g(x) - g(y)]^2 \frac{1}{|x-y|} \log^{\delta}\left[\frac{1}{|x-y|}\right] dx\,dy < \infty \tag{18}$$

for some $\delta > 0$, *then* (16) *holds.*

PROOF. If $g \in L_0^2(\lambda)$, then

$$\begin{aligned} Q^k g(x) &= 2^{-k} \sum_{z \in D_k} g\left(\frac{x}{2^k} + z\right) \\ &= 2^{-k} \sum_{z \in D_k} \int_0^1 \left[g\left(\frac{x}{2^k} + z\right) - g\left(\frac{y}{2^k} + z\right)\right] dy, \end{aligned}$$

where $D_k = \{j2^{-k}\colon j = 0, \ldots, 2^k - 1\}$ and the second equality follows from $\int_0^1 Q^k g(y)\,dy = 0$. Thus,

$$\begin{aligned} \|Q^k g\|^2 &\le 2^{-k} \sum_{z \in D_k} \int_0^1 \int_0^1 \left[g\left(\frac{x}{2^k} + z\right) - g\left(\frac{y}{2^k} + z\right)\right]^2 dy\,dx \\ &\le 2^k \int\int_{|x-y| \le 2^{-k}} [g(x) - g(y)]^2\,dx\,dy \end{aligned}$$

and

$$\sum_{k=1}^{\infty} \log^{1+\delta}(k) \|Q^k g\|^2 \le \int_0^1 \int_0^1 J(|x-y|)[g(x) - g(y)]^2\,dx\,dy,$$

where

$$J(z) = \sum_{k: 2^{-k} \ge z} 2^k \log^{1+\delta}(k) \le \frac{C}{z} \log^{1+\delta}\left[\log\left(\frac{1}{z}\right)\right]$$

for $0 < z < 1$ for some constant C. The first assertion of the proposition follows easily, and the second may be established similarly.

The condition (18) is not very restrictive. For example, if

$$g(x) = \frac{1}{x^\alpha} \sin\left(\frac{1}{x}\right), \qquad 0 < x \le 1,$$

where $0 < \alpha < 1/2$, then (18) holds. This may be verified by making the change of variables $x = 1/x'$ and $y = 1/y'$ in (18) and dividing the resulting integral into regions where $|x' - y'| \le 1$ and $|x' - y'| > 1$. The details are straightforward and have been omitted. That $[S_n - E(S_n|X_0)]/\sqrt{n}$ is asymptotically normal for this example was shown by Woodroofe (1992) using Fourier based techniques. $\square$

Lebesgue shifts. Now let U_k, $k = 0, \pm 1, \pm 2, \dots$ be i.i.d. random variables that are uniformly distributed over $I = [0, 1]$ and let

$$X_n = (\cdots U_{n-2}, U_{n-1}, U_n), \qquad n \geq 0.$$

Then X_n, $n \geq 0$, is a stationary Markov process taking values in $\mathscr{X} = I^M$, where M denotes the nonpositive integers. The stationary initial distribution π here is the countable product of copies of Lebesgue measure. Processes of the form $g(X_n)$, $n = 0, 1, 2, \dots$, include a wide class of stationary sequences.

Define measures Γ_δ^i on $\mathscr{X} \times \mathscr{X}$, $i = 0, 1$ by

$$\Gamma_\delta^0\{B\} = \sum_{k=1}^{\infty} \log^{1+\delta}(k)(\pi \times \pi \times \lambda^k)\big\{(x, y, z) \colon [(x, z), (y, z)] \in B\big\},$$

$$\Gamma_\delta^1\{B\} = \sum_{k=1}^{\infty} k^{\delta}(\pi \times \pi \times \lambda^k)\big\{(x, y, z) \colon [(x, z), (y, z)] \in B\big\}$$

for Borel sets $B \subseteq \mathscr{X} \times \mathscr{X}$, where x and (x, z) denote the sequences $x = (\dots, u_{-2}, u_{-1}, u_0)$ and $(x, z) = (\dots, u_{-1}, u_0, z_1, \dots, z_k)$ and λ^k denotes Lebesgue measure on I^k.

PROPOSITION 4. *If* $g \in L_0^2(\pi)$ *and*

$$\int_{\mathscr{X}} \int_{\mathscr{X}} [g(x) - g(y)]^2 \, \Gamma_\delta^0\{dx\, dy\} < \infty, \tag{19}$$

for some $\delta > 0$, *then* (11) *holds*; *and if*

$$\int_{\mathscr{X}} \int_{\mathscr{X}} [g(x) - g(y)]^2 \, \Gamma_\delta^1\{dx\, dy\} < \infty, \tag{20}$$

for some $\delta > 0$, *then* (16) *holds*.

PROOF. Clearly,

$$Q^k g(x) = \int_{\mathscr{X}} \int_{I^k} [g(x, z) - g(y, z)] \; \lambda^k\{dz\} \, \pi\{dy\}$$

for a.e. $x \in I^M$ and, therefore,

$$\|Q^k g\|^2 \leq \int_{\mathscr{X}} \int_{\mathscr{X}} \int_{i^k} [g(x, z) - g(y, z)]^2 \, \lambda^k\{dz\} \pi\{dx\} \pi\{dy\}$$

and all $k \geq 1$. The first assertion then follows from multiplying $\|Q^k g\|^2$ by $\log^{1+\delta}(k)$ and summing over $k = 1, 2, \dots$, and the second may be established similarly. □

To illustrate the use of (19) and (20), observe that any $g \in L_0^2(\pi)$ may be written in the form

$$g(x) = \sum_{k=0}^{\infty} g_k(u_{-k}, \dots, u_0),$$

where $x = (\ldots, u_{-2}, u_{-1}, u_0)$, $g_k: \Re^{k+1} \to \Re$ are measurable, $\int_{\Re} g_k(u_{-k}, \ldots, u_0) \times du_{-k} = 0$ for a.e. $(u_{-k+1}, \ldots, u_0)$, $\sum_{k=1}^{\infty} \int g_k^2 \, d\lambda^{k+1} < \infty$, and λ^k denotes k-dimensional Lebesgue measure.

COROLLARY 6. *Relation* (11) *holds if*

$$\sum_{k=1}^{\infty} k \log^{1+\delta}(k) \int g_k^2 \, d\lambda^{k+1} < \infty$$

for some $\delta > 0$, *and* (16) *holds if for some* $\delta > 0$,

$$\sum_{k=1}^{\infty} k^{1+\delta} \int g_k^2 \, d\lambda^{k+1} < \infty.$$

PROOF. If $y = (\ldots, w_{-1}, w_0)$ and $z \in \Re^k$, then

$$g(x, z) - g(y, z) = \sum_{j=k}^{\infty} [g_j(u_{-j+k}, \ldots, u_0, z) - g_j(w_{-j+k}, \ldots, w_0, z)]$$

and, therefore,

$$\int [g(x) - g(y)]^2 \Gamma_\delta^1\{dx\, dy\} \le 4 \sum_{k=1}^{\infty} \sum_{j=k}^{\infty} k^{\delta} \int g_j^2 \, d\lambda^{j+1}.$$

The second assertion of the corollary now follows from routine manipulations, and the first may be established similarly. □

Acknowledgments. Thanks to Persi Diaconis, Bob Keener and Walter Phillip for helpful conversations. Theorem 1 is adapted from Maxwell (1997).

REFERENCES

BHATTACHARRYA, R. and LEE, O. (1988). Asymptotics for a class of Markov processes that are not in general irreducible. *Ann. Probab.* **16** 1333–1347. [Correction (1997) *Ann. Probab.* **25** 1541–1543.]

BILLINGSLEY, P. (1968). *Convergence of Probability Measures*. Wiley, New York.

DEHLING, H., DENKER, M. and PHILLIPP, W. (1986). Central limit theorems for mixing sequences of random variables under minimal conditions. *Ann. Probab.* **14** 1359–1370.

R. DURRETT, and RESNICK, S. (1978). Functional limit theorems for dependent variables. *Ann. Probab.* **6** 829–846.

GORDIN, M. I. and LIFSIC, B. A. (1978). Central limit theorems for stationary Markov processes. *Dokl. Akad. Nauk. SSSR* **239** 766–767.

HALL, P. and HEYDE, C. (1981). *Martingale Limit Theory and Its Applications*. Academic Press, New York.

KIPNIS, C. and VARADHAN, S. R. S. (1986). Central limit theorems for additive functionals of reversible Markov processes and applications to simple excursions. *Comm. Math. Phys.* **104** 1–19.

MAXWELL, M. (1997). Local and global central limit theorems for stationary ergodic sequences. Ph.D. dissertation, Univ. Michigan.

MEYN, S. P. and TWEEDIE, R. L. (1993). *Markov Chains and Stochastic Stability*. Springer, New York.

PELIGRAD, M. (1986). Recent advances in the central limit theorem and its weak invariance principle for mixing sequences of random variables. In *Dependence in Probability and Statistics* (E. Eberlein and M. S. Taqqu, eds.) 193–224. Birkhäuser, Boston.

TOTH, B. (1986). Persistent random walks in random environment. *Probab. Theory Related Fields* **71** 615–625.

WOODROOFE, M. (1992). A central limit theorem for functions of a Markov chain with applications to shifts. *Stochastic. Process. Appl*. **41** 31–42.

DEPARTMENT OF MATHEMATICS AND COMPUTER SCIENCE
SAN JOSE STATE UNIVERSITY
ONE WASHINGTON SQUARE
SAN JOSE, CALIFORNIA 95192-0103
E-MAIL: maxwell@mathcs.sjsu.edu

DEPARTMENT OF STATISTICS
UNIVERSITY OF MICHIGAN
4062 FRIEZE BUILDING
ANN ARBOR, MICHIGAN 48109
E-MAIL: michaelw@umich.edu