Econometrica, Vol. 76, No. 2 (March, 2008), 443–450

COMPUTING THE DISTRIBUTIONS OF ECONOMIC MODELS VIA SIMULATION

BY JOHN STACHURSKI AND VANCE MARTIN[1]

We study a Monte Carlo algorithm for computing marginal and stationary densities of stochastic models with the Markov property, establishing global asymptotic normality and $O_P(n^{-1/2})$ convergence. Asymptotic normality is used to derive error bounds in terms of the distribution of the norm deviation.

KEYWORDS: Markov processes, simulation, ergodicity, numerical methods.

1. INTRODUCTION

WHEN ANALYZING THE DYNAMICS of economic and econometric models, one often wishes to study the marginal and stationary distributions associated with the vector of state variables. For many models no closed form solution for these distributions exists, and numerical methods form the main bridge to quantitative applications. This paper studies one such method, proposed first by Glynn and Henderson (2001).

The problem can be introduced as follows. Let $\mathbb{X} \subset \mathbb{R}^k$ and let $p : \mathbb{X} \times \mathbb{X} \to \mathbb{R}$ be a *density kernel* on $\mathbb{X}$. That is, p is jointly measurable and $p(x, y)\,\mathrm{d}y$ is a density on $\mathbb{X}$ for each $x \in \mathbb{X}$. Taking X_1 as given and recursively drawing

$$X_{t+1} \sim p(X_t, y)\,\mathrm{d}y \quad (t \geq 1)$$

yields a discrete time Markov process $(X_t)_{t\geq 1}$ on $\mathbb{X}$.[2] It is well known that for such a process, the (marginal) distribution of X_t can be represented by a density ψ_t on $\mathbb{X}$ and, moreover, the sequence $(\psi_t)_{t\geq 1}$ satisfies

$$\psi_{t+1}(y) = \int p(x, y)\psi_t(x)\,\mathrm{d}x \quad (y \in \mathbb{X}, t \geq 1). \tag{1}$$

Furthermore, a density ψ_∞ on $\mathbb{X}$ is called *stationary* for the kernel p if

$$\psi_\infty(y) = \int p(x, y)\psi_\infty(x)\,\mathrm{d}x \quad (y \in \mathbb{X}). \tag{2}$$

[1]This project has benefited from the comments of Peter Glynn, Shane Henderson, Sean Meyn, Kazuo Nishimura, Yoshihiko Nishiyama, Roberto Raimondo, a co-editor, an anonymous referee, and participants at the 12th International Conference on Computing in Economics and Finance, Cyprus 2006.

[2]Given X_1 and p, such a process $(X_t)_{t\geq 1}$ exists on some probability space $(\Omega, \mathcal{F}, \mathbb{P})$. Conversely, given a model which defines the random process $(X_t)_{t\geq 1}$ directly, let $p(x, \mathrm{d}y)$ be the conditional distribution of X_{t+1} given $X_t = x$. We require that $p(x, \mathrm{d}y)$ can be represented by a density $p(x, y)\,\mathrm{d}y$ for all $x \in \mathbb{X}$.

It is an equilibrium in the sense that if $X_1 \sim \psi_\infty$, then $X_t \sim \psi_\infty$ for all t, and in fact one can show that $(X_t)_{t\geq 1}$ is (in the strict sense) stationary.

In this paper, we study how to compute numerical approximations to ψ_T (for some given $T \in \mathbb{N}$) and ψ_∞ when analytical expressions are unavailable. Previously a number of techniques have been suggested, including (i) discretization and (ii) simulation combined with histograms or nonparametric kernel density estimates. In what follows we analyze an alternative simulation-based technique which is both intuitively simple and computationally efficient.

To compute ψ_T, Glynn and Henderson (2001) proposed the *marginal density look ahead estimator* (MDLAE) defined by

$$\psi_T^n(y) := \frac{1}{n}\sum_{i=1}^{n} p(X_{T-1}^i, y) \quad (y \in \mathbb{X}), \tag{3}$$

where $(X_{T-1}^i)_{i=1}^n$ is n independent draws of the *lagged* state X_{T-1}. The intuition behind the estimator is straightforward: In view of (1) we have $\mathbb{E}p(X_{T-1}, y) = \psi_T(y)$. As $\psi_T^n(y)$ in (3) is by definition the sample mean of independent observations of $p(X_{T-1}, y)$, it follows that $\psi_T^n(y)$ is unbiased and consistent for $\mathbb{E}p(X_{T-1}, y) = \psi_T(y)$. Moreover, when $\mathbb{E}p(X_{T-1}, y)^2$ is finite, the central limit theorem (CLT) implies that $\psi_T^n(y)$ is also $\sqrt{n}$-consistent for $\psi_T(y)$.[3]

The following example helps illustrate how ψ_T^n can be constructed in applications. Consider a model of the form

$$X_{t+1} = \mu(X_t) + \Sigma U_{t+1}, \quad (U_t)_{t\geq 1} \overset{\text{IID}}{\sim} N(0, \mathbb{I}_k), \tag{4}$$

where $\Gamma := \Sigma\Sigma^\top$ has positive determinant. The corresponding density kernel (i.e., conditional density of X_{t+1} given $X_t = x$) is

$$p(x, y) := \frac{1}{(2\pi)^{k/2}|\Gamma|^{1/2}} \times \exp\left\{-\frac{1}{2}(y - \mu(x))^\top \Gamma^{-1}(y - \mu(x))\right\}. \tag{5}$$

[3]In comparison, the nonparametric kernel density estimator generated from observations of X_T is biased and the error is $O_P((n\delta_n^k)^{-1/2})$, where $\delta_n \to 0$ is the bandwidth and k is the dimension of $\mathbb{X}$ (Yakowitz (1985)). The intuition behind the superior performance of the MDLAE is that the conditional density p in (3) subsumes the role of the kernel in the nonparametric estimator. While p always incorporates the dynamic structure contained in the original model, the nonparametric kernel and bandwidth do not.

An observation of $\psi_T^n(y)$ for this model can be generated using the algorithm

for i in 1 to n **do**
  draw X from the distribution of X_1 (which is given);
  for t in 2 to $T-1$ **do**
    draw $U \sim N(0, \mathbb{I}_k)$;
    set $X \leftarrow \mu(X) + \Sigma U$;
  end
  set $X_{T-1}^i \leftarrow X$;
end
return $\psi_T^n(y) := \frac{1}{n}\sum_{i=1}^n p(X_{T-1}^i, y)$, where p is defined in (5).

Next let us consider approximating the stationary density ψ_∞. Under the conditions on p in Section 3, a unique stationary density exists, and the associated Markov process $(X_t)_{t\geq 1}$ is ergodic in the sense that

$$\lim_{n\to\infty} \frac{1}{n}\sum_{t=1}^n h(X_t) = \int h(x)\psi_\infty(x)\,\mathrm{d}x \quad \text{with probability 1} \tag{6}$$

for any initial X_1 and any ψ_∞-integrable function h.[4] Ergodicity implies that sample moments contain information about ψ_∞. Based on this intuition, Glynn and Henderson (2001) proposed approximating ψ_∞ via the *stationary density look ahead estimator* (SDLAE)

$$\psi_\infty^n(y) := \frac{1}{n}\sum_{t=1}^n p(X_t, y) \quad (y \in \mathbb{X}), \tag{7}$$

where $(X_t)_{t=1}^n$ is a *time series* simulated from p and arbitrary X_1. Condition (6) now implies that with probability 1,

$$\lim_{n\to\infty} \psi_\infty^n(y) = \lim_{n\to\infty} \frac{1}{n}\sum_{t=1}^n p(X_t, y) = \int p(x, y)\psi_\infty(x)\,\mathrm{d}x.$$

In light of (2) this reads $\lim \psi_\infty^n(y) = \psi_\infty(y)$, and hence $\psi_\infty^n(y)$ is consistent for all $y \in \mathbb{X}$, independent of the initial condition X_1. Under some additional mixing conditions, $\psi_\infty^n(y)$ is also $\sqrt{n}$-consistent for $\psi_\infty(y)$.

Returning to the model (4), with a growth restriction on μ (see below) the model is ergodic with unique stationary density ψ_∞. To approximate $\psi_\infty(y)$

[4]That is, any measurable $h: \mathbb{X} \to \mathbb{R}$ with $\int |h(x)|\psi_\infty(x)\,\mathrm{d}x < \infty$.

using the SDLAE, one can apply the algorithm

set $X_1 \leftarrow x$, where x is an arbitrary point in $\mathbb{X}$;
for t in $1, \dots, n-1$ **do** `//` generate $X_{t+1} \sim p(X_t, y)\,\mathrm{d}y$
  draw $U \sim N(0, \mathbb{I}_k)$;
  set $X_{t+1} \leftarrow \mu(X_t) + \Sigma U$;
end
return $\psi_\infty^n(y) := \frac{1}{n}\sum_{t=1}^n p(X_t, y)$, where p is defined in (5).

In this paper we extend Glynn and Henderson's analysis of the look ahead estimator, analyzing *global* convergence of ψ_T^n to ψ_T and ψ_∞^n to ψ_∞. Using a Hilbert space CLT, we show that, when viewed as random functions, the deviations $\psi_T^n - \psi_T$ and $\psi_\infty^n - \psi_\infty$ are asymptotically normally distributed over a certain function space and are $\sqrt{n}$-consistent in the sense that the *norm* deviation is $O_P(n^{-1/2})$.

2. GLOBAL CONVERGENCE, MARGINAL DISTRIBUTION

First let us consider global convergence of ψ_T^n to ψ_T. We use some facts concerning probability in Hilbert space. In what follows, let $\mathcal{H}$ be a separable Hilbert space with inner product $\langle g, h\rangle$ and norm $\|h\|_{\mathcal{H}} := \langle h, h\rangle^{1/2}$. If Y is a random variable taking values in $\mathcal{H}$ and $\mathbb{E}\|Y\|_{\mathcal{H}}$ is finite, we can define $\mathcal{E}Y \in \mathcal{H}$ by the expression $\langle \mathcal{E}Y, h\rangle = \mathbb{E}\langle Y, h\rangle$, all $h \in \mathcal{H}$. This vector $\mathcal{E}Y$ is called the *expectation* of Y and is necessarily unique.[5]

The CLT extends from $\mathbb{R}^k$ to general $\mathcal{H}$ almost unchanged: If $(Y_n)_{n\geq 1}$ is independent and identically distributed (IID) and $\mathbb{E}\|Y_1\|_{\mathcal{H}}^2$ is finite, then $\bar{Y}_n := n^{-1}\sum_{i=1}^n Y_i$ satisfies

$$\sqrt{n}(\bar{Y}_n - \mathcal{E}Y_1) \stackrel{\mathcal{D}}{\to} W \quad (n \to \infty), \tag{8}$$

where the random variable W is centered Gaussian on $\mathcal{H}$.[6] A corollary of this convergence in distribution is that $\|\bar{Y}_n - \mathcal{E}Y_1\|_{\mathcal{H}} = O_P(n^{-1/2})$.

The Hilbert space CLT can be used to study convergence of ψ_T^n to ψ_T. Let X_{T-1} be a random variable distributed according to ψ_{T-1} and let $Y := p(X_{T-1}, \cdot)$ be the random function $y \mapsto p(X_{T-1}, y)$ from $\mathbb{X}$ to $\mathbb{R}$. An immediate consequence of this definition is that if $(X_{T-1}^i)_{i=1}^n$ are IID copies of X_{T-1},

[5]By the Cauchy–Schwarz inequality, $|\mathbb{E}\langle Y, h\rangle| \leq \mathbb{E}\|Y\|_{\mathcal{H}}\|h\|_{\mathcal{H}}$, and since $\mathbb{E}\|Y\|_{\mathcal{H}}$ is finite, $h \mapsto \mathbb{E}\langle Y, h\rangle$ is a bounded linear functional on $\mathcal{H}$. By the Riesz representation theorem, to such a functional there corresponds a vector $\mathcal{E}Y \in \mathcal{H}$ satisfying $\langle \mathcal{E}Y, h\rangle = \mathbb{E}\langle Y, h\rangle$, $h \in \mathcal{H}$. This $\mathcal{E}Y$ is defined to be the expectation of Y. In the present context all standard notions of vector-valued integration coincide (cf., e.g., Bosq (2000)).

[6]W is called centered Gaussian on $\mathcal{H}$ if, for every $h \in \mathcal{H}$, the real-valued random variable $\langle W, h\rangle$ has Gaussian distribution $N(0, \sigma_h^2)$ on $\mathbb{R}$ for some $\sigma_h^2 \geq 0$.

then the sample mean

$$\bar{Y}_n := \frac{1}{n}\sum_{i=1}^{n} Y_i = \frac{1}{n}\sum_{i=1}^{n} p(X_{T-1}^i, \cdot) \tag{9}$$

is precisely ψ_T^n. Our asymptotic normality proof applies the CLT in (8) to $\bar{Y}_n = \psi_T^n$ in (9).

To employ the CLT in (8), three steps are necessary, the details of which are deferred to the Appendix. The first step is to ensure that $Y = p(X_{T-1}, \cdot)$ does in fact take values in a separable Hilbert space; in particular,

$$\mathcal{H} = L_2(\mathbb{X}) := \left\{\text{all measurable } h : \mathbb{X} \to \mathbb{R} \text{ s.t.} \int h(x)^2 \,\mathrm{d}x < \infty\right\}$$

with inner product $\langle g, h\rangle = \int gh$. This is done by placing a restriction on p in Theorem 1 below. The second step is to show that the moment condition $\mathbb{E}\|Y\|^2 < \infty$ is satisfied, where $\|\cdot\|$ is the norm on $L_2(\mathbb{X})$. The third step is to show that the expectation $\mathcal{E}Y$ of Y is ψ_T, in which case we have

$$\sqrt{n}(\bar{Y}_n - \mathcal{E}Y) = \sqrt{n}(\psi_T^n - \psi_T) \tag{10}$$

and the CLT in (8) can be applied:

THEOREM 1: *Let $(X_{T-1}^i)_{i=1}^n$ be IID copies of X_{T-1} and let ψ_T^n be the MDLAE. If there exists a ψ_{T-1}-integrable function $V : \mathbb{X} \to \mathbb{R}$ such that*

$$\int p(x, y)^2 \,\mathrm{d}y \le V(x) \quad (x \in \mathbb{X}), \tag{11}$$

then $\sqrt{n}(\psi_T^n - \psi_T)$ converges in distribution to a centered Gaussian random variable W taking values in $L_2(\mathbb{X})$.[7]

As a consequence we obtain the rate $\|\psi_T^n - \psi_T\| = O_P(n^{-1/2})$.

3. GLOBAL CONVERGENCE, STATIONARY DISTRIBUTION

Next we consider convergence of the SDLAE ψ_∞^n in (7) to ψ_∞. As for the case of local convergence (i.e., $\psi_\infty^n(y) \to \psi_\infty(y)$ for fixed y), global convergence of ψ_∞^n to ψ_∞ requires a form of ergodicity. We suppose that p is V-uniformly ergodic (V-UE); namely, there exists a measurable function $V : \mathbb{X} \to [1, \infty)$ and positive constants $\alpha < 1$ and $R < \infty$ with

$$\sup_{|h|\le V}\left|\int h(y) p^t(x, y)\,\mathrm{d}y - \int h(y)\psi_\infty(y)\,\mathrm{d}y\right| \le \alpha^t R V(x)$$

for all $x \in \mathbb{X}$ and all $t \ge 1$. Here p^t refers to the tth order kernel: $p^t(x, \cdot)$ is

[7]For example, if $x \mapsto \int p(x, y)^2\,\mathrm{d}y$ is bounded on $\mathbb{X}$, then the conditions of the theorem are always satisfied.

the density of X_{k+t} when $X_k = x$.[8] Thus, $\int h(y) p^t(x, y)\, dy$ is the expectation of $h(X_{t+1})$ conditional on $X_1 = x$.

V-UE implies that $\int h(y) p^t(x, y)\, \mathrm{d}y$ converges geometrically to the expectation of h with respect to the stationary distribution. It also implies total variation (and hence L_1) convergence of $p^t(x, \cdot)$ to ψ_∞, as well as uniqueness of ψ_∞ and ergodicity as in (6).[9]

The V-UE property is closely related to geometric ergodicity, and sufficient conditions are well understood. For example, the model given by (4) and (5) is V-UE whenever μ satisfies

$$\exists a \in [0, 1) \text{ and } b \in \mathbb{R}_+ \text{ s.t. } \|\mu(x)\| \leq a\|x\| + b \quad (x \in \mathbb{X}) \tag{12}$$

for some norm $\|\cdot\|$ on $\mathbb{X}$. Kristensen (2006, Theorem 2) gave a useful set of sufficient conditions for geometric ergodicity, which he applied to linear and nonlinear autoregressive moving average, random coefficients and generalized autoregressive conditional heteroskedasticity models. These conditions are, in fact, sufficient for the V-UE property.

With some modifications, the Hilbert space CLT in (8) can be used to prove asymptotic normality of the SDLAE. Let

$$L_2(\mathbb{X}, \psi_\infty) := \left\{ \text{all measurable } h : \mathbb{X} \to \mathbb{R} \text{ s.t. } \int h(x)^2 \psi_\infty(x)\, \mathrm{d}x < \infty \right\},$$

let $\langle g, h\rangle_{\psi_\infty} = \int g(x)h(x)\psi_\infty(x)\, \mathrm{d}x$ be the inner product on $L_2(\mathbb{X}, \psi_\infty)$ and let $\|\cdot\|_{\psi_\infty}$ denote the norm. Adding mild restrictions to p (see below), the densities $p(x, \cdot)$, ψ_∞^n, and ψ_∞ all take values in $L_2(\mathbb{X}, \psi_\infty)$.

Now let $(X_t)_{t\geq 1}$ be a time series generated by p and let Y_t be the $L_2(\mathbb{X}, \psi_\infty)$-valued random variable $p(X_t, \cdot)$. It follows that the sample mean $\bar{Y}_n$ is precisely ψ_∞^n. As discussed in the Appendix, if $(X_t)_{t\geq 1}$ is stationary, then the expectation $\mathcal{E}Y_1 = \mathcal{E}p(X_1, \cdot)$ is equal to ψ_∞, which yields

$$\sqrt{n}(\bar{Y}_n - \mathcal{E}Y_1) = \sqrt{n}(\psi_\infty^n - \psi_\infty). \tag{13}$$

The Hilbert space CLT in (8) does not immediately apply, as $(Y_t)_{t\geq 1}$ is now a correlated process. However, it is known that for Hilbert-space-valued functions of V-UE processes, the CLT continues to hold (Stachurski (2006)). This gives the foundations of the following result:

[8]The kernels are defined by $p^1 = p$ and $p^{t+1}(x, y) = \int p(x, z)p^t(z, y)\, \mathrm{d}z$.

[9]In addition, V-UE implies aperiodicity, irreducibility, and geometric mixing. Interested readers should consult Meyn and Tweedie (1993, Chap. 16).

THEOREM 2: *Let $(X_t)_{t\geq 1}$ be a Markov process on $\mathbb{X}$ with V-UE density kernel p. If*

$$\int p(x,y)^2\psi_\infty(y)\,\mathrm{d}y \leq V(x) \quad (x\in\mathbb{X}), \tag{14}$$

then $\sqrt{n}(\psi_\infty^n-\psi_\infty)$ converges in distribution to a centered Gaussian random variable W on $L_2(\mathbb{X},\psi_\infty)$ with covariance function

$$\begin{aligned}\Gamma(y,y') &= \int p(x,y)p(x,y')\psi_\infty(x)\,\mathrm{d}x-\psi_\infty(y)\psi_\infty(y')\\ &\quad+\sum_{t\geq 1}^{\infty}\left[\int p(x,y)p^{t+1}(x,y')\psi_\infty(x)\,\mathrm{d}x-\psi_\infty(y)\psi_\infty(y')\right]\\ &\quad+\sum_{t\geq 1}^{\infty}\left[\int p(x,y')p^{t+1}(x,y)\psi_\infty(x)\,\mathrm{d}x-\psi_\infty(y)\psi_\infty(y')\right].\end{aligned}$$

The covariance function $\Gamma(y,y')$ can be viewed as the infinite dimensional analogue of a variance–covariance matrix.[10]

From Theorem 2, we obtain the asymptotic distribution of the error, measured in terms of the norm distance between ψ_∞^n and ψ_∞.

COROLLARY 1: *Under the hypotheses of Theorem 2, we have*

$$n\|\psi_\infty^n-\psi_\infty\|_{\psi_\infty}^2 \xrightarrow{\mathcal{D}} \sum_{\ell\geq 1}^{\infty}\lambda_\ell Z_\ell^2 \quad (n\to\infty),$$

where $(\lambda_\ell)_{\ell\geq 1}$ are the eigenvalues of the covariance function Γ in Theorem 2, and $(Z_\ell)_{\ell\geq 1}$ are independent standard normal.[11]

Here $n\|\psi_\infty^n-\psi_\infty\|_{\psi_\infty}^2$ is the square of $\|\sqrt{n}(\psi_\infty^n-\psi_\infty)\|_{\psi_\infty}$, and Corollary 1 is an infinite dimensional version of the well-known fact that if $Y\sim N(0,C)$ in $\mathbb{R}^k$, then $\|Y\|^2$ has the same distribution as $\sum_{\ell=1}^k\lambda_\ell Z_\ell^2$, where $\|\cdot\|$ is the norm on $\mathbb{R}^k$, λ_ℓ is the ℓth eigenvalue of C, and $(Z_\ell)_{\ell=1}^k$ are IID and $N(0,1)$. An immediate consequence of Corollary 1 is global $\sqrt{n}$-consistency. In particular, $\|\psi_\infty^n-\psi_\infty\|_{\psi_\infty}=O_P(n^{-1/2})$.

A final remark on Theorem 2 is that if p is V-UE and bounded, then the conclusion of the theorem holds without (14). For example, p in (5) satisfies all the conditions of the theorem when (12) holds.

[10]Note that, in fact, we do not need $X_1\sim\psi_\infty$. The result holds for $X_1=x\in\mathbb{X}$, where x is arbitrary. This is important for implementation. It means that when simulating $(X_t)_{t\geq 1}$ to construct ψ_∞^n one can start at any $x\in\mathbb{X}$.

[11]More correctly, $(\lambda_\ell)_{\ell\geq 1}$ are the eigenvalues of the covariance operator C defined by the function Γ. For $h\in L_2(\mathbb{X},\psi_\infty)$, Ch is given by $Ch(y'):=\int\Gamma(y,y')h(y)\psi_\infty(y)\,\mathrm{d}y$.

APPENDIX

Regarding Theorem 1, to employ the CLT in (8), we must establish (i) that $Y = p(X_{T-1}, \cdot)$ takes values $L_2(\mathbb{X})$, (ii) that $\mathbb{E}\|Y\|^2 < \infty$, and (iii) that $\mathcal{E}Y = \psi_T$. In fact, (i) is immediate from (11), as is (ii) because

$$\|Y\|^2 = \int p(X_{T-1}, y)^2 \, \mathrm{d}y \leq V(X_{T-1})$$

and $\mathbb{E}V(X_{T-1})$ is finite by assumption. To prove (iii) we must show that $\langle \psi_T, h \rangle = \mathbb{E}\langle Y, h \rangle$ for any $h \in L_2(\mathbb{X})$. Since $\psi_T(y) = \mathbb{E}p(X_{T-1}, y)$, for such an h we have

$$\langle \psi_T, h \rangle := \int \psi_T(y) h(y) \, \mathrm{d}y = \int \mathbb{E}p(X_{T-1}, y) h(y) \, \mathrm{d}y.$$

On the other hand, an application of Fubini's theorem gives

$$\mathbb{E}\langle Y, h \rangle = \mathbb{E} \int p(X_{T-1}, y) h(y) \, \mathrm{d}y = \int \mathbb{E}p(X_{T-1}, y) h(y) \, \mathrm{d}y.$$

Hence $\langle \psi_T, h \rangle = \mathbb{E}\langle Y, h \rangle$ for all $h \in L_2(\mathbb{X})$ and $\mathcal{E}Y = \psi_T$ as claimed.

Regarding Theorem 2, the fact that $\mathcal{E}Y_1 = \mathcal{E}p(X_1, \cdot) = \psi_\infty$ when $(X_t)_{t \geq 1}$ is stationary (and hence $X_1 \sim \psi_\infty$) can be proved in an almost identical manner to the proof of (iii) above. The sufficiency of (14) and the expression for Γ follow directly from Stachurski (2006, Theorem 3.1).

Institute of Economic Research, Kyoto University, Yoshida honmachi, Sakyo-ku, Kyoto 606-8501, Japan; john@keir.kyoto-u.ac.jp

and

Dept. of Economics, University of Melbourne, Melbourne VIC 3010, Australia; vance@unimelb.edu.au.

Manuscript received November, 2005; final revision received August, 2007.

REFERENCES

BOSQ, D. (2000): *Linear Processes in Function Space*. New York: Springer-Verlag. [446]

GLYNN, P. W., AND S. G. HENDERSON (2001): "Computing Densities for Markov Chains via Simulation," *Mathematics of Operations Research*, 26, 375–400. [443-445]

KRISTENSEN, D. (2006): "Geometric Ergodicity of a Class of Markov Chains with Applications to Time Series Models," Mimeo, University of Wisconsin. [448]

MEYN, S. P., AND R. L. TWEEDIE (1993): *Markov Chains and Stochastic Stability*. London: Springer-Verlag. [448]

STACHURSKI, J. (2006): "A Hilbert Space Central Limit Theorem for Geometrically Ergodic Markov Chains," Mimeo, Kyoto University. [448,450]

YAKOWITZ, S. J. (1985): "Nonparametric Density Estimation, Prediction and Regression for Markov Sequences," *Journal of the American Statistical Association*, 80, 215–221. [444]