APPLICATIONES MATHEMATICAE
38,1 (2011), pp. 93–105

Samir Benaissa, Abbes Rabhi and Belaid Mechab (Sidi Bel Abbes)

ASYMPTOTIC NORMALITY OF THE KERNEL ESTIMATE FOR THE MARKOVIAN TRANSITION OPERATOR

Abstract. We build a kernel estimator of the Markovian transition operator as an endomorphism on L^1 for some discrete time continuous states Markov processes which satisfy certain additional regularity conditions. The main result deals with the asymptotic normality of the kernel estimator constructed.

1. Introduction. Let the real-valued random variables $(X_n)_{n\in\mathbb{N}}$ be defined on a probability space $(\Omega, \mathcal{A}, P)$ and suppose they constitute a strictly stationary and homogeneous Markov process satisfying some additional requirements. Let $\mathcal{F}_a^b$ denote the σ-field generated by the random variables $X_a, X_{a+1}, \ldots, X_b$. For any two σ-fields $\mathcal{A}, \mathcal{B} \subset F$ put

$$\phi(\mathcal{A}, \mathcal{B}) = \sup\{|P(B|A) - P(B)| : P(A) \neq 0,\ A \in \mathcal{A},\ B \in \mathcal{B}\}.$$

The *mixing coefficient* of the sequence $\{X_n, n \geq 1\}$ is defined as usual:

$$\phi(n) = \sup\nolimits_{k\geq 1} \phi(\mathcal{F}_1^k, \mathcal{F}_{k+n}^\infty).$$

The sequence $\{X_n, n \geq 1\}$ is called ϕ-*mixing* or *uniformly mixing* if $\phi(n) \to 0$.

Suppose the initial law ν and the one-step transition distribution have probability density function $f(\cdot)$ and $P(\cdot,\cdot)$ respectively, relative to Lebesgue measure μ.

We define the one-step transition operator $H : L^1(\nu) \to L^1(\nu)$ by

$$\begin{aligned} &g \mapsto H(g) : (\mathbb{R}, \mathcal{B}_\mathbb{R}) \to\ (\mathbb{R}, \mathcal{B}_\mathbb{R}), \\ &x \mapsto Hg(x) = E(g(X_{t+1}) \mid X_t = x) = \int_{-\infty}^{+\infty} g(y)P(x,y)\,dy, \end{aligned} \tag{1}$$

which is an idempotent operator on $L^1(\nu)$.

2010 *Mathematics Subject Classification*: 62M05, 62M15.
Key words and phrases: asymptotic normality, kernel estimate, Markov processes, transition operator.

DOI: 10.4064/am38-1-7

The one-step transition operator H has been studied by Györfi et al. [8]. The authors treated this functional parameter as an operator from autoregression, and obtained almost complete convergence under the Doeblin condition. In this paper, we study this parameter as an endomorphism on L^1. For this we estimate the quantities $P(x,y)$ nonparametrically by

$$P_n(x,y) = \frac{1}{h_n} \frac{\sum_{i=1}^n K\big(\frac{x-X_i}{h_n}\big) K\big(\frac{y-X_{i+1}}{h_n}\big)}{\sum_{i=1}^n K\big(\frac{x-X_i}{h_n}\big)} \tag{2}$$

where $0 < h_n \to 0$ as $n \to \infty$ (see, e.g., Youndjé [19]). Then the natural nonparametric estimator of the one-step transition operator H is

$$\forall g \in L^1(\nu) \qquad H_n\, g(x) = \frac{1}{h_n} \int_{-\infty}^{\infty} \frac{\sum_{i=1}^n K\big(\frac{x-X_i}{h_n}\big) K\big(\frac{y-X_{i+1}}{h_n}\big)}{\sum_{i=1}^n K\big(\frac{x-X_i}{h_n}\big)} g(y)\, dy. \tag{3}$$

Here K is a nonnegative function and $(h_n)_{n\in\mathbb{N}}$ is a nonnegative sequence that converges to zero as n tends to infinity. For further use, let $K_{h_n}(\cdot) = (1/h_n)K(\cdot/h_n)$.

We note that if g is continuous and K is a Rosenblatt kernel, according to the Bochner lemma, we have $\lim_{h_n\to 0} K_{h_n} * g = g$ (see Bosq and Lecoutre [3]). So, for all fixed x with $\lim_{h_n\to 0}(K_{h_n} * g)(x) = g(x)$, we can build another estimator of H defined by

$$\tilde{H}_n g(x) = \frac{\sum_{i=1}^n K_{h_n}(x - X_i) g(X_{i+1})}{\sum_{i=1}^n K_{h_n}(x - X_i)}. \tag{4}$$

The estimator $\tilde{H}_n$ was introduced by Collomb and Doukhan [4] and was used in forecast (see Ferraty, Goïa and Vieu [6] for the most recent references).

Historically, Roussas [15] was one of the first authors who tackled the problem of estimation by using observation with Markovian character. He established the convergence of the kernel estimate of the transition density using the L^2 norm. Other authors were interested in the functional estimation Markovian process by treating the case of the stationary density of a stationary Markov chain satisfying the G_2 condition (see Rosenblatt [14]). Under Doeblin's condition, Gillert and Wartenberg [7] and Liebscher [11] studied the asymptotic behaviour of the mean square error of kernel density estimate for the stationary density of a Markov chain. Recently, Laksaci and Yousfate [9] considered the L_p-convergence of the kernel estimate of the transition operator density. Among the numerous papers concerning the asymptotic normality and estimation for stationary mixing sequences, we only mention Collomb and Doukhan [4], Yakowitz [18], Ango Nzé and Rios [1], Bosq [2], Louani and Ouled-Said [10], Liebscher [12], and Delecroix et al. [5].

The goal of this paper is to establish the asymptotic normality of a kernel estimator (3) of the one-step transition operator under the G_1 condition. We note that the main difficulty in this context is that the asymptotic normality of the conditional density at all points is not enough to obtain the asymptotic normality of our estimator. It is worth noting that a particular case (with $g(y) = \mathcal{I}_{]-\infty,z]}(y)$) has been studied by Roussas [16]. The main feature of our approach is to build an estimator and derive the asymptotic normality for a class of parameters. Note that $Hg(x)$ can be identified with some useful statistics where g is known. For example, if $g(y) = \mathcal{I}_{]-\infty,z]}(y)$, then $Hg(x)$ is identified with a one-step transition distribution function, and if $g(y) = e^{ity}$, then $Hg(x)$ represents the one-step transition characteristic function.

The paper is organized as follows. After establishing the notation and listing the required assumptions in Section 2, we state our main results in Section 3. In the last section we list some preliminary results needed to prove the main result.

2. Notation and assumptions. We denote by $\varphi_m(x_1, \ldots, x_m)$ the joint density of the random variables $X_1, \ldots, X_m$. We need the following assumptions:

(H.1) The process $(X_k)_{k\in\mathbb{N}}$ satisfies the G_1 *condition*:

$$\sup_{\|g\|_1\le 1} \frac{\|Hg^s\|_1}{\|g\|_1} \to 0 \quad \text{as } s \to \infty,$$

where $Hg^s(x) = E(g(X_{t+s}) \mid X_t = x)$.

(H.2) K is a p.d.f. defined on $\mathbb{R}$ such that:
- (a) $|x|K(x) \to 0$ as $|x| \to \infty$,
- (b) $\int xK(x)\,dx = 0$ and $\int x^2K(x)\,dx < \infty$.

(H.3) h_n is a sequence of real numbers such that $nh_n \to \infty$ and $nh_n^5 \to 0$ whenever $n \to \infty$.

(H.4)
- (a) $f := \varphi_1$ is bounded,
- (b) $f(\cdot)$ has a continuous and bounded second order derivative,
- (c) $\varphi_2(\cdot,\cdot)$ has continuous second order partial derivatives, denoted by $\varphi''_{2ij}(\cdot,\cdot)$, $i,j = 1,2$, such that

$$\int \varphi''_{2ij}(x,y)\,dy \le C,$$

- (d) $|\varphi_4(x_1,x_2,x_3,x_4) - \varphi_2(x_1,x_2)\varphi_2(x_3,x_4)| \le C$ for $x_1,x_2,x_3,x_4 \in \mathbb{R}$.

3. Main results. For convenience we set

(5) $$\forall x \in \mathbb{R} \quad G_n(x) = \frac{1}{n}\sum_{i=1}^{n} K_{h_n}(x - X_i)\int K_{h_n}(y - X_{i+1})g(y)\,dy,$$

and

$$\forall x \in \mathbb{R} \quad f_n(x) = \frac{1}{n}\sum_{i=1}^{n} K_{h_n}(x - X_i).$$

It follows that $H_n g(x) = G_n(x)/f_n(x)$.

To study asymptotic normality of $H_n g(x)$, we show that $G_n(x) - E(G_n(x))$ suitably normalized is asymptotically normally distributed and that $f_n(x)$ (resp. $E(G_n(x))/f_n(x)$) converges in probability to a constant.

Our main result is given in the following theorem:

THEOREM 3.1. *Let the assumptions* (H.1)–(H.4) *be satisfied. Then for any $x \in \mathbb{R}$, and g integrable with respect to Lebesgue measure,*

$$\sqrt{nh_n}(H_n g(x) - Hg(x)) \to N(0, \sigma(x)),$$

where

$$\sigma^2(x) := Hg^2(x)\int K^2(z)\,dz.$$

4. Some preliminary results. In this section, we present several intermediate results used for the proof of the main result.

PROPOSITION 4.1. *Let the assumptions* (H.1), (H.2)(b), (H.3), *and* (H.4)(d) *be satisfied. Then for any $x \in \mathbb{R}$ and g integrable with respect to Lebesgue measure,*

(6) $$nh_n \operatorname{Var} G_n(x)) \to \sigma_1(x) := Hg^2(x) f(x)\int K^2(z)\,dz.$$

Proof. We put

$$K_i(x) = K\left(\frac{x - X_i}{h_n}\right).$$

By (5) we have

$$nh_n \operatorname{Var} G_n(x) = nh_n E[G_n(x) - EG_n(x)]^2 = I_{1n}(x) + \frac{1}{nh_n^3} I_{2n}(x)$$

where

$$I_{1n}(x) = \frac{1}{h_n^3}\operatorname{Var}\left[K_1(x)\int K_2(y)g(y)\,dy\right]$$

and

(7) $$I_{2n}(x) = \sum_{1\le i<j\le n}\operatorname{Cov}\Big(K_i(x)\int K_{i+1}(y)g(y)\,dy, K_j(x)\int K_{j+1}(y)g(y)\,dy\Big).$$

a) By a suitable application of Bochner's theorem and the dominated convergence theorem we show

$$I_{1n}(x) \to f(x)Hg^2(x)\int K^2(z)\,dz. \tag{8}$$

Indeed,

$$\begin{aligned}
&I_{1n}(x)\\
&\quad= \frac{1}{h_n^3}\int K^2\Big(\frac{x-x_1}{h_n}\Big)\Big[\int K\Big(\frac{y-x_2}{h_n}\Big)g(y)\,dy\Big]^2\varphi_2(x_1,x_2)\,dx_1\,dx_2\\
&\qquad - h_n\Big(\frac{1}{h_n^2}\int K\Big(\frac{x-x_1}{h_n}\Big)\Big[\int K\Big(\frac{y-x_2}{h_n}\Big)g(y)\,dy\Big]\varphi_2(x_1,x_2)\,dx_1\,dx_2\Big)^2\\
&\quad= \int K^2(z_1)\Big[\int K(z_2)g(x_2+h_nz_2)\,dz_2\Big]^2\varphi_2(x-h_nz_1,x_2)\,dz_1\,dx_2\\
&\qquad - h_n\Big(\int K(z_1)\Big[\int K(z_2)g(x_2+h_nz_2)\,dz_2\Big]\varphi_2(x-h_nz_1,x_2)\,dz_1\,dx_2\Big)^2.
\end{aligned}$$

Now, the first term on the right hand side tends to $f(x)Hg^2(x)\int K^2(z)\,dz$, and the second term tends to zero.

b) Using hypothesis (H.4)(d), it follows that

$$\Big|\operatorname{Cov}\Big[K_i(x)\Big(\int K_{i+1}(y)g(y)\,dy\Big), K_j(x)\Big(\int K_{j+1}(y)g(y)\,dy\Big)\Big]\Big| \le Ch_n^4. \tag{9}$$

Indeed,

$$\begin{aligned}
&\operatorname{Cov}\Big[K_i(x)\Big(\int K_{i+1}(y)g(y)\,dy\Big), K_j(x)\Big(\int K_{j+1}(y)g(y)\,dy\Big)\Big]\\
&= \int K\Big(\frac{x-x_1}{h_n}\Big)\Big(\int K\Big(\frac{y-x_2}{h_n}\Big)g(y)\,dy\Big)\\
&\quad\times K\Big(\frac{x-x_3}{h_n}\Big)\Big(\int K\Big(\frac{y-x_4}{h_n}\Big)g(y)\,dy\Big)\varphi_4(x_1,x_2,x_3,x_4)\,dx_1\,dx_2\,dx_3\,dx_4\\
&\quad-\int K\Big(\frac{x-x_1}{h_n}\Big)\Big(\int K\Big(\frac{y-x_2}{h_n}\Big)g(y)\,dy\Big)\varphi_2(x_1,x_2)\,dx_1\,dx_2\\
&\quad\times\int K\Big(\frac{x-x_3}{h_n}\Big)\Big(\int K\Big(\frac{y-x_4}{h_n}\Big)g(y)\,dy\Big)\varphi_2(x_3,x_4)\,dx_3\,dx_4\\
&= h_n^4\int K(z_1)\Big(\int K(x_2)g(x_2+h_nz_2)\,dy\Big)K(z_3)\\
&\quad\times\Big(\int K(z_4)g(x_4+h_nz_4)\,dy\Big)\varphi_4(x-h_nz_1,x_2,x-h_nz_3,x_4)dz_1\,dx_2\,dz_3\,dx_4\\
&\quad- h_n^4\int K(z_1)\Big(\int K(z_2)g(x_2+h_nz_2)\,dz_2\Big)K(z_3)\\
&\quad\times\Big(\int K(z_4)g(x_4+h_nz_4)\,dz_4\Big)\varphi_2(x-h_nz_1,x_2)\\
&\quad\times\varphi_2(x-h_nz_3,x_4)\,dz_1\,dx_2\,dz_3\,dx_4
\end{aligned}$$

$$= h_n^4 \iint \Big[K(z_1)K(z_3)\Big(\int K(z_2)g(x_2+h_nz_2)\,dz_2\Big)\Big(\int K(z_4)g(x_4+h_nz_4)\,dz_4\Big)\Big]$$
$$\times \varphi_4(x-h_nz_1, x_2, x-h_nz_3, x_4)$$
$$-\varphi_2(x-h_nz_1,x_2)\varphi_2(x-h_nz_3,x_4)\,dz_1\,dx_2\,dz_3\,dx_4.$$

Since K and g are integrable functions, we have (9).

In the following, we use the technique developed by Masry [13]. Define

$$S_1 = \{(i,j) : 1 \le j-i \le m_n\}, \qquad S_2 = \{(i,j) : m_n+1 \le j-i \le n-1\},$$

where $(m_n)_n$ is any sequence of positive integers such that $m_n \to \infty$ and $m_nh_n \to 0$. We can take $m_n = [1/h_n^{1-\lambda}]$, $0<\lambda<1$, where $[x]$ indicates the integral part of x. Next, let $J_{1,n}(x)$ and $J_{2,n}(x)$ be the sums of the covariances over S_1 and S_2, respectively. Then

$$J_{1,n}(x) \le Ch_n^4 nm_n.$$

To bound the sum over S_2, we use a moment inequality for ϕ-mixing (see Roussas and Ioannides [17, p. 64]) (since $(X_k)_{k\in\mathbb{N}}$ satisfy the (H.1) condition):

$$\operatorname{Cov}\bigg[K\Big(\frac{x-X_i}{h_n}\Big)\Big(\int K\Big(\frac{y-X_{i+1}}{h_n}\Big)g(y)\,dy\Big),$$
$$K\Big(\frac{x-X_j}{h_n}\Big)\Big(\int K\Big(\frac{y-X_{j+1}}{h_n}\Big)g(y)\,dy\Big)\bigg]$$
$$\le 2\phi^{1/2}(j-i)\bigg\|K\Big(\frac{x-X_i}{h_n}\Big)\Big(\int K\Big(\frac{y-X_{i+1}}{h_n}\Big)g(y)\,dy\Big)$$
$$-E\bigg[K\Big(\frac{x-X_i}{h_n}\Big)\Big(\int K\Big(\frac{y-X_{i+1}}{h_n}\Big)g(y)\,dy\Big)\bigg]\bigg\|_2$$
$$\times\bigg\|K\Big(\frac{x-X_j}{h_n}\Big)\Big(\int K\Big(\frac{y-X_{j+1}}{h_n}\Big)g(y)\,dy\Big)$$
$$-E\bigg[K\Big(\frac{x-X_j}{h_n}\Big)\Big(\int K\Big(\frac{y-X_{j+1}}{h_n}\Big)g(y)\,dy\Big)\bigg]\bigg\|_2.$$

We define the norm $\|\cdot\|_2$ by

$$\|X\|_2 = \Big(\int_\Omega |X|^2\,dP\Big)^{1/2}.$$

It is clear that

$$\bigg\|K\Big(\frac{x-X_j}{h_n}\Big)\Big(\int K\Big(\frac{y-X_{j+1}}{h_n}\Big)g(y)\,dy\Big)$$
$$-E\bigg[K\Big(\frac{x-X_j}{h_n}\Big)\Big(\int K\Big(\frac{y-X_{j+1}}{h_n}\Big)g(y)\,dy\Big)\bigg]\bigg\|_2 = h_n^3I_{1n} \le h_n^3C'(x).$$

Therefore

$$\sum_{S_2} \mathrm{Cov}\left[K\left(\frac{x-X_i}{h_n}\right)\left(\int K\left(\frac{y-X_{i+1}}{h_n}\right)g(y)\,dy\right),\right.$$
$$\left. K\left(\frac{x-X_j}{h_n}\right)\left(\int K\left(\frac{y-X_{j+1}}{h_n}\right)g(y)\,dy\right)\right]$$
$$\le N(x)h_n^3\sum_{l=m_n}^{n-1}\phi^{1/2}(l).$$

In this case, there exist $s \in (0,\infty)$ and $\rho \in (0,1)$ such that $\phi(l) \le s\rho^l$. Finally,

$$\frac{1}{nh_n^3}\sum_{1\le i<j\le n} \mathrm{Cov}\left[K\left(\frac{x-X_i}{h_n}\right)\left(\int K\left(\frac{y-X_{i+1}}{h_n}\right)g(y)\,dy\right),\right.$$
$$\left. K\left(\frac{x-X_j}{h_n}\right)\left(\int K\left(\frac{y-X_{j+1}}{h_n}\right)g(y)\,dy\right)\right]$$
$$\le C(x)\left[m_nh_n+\frac{1}{n}\sum_{l=m_n}^{n}\rho^l\right].$$

Then

$$\frac{1}{nh_n^3}I_{2n}(x)\to 0. \tag{10}$$

PROPOSITION 4.2. *Under the assumptions of Proposition* 4.1, *for* $x\in\mathbb{R}$,

$$\sqrt{nh_n}(G_n(x)-EG_n(x))\to N(0,\sigma_1(x)). \tag{11}$$

Proof. We can write

$$\sqrt{nh_n}(G_n(x)-EG_n(x))=\frac{1}{\sqrt{n}}\sum_{i=1}^{n}L_i(x)$$

where

$$L_i(x)=\frac{1}{\sqrt{h_n^3}}\left(K\left(\frac{x-X_i}{h_n}\right)\left(\int K\left(\frac{y-X_{i+1}}{h_n}\right)g(y)\,dy\right)\right.$$
$$\left.-E\left[K\left(\frac{x-X_i}{h_n}\right)\left(\int K\left(\frac{y-X_{i+1}}{h_n}\right)g(y)\,dy\right)\right]\right).$$

In order to establish this result, we use Doob's technique: the sum $\sum_{i=1}^{n}L_i(x)$ is split up into large (M_n) blocks and small (N_n) blocks where M_n and N_n are positive integers tending to infinity and $M_n+N_n\le n$, and k_n is the largest integer for which $k_n(M_n+N_n)\le n$ (we can take $M_n=[n^\delta], N_n=[n^{\delta'}]$ and $k_n=[n/(M_n+N_n)]$ where $0<\delta'<\delta<1$). Set

$$S_n(x) = \sum_{j=0}^{k_n} Y_{nj}(x) \quad \text{with} \quad Y_{nj}(x) = \sum_{i=j(M_n+Nn)+1}^{j(M_n+N_n)+M_n} L_i(x),$$

$$T_n(x) = \sum_{j=0}^{k_n} Y'_{nj}(x) \quad \text{with} \quad Y'_{nj}(x) = \sum_{i=j(M_n+Nn)+M_n+1}^{(j+1)(M_n+N_n)} L_i(x),$$

$$T'_n(x) = \sum_{j=(M_n+N_n)k_n+1}^{n} L_j(x),$$

so that

$$\sum_{i=1}^{n} L_i(x) = S_n(x) + T_n(x) + T'_n(x).$$

We will show

Proposition 4.3. *Under the assumptions of Proposition* 4.2, *for* $x \in \mathbb{R}$,

$$\frac{1}{n}[E(T_n^2(x)) + E(T_n'^2(x))] \to 0 \tag{12}$$

and

$$\frac{1}{\sqrt{n}} S_n(x) \to N(0, \sigma_1(x)). \tag{13}$$

The following result is easily established:

Lemma 4.1. *Under the assumptions of Proposition* 4.3, *for* $x \in \mathbb{R}$,

(a) $\operatorname{Var} L_i(x) \le C(x)$,

(b) $n^{-1} \sum_{j=0}^{k_m} \operatorname{Var} Y'_{nj}(x) \to 0$.

Proof. (a) Indeed,

$$\begin{aligned}
&\operatorname{Var}(L_1(x)) \\
&\quad = \frac{1}{h_n^3}\left[E\left(K^2\left(\frac{x-X_1}{h_n}\right)\left[\int K\left(\frac{y-X_2}{h_n}\right) g(y)\,dy\right]^2\right)\right] \\
&\qquad - \frac{1}{h_n^3}\left[E\left(K\left(\frac{x-X_1}{h_n}\right)\left[\int K\left(\frac{y-X_2}{h_n}\right) g(y)\,dy\right]\right)\right]^2 \\
&\quad = \frac{1}{h_n^3}\int K^2\left(\frac{x-z_1}{h_n}\right)\left(\int K\left(\frac{y-z_2}{h_n}\right) g(y)\,dy\right)^2 \varphi_2(z_1,z_2)\,dz_1\,dz_2 \\
&\qquad - \frac{1}{h_n^3}\left(\int K\left(\frac{x-z_1}{h_n}\right)\left(\int K\left(\frac{y-z_2}{h_n}\right) g(y)\,dy\right)\varphi_2(z_1,z_2)\,dz_1\right)^2
\end{aligned}$$

$$= \int K^2(x_1)\left(\int K(x_2)g(z_2+h_nx_2)\,dx_2\right)\varphi_2(x-h_nx_1,z_2)\,dx_1\,dz_2$$
$$- h_n\left[\int K(x_1)\left(\frac{1}{h_n}\int K(x_2)g(x_2h_nz_2)\,dy\right)\varphi_2(x-h_nx_1,z_2)\,dx_1\,dz_2\right]^2$$
$$\le C(x)$$

since $h_n \to 0$, and g is an integrable function.

(b) From (10),

$$\frac{1}{n}\sum_{1\le i<j\le n}\operatorname{Cov}(L_i(x),L_j(x)) = \frac{1}{nh_n^3}I_{2n}(x) \to 0. \tag{14}$$

By (a),

$$\frac{1}{n}\sum_{m=0}^{k_n}\operatorname{Var}Y'_{nm}(x) = \frac{1}{n}\sum_{m=0}^{k_n}\sum_{i=j(M_n+N_n)+M_n+1}^{(j+1)(M_n+N_n)}\operatorname{Var}L_i(x)$$
$$+\frac{2}{n}\sum_{m=0}^{k_n}\sum_{j(M_n+N_n)+M_n+1\le i<j\le(j+1)(M_n+N_n)}\operatorname{Cov}(L_i(x),L_j(x))$$
$$\le C\left[\frac{k_nN_n}{n}+\frac{2}{n}\sum_{1\le i<j\le n}\operatorname{Cov}(L_i(x),L_j(x))\right] \to 0.$$

The proof is then completed by means of (14).

Proof of Proposition 4.3. By Lemma 4.1,

$$n^{-1}E(T_n^2(x)) = n^{-1}\sum_{j=0}^{k_n}\operatorname{Var}Y'_{nj}(x)+2n^{-1}\sum_{0\le i<j\le k_n}\operatorname{Cov}(Y'_{ni}(x),Y'_{nj}(x))$$
$$\le n^{-1}\sum_{j=0}^{k_n}\operatorname{Var}Y'_{nj}(x)+2n^{-1}\sum_{1\le i<j\le n}\operatorname{Cov}(L_i(x),L_j(x)) \to 0.$$

Similarly

$$n^{-1}E(T_n'^2)(x)$$
$$= n^{-1}\sum_{j=k_n(M_n+N_n)+1}^{n}\operatorname{Var}L_j(x)+2n^{-1}\sum_{1\le i<j\le k_n}\operatorname{Cov}(L_i(x),L_j(x))$$
$$\le Cn^{-1}[n-k_n(N_n+M_n)]+2n^{-1}\sum_{1\le i<j\le n}\operatorname{Cov}(L_i(x),L_j(x)) \to 0.$$

We use the definition of N_n and M_n to complete the proof of the convergence (we note that L^2-norm convergence implies convergence in probability).

To establish the last desired result, we proceed as follows. Let $\widehat{W}_{n1}(x), \ldots, \widehat{W}_{nk_n}(x)$ be independent r.v.'s with $\widehat{W}_{nj}(x)$ distributed as $(1/\sqrt{n})Y_{n1}(x)$. Let Φ_n be the characteristic function of $(1/\sqrt{n})Y_{n1}(x)$ so that the characteristic function of $\sum_{j=1}^{k_n} \widehat{W}_{nj}(x)$ is

$$\Phi_n^{k_n}(t/\sqrt{n}) = \prod_{j=1}^{k_n} E(e^{itY_{nj}(x)/\sqrt{n}}).$$

Now, by Theorem 5.3 of Roussas and Ioannides [17, p. 97], with $\xi_j = e^{itY_{nj}(x)/\sqrt{n}} \in \mathcal{F}_{t_j}^{s_j} |\xi_j| = 1$, we get

$$\Big| E\Big(\prod_{j=1}^{k_n} e^{itY_{nj}(x)/\sqrt{n}}\Big) - \Phi_n^{k_n}(t/\sqrt{n})\Big| \le C(k_n - 1)\Phi_2(M_n) \to 0.$$

It remains to show that $\Phi_n^{k_n}(t/\sqrt{n})$ converges to the ch.f. of $N(0, \sigma_1(x))$. To this end, from (6),

$$\begin{aligned} s_n^2(x) &= \sum_{j=1}^{k_n} \operatorname{Var} \widehat{W}_{nj}(x) = k_n n^{-1} \sigma^2 Y_{n1}(x) \\ &= k_n n^{-1} M_n M_n^{-1} \sigma^2 Y_{n1}(x) \to \sigma_1(x). \end{aligned}$$

To prove the asymptotic normality, we have to show that the Lindberg condition is satisfied for the sequence

$$\tilde{W}_{nj}(x) = \frac{\widehat{W}_{nj}(x)}{s_n} \qquad \Big(E(\tilde{W}_{ni}(x)) = 0, \sum_{i=1}^{k_n} \operatorname{Var}(\tilde{W}_{nj}(x)) = 1\Big),$$

that is, for all $\varepsilon > 0$,

$$\Psi_n(\varepsilon) = \sum_{j=1}^{k_n} \int_{|x|>\epsilon} x^2\, dF_{nj}(x) \to 0 \quad \text{as } n \to \infty,$$

where F_{nj} is the distribution function of $\tilde{W}_{nj}$. But

$$\begin{aligned} \Psi_n(\varepsilon) &= k_n E(\tilde{W}_{n1}^2(x) I_{\{\tilde{W}_{n1}(x) > \varepsilon\}}) = k_n E\bigg[\bigg(\frac{\widehat{W}_{n1}(x)}{s_n}\bigg)^2 I_{(\widehat{W}_{n1}(x)/s_n)^2 > \varepsilon}\bigg] \\ &= \frac{k_n}{n s_n^2} E[Y_{n1}^2(x) I_{\{(Y_{n1}^2(x)/\sqrt{n}) > \varepsilon s_n\}}]. \end{aligned}$$

Since $Y_{n1}(x) < C M_n/\sqrt{h_n^3}$, we have

$$\Psi_n(\varepsilon) \le \frac{k_n C^2 M_n^2}{n s_n^2 h_n^3} P(|Y_{n1}(x)| > \varepsilon n^{-1/2} s_n).$$

From the Chebyshev inequality, it follows that

$$\Psi_n(\varepsilon) \le \frac{k_n c^2 M_n^2}{\epsilon^2 n^2 s_n^4 h_n^3}\sigma^2 Y_{n1}(x) = \left(\frac{M_n}{\sqrt{nh_n^3}}\right)^2 \left(\frac{c}{\epsilon}\right)^2 \left(\frac{k_n}{ns_n^4}\sigma^2 Y_{n1}(x)\right).$$

Since $s_n^2 \to \sigma_1(x)$, with a suitable choice of M_n, the right-hand side above converges to zero. This completes the proof of the proposition.

PROPOSITION 4.4. *Under the assumptions* (H.1), (H.2), (H.3), *and* (H.4)(a)&(b), *for any* $x \in \mathbb{R}$,

$$f_n(x) - f(x) = O\left(h_n^2 + \frac{1}{nh_n}\right) \quad a.s.$$

Proof. First we evaluate the bias term. It is clear that

$$\frac{1}{nh_n}\sum_{i=1}^{n} E\left[K\left(\frac{x - X_i}{h_n}\right)\right] - f(x) = \frac{1}{h_n}E\left[K\left(\frac{x - X_1}{h_n}\right)\right] - f(x)$$

$$= \frac{1}{h_n}\int K\left(\frac{x-z}{h_n}\right) f(z)\,dz - f(x)\int K(z)\,dz \quad (K \text{ is p.d.f.}).$$

Let $z_1 = (x - z)/h_n$. Then from this change of variable, the Taylor expansion to order 2 under (H.2)(b) and the dominated convergence theorem, we obtain

$$Ef_n(x) - f(x) = O(h_n^2) \quad \text{a.s.} \tag{15}$$

Let us now examine the variance of $f_n(x)$. By repeating the arguments employed in the proof of Proposition 4.1 ($G_n(x) :\leftrightarrow: f_n(x)$),

$$nh_n \operatorname{Var} f_n(x) \to f(x)\int K^2(z)\,dz,$$

so that

$$\operatorname{Var} f_n(x) = O\left(\frac{1}{nh_n}\right) \quad \text{a.s.} \tag{16}$$

By combining (15) and (16), we deduce that

$$f_n(x) - f(x) = O\left(h_n^2 + \frac{1}{nh_n}\right) \quad \text{a.s.} \tag{17}$$

Proof of the main result. It suffices to combine the results of Proposition 4.2, 4.4 and show that

$$\sqrt{nh_n}\left(\frac{EG_n(x)}{f_n(x)} - Hg(x)\right) \to 0 \quad \text{as } n \to \infty.$$

To this end, observe that

$$\frac{EG_n(x)}{f_n(x)} - Hg(x) = \frac{1}{f_n(x)}[E(G_n(x)) - G(x)] + \frac{G(x)}{f_n(x)f(x)}[f_n(x) - f(x)]$$

where $G(x) := \int g(y)\varphi_2(x, y)\,dy$.

Similarly for the the bias term of f_n, we have

$$E(G_n(x)) - G(x) = O(h_n^2) \quad \text{a.s.} \tag{18}$$

and

$$\text{there exists } \eta > 0 \text{ such that } \sum_{i=1}^{\infty} P(\inf |f_n(x)| < \eta) < \infty.$$

Indeed,

$$\sum_{i=1}^{\infty} P(|f_n(x)| \leq 1/2) \leq \sum_{i=1}^{\infty} P(|f_n(x) - \mathbb{E}f_n(x)| > 1/2) < \infty. \tag{19}$$

Thus by (17)–(19) and (H.3), we get the convergence, which completes the proof.

References

[1] P. Ango-Nzé and R. Rios, *Density estimation in the L^∞ norm for mixing processes*, C. R. Acad. Sci. Paris Sér. I 320 (1995), 1259–1262.

[2] D. Bosq, *Nonparametric Statistics for Stochastic Processes: Estimation and Prediction*, Lecture Notes in Statist. 110, Springer, New York, 1996.

[3] D. Bosq et J. P. Lecoutre, *Théorie de l'estimation fonctionnelle*, Economica, Paris, 1987.

[4] G. Collomb et P. Doukhan, *Estimation non paramétrique de la fonction d'autorégression d'un processus stationnaire et ϕ-mélangeant: risques quadratiques par la méthode du noyau*, C. R. Acad. Sci. Paris Sér. I 296 (1983), 859–862.

[5] M. Delecroix, M. E. Nogueira et A. C. Rosa, *Sur l'estimation de la densité d'observations ergodiques*, Statist. Anal. Données 16 (1991), no. 3, 25–38.

[6] F. Ferraty, A. Goïa and P. Vieu, *Functional nonparametric model for time series: a fractal approach to dimension reduction*, Test 11 (2002), 317–344.

[7] H. Gillert and A. Wartenberg, *Density estimation for non-stationary Markov processes*, Math. Operationsforsch. Statist. Ser. Statist. 15 (1984), 263–275.

[8] L. Györfi, W. Härdle, P. Sarda and P. Vieu, *Nonparametric Curve Estimation for Time Series*, Springer, Berlin, 1989.

[9] A. Laksaci et A. Yousfate, *Estimation fonctionnelle de la densité de l'opérateur de transition d'un processus de Markov à temps discret*, C. R. Math. Acad. Sci. Paris 334 (2002), 1035–1038.

[10] D. Louani and E. Ould-Said, *Asymptotic normality of kernel estimators of the conditional mode under strong mixing hypothesis*, J. Nonparametric Statist. 11 (1999), 413–442.

[11] E. Liebscher, *Density estimation for Markov chains*, Statistics 23 (1992), 27–48.

[12] —, *Asymptotic normality of nonparametric estimators under α-mixing condition*, Statist. Probab. Lett. 43 (1999), 243–250.

[13] E. Masry, *Recursive probability density estimation for weakly dependent stationary processes*, IEEE Trans. Inform. Theory 32 (1986), 254–267.

[14] M. Rosenblatt, *Density estimates and Markov sequences*, in: Nonparametric Techniques in Statistical Inference (Bloomington, IN, 1969), M. Pur (ed.), Cambridge Univ. Press, 1970, 199–213.

[15] G. Roussas, *Nonparametric estimation of the transition distribution function of a Markov process*, Ann. Math. Statist. 40 (1969), 1386–1400.
[16] —, *Recursive estimation of the transition distribution function of a Markov process: Asymptotic normality*, Statist. Probab. Lett. 11 (1991), 435–447.
[17] G. Roussas and O. Ioannides, *Moment inequalities for mixing sequences of random variables*, Stoch. Anal. Appl. 5 (1987), 60–120.
[18] S. Yakowitz, *Nonparametric density and regression estimation for Markov sequences without mixing assumptions*, J. Multivariate Anal. 30 (1989), 124–136.
[19] E. Youndjé, *Propriétés de convergence de l'estimateur à noyau de la densité conditionnelle*, Rev. Roumaine Math. Pures Appl. 41 (1996), 535–566.

Samir Benaissa, Abbes Rabhi
Laboratoire de Mathématiques
Université de Sidi Bel Abbes
BP 89, Sidi Bel Abbes 22000, Algeria
E-mail: benaissamir@yahoo.fr
rabhi_abbes@yahoo.fr

Belaid Mechab
LMPM, Department of Mechanical Engineering
University of Sidi Bel Abbes
BP 89, Cité Ben M'hidi
Sidi Bel Abbes 22000, Algeria

Received on 10.3.2010;
revised version on 2.11.2010 (2036)