Int. J. Open Problems Compt. Math., Vol. 1, No. 1, June 2008

Drift Estimation From $\widetilde{\rho}$ -Mixing Sequences

MOUNIR ARFI
Dept. of Mathematics, University of Bahrain, Kingdom of Bahrain
e-mail : m-arfi@hotmail.com

Abstract

we obtain the almost sure convergence for a kernel estimate of the drift coefficient in the diffusion equation for $\widetilde{\rho}$ mixing sequences over a sequence of compact sets which increases to $\Re$ when n approaches infnity.

Keywords: *Almost sure convergence, Diffusion equation, Drift coefficient, Kernel estimate, $\widetilde{\rho}$-mixing sequences.*

1 Introduction

Let X_t be a diffusion solution of the stochastic differential equation:

$$dX_t = \mu(X_t)dt + \sigma(X_t)dW_t \qquad t \in \Re^+$$

$(W_t\,;\, t \in \Re^+)$ is a standard Brownian motion; μ and σ are two Lipschitz and unknown functions of class C^1 with σ strictly positive.We know that under Lipschitz conditions on μ and σ , there exists for any given initial X_0 independent of $(W_t;\, t \geq 0)$ a unique, with probability one, solution to the equation above and this solution is a measurable Markov process (Wong [11]) .

This unique solution must have a stationary transition density, say $f_{X_t|X_0}(.)$ satisfying the forward equation of Kolmogorov:

$$\frac{\partial^2}{\partial x^2}\left(\frac{1}{2}\sigma^2(x) f_{X_t|X_0}(x)\right) - \frac{\partial}{\partial x}\left(\mu(x) f_{X_t|X_0}(x)\right) = \frac{\partial}{\partial t} f_{X_t|X_0}(x)$$

with $f_{X_t|X_0}(.)$ tending to a limiting density, say $f(.)$ as t goes to infinity.

For simplicity, we shall suppose that the initial distribution of X_0 has density $f(.)$ so that $(X_t)_{t\geq 0}$ is a stationary process and we are interested in estimation of $\mu(x)$ for each $x \in S$ where S is the nonempty set $\{x \in \Re \,/\, f(x) > 0\}$.

Moreover, under conditions of existence and uniqueness of the solution to the stochastic differential equation , the stationary diffusion X is ergodic (*see* Brown and Hewitt [7]).

This problem has been considered by several authors, among others Pham [10] gave a convergence in quadratic mean for the kernel estimate of the drift coefficient from the regression equation $E(X_{t+p}|X_t = .)$; $p \geq 1$, Arfi [1] established the almost sure convergence when the observed process is ergodic, Arfi and Lecoutre [3] established the almost sure convergence for a kernel estimate of the diffusion coefficient, and lately, Arfi [2] studied the almost sure convergence for a kernel estimate of the drift coefficient when the observed process is mixing over a sequence of compact sets which increases to $\Re$.

In this paper we give the almost sure convergence for the kernel estimate of the drift coefficient when the observed sequences are $\widetilde{\rho} - mixing$ over a sequence of compact sets C_n which increases to $\Re$ when $n \to \infty$.

Let $(\Omega, \mathcal{F}, \mathcal{P})$ be a probability space. Given the σ-algebras $\mathcal{B}$ and $\mathcal{R}$ in $\mathcal{F}$, let $\rho(\mathcal{B}, \mathcal{R}) = \sup\{corr(X,Y), X \in L_2(\mathcal{B}), Y \in L_2(\mathcal{R})\}$ where $corr(X,Y) = (EXY - EXEY)/\sqrt{varXvarY}$.

Bradley [5] introduced the following coefficients of dependence $\widetilde{\rho}(k) = \sup\{\rho(\mathcal{F}_S, \mathcal{F}_T)\}$, $k \geq 0$ where the supermum is taken over all finite subsets $S, T \subset N$ such that $dist(S,T) \geq k$.

Obviously,

$$0 \leq \widetilde{\rho}(k+1) \leq \widetilde{\rho}(k) \leq 1, \quad k \geq 0, \quad \text{and} \quad \widetilde{\rho}(0) = 1.$$

Definition
A random variable sequence $(X_t, t \geq 1)$ is said to be $\widetilde{\rho}$-mixing sequence if there exists k$\in$ N such that $\widetilde{\rho}$(k) ¡ 1.

Without loss of generality we may assume that the observed process is such that

$\widetilde{\rho}$(k) ¡ 1 (see Bryc and Smolenski [8]).

In the study of $\widetilde{\rho}$-mixing sequences we refer to Bradley [5], [6] for the central limit theorem, Bryc and Smolenski [8] for moment inequalities and almost sure convergence, Peligrad and Gut [9] for almost sure results.

2. The Model, the Notation, Some definitions

Let d be positive and fixed and $n \in \mathrm{N}$, the Markov observation $(X_{jd}\ ;\ 0 \leq j \leq n-1)$ permit to write:

$$X_{jd+d} - X_{jd} = \mu_d(X_{jd}) + \sigma_d(X_{jd})Y_{jd+d}$$

where $\mu_d(X_j) = E(X_{j+d} - X_j \,|\, X_j)$ and $\sigma_d^2(X_j) = V(X_{j+d} \,|\, X_j)$ are supposed to exist and define discrete versions of μ and σ^2, (Y_j) being a stationary Gaussian process such that :

$E(Y_{j+d} \,|\, X_s\ ;\ s \leq j) = 0$ and $E(Y_{j+d}^2 \mid X_s\ ;\ s \leq j) = 1.$

A natural estimator of μ_d is :

$$\mu_{d,n}(x) = \frac{\sum_{j=0}^{n-1} K\left(\frac{x-X_{jd}}{h_n}\right)(X_{jd+d} - X_{jd})}{\sum_{j=0}^{n-1} K\left(\frac{x-X_{jd}}{h_n}\right)} \qquad \forall\ x \in S$$

where (h_n) is a positive sequence of real numbers such that $h_n \to 0$, and $nh_n \to \infty$ when $n \to \infty$, and K a Parzen Rosenblatt kernel type, that is a bounded function satisfying $\int K(x)dx = 1$ and $\lim |x|K(x) = 0$ when $|x| \to \infty$, in addition we will assume it to be strictly positive and with bounded variation .

The almost sure convergence of $\mu_{d,n}$ to μ_d is established under the $\widetilde{\rho}$-mixing condition and using the fact that : $\mu(x) = \lim_{d\to 0} d^{-1}E(X_{j+d} - X_j \mid X_j = x)$, we deduce an estimate $(\mu_{d,n}/d)$ of μ , if $d = d(n)$ such that $N = nd \to \infty$, which is a necessary condition for both $Nh_n \to \infty$ and the $\widetilde{\rho}$-mixing condition .

We make the following assumptions:

(A.1) The process (X_{jd}) , $j \in \mathrm{N}$ is strictly stationary and $\widetilde{\rho}$-mixing.

(A.2) The initial random variable X_0 is of second order : $E(X_0^2) < \infty$.

(A.3) The kernel K is Lipschitz of order γ_1.

(A.4) The functions $\mu(.)$ and $\sigma(.)$ are Borel measurable on $\Re$ satisfying for $x, y \in \Re$ the

uniform Lipschitz condition:

$$\begin{aligned} |\mu(x) - \mu(y)| &\leq c\,|x-y| \\ |\sigma(x) - \sigma(y)| &\leq c\,|x-y| \end{aligned}$$

and the linear growth condition

$$\begin{aligned}|\mu(x)| &\leq c\sqrt{1+x^2}\\ |\sigma(x)| &\leq c\sqrt{1+x^2}\end{aligned}$$

where c is a positive constant.

(A.5) $\exists\,\Gamma < \infty, \quad \forall\, x \in \Re \quad f(x) \leq \Gamma$
and
$\exists\, \gamma_n > 0, \quad \forall\, x \in C_n \quad f(x) \geq \gamma_n.$

where C_n is a sequence of compact sets such that $C_n = \{x : ||x|| \leq c_n\}$ with $c_n \longrightarrow \infty$.

(A.6) The density f is twice differentiable and its derivatives are bounded.

3. Main Results

The main results of this paper are the following theorem and corollary.

Theorem

Suppose that h_n *is a positive sequence of real numbers such that* $h_n = o(\gamma_n)$ that *satisfying* $\lim\limits_{n\to\infty} \frac{n^{1-\xi}h_n}{Logn} = \infty$ *for some* $\xi \in]0, 1/2[$; *and let* K *to be Lipschitz kernel with bounded variation; i.e.* $\int z^2K(z)dz < \infty$, *then under assumptions A1 - A6, and for a compact sets* C_n *we have*

$$\sup_{x\in C_n} |\mu_{d,n}(x) - \mu_d(x)| \longrightarrow 0, \quad a.s. \quad n \to \infty.$$

Corollary *Under assumptions of Theorem 1, if we choose* h_n and d such that :

$$d \to 0, \quad \lim_{n\to\infty} \frac{n^{1-\xi}dh_n}{Logn} = \infty, \quad h_n = o(d),$$

then we have:

$$\sup_{x\in C_n} |\frac{\mu_{d,n}(x)}{d} - \mu(x)| \longrightarrow 0, \quad a.s. \quad n \to \infty.$$

Remark If we assume that the initial condition X_0 is independent of $(W_j \; ; \; j \in \Re^+)$ with density f , then a condition such as : for all $x \in \Re \quad |\mu(x)| + \sigma(x) \leq c(1+x^2)^{1/2}$ where c is a strictly positive constant, implies that the process (X_j) is stationary (Wong [11]).

Remark As sequences c_n and h_n defined in the Theorem 1, we can choose $c_n = O((Logn)^{1/\gamma_1})$ and $h_n = O(n^{-\tau})$ with $0 < \tau < 1$.*On the other part, the construction of the estimator requires a choice of* K *and* h_n. If the choice of K does not much influence the asymptotic behavior of $\mu_{d,n}$,on the contrary the choice of h$_n$ turns to be crucial for the estimator's accuracy. One can employ a cross-validation or plug-in method. In a forthcoming paper using simulations,we give comparisons of the results between two methods of estimation.

4. Preliminary Results

We make use of the following decomposition:

$$\mu_{d,n}(x) - \mu_d(x) = A_n(x) + B_n(x)$$

with

$$\begin{aligned} A_n(x) &= \frac{1}{f(x)} \left\{ [g_n(x) - \mu_d(x)f(x)] - W_{n,d}(x)\,[f_n(x) - f(x)] \right\} \\ B_n(x) &= \frac{1}{f(x)} \left\{ G_n(x) - T_n(x)\,[f_n(x) - f(x)] \right\} \end{aligned}$$

where

$$\begin{aligned} g_n(x) &= \frac{1}{nh_n} \sum_{j=0}^{n-1} K\left(\frac{x - X_{jd}}{h_n}\right) \mu_d(X_{jd}) \\ f_n(x) &= \frac{1}{nh_n} \sum_{j=0}^{n-1} K\left(\frac{x - X_{jd}}{h_n}\right) \\ W_{n,d}(x) &= \frac{\sum_{j=0}^{n-1} K\left(\frac{x-X_{jd}}{h_n}\right) \mu_d(X_{jd})}{\sum_{j=0}^{n-1} K\left(\frac{x-X_{jd}}{h_n}\right)} \\ G_n(x) &= \frac{1}{nh_n} \sum_{j=0}^{n-1} K\left(\frac{x - X_{jd}}{h_n}\right) \sigma_d(X_{jd}) Y_{jd+d} \end{aligned}$$

$$T_n(x) = \frac{\sum_{j=0}^{n-1} K\left(\frac{x-X_{jd}}{h_n}\right)\sigma_d(X_{jd})Y_{jd+d}}{\sum_{j=0}^{n-1} K\left(\frac{x-X_{jd}}{h_n}\right)}$$

If $|Y_{jd+d}| < M_n$ then $|T_n(x)| <$ constant $\times M_n$ *a.s.* where $M_n \to \infty$ is a sequence to be defined later.

And we can write:

$$\sup_{x\in C_n} |A_n(x)| \leq \frac{1}{\gamma_n}\left\{\sup_{x\in C_n} |g_n(x) - \mu_d(x)f(x)| + \sup_{x\in C_n} |W_{n,d}(x)||f_n(x) - f(x)|\right\}$$

$$\sup_{x\in C_n} |B_n(x)| \leq \frac{1}{\gamma_n}\left\{\sup_{x\in C_n} |G_n(x)| + \rho_2 M_n \sup_{x\in C_n} |f_n(x) - f(x)|\right\}$$

where ρ_2 is an upperbound of $\sigma_d(.)$

Lemma

Under hypotheses of Theorem 1, we have:

$$\gamma_n^{-1} \sup_{x\in C_n} |g_n(x) - \mu_d(x)f(x)| \to 0, \ a.s. \qquad n \to \infty.$$

proof

We have $C_n = \{x : ||x|| \leq c_n\}$ where $c_n \to \infty$ and

$$g_n(x) = \frac{1}{nh_n}\sum_{j=0}^{n-1} K\left(\frac{x - X_{jd}}{h_n}\right)\mu_d(X_{jd})$$

then we write

$$g_n(x) - \mu_d(x)f(x) = (g_n(x) - Eg_n(x)) + (Eg_n(x) - \mu_d(x)f(x))\,.$$

We put $g_n(x) - Eg_n(x) = \sum_{j=0}^{n-1} Z_j$ with

$$Z_j = \frac{1}{nh_n}\left\{K\left(\frac{x - X_{jd}}{h_n}\right)\mu_d(X_{jd}) - E\left(K\left(\frac{x - X_{jd}}{h_n}\right)\mu_d(X_{jd})\right)\right\}$$

by construction $EZ_j = 0$.

If $\overline{K}$ and $\rho_1$are upperbounds of K and μ_d respectively, we have: $|Z_j| \leq (2\overline{K}\,\rho_1)/(nh_n)$ and $E|Z_j| \leq (2\overline{K}\,\rho_1)/n$.

Now, let us write

$$\sum_{n=1}^{\infty} P(\gamma_n^{-1}|g_n(x) - Eg_n(x)| > \varepsilon) = \sum_{n=1}^{\infty} P(\gamma_n^{-1}|\sum_{j=0}^{n-1} Z_j| > \varepsilon).$$

Now we write for $\alpha > 1$

$\psi_{nj} = Z_j I_{(|Z_j| \leq n^\alpha)}$ and $V_{nj} = Z_j I_{(|Z_j| > n^\alpha)}$ for $0 \leq j \leq n-1$.

Then,

$|\sum_{j=0}^{n-1} Z_j| \leq |\sum_{j=0}^{n-1}(\psi_{nj} - E\psi_{nj})| + |\sum_{j=0}^{n-1} V_{nj}| + |\sum_{j=0}^{n-1} E\psi_{nj}|$

We need to show the following:

$$\sum_{n=1}^{\infty} P(\gamma\,_n^{-1}|\sum_{j=0}^{n-1}(\psi_{nj} - E\psi_{nj})| > \varepsilon n^\alpha/3) < \infty$$

$$\sum_{n=1}^{\infty} P(\gamma_n^{-1}|\sum_{j=0}^{n-1} V_{nj}| > \varepsilon n^\alpha/3) < \infty$$

$$\gamma_n^{-1}|\sum_{j=0}^{n-1} E\psi_{nj}|/n^\alpha \longrightarrow 0, \quad n \to \infty.$$

The Markov inequality and Chebyshev's inequality lead to:

$$\sum_{n=1}^{\infty} P(\gamma_n^{-1}|\sum_{j=0}^{n-1}(\psi_{nj}-E\psi_{nj})| > \varepsilon n^{\alpha}/3) \leq c_1 \sum_{n=1}^{\infty}\sum_{j=0}^{n-1} E|\psi_{nj}|^b \diagup \gamma_n^b n^{\alpha b} \leq c_2 \sum_{n=1}^{\infty} \gamma_n^{-b} n^{-\alpha b} < \infty$$

if we choose $\gamma_n = n^{-a}$ with $\alpha > a > 0$ and where $c_1$and c_2 are two positive constants and b such that $b > 1/(\alpha - a)$. The Borel-Cantelli lemma permits to conclude for (4.2).

Now, note that

$$(|\sum_{j=0}^{n-1} V_{nj}| > \varepsilon n^{\alpha}/3) \subset \bigcup_{j=0}^{n-1}(|Z_j| > n^{\alpha})$$

then,

$$\sum_{n=1}^{\infty} P(\gamma_n^{-1}|\sum_{j=0}^{n-1} V_{nj}| > \varepsilon n^{\alpha}/3) \leq \sum_{n=1}^{\infty} nP(|Z_j| > n^{\alpha}\gamma_n/3) \leq \sum_{n=1}^{\infty} nE|Z_j|^b \diagup n^{\alpha b}\gamma_n^b \leq c_3 \sum_{n=1}^{\infty} n^{-\alpha b}\gamma_n^{-b} <$$

with $\gamma_n = n^{-a}$ for $a > 0$ and such that $b(\alpha - a) > 1$ and where c_3 is a positive constant.

Lastly, we can write for $\alpha > a$:

$$\gamma_n^{-1} n^{-\alpha}|\sum_{j=0}^{n-1} E\psi_{nj}| \leq \gamma_n^{-1} n^{-\alpha} \sum_{j=0}^{n-1} |EV_{nj}| = \gamma_n^{-1} n^{-\alpha} \sum_{j=0}^{n-1} E|Z_j| I_{(|Z_j| > n^{\alpha})} = n^{a-\alpha} E|Z_j| I_{(||Z_j| > n^{\alpha})} \longrightarrow 0.$$

Next we cover C_n by δ_n spheres in the shape of $\{x : ||x - x_{nk}|| \leq c_n \delta_n^{-1}\}$ for $1 \leq k \leq \delta_n$,
$c_n \to \infty$ and δ_n chosen such that $\delta_n \to \infty$ to be defined later, and we make use of the following decomposition.

$$\left|\sum_{j=0}^{n-1} Z_j\right| \leq |g_n(x) - g_n(x_{nk})| + |E\left[g_n(x) - g_n(x_{nk})\right]| + |g_n(x_{nk}) - Eg_n(x_{nk})|\,.$$

The first and the second component in the right-hand side of the inequality above, will be considered in the same manner.

The kernel K being Lipschitz, we obtain

$$\sup_{x\in C_n}|g_n(x)-g_n(x_{nk})|\le\frac{L_K\rho_1}{h_n^{1+\gamma_1}}||x-x_{nk}||^{\gamma_1}\le\frac{L_K\rho_1}{h_n^{1+\gamma_1}}c_n^{\gamma_1}\delta_n^{-\gamma_1}=\frac{1}{Logn}$$

δ_n is chosen such that :

$$\delta_n=\frac{L_K^{1/\gamma_1}\rho_1^{1/\gamma_1}(Logn)^{1/\gamma_1}c_n}{h_n^{(1+\gamma_1)/\gamma_1}}\to\infty.$$

Then

$$\left|\sum_{j=0}^{n-1}Z_j\right|\le\sup_{1\le k\le\delta_n}|g_n(x_{nk})-Eg_n(x_{nk})|+\frac{2}{Logn}$$

so that for all $n\ge n_1(\varepsilon_n)$, $\forall\,\varepsilon_n>0$ we have

$$P\left(\gamma_n^{-1}\sup_{x\in C_n}\left|\sum_{j=0}^{n-1}Z_j\right|>2\varepsilon_n\right)\le\sum_{k=1}^{\delta_n}P\left\{\gamma_n^{-1}\left|g_n(x_{nk})-Eg_n(x_{nk})\right|>\varepsilon_n\right\}\ .$$

Now, using similar decomposition as in (4.1) δ_n times; the use of $\delta_n n^\alpha\gamma_n^{-1}$ instead of $\gamma_n^{-1}n^\alpha$ and hypotheses of Theorem 1 permit to conclude that

$$\gamma_n^{-1}\sup_{x\in C_n}\left|\sum_{j=0}^{n-1}Z_j\right|\longrightarrow 0,\quad a.s., n\to\infty.$$

It remains to show that : $\gamma_n^{-1}\sup_{x\in C_n}|Eg_n(x)-\mu_d(x)f(x)|\to 0,\qquad n\to\infty.$

We write

$$\gamma_n^{-1}\sup_{x\in C_n}|Eg_n(x)-\mu_d(x)f(x)|\le\gamma_n^{-1}h_n^{-1}\sup_{x\in C_n}\int K(h_n^{-1}(u-x))|\mu_d(u)-\mu_d(x)|f(u)du$$

$$+\gamma_n^{-1}h_n^{-1}\sup_{x\in C_n}|\mu_d(x)|\int K(h_n^{-1}(u-x))|f(u)-f(x)|du=I_1+I_2.$$

Now if we put $z=h_n^{-1}(u-x)$, the fact that μ_d is Lipschitz provides

$$I_1\le\gamma_n^{-1}h_n\sup_{x\in C_n}\int|z|K(z)f(zh_n+x)dz$$

then a choice such as $\gamma_n^{-1}h_n\longrightarrow 0$ conclude that $I_1\longrightarrow 0$ when $n\longrightarrow\infty$.

It remains to show that $I_2 \longrightarrow 0$.

$$I_2 = \gamma_n^{-1} \sup_{x \in C_n} |\mu_d(x)| \int K(z)|f(zh_n + x) - f(x)|dz$$

A taylor expansion gives:

$$I_2 \le \rho_1 \gamma_n^{-1} h_n \int |z| K(z) f'(x) dz + 0.5 \rho_1 \gamma_n^{-1} h_n^2 \int z^2 K(z) f^{''(x)dz} \longrightarrow 0, \; n \to \infty.$$

where ρ_1 is an upper bound of μ_d.

Lemma

Under hypotheses of Theorem 1, we have:

$$\gamma_n^{-1} \sup_{x \in C_n} |f_n(x) - f(x)| \to 0, \; a.s. \quad when \quad n \to \infty.$$

proof

This is a particular case of Lemma 3 when $\mu_d(x) = 1$.

Now, the kernel K being positive, we get $\sup_{x \in C_n} |W_{n,d}(x)| < \rho_1$ where ρ_1 is an upperbound of μ_d.

And we conclude that :

$$\sup_{x \in C_n} |A_n(x)| \le \frac{1}{\gamma_n} \sup_{x \in C_n} |g_n(x) - \mu_d(x) f(x)| + \frac{\rho_1}{\gamma_n} \sup_{x \in C_n} |f_n(x) - f(x)|.$$

Lemma

Under hypotheses of Theorem 1, we have:

$$\gamma_n^{-1} \sup_{x \in C_n} |G_n(x)| \to 0 \quad a.s. \quad n \to \infty.$$

proof

The study of $G_n(x)$ cannot be made directly because of the possible large values of the variables Y_{jd+d} so we use a truncation technique which consists in decomposing $G_n(x)$ in $G_n^+(x)$ and $G_n^-(x)$ where

$$G_n^+(x) = \frac{1}{nh_n} \sum_{j=0}^{n-1} K\left(\frac{x - X_{jd}}{h_n}\right) \sigma_d(X_{jd}) Y_{jd+d} I_{[|Y| > M_n]}$$

and $G_n^-(x) = G_n(x) - G_n^+(x)$ with M_n a nondecreasing and unbounded sequence.

We write:

$$\gamma_n^{-1} \sup_{x \in C_n} |G_n^+(x) - EG_n^+(x)| \leq \mathrm{E}_n + \mathrm{F}_n$$

with :

$$\mathrm{E}_n = \frac{1}{n\gamma_n h_n} \sup_{x \in C_n} \sigma_d(X_{jd}) \sum_{j=0}^{n-1} K\left(\frac{x - X_{jd}}{h_n}\right) |Y_{jd+d}| I_{[|Y|>M_n]}$$

we have $(\mathrm{E}_n \neq 0) \subset \{\exists\, j_0 \in [0, 1, ..., n-1]$ such that $|Y_{j_0}| > M_n\}$

$$(\mathrm{E}_n \neq 0) \subset \bigcup_{j=0}^{n-1} \{|Y_{jd+d}| > M_n\}$$

$$P(\mathrm{E}_n \neq 0) \leq \textstyle\sum_{j=0}^{n-1} P\{|Y_{jd+d}| > M_n\} = nP\{|Y_0| > M_n\}$$

$$\sum_n P(\mathrm{E}_n \neq 0) \leq \sum_n n\frac{E|Y_0|^\beta}{M_n^\beta} \leq M \sum_n n M_n^{-\beta}$$

with M being a positive constant and β such that $\beta > (2/\xi)$. Then it is sufficient to choose $M_n = n^\xi$ for some $\xi \in]0, 1/2[$ to get $\sum_n P(\mathrm{E}_n \neq 0) < \infty$.

We conclude with Borel-Cantelli Lemma that $\mathrm{E}_n \to 0,\ a.s. \quad n \to \infty$ and $\sup_{0 \leq j \leq n-1} |Y_{jd+d}| \leq M_n \ \ a.s.$

Then the kernel K being strictly positive , we deduce that $|T_n(x)| \leq \rho_2 M_n \ \ a.s.$

Now,

$$\mathrm{F}_n = \frac{1}{n\gamma_n h_n} \sup_{x \in C_n} \left| E \sum_{j=0}^{n-1} K\left(\frac{x - X_{jd}}{h_n}\right) \sigma_d(X_{jd}) Y_{jd+d} I_{[|Y|>M_n]} \right|$$

$$E(\mathrm{F}_n) \leq \frac{\overline{K}\rho_2}{\gamma_n h_n} E\left(|Y|\, I_{[|Y|>M_n]}\right)$$

where $\overline{K}$ and ρ_2 are upperbounds of K and σ_d respectively.

Then

$$E(\mathrm{F}_n) \leq \frac{\overline{K}\rho_2}{\gamma_n h_n} \left(E\left(Y^2\right)\right)^{1/2} \left(P\left(|Y| > M_n\right)\right)^{1/2} \leq \frac{c_3}{\gamma_n h_n M_n^{\beta/2}}$$

where c_3 is a positive constant and M_n is the sequence defined above.

This leads to $E(\mathrm{F}_n) \to 0,\ n \to \infty \Longrightarrow \mathrm{F}_n \to 0,\ a.s.$ when $n \to \infty$, with the choice $\gamma_n = n^{-a}$ for $a > 0$, $h_n = n^{-\tau}$ for $0 < \tau < 1$ and $1 < \beta < 2(a + \tau)$.

It remains to show that :

$$\gamma_n^{-1} \sup_{x \in C_n} |G_n^-(x) - EG_n^-(x)| \longrightarrow 0, \quad a.s. \quad n \to 0$$

we write:

$$G_n^-(x) - EG_n^-(x) = \sum_{j=0}^{n-1} T_j$$

with

$$T_j = \frac{1}{nh_n} \left\{ K\left(\frac{x - X_{jd}}{h_n}\right) \sigma_d(X_{jd}) Y_{jd+d} I_{[|Y| \leq M_n]} - E\left[K\left(\frac{x - X_{jd}}{h_n}\right) \sigma_d(X_{jd}) Y_{jd+d} I_{[|Y| \leq M_n]} \right] \right\}$$

$|T_j| \leq (c_4 M_n) / (nh_n)$, where c_4 is a positive constant.
Now let us write

$$\sum_{n=1}^{\infty} P(\gamma_n^{-1} |G_n^-(x) - EG_n^-(x)| > \varepsilon) = \sum_{n=1}^{\infty} P(|G_n^-(x) - EG_n^-(x)| > \gamma_n \varepsilon)$$

same arguments as in the proof of lemma 3 permit to conclude that

$$\gamma_n^{-1} \sup_{x \in C_n} |G_n^-(x) - EG_n^-(x)| \to 0, \ a.s. \quad n \to \infty.$$

In the end , the fact that :

$$G_n(x) = G_n(x) - EG_n(x)$$

permit to conclude that

$$\gamma_n^{-1} \sup_{x \in C_n} |G_n(x)| \to 0, \quad a.s., \qquad n \to \infty.$$

Finally, similar works to those used in Lemma 4 with the use of $\gamma_n^{-1} M_n$ instead of γ_n^{-1} permit to conclude that :

$$\gamma_n^{-1} M_n \sup_{x \in C_n} |f_n(x) - f(x)| \longrightarrow 0, \quad a.s. \quad n \longrightarrow \infty.$$

5. Proof of the Main Results

5.1 Proof of Theorem 1

Lemmas 3, 4 and 5 permit to conclude.

5.2 Proof of Corollary 2

It suffices to write:

$$\frac{\mu_{d,n}(x)}{d} - \mu(x) = \frac{\mu_{d,n}(x) - \mu(x)}{d} + \left[\frac{\mu_d(x)}{d} - \mu(x) \right]$$

Then, similar techniques to those of Theorem 1 with the conditions of Corollary 2 permit to conclude.

Acknowledgements

The author is grateful to the referees for their comments.

References

[1] Arfi, M. Nonparametric drift estimation from ergodic samples. *J. Nonparametric Statistics.* 5, 1995, pp.381-389

[2] Arfi, M. Nonparametric estimation in the diffusion equation. *Jour. of Math. Sciences.* 12, pp. 129-136

[3] Arfi,M., Lecoutre, J.P. Estimation of the diffusion coefficient under strong mixing. *Statistics* 33,1999, pp. 73-84

[4] Banon, G. *Estimation non parametrique de densite de probabilite pour les processus de Markov.*These. Univ. Paul Sabatier de Toulouse, 1977

[5] Bradley, R.C. *Equivalent mixing conditions of random fields*, Technical Report 336. Center for Stochastic Processes. University of North Carolina, 1990

[6] Bradley, R.C. On the spectral density and asymptotic normality of weakly dependent random fields. *J. Theoret. Probab.* 5, , 1992, pp. 355-374

[7] Brown, B., Hewitt, J.I. Asymptotic likelihood theory for diffusion processes, *J. Appl. Prob.* 12, 1975, pp. 228-238

[8] Bryc, W., Smolenski, W. Moment conditions for almost sure convergence of weakly correlated random variables. *Proc. mer. Math. Soc.* 119, 1993, pp. 629-635

[9] Peligrad, M., Gut, A. Almost sure results for a class of dependent random variables. *J. Theoret. Probab.* 12, 1999, pp. 87-104

[10] Pham, D.T. Nonparametric estimation of the drift coefficient in the diffusion equation. *Math. Oper. Statist., ser. Statistics* 12, 1981 pp. 61-73

[11] Wong, E. *Stochastic process in information and dynamical system.* Mc Graw Hill, 1971