Proceedings of The 9th Islamic Countries Conference on Statistical Sciences 2007
ICCS-IX 12-14 Dec 2007

Wavelet-Based Estimatirs of the Integrated Squared Density Derivatives for a mixing sequances

N. Hosseinioun na_ho8@stu-math.um.ac.ir
H. Doosti doosti@math.um.ac.ir
H.A. Nirumand Nirumand@math.um.ac.ir

Department of Statistics, School of Mathematical Sciences, Ferdowsi University of Mashhad, Iran

Abstract

The problem of estimation of the squared derivative of a probability density f is considered using wavelet orthogonal bases. We obtain the precise asymptotic expression for the mean integrated error of the wavelet estimators when the process is strongly mixing.

*Keywords :*Nonparametric estimation of a density, Wavelet, Mixing process .

1 Introduction

The motivation for estimation $I_d(f) = \int f^{(d)^2}(x)dx$ where f is a probability density and $f^{(d)}$ is the d-th derivative is well known. Kernel-type estimation the functional$I_2(f)$ has been investigated by Hall and Marron(1987). and Bickel and Ritov(1988) among others. In Prakasa Roa(1996), we have studied nonparametric estimation of the derivative of a density by wavelets and obtained a precise asymptotic expression for the mean integrated squared error following techniques of Masry(1994). Estimation of the integrated of squared density was discussed in Prakasa Roa(1997) by the method of wavelets and a precise asymptotic expression for the mean squared error had been obtained. Prakasa Roa(1999) also obtained the precise asymptotic expression integrated squared error of the wavelet estimators.
We now extend the result to the case of strongly mixing process . We show that the L_p error of the proposed estimator attains the same rate as when the observations are independent. Certain week dependence conditions are imposed to the x_i defined in $\{\Omega, N, P\}$.

Let N_k^m denote the σ-algebra generated by events $X_k \in A_k, ..., X_m \in A_m$. We consider the following classical mixing conditions:

1. Strong mixing (s.m) also called α-mixing:

$$\sup\sup |p(AB) - p(A)p(B)| = \alpha(s) \to 0 \qquad as \;\; s \to \infty$$

2. Complete regularity (c.r.), also called β-mixing:

$$\sup E\{var|p(B|N_1^m) - p(B)|\} = \beta(s) \to 0 \qquad as \;\; s \to \infty$$

3. Uniformely strong mixing (u.s.m.), also called $\phi - mixing$:

$$\sup\sup \frac{|p(AB) - p(A)p(B)|}{p(A)} = \phi(s) \to 0 \qquad as \;\; s \to \infty$$

4. ρ-mixing:

$$\sup\sup |corr(X, Y)| = \rho(s) \to 0 \qquad as \;\; s \to \infty$$

The problem of density estimation from dependent samples is often considered. For instance quadratic losses were considered by Ango Nze and Doukhan(1993). Bosq (1995), Castella and Leadbetter(1998) and Doukhan and Loen(1990).
Linear wavelet estimators were also used in context: Doukhan (1988) and doukhan and Loen(1990). Leblance (1995)also established that the $L_{\acute{p}}$ -loss $(2 \leq \acute{p} < \infty)$ of the linear wavelet density estimators for a stochastic process converges at the rate $N^{\frac{-s}{(2s+1)}}$ $(s = 1/p + 1/\acute{p})$, when the density of f belongs to the Besov space $B^s_{p,\acute{p}}$. Dooti , Niroumand and Afshari (2006) extended the above result for derivative of a density.

2 Discussion of Theorem's Assumptions

Consider the following conditions:

C_1:The distribution of (X_m, X_t) has a joint density $f_{m,t}$ such that for all m and t, $m \neq t$

$$(\int |f_{m,t}(x,y)|^v dxdy)^{1/v} = \|f_{m,t}(.,.)\| \leq F_v < \infty \qquad forsome\ v > 2$$

M_1:The process is ρ-mixing and $\sum_{t=1}^{\infty} \rho(t) \leq R < \infty$.

M_2:The process is ϕ-mixing and $\sum_{t=1}^{\infty} \phi^{1/2}(t) \leq \phi < \infty$.

Since the inequality $\rho(t) \leq 2\phi^{1/2}(t)$ holds (see Doukhan 1994), M_2 implies M_1 . Also note that if X and Y are random variables , then the following covariance inequalities hold.(see Doukhan , 1999, section 1.2.2)

$$cov(X_i, Y_j) \leq 2\rho(j-i)\|X\|_2.\|Y\|_2 \tag{2.1}$$

$$cov(X_i, Y_j) \leq 2\phi^{1/p}(j-i)\|X\|_p.\|Y\|_q \tag{2.2}$$

for any $p, q \geq 1$ and $1/p + 1/q = 1$.

3 Introduction to Wavelet

A wavelet system is an infinite collection of translated and scaled versions of functions ϕ and ψ called the *scaling function* and the *primary wavelet function*respectively. The function $\phi(x)$ is a solution of the equation

$$\phi(x) = \sum_{k=-\infty}^{\infty} C_k \phi(2x - k) \tag{3.1}$$

with

$$\int_{-\infty}^{\infty} \phi(x)dx = 1 \tag{3.2}$$

and the function $\psi(x)$ is defined by

$$\psi(x) = \sum_{-\infty}^{\infty} (-1)^k C_{-k+1} \psi(2x - k) \tag{3.2}$$

Note that the choice of the sequence C_k determines the wavelet system. It is easy to see that

$$\sum_{k=-\infty}^{\infty} C_k = 2 \tag{3.3}$$

Define

$$\phi_{j,k}(x) = 2^{j/2}\phi(2^j x - k), \quad -\infty < j, k < \infty \tag{3.5}$$

and

$$\psi_{j,k}(x) = 2^{j/2}\psi(2^j x - k), \quad -\infty < j, k < \infty \tag{3.6}$$

Suppose that the coefficients C_k satisfy the condition

$$\begin{aligned} \sum_{-\infty}^{\infty} c_K c_{k+2l} &= 2 \quad if \quad l = 0 \\ &= 0 \quad if \quad l \neq 0 \end{aligned}$$

It is known that , under some additional condition on ψ, the collection $\{\psi_{j,k}, -\infty < j, k < \infty\}$ is an orthonormal basis for $L^2(R)$ and $\{\psi_{j,k}, -\infty < k < \infty\}$ is an orthonormal system in $L^2(R)$ for each $-\infty < j < \infty$ (cf. Doubachies (1990)).

Definition 3.1. A scaling function $\phi \in c^{(r)}$is said to be *r-regular* for an integer $r \geq 1$ if for every non-negative integer $l \leq r$ and for any integer k,

$$|\phi^{(l)}(x)| \leq c_k(1 + |x|)^{-k}, \quad -\infty < x < \infty \tag{3.8}$$

for some $c_k \geq 0$ depending only on k where $\phi^{(l)}(.)$ denotes the l-th derivative of ϕ.

Definition 3.2. A *multiresolution analysis* of $L^2(R)$ contains of increasing sequences of closed subspaces V_j of $L^2(R)$ such that

(i)$\bigcap_{j=-\infty}^{\infty} V_j = \{0\}$;

(ii)$\overline{\bigcup}_{j=-\infty}^{\infty} V_j = L^2(R)$;

(iii)there is a scaling function $\phi \in V_0$ such that

$$\phi(x - k), \quad -\infty < k < \infty$$

is an orthonormal basis for V_0; and for all $h \in L^2(R)$,

(iv) For all $-\infty < k < \infty, \; h(x) \in V_0 \Rightarrow h(x - k) \in V_0$

(v) $h(x) \in V_j \Rightarrow h(2x) \in V_{j+1}$.

Let $\acute{H}_2$ denote the space of all functions $g(.)$ in $L^2(R)$ whose first $(S - 1)$ derivatives are absolutely continuous and define the norm

$$\|g\|_{\acute{H}_2} = \sum_{-\infty}^{\infty}[\int |g^{(j)}(t)|^2 dt]^{1/2}$$

Lemma 3.1.(Mallat(1989)) Let a *multiresolution analysi* be *r-regular.* Then for every $0 < s < r$, any function $g \in L^2(R)$ belongs to $\acute{H}_2$ iff

$$\sum_{t=-\infty}^{\infty} e_t^2 e^{2sl} < \infty \tag{3.9}$$

where $e_l^2 = \|g - g_l\|_2^2$ and g_l is the orthogonal projection of g on V_t.

Remarks. The above introduction is based on Antoniadis et al. (1994). For a detailed introduction to wavelet, see Chui (1992) or Daubechies (1992). For a brief survey, see Strang (1989).

4 Estimation by the Methods of Wavelets

Suppose $X_1, ...X_n$ are independent and identically distributed random variables with density f. that f is d-times differentiable and that $f^{(d)}$ denotes the d-th derivative of f We interpret $f^{(0)}$ as f. The problem of interest is the estimation of

$$I_d(f) = \int_{-\infty}^{\infty} f^{(d)^2}(x)dx \tag{4.1}$$

Assume that $f_{(d)} \in L^2(R)$ and there exist $D_j \geq 0$, $\beta_j \geq 0$ such that

$$|f^{(j)}(x)| \leq D_j |x|^{-\beta_j} \quad for\ |x| \geq 1, 0 \leq j \leq d \tag{4.2}$$

where $\beta > 1$.
Consider a mulitiresolution as discussed in Section 3. Let ϕ be the corresponding scaling function. Suppose that the multiresolution is *r-regular* for some $r \geq d$. Then by definition, $\phi \in C^{(r)}$, ϕ and its derivative $\phi^{(j)}$ up to order r are rapidly decreasing i.e., for every integer $m \geq 1$, there exists a constant $A_m > 0$ such that

$$|\phi^{(j)}(x)| \leq \frac{A_m}{(1+|x|)^m},\ 0 \leq j \leq r \tag{4.2}$$

Let

$$\phi_{l,k} = 2^{l/2}\phi(2^l x - k),\ \ -\infty < k, t < \infty \tag{4.3}$$

Then

$$\phi^{(j)}_{l,k} = 2^{l/2+lj}\phi^{(j)}(2^l x - k),\ \ -0 \leq j \leq r \tag{4.4}$$

and

$$|\phi^{(j)}_{l,k}(x)| \leq \frac{2^{(l/2)+lj}A_m}{(1+|x|)^m},\ 0 \leq j \leq r \tag{4.5}$$

If $d \geq 1$, then it is clear that

$$\lim_{|x| \longrightarrow \infty} \phi^{(j)}_{l,k} f^{(d-j-1)}(x) = 0,\ 0 \leq j \leq d-1 \tag{4.6}$$

for any fixed l and k. Let f_{ld} is the orthogonal projection of $f^{(d)}$ on V_l. Note that

$$f_{ld}(x) = \sum_{j=-\infty}^{\infty} a_{l,j}\phi_{lj}(x) \tag{4.7}$$

where

$$\begin{aligned} a_{lj} &= \int_{-\infty}^{\infty} f^{(d)}(u)\phi_{l,j}(u)du \\ &= (-1)^d \int_{-\infty}^{\infty} f(u)\phi^{(d)}_{l,j}(u)du \end{aligned} \tag{4.8}$$

by (3.4) for $d \leq 1$. Clearly the equation (4.9) holds for $d = 0$. Hence for all $d \geq 0$

$$a_{lj} = (-1)^d E[\phi^{(d)}_{l,j}(X_1)] \tag{4.10}$$

Further more

$$e_l^2 \equiv \|f^{(d)} - f_{ld}\|_2^2 = \|f^{(d)}\|_2^2 - \sum_{k=-\infty}^{\infty} a_{lk}^2 \longrightarrow 0 \ \ as \ \ l \longrightarrow \infty \tag{4.11}$$

by the properties of multiresolution decomposition. Hence $\|g\|_p = \int_{-\infty}^{\infty} |g|^p dx^{1/p},\ p \geq 1$. Note that

$$I_d(f) = \|f^{(d)}\|_2^2 \tag{4.12}$$

Let

$$f_{K,l,d}(x) = \sum_{k=-K}^{K} a_{lk}\phi_{l,k}(x) \tag{4.13}$$

where $K = K_n$ is a sequence of positive integers depending on $l = l_n$ tending to infinity as $n \longrightarrow \infty$ and $l = l_n \longrightarrow \infty$ as $n \longrightarrow \infty$. Note that $f_{K,l,d}(x)$ is a truncated projection of $f^{(d)}$ on V_t. Given an i.i.d sample $X_1, ...X_n$, let

$$A_{lk} = \frac{1}{n(n-1)} \sum_{i=1 \neq j}^{n} \sum_{j=1}^{n} \phi_{lk}^{(d)}(x_i)\phi_{lk}^{(d)}(x_j) \tag{4.14}$$

and we estimate $I_d(f)$ by

$$\hat{I}_d(f) = \sum_{k=-K}^{K} A_{lk} \tag{4.15}$$

Note that

$$E(A_{lk}) = a_{lk}^2 \tag{4.16}$$

and

$$E(\hat{I}_d(f)) = \sum_{k=-K}^{K} a_{lk}^2 \tag{4.17}$$

5 Main Results

Suppose that as $l_n \longrightarrow \infty$

$$k_n = 2^{\{(2d-1)+2\beta_0+2s\}\{l_n/(2\beta_0-1)\}} \log n$$

Define $\hat{I}_d(f)$ as an estimator of $I_d(f)$ where $\hat{I}_d(f)$ is given by the equation (4.15), then we have the following two results:

Theorem 5.1. If $\{X_n\}$ satisfies the condition c_1, then

$$\frac{n(n-1)}{2^{2l_n(1+2d)}} E|\hat{I}_d(f) - I_d(f)| \longrightarrow \int {\phi^{(d)}}^2(x)dx^2 \qquad as \quad n \longrightarrow \infty$$

Theorem 5.2. If $\{X_n\}$ satisfies the condition M_1, then

$$\frac{n(n-1)}{2^{2l_n(1+2d)}} E|\hat{I}_d(f) - I_d(f)| \longrightarrow \int {\phi^{(d)}}^2(x)dx^2 \qquad as \quad n \longrightarrow \infty$$

6 Proofs

Let

$$\begin{aligned} J_n^2 &= E|\hat{I}_d(f) - I_d(f)|^2 = Var(\hat{I}_d(f)) + \{E\hat{I}_d(f) - I_d(f)\}^2 \\ &= Var(\hat{I}_d(f)) + (\sum a_{lk}^2 - \int {f^{(d)}}^2(x)dx)^2 \\ &= Var(\hat{I}_d(f)) + (\|f_{k,l,d}\|_2^2 - \|f^{(d)}\|_2^2)^2 \end{aligned}$$

Following along the lines of Roa(1999), we get

$$(\|f_{k,l,d}\|_2^2 - \|f^{(d)}\|_2^2)^2 = o(2^{-4sl_n}) \tag{6.1}$$

proof of Theorem 5.1. Observe that

$$Var(\hat{I}_d(f)) = Var(\sum_{-k}^{k} A_{lk}) = \sum_{k}\sum_{\acute{k}} cov(A_{lk}, A_{l\acute{k}}) \tag{6.2}$$

where $cov(X,Y)$ is interpreted as $var(X)$. It is straightforward to check that

$$\sum_{k}\sum_{\acute{k}} EA_{lk}A_{l\acute{k}} = \frac{1}{n^2(n-1)^2}\sum_{k}\sum_{\acute{k}}\sum E\phi_{lk}^{(d)}(x_i)\phi_{l\acute{k}}^{(d)}(x_j)\phi_{lk}^{(d)}(\acute{x}_i)\phi_{l\acute{k}}^{(d)}(\acute{x}_j) \tag{6.3}$$

where the last summation runs over all $i, j, \acute{i}, \acute{j}$. Using (2.1) in (6.2) leads to

$$\sum_{k}\sum_{\acute{k}} EA_{lk}A_{l\acute{k}} =$$

$$\frac{1}{n^2(n-1)^2}\sum_{1\le i\le j\le n}\rho(j-i)\sum_{k}(\int \phi_{lk}^{(d^4)}(x_i)f(x_i)dx_i)^{1/2}\sum_{\acute{k}}(\int \phi_{l\acute{k}}^{(d^4)}(x_i)f(x_i)dx_i)^{1/2}+$$

$$\frac{1}{n^2(n-1)^2}\sum_{i<j}\sum_{k} E\phi_{lk}^{(d^2)}(x_i)\sum_{\acute{k}} E\phi_{l\acute{k}}^{(d^2)}(x_i) \tag{6.4}$$

Note that it suffices to bound the right-hand side of (6.3). By (4.5) and Masry(1994), one may easily get

$$\sum_{k}(\int \phi_{lk}^{(d)^4}(x_i)f(x_i)d(x_i))^{1/2}\sum_{\acute{k}}(\int \phi_{l\acute{k}}^{(d)^4}(x_i)f(x_i)d(x_i))^{1/2}$$

$$\le \sum_{k}(2^{l+4ld}\int \phi_{lk}^{(d)^4}(x_i)f(\frac{k+u}{2l})d(u))^{1/2}\sum_{\acute{k}}(2^{l+4ld}\int \phi_{l\acute{k}}^{(d)^4}(x_i)f(\frac{\acute{k}+v}{2l})d(v))^{1/2}$$

$$= 2^{2l+4lk}\sum_{k}\int \phi^{(d)^4}(u)f(\frac{u+k}{2^l})du$$

$$= 2^{2l+4lk}\int \phi^{(d)^4}(u)du(1+O(2^{-l})) \tag{6.5}$$

By similar argument as in Rao(1999), we get

$$\sum_{k} E\phi_{lk}^{(d)^2}(x_i)\sum_{\acute{k}} E\phi_{l\acute{k}}^{(d)^2}(x_i) \le 2^{2l(1+2d)}\{\int \phi^{(d)^4}(u)du\}^2$$

$$+2^{-2l(1+2d)}\sum_{k}\sum_{\acute{k}} a_{lk}^2 a_{l\acute{k}}^2 + O(\frac{1}{2^{2l(1+2d)}}) \tag{6.6}$$

Substituting (6.5) and (6.6) in (6.4), one may easily obtain

$$\sum_{k}\sum_{\acute{k}} EA_{lk}A_{l\acute{k}} \le \frac{2^{2l+4ld}}{n^2(n-1)^2}\frac{2}{n}\sum_{k}\rho(k)\int \phi^{(d)^4}(u)du(1+O(2^{-l}))$$

$$+\frac{2}{n(n-1)}[2^{2l+(1+2d)}]\{\int \phi^{(d)^4}(u)du\}^2 + \frac{1}{2^{2l(1+2d)}}\sum_k \sum_{\acute{k}} a_{lk}^2 a_{l\acute{k}}^2 + O(\frac{1}{2^{2l(1+2d)}})$$

Since $\sum_k \rho(k) < \infty$ and $\frac{1}{2^{2l(1+2d)}}\sum_k \sum_{\acute{k}} a_{lk}^2 a_{l\acute{k}}^2 = o(1)$,(Roa(1999)),

$$\frac{1}{2^{2l(1+2d)}}\sum_k \sum_{\acute{k}} EA_{lk}A_{l\acute{k}} = O(n^{-3}) + \frac{1}{n^2(n-1)^2}\{\int \phi^{(d)^4}(u)du\}^2 + o(1) + O(1) \quad (6.7)$$

So we may easily conclude

$$\frac{n(n-1)}{2^{2l(1+2d)}} Var\hat{I}_d(f) = O(n^{-2}) + \{\int \phi^{(d)^4}(u)du\}^2 + o(1) + O(\frac{1}{2^{2l(1+2d)}}) \quad (6.8)$$

Applying (6.8) in (6.1), yields the desired result.

proof of Theorem 5.2. Applying Holder inequality for v and $\acute{v}$ with $1/v + 1/\acute{v} = 1$, one may obtain

$$\int \phi_{lk}^{(d)^2}(x_i)\phi_{l\acute{k}}^{(d)^2}(x_j)f(x_i,x_j)dx_i dx_j$$

$$\leq F_v 2^{l+4ld}(\int \phi_{lk}^{(d)^4 \acute{v}}(u)du)^{1/2\acute{v}}(\int \phi_{lk}^{(d)^4 \acute{v}}(v)dv)^{1/2\acute{v}}$$

$$\leq F_v 2^{l+4ld}(\int \frac{A_m^{4\acute{v}}}{(1+u)^{4m\acute{v}}}du)^{1/2\acute{v}}(\int \frac{A_m^{4\acute{v}}}{(1+v)^{4m\acute{v}}}dv)^{1/2\acute{v}}$$

So it is easy to obtain

$$\sum_k \sum_{\acute{k}} \int \phi_{lk}^{(d)^2}(x_i)\phi_{lk}^{(d)^2}(x_j)f(x_i,x_j)dx_i dx_j$$

$$\leq F_v 2^{l+4ld}A_m^{4\acute{v}} \sum_k (\int \frac{du}{u^{4m\acute{v}}}du)^{1/2\acute{v}} \sum_k (\int \frac{dv}{v^{4m\acute{v}}}dv)^{1/2\acute{v}}$$

$$= F_v 2^{l+4ld}A_m^{4\acute{v}} \sum_u (\int \frac{du}{u^{4m\acute{v}}}du)^{1/2\acute{v}} \sum_v (\int \frac{dv}{v^{4m\acute{v}}}dv)^{1/2\acute{v}}$$

$$\leq F_v 2^{l+4ld}A_m^{4\acute{v}} \sum_u \frac{u^{(-4m\acute{v})^{1/2\acute{v}}}}{1-4m\acute{v}} \sum_v \frac{v^{(-4m\acute{v})^{1/2\acute{v}}}}{1-4m\acute{v}}$$

$$\leq F_v 2^{l+4ld}A_m^{4\acute{v}} \int_1^k \frac{u^{-2m+1/\acute{v}}}{1-4m\acute{v}}du \int_1^k \frac{v^{-2m+1/\acute{v}}}{1-4m\acute{v}}dv$$

$$= F_v 2^{2+4ld}A_m^{4\acute{v}}[\frac{-k^{-2m+\frac{1}{2\acute{v}}+1}}{(1-4mv)(2m+\frac{1}{2\acute{v}}+1)}]^2$$

$$= O(2^{l+4ld}) = o(1) \quad (69)$$

Using (6.6), (6.9) and (6.2) in (6.1), conclude the result.

References

[1] Ango Nze, P. and P. Doubachies (1993), Functional estimation for time series: a general approach, Prépublication de I'Université Paris-Sun no. 93-43.
[2] Bosq, D. (1995). Optimal asymptotic quadratic error of density estimators for strong mixing or chaotic data. *Statist. Probab. Lett.* to appear.
[3] Bickel, P. and Ritrov, Y. (1988), Estimation of integrated squared density derivatives; sharp best order of convergence estimate, *sankhya A*, 50, 381-393.
[4] Doukhan, P. (1988), Forme de Toeplitz associée a une analyse multi-échelle, C.R. Acad *Sci. Paris*, t306, Série 1, 663-666.
[5] Doukhan, P. and J.R. Loen (1990), Une note sur la déviation quadratique d'estimateurs de densités par projections orthogonales, C.R. *Acad Sci. Paris,* t310, série 1, 425-430.
[6] Hall,P. and Marron, J.S., (1987), Estimation of integrated squared density derivatives, *statist. Prob. Lett.* 6, 109-115.
[7] Leblance, F.(1994), L_p-risk of the wavelet linear density estimator for a stochastic process, Rapport Technique no. 9402, L.S.T.A Paris 6.
[8] Leblance, F., (1996), Wavelet linear density estimator for a discrete-time stochastic process: L_P-losses, Statistics and probability letters 27, 71-84.
[9] Masry, E. (1994), Probability density estimation from dependent observation using wavelet orthonormal bases, *Statist. Probab. Lett.* 21,181-194.
[10] Prakasa Roa. B.L.S. (1996), Nonparametric estimation of the derivatives of a density by the method of wavelets, *Bull, Inform. Cyb.* 28, 91-100
[11] Prakasa Roa. B.L.S. (1989), Wavelets and dillation equation; a brief introduction, *SIAM Review*, 31, 614-627.
[12] Prakasa Roa. B.L.S. (1999), Estimation of the integrated squared density derivatives by wavelets, Bulletin of Information and Cybernetics, Vol 31, no. 1.