Wavelet Based Estimation of the Derivatives of a Density for m-Dependent Random Variables

Yogendra P. Chaubey¹, Hassan Doosti²

¹Concordia University, Montreal, Canada.
²Ferdowsi University, Mashhad, Iran.

Abstract. Here, we propose a method of estimation of the derivatives of probability density based wavelets methods for a sequence of m—dependent random variables with a common one-dimensional probability density function and obtain an upper bound on \( L_p \)-losses for the such estimators.

1 Introduction

Estimation of density and its derivatives using wavelets has generated a lot of interest in recent years. We refer to Härdle et al. (1998) and Vidakovic (1999) for a detailed coverage of wavelet theory in statistics and to Prakasa Rao (1999a) for a recent comprehensive review of nonparametric functional estimation.

For the iid case, Prakasa Rao (1996) considered the use of wavelets for estimating the derivatives of a density and obtained an upper bound on the \( L_2 \)-losses for the proposed estimator. Prakasa Rao (1999b) further investigated the use of wavelets for estimating the
integrated squared density. For the non iid case Prakasa Rao (2003) considered the case of associated sequences for estimation of density using wavelets. Recently, Chaubey et al. (2006a, 2006b) have extended the results in Prakasa Rao (1996) for estimation of derivatives of a density for negatively and positively associated sequences, respectively. Here we consider yet another form of a dependence, namely, $m-$ dependence, described below.

Let $\zeta = \{X_i, i \geq 1\}$ denote a sequence of stationary random variables defined on a common probability space such that $\{X_i, 1 \leq i \leq k\}$ is independent of $\{X_i, i \geq k + m + 1\}$ for all $k \geq 1$. Then such a sequence $\zeta$ is called dependent of order $m$ or in short $m-$ dependent.

This note concerns with estimating the common one-dimensional density $f$ and its derivatives based on $n$ observations $\{X_1, ..., X_n\}$.

The organization of the paper is as follows. In section 2, we discuss the preliminaries of the wavelet based estimation of the derivatives of the density along with the necessary underlying setup considered in Prakasa Rao (1996). Section 3 provides the bounds on the $L_p-$losses for the proposed estimator.

2 Preliminaries

Let $\{X_n, n \geq 1\}$ be a sequence of random variables on the probability space $(\Omega, \mathcal{F}, P)$. We suppose that $X_i$ has a bounded and compactly supported marginal density $f$, with respect to Lebesgue measure, which does not depend on $i$. We estimate this density from $n$ observations $X_i, i = 1, ..., n$. For any function $f \in L^2(\mathbb{R})$, we can write a formal expansion (see Daubechies (1992)):

$$f = \sum_{k \in \mathbb{Z}} \alpha_{j_0,k} \varphi_{j_0,k} + \sum_{j \geq j_0} \sum_{k \in \mathbb{Z}} \delta_{j,k} \psi_{j,k} = P_{j_0} f + \sum_{j \geq j_0} D_j f$$

where the functions

$$\varphi_{j_0,k}(x) = 2^{j_0/2} \phi(2^{j_0} x - k)$$

and

$$\psi_{j,k}(x) = 2^{j/2} \psi(2^j x - k)$$

constitute an (inhomogeneous) orthonormal basis of $L^2(\mathbb{R})$. Here $\phi(x)$ and $\psi(x)$ are the scale function and the orthogonal wavelet,
respectively. Wavelet coefficients are given by the integrals

$$\alpha_{j_0, k} = \int f(x) \phi_{j_0, k}(x) dx, \delta_{j, k} = \int f(x) \psi_{j, k} dx$$

We suppose that both $\phi$ and $\psi \in C^{r+1}, r \in \mathbb{N}$, have compact supports included in $[-\delta, \delta]$. Note that, by corollary 5.5.2 in Daubechies (1988), $\psi$ is orthogonal to polynomials of degree $\leq r$, i.e.

$$\int \psi(x) x^l dx = 0, \forall l = 0, \ldots, r$$

We suppose that $f$ belongs to the Besov class (see Meyer (1990), §VI.10),

$$\|f\|_{B_p^q} = \{f \in B_p^q, \|f\|_{B_p^q} \leq M\}$$

for some $0 \leq s \leq r + 1, p \geq 1$ and $q \geq 1$, where

$$\|f\|_{B_p^q} = \|P_{j_0} f\|_p + \left( \sum_{j \geq j_0} (\|D_j f\|_p 2^{js})^q \right)^{1/q}$$

We may also say $f \in B_p^q$ if and only if

$$\|\alpha_{j_0, k}\|_p < \infty, \quad \text{and} \quad \left( \sum_{j \geq j_0} (\|\delta_{j, k}\|_p 2^{j(s+1/2-1/p)})^q \right)^{1/q} < \infty \quad (2.1)$$

where $\|\gamma_{j, k}\|_p = (\sum_{k \in \mathbb{Z}} \gamma_{j, k}^p)^{1/p}$. We consider Besov spaces essentially because of their executonal expressive power [see Triebel (1992) and the discussion in Donoho et al. (1995)]. We construct the density estimator [see Prakasa Rao (2003)]

$$\hat{f} = \sum_{k \in K_{j_0}} \hat{\alpha}_{j_0, k} \phi_{j_0, k}, \quad \text{with} \quad \hat{\alpha}_{j_0, k} = \frac{1}{n} \sum_{i=1}^{n} \phi_{j_0, k}(X_i), \quad (2.2)$$

where $K_{j_0}$ is the set of $k$ such that $\text{supp}(f) \cap \text{supp}\phi_{j_0, k} \neq \emptyset$. The fact that $\phi$ has a compact support implies that $K_{j_0}$ is finite and $\text{card}(K_{j_0}) = O(2^{j_0})$. Wavelet density estimators aroused much interest in the recent literature, see Donoho et al. (1996) and Doukhan and Leon (1990). In the case of independent samples the properties of the linear estimator (2.2) have been studied for a variety of error measures and density classes [see Kerkyacharian and Picard (1992), Leblanc (1996) and Tribouley (1995)].

In the setup considered by Prakasa Rao (1996) we assume $\phi$ is a scaling function generating an $r$–regular multiresolution analysis and
$f^{(d)} \in L^2(\mathbb{R})$. Furthermore, we assume that there exists $C_m \geq 0$ and $eta_m \geq 0$ such that
\[
|f^{(m)}(x)| \leq C_m (1 + |x|)^{-\beta_m}, 0 \leq m \leq d. \tag{2.3}
\]
Prakasa Rao (1996) showed that the projection of $f^{(d)}$ on $V_{j_0}$ is
\[
f^{(d)}_{n,d}(x) = \sum_k a_{j_0,k} \phi^{(d)}_{j_0,k}(x),
\]
where
\[
a_{j_0,k} = (-1)^d \int \phi^{(d)}_{j_0,k}(x) f(x) dx.
\]
Hence an estimator of $f^{(d)}$ may be given by
\[
\hat{f}^{(d)}_{n,d}(x) = \sum_k \hat{a}_{j_0,k} \phi^{(d)}_{j_0,k}(x), \tag{2.4}
\]
where
\[
\hat{a}_{j_0,k} = \frac{(-1)^d}{n} \sum_{i=1}^n \phi^{(d)}_{j_0,k}(X_i).
\]
For the estimator in Eq. (2.4), the sum is considered for $k \in K_{j_0}$.

3 Main Results

Theorem 3.1 given below provides bounds on $E \| \hat{f}^{(d)}_{n,d}(x) - f^{(d)}(x) \|_{p'}^2$ for $p' \geq \max(2, p)$, similar to one obtained in the iid case by Prakasa Rao (1996).

**Theorem 3.1.** Let $f^{(d)}(x) \in F_{s,p,q}$ with $s \geq 1/p, p \geq 1, \text{ and } q \geq 1$ then for all $n \geq 2m$ and $p' \geq \max(2, p)$, there exists a constant $C$ such that
\[
E \| \hat{f}^{(d)}_{n,d}(x) - f^{(d)}(x) \|_{p'}^2 \leq C \left( \frac{n}{m} \right)^{-\frac{2(s-1)}{1+2s'}}
\]
where $s' = s + 1/p' - 1/p$ and $2^{j_0} = \left( \frac{n}{m} \right)^{\frac{1}{1+2s'}}$.

**Proof.** First, we decompose $E \| \hat{f}^{(d)}_{n,d}(x) - f^{(d)}(x) \|_{p'}^2$ into a bias term and a stochastic term
\[
E \| \hat{f}^{(d)}_{n,d}(x) - f^{(d)}(x) \|_{p'}^2 \\
\leq 2(\| f^{(d)}_{n,d} - f^{(d)} \|_{p'}^2 + E \| \hat{f}^{(d)}_{n,d} - f^{(d)} \|_{p'}^2) = 2(T_1 + T_2). \tag{3.1}
\]
Now, we find upper bounds for $T_1$ and $T_2$, separately. Note that

$$\sqrt{T_1} = \| \sum_{j \geq j_0} D_j f(d) \|_{p'} \leq \sum_{j \geq j_0} \| D_j f(d) \|_{p'} 2^{-j s'} \leq \{ \sum_{j \geq j_0} (\| D_j f(d) \|_{p'} 2^{j s'})^q \}^{1/q} \{ \sum_{j \geq j_0} 2^{-j s' q'} \}^{1/q'}.$$

Using the Holder’s inequality, with $1/q + 1/q' = 1$, the above equation implies

$$T_1 \leq C \| f(d) \|_{B_{p',q}^s} 2^{-s j_0}. \quad (3.2)$$

Now using the continuous Sobolev injection [see Triebel (1992) and the discussion in Donoho et al. (1996)] implies that $B_{p,q}^s \subset B_{p',q}^{s'}$. Hence one gets,

$$\| f(d) \|_{B_{p',q}^{s'}} \leq \| f(d) \|_{B_{p,q}^s},$$

and in turn, we get from Eq. (3.2)

$$T_1 \leq K 2^{-s j_0}. \quad (3.3)$$

Next, we have

$$T_2 = E \| \hat{f}_{n,d}^{(d)} - f_{n,d}^{(d)} \|_{p'}^2 = E \sum_{k \in K_{j_0}} (\hat{a}_{j_0,k} - a_{j_0,k}) \phi_{j_0,k}(x) \|_{p'}^2.$$

Using Lemma 1 in Leblanc (1996), p. 82 (using Meyer (1990)), the above equation implies,

$$T_2 \leq C E \{ \| \hat{a}_{j_0,k} - a_{j_0,k} \|_{l_{p'}}^2 \} 2^{2j_0(1/2-1/p')}.$$

Further, by using Jensen’s inequality the above equation implies,

$$T_2 \leq C 2^{2j_0(1/2-1/p')} \{ \sum_{k \in K_{j_0}} E |\hat{a}_{j_0,k} - a_{j_0,k}|_{p'} \}^{2/p'}. \quad (3.4)$$

Now, it is enough to find a bound for $E |\hat{a}_{j_0,k} - a_{j_0,k}|_{p'}$ to complete the proof. We know that

$$\hat{a}_{j_0,k} - a_{j_0,k} = \frac{1}{n} \sum_{i=1}^{n} \{ \phi_{j_0,k}^{(d)}(X_i) - a_{j_0,k} \} = \frac{1}{n} \sum_{i=1}^{n} \xi_i,$$

where $\xi_i = [\phi_{j_0,k}^{(d)}(X_i) - a_{j_0,k}]$. 

Wavelet Based Estimation of the Derivatives of ...
Note that $\|\xi_i\|_\infty \leq K, 2^{j_0(1/2+d)}\|\phi\|_\infty, E\xi_i = 0,$

\[ E\xi_i^2 \leq \|f\|_\infty 2^{2j_0d} \int_{-\infty}^{\infty} \phi^2(d)(v)dv \]

and

\[ \hat{\alpha}_{j_0,k}^{(d)} - \alpha_{j_0,k} = \frac{(-1)^d}{n} \sum_{i=1}^{n} \xi_i. \]

Now we need the following result which will be required in the rest of the proof.

**Lemma 3.1.** [Romano and Wolfe (2000), Corollary A.1, p. 121] Let $\{X_i\}$ be a m-dependent sequence of mean zero. Assume $E|X_i|^q \leq \Delta,$ for some $q \geq 2$ and all $i$. Then, for all $n \geq 2m$

\[ E(|\sum_{i=1}^{n} X_i|^q) \leq C_q \Delta (4mn)^{q/2}, \]

where $C_q$ is a positive constant depending only upon $q$.

Using the above result and the fact that $card(K_{j_0}) = O(2^{j_0})$ we have,

\[ \{ \sum_{k \in K_{j_0}} E|\hat{a}_{j_0,k} - a_{j_0,k}|^{p'/p'} \}^{2/p'} \leq \{ C2^{j_0} \frac{1}{n^{p'}} (2^{2j_0d} (4mn)^{p'/2}) \}^{2/p'} \]

\[ \leq K_1 \{ \frac{m}{n} 2^{j_0(1/p'+2d/p')} \}. \]

Now by substituting above inequality in (3.4), we get

\[ T_2 \leq K_2 2^{2j_0(1/2-1/p')} \{ \frac{m}{n} 2^{2j_0(1/p'+2d/p')} \} \]

\[ = K_2 \{ 2^{2j_0(1/2+2d/p')} \frac{m}{n} \} \]

\[ \leq K_3 \{ 2^{2j_0(1+2d)} \frac{m}{n} \}. \] \hspace{1cm} (3.5)

By Substituting (3.3), (3.5), and $2^{j_0} = (\frac{n}{m})^{\frac{1}{1+2d}}$ in (3.1), theorem is proved.

**Remark 3.1.** Letting $d = 0$ in Theorem 3.1 the results of Doosti and Nezakati (2006) are obtained.
Remark 3.2. If one considers $m$ as a fixed integer, then it can be shown that the upper bound in Theorem 3.1 is similar to the result of Chaubey et al. (2006a, 2006b).

Remark 3.3. Suppose $1 < p' \leq 2$. One can get upper bounds similar to those given in Theorem 3.1 for the expected loss $E\|\hat{f}_{n,d}^{(d)} - f(d)\|_{p'}$, as explained below. Observe that

$$E\|\hat{f}^{(d)}\|_{p'} \leq 2^{p'-1}(\|f_{n,d}^{(d)} - f(d)\|_{p'} + E\|\hat{f}_{n,d}^{(d)} - f_{n,d}^{(d)}\|_{p'}) \quad (3.6)$$

$$\|f_{n,d}^{(d)} - f(d)\|_{p'} \leq C_1 2^{-p' s_j} \quad (3.7)$$

$$E\|\hat{f}_{n,d}^{(d)} - f_{n,d}^{(d)}\|_{p'} \leq C_2 2^{j_0(p' / 2 - 1)} \left\{ \sum_{k \in K_{j_0}} E|\hat{a}_{j_0,k} - a_{j_0,k}|_{p'} \right\}$$

$$\leq C_2 2^{j_0(p' / 2 - 1)} \left\{ \sum_{k \in K_{j_0}} \sqrt{E|\hat{a}_{j_0,k} - a_{j_0,k}|_{2p'}} \right\}$$

$$\leq C_3 2^{j_0(p' / 2 - 1)} \left\{ 2^{j_0} \sqrt{2^{j_0 d} \frac{(4mn)^{p'}}{n^{2p'}}} \right\}$$

$$= C_3 2^{j_0(p' - 1 + d)} \frac{m}{n^{p'/2}}. \quad (3.8)$$

for some positive constant $C_1$, $C_2$ and $C_3$.

References


Chaubey and Doosti


