Hindawi
Journal of Function Spaces
Volume 2019, Article ID 7102346, 6 pages
https://doi.org/10.1155/2019/7102346

Research Article

Strong Uniform Convergence Rates of Wavelet Density Estimators with Size-Biased Data

Huijun Guo and Junke Kou

School of Mathematics and Computational Science, Guilin University of Electronic Technology, Guilin, Guangxi 541004, China

Correspondence should be addressed to Junke Kou; kjkou@guet.edu.cn

Received 15 December 2018; Accepted 20 February 2019; Published 6 March 2019

Academic Editor: Raúl E. Curto

This paper considers the strong uniform convergence of multivariate density estimators in Besov space $B^s_{p,q}(\mathbb{R}^d)$ based on size-biased data. We provide convergence rates of wavelet estimators when the parametric μ is known or unknown, respectively. It turns out that the convergence rates coincide with that of Giné and Nickl's (*Uniform Limit Theorems for Wavelet Density Estimators, Ann. Probab., 37(4), 1605-1646, 2009*), when the dimension $d = 1$, $p = q = \infty$, and $\omega(y) \equiv 1$.

1. Introduction

Let $Y_1, Y_2, \ldots, Y_n$ be independent and identically distributed ($i.i.d$) continuous random variables defined on a probability space $(\Omega, \mathscr{F}, P)$ with the common density function

$$g(y) = \frac{\omega(y) f(y)}{\mu}, \quad y \in \mathbb{R}^d, \tag{1}$$

where ω denotes a known positive function and f stands for an unknown density function of the unobserved continuous random variable X and $\mu = E\omega(X) = \int_{\mathbb{R}^d} \omega(y) f(y) dy < +\infty$. In this setup f and g mean the target density and weighted density function, respectively, and the resulting data are size-biased data. Then we want to estimate the unknown density function f from a sequence of biased data $Y_1, Y_2, \ldots, Y_n$.

Wavelet methods are of interest in nonparametric statistics thanks to their ability to estimate efficiently a wide variety of unknown functions, especially for those with discontinuities or sharp spikes. Hence, wavelet methods have been widely used for this density estimation model (1). Ramírez and Vidakovic [1] propose a linear wavelet estimator and show it to be L^2 consistent. Shirazi and Doosti [2] expand their work to multivariate case. Chesneau, Dewan, and Doosti [3] extend the independence to both positively and negatively associated cases. They show a convergence rate for mean integrated squared error (MISE). An upper bound of wavelet estimation on $L^p (1 \le p < +\infty)$ risk in negatively associated case is given by Liu and Xu [4]. Kou and Guo [5] discuss the MISE of wavelet estimators in strong mixing case. For the strong convergence of density estimation, Masry [6] studies the strong convergence rates over a compact subset in Besov space $B^s_{p,q}(\mathbb{R}^d)$, when $\omega(y) \equiv 1$ (the model (1) reduces to the classical density estimation) and the sample is strong mixing. Recently, Giné and Nickl [7] investigate the same problem by wavelet method and obtain the optimal strong convergence rates in Besov space $B^s_{\infty,\infty}(\mathbb{R})$, when the data is $i.i.d$. To our knowledge, there does not exist research on the strong uniform convergence for the model (1).

The aim of this paper is to discuss the strong uniform convergence rates of wavelet estimators in Besov space $B^s_{p,q}(\mathbb{R}^d)$ based on size-biased data. First of all, we construct a linear wavelet estimator $\widetilde{f}_n$ when the parametric μ is known and give its convergence rate. However, people always do not know μ in many practical applications. For this reason, an estimator $\widehat{\mu}$ of μ is given. Then we develop a new linear wavelet estimator $\widehat{f}_n$ in which the parametric μ is replaced by $\widehat{\mu}$. Finally, we establish the convergence rate of estimator $\widehat{f}_n$.

2. Wavelets and Besov Spaces

As a central notion in wavelet analysis, Multiresolution Analysis (MRA, [8]) plays an important role for constructing a wavelet basis, which means a sequence of closed subspaces $\{V_j\}_{j\in\mathbb{Z}}$ of the square integrable function space $L^2(\mathbb{R}^d)$ satisfying the following properties:

(i) $V_j \subseteq V_{j+1}$, $j \in \mathbb{Z}$. Here and after, $\mathbb{Z}$ denotes the integer set and $\mathbb{N} := \{n \in \mathbb{Z}, n \geq 0\}$;

(ii) $\overline{\bigcup_{j\in\mathbb{Z}} V_j} = L^2(\mathbb{R}^d)$. This means the space $\bigcup_{j\in\mathbb{Z}} V_j$ being dense in $L^2(\mathbb{R}^d)$;

(iii) $f(2\cdot) \in V_{j+1}$ if and only if $f(\cdot) \in V_j$ for each $j \in \mathbb{Z}$;

(iv) There exists a scaling function $\varphi \in L^2(\mathbb{R}^d)$ such that $\{\varphi(\cdot - k), k \in \mathbb{Z}^d\}$ forms an orthonormal basis of $V_0 = \overline{\text{span}}\{\varphi(\cdot - k)\}$.

When $d = 1$, there is a simple way to define an orthonormal wavelet basis. Examples include the Daubechies wavelets with compact supports. For $d \geq 2$, the tensor product method gives an MRA $\{V_j\}$ of $L^2(\mathbb{R}^d)$ from one-dimensional MRA. In fact, with a scaling function φ of tensor products, we find $M = 2^d - 1$ wavelet functions $\psi^\ell(\ell = 1, 2, \ldots, M)$ such that, for each $f \in L^2(\mathbb{R}^d)$, the decomposition

$$f = \sum_{k\in\mathbb{Z}^d} \alpha_{j_0,k}\varphi_{j_0,k} + \sum_{j=j_0}^{\infty}\sum_{\ell=1}^{M}\sum_{k\in\mathbb{Z}^d} \beta_{j,k}^\ell \psi_{j,k}^\ell \quad (2)$$

holds in $L^2(\mathbb{R}^d)$ sense, where $\alpha_{j_0,k} = \langle f, \varphi_{j_0,k}\rangle$, $\beta_{j,k}^\ell = \langle f, \psi_{j,k}^\ell\rangle$, and

$$\begin{aligned} \varphi_{j_0,k}(y) &= 2^{j_0 d/2}\varphi\left(2^{j_0}y - k\right), \\ \psi_{j,k}^\ell(y) &= 2^{jd/2}\psi^\ell\left(2^{j}y - k\right). \end{aligned} \quad (3)$$

Let P_j be the orthogonal projection operator from $L^2(\mathbb{R}^d)$ onto the space V_j with the orthonormal basis $\{\varphi_{j,k}(\cdot) = 2^{jd/2}\varphi(2^j \cdot -k), k \in \mathbb{Z}^d\}$. Then for $f \in L^2(\mathbb{R}^d)$, $P_j f = \sum_{k\in\mathbb{Z}^d} \alpha_{j,k}\varphi_{j,k}$.

If a scaling function φ satisfies Condition (θ), i.e.,

$$\sum_{k\in\mathbb{Z}^d} |\varphi(y-k)| \in L^\infty\left(\mathbb{R}^d\right), \quad (4)$$

then the function $\varphi \in L(\mathbb{R}^d)\bigcap L^\infty(\mathbb{R}^d)$(so that $\varphi \in L^p$ for $1 \leq p \leq \infty$) and $\sum_{k\in\mathbb{Z}^d}\varphi(x-k)\overline{\varphi(y-k)}$ converges absolutely almost everywhere. It can be shown that, for $f \in L^p(\mathbb{R}^d)$ $(1 \leq p \leq \infty)$,

$$P_j f(y) = \sum_{k\in\mathbb{Z}^d} \alpha_{j,k}\varphi_{j,k}(y) \quad (5)$$

holds almost everywhere on $\mathbb{R}^d$ [9]. In this paper, we also need another concept, which is a little stronger than Condition (θ).

A function φ is said to satisfy Condition (S), if there exists a bounded and radical nonincreasing function Φ such that

$$|\varphi(y)| \leq \Phi(|y|)\ (a.e.) \quad \text{and} \quad \int_{\mathbb{R}^d} \Phi(|y|)\,dy < \infty. \quad (6)$$

Condition (S) is not very restrictive. Examples include bounded and compactly supported measurable functions. Daubechies scaling functions satisfy Condition (S).

A wavelet basis can be used to characterize Besov spaces. The next lemma provides equivalent definitions for those spaces, for which we need one more notation: a scaling function φ is called m-regular, if $\varphi \in C^m(\mathbb{R}^d)$ and $|D^\alpha\varphi(y)| \leq c(1+|y|^2)^{-\ell}$ for each $\ell \in \mathbb{Z}$ and each multi-index $\alpha \in \mathbb{N}^d$ with $|\alpha| \leq m$.

Lemma 1 ([8]). *Let φ be m-regular, ψ^ℓ $(\ell = 1, 2, \ldots, M, M = 2^d - 1)$ be the corresponding wavelets and $f \in L^p(\mathbb{R}^d)$. If $\alpha_{j,k} = \langle f, \varphi_{j,k}\rangle$, $\beta_{j,k}^\ell = \langle f, \psi_{j,k}^\ell\rangle$, $p, q \in [1, +\infty]$, and $0 < s < m$, then the following assertions are equivalent:*

(1) $f \in B_{p,q}^s(\mathbb{R}^d)$;

(2) $\{2^{js}\|P_j f - f\|_p\} \in l_q$;

(3) $\|(\alpha_{j_0})\|_p + \|(2^{j(s-d/p+d/2)}\|\beta_j\|_p)_{j\geq j_0}\|_q < +\infty$.

The Besov norm of f can be defined by

$$\|f\|_{B_{p,q}^s} := \left\|\left(\alpha_{j_0}\right)\right\|_p + \left\|\left(2^{j(s-d/p+d/2)}\left\|\beta_j\right\|_p\right)_{j\geq j_0}\right\|_q \quad (7)$$

with $\|(\alpha_{j_0})\|_p^p = \sum_{k\in\mathbb{Z}^d}|\alpha_{j_0,k}|^p$ and $\|\beta_j\|_p^p = \sum_{\ell=1}^M\sum_{k\in\mathbb{Z}^d}|\beta_{j,k}^\ell|^p$.

We also need the following classical inequality in the proof of our theorems.

Bernstein's inequality. Let $Y_1, \ldots, Y_n$ be independent random variables such that $EY_i = 0$, $|Y_i| \leq M$, and $EY_i^2 = \sigma^2$. Then for each $v \geq 0$,

$$\mathbb{P}\left\{\frac{1}{n}\left|\sum_{i=1}^n Y_i\right| \geq v\right\} \leq 2\cdot\exp\left\{-\frac{nv^2}{2\left(\sigma^2 + vM/3\right)}\right\}. \quad (8)$$

3. Estimation with Known μ

In this paper, we require supp $\mathrm{Y_i} \subseteq [0,1]^d$ in the model (1). This is similar to Chesneau, Dewan, and Doosti [3], Liu and Xu [4], and Kou and Guo [5]. We choose d-dimensional scaling function

$$\varphi(y) = \varphi(y_1, \ldots, y_d) := D_{2N}(y_1)\cdots D_{2N}(y_d) \quad (9)$$

with $D_{2N}(\cdot)$ being the one-dimensional Daubechies scaling function. Then φ is m-regular $(m > 0)$ when N gets large enough. Note that D_{2N} has compact support $[0, 2N-1]$ and the corresponding wavelet has compact support $[-N+1, N]$. Then for $f \in L^2(\mathbb{R}^d)$ with supp $f \subseteq [0,1]^d$ and $M = 2^d - 1$,

$$f(y) = \sum_{k\in\Lambda_{j_0}} \alpha_{j_0,k}\varphi_{j_0,k}(y) + \sum_{j=j_0}^{\infty}\sum_{\ell=1}^{M}\sum_{k\in\Lambda_j} \beta_{j,k}^\ell\psi_{j,k}^\ell(y), \quad (10)$$

where $\Lambda_{j_0} = \{1-2N, 2-2N, \ldots, 2^{j_0}\}^d$, $\Lambda_j = \{-N, -N+1, \ldots, 2^j + N - 1\}^d$, and

$$\begin{aligned} \alpha_{j_0,k} &= \int_{[0,1]^d} f(y)\varphi_{j_0,k}(y)\,dy, \\ \beta_{j,k}^\ell &= \int_{[0,1]^d} f(y)\psi_{j,k}^\ell(y)\,dy. \end{aligned} \quad (11)$$

A linear wavelet estimator is defined by

$$\widetilde{f}_n(y) = \sum_{k\in\Lambda_{j_0}} \widetilde{\alpha}_{j_0,k}\varphi_{j_0,k}(y), \tag{12}$$

where

$$\widetilde{\alpha}_{j_0,k} = \frac{\mu}{n}\sum_{i=1}^{n}\frac{\varphi_{j_0,k}(Y_i)}{\omega(Y_i)}. \tag{13}$$

It follows from (1) that

$$\begin{aligned} E\left[\frac{\mu}{n}\sum_{i=1}^{n}\frac{\varphi_{j_0,k}(Y_i)}{\omega(Y_i)}\right] &= E\left[\frac{\mu\varphi_{j_0,k}(Y_i)}{\omega(Y_i)}\right] \\ &= \int_{[0,1]^d}\frac{\mu\varphi_{j_0,k}(y)}{\omega(y)}g(y)\,dy \\ &= \int_{[0,1]^d}\varphi_{j_0,k}(y)f(y)\,dy = \alpha_{j_0,k}. \end{aligned} \tag{14}$$

This means $\widetilde{\alpha}_{j_0,k}$ is an unbiased estimate of $\alpha_{j_0,k}$. The following notations are needed to state our theorems. $A \lesssim B$ denotes $A \le cB$ for some constant $c > 0$; $A \gtrsim B$ means $B \lesssim A$; $A \sim B$ stands for both $A \lesssim B$ and $B \lesssim A$.

Theorem 2. *Consider the problem (1) with $\omega(y) \sim 1$. Let $f \in B^s_{p,q}(\mathbb{R}^d)(p, q \in [1,\infty], s > d/p)$ and $\operatorname{supp} f \subseteq [0,1]^d$. Then the linear wavelet estimator $\widetilde{f}_n$ defined in (12) with $2^{j_0} \sim (n/\ln n)^{1/(2(s-d/p)+d)}$ satisfies*

$$\begin{aligned} &\sup_{y\in[0,1]^d}\left|\widetilde{f}_n(y) - f(y)\right| \\ &\quad = O_{a.s.}\left(\frac{\ln n}{n}\right)^{(s-d/p)/(2(s-d/p)+d)}. \end{aligned} \tag{15}$$

Remark 3. When $\omega(y) \equiv 1$, our model reduces to the classical nonparametric density estimation. Then our result is same as the convergence rate in Masry [6]. On the other hand, we find that

$$\sup_{y\in[0,1]}\left|\widetilde{f}_n(y) - f(y)\right| = O_{a.s.}\left(\frac{\ln n}{n}\right)^{s/(2s+1)} \tag{16}$$

with $d = 1$ and $p = q = \infty$. This coincides with the convergence rate in Theorem 3 of Giné and Nickl [7].

Proof. It is easy to see that

$$\begin{aligned} &\sup_{y\in[0,1]^d}\left|\widetilde{f}_n(y) - f(y)\right| \\ &\quad \le \sup_{y\in[0,1]^d}\left|\widetilde{f}_n(y) - E\widetilde{f}_n(y)\right| \\ &\qquad + \sup_{y\in[0,1]^d}\left|E\widetilde{f}_n(y) - f(y)\right|. \end{aligned} \tag{17}$$

By $s > d/p$, $B^s_{p,q}(\mathbb{R}^d) \subseteq B^{s-d/p}_{\infty,\infty}(\mathbb{R}^d)$. Then it follows from (14) and Lemma 1 ($p = \infty$) that

$$\begin{aligned} \sup_{y\in[0,1]^d}\left|E\widetilde{f}_n(y) - f(y)\right| &\le \sup_{y\in\mathbb{R}^d}\left|E\widetilde{f}_n(y) - f(y)\right| \\ &= \left\|P_{j_0}f - f\right\|_\infty \lesssim 2^{-j_0(s-d/p)}. \end{aligned} \tag{18}$$

This with the choice $2^{j_0} \sim (n/\ln n)^{1/(2(s-d/p)+d)}$ leads to

$$\sup_{y\in[0,1]^d}\left|E\widetilde{f}_n(y) - f(y)\right| \lesssim \left(\frac{\ln n}{n}\right)^{(s-d/p)/(2(s-d/p)+d)}. \tag{19}$$

To estimate the other term of (17), by splitting the interval $[0,1]$ equally into

$$L_n = \left\lceil\left(\frac{n2^{j_0(3d+2)}}{\ln n}\right)^{1/2}\right\rceil \tag{20}$$

($\lceil x\rceil$ standing for the smallest integer greater than or equal to x) subintervals, one receives L_n^d sub-cubes $I^{(\ell)}(\ell = 1, 2, \ldots, L_n^d)$ of $[0,1]^d$. Clearly, the side length l_n of $I^{(\ell)}$ satisfies that

$$l_n \lesssim \left(\frac{n2^{j_0(3d+2)}}{\ln n}\right)^{-1/2}. \tag{21}$$

Note that

$$\begin{aligned} &\sup_{y\in[0,1]^d}\left|\widetilde{f}_n(y) - E\widetilde{f}_n(y)\right| \\ &\quad = \max_{1\le\ell\le L_n^d}\sup_{y\in I^{(\ell)}}\left|\widetilde{f}_n(y) - E\widetilde{f}_n(y)\right|. \end{aligned} \tag{22}$$

Then with the center point $y^{(\ell)}$ of $I^{(\ell)}$,

$$\sup_{y\in[0,1]^d}\left|\widetilde{f}_n(y) - E\widetilde{f}_n(y)\right| \le Q_1 + Q_2 + Q_3, \tag{23}$$

where

$$\begin{aligned} Q_1 &:= \max_{1\le\ell\le L_n^d}\sup_{y\in I^{(\ell)}}\left|\widetilde{f}_n(y) - \widetilde{f}_n\left(y^{(\ell)}\right)\right|, \\ Q_2 &:= \max_{1\le\ell\le L_n^d}\left|\widetilde{f}_n\left(y^{(\ell)}\right) - E\widetilde{f}_n\left(y^{(\ell)}\right)\right|, \\ Q_3 &:= \max_{1\le\ell\le L_n^d}\sup_{y\in I^{(\ell)}}\left|E\widetilde{f}_n(y) - E\widetilde{f}_n\left(y^{(\ell)}\right)\right|. \end{aligned} \tag{24}$$

By the definition of $\widetilde{f}_n(y)$,

$$Q_1 \le \max_{1\le\ell\le L_n^d}\sup_{y\in I^{(\ell)}}\sum_{k\in\Lambda_{j_0}}\left|\widetilde{\alpha}_{j_0,k}\right|\left|\varphi_{j_0,k}(y) - \varphi_{j_0,k}\left(y^{(\ell)}\right)\right|. \tag{25}$$

Since $\omega(y) \sim 1$ the properties of φ imply $|\widetilde{\alpha}_{j_0,k}| \lesssim 2^{j_0 d/2}$. On the other hand, the Daubechies function D_{2N} satisfies

Lipschitz condition ($|D_{2N}(x) - D_{2N}(y)| \lesssim |x - y|$) for larger N. Then for $\varphi(y) = \prod_{i=1}^{d} D_{2N}(y_i)$,

$$\begin{aligned} &\left|\varphi(x) - \varphi(y)\right| \\ &\quad \le d \left\|D_{2N}\right\|_{\infty}^{d-1} \sup_{1\le i\le d} \left|D_{2N}(x_i) - D_{2N}(y_i)\right| \\ &\quad \lesssim \sup_{1\le i\le d} \left|x_i - y_i\right|. \end{aligned} \tag{26}$$

Hence, for any $y \in I^{(\ell)}$,

$$\begin{aligned} &\left|\varphi_{j_0,k}(y) - \varphi_{j_0,k}\left(y^{(\ell)}\right)\right| \lesssim 2^{j_0 d/2} \sup_{1\le i\le d} 2^{j_0} \left|y_i - y_i^{(\ell)}\right| \\ &\quad \lesssim 2^{j_0(d/2+1)} \ell_n. \end{aligned} \tag{27}$$

Combining this with (25) and $|\Lambda_{i_0}| \lesssim 2^{j_0 d}$, one finds that

$$Q_1 \lesssim 2^{j_0(2d+1)} \ell_n. \tag{28}$$

Recalling that $l_n \lesssim (n 2^{j_0(3d+2)}/\ln n)^{-1/2}$ and $2^{j_0} \sim (n/\ln n)^{1/(2(s-d/p)+d)}$, then

$$Q_1 \lesssim \left(\frac{\ln n}{n}\right)^{(s-d/p)/(2(s-d/p)+d)}. \tag{29}$$

By $|X| \le \widetilde{M}$, $|EX| \le \widetilde{M}$. Furthermore, it follows from the proof of (29) that

$$\begin{aligned} Q_3 &= \max_{1\le\ell\le L_n^d} \sup_{y\in I^{(\ell)}} \left|E\widetilde{f}_n(y) - E\widetilde{f}_n\left(y^{(\ell)}\right)\right| \\ &= \max_{1\le\ell\le L_n^d} \sup_{y\in I^{(\ell)}} \left|E\left[\widetilde{f}_n(y) - \widetilde{f}_n\left(y^{(\ell)}\right)\right]\right| \\ &\lesssim \left(\frac{\ln n}{n}\right)^{(s-d/p)/(2(s-d/p)+d)}. \end{aligned} \tag{30}$$

The main work for the proof of Theorem 2 is to estimate

$$Q_2 = \max_{1\le\ell\le L_n^d} \left|\widetilde{f}_n\left(y^{(\ell)}\right) - E\widetilde{f}_n\left(y^{(\ell)}\right)\right|. \tag{31}$$

Set $\eta_n := (\ln n/n)^{(s-d/p)/(2(s-d/p)+d)}$ and $c_* > 0$ is constant which will be chosen later. Then note that

$$\begin{aligned} &\mathbb{P}\left\{Q_2 \ge c_*\eta_n\right\} \\ &\quad = \mathbb{P}\left\{\max_{1\le\ell\le L_n^d} \left|\widetilde{f}_n\left(y^{(\ell)}\right) - E\widetilde{f}_n\left(y^{(\ell)}\right)\right| \ge c_*\eta_n\right\} \\ &\quad \le \sum_{\ell=1}^{L_n^d} \mathbb{P}\left\{\left|\widetilde{f}_n\left(y^{(\ell)}\right) - E\widetilde{f}_n\left(y^{(\ell)}\right)\right| \ge c_*\eta_n\right\} \\ &\quad \le L_n^d \sup_{y\in[0,1]^d} \mathbb{P}\left\{\left|\widetilde{f}_n(y) - E\widetilde{f}_n(y)\right| \ge c_*\eta_n\right\}. \end{aligned} \tag{32}$$

According to the definition of $\widetilde{f}_n(y)$, one concludes

$$\begin{aligned} &\widetilde{f}_n(y) - E\widetilde{f}_n(y) = \sum_{k\in\Lambda_{j_0}} \left(\widetilde{\alpha}_{j_0,k} - E\widetilde{\alpha}_{j_0,k}\right)\varphi_{j_0,k}(y) = \frac{1}{n} \\ &\quad \cdot \sum_{i=1}^{n} \sum_{k\in\Lambda_{j_0}} \left[\frac{\mu\varphi_{j_0,k}(Y_i)}{\omega(Y_i)} - E\left(\frac{\mu\varphi_{j_0,k}(Y_i)}{\omega(Y_i)}\right)\right] \\ &\quad \cdot \varphi_{j_0,k}(y). \end{aligned} \tag{33}$$

Denote

$$\begin{aligned} &Z_i(y) \\ &\quad := \sum_{k\in\Lambda_{j_0}} \left[\frac{\mu\varphi_{j_0,k}(Y_i)}{\omega(Y_i)} - E\left(\frac{\mu\varphi_{j_0,k}(Y_i)}{\omega(Y_i)}\right)\right]\varphi_{j_0,k}(y) \end{aligned} \tag{34}$$

for $i = 1, 2, \ldots, n$. Then $Z_1, Z_2, \ldots, Z_n$ are *i.i.d*, $E(Z_i) = 0$. By $\omega(y) \sim 1$ and Condition (θ), $|Z_i(y)| \lesssim 2^{j_0 d}$ and

$$\begin{aligned} E(Z_i)^2 &= \operatorname{var}\left[\sum_{k\in\Lambda_{j_0}} \frac{\mu\varphi_{j_0,k}(Y_i)}{\omega(Y_i)}\varphi_{j_0,k}(y)\right] \\ &\le E\left[\sum_{k\in\Lambda_{j_0}} \frac{\mu\varphi_{j_0,k}(Y_i)}{\omega(Y_i)}\varphi_{j_0,k}(y)\right]^2 \\ &\lesssim \int_{[0,1]^d} \left|\sum_{k\in\Lambda_{j_0}} \varphi_{j_0,k}(x)\varphi_{j_0,k}(y)\right|^2 g(x)\,dx \\ &\lesssim 2^{j_0 d}. \end{aligned} \tag{35}$$

This with Bernstein's inequality (Härdle et al., 1998) and $\eta_n = (\ln n/n)^{(s-d/p)/(2(s-d/p)+d)}$ leads to

$$\begin{aligned} \mathbb{P}\left\{\left|\widetilde{f}_n(y) - E\widetilde{f}_n(y)\right| \ge c_*\eta_n\right\} &= \mathbb{P}\left\{\left|\frac{1}{n}\sum_{i=1}^{n} Z_i\right| \ge c_*\eta_n\right\} \le 2\exp\left\{-\frac{n(c_*\eta_n)^2}{2\left(2^{j_0 d} + 2^{j_0 d}c_*\eta_n/3\right)}\right\} \\ &\le 2\exp\left\{-\frac{nc_*^2(\ln n/n)^{2(s-d/p)/(2(s-d/p)+d)}}{2\left[(n/\ln n)^{d/(2(s-d/p)+d)} + (1/3)\,c_*\,(n/\ln n)^{d/(2(s-d/p)+d)}\right]}\right\} \\ &\le 2\exp\left\{-c_*^2 \ln n/\left(2 + (1/3)\,c_*\right)\right\} = 2n^{-c_*^2/(2+(1/3)c_*)}. \end{aligned} \tag{36}$$

It follows from (32), (36), $2^{j_0} \sim (n/\ln n)^{1/(2(s-d/p)+d)}$, and the definition of L_n that

$$\begin{aligned}\mathbb{P}\{Q_2 \geq c_*\eta_n\} &\leq L_n^d \sup_{y\in[0,1]^d} \mathbb{P}\left\{\left|\widetilde{f}_n(y) - E\widetilde{f}_n(y)\right| \geq c_*\eta_n\right\} \\ &\leq \left(\frac{n2^{j_0(3d+2)}}{\ln n}\right)^{d/2} 2n^{-c_*^2/(2+(1/3)c_*)} \\ &\leq 2n^{d(s-d/p+2d+1)/(2(s-d/p)+d)-c_*^2/(2+(1/3)c_*)}.\end{aligned} \tag{37}$$

Obviously, there exists sufficiently large $c_* > 0$ such that $n^{d(s-d/p+2d+1)/(2(s-d/p)+d)-c_*^2/(2+(1/3)c_*)} \leq 2n^{-2}$. Then $\mathbb{P}\{Q_2 \geq c_*\eta_n\} \leq 2n^{-2}$ and

$$\sum_{n=1}^{\infty} \mathbb{P}\{Q_2 \geq c_*\eta_n\} < +\infty. \tag{38}$$

Hence,

$$\begin{aligned}Q_2 &= \max_{1\leq \ell \leq L_n^d} \left|\widetilde{f}_n\left(y^{(\ell)}\right) - E\widetilde{f}_n\left(y^{(\ell)}\right)\right| \\ &= O_{a.s.}\left(\frac{\ln n}{n}\right)^{(s-d/p)/(2(s-d/p)+d)}\end{aligned} \tag{39}$$

thanks to Borel-Cantelli lemma. This with (23), (29), and (30) shows

$$\begin{aligned}&\sup_{y\in[0,1]^d} \left|\widetilde{f}_n(y) - E\widetilde{f}_n(y)\right| \\ &\quad = O_{a.s.}\left(\frac{\ln n}{n}\right)^{(s-d/p)/(2(s-d/p)+d)}.\end{aligned} \tag{40}$$

Combining this with (17) and (19), one knows that

$$\begin{aligned}&\sup_{y\in[0,1]^d} \left|\widetilde{f}_n(y) - f(y)\right| \\ &\quad = O_{a.s.}\left(\frac{\ln n}{n}\right)^{(s-d/p)/(2(s-d/p)+d)}.\end{aligned} \tag{41}$$

□

A careful observation of (12) shows the construction of $\widetilde{f}_n(y)$ strictly depends on μ, which needs μ known. However, the parametric μ is always unknown in many practical applications. So we will deal with the unknown case in the following section.

4. Estimation with Unknown μ

In this section, we provide a strong convergence rate of wavelet estimator for the model (1) with unknown parametric μ. A first step is to give an estimator of μ from the given data $Y_1, Y_2, \ldots, Y_n$. Similar to Chesneau, Dewan, and Doosti [3] and Liu and Xu [4], we introduce

$$\widehat{\mu}_n = \left[\frac{1}{n}\sum_{i=1}^{n} \frac{1}{\omega(Y_i)}\right]^{-1}. \tag{42}$$

By (1),

$$\begin{aligned}E\left(\frac{1}{\widehat{\mu}_n}\right) &= E\left[\frac{1}{n}\sum_{i=1}^{n}\frac{1}{\omega(Y_i)}\right] = E\left[\frac{1}{\omega(Y_i)}\right] \\ &= \int_{[0,1]^d} \frac{g(y)}{\omega(y)} dy = \frac{1}{\mu}\int_{[0,1]^d} f(y)\, dy = \frac{1}{\mu}.\end{aligned} \tag{43}$$

Now, we define a practical linear wavelet estimator

$$\widehat{f}_n(y) = \sum_{k\in\Lambda_{j_0}} \widehat{\alpha}_{j_0,k}\varphi_{j_0,k}(y) \tag{44}$$

with

$$\widehat{\alpha}_{j_0,k} = \frac{\widehat{\mu}_n}{n}\sum_{i=1}^{n}\frac{\varphi_{j_0,k}(Y_i)}{\omega(Y_i)}. \tag{45}$$

Theorem 4 investigates the strong uniform convergence rate of practical wavelet estimator $\widehat{f}_n(y)$.

Theorem 4. *Consider the problem (1) with $\omega(y) \sim 1$. Let $f \in B^s_{p,q}(\mathbb{R}^d)(p, q \in [1,\infty], s > d/p)$ and $\operatorname{supp} f \subseteq [0,1]^d$. Then the linear wavelet estimator $\widehat{f}_n$ defined in (44) with $2^{j_0} \sim (n/\ln n)^{1/(2(s-d/p)+d)}$ satisfies*

$$\begin{aligned}&\sup_{y\in[0,1]^d} \left|\widehat{f}_n(y) - f(y)\right| \\ &\quad = O_{a.s.}\left(\frac{\ln n}{n}\right)^{(s-d/p)/(2(s-d/p)+d)}.\end{aligned} \tag{46}$$

Remark 5. Note that the convergence rate of wavelet estimator $\widehat{f}_n(y)$ in Theorem 4 remains same as that of $\widetilde{f}_n(y)$ in Theorem 2. However, the estimator $\widehat{f}_n(y)$ does not depend on the parametric μ, which means it is more practical.

Proof. By the definition of $\widehat{\alpha}_{j_0,k}$ and $\widetilde{\alpha}_{j_0,k}$,

$$\begin{aligned}\widehat{\alpha}_{j_0,k} &= \frac{\widehat{\mu}_n}{\mu}\widetilde{\alpha}_{j_0,k} \\ \text{and } \widehat{f}_n(y) &= \frac{\widehat{\mu}_n}{\mu}\widetilde{f}_n(y).\end{aligned} \tag{47}$$

Then one observes that

$$\begin{aligned}\sup_{y\in[0,1]^d}\left|\widehat{f}_n(y) - f(y)\right| &= \sup_{y\in[0,1]^d}\left|\frac{\widehat{\mu}_n}{\mu}\widetilde{f}_n(y) - f(y)\right| \\ &\leq T_1 + T_2 + T_3\end{aligned} \tag{48}$$

where

$$\begin{aligned}T_1 &:= \sup_{y\in[0,1]^d}\left|\frac{\widehat{\mu}_n}{\mu}\left[\widetilde{f}_n(y) - E\widetilde{f}_n(y)\right]\right|, \\ T_2 &:= \sup_{y\in[0,1]^d}\left|\left(\frac{\widehat{\mu}_n}{\mu} - 1\right)E\widetilde{f}_n(y)\right|, \\ T_3 &:= \sup_{y\in[0,1]^d}\left|f(y) - E\widetilde{f}_n(y)\right|.\end{aligned} \tag{49}$$

It follows from (19) that

$$T_3 = \sup_{y\in[0,1]^d} \left| f(y) - E\widetilde{f}_n(y) \right| \lesssim \left(\frac{\ln n}{n}\right)^{(s-d/p)/(2(s-d/p)+d)}. \quad (50)$$

According to $\omega(y) \sim 1$ and the definition of $\widehat{\mu}_n$ in (42), one gets $|\widehat{\mu}_n/\mu| \lesssim 1$ and

$$\sup_{y\in[0,1]^d} \left| \frac{\widehat{\mu}_n}{\mu} \left[\widetilde{f}_n(y) - E\widetilde{f}_n(y) \right] \right| \lesssim \sup_{y\in[0,1]^d} \left| \widetilde{f}_n(y) - E\widetilde{f}_n(y) \right|. \quad (51)$$

Then it is easy to see from (40) that

$$T_1 = O_{a.s.}\left(\frac{\ln n}{n}\right)^{(s-d/p)/(2(s-d/p)+d)}. \quad (52)$$

Finally, one estimate T_2. Since $\omega(y) \sim 1$, $|\widehat{\mu}_n| \lesssim 1$. On the other hand, $E|\widetilde{f}_n(y)| \lesssim 1$ thanks to Condition (S) and Lemma 3.3 in Liu and Xu [10]. Hence,

$$\begin{aligned} T_2 &= \sup_{y\in[0,1]^d} \left| \left(\frac{\widehat{\mu}_n}{\mu} - 1 \right) E\widetilde{f}_n(y) \right| \\ &\leq \sup_{y\in[0,1]^d} |\widehat{\mu}_n| \left| \frac{1}{\widehat{\mu}_n} - \frac{1}{\mu} \right| E\left|\widetilde{f}_n(y)\right| \lesssim \left| \frac{1}{\widehat{\mu}_n} - \frac{1}{\mu} \right| \\ &= \left| \frac{1}{n} \sum_{i=1}^{n} \left(\frac{1}{\omega(Y_i)} - \frac{1}{\mu} \right) \right|. \end{aligned} \quad (53)$$

Take $W_i = 1/\omega(Y_i) - 1/\mu$. Then $W_1, W_2, \ldots, W_n$ is $i.i.d$ and $EW_i = 0$ (see (43)). By $\omega(y) \sim 1$, $|W_i| \lesssim 1$ and $E(W_i)^2 \lesssim 1$. Similar to the arguments of (39),

$$T_2 = O_{a.s.}\left(\frac{\ln n}{n}\right)^{(s-d/p)/(2(s-d/p)+d)}. \quad (54)$$

This with (48), (50), and (52) shows

$$\sup_{y\in[0,1]^d} \left| \widehat{f}_n(y) - f(y) \right| = O_{a.s.}\left(\frac{\ln n}{n}\right)^{(s-d/p)/(2(s-d/p)+d)}. \quad (55)$$

□

Data Availability

No data were used to support this study.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This paper is supported by Guangxi Natural Science Foundation (Nos. 2017GXNSFAA198194 and 2018GXNSFBA281076), Guangxi Science and Technology Project (Nos. AD18281058 and AD18281019), the Guangxi Young Teachers Basic Ability Improvement Project (Nos. 2018KY0212 and 2019KY0218), and Guangxi Colleges and Universities Key Laboratory of Data Analysis and Computation.

References

[1] P. Ramirez and B. Vidakovic, "Wavelet density estimation for stratified size-biased sample," *Journal of Statistical Planning and Inference*, vol. 140, no. 2, pp. 419–432, 2010.

[2] E. Shirazi and H. Doosti, "Multivariate wavelet-based density estimation with size-biased data," *Statistical Methodology*, vol. 27, pp. 12–19, 2015.

[3] C. Chesneau, I. Dewan, and H. Doosti, "Wavelet linear density estimation for associated stratified size-biased sample," *Journal of Nonparametric Statistics*, vol. 24, no. 2, pp. 429–445, 2012.

[4] Y. Liu and J. Xu, "Wavelet density estimation for negatively associated stratified size-biased sample," *Journal of Nonparametric Statistics*, vol. 26, no. 3, pp. 537–554, 2014.

[5] J. Kou and H. Guo, "Wavelet density estimation for mixing and size-biased data," *Journal of Inequalities and Applications*, vol. 189, 2018.

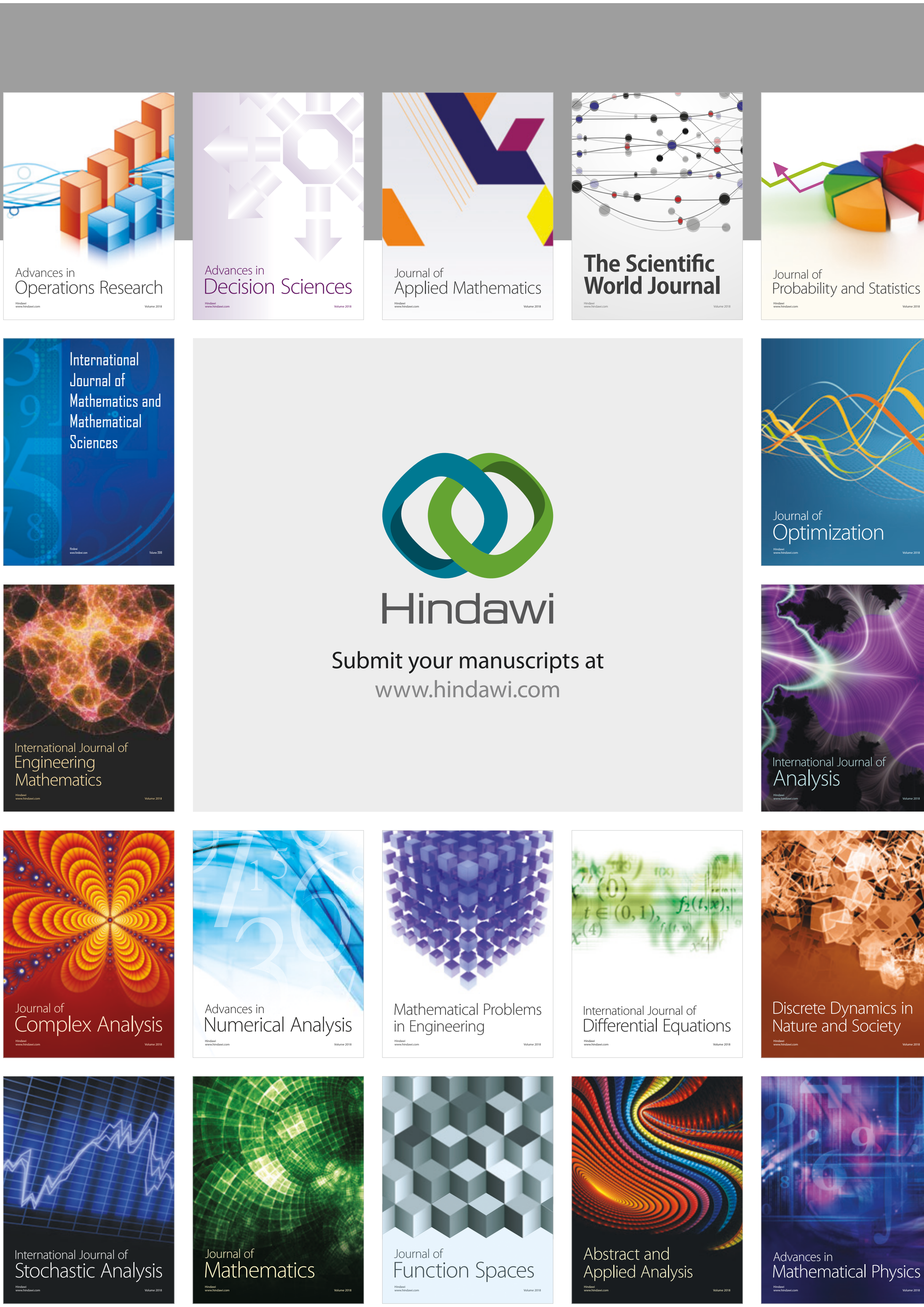
[6] E. Masry, "Multivariate probability density estimation by wavelet methods: strong consistency and rates for stationary time series," *Stochastic Processes and Their Applications*, vol. 67, no. 2, pp. 177–193, 1997.

[7] E. Giné and R. Nickl, "Uniform limit theorems for wavelet density estimators," *Annals of Probability*, vol. 37, no. 4, pp. 1605–1646, 2009.

[8] Y. Meyer, *Wavelets and Operators*, Hermann, Paris, France, 1990.

[9] W. Härdle, G. Kerkyacharian, D. Picard, and A. Tsybakov, *Wavelets, Approximation and Statistical Application*, Springer-Verlag, New York, NY, USA, 1997.

[10] Y. M. Liu and J. L. Xu, "On the L^p-consistency of wavelet estimators," *Acta Mathematica Sinica*, vol. 32, no. 7, pp. 765–782, 2016.


Advances in Operations Research
Advances in Decision Sciences
Journal of Applied Mathematics
The Scientific World Journal
Journal of Probability and Statistics
International Journal of Mathematics and Mathematical Sciences
Hindawi
Submit your manuscripts at
www.hindawi.com
Journal of Optimization
International Journal of Engineering Mathematics
International Journal of Analysis
Journal of Complex Analysis
Advances in Numerical Analysis
Mathematical Problems in Engineering
International Journal of Differential Equations
Discrete Dynamics in Nature and Society
International Journal of Stochastic Analysis
Journal of Mathematics
Journal of Function Spaces
Abstract and Applied Analysis
Advances in Mathematical Physics