

RESEARCH

Open Access



Pointwise density estimation based on negatively associated data

Huijun Guo¹ and Junke Kou^{1*} 

*Correspondence:

kjkou@guet.edu.cn

¹School of Mathematics and
Computational Science, Guilin
University of Electronic Technology,
Guilin, P.R. China

Abstract

In this paper, we consider pointwise estimation over L^p ($1 \leq p < \infty$) risk for a density function based on a negatively associated sample. We construct linear and nonlinear wavelet estimators and provide their convergence rates. It turns out that those wavelet estimators have the same convergence rate up to the $\ln n$ factor. Moreover, the nonlinear wavelet estimator is adaptive.

Keywords: Pointwise density estimation; L^p risk; Negatively associated; Wavelets

1 Introduction

In practical problems, due to the existence of noise, it is possible to obtain real measurement data only with bias (noise). This paper considers the following density estimation model. Let Y_1, Y_2, \dots, Y_n be identically distributed continuous random variables with the density function

$$g(y) = \frac{\omega(y)f(y)}{\mu}, \quad y \in [0, 1]. \quad (1)$$

In this equation, ω is a known biasing function, f denotes the unknown density function of unobserved random variable X , and $\mu := \mathbb{E}[\omega(X)] < \infty$. The aim of this model is to estimate the unknown density function f by the observed negatively associated data Y_1, Y_2, \dots, Y_n .

This model has many applications in industry [4] and economics [8]. Since wavelet bases have a good local property in both time and frequency domains, the wavelet method has been widely used for density estimation problem. When the observed data Y_1, Y_2, \dots, Y_n are independent, Ramírez and Vidakovic [13] constructed a linear wavelet estimator and study the L^2 consistency of this wavelet estimator. Shirazi and Doosti [15] expanded their work to the multivariate case. Because the definition of linear wavelet estimator depends on the smooth parametric of the density function f , the linear estimator is not adaptive. To overcome this shortage, Chesneau [3] proposed a nonlinear wavelet estimator by hard thresholding method. Moreover, an optimal convergence rate over L^p ($1 \leq p < \infty$) risk is considered. When the independence of data is relaxed to the strong mixing case, Kou and Guo [10] studied the L^2 risk of linear and nonlinear wavelet estimators in the Besov space. Note that all those studies all focus on the global error. There is a lack of theoretical results on pointwise wavelet estimation for this density estimation model (1).

In this paper, we establish wavelet estimations on pointwise l^p ($1 \leq p < \infty$) risk for a density function based on a negatively associated sample. Upper bounds of linear and nonlinear wavelet estimators are considered in the Besov space $B_{r,q}^s(\mathbb{R})$. It turns out that the convergence rate of our estimators coincides with the optimal convergence rate for pointwise estimation [2]. Furthermore, our theorem reduces to the corresponding results of Rebelles [14] when $\omega(y) \equiv 1$ and the sample is independent.

1.1 Negative association and wavelets

We first introduce the definition of negative association [1].

Definition 1.1 A sequence of random variables Y_1, Y_2, \dots, Y_n is said to be negatively associated if for each pair of disjoint nonempty subsets A and B of $\{i = 1, 2, \dots, n\}$,

$$\text{Cov}(f(X_i, i \in A), g(X_j, j \in B)) \leq 0,$$

where f and g are real-valued coordinatewise nondecreasing functions and the corresponding covariances exist.

It is well known that $\text{Cov}(Y_i, Y_j) \equiv 0$ when the random variables are independent. Hence the independent and identically distributed data must be negatively associated. Next, we give an important property of negative association, which will be needed in the later discussion.

Lemma 1.1 ([9]) *Let Y_1, Y_2, \dots, Y_n be a sequence of negatively associated random variables, and let A_1, A_2, \dots, A_m be pairwise disjoint nonempty subsets of $\{i = 1, 2, \dots, n\}$. If f_i ($i = 1, 2, \dots, m$) are m coordinatewise nondecreasing (nonincreasing) functions, then $f_1(Y_i, i \in A_1), f_2(Y_i, i \in A_2), \dots, f_m(Y_i, i \in A_m)$ are also negatively associated.*

To construct our wavelet estimators, we provide the basic theory of wavelets.

Throughout this paper, we work with the wavelet basis described as follows. Let $\{V_j, j \in \mathbb{Z}\}$ be a classical orthonormal multiresolution analysis of $L^2(\mathbb{R})$ with a scaling function φ . Then for each $f \in L^2(\mathbb{R})$,

$$f = \sum_{k \in \mathbb{Z}} \alpha_{j_0,k} \varphi_{j_0,k} + \sum_{j=j_0}^{\infty} \sum_{k \in \mathbb{Z}} \beta_{j,k} \psi_{j,k},$$

where $\alpha_{j_0,k} = \langle f, \varphi_{j_0,k} \rangle$, $\beta_{j,k} = \langle f, \psi_{j,k} \rangle$ and

$$\varphi_{j_0,k} = 2^{j_0/2} \varphi(2^{j_0}x - k), \quad \psi_{j,k} = 2^{j/2} \psi(2^jx - k).$$

Let P_j be the orthogonal projection operator from $L^2(\mathbb{R})$ onto the space V_j with orthonormal basis $\{\varphi_{j,k}, k \in \mathbb{Z}\}$. If the scaling function φ satisfies Condition θ , that is,

$$\sum_{k \in \mathbb{Z}} |\varphi(x - k)| \in L^\infty(\mathbb{R}),$$

then it can be shown that for each $f \in L^p(\mathbb{R})$ ($1 \leq p < \infty$),

$$P_j f = \sum_{k \in \mathbb{Z}} \alpha_{j,k} \varphi_{j,k}. \quad (2)$$

On the other hand, a scaling function φ is called m regular if $\varphi \in C^m(\mathbb{R})$ and $|D^l \varphi(y)| \leq c(1+y^2)^{-l}$ for each $l \in \mathbb{Z}$ ($\delta = 0, 1, 2, \dots, m$). In this paper, we choose the Daubechies scaling function D_{2N} [5]. It is easy to see that D_{2N} satisfies m regular when N gets large enough.

Note that a wavelet basis can characterize a Besov space. These spaces contain many well-known function spaces, such as the Hölder and L^2 Sobolev spaces. The following lemma gives equivalent definition of Besov spaces.

Lemma 1.2 ([7]) *Let $f \in L^r(\mathbb{R})$ ($1 \leq r \leq +\infty$), let the scaling function φ be m -regular, and let $0 < s < m$. Then the following statements are equivalent:*

- (i) $f \in B_{r,q}^s(\mathbb{R})$, $1 \leq q \leq +\infty$;
- (ii) $\{2^{js} \|P_j f - f\|_r\} \in l_q$;
- (iii) $\{2^{j(s-\frac{1}{r}+\frac{1}{2})} \|\beta_j\|_r\} \in l_q$.

The Besov norm of f can be defined as

$$\|f\|_{B_{r,q}^s} := \|(\alpha_{j_0})\|_r + \left\| \left(2^{j(s-\frac{1}{r}+\frac{1}{2})} \|\beta_j\|_r \right)_{j \geq j_0} \right\|_q$$

with $\|(\alpha_{j_0})\|_r^r := \sum_{k \in \mathbb{Z}} |\alpha_{j_0,k}|^r$ and $\|\beta_j\|_r^r := \sum_{k \in \mathbb{Z}} |\beta_{j,k}|^r$.

In this paper, we assume that the density function f belongs to the Besov ball with radius $H > 0$, that is,

$$f \in B_{r,q}^s(H) := \{f \in B_{r,q}^s(\mathbb{R}), \|f\|_{B_{r,q}^s} \leq H\}.$$

1.2 Wavelet estimators and theorem

Define our linear wavelet estimator as follows:

$$\hat{f}_n^{\text{lin}}(y) := \sum_{k \in \Lambda} \hat{\alpha}_{j_0,k} \varphi_{j_0,k}(y) \quad (3)$$

with

$$\hat{\alpha}_{j_0,k} = \frac{\hat{\mu}_n}{n} \sum_{i=1}^n \frac{\varphi_{j_0,k}(Y_i)}{\omega(Y_i)} \quad (4)$$

and

$$\hat{\mu}_n = \left[\frac{1}{n} \sum_{i=1}^n \frac{1}{\omega(Y_i)} \right]^{-1}. \quad (5)$$

Using the hard thresholding method, a nonlinear wavelet estimator is defined by

$$\hat{f}_n^{\text{non}}(y) := \sum_{k \in \Lambda} \hat{\alpha}_{j_0,k} \varphi_{j_0,k}(y) + \sum_{j=j_0}^{j_1} \sum_{k \in \Lambda_j} \hat{\beta}_{j,k} I_{\{|\hat{\beta}_{j,k}| \geq \kappa t_n\}} \psi_{j,k}(y), \quad (6)$$

where $t_n = \sqrt{\frac{\ln n}{n}}$ and

$$\hat{\beta}_{j,k} = \frac{\hat{\mu}_n}{n} \sum_{i=1}^n \frac{\psi_{j,k}(Y_i)}{\omega(Y_i)}. \quad (7)$$

In these definitions, $\Lambda := \{k \in \mathbb{Z}, \text{supp} f \cap \text{supp} \varphi_{j_0,k} \neq \emptyset\}$ and $\Lambda_j := \{k \in \mathbb{Z}, \text{supp} f \cap \text{supp} \psi_{j,k} \neq \emptyset\}$. Note that the cardinality of Λ (Λ_j) satisfies $|\Lambda| \sim 2^{j_0}$ ($|\Lambda_j| \sim 2^j$) due to the compactly supported properties of the functions f and $\varphi_{j_0,k}$ ($\psi_{j,k}$). Here and further, $A \sim B$ stands for both $A \lesssim B$ and $B \lesssim A$, where $A \lesssim B$ denotes $A \leq cB$ with a positive constant c that is independent of A and B . In addition, the constant κ will be chosen in later discussion.

We are in position to state our main theorem.

Theorem 1 *Let $f \in B_{r,q}^s(H)$ ($r, q \in [1, \infty), s > \frac{1}{r}$), and let $\omega(y)$ be a nonincreasing function such that $\omega(y) \sim 1$. Then for each $1 \leq p < \infty$, the linear wavelet estimator with $2^{j_0} \sim n^{\frac{1}{2(s-1/r)+1}}$ satisfies*

$$\mathbb{E}[|\widehat{f}_n^{\text{lin}}(y) - f(y)|]^p \lesssim n^{-\frac{(s-1/r)p}{2(s-1/r)+1}}, \quad (8)$$

and the nonlinear wavelet estimator with $2^{j_0} \sim n^{\frac{1}{2m+1}}$ ($m > s$) and $2^{j_1} \sim \frac{n}{\ln n}$ satisfies

$$\mathbb{E}[|\widehat{f}_n^{\text{non}}(y) - f(y)|]^p \lesssim (\ln n)^{\frac{3p}{2}} n^{-\frac{(s-1/r)p}{2(s-1/r)+1}}. \quad (9)$$

Remark 1 Note that $n^{-\frac{(s-1/r)p}{2(s-1/r)+1}}$ is the optimal convergence rate in the minimax sense for pointwise estimation in a Besov space [2]. Moreover, our theorem reduces to the results of Rebelles [14] when $\omega(y) \equiv 1$ and the random sample is independent.

Remark 2 In contrast to the linear wavelet estimator, the convergence rate of the nonlinear estimator remains the same as that of the linear one up to the $\ln n$ factor. However, the nonlinear one is adaptive, which means that both j_0 and j_1 do not depend on s .

2 Auxiliary lemmas

In this section, we give some lemmas, which are very useful for proving Theorem 1.

Lemma 2.1 *For the model defined by (1), we have*

$$\mathbb{E}\left[\frac{1}{\omega(Y_i)}\right] = \frac{1}{\mu}, \quad \mathbb{E}\left[\frac{\mu\varphi_{j,k}(Y_i)}{\omega(Y_i)}\right] = \alpha_{j,k} \quad \text{and} \quad \mathbb{E}\left[\frac{\mu\psi_{j,k}(Y_i)}{\omega(Y_i)}\right] = \beta_{j,k}.$$

Proof This lemma can be proved by the same arguments of Kou and Guo [10]. \square

Lemma 2.2 *Let $f \in B_{r,q}^s$ ($1 \leq r, q < +\infty, s > 1/r$), and let $\omega(y)$ be a nonincreasing function such that $\omega(y) \sim 1$. If $2^j \leq n$ and $1 \leq p < +\infty$, then*

$$\mathbb{E}[|\widehat{\alpha}_{j_0,k} - \alpha_{j_0,k}|^p] \lesssim n^{-\frac{p}{2}}, \quad \mathbb{E}[|\widehat{\beta}_{j,k} - \beta_{j,k}|^p] \lesssim n^{-\frac{p}{2}}.$$

Proof Because the proofs of both inequalities are similar, we only prove the second one. By the definition of $\widehat{\beta}_{j,k}$ we have

$$|\widehat{\beta}_{j,k} - \beta_{j,k}| \leq \left| \frac{\widehat{\mu}_n}{\mu} \left(\frac{\mu}{n} \sum_{i=1}^n \frac{\psi_{j,k}(Y_i)}{\omega(Y_i)} - \beta_{j,k} \right) \right| + \left| \beta_{j,k} \widehat{\mu}_n \left(\frac{1}{\mu} - \frac{1}{\widehat{\mu}_n} \right) \right|.$$

Note that the definition of $\widehat{\mu}_n$ and $\omega(y) \sim 1$ imply $|\widehat{\mu}_n| \lesssim 1$. We have $B_{r,q}^s(\mathbb{R}) \subseteq B_{\infty,\infty}^{s-1/r}(\mathbb{R})$ in the case of $s > \frac{1}{r}$; then $f \in B_{\infty,\infty}^{s-1/r}(\mathbb{R})$ and $\|f\|_\infty \lesssim 1$. Moreover, $|\beta_{j,k}| = |\langle f, \psi_{j,k} \rangle| \lesssim 1$ by the Cauchy–Schwarz inequality and the orthonormality of wavelet functions. Hence, we have the following conclusion:

$$\mathbb{E}[|\widehat{\beta}_{j,k} - \beta_{j,k}|^p] \lesssim \mathbb{E}\left[\left|\frac{1}{n} \sum_{i=1}^n \frac{\mu \psi_{j,k}(Y_i)}{\omega(Y_i)} - \beta_{j,k}\right|^p\right] + \mathbb{E}\left[\left|\frac{1}{\mu} - \frac{1}{\widehat{\mu}_n}\right|^p\right]. \quad (10)$$

Then we need to estimate $T_1 := \mathbb{E}[|\frac{1}{n} \sum_{i=1}^n \frac{\mu \psi_{j,k}(Y_i)}{\omega(Y_i)} - \beta_{j,k}|^p]$ and $T_2 := \mathbb{E}[|\frac{1}{\mu} - \frac{1}{\widehat{\mu}_n}|^p]$.

- An upper bound for T_1 . Taking $\eta_i := \frac{\mu \psi_{j,k}(Y_i)}{\omega(Y_i)} - \beta_{j,k}$, we get

$$T_1 = \mathbb{E}\left[\left|\frac{1}{n} \sum_{i=1}^n \eta_i\right|^p\right] = \left(\frac{1}{n}\right)^p \mathbb{E}\left[\left|\sum_{i=1}^n \eta_i\right|^p\right].$$

Note that ψ is a function of bounded variation (see Liu and Xu [12]). We can get $\psi := \widetilde{\psi} - \overline{\psi}$, where $\widetilde{\psi}$ and $\overline{\psi}$ bounded nonnegative nondecreasing functions. Define

$$\widetilde{\eta}_i := \frac{\mu \widetilde{\psi}_{j,k}(Y_i)}{\omega(Y_i)} - \widetilde{\beta}_{j,k}, \quad \overline{\eta}_i := \frac{\mu \overline{\psi}_{j,k}(Y_i)}{\omega(Y_i)} - \overline{\beta}_{j,k}$$

with $\widetilde{\beta}_{j,k} := \langle f, \widetilde{\psi}_{j,k} \rangle$ and $\overline{\beta}_{j,k} := \langle f, \overline{\psi}_{j,k} \rangle$. Then $\eta_i = \widetilde{\eta}_i - \overline{\eta}_i$, $\beta_{j,k} = \widetilde{\beta}_{j,k} - \overline{\beta}_{j,k}$, and

$$T_1 = \left(\frac{1}{n}\right)^p \mathbb{E}\left[\left|\sum_{i=1}^n (\widetilde{\eta}_i - \overline{\eta}_i)\right|^p\right] \lesssim \left(\frac{1}{n}\right)^p \left\{ \mathbb{E}\left[\left|\sum_{i=1}^n \widetilde{\eta}_i\right|^p\right] + \mathbb{E}\left[\left|\sum_{i=1}^n \overline{\eta}_i\right|^p\right] \right\}. \quad (11)$$

Similar arguments as in Lemma 2.1 show that $\mathbb{E}[\widetilde{\eta}_i] = 0$. The function $\frac{\widetilde{\psi}_{j,k}(y)}{\omega(y)}$ is nondecreasing by the monotonicity of $\widetilde{\psi}(y)$ and $\omega(y)$. Furthermore, we get that $\{\widetilde{\eta}_i, i = 1, 2, \dots, n\}$ is negatively associated by Lemma 1.1. On the other hand, it follows from (1) and $\omega(y) \sim 1$ that

$$\mathbb{E}[|\widetilde{\eta}_i|^p] \lesssim \mathbb{E}\left[\left|\frac{\mu \widetilde{\psi}_{j,k}(Y_i)}{\omega(Y_i)}\right|^p\right] \lesssim \int_{[0,1]} |\widetilde{\psi}_{j,k}(y)|^p f(y) dy \lesssim 2^{j(p/2-1)}. \quad (12)$$

In particular, $\mathbb{E}[|\widetilde{\eta}_i|^2] \lesssim 1$. Recall Rosenthal's inequality [12]: if Y_1, Y_2, \dots, Y_n are negatively associated random variables such that $\mathbb{E}[Y_i] = 0$ and $\mathbb{E}[|Y_i|^p] < \infty$, then

$$\mathbb{E}\left[\left|\sum_{i=1}^n Y_i\right|^p\right] \lesssim \begin{cases} \sum_{i=1}^n \mathbb{E}[|Y_i|^p] + (\sum_{i=1}^n \mathbb{E}[|Y_i|^2])^{p/2}, & p > 2; \\ (\sum_{i=1}^n \mathbb{E}[|Y_i|^2])^{p/2}, & 1 \leq p \leq 2. \end{cases}$$

From this we clearly have

$$\mathbb{E}\left[\left|\sum_{i=1}^n \widetilde{\eta}_i\right|^p\right] \lesssim \begin{cases} n2^{j(p/2-1)} + n^{p/2}, & p > 2; \\ n^{p/2}, & 1 \leq p \leq 2. \end{cases}$$

This, together with $2^j \leq n$, shows that $\mathbb{E}[|\sum_{i=1}^n \tilde{\eta}_i|^p] \lesssim n^{p/2}$. Similarly, $\mathbb{E}[|\sum_{i=1}^n \bar{\eta}_i|^p] \lesssim n^{p/2}$. Combining these with (11), we get that

$$T_1 \lesssim \left(\frac{1}{n}\right)^p \left\{ \mathbb{E} \left[\left| \sum_{i=1}^n \tilde{\eta}_i \right|^p \right] + \mathbb{E} \left[\left| \sum_{i=1}^n \bar{\eta}_i \right|^p \right] \right\} \lesssim n^{-\frac{p}{2}}. \quad (13)$$

• An upper bound for T_2 . It is easy to see from the definition of $\hat{\mu}_n$ that

$$T_2 = \mathbb{E} \left[\left| \frac{1}{\mu} - \frac{1}{\hat{\mu}_n} \right|^p \right] = \left(\frac{1}{n}\right)^p \mathbb{E} \left[\left| \sum_{i=1}^n \left(\frac{1}{\omega(Y_i)} - \frac{1}{\mu} \right) \right|^p \right]. \quad (14)$$

Defining $\xi_i := \frac{1}{\omega(Y_i)} - \frac{1}{\mu}$, we obtain that $\mathbb{E}[\xi_i] = 0$ and $\mathbb{E}[|\xi_i|^p] \lesssim 1$ by Lemma 2.1 and $\omega(y) \sim 1$. In addition, by the monotonicity of $\omega(y)$ and Lemma 1.1 we know that $\xi_1, \xi_2, \dots, \xi_n$ are also negatively associated. Then using Rosenthal's inequality, we get

$$\mathbb{E} \left[\left| \sum_{i=1}^n \xi_i \right|^p \right] \lesssim \begin{cases} n + n^{p/2}, & p > 2; \\ n^{p/2}, & 1 \leq p \leq 2. \end{cases}$$

Hence

$$T_2 = \left(\frac{1}{n}\right)^p \mathbb{E} \left[\left| \sum_{i=1}^n \xi_i \right|^p \right] \lesssim n^{-\frac{p}{2}}. \quad (15)$$

Finally, by (10), (13), and (15) we have

$$\mathbb{E} \left[|\hat{\beta}_{j,k} - \beta_{j,k}|^p \right] \lesssim n^{-\frac{p}{2}}.$$

This ends the proof. \square

Lemma 2.3 Let $f \in B_{r,q}^s$ ($1 \leq r, q < +\infty, s > 1/r$) and $\hat{\beta}_{j,k}$ be defined by (7). If $\omega(y)$ is a non-increasing function, $\omega(y) \sim 1$, and $2^j \leq \frac{n}{\ln n}$, then for each $\lambda > 0$, there exists a constant $\kappa > 1$ such that

$$\mathbb{P} \{ |\hat{\beta}_{j,k} - \beta_{j,k}| \geq \kappa t_n \} \lesssim 2^{-\lambda j}.$$

Proof By the same arguments of (10) we can obtain that

$$\begin{aligned} \mathbb{P} \{ |\hat{\beta}_{j,k} - \beta_{j,k}| \geq \kappa t_n \} &\leq \mathbb{P} \left\{ \left| \frac{1}{n} \sum_{i=1}^n \left(\frac{1}{\omega(Y_i)} - \frac{1}{\mu} \right) \right| \geq \frac{\kappa t_n}{2} \right\} \\ &\quad + \mathbb{P} \left\{ \left| \frac{1}{n} \sum_{i=1}^n \left(\frac{\mu \psi_{j,k}(Y_i)}{\omega(Y_i)} - \beta_{j,k} \right) \right| \geq \frac{\kappa t_n}{2} \right\}. \end{aligned} \quad (16)$$

To estimate $\mathbb{P} \{ |\frac{1}{n} \sum_{i=1}^n (\frac{1}{\omega(Y_i)} - \frac{1}{\mu})| \geq \frac{\kappa t_n}{2} \}$, we also define $\xi_i := \frac{1}{\omega(Y_i)} - \frac{1}{\mu}$. Then Lemma 2.1 implies that $\mathbb{E}[\xi_i] = 0$. Moreover, $|\xi_i| \lesssim 1$ and $\mathbb{E}[|\xi_i|^2] \lesssim 1$ thanks to $\omega(y) \sim 1$. On the other hand, because of the monotonicity of $\omega(y)$ and Lemma 1.1, $\xi_1, \xi_2, \dots, \xi_n$ are also negatively associated.

Recall Bernstein's inequality [12]: If Y_1, Y_2, \dots, Y_n are negatively associated random variables such that $\mathbb{E}[Y_i] = 0$, $|Y_i| \leq M < \infty$, and $\mathbb{E}[|Y_i|^2] = \sigma^2$, then for each $\varepsilon > 0$,

$$\mathbb{P}\left\{\left|\frac{1}{n} \sum_{i=1}^n Y_i\right| \geq \varepsilon\right\} \lesssim \exp\left(-\frac{n\varepsilon^2}{2(\sigma^2 + \varepsilon M/3)}\right).$$

Therefore, by the previous arguments for ξ_i and $t_n = \sqrt{\frac{\ln n}{n}}$, we derive

$$\mathbb{P}\left\{\left|\frac{1}{n} \sum_{i=1}^n \left(\frac{1}{\omega(Y_i)} - \frac{1}{\mu}\right)\right| \geq \frac{\kappa t_n}{2}\right\} \lesssim \exp\left(-\frac{(\ln n)\kappa^2/4}{2(\sigma^2 + \kappa/6)}\right).$$

Then there exists $\kappa > 1$ such that $\exp\left(-\frac{(\ln n)\kappa^2/4}{2(\sigma^2 + \kappa/6)}\right) \lesssim 2^{-\lambda j}$ with fixed $\lambda > 0$. Hence

$$\mathbb{P}\left\{\left|\frac{1}{n} \sum_{i=1}^n \left(\frac{1}{\omega(Y_i)} - \frac{1}{\mu}\right)\right| \geq \frac{\kappa t_n}{2}\right\} \lesssim 2^{-\lambda j}. \quad (17)$$

Next, we estimate $\mathbb{P}\{|\frac{1}{n} \sum_{i=1}^n (\frac{\mu \psi_{j,k}(Y_i)}{\omega(Y_i)} - \beta_{j,k})| \geq \frac{\kappa t_n}{2}\}$. By the same arguments of (11) we get

$$\mathbb{P}\left\{\left|\frac{1}{n} \sum_{i=1}^n \eta_i\right| \geq \frac{\kappa t_n}{2}\right\} \leq \mathbb{P}\left\{\left|\frac{1}{n} \sum_{i=1}^n \tilde{\eta}_i\right| \geq \frac{\kappa t_n}{4}\right\} + \mathbb{P}\left\{\left|\frac{1}{n} \sum_{i=1}^n \bar{\eta}_i\right| \geq \frac{\kappa t_n}{4}\right\}. \quad (18)$$

It is easy to see from the definition of $\tilde{\eta}_i$ and Lemma 2.1 that $\mathbb{E}[\tilde{\eta}_i] = 0$. Moreover, $\mathbb{E}[|\tilde{\eta}_i|^2] \lesssim 1$ by (12) with $p = 2$. Using $\omega(y) \sim 1$, we get $|\frac{\mu \tilde{\psi}_{j,k}(Y_i)}{\omega(Y_i)}| \lesssim 2^{j/2}$ and $|\tilde{\eta}_i| \leq |\frac{\mu \tilde{\psi}_{j,k}(Y_i)}{\omega(Y_i)}| + \mathbb{E}[|\frac{\mu \tilde{\psi}_{j,k}(Y_i)}{\omega(Y_i)}|] \lesssim 2^{j/2}$. Then it follows from Bernstein's inequality, $2^j \leq \frac{n}{\ln n}$, and $t_n = \sqrt{\frac{\ln n}{n}}$ that

$$\mathbb{P}\left\{\left|\frac{1}{n} \sum_{i=1}^n \tilde{\eta}_i\right| \geq \frac{\kappa t_n}{4}\right\} \lesssim \exp\left(-\frac{n(\kappa t_n/4)^2}{2(\sigma^2 + \kappa t_n 2^{j/2}/12)}\right) \lesssim \exp\left(-\frac{(\ln n)\kappa^2/16}{2(\sigma^2 + \kappa/12)}\right).$$

Clearly, we can take $\kappa > 1$ such that $\mathbb{P}\{|\frac{1}{n} \sum_{i=1}^n \tilde{\eta}_i| \geq \frac{\kappa t_n}{4}\} \lesssim 2^{-\lambda j}$. Then similar arguments show that $\mathbb{P}\{|\frac{1}{n} \sum_{i=1}^n \bar{\eta}_i| \geq \frac{\kappa t_n}{4}\} \lesssim 2^{-\lambda j}$. Combining those with (18), we obtain

$$\mathbb{P}\left\{\left|\frac{1}{n} \sum_{i=1}^n \left(\frac{\mu \psi_{j,k}(Y_i)}{\omega(Y_i)} - \beta_{j,k}\right)\right| \geq \frac{\kappa t_n}{2}\right\} \lesssim 2^{-\lambda j}. \quad (19)$$

By (16), (17), and (19) we get

$$\mathbb{P}\{|\hat{\beta}_{j,k} - \beta_{j,k}| \geq \kappa t_n\} \lesssim 2^{-\lambda j}.$$

This ends the proof. \square

3 Proof of theorem

In this section, we prove the Theorem 1.

Proof of (8) It is easy to see that

$$\mathbb{E}[\widehat{f}_n^{\text{lin}}(y) - f(y)|^p] \lesssim \mathbb{E}[\widehat{f}_n^{\text{lin}}(y) - P_{j_0}f(y)|^p] + |P_{j_0}f(y) - f(y)|^p. \quad (20)$$

Then we need to estimate $\mathbb{E}[\widehat{f}_n^{\text{lin}}(y) - P_{j_0}f(y)|^p]$ and $|P_{j_0}f(y) - f(y)|^p$.

By (2) and (3) we get that

$$\mathbb{E}[\widehat{f}_n^{\text{lin}}(y) - P_{j_0}f(y)|^p] = \mathbb{E}\left[\left|\sum_{k \in \Lambda} (\widehat{\alpha}_{j_0,k} - \alpha_{j_0,k}) \varphi_{j_0,k}(y)\right|^p\right].$$

Using the Hölder inequality ($1/p + 1/p' = 1$), we see that

$$\mathbb{E}[\widehat{f}_n^{\text{lin}}(y) - P_{j_0}f(y)|^p] \leq \mathbb{E}\left[\left(\sum_{k \in \Lambda} |\widehat{\alpha}_{j_0,k} - \alpha_{j_0,k}|^p |\varphi_{j_0,k}(y)|\right) \left(\sum_{k \in \Lambda} |\varphi_{j_0,k}(y)|\right)^{\frac{p}{p'}}\right].$$

Then it follows from Condition θ and Lemma 2.2 that

$$\mathbb{E}[\widehat{f}_n^{\text{lin}}(y) - P_{j_0}f(y)|^p] \lesssim \sum_{k \in \Lambda} \mathbb{E}[|\widehat{\alpha}_{j_0,k} - \alpha_{j_0,k}|^p] |\varphi_{j_0,k}(y)| 2^{\frac{j_0 p}{2p'}} \lesssim \left(\frac{2^{j_0}}{n}\right)^{\frac{p}{2}}. \quad (21)$$

This, together with $2^{j_0} \sim n^{\frac{1}{2(s-1/r)+1}}$, shows that

$$\mathbb{E}[\widehat{f}_n^{\text{lin}}(y) - P_{j_0}f(y)|^p] \lesssim n^{-\frac{(s-1/r)p}{2(s-1/r)+1}}. \quad (22)$$

Note that $B_{r,q}^s(\mathbb{R}) \subseteq B_{\infty,\infty}^{s-1/r}(\mathbb{R})$ in the case $s > 1/r$. It should be pointed out that $B_{\infty,\infty}^{s-1/r}(\mathbb{R})$ is also a Hölder space. Then by Lemma 1.2, $f \in B_{r,q}^s(\mathbb{R})$, and $2^{j_0} \sim n^{\frac{1}{2(s-1/r)+1}}$ we obtain that

$$|P_{j_0}f(y) - f(y)|^p \lesssim 2^{-j_0(s-1/r)p} \lesssim n^{-\frac{(s-1/r)p}{2(s-1/r)+1}}. \quad (23)$$

Combining this with (20) and (22), we get

$$\mathbb{E}[\widehat{f}_n^{\text{lin}}(y) - f(y)|^p] \lesssim n^{-\frac{(s-1/r)p}{2(s-1/r)+1}}. \quad \square$$

Proof of (9) Using the definitions of $\widehat{f}_n^{\text{lin}}(y)$ and $\widehat{f}_n^{\text{non}}(y)$, we get that

$$\mathbb{E}[\widehat{f}_n^{\text{non}}(y) - f(y)|^p] \lesssim W_1 + W_2 + G, \quad (24)$$

where $W_1 := \mathbb{E}[\widehat{f}_n^{\text{lin}}(y) - P_{j_0}f(y)|^p]$, $W_2 := |P_{j_1+1}f(y) - f(y)|^p$, and

$$G := \mathbb{E}\left[\left|\sum_{j=j_0}^{j_1} \sum_{k \in \Lambda_j} (\widehat{\beta}_{j,k} I_{\{|\widehat{\beta}_{j,k}| \geq \kappa t_n\}} - \beta_{j,k}) \psi_{j,k}(y)\right|^p\right].$$

It follows from (21), $2^{j_0} \sim n^{\frac{1}{2m+1}}$ ($m > s$), and $s > 1/r$ that

$$W_1 \lesssim \left(\frac{2^{j_0}}{n}\right)^{\frac{p}{2}} \sim n^{-\frac{mp}{2m+1}} < n^{-\frac{(s-1/r)p}{2(s-1/r)+1}}. \quad (25)$$

On the other hand, by the same arguments as for (23), we can obtain that $W_2 \lesssim 2^{-j_1(s-1/r)p}$. This with the choice of $2^{j_1} \sim \frac{n}{\ln n}$ shows

$$W_2 \lesssim 2^{-j_1(s-1/r)p} \sim \left(\frac{\ln n}{n}\right)^{(s-1/r)p} < \left(\frac{\ln n}{n}\right)^{\frac{(s-1/r)p}{2(s-1/r)+1}}. \quad (26)$$

Then the remaining task is to estimate G .

Using the classical technique in [6], we get that

$$G \lesssim (\ln n)^{p-1} (G_1 + G_2 + G_3), \quad (27)$$

where

$$\begin{aligned} G_1 &:= \mathbb{E} \left[\sum_{j=j_0}^{j_1} \left(\sum_{k \in \Lambda_j} |\widehat{\beta}_{j,k} - \beta_{j,k}| I_{\{|\widehat{\beta}_{j,k} - \beta_{j,k}| \geq \frac{\kappa t_n}{2}\}} |\psi_{j,k}(y)| \right)^p \right], \\ G_2 &:= \mathbb{E} \left[\sum_{j=j_0}^{j_1} \left(\sum_{k \in \Lambda_j} |\widehat{\beta}_{j,k} - \beta_{j,k}| I_{\{|\beta_{j,k}| \geq \frac{\kappa t_n}{2}\}} |\psi_{j,k}(y)| \right)^p \right], \\ G_3 &:= \sum_{j=j_0}^{j_1} \left(\sum_{k \in \Lambda_j} |\beta_{j,k}| I_{\{|\beta_{j,k}| \leq 2\kappa t_n\}} |\psi_{j,k}(y)| \right)^p. \end{aligned}$$

• An upper bound for G_1 . By the definition of $\widehat{\beta}_{j,k}$, $\omega(y) \sim 1$, and Lemma 2.1, $|\widehat{\beta}_{j,k}| \lesssim 2^{j/2}$ and $|\widehat{\beta}_{j,k} - \beta_{j,k}| \lesssim 2^{j/2}$. Furthermore, we obtain that

$$G_1 \lesssim \mathbb{E} \left[\sum_{j=j_0}^{j_1} \left(\sum_{k \in \Lambda_j} 2^{j/2} I_{\{|\widehat{\beta}_{j,k} - \beta_{j,k}| \geq \frac{\kappa t_n}{2}\}} |\psi_{j,k}(y)| \right)^p \right].$$

On the other hand, it follows from the Hölder inequality and Condition θ that

$$\begin{aligned} & \left(\sum_{k \in \Lambda_j} I_{\{|\widehat{\beta}_{j,k} - \beta_{j,k}| \geq \frac{\kappa t_n}{2}\}} |\psi_{j,k}(y)| \right)^p \\ & \lesssim \left(\sum_{k \in \Lambda_j} I_{\{|\widehat{\beta}_{j,k} - \beta_{j,k}| \geq \frac{\kappa t_n}{2}\}} |\psi_{j,k}(y)| \right) \left(\sum_{k \in \Lambda_j} |\psi_{j,k}(y)| \right)^{\frac{p}{p'}} \\ & \lesssim \left(\sum_{k \in \Lambda_j} I_{\{|\widehat{\beta}_{j,k} - \beta_{j,k}| \geq \frac{\kappa t_n}{2}\}} |\psi_{j,k}(y)| \right) 2^{\frac{jp}{2p'}}. \end{aligned}$$

Then using Condition θ and Lemma 2.3, we derive that

$$\begin{aligned} G_1 &\lesssim \mathbb{E} \left[\sum_{j=j_0}^{j_1} 2^{jp/2} \left(\sum_{k \in \Lambda_j} I_{\{|\widehat{\beta}_{j,k} - \beta_{j,k}| \geq \frac{\kappa t_n}{2}\}} |\psi_{j,k}(y)| \right)^{\frac{jp}{2p'}} \right] \\ &\lesssim \sum_{j=j_0}^{j_1} 2^{\frac{j}{2}(p+\frac{p}{p'})} \sum_{k \in \Lambda_j} |\psi_{j,k}(y)| \mathbb{E} [I_{\{|\widehat{\beta}_{j,k} - \beta_{j,k}| \geq \frac{\kappa t_n}{2}\}}] \lesssim \sum_{j=j_0}^{j_1} 2^{j(p-\lambda)}. \end{aligned} \quad (28)$$

Clearly, there exists $\kappa > 1$ such that $\lambda > p + mp$ in Lemma 2.3. Then $G_1 \lesssim \sum_{j=j_0}^{j_1} 2^{j(p-\lambda)} \lesssim \sum_{j=j_0}^{j_1} 2^{-jmp}$. This with the choice of $2^{j_0} \sim n^{\frac{1}{2m+1}}$ ($m > s$) shows that

$$G_1 \lesssim \sum_{j=j_0}^{j_1} 2^{-jmp} \lesssim 2^{-j_0 mp} \sim n^{-\frac{mp}{2m+1}} < n^{-\frac{(s-1/r)p}{2(s-1/r)+1}}. \quad (29)$$

• An upper bound for G_2 . Taking $2^{j_*} \sim n^{\frac{1}{2(s-1/r)+1}}$, we get that $2^{j_0} < 2^{j_*} < 2^{j_1}$. It is easy to see that

$$\begin{aligned} G_{21} &:= \mathbb{E} \left[\sum_{j=j_0}^{j_*} \left(\sum_{k \in \Lambda_j} |\widehat{\beta}_{j,k} - \beta_{j,k}| I_{\{|\beta_{j,k}| \geq \frac{\kappa t_n}{2}\}} |\psi_{j,k}(y)| \right)^p \right] \\ &\lesssim \sum_{j=j_0}^{j_*} \mathbb{E} \left[\left(\sum_{k \in \Lambda_j} |\widehat{\beta}_{j,k} - \beta_{j,k}| |\psi_{j,k}(y)| \right)^p \right]. \end{aligned}$$

Similarly to the arguments of (21), we get

$$G_{21} \lesssim \sum_{j=j_0}^{j_*} \left(\frac{2^j}{n} \right)^{\frac{p}{2}} \lesssim \left(\frac{2^{j_*}}{n} \right)^{\frac{p}{2}} \sim n^{-\frac{(s-1/r)p}{2(s-1/r)+1}} \quad (30)$$

by Lemma 2.2 and $2^{j_*} \sim n^{\frac{1}{2(s-1/r)+1}}$.

On the other hand,

$$\begin{aligned} G_{22} &:= \mathbb{E} \left[\sum_{j=j_*+1}^{j_1} \left(\sum_{k \in \Lambda_j} |\widehat{\beta}_{j,k} - \beta_{j,k}| I_{\{|\beta_{j,k}| \geq \frac{\kappa t_n}{2}\}} |\psi_{j,k}(y)| \right)^p \right] \\ &\lesssim \mathbb{E} \left[\sum_{j=j_*+1}^{j_1} \left(\sum_{k \in \Lambda_j} |\widehat{\beta}_{j,k} - \beta_{j,k}| \left| \frac{\beta_{j,k}}{\kappa t_n} \right| |\psi_{j,k}(y)| \right)^p \right]. \end{aligned}$$

Using the Hölder inequality and Lemma 2.2, we have

$$\begin{aligned} G_{22} &\lesssim \mathbb{E} \left[\sum_{j=j_*+1}^{j_1} \left(\frac{1}{t_n} \right)^p \left(\sum_{k \in \Lambda_j} |\widehat{\beta}_{j,k} - \beta_{j,k}|^p |\beta_{j,k} \psi_{j,k}(y)| \right) \left(\sum_{k \in \Lambda_j} |\beta_{j,k} \psi_{j,k}(y)| \right)^{\frac{p}{p'}} \right] \\ &\lesssim \sum_{j=j_*+1}^{j_1} \left(\frac{1}{t_n} \right)^p n^{-\frac{p}{2}} \left(\sum_{k \in \Lambda_j} |\beta_{j,k} \psi_{j,k}(y)| \right)^p. \end{aligned}$$

When $s > 1/r$, $B_{r,q}^s(\mathbb{R}) \subseteq B_{\infty,\infty}^{s-1/r}(\mathbb{R})$. Clearly, $B_{\infty,\infty}^{s-1/r}(\mathbb{R})$ is a Hölder space. Then we can derive that $\sum_{k \in \Lambda_j} |\beta_{j,k} \psi_{j,k}(y)| \lesssim 2^{-j(s-1/r)}$ as in [11]. Hence it follows from the choice of 2^{j_*} that

$$G_{22} \lesssim \sum_{j=j_*+1}^{j_1} 2^{-j(s-1/r)p} \lesssim 2^{-j_*(s-1/r)p} \sim n^{-\frac{(s-1/r)p}{2(s-1/r)+1}}. \quad (31)$$

Therefore we have

$$G_2 = G_{21} + G_{22} \lesssim n^{-\frac{(s-1/r)p}{2(s-1/r)+1}}. \quad (32)$$

- An upper bound for G_3 . Clearly, we can obtain that

$$\begin{aligned} G_{31} &:= \sum_{j=0}^{j_*} \left(\sum_{k \in \Lambda_j} |\beta_{j,k}| I_{\{|\beta_{j,k}| \leq 2\kappa t_n\}} |\psi_{j,k}(y)| \right)^p \lesssim \sum_{j=0}^{j_*} \left(\sum_{k \in \Lambda_j} t_n |\psi_{j,k}(y)| \right)^p \\ &\lesssim \sum_{j=0}^{j_*} \left(\frac{\ln n}{n} \right)^{\frac{p}{2}} 2^{jp/2} \lesssim \left(\frac{\ln n}{n} \right)^{\frac{p}{2}} 2^{j_*p/2} \lesssim (\ln n)^{p/2} n^{-\frac{(s-1/r)p}{2(s-1/r)+1}}. \end{aligned} \quad (33)$$

In addition, it follows from the Hölder inequality ($1/r + 1/r' = 1$), Condition θ , and Lemma 1.2 that

$$\begin{aligned} G_{32} &:= \sum_{j=j_*+1}^{j_1} \left(\sum_{k \in \Lambda_j} |\beta_{j,k}| I_{\{|\beta_{j,k}| \leq 2\kappa t_n\}} |\psi_{j,k}(y)| \right)^p \lesssim \sum_{j=j_*+1}^{j_1} \left(\sum_{k \in \Lambda_j} |\beta_{j,k}| |\psi_{j,k}(y)| \right)^p \\ &\lesssim \sum_{j=j_*+1}^{j_1} \left(\sum_{k \in \Lambda_j} |\beta_{j,k}|^r \right)^{\frac{p}{r}} \left(\sum_{k \in \Lambda_j} |\psi_{j,k}(y)|^{r'} \right)^{\frac{p}{r'}} \lesssim \sum_{j=j_*+1}^{j_1} 2^{-j(s-1/r)p}. \end{aligned}$$

This with $2^{j_*} \sim n^{\frac{1}{2(s-1/r)+1}}$ shows that

$$G_{32} \lesssim \sum_{j=j_*+1}^{j_1} 2^{-j(s-1/r)p} \lesssim 2^{-j_*(s-1/r)p} \sim n^{-\frac{(s-1/r)p}{2(s-1/r)+1}}. \quad (34)$$

Therefore

$$G_3 = G_{31} + G_{32} \lesssim (\ln n)^{p/2} n^{-\frac{(s-1/r)p}{2(s-1/r)+1}}. \quad (35)$$

By (27), (29), (32), and (35) we have $G \lesssim (\ln n)^{\frac{3p}{2}} n^{-\frac{(s-1/r)p}{2(s-1/r)+1}}$. Then it is easy to see from (24), (25), and (26) that

$$\mathbb{E} \left[|\widehat{f}_n^{\text{non}}(y) - f(y)|^p \right] \lesssim (\ln n)^{\frac{3p}{2}} n^{-\frac{(s-1/r)p}{2(s-1/r)+1}}.$$

This ends the proof. \square

Acknowledgements

The authors would like to thank the referees and editor for their important comments and suggestions.

Funding

This paper is supported by Guangxi Natural Science Foundation (No. 2018GXNSFBA281076), Guangxi Science and Technology Project (Nos. Guike AD18281058 and Guike 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.

Competing interests

The authors declare that they have no competing interests.

Authors' contributions

Both authors contributed equally to the writing of this paper. Both authors read and approved the final manuscript.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Received: 29 March 2019 Accepted: 17 July 2019 Published online: 25 July 2019

References

1. Alam, K., Saxena, K.M.L.: Positive dependence in multivariate distributions. *Commun. Stat., Theory Methods* **10**, 1183–1196 (1981)
2. Cai, T.T.: Rates of convergence and adaptation over Besov space under pointwise risk. *Stat. Sin.* **13**, 881–902 (2003)
3. Chesneau, C.: Wavelet block thresholding for density estimation in the presence of bias. *J. Korean Stat. Soc.* **39**, 43–53 (2010)
4. Cox, D.R.: *Some Sampling Problems in Technology*. Wiley, New York (1969)
5. Daubechies, I.: *Ten Lectures on Wavelets*. SIAM, Philadelphia (1992)
6. Donoho, D.L., Johnstone, I.M., Keryacharian, G., Picard, D.: Density estimation by wavelet thresholding. *Ann. Stat.* **2**, 508–539 (1996)
7. Härdle, W., Kerkycharian, G., Picard, D., Tsybakov, A.: *Wavelets, Approximation and Statistical Application*. Springer, New York (1997)
8. Heckman, J.: *Selection Bias and Self-Selection*. Macmillan Press, New York (1985)
9. Joag-Dev, K., Proschan, F.: Negative association of random variables with applications. *Ann. Stat.* **11**, 286–295 (1983)
10. Kou, J.K., Guo, H.J.: Wavelet density estimation for mixing and size-biased data. *J. Inequal. Appl.* **2018**, 189 (2018)
11. Liu, Y.M., Wu, C.: Point-wise estimation for anisotropic densities. *J. Multivar. Anal.* **17**, 112–125 (2019)
12. Liu, Y.M., Xu, J.L.: Wavelet density estimation for negatively associated stratified size-biased sample. *J. Nonparametr. Stat.* **26**, 537–554 (2014)
13. Ramírez, P., Vidakovic, B.: Wavelet density estimation for stratified size-biased sample. *J. Stat. Plan. Inference* **140**, 419–432 (2010)
14. Rebelles, G.: Pointwise adaptive estimation of a multivariate density under independence hypothesis. *Bernoulli* **21**, 1984–2023 (2015)
15. Shirazi, E., Doosti, H.: Multivariate wavelet-based density estimation with size-biased data. *Stat. Methodol.* **27**, 12–19 (2015)

Submit your manuscript to a SpringerOpen[®] journal and benefit from:

- Convenient online submission
- Rigorous peer review
- Open access: articles freely available online
- High visibility within the field
- Retaining the copyright to your article

Submit your next manuscript at ► [springeropen.com](https://www.springeropen.com)