Adv. & Appl. in Stat. **6**(3) (2006), 353-359

UNIFORM COVERAGE RATE OF NONPARAMETRIC KERNEL ESTIMATE UNDER MIXING DEPENDENT DATA

FULONG HAN[1], RIQUAN ZHANG[1, 2], JINGYAN FENG[1] and ZHIQIANG ZHANG[1]

[1]Department of Mathematics
Shanxi Datong University
Datong, Shanxi, 037009, P. R. China
e-mail: zhangriquan@163.com

[2]Department of Statistics
East China Normal University
Shanghai, 20062, P. R. China

Abstract

In nonparametric models or semi-parametric models, the estimates of unknown mean functions or condition mean of unknown functions are always considered either under independent data or under dependent data. In this paper, under mild conditions we give a solution for this problem and uniform coverage rates of nonparametric Kernel estimates under mixing dependent data.

1. Introduction

In the case of studying the nonparametric model $m(x) = E(Y \mid X = x)$ with unknown mean function $m(x)$ for statistics analysis, the estimates of $m(x)$ are always considered; and the estimates of $m(x, \theta)$ are always

2000 Mathematics Subject Classification: Primary 62G08, Secondary 62G20.

Keywords and phrases: non-parameter, semi-parameter, Kernel estimate, α-mixing dependent, uniform coverage rate.

Received June 17, 2006

considered in studying the semi-parameter model $m(x, \theta) = E(Y \mid X = x)$ with unknown mean function $m(x, \theta)$ for estimating θ and statistics analysis. Generally, suppose that $G(z, \theta)$ is a known function up to θ, where $z = (y, x)$ and θ is a parameter. To estimate θ, or some function of conditional mean of $G(Z, \theta)$ under X, the estimate of $E(G(Z, \theta) \mid X = x)$ is necessary. So the estimate of $E(G(Z, \theta) \mid X = x)$ is important to study nonparametric and semi-parametric models for statistics analysis. The above problem can be summarized as the model,

$$m(x, \theta) = E(G(Z, \theta) \mid X = x), \tag{1}$$

where $G(Z, \theta)$ is a measurable known function: $\mathbf{Z} \times \Theta \to R$, $Z = (Y, X) \in \mathbf{Z}$ $= (R, \mathbf{X})$, $\mathbf{X}$ has a compact subset in R^s, $\theta \in \Theta$, and Θ is a compact subset in R^b. In this paper we consider the estimate of $m(x, \theta)$ under the case where the sample is α-mixing dependent.

2. Estimation and Main Result

Suppose that $\{Z_1, \ldots, Z_n\}$ is from the model (1), and $f(x)$ is the marginal density of an s-dimensional random variable X. Define the estimate of $f(x)$ as

$$\hat{f}(x) = \frac{1}{nh_n^s} \sum_{j=1}^{n} K\left(\frac{X_j - x}{h_n}\right), \tag{2}$$

where $K(u)$ is a Kernel function, h_n is a bandwidth, and $h_n \to 0$ as $n \to \infty$. $\hat{f}(x)$ is called a *Kernel estimate* of $f(x)$.

Let $g(x, \theta) = m(x, \theta) f(x)$. Then define the Kernel estimate of $g(x, \theta)$ as

$$\hat{g}(x, \theta) = \frac{1}{nh_n^s} \sum_{j=1}^{n} G(Z_j, \theta) K\left(\frac{X_j - x}{h_n}\right). \tag{3}$$

Furthermore, the *Kernel estimate* of $m(x, \theta)$ is defined as

$$\hat{m}(x, \theta) = \hat{g}(x, \theta) / \hat{f}(x). \tag{4}$$

Conditions:

C1. $\{Z_t\}$ is α-mixing dependent strictly stationary random sequence with the mixing coefficient $\alpha(n)$ satisfying $\sum_{j=1}^{\infty} \alpha(j)^{\delta/(2+\delta)} < \infty$, where δ is a positive number;

C2. $E[G(Z, \theta) | X = x]$ has a continuous J-th derivative with respect to x, and these derivatives are uniformly bounded in $\mathbf{X} \times \Theta$, and the $2 + \delta$ order moment of $G(Z, \theta)$ is bounded, where $(J \geq 2)$.

C3. There are two positive numbers c_1 and c_2 such that $c_1 \leq f(x) \leq c_2$ and $f(x)$ has a continuous J-th derivative.

C4. $K(u)$ is a continuous symmetrical probability density function with bounded compact support and J order.

C5. $h_n \to 0$, $nh_n^{2s} \to \infty$, as $n \to \infty$.

Lemma 1. *Suppose that conditions* C1-C5 *hold. Then*

$$\sup_{(x,\, \theta) \in \mathbf{X} \times \Theta} | E[\hat{g}(x, \theta)] - g(x, \theta) | = O(h_n^J). \tag{5}$$

Lemma 2. *Suppose that conditions* C1-C5 *hold. Then, for* $\forall \theta \in \Theta$,

$$\sup_{x \in \mathbf{X}} | \hat{g}(x, \theta) - E[\hat{g}(x, \theta)] | = O_p(n^{-1/2} h_n^{-s}). \tag{6}$$

Theorem 1. *Suppose that conditions* C1-C5 *hold. Then, for* $\forall \theta \in \Theta$,

$$\sup_{x \in \mathbf{X}} | \hat{g}(x, \theta) - g(x, \theta) | = O_p(h_n^J + n^{-1/2} h_n^{-s}). \tag{7}$$

Corollary 1. *Suppose that conditions* C1-C5 *hold. Then*

$$\sup_{x \in \mathbf{X}} | \hat{f}(x) - f(x) | = O_p(h_n^J + n^{-1/2} h_n^{-s}). \tag{8}$$

Theorem 2. *Suppose that conditions* C1-C5 *hold. Then, for* $\forall \theta \in \Theta$,

$$\sup_{x \in \mathbf{X}} | \hat{m}(x, \theta) - m(x, \theta) | = O_p(h_n^J + n^{-1/2} h_n^{-s}). \tag{9}$$

Remark. From Theorem 1, Corollary 1 and Theorem 2, we can see that the uniform coverage rate of $\hat{g}(x, \theta)\hat{f}(x)$ or $\hat{m}(x, \theta)$ has two parts. The first part is the bias $O_p(h_n^J)$, which depends on the order of the Kernel function $K(u)$, and the smooth degree of $f(x)$. The higher the order of the Kernel and the smooth degree of $f(x)$ are, the faster the coverage rate is, otherwise the slower the coverage rate is. The second part is $O_p(n^{-1/2}h_n^{-s})$, which depends on the dimensionality of x. The higher it is, the slower the coverage rate is, which is called *curse of dimensionality*. Since the β-mixing dependent and γ-mixing dependent are two special cases of the α-mixing dependent. So the above results hold in the case where the sample is β-mixing dependent or γ-mixing dependent.

3. Proof

Proof of Lemma 1. Since $\{Z_j\}$ is identical distribution,

$$E[\hat{g}(x, \theta)] - g(x, \theta) = h_n^{-s}E[G(Z, \theta)K(h_n^{-1}(X - x))] - m(x, \theta)f(x)$$

$$= h_n^{-s}\int E[G(Z, \theta)|X = t]K(h_n^{-1}(t - x))f(t)dt - m(x, \theta)f(x).$$

Let

$$u = (t - x)/h_n.$$

Then

$$E[\hat{g}(x, \theta)] - g(x, \theta)$$

$$= \int E(G(Z, \theta)|X = x + uh_n)K(u)f(x + uh_n)du - m(x, \theta)f(x).$$

By using Taylor's expansions in $E(G(Z, \theta)|X = x + uh_n)$ and $f(x + uh_n)$ at x, and conditions C2-C4, we have

$$\sup_{(x, \theta)\in \mathbf{X}\times\Theta} |E[\hat{g}(x, \theta)] - g(x, \theta)| = O(h_n^J).$$

Proof of Lemma 2. Let $\varphi(t) = \int \exp\{it'u\}K(u)du$ be a Fourier transformation of $K(u)$, and $v = t/h_n$. Then, for any $\theta \in \Theta$, we have

$$E[\sup_{x\in\mathbf{X}} |\hat{g}(x, \theta) - E(\hat{g}(x, \theta))|]$$

$$= (nh_n^s)^{-1} E\left[\sup_{x\in\mathbf{X}} \left| \sum_{j=1}^{n} \left(G(Z_j, \theta) K\left(\frac{X_j - x}{h_n}\right) - E\left(G(Z_j, \theta) K\left(\frac{X_j - x}{h_n}\right)\right)\right)\right|\right]$$

$$\leq (nh_n^s)^{-1} E\left[\sup_{x\in\mathbf{X}} \left| \sum_{j=1}^{n} \left(G(Z_j, \theta) \int \varphi(t) \exp\left(\frac{it'(X_j - x)}{h_n}\right) dt \right.\right.\right.$$

$$\left.\left.\left. - E\left(G(Z_j, \theta) \int \varphi(t) \exp\left(\frac{it'(X_j - x)}{h_n}\right) dt\right)\right)\right|\right]$$

$$\leq (nh_n^s)^{-1} E\left[\left| \sum_{j=1}^{n} \left(G(Z_j, \theta) \int \varphi(t) \exp\left(\frac{it'X_j}{h_n}\right) dt \right.\right.\right.$$

$$\left.\left.\left. - E\left(G(Z_j, \theta) \int \varphi(t) \exp\left(\frac{it'X_j}{h_n}\right) dt\right)\right)\right|\right]$$

$$\leq \int \varphi(t) (nh_n^s)^{-1} E\left| \sum_{j=1}^{n} \left(G(Z_j, \theta) \exp\left(\frac{it'X_j}{h_n}\right) \right.\right.$$

$$\left.\left. - E\left(G(Z_j, \theta) \exp\left(\frac{it'X_j}{h_n}\right)\right)\right)\right| dt$$

$$= \int \varphi(h_n v) E\left| n^{-1} \sum_{j=1}^{n} [G(Z_j, \theta) \exp(iv'X_j) - E(G(Z_j, \theta) \exp(iv'X_j))]\right| dv. \quad (10)$$

Let

$$\Delta = n^{-1} E\left| \sum_{j=1}^{n} [G(Z_j, \theta)\exp(iv'X_j) - E(G(Z_j, \theta)\exp(iv'X_j))\right|.$$

Then

$$\Delta^2 = n^{-1}E[G(Z_1, \theta)\exp(iv'X_1) - E(G(Z_1, \theta)\exp(iv'X_1))]^2$$

$$+ n^{-2}\sum_{j \neq k} E[G(Z_j, \theta)\exp(iv'X_j)G(Z_k, \theta)\exp(iv'X_k)$$

$$- E(G(Z_j, \theta)\exp(iv'X_j))E(G(Z_k, \theta)\exp(iv'X_k))]$$

$$\leq \frac{M_1}{n} + \frac{M_2}{n}\sum_{i=1}^{n-1} \alpha(i)^{\frac{\delta}{2+\delta}},$$

which implies $\Delta^2 = O(n^{-1})$, where M_1 and M_2 are two positive numbers. Furthermore,

$$\Delta = O(n^{-1/2}). \tag{11}$$

By Replacing (10) by (11), we have

$$E[\sup_{x \in \mathbf{X}} | \hat{g}(x, \theta) - E(\hat{g}(x, \theta)) |] = O(n^{-1/2}h_n^{-s}),$$

therefore,

$$\sup_{x \in \mathbf{X}} | \hat{g}(x, \theta) - E(\hat{g}(x, \theta)) | = O_p(n^{-1/2}h_n^{-s}).$$

Proof of Theorem 1. Since

$$\sup_{x \in \mathbf{X}} | \hat{g}(x, \theta) - g(x, \theta) | \leq \sup_{x \in \mathbf{X}} | \hat{g}(x, \theta) - E(\hat{g}(x, \theta)) |$$

$$+ \sup_{x \in \mathbf{X}} | E(\hat{g}(x, \theta)) - g(x, \theta) |,$$

we get the result of Theorem 1 from Lemmas 1 and 2.

Proof of Corollary 1. Let $G(z, \theta) = 1$ in Theorem 1. Then the result of corollary can be obtained.

Proof of Theorem 2. Notice that

$$f(x) - \hat{f}(x) \leq \sup_{x \in \mathbf{X}} | \hat{f}(x) - f(x) |,$$

then

$$f(x)\hat{f}(x) \geq f^2(x) - f(x)\sup_{x \in \mathbf{X}} | \hat{f}(x) - f(x) | \geq c_1^2 - O_p(n^{-1/2}h_n^{-s}) = O_p(1). \quad (12)$$

Furthermore,

$$\sup_{x \in \mathbf{X}} | \hat{m}(x, \theta) - m(x, \theta) |$$

$$= \sup_{x \in \mathbf{X}} \left| \frac{f(x)(\hat{g}(x, \theta) - g(x, \theta)) + (\hat{f}(x) - f(x))g(x, \theta)}{\hat{f}(x)f(x)} \right|$$

$$\leq \frac{\sup_{x \in \mathbf{X}} | f(x)(\hat{g}(x, \theta) - g(x, \theta)) | + \sup_{x \in \mathbf{X}} | (\hat{f}(x) - f(x))g(x, \theta) |}{\inf_{x \in \mathbf{X}} | \hat{f}(x)f(x) |}$$

from (12), Theorem 1 and Corollary 1 we have

$$\sup_{x \in \mathbf{X}} | \hat{m}(x, \theta) - m(x, \theta) | = O_p(n^{-1/2}h_n^{-s}).$$

References

[1] P. K. Sen and J. M. Singer, Large Sample Methods in Statistics: An Introduction with Applications, Chapman & Hall, New York, 1993.

[2] G. S. Watson, Smooth regression analysis, Sankhya Ser. A 26 (1964), 359-372.