Statistics and Probability Letters 80 (2010) 169–176

Contents lists available at ScienceDirect

Statistics and Probability Letters

journal homepage: www.elsevier.com/locate/stapro

Moderate deviations for deconvolution kernel density estimators with ordinary smooth measurement errors

Weixing Song

Department of Statistics, Kansas State University, Manhattan, 108F, Dickens Hall KS, 66502, United States

ARTICLE INFO

Article history:
Received 18 May 2009
Received in revised form 9 October 2009
Accepted 9 October 2009
Available online 17 October 2009

MSC:
primary 62G08
secondary 62G10

ABSTRACT

In this paper, we establish the pointwise and uniform moderate deviations limit results for the deconvolution kernel density estimator in the errors-in-variables model, when the measurement error possesses an ordinary smooth distribution. The results are similar to the moderate deviations theorems for the classical kernel density estimators, but a factor related to the ordinary smooth order is needed to account for the measurement errors.

1. Introduction

In many practical situations, the random variable X cannot be observed directly, instead, a surrogate Y can be obtained which is related to X in the manner $Y = X + \varepsilon$, where ε is the measurement error with a known density function, X and ε are independent. This is the so-called errors-in-variables or the convolution model. The errors-in-variables modeling is widely used in microfluorimetry, nutrition study, electrophoresis, biostatistics, and some other areas. The investigations on its theories and applications have received much attention in the past decades. A typical example of the errors-in-variables model is given in Fan (1991) about the AIDS study, where Y may be the time from some starting point to the time that symptoms appear, ε may be the time from the start point to the time that infection occurs, and X is the time from the occurrence of infection to the time of symptoms. A comprehensive introduction to the errors-in-variables model, as well as many other real data examples, can be found in Fuller (1987), Carroll et al. (1995) and the references therein.

The problem of interest in errors-in-variables model is to estimate the density function of X. Because X is not observable and its density is unknown, the kernel type estimator cannot be applied here. The most commonly used estimator for the density f_X of X is the deconvolution kernel density estimator, which is based on the additive structure of X and ε. It is well known that this additivity structure, and the independence of X and ε imply that $\phi_Y(t) = \phi_X(t)\phi_\varepsilon(t)$ for $t \in \mathbb{R}$, where $\phi_\cdot(t)$ denotes the characteristic function of a random variable which is specified by the subscript. Since the density function of ε, denoted by f_ε, is known hence ϕ_ε is known, and Y is observable hence ϕ_Y can be estimated, so the inverse Fourier transformation on $\hat{\phi}_Y(t)/\phi_\varepsilon(t)$ can provide us an estimator for f_X, where $\hat{\phi}_Y(t)$ is an estimate of the characteristic function of Y. Using a kernel function K with bandwidth h_n, Stefanski and Carroll (1991) consider the following so-called deconvolution kernel density estimator of X:

$$\hat{f}_n(x) = \frac{1}{2\pi}\int_{-\infty}^{\infty} \exp(-\mathbf{i}tx)\frac{\phi_K(th_n)\hat{\phi}_n(t)}{\phi_\varepsilon(t)}\mathrm{d}t, \qquad \mathbf{i} = \sqrt{-1} \tag{1.1}$$

where $\phi_K(t)$ is the characteristic function of some kernel function K, $\hat{\phi}_n(t)$ is the empirical characteristic function of the sample $Y_1, \dots, Y_n$ defined by $\hat{\phi}_n(t) = n^{-1}\sum_{i=1}^{n}\exp(-\mathbf{i}tY_i)$, and h_n has the same meaning as the bandwidth in the classical

E-mail address: weixing@ksu.edu.

0167-7152/$ – see front matter
doi:10.1016/j.spl.2009.10.003

kernel density estimators. Let

$$g_n(y) = \frac{1}{2\pi}\int_{-\infty}^{\infty} \exp(-\mathbf{i}ty)\frac{\phi_K(t)}{\phi_\varepsilon(t/h_n)}\mathrm{d}t,$$

then (1.1) has the following form which resembles the classical kernel density estimator,

$$\hat{f}_n(x) = \frac{1}{n}\sum_{i=1}^{n}\frac{1}{h_n}g_n\left(\frac{x-Y_i}{h_n}\right).$$

g_n is significantly different from the classical kernel functions in that g_n might be a complex function, and it depends on the bandwidth h_n, hence on the sample size.

The large sample behavior of $\hat{f}_n(x)$ strongly depends on the smoothness of the distribution of the measurement error ε. Using the terms in Fan and Truong (1993), a distribution is called ordinary smooth if the tails of its characteristic function decay to 0 at an algebraic rate; it is called supersmooth if its characteristic function has tails approaching 0 exponentially fast. It is known that the local and global rates of convergence of the sequences of deconvolution kernel density estimators are slower than that of the classical kernel density estimators. Moreover, the convergence rate are much slower in the supersmooth cases than in the ordinary smooth cases. Related works on the consistency and the asymptotic normality of the deconvolution kernel density estimator $\hat{f}_n(x)$ can be found in Carroll and Hall (1988), Devroye (1989), Fan (1991), and Fan and Liu (1997) and the references therein. But the literature seems scant in the discussion of the moderate and large deviation results for the deconvolution kernel density estimators, while there are many such works on the classical kernel estimators, see Louani (1998) and Gao (2003) for more details. The current paper will try to fill this void partly. In particular, we will focus on the moderate deviation limit theorems when the measurement error has an ordinary smooth distribution. Further research will be conducted in the future for the large deviation results for the ordinary smooth measurement models, and the moderate and large deviation results for the supersmooth cases. Another interesting topic will be the moderate and large deviation results for the nonparametric prediction in measurement error regression models based on the work of Carroll et al. (2009). Since a parametric rate $n^{-1/2}$ can be achieved for the nonparametric prediction, so the deviation results should be similar to the nonparametric prediction in the classical regression models.

The paper is organized as follows. Section 2 will present the necessary assumptions for the theory, the pointwise and uniform moderate deviation results will be also introduced; the proofs of the main results will be provided in Section 3.

2. Assumptions and main results

The ordinary smooth condition on ε often takes the following form

(C1) $t^\beta\phi_\varepsilon(t) \to c$, $t^{\beta+1}\phi'_\varepsilon(t) \to -\beta c$, as $t \to +\infty$, with some constant $c \neq 0$ and $\beta \geq 0$. Moreover, $\phi_\varepsilon(t) \neq 0$, for all t,

where β is called the smooth order.

For example, the Gamma density function $\alpha^p x^{p-1}\exp(-\alpha x)I(x > 0)/\Gamma(p)$ is ordinary smooth with $\beta = p$, and the double exponential density function $\exp(-|x|)/2$ is ordinary smooth with $\beta = 2$.

The following condition is needed to guarantee the boundedness of the density function Y, which is similar to condition (A1) in Gao (2003).

(C2) The density function of ε is continuous and $f_\varepsilon(t) \to 0$ as $t \to \infty$.

Obviously, the above mentioned Gamma and double exponential distributions satisfy (C2).

As for the kernel function, we shall assume

(C3) $\phi_K(t)$ is a symmetric function, having second bounded integrable derivatives.
(C4) $\int_{-\infty}^{\infty}[|\phi_K(t)| + |\phi'_K(t)|]|t|^\beta \mathrm{d}t < \infty$, $\int_{-\infty}^{\infty}|t|^{2\beta}|\phi_K(t)|^2\mathrm{d}t < \infty$.

The following condition will be imposed on the density function of X.

(C5) The second derivative of the unknown f_X exists and is continuous.

For the bandwidth h_n, we shall assume

(C6) h_n, b_n are sequence of positive numbers such that,

$$nh_n \to \infty, \qquad \frac{b_n}{nh_n} \to 0, \qquad \frac{nh_n \log h_n}{b_n^2} \to 0.$$

Similar to Giné and Guillou (2001) and Gao (2003), the following condition is needed when using Giné and Guillou (2001)'s exponential inequality to prove the uniform moderate deviation result,

(C7) $g_n(y)$ is a bounded, square integrable function in the linear span (the set of finite linear combinations) of functions $q \geq 0$ satisfying the following property: the subgraph of q, $\{(s,u) : q(s) \geq u\}$, can be represented as a finite number of Boolean operations among sets of the form $\{(s,u) : w(s,u) \geq \psi(u)\}$, where w is a polynomial on $\mathbb{R}^2$ and ψ is an arbitrary real function.

As an example, if ε has a double exponential distribution $f_\varepsilon(x) = (\sqrt{2}\sigma_0^2)\exp(-\sqrt{2}|x|/\sigma_0)$ with the characteristic function $\phi_\varepsilon(t) = (1+\sigma_0^2 t^2/2)^{-1}$, and K is further chosen to be the Gaussian kernel $K(x) = (\sqrt{2\pi})^{-1}\exp(-x^2/2)$, then one can show that

$$g_n(y) = \frac{1}{\sqrt{2\pi}}\exp\left(-\frac{y^2}{2}\right)\left[1-\frac{\sigma_0^2}{2h_n^2}(y^2-1)\right].$$

Easy to see that $g_n(y)$ is a bounded real function of bounded variation, hence $g_n(y)$ satisfies (C7).

Under above conditions, we can show the following pointwise moderate deviation limit result for the estimator $\hat{f}_n(x)$.

Theorem 2.1. *Suppose* (C1)–(C6) *hold. Then for any* $x \in \mathbb{R}$, *any closed set* $F \subset \mathbb{R}$, *any open set* $G \subset \mathbb{R}$, *we have*

$$\limsup_{n\to\infty}\frac{nh_n}{b_n^2}\log P\left(\frac{nh_n^{\beta+1}}{b_n}[\hat{f}_n(x)-E\hat{f}_n(x)] \in F\right) \le -\inf_{\lambda\in F} I_x(\lambda),$$

$$\liminf_{n\to\infty}\frac{nh_n}{b_n^2}\log P\left(\frac{nh_n^{\beta+1}}{b_n}[\hat{f}_n(x)-E\hat{f}_n(x)] \in G\right) \ge -\inf_{\lambda\in G} I_x(\lambda),$$

where

$$I_x(t) = \frac{\lambda^2}{2G(x)}, \quad G(x) = \frac{f_Y(x)}{2\pi|c|^2}\int_{-\infty}^{\infty}|t|^{2\beta}|\phi_K(t)|^2\mathrm{d}t.$$

and f_Y *is the density function of* Y.

Furthermore, we also have the following uniform moderate deviation limit theorem for the estimator $\hat{f}_n(x)$.

Theorem 2.2. *If assumptions* (C1)–(C7) *hold, then for any* $\lambda > 0$,

$$\lim_{n\to\infty}\frac{nh_n}{b_n^2}\log P\left(\frac{nh_n^{\beta+1}}{b_n}\left\|\hat{f}_n - E\hat{f}_n\right\|_\infty > \lambda\right) = -I(\lambda),$$

where $\|\cdot\|_\infty$ *denotes the supreme norm of a function and*

$$I(\lambda) = \frac{\lambda^2}{2\|f_Y\|_\infty L}, \quad L = \frac{1}{2\pi|c|^2}\int_{-\infty}^{\infty}|t|^{2\beta}|\phi_K(t)|^2\mathrm{d}t.$$

If we further assume

(C8) $\phi_K(t) = 1 + O(|t|^2)$ as $t \to 0$,
(C9) $nh_n^{\beta+3}/b_n \to 0$ as $n \to \infty$,
(C10) $f_X(x)$ and its second derivative is bounded,

then the following corollary holds.

Corollary 2.1. *Suppose* (C1)–(C10) *hold, then for any* $x \in \mathbb{R}$, *any closed set* $F \subset \mathbb{R}$, *any open set* $G \subset \mathbb{R}$,

$$\limsup_{n\to\infty}\frac{nh_n}{b_n^2}\log P\left(\frac{nh_n^{\beta+1}}{b_n}[\hat{f}_n(x)-f_X(x)] \in F\right) \le -\inf_{\lambda\in F} I_x(\lambda),$$

$$\liminf_{n\to\infty}\frac{nh_n}{b_n^2}\log P\left(\frac{nh_n^{\beta+1}}{b_n}[\hat{f}_n(x)-f_X(x)] \in G\right) \ge -\inf_{\lambda\in G} I_x(\lambda),$$

and for any $\lambda > 0$,

$$\lim_{n\to\infty}\frac{nh_n}{b_n^2}\log P\left(\frac{nh_n^{\beta+1}}{b_n}\left\|\hat{f}_n - f_X(x)\right\|_\infty > \lambda\right) = -I(\lambda),$$

where $I_x(\lambda)$ *and* $I(\lambda)$ *are the same as in Theorems 2.1 and 2.2, respectively.*

Condition (C8) is also used in Fan (1991). A typical example of K satisfying (C8) is the standard normal density function. Condition (C9) guarantees the bias of $\hat{f}_n(x)$ converges to 0 sufficiently fast, while the condition (C10) ensures this convergence is uniform in $x \in \mathbb{R}$.

3. The proofs of main results

Proof of Theorem 2.1. From Fan (1991), conditions (C1), (C3)–(C5) imply

$$|h_n^\beta g_n(y)| \le B(y) \hat{=} \min(C_2, C_1/y), \tag{3.1}$$

where C_1, C_2 are two constants, and

$$h_n^\beta g_n(y) \to C(y) \hat{=} \frac{1}{2\pi c}\int_{-\infty}^{\infty} \exp(-\mathbf{i}ty)t^\beta \phi_K(t)\mathrm{d}t \tag{3.2}$$

as $n \to \infty$. For convenience, let $Z_{ni} = h_n^\beta g_n\left((x - Y_i)/h_n\right)$. Note that assumption (C1) implies that the density function $f_Y(\cdot)$ of Y is bounded. Therefore, for any $k \ge 3$, we have

$$\begin{aligned} E|Z_{ni}|^k &= \int_{-\infty}^{\infty}\left|h_n^\beta g_n\left(\frac{x-y}{h_n}\right)\right|^k f_Y(y)\mathrm{d}y = h_n\int_{-\infty}^{\infty}\left|h_n^\beta g_n(y)\right|^k f_Y(x-yh)\mathrm{d}y \\ &\le h_n\|f_Y(y)\|_\infty\cdot\int_{-\infty}^{\infty}\left|h_n^\beta g_n(y)\right|^k\mathrm{d}y \le h_n\|f_Y(y)\|_\infty\int_{-\infty}^{\infty}|B(y)|^k\,\mathrm{d}y. \end{aligned}$$

To proceed, we need an upper bound for the above integration. Let $d = \max\{C_1, C_2\}$. Then

$$\int_{-\infty}^{\infty}|B(y)|^k\,\mathrm{d}y = 2d^k\int_0^\infty[1\wedge 1/y^k]\mathrm{d}y = 2d^k\int_0^1\mathrm{d}y + 2d^k\int_1^\infty\frac{1}{y^k}\mathrm{d}y = 2d^k\left[1+\frac{1}{k-1}\right] \le 3d^k.$$

Therefore, $E|Z_{n1} - EZ_{n1}|^k \le 3\cdot 2^k d^k h\|f_Y(y)\|_\infty$ and

$$\begin{aligned} \sum_{k=3}^{\infty}\frac{1}{k!}E\left|\frac{tb_n}{nh_n}(Z_n - EZ_n)\right|^k &\le 3h_n\|f_Y(y)\|_\infty\sum_{k=3}^{\infty}\frac{1}{k!}\frac{2^k d^k t^k b_n^k}{n^k h^k} \\ &\le 3h_n\|f_Y(y)\|_\infty\left(\frac{2tdb_n}{nh_n}\right)^3\exp\left(\frac{2tdb_n}{nh_n}\right). \end{aligned}$$

By assumption (C6), we have

$$\sum_{k=3}^{\infty}\frac{1}{k!}E\left|\frac{tb_n}{nh_n}(Z_n - EZ_n)\right|^k = O\left(\left(\frac{2tdb_n}{nh_n}\right)^3 h_n\right).$$

Since $Y_1, \ldots, Y_n$ are i.i.d., so

$$\begin{aligned} \Phi_{n,x}(t) &= E\left(\exp[tb_n h_n^\beta(\hat{f}_n(x) - E\hat{f}_n(x))]\right) \\ &= E\left(\exp\left[\frac{tb_n}{nh_n}\sum_{i=1}^{n}(Z_{ni} - EZ_{ni})\right]\right) = \left[E\left(\exp\left[\frac{tb_n}{nh_n}(Z_{n1} - EZ_{n1})\right]\right)\right]^n \\ &= \left[1 + \frac{tb_n}{nh_n}E(Z_{n1} - EZ_{n1}) + \frac{t^2 b_n^2}{2n^2 h_n^2}E(Z_{n1} - EZ_{n1})^2 + O\left(\left(\frac{2tdb_n}{nh_n}\right)^3 h_n\right)\right]^n. \end{aligned}$$

By (3.1) and (3.2), we can obtain

$$h_n^{-1}E(Z_{n1} - EZ_{n1})^2 \to \frac{f_Y(x)}{2\pi|c|^2}\int_{-\infty}^{\infty}|t|^{2\beta}|\phi_K(t)|^2\mathrm{d}t \hat{=} G(x).$$

Therefore, we have

$$\Phi_{n,x}(t) = \left[1 + \frac{t^2 b_n^2}{2n^2 h_n}G(x)(1+o(1)) + O\left(\left(\frac{2tdb_n}{nh_n}\right)^3 h_n\right)\right]^n.$$

Finally, by letting $n \to \infty$,

$$\Phi_x(t) = \lim_{n\to\infty}\frac{nh_n}{b_n^2}\log\Phi_{n,x}(t) = \frac{t^2}{2}G(x),$$

and the dual of $\Phi_x(t)$ is given by

$$I_x(t) \hat{=} \sup_{t\in R}[t\lambda - \Phi_x(t)] = \frac{\lambda^2}{2G(x)}.$$

Note that $\Phi_x(t)$ is differentiable with respect to $t \in R$, from Gärtner–Ellis Theorem, we have the desired result. $\square$

To show the uniform moderate deviation result, we need an exponential inequality for the empirical process from Giné and Guillou (2001). For the sake of completeness, we state the exponential inequality below.

Lemma 3.1. *Let $\mathscr{F}$ be a measurable uniformly bounded VC class of functions, and let σ^2 and U be any numbers such that $\sigma^2 \ge \sup_{f\in\mathscr{F}} Var(f)$, $U \ge \sup_{f\in\mathscr{F}} \|f\|_\infty$, and $0 < \sigma \le U/2$. Then, there exist constants C and L depending only on the characteristic of the class $\mathscr{F}$, such that the inequality*

$$P\left(\left\|\sum_{i=1}^n [f(\xi_i) - Ef(\xi_i)]\right\|_\infty > t\right) \le L\exp\left[-\frac{t}{LU}\log\left(1+\frac{tU}{L(\sqrt{n}\sigma + U\sqrt{\log(U/\sigma)})^2}\right)\right]$$

holds for all

$$t \ge C\left(U\log\frac{U}{\sigma} + \sqrt{n}\sigma\sqrt{\log\frac{U}{\sigma}}\right).$$

More about the VC (Vapnik–Ȟervonenkis) class of functions, please see Pollard (1984). We also need the following lemmas which are similar to the lemmas 2.4 and 2.5 in Gao (2003).

Lemma 3.2. *Let assumptions* (C1)–(C7) *hold. For any $0 < \delta < 1$, let $I_{n,k}$, $k = 1, 2, \ldots, l_n$, be l_n intervals with length δh_n, such that $\{I_{n,k}, k = 1, \ldots, l_n\}$ is a covering of $[-h_n^{-1}, h_n^{-1}]$ and $l_n \le 2/\delta h_n^2 + 1$. Let $x_{n,k} \in I_{n,k}$, $1 \le k \le l_n$, $n \ge 1$. Then for any $\varepsilon > 0$,*

$$\lim_{\delta\to 0}\limsup_{n\to\infty}\frac{nh_n}{b_n^2}\log P\left(\frac{nh_n^{\beta+1}}{b_n}\sup_{1\le k\le l_n}\sup_{x\in I_{n,k}}|\hat{f}_{n,k}(x) - E\hat{f}_{n,k}(x)| \ge \varepsilon\right) = -\infty,$$

where $\hat{f}_{n,k}(x) = \hat{f}_n(x) - \hat{f}_n(x_{n,k})$.

Proof. By (C7), the class of functions $\mathscr{F} = \{h_n^\beta g_n((x-\cdot)/h_n), x \in \mathbb{R}, h_n \ne 0\}$ is a bounded measurable VC class of functions, so the following classes of functions

$$\mathscr{F}_{n,k} = \left\{h_n^\beta g_n\left(\frac{x-\cdot}{h_n}\right) - h_n^\beta g_n\left(\frac{x_{n,k}-\cdot}{h_n}\right); x \in I_{n,k}\right\}, \quad k = 1, \ldots, l_n;\ n \ge 1$$

are measurable VC classes of functions. Moreover, there is a common VC characteristic for all these classes that does not depend on k and n. Note that

$$\begin{aligned}&\int\left[h_n^\beta g_n\left(\frac{x-y}{h_n}\right) - h_n^\beta g_n\left(\frac{x_{n,k}-y}{h_n}\right)\right]^2 f_Y(y)\mathrm{d}y\\ &= h_n\int\left[h_n^\beta g_n(z) - h_n^\beta g_n\left(\frac{x_{n,k}-x}{h_n}+z\right)\right]^2 f_Y(x-hz)\mathrm{d}z \le 2\|f_Y\|_\infty \eta h_n.\end{aligned}$$

Now let us take $U_n = 2\|h_n^\beta g_n(y)\|_\infty$, $\mathscr{F} = \mathscr{F}_{n,k}$, and $\sigma_n^2 = 2\|f_Y\|_\infty\eta h_n$. Then for n large enough, we have $\sigma_n \le U_n/2$. Too see this point, note that, from (3.1), $U_n = 2\|h_n^\beta g_n(y)\|_\infty = O(1)$, and $\sigma_n = O(h_n^{1/2})$. We also have $\sqrt{n}\sigma_n \ge U_n\sqrt{\log(U_n/\sigma_n)}$ because of $h_n \to 0$ and $nh_n/\log(1/h_n) \to \infty$. Moreover, from (C6), one has

$$\limsup_{n\to\infty}\frac{U_n\log(U_n/\sigma_n) + \sqrt{n}\sigma_n\sqrt{\log(U_n/\sigma_n)}}{b_n} \le \limsup_{n\to\infty}\frac{2\sqrt{n}\sigma_n\sqrt{\log(U_n/\sigma_n)}}{b_n} = 0.$$

Consequently, there exists a positive integer number n_0 such that for any $n \ge n_0$,

$$b_n \ge C\left(U_n\log(U_n/\sigma_n) + \sqrt{n}\sigma_n\sqrt{\log(U_n/\sigma_n)}\right)$$

and $\sqrt{n}\sigma_n + U_n\sqrt{\log(U_n/\sigma_n)} \le 2\sqrt{2\eta nh_n\|f_Y\|_\infty}$.

Now applying Giné and Guillou (2001)'s lemma, for $n \ge n_0$, one has

$$P\left(\frac{nh_n^{\beta+1}}{b_n}\sup_{1\le k\le l_n}\sup_{x\in I_{n,k}}|\hat{f}_{n,k}(x) - E\hat{f}_{n,k}(x)| \ge \varepsilon\right) \le Ll_n\exp\left\{-\frac{b_n\varepsilon}{2LU_n}\log\left(1+\frac{b_n\varepsilon U_n}{4Lnh\|f_Y\|_\infty\eta}\right)\right\}.$$

Therefore,

$$\lim_{\delta\to 0}\lim\sup_{n\to\infty}\frac{nh_n}{b_n^2}\log P\left(\frac{nh_n^{\beta+1}}{b_n}\sup_{1\le k\le l_n}\sup_{x\in I_{n,k}}|\hat{f}_{n,k}(x)-E\hat{f}_{n,k}(x)|\ge\varepsilon\right)$$
$$\le\lim\sup_{n\to\infty}\frac{nh_n}{b_n^2}\log Ll_n-\lim_{n\to\infty}\frac{nh_n\varepsilon}{2LU_nb_n}\log\left(1+\frac{U_nb_n\varepsilon}{4nhL\eta\|f_Y\|_\infty}\right)$$
$$\le-\frac{\varepsilon^2}{8L^2\eta\|f_Y\|_\infty}.$$

Letting $\eta\to 0$ implies the desired result. □

Lemma 3.3. *Assume* (C1)–(C7) *hold. Then for any* $\varepsilon>0$,

$$\lim\sup_{n\to\infty}\frac{nh_n}{b_n^2}\log P\left(\frac{nh_n^{\beta+1}}{b_n}\sup_{x\notin[-h_n^{-1},h_n^{-1}]}\left|\hat{f}_n(x)-E(\hat{f}_n(x))\right|\ge\varepsilon\right)=-\infty.$$

Proof. Set

$$\mathscr{F}_n=\left\{h_n^\beta g_n\left(\frac{x-\cdot}{h_n}\right);x\notin[-h_n^{-1},h_n^{-1}]\right\}.$$

Then $\mathscr{F}_n$, $n\ge 1$ are measurable VC classes of functions with VC characteristic (A,v) that does not depend on n. Note that (C1) implies $f_Y(y)\to 0$ as $y\to\infty$. So, for any $\eta\in(0,\varepsilon)$, there exists an $n_0>0$ such that for any $n\ge n_0$,

$$\sup_{x\notin[-h_n^{-1},h_n^{-1}]}h_n^{2\beta}\int g_n^2\left(\frac{x-y}{h_n}\right)f_Y(y)\mathrm{d}y\le\eta h_n.$$

Now let $U_n=\|h_n^\beta g_n(x)\|_\infty$, $\mathscr{F}=\mathscr{F}_n$, and $\sigma_n^2=\eta h_n$. Then for n large enough, $\sigma_n\le U_n/2$, $\sqrt{n}\sigma_n\ge U_n\sqrt{\log(U_n/\sigma_n)}$. From (C6), one has,

$$\lim\sup_{n\to\infty}\frac{U_n\log(U_n/\sigma_n)+\sqrt{n}\sigma_n\sqrt{\log(U_n/\sigma_n)}}{b_n}\le\lim\sup_{n\to\infty}\frac{2\sqrt{n}\sigma_n\sqrt{\log(U_n/\sigma_n)}}{b_n}=0.$$

Consequently, there exists a positive integer number n_0 such that for any $n\ge n_0$,

$$b_n\ge C\left(U_n\log(U_n/\sigma_n)+\sqrt{n}\sigma_n\sqrt{\log(U_n/\sigma_n)}\right)$$

and $\sqrt{n}\sigma_n+U_n\sqrt{\log(U_n/\sigma_n)}\le 2\sqrt{\eta nh_n}$. Now applying Giné and Guillou's lemma, for $n\ge n_0$, one has

$$P\left(\frac{nh_n^{\beta+1}}{b_n}\sup_{x\notin[-h_n^{-1},h_n^{-1}]}|\hat{f}_n(x)-E\hat{f}_n(x)|\ge\varepsilon\right)\le L\exp\left\{-\frac{b_n\varepsilon}{LU_n}\log\left(1+\frac{U_nb_n\varepsilon}{4Lnh\eta}\right)\right\}.$$

Therefore,

$$\lim\sup_{n\to\infty}\frac{nh_n}{b_n^2}\log P\left(\frac{nh_n^{\beta+1}}{b_n}\sup_{x\notin[-h_n^{-1},h_n^{-1}]}|\hat{f}_n(x)-E\hat{f}_n(x)|\ge\varepsilon\right)$$
$$\le\lim_{n\to\infty}\frac{nh_n}{b_n^2}\log L-\lim_{n\to\infty}\frac{nh_n\varepsilon}{LU_nb_n}\log\left(1+\frac{U_nb_n\varepsilon}{4nhL\eta}\right)\le-\frac{\varepsilon^2}{4L^2\eta}.$$

Letting $\eta\to 0$ implies the desired result. □

Proof of Theorem 2.2. The lower bound is easy. In fact, for any $x\in\mathbb{R}$, from Theorem 2.1 we have

$$\lim\inf_{n\to\infty}\frac{nh_n}{b_n^2}\log P\left(\frac{nh_n^{\beta+1}}{b_n}\|\hat{f}_n-E\hat{f}_n\|_\infty\ge\lambda\right)$$
$$\ge\lim\inf_{n\to\infty}\frac{nh_n}{b_n^2}\log P\left(\frac{nh_n^{\beta+1}}{b_n}\|\hat{f}_n(x)-E\hat{f}_n(x)\|\ge\lambda\right)\ge-I_x(\lambda).$$

Hence

$$\lim\inf_{n\to\infty}\frac{nh_n}{b_n^2}\log P\left(\frac{nh_n^{\beta+1}}{b_n}\|\hat{f}_n-E\hat{f}_n\|_\infty\ge\lambda\right)\ge-I(\lambda).$$

To show the opposite inequality, note that $\|\hat{f}_n - E(\hat{f}_n)\|_\infty$ equals to

$$\max\left\{\sup_{x\in[-h_n^{-1},h_n^{-1}]}|\hat{f}_n(x) - E\hat{f}_n(x)|, \sup_{x\notin[-h_n^{-1},h_n^{-1}]}|\hat{f}_n(x) - E\hat{f}_n(x)|\right\},$$

and $\sup_{x\in[-h_n^{-1},h_n^{-1}]}|\hat{f}_n(x) - E\hat{f}_n(x)|$ is bounded above by

$$\max_{1\le k\le l_n}\{\sup_{x\in I_{n,k}}|\hat{f}_{n,k}(x) - E\hat{f}_{n,k}(x)| + |\hat{f}_n(z_{n,k}) - E\hat{f}_n(z_{n,k})|\}.$$

By Lemmas 3.2 and 3.3 and using inequality: $\max\{\log a, \log b\} \le \log(a+b) \le \log 2 + \max\{\log a, \log b\}$, $a \ge 0$, $b \ge 0$, we have that for any $0 < \varepsilon < \lambda$,

$$\begin{aligned}
&\limsup_{n\to\infty}\frac{nh_n}{b_n^2}\log P\left(\frac{nh_n^{\beta+1}}{b_n}\|\hat{f}_n - E\hat{f}_n\|_\infty > \lambda\right)\\
&\le \limsup_{n\to\infty}\frac{nh_n}{b_n^2}\log P\left(\frac{nh_n^{\beta+1}}{b_n}\sup_{x\in[-h_n^{-1},h_n^{-1}]}|\hat{f}_n(x) - E\hat{f}_n(x)| > \lambda\right)\\
&\le \lim_{\delta\to 0}\limsup_{n\to\infty}\frac{nh_n}{b_n^2}\log\left[P\left(\frac{nh_n^{\beta+1}}{b_n}\max_{1\le k\le l_n}\sup_{x\in I_{n,k}}|\hat{f}_{n,k}(x) - E\hat{f}_{n,k}(x)| \ge \varepsilon\right)\right.\\
&\quad\left. + P\left(\frac{nh_n^{\beta+1}}{b_n}\max_{1\le k\le l_n}|\hat{f}_n(z_{n,k}) - E\hat{f}_n(z_{n,k})| > \lambda - \varepsilon\right)\right]\\
&\le \lim_{\delta\to 0}\limsup_{n\to\infty}\frac{nh_n}{b_n^2}\log P\left(\frac{nh_n^{\beta+1}}{b_n}\max_{1\le k\le l_n}|\hat{f}_{n,k}(z_{n,k}) - E\hat{f}_{n,k}(z_{n,k})| > \lambda - \varepsilon\right).
\end{aligned}$$

By Chebyshev's inequality,

$$\begin{aligned}
&P\left(\frac{nh_n^{\beta+1}}{b_n}\max_{1\le k\le l_n}|\hat{f}_{n,k}(z_{n,k}) - E\hat{f}_{n,k}(z_{n,k})| > \lambda - \varepsilon\right)\\
&\le l_n\max_{1\le k\le l_n}\left\{\exp\left(-\frac{b_n^2}{nh_n}(\lambda-\varepsilon)t\right)\Phi_{z_{n,k}}(t)\right\}\\
&\le l_n\exp\left(-\frac{b_n^2}{nh_n}(\lambda-\varepsilon)t\right)\left[1 + \frac{b_n^2}{n^2h_n}\max_{x\in R}\Phi_x(t)(1+o(1)) + O\left(\left(\frac{2tdb_n}{nh_n}\right)^3 h_n\right)\right]^n.
\end{aligned}$$

Therefore,

$$\limsup_{n\to\infty}\frac{nh_n}{b_n^2}\log P\left(\frac{nh_n^{\beta+1}}{b_n}\|\hat{f}_n - E\hat{f}_n\|_\infty > \lambda\right) \le -\left((\lambda-\varepsilon)t - \max_{x\in\mathbb{R}}\Phi_x(t)\right).$$

The desired result follows from the fact that the last inequality holds for arbitrary $0 < \varepsilon < \lambda$ and any $t > 0$. □

Proof of Corollary 2.1. To show the first two claims one only has to investigate the term $\Phi_{n,x}(t)$. Now $\Phi_{n,x}(t) = E\exp(tb_nh_n^\beta[\hat{f}_n(x) - f_X(x)])$, it can be rewritten as

$$\Phi_{n,x}(t) = E\exp(tb_nh_n^\beta[\hat{f}_n(x) - E\hat{f}_n(x)]) \cdot \exp(tb_nh_n^\beta[E\hat{f}_n(x) - f_X(x)]).$$

So, it suffices to show that

$$\frac{nh_n}{b_n^2}\cdot tb_nh_n^\beta[E\hat{f}_n(x) - f_X(x)] \to 0$$

as $n \to \infty$. The conditions (C8) and (C10) imply

$$E\hat{f}_n(x) - f_X(x) = O(h_n^2)\quad \text{uniformly in } x \in \mathbb{R}, \tag{3.3}$$

so the desired result follows from (C9).

The third claim follows from similar argument, hence omitted here for brevity. □

References

Carroll, R.J., Hall, P., 1988. Optimal rates of convergence for deconvoluting a density. JASA 83, 1184–1186.
Carroll, R.J., Delaigle, A., Hall, P., 2009. Nonparametric prediction in measurement error models. JASA 104, 993–1003.
Carroll, R.J., Ruppert, D., Stefanski, L.A., 1995. Measurement Error in Nonlinear Models. Chapman & Hall/CRC, Boca Raton.
Devroye, L., 1989. A note on consistent deconvolution in density estimation. Canad. J. Statist. 17, 235–239.
Fan, J., 1991. Asymptotic normality for deconvolution kernel density estimators. Sankhyā 53 (Series A), 97–110.
Fan, J., Truong, K.T., 1993. Nonparametric regression with errors in variables. Ann. Statist. 21, 1900–1925.
Fan, Y.Q., Liu, Y.J., 1997. A note on asymptotic normality for deconvolution kernel density estimators. Sankhyā 59 (Series A), 138–141.
Fuller, W.A., 1987. Measurement Error Models. Wiley, New York.
Gao, F.Q., 2003. Moderate deviations and large deviations for kernel density estimators. J. Theoret. Probab. 16 (2), 401–418.
Giné, E., Guillou, A., 2001. On consistency of kernel density estimators for randomly censored data: rates holding uniform over adaptive intervals. Ann. Inst. H. Poincaré Probab. Statist. 37, 503–522.
Louani, D., 1998. Large deviations limit theorems for the kernel density estimator. Scand. J. Statist. 25, 243–253.
Pollard, D., 1984. Convergence of Stochastic Processes. Springer-Verlag.
Stefanski, L.A., Carroll, R.J., 1991. Deconvolution kernel density estimators. Statistics 21, 169–184.