PROBABILITY
AND
MATHEMATICAL STATISTICS
Vol. 39, Fasc. 1 (2019), pp. 127–138
doi:10.19195/0208-4147.39.1.9

COMPLETE CONSISTENCY FOR RECURSIVE PROBABILITY DENSITY ESTIMATOR OF WIDELY ORTHANT DEPENDENT SAMPLES*

BY

CHENLU ZHUANSUN (HEFEI) AND XIAOXIN LI (CHIZHOU)

Abstract. In this paper, we will study the recursive density estimators of the probability density function for widely orthant dependent (WOD) random variables. The complete consistency and complete convergence rate are established under some general conditions.

2010 AMS Mathematics Subject Classification: Primary: 62G05; Secondary: 62G20.

Key words and phrases: Complete consistency, complete convergence rate, recursive density estimator, widely orthant dependent random variable.

1. INTRODUCTION

The random variables in many statistical applications are assumed to be independent. However, that is often not a very realistic assumption. Therefore, many statisticians extended this condition to various dependence structures. In this paper, we will consider a rather weak and applicable dependence structure, i.e., a widely orthant dependence structure, the concept of which was first introduced by Wang et al. [16] as follows.

DEFINITION 1.1. A finite set of random variables $X_1, X_2, \ldots, X_n$ is said to be *widely upper orthant dependent* (WUOD) if there exists a finite real number $g_U(n)$ such that for all finite real numbers $x_i, 1 \leqslant i \leqslant n$,

$$P(X_1 > x_1, X_2 > x_2, \ldots, X_n > x_n) \leqslant g_U(n) \prod_{i=1}^{n} P(X_i > x_i). \tag{1.1}$$

A finite set of random variables $X_1, X_2, \ldots, X_n$ is said to be *widely lower orthant dependent* (WLOD) if there exists a finite real number $g_L(n)$ such that for all finite

* Supported by the Key Project of the Outstanding Young Talent Support Program of the University of Anhui Province (gxyqZD2016367).

real numbers $x_i, 1 \leqslant i \leqslant n$,

$$(1.2) \qquad P(X_1 \leqslant x_1, X_2 \leqslant x_2, \ldots, X_n \leqslant x_n) \leqslant g_L(n) \prod_{i=1}^{n} P(X_i \leqslant x_i).$$

If the set $X_1, X_2, \ldots, X_n$ is both WUOD and WLOD, then we say that $X_1, X_2, \ldots, X_n$ are *widely orthant dependent* (WOD) random variables, and $g_U(n), g_L(n)$ are called *dominating coefficients*. A sequence of random variables $\{X_n, n \geqslant 1\}$ is said to be *WOD* if every its finite subset is WOD.

With various dominating coefficients, the WOD structure contains many other dependence structures. Wang et al. [16] presented some examples showing that WOD random variables contain negatively dependent random variables, positively dependent random variables, and some other classes of dependent random variables; moreover, they also presented some examples of WOD random variables which do not satisfy these other dependence structures.

It can be easily checked that $g_U(n) \geqslant 1$ and $g_L(n) \geqslant 1$. If both (1.1) and (1.2) hold with $g_U(n) = g_L(n) = M$ for all $n \geqslant 1$, where M is a positive constant, then the random variables are called *END*, which was introduced by Liu [9]. If $M = 1$ for all $n \geqslant 1$, then the random variables are called *NOD*, which was introduced by Lehmann [4] (cf. also Joag-Dev and Proschan [3]). As is well known, negatively associated (NA) random variables are NOD. Furthermore, Hu [2] pointed out that negatively superadditive dependent (NSD, for short) random variables are NOD. Hence, the class of WOD random variables includes independent sequences, NA sequences, NSD sequences, NOD sequences and END sequences as special cases. Thus, studying the limit behavior of WOD random variables is of general interest. There are many results investigating the WOD random variables. For example, Wang and Cheng [19] studied the basic renewal theorems for random walks with widely dependent increments; Chen et al. [1] investigated the uniform asymptotics for the finite-time ruin probabilities of two kinds of nonstandard bidimensional renewal risk models with constant interest forces and diffusion generated by Brownian motions; Shen [14] established the Bernstein-type inequality for WOD random variables with its application to nonparametric regression models; Shen [15] studied the asymptotic approximation of inverse moments for a class of nonnegative random variables including WOD random variables as a special case; Qiu and Chen [12] obtained some results on complete convergence and complete moment convergence for weighted sums of WOD random variables; Wang et al. [17] established complete convergence for arrays of rowwise WOD random variables with application to complete consistency for the estimator in a nonparametric regression model based on WOD errors; Yang et al. [22] presented the Bahadur representation of sample quantiles for WOD random variables, and so forth.

Estimating a probability density function is a fundamental problem in statistics. Let $\{X_n, n \geqslant 1\}$ be a sequence of random variables with probability density function $f(x)$. Rosenblatt [13] and Parzen [11] introduced the following classical kernel estimator of $f(x)$:

$$f_n(x) = \frac{1}{nh_n} \sum_{i=1}^{n} K\left(\frac{x - X_i}{h_n}\right).$$

Wolverton and Wagner [20] introduced the following recursive kernel estimator of $f(x)$:

$$\hat{f}_n(x) = \frac{1}{n} \sum_{i=1}^{n} \frac{1}{h_i} K\left(\frac{x - X_i}{h_i}\right), \tag{1.3}$$

where $0 < h_n \downarrow 0$ are bandwidths and K is some kernel function. Note that (1.3) can be computed recursively by

$$\hat{f}_n(x) = \frac{n-1}{n} \hat{f}_{n-1}(x) + \frac{1}{nh_n} K\left(\frac{x - X_n}{h_n}\right). \tag{1.4}$$

This recursive property is particularly useful in large sample sizes since $\hat{f}_n(x)$ can be easily updated with each additional observation. This is especially relevant in a time series context, where there has been an interest in the use of nonparametric estimates in very long financial time series. Also, under certain circumstances, the recursive estimator is more efficient than its nonrecursive estimator $f_n(x)$ when efficiency is measured in terms of the variance of an appropriate asymptotic distribution. Moreover, the estimator can be applied in estimating the hazard rate function, which is defined as $r(x) = f(x)/\big(1 - F(x)\big)$, where $f(x)$ is an unknown marginal probability density function and $F(x)$ is a distribution function. A general hazard rate estimator for $r(x)$ is

$$\hat{r}_n(x) = \frac{\hat{f}_n(x)}{1 - F_n(x)}, \tag{1.5}$$

where $F_n(x)$ is an empirical distribution of $X_1, X_2, \ldots, X_n$. Therefore, the properties of $\hat{f}_n(x)$ are extensively discussed by some authors. For example, Liang and Baek [8] discussed the point asymptotic normality for $\hat{f}_n(x)$ under NA random variables. Masry [10] obtained the quadratic mean convergence and asymptotic normality of the recursive estimator under various assumptions on the dependence of X_i; Li et al. [6] discussed the asymptotic bias, quadratic-mean convergence and established the pointwise asymptotic normality of $\hat{f}_n(x)$ for a stationary sequence of NA sequences. Li and Yang [7] studied the strong convergence rate of recursive probability density estimator $\hat{f}_n(x)$ based on NA samples. Li [5] extended the results of Li and Yang [7] from NA samples to END samples.

In this paper, we will consider the complete convergence rate of recursive probability density estimator (1.3) under strictly stationary WOD random variables. The results obtained in the paper improve and extend the corresponding ones of Li [5] for END samples and of Li and Yang [7] for NA samples. We will also study the complete consistency for the estimator (1.3) under some mild conditions.

The paper is organized as follows. The main results are presented in Section 2. Some lemmas are provided in Section 3. The proofs are given in Section 4. Throughout the paper, $C, c_0, c_1, \ldots$ denote some positive constants whose value may be different in different places; $a = O(b)$ implies that $a \leqslant Cb$; $C(f)$ denotes all the continuity points of a function f; and $C^2(f)$ stands for a point set for which the second-order derivative $f^{''}$ exists and is bounded and continuous.

2. MAIN RESULTS

In this section, we will present the strong convergence rate for the recursive kernel estimator $\hat{f}_n(x)$. We adopt the following assumptions which were also used in Li and Yang [7] and Li [5]:

(A$_1$) $\int_{-\infty}^{\infty} K(u)du = 1$, $\int_{-\infty}^{\infty} uK(u)du = 0$, $\int_{-\infty}^{\infty} u^2K(u)du < \infty$, $K(\cdot) \in L_1$.

(A$_2$) The bandwidths h_n are such that $0 < h_n \downarrow 0$ and $nh_n \to \infty$ as $n \to \infty$.

Now we state our main results as follows.

THEOREM 2.1. *Suppose that* (A$_1$) *and* (A$_2$) *hold. Let* $\{X_n, n \geqslant 1\}$ *be a sequence of strictly stationary WOD random variables with* $g(n) = O(n^\delta)$ *for some* $\delta \geqslant 0$. *Suppose that the kernel* $K(\cdot)$ *is a bounded monotone density function and the bandwidth* $h_n = O(n^{-1/5}\log^{1/5} n)$. *Then for any* $x \in C^2(f)$,

$$|\hat{f}_n(x) - f(x)| = O\big([\log n/(nh_n)]^{1/2}\big), \text{ completely.} \tag{2.1}$$

REMARK 2.1. Li and Yang [7] and Li [5] obtained similar results under NA and END samples, respectively. The convergence rate obtained in their result is $o\left([\log n(\log\log n)^l/(nb_n)]^{1/2}\right)$ for some $l > 0$ under the meaning of almost surely (a.s.). Noting that WOD contains END and NA, complete convergence is stronger than a.s. convergence (by the Borel–Cantelli lemma), and the rate in our result is slightly faster, thus our result improves and extends the corresponding ones of Li [5] as well as Li and Yang [7].

Relaxing the restriction on the dominating coefficients $g(n)$, we have the following more general result.

THEOREM 2.2. *Suppose that* (A$_1$) *and* (A$_2$) *hold. Let*

$$\gamma_n = \left[\log\left(ng(n)\right)/(nh_n)\right]^{1/2} \to 0.$$

Let $\{X_n, n \geqslant 1\}$ *be a sequence of strictly stationary WOD random variables. Suppose that the kernel* $K(\cdot)$ *is a bounded monotone density function and the bandwidth* h_n *is such that* $h_n = O(n^{-1/5}\log^{1/5} n)$. *Then for any* $x \in C^2(f)$,

$$|\hat{f}_n(x) - f(x)| = O(\gamma_n), \text{ completely.} \tag{2.2}$$

REMARK 2.2. In Theorem 2.1, we required that the dominating coefficients $g(n)$ are polynomially increasing, which is always assumed in many papers. However, Theorem 2.2 allows the dominating coefficients $g(n)$ to be geometrically increasing. If $g(n) = O(n^{\delta})$ for some $\delta \geqslant 0$, the strong convergence rate is the same as that in Theorem 2.1. Consequently, Theorem 2.2 is much more general and applicable.

Furthermore, by relaxing the restriction on the bandwidth h_n, we have the following result.

THEOREM 2.3. *Suppose that* (A_1) *and* (A_2) *hold. Let* $\{X_n, n \geqslant 1\}$ *be a sequence of strictly stationary WOD random variables. Suppose that the kernel* $K(\cdot)$ *is a bounded monotone density function and* $\log\big(ng(n)\big)/(nh_n) \to 0$*. Then for any* $x \in C^2(f)$,

$$\hat{f}_n(x) - f(x) \to 0, \ \textit{completely}. \tag{2.3}$$

As an application of the results above, we obtain the complete consistency and the rate of the complete consistency for the hazard rate estimator $\hat{r}_n(x)$ as follows.

THEOREM 2.4. *Suppose that* (A_1) *and* (A_2) *hold. Let* $\{X_n, n \geqslant 1\}$ *be a sequence of strictly stationary WOD random variables with* $g(n) = O(n^{\delta})$ *for some* $\delta \geqslant 0$*. Suppose that the kernel* $K(\cdot)$ *is a bounded monotone density function and the bandwidth* $h_n = O(n^{-1/5} \log^{1/5} n)$*. If there exists a point* x_0 *such that* $F(x_0) < 1$*, then for any* $x \in C^2(f)$ *and* $x \leqslant x_0$,

$$|\hat{r}_n(x) - r(x)| = O\big([\log n/(nh_n)]^{1/2}\big), \ \textit{completely}. \tag{2.4}$$

THEOREM 2.5. *Suppose that* (A_1) *and* (A_2) *hold. Let*

$$\gamma_n = \big[\log\big(ng(n)\big)/(nh_n)\big]^{1/2} \to 0.$$

Let $\{X_n, n \geqslant 1\}$ *be a sequence of strictly stationary WOD random variables. Suppose that the kernel* $K(\cdot)$ *is a bounded monotone density function and the bandwidth* h_n *is such that* $h_n = O(n^{-1/5} \log^{1/5} n)$*. If there exists a point* x_0 *such that* $F(x_0) < 1$*, then for any* $x \in C^2(f)$ *and* $x \leqslant x_0$,

$$|\hat{r}_n(x) - r(x)| = O(\gamma_n), \ \textit{completely}. \tag{2.5}$$

THEOREM 2.6. *Suppose that* (A_1) *and* (A_2) *hold. Let* $\{X_n, n \geqslant 1\}$ *be a sequence of strictly stationary WOD random variables. Suppose that the kernel* $K(\cdot)$ *is a bounded monotone density function and* $\log\big(ng(n)\big)/(nh_n) \to 0$*. If there exists a point* x_0 *such that* $F(x_0) < 1$*, then for any* $x \in C^2(f)$ *and* $x \leqslant x_0$,

$$\hat{r}_n(x) - r(x) \to 0, \ \textit{completely}. \tag{2.6}$$

3. SOME LEMMAS

In this section, we will present some lemmas which will be used in proving our main results.

LEMMA 3.1 (Wang et al. [18]). *Let* $\{X_n, n \geqslant 1\}$ *be a sequence of WOD random variables.*

(i) *If* $\{f_n(\cdot), n \geqslant 1\}$ *are all nondecreasing* (*or nonincreasing*), *then* $\{f_n(X_n), n \geqslant 1\}$ *are still WOD.*

(ii) *For each* $n \geqslant 1$ *and any* $t \in \mathbb{R}$,

$$E \exp\big\{t \sum_{i=1}^{n} X_i\big\} \leqslant g(n) \prod_{i=1}^{n} E \exp\{tX_i\}.$$

LEMMA 3.2. *Let* $\{X_n, n \geqslant 1\}$ *be a sequence of WOD random variables with* $EX_n = 0$ *and* $\max_{1 \leqslant i \leqslant n} |X_i| \leqslant b_n$ *a.s. for each* $n \geqslant 1$, *where* $\{b_n, n \geqslant 1\}$ *is a sequence of positive numbers. Suppose that there exists some* $t > 0$ *such that* $tb_n \leqslant 1$. *Then for any* $\varepsilon > 0$,

$$P\big(\big|\sum_{i=1}^{n} X_i\big| \geqslant \varepsilon\big) \leqslant 2g(n) \exp\big\{-t\varepsilon + t^2 \sum_{i=1}^{n} EX_i^2\big\}.$$

Proof. Noting that $|tX_i| \leqslant 1$ a.s. and $EX_i = 0$ for each $i \geqslant 1$, we have

$$\begin{aligned} E\exp\{tX_i\} &= 1 + \sum_{k=2}^{\infty} \frac{E(tX_i)^k}{k!} \leqslant 1 + t^2 EX_i^2 \sum_{k=2}^{\infty} \frac{1}{k!} \\ &\leqslant 1 + t^2 EX_i^2 \leqslant \exp\{t^2 EX_i^2\}. \end{aligned} \tag{3.1}$$

By Markov's inequality, Lemma 3.1 (ii) and (3.1), we can see that

$$\begin{aligned} P\big(\sum_{i=1}^{n} X_i \geqslant \varepsilon\big) &\leqslant e^{-t\varepsilon} E\exp\big\{t\sum_{i=1}^{n} X_i\big\} \leqslant g(n) e^{-t\varepsilon} \prod_{i=1}^{n} E\exp\{tX_i\} \\ &\leqslant g(n) \exp\big\{-t\varepsilon + t^2 \sum_{i=1}^{n} EX_i^2\big\}. \end{aligned} \tag{3.2}$$

The desired result follows by replacing X_i by $-X_i$ in (3.2). This completes the proof of the lemma. ∎

LEMMA 3.3 (Li and Yang [7]). *Suppose that* (A_1) *holds*; *then for all* $x \in C^2(f)$,

$$\lim_{h \to 0} \int_{\mathbb{R}} K(u) f(x - hu) du = f(x).$$

LEMMA 3.4 (Li and Yang [7]). *Suppose that* (A_1) *holds*; *then for all* $x \in C^2(f)$,

$$\Big(\frac{1}{n}\sum_{i=1}^{n} h_i^2\Big)^{-1} |E\hat{f}_n(x) - f(x)| \leqslant C < \infty.$$

LEMMA 3.5. *Let* $\{X_n, n \geqslant 1\}$ *be a sequence of WOD random variables with unknown distribution function* $F(x)$ *and bounded probability density function* $f(x)$. *Let* $F_n(x)$ *be an empirical distribution function. If*

$$\mu_n =: \big[\log\big(ng(n)\big)/n\big]^{1/2} \to 0,$$

then

$$\sup_x |F_n(x) - F(x)| = O(\mu_n), \text{ completely.}$$

In particular, *if* $g(n) = O(n^\delta)$ *for some* $\delta \geqslant 0$, *then*

$$\sup_x |F_n(x) - F(x)| = O\big((\log n/n)^{1/2}\big), \text{ completely.}$$

Proof. Let $F(x_{ni}) = i/n$ for $n \geqslant 3$ and $1 \leqslant i \leqslant n-1$. By Lemma 2 in Yang [21] we have

$$\sup_{-\infty<x<\infty} |F_n(x) - F(x)| \leqslant \max_{1\leqslant j\leqslant n-1} |F_n(x_{nj}) - F(x_{nj})| + 2/n. \tag{3.3}$$

Noting that $n\mu_n \to \infty$, for any positive constant D_1, we have $2/n < D_1\mu_n/2$ for all n large enough. Then it follows from (3.3) that

$$\begin{aligned} P\big(\sup_{-\infty<x<\infty} |F_n(x) - F(x)| > D_1\mu_n\big) &\leqslant P\big(\max_{1\leqslant j\leqslant n-1} |F_n(x_{nj}) - F(x_{nj})| > D_1\mu_n/2\big) \\ &\leqslant \sum_{j=1}^{n-1} P\big(|F_n(x_{nj}) - F(x_{nj})| > D_1\mu_n/2\big). \end{aligned} \tag{3.4}$$

Let $\xi_i = I(X_i < x_{nj}) - EI(X_i < x_{nj})$. By Lemma 3.1, $\{\xi_i, i \geqslant 1\}$ is still a sequence of WOD random variables with $E\xi_i = 0$, $|\xi_i| \leqslant 2$ and $E\xi_i^2 \leqslant 1$. Thus, choosing $t = D_1\mu_n/4$ in Lemma 3.2, we see that, for all n large enough,

$$\begin{aligned} &P\big(|F_n(x_{nj}) - F(x_{nj})| > D_1\mu_n/2\big) = P\big(\big|\sum_{i=1}^{n}\xi_i\big| > D_1 n\mu_n/2\big) \\ &\leqslant 2g(n)\exp\Big\{-D_1 n\mu_n t/2 + t^2\sum_{i=1}^{n} E\xi_i^2\Big\} \leqslant 2g(n)\exp\{-D_1 n\mu_n t/2 + nt^2\} \\ &\leqslant 2g(n)\exp\{-D_1^2 n\mu_n^2/16\} \leqslant 2g(n)\exp\big\{-D_1^2 c_0\log\big(ng(n)\big)/16\big\} \\ &\leqslant 2g(n)\big(ng(n)\big)^{-D_1^2 c_0/16}. \end{aligned} \tag{3.5}$$

Recall that $g(n) \geqslant 1$. Taking D_1 sufficiently large such that $D_1^2 c_0/16 > 3$, by (3.4) and (3.5) we have

$$\sum_{n=1}^{\infty} P\big(\sup_{-\infty<x<\infty} |F_n(x) - F(x)| > D_1\mu_n\big) \leqslant C \sum_{n=1}^{\infty}\sum_{j=1}^{n-1} g(n)\big(ng(n)\big)^{-D_1^2 c_0/16} < \infty.$$

This completes the proof of the lemma. ■

4. PROOF OF THE MAIN RESULTS

Proof of Theorem 2.1. Set

$$\eta_i = h_i^{-1}\left[K\left(\frac{x - X_i}{h_i}\right) - EK\left(\frac{x - X_i}{h_i}\right)\right] \quad \text{for } 1 \leqslant i \leqslant n.$$

Since $K(\cdot)$ is bounded and monotone, $\{\eta_i, i \geqslant 1\}$ is still a sequence of WOD random variables. Moreover, it follows from $0 < h_n \downarrow 0$ that there exists some positive constant c_1 such that $\max_{1\leqslant i\leqslant n} |\eta_i| \leqslant c_1/h_n$. By Lemma 3.2 we have

$$\sum_{i=1}^{n} E\eta_i^2 \leqslant \sum_{i=1}^{n} h_i^{-2} EK^2\left(\frac{x - X_i}{h_i}\right) = \sum_{i=1}^{n} h_i^{-2} \int_{\mathbb{R}} K^2\left(\frac{x-u}{h_i}\right) f(u)du$$
$$= \sum_{i=1}^{n} h_i^{-1} \int_{\mathbb{R}} K^2(u) f(x - h_i u)du \leqslant c_2 n h_n^{-1}.$$

Set $\lambda_n = [\log n/(nh_n)]^{1/2}$. Applying Lemma 3.2 with $t = D_2 h_n \lambda_n/(2c_2)$, where D_2 is some positive constant which will be specified later. It is easy to check that $t \cdot c_1/h_n \leqslant 1$ for all n large enough. Then we get

$$P\big(|\hat f_n(x) - E\hat f_n(x)| > D_2\lambda_n\big) = P\big(\big|\sum_{i=1}^{n}\eta_i\big| > D_2 n\lambda_n\big)$$
$$\leqslant 2g(n)\exp\big\{-D_2 n\lambda_n t + t^2\sum_{i=1}^{n} E\eta_i^2\big\} \leqslant 2g(n)\exp\{-D_2 n\lambda_n t + c_2 n h_n^{-1} t^2\}$$
$$\leqslant 2g(n)\exp\{-D_2^2 n h_n \lambda_n^2/(4c_2)\} \leqslant 2g(n)\exp\{-\log n \cdot D_2^2/(4c_2)\}$$
$$\leqslant Cn^{\delta - D_2^2/(4c_2)}.$$

Taking D_2 large enough such that $\delta - D_2^2/(4c_2) < -2$, we have

$$\sum_{n=1}^{\infty} P\big(|\hat f_n(x) - E\hat f_n(x)| > D_2\lambda_n\big) < \infty,$$

that is,

$$|\hat{f}_n(x) - E\hat{f}_n(x)| = O\big([\log n/(nh_n)]^{1/2}\big), \text{ completely.} \tag{4.1}$$

On the other hand, using $h_n = O(n^{-1/5}\log^{1/5} n)$, we have by Lemma 3.4

$$\begin{aligned}&[\log n/(nh_n)]^{-1/2}|E\hat{f}_n(x) - f(x)| \leqslant C[\log n/(nh_n)]^{-1/2}\frac{1}{n}\sum_{i=1}^{n} h_i^2\\ &\leqslant C\big(h_n/(n\log n)\big)^{1/2}\sum_{i=1}^{n} h_i^2 \leqslant Cn^{-3/5}(\log^{-2/5} n)\sum_{i=1}^{n} i^{-2/5}\log^{2/5} i \leqslant C,\end{aligned}$$

which implies that

$$|E\hat{f}_n(x) - f(x)| = O\big([\log n/(nh_n)]^{1/2}\big). \tag{4.2}$$

Note that

$$|\hat{f}_n(x) - f(x)| \leqslant |\hat{f}_n(x) - E\hat{f}_n(x)| + |E\hat{f}_n(x) - f(x)|. \tag{4.3}$$

Therefore, the desired result (2.1) follows immediately by (4.1)–(4.3). The proof is completed. ■

P r o o f o f T h e o r e m 2.2. In view of the proof of Theorem 2.1, we only need to show that

$$|\hat{f}_n(x) - E\hat{f}_n(x)| = O\Big(\big[\log\big(ng(n)\big)/(nh_n)\big]^{1/2}\Big), \text{ completely,} \tag{4.4}$$

and

$$|E\hat{f}_n(x) - f(x)| = O\Big(\big[\log\big(ng(n)\big)/(nh_n)\big]^{1/2}\Big). \tag{4.5}$$

Noting that $g(n) \geqslant 1$, we obtain (4.5) from (4.2) immediately and thus we only need to prove (4.4). As in the proof of (4.1), set $\gamma_n = \big[\log\big(ng(n)\big)/(nh_n)\big]^{1/2}$. Let us apply Lemma 3.2 with $t = D_3 h_n\gamma_n/(2c_2)$ to see that for all n large enough,

$$\begin{aligned}&P\big(|\hat{f}_n(x) - E\hat{f}_n(x)| > D_3\gamma_n\big) = P\big(\big|\sum_{i=1}^{n}\eta_i\big| > D_3 n\gamma_n\big)\\ &\leqslant 2g(n)\exp\Big\{-D_3 n\gamma_n t + t^2\sum_{i=1}^{n} E\eta_i^2\Big\} \leqslant 2g(n)\exp\{-D_3 n\gamma_n t + c_2 nh_n^{-1}t^2\}\\ &\leqslant 2g(n)\exp\{-D_3^2 nh_n\gamma_n^2/(4c_2)\} \leqslant 2g(n)\exp\big\{-D_3^2/(4c_2)\cdot\log\big(ng(n)\big)\big\}\\ &\leqslant 2g(n)\big(ng(n)\big)^{-D_3^2/(4c_2)}.\end{aligned}$$

Taking D_3 sufficiently large such that $D_3^2/(4c_2) > 2$, we have

$$\sum_{n=1}^{\infty} P\big(|\hat{f}_n(x) - E\hat{f}_n(x)| > D_3\gamma_n\big) < \infty,$$

which is equivalent to (4.4). The proof is completed. ∎

Proof of Theorem 2.3. In view of the proof of Theorem 2.2, by (4.4) and $\log\big(ng(n)\big)/(nh_n) \to 0$ we have

$$\hat{f}_n(x) - E\hat{f}_n(x) \to 0, \text{ completely.}$$

Therefore, we only need to show that

$$|E\hat{f}_n(x) - f(x)| \to 0 \tag{4.6}$$

without the condition $h_n = O(n^{-1/5}\log^{1/5} n)$ in Theorem 2.2. Actually, by Lemma 3.4 and Stolz's theorem we have

$$\lim_{n\to\infty} |E\hat{f}_n(x) - f(x)| \leqslant C \lim_{n\to\infty} \frac{1}{n}\sum_{i=1}^{n} h_i^2 = C \lim_{n\to\infty} h_n^2 = 0.$$

Consequently, (4.6) is proved and thus the proof of the theorem is completed. ∎

Since the proofs of Theorems 2.4–2.6 are similar, we present only the proof of Theorem 2.4 as follows.

Proof of Theorem 2.4. Set $\bar{F}_n(x) = 1 - F_n(x)$ and $\bar{F}(x) = 1 - F(x)$. It follows from (1.5) that

$$|\hat{r}_n(x) - r(x)| \leqslant \frac{\bar{F}(x)|\hat{f}_n(x) - f(x)| + |F_n(x) - F(x)|f(x)}{\bar{F}_n(x)\bar{F}(x)}. \tag{4.7}$$

From $0 \leqslant F(x) \leqslant F(x_0) < 1$ for all $x \leqslant x_0$, $\sup_x f(x) \leqslant C < \infty$, applying Theorem 2.1 and taking $\mu_n = (\log n/n)^{1/2}$ in Lemma 3.5, we can see that

$$|\hat{f}_n(x) - f(x)| = O\big([\log n/(nh_n)]^{1/2}\big), \text{ completely,} \tag{4.8}$$

and

$$\sup_{x\leqslant x_0} |F_n(x) - F(x)| = O\big((\log n/n)^{1/2}\big), \text{ completely.} \tag{4.9}$$

On the other hand, we infer from (4.9) that for $x \leqslant x_0$ and all n large enough,

$$\bar{F}_n(x) \geqslant \bar{F}(x)/2 \geqslant \bar{F}(x_0)/2 > 0. \tag{4.10}$$

Consequently, the desired result (2.4) follows from (4.7)–(4.10). The proof is completed. ∎

REFERENCES

[1] Y. Chen, L. Wang, and Y. Wang, *Uniform asymptotics for the finite-time ruin probabilities of two kinds of nonstandard bidimensional risk models*, J. Math. Anal. Appl. 401 (2013), pp. 114–129.

[2] T. Hu, *Negatively superadditive dependence of random variables with applications*, Chinese J. Appl. Probab. Statist. 16 (2000), pp. 133–144.

[3] K. Joag-Dev and F. Proschan, *Negative association of random variables with applications*, Ann. Statist. 11 (1983), pp. 286–295.

[4] E. Lehmann, *Some concepts of dependence*, Ann. Math. Statist. 37 (1966), pp. 1137–1153.

[5] Y. Li, *On the rate of strong convergence for a recursive probability density estimator of END samples and its applications*, J. Math. Inequal. 11 (2) (2017), pp. 335–343.

[6] Y. Li, C. Wei, and S. Yang, *The recursive kernel distribution function estimator based on negatively and positively associated sequences*, Comm. Statist. Theory Methods 39 (20) (2010), pp. 3585–3595.

[7] Y. Li and S. Yang, *Strong convergence rate of recursive probability density estimators for NA sequences*, Chinese J. Engrg. Math. 22 (4) (2005), pp. 659–665.

[8] H. Liang and J. Baek, *Asymptotic normality of recursive density estimates under some dependence assumptions*, Metrika 60 (2004), pp. 155–166.

[9] L. Liu, *Precise large deviations for dependent random variables with heavy tails*, Statist. Probab. Lett. 79 (2009), pp. 1290–1298.

[10] E. Masry, *Recursive probability density estimation for weakly dependent stationary processes*, IEEE Trans. Inform. Theory 32 (2) (1986), pp. 254–267.

[11] E. Parzen, *On estimation of a probability density function and mode*, Ann. Math. Statist. 33 (1962), pp. 1065–1076.

[12] D. Qiu and P. Chen, *Complete and complete moment convergence for weighted sums of widely orthant dependent random variables*, Acta Math. Sin. (Engl. Ser.) 30 (2014), pp. 1539–1548.

[13] M. Rosenblatt, *A central limit theorem and a strong mixing condition*, Proc. Natl. Acad. Sci. USA 42 (1956), pp. 43–47.

[14] A. Shen, *Bernstein-type inequality for widely dependent sequence and its application to nonparametric regression models*, Abstr. Appl. Anal. (2013), Article ID 862602.

[15] A. Shen, *On asymptotic approximation of inverse moments for a class of nonnegative random variables*, Statistics 48 (2014), pp. 1371–1379.

[16] K. Wang, Y. Wang, and Q. Gao, *Uniform asymptotics for the finite-time ruin probability of a new dependent risk model with a constant interest rate*, Methodol. Comput. Appl. Probab. 15 (2013), pp. 109–124.

[17] X. Wang, Y. Wu, and A. Rosalsky, *Complete convergence for arrays of rowwise widely orthant dependent random variables and its applications*, Stochastics 89 (8) (2017), pp. 1228–1252.

[18] X. Wang, C. Xu, T.-C. Hu, A. Volodin, and S. Hu, *On complete convergence for widely orthant-dependent random variables and its applications in nonparametric regression models*, TEST 23 (2014), pp. 607–629.

[19] Y. Wang and D. Cheng, *Basic renewal theorems for random walks with widely dependent increments*, J. Math. Anal. Appl. 384 (2011), pp. 597–606.

[20] C. Wolverton and T. Wagner, *Asymptotically optimal discriminant functions for pattern classification*, IEEE Trans. Inform. Theory 15 (1969), pp. 258–265.

[21] S. Yang, *Consistency of nearest neighbor estimator of density function for negative associated samples*, Acta Math. Appl. Sin. 26 (3) (2003), pp. 385–395.

[22] W. Yang, T. Liu, X. Wang, and S. Hu, *On the Bahadur representation of sample quantiles for widely orthant dependent sequences*, Filomat 28 (2014), pp. 1333–1343.

Chenlu Zhuansun
School of Mathematical Sciences
Anhui University, P.R. China
E-mail: zhuansuncl@163.com

Xiaoxin Li
Department of Mathematics
and Computer Sciences
Chizhou University, P.R. China
E-mail: lxx@czu.edu.cn

Received on 5.5.2017;
revised version on 4.10.2017