Li et al. *Journal of Inequalities and Applications* (2018) 2018:71
https://doi.org/10.1186/s13660-018-1659-1

Journal of Inequalities and Applications
a SpringerOpen Journal

RESEARCH Open Access

On the convergence rates of kernel estimator and hazard estimator for widely dependent samples

Yongming Li[1,2*], Yong Zhou[1] and Chao Liu[2]

*Correspondence: lym1019@163.com
[1]School of Statistics and Management, Shanghai University of Finance and Economics, Shanghai, China
[2]School of Mathematics and Computer Science, Shangrao Normal University, Shangrao, China

Abstract

In this paper, we establish a Bernstein-type inequality for widely orthant dependent random variables, and obtain the rates of strong convergence for kernel estimators of density and hazard functions, under some suitable conditions.

MSC: 62G05; 62G20

Keywords: Widely orthant dependent; Kernel density estimator; Hazard rate; Strong convergence

1 Introduction

Let $\{X_n, n \geq 1\}$ be a sequence of random variables with an unknown marginal probability density function $f(x)$ and distribution function $F(x)$. Assume that $K(x)$ is a known kernel function, the kernel estimate of $f(x)$ and the empirical distribution function of $F(x)$ are given by

$$\hat{f}_n(x) = \frac{1}{nh_n}\sum_{j=1}^{n} K\left(\frac{x - X_j}{h_n}\right), \qquad F_n(x) = n^{-1}\sum_{j=1}^{n} I(X_j < x), \tag{1.1}$$

where $\{h_n, n \geq 1\}$ is a sequence of positive bandwidths tending to zero as $n \to \infty$, and $I(\cdot)$ is the indicator of the event specified in the parentheses. Denote the hazard rate of distribution $F(x)$ by $\lambda(t) = f(t)/(1 - F(t))$, and it can be estimated by

$$\hat{\lambda}_n(x) = \frac{\hat{f}_n(x)}{1 - F_n(x)}. \tag{1.2}$$

Works devoted to the estimation of probability density and hazard rate functions include the following. Izenman and Tran ([1], 1990) discussed the uniform consistency and sharp rates of convergence under strong mixing and absolute regularity conditions. Cai ([2], 1998) established the asymptotic normality and the uniform consistency with rates of the kernel estimators for density and hazard functions under a censored dependent model. Liebscher ([3], 2002) derived the rates of uniform strong convergence for density and hazard rate estimators for right censoring based on a stationary strong mixing sequence. Liang et al. ([4], 2005) obtained the optimal convergence rates of the nonlinear wavelet estimators of the hazard rate function when the survival times form a stationary strong mixing sequence. Bouezmarni et al. ([5], 2011) proposed new estimators based on the gamma kernels for density and hazard rate functions which are free of bias, and achieved the optimal rate of convergence in terms of integrated mean squared error, and so on.

On the other hand, different from strong mixing and negatively associated random variables, widely orthant dependence random variables (defined below) were introduced by Wang and Cheng ([6], 2011). Chen et al. ([7], 2016) proved a new type of Nagaev's inequality and a refined inequality of widely dependent random variables, and as applications, investigated elementary renewal theorems and weighted elementary renewal theorem. Now, let us recall the following definition of widely orthant dependence.

Definition 1.1 For random sequence $\{X_n, n \geq 1\}$, if there exists a finite real sequence $\{g_U(n), n \geq 1\}$ satisfying, for each $n \geq 1$ and for all $x_i \in (-\infty, \infty)$, $1 \leq i \leq n$,

$$P(X_1 > x_1, X_2 > x_2, \ldots, X_n > x_n) \leq g_U(n) \prod_{i=1}^{n} P(X_i > x_i), \tag{1.3}$$

and there also exists a finite real sequence $\{g_L(n), n \geq 1\}$ satisfying, for each $n \geq 1$ and for all $x_i \in (-\infty, \infty)$, $1 \leq i \leq n$,

$$P(X_1 \leq x_1, X_2 \leq x_2, \ldots, X_n \leq x_n) \leq g_L(n) \prod_{i=1}^{n} P(X_i \leq x_i), \tag{1.4}$$

then a random sequence $\{X_n, n \geq 1\}$ is called widely orthant dependent (*WOD*) with dominating coefficients $g(n) = \max\{g_U(n), g_L(n)\}$.

Now, we will give two real examples of WOD sequences. The first example of WOD that satisfies the conditions of the main results is given as follows (see Example 1.1) by the framework of Farlie–Gumbel–Morgenstern (FGM) dependence (see Cambanis [8], 1991). The second example (see Example 1.2) does not satisfy the conditions of the paper, but it is useful as an example in the article.

Example 1.1 A sequence $\{X_n, n \geq 1\}$ of random variables on (Ω, B, P) is called FGM if for any $n \in N$ and $(x_1, \ldots, x_n) \in R_n$,

$$P\{X_1 \leq x_1, \ldots, X_n \leq x_n\} = \prod_{i=1}^{n} F_i(x_i)\left(1 + \sum_{1 \leq j < k \leq n} a(j,k)\overline{F}_j(x_j)\overline{F}_k(x_k)\right), \tag{1.5}$$

the constants $a(\cdot,\cdot)$ are admissible if the 2^n inequalities $1 + \sum_{1 \leq j<k \leq n} a(j,k)\varepsilon_j\varepsilon_k \geq 0$ for all $\varepsilon_j = -M_j$ or $1 - m_j$ hold, where M_j and m_j are the supremum and the infimum of the set $\{\{F_i(x), x \in R\} \setminus \{0,1\}\}$. If for some integer i the marginal $F_i(\cdot)$ is absolutely continuous, then $M_j = 1$ and $m_j = 0$, hence $\varepsilon_i = \pm 1$. Next, by Hashorva and Hüsler ([9], 1999), we have $\sum_{1 \leq j<k \leq n} a(j,k) = O(n)$, $n \in N$. Hence, from (1.4) and (1.5), we can take $g_L(n) = O(1 + \sum_{1 \leq j<k \leq n} a(j,k)) = O(n)$, then $P\{X_1 \leq x_1, \ldots, X_n \leq x_n\} \leq g_L(n) \prod_{i=1}^{n} F_i(x_i)$. This implies that the FGM sequence is a WOD sequence, and the conditions of the main results and lemmas are satisfied.

Example 1.2 Assume that the random vectors (ξ_n, η_n), $n = 1, 2, \dots$, are independent and for each integer $n \geq 1$, the random variables ξ_n and η_n are dependent according to the Farlie–Gumbel–Morgenstern copula with the parameter $a_n \in [-1, 1]$. Suppose that the distributions of ξ_n and η_n, $n = 1, 2, \dots$, are absolutely continuous, denoted by F_{ξ_n} and F_{η_n}, $n = 1, 2, \dots$, respectively. By Sklar's theorem (see Chap. 2 of Nelsen RB ([10], 2006)), for each integer $n \geq 1$ and any $x_n, y_n \in (-\infty, +\infty)$, we can construct the cumulative distribution function of (ξ_n, η_n) as follows:

$$P(\xi_n \leq x_n, \eta_n \leq y_n) = F_{\xi_n}(x_n)F_{\eta_n}(y_n)\bigl[1 + a_n\overline{F_{\xi_n}}(x_n)\overline{F_{\eta_n}}(y_n)\bigr]$$

and

$$P(\xi_n > x_n, \eta_n > y_n) = \overline{F_{\xi_n}}(x_n)\overline{F_{\eta_n}}(y_n)\bigl[1 + a_nF_{\xi_n}(x_n)F_{\eta_n}(y_n)\bigr].$$

Therefore, for each $n \geq 1$, we have

$$\begin{aligned} &P(\xi_n \leq x_n, \eta_n \leq y_n) \leq 2P(\xi_n \leq x_n)P(\eta_n \leq y_n), \\ &P(\xi_n > x_n, \eta_n > y_n) \leq 2P(\xi_n > x_n)P(\eta_n > y_n). \end{aligned}$$

Then $\{(\xi_n, \eta_n)\}$ is a sequence of independent bivariate WOD random variables. Thus, for all $x_j, y_j \in R$,

$$P\Biggl[\bigcap_{j=1}^{n}(\xi_j \leq x_j, \eta_j \leq y_j)\Biggr] \leq 2^n\prod_{j=1}^{n}P(\xi_j \leq x_j)P(\eta_j \leq y_j)$$

and

$$P\Biggl[\bigcap_{j=1}^{n}(\xi_j > x_j, \eta_j > y_j)\Biggr] \leq 2^n\prod_{j=1}^{n}P(\xi_j > x_j)P(\eta_j > y_j).$$

From this, and by Definition 1.1, it is easy to see that the sequence $\{\xi_1, \eta_1, \dots, \xi_n, \eta_n, \dots\}$ is WOD with $g_L(n) = g_U(n) = 2^n$, but the condition of Lemma 2.5 is not satisfied.

We can further refer to some large sample properties of nonparameter estimate based on WOD samples. For instance, Wang et al. ([11], 2013) studied the strong consistency of estimator of fixed design regression model for WOD samples. Shi and Wu ([12], 2014) discussed the strong consistency of kernel density estimator for identically distributed WOD samples. Li et al. ([13], 2015) studied the pointwise strong consistency for a kind of recursive kernel estimator based WOD samples.

In this paper, we attempt to establish a Bernstein-type inequality and to derive the rates of strong convergence for the estimators of density and hazard rate functions for WOD samples. Throughout the paper, all limits are taken as n tends to ∞, and $c, C, C_1, C_2, \dots$ denote positive constants whose values may change from one place to another, unless specified otherwise.

2 Assumptions and some auxiliary results

For the sake of simplicity, some assumptions on kernel function $K(\cdot)$ and density function $f(x)$ are listed below.

(A1) $K(u) \in L_1$, $\int_{-\infty}^{+\infty} K(u)\,du = 1$, $\sup_{x\in R}(1+|x|)|K(x)| \le c < \infty$.

(A2) $\int_{-\infty}^{+\infty} u^r K(u)\,du = 0$, $r = 1, 2, \ldots, s-1$, $\int_{-\infty}^{+\infty} u^s K(u)\,du = M \neq 0$, where M is a positive constant, $s \ge 2$ is some positive integer.

(A3) $f(x) \in C_{2,\alpha}$, where α is a positive constant, $C_{2,\alpha}$ implies that $f(x)$ is 2nd differentiable, $f''(x)$ is a continuous function, and $|f''(x)| \le \alpha$.

The following proposition for WOD random sequence comes from Corollary 3 in Shen ([3], 2013), which will be used in the following.

Lemma 2.1

(i) *Let a random sequence $\{X_n, n \ge 1\}$ be WOD with dominating coefficients $g(n)$. If $\{G_n(\cdot), n \ge 1\}$ is a nondecreasing (or nonincreasing) function sequence, then the random sequence $\{G_n(X_n), n \ge 1\}$ is still WOD with the same dominating coefficients $g(n)$.*

(ii) *Let a random sequence $\{X_n, n \ge 1\}$ be WOD, then for each $n \ge 1$ and any $s > 0$,*

$$E\exp\left\{s\sum_{i=1}^{n} X_i\right\} \le g(n)\prod_{i=1}^{n} E\exp\{sX_i\}.$$

Remark 2.1 Condition (A1) is a reasonable condition, we can refer to condition (2) in Cai ([14], 1993) and condition (II) in Theorem S of Lin ([15], 1987). And by (A1), we can get the following lemma.

Lemma 2.2 (see Lemma 3, [14], 1993, or [15], 1987) *Let $K(x)$ be satisfied* (A1), *and $f(\cdot) \in L_1$, then*

(i) *for the continuous point of $f(x)$,*

$$\lim_{h_n\to 0} h_n^{-1}\int_R K\left(\frac{x-u}{h_n}\right) f(u)\,du = f(x),$$
$$\lim_{h_n\to 0} h_n^{-1}\int_R K^2\left(\frac{x-u}{h_n}\right) f(u)\,du = f(x)\int_R K^2(u)\,du;$$

(ii) *for all $x, y \in R$, $x \neq y$,*

$$\lim_{h_n\to 0} h_n^{-1}\int_R K\left(\frac{x-u}{h_n}\right) K\left(\frac{y-u}{h_n}\right) du = 0.$$

Lemma 2.3 *By Lemma* 3.4 *of Li* ([16], 2017), *we obtain that*

$$h_n^{-2}\left|E\hat{f}_n(x) - f(x)\right| \le h_n^{-2}\left(\frac{h_n^2}{2}\left|\int_R K(u) f''(x - \xi h_n u) u^2\,du\right|\right) \le C.$$

Now, we will establish a Bernstein-type inequality for a WOD random sequence as follows.

Lemma 2.4 *Let a random sequence $\{X_n, n \geq 1\}$ be WOD with dominating coefficients $g(n)$, and $EX_i = 0$, $|X_i| \leq d_i$ for $i = 1, \ldots, n$, where $\{d_i, 1 \leq i \leq n\}$ is a sequence of positive constants. For $t > 0$, if $t \cdot \max_{1\leq i\leq n} d_i \leq 1$, then for any $\varepsilon > 0$,*

$$P\left(\left|\sum_{i=1}^{n} X_i\right| > \varepsilon\right) \leq 2g(n)\exp\left\{-t\varepsilon + t^2 \sum_{i=1}^{n} EX_i^2\right\}.$$

Proof By $1 \leq i \leq n$, $|tX_i| \leq 1$ a.s. and noting that $1 + x \leq e^x$ for $x \in R$, we have that

$$\begin{aligned} E\exp(tX_i) &= \sum_{k=0}^{\infty} \frac{E(tX_i)^k}{k!} \leq 1 + E(tX_i)^2\left\{\frac{1}{2!} + \frac{1}{3!} + \cdots\right\} \\ &\leq 1 + t^2 EX_i^2 \leq \exp\left\{t^2 EX_i^2\right\}. \end{aligned}$$

For any $\varepsilon > 0$, using Markov's inequality and Lemma 2.1(ii), we can get

$$\begin{aligned} P\left(\sum_{i=1}^{n} X_i > \varepsilon\right) &\leq \exp(-t\varepsilon)E\exp\left(t\sum_{i=1}^{n} X_i\right) \leq g(n)\exp(-t\varepsilon)\prod_{i=1}^{n} E\exp(tX_i) \\ &\leq g(n)\exp\left\{-t\varepsilon + t^2\sum_{i=1}^{n} EX_i^2\right\}. \end{aligned} \tag{2.1}$$

By Lemma 2.1(i), we see that the random sequence $\{-X_n, n \geq 1\}$ is still WOD with dominating coefficients $g(n)$, then

$$P\left(\sum_{i=1}^{n} X_i \leq -\varepsilon\right) = P\left(\sum_{i=1}^{n} (-X_i) \geq \varepsilon\right) \leq g(n)\exp\left\{-t\varepsilon + t^2\sum_{i=1}^{n} EX_i^2\right\}. \tag{2.2}$$

Combining (2.1) and (2.2), we complete the proof of Lemma 2.4. □

Corollary 2.1 *Let a random sequence $\{X_n, n \geq 1\}$ be WOD with dominating coefficients $g(n)$, and $EX_i = 0$, $|X_i| \leq d$, a.s. for $i = 1, \ldots, n$, where d is a positive constant. Set $\sigma_n^2 = n^{-1}\sum_{i=1}^{n} EX_i^2$, then for any $\varepsilon > 0$,*

$$P\left(\frac{1}{n}\left|\sum_{i=1}^{n} X_i\right| > \varepsilon\right) \leq 2g(n)\exp\left\{-\frac{n\varepsilon^2}{2(2\sigma_n^2 + d\varepsilon)}\right\}.$$

Proof By Lemma 2.4, taking $t = \varepsilon/(2\sigma_n^2 + d\varepsilon)$, we can get Corollary 2.1 immediately. □

Remark 2.2 By comparison, the conditions of Lemma 2.4 are weaker than those of Theorem 4 in Shen ([17], 2013). So, Lemma 2.4 is a generalization of Theorem 4 in Shen ([17], 2013), and it extends the Bernstein-type inequality in Wang ([18], 2015, Lemma 2.2) from END to WOD sequence.

Lemma 2.5 *Let a random sequence $\{X_n, n \geq 1\}$ be WOD with dominating coefficients $g(n)$. $F(x)$ is a continuous distribution function. If there exists a positive constant $\{\tau_n\}$ such that $\tau_n \to 0$ and $n\tau_n^2/[\log(ng^3(n))] \to \infty$, then*

$$\sup_x \left|F_n(x) - F(x)\right| = o(\tau_n), \quad a.s.$$

In particular, taking $\tau_n = n^{-1/2}[\log(ng^3(n))(\log\log n)^\delta]^{1/2}$ *for some* $\delta > 0$, *we have*

$$\sup_x \bigl|F_n(x) - F(x)\bigr| = o\bigl(n^{-1/2}\bigl[\log\bigl(ng^3(n)\bigr)(\log\log n)^\delta\bigr]^{1/2}\bigr), \quad a.s.$$

Proof The proof is based on a modification of the proof of Lemma 3.5 in Li ([16], 2017). We only outline the difference. Let $x_{n,k}$ satisfy $F(x_{n,k}) = k/n$, $k = 1, 2, \dots, n-1$, and $\xi_{ik} = I(X_i \le x_{n,k}) - EI(X_i \le x_{n,k})$, then

$$P\Bigl(\sup_{x\in R} \bigl|F_n(x) - F(x)\bigr| > \varepsilon\tau_n\Bigr) \le \sum_{k=1}^{n-1} P\left(\frac{1}{n}\left|\sum_{i=1}^{n}\xi_{ik}\right| > \frac{\varepsilon\tau_n}{2}\right). \tag{2.3}$$

By Lemma 2.1, it is easy to see that a random sequence $\{\xi_{ik}, i \ge 1\}$ is WOD with $E\xi_{ik} = 0$, $|\xi_{ik}| \le 2$ for fixed k. Thus, by Lemma 2.4, and by choosing $t = \varepsilon\tau_n/4$ for n large enough, we have

$$P\left(\frac{1}{n}\left|\sum_{i=1}^{n}\xi_{ik}\right| > \frac{\varepsilon\tau_n}{2}\right) \le 2g(n)\exp\left\{-\frac{\varepsilon^2 n\tau_n^2}{16}\right\} \le 2g(n)\bigl[ng^3(n)\bigr]^{-3} \le 2n^{-3}. \tag{2.4}$$

By (2.3) and (2.4), we have

$$P\Bigl(\sup_{x\in R} \bigl|F_n(x) - F(x)\bigr| > \varepsilon\tau_n\Bigr) \le 2\sum_{k=1}^{n-1} n^{-3} \le C_3 n^{-2}.$$

Therefore, by the Borel–Cantelli lemma, we obtain the result of Lemma 2.5. □

3 Main results and proofs

Theorem 3.1 *Let a random sequence* $\{X_n, n \ge 1\}$ *be WOD with dominating coefficients* $g(n)$, *and let* (A1)–(A3) *hold true. Let* $K(\cdot)$ *be a bounded monotonic density function, and* $nh_n^6/[\log(ng^2(n))(\log\log n)^\delta] \to 0$ *for some* $\delta > 0$. *Then, for* $f(x) \in C_{2,\alpha}$,

$$\bigl[nh_n^2/\bigl[\log\bigl(ng^2(n)\bigr)(\log\log n)^\delta\bigr]\bigr]^{1/2}\bigl(\hat{f}_n(x) - f(x)\bigr) \to 0, \quad a.s.$$

Weakening the condition of density kernel $K(\cdot)$ from bounded monotonic density function to bounded variation function, we can get the result as follows.

Corollary 3.1 *Let a random sequence* $\{X_n, n \ge 1\}$ *be WOD with dominating coefficients* $g(n)$, *and let* (A1)–(A3) *hold true. If the density kernel* $K(\cdot)$ *is a Borel measurable and bounded variation function and* $nh_n^6/[\log(ng^2(n))(\log\log n)^\delta] \to 0$ *for some* $\delta > 0$, *then, for* $f(x) \in C_{2,\alpha}$,

$$\bigl[nh_n^2/\bigl[\log\bigl(ng^2(n)\bigr)(\log\log n)^\delta\bigr]\bigr]^{1/2}\bigl(\hat{f}_n(x) - f(x)\bigr) \to 0, \quad a.s.$$

Theorem 3.2 *If, in addition to the assumptions of Theorem* 3.1 *and Lemma* 2.5, *the distribution function* $F(x_0) < 1$, *then for* $x \le x_0$,

$$\bigl[nh_n^2/\bigl[\log\bigl(ng^2(n)\bigr)(\log\log n)^\delta\bigr]\bigr]^{1/2}\bigl(\hat{\lambda}_n(x) - \lambda(x)\bigr) \to 0, \quad a.s.$$

Corollary 3.2 *If, in addition to the assumptions of Corollary* 3.1 *and Lemma* 2.5, *the distribution function* $F(x_0) < 1$, *then for* $x \le x_0$,

$$\left[nh_n^2/\left[\log\left(ng^2(n)\right)(\log\log n)^\delta\right]\right]^{1/2}\left(\hat{\lambda}_n(x) - \lambda(x)\right) \to 0, \quad a.s.$$

Remark 3.1 From Theorems 3.1 and 3.2, Corollaries 3.1 and 3.2, the rates of strong convergence can nearly reach $O(n^{-2/5})$ by choosing bandwidth $h_n = O(n^{-1/5})$. So the rates in here reach the same order of convergence as those in Li ([16], 2017) and Li and Yang ([19], 2005). Note that negative association implies WOD, and END also implies WOD, but the converse does not hold. Then the results of this paper are the generalization of those in Li ([16], 2017) and Li and Yang ([19], 2005).

Proof of Theorem 3.1 Writing

$$nh_n\left[\hat{f}_n(x) - E\hat{f}_n(x)\right] = \sum_{j=1}^n \left[K\left(\frac{x - X_j}{h_n}\right) - EK\left(\frac{x - X_j}{h_n}\right)\right] = \sum_{j=1}^n Y_j, \tag{3.1}$$

where

$$Y_j = K\left(\frac{x - X_j}{h_n}\right) - EK\left(\frac{x - X_j}{h_n}\right).$$

By Lemma 2.1, we see that a random sequence $\{Y_j, 1 \le j \le n\}$ is still WOD with dominating coefficients $g(n)$. And by $K(x)$ is bounded, we can get that $EY_j = 0$, $|Y_j| \le C_3$,

$$EY_j^2 = EK\left(\frac{x - X_j}{h_n}\right)^2 \le C_3 \quad \text{and} \quad \sigma_n^2 = n^{-1}\sum_{j=1}^n EY_j^2 \le C_3.$$

Set $\theta_n = \{nh_n^2/[\log(ng^2(n))(\log\log n)^\delta]\}^{-1/2}$, by Corollary 2.1, for any $\varepsilon > 0$,

$$\begin{aligned}
&P\left(\frac{1}{n}\left|\sum_{j=1}^n Y_j\right| > h_n\theta_n\varepsilon\right) \\
&\quad \le 2g(n)\exp\left\{-\frac{n(\theta_n h_n\varepsilon)^2}{2[2\sigma_n^2 + C_3(\theta_n h_n\varepsilon)]}\right\} \\
&\quad \le 2g(n)\exp\left\{-\frac{\log(ng^2(n))(\log\log n)^\delta\varepsilon^2}{2\{2C_3 + n^{-1/2}[\log(ng^2(n))(\log\log n)^\delta]^{1/2}C_3\varepsilon\}}\right\} \\
&\quad \le 2g(n)\left(ng^2(n)\right)^{-2} = 2n^{-2}.
\end{aligned}$$

Thus, by the Borel–Cantelli lemma, we have

$$\frac{1}{nh_n\theta_n}\left|\sum_{j=1}^n Y_j\right| \to 0, \quad \text{a.s.} \tag{3.2}$$

Hence, using (3.1) and (3.2), we get

$$\theta_n^{-1}\left[\hat{f}_n(x) - E\hat{f}_n(x)\right] \to 0, \quad \text{a.s.} \tag{3.3}$$

Note that

$$\theta_n^{-1}\big[\hat{f}_n(x)-f(x)\big]=\theta_n^{-1}\big[\hat{f}_n(x)-E\hat{f}_n(x)\big]+\theta_n^{-1}\big[E\hat{f}_n(x)-f(x)\big] \tag{3.4}$$

and

$$\theta_n^{-1}h_n^2=\frac{\sqrt{nh_n^6}}{\sqrt{\log(ng^2(n))(\log\log n)^\delta}}\to 0.$$

Then, by Lemma 2.3, we have

$$\theta_n^{-1}\big[E\hat{f}_n(x)-f(x)\big]=\frac{h_n^2}{\theta_n}\cdot\frac{1}{h_n^2}\cdot\big[E\hat{f}_n(x)-f(x)\big]\to 0. \tag{3.5}$$

It follows by using (3.3)–(3.5) that

$$\big\{nh_n^2/\big[\log\big(ng^2(n)\big)(\log\log n)^\delta\big]\big\}^{1/2}\big[\hat{f}_n(x)-f(x)\big]\to 0,\quad \text{a.s.}$$

This completes the proof of Theorem 3.1. □

Proof of Corollary 3.1 By $K(x)$ is a bounded variation function, we can write that $K(x)=K_1(x)-K_2(x)$, where $K_1(x)$ and $K_2(x)$ are two monotone increasing functions. Then

$$nh_n\big[\hat{f}_n(x)-E\hat{f}_n(x)\big]=\sum_{j=1}^n Y_{1j}-\sum_{j=1}^n Y_{2j}, \tag{3.6}$$

where

$$Y_{ij}=K_i\left(\frac{x-X_j}{h_n}\right)-EK_i\left(\frac{x-X_j}{h_n}\right)\quad \text{for } i=1,2.$$

Then the following proof of Corollary 3.1 is the same as the proof of Theorem 3.1, here it is omitted. □

Proof of Theorem 3.2 Set $S(x)=1-F(x)$, $S_n(x)=1-F_n(x)$. By the hazard rate estimator (1.2), we get

$$\big|\hat{\lambda}_n(x)-\lambda(x)\big|\le\frac{S(x)|\hat{f}_n(x)-f(x)|+f(x)|S_n(x)-S(x)|}{S(x)S_n(x)}. \tag{3.7}$$

Note that $0<S(x_0)\le S(x)\le 1$ for $x\le x_0$, and $\sup_x f(x)\le C_5$. It follows by Theorem 3.1 and Lemma 2.5 that

$$\big[nh_n^2/\big[\log\big(ng^2(n)\big)(\log\log n)^\delta\big]\big]^{1/2}\big(\hat{f}_n(x)-f(x)\big)\to 0,\quad \text{a.s.} \tag{3.8}$$

and

$$n^{1/2}/\big[(\log\big(ng^3(n)\big)(\log\log n)^\delta\big]^{1/2}\sup_{x\le x_0}\big|S_n(x)-S(x)\big|\to 0,\quad \text{a.s.} \tag{3.9}$$

On the other hand, for n large enough, as $x \le x_0$, we have

$$S_n(x) > S(x) - S(x_0) > \frac{1}{2}S(x_0) > 0.$$

Hence, from (3.7), (3.8), and (3.9), we have

$$\bigl[nh_n^2/\bigl[\log\bigl(ng^2(n)\bigr)(\log\log n)^\delta\bigr]\bigr]^{1/2}\bigl(\hat{\lambda}_n(x) - \lambda(x)\bigr) \to 0, \quad \text{a.s.}$$

Then we obtain Theorem 3.2. □

Proof of Corollary 3.2 The proof is analogous to the one of Theorem 3.2 by Corollary 3.1, and here it is also omitted. □

4 Conclusion

In Sect. 2, we give a Bernstein-type inequality for WOD random sequence which extends the Bernstein-type inequality based END sequence. In Sect. 3, using the Bernstein-type inequality, we obtain the rates of strong convergence for the estimators of density and hazard rate functions. By choosing the bandwidth $h_n = O(n^{-1/6})$, the rates of strong convergence can nearly reach $O(n^{-1/3})$.

Acknowledgements
The authors thank the referee and the editor for their very valuable comments on an earlier version of this paper. Li's work was supported by the National Natural Science Foundation of China (NSFC) (11461057), the National Science Foundation of Jiangxi (20161BAB201003). Zhou's work was supported by the State Key Program of National Natural Science Foundation of China (71331006), the State Key Program in the Major Research Plan of National Natural Science Foundation of China (91546202), National Center for Mathematics and Interdisciplinary Sciences (NCMIS), Key Laboratory of RCSDS, AMSS, CAS (2008DP173182) and Innovative Research Team of Shanghai University of Finance and Economics (IRTSHUFE13122402).

Competing interests
The authors declare that they have no competing interests.

Authors' contributions
The authors contributed equally and significantly in writing this paper. All authors read and approved the final manuscript.

Publisher's Note

Received: 21 June 2017 Accepted: 20 March 2018 Published online: 03 April 2018

References

1. Izenman, A.J., Tran, L.T.: Kernel estimation of the survival function and hazard rate under weak dependence. J. Stat. Plan. Inference **24**(2), 233–247 (1990)
2. Cai, Z.W.: Kernel density and hazard rate estimation for censored dependent data. J. Multivar. Anal. **67**(1), 23–34 (1998)
3. Liebscher, E.: Kernel density and hazard rate estimation for censored data under α-mixing condition. Ann. Inst. Stat. Math. **54**(1), 19–28 (2002)
4. Liang, H.Y., Mammitzsch, V., Steinebach, J.: Nonlinear wavelet density and hazard rate estimation for censored data under dependent observations. Stat. Decis. **23**(3), 161–180 (2005)
5. Bouezmarni, T., Ghouch, A.E., Mesfioui, M.: Gamma kernel estimators for density and hazard rate of right-censored data. J. Probab. Stat. **2011**, Article ID 937574 (2011)
6. Wang, Y.B., Cheng, D.Y.: Basic renewal theorems for random walks with widely dependent increments. J. Math. Anal. Appl. **384**(2), 597–606 (2011)
7. Chen, W., Wang, Y.B. Cheng, D.Y.: An inequality of widely dependent random variables and its applications. Lith. Math. J. **56**(1), 16–31 (2016)
8. Cambanis, S.: On Eyraud–Farlie–Gumbel–Morgenstern random processes. In: Dall'Aglio, G., Kotz, S., Salinetti, G. (eds.) Proceedings of the Symposium on Distributions with Given Marginals (Fréchet Classes), pp. 207–222. Kluwer Academic, Dordrecht (1991)

9. Hashorva, E., Hüsler, J.: Extreme values in FGM random sequences. J. Multivar. Anal. **68**(2), 212–225 (1999)
10. Nelsen, R.B.: An Introduction to Copulas, 2nd edn. Springer, New York (2006)
11. Wang, K.Y., Wang, Y.B. Gao, Q.W.: Uniform asymptotics for the finite-time ruin probability of a new dependent risk model with a constant interest rate. Methodol. Comput. Appl. Probab. **15**(1), 109–124 (2013)
12. Shi, S.T., Wu, Q.Y.: The strong consistency for kernel-type density estimation in the case of widely orthant dependent samples. J. Zhejiang Univ. Sci. Ed. **41**(1), 26–28 (2014)
13. Li, Y.M., Ying, R., Cai, J.P., Yao, J.: Pointwise strong consistency of recursive kernel estimator for probability density and failure rate function under WOD sequence. J. Jilin Univ. Sci. Ed. **53**(6), 1134–1138 (2015)
14. Cai, Z.W.: Asymptotic normality of recursive kernel density estimates under dependent assumptions. Chinese J. Appl. Probab. Statist. **9**(2), 123–129 (1993)
15. Lin, Z.Y.: Asymptotic normality of kernel regression function estimates. Adv. Math. **16**(1), 97–102 (1987)
16. Li, Y.M.: On the rate of strong convergence for a recursive probability density estimator of END samples and its applications. J. Math. Inequal. **11**(2), 335–343 (2017)
17. Shen, A.: Bernstein-type inequality for widely dependent sequence and its application to nonparametric regression models. Abstr. Appl. Anal. **2013**, Article ID 862602 (2013)
18. Wang, X.J., Zheng, L.L., Xu, C., Hu, S.H.: Complete consistency for the estimator of nonparametric regression models based on extended negatively dependent errors. Statistics **49**(2), 396–407 (2015)
19. Li, Y.M., Yang, S.C.: Strong convergence rate of recursive probability density estimator for NA sequences. Chin. J. Eng. Math. **22**(4), 659–665 (2005)