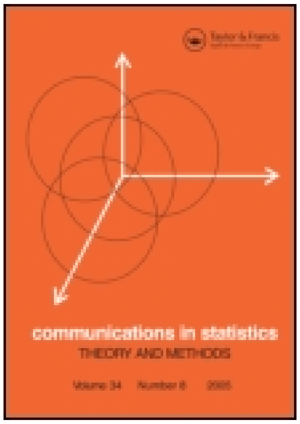

Communications in Statistics - Theory and Methods
Publication details, including instructions for authors and subscription information:
http://www.tandfonline.com/loi/lsta20

Central Limit Theorem for ISE of Kernel Density Estimators in Censored Dependent Model

Sarah Jomhoori [a], Vahid Fakoor [a] & Hasanali Azarnoosh [a]

[a] Department of Statistics, Ferdowsi University of Mashhad, Masshad, Iran

Available online: 12 Mar 2012

To cite this article: Sarah Jomhoori, Vahid Fakoor & Hasanali Azarnoosh (2012): Central Limit Theorem for ISE of Kernel Density Estimators in Censored Dependent Model, Communications in Statistics - Theory and Methods, 41:8, 1334-1349

To link to this article: http://dx.doi.org/10.1080/03610926.2010.542849

Communications in Statistics—Theory and Methods, 41: 1334–1349, 2012

ISSN: 0361-0926 print/1532-415X online
DOI: 10.1080/03610926.2010.542849

Central Limit Theorem for ISE of Kernel Density Estimators in Censored Dependent Model

SARAH JOMHOORI, VAHID FAKOOR,
AND HASANALI AZARNOOSH

Department of Statistics, Ferdowsi University of Mashhad, Masshad, Iran

In some long-term studies, a series of dependent and possibly censored failure times may be observed. Suppose that the failure times have a common continuous distribution function F. A popular stochastic measure of the distance between the density function f of the failure times and its kernel estimate f_n is the integrated square error(ISE). In this article, we derive a central limit theorem for the integrated square error of the kernel density estimators under a censored dependent model.

Keywords α-mixing; Bandwidth; Censored dependent data; Integrated square error; Kaplan–Meier estimator; Kernel density estimator.

Mathematics Subject Classification 62G07; 62G20.

1. Introduction and Main Result

Let $X_1, \ldots, X_n$ be a sequence of failure times, having a common unknown continuous marginal distribution function F with a probability density function $f = F'$. The random variables are not assumed to be mutually independent (see assumption (1) for the kind of dependence stipulated). Let the random variable X_i be censored on the right by the random variable Y_i, so that one observes only

$$Z_i = X_i \wedge Y_i \quad \text{and} \quad \delta_i = I(X_i \leq Y_i),$$

where $\wedge$ denotes minimum and $I(.)$ is the indicator of the event specified in parentheses. In this random censorship model, the censoring times $Y_1, \ldots, Y_n$ are assumed to be independently and identically distributed and they are also assumed to be independent of the X_i's. For easy reference, denote with G the distribution of the Y_i's. Since censored data traditionally occur in lifetime analysis, we assume that X_i and Y_i are nonnegative. The actually observed Z_i's have a distribution function H satisfying

$$\overline{H}(t) = 1 - H(t) = (1 - F(t))(1 - G(t)).$$

Received February 3, 2008; Accepted November 20, 2010
Address correspondence to Sarah Jomhoori, Department of Statistics, Ferdowsi University of Mashhad, Mashhad, Iran; E-mail: sjomhoori@birjand.ac.ir

Let

$$H^*(t) = P(Z \leq t, \delta = 1).$$

Define

$$N_n(t) = \sum_{i=1}^{n} I(Z_i \leq t, \delta = 1) = \sum_{i=1}^{n} I(X_i \leq t \wedge Y_i),$$

the number of uncensored observations less than or equal to t, and

$$Y_n(t) = \sum_{i=1}^{n} I(Z_i \geq t),$$

the number of censored or uncensored observations greater than or equal to t and also the empirical distribution function of $\overline{H}(t)$ and $H^*(t)$ are, respectively, defined as

$$\overline{Y}_n(t) = n^{-1} Y_n(t), \quad \overline{N}_n(t) = n^{-1} N_n(t).$$

Then the Kaplan–Meier (K-M) estimator for $1 - F(t)$, based on the censored data is

$$1 - \widehat{F}_n(t) = \prod_{s \leq t} \left(1 - \frac{dN_n(s)}{Y_n(s)}\right), \quad t < Z_{(n)},$$

where $Z_{(i)}$'s are the order statistics of Z_i's and $dN_n(t) = N_n(t) - N_n(t-)$. The empirical cumulative hazard function for the underlying cumulative hazard function $\Lambda(t) = -log(1 - F(t))$ is

$$\widehat{\Lambda}_n(t) = \int_{-\infty}^{t} \frac{dN_n(s)}{Y_n(s)}.$$

For the case that the failure time observations are mutually independent, based on the Kaplan–Meier estimator, Blum and Susarla (1980) proposed to estimate f by a sequence of kernel estimators f_n defined by

$$f_n(t) = \frac{1}{h_n} \int_{0}^{\infty} K\left(\frac{t - x}{h_n}\right) d\widehat{F}_n(x),$$

where K is a smooth kernel function and h_n is a sequence of positive bandwidths tending to zero. The properties of the kernel estimator f_n have been examined by Blum and Susarla (1980), Burke and Horváth (1984), and Földes et al. (1981), among others.

It is well known that the most widely accepted stochastic measure of the global performance of a kernel estimator is its integrated square error, defined by

$$ISE(f_n) = \int (f_n(t) - f(t))^2 w(t) dt,$$

where w is a non negative weighted function. The corresponding deterministic measure of the accuracy of f_n is the mean integrated square error given by

$$MISE(f_n) = \int E\left(f_n(t) - f(t)\right)^2 w(t)dt,$$

Integrated square error is often used in simulation studies to measure the performance of density estimators. It is also used implicitly in adaptive constructions of estimators, when the aim is to minimize mean integrated square error in some sense. Both these applications involve the assumption that integrated square error is somehow close to mean integrated square error. The central limit theorem for *ISE* provide an explicit description of the order of this closeness, by showing that

$$c(n)\{ISE(f_n) - MISE(f_n)\} \xrightarrow{\mathscr{L}} N(0, 1), \tag{1.1}$$

as $n \to \infty$, where $\xrightarrow{\mathscr{L}}$ denotes convergence in distribution and $c(n), n \geq 1$ is a sequence of positive constants diverging to infinity. The asymptotic behavior of *ISE* has been studied extensively by many authors. In the uncensored case, Bickel and Rosenblatt (1973) employed the uniform strong approximation of the empirical process by the Brownian bridge to obtain a central limit theorem for the *ISE* of the Rosenblatt–Parzen kernel estimators of a density function. Hall (1984) derived central limit theorem for the *ISE* of density estimator using martingale theory and U-statistics approach. In general, central limit theorem for L_p deviation between f_n and f, i.e.,

$$I_n(f_n, p) = \int |f_n(t) - f(t)|^p dt, \quad 1 \leq p < \infty,$$

was studied by Csörgő and Horváth (1988) based on complete samples.

In the right-censored case, Csörgő et al. (1991) obtained central limit theorem for L_p distances of kernel estimators. Yang (1993) employed the martingale techniques by Gill (1983) to get a central limit theorem for the *ISE* of the product-limit kernel density estimators. Zhang (1996) obtained a simple asymptotic expression for the mean integrated square error of the kernel estimator f_n, and then derived an asymptotically optimal bandwidth for f_n. Zhang (1998) applied the technique of strong approximation to establish an asymptotic expansion for *ISE* of the kernel density estimate f_n. Sun and Zheng (1999) proved a central limit theorem for the *ISE* of the kernel hazard rate estimators and also presented an asymptotic representation of the *MISE* for kernel hazard rate estimators in left-truncated and right-censored data.

However, for the case that censored observations are dependent, there are no results available. The main aim of this article is to derive a central limit theorem for *ISE* of kernel density estimator when the censored data exhibit some kind of dependence.

Among various mixing conditions used in the literature, α-mixing, whose definition is given below is reasonably weak and has many practical applications. Many stochastic processes and time series are known to be α-mixing. In particular, the stationary autoregressive-moving average (ARMA) processes, which are widely applied in time series analysis, are α-mixing with exponential mixing coefficient, i.e., $\alpha(n) = e^{-\nu n}$ for some $\nu > 0$.

Definition 1.1. Let $\{X_i, i \geq 1\}$ denote a sequence of random variables. Given a positive integer m, set

$$\alpha(m) = \sup_{k\geq 1}\{|P(A \cap B) - P(A)P(B)|;\ A \in \mathscr{F}_1^k,\ B \in \mathscr{F}_{k+m}^{\infty}\}, \tag{1.2}$$

where $\mathscr{F}_i^k$ denote the σ-field of events generated by $\{X_j;\ i \leq j \leq k\}$. The sequence is said to be α-mixing (strongly mixing) if the mixing coefficient $\alpha(m) \to 0$ as $m \to \infty$.

For such mixing condition, the following basic inequality (see Kim, 1994) is well known. Let ξ and η be measurable with respect to $\mathscr{F}_1^k$ and $\mathscr{F}_{k+n}^{\infty}$, respectively. Then, under α-mixing,

$$|E(\xi\eta) - E(\xi)E(\eta)| \leq 2\pi[\alpha(n)]^{1-(1/p+1/q)}E^{1/p}|\xi|^p E^{1/q}|\eta|^q, \tag{1.3}$$

where $E|\xi|^p, E|\eta|^q < \infty$ for $1 \leq p, q \leq \infty$ with $1/p + 1/q = 1 - \delta$ and $0 < \delta < 1$.

For the sake of simplicity, the assumptions used in this article are as follows.

Assumptions.

1. Suppose that $\{X_i, i \geq 1\}$ is a sequence of stationary strongly mixing random variables with continuous distribution function F and mixing coefficient α as defined on (1.2).
2. Suppose that the censoring time variables $\{Y_i, i \geq 1\}$ are i.i.d. random variables with continuous distribution function G and are independent of $\{X_i, i \geq 1\}$.
3. $\alpha(n) = O(n^{-v})$ for some $v > 4$.
4. K is a bounded, non negative and continuously differentiable function on $[-1, 1]$ and satisfies the following conditions:

$$\int_{-1}^{1} K(t)dt = 1, \quad \int_{-1}^{1} tK(t)dt = 0, \quad \int_{-1}^{1} t^2K(t)dt = k \neq 0,$$

5. The joint probability density function (pdf) of X_1 and X_j, $f_{1,j}(.,.)$ exists and $|f_{1,j}(u, v) - f(u)f(v)| \leq M$ for all $j \geq 2$ and $u, v \in \mathscr{R}$ and some constant M.
6. f is twice continuously differentiable at t and $\inf_{0\leq t\leq\tau} f(t) > 0$, where τ is such that $\overline{H}(\tau) > 0$.
7. The weight function w is continuously differentiable and supported on $[0, \tau]$.

Remark 1.1. In uncensored case, Rosenblatt (1985) established asymptotic normality, under α-mixing including the assumption that $\sum_{j=2}^{n}|f_{1,j}(u, v) - f(u)f(v)| < \infty$ for $u, v \in \mathscr{R}$. The assumption $|f_{1,j}(u, v) - f(u)f(v)| \leq M < \infty$ has been considered by Masry (1986) to derive asymptotic normality of density estimators for strong mixing and asymptotically uncorrelated processes. This assumption has been also considered by Roussas (1990) to derive asymptotic normality of kernel estimators in a different way. Cai (1998b) applied this assumption to establish the asymptotic normality and the uniform consistency of the kernel estimators for density and hazard function under a censored dependent model.

The main result of this article is the following theorem, which presents a central limit theorem for $ISE(f_n)$. The proof is deferred to Sec. 2. The integrated square error of f_n on the interval $[0, \tau]$ is defined by

$$ISE(f_n) = \int_0^\tau \left(f_n(t) - f(t)\right)^2 w(t)dt. \tag{1.4}$$

Let

$$d_{n1} = \int_0^\tau (\overline{f}_n(t) - f(t))^2 w(t)dt + (nh_n)^{-1}\left(\int K^2(u)du\right)\int \frac{f(t)}{\overline{G}(t)} w(t)dt,$$

where

$$\overline{f}_n(t) = h_n^{-1}\int_0^\infty K\left(\frac{t-u}{h_n}\right) f(u)du,$$

and

$$\begin{aligned}\sigma_{01}^2 &= \left(\int_{-1}^1 u^2K(u)du\right)^2 \int (f''(t)w(t))^2 \frac{f(t)}{\overline{G}(t)}dt \\ &\quad + 16\pi\sum_{i=1}^\infty \alpha^{1/2}(i)\left(\int_{-1}^1 u^2K(u)du\right)^2\left(\int\left(\frac{f''(t)w(t)}{\overline{G}(t)}\right)^4 f(t)dt\right)^{1/2}.\end{aligned} \tag{1.5}$$

Theorem 1.1. *Let h_n be a sequence of positive bandwidth which satisfies $nh_n^6 \to \infty$ as $n \to \infty$. Under stated assumptions, we have*

$$h_n^{-2}\sqrt{n}(ISE(f_n) - d_{n1}) \xrightarrow{\mathscr{L}} N(0, \sigma_1^2),$$

where $\sigma_1^2 \in (0, \sigma_{01}^2)$.

Remark 1.2. The above theorem in iid case has been proved by Hall (1984), Yang (1993) and Sun and Zheng (1999), respectively. The condition imposed on the bandwidth in those papers is $nh_n^5 \to \infty$ as $n \to \infty$. Csörgő et al. (1991) established central limit theorem for L_p deviation between f_n and f with the optimal bandwidth choice which satisfies $nh_n^5 \to c > 0$, where c is a positive constant. However, this conditions does not seem to be achievable in α-mixing setting.

Another stochastic measure of accuracy is Hellinger distance defined by

$$HD(f_n) = \int_0^\tau \left(\sqrt{f_n(t)} - \sqrt{f(t)}\right)^2 dt. \tag{1.6}$$

Let

$$d_{n2} = \int_0^\tau \frac{(\overline{f}_n(t) - f(t))^2}{4f(t)}dt + (nh_n)^{-1}\left(\int K^2(u)du\right)\int \frac{1}{4\overline{G}(t)}dt,$$

$$\begin{aligned}\sigma_{02}^2 &= \frac{1}{16}\left(\int_{-1}^1 u^2K(u)du\right)^2\int \frac{(f''(t))^2}{f(t)\overline{G}(t)}dt \\ &\quad + \pi\sum_{i=1}^\infty \alpha^{1/2}(i)\left(\int_{-1}^1 u^2K(u)du\right)^2\left(\int \frac{(f''(t))^4}{f^3(t)\overline{G}^4(t)}dt\right)^{1/2}.\end{aligned}$$

Corollary 1.1. *Under the stipulated assumptions on the theorem*

$$h_n^{-2}\sqrt{n}(HD(f_n) - d_{n2}) \xrightarrow{\mathscr{L}} N(0, \sigma_2^2),$$

where $\sigma_2^2 \in (0, \sigma_{02}^2)$.

2. Proofs

The proof of theorem is based on the following lemmas. We begin with introducing some further notations. We define

$$f_n^*(t) = h_n^{-1}\int_0^\infty \frac{K\left(\frac{t-u}{h_n}\right)}{\overline{G}(u)} d\overline{N}_n(u),$$

$$Q_{n1} = 2\int_0^\tau (f_n(t) - \overline{f}_n(t))(\overline{f}_n(t) - f(t))w(t)dt, \tag{2.1}$$

$$Q_{n2} = \int_0^\tau (f_n(t) - \overline{f}_n(t))^2 w(t)dt. \tag{2.2}$$

Let $g(x) = \int_0^x (\overline{H}(s))^{-2} dH^*(s)$, and for positive real z and x and $\delta = 0$ or 1, let

$$\xi(z, \delta, x) = g(z \wedge x) - I(z \leq x, \delta = 1)/\overline{H}(z).$$

Observe that

$$E(\xi(Z_i, \delta_i, x)) = 0, \quad Cov(\xi(Z_i, \delta_i, s), \xi(Z_i, \delta_i, t)) = g(s \wedge t).$$

Lemma 2.1 in Cai (1998a) ensures that $\{\xi(Z_i, \delta_i, t)\}$ is a sequence of stationary α-mixing bounded random variables.

Lemma 2.1. *Under the assumptions of Theorem* 1.1, *we have*

$$h_n^{-2}\sqrt{n}Q_{n1} \xrightarrow{\mathscr{L}} N(0, \sigma_1^2), \tag{2.3}$$

where $\sigma_1^2 \in (0, \sigma_{01}^2)$, *and* σ_{01}^2 *is defined in* (1.5).

Proof. It has been shown by Cai (1998a) that

$$\widehat{F}_n(t) - F(t) = \overline{F}(t)(\widehat{\Lambda}_n(t) - \Lambda(t)) + R_n(t), \tag{2.4}$$

where

$$\sup_{t\geq 0} |R_n(t)| = O(a_n^2) \quad a.s., \tag{2.5}$$

and

$$a_n = \left(\frac{\log\log n}{n}\right)^{1/2}.$$

Simple algebra shows

$$
\begin{aligned}
\widehat{\Lambda}_n(t) - \Lambda(t) = {} & \left(\int_0^t \frac{d\overline{N}_n(s)}{\overline{H}(s)} - \int_0^t \frac{\overline{Y}_n(s)}{\overline{H}^2(s)} dH_*(s) \right) \\
& + \int_0^t \left(\frac{1}{\overline{Y}_n(s)} - \frac{1}{\overline{H}(s)} \right) d\overline{N}_n(s) \\
& + \int_0^t \frac{\overline{Y}_n(s) - \overline{H}(s)}{\overline{H}^2(s)} dH_*(s). \qquad (2.6)
\end{aligned}
$$

Integration by parts in conjunction with (2.4) and (2.6) implies

$$
\begin{aligned}
f_n(t) - \overline{f}_n(t) = {} & -h_n^{-1} \int \overline{F}(x)(\widehat{\Lambda}_n(x) - \Lambda(x)) dK\left(\frac{t-x}{h_n}\right) - h_n^{-1} \int R_n(x) dK\left(\frac{t-x}{h_n}\right) \\
= {} & (nh_n)^{-1} \sum_{i=1}^{n} \int \overline{F}(x) \xi(Z_i, \delta_i, x) dK\left(\frac{t-x}{h_n}\right) \\
& - h_n^{-1} \int \overline{F}(x) \int_0^x \left(\frac{1}{\overline{Y}_n(s)} - \frac{1}{\overline{H}(s)} \right) d\overline{N}_n(s) dK\left(\frac{t-x}{h_n}\right) \\
& - h_n^{-1} \int \overline{F}(x) \int_0^x \left(\frac{\overline{Y}_n(s) - \overline{H}(s)}{\overline{H}^2(s)} \right) \; dH_*(s) dK\left(\frac{t-x}{h_n}\right) + r_n(t),
\end{aligned}
$$

where

$$
\sup_{t \geq 0} |r_n(t)| = O(h_n^{-1} a_n^2) \quad a.s. \qquad (2.7)
$$

Denote the second and the third terms in the above equality I_{n1} and I_{n2}, respectively. To estimate I_{n1}, divide the interval $[0, \tau]$ into subintervals $[x_i, x_{i+1}]$, $i = 1, \ldots, m_n$, where $m_n = O(a_n^{-1/2})$ and $0 = x_1 < x_2 < \cdots < x_{m_n+1} = \tau$ are such that $H_*(x_{i+1}) - H_*(x_i) = O(a_n)$. Using Theorem 1 of Cai (1998a), we conclude that

$$
\begin{aligned}
|I_{n1}| = {} & h_n^{-1} \left| \int \overline{F}(x) \int_0^x \left(\frac{1}{\overline{Y}_n(s)} - \frac{1}{\overline{H}(s)} \right) d\overline{N}_n(s) dK\left(\frac{t-x}{h_n}\right) \right| \\
= {} & h_n^{-1} \left| \int_{-1}^{1} \overline{F}(t - h_n u) \int_0^{t-h_n u} \left(\frac{1}{\overline{Y}_n(s)} - \frac{1}{\overline{H}(s)} \right) d\overline{N}_n(s) dK(u) \right| \\
\leq {} & h_n^{-1} \max_{1 \leq i \leq m_n} \sup_{x_i \leq t - h_n u \leq x_{i+1}} \int_{-1}^{1} \left| \int_{x_i}^{t-h_n u} \left(\frac{1}{\overline{Y}_n(s)} - \frac{1}{\overline{H}(s)} \right) d\overline{N}_n(s) \right| |dK(u)| \\
& + h_n^{-1} \max_{1 \leq i \leq m_n} \int_{-1}^{1} \left| \int_0^{x_i} \left(\frac{1}{\overline{Y}_n(s)} - \frac{1}{\overline{H}(s)} \right) d\overline{N}_n(s) \right| |dK(u)| \\
\leq {} & O(h_n^{-1} a_n) \max_{1 \leq i \leq m_n} \int_{-1}^{1} |\overline{N}_n(x_{i+1}) - \overline{N}_n(x_i)| |dK(u)| \\
& + O(h_n^{-1} a_n) \max_{1 \leq i \leq m_n} \int_{-1}^{1} \left| \sum_{j=1}^{i-1} \int_{x_j}^{x_{j+1}} d\overline{N}_n(s) \right| |dK(u)| \\
= {} & O(h_n^{-1} a_n^{3/2}) \quad a.s. \qquad (2.8)
\end{aligned}
$$

Similarly, we conclude that

$$|I_{n2}| = O(h_n^{-1} a_n^{3/2}) \quad a.s. \tag{2.9}$$

Therefore,

$$Q_{n1} = \frac{1}{nh_n} \sum_{i=1}^{n} V_{ni} + J_{n1} + J_{n2} + O(h_n a_n^2) \quad a.s., \tag{2.10}$$

where

$$V_{ni} = 2 \int_0^{\tau} (\bar{f}_n(t) - f(t)) \int \overline{F}(x) \xi(Z_i, \delta_i, x) dK\left(\frac{t-x}{h_n}\right) w(t) dt,$$

and

$$J_{n1} = -2h_n^{-1} \int_0^{\tau} \big(\bar{f}_n(t) - f(t)\big) w(t) \int \overline{F}(x) \int_0^x \left(\frac{1}{\overline{Y}_n(s)} - \frac{1}{\overline{H}(s)}\right) d\overline{N}_n(s) dK\left(\frac{t-x}{h_n}\right) dt,$$

$$J_{n2} = -2h_n^{-1} \int_0^{\tau} \big(\bar{f}_n(t) - f(t)\big) w(t) \int \overline{F}(x) \int_0^x \left(\frac{\overline{Y}_n(s) - \overline{H}(s)}{\overline{H}^2(s)}\right) dH_*(s) dK\left(\frac{t-x}{h_n}\right) dt.$$

Using Taylor's expansion

$$\bar{f}_n(t) - f(t) = \frac{h_n^2}{2} f''(t) \int_{-1}^{1} u^2 K(u) du + o(h_n^2). \tag{2.11}$$

Therefore, (2.11) in conjunction with (2.8) and (2.9) implies that

$$J_{n1} = O(a_n^{3/2} h_n) \quad a.s., \tag{2.12}$$

$$J_{n2} = O(a_n^{3/2} h_n) \quad a.s. \tag{2.13}$$

It is clear to see that $\{V_{ni}\}$ is a sequence of stationary α-mixing bounded random variables. It can be easily checked that

$$\begin{aligned}
E(V_{ni}) &= 0, \\
E(V_{ni}^2) &= 4 \int_0^{\tau} \int_0^{\tau} \big(\bar{f}_n(t) - f(t)\big)\big(\bar{f}_n(s) - f(s)\big) \Big[\int \int \overline{F}(x) \overline{F}(y) \\
&\qquad \times g(x \wedge y) dK\left(\frac{t-x}{h_n}\right) dK\left(\frac{s-y}{h_n}\right) \Big] w(t) w(s) dt ds \\
&\le 4 \int_0^{\tau} \int_0^{\tau} \big(\bar{f}_n(t) - f(t)\big)\big(\bar{f}_n(s) - f(s)\big) \Big[\int K\left(\frac{t-x}{h_n}\right) K\left(\frac{s-x}{h_n}\right) \\
&\qquad \times \frac{f(x)}{1 - G(x)} dx \Big] w(t) w(s) dt ds \\
&\le h_n^6 \left(\int_{-1}^{1} u^2 K(u) du\right)^2 \int \big(f''(x) w(x)\big)^2 \frac{f(x)}{1 - G(x)} dx + o(h_n^6).
\end{aligned}$$

We may write, $V_{ni} = V_{ni1} - V_{ni2}$, where

$$V_{ni1} = 2\int_0^\tau \left(\bar{f}_n(t) - f(t)\right) \int \overline{F}(x) \int_0^x \frac{I(Z_i \geq u)}{\overline{H}^2(u)} dH^*(u) dK\left(\frac{t-x}{h_n}\right) w(t)dt,$$

and

$$V_{ni2} = 2\int_0^\tau \left(\bar{f}_n(t) - f(t)\right) \int_0^\infty \frac{\overline{F}(x)}{\overline{H}(Z_i)} I(Z_i \leq x, \delta_i = 1) dK\left(\frac{t-x}{h_n}\right) w(t)dt.$$

By easy computations, we derive

$$E|V_{ni1}|^m \leq h_n^{3m} \left|\int_{-1}^1 u^2 K(u)du\right|^m \int |f''(x)w(x)|^m \frac{f(x)}{\overline{G}^m(x)} dx + o(h_n^{3m}),$$

$$E|V_{ni2}|^m \leq h_n^{3m} \left|\int_{-1}^1 u^2 K(u)du\right|^m \int |f''(x)w(x)|^m \frac{f(x)}{\overline{H}^m(x)} dx + o(h_n^{3m}).$$

Therefore,

$$E^{1/m}|V_{ni}|^m \leq 2h_n^3 \left|\int_{-1}^1 u^2 K(u)du\right| \left(\int \left|\frac{f''(x)w(x)}{\overline{H}(x)}\right|^m f(x)dx\right)^{1/m} + o(h_n^3).$$

Applying Eq. (1.3) with $p = q = 4$, we get

$$|E(V_{n1}V_{nj})| \leq 8\pi h_n^6 \alpha^{1/2}(j-1) \left(\int_{-1}^1 u^2 K(u)du\right)^2 \left(\int \left(\frac{f''(x)w(x)}{\overline{H}(x)}\right)^4 f(x)dx\right)^{1/2} + o(h_n^6).$$

So, we have

$$Var\left(\sum_{i=1}^n V_i\right) = n\sigma^2(1 + o(1)),$$

where

$$\begin{aligned}
\sigma^2 &= E(V_{nj}^2) + 2\sum_{j=2}^\infty E(V_{n1}V_{nj}) \\
&\leq h_n^6 \left(\int_{-1}^1 u^2 K(u)du\right)^2 \int (f''(x)w(x))^2 \frac{f(x)}{1-G(x)} dx \\
&\quad + 16\pi h_n^6 \sum_{j=2}^\infty \alpha^{1/2}(j-1) \left(\int_{-1}^1 u^2 K(u)du\right)^2 \left(\int \left(\frac{f''(x)w(x)}{\overline{H}(x)}\right)^4 f(x)dx\right)^{1/2} + o(h_n^6).
\end{aligned}$$

Applying Theorem 18.5.4 in Ibragimov and Linnik (1971), we obtain the result.

Lemma 2.2. *Under assumptions of Theorem* 1.1, *we have*

$$\int_0^\tau (f_n^*(t) - \overline{f}_n(t))^2 w(t)dt = (nh_n)^{-1} \left(\int K^2(u)du\right) \left(\int \frac{f(x)}{\overline{G}(x)} w(x)dx\right) + o_p(n^{-1/2}h_n^2).$$

Proof. It can be checked easily that

$$\int_0^\tau (f_n^*(t) - \overline{f}_n(t))^2 w(t)dt = \frac{1}{(nh_n)^2}\int_0^\tau \left(\sum_{i=1}^n X_i(t)\right)^2 w(t)dt,$$

where $X_i(t)$'s are bounded random variables defined by

$$X_i(t) = K\left(\frac{t-Z_i}{h_n}\right)\frac{\delta_i}{\overline{G}(Z_i)} - E\left(K\left(\frac{t-Z_i}{h_n}\right)\frac{\delta_i}{\overline{G}(Z_i)}\right).$$

It is easy to find out

$$\begin{aligned}\int_0^\tau E\big(f_n^*(t) - \overline{f}_n(t)\big)^2 w(t)dt &\le \frac{1}{nh_n^2}\int_0^\tau E\big(X_i^2(t)\big)w(t)dt \\ &\quad + \frac{2}{n^2h_n^2}\int_0^\tau \sum_{1\le i<j\le n} E\big(X_i(t)X_j(t)\big)w(t)dt \\ &= I_{n1} + I_{n2}.\end{aligned}$$

Simple algebra shows

$$I_{n1} = \frac{1}{nh_n}\left(\int K^2(u)du\right)\int \frac{f(x)}{\overline{G}(x)}w(x)dx + o(n^{-1}). \tag{2.14}$$

To estimate the second term, define A_1 and A_2 as follows:

$$\begin{aligned}A_1 &= \{(i,j)\,|\, i,j\in\{1,\ldots,n\}, 1\le j-i\le \varphi_n\},\\ A_2 &= \{(i,j)\,|\, i,j\in\{1,\ldots,n\}, \varphi_n+1\le j-i\le n-1\},\end{aligned}$$

where $\varphi_n = h_n^{-1}\log n$. Assumption (5) implies that for all $i<j$

$$\begin{aligned}\big|E\big(X_i(t)X_j(t)\big)\big| &= \left|\int\int K\left(\frac{t-x}{h_n}\right)K\left(\frac{t-y}{h_n}\right)f_{1,j-i+1}(x,y)dxdy\right. \\ &\quad \left. - \int K\left(\frac{t-x}{h_n}\right)f(x)dx\int K\left(\frac{t-y}{h_n}\right)f(y)dy\right| \\ &\le h_n^2\int\int K(u)K(v)|f_{1,j-i+1}(t-h_nu, t-h_nv) \\ &\quad - f(t-h_nu)f(t-h_nv)|du\,dv \\ &\le Mh_n^2.\end{aligned} \tag{2.15}$$

We have

$$\sum_{1\le i<j\le n}\big|E\big(X_i(t)X_j(t)\big)\big| = \sum_{(i,j)\in A_1}\big|E\big(X_i(t)X_j(t)\big)\big| + \sum_{(i,j)\in A_2}\big|E\big(X_i(t)X_j(t)\big)\big|.$$

Using (2.15), we can write

$$\sum_{(i,j)\in A_1} \left|E\big(X_i(t)X_j(t)\big)\right| = \sum_{i=1}^{n-\varphi_n} \sum_{j=i+1}^{i+\varphi_n} |E\big(X_i(t)X_j(t)\big)|$$
$$= O(nh_n \log n). \tag{2.16}$$

We conclude from Eq. (1.3) that

$$|E\left(X_i(t)X_j(t)\right)| = O\left(\alpha^{\delta}(j-i)\left\{E(X_i^{\frac{2}{1-\delta}}(t))E(X_j^{\frac{2}{1-\delta}}(t))\right\}^{\frac{1-\delta}{2}}\right)$$
$$= O(h_n^{1-\delta}\alpha^{\delta}(j-i)). \tag{2.17}$$

Therefore, we have

$$\sum_{(i,j)\in A_2} \left|E\big(X_i(t)X_j(t)\big)\right| = O\left(h_n^{1-\delta}\sum_{q=\varphi_n}^{n-1}\sum_{i=1}^{n-q}\alpha^{\delta}(q)\right) = O\left(nh_n^{1-\delta}\sum_{q=\varphi_n}^{n-1}\alpha^{\delta}(q)\right)$$
$$= O\left(nh_n^{1-\delta}\sum_{q=\varphi_n}^{n-1}\left(\frac{q}{\varphi_n}\right)^{\delta}\alpha^{\delta}(q)\right). \tag{2.18}$$

Now choose δ such that $(v-1)\delta > 1$, then (2.18) implies that

$$\sum_{(i,j)\in A_2} \left|E\big(X_i(t)X_j(t)\big)\right| = o(nh_n). \tag{2.19}$$

Hence, (2.16) and (2.19) implies

$$I_{n2} = o(n^{-1/2}h_n^2). \tag{2.20}$$

By means of Holder's inequality,

$$Var\left(\int_0^{\tau}\left(\sum_{i=1}^{n}X_i(t)\right)^2 w(t)dt\right)$$
$$\leq \int_0^{\tau}\int_0^{\tau} E\left(\left[\sum_{i=1}^{n}X_i(t)\right]^2\left[\sum_{i=1}^{n}X_i(s)\right]^2\right) w(t)w(s)dt\,ds$$
$$\leq \int_0^{\tau}\int_0^{\tau} E^{1/2}\left(\sum_{i=1}^{n}X_i(t)\right)^4 E^{1/2}\left(\sum_{i=1}^{n}X_i(s)\right)^4 w(t)w(s)dt\,ds.$$

Obviously, we could write

$$E\left(\sum_{i=1}^{n}X_i(t)\right)^4 = \sum_{i=1}^{n}E\big(X_i^4(t)\big) + \sum_{i\neq j}E\big(X_i^2(t)X_j^2(t)\big) + \sum_{i\neq j}E\big(X_i^3(t)X_j(t)\big)$$
$$+ \sum_{i\neq j\neq k}E\big(X_i^2(t)X_j(t)X_k(t)\big) + \sum_{i\neq j\neq k\neq l}E\big(X_i(t)X_j(t)X_k(t)X_l(t)\big). \tag{2.21}$$

Note that $|X_i(t)| < C$ for all $i \geq 1$, where C is a positive constant. Therefore,

$$\sum_{i=1}^{n} E\big(X_i^4(t)\big) = O(n). \tag{2.22}$$

On the other hand,

$$E\big(X_i^2(t)X_j^2(t)\big) \leq C^2|E(X_i(t)X_j(t))|. \tag{2.23}$$

To the second expectation in (2.21) above, we will apply the same method of computing (2.16) and (2.19) on (2.23), therefore

$$\sum_{i\neq j} E\big(X_i^2(t)X_j^2(t)\big) = O(nh_n \log n). \tag{2.24}$$

Similarly, we conclude

$$\sum_{i\neq j} E\big(X_i^3(t)X_j(t)\big) = O(nh_n \log n). \tag{2.25}$$

To estimate the forth expectation, we proceed as follows:

$$\sum_{i\neq j\neq k} E\big(X_i^2(t)X_j(t)X_k(t)\big) = O\left(\sum_{i<j<k} |E\left(X_i^2(t)X_k(t)\right)|\right). \tag{2.26}$$

By (2.17) we have

$$\begin{aligned}
\sum_{i<j<k} |E\left(X_i^2(t)X_k(t)\right)| &= O\left(h_n^{1-\delta}\sum_{i=1}^{n}\sum_{j=i+1}^{n}\sum_{k=j+1}^{n}\alpha^{\delta}(k-i)\right)\\
&= O\left(h_n^{1-\delta}\sum_{i=1}^{n}\sum_{p=1}^{n}\sum_{q=p+1}^{n}\alpha^{\delta}(q)\right)\\
&= O\left(h_n^{1-\delta}\sum_{i=1}^{n}\sum_{p=1}^{n}\sum_{q=p+1}^{n}\left(\frac{q}{p}\right)^2\alpha^{\delta}(q)\right)\\
&= O\left(nh_n^{1-\delta}\sum_{p=1}^{n}\frac{1}{p^2}\sum_{q=p+1}^{n}q^2\alpha^{\delta}(q)\right).
\end{aligned} \tag{2.27}$$

Now, choose δ such that $v\delta > 3$, then (2.27) implies that

$$\sum_{i\neq j\neq k} E\left(X_i^2(t)X_j(t)X_k(t)\right) = O(nh_n^{1-\delta}). \tag{2.28}$$

We apply similar procedure to derive fifth term in (2.21),

$$\sum_{i\neq j\neq k\neq l} E\big(X_i(t)X_j(t)X_k(t)X_l(t)\big) = O\left(\sum_{i<j<k<l} |E\big(X_i(t)X_l(t)\big)|\right)$$

$$= O\left(h_n^{1-\delta}\sum_{i=1}^{n}\sum_{j=i+1}^{n}\sum_{k=j+1}^{n}\sum_{l=k+1}^{n}\alpha^{\delta}(l-i)\right)$$

$$= O\left(h_n^{1-\delta}\sum_{i=1}^{n}\sum_{m=1}^{n}\sum_{p=m+1}^{n}\sum_{q=p+1}^{n}\left(\frac{q}{p}\right)^{3}\alpha^{\delta}(q)\right)$$

$$= O\left(nh_n^{1-\delta}\sum_{m=1}^{n}\frac{1}{m^{3/2}}\sum_{p=m+1}^{n}\frac{1}{p^{3/2}}\sum_{q=p+1}^{n}q^{3}\alpha^{\delta}(q)\right). \quad (2.29)$$

Now, choose δ such that $v\delta > 4$, then (2.29) implies that

$$\sum_{i\neq j\neq k\neq l} E\left(X_i(t)X_j(t)X_k(t)X_l(t)\right) = O(nh_n^{1-\delta}). \quad (2.30)$$

Equations (2.22), (2.24), (2.25), (2.28), and (2.30) imply that

$$(nh_n)^{-4}Var\left(\int_0^{\tau}\left(\sum_{i=1}^{n}X_i(t)\right)^2 w(t)dt\right) = o(n^{-1}h_n^4). \quad (2.31)$$

Finally, the result follows from (2.14), (2.20), and (2.31).

The next lemma exhibits an asymptotic expansion for Q_{n2}.

Lemma 2.3. *Under the assumptions of Theorem* 1.1, *we have*

$$Q_{n2} = (nh_n)^{-1}\left(\int K^2(u)du\right)\int \frac{f(t)}{\overline{G}(t)}w(t)dt + o_p(n^{-1/2}h_n^2). \quad (2.32)$$

Proof. Equation (2.4), in conjunction with simple computations, implies

$$Q_{n2} = D_{n1} + D_{n2} + D_{n3} + D_{n4} + D_{n5} + D_{n6}, \quad (2.33)$$

where

$$D_{n1} = \int_0^{\tau}(f_n^*(t) - \overline{f}_n(t))^2 w(t)dt,$$

$$D_{n2} = \int_0^{\tau}\left[h_n^{-1}\int \frac{\overline{H}(x) - \overline{Y}_n(x)}{\overline{Y}_n(x)\overline{G}(x)}K\left(\frac{t-x}{h_n}\right)d\overline{N}_n(x)\right]^2 w(t)dt,$$

$$D_{n3} = \int_0^{\tau}\left[h_n^{-1}\int R_n(x)dK\left(\frac{t-x}{h_n}\right)\right]^2 w(t)dt,$$

$$D_{n4} = 2\int_0^{\tau}[f_n^*(t) - \overline{f}_n(t)]\left[h_n^{-1}\int R_n(x)dK\left(\frac{t-x}{h_n}\right)\right]w(t)dt,$$

$$D_{n5} = 2\int_0^{\tau}[f_n^*(t) - \overline{f}_n(t)]\left[h_n^{-1}\int \frac{\overline{H}(x) - \overline{Y}_n(x)}{\overline{Y}_n(x)\overline{G}(x)}K\left(\frac{t-x}{h_n}\right)d\overline{N}_n(x)\right]w(t)dt,$$

$$D_{n6} = 2\int_0^{\tau}\left[h_n^{-1}\int R_n(x)dK\left(\frac{t-x}{h_n}\right)\right]\left[h_n^{-1}\int \frac{\overline{H}(x) - \overline{Y}_n(x)}{\overline{Y}_n(x)\overline{G}(x)}K\left(\frac{t-x}{h_n}\right)d\overline{N}_n(x)\right]w(t)dt.$$

Lemma 2.2 implies

$$D_{n1} = (nh_n)^{-1}\left(\int K^2(u)du\right)\left(\int \frac{f(x)}{\overline{G}(x)}w(x)dx\right) + o_p(n^{-1/2}h_n^2). \tag{2.34}$$

Applying Theorem 1 in Cai (1998a), we can get easily that

$$\begin{aligned} |D_{n2}| &\leq \left(\sup_{x\geq 0}|\overline{H}(x) - \overline{Y}_n(x)|\right)^2 \int_0^\tau \left[h_n^{-1}\int K\left(\frac{t-x}{h_n}\right)\frac{d\overline{N}_n(x)}{\overline{Y}_n(x)\overline{G}(x)}\right]^2 w(t)dt \\ &= o(n^{-1/2}h_n^2) \quad a.s. \end{aligned} \tag{2.35}$$

Equation (2.5) implies

$$\begin{aligned} |D_{n3}| &\leq \left(\sup_{x\geq 0}|R_n(x)|\right)^2 \int_0^\tau \left[h_n^{-1}\int dK\left(\frac{t-x}{h_n}\right)\right]^2 w(t)dt \\ &= o(n^{-1/2}h_n^2) \quad a.s., \end{aligned} \tag{2.36}$$

$$|D_{n4}| \leq 2|D_{n1}|^{1/2}|D_{n3}|^{1/2} = o_p(n^{-1/2}h_n^2) \quad a.s. \tag{2.37}$$

Likewise, applying Theorem 1 in Cai (1998a), we obtain

$$|D_{n5}| \leq 2|D_{n1}|^{1/2}|D_{n2}|^{1/2} = o_p(n^{-1/2}h_n^2) \quad a.s., \tag{2.38}$$

$$|D_{n6}| \leq 2|D_{n2}|^{1/2}|D_{n3}|^{1/2} = o_p(n^{-1/2}h_n^2) \quad a.s. \tag{2.39}$$

The result follows from (2.33)–(2.39).

Proof of Theorem 1.1. By expanding the square in (1.4), we have

$$ISE(f_n(t)) = Q_{n1} + Q_{n2} + \int_0^\tau (\overline{f}_n(t) - f(t))^2 w(t)dt,$$

where Q_{n1} and Q_{n2} are defined in (2.1) and (2.2). Applying Lemmas 2.1 and 2.3, we obtain the result.

Proof of Corollary 1.1. Let

$$\varepsilon_n(t) = \frac{\sqrt{f_n(t)} - \sqrt{f(t)}}{\sqrt{f_n(t)} + \sqrt{f(t)}}.$$

It follows from (2.4) that

$$\begin{aligned} \sup_{0\leq t\leq\tau}|\varepsilon_n(t)| &= \sup_{0\leq t\leq\tau}\frac{|f_n(t) - f(t)|}{(\sqrt{f_n(t)} + \sqrt{f(t)})^2} \\ &\leq \sup_{0\leq t\leq\tau}\frac{|f_n(t) - \bar{f}_n(t)|}{(\sqrt{f_n(t)} + \sqrt{f(t)})^2} + \sup_{0\leq t\leq\tau}\frac{|\bar{f}_n(t) - f(t)|}{(\sqrt{f_n(t)} + \sqrt{f(t)})^2} \\ &\leq \sup_{0\leq t\leq\tau}\frac{h_n^{-1}|\int(F_n(x) - F(x))dK(\frac{t-x}{h_n})|}{f(t)} + \sup_{0\leq t\leq\tau}\frac{|\bar{f}_n(t) - f(t)|}{f(t)} \end{aligned}$$

$$\leq \left(\sup_{x\geq 0} |\widehat{\Lambda}_n(x) - \Lambda(x)| \right) \sup_{0\leq t\leq \tau} \frac{h_n^{-1} |\int \overline{F}(x) dK\left(\frac{t-x}{h_n}\right)|}{f(t)} + O\left(\frac{a_n^2}{h_n}\right) + O(h_n^2)$$

$$\leq O\left(\frac{a_n}{h_n}\right) + O\left(\frac{a_n^2}{h_n}\right) + O(h_n^2) = o(1) \quad a.s.$$

Therefore,

$$\begin{aligned} HD(f_n) &= \int_0^\tau \frac{(f_n(t) - f(t))^2}{4f(t)} dt + \int_0^\tau \frac{\varepsilon_n^2(t) - 2\varepsilon_n(t)}{4f(t)} (f_n(t) - f(t))^2 dt \\ &= \int_0^\tau \frac{(f_n(t) - f(t))^2}{4f(t)} dt + o(1) \int_0^\tau \frac{(f_n(t) - f(t))^2}{4f(t)} dt. \end{aligned} \tag{2.40}$$

Equation (2.40), in conjunction with the main theorem, in the case that the weight function is $\frac{1}{4f(t)}$, completes the proof.

References

Bickel, P., Rosenblatt, M. (1973). On some global measures of the deviations of density function estimates. *Ann. Statist.* 1:1071–1095.

Blum, J. R., Susarla, V. (1980). Maximal deviation theory of density and failure rate function estimates based on censored data. *J. Multivariate Anal.* 5:213–222.

Burke, M. D., Horváth, L. (1984). Density and failure rate estimation in a competing risks model. *Sankhyā Ser. A* 64:135–154.

Cai, Z. (1998a). Asymptotic properties of Kaplan–Meier estimator for censored dependent data. *Statist. Probab. Lett.* 37:381–389.

Cai, Z. (1998b). Kernel density and hazard rate estimation for censored dependent data. *J. Multivariate Anal.* 67:23–34.

Csörgő, M., Horváth, L. (1988). Central limit theorem for L_p norms of density estimators. *Probab. Theor. Rel. Fields* 80:269–291.

Csörgő, M., Gombay, E., Horváth, L. (1991). Central limit theorems for L_p distances of kernel estimators of densities under random censorship. *Ann. Statist.* 19:1813–1831.

Földes, A., Rejtö, L., Winter, B. B. (1981). Strong consistancy properties of nonparametric estimators for randomly censored data, II: Estimation of density and failure rate. *Period. Math. Hungar.* 12:15–29.

Gill, R. (1983). Large sample behavior of the product-limit estimator on the whole line. *Ann. Statist.* 11:49–58.

Hall, P. (1984). Central limit theorem for integrated square error of multivariate nonparametric density estimators. *J. Multivariate Anal.* 14:1–16.

Ibragimov, I. A., Linnik, Yu. V. (1971). *Independent and Stationary Sequences of Random Variables*. Groningen: Walters-Noordhoff.

Kim, T. Y. (1994). Moment bounds for non-stationary dependent sequences. *J. Appl. Probab.* 31:731–742.

Masry, E. (1986). Recursive probability density estimation for weakly dependent stationary processes. *IEEE Trans. Inform. Theor.* 32:254–267.

Rosenblatt, M. (1985). *Stationary Sequences and Random Fields*. Boston: Birkhäuser.

Roussas, G. G. (1990). Asymptotic normality of the Kernel estimate under dependence conditions: Application to hazard rate. *J. Statist. Plann. Inf.* 25:81–104.

Sun, L., Zheng, Z. (1999). The asymptotics of the integrated square error for the kernel hazard rate estimators with left truncated and right censored data. *Syst. Sci. Mathemat. Sci.* 12:251–262.

Yang, S. (1993). A central limit theorem for the integrated square error of the kernel density estimators with randomly censored data. *J. Statist. Plann. Infer.* 37:127–143.

Zhang, B. (1996). Some asymptotic results for kernel density estimation under random censorship. *Bernoulli* 2(2):183–198.

Zhang, B. (1998). A note on the integrated square errors of kernel density estimators under random censorship. *Stoch. Proc. Appl.* 75:225–234.