Statistica Sinica **8**(1998), 571-587

THE EDGEWORTH EXPANSIONS AND SMOOTHED BOOTSTRAP APPROXIMATION FOR THE STUDENTIZED KAPLAN-MEIER ESTIMATOR

Qihua Wang and Zhongguo Zheng

Peking University

Abstract: In this paper, the asymptotic accuracies of the one-term Edgeworth expansions and the smoothed bootstrap approximation for the studentized Kaplan-Meier (KM) estimator are investigated, respectively. It is shown that the Edgeworth expansions and the smoothed bootstrap approximation are asymptotically close to the exact distribution of the KM estimator with a remainder $o(n^{-\frac{1}{2}})$ under some mild conditions. A simulation study has been used to compare the asymptotic performance of the Edgeworth expansion with that of smoothed bootstrap and unsmoothed bootstrap.

Key words and phrases: Asymptotic normality, asymptotic representation, asymptotic variance, Edgeworth expansion, Jackknife estimator, studentized Kaplan-Meier estimator.

1. Introduction

Let $T_1, \ldots, T_n$ be nonnegative independent random variables with common continuous distribution function F. In the right random censoring model, associated with each T_i there is an independent nonnegative censoring time C_i, and here $C_1, \ldots, C_n$ are assumed to be i.i.d. random variables with continuous distribution function G. The observations in this model are the pairs (Z_i, δ_i), where $Z_i = min(T_i, C_i)$ and $\delta_i = I[T_i \le C_i], i = 1, \ldots, n$. Clearly, the $Z_1, \ldots, Z_n$ are i.i.d. with continuous distribution function $H = 1 - (1 - F)(1 - G)$. Let $Z_{(1)} \le Z_{(2)} \le \cdots \le Z_{(n)}$ be the ordered $Z_i's$ and $\delta_{(i)}$ be the δ corresponding to $Z_{(i)}$. The Kaplan-Meier (KM) estimator of the survival function $\bar{F}(t) = 1 - F(t)$ is defined as

$$\hat{\bar{F}}_n(t) = \prod_{i:Z_{(i)} \le t} \left(\frac{n-i}{n-i+1} \right)^{\delta_{(i)}} I[Z_{(n)} > t]. \tag{1.1}$$

The asymptotic properties of $\hat{\bar{F}}_n$ have been investigated by many authors, (see, for example, Breslow and Crowley (1974), Földes and Rejtö (1981), Földes (1981), Phadia and Van Ryzin (1980), Wang (1987), Zheng (1989)). Csörgö and Horváth (1983) proved its strong uniform consistency and obtained the convergence rate. Lo and Singh (1986) established the i.i.d. representation of the KM estimator with remainders, and gave a bootstrap version of the representation. Recently Chang and Rao (1989) and Chang (1990) derived a Berry-Essen bound and established an Edgeworth expression with a remainder $o(n^{-\frac{1}{2}})$, respectively, for the standard Kaplan-Meier estimator $\sqrt{n}(\hat{\bar{F}}_n - \bar{F})/\sigma$, where σ^2 is the asymptotic variance of $\sqrt{n}(\hat{\bar{F}}_n - \bar{F})$.

In the present paper, we pursue Edgeworth expansion of the distribution of the studentized KM estimator $\sqrt{n}(\hat{\bar{F}}_n - \bar{F})/\hat{\sigma}_{nJ}$, where $\hat{\sigma}^2_{nJ}$ is the Jackknife estimator of the variance of $\sqrt{n}\hat{\bar{F}}$ derived by Singh and Liu (1990). This result is used to investigate the asymptotic accuracy of the estimated Edgeworth expansion and the smoothed bootstrap approximation. The main results are stated in section 2 and are proved in section 4. A simulation study was performed to compare their performances in Section 3. One will see that the results can be used to construct point-wise confidence intervals of the survival function $\bar{F}(t)$.

Let $\bar{Q} = 1-Q, Q^{-1} = \frac{1}{Q}$ for any function Q, and let us define $\bar{H} = \bar{F}\,\bar{G}, \bar{H}_n = n^{-1}\sum_{i=1}^{n} I[Z_i > t], \widetilde{H}_1 = P(Z_1 > t, \delta_1 = 1), \widetilde{H}_{n1}(t) = n^{-1}\sum_{i=1}^{n} I[Z_i > t, \delta_i = 1], \widetilde{H}_2(t) = P(Z_1 > t, \delta_1 = 0), \widetilde{H}_{n2} = n^{-1}\sum_{i=1}^{n} I[Z_i > t, \delta_i = 0], \sigma_0^2 = \bar{F}^{-2}\sigma^2$. By Chang (1990) the asymptotic variance σ^2 is given by $\sigma^2 = -\bar{F}^2(t)\int_0^t \bar{H}^{-2}\, d\widetilde{H}_1$.

2. The Main Results

Throughout this paper, we suppose the supports of F and G are $[0,\infty)$.

Let $\hat{\bar{F}}_n^{(-i)}$ be the Kaplan-Meier estimator based on $\{(Z_1,\delta_1),\ldots,(Z_n,\delta_n)\} - \{(Z_i,\delta_i)\}$. The stochastic process of pseudo-values can be defined as

$$J_{ni} = n\hat{\bar{F}}_n(t) - (n-1)\hat{\bar{F}}_n^{(-i)}(t), \qquad t \geq 0, 1 \leq i \leq n. \tag{2.1}$$

By Singh and Liu (1990), the Jackknife estimate for σ^2 is

$$\hat{\sigma}^2_{nJ} = n^{-1}\sum_{i=1}^{n}(J_{ni} - \bar{J}_n)^2, \tag{2.2}$$

where $\bar{J}_n = n^{-1}\sum_{i=1}^{n} J_{ni}$. The Edgeworth expansions for the studentized KM estimator $\sqrt{n}(\hat{\bar{F}}_n - \bar{F})/\hat{\sigma}_{nJ}$ are established in the following Theorem 1 and Theorem 2.

Theorem 1. *Assuming that F and G are continuous, we have, as $n \to +\infty$,*

$$\sup_x |P\left(\sqrt{n}\hat{\sigma}_{nJ}^{-1}(\hat{\bar{F}}_n(t) - \bar{F}(t)) \leq x\right) - K_n(x)| = o(n^{-\frac{1}{2}}), \tag{2.3}$$

for any $t > 0$, where

$$K_n(x) = \Phi(x) - \frac{\kappa_3}{6} n^{-\frac{1}{2}}\phi(x)(x^2 + \frac{1}{2}), \tag{2.4}$$

$$\kappa_3 = -2\sigma_0^{-3}\left(\int_0^t \bar{H}^{-3}\, d\widetilde{H}_1 + \frac{3}{2}\sigma_0^4\right),$$

$\Phi(x)$ *is the standard normal distribution function,and* $\phi(x)$ *is its probability density.*

The proof of Theorem 1 (cf. Section 4) depends heavily on the results of Bickel, Götze and Van Zwet (1986) and Chang(1990).

When F is completely unknown, one does not know κ_3 and σ_0 appearing in the expansion (2.3). We must estimate κ_3 and σ_0^2 from the observation $(Z_i, \delta_i), i = 1, \ldots, n$. One way of doing this is to replace $\bar{H}$ and $\widetilde{H}_1$ in κ_3 and σ_0^2 by $\bar{H}_n$ and $\widetilde{H}_{n1}$. The estimates of κ_3 and σ_0^2 are given by

$$\begin{aligned}\kappa_{3n} &= -2\sigma_{0n}^{-3}\Big(\int_0^t \bar{H}_n^{-3}\, d\widetilde{H}_{n1} + \frac{3}{2}\sigma_{0n}^4\Big)\\ \sigma_{0n}^2 &= \int_0^t \bar{H}_n^{-2}\, d\widetilde{H}_{n1}.\end{aligned}$$

Denote

$$\tilde{K}_n(x) = \Phi(x) - \frac{\kappa_{3n}}{6} n^{-\frac{1}{2}}\phi(x)(x^2 + \frac{1}{2}).$$

In our second theorem, we shall show that we may replace $K_n(x)$ in (2.3) by $\tilde{K}_n(x)$ without affecting the asymptotic accuracy of the expansion.

Theorem 2. *Suppose that the assumptions of Theorem* 1 *are satisfied. Then, with probability* 1, *as* $n \to +\infty$

$$\sup_x |P(\sqrt{n}\hat{\sigma}_{nJ}^{-1}(\hat{\bar{F}}_n(t) - \bar{F}(t)) \le x) - \tilde{K}_n(x)| = o(n^{-\frac{1}{2}}),$$

for any $t > 0$.

Another way to obtain an approximation to the distribution of studentized KM estimator is to employ the bootstrap method. Efron (1981) introduced two different ways of bootstrapping $\hat{\bar{F}}$. But, we do not apply his method directly since the unsmoothed bootstrap by Efron (1981) can not be proved to be of the same asymptotic accuracy as the Edgeworth expansions for technical reasons. Motivated by Efron (1979)'s well known smoothed bootstrap method in the uncensored case, we introduce the smoothed bootstrap to the censoring case. Now let us state the method as follows. We construct the KM estimator of G

$$\widehat{G}_n(t) = 1 - \prod_{i:Z_{(i)}\le t}\Big(\frac{n-i}{n-i+1}\Big)^{1-\delta_{(i)}} I[Z_{(n)} > t].$$

Let

$$\begin{aligned}\widehat{F}_n(t) &= 1 - \hat{\bar{F}}_n(t),\\ \widehat{f}_{nh}(t) &= \frac{1}{h_n}\int_0^{+\infty} k\Big(\frac{t-s}{h_n}\Big)\, d\widehat{F}_n,\end{aligned}$$

$$\widehat{g}_{nh}(t)=\frac{1}{h_n}\int_0^{+\infty}k\Big(\frac{t-s}{h_n}\Big)\,d\widehat{G}_n,$$

where h_n is a constant sequence satisfying $h_n\to 0$ and $k(\cdot)$ is a probability density kernel function. The estimated populations $\widehat{F}_{nh}$ and $\widehat{G}_{nh}$, corresponding to $\widehat{f}_{nh}$ and $\widehat{g}_{nh}$, are then treated as the true survival and censoring populations for the purpose of drawing the second stage samples. Conditionally, given $(Z_1,\delta_1),\ldots,(Z_n,\delta_n)$, let $(Z_1^*,\delta_1^*),\ldots,(Z_n^*,\delta_n^*)$ be n independent r.v.'s with $Z_i^*=T_i^*\wedge C_i^*$, $\delta_i^*=I[T_i^*\le C_i^*]$ where $T_1^*,\ldots,T_n^*$ follow the distribution $\widehat{F}_{nh}$ independently, and $C_1^*,\ldots,C_n^*$ follow the distribution $\widehat{G}_{nh}$ independently. The smoothed bootstrapping studentized KM estimator is given by $\sqrt{n}\widehat{\sigma}^*_{nJ}(\widehat{\bar{F}}{}^*_n-\widehat{\bar{F}}_{nh})$, where $\widehat{\sigma}^*_{nJ}$ and $\widehat{\bar{F}}{}^*_n$ are obtained from $\widehat{\sigma}_{nJ}$ and $\widehat{\bar{F}}_n$ simply by replacing (Z_i,δ_i) by the (Z_i^*,δ_i^*)in the formula (2.2) and (1.1), $i=1,\ldots,n$. Let us denote by P^* the bootstrap probability below.

Theorem 3. *Suppose that T_i and C_i have continuous and bounded probability density f and g with respect to the Lebesgue measure on R^1. If $k(\cdot)$ is a probability density kernel function on $[0,+\infty)$ satisfying $\int_0^\infty uk(u)\,du<+\infty$, then, with probability* 1*, we have*

$$\sup_x|P(\sqrt{n}\widehat{\sigma}_{nJ}^{-1}(\widehat{\bar{F}}_n(t)-\bar{F}(t))\le x)-P^*(\sqrt{n}\widehat{\sigma}_{nJ}^{*-1}(\widehat{\bar{F}}{}^*_n(t)-\widehat{\bar{F}}_{nh}(t))\le x)|=o(n^{-\frac{1}{2}}),$$

for any $t>0$.

Theorem 2 and Theorem 3 tell us that the Edgeworth expansion estimate $\tilde{K}_n(x)$ and the smoothed bootstrap distribution function (df) are asymptotically close to the exact df of studentized KM estimator with rate $o(n^{-\frac{1}{2}})$. The same result is perhaps valid under proper conditions for the unsmoothed bootstrap.

3. Simulation Results

In this section, the three methods, including the unsmoothed bootstrap, were investigated by simulation. All the simulation was carried out under the assumptions of the model that $F(t)=1-e^{-t},G(t)=1-e^{-\frac{t}{2}},,t\ge 0$. The sample size n was taken to be 10, 40 and 60 respectively. For the smoothed bootstrap, h_n was taken to be $\frac{0.1}{\sqrt{n}}$ and $\frac{0.01}{\sqrt{n}}$ respectively, and $k(x)$ to be $e^{-x}I[x\ge 0]$. For each combination of these parameters, we calculated the distribution curves of studentized KM estimator, unsmoothed bootstrap and the smoothed bootstrap counterparts at $t=1$ by simulation, respectively, and the curves of the Edgeworth expansion were obtained by calculating $K_n(x)$ directly. All these curves are given in the following Figures.

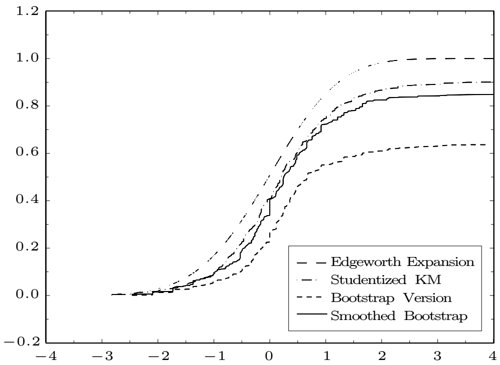

Figure 1. $n = 10$, $h_n = 0.1/\sqrt{n}$, $t = 1$.

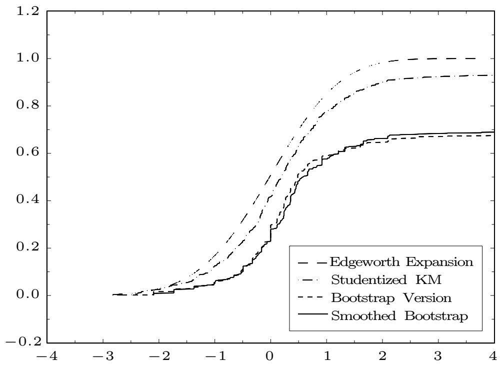

Figure 2. $n = 10$, $h_n = 0.01/\sqrt{n}$, $t = 1$.

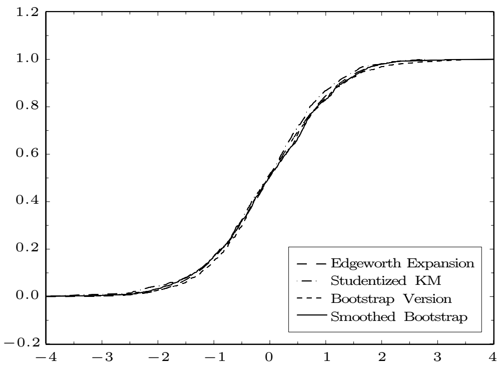

Figure 3. $n = 40$, $h_n = 0.1/\sqrt{n}$, $t = 1$.

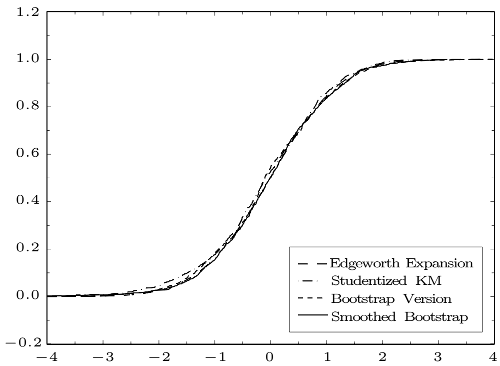

Figure 4. $n = 40$, $h_n = 0.01/\sqrt{n}$, $t = 1$.

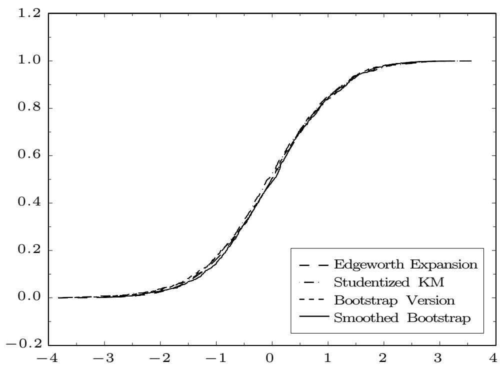

Figure 5. $n = 60$, $h_n = 0.1/\sqrt{n}$, $t = 1$.

From these figures, we see that all the curves are closer to each other as n increases, and that the curve of the smoothed bootstrap is also closer to that of unsmoothed bootstrap when h_n is chosen smaller. For small sample size ($n = 10$), Figue 1 and Figure 2 show that all the three methods perform poorly. But relatively speaking, the smoothed bootstrap with h_n chosen appropriately may perform better than the other two methods. Note that $\widehat{\sigma}_{nJk}$ and its smoothed bootstrap and unsmoothed bootstrap versions take zero with positive probability for small sample size ($n = 10$). This is why we see that the distribution curves of the studentized KM estimator and its smoothed and unsmoothed bootstrap versions in Figure 1 and Figure 2 are far lower than those of the Edgeworth expansion even at the points at which the function values of the curves of the Edgeworth expansion are close to 1. Such phenomenon can not be seen in Figure 3, Figure 4 and Figure 5 because $\widehat{\sigma}_{nJk}$ and its bootstrap versions, which tend to $\sigma > 0$ with probability 1, take zero with negligible probability for $n = 40$ or $n = 60$. Indeed, this is also the one of the reasons that smoothed bootstrap and unsmoooothed bootstrap perform better in Figure 3, Figure 4 and Figure 5 than in Figure 1 and Figure 2. For large sample size($n = 40$ or $n = 60$), closeness of these curves in Figure 4 and Figure 5 vindicates our results in section 2, and shows that the unsmoothed bootstrap perhaps works well for large sample in some cases, for example, in the case that F and G are sufficiently smooth.

It is worthwhile pointing out that it is important and diffcult to choose appropriate h_n for the smoothed bootstrap. In what follows, we only discuss the problem roughly. If h_n is chosen too large, for example $h_n = n^{-\frac{1}{k}}$ for some large $k > 0$, the smoothed bootstrap sample will contain less information provided by $\widehat{F}_n$ and $\widehat{G}_n$ and much more information provided by kernel density function $k(\cdot)$ which is used in the smoothed bootstrap. In this case, the man-made factor $k(\cdot)$ will affect the asymptotic performances of the smoothed bootstreap greatly. If we choose h_n close to zero, for example $h_n = n^{-k}$ for some large $k > 0$, under some conditions the difference between the smoothed bootstrap and the unsmoothed bootstrap disappears. This implies that if the unsmoothed bootstrap has good asymptotic accuracy we may choose small h_n for the smoothed bootstrap, otherwise we choose moderate h_n (not too large and not too small). For example, from Figure 1 and Figure 2 it is seen that unsmoothed bootstrap is considerably worse, and in Figure 2 the curve of the smoothed bootstrap is closer to that of unsmoothed bootstrap and hence further from that of the studentized KM estimator. Therefore, we should choose a moderate h_n in the case of small sample size ($n = 10$). From Figure 3, Figure 4 and Figure 5, we see that the unsmoothed bootstrap performs very well and the curves of the smoothed bootstrap are closer to those of the studentized KM estimator and the unsmoothed bootstrap in Figure 4 than in Figure 3. Hence, we may choose small h_n for the smoothed bootstrap in the case of large samples. Though the curve of the smoothed bootstrap departs from that of the studentized KM estimator slightly in Figure 3, in Figure 5 they are close to each other very much because n increases. This shows that if h_n is chosen larger, the smoothed bootstrap also performs fairly well as long as we increase n appropriately.

4. Proof of Theorems

For convenience, we adopt a convention that C may represent any constants needed, even if C's appear in the same formula they may represent different constants.

As pointed out by Chang (1990), for any sample size n $\hat{\bar{F}}_n(t) = 0$ and $\log \hat{\bar{F}}_n(t) = -\infty$ with positive probability. To overcome this diffculty, Chang partitioned the sample space Ω into two parts: $\Omega_0^{(n)}$ and $\Omega_1^{(n)} = \Omega - \Omega_0^{(n)}$, where $\Omega_0^{(n)}$ is defined as in Chang (1990), and showed that $P(\Omega_1^{(n)}) = \mathrm{o}(n^{-k})$ for any $k > 0$, and for $\omega \in \Omega_0^{(n)}$,$-\infty < \log \hat{\bar{F}}_n(t) < +\infty$. This implies that we can focus our discussion on the sub-sample space $\Omega_0^{(n)}$

In what follows we engage in our studies on $\Omega_0^{(n)}$ as in Chang (1990).

Lemma 1. *Assuming F and G are continuous, one can express*

$$\log \hat{\bar{F}}_n(t) - \log \bar{F}(t) = n^{-1} \sum_{i=1}^{n} \varphi(Z_i, \delta_i; t) + r_{n1}(t), \tag{4.1}$$

$$\hat{\bar{F}}_n(t) - \bar{F}(t) = n^{-1} \sum_{i=1}^{n} \bar{F}(t) \varphi(Z_i, \delta_i; t) + r_{n2}(t), \tag{4.2}$$

for any $t > 0$, where

$$\varphi(Z, \delta; t) = -\Big(\int_0^{z \wedge t} [\bar{H}(s)]^{-2} \, d\widetilde{H}_1(s) + \bar{H}^{-1}(z) I[z \le t, \delta = 1] \Big)$$

and $r_{n1}(t), r_{n2}(t)$ satisfy that

$$P(|r_{n1}(t)| > \epsilon_n) \le C(n\epsilon_n)^{-k}, \tag{4.3}$$

$$P(|r_{n2}(t)| > \epsilon_n) \le C(n\epsilon_n)^{-\frac{k}{2}}, \tag{4.4}$$

for any $\epsilon_n > 0, k > 0$ and $t > 0$.

Proof. By Lo and Singh (1986), we have

$$\begin{aligned} P(|r_{n1}| > \epsilon_n) \le C\, P\Big(\big| \int_0^t (\bar{H}_n - \bar{H}) \bar{H}^{-2} \, d(\widetilde{H}_{n1} - \widetilde{H}_1) \big| > \frac{\epsilon_n}{2} \Big) \\ + P\Big(\sup_{0 \le s \le t} |\bar{H}_n - \bar{H}| > \frac{1}{2} \bar{H}(t) \Big) \\ + C P\Big(\big| \log \hat{\bar{F}}_n(t) - \int_0^t [\bar{H}_n(s)]^{-1} d\widetilde{H}_{n1}(s) \big| > \frac{\epsilon_n}{2} \Big) \end{aligned} \tag{4.5}$$

for any $t > 0$ and $\epsilon_n > 0$ as n is sufficiently large.

Using the Tchebyschev inequality and the Lemma in the appendix of Chang (1989), it follows that

$$P(| \triangle_n | > \epsilon_n) \leq C(n\epsilon_n)^{-k}, \tag{4.6}$$

for any $t > 0, k > 0$ and $\epsilon_n > 0$. It is well known that

$$P\Big(\sup_{0\leq s\leq t} |\bar{H}_n - \bar{H}| > \frac{1}{2}\bar{H}(t)\Big) \leq \exp\{-\frac{1}{2}n\bar{H}^2(t)\}. \tag{4.7}$$

Clearly

$$\begin{aligned}\triangle_n' &\stackrel{\triangle}{=} \log \hat{\bar{F}}_n - \int_0^t [\bar{H}_n(s)]^{-1} d\widetilde{H}_{n1}(s) \\ &= \sum_{i=1}^n I[Z_i \leq t, \delta_i = 1]\Big\{ \log\Big(1 - \frac{1}{1+n\bar{H}_n(Z_i)}\Big) + \frac{1}{n\bar{H}_n(Z_i)}\Big\}. \end{aligned} \tag{4.8}$$

Let $A_{n\epsilon} = \cap_{i=1}^n \{\{\omega : n(1+\bar{H}_n(Z_i))I[Z_i \leq t] \geq 2\} \cup \{Z_i > t\}\}$, and $A_{n\epsilon}^c$ be the complement of $A_{n\epsilon}$. By (4.8) and the inequality $|\log(1-x)+x| \leq x^2$, for $0 \leq x \leq \frac{1}{2}$, we have on $A_{n\epsilon}$

$$\triangle_n' \leq 2\sum_{i=1}^n I[Z_i \leq t, \delta_i = 1]\frac{1}{(n\bar{H}_n(Z_i))^2}, \tag{4.9}$$

and

$$P(A_{n\epsilon}^c) \leq \sum_{i=1}^n P((1+n\bar{H}_n(Z_i))I[Z_i \leq t] = 1) \leq ne^{-(n-1)\bar{H}(t)}. \tag{4.10}$$

The inequality (4.3) is trival for any $k > 0$ if $n\epsilon_n < 16\bar{H}^{-2}(t)$. Hence, in the following we need only consider the case that $n\epsilon_n \geq 16\bar{H}^{-2}(t)$. In this case, we have

$$\begin{aligned} P(| \triangle_n' | > \frac{\epsilon_n}{2}) &\leq P\Big(\frac{1}{n}\sum_{i=1}^n I[Z_i \leq t, \delta_i = 1] > \frac{n\epsilon_n\bar{H}^2(t)}{16}\Big) \\ &\qquad + P(\bar{H}_n(t) < \frac{1}{2}\bar{H}(t)) + P(A_{n\epsilon}^c) \\ &\leq e^{-n\bar{H}^2(t)} + ne^{-(n-1)\bar{H}(t)} \end{aligned} \tag{4.11}$$

by (4.9) and (4.10).

Combining(4.5), (4.6), (4.7) and (4.11), (4.3) is proved. To prove (4.4), at first we prove that

$$P(|\log \hat{\bar{F}}_n(t) - \log \bar{F}(t)| > \epsilon_n) \leq C(n^{\frac{1}{2}}\epsilon_n)^{-k}, \tag{4.12}$$

for any $t > 0, k > 0$ and $\epsilon_n > 0$. It is easy to prove (4.12) by (4.1), (4.3) and the Tchebyschev inequality and the Dharmadhikari-Jodgeo (D-J) inequality (e.g., see, Rao (1987)).

By Taylor's formula for $\widehat{F}_n(t) - F(t)$, (4.1) and (4.2), it is easily seen that

$$P(|r_{n2}(t)| > \epsilon_n) \leq P(|r_{n1}(t)| > \frac{\epsilon_n}{2\bar{F}}) + P\Big(\frac{\bar{F}(t)e^{\theta(\log \widehat{\bar{F}}_n - \log \bar{F})}}{2}(\log \widehat{\bar{F}}_n(t) - \log \bar{F}(t))^2 > \frac{\epsilon_n}{2}\Big). \quad (4.13)$$

Let P_n represent the second term in the right hand side of (4.13). From (4.12), we have

$$P_n \leq P(|\log \widehat{\bar{F}}_n - \log \bar{F}| > \epsilon_n{}^{\frac{1}{2}}) + P(|\log \widehat{\bar{F}}_n - \log \bar{F}| > -\log \bar{F}) \leq C(n\epsilon_n)^{-\frac{k}{2}}. \quad (4.14)$$

Hence, (4.4) is proved by combining (4.13), (4.14) and (4.3).

For the simplicity, let us denote $\varphi(z, \delta; t)$ by $\varphi(z, \delta)$ and $r_{ni}(t)$ by r_{ni}, where $i = 1, 2$.

Lemma 2. *Under the assumptions of Lemma* 1*, we have*

$$\widehat{\sigma}^2_{nJ} = n^{-1}\sum_{i=1}^{n} \bar{F}^2(\varphi(Z_i, \delta_i) - n^{-1}\sum_{i=1}^{n}\varphi(Z_i, \delta_i))^2 + R_n(t), \quad (4.15)$$

with

$$P(|R_n(t)| > \epsilon_n) \leq Cn^{1-\frac{k}{2}}\epsilon_n^{-k}, \quad (4.16)$$

for any $t > 0, k > 0$ and $\epsilon_n > 0$.

The proof of Lemma 2 is based on a standard type of argument, and hence is omitted.

Lemma 3. *Under the assumptions of Lemma* 1*, we have $P(|\widehat{\sigma}^2_{nJ} - \sigma^2| > \epsilon_n) \leq Cn^{1-\frac{k}{2}}\epsilon_n{}^{-k}$, for any $t > 0, k > 0$ and $\epsilon_n > 0$.*

By Lemma 2, it is easy to prove Lemma 3.

Lemma 4. *If F and G are continuous, one can express $\log \widehat{\bar{F}}_n - \log \bar{F} = U_{n0} - \frac{1}{2}n^{-1}\sigma_0^2 + \triangle_{n1}$ with $P(\sqrt{n}| \triangle_{n1} | > n^{-\frac{1}{2}}\log n^{-1}) = o(n^{-\frac{1}{2}})$, for any $t > 0$, where*

$$U_{n0} = n^{-2}\sum_{i<j} h_0(Z_i, \delta_i; Z_j, \delta_j),$$
$$h_0(Z_1, \delta_1; Z_2, \delta_2) = g_0(Z_1, \delta_1) + g_0(Z_2, \delta_2) + \psi_0(Z_1, \delta_1; Z_2, \delta_2),$$
$$g_0(Z, \delta) = -A_{10}(Z)\delta - A_{20}(Z \wedge t),$$

$$
\begin{aligned}
&A_{10}(s) = \bar{H}^{-1}(s)I[0 \le s \le t],\\
&A_{20}(s) = \int_0^s \bar{H}^{-2}\, d\widetilde{H}_1,\\
&\psi_0(Z_1,\delta_1;Z_2,\delta_2) = \eta(Z_1,\delta_1;Z_2,\delta_2) + \eta(Z_2,\delta_2;Z_1,\delta_1),\\
&\eta(Z_1,\delta_1;Z_2,\delta_2) = B_1(Z_1,Z_2)\delta_1 + B_2(Z_1,Z_2)\\
&\qquad\qquad\qquad\qquad -E[B_1(Z_1,Z_2)\delta_1 + B_2(Z_1,Z_2)|Z_1,\delta_1]\\
&B_1(s,u) = \bar{H}^{-2}(s)I[0 \le s \le t, s < u],\\
&B_2(s,u) = \int_0^{t\wedge s\wedge u} \bar{H}^{-3}\, d\widetilde{H}_1
\end{aligned}
$$

Proof. See Chang (1990).

Remark. $h_0(Z_1,\delta_1;Z_2,\delta_2)$, $g_0(Z_1,\delta_1)$ and $\psi_0(Z_1,\delta_1;Z_2,\delta_2)$ here are the same as $-h(Z_1,\delta_1;Z_2,\delta_2)$, $-g(Z_1,\delta_1)$ and $-\psi(Z_1,\delta_1;Z_2,\delta_2)$ in Chang (1990). In the following, we will often make use of the facts directly.

Lemma 5. *For any random variables X, Y and constant $a > 0$, there exists a constant α such that*

$$\sup_x |P(X+Y \le x) - K_n(x)| = \sup_x |P(X \le x) - K_n(x)| + \alpha a + P(|Y| > a),$$

where $K_n(x)$ is defined as in Theorem 1.

The proof of Lemma 5 is similar to that of Lemma 2 in Chang (1989), hence we omit it here.

Denote

$$U_n = n^{-2}\sum\sum\nolimits_{i<j} h(Z_i,\delta_i;Z_j,\delta_j) \tag{4.17}$$

where

$$
\begin{aligned}
h(Z_i,\delta_i;Z_j,\delta_j) = {}& h_0(Z_i,\delta_i;Z_j,\delta_j)\\
& -\frac{\varphi^2(Z_i,\delta_i)\varphi(Z_j,\delta_j) + \varphi(Z_i,\delta_i)\varphi^2(Z_j,\delta_j)}{2\sigma_0^2}\\
& +\frac{\varphi(Z_i,\delta_i) + \varphi(Z_j,\delta_j)}{2} + \varphi(Z_i,\delta_i)\varphi(Z_j,\delta_j).
\end{aligned} \tag{4.18}
$$

Lemma 6. *Let σ_n^2 be the variance of U_n; then as $n \to +\infty$ $\sup_x |P(\sqrt{n}\sigma_n^{-1}U_n \le x) - K_{n0}(x)| = o(n^{-\frac{1}{2}})$, where*

$$
\begin{aligned}
K_{n0}(x) &= \Phi(x) - \frac{\kappa_3}{6} n^{-\frac{1}{2}}\phi(x)(x^2-1),\\
\kappa_3 &= -2\sigma_0^{-3}\left(\int_0^t \bar{H}^{-3}\, d\widetilde{H}_1 + \frac{3}{2}\sigma_0^4\right).
\end{aligned}
$$

The same arguments as in the proof of Lemma 2 of Chang (1990) can be employed to prove Lemma 6.

Lemma 7. *Under the assumptions of Lemma* 6, *we have* $\sup_x |P(\sqrt{n}\sigma_0^{-1}U_n \leq x) - K_{n0}(x)| = o(n^{-\frac{1}{2}})$, *as* $n \longrightarrow +\infty$.

The same method as that of proving Lemma 3 in Chang (1990) can be used to prove Lemma 7 in virtue of Lemma 5 and Lemma 6.

Lemma 8. *Under the assumptions of Lemma* 6, *we have* $\sup_x |P(\sqrt{n}\sigma_0^{-1}U_n - \nu n^{-\frac{1}{2}} \leq x) - K_n(x)| = o(n^{-\frac{1}{2}})$, *as* $n \to +\infty$, *where* $K_n(x)$ *is defined as in Theorem* 1, *and* $\nu = \frac{1}{2}\sigma_0^{-3}(\int_0^t \bar{H}^{-3}\, d\widetilde{H}_1 + \frac{3}{2}\sigma_0^4)$.

The proof is similar to that of Theorem 2 in Chang (1990).

The proof of Theorem 1. In the following, we fix t such that $t > 0$. Let $Q(x) = \frac{1}{1+x} - (1-x)$. Then

$$\sqrt{n}\widehat{\sigma}_{nJ}^{-1}(\widehat{\bar{F}}_n - \bar{F}) = \sqrt{n}\sigma^{-1}(\widehat{\bar{F}}_n - \bar{F})(1 - \frac{\widehat{\sigma}_{nJ} - \sigma}{\sigma}) + \sqrt{n}\sigma^{-1}(\widehat{\bar{F}}_n - \bar{F})Q(\frac{\widehat{\sigma}_{nJ} - \sigma}{\sigma}) \overset{\triangle}{=} R_{1n} + R_{2n}. \quad (4.19)$$

Clearly

$$P(|R_{2n}| > (n\log n)^{-\frac{1}{2}}) \leq P(|\sqrt{n}\sigma^{-1}(\widehat{\bar{F}}_n - \bar{F}| > \log^{-\frac{1}{2}} n) + P(|Q(\frac{\widehat{\sigma}_{nJ} - \sigma}{\sigma})| > n^{-\frac{1}{2}}\log^{-1} n) \overset{\triangle}{=} R_{21n} + R_{22n}. \quad (4.20)$$

Denote

$$\Psi_n(x) = \Phi(x) - \frac{\widetilde{\kappa}_3}{6}n^{-\frac{1}{2}}\phi(x)(x^2 - 1),$$

where

$$\widetilde{\kappa}_3 = -\sigma_0{}^{-3}(-\int_0^t \bar{H}^{-3}\, d\widetilde{H}_1 + \frac{3}{2}\sigma_0^4) + 3\sigma_0.$$

It is easy to see that

$$R_{21n} \leq |P(\sqrt{n}\sigma^{-1}(\widehat{\bar{F}}_n - \bar{F}) > \log^{\frac{1}{2}} n) - (1 - \Psi_n(\log^{\frac{1}{2}} n))| + |P(\sqrt{n}\sigma^{-1}(\widehat{\bar{F}}_n - \bar{F}) < -\log^{-\frac{1}{2}} n) - \Psi_n(-\log^{\frac{1}{2}} n)| + |1 - \Psi_n(\log^{\frac{1}{2}} n)| + |\Psi_n(-\log^{\frac{1}{2}} n)|. \quad (4.21)$$

In virtue of Lemma 3 in Chang (1990) and Lemma 3 in page 49 of Chow and Teicher (1978), it follows that

$$R_{21n} \leq C(n\log n)^{-\frac{1}{2}}. \quad (4.22)$$

Using the inequality $|Q(x)| \le 2x^2$, for $|x| < \frac{1}{2}$, by Lemma 3 we have

$$R_{22n} \le P(|\widehat{\sigma}_{nJ}^2 - \sigma^2| > \frac{3}{2\sqrt{2}}\sigma^2 n^{-\frac{1}{4}} \log^{-\frac{1}{2}} n, |\frac{\widehat{\sigma}_{nJ} - \sigma}{\sigma}| \le \frac{1}{2}) \\ +P(|\widehat{\sigma}_{nJ}^2 - \sigma^2| > \frac{1}{2}\sigma^2) \le C((n\log n)^{-\frac{1}{2}} + n^{1-\frac{k}{4}} \log^{\frac{k}{2}} n). \quad (4.23)$$

Hence, by (4.20), (4.22), (4.23), it follows that

$$P(|R_{2n}| > (n\log n)^{-\frac{1}{2}}) \le C((n\log n)^{-\frac{1}{2}} + n^{1-\frac{k}{4}} \log^{\frac{k}{2}} n). \quad (4.24)$$

By Taylor's formula, we have

$$R_{1n} = \sqrt{n}\sigma^{-1}\bar{F}\Big[(\log \widehat{\bar{F}}_n - \log \bar{F})(1 - \frac{\widehat{\sigma}_{nJ} - \sigma}{\sigma}) + \frac{(\log \widehat{\bar{F}}_n - \log \bar{F})^2}{2}\Big] \\ -\frac{\sqrt{n}\bar{F}(\widehat{\sigma}_{nJ} - \sigma)(\log \widehat{\bar{F}}_n - \log \bar{F})^2}{2\sigma^2} \\ +\sqrt{n}\sigma^{-1}\bar{F}\frac{e^{\theta'(\log \widehat{\bar{F}}_n - \log \bar{F})}}{6}(\log \widehat{\bar{F}}_n - \log \bar{F})^3\Big(1 - \frac{\widehat{\sigma}_{nJ} - \sigma}{\sigma}\Big) \\ \stackrel{\triangle}{=} R_{11n} + R_{12n} + R_{13n}. \quad (4.25)$$

Using the same arguments as in the proof of (4.22) in virtue of Theorem 2 in Chang (1990), it can be proved that

$$P(\sqrt{n}\sigma^{-1}\bar{F}|(\log \widehat{\bar{F}}_n - \log \bar{F})| > \log^{\frac{1}{2}} n) \le C(n\log n)^{-\frac{1}{2}}. \quad (4.26)$$

Thus, by (4.12) and Lemma 3 we have

$$P(|R_{12n}| > (n\log n)^{-\frac{1}{2}}) \\ \le P(|\widehat{\sigma}_{nJ}^2 - \sigma^2| > \sigma n^{-\frac{1}{4}} \log^{-\frac{1}{2}} n) + P(|\log \widehat{\bar{F}}_n - \log \bar{F}| > \sigma n^{-\frac{1}{4}} \log^{-\frac{1}{2}} n) \\ +P(\widehat{\sigma}_{nJ} \le \frac{1}{2}\sigma) + P(|\sqrt{n}\sigma^{-1}\bar{F}(\log \widehat{\bar{F}}_n - \log \bar{F})| > \log^{\frac{1}{2}} n) \\ \le C\left(n^{1-\frac{k}{4}} \log^{\frac{k}{2}} n + (n\log n)^{-\frac{1}{2}}\right). \quad (4.27)$$

Furthermore, we have by (4.12) and Lemma 3

$$P(|R_{13n}| > (n\log n)^{-\frac{1}{2}}) \\ \le P\Big(|\log \widehat{\bar{F}}_n - \log \bar{F}| \le 1, \Big|\frac{\widehat{\sigma}_{nJ} - \sigma}{\sigma}\Big| \le \frac{1}{2}, |\log \widehat{\bar{F}}_n - \log \bar{F}|^3 > \frac{4\sigma}{e}(n\log n)^{-1}\Big) \\ +P(|\log \widehat{\bar{F}}_n - \log \bar{F}| > 1) + P\Big(\Big|\frac{\widehat{\sigma}_{nJ} - \sigma}{\sigma}\Big| > \frac{1}{2}\Big) \\ \le C(n^{-\frac{k}{6}} \log^{\frac{k}{3}} n + n^{1-\frac{k}{2}}). \quad (4.28)$$

Since

$$\frac{\widehat{\sigma}_{nJ}-\sigma}{\sigma}=\frac{\widehat{\sigma}_{nJ}^2-\sigma^2}{2\sigma^2}\Big[1-\frac{\widehat{\sigma}_{nJ}-\sigma}{2\sigma}+Q\Big(\frac{\widehat{\sigma}_{nJ}-\sigma}{2\sigma}\Big)\Big],$$

we get

$$\begin{aligned}R_{11n}=&\sqrt{n}\sigma^{-1}\bar{F}\Big[(\log\widehat{\bar{F}}_n-\log\bar{F})-(\log\widehat{\bar{F}}_n-\log\bar{F})\frac{\widehat{\sigma}_{nJ}^2-\sigma^2}{2\sigma^2}+\frac{(\log\widehat{\bar{F}}_n-\log\bar{F})^2}{2}\Big]\\&+\sqrt{n}\sigma^{-1}\bar{F}(\log\widehat{\bar{F}}_n-\log\bar{F})\Big[\frac{\widehat{\sigma}_{nJ}-\sigma}{2\sigma}-Q\Big(\frac{\widehat{\sigma}_{nJ}-\sigma}{2\sigma}\Big)\Big]\frac{\widehat{\sigma}_{nJ}^2-\sigma^2}{2\sigma^2}\\&\stackrel{\triangle}{=}R_{111n}+R_{112n}.\end{aligned}\tag{4.29}$$

Using the methods similar to proving (4.28) and (4.27), we can prove that

$$P(|R_{112n}|>(n\log n)^{-\frac{1}{2}})\le C(n^{1-\frac{k}{4}}\log^{\frac{k}{2}}n+(n\log n)^{-\frac{1}{2}}).\tag{4.30}$$

In virtue of Lemma 4 and Lemma 1, we have

$$\begin{aligned}R_{111n}=&\sqrt{n}\sigma_0^{-1}\Big[n^{-2}\sum\sum\nolimits_{i<j}h_0(Z_i,\delta_i;Z_j,\delta_j)-n^{-1}\sum_{i=1}^n\varphi(Z_i,\delta_i)\frac{\widehat{\sigma}_{nJ}^2-\sigma^2}{2\sigma^2}\\&+\frac{1}{2}n^{-2}\sum_{i=1}^n\sum_{j=1}^n\varphi(Z_i,\delta_i)\varphi(Z_j,\delta_j)-\frac{1}{2}\sigma_0n^{-1}\Big]\\&+\sqrt{n}\sigma_0^{-1}\triangle_{1n}-\sqrt{n}\sigma_0^{-1}r_{n1}\frac{\widehat{\sigma}_{nJ}^2-\sigma^2}{2\sigma^2}\\&+\sqrt{n}\sigma_0^{-1}r_{n1}\frac{1}{n}\sum_{i=1}^n\varphi(Z_i,\delta_i)+\frac{\sqrt{n}\sigma_0^{-1}r_{n1}^2}{2}\\&\stackrel{\triangle}{=}R_{111n}^{(1)}+R_{111n}^{(2)}+R_{111n}^{(3)}+R_{111n}^{(4)}+R_{111n}^{(5)}.\end{aligned}\tag{4.31}$$

From Lemma 4, it follows that

$$P(|R_{111n}^{(2)}|>(n\log n)^{-\frac{1}{2}})=\mathrm{o}(n^{-\frac{1}{2}}).\tag{4.32}$$

Using Lemma 1 and Lemma 3, we get

$$\begin{aligned}&P(|R_{111n}^{(3)}|>(n\log n)^{-\frac{1}{2}})\le P(|r_{n1}|>Cn^{-\frac{3}{4}}(\log n)^{-\frac{1}{2}})\\&+P(|\widehat{\sigma}_{nJ}^2-\sigma^2|>Cn^{-\frac{1}{4}})\le C(n^{-\frac{k}{4}}\log^{\frac{k}{2}}n+n^{1-\frac{k}{4}}).\end{aligned}\tag{4.33}$$

Analogously, we have by Lemma 1 and D-J inequality

$$\begin{aligned}&P(|R_{111n}^{(4)}|>(n\log n)^{-\frac{1}{2}})\\&\le P(|r_{n1}|>Cn^{-\frac{3}{4}}\log^{-\frac{1}{2}}n)+P(|n^{-1}\sum_{i=1}^n\varphi(Z_i,\delta_i)|>Cn^{-\frac{1}{4}})\le Cn^{-\frac{k}{4}}\log^{\frac{k}{2}}n.\end{aligned}\tag{4.34}$$

and

$$P(|R_{111n}^{(5)}| > (n \log n)^{-\frac{1}{2}}) \leq P(|r_{n1}| > Cn^{-\frac{1}{2}} \log^{-\frac{1}{4}} n) \leq Cn^{-\frac{k}{2}} \log^{\frac{k}{4}} n. \quad (4.35)$$

In virtue of Lemma 2, we get

$$\begin{aligned}
&\Big(n^{-1}\sum_{i=1}^{n}\varphi(Z_i,\delta_i)\Big)(\widehat{\sigma}_{nJ}^2-\sigma^2)\\
&= n^{-2}\bar{F}^2\sum\sum\nolimits_{i<j}\Big[\varphi^2(Z_i,\delta_i)\varphi(Z_j,\delta_j)+\varphi(Z_i,\delta_i)\varphi^2(Z_j,\delta_j)\\
&\quad -(\varphi(Z_i,\delta_i)+\varphi(Z_j,\delta_j))\sigma_0^2\Big]+n^{-2}\bar{F}^2\sum_{i=1}^{n}\varphi^3(Z_i,\delta_i)\\
&\quad -\frac{\sigma^2}{(n-1)n^2}\sum\sum\nolimits_{i<j}(\varphi(Z_i,\delta_i)+\varphi(Z_j,\delta_j))\\
&\quad -\bar{F}^2\Big(n^{-1}\sum_{i=1}^{n}\varphi(Z_i,\delta_i)\Big)^3+R_n n^{-1}\sum_{i=1}^{n}\varphi(Z_i,\delta_i). \qquad (4.36)
\end{aligned}$$

Recalling the definitions of $R_{111n}^{(1)}$ and U_n in (4.31) and in (4.18), we have

$$\begin{aligned}
R_{111n}^{(1)} &= \sqrt{n}\sigma_0^{-1}U_n-\frac{1}{2}n^{-\frac{1}{2}}\sigma_0^{-3}E\varphi^3(Z_1,\delta_1)\\
&\quad -(2\sigma_0^3 n^{\frac{3}{2}})^{-1}\sum_{i=1}^{n}\Big(\varphi^3(Z_i,\delta_i)-E\varphi^3(Z_i,\delta_i)\Big)\\
&\quad +(2\sigma_0 n^{\frac{3}{2}})^{-1}\sum_{i=1}^{n}\varphi(Z_i,\delta_i)+\sqrt{n}(2\sigma_0^3)^{-1}\Big(n^{-1}\sum_{i=1}^{n}\varphi(Z_i,\delta_i)\Big)^3\\
&\quad -\sqrt{n}(2\bar{F}\sigma_0^3)^{-1}R_n n^{-1}\sum_{i=1}^{n}\varphi(Z_i,\delta_i)\\
&\quad +(2n^{\frac{3}{2}}\sigma_0)^{-1}\sum_{i=1}^{n}\Big(\varphi_0^2(Z_i,\delta_i)-E\varphi_0^2(z_i,\delta_i)\Big)\\
&\triangleq \sqrt{n}\sigma_0^{-1}U_n+e_{1n}+e_{2n}+e_{3n}+e_{4n}+e_{5n}+e_{6n}. \qquad (4.37)
\end{aligned}$$

Using the Tchebychev inequality, the D-J inequality and noting that $\varphi(Z_i,\delta_i), 1 \leq i \leq n$ are bounded, independent and identically distributed random variables with $E\varphi(Z_1,\delta_1)=0$, it is easy to prove that

$$P(|e_{in}| > C(n \log n)^{-\frac{1}{2}}) = o(n^{-\frac{1}{2}}), \quad i = 2,3,4,5,6. \qquad (4.38)$$

Combining (4.19), (4.24),(4.25), (4.27), (4.28) – (4.35), (4.37), (4.38) and noting that k is an arbitrary constant, we have $\sqrt{n}\widehat{\sigma}_{nJ}^{-1}(\widehat{\bar{F}}_n-\bar{F}) = \sqrt{n}\sigma_0^{-1}U_n-\nu n^{-\frac{1}{2}}+E_n$, with $P(|E_n| > C(n\log n)^{-\frac{1}{2}}) = \mathrm{o}(n^{-\frac{1}{2}})$, where ν is defined as in Lemma 8. By Lemma 8 and Lemma 5, Theorem 1 is proved.

The proof of Theorem 2 is easy, so we omit it here. Indeed the method to show Theorem 2 is included in the proof of Theorem 3. In the following we will prove Theorem 3. In order to do this we need the following Lemma 9.

Lemma 9. *Under the assumptions of Theorem* 3, *with probability* 1, *we have*

$$\sup_{t\le s<\infty} |\widehat{F}_{nh}(s) - F(s)| \longrightarrow 0 \tag{4.39}$$

$$\sup_{t\le s<\infty} |\widehat{G}_{nh}(s) - G(s)| \longrightarrow 0 \tag{4.40}$$

for any $t > 0$.

Proof. Now we prove (4.39) only; (4.40) can be proved analogously. Since $\widehat{F}_{nh}$ is the probability distribution function corresponding to $\widehat{f}_n$ which is defined in section 2, one can express $\widehat{F}_{nh}$ as follows:

$$\widehat{F}_{nh}(t) = \int_0^{+\infty} K\left(\frac{t-y}{h_n}\right) d\widehat{F}_n(y),$$

where $K(x) = \int_0^x k(t)\,dt$. Hence, we have for any $t > 0$

$$\begin{aligned}
&\sup_{t\le s<\infty} |\widehat{F}_{nh}(s) - F(s)| \\
&\le \sup_{t\le s<\infty} \left|\int_0^{+\infty} K\left(\frac{s-y}{h_n}\right) d\widehat{F}_n(y) - \int_0^{+\infty} K\left(\frac{s-y}{h_n}\right) dF(y)\right| \\
&\quad + \sup_{t\le s<\infty} \left|\int_0^{+\infty} K\left(\frac{s-y}{h_n}\right) dF(y) - F(y)\right| \triangleq I_1 + I_2.
\end{aligned} \tag{4.41}$$

Integrating by parts and employing corrollary 2(ii) on strong uniform consistency of $\widehat{F}_n$ in Csörgö and Horváth (1983), we have

$$I_1 \le \sup_{t\le s<\infty} |\widehat{F}_n(s) - F(s)| \xrightarrow{a.s.} 0. \tag{4.42}$$

Again integrating by parts, and applying the Mean Value Theorem and the boundedness of f to the following proof, we get

$$\begin{aligned}
I_2 &\le \sup_{t\le s<\infty} \Big| \int_0^{sh_n^{-1}} F(s - h_n u)k(u)\,du - \int_0^{+\infty} k(u)F(s)\,du \Big| \\
&\le \sup_{t\le s<\infty} \Big| \int_0^{sh_n^{-1}} f(\zeta_s(u))h_n u k(u)\,du \Big| + \sup_{t\le s<\infty} \int_{sh_n^{-1}}^{+\infty} k(u)\,du \\
&\le Ch_n \int_0^{+\infty} uk(u)\,du + \sup_{t\le s<\infty} \Big[s^{-1}h_n \int_{sh_n^{-1}}^{+\infty} uk(u)\,du \Big] \longrightarrow 0,
\end{aligned} \tag{4.43}$$

where $\varsigma_s(u)$ lies between $s - h_n u$ and u. Therefore, (4.39) follows from (4.41), (4.42) and (4.43).

Proof of Theorem 3. By the continuity of $\widehat{F}_{nh}$ and $\widehat{G}_{nh}$ (this is just why we use the smoothed bootstrap method) and Theorem 1, we have

$$\sup_x |P(\sqrt{n}\widehat{\sigma}_{nJ}^{*-1}(\widehat{\bar{F}}{}_n^*(t) - \widehat{\bar{F}}_{nh}(t)) \le x) - K_n^*(x)| = \mathrm{o}(n^{-\frac{1}{2}}), \tag{4.44}$$

where

$$\begin{aligned}
&K_n^*(x) = \Phi(x) - \frac{\kappa_3^*}{6} n^{-\frac{1}{2}}\phi(x)(x^2 + \frac{1}{2}) \\
&\kappa_3^* = -2\sigma_0^*\Big(\int_0^t \bar{H}^{*-3}\, d\widetilde{H}_1^* + \frac{3}{2}\sigma_0^{*4}\Big) \\
&\sigma_0^{*2} = \bar{F}_{nh}^2 \int_0^t \bar{H}^{*-2}\, d\widetilde{H}_1 \\
&\bar{H}^* = P(Z_1^* > t) \\
&\widetilde{H}_1 = P(Z_1^* > t, \delta_1^* = 1).
\end{aligned}$$

In virtue of Lemma 9, we get

$$\bar{H}^* = (1 - \widehat{F}_{nh}(t))(1 - \widehat{G}_{nh}(t)) \xrightarrow{a.s.} \bar{H}, \tag{4.45}$$

$$\widetilde{H}_1^* = \int_t^{\tau_H \wedge \tau_{H^*}} (1 - \widehat{G}_{nh})\, d\widehat{F}_{nh}(s) \xrightarrow{a.s.} \widetilde{H}_1, \tag{4.46}$$

where $\tau_H^* = \inf\{t : \bar{H}^*(t) = 0\}$. By (4.45) and (4.46), it is easy to obtain $\kappa_3^* \xrightarrow{a.s.} \kappa_3, \sigma_0^{*2} \xrightarrow{a.s.} \sigma_0^2$, which yields $\sup_x |K_n^*(x) - K_n(x)| = \mathrm{o}(n^{-\frac{1}{2}})$, a.s. By Theorem 1 and (4.44), Theorem 3 is proved.

Acknowledgement

The authors are grateful to a referee for his valuable comments and suggestions which have greatly improved the paper.

References

Bickel, P. J., Götze, F. and Van Zwet, W. R. (1986). The Edgeworth expansion for U-statistics of degree two. *Ann. Statist.* **14**, 1463-1481.

Bresley, N. E. and Crowley, J. (1974). A large sample study of the life table and product-limit estimates under random censorship. *Ann. Statist.* **2**, 437-453.

Chang, M. N. and Rao, P. V. (1989). Berry-Esseen bound for the Kaplan-Meier estimator. *Comm. Statist. Theory. Methods* **18**, 4647-4664.

Chang, M. N. (1990). Edgeworth expansion for the Kaplan-Meier estimator. *Commun. Statist. Theory Methods* **20**, 2479-2494.

Chang, M. N. (1991). Moments of the Kaplan-Meier estimator. *Sankhyā: The Indian J. Statist. Ser. A* **53**, 27-50.

Chow, Y. S. and Teicher, H. (1978). *Probability Theory.* Springer-Verlag, New York.

Csörgő, S. and Horváth, L. (1983). The rate of strong uniform consistency for the product-limit estimator. Z. Wahrscheinlichkeitstheorie verw. *Gebiete* **62**, 411-426.

Efron, B. (1979). Bootstrap methods: another look at the jackknife. *Ann. Statist.* **7**, 1-26.

Efron, B. (1981). Censored data and the bootstrap. *J. Amer. Statist. Assoc.* **76**, 312-319.

Földes, A. and Rejtö, L. (1981). Strong uniform consistency for nonparametric survival curve estimators from randomly censored data. *Ann. Statist.* **9**, 122-129.

Földes, A. (1981). A LIL type result for product limit estimator. *Z. Wahrsch. Verw Gebiete* **56**, 75-86.

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

Kaplan, E. L. and Meier, P. (1958). Nonparametric estimation from incomplete observations. *J. Amer. Statist. Assoc.* **53**, 457-481.

Lo, S. H. and Singh, K. (1986). The product-limit estimator and the bootstrap: some asymptotic representations. *Probab. Theory Related Fields* **11**, 455-465.

Phadia, E. G. and Van Ryzin, J. (1980). A note on convergence rates for the product-limit estimator. *Ann. Statist.* **8**, 673-678.

Rao, B. L. S. and Prakasa (1987). *Asymptotic Theory of Statistical Inference.* John Wiley, New York. 128-129.

Singh, K. and Liu, R. Y. (1990). On the validity of the Jack-knife procedure. *Scand J. Statist.* **17**, 11-21.

Wang, J. G. (1987). A note on the uniform consistency of the Kaplan-Meier estimator. *Ann. Statist.* **15**, 1313-1316.

Wang, Q. H. (1994). Some asymptotic properties and smoothed bootstrap approximation for L-estimator in linear models. *Acata. Math. Sci.* **1**, 73-84.

Wang, Q. H. (1995). Strong embedding of product-limit estimator of bivariate survival distribution function under random censorship. *Acta. Math. Sci.* **15**, 123-132.

Wang, Q. H. (1996). Some asymptotic properties for bivariate product-limit estimator. *Chinese Science Bulletin* **40**, 3-5.

Zheng, Z. K. (1989). On the limit points of the Kaplan-Meier estimator. The Third Japan-China Symposium on Statistics. Tokyo, Japan. 167-170.

Department of Probability and Statistics, Peking University, Beijing 100871.

E-mail: qhwang@amath4.amt.ac.cn

(Received December 1995; accepted June 1997)