Statistica Sinica **11**(2001), 1069-1079

EDGEWORTH EXPANSIONS FOR THE PRODUCT-LIMIT ESTIMATOR UNDER LEFT-TRUNCATION AND RIGHT-CENSORING WITH THE BOOTSTRAP

Yi-Ting Hwang

Georgetown University School of Medicine

Abstract: An Edgeworth expansion for the distribution function of the product-limit estimator of survival time under the left-truncation and right-censoring model is derived. This expansion gives more accurate approximations than the usual normal approximation from weak convergence. In addition, by constructing the bootstrap sample from left-truncation and right-censored data, the Edgeworth expansion for the bootstrap statistic is given, allowing a bootstrap base confidence interval with better coverage probability.

Key words and phrases: Bootstrap, censoring, edgeworth expansion, truncation, U-statistic.

1. Introduction

Censored or truncated data occur frequently in many fields such as epidemiology, astronomy and engineering life tests. For instance, consider a prevalent cohort study in epidemiology, which recruits a group of individuals at a specific time with a certain disease status and follows them over time. The variable of interest is the survival time defined as an individual's age at death. Censoring occurs when an individual loses to follow-up, whereas truncation happens when an individual dies before the beginning of the follow-up study. See Hyde (1976), Tasi, Jewell and Wang (1987) and Wang (1991).

Consider an infinite sequence of random vectors $\{(X_m, T_m, C_m), m = 1, 2, \ldots\}$, where the random variables X, T and C are nonnegative and independent with continuous distribution functions, F_X, F_T and F_C, respectively. Suppose one can only observe the pair $Z = \min(X, C)$ and $\delta = I[X \leq C]$, where $I[A]$ denotes the indicator of the event A. Under this restriction, the random variable X is called right-censored by C. Furthermore, if (Z, T, δ) can be observed only when $Z \geq T$, then the triple (Z, T, δ) is called a left-truncated and right-censored observation of X. For convenience, we denote the observable subsequence of (Z_m, T_m, δ_m), $m = 1, 2, \ldots$, by (U_i, V_i, η_i), $i = 1, 2, \ldots$, where $U_i = Z_i$, $V_i = T_i$, and $\eta_i = \delta_i$ if $Z_i \geq T_i$. The conditional distribution $\mathrm{P}[Z \leq z, T \leq t, \delta = d | Z \geq T]$,

for $z, t \in [0, \infty)$, $d = 0$ or 1 defines the left-truncation and right-censoring model (hereafter the TC model). If $T \equiv 0$ with probability 1, the TC model specializes to the right-censoring model; if $C \equiv \infty$ with probability 1, then the TC model becomes the random truncation model.

The general problem is to make inference about the unknown distribution function $F_X(t)$ based on a sample of n independent identically distributed random vectors $\{(U_j, V_j, \eta_j),\ j = 1, \ldots, n\}$ from $\{(Z_m, T_m, \delta_m),\ m = 1, \ldots, m_n\}$, where $n \leq m_n$ and m_n is unknown.

A product-limit estimator $\widehat{F}_X$ as defined in (1) is a well-known estimator for F_X. The asymptotic properties of the PLE have been studied by Gijbels and Wang (1993) and He and Yang (2000). The weak convergence rate is $O(n^{-1/2})$ which offers a confidence interval with coverage probability of $O(n^{-1/2})$. However, under some regularity conditions, the bootstrap approximation is better than the normal approximation for a broad class of studentized statistics. Chen and Lo (1996) showed the bootstrap approximation for the studentized Kaplan-Meier estimator performs better than the normal approximation. Also, see Hall (1992) and Helmers (1991). The objective is to establish an Edgeworth expansion of a studentized PLE and find the bootstrap approximation of a studentized PLE under the TC model to improve the normal approximation.

In Section 3 we establish the Edgeworth expansion of the studentized PLE, which provides an accuracy of $o(n^{-1/2})$. However, since the PLE takes the product form, this is not done directly. By converting the target statistic into a U-statistic, we derive the Edgeworth expansion for the U-statistic first, then for the studentized PLE. In Section 4, we consider bootstrap samples obtained by simple random sampling with replacement from data. This allows us to obtain the expansions for the bootstrap of the studentized PLE. Using this expansion, we construct the coverage probability for the bootstrap approximation. Finally, simulations are conducted to provide numerical supports for the theoretical findings.

2. Assumptions and Notations

Under left-truncation and right-censoring, the range of x for which $F_X(x)$ can be estimated needs to be carefully specified. Let a_X and b_X denote the lower and upper boundaries of the support of F_X: $a_X = \inf\{z : F_X(z) > 0\}$ and $b_X = \sup\{z : F_X(z) < 1\}$. Similar notation will be used for other distributions. It is well known that the non-parametric estimable range of F_X is (a_T, b_Z), where $a_T < b_Z$ and $b_Z = \min(b_X, b_C)$. To ensure the finiteness of moments, we assume $a_T < \min(a_X, a_C)$ such that $F_T(\max(a_X, a_C)) > 0$. Moreover, we set $a_T = 0$ to simplify the notation.

Let the survival function of F_X be $\bar{F}_X = 1 - F_X$. $\bar{F}_C$ and $\bar{F}_T$ are defined similarly. Denote marginal subdistributions by

$$H_U^0(u) = \mathrm{P}[U \le u, \eta = 0] = \alpha^{-1} \int_0^u \bar{F}_X(z) F_T(z) dF_C(z),$$
$$H_U^1(u) = \mathrm{P}[U \le u, \eta = 1] = \alpha^{-1} \int_0^u \bar{F}_C(z) F_T(z) dF_X(z),$$
$$H_V(v) = \mathrm{P}[V \le v] = \alpha^{-1} \int_0^v \bar{F}_C(s) \bar{F}_X(s) dF_T(s),$$

where $\alpha = \mathrm{P}[T \le Z]$. An important quantity for estimation is the coverage probability $R(x) = \mathrm{P}[V \le x \le U] = \alpha^{-1} F_T(x) \bar{F}_C(x) \bar{F}_X(x)$. Let the corresponding empirical subdistributions of H_U^0, H_U^1 and H_V be given by $\widehat{H}_U^0(s) = n^{-1} \sum_{i=1}^n I[U_i \le s, \eta_i = 0]$, $\widehat{H}_U^1(s) = n^{-1} \sum_{i=1}^n I[U_i \le s, \eta_i = 1]$, $\widehat{H}_V(s) = n^{-1} \sum_{i=1}^n I[V_i \le s]$. Thus, the corresponding estimator for $R(s)$ is $R_n(s) = n^{-1} \sum_{i=1}^n I[V_i \le s \le U_i] = \widehat{H}_V(s) - \widehat{H}_U^0(s-) - \widehat{H}_U^1(s-)$ where, for any function $h(x)$, $h(x-)$ denotes the left-continuous version of $h(x)$. Then the well-known product-limit estimator (PLE) of F_X is given by

$$\widehat{F}_X(t) = 1 - \prod_{s \le t} \Big\{ 1 - \frac{\Delta \widehat{H}_U^1(s)}{R_n(s)} \Big\}, \qquad t \in (0, b_Z), \tag{1}$$

where $\Delta \widehat{H}_U^1(s)$ is the difference $\widehat{H}_U^1(s) - \widehat{H}_U^1(s-)$. Here we use the convention that $0/0 = 0$.

Let the cumulative hazard function of F_X be denoted as

$$\Lambda_X(t) = \int_0^t \frac{dF_X(s)}{1 - F_X(s-)} = -\ln \bar{F}_X(t). \tag{2}$$

Replacing F_X by the PLE $\widehat{F}_X$ in (2), we obtain an estimator $-\ln \hat{\bar{F}}_X(t)$ for the cumulative hazard function. Since $\hat{\bar{F}}_X$ is complicated, we use the cumulative hazard estimator as an auxiliary estimator for deriving an Edgeworth expansion for the PLE. As discussed in Hwang (2000), $\widehat{F}_X(t) = 0$ with positive probability. To avoid this difficulty, we partition the sample space Ω as follows: $\Omega_0 = \{\omega \in \Omega : \sup_{0 \le s \le t} \max(|\widehat{H}_V - H_V|, |\widehat{H}_U^0 - H_U^0|, |\widehat{H}_U^1 - H_U^1|) < \gamma\}$ and $\Omega_1 = \Omega - \Omega_0$, where for a fixed t, $\tau \in (0, t]$ is chosen such that $\Theta = \mathrm{P}[V \le \tau, U \ge t] > 0$ and γ is a real number with $0 < \gamma < \Theta/3$ and $6\gamma(\Theta - 3\gamma)^{-2} < 1$. Since the coverage probability $R(x)$ is not necessarily monotone, it is necessary to introduce Θ to compute the probability bound. Clearly, $\Theta \le R(s)$, for any $\tau \le s \le t$. Then, for $\omega \in \Omega_0$, we have $|-\ln \hat{\bar{F}}_X(t) + \ln \bar{F}_X(t)| < \infty$. Also, from the Dvoretzky, Kiefer and Wolfowitz (1956) inequality (DKW inequality hereafter), we have $\mathrm{P}[\Omega_1] = o(n^{-k})$. Thus, we focus discussion on the subspace Ω_0.

3. The Edgeworth Expansion for the Studentized PLE

Let the asymptotic variance of $\sqrt{n}(-\ln \hat{\bar{F}}_X(t)+\ln \bar{F}_X(t))$ be written as $\sigma_0^2 = \int_0^t R^{-2} dH_U^1$ (see Gijbel and Wang (1993)) and the empirical variance estimator of σ_0^2 be $\widehat{\sigma}_0^2 = \int_0^t R_n^{-2} d\widehat{H}_U^1$. Let $\Phi(x)$ and $\phi(x)$ be the standard normal distribution function and the standard normal density function, respectively. The following theorem derives the Edgeworth expansion for the cumulative hazard estimator:

Theorem 3.1. *We have*

$$\sup_x \Big| P\Big[\frac{n^{1/2}}{\widehat{\sigma}_0}\Big(-\ln \hat{\bar{F}}_X(t) + \ln \bar{F}_X(t)\Big) \le x\Big] - \widetilde{\Psi}_n(x)\Big| = o(n^{-1/2}) \tag{3}$$

uniformly in x, where $\widetilde{\Psi}_n(x) = \Phi(x) + n^{-1/2}\phi(x)[\widetilde{\kappa}_1 x^2 + \widetilde{\kappa}_2 + \sigma_0^2(2n^{1/2})^{-1}]$, and

$$\widetilde{\kappa}_1 = \frac{1}{3\sigma_0^3}\int_0^t R^{-3}(u) dH_U^1(u), \tag{4}$$

$$\widetilde{\kappa}_2 = \frac{1}{6\sigma_0^3}\int_0^t R^{-3}(u) dH_U^1(u) + \frac{1}{2\sigma_0^3\alpha}\int_0^t R^{-2}(u)\bar{F}_C(u)\int_0^u R^{-2}(s)F_T(s)dH_U^1(s)dF_X(u). \tag{5}$$

Now we are ready to present the main theorem in this section. Let the asymptotic variance of $\sqrt{n}(\hat{\bar{F}}_X(t) - \bar{F}_X(t))$ be $\sigma^2 = \bar{F}_X^2(t)\sigma_0^2$ and the empirical variance estimator of σ^2 be $\widehat{\sigma}^2 = \hat{\bar{F}}_X^2\widehat{\sigma}_0^2$.

Theorem 3.2. *We have*

$$\sup_x \Big| P\Big[\frac{n^{1/2}}{\widehat{\sigma}}\Big(\hat{\bar{F}}_X(t) - \bar{F}_X(t)\Big) \le x\Big] - \Psi_n(x)\Big| = o(n^{-1/2})$$

uniformly in x, where

$$\Psi_n(x) = \Phi(x) + n^{-1/2}\phi(x)\left(\kappa_1 x^2 + \kappa_2\right), \tag{6}$$

$$\kappa_1 = -\frac{\sigma_0^{-3}}{3}\int_0^t R^{-3} dH_U^1 + \frac{\sigma_0}{2}, \tag{7}$$

$$\kappa_2 = -\frac{\sigma_0^{-3}}{6}\int_0^t R^{-3} dH_U^1 - \frac{\sigma_0}{2} - \frac{\sigma_0^{-3}}{2\alpha}\int_0^t R^{-2}(u)\bar{F}_C(u)\int_0^u R^{-2}(s)F_T(s)dH_U^1(s)dF_X(u). \tag{8}$$

4. The Bootstrap Statistic

From the normal approximation, a confidence interval of $F_X(t)$ can be constructed with a coverage probability accurate to $O(n^{-1/2})$. By means of the bootstrap, we show that the coverage probability is accurate to order $o(n^{-1/2})$.

A bootstrap sample is obtained by simple random sampling with replacement from $\{(U_i, V_i, \eta_i), i = 1, \ldots, n\}$. Let $\{(U_i^*, V_i^*, \eta_i^*), i = 1, \ldots, n\}$ denote the bootstrap sample. The symbol $*$ represents statistics associated with the bootstrap sample. For instance, P^* is the probability measure on the bootstrap sample,

$$\begin{aligned}\widehat{H}_U^{1*}(u) &= \frac{1}{n}\sum_{i=1}^{n} I[U_i^* \le u, \eta_i^* = 1],\\ g^*(U_i^*, V_i^*, \eta_i^*) &= R_n^{-1}(U_i^*)\eta_i^* I[0 \le U_i^* \le t]\\ &\quad + \int_0^t R_n^{-2}(s)(I[s \le V_i^*] - I[s < U_i^*])d\widehat{H}_U^1(s).\end{aligned}$$

Now the process $\left\{\widehat{H}_U^{1*}(s), s \in [0,\infty)\right\}$ is the empirical process with the parent distribution $\{\widehat{H}_U^1(s), s \in [0,\infty)\}$.

Let $\Psi_n^*(x) = \Phi(x) + n^{-1/2}\phi(x)(\kappa_1^* x^2 + \kappa_2^*)$, where κ_1^* and κ_2^* are the corresponding bootstrap estimates of κ_1 and κ_2 as defined in (7) and (8). Note that κ_1^* and κ_2^* depend only on the sample $(U_i, V_i, \eta_i), i = 1, \ldots, n$, and not on the bootstrap samples. The following theorem gives the bootstrap accuracy for the studentized PLE. The proof is similar to that of Theorem 3 in Helmers (1991) and is there fore omitted .

Theorem 4.1. *For $\omega \in \Omega_0$ we have*

$$\sup_x \left|\mathrm{P}^*\Big[\frac{n^{1/2}}{\widehat{\sigma}^*}\Big(\widehat{\bar{F}}_X^*(t) - \widehat{\bar{F}}_X(t)\Big) \le x\Big] - \Psi_n^*(x)\right| = o(n^{-1/2}),$$

$$\sup_x \left|\mathrm{P}^*\Big[\frac{n^{1/2}}{\widehat{\sigma}^*}\Big(\widehat{\bar{F}}_X^*(t) - \widehat{\bar{F}}_X(t)\Big) \le x\Big] - \widehat{Q}_n(x)\right| = o(n^{-1/2}),$$

where $\widehat{Q}_n(x) = \mathrm{P}[\sqrt{n}\widehat{\sigma}^{-1}(\widehat{\bar{F}}_X(t) - \bar{F}_X(t)) \le x]$.

From the Edgeworth expansion for the bootstrap statistic, we can now construct confidence intervals with better coverage probabilities. Let $z_\alpha = \Phi^{-1}(\alpha)$. The normal approximation yields the following one-sided confidence interval for $\bar{F}_X(t)$: $(-\infty, \widehat{\bar{F}}_X - z_\alpha\widehat{\sigma}n^{-1/2})$. It is easy to see that by (6), we have

$$\mathrm{P}\Big[\bar{F}_X(t) \le \widehat{\bar{F}}_X(t) - z_\alpha\widehat{\sigma}n^{-1/2}\Big] = 1 - \alpha + n^{-1/2}\phi(z_\alpha)(\kappa_1 z_\alpha^2 + \kappa_2) + o(n^{-1/2}).$$

Therefore, the error in the coverage probability for the normal based confidence interval is of order $O(n^{-1/2})$. Let q_α denote the α-quantile of $\widehat{Q}_n$. According to

the inversion formula for the Edgeworth expansion (Hall (1992, p.88)), we have $z_\alpha = q_\alpha + n^{-1/2}(-\kappa_1 z_\alpha^2 - \kappa_2)$. Thus, the error in approximating the quantile z_α by q_α is of order $O(n^{-1/2})$.

Let q_α^* be the α-quantile of the distribution of $\sqrt{n}(\widehat{\sigma}^*)^{-1}(\widehat{\bar{F}}_X^*(t) - \widehat{\bar{F}}_X(t))$. Then, the following theorem shows that the coverage probability for the bootstrap -based confidence interval is accurate to $o(n^{-1/2})$. Also, the error in estimating the quantile q_α^* by q_α is of order $o(n^{-1/2})$.

Theorem 4.2. *For fixed $0 < \alpha < 1$, we have $q_\alpha^* = q_\alpha + o(n^{-1/2})$ a.s. and*

$$\mathrm{P}\Big[\bar{F}_X(t) \le \widehat{\bar{F}}_X(t) - q_\alpha^* \widehat{\sigma} n^{-1/2}\Big] = 1 - \alpha + o(n^{-1/2}). \tag{9}$$

The proof, based on a standard delta method, can be taken from Theorem 7 in Chen and Lo (1996).

Example 4.1. To examine the result of Theorem 4.2, consider the distribution functions $F_X(s) = 1 - exp(-(s-a))$, $F_C(s) = 1 - exp(-\beta_C(s-a))$ and $F_T(s) = 1 - exp(\beta_T s)$, where a constant a is chosen so that $F_T(a) > 0$. Here we set $a = 0.01$. The truncation probability is 0.351. The coverage probabilities of one-sided confidence intervals constructed on normal approximations and bootstrap approximations are shown in Table 1 for three fixed time points, $t = 0.7, t = 0.6$ and $t = 0.5$. All three bootstrap approximations perform better than the normal approximatioin at all nominal levels. At time $t = 0.7$ and nominal level 0.975, the improvement for the bootstrap approximation is 100%. However at time 0.5, the improvement does not look as dramatic. The normal and bootstrap approximations are rougher at $t = 0.5$ due to fewer observations in calculating the PLE.

Table 1. Coverage probability of confidence interval*

	Time $t = 0.7$		Time $t = 0.6$		Time $t = 0.5$	
	$n^{**} = 28$		$n = 25$		$n = 29$	
Nominal	Normal	Bootstrap	Normal	Bootstrap	Normal	Bootstrap
0.975	0.932	0.975	0.920	0.974	0.912	0.973
0.95	0.888	0.949	0.872	0.935	0.863	0.921
0.90	0.816	0.885	0.797	0.850	0.788	0.828
0.85	0.753	0.799	0.735	0.768	0.728	0.742

*Sample size=50. The bootstrap approximations are based on 1000 repetitions.

**n represents the truncated sample size.

Remark 4.1. During the revision, the author was made aware of a recent unpublished manuscript by Drs. Wang and Jing (2000) which addresses the same

problems as presented in this article. Both papers use similar methodologies; however, the results on Edgeworth expansions for the studentized PLE are different. It is the author's observation that while implicating Theorem 1.2 in Bickel, Götze and van Zwet (1986), the coefficient κ_3 in Lemma 5.5 in Wang and Jing is derived incorrectly.

5. Extensions

The independence assumption on (X, T, C) can be further relaxed by assuming that X is independent of the pair (T, C), with possible dependence between the random variables T and C.

Acknowledgements

The author would like to thank her dissertation advisor Professor Grace L. Yang. The author is also grateful for valuable comments from the Associate Editor and the referee.

This research was partially supported by Agency for HealthCare Research and Quality Grant #HS08395 and Contract DAMD 17-96-C-6069 from the Department of the Army.

Appendix A. Proofs

To derive the U-statistic representation for $-\ln \widehat{\bar{F}}_X(t) + \ln \bar{F}_X(t)$, we need the following notation. For $1 \le j, k \le n$, let

$$B_1(U_j, U_k, V_k) = R^{-2}(U_j) I[0 \le U_j \le t](I[U_j \le V_k] - I[U_j < U_k]),$$
$$B_2(U_j, V_j, U_k, V_k) = \int_0^t R^{-3}(s) \prod_{i=j,k} (I[s \le V_i] - I[s < U_i]) dH_U^1(s).$$

In particular, when $j = k$, it is easy to see that

$$\widetilde{B}_1(U_j) = B_1(U_j, V_j, U_j, V_j) = -\frac{1}{2} R^{-2}(U_j) I[0 \le U_j \le t],$$
$$\widetilde{B}_2(U_j, V_j) = B_2(U_j, V_j, U_j, V_j) = -\int_0^t R^{-3}(s)(I[s \le V_j] - I[s < U_j]) dH_U^1(s).$$

To simplify the notation, set $\mathbf{U}_i = (U_i, V_i, \eta_i)$ and let

$$g(\mathbf{U}_j) = R^{-1}(U_j) I[0 \le U_j \le t] + \int_0^t R^{-2}(s)(I[s \le V_j] - I[s < U_j]) dH_U^1(s),$$
$$\begin{aligned}\psi(\mathbf{U}_j, \mathbf{U}_k) &= B_1(U_j, U_k, V_k)\eta_j + B_2(U_j, V_j, U_k, V_k) \\ &\quad - \mathrm{E}[B_1(U_j, U_k, V_k)\eta_j + B_2(U_j, V_j, U_k, V_k) | U_j, V_j, \eta_j],\end{aligned}$$
$$h(\mathbf{U}_j, \mathbf{U}_k) = g(\mathbf{U}_j) + g(\mathbf{U}_k) + \psi(\mathbf{U}_j, \mathbf{U}_k) + \psi(\mathbf{U}_k, \mathbf{U}_j).$$

For $\omega \in \Omega_0$, Lemma 4.1 in Hwang (2000) yields

$$-\ln \hat{\bar{F}}_X(t) + \ln \bar{F}_X(t) = \mathcal{U}_n + \text{Rem} + \frac{\sigma_0^2}{2n}, \tag{10}$$

where $\widetilde{R}_n(s) = (\widehat{H}_V - \widehat{H}_U^0 - \widehat{H}_U^1)(s)$, $\mathcal{U}_n = n^{-2}\sum_{i<j} h(\mathbf{U_i}, \mathbf{U_j})$ and

$$\text{Rem} = \int_0^t R^{-3}[\widetilde{R}_n - R]^2 d[\widehat{H}_U^1 - H_U^1] + \sum_{k=3}^{\infty}\Big\{(-1)^k \int_0^t R^{-(k+1)}[\widetilde{R}_n - R]^k d\widehat{H}_U^1\Big\} \tag{11}$$

$$+ \sum_{k=3}^{\infty}\Big\{(-1)^{k-1}k^{-1}\sum_{i=2}^{k} n^{-i+1}\binom{k}{i}\int_0^t R^{-k}(\widetilde{R}_n - R)^{k-i} d\widehat{H}_U^1\Big\} \tag{12}$$

$$+ n^{-2}\sum_{i=1}^{n}\Big\{2g(\mathbf{U}_i) + \widetilde{B}_1(U_i)\eta_i + \widetilde{B}_2(U_i, V_i) - \frac{\sigma_0^2}{2}\Big\}. \tag{13}$$

We can show that

$$\mathrm{P}[\sqrt{n}\sigma_0^{-1}|\text{Rem}| > (\log n)^{-1} n^{-1/2}] = o(n^{-1/2}) \tag{14}$$

using Lemma 3.2 in Hwang (2000) and the DKW inequality. Details are in Lemma 5.1.2 in Hwang (1999).

Since $\mathcal{U}_n$ is a U-statistic of order n^{-2} with a symmetric kernel h, $\mathcal{U}_n$ can be expressed as $\mathcal{U}_n = n^{-2}\{(n-1)\sum_{i=1}^n g(\mathbf{U}_i) + \sum_{i<j}[\psi(\mathbf{U}_i, \mathbf{U}_j) + \psi(\mathbf{U}_j, \mathbf{U}_i)]\}$, which has the form of (1.5) in Bickel, Götze and van Zwet (1986). From their Theorem 1.2, we obtain the following lemma. The proof is similar to that of Lemma 1 in Chang (1991). Note that, because of the dependence between U and V, the coefficient K_3 is different from that in Chang's Lemma 1.

Lemma A.1 *We have* $\sup_x \left|P[\sqrt{n}\sigma_0^{-1}\mathcal{U}_n \le x] - G_n(x)\right| = o(n^{-1/2})$, $G_n(x) = \Phi(x) - K_3(6n^{1/2})^{-1}\phi(x)(x^2-1)$ *and*

$$K_3 = \sigma_0^{-3}\int_0^t R^{-3}(u) dH_U^1(u) + \frac{3}{\sigma_0^3\alpha}\int_0^t R^{-2}(s)\bar{F}_C(s)\int_0^s R^{-2}(u)F_T(u)dH_U^1(u)dF_X(s). \tag{15}$$

Proof of Theorem 3.1. To use Lemma A.1, write $\widehat{\sigma}_0^2 - \sigma_0^2 = \int_0^t R_n^{-2} d\widehat{H}_U^1 - \int_0^t R^{-2} dH_U^1 = n^{-1}\sum_{i=1}^n f(\mathbf{U}_i) + \xi_1$, where $f(\mathbf{U}_i) = -2\widetilde{B}_1(\mathbf{U}_i)\eta_i - \widetilde{B}_2(\mathbf{U}_i, \mathbf{U}_j) + \sigma_0^2$ and $\xi_1 = -2\int_0^t R^{-3}(R_n - R)d(\widehat{H}_U^1 - H_U^1) + \int_0^t R_n^{-2}R^{-3}(R_n - R)^2(R + 2R_n)d\widehat{H}_U^1$. Thus, we have

$$\frac{\sigma_0}{\widehat{\sigma}_0} = 1 - \frac{1}{2n\sigma_0^2}\sum_{i=1}^n f(\mathbf{U}_i) + \xi_2, \tag{16}$$

where $\xi_2 = -\xi_1(2\sigma_0^2)^{-1} + (\widehat{\sigma}_0 - \sigma_0)^2(\widehat{\sigma}_0 + 2\sigma_0)(2\widehat{\sigma}_0\sigma_0^2)^{-1}$. By (10) and (16), we have

$$\frac{n^{1/2}}{\widehat{\sigma}_0}\Big(- \ln \widehat{\bar{F}}_X(t) + \ln \bar{F}_X(t)\Big) = \zeta + \xi_3 - \frac{\sigma_0}{2n^{1/2}}, \tag{17}$$

where

$$\zeta = \frac{n^{1/2}}{2\sigma_0}\Big(\frac{2}{n^2}\sum_{i<j} h(\mathbf{U}_i, \mathbf{U}_j)\Big)\Big(1 - \frac{1}{2n\sigma_0^2}\sum_{i=1}^{n} f(\mathbf{U}_i)\Big),$$

$$\xi_3 = \frac{n^{1/2}\xi_2}{2\sigma_0}\Big(\frac{2}{n^2}\sum_{i<j} h(\mathbf{U}_i, \mathbf{U}_j)\Big) + \frac{n^{1/2}}{\widehat{\sigma}_0}\mathrm{Rem} + \frac{1}{2n^{3/2}\sigma_0}\sum_{i=1}^{n} f(\mathbf{U}_i) - \frac{\sigma_0\xi_2}{2n^{1/2}}.$$

The quantity ζ can be rewritten as

$$\zeta = \frac{n-1}{n}\frac{n^{1/2}}{2\sigma_0}\binom{n}{2}^{-1}\sum_{i<j}\Big\{\widetilde{h}(\mathbf{U}_i, \mathbf{U}_j) - \frac{1}{n\sigma_0^2}\mathrm{E}[g(\mathbf{U}_i)f(\mathbf{U}_i)]\Big\} + \mathrm{Rem}_1, \tag{18}$$

where $\widetilde{h}(\mathbf{U}_i, \mathbf{U}_j) = h(\mathbf{U}_i, \mathbf{U}_j) - (2\sigma_0^2)^{-1}[g(\mathbf{U}_i)f(\mathbf{U}_j) + g(\mathbf{U}_j)f(\mathbf{U}_i)]$ and

$$\begin{aligned}\mathrm{Rem}_1 = &-\frac{1}{2n^{5/2}\sigma_0^3}\sum_{i<j}\big\{\varphi(\mathbf{U}_i, \mathbf{U}_j)(f(\mathbf{U}_i)+f(\mathbf{U}_j)) - g(\mathbf{U}_i)f(\mathbf{U}_j) + g(\mathbf{U}_j)f(\mathbf{U}_i)\big\}\\ &-\frac{n-1}{2n^{5/2}\sigma_0^3}\sum_{i=1}^{n}\big\{g(\mathbf{U}_i)f(\mathbf{U}_i) - \mathrm{E}[g(\mathbf{U}_i)f(\mathbf{U}_i)]\big\}\\ &-\frac{1}{2n^{5/2}\sigma_0^3}\sum_{i=1}^{n}\Big\{f(\mathbf{U}_i)\sum_{\substack{k<m\\ k\neq i\neq m}}\varphi(\mathbf{U}_k, \mathbf{U}_m)\Big\}.\end{aligned}$$

Clearly, $\sum_{i<j}\widetilde{h}(\mathbf{U}_i, \mathbf{U}_j)$ is a U-statistic of degree 2 with an expected value of zero. Also, we have $\mathrm{E}[\widetilde{h}(\mathbf{U}_j, \mathbf{U}_k)|\mathbf{U}_j] = g(\mathbf{U}_j)$ and $\mathrm{E}[g(\mathbf{U}_j)] = \mathrm{E}[f(\mathbf{U}_j)] = 0$ from Lemma 3.1 in Chang and Hwang (2000). Thus, by Lemma A.1 and Lemma 2 in Chang (1991), we have

$$\begin{aligned}&\mathrm{P}\Big[\frac{n^{1/2}}{\sigma_0 n^2}\sum_{i<j}\Big(\widetilde{h}(\mathbf{U}_i, \mathbf{U}_j) - \frac{\mathrm{E}[g(\mathbf{U}_i)f(\mathbf{U}_i)]}{n\sigma_0^2}\Big) \le x\Big]\\ &= \Phi(x) + n^{-1/2}\phi(x)(\widetilde{\kappa}_1 x^2 + \widetilde{\kappa}_2) + o(n^{-1/2}),\end{aligned}$$

where $\widetilde{\kappa}_1$ and $\widetilde{\kappa}_2$ are as defined in (4) and (5).

To finish the proof, we must show that the error term Rem_1 and ξ_3 are of order $o(n^{-1/2})$. Calculations for the moments of Rem_1 can be found in Callaert and Veraverbeke (1981). Then, applying Chebyshev's inequality, one can show that

$$\mathrm{P}[|\mathrm{Rem}_1| > (n\log n)^{-1/2}] = o(n^{-1/2}). \tag{19}$$

Moreover, we have

$$\mathrm{P}[|\xi_3| > (n\log n)^{-1/2}] = o(n^{-1/2}). \tag{20}$$

The proof applies Hoeffding's inequality, Bernstein's inequality (Serfling (1980, p.85)) and (14). Details are described in Lemma A.5.2. in Hwang (1999). The proof is therefore complete.

Proof of Theorem 3.2. From a Taylor expansion and (10), we obtain

$$\begin{aligned}\frac{n^{1/2}}{\widehat{\sigma}}\Big(\widehat{\bar{F}}_X(t) - \bar{F}_X(t)\Big) = &-\frac{n^{1/2}}{\widehat{\sigma}_0}\Big(-\ln\widehat{\bar{F}}_X(t) + \ln\bar{F}_X(t)\Big)\Big(1 + \frac{1}{2n}\sum_{i=1}^{n} g(\mathbf{U}_i)\Big)\\ &-\frac{n^{1/2}}{\widehat{\sigma}_0}\Big(-\ln\widehat{\bar{F}}_X(t) + \ln\bar{F}_X(t)\Big) \times \mathrm{Rem}_2, \end{aligned}\tag{21}$$

where $|\epsilon^{'} - 1| \le \exp(-\ln\widehat{\bar{F}}_X(t) + \ln\bar{F}_X(t))$ and

$$\mathrm{Rem}_2 = \frac{1}{2n^2}\Big(\sum_{i<j}\varphi(\mathbf{U}_i,\mathbf{U}_j) - \sum_{i=1}^{n} g(\mathbf{U}_i)\Big) + \mathrm{Rem} + \frac{\epsilon^{'}}{6}\Big(-\ln\widehat{\bar{F}}_X(t) + \ln\bar{F}_X(t)\Big)^2 + \frac{\sigma_0^2}{2n}.$$

Applying (14), Lemma 3.2 in Hwang (2000) and Lemma 3.1 in Chang and Hwang (2000), the quantity Rem_2 can be shown to have a second moment of order $O(n^{-2})$. Therefore, combining Theorem 3.1 and the fact that $\mathrm{P}[|\mathrm{Rem}_2| > (n\log n)^{-1/2}] = o(n^{-1/2})$, the second term of the left side of (21) is clearly negligible. From (17), the first term in (21) can be rewritten as $-\zeta(1+(2n)^{-1}\sum_{i=1}^{n} g(\mathbf{U}_i)) + \sigma_0(2\sqrt{n})^{-1} + \mathrm{Rem}_3$, where $\mathrm{Rem}_3 = -\xi_3(1+(2n)^{-1}\sum_{i=1}^{n} g(\mathbf{U}_i)) + \sigma_0(4n^{3/2})^{-1}\sum_{i=1}^{n} g(\mathbf{U}_i)$. By (20) and Bernstein's inequality, we have $\mathrm{P}[|\mathrm{Rem}_3| > (n\log n)^{-1/2}] = o(n^{-1/2})$. To complete the proof, it suffices to show that $-\zeta(1 + (2n)^{-1}\sum_{i=1}^{n} g(\mathbf{U}_i)) + \sigma_0(2\sqrt{n})^{-1}$ has the same Edgeworth expansion as $\Psi_n(x)$, as defined in (6). From the proof of (18) and (19), ζ can be reexpressed as $(n^{3/2}\sigma_0)^{-1}\sum_{i<j}\{\widetilde{h}(\mathbf{U}_i,\mathbf{U}_j) - (n\sigma_0^2)^{-1}\mathrm{E}[g(\mathbf{U}_i)f(\mathbf{U}_i)]\} + o(n^{-1/2})$, where the first term can be represented as a U-statistic by using the technique employed in (18). We complete the proof by applying Lemma A.1 and Lemma 2 in Chang (1991).

References

Bickel, P. J., Götze, F. and van Zwet, W. R. (1986). The Edgeworth expansion for U-statistics of degree two. *Ann. Statist.* **14**, 1463-1484.

Callaret, H. and Veraverbeke, N. (1981). The order of the normal approximation for a studentized U-statistics. *Ann. Statist.* **9**, 194-200.

Chang, D. C. and Hwang, Y. T. (2000). A Berry-Esséen bound for the product-limit estimator under left-truncation and right-censoring. To appear in *Appl. Anal.*

Chang, M. N. (1991). Edgeworth expansion for the Kaplan-Meier estimator. *Commun. Statist. Theory Meth.* **20**, 2479-2494.

Chen, K. and Lo, S. H. (1996). On bootstrap accuracy with censored data. *Ann. Statist.* **24**, 569-595.

Dvoretzky, A., Kiefer, J. and Wolfowitz, J. (1956). Asymptotic minimax character of the sample distribution function and of the classical multinomial estimator. *Ann. Math. Statist.* **27**, 642-669.

Gijbels, I. and Wang, J. L. (1993). Strong representations of the survival function estimator for truncated and censored data with applications. *J. Mult. Anal.* **47**, 210-229.

Hall, P. (1992). *The Bootstrap and Edgeworth Expansion.* Springer-Verlag, New York.

Helmers, R. (1991). On the Edgeworth expansion and the bootstrap approximation for a studentized U-statistics. *Ann. Statist.* **19**, 470-484.

He, S. and Yang, G. L. (2000). On the strong convergence of the product-limit estimator and its integrals under censoring and random truncation. *Statist. Probab. Lett.* **49**, 235-244.

Hwang, Y. T. (1999). Finite moments for the left-truncation and right-censoring model. Ph.D. Dissertation, Dept. of Mathematics, University of Maryland.

Hwang, Y. T. (2000). Moments of the product-limit estimator under left-truncation and right-censoring. Submitted for publication.

Hyde, J. (1977). Testing survival under right censoring and left truncaton. *Biometrika* **64**, 225-230.

Serfling, R. J. (1980). *Approximation Theorems of Mathematical Statistics.* John Wiley, New York.

Tsai, W. Y., Jewell, N. P. and Wang, W. C. (1987). A note on the product-limit estimtor under right censoring and left truncation. *Biometrika* **74**, 883-886.

Wang, M. C. (1991). Nonparametric estimation from cross-sectional survival data. *J. Amer. Statist. Assoc.* **86**, 130-143.

Wang, Q. H. and Jing, B. Y. (2000). Edgeworth expansion and bootstrap approximation for studentized product-limit estimator with truncated and censored data. Unpublished.

Department of Oncology, Georgetown University School of Medicine, Lombardi Cancer Center, Washington, DC, U.S.A.

E-mail: yh4@gunet.georgetown.edu

(Received April 2000; accepted April 2001)