Statistica Sinica **10**(2000), 281-296

STRONG GAUSSIAN APPROXIMATIONS IN THE RANDOM TRUNCATION MODE

SzeMan Tse

National Donghua University

Abstract: In the random left-truncation model, one observes (X_i, Y_i) only if $X_i \geq Y_i$, $i = 1, \ldots, N$. The nonparametric maximum likelihood estimator aims at reconstructing the distribution function of X from the observed empirical data. In this paper, strong approximations of the cumulative hazard process and product-limit process on increasing sets by sequences of copies of corresponding Gaussian limiting processes are constructed. The convergence rates are $N^{-1/6}\log N$ on fixed sets. Futhermore, strong approximations with two-parameter Gaussian processes are obtained with convergence rates $N^{-1/8}(\log N)^{3/2}$ on fixed sets.

Key words and phrases: Cumulative hazard, Gaussian approximations, product-limit, random truncation.

1. Introduction

Let X and Y be independent, positive random variables with distribution functions F and G respectively, both assumed to be continuous. Let $(X_1, Y_1), \ldots,$ (X_N, Y_N) be i.i.d. as (X, Y), where the population size N is fixed, but unknown. Suppose one observes only those pairs (X_i, Y_i), $i \leq N$ for which $Y_i \leq X_i$ and at least one such pair exists. Let $(x_1, y_1), \ldots, (x_n, y_n)$ denote these pairs. The problem is to estimate F, G and N using the observed empirical data. By convention, X is the variable of interest. Hence, this model is called the left truncation model.

As a consequence of truncation, the number of observed pairs, n, is a $\mathrm{Bin}(N, \alpha)$ random variable, with $\alpha := P(Y \leq X)$. By the strong law of large numbers, $n/N \to \alpha$ almost surely as $N \to \infty$. Conditional on the value of n, (x_i, y_i), $i = 1, \ldots, n$ are still i.i.d. but with conditional joint distribution

$$\begin{aligned} H_*(x, y) &= P\{\, x_1 \leq x, y_1 \leq y \,\} = P\{\, X \leq x, Y \leq y \mid Y \leq X \,\} \\ &= \alpha^{-1} \int_0^x G(y \wedge z)\, dF(z) \end{aligned}$$

for $x, y > 0$. The marginal distribution functions are denoted by

$$F_*(x) := H_*(x, \infty) = \alpha^{-1} \int_0^x G(z)\, dF(z),$$

$$G_*(y) := H_*(\infty, y) = \alpha^{-1}\int_0^\infty G(y \wedge z)\, dF(z).$$

Here and in the following, $\int_a^b = \int_{(a,b]}$ for $0 \le a < b \le \infty$. Empirical counterparts of these distribution functions are denoted by $H_n^*(x,y)$, $F_n^*(x)$ and $G_n^*(y)$, respectively.

Woodroofe (1985) studied the nonparametric maximum likelihood estimator (MLE), and reviewed examples where such data may occur. Let $a_F = \inf\{\, z > 0 : F(z) > 0 \,\}$ and $b_F = \sup\{\, z > 0 : F(z) < 1 \,\}$, and similarly for a_G and b_G. It was shown that F can be reconstructed only when $a_G \le a_F$. The interesting case is $a_G = a_F$. For the sake of simplicity, we suppose that $a_G = a_F = 0$ throughout. Woodroofe's analysis makes use of the one-to-one correspondence between F and its associated cumulative hazard function

$$\Lambda(t) := \int_0^t \frac{dF(z)}{1 - F(z-)} = \int_0^t \frac{dF_*(z)}{C(z)}, \qquad 0 \le t < \infty,$$

where, for $0 \le z < \infty$, $C(z) = G_*(z) - F_*(z-) = \alpha^{-1}G(z)\,[1 - F(z-)]$. With $C_n(z) = G_n^*(z) - F_n^*(z-)$, a natural estimator of $\Lambda(t)$ is

$$\Lambda_n(t) := \int_0^t \frac{dF_n^*(z)}{C_n(z)} = \frac{1}{n}\sum_{i:x_i \le t} \frac{1}{C_n(x_i)}.$$

The nonparametric MLE of F, originally proposed by Lynden-Bell (1971), is

$$1 - F_n(t) = \prod_{i:x_i \le t}\left[\frac{nC_n(x_i) - 1}{nC_n(x_i)}\right],$$

assuming no ties in the data. This is analogous to the Kaplan-Meier product-limit estimator in the random censorship model. The term $nC_n(z)$ can be thought of as the number of pairs at risk at time z. Note that, in contrast with the censorship model, $nC_n(z)$ is not a monotonic function of z in general.

Woodroofe (1985) proved the weak convergence of the cumulative hazard process $\hat{Z}_n(t) = \sqrt{n}\,[\,\Lambda_n(t) - \Lambda(t)\,]$ and product-limit process $Z_n(t) = \sqrt{n}\,[\,F_n(t) - F(t)\,]$ to certain Gaussian processes in the Skorohod space $D[a,b]$ for any interval (a,b) such that $a_G < a < b < b_F$. No further assumption on F and G is needed. However, as $a \to a_G$, the variances of the limiting processes diverge to infinity unless F and G satisfy the condition

$$\int_{a_F}^\infty \frac{dF(z)}{G(z)} < \infty. \tag{1.1a}$$

Chao and Lo (1988) initiated almost sure representations of $\hat{Z}_n$ and Z_n in terms of sums of normed i.i.d. random processes with uniformly valid remainder terms of order $o(1)$. Stute (1993) provided a rigorous proof and improved the rates of convergence. By assuming the stronger condition

$$\int_{a_F}^{\infty} \frac{dF(z)}{G^2(z)} < \infty, \tag{1.2}$$

Stute showed that the respective remainder terms are $o(n^{-1/2}(\log n)^{(3+\eta)/2})$ for any $\eta > 0$, and $O(n^{-1/2}(\log n)^3)$ uniformly over $[a_F, b]$, $a_F < b < b_F$. The usefulness of the representations is thus greatly enhanced. In applications, these are used by invoking asymptotic results for sums of i.i.d. random processes. While the representations are illuminating, the summands are complicated. In contrast, the limiting Gaussian processes are easily characterized by their respective covariance functions. The question then arises: can the sums of i.i.d. processes themselves be approximated in the almost sure sense by their respective Gaussian limits? And if so, at what rates? A positive answer to these questions will provide a neat representation of the processes and eliminate the need to invoke asymptotic results for sums of i.i.d. processes. Also, with good approximation rates, almost sure statements like the functional law of the iterated logarithm are directly inherited from the same for the limiting processes.

In this paper, we construct strong Gaussian approximations for the cumulative hazard process and product-limit process. The results are meant as large-sample Gaussian approximations for a large finite population of partially hidden objects. The basis of our work is Borisov's (1982) extension of Komlós, Major and Tusnády's (KMT) approximation theorem for the univariate empirical process to higher dimensions. In our case, the appropriate dimension is two. In applying Borisov's theorem, we bring forth the known two-dimensional nature of the estimation process involved in a form which is more explicit than before. In contrast, Burke, S. Csörgő and Horváth (1981, 1988) used the KMT theorem to obtain strong Gaussian approximations of the estimation processes in the random censorship model, thus revealing the true one-dimensional nature of the problem there. While the KMT approximation rate is known to be the best possible, the status of Borisov's rate is much less certain. Any improvement on that rate implies the same here.

Woodroofe (1985) suggested using $\int G_n dF_n$ as an estimator for α. In that case, both X and Y are of interest simultaneously. The interesting case is $b_G = b_F$ which we also assume. Thus, even though we follow the literature in presenting the results for the random variable X only, analogs for the random variable Y hold jointly under the same condition. The counterpart of Woodroofe's condition (1.1a) is

$$\int_0^{b_G} \frac{dG}{1-F} < \infty, \tag{1.1b}$$

when Y is the variable of interest. Both conditions will be assumed in this paper.

2. Notation and Basic Auxiliary Results

In contrast to the usual practice in current literature, we formulate the results in terms of the non-random population size N rather than the random number of observed items n. Define

$$NF_N^*(x) =: \sum_{i=1}^{N} I\{X_i \le x, Y_i \le X_i\},$$
$$NG_N^*(y) =: \sum_{i=1}^{N} I\{Y_i \le y, Y_i \le X_i\}$$

in terms of the possibly unobserved (X_i, Y_i), $i \le N$. Note that $NF_N^* = nF_n^*$ and $NG_N^* = nG_n^*$. Also, letting $C_*(z) = \alpha C(z)$ and $C_N(z) = G_N^*(z) - F_N^*(z-)$, we have $NC_N = nC_n$. Correspondingly, Λ_n and F_n are now rewritten as

$$\Lambda_N(t) := \int_0^t \frac{dF_N^*(z)}{C_N(z)} = \frac{1}{N} \sum_{i:x_i \le t} \frac{1}{C_N(x_i)},$$
$$1 - F_N(t) = \prod_{i:x_i \le t} \left[\frac{NC_N(x_i) - 1}{NC_N(x_i)} \right],$$

with associated processes $\hat{Z}_N(t) = \sqrt{N}\,[\,\Lambda_N(t) - \Lambda(t)\,]$ and $Z_N(t) = \sqrt{N}\,[\,F_N(t) - F(t)\,]$, respectively. Note that $\Lambda_N = \Lambda_n$, $F_N = F_n$, $\hat{Z}_N = \sqrt{N/n}\,\hat{Z}_n$ and $Z_N = \sqrt{N/n}\,Z_n$. Keeping α fixed, the observed sample size n increases as N increases.

Our aim is to construct strong Gaussian approximations for $\hat{Z}_N$ and Z_N. For that purpose, we first recall the following version of Borisov's result as specialized to R^2 random variables. Let U_1, U_2, $\ldots$, be i.i.d. random variables in R^2, with distribution function denoted by $J(s)$, $s = (s_1, s_2) \in R^2$, and set $\alpha_N(s) := \sqrt{N}\,[\,J_N(s) - J(s)\,]$, where $J_N(s)$ is the empirical distribution function of U_1, $\ldots$, U_N. Also, let $W_J(\cdot)$ be the weak limit of $\alpha_N(\cdot)$, which is a random Gaussian field with zero mean and covariance $E[\,W_J(t)\,W_J(s)\,] = J(\min(t,s)) - J(t)\,J(s)$, for $t = (t_1, t_2)$, $s = (s_1, s_2) \in R^2$, where $\min(t,s) := (\,\min(t_1, s_1), \min(t_2, s_2)\,)$.

Theorem 2.A. (*Borisov* 1982) *On a rich enough probability space, there exist distributionally equivalent copies $W_{J,N}$ of W_J such that*

$$P\left\{ \sup_{s \in R^2} |\, \alpha_N(s) - W_{J,N}(s)\,| > C_1(\delta)\,\frac{\log N}{N^{1/6}} \right\} \le C_2\, N^{-\delta}$$

for any $\delta > 0$, where $C_1(\delta)$ and C_2 are constants.

For our case, $U = (X, Y)$ such that $Y \le X$, so $J(s) = \alpha H_*(s)$ is a sub-distribution function, $J_N(s) = H_N^*(s)$ and $\alpha_N(s) = \sqrt{N}\,[\,H_N^*(s) - \alpha H_*(s)\,]$ is a sub-empirical process. Note also that

$$\begin{aligned}\alpha_N^1(s_1) &: = \sqrt{N}\,[\,F_N^*(s_1) - \alpha F_*(s_1)\,] = \alpha_N(s_1, \infty),\\ \alpha_N^2(s_2) &: = \sqrt{N}\,[\,G_N^*(s_2) - \alpha G_*(s_2)\,] = \alpha_N(\infty, s_2).\end{aligned}$$

Let $W_N^1(s_1) := W_{J,n}(s_1, \infty)$ and $W_N^2(s_2) := W_{J,N}(\infty, s_2)$. Then, for all $\delta > 0$,

$$\begin{aligned}&P\Big\{\sup_{s_i \in R} \mid \alpha_N^i(s_i) - W_N^i(s_i) \mid > \, C_1(\delta)\,\frac{\log N}{N^{1/6}}, \quad i = 1, 2\Big\}\\ &\le P\Big\{\sup_{s \in R^2} \mid \alpha_N(s) - W_{J,N}(s) \mid > \, C_1(\delta)\,\frac{\log N}{N^{1/6}}\Big\} \le \frac{C_2}{N^\delta}.\end{aligned}$$

This can be arranged on a single rich enough probability space (Ω, A, P) by inserting one of the two suitably adjusted proofs of the KMT theorem, given by Mason and van Zwet (1987) and Bretagnolle and Massart (1989), into Borisov's proof. The Borel-Cantelli lemma then gives, almost surely,

$$\sup_{s_i \in R} \mid \alpha_N^i(s_i) - W_N^i(s_i) \mid = O\Big(\frac{\log N}{N^{1/6}}\Big), \qquad \text{jointly for } i = 1, 2. \tag{2.1}$$

Note also that the Gaussian processes $W_N^1(s_1)$, $W_N^2(s_2)$ can be expressed in terms of Brownian Bridges $\tilde{W}_N^1$, $\tilde{W}_N^2$ as $W_N^1(s_1) = \tilde{W}_N^1(\alpha F_*(s_1))$, and $W_N^2(s_2) = \tilde{W}_N^2(\alpha G_*(s_2))$.

Borisov's theorem provides strong approximation by a sequence of copies of the Gaussian limit. The next result, due to Csörgő and Horváth (1988), builds upon it to give an approximation by a two-parameter Gaussian process.

Theorem 2.B. *On a rich enough probability space, one can define a Gaussian process* $\{\,\Gamma(s, u),\ s \in R^2,\ u > 0\,\}$, *with* $E\,\Gamma(s, N) = 0$ *and* $E\,\Gamma(s, N)\,\Gamma(t, M) = (N \wedge M)\,[\,J(s \wedge t) - J(s)J(t)\,]$, $s, t \in R^2$, $N, M = 1, 2, \ldots$, *such that*

$$\sup_{s \in R^2} \Big|\ \alpha_N(s) - \frac{\Gamma(s, N)}{\sqrt{N}}\ \Big| = O\Big(\frac{(\log N)^{3/2}}{N^{1/8}}\Big)$$

almost surely.

We also use Bennett's inequality, as stated in Pollard (1984, p.192).

Theorem 2.C. *Let* $Z_1, \ldots, Z_N$ *be i.i.d. mean zero random variables with* $|Z_i| \le M$, *where* $M > 0$ *is a constant, and* $V = \sum_{i=1}^N \mathrm{Var}(Z_i)$. *Then for* $\eta \ge 0$,

$$P\,\{\,|\,Z_1 + \cdots + Z_N\,| > \,\eta\,\} \,\le\, 2\exp\Big\{-\frac{\eta^2}{2V}\,B\Big(\frac{M\eta}{V}\Big)\Big\},$$

where $B(\lambda) = 2\lambda^{-2}[\,(1+\lambda)\,\log(1+\lambda) - \lambda\,]$ *for* $\lambda > 0$.

3. Main Results

As mentioned in the introduction, in the presence of left truncation, the variances of the limiting processes may diverge as $a \to 0$. This reflects the uncertainty near 0 where an item X is very likely to be unobserved. To help control possible divergences as $a \to 0$, Stute (1993) assumed condition (1.2) and simplified the proof a great deal. On the other hand, Woodroofe (1985) considered the weak convergence of the modified process $\hat{Z}_n^a(t) := \hat{Z}_n(t) - \hat{Z}_n(a)$ and showed that condition (1.1a) is sufficient to guarantee finite variance of the limiting process even as $a \to 0$. Here we follow up and develop this line of thought. Instead of restricting consideration to fixed intervals $[a, b]$, as in all previous work under (1.1a), our construction will be over increasing sets $[a_N, b_N]$, where $a_N \to 0$ and $b_N \to b_F$ at appropriate rates. Modification of any function to $[a_N, b_N]$ will be denoted by affixing a_N as superscript. Thus, for example,

$$\hat{Z}_N^{a_N}(t) := \sqrt{N}\,[\,\Lambda_N^{a_N}(t) - \Lambda^{a_N}(t)\,] := \sqrt{N}\int_{a_N}^{t} \frac{dF_N^*}{C_N} - \sqrt{N}\int_{a_N}^{t} \frac{d\,\alpha F_*}{C_*},$$

and $Z_N^{a_N}(t) := \sqrt{N}\,[\,F_N^{a_N}(t) - F^{a_N}(t)\,]$, where

$$1 - F^{a_N}(t) := \begin{cases} \frac{1-F(t)}{1-F(a_N)}, & \text{if } t \geq a_N; \\ 1, & \text{otherwise,} \end{cases}$$
$$1 - F_N^{a_N}(t) = \prod_{i: a_N < x_i \leq t} \left[\frac{NC_N(x_i) - 1}{NC_N(x_i)} \right].$$

Here $\Lambda_N^{a_N}$ and $F_N^{a_N}$ are non-parametric MLE for Λ^{a_N} and F^{a_N}, respectively. Our modifications of Λ_N and F_N are analogous to those of Gu and Lai (1990).

Define, for $t \in [a_N, b_N]$, the Gaussian process

$$H_N^{a_N}(t) := \int_{a_N}^{t} \frac{W_N^1}{C_*^2}\, d\,\alpha G_* - \int_{a_N}^{t} \frac{W_N^2}{C_*^2}\, d\,\alpha F_* + \left.\frac{W_N^1}{C_*}\right|_{a_N}^{t}.$$

Clearly, $E\,H_N^{a_N}(t) = 0$. A lengthy but otherwise straightforward calculation (see Tse (1995)) gives the covariance of $H_N^{a_N}(t)$: $\mathrm{Cov}\,[\,H_N^{a_N}(s),\, H_N^{a_N}(t)\,] = l_{a_N}(s \wedge t) = l_{a_N}(s) \wedge l_{a_N}(t)$ for $s, t > a_N$, where $l_a(t) := \int_a^t (1/C_*^2(u))\,d\alpha F_*(u)$. Define $l_a^{-1}(t)$ as the generalized inverse function of $l_a(t)$ as in Burke, S. Csörgő and Horváth (1981) for the random censorship model. Then the covariance formula above implies that $W_N(\cdot) = H_N^{a_N}(l_{a_N}^{-1}(\cdot))$ is a standard Wiener process on $[0, \infty]$ for each N, and hence $H_N^{a_N}(t) = W_N(l_{a_N}(t))$, for $t \geq a_N$.

To parametrize the dependence of the approximation rates on the increasing sets $[a_N, b_N]$, we define $d_N := [C_*(l_N)]^{-1} = \{G(l_N)[1 - F(l_N-)]\}^{-1}$ where $l_N \in [\,a_N,\, b_N\,]$ such that $C_*(l_N) = \min_{a_N \le t \le b_N} C_*(t)$, and assume that

$$C_*(l_N) = G(l_N)[1 - F(l_N)] \ge \left(8\delta\, \frac{\log N}{N}\right)^{1/2}, \qquad \delta > 1. \tag{3.1}$$

This condition arises as a technical criterion in Lemma 4.1 below. It prevents d_N from increasing too fast, and thus limits the rate at which $a_N \to 0$ and $b_N \to b_F$. For instance, suppose that F and G have the expansions $F(z) = k_1 z^2 + O(z^3)$ and $G(z) = k_2 z + k_3 z^2 + O(z^3)$ as $z \downarrow 0$, where k_1, k_2 and k_3 are constants with $k_1, k_2 > 0$, so that condition (1.1a) is satisfied (but condition (1.2) is *not* satisfied in general), and $k_3 \ge 0$. Then, for condition (3.1) to be satisfied, it suffices to choose $a_N := (\,\log N/N\,)^{-\eta+1/2}$ for $1/8 > \eta > 0$. It is also easy to see that under these conditions F^{a_N} can be replaced by F in the process $Z_N^{a_N}$ in Theorems 3.2 and 3.4 below; analogous statements may be made for Theorems 3.1 and 3.3. Similar comments hold for b_N with the role of (1.1a) replaced by that of (1.1b).

The statements of the next two theorems hold in the probability space of (2.A).

Theorem 3.1. *If conditions* (1.1) *and* (3.1) *are satisfied, then, almost surely,*

$$\sup_{a_N \le t \le b_N} |\, \hat{Z}_N^{a_N}(t) - W_N(l_{a_N}(t))\,| = O\left(\, d_{2N}\, \frac{\log N}{N^{1/6}}\,\right).$$

Next, to approximate the product-limit process, we define the Gaussian process $K_N^{a_N}(t) := [1 - F^{a_N}(t)]\, H_N^{a_N}(t) = [1 - F^{a_N}(t)]\, W_N(l_{a_N}(t))$.

Theorem 3.2. *If conditions* (1.1) *and* (3.1) *are satisfied, then, almost surely,*

$$\sup_{a_N \le t \le b_N} |\, Z_N^{a_N}(t) - K_N^{a_N}(t)\,| = O\Big(\, d_{2N}\, \frac{\log N}{N^{1/6}}\Big).$$

Of course, these approximations are meaningful only if the rates go to zero as $N \to \infty$. This further restricts the speed at which $a_N \to 0$ and $b_N \to b_F$ even if Borisov's bivariate rate $\log N/\, N^{1/6}$ is improved. (A lower bound of the latter improvement is KMT's rate $\log N/\, N^{1/2}$; c.f. Beck (1985)).

Theorems 3.1 and 3.2 approximate $\hat{Z}_N$ and Z_N by sequences of their Gaussian limits. Weak convergence results follow immediately. However, almost sure statements cannot be obtained from them since the covariances between members in the sequences are not specified. In the next two theorems, these sequences are replaced by single two-parameter Gaussian processes. Almost sure statements, like the law of the iterated logarithm, can be derived from the results. Let $s = (s_1, s_2) \in R^2$. Define a two-parameter Gaussian process $\{ H^{a_N}(t,u),\ t \geq a_N,\ u \geq 0 \}$ in terms of $\Gamma(s,u)$ in Theorem 2.B such that

$$H^{a_N}(t,u) := \int_{a_N}^{t} \frac{\Gamma((s_1,\infty),u)}{\sqrt{u}\, C_*^2(s_1)}\, d\,\alpha G_*(s_1) - \int_{a_N}^{t} \frac{\Gamma((\infty,s_2),u)}{\sqrt{u}\, C_*^2(s_2)}\, d\,\alpha F_*(s_2) + \frac{\Gamma((s_1,\infty),u)}{\sqrt{u}\, C_*(s_1)} \Big|_{a_N}^{t}.$$

Also, define $K^{a_N}(t,u) := [\,1 - F^{a_N}(t)\,]\, H^{a_N}(t,u)$. It is straightforward to check (Tse (1995)) that the covariances of these processes are, for $N \leq M,\ s \leq t$,

$$\operatorname{Cov}[\, H^{a_N}(s,N),\ H^{a_M}(t,M)\,] = \sqrt{\frac{N}{M}}\, l_{a_N}(s),$$

$$\operatorname{Cov}[\, K^{a_N}(s,N),\ K^{a_M}(t,M)\,] = [\,1 - F^{a_N}(s)\,][\,1 - F^{a_M}(t)\,]\sqrt{\frac{N}{M}}\, l_{a_N}(s).$$

Hence, $W(v,u) = \sqrt{u}\, H^{a_N}(l_{a_N}^{-1}(v), u)$, $v,u \geq 0$, is a bivariate Wiener process. The statements of the next two theorems hold in the probability space of (2.B).

Theorem 3.3. *If conditions* (1.1) *and* (3.1) *are satisfied, then, almost surely,*

$$\sup_{a_N \leq t \leq b_N} |\, \hat{Z}_N^{a_N}(t) - \frac{W(l_{a_N}(t),N)}{\sqrt{N}}\,| = O\Big(d_{2N}\, \frac{(\log N)^{3/2}}{N^{1/8}} \Big).$$

Theorem 3.4. *If conditions* (1.1) *and* (3.1) *are satisfied, then, almost surely,*

$$\sup_{a_N \leq t \leq b_N} |\, Z_N^{a_N}(t) - K^{a_N}(t,N)\,| = O\Big(d_{2N}\, \frac{(\log N)^{3/2}}{N^{1/8}} \Big).$$

4. Proofs

For the sake of simplicity, we often denote $\sup_{a_N \leq t \leq b_N} |f(t)|$ by $\|f(\cdot)\|$, or even simpler, $\|f\|$. We start with the usual decomposition of $\hat{Z}_N^{a_N}(t)$:

$$\begin{aligned} \hat{Z}_N^{a_N}(t) &= \int_{a_N}^{t} \frac{d\alpha_N^1}{C_*} + \int_{a_N}^{t} \frac{\sqrt{N}\,(C_* - C_N)}{C_*^2}\, d\,\alpha F_* + R_{1N}(t) + R_{2N}(t) \\ &= \int_{a_N}^{t} \frac{\alpha_N^1\, d\,\alpha G_*}{C_*^2} - \int_{a_N}^{t} \frac{\alpha_N^2\, d\,\alpha F_*}{C_*^2} + \frac{\alpha_N^1}{C_*}\Big|_{a_N}^{t} + R_{1N}(t) + R_{2N}(t), \quad (4.1) \end{aligned}$$

where

$$R_{1N}(t) = \int_{a_N}^{t} \frac{\sqrt{N}\,(C_* - C_N)}{C_*^2}\, d(F_N^* - \alpha F_*),$$

$$R_{2N}(t) = \int_{a_N}^{t} \frac{\sqrt{N}\,(C_* - C_N)^2}{C_N\, C_*^2}\, dF_N^*.$$

Theorem 3.1 is about the order of $\|\hat{Z}_N^{a_N} - H^{a_N}\| = \|R_{1N} + R_{2N} + R_{3N}\|$, where

$$R_{3N}(t) = -\int_{a_N}^{t} \frac{\alpha_N^2 - W_N^2}{C_*^2}\, d\,\alpha F_* + \left(\int_{a_N}^{t} \frac{\alpha_N^1 - W_N^1}{C_*^2}\, d\,\alpha G_* + \frac{\alpha_N^1 - W_N^1}{C_*} \Bigg|_{a_N}^{t} \right).$$

For ease of comparison, note that our definitions of R_{1N} and R_{2N} are analogous to $\sqrt{n}R_{n1}$ and $\sqrt{n}R_{n2}$ in Stute (1993), except that we have N in place of n and varying a_N. The treatment of these two terms differs from Stute's on two accounts. First, the approximation rates depend on the increasing sets $[a_N, b_N]$ through d_N, which is a constant for fixed intervals strictly contained in $(0, b_F)$. Second, we assume the weaker condition (1.1). To lower the exponent of the d_N factor in the rate, we invoke Lemmas 4.3 and 4.4 below. For R_{1N}, we adapt Stute's results (1994) on U-statistics processes to the present case. The relevant analysis is in Lemma 4.6. For R_{2N}, we start by using the Dvoretzky-Kiefer-Wolfowitz (DKW) inequality to obtain, in Lemma 4.1 under (3.1), a bound for C/C_N on $\{w_i : a_N \le w_i \le b_N\}$, $w_i = x_i$, or y_i. The remaining factors in R_{2N} are dealt with by Lemmas 4.4 and 4.5; the latter is a suitable form of Csáki's (1975) asymptotic result for a standardized empirical process. Finally, we use (2.1) to treat R_{3N}. For a fixed interval $[a, b]$ strictly contained in $(0, b_F)$, d_N is a finite constant and the theorems are valid under condition (1.1) only. If we are interested in the random variable X only, we can set $a = 0$ by replacing Lemma 4.1 with Stute's Corollary 1.3. The statements of the theorems then remain valid under condition (1.2), with the supremum taken over $(0, b]$ and the approximation rates multiplied by a factor of $(\log N)^{1/2}$.

Let w_i denote either x_i or y_i, and note that since F and G are assumed to be continuous, $F_*(t-) = F_*(t)$ for every t.

Lemma 4.1. *If condition* (3.1) *is satisfied, then for any* $\delta > 0$,

$$P\left\{ \sup_{i: a_N \le w_i \le b_N} \frac{1}{C_N(w_i)} > \frac{2}{C_*(l_N)} \right\} \le 4\, N^{-\delta}$$

and $\sup_{i: a_N \le w_i \le b_N} C_*(w_i)/C_N(w_i) \le 2$ *almost surely.*

Proof. By condition (3.1) and the DKW inequality as specified by Massart (1990), the left-hand side is

$$P\left\{ \sup_{i: a_N \le w_i \le b_N} [\, F_N^*(w_i-) - G_N^*(w_i) \,] > \frac{-C_*(l_N)}{2} \right\}$$

$$\leq P\left\{\|F_N^* - \alpha F_*\| > \frac{C_*(l_N)}{4}\right\} + P\left\{\|G_N^* - \alpha G_*\| > \frac{C_*(l_N)}{4}\right\} \leq 4\,N^{-\delta}.$$

The Borel-Cantelli lemma then gives the second statement.

Lemma 4.2. *For $s \leq t_0 \leq t$ and $k > 1$ we have*

$$\int_{t_0}^{t} \frac{dF}{(1-F)^k} \leq \frac{1}{(k-1)\,[\,1-F(t)\,]^{k-1}} \quad \text{and} \quad \int_{s}^{t_0} \frac{dG}{G^k} \leq \frac{1}{(k-1)\,G^{k-1}(s)}.$$

Proof. The statements follow from monotonicity and the assumed continuity of F and G respectively.

Lemma 4.3. *If F and G satisfy the conditions in* (1.1)*, then for $k > 1$, we have*

$$\Big\| \int_{a_N}^{\cdot} \frac{dF}{C_*^k} \Big\| = O(d_N^{k-1}) \quad \text{and} \quad \Big\| \int_{a_N}^{\cdot} \frac{dG}{C_*^k} \Big\| = O(d_N^{k-1}).$$

Proof. Recall that $C_* = G(1-F)$. For the first statement, write

$$\int_{a_N}^{t} \frac{dF}{C_*^k} = \left(\int_{a_N}^{t_0} + \int_{t_0}^{t}\right) \frac{dF}{C_*^k}, \qquad a_N < t_0 < t \leq b_N,$$

with a t_0 to be specified. Now, using Lemma 4.2,

$$\int_{t_0}^{t} \frac{dF}{C_*^k} \leq \frac{1}{G^k(t_0)} \int_{t_0}^{t} \frac{dF}{(\,1-F\,)^k} < \frac{1}{(k-1)\,G^k(t_0)\,C_*^{k-1}(t)}.$$

Also,

$$\int_{a_N}^{t_0} \frac{dF}{C_*^k} \leq \left(\frac{1}{[\,1-F(t_0)\,]^k} \int_{a_N}^{t_0} \frac{dF}{G} \right) \frac{1}{C_*^{k-1}(a_N)}.$$

By assumption, $\int_{a_N}^{t_0} \frac{dF}{G} < \infty$, and t_0 can be chosen so that $1/G(t_0)$ and $1/[1 - F(t_0)]$ are finite. For example, take t_0 to be the average of the medians of F and G. This is possible since $a_F = a_G = 0$ and $b_G \leq b_F$. The first statement follows. The proof for the second statement is analogous.

Lemma 4.4. *If conditions* (1.1) *and* (3.1) *are satisfied, then, almost surely,*

$$\Big\| \int_{a_N}^{\cdot} \frac{dF_N^*}{C_*^2} \Big\| = O\,(\,d_N\,).$$

Proof. Write

$$\Big\| \int_{a_N}^{\cdot} \frac{dF_N^*}{C_*^2} \Big\| \leq \Big\| \int_{a_N}^{\cdot} \frac{d(F_N^* - \alpha F_*)}{C_*^2} \Big\| + \Big\| \int_{a_N}^{\cdot} \frac{d\,\alpha F_*}{C_*^2} \Big\|.$$

The second term is $O(d_N)$ by Lemma 4.3. With a later application in mind, we prove a more general statement for the first term. For $k = 1, 2$, let

$$Z_i^k(t) = \frac{I(a_N \le X_i \le t, Y_i \le X_i)}{C_*^k(X_i)},$$

so that $|Z_i^k(t)| \le d_N^k$ and, by Lemma 4.3, Var $Z_i^k(t) = K\, d_N^{2k-1}$ where K is a finite constant. Now

$$\Big\| \int_{a_N}^{\cdot} \frac{d\,\alpha_N^1}{C_*^k} \Big\| = \Big\| \frac{1}{\sqrt{N}} \sum_{i=1}^{N} [\, Z_i^k(\cdot) - E\, Z_i^k(\cdot)\,] \Big\| .$$

Bennett's inequality gives, for any $\gamma > 0$,

$$P\Big(\Big| \int_{a_N}^{t} \frac{d\alpha_N^1}{C_*^k} \Big| > \sqrt{\gamma d_N^{2k-1} \log N}\Big) \le 2 \exp\Big\{ -\frac{\gamma \log N}{2\,K} B\Big(\sqrt{\frac{\gamma d_N \log N}{N}} \frac{1}{K}\Big)\Big\}.$$

Since $d_N \log N / N = o(1)$ by condition (3.1), and $1 > B(\lambda) > 0$ for any finite $\lambda > 0$, we can choose $\gamma > 0$ such that the last expression is $O(N^{-\gamma})$. The Borel-Cantelli lemma then gives, almost surely,

$$\max_{t \in A_N} \Big| \int_{a_N}^{t} \frac{d\alpha_N^1}{C_*^k} \Big| = O\Big(\sqrt{d_N^{2k-1} \log N}\Big)$$

for every finite set A_N of $t's$ with $a_N \le t \le b_N$. Since F_* is continuous and

$$\int_{a_N}^{\cdot} \frac{\sqrt{N}\, d\,\alpha F_*}{C_*^k} \;=\; O(1) \sqrt{d_N^{2k-1} \log N}\,,$$

we may choose $A_N = \{t_{Nj} : 1 \le j \le M\}$ such that $a_N = t_{N1} < \cdots < t_{NM} = b_N$ and

$$\int_{t_j}^{t_{j+1}} \frac{\sqrt{N}\, d\,\alpha F_*}{C_*^k} \;\le\; \sqrt{d_N^{2k-1} \log N}$$

for $j = 1, \ldots, M-1$. Monotonicity of Λ^{a_N} and $\Lambda_N^{a_N}$ then gives

$$\Big\| \int_{a_N}^{\cdot} \frac{d\alpha_N^1}{C_*^k} \Big\| \le \sup_{t \in A_N} \Big| \int_{a_N}^{t} \frac{d\alpha_N^1}{C_*^k} \Big| + O\Big(\sqrt{d_N^{2k-1} \log N}\Big) = O\Big(\sqrt{d_N^{2k-1} \log N}\Big). \tag{4.2}$$

With condition (3.1), the $k = 2$ case gives the first term of the lemma.

Lemma 4.5. *If conditions* (1.1) *are satisfied, then, almost surely, for every* $\epsilon > 0$, *we have*

$$\Big\| \frac{\sqrt{N}\,(C_* - C_N)}{\sqrt{C_*}} \Big\| = O\Big(\sqrt{(\log N)^{1+\epsilon}}\Big).$$

Proof. Recalling that $\sqrt{N}\,(C_* - C_N) = \alpha_N^1 - \alpha_N^2$ and $C_* = G(1-F)$, we have

$$\left\| \frac{\sqrt{N}\,(C_* - C_N)}{\sqrt{C_*}} \right\| \le \left\| \frac{\alpha_N^1}{\sqrt{G(1-F)}} \right\| + \left\| \frac{\alpha_N^2}{\sqrt{G(1-F)}} \right\|.$$

The first term is bounded by

$$\left\| \sqrt{\frac{F}{G}} \right\| \left\| \sqrt{\frac{\alpha F_*\,(1-\alpha F_*)}{F\,(1-F)}} \right\| \left\| \frac{\alpha_N^1}{\sqrt{\alpha F_*\,(1-\alpha F_*)}} \right\|.$$

By (1.1a) and the definition of F_*, the first two factors are bounded. The last factor is, almost surely, $O(\,(\log N)^{(1+\epsilon)/2}\,)$ for each $\epsilon > 0$ by Csáki's (1975) result. The treatment of the second term is analogous.

Proof of Theorem 3.1. Recall $\hat{Z}_N^{a_N}(t) - H^{a_N}(t) = R_{1N}(t) + R_{2N}(t) + R_{3N}(t)$. By (2.1) and Lemma 4.3, $\|R_{3N}\| = O(\,d_N \log N/N^{1/6}\,)$ almost surely. By Lemmas 4.1, 4.5 and 4.4, we also have, almost surely,

$$\|R_{2N}\| \le \frac{1}{\sqrt{N}} \left\| \frac{\sqrt{N}\,(C_* - C_N)}{\sqrt{C_*}} \right\|^2 \left\| \int_{a_N}^{\cdot} \frac{2\,dF_N^*}{C_*^2} \right\| = O\left(d_N \, \frac{(\log N)^{1+\epsilon}}{\sqrt{N}} \right)$$

for any $\epsilon > 0$. Lastly, Stute (1993) noted that there is a double sum of the form

$$\frac{1}{N} \sum_i \sum_j h_N(X_i, Y_i; X_j, Y_j)$$

in $R_{1N}(t)$ which may be decomposed into its diagonal and off-diagonal parts. Thus,

$$\sqrt{N}\,R_{1N}(t) = \int_{a_N}^t \frac{dF_N^*}{C_*} - \int_{a_N}^t \frac{dF_N^*}{C_*^2} + \int_{a_N}^t \frac{C_N - C_*}{C_*^2} d\,\alpha F_* + \frac{1}{N}\,[D_N^1(t) - D_N^2(t)],$$

where $D_N^i(t)$, $i = 1, 2$, is a degenerate U-statistic for each t. The first three terms and the last term correspond to the diagonal and the off-diagonal part of the double sum respectively. By Lemmas 4.3 to 4.5, the sup norm of the first three terms over $[a_N, b_N]$ are $O(d_N)$ almost surely. The last term is $o(d_{2N}(\log N)^{\epsilon+1/2})$ by Lemma 4.6 below. This completes the proof.

Lemma 4.6. *If conditions* (1.1) *and* (3.1) *are satisfied, then, almost surely,*

$$\left\| \frac{D_N^i}{N} \right\| = o\,(\,d_{2N}\,(\log N)^{\epsilon+1/2}\,), \qquad i = 1, 2.$$

Proof. By Theorem 5 of Stute (1994), with $p = 2$, we have $E\,[\,\|\,D_N^i\,\|^2\,] \le (\,K_N^i\,N\,)^2$ for $i = 1$, 2, with

$$K_N^1 \le A \left[\int_{a_N}^{b_N} \int_{a_N}^{b_N} \frac{I(z \ge y)}{C_*^4(y)} d\,\alpha F_*(y) d\,\alpha F_*(z) \right]^{1/2} \le M\,d_N,$$

$$K_N^2 \le A \left[\int_{a_N}^{b_N} \int_{a_N}^{b_N} \frac{I(z \ge y)}{C_*^4(y)} d\,\alpha F_*(y) d\,\alpha G_*(z) \right]^{1/2} \le M\, d_N,$$

where A and M are finite constants. Consider the blocks $2^k \le N \le 2^{k+1}$ for $k \ge 1$. For each fixed t, $D_N^i(t)/N\,(N-1)$ is a reverse time martingale in N (Berk (1966)). Hence, $\|D_N^i\|/N\,(N-1)$ is a non-negative submartingale in reverse time for $2^k \le N \le 2^{k+1}$. Setting $B_N := N/\,(\log N)^\beta$, since d_N is non-decreasing, we obtain

$$\begin{aligned} P\Big\{ \frac{\|D_N^i\|}{N^2} \ge M\, \frac{d_{2N}\,\epsilon}{B_N} \Big\} &\le P\Big\{ B_{2^{k+1}} \max_{2^k \le N \le 2^{k+1}} \frac{\|D_N^i\|}{N(N-1)\, M\, d_{2^{k+1}}} \ge \epsilon \Big\} \\ &\le \frac{B_{2^{k+1}}^2}{\epsilon^2\,[\,2^k\,(2^k-1)\,]^2}\, E\left[\sup_{a_{2^k} \le t \le b_{2^k}} \left| \frac{D_{2^k}^i(t)}{M\, d_{2^{k+1}}} \right|^2 \right] \\ &\le \frac{B_{2^{k+1}}^2}{\epsilon^2} \frac{2^{2k}}{[\,2^k\,(2^k-1)\,]^2} = O\left(k^{-2\,\beta} \right) \end{aligned}$$

for each $\epsilon > 0$. Hence, if $\beta > 1/2$, Borel-Cantelli lemma gives the result.

Note that the rate in Theorem 3.1 is $O(\,d_{2N}\ \log N/\,N^{1/6}\,)$, whereas $\|R_{1N}\| = o(\,d_{2N}\,(\log N)^{1/2}/\sqrt{N}\,)$ and $\|R_{2N}\| = O(\,d_N\,(\log N)^{1+\epsilon}/\sqrt{N}\,)$. The $N^{-1/6}$ factor comes from (2.1). Any improvement in Borisov's rate will result in the same here.

Next, we turn to the proof of Theorem 3.2. Noting that $\log[\,1 - F_N^{a_N}(t)\,]$ may yield $\log 0$, we define a slight modification of $F_N^{a_N}$. Let

$$1 - \bar F_N^{a_N}(t) := \prod_{a_N \le x_i \le t} \frac{NC_N(x_i)}{NC_N(x_i)+1}, \qquad a_N \le t \le b_N.$$

Lemma 4.7. *Suppose conditions* (1.1) *and* (3.1) *are satisfied. Then, almost surely,* $\|\,\bar F_N^{a_N} - F_N^{a_N}\,\| = O(\,d_N/N\,)$.

Proof. Using Lemmas 4.1, 4.4 and $|\prod_{i=1}^n a_i - \prod_{i=1}^n b_i| \le \sum_{i=1}^n |a_i - b_i|$ for $|a_i|, |b_i| \le 1$, we have,

$$\|\,\bar F_N^{a_N} - F_N^{a_N}\,\| \le \left\| \int_{a_N}^{\cdot} \frac{dF_N^*(u)}{C_N(u)\,[\,N\,C_N(u)+1\,]} \right\| = O\left(\frac{d_N}{N} \right)$$

almost surely.

Lemma 4.8. *Suppose conditions* (1.1) *and* (3.1) *are satisfied. Then, almost surely,* $\|-\log[\,1-\bar F_N^{a_N}\,] - \Lambda_N^{a_N}\,\| = O(\,d_N/N)$.

Proof. Using the definitions of $\bar F_N^{a_N}$ and $\Lambda_N^{a_N}$, the left hand side is bounded by

$$\left\| -\log \prod_{a_N \le x_i \le \cdot} \Big[1 - \frac{1}{N\,C_N(x_i)+1} \Big] - \sum_{a_N \le x_i \le \cdot} \frac{1}{N\,C_N(x_i)+1} \right\|$$

$$+ \Big\| \sum_{a_N \le x_i \le \cdot} \Big[\frac{1}{N\, C_N(x_i) + 1} - \frac{1}{N\, C_N(x_i)} \Big] \Big\| .$$

Applying the algebraic inequality

$$0 \ < \ -\log\left[1 - \frac{1}{1+x}\right] - \frac{1}{1+x} \ < \ \frac{1}{x(x+1)}, \qquad \text{for } x > 0,$$

the last expression is bounded by $O(\, d_N/N\,)$ almost surely in view of Lemmas 4.1 and 4.4.

Next, for $a_N \le t \le d_N$, let $L_N^{a_N}(t) = \sqrt{N}\{\,-\log[\,1 - \bar{F}_N^{a_N}(t)\,] - \Lambda^{a_N}(t)\,\}$.

Lemma 4.9. *Suppose conditions* (1.1) *and* (3.1) *are satisfied. Then, almost surely,* $\|L_N^{a_N} - W_N(\, l_{a_N}\,)\| = O(\, d_{2N} \log N/N^{1/6}\,)$.

Proof. This follows from Lemma 4.8 and Theorem 3.1.

Proof of Theorem 3.2. Noting that $1 - F^{a_N} = \exp(\,-\Lambda^{a_N}\,)$, we have

$$\begin{aligned} \|\, Z_N^{a_N} - K_N^{a_N}\,\| &= \Big\| \ \sqrt{N}\,[\,F_N^{a_N} - F^{a_N}\,] - (1 - F^{a_N})\, W_N(\, l_{a_N}\,) \ \Big\| \\ &= \Big\| \ \sqrt{N}\,\Big[[\,1 - \bar{F}_N^{a_N}\,] - \exp[\,-\Lambda^{a_N}\,]\Big] + (1 - F^{a_N})\, L_N^{a_N} \ \Big\| \\ &\qquad + O\left(\frac{d_N}{\sqrt{N}}\right) + O\left(d_{2N}\,\frac{\log N}{N^{1/6}}\right) \end{aligned}$$

almost surely by Lemmas 4.7 and 4.9. Now apply a Taylor expansion around $\exp(-\Lambda^{a_N}(t))$ for the first term, with an intermediate point $\theta_N(t)$:

$$\min\Big(\log[\,1 - \bar{F}_N^{a_N}(t)\,],\ -\Lambda^{a_N}(t)\Big) \ \le\ \theta_N(t) \ \le\ \max\Big(\log[\,1 - \bar{F}_N^{a_N}(t)\,],\ -\Lambda^{a_N}(t)\Big),$$

so that $\theta_N(t) \le 0$. Note that the power one term in the expansion cancels with $(1 - F^{a_N}(t))\, L_N^{a_N}(t)$. Hence the last expression is bounded by

$$\Big\| \ \frac{\sqrt{N}}{2}\,[\,\Lambda_N^{a_N} - \Lambda^{a_N}\,]^2 \ \Big\| + O\Big(\, d_{2N}\,\frac{\log N}{N^{1/6}}\Big).$$

Applying Lemma 4.10 below, we get the desired result.

Lemma 4.10. *Suppose conditions* (1.1) *and* (3.1) *are satisfied. Then, almost surely,* $\|\hat{Z}_N^{a_N}\| = O(\, d_N\,(\log N)^{(1+\epsilon)/2}\,)$.

Proof. By (4.1),

$$\hat{Z}_N^{a_N}(t) = \int_{a_N}^{t} \frac{d\alpha_N^1}{C_*} + \int_{a_N}^{t} \frac{\sqrt{N}\,(C_* - C_N)}{C_*^2}\, d\,\alpha F_* + R_{1N}(t) + R_{2N}(t).$$

The first term corresponds to the $k = 1$ case of (4.2). Also, by Lemmas 4.5 and 4.3,

$$\Big\| \int_{a_N}^{\cdot} \frac{\sqrt{N}\,(C_* - C_N)}{C_*^2} d\,\alpha F_* \ \Big\| = O\Big(\ \sqrt{d_N (\log N)^{1+\epsilon}}\ \Big)$$

almost surely. Finally, from the proof of Theorem 3.1 and condition (3.1), we have $\| R_{1N} + R_{2N} \| = O([d_N (\log N)^{1+\epsilon}]^{1/2})$ almost surely.

Proofs of Theorems 3.3 and 3.4. These are the same as the proofs of Theorems 3.1 and 3.2, with the roles of Theorem 2.A and Theorem 3.1 taken by Theorem 2.B and Theorem 3.3, respectively.

Acknowledgement

The author is greatly indebted to Professor Sándor Csörgő for suggesting this problem and providing much helpful advice. The author would also like to thank the editor, two referees and Professor Michael Woodroofe for their valuable comments. This work is partially supported by NSF Grant DMS-9203357.

References

Beck, J. (1985). Lower bounds on the approximation of the multivariate process. *Z. Wahrsch. verw. Gebiete* **70**, 289-306.

Berk, R. H. (1966). Limiting behavior of posterior distributions when the model is incorrect. *Ann. Math. Statist.* **37**, 51-58.

Borisov, I. S. (1982). Approximation of empirical fields constructed with respect to vector observations with dependent components. *Sibirskii Matematicheskii Zhurnal* **23**, 31-41.

Bretagnolle, J. and Massart, P. (1989). Hungarian constructions from the nonasymptotic viewpoint. *Ann. Statist.* **1**, 239-256.

Burke, M. D., Csörgő, S. and Horváth, L. (1981). Strong approximations of some biometric estimates under random censorship. *Z. Wahrsch. verw. Gebiete* **56**, 87-112.

Burke, M. D., Csörgő, S. and Horváth, L. (1988). A correction to and improvement of "Strong approximations of some biometric estimates under random censorship". *Probab. Theory Related Fields* **79**, 51-57.

Chao, M. T. and Lo, S. H. (1988). Some representations of the nonparametric maximum likelihood estimators with truncated data. *Ann. Statist.* **16**, 661-668.

Csáki, E. (1975). Some notes on the law of the iterated-logarithm for empirical distribution function. In *Colloq. Math. Soc. János Bolyai* **11**, *Limit Theorems of Probability Theory* (Edited by P. Révész), 47-57. North Holland, Amsterdam.

Csörgő, M. and Horváth, L. (1988). A note on strong approximations of multivariate empirical processes. *Stoch. Process and Their Applications* **27**, 101-109.

Csörgő, M. and Révész, P. (1981). *Strong Approximations in Probability and Statistics.* Academic Press, New York.

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

Gu, M. G. and Lai, T. L. (1990). Functional laws of the iterated logarithm for the product-limit estimator of a distribution function under random censorship or truncation. *Ann. Probab.* **18**, 160-189.

Komlós, J., Major, P. and Tusnády, G. (1975). An approximation of partial sums of independent r.v.'s and the sample d.f. *Z. Wahrsch. verw. Gebiete* **32**, 111-132.

Lynden-Bell, D. (1971). A method of allowing for known observational selection in small samples applied to 3CR quasars. *Monthly Notices Roy. Astronom. Soc.* **155**, 95-118.

Mason, D. M. and van Zwet, W. R. (1987). A refinement of the KMT inequality for the uniform empirical process. *Ann. Probab.* **15**, 871-884.

Massart, P. (1990). The tight constant in the Dvoretzky-Kiefer-Wolfowitz inequality. *Ann. Probab.* **18**, 1269-1283.

Pollard, D. (1984). *Convergence of Stochastic Processes*. Springer-Verlag, New York.

Stute, W. (1993). Almost sure representations of the product-limit estimator for truncated data. *Ann. Statist.* **21**, 146-156.

Stute, W. (1994). Integrated U-statistic processes: a martingale approach. *Ann. Probab.* **22**, 1725-1744.

Tse, S. M. (1995). *Gaussian Approximations in the Random Truncation Model*. PhD dissertation, U. of Michigan.

Wang, M. C., Jewel, N. P. and Tsai, W. Y. (1986). Asymptotic properties of the product-limit estimate under random truncation. *Ann. Statist.* **14**, 1597-1605.

Woodroofe, M. (1985). Estimating a distribution function with truncated data. *Ann. Statist.* **13**, 163-177.

Department of Applied Mathematics, National Donghua University, Hualien, Taiwan.

E-mail: tse@cc.ndhu.edu.tw

(Received May 1997; accepted March 1999)