STRONG REPRESENTATION OF THE QUANTILE FUNCTION FOR LEFT TRUNCATED AND DEPENDENT DATA

M Lemdani, E Ould-Saïd, Nicolas Poulin

Volume 14, No. 3 (2005), pp. Allerton Press, Inc.

MATHEMATICAL METHODS OF STATISTICS

STRONG REPRESENTATION OF THE QUANTILE FUNCTION FOR LEFT TRUNCATED AND DEPENDENT DATA

M. Lemdani[1], E. Ould-Saïd[2], and N. Poulin[2]

[1]Lab. de Biomathématiques, EA 3614, Univ. de Lille 2, Fac. de Pharmacie
3, rue du Pr. Laguesse, 59006 Lille, France
E-mail: mlemdani@pharma.univ-lille2.fr

[2]L.M.P.A. J. Liouville, Univ. du Littoral Côte d'Opale
BP 699, 62228 Calais, France
E-mail: ouldsaid@lmpa.univ-littoral.fr, poulin@lmpa.univ-littoral.fr

Let $(\mathbf{X}_i)_{i\geq 1}$ be a sequence of strong-mixing random variables with common unknown absolutely continuous distribution function F subject to random left truncation. Let $F^{-1}(p)$ denote the pth $(p \in]0,1[)$ quantile function of the marginal distribution function of the $\mathbf{X}_i$'s which is estimated by the sample quantile $\widehat{F}_n^{-1}(p)$. In this paper, we derive the strong consistency and a strong representation for $\widehat{F}_n^{-1}(p)$, the quantile function of the Lynden-Bell estimator of F for strong-mixing processes.

Key words: quantile function, strong-mixing, Lynden-Bell estimator, truncated data, strong representation, strong consistency.

2000 Mathematics Subject Classification: 62G05, 62G20.??????

1. Introduction

Consider a sequence of random variables (rv's) $\mathbf{X}_1, \mathbf{X}_2, \ldots, \mathbf{X}_N$, which may not be mutually independent, with common unknown distribution function (df) F with continuous density f. These rv's are regarded as the lifetimes of the items under study. Among the various forms in which incomplete data appear, right censoring and left truncation are two common ones. Left truncation may occur if the time origin of the lifetime precedes the time origin of the study. Only the subjects that fail after the start of the study are to be followed, otherwise they are truncated. This model arises in various fields, e.g., astronomy, economy, and medical studies (see, e.g., Woodroofe (1985)). Let $\mathbf{T}_1, \mathbf{T}_2, \ldots, \mathbf{T}_N$ be a sequence of independent

and identically distributed (iid) truncation rv's with continuous df G; they are also assumed to be independent of the $\mathbf{X}_i$'s. In the left truncation model $(\mathbf{X}_i, \mathbf{T}_i)$ is observed only when $\mathbf{X}_i \geq \mathbf{T}_i$ (otherwise neither $\mathbf{X}_i$ nor $\mathbf{T}_i$ are observed). Let $(X_1, T_1), \ldots, (X_n, T_n)$ be the actually observed sample (i.e., $X_i \geq T_i$) and put $\gamma := \mathbf{P}(\mathbf{T}_1 \leq \mathbf{X}_1) > 0$, where $\mathbf{P}$ is the absolute probability (related to the N-sample). Note that n itself is a rv and that γ can be estimated by n/N (although this estimator cannot be calculated since N is unknown). For any df W we denote the left and right endpoints of its support by $a_W = \inf\{x : W(x) > 0\}$ and $b_W = \sup\{x \colon W(x) < 1\}$, respectively. As discussed by Woodroofe (1985), we assume that $a_G \leq a_F$ and $b_G \leq b_F$, in order to ensure the identifiability of the model. Furthermore, we will require the (slightly) stronger condition $a_G < a_F$ (see Remark 1). Finally, without loss of generality, we put $a_F = 0$ and then suppose the X_i's positive.

Define

$$C(x) = \mathbf{P}\big(\mathbf{T}_1 \leq x \leq \mathbf{X}_1 \mid \mathbf{T}_1 \leq \mathbf{X}_1\big) = \mathbb{P}(T_1 \leq x \leq X_1) = \gamma^{-1} G(x)(1 - F(x)), \tag{1}$$

where $\mathbb{P}(\cdot) = \mathbf{P}(\cdot \mid n)$ is the conditional probability (related to the n-sample) and consider the empirical estimate

$$C_n(x) = \frac{1}{n} \sum_{i=1}^{n} \mathbf{1}_{\{T_i \leq x \leq X_i\}}, \tag{2}$$

where $\mathbf{1}_{\{\cdot\}}$ is the indicator function.

Then the product-limit (PL) estimator $\widehat{F}_n$ of F was obtained by Lynden-Bell (1971),

$$\widehat{F}_n(x) = 1 - \prod_{X_i \leq x}{}' \left(1 - \frac{r_n(X_i)}{n C_n(X_i)}\right), \tag{3}$$

where $r_n(x) = \#\{j \leq n,\, X_j = x\}$ and $\prod'$ is the product over pairwise distinct X_i's. The almost sure convergence of this estimator was given by Woodroofe (1985), Wang, Jewell and Tsai (1986), Gu and Lai (1990), Keiding and Gill (1990), and Lai and Ying (1991) under various models. Strong representations of $\widehat{F}_n$ as a mean of some rv's with rate were stated by Chao and Lo (1988), Stute (1993), Gijbels and Wang (1993), Zhou (1996), Zhou and Yip (1999), and Tse (2000) in the iid framework. Note that some of these papers consider data that are both truncated and censored.

The cumulative hazard function $\Lambda(x)$ is defined by

$$\Lambda(x) = \int_0^x \frac{dF(u)}{1 - F(u)}. \tag{4}$$

Let

$$F^*(x) = \mathbf{P}\big(\mathbf{X}_1 \leq x \mid \mathbf{T}_1 \leq \mathbf{X}_1\big) = \mathbb{P}(X_1 \leq x) = \gamma^{-1} \int_0^x G(u)\, dF(u) \tag{5}$$

be the df of the observed lifetimes. Its empirical estimator is given by

$$F_n^*(x) = \frac{1}{n}\sum_{i=1}^{n} \mathbf{1}_{\{X_i \le x\}}. \tag{6}$$

The density of F^* is given by

$$f^*(x) = \gamma^{-1} G(x) f(x).$$

On the other hand, the df of the observed T_i's is given by

$$G^*(x) = \mathbf{P}\big(\mathbf{T}_1 \le x \mid \mathbf{T}_1 \le \mathbf{X}_1\big) = \mathbb{P}(T_1 \le x) = \gamma^{-1} \int_0^{+\infty} G(x \wedge u)\, dF(u)$$

and is estimated by

$$G_n^*(x) = \frac{1}{n}\sum_{i=1}^{n} \mathbf{1}_{\{T_i \le x\}}.$$

It then follows from (1) and (2) that

$$C(x) = G^*(x) - F^*(x^-) \qquad \text{and} \qquad C_n(x) = G_n^*(x) - F_n^*(x^-). \tag{7}$$

Finally (1), (4), and (5) give

$$\Lambda(x) = \int_0^x \frac{dF^*(u)}{C(u)}.$$

Hence, a natural estimator of Λ is given by

$$\widehat{\Lambda}_n(x) = \int_0^x \frac{dF_n^*(u)}{C_n(u)} = \sum_{i=1}^{n} \frac{\mathbf{1}_{\{X_i \le x\}}}{nC_n(X_i)},$$

which is the usual so-called Nelson–Aalen estimator of Λ. Moreover, $\widehat{\Lambda}_n$ is the cumulative hazard function of the PL estimator $\widehat{F}_n$ defined in (3).

Now, for $0 < p < 1$, the pth quantile of $F(t)$ is defined by

$$F^{-1}(p) = \inf\big\{x \in \mathbb{R}\colon F(x) \ge p\big\} \tag{8}$$

and its sample estimator is given by

$$\widehat{F}_n^{-1}(p) = \inf\big\{x \in \mathbb{R}\colon \widehat{F}_n(x) \ge p\big\}. \tag{9}$$

In this paper we consider the strong-mixing dependence, which amounts to a form of asymptotic independence between the past and the future as shows its definition.

Definition 1. Let $\{Z_i;\, i \ge 1\}$ denote a sequence of stationary rv's and consider the family of σ-fields $\mathcal{F}_j^l := \sigma\{Z_i;\, j \le i \le l\}$. Given a positive integer ℓ, set for all fixed $k \ge 1$,

$$\alpha(\ell) = \sup\big\{|\mathbb{P}(A \cap B) - \mathbb{P}(A)\mathbb{P}(B)| :\, A \in \mathcal{F}_1^k,\, B \in \mathcal{F}_{k+\ell}^{\infty}\big\}. \tag{10}$$

The sequence of rv's is said to be α-mixing (strongly mixing) if the mixing coefficients satisfy

$$\lim_{\ell\to\infty} \alpha(\ell) = 0. \tag{11}$$

Among various mixing conditions used in the literature, α-mixing is reasonably weak and has many practical applications. In particular, the stationary autoregressive–moving average (ARMA) processes, which are widely applied in time series analysis, are α-mixing with an exponential decreasing rate.

Sun and Zhou (2001) obtained strong representations for both the PL and the Nelson–Aalen estimators in the case of truncated dependent data. Gűrler, Stute and Wang (1993) obtained weak and strong representations in the iid case for the quantile function $\widehat{F}_n^{-1}$. As our aim is to generalize their results to the dependent case, we emphasize here the main differences from the iid case: first, we have to deal with the problem of ties which can happen in the α-mixing case with a nonzero probability (although we prove in Lemma 1 that there are not "too many" of them). Then, since Bennett's inequality cannot be used, we recall a large deviation result due to Bosq (Lemma 4). Finally, the strong approximation results which are available for the α-mixing case are not as precise as in the iid case (the remainder term rate is $n^{-1/2}\log^{-\zeta} n$ whereas it is almost n^{-1} for iid data), therefore our rate for the quantile does not achieve the classical Bahadur's rate (with magnitude of almost $n^{-3/4}$). Anyway it is enough if we want to derive practical results such as asymptotic normality (see Theorem 2).

In this paper, the assumptions and results will be formulated with respect to the actually observed sample (of size n), that is using the probability $\mathbb{P}$. In the following section we give the assumptions under which we state our results. The proofs are given in the last section. The convergence of $\widehat{F}_n^{-1}(p)$ to $F^{-1}(p)$ is established in Proposition 2, whereas Theorem 1 states the representation results.

2. Assumptions and Main Results

Our main assumption is the following.

H. $(X_i)_{i\geq 1}$ is a sequence of stationary α-mixing rv's with mixing coefficient

$$\alpha(\ell) = O(e^{-\log^{1+\nu}\ell})$$

for some $\nu > 0$.

Now in order to state our representation results for the quantile estimator we define

$$\xi(x,t,y) = \frac{\mathbf{1}_{\{x\leq y\}}}{C(x)} - \int_0^y \frac{\mathbf{1}_{\{t\leq u\leq x\}}}{C^2(u)}\, dF^*(u). \tag{12}$$

Then $\xi(X_1, T_1, \cdot)$ is a centered process with covariance function

$$\mathrm{Cov}\left(\xi(X_1,T_1,x), \xi(X_1,T_1,y)\right) = \int_0^{x\wedge y} \frac{dF^*(u)}{C^2(u)}.$$

Proposition 1 (Theorem 1 in Sun and Zhou (2001)). *Let* $0 \leq b < b_F$. *Under* **H** *we have*

$$\widehat{\Lambda}_n(x) - \Lambda(x) = \frac{1}{n}\sum_{i=1}^{n} \xi(X_i, T_i, x) + R_{n1}(x) \tag{13}$$

and

$$\widehat{F}_n(x) - F(x) = (1 - F(x))\frac{1}{n}\sum_{i=1}^{n} \xi(X_i, T_i, x) + R_{n2}(x), \tag{14}$$

where

$$\sup_{0\leq x\leq b} |R_{ni}(x)| = O(n^{-1/2}(\log n)^{-\zeta}) \quad a.s. \quad for \quad i = 1, 2$$

for some $\zeta > 0$.

Remark 1. Sun and Zhou (2001) consider the rate $\alpha'(\ell) = O(\ell^{-\omega})$ for some $\omega > 3$, which is a whisker milder than ours. Needless to say, their representation result is valid in our case. Moreover, their ζ depends on ω. Since our $\alpha(\ell)$ is $o(\ell^{-\omega})$ for any $\omega > 0$, in our case ζ is "universal" (even if the "exact" rate may depend on ν). Note finally that the previous result is stated in the case $a_G < a_F$, whereas if $a_G = a_F$, the remainder term is uniformly bounded only on intervals of the form $[a, b]$ for $a > 0$.

Corollary 1. *Under* **H**, *we have for* $0 \leq b < b_F$

$$\sup_{0\leq x\leq b} |\widehat{F}_n(x) - F(x)| = O(\eta_n) \quad a.s., \tag{15}$$

where $\eta_n = n^{-1/2}(\log\log n)^{1/2}$.

Proposition 2. *Under* **H**, *if* $F^{-1}(p)$ *is the unique solution of* $F(x) = p$, *then*

$$\widehat{F}_n^{-1}(p) \overset{n\to\infty}{\longrightarrow} F^{-1}(p) \quad a.s.$$

The following theorem is our main result.

Theorem 1. *Assume* F *is Lipschitz continuous. Let* $0 < p < 1$ *and* $(F^{-1})'(p) = f(F^{-1}(p)) > 0$. *Then, under* **H** *we have the representations*

$$\begin{aligned}\widehat{F}_n^{-1}(p) - F^{-1}(p) &= \frac{p - \widehat{F}_n(F^{-1}(p))}{f(F^{-1}(p))} + R_{n3}(p) \\ &= -\frac{1-p}{f(F^{-1}(p))} \cdot \frac{1}{n}\sum_{i=1}^{n} \xi(X_i,\, T_i,\, F^{-1}(p)) + R_{n4}(p),\end{aligned}$$

where

$$R_{ni}(p) = o(\eta_n) \quad a.s. \quad for \quad i = 3, 4. \tag{16}$$

In addition, if F is twice continuously differentiable at $F^{-1}(p)$, then

$$R_{ni}(p) = O\big(n^{-1/2}(\log n)^{-\zeta}\big) \quad a.s \quad for \quad i = 3, 4, \tag{17}$$

for some $\zeta > 0$.

Finally, if F is continuously (resp., twice continuously) differentiable on $[F^{-1}(p_0) - \rho,\, F^{-1}(p_1) + \rho]$ for some $\rho > 0$ and $0 < p_0 \leq p_1 < 1$ such that $f = F'$ is bounded away from zero there, the error bounds (16) (*resp.,* (17)) *hold uniformly in $p_0 \leq p \leq p_1$.*

Now we give a consequence of the strong representation of the quantile function. It is clear that the sequence $\{\xi(X_i, T_i, F^{-1}(p))\}$ (where ξ is defined in (12)) is formed of stationary α-mixing bounded random variables. Applying Theorem 1 and Theorem 18.5.4 in Ibragimov and Linnik (1971) we have

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

$$\sqrt{n}\left(\widehat{F}_n^{-1}(p) - F^{-1}(p)\right) \xrightarrow{\mathcal{D}} \mathcal{N}(0, \Sigma),$$

where

$$\Sigma^2 = \frac{(1-p)^2}{\{f[F^{-1}(p)]\}^2}\Big\{ \operatorname{Var}\{\xi(X_1, T_1, F^{-1}(p))\} + 2\sum_{j=2}^{\infty} \operatorname{cov}\{\xi(X_1, T_1, F^{-1}(p)), \xi(X_j, T_j, F^{-1}(p))\}\Big\},$$

with $\operatorname{Var}(\xi(X_1, T_1, F^{-1}(p)) = \int_0^{F^{-1}(p)} \frac{F^*(du)}{C^2(u)}$ *and* $\xrightarrow{\mathcal{D}}$ *denoting convergence in distribution.*

3. Proofs

In what follows, $K_1, K_2, \ldots$ denote universal positive constants.

Proof of Corollary 1. From Proposition 1 we have

$$\widehat{F}_n(x) - F(x) = (1 - F(x))(\widehat{\Lambda}_n(x) - \Lambda(x)) + O(n^{-1/2}(\log n)^{-\zeta})$$

on $0 \leq x \leq b < b_F$. Applying Theorem 3.2 in Cai and Roussas (1992) (which is a law of iterated logarithm) to (A11) in Sun and Zhou (2001) we get

$$\sup_{x \in [0,b]} \left|\widehat{\Lambda}_n(x) - \Lambda(x)\right| = O\left(\eta_n\right) \ a.s. \tag{18}$$

which then gives (15). □

Proof of Proposition 2. Let $\varepsilon > 0$. Since $F^{-1}(p)$ is the unique solution of $F(x) = p$, we have

$$F(F^{-1}(p) - \varepsilon) < p < F(F^{-1}(p) + \varepsilon).$$

By Corollary 1 we have

$$\widehat{F}_n(F^{-1}(p)-\varepsilon) \overset{n\to\infty}{\longrightarrow} F(F^{-1}(p)-\varepsilon) \quad \text{a.s.}$$

and

$$\widehat{F}_n(F^{-1}(p)+\varepsilon) \overset{n\to\infty}{\longrightarrow} F(F^{-1}(p)+\varepsilon) \quad \text{a.s.}$$

Then

$$\mathbb{P}\big\{\widehat{F}_m(F^{-1}(p)-\varepsilon) < p < \widehat{F}_m(F^{-1}(p)+\varepsilon),\, \forall m \geq n\big\} \overset{n\to\infty}{\longrightarrow} 1.$$

Now, for any df W, we have

$$W(x) \geq p \iff x \geq W^{-1}(p).$$

Therefore

$$\mathbb{P}\big(F^{-1}(p)-\varepsilon < \widehat{F}_m^{-1}(p) \leq F^{-1}(p)+\varepsilon,\, \forall m \geq n\big) \overset{n\to\infty}{\longrightarrow} 1,$$

that is

$$\mathbb{P}\big(\sup_{m\geq n} |\widehat{F}_m^{-1}(p) - F^{-1}(p)| > \varepsilon\big) \overset{n\to\infty}{\longrightarrow} 0,$$

which permits us to conclude. □

To prove Theorem 1, we will need some lemmas. The first lemma deals with ties in the α-mixing case.

Lemma 1. *Under* **H** *we have*

$$\sup_{x\in[0,b]} \sum_{i=1}^{n} \mathbf{1}_{\{X_i=x\}} = o(n^{\kappa}) \quad \textit{a.s.}$$

for any $\kappa > 0$.

Proof. Let $\kappa > 0$ and consider $E_\kappa = \{\sup_{x\in[0,b]} \sum_{i=1}^{n} \mathbf{1}_{\{X_i=x\}} \geq n^{\kappa}\}$. Put $[0,b] = \bigcup_{l=0}^{u_n-1} A_l$, where $A_l = \left[\frac{lb}{u_n};\ \frac{(l+1)b}{u_n}\right]$ and the number u_n of elements A_l will be precised below. We have

$$E_\kappa \subset \bigcup_{j=1}^{n} \bigcup_{k\geq j+n^{\kappa}} \{X_j = X_k\},$$

which implies

$$\mathbb{P}(E_\kappa) \leq \sum_{j=1}^{n} \sum_{k\geq j+n^{\kappa}} \mathbb{P}(X_j = X_k). \tag{19}$$

Furthermore, we have

$$\{X_j = X_k\} \subset \bigcup_{l=0}^{u_n-1} \big\{X_j \in A_l \quad \text{and} \quad X_k \in A_l\big\}.$$

Then, for $k \geq j + n^{\kappa}$,

$$\mathbb{P}(X_j = X_k) \leq \sum_{l=0}^{u_n - 1} \mathbb{P}\big(\{X_j \in A_l\} \cap \{X_k \in A_l\}\big)$$
$$\leq \sum_{l=0}^{u_n - 1} \left[\{\mathbb{P}(X_1 \in A_l)\}^2 + \alpha(n^{\kappa})\right] \leq \left[\frac{b \sup_{[0,b]} f^*}{u_n}\right]^2 u_n + \alpha(n^{\kappa}) u_n.$$

The last term achieves its minimum at $u_n = b \sup_{[0,b]} f^* / \sqrt{\alpha(n^{\kappa})}$ (we take the integer part), which gives

$$\mathbb{P}(X_j = X_k) \leq 2b \sup_{[0,b]} f^* \cdot \sqrt{\alpha(n^{\kappa})}.$$

Now (19) yields

$$\mathbb{P}(E_{\kappa}) \leq 2b \sup_{[0,b]} f^* \cdot n^2 \sqrt{\alpha(n^{\kappa})} = O\big(n^2 \exp\{-\tfrac{1}{2}\kappa^{1+\nu} \log^{1+\nu} n\}\big).$$

The latter being the general term of a convergent series, Borel–Cantelli's lemma permits us to conclude. □

The next lemma deals with the behavior of the difference between $\widehat{F}_n\left(\widehat{F}_n^{-1}(p)\right)$ and p.

Lemma 2. *Let $0 < p_0 \leq p_1 < 1$. Under* **H** *we have*

$$\sup_{p_0 \leq p \leq p_1} |\widehat{F}_n(\widehat{F}_n^{-1}(p)) - p| = O(n^{-1/2} (\log n)^{-\zeta}) \quad a.s. \tag{20}$$

Proof. Let $x = \widehat{F}_n^{-1}(p)$. Using (14) we get

$$\sup_{p_0 \leq p \leq p_1} |\widehat{F}_n(\widehat{F}_n^{-1}(p)) - p| \leq \sup_{\widehat{F}_n^{-1}(p_0) \leq x \leq \widehat{F}_n^{-1}(p_1)} |\widehat{F}_n(x) - \widehat{F}_n(x^-)|$$
$$\leq \sup_{\widehat{F}_n^{-1}(p_0) \leq x \leq \widehat{F}_n^{-1}(p_1)} \left| \frac{1}{n} \sum_{i=1}^{n} \xi\left(X_i, T_i, x\right) - \frac{1}{n} \sum_{i=1}^{n} \xi(X_i, T_i, x^-) \right|$$
$$+ \sup_{\widehat{F}_n^{-1}(p_0) \leq x \leq \widehat{F}_n^{-1}(p_1)} (R'_{n2}(x)),$$

where $R'_{n2}(x) = |R_{n2}(x) - R_{n2}(x^-)|$, so that

$$\sup_{\widehat{F}_n^{-1}(p_0) \leq x \leq \widehat{F}_n^{-1}(p_1)} (R'_{n2}(x)) = O(n^{-1/2} (\log n)^{-\zeta}) \quad \text{a.s.} \tag{21}$$

On the other hand, since F^* is absolutely continuous, we get from (12)

$$\left| \frac{1}{n} \sum_{i=1}^{n} \xi(X_i, T_i, x) - \frac{1}{n} \sum_{i=1}^{n} \xi(X_i, T_i, x^-) \right|$$
$$= \left| \frac{1}{n} \sum_{i=1}^{n} \left(\frac{\mathbf{1}_{\{X_i \leq x\}} - \mathbf{1}_{\{X_i < x\}}}{C(X_i)} \right) \right| = \frac{1}{nC(x)} \sum_{i=1}^{n} \mathbf{1}_{\{X_i = x\}}. \tag{22}$$

Now from Proposition 2, for n large enough we have $\inf_{\widehat{F}_n^{-1}(p_0)\le x\le \widehat{F}_n^{-1}(p_1)} C(x) > 0$ a.s. This with Lemma 1 (with $\kappa = 1/4$) and (22) give

$$\sup_{\widehat{F}_n^{-1}(p_0)\le x\le \widehat{F}_n^{-1}(p_1)} \left| \frac{1}{n}\sum_{i=1}^{n} \xi(X_i, T_i, x) - \frac{1}{n}\sum_{i=1}^{n} \xi(X_i, T_i, x^-) \right| = o(n^{-3/4}) \quad \text{a.s.}$$

which, combined with (21) gives (20). □

Lemma 3. *Let $0 < p_0 \le p_1 < 1$. Under* **H**, *assuming that $F' = f$ is bounded away from zero on $[F^{-1}(p_0) - \rho, F^{-1}(p_1) + \rho]$ for some $\rho > 0$, we have*

$$\sup_{p_0\le p\le p_1} |\widehat{F}_n^{-1}(p) - F^{-1}(p)| = O(\eta_n) \quad \textit{a.s.} \tag{23}$$

Proof. From (15) and Proposition 2, using the Taylor expansion in a neighborhood of $F^{-1}(p)$, we have, for n large enough

$$\begin{aligned} \widehat{F}_n(\widehat{F}_n^{-1}(p)) &= F(\widehat{F}_n^{-1}(p)) + R_{n5}(p) \\ &= F(F^{-1}(p)) + [\widehat{F}_n^{-1}(p) - F^{-1}(p)] f(\vartheta_n(p)) + R_{n5}(p) \end{aligned}$$

for some $\vartheta_n(p)$ between $F^{-1}(p)$ and $\widehat{F}_n^{-1}(p)$, where $R_{n5}(p) = O(\eta_n)$ a.s. uniformly in $p_0 \le p \le p_1$.

Hence

$$\widehat{F}_n^{-1}(p) - F^{-1}(p) = \frac{\widehat{F}_n(\widehat{F}_n^{-1}(p)) - p}{f(\vartheta_n(p))} - \frac{R_{n5}(p)}{f(\vartheta_n(p))}.$$

Now, by Proposition 2 we have for n large enough $\vartheta_n(p) \in [F^{-1}(p_0)-\rho, F^{-1}(p_1)+\rho]$ a.s. Since f is bounded away from zero over this interval we get by Lemma 2

$$\begin{aligned} \sup_{p_0\le p\le p_1} |\widehat{F}_n^{-1}(p) - F^{-1}(p)| &= O\big(n^{-1/2}(\log n)^{-\zeta}\big) + O(\eta_n) \quad \text{a.s.} \\ &= O\big(\eta_n\big) \quad \text{a.s.} \qquad \square \end{aligned}$$

We now introduce the following lemma, which is adapted from a result by Bosq (1998).

Lemma 4. *Let $(Z_i)_{i\in\mathbb{N}^*}$ be a sequence of stationary strongly mixing centered rv's such that $\sup_{i\ge1} |Z_i| \le d$. Then, for each integer $q \in [1; n/2]$, for each $\varepsilon > 0$*

$$\mathbb{P}\left(\frac{1}{n}\left|\sum_{i=1}^{n} Z_i\right| > \varepsilon\right) \le 4\exp\left(-\frac{\varepsilon^2 q}{8}\left[\frac{8q^2\sigma^2(q)}{n^2} + \frac{\varepsilon d}{2}\right]^{-1}\right) + 22\left(1+\frac{4d}{\varepsilon}\right)^{1/2} q\alpha\left(\left[\frac{n}{2q}\right]\right),$$

where $\sigma^2(q) = K_1 \frac{n}{q}\mathbb{E}(Z_1^2)$, $\alpha(\cdot)$ is the mixing coefficient, and $[\cdot]$ is the integer part.

Proof. The same result is stated in Theorem 1.3 (2) in Bosq (1998) with

$$\begin{aligned} \sigma_0^2(q) = \max_{0\le j\le 2q-1} \mathbb{E}\Bigg\{ &\left(\left[\frac{jn}{2q}\right] + 1 - \frac{jn}{2q}\right) Z_{[\frac{jn}{2q}+1]} + Z_{[\frac{jn}{2q}+2]} \\ &+ \ldots + Z_{[\frac{(j+1)n}{2q}]} + \left(\left[(j+1)\frac{n}{2q}\right] - (j+1)\frac{n}{2q}\right) Z_{[(j+1)\frac{n}{2q}]+1} \Bigg\}^2 \end{aligned}$$

instead of $\sigma^2(q)$. Since (Z_i) is stationary, we can take $j = 0$. Then, we have

$$\sigma_0^2(q) \leq \left[\frac{n}{2q}\right] \mathbb{E}[Z_1^2] + 2 \sum_{j=2}^{[n/2q]} \left(\left[\frac{n}{2q}\right] + 1 - j \right) |\mathbb{E}(Z_1 Z_j)|.$$

Now, Theorem 17.2.3 in Ibragimov and Linnik (1981) gives

$$|\mathbb{E}(Z_1 Z_j)| \leq 2\alpha(j)^{1/2} \mathbb{E}(Z_1^2).$$

Then

$$\sigma_0^2(q) \leq \left[\frac{n}{2q}\right] \mathbb{E}(Z_1^2) \left\{ 1 + 2 \sum_{j=2}^{[n/2q]} \alpha(j)^{1/2} \right\} \leq \frac{n}{2q}\, \mathbb{E}(Z_1^2)\{1 + 2\mathcal{S}\} := \sigma^2(q),$$

where $\mathcal{S}$ is the sum of the series, which gives the result. □

Lemma 5. *Let $K_2 > 0$ and $0 \leq b < b_F$. Under* **H**, *we have*

$$\sup_{|t-s| \leq K_2 \eta_n,\, 0 \leq s,t \leq b} \left| \widehat{F}_n(t) - F(t) - \widehat{F}_n(s) + F(s) \right| = O(n^{-1/2} (\log n)^{-\zeta}) \quad a.s. \tag{24}$$

for some $\zeta > 0$.

Proof. We first show the statement for $\widehat{\Lambda}_n - \Lambda$ before considering $\widehat{F}_n - F$. Let $s, t \in [0, b]$ such that $|t - s| \leq K_2 \eta_n$. In view of (6) we have from (12) and (13)

$$\begin{aligned} \widehat{\Lambda}_n(t) - \Lambda(t) - \widehat{\Lambda}_n(s) + \Lambda(s) &= \int_{]s,t]} \frac{d[F_n^*(u) - F^*(u)]}{C(u)} \\ &\quad - \int_{]s,t]} \frac{C_n(u) - C(u)}{C^2(u)}\, dF^*(u) + O(n^{-1/2}(\log n)^{-\zeta}) \\ &:= I_1 + I_2 + O(n^{-1/2}(\log n)^{-\zeta}) \end{aligned} \tag{25}$$

with a uniform remainder term for $|t - s| \leq K_2 \eta_n$. We now show that uniformly in $|t - s| \leq K_2 \eta_n$ both integrals are of the stated order. First, we have

$$|I_2| \leq \sup_{0 \leq x \leq b} \left| \frac{C_n(x) - C(x)}{C^2(x)} \right| \int_{]s,t]} dF^*(u) \leq \sup_{0 \leq x \leq b} \left| \frac{C_n(x) - C(x)}{C^2(x)} \right| \cdot K_3 \eta_n.$$

Applying Theorem 3.2 in Cai and Roussas (1992) to (7) gives

$$\sup_{0 \leq x \leq b} \left| \frac{C_n(x) - C(x)}{C^2(x)} \right| = O(\eta_n) \quad \text{a.s.}$$

and then

$$I_2 = O(n^{-1} \log\log n) \quad \text{a.s.} \tag{26}$$

Now in order to estimate I_1 put

$$\delta_n = K_0 n^{-\beta} \tag{27}$$

with $5/8 < \beta < 3/4$ and consider the grid

$$x_{n,j} = j\delta_n, \quad j = 0, 1, \dots, k_n = \left[\frac{b}{\delta_n}\right], \qquad x_{n,k_n+1} = b.$$

Then, for any $y \in [0, b)$ there exists j such that $x_{n,j} \le y < x_{n,j+1}$. Using easy notation, we put $\underline{y} := x_{n,j}$ and $\overline{y} := x_{n,j+1}$.

Since C is nonnegative and F is Lipschitz continuous, we can write

$$I_1 \le \int_{]\underline{s},\,\overline{t}\,]} \frac{d[F_n^*(u) - F^*(u)]}{C(u)} + O(\delta_n) := \mu_n + O(\delta_n). \tag{28}$$

Now each μ_n can be written as a mean of α-mixing rv's, that is, $\mu_n = n^{-1}\sum_{i=1}^n Y_i$, where

$$Y_i = \int_{]\underline{s},\,\overline{t}\,]} \frac{d[\mathbf{1}_{\{X_i \le u\}} - F^*(u)]}{C(u)} = \frac{\mathbf{1}_{\{\underline{s} \le X_i \le \overline{t}\}}}{C(X_i)} - \int_{]\underline{s},\,\overline{t}\,]} \frac{dF^*(u)}{C(u)}.$$

We have

$$\begin{aligned} |Y_i| &\le \int_{]\underline{s},\,\overline{t}\,]} \left| \frac{d[\mathbf{1}_{\{X_i \le u\}} - F^*(u)]}{C(u)} \right| \\ &\le \frac{1}{\inf_{x\in[0,b]} C(x)} \big[1 + (K_2\eta_n + 2\delta_n) \sup_{[0,b]} f^*\big] \le \frac{K_4}{\inf_{x\in[0,b]} C(x)} =: d. \end{aligned} \tag{29}$$

It is easy to see that $\mathbb{E}[Y_i] = 0$. Moreover,

$$\begin{aligned} \mathrm{Var}(Y_i) &= \mathrm{Var}\left(\int_{]\underline{s},\,\overline{t}\,]} \frac{d[\mathbf{1}_{\{X_i \le u\}} - F^*(u)]}{C(u)} \right) \le \int_{]\underline{s},\,\overline{t}\,]} \frac{\mathrm{Var}(d\mathbf{1}_{\{X_i \le u\}})}{C^2(u)} \\ &\le \int_{]\underline{s},\,\overline{t}\,]} \frac{dF^*(u)}{C^2(u)} \le \frac{1}{\inf_{0\le x\le b}(C^2(x))} \int_{]\underline{s},\,\overline{t}\,]} dF^*(u) \le K_5\eta_n \end{aligned}$$

by the Lipschitz property.

Now we apply Lemma 4 to $Z_i = Y_i$ with $\varepsilon = \delta_n$, d given in (29), and $q = q_n := [n^{1/4+\beta}]$, where β is the same as in (27):

$$\begin{aligned} \mathbb{P}\left(\frac{1}{n} \left| \sum_{i=1}^n Y_i \right| > \delta_n \right) &\le 4\exp\left\{ -\frac{\delta_n^2 q_n}{8} \left[\frac{8q_n^2\sigma^2(q_n)}{n^2} + \frac{d\delta_n}{2} \right]^{-1} \right\} \\ &\quad + 22q_n \left(1 + \frac{4d}{\delta_n}\right)^{1/2} \alpha\left(\left[\frac{n}{2q_n} \right] \right) := II_1 + II_2. \end{aligned} \tag{30}$$

Now $q_n^2\sigma^2(q_n)/n^2$ is of order $n^{\beta-5/4}(\log\log n)^{1/2}$ and is therefore asymptotically larger than $d\delta_n/2$. Then, for n large enough

$$II_1 \le 4\exp\left\{ -K_6 \frac{n^{3/2-2\beta}}{\sqrt{\log\log n}} \right\} = o(n^{-5/2}). \tag{31}$$

On the other hand,

$$II_2 \leq K_7 n^{1/4+3\beta/2} \exp\left\{ -(3/4-\beta)^{1+\nu}(\log n)^{1+\nu} \right\} = o(n^{-5/2}). \tag{32}$$

Now we come back to (28). Since the number of such μ_n is

$$\binom{k_n+2}{2} = O(k_n^2) = O(\delta_n^{-2}) = O(n^{2\beta}),$$

(30)–(32) in conjunction with Borel–Cantelli's Lemma (since $2\beta - 5/2 < -1$) imply

$$\max|\mu_n| \leq \delta_n \quad \text{a.s.},$$

which added to (25), (26), and (28) allows us to conclude for the upper bound of the oscillation modulus of $\widehat{\Lambda}_n - \Lambda$. The lower bound can be derived similarly by considering in (28) $\overline{s}$ and $\underline{t}$ instead of $\underline{s}$ and $\overline{t}$, respectively.

Now we come back to $\widehat{F}_n - F$. From Sun and Zhou (2001), we have for $0 \leq s, t \leq b$ and $|t-s| \leq K_2\eta_n$

$$\begin{aligned}
&\widehat{F}_n(t) - F(t) - \widehat{F}_n(s) + F(s) \\
&\quad = (1-F(t))(\widehat{\Lambda}_n(t) - \Lambda(t)) - (1-F(s))(\widehat{\Lambda}_n(s) - \Lambda(s)) + O(\eta_n^2) \\
&\quad = \widehat{\Lambda}_n(t) - \Lambda(t) - \widehat{\Lambda}_n(s) + \Lambda(s) - F(t)(\widehat{\Lambda}_n(t) - \Lambda(t)) \\
&\qquad + F(s)(\widehat{\Lambda}_n(s) - \Lambda(s)) + O(\eta_n^2).
\end{aligned}$$

The result established for the oscillation modulus of $\widehat{\Lambda}_n - \Lambda$ (when $|t-s| \leq K_2\eta_n$) is then easily extended to $\widehat{F}_n - F$, since F is Lipchitz continuous and using (18). □

Proof of Theorem 2. From Lemmas 3 and 5 we have, for n large enough

$$\begin{aligned}
&\widehat{F}_n(\widehat{F}_n^{-1}(p)) - \widehat{F}_n(F^{-1}(p)) \\
&\quad = \widehat{F}_n(\widehat{F}_n^{-1}(p)) - F(\widehat{F}_n^{-1}(p)) - \widehat{F}_n(F^{-1}(p)) + F(F^{-1}(p)) \\
&\qquad + F(\widehat{F}_n^{-1}(p)) - F(F^{-1}(p)) \\
&\quad = F(\widehat{F}_n^{-1}(p)) - F(F^{-1}(p)) + O\left(\frac{1}{\sqrt{n}\,(\log n)^{\zeta}}\right) \quad \text{a.s.}
\end{aligned}$$

Then the Taylor expansion in a neighborhood of $F^{-1}(p)$ gives

$$\begin{aligned}
\widehat{F}_n(\widehat{F}_n^{-1}(p)) - \widehat{F}_n(F^{-1}(p)) &= f(F^{-1}(p))(\widehat{F}_n^{-1}(p) - F^{-1}(p)) \\
&\quad + o(\widehat{F}_n^{-1}(p) - F^{-1}(p)) + O\left(\frac{1}{\sqrt{n}(\log n)^{\zeta}}\right) \quad \text{a.s.}
\end{aligned} \tag{33}$$

Now the first representation of Theorem 1 gives

$$\begin{aligned}
\widehat{F}_n(\widehat{F}_n^{-1}(p)) - \widehat{F}_n(F^{-1}(p)) &= p - \widehat{F}_n(F^{-1}(p)) + f(F^{-1}(p))R_{n3}(p) \\
&\quad + o(\widehat{F}_n^{-1}(p) - F^{-1}(p)) + O\left(\frac{1}{\sqrt{n}(\log n)^{\zeta}}\right) \quad \text{a.s.},
\end{aligned}$$

that is

$$R_{n3}(p) = \frac{\widehat{F}_n(\widehat{F}_n^{-1}(p)) - p}{f(F^{-1}(p))} + o\big(\widehat{F}_n^{-1}(p) - F^{-1}(p)\big) + O\bigg(\frac{1}{\sqrt{n}(\log n)^{\zeta}}\bigg) \quad \text{a.s.} \tag{34}$$

Since $f(F^{-1}(p)) > 0$ we get, using Lemmas 2 and 3,

$$R_{n3}(p) = o(\eta_n) \quad \text{a.s.}$$

If F is twice differentiable at $F^{-1}(p)$, the $o(\widehat{F}_n^{-1}(p) - F^{-1}(p))$ term in (33) is in fact an $O\big((\widehat{F}_n^{-1}(p) - F^{-1}(p))^2\big)$. The remainder term in (34) is then $O(n^{-1/2}(\log n)^{-\zeta})$.

Finally, since the bounds leading to (33) are already uniform in $p_0 \le p \le p_1$, we only have to note that the error bounds in the Taylor expansion also hold uniformly under the stated regularity assumptions on F which completes the proof for R_{n3}.

Now using (14) with $x = F^{-1}(p)$, we get the same representation for $R_{n4}(p)$ as in (34). Therefore the results proved for R_{n3} extend to R_{n4}. □

Acknowledgment. The authors are grateful to an anonymous referee whose comments gave the way to an impoved version of the paper.

References

[1] R. R. Bahadur (1966), *A note on quantiles in large samples*, Ann. Math. Statist., 37 , 577–580.

[2] D. Bosq (1998), *Nonparametric Statistics for Stochastic Processes. Estimation and Prediction*, Lect. Notes in Statist., Vol. 110, Springer.

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

[4] Z. Cai and G. G. Roussas (1992), *Uniform strong estimation under α-mixing, with rates*, Statist. Probab. Lett., 15, 47–55.

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

[6] I. Gijbels and J. L. Wang (1993), *Strong representations of survival function estimator for truncated and censored data with applications*, J. Multivariate Anal., 47, 210–229.

[7] M. G. Gu and T. L. Lai (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.

[8] U. Gürler, W. Stute, and J. L. Wang (1993), *Weak and strong quantile representations for randomly truncated data with applications*, Statist. Probab.Lett., 17, 139–148.

[9] I. A. Ibragimov and Y. V. Linnik (1981), *Independent and Stationary Processes for Random Variables*, Wiley, New York.

[10] N. Keiding and R. D. Gill (1990), *Random truncation models and Markov processes*, Ann. Statist., 18, 582–602.

[11] T. L. Lai and L. Ying (1991), *Estimating a distribution function with truncated and censored data*, Ann. Statist., 19, 417–442.

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

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

[14] G. Shorack and J. A. Wellner (1986), *Empirical Processes with Application to Statistics*, Wiley, New York.

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

[16] L. Sun and X. Zhou (2001), *Survival function and density estimation for truncated dependent data*, Statist. Probab. Lett., 52, 47–57.
[17] S. M. Tse (2000), *Strong Gaussian approximations in the random truncation model*, Statist. Sinica, 10, 281–296.
[18] M. C. Wang, N. P. Jewell, and M. Y. Tsai (1986), *Asymptotic properties of the product-limit estimate under random truncation*, Ann. Statist., 14, 1599–1605.
[19] M. Woodroofe (1985), *Estimating a distribution function with truncated data*, Ann. Statist., 13, 163–177.
[20] Y. Zhou (1996), *A note on the TJW product-limit estimator for truncated and censored data*, Statist. Probab. Lett., 26, 381–387.
[21] Y. Zhou and P. S. F. Yip (1999), *Strong representation of the product-limit estimator for the left truncated and right censored data*, J. Multivariate Anal., 69, 261–280.

[*Received August* 2004; *revised July* 2005]