Weak and strong quantile representations for randomly truncated data with applications

Ülkü Gürler *

Faculty of Engineering and Sciences, Bilkent University, Ankara, Turkey

Winfried Stute **

Mathematics Institute, Justus-Liebig-University, Giessen, Germany

Jane-Ling Wang ***

Division of Statistics, University of California, Davis, CA, USA

Abstract: Suppose that we observe bivariate data (X_i, Y_i) only when $Y_i \leqslant X_i$ (left truncation). Denote with F the marginal d.f. of the X's. In this paper we derive a Bahadur-type representation for the quantile function of the pertaining product-limit estimator of F. As an application we obtain confidence intervals and bands for quantiles of F.

AMS 1991 Subject Classifications: Primary 62G05; Secondary 62G30, 62G15.

Keywords: Truncated data; Bahadur representation; confidence interval; product-limit estimator; survival function.

1. Introduction

Let (X_i, Y_i), $1 \leqslant i \leqslant N$, be a random sample (i.i.d.) from some population such that X_i is independent of Y_i. Denote with

$$F(t) = \mathbb{P}(X \leqslant t) \quad \text{and} \quad G(t) = \mathbb{P}(Y \leqslant t)$$

the marginal distribution functions (d.f.'s) of X and Y, respectively. In the random left truncation model one observes only those pairs (X_i, Y_i) for which $X_i \geqslant Y_i$ but the label i is not observed. This model arises in various fields, e.g., astronomy, economics and medical studies. See, e.g., Woodroofe (1985). Let (x_i, y_i), $1 \leqslant i \leqslant n$, denote the observed values of the sample. Note that n is a random variable itself. The problem now becomes one of reconstructing F and G from (x_i, y_i), $1 \leqslant i \leqslant n$. In most applications the

main interest is in the X-variable. Since in many examples X turns out to be nonnegative we shall restrict ourselves to this case though this assumption in no way limits the method.

Now, given n, we may look at the data as the outcomes of an i.i.d. sample with d.f.

$$H^*(x, y) = \mathbb{P}(X \leqslant x, Y \leqslant y \mid Y \leqslant X).$$

where we assume that

$$\alpha \equiv \mathbb{P}(Y \leqslant X) > 0.$$

Denote with

$$F^*(x) = H^*(x, \infty) \quad \text{and} \quad G^*(y) = H^*(\infty, y)$$

the marginals of H^*. The actually observed X's and Y's thus have d.f. F^* and G^*, respectively. The nonparametric maximum likelihood estimator (NPMLE) of F was derived by Lynden-Bell (1971) and is of the form

$$1 - \hat{F}_n(x) = \prod_{\{i: X_i \leqslant x\}}' \left[1 - \frac{r_n(X_i)}{nC_n(X_i)} \right], \quad x \geqslant 0, \tag{1.1}$$

where

$$r_n(x) = \#\{j \leqslant n : X_j = x\}, \qquad C_n(x) = n^{-1}\#\{j \leqslant n : Y_j \leqslant x \leqslant X_j\},$$

and $\prod'$ extends over all pairwise distinct X's. (1.1) may be motivated as follows: Denote with

$$\Lambda(x) = \int_{[0, x]} \frac{F(\mathrm{d}z)}{1 - F(z-)}, \quad 0 \leqslant x < \infty,$$

the so-called cumulative hazard function of F. Then we have the representation

$$1 - F(x) = \left\{ \prod_{z \in A, z \leqslant x} [1 - \lambda(z)] \right\} \exp[-\Lambda_c(x)],$$

in which

$$\lambda(x) = \Lambda(x) - \Lambda(x-), \qquad \Lambda_c(x) = \Lambda(x) - \sum_{z \in A, z \leqslant x} \lambda(z)$$

and $A = \{z : \lambda(z) > 0\}$ is the set of atoms of Λ. Putting

$$a_F = \inf\{x : F(x) > 0\} \quad \text{and} \quad b_F = \sup\{x : F(x) < 1\},$$

and similarly for G, Woodroofe (1985) observed that, when $a_G \leqslant a_F$ and $b_G \leqslant b_F$,

$$\Lambda(x) = \int_{[0, x]} \frac{F^*(\mathrm{d}z)}{C(z)},$$

with

$$C(z) = G^*(z) - F^*(z-) = \alpha^{-1} G(z)[1 - F(z-)].$$

The function C may be consistently estimated by C_n, uniformly on $[0, \infty)$. Write F_n^* for the empirical d.f. of $X_1, \ldots, X_n$ and set

$$\Lambda_n(x) = \int_{[0, x]} \frac{F_n^*(\mathrm{d}z)}{C_n(z)}.$$

Since Λ_n is a pure step function (so that $\Lambda_{nc} \equiv 0$), we may define $\hat{F}_n$, in obvious notation, by

$$1 - \hat{F}_n(x) = \prod_{z \leqslant x} [1 - \lambda_n(z)],$$

which is identical to (1.1).

The distributional convergence of this estimator has been studied by Woodroofe (1985), Wang, Jewell and Tsai (1986), Gu and Lai (1990), Keiding and Gill (1990) and Lai and Ying (1991). Chao and Lo (1986) and Stute (1993) obtained almost sure representations with rate. In this article we explore properties of its quantile function $\hat{F}_n^{-1}$, where for any d.f. L,

$$L^{-1}(p) = \inf\{x \in \mathbb{R}: L(x) \geqslant p\}, \quad 0 < p < 1.$$

Elementary properties of a quantile function are listed on p. 5 of Shorack and Wellner (1986). We show in Theorem 1 that almost surely (a.s.)

$$\hat{F}_n^{-1}(p) \to F^{-1}(p),$$

if $F^{-1}(p)$ is the unique solution of $F(x) = p$, i.e. if F is strictly increasing at $F^{-1}(p)$. Finer results on the difference $\hat{F}_n^{-1}(p) - F^{-1}(p)$ are studied in Theorem 2, via asymptotic representations of the form

$$\hat{F}_n^{-1}(p) - F^{-1}(p) = \frac{p - \hat{F}_n(F^{-1}(p))}{f(F^{-1}(p))} + R_n(p), \tag{1.2}$$

where $f = F'$ and $R_n(p) = O((\ln n/n)^{3/4})$ a.s. or $o(n^{-1/2})$ in probability. The a.s. part of (1.2) is the analogue of the Bahadur (1966) representation for quantiles of the empirical d.f. for i.i.d. data. The in-probability part for i.i.d. data is due to Ghosh (1971). Note that for randomly censored data the Bahadur representation has been derived by Cheng (1984), Aly, Csörgő and Horváth (1985) and Lo and Singh (1986). Ghosh's representation has been extended to the censored case by Gastwirth and Wang (1988) and Gijbels and Veraverbeke (1988). For truncated data no results are available in the literature so far.

The representation (1.2) also holds uniformly for p in any interval $[p_0, p_1]$ contained in $(0, 1)$. Weak convergence of the quantile process $n^{1/2}[\hat{F}_n^{-1}(p) - F^{-1}(p)]$, $p_0 \leqslant p \leqslant p_1$, and confidence intervals respectively bands for quantiles are illustrated in Section 4 as applications of the weak and strong representation.

2. Preliminaries and assumptions

In this section we present some preliminary results which are needed in the next section.

Recall the definition of a_F, b_F, a_G and b_G. Woodroofe (1985) pointed out that F can be estimated on $[a_F, b_F]$ only if $a_G \leqslant a_F$ and $b_G \leqslant b_F$. Assuming that this holds Stute (1993) implies that for continuous F, uniformly in $a_G < a \leqslant x \leqslant b < b_F$, Λ_n admits the representation

$$\Lambda_n(x) - \Lambda_n(a) - \Lambda(x) + \Lambda(a)$$

$$= \int_{(a, x]} C^{-1}(z) [F_n^*(\mathrm{d}z) - F^*(\mathrm{d}z)] - \int_{(a, x]} \frac{C_n(z) - C(z)}{C^2(z)} F^*(\mathrm{d}z) + R_n'(a, x)$$

$$\equiv S_n(a, x) + R_n'(a, x), \tag{2.1}$$

where for any $\delta > \frac{3}{2}$,

$$\sup_{a \leqslant x \leqslant b} |R_n'(a, x)| = o\left(n^{-1}(\ln n)^\delta\right) \quad \text{w.p. 1.}$$

Furthermore, if $a_G < a_F$, one can choose $a_G < a < a_F$ in which case

$$\Lambda_n(x) - \Lambda(x) = S_n(0, x) + R'_n(0, x). \tag{2.2}$$

Finally, if $a_F = a_G$, then (2.2) still remains true with the same bound on R'_n provided that $\int dF/G^2 < \infty$. Informally speaking, the last integrability condition is needed to control the effect of truncation in a neighborhood of the critical point $a_F = a_G$. Observe that $S_n(a, x)$ is an average of i.i.d. random variables with expectation zero to which both the SLLN and the CLT apply. As to $\hat{F}_n$, one has almost surely

$$\hat{F}_n(x) - F(x) = (1 - F(x))S_n(0, x) + R_n(0, x),$$

with

$$\sup_{a_F \leqslant x \leqslant b} |R_n(0, x)| = O(\ln^3 n/n) \quad \text{w.p. 1}. \tag{2.3}$$

The following theorem is similar to Theorem 2.3.1. of Serfling (1980, p. 75). Therefore the proof is omitted.

Theorem 1. *Assume that F is continuous; also suppose that $a_G \leqslant a_F$, $b_G \leqslant b_F$. If $\int dF/G^2 < \infty$ and $F^{-1}(p)$ is the unique solution of $F(x) = p$, then*

$$\hat{F}_n^{-1}(p) \to F^{-1}(p) \quad a.s. \qquad \square$$

3. Weak and strong representations of p-quantiles

In this section we shall derive the representation (1.2). A series of lemmas which are of independent interest will now be derived. Lemma 3.1 shows that $\hat{F}_n$ composed with $\hat{F}_n^{-1}$ yields the identity on $[p_0, p_1]$, up to an error $O(n^{-1})$. Lemmas 3.2 and 3.3 provide global and local bounds for the deviation between $\hat{F}_n^{-1}$ and F^{-1}. For the classical empirical process, a similar analysis may be found, e.g., in Stute (1982, p. 99). By iteration he also derived higher-order representations of (uniform) quantiles. The notation of Sections 1 and 2 as well as the assumptions $a_G \leqslant a_F$, $b_G \leqslant b_F$ and $\int dF/G^2 < \infty$ will be adopted throughout.

Lemma 3.1. *For continuous F, for each $0 < p_0 \leqslant p_1 < 1$,*

$$\sup_{p_0 \leqslant p \leqslant p_1} |\hat{F}_n \hat{F}_n^{-1}(p) - p| = O(n^{-1}) \quad a.s.$$

Proof. First observe that $\hat{F}_n^{-1}(p) = X_i$ for some $1 \leqslant i \leqslant n$. Thus

$$|\hat{F}_n \hat{F}_n^{-1}(p) - p| = \hat{F}_n \hat{F}_n^{-1}(p) - p \leqslant \hat{F}_n(X_i) - \hat{F}_n(X_i-).$$

Hence it follows from (1.1) that a.s.,

$$\sup_{p_0 \leqslant p \leqslant p_1} |\hat{F}_n \hat{F}_n^{-1}(p) - p|$$

$$\leqslant \sup_{\hat{F}_n^{-1}(p_0) \leqslant x \leqslant \hat{F}_n^{-1}(p_1)} [\hat{F}_n(x) - \hat{F}_n(x-)] = \sup_{\hat{F}_n^{-1}(p_0) \leqslant x \leqslant \hat{F}_n^{-1}(p_1)} [1 - \hat{F}_n(x-)][nC_n(x)]^{-1}$$

$$\leqslant \sup_{\hat{F}_n^{-1}(p_0) \leqslant x \leqslant \hat{F}_n^{-1}(p_1)} [nC_n(x)]^{-1} \leqslant \sup_{F^{-1}(p_0)-\varepsilon \leqslant x \leqslant F^{-1}(p_1)+\varepsilon} [nC_n(x)]^{-1},$$

where the last inequality holds for all small enough $\varepsilon > 0$ and all large n, according to Theorem 1. The lemma now follows from the uniform convergence of C_n to C and

$$\inf_{F^{-1}(p_0)-\varepsilon \leqslant x \leqslant F^{-1}(p_1)+\varepsilon} C(x) > 0. \qquad \square$$

Lemma 3.2. *Suppose that $F' = f$ is continuous and bounded away from zero on $[F^{-1}(p_0) - \delta, F^{-1}(p_1) + \delta]$, for some $\delta > 0$. Then*

$$\sup_{p_0 \leqslant p \leqslant p_1} \left| \hat{F}_n^{-1}(p) - F^{-1}(p) \right| = O\big((\ln n/n)^{1/2}\big) \quad a.s.$$

$$= O_{\mathbb{P}}(n^{-1/2}).$$

Proof. We shall only deal with the a.s. part. From (2.3) and the LIL for empirical d.f.'s we have

$$\sup_{a \leqslant x \leqslant b} \left| \hat{F}_n(x) - F(x) \right| = O\big((\ln n/n)^{1/2}\big).$$

Together with Theorem 1 this implies

$$\hat{F}_n \hat{F}_n^{-1}(p) = F\hat{F}_n^{-1}(p) + R_n(p) = FF^{-1}(p) + \left[\hat{F}_n^{-1}(p) - F^{-1}(p) \right] f(\xi_n(p)) + R_n(p),$$

for some $\xi_n(p)$ between $F^{-1}(p)$ and $\hat{F}_n^{-1}(p)$, and where almost surely

$$R_n(p) = O\big((\ln n/n)^{1/2}\big) \quad \text{uniformly in } p_0 \leqslant p \leqslant p_1.$$

Hence

$$\hat{F}_n^{-1}(p) - F^{-1}(p) = \left[\hat{F}_n \hat{F}_n^{-1}(p) - FF^{-1}(p) \right] / f(\xi_n(p)) - R_n(p)/f(\xi_n(p)).$$

Thus the lemma is an immediate consequence of Lemma 3.1. $\square$

The next lemma provides a special version of the oscillation behavior of $\hat{F}_n$.

Lemma 3.3. *Let $\lambda_n = \text{const}(\ln n/n)^{1/2}$, and $a_G < a < b < b_F$. If F is Lipschitz continuous on $[a, b]$, then*

$$\sup_{\substack{|s-t| \leqslant \lambda_n \\ a \leqslant s, t \leqslant b}} \left| \hat{F}_n(s) - F(s) - \hat{F}_n(t) + F(t) \right| = O\big((\ln n/n)^{3/4}\big) \quad a.s.$$

Proof. We first show the statement for $\Lambda_n - \Lambda$ rather than $\hat{F}_n - F$. According to (2.1) it remains to show that uniformly in $|s - t| \leqslant \lambda_n$ both

$$\int_{(s,t]} C^{-1}(z)\left[F_n^*(\mathrm{d}z) - F^*(\mathrm{d}z)\right] \quad \text{and} \quad \int_{(s,t]} \frac{C_n(z) - C(z)}{C^2(z)} F^*(\mathrm{d}z)$$

are of the stated order. From the LIL for empirical d.f.'s and the Lipschitz continuity of F, the second integral is even bounded by $O(\sqrt{\ln n \ln \ln n}/n) = O(\ln n/n)$.

As to the first integral, put $\delta_n = K_0[\ln n/n]^{3/4}$. Introduce the grid

$$x_{n,j} = a + j\delta_n, \quad j = 0, 1, \ldots, k_n = [(b-a)/\delta_n], \qquad x_{n,k_{n+1}} = b.$$

For each x choose $x_{n,j}$ such that $x_{n,j} \leqslant x \leqslant x_{n,j+1}$. Since C is nonnegative and F is Lipschitz continuous,

$$\int_{(s, t]} C^{-1}(z)\left[F_n^*(dz) - F^*(dz)\right] \leqslant \int_{s_{n,j}}^{t_{n,j+1}} C^{-1}(z)\left[F_n^*(dz) - F^*(dz)\right] + O(\delta_n)$$

$$\equiv \eta_{nj} + O(\delta_n). \tag{3.1}$$

Each η_{nj} is an average of i.i.d. random variables with expectation zero, being bounded by some common constant M. Furthermore,

$$n \operatorname{Var}(\eta_{nj}) \leqslant \int_{s_{n,j}}^{t_{n,j+1}} C^{-2}(z) F^*(dz) = O(\lambda_n).$$

Bennett's inequality (cf. Shorack and Wellner, 1986, p. 851) yields

$$\mathbb{P}(\eta_{nj} \geqslant \delta_n) \leqslant \exp\left[-\frac{K_1 n \delta_n^2}{2\lambda_n + \frac{2}{3} M \delta_n}\right], \tag{3.2}$$

where $K_1 = K_1(K_0)$ increases with K_0. In particular, this can be made $O(n^{-3})$ if we choose the constant large enough. Since there are at most $O(n^{3/2})$ η_{nj}'s, (3.2) together with Borel–Cantelli implies that with probability one,

$$\max_j \eta_{nj} \leqslant \delta_n \quad \text{eventually.}$$

Recall (3.1) to obtain the desired upper bound for the oscillation modulus of $\Lambda_n - \Lambda$. The lower bound is derived similarly. The corresponding result for $\hat{F}_n - F$ follows by taking logarithms of $1 - \hat{F}_n$ and $1 - F$, then using a Taylor expansion of the logarithm and finally applying the Lipschitz continuity of F. $\quad\square$

We are now ready to state the main results on the quantile representation for $\hat{F}_n$.

Theorem 2. *Assume $a_G \leqslant a_F$ and $b_G \leqslant b_F$ and that F is Lipschitz continuous. Let $0 < p < 1$, and suppose that F is continuously differentiable at $F^{-1}(p)$ with $F'(F^{-1}(p)) = f(F^{-1}(p)) > 0$. Consider the representation*

$$\hat{F}_n^{-1}(p) - F^{-1}(p) = \frac{p - \hat{F}_n F^{-1}(p)}{F'(F^{-1}(p))} + R_{n1}(p) = -\frac{1-p}{F'(F^{-1}(p))} S_n(0, F^{-1}(p)) + R_{n2}(p),$$

where S_n is defined in (2.1). Then, if $\int dF / G^2 < \infty$,

$$R_{ni}(p) = o\left((\ln n / n)^{1/2}\right) \quad a.s. \quad \text{and} \quad R_{ni}(p) = o_{\mathbb{P}}(n^{-1/2}), \quad i = 1, 2.$$

If, in addition, F is twice continuously differentiable at $F^{-1}(p)$,

$$R_{ni}(p) = O\left((\ln n / n)^{3/4}\right) \quad a.s., \quad i = 1, 2. \tag{3.3}$$

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

Proof. Lemmas 3.2 and 3.3 imply that with probability one,

$$\hat{F}_n \hat{F}_n^{-1}(p) - \hat{F}_n F^{-1}(p)$$

$$= F\hat{F}_n^{-1}(p) - F F^{-1}(p) + O\left((\ln n / n)^{3/4}\right)$$

$$= f(F^{-1}(p))\left[\hat{F}_n^{-1}(p) - F^{-1}(p)\right] + o\left(\hat{F}_n^{-1}(p) - F^{-1}(p)\right) + O\left((\ln n / n)^{3/4}\right). \tag{3.4}$$

The assertion now follows from Lemma 3.1. As to the representation in terms of $S_n(0, x)$, apply (2.3). If F is twice differentiable at $F^{-1}(p)$, one can further expand (3.4) one more term to get

$$O\left(\hat{F}_n^{-1}(p) - F^{-1}(p)\right)^2 \text{ instead of } o\left(\hat{F}_n^{-1}(p) - F^{-1}(p)\right).$$

The remainder is therefore $O((\ln n/n)^{3/4})$. Finally, since the bounds leading to (3.4) already are uniform in $p_0 \leqslant p \leqslant p_1$, we only have to note that also in the Taylor expansion the error bounds hold uniformly under the stated regularity assumptions on F.   □

4. Applications

We will show in this section some implications of the quantile representations in Theorem 2: Asymptotic normality and LIL of quantiles, two types of confidence intervals for $F^{-1}(p)$ together with their asymptotic lengths and, finally, a confidence band for F^{-1}.

An immediate consequence of Theorem 2 is the following:

Theorem 3 (Asymptotic normality and LIL). *Under the assumptions of Theorem 2 guaranteeing*

$$R_{n2}(p) = o_{\mathbb{P}}(n^{-1/2})$$

we have

(a) $n^{1/2}\left[\hat{F}_n^{-1}(p) - F^{-1}(p)\right] \to \mathcal{N}(0, \rho^2),$

where

$$\rho^2 = \sigma^2\left(F^{-1}(p)\right)\left[f\left(F^{-1}(p)\right)\right]^{-2}$$

and

$$\sigma^2(t) = \alpha[1 - F(t)]^2 \int_0^t G^{-1}(z)[1 - F(z)]^{-2} F(\mathrm{d}z) = [1 - F(t)]^2 \int_0^t \frac{F^*(\mathrm{d}z)}{C^2(z)}.$$

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

(b) $\displaystyle \limsup_{n \to \infty} \sqrt{n/\ln \ln n}\left[\hat{F}_n^{-1}(p) - F^{-1}(p)\right] = \sqrt{2\rho^2}$ *a.s.*   □

Remark 4.1. Also, a multivariate version of Theorem 3 is available upon applying the standard Cramér–Wold device. Tightness of the mean process $n^{1/2}f(F^{-1}(p))S_n(0, F^{-1}(p)) \equiv Z_n(p)$ on $0 < p_0 \leqslant p \leqslant p_1 < 1$, can be shown by verifying the moment condition in Billingsley (1968, p. 128). The weak convergence of the quantile process

$$Q_n(p) = n^{1/2}f\left(F^{-1}(p)\right)\left[\hat{F}_n^{-1}(p) - F^{-1}(p)\right], \quad p_0 \leqslant p \leqslant p_1,$$

to a Gaussian process thus follows.

Theorem 4. *Under the assumptions of Theorem 2,*

$$Q_n \to Z \quad \text{in distribution}$$

in the space $D[\,p_0,\ p_1]$ of left-continuous functions with right-hand limits, where Z is a zero mean Gaussian process with continuous sample paths and covariance function

$$\mathrm{Cov}\big[Z(\,p_1),\ Z(\,p_2)\big] = \alpha(1-p_1)(1-p_2)\int_0^{F^{-1}(p_1)\wedge F^{-1}(p_2)} G^{-1}(u)\big[1-F(u)\big]^{-2}F(\mathrm{d}u). \qquad \square$$

Let $0 < p < 1$. An approximate confidence interval for $F^{-1}(p)$ can be established immediately from Theorem 3. Under the conditions of Theorem 3, for $0 < \gamma < 1$, let $Z_\gamma = \Phi^{-1}(1-\gamma)$ denote the $1-\gamma$ quantile of the standard normal distribution. Then

$$\hat{F}_n^{-1}(\,p) \pm Z_{\gamma/2} n^{-1/2} \hat{\sigma}_n / \hat{f}_n\big(\hat{F}_n^{-1}(\,p)\big) \qquad (4.1)$$

is an approximate level $1-\gamma$ interval for $F^{-1}(p)$. Here $\hat{f}_n$ is some nonparametric estimate of f and $\hat{\sigma}_n^2$ is some consistent estimate of $\sigma^2(F^{-1}(p))$. Although estimation of f is feasible it can be avoided. The next method which constructs a confidence interval based on the order statistics of the X's eliminates this drawback. This interval is of the form

$$\Big[\hat{F}_n^{-1}(\underline{p}_n),\ \hat{F}_n^{-1}(\bar{p}_n)\Big], \qquad (4.2)$$

where $p - \underline{p}_n$ and $\bar{p}_n - p$ are approximately of the order $n^{-1/2}$. Note that the upper and lower bounds of this interval correspond to order statistics of X. More precisely, write $\gamma = \gamma_1 + \gamma_2$ and let

$$\underline{p}_n = p - Z_{\gamma_1} n^{-1/2} \hat{\sigma}_n, \qquad \bar{p}_n = p + Z_{\gamma_2} n^{-1/2} \hat{\sigma}_n. \qquad (4.3)$$

Observe that here we did not require $\gamma_1 = \gamma_2 = \frac{1}{2}\gamma$ as in (4.1), since such a requirement does not necessarily yield an interval of minimal length. However, the asymptotic length of the interval (4.2) is shortest for $\gamma_1 = \gamma_2 = \frac{1}{2}\gamma$, since by the uniform version of (3.3) and Lemma 3.3,

$$n^{1/2}\Big[\hat{F}_n^{-1}(\bar{p}_n) - \hat{F}_n^{-1}(\underline{p}_n)\Big] \to \big(Z_{\gamma_1} + Z_{\gamma_2}\big)\sigma\big(F^{-1}(p)\big)/f\big(F^{-1}(p)\big) \quad \text{a.s.}$$

If one chooses $\gamma_1 = \gamma_2 = \frac{1}{2}\gamma$ in (4.3) then the asymptotic length is the same as that of (4.1), namely $2Z_{\gamma/2}\sigma(F^{-1}(p))/f(F^{-1}(p))$.

Confidence intervals of the form (4.1) and (4.2) have been studied for randomly censored data independently by Gijbels and Veraverbeke (1988) and Wang and Hettmansperger (1990).

Remark 4.2. Several choices of consistent estimators for $\sigma^2(F^{-1}(p))$ are available. Let

$$V_{np} = \sum_{X_i \le \hat{F}_n^{-1}(p)} \big[nC_n^2(X_i)\big]^{-1}.$$

Two choices of $\hat{\sigma}_n^2$ are given below:

$$\hat{\sigma}_{n1}^2 = (1-p)^2 V_{np}, \qquad \hat{\sigma}_{n2}^2 = \Big[1 - \hat{F}_n \hat{F}_n^{-1}(p)\Big]^2 V_{np}.$$

Using Theorem 1 and standard empirical arguments, it can be shown that both estimators are consistent.

We now address the possibility of constructing confidence bands for $F^{-1}(p)$, $p_0 \le p \le p_1$. For this, let $(d_n)_n$ be a sequence of nonnegative real numbers to be specified later, with limit $d_0 \ge 0$. Put, for $p_0 \le p \le p_1$,

$$I_n(p) = \Big[\hat{F}_n^{-1}\big(p - (1-p)d_n n^{-1/2}\big),\ \hat{F}_n^{-1}\big(p + (1-p)d_n n^{-1/2}\big)\Big],$$

where for the sake of definiteness $\hat{F}_n^{-1}(u)$ is the least upper respectively largest lower bound for the

support of $\hat{F}_n$ if $u \geqslant 1$ respectively $u \leqslant 0$. Utilize representation (2.3) and a continuity argument to show that

$$\mathbb{P}\big(F^{-1}(p) \in I_n(p) \text{ for } p_0 \leqslant p \leqslant p_1\big) \sim \mathbb{P}\bigg(\sup_{p_0 \leqslant p \leqslant p_1} |S_n(0, F^{-1}(p))| \leqslant d_n\bigg)$$

$$\to \mathbb{P}\bigg(\sup_{p_0 \leqslant p \leqslant p_1} |W(d(F^{-1}(p)))| \leqslant d_0 \alpha^{-1/2}\bigg) \quad \text{as } n \to \infty.$$

Here the function $d(t)$ is defined by

$$d(t) = \int_0^t G^{-1}(u)[1 - F(u)]^{-2} F(\mathrm{d}u)$$

and W denotes a standard Wiener process. Clearly, the above probability equals

$$\mathbb{P}\bigg(\sup_{d(F^{-1}(p_0)) \leqslant s \leqslant d(F^{-1}(p_1))} |W(s)| \leqslant d_0 \alpha^{-1/2}\bigg). \tag{4.4}$$

Note that

$$\alpha d(t) = \int_0^t \frac{F^*(\mathrm{d}z)}{C^2(\mathrm{d}z)} \equiv e(t)$$

so that (4.4) reduces to

$$\mathbb{P}\bigg(\sup_{e(F^{-1}(p_0)) \leqslant s \leqslant e(F^{-1}(p_1))} |W(s)| \leqslant d_0\bigg).$$

Chung (1987) provided a computer package which computes the probabilities

$$\mathbb{P}\bigg(\sup_{s_0 \leqslant s \leqslant s_1} |W(s)| \leqslant d_0\bigg). \tag{4.5}$$

If we choose d_0 such that for $s_i = e(F^{-1}(p_i))$, $i = 0, 1$, the last probability equals γ, this would lead to a confidence band with (asymptotic) coverage level γ. Since in practice s_0 and s_1 are unknown they need to be replaced by

$$s_{in} = \int_0^{\hat{F}_n^{-1}(p_i)} \frac{F_n^*(\mathrm{d}z)}{C_n^2(z)}, \quad i = 0, 1.$$

Finally, choose d_n such that given s_{0n} and s_{1n},

$$\mathbb{P}\bigg(\sup_{s_{0n} \leqslant s \leqslant s_{1n}} |W(s)| \leqslant d_n\bigg) = \gamma.$$

Remark 4.3. For the random censorship model confidence bands for F^{-1} which are similar in spirit were proposed by Aly et al. (1985).

References

Aly, E.-E.A.A., M. Csörgő and L. Horváth (1985), Strong approximations of the quantile process of the product-limit estimator, *J. Multivariate Anal.* 16, 185–210.

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

Billingsley, P. (1968), *Convergence of Probability Measures* (Wiley, New York).

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

Cheng, K.-F. (1984), On almost sure representation for quantities of the product limit estimator with applications, *Sankhyā Ser. A* **46**, 426–443.

Chung, C.-J.F. (1987), Wiener pack – A subroutine package for computing probabilities associated with Wiener and Brownian bridge processes, Paper 87-12, Geol. Surv. of Canada (Ottawa, Ont.).

Gastwirth, J.L. and J.-L. Wang (1988), Control percentile test procedures for censored data, *J. Statist. Plann. Inference* **18**, 267–276.

Ghosh, J.K. (1971), A new proof of the Bahadur representation of quantiles and an application, *Ann. Math. Statist.* **42**, 1957–1961.

Gijbels, I. and N. Veraverbeke (1988), Weak asymptotic representations for quantiles of the product-limit estimator, *J. Statist. Plann. Inference* **18**, 151–160.

Gu, M.G. 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.

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

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

Lo, S.-H. and K. Singh (1986), The product-limit estimator and the bootstrap: some asymptotic representations, *Probab. Theory Rel. Fields* **71**, 455–465.

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

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

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

Stute, W. (1982), The oscillation behavior of empirical processes, *Ann. Probab.* **10**, 86–107.

Stute, W. (1993), Almost sure representations of the product-limit estimator for truncated data, to appear in: *Ann. Statist.*

Wang, J.-L. and T.P. Hettmansperger (1990), Two-sample inference for median survival times based on one-sample procedures for censored survival data, *J. Amer. Statist. Assoc.* **85**, 529–536.

Wang, M.-C., N.P. Jewell and W.-Y. Tsai (1986), Asymptotic properties of the product limit estimate under random truncation, *Ann. Statist.* **14**, 1599–1605.

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