Inference for Shift Functions in the Two-Sample Problem With Right-Censored Data: With Applications

Henry H. S. Lu, Martin T. Wells, and Ram C. Tiwari*

For two distribution functions, F and G, the shift function is defined by $\Delta(t) \equiv G^{-1} \circ F(t) - t$. The shift function is the distance from the 45° line and the quantity plotted in Q-Q plots. In the analysis of lifetime data, Δ represents the difference between two treatments. The shift function can also be used to find crossing points of two distribution functions. The large-sample distribution theory for estimates of Δ is studied for right-censored data. It turns out that the asymptotic covariance function depends on the unknown distribution functions F and G; hence simultaneous confidence bands cannot be directly constructed. A construction of simultaneous confidence bands for Δ is developed via the bootstrap. Construction and application of such bands are explored for the Q-Q plot.

KEY WORDS: Bootstrap; Censored data; Crossing points; Q-Q plots; Shift function; Treatment effect; Two-sample problems.

1. INTRODUCTION

Suppose that F and G denote the distribution functions of random variables X and Y. For the distributions F and G, the horizontal shift (or translation) function is defined by

$$\Delta(t) \equiv G^{-1} \circ F(t) - t. \tag{1}$$

Analogously, we can define the vertical shift as

$$\Delta^{\dagger}(t) \equiv G \circ F^{-1}(t) - t. \tag{2}$$

These shifts are important measures in the analysis of survival data. In the analysis of lifetime data, it is often necessary to estimate the difference between treatments. Suppose that X is the control and Y is the treatment; in such a situation, the lifetimes of the treatment and control groups can be compared. Doksum (1974) proved that the shift function $\Delta(\cdot)$ in (1) is the unique function such that $X + \Delta(X) =_d Y$; that is, $X + \Delta(X)$ equals Y in distribution. If $\Delta(X) \equiv \Delta$, a constant, then the model reduces to a linear model; otherwise, it is a nonlinear model. Obviously, $\Delta(X) = 0$ if and only if $F = G$. Doksum proposed an empirical estimator, $\hat{\Delta}_N(t) = G_m^{-1} \circ F_n(t) - t$, of Δ and proved its weak convergence; where $N = m + n$, F_n is the empirical distribution function (edf) of a sample of size n from F and G_m^{-1} is the empirical quantile function (eqf), the inverse of G_m. Doksum and Sievers (1976) constructed simultaneous confidence bands for Δ in the case of no censoring. The derivation of their bands depends on certain approximating assumptions. Hollander and Korwar (1982) and Wells and Tiwari (1989a) extended the results of Doksum (1974) to the nonparametric Bayesian framework. These results seem to lead us to the possibility of constructing simultaneous confidence bands for Δ.

When $\Delta(\cdot) \equiv \Delta$, it is easy to see that Δ is the median of the distribution of $Y - X$. Meng, Bassiakos, and Lo (1991) studied large-sample properties of the censored data analog of the Hodges and Lehmann (1963) estimator of Δ. Their results are somewhat restrictive, as they assumed that $\Delta(\cdot)$ is constant, which is the case in the analysis of a treatment effect but not in general, however. Padgett and Wei (1980) and Wei and Gail (1983) studied the two-sample scale problem with censored data. Wang and Hettmansperger (1990) gave related results on two-sample inference for median survival times based on one-sample procedures for censored data.

The functions $\Delta(t)$ and $\Delta^{\dagger}(t)$ also play an important role in graphical statistics. Graphical methods are very powerful tools for data analysis. The compatibility of the proposed model with the observed data may be determined easily by a graphical examination of the fit. Two well-known plots for goodness of fit are the quantile (Q-Q) plots and the probability (P-P) plots. These were discussed in detail by Wilk and Gnanadesikan (1968). The function Δ is the distance from the 45° line and the quantity plotted in the Q-Q plots. If $F = G$, then the Q-Q plot is a 45° straight line; otherwise, the plot is in a different shape. Such a plot would be useful when comparing survival functions. A simultaneous confidence band would display more information than the simple Q-Q plot. But the limiting distribution theory shows that the asymptotic covariance function depends on the unknown distribution functions F and G; hence the band cannot be directly constructed. A bootstrap solution to this problem is discussed in Section 2.

Assume that the sequence $\{X_i^0\}_{i=1}^n$ of survival times is iid from a continuous df F on $[0, \infty)$ and that the sequence $\{C_i\}_{i=1}^n$ of censoring times is iid from a continuous df H_1 on $[0, \infty)$. Furthermore, suppose that the C_i's are independent of the X_i^0's. The observed data are $\{X_i, \delta_i\}_{i=1}^n$, where $\delta_i = \mathbb{I}[X_i^0 \le C_i]$ is the indicator function for the event and $X_i = X_i^0 \wedge C_i = \min\{X_i^0, C_i\}$. This model is useful when the observations are incomplete due to random censoring. The product limit (PL) estimator of the survival function, $\bar{F}(=1 - F)$, proposed by Kaplan and Meier (1958) is defined as

* Henry H. S. Lu is a graduate student and Martin T. Wells is Associate Professor, Department of Economic and Social Statistics and the Statistics Center, Cornell University, Ithaca, NY 14853. Ram C. Tiwari is Professor of Mathematics, University of North Carolina, Charlotte, NC 28223. This research was partially supported by the School of Industrial and Labor Relations at Cornell and by National Institutes of Health Grant R01 CA61120. The authors thank the associate editor and the two referees for helpful discussion and suggestions.

Journal of the American Statistical Association
September 1994, Vol. 89, No. 427, Theory and Methods

$$1 - \hat{F}_n(t) = \prod_{\{X_{n:i} \le t\}} \left[1 - \frac{1}{n-i+1}\right]^{\delta_{n:i}}$$

$$= \prod_{\{X_{n:i} \le t\}} \left[1 - \frac{\delta_{n:i}}{n-i+1}\right], \quad (3)$$

where $X_{n:1} \le X_{n:2} \le \cdots \le X_{n:n}$ are the order statistics of $X_1, X_2, \ldots, X_n$ and $\delta_{n:i}$ are the induced order statistics (i.e., $\delta_{n:i}$ is the δ_i corresponding to $X_{n:i}$). When there is no censoring, $\delta_i = 1$ for all i, and the PL estimator in (3) reduces to the empirical survival function $(1 - F_n)$. Special attention shall be paid to the case $\lim_{t\to\infty} [1 - \hat{F}_n(t)] \ne 0$ when the last ordered observation $X_{n:n}$ is censored; that is, when $\delta_{n:n} = 0$. We will redefine $[1 - \hat{F}_n(t)] = 0$ for $t \ge X_{n:n}$ in this case. The PL estimator has many desirable properties, including consistency and asymptotic normality, and is the generalized maximum likelihood estimator (see Miller 1981, p. 57).

In this article we construct a nonparametric method to test the difference of characteristics of two independently censored samples using an estimate of $\Delta(\cdot)$. First, we have to discuss a model of random censorship for the two-sample case. In the first sample, the observed data are $\{(X_i, \delta_i)\}_{i=1}^n$, where $\delta_i = \mathbb{I}[X_i^0 \le C_i]$ and $X_i = X_i^0 \wedge C_i$ as discussed previously. In the second sample, assume that the sequence $\{Y_j^0\}_{j=1}^m$ of survival times is iid from a continuous df G on $[0, \infty)$ and that the sequence $\{D_j\}_{j=1}^m$ of censoring times is iid from a continuous df H_2 on $[0, \infty)$. Furthermore, suppose that the D_j's are independent of the Y_j^0's. The observed data for the second sample is $\{(Y_j, \gamma_j)\}_{j=1}^m$, where $\gamma_j = \mathbb{I}[Y_j^0 \le D_j]$ and $Y_j = Y_j^0 \wedge D_j$. In addition, assume that these two samples are independent. We develop a procedure to test whether $F = G$, based on the censored data $\{(X_i, \delta_i)\}_{i=1}^n$ and $\{(Y_j, \gamma_j)\}_{j=1}^m$. We use the bootstrap methodology to construct simultaneous confidence bands for the function $\Delta(t)$.

The validity of the bootstrap has been demonstrated by Bickel and Freedman (1981) and Gill (1989). Akritas (1986) extended the results of Bickel and Freedman (1981) to cover censored data problems. Wells and Tiwari (1989b) showed the asymptotic consistency of the Bayesian bootstrap with censored data. These results show that the bootstrap of the PL estimator and nonparametric Bayes estimator are both valid and asymptotically equivalent. There are two possible resampling schemes for this problem, one proposed by Efron (1981) and the other suggested by Reid (1981). Efron (1981) resampled $\{X_j^{0*}\}_{j=1}^{n*}$ iid from F_n (the edf of $\{X_i^0\}_{i=1}^n$), $\{C_j^*\}_{j=1}^{n*}$ iid from H_{1n} (the edf of $\{C_i\}_{i=1}^n$) and then used the Kaplan–Meier estimator for F_n corresponding to the new data $\{(X_j^*, \delta_j^*)\}_{j=1}^{n*}$, where $X_i^* = X_i^{0*} \wedge Y_i^{0*}$ and $\delta_i^* = \mathbb{I}[X_i^{0*} \ge Y_i^{0*}]$. Efron showed that his scheme is equivalent to resampling $\{(X_j^*, \delta_j^*)\}_{j=1}^{m*}$ with replacement from $\{(X_i, \delta_i)\}_{i=1}^n$. Based on this new data $\{(X_j^*, \delta_j^*)\}_{j=1}^{m*}$, one can construct the bootstrap PL estimator. Reid (1981) resampled the new data iid from the PL estimator and then used them to construct the bootstrap estimator. Reid's plan used a limit process different from that of the original PL estimator. Akritas (1986) showed that Reid's approach does not generate correct asymptotic confidence bands but that Efron's approach does. Consequently, Efron's approach is the procedure to follow. Akritas (1986) also showed how to construct appropriate confidence bands under Efron's sampling scheme.

In this article we give the construction of simultaneous confidence bands for the horizontal distance between two distribution functions, $\Delta(\cdot)$. In Section 2 we provide the necessary distribution theory and construction. In Section 3 we illustrate the proposed procedure with some examples and study the size and power of the corresponding goodness-of-fit procedure. In the Appendix we present most of the proofs along with some general results about bootstrapping functional of the PL estimator.

2. BOOTSTRAPPED SHIFT FUNCTIONS AND Q-Q PLOTS FOR CENSORED DATA

In this section we develop a nonparametric graphical procedure to test the difference between two independent censored samples. Let F and G be continuous distribution functions on $[0, \infty)$. The horizontal distance between the two distribution functions is defined in (1). As mentioned earlier, Doksum (1974) proved that the shift function $\Delta(\cdot)$ in (1) is the unique function such that $X + \Delta(X)$ equals Y in distribution. Hence this distance gives a measure that characterizes the difference between the distribution functions F and G. A common approach to assessing the magnitude of Δ is by an inspection of a graphical procedure, as in Q-Q plots. But as with any graphical procedure, any inference about the parameter of interest may be influenced by the viewer's interpretation. Therefore, it is of interest to study the significant deviations of estimates of Δ from a particular value (such as 0). In this section we construct simultaneous confidence bands that accomplish this goal. Doksum and Sievers (1976) constructed approximate simultaneous confidence bands for Δ in the case of no censoring; we extend these results to the case of censored data. This extension is not at all trivial. The limiting distribution theory shows that the asymptotic covariance function depends on the unknown distribution functions F and G; hence the simultaneous confidence bands cannot be directly constructed. A bootstrap solution to this problem is proposed. These simultaneous confidence bands also give a method to assess whether a treatment effect is constant or is varying as a function of time. This new method is applied in Section 3.

We suppose that the data $\{(X_i, \delta_i)\}_{i=1}^n$ and $\{(Y_j, \gamma_j)\}_{j=1}^m$ are randomly right-censored data from F and G. Let $\hat{F}_n$ and $\hat{G}_m$ be the corresponding PL estimators of F and G. Define the PL quantile estimator of G as $\hat{G}_m^{-1}(t) = \inf\{x: \hat{G}_m(x) > t\}$. Hence define the PL shift estimator of Δ as

$$\hat{\Delta}_{mn}(t) = \hat{G}_m^{-1} \circ \hat{F}_n(t) - t. \quad (4)$$

Following the approach of Efron (1982), we construct the bootstrap samples $(X_j^*, \delta_j^*)\}_{j=1}^{n*}$ and $\{(Y_j^*, \gamma_j^*)\}_{j=1}^{m*}$ from $\{(X_i, \delta_i)\}_{i=1}^n$ and $\{(Y_j, \gamma_j)\}_{j=1}^m$. Based on the bootstrap samples, define the corresponding PL estimators of F and G by $\hat{F}_n^*$ and $\hat{G}_m^*$. Similarly, define the bootstrap PL quantile estimator of G as $\hat{G}_m^{*-1}$. Hence define the bootstrap PL shift estimator of Δ as

$$\hat{\Delta}_{mn}^*(t) = \hat{G}_m^{*-1} \circ \hat{F}_n^*(t) - t. \quad (5)$$

Before constructing the simultaneous confidence bands for Δ in the case of censoring, we need to prove the weak convergence of the bootstrapped shift process

$$\hat{\mathbb{D}}_N = \sqrt{\frac{mn}{N}}\,[\hat{\Delta}_{mn}^*(t) - \hat{\Delta}_{mn}(t)]$$

$$= \sqrt{\frac{mn}{N}}\,[\hat{G}_m^{*-1} \circ \hat{F}_n^*(t) - \hat{G}_m^{-1} \circ \hat{F}_n(t)], \quad (6)$$

where $N = m + n$. The preliminary tools for the proof of the convergence of this bootstrapped process are given in the Appendix. These results are complicated and involve some abstract concepts from topological vector space theory and empirical process theory. The interested reader may find the methodology in the proofs quite illuminating.

The weak convergence of this bootstrapped process with censored data is as follows. Let T_1 and T_2 be finite constants such that $[1 - F(T_1)][1 - H_1(T_1)] > 0$ and $[1 - G(T_2)][1 - H_2(T_2)] > 0$. As convention, we use the following notations: $=_d$, $\rightarrow_{\|\cdot\|}$, $\rightarrow_d$, $\rightarrow_{\text{a.s.}}$, $\rightarrow_p$, $\Rightarrow$ to mean equal in distribution, convergence in sup norm $\|\cdot\|$, converge in distribution, converge almost surely, converge in probability and weak convergence. Let $D[a, b]$, the space of cadlag real-valued functions on the interval $[a, b]$. The first result is the weak convergence of the bootstrapped shift process.

Theorem 2.1. Let g be the probability density function of G. Assume that $g(G^{-1}(\cdot))$ is continuous and bounded away from 0 on $[0, T_2]$. Then

$$\hat{\mathbb{D}}_N(t) \Rightarrow Z(t)/g \circ G^{-1} \circ F(t) \quad \text{as} \quad m \wedge n \rightarrow \infty$$
$$\text{and} \quad n/N \rightarrow \theta \in [0, 1] \quad (7)$$

on $D[0, T_1]$, where $Z(\cdot)$ is a mean 0 Gaussian process with covariance function

$$C(s,t) = (1-\theta)C_1(s,t) + \theta C_2(G^{-1} \circ F(s), G^{-1} \circ F(t), \quad \text{for} \quad s, t \in [0, T_1],$$

where $C_1(\cdot, \cdot)$ and $C_2(\cdot, \cdot)$ are defined in Lemma A.3 and Corollary A.6.

Using Theorem 2.1, we can construct the bootstrap confidence bands for the shift process $\Delta(\cdot)$ based on $\hat{\Delta}_{mn}$. The construction for two-sample case is analogous to that for the one-sample case of Akritas (1986). We outline this construction in the remainder of this section. The first matter to address is estimating the denominator in (7). At first glance, it seems difficult to estimate $g \circ G^{-1} \circ F$. But from the proofs of Theorem 2.1 and Theorem A.5, note that $d(Q \circ F) = dG^{-1} \circ F = 1/g \circ G^{-1} \circ F$, where $\phi(\cdot) = (\cdot)^{-1}$ is the inverse function. To have the needed differentiability, we must develop a smooth version of $\hat{G}_m^{-1}$. Padgett (1986) and Lio and Padgett (1992) studied a kernel-type smoothed estimator of the quantile function of the PL estimator for censored data, deduced its asymptotic convergence, and addressed the bandwidth selection issue. The following methodology was suggested. Choose a kernel K that is a probability density function with a finite support that is symmetric about 0 and satisfies a Lipschitz condition. Let $\{h_m\}$ be the bandwidth sequence of positive numbers such that $h_m \rightarrow 0$ as $m \rightarrow \infty$. Let $Q \equiv G^{-1}$ and $\hat{Q}_m \equiv \hat{G}_m^{-1}$. Define the kernel-type quantile function estimator for $0 < p < 1$ as

$$Q_m(p) = h_m^{-1} \int_0^1 \hat{Q}_m(t) K((t-p)/h_m)\,dt$$

$$= h_m^{-1} \sum_{j=1}^{m} Y_{m:j} \int_{G_{j-1}}^{G_j} K((t-p)/h_m)\,dt,$$

where $G_0 = 0$ and $G_j = \hat{G}_m(Y_{m:j})$ for $j = 1, 2, \ldots, m$. Furthermore,

$$\int_{G_{j-1}}^{G_j} K((t-p)/h_m)\,dt = 0, \quad \text{if } Y_{m:j} \text{ is censored},$$

$$= h_m[K^*((G_j - p)/h_m) - K^*((G_j - p)/h_m)] \quad \text{otherwise},$$

where $K^*(\cdot)$ denotes the cdf of K.

For our application, we need to differentiate this estimate, following Sheather and Marron (1990). The estimate of the derivative of Q_m is

$$dQ_m(p) = -h_m^{-2} \int_0^1 \hat{Q}_m(t) K^{(1)}((t-p)/h_m)\,dt$$

$$= -h_m^{-2} \sum_{j=1}^{m} Y_{m:j} \int_{G_{j-1}}^{G_j} K^{(1)}((t-p)/h_m)\,dt, \quad (8)$$

where $K^{(1)}$, the first derivative of K, is assumed to exist. It may be shown using the foregoing results that if $h_m = O(m^{-1/5})$, then $dQ_m(\cdot) \rightarrow_p dQ(\cdot)$. Hence, using these arguments, we propose to estimate $dQ \circ F = dG^{-1} \circ F = 1/g \circ G^{-1} \circ F$ by $dQ_m \circ \hat{F}_n$. The consistency of this estimator follows from the results of Sheather and Marron (1990). Using the foregoing, we can construct the bootstrap confidence bands of $\hat{\Delta}_{mn}$ as follows.

Theorem 2.2. Suppose that $dQ_m \circ \hat{F}_n(t) \neq 0$ for any $t \in [0, T_1]$, the bandwidth $h_m = O(m^{-1/5})$, and $c_{mn}^*(\Delta)$ is chosen so that for some fixed α, $0 < \alpha < 1$,

$$\Pr\left\{ \sqrt{\frac{mn}{N}} \sup_{0 \le t \le T_1} (|[\hat{\Delta}_{mn}^*(t) - \hat{\Delta}_{mn}(t)]/[dQ_m \circ \hat{F}_n(t)]|) \le c_{mn}^*(\Delta) \,|\, \{(X_i, \delta_i)\}_{i=1}^n, \{(Y_j, \gamma_j)\}_{j=1}^m \right\} = 1 - \alpha.$$

Then

$$\Pr\left\{ \hat{\Delta}_{mn}(t) - \sqrt{\frac{N}{mn}}\, c_{mn}^*(\Delta)[dQ_m \circ \hat{F}_n(t)] \le \Delta(t) \le \hat{\Delta}_{mn}(t) + \sqrt{\frac{N}{mn}}\, c_{mn}^*(\Delta)[dQ_m \circ \hat{F}_n(t)], \ \forall\, t \in [0, T_1] \right\} \rightarrow 1 - \alpha.$$

From this theorem, it is clear that

$$\left\{\hat{\Delta}_{mn}(t) - \sqrt{\frac{N}{mn}}\, c^*_{mn}(\Delta)[dQ_m \circ \hat{F}_n(t)],\right.$$
$$\left.\hat{\Delta}_{mn}(t) + \sqrt{\frac{N}{mn}}\, c^*_{mn}(\Delta)[dQ_m \circ \hat{F}_n(t)]\right\} \quad (9)$$

gives the simultaneous confidence bands for $\Delta(t)$ with asymptotic coverage probability $(1 - \alpha)$. We show the usefulness of this method in the next section. Furthermore, it is not difficult to see that these results are also applicable to one-sided tests as well as to the two-sided tests in the theorem, such as testing $F > G$ or $F < G$.

The methodology developed here could be easily modified to work for any increasing function estimate. Hence our techniques may be applied in a variety of situations. Dabrowska, Doksum, and Song (1989) studied a graphical comparison of cumulative hazards (an increasing function) for the two-sample problem with censored data. We are currently investigating the application of the technique to the comparison of two Lorenz curves.

The results of this section may be used to extend the work of Hawkins and Kochar (1991) on the estimation of the crossing point of two cdf's to the case of censored data. Suppose that there is a unique estimated crossing point, say at $t = t^*_{mn}$; then $\hat{\Delta}_{mn}(t^*_{mn}) = 0$. It is easy to deduce the properties of t^*_{mn} via the mean value theorem by noting that if t^* is the true unique crossing point, then

$$\hat{\Delta}_{mn}(t^*_{mn}) = \hat{\Delta}_{mn}(t^*) + (t^*_{mn} - t^*)\Delta'(\xi) + o_p(N^{-1/2}) = 0,$$

where $|\xi - t^*| \leq |t^*_{mn} - t^*|$. Hence it follows that the behavior of $(t^*_{mn} - t^*)$ can be expressed in terms of the process $\hat{\Delta}_{mn}(t^*)$, which has already been analyzed in Theorem 2.1. The measures developed by Hawkins and Kochar (1991) are essentially continuous functionals of $\hat{\Delta}_{mn}(\cdot)$; therefore, an application of the continuous mapping theorem to our weak convergence results will extend their results to the case of censored data. Granted, this argument is extremely heuristic; however, it could be made rigorous with a bit of work.

3. NUMERICAL STUDIES

As a demonstration of the proposed methodology, we consider one real data illustration and three simulated examples. We also examine the level and power of the Q-Q procedure when viewed as a formal goodness-of-fit test.

The first problem encountered in these Monte Carlo simulation studies is the choice of a smooth kernel and bandwidth used in estimating dQ_m in (8). The kernel K must satisfy the particular conditions. The Gaussian kernel is a possible choice, and we use it here. Because the precise mean squared error of Q_m for censored case is not available, Padgett and Thombs (1986) used the bootstrap method to find an optimal bandwidth for Q_m. It can be shown that the rates $h_m = O(m^{-1/3})$ and $h_m = O(m^{-1/5})$ are the optimal rates for Q_m and dQ_m. The optimal values of the bandwidths can also be found via the method of cross-validation. Note that we use $dQ_m \circ \hat{F}_n$ to estimate $dQ \circ F = 1/g \circ G^{-1} \circ F$ and the optimal rate is $O(m^{-1/5})$, the same rate as that used to smooth a probability density function (pdf), like g (see Silverman 1986, eq. 3.21). Therefore, for simplicity we will use the bandwidth referenced to a standard distribution for the Gaussian kernel, as suggested by Silverman (1986, eq. 3.31). The bandwidth is $h_m = .9Am^{-1/5}$, where $A = \min\{$standard deviation, interquartile range$/1.34\}$.

Table 1. The Three Examples Used in the Monte Carlo Studies

	First sample		Second sample	
	F	H_1	G	H_2
Example 1	exp(1)	$U[0, 2.2316]$	exp(1)	$U[0, 2.2316]$
Example 2	exp(1)	$U[0, 2.2316]$	exp(2)	$U[0, 1.1158]$
Example 3	exp(1)	$U[0, 2.2316]$	exp(1)	$U[0, 4.9651]$

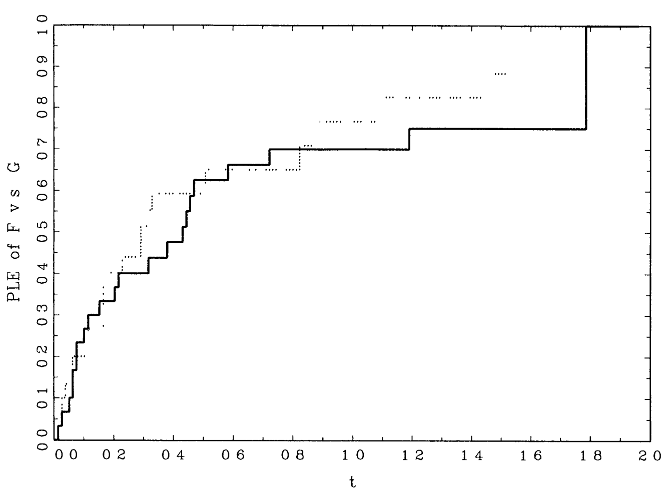

Figure 1. Estimated Distributions of the First Sample From Exponential (1). The 40% uniform censoring times are denoted by the solid line; the second sample from the same situation is denoted by the dotted line.

The models for generating the survival times and the censoring times were adapted from Akritas (1986) and are listed in Table 1 for an easy comparison. In these examples we used sample sizes $m = n = 30$ and used a common (linear congruential) random number generator with multiplier equal to 16,807 and the modulus equal to $2^{31} - 1$. Different initial seeds, 2 and 3, were used to generate the first and second samples. The other initial seed, 1, was used to generate bootstrap resamplings of these two samples. The number of bootstrap resampling was set at 200.

In Example 1, the two samples came from the same model; the survival times came from an exponential distribution with the scale parameter $\beta = 1$, and the censoring times came from a uniform $[0, b]$ distribution with 40% censoring (i.e., $b = 2.2316$ for this case). The only difference in sample data was due to the different initial seeds, 2 versus 3. The resulting PL estimators are shown in Figure 1, corresponding to a solid line and a dotted line. It seems difficult to tell whether $F = G$ if we can judge only from Figure 1. Although the Q-Q plot hits the 45° straight line in several places, it is still hard to decide whether $F = G$ without the confidence

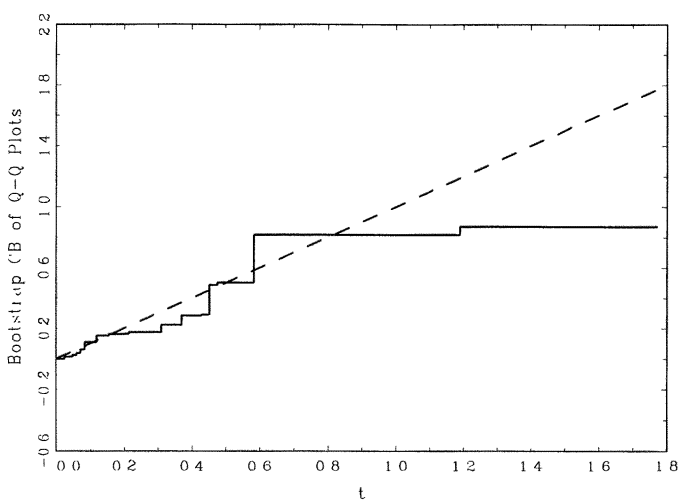

Figure 2. Estimated Q-Q plot of Example 1 (Solid Line) and 90% Bootstrap Confidence Band (Dotted Line). The estimated confidence band includes the 45° dashed line.

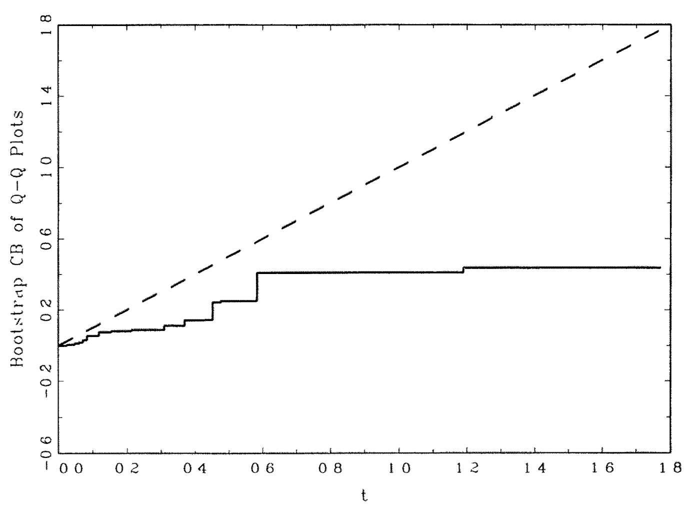

Figure 4. Estimated Q-Q Plot of Example 2 (Solid Line) and 90% Bootstrap Confidence Band (Dotted Line). The estimated confidence band does not include the 45° dashed line.

band. If we choose $\alpha = .10$ and use Theorem 2.2, we can plot the approximate confidence bands of the bootstrap Q-Q plots based on the PL estimators of this example, for $0 \le t \le T_1 < X_{n:n}$, where the upper and lower bands are drawn as dotted lines. Though the lower bands shall be truncated at 0, we keep the original shape for easy viewing. It is clear from Figure 2 that the simultaneous confidence band contains the 45° straight line entirely. Hence we can conclude that $F = G$ at approximately 90% confidence.

In Example 2, the two samples came from two different survival time distributions but with 40% uniform censoring distributions. One was exponential ($\beta = 1$) survival distribution with uniform [0, 2.2316] censoring distribution, and the other was exponential ($\beta = 2$) survival distribution with uniform [0, 1.1158] censoring distribution. The resulting PL estimators are shown in Figure 3. Because the two plots in Figure 3 do not hit each other, the figure shows that $F \neq G$, but the confidence level is unknown. The approximate confidence band of bootstrap Q-Q plot in Figure 4 does not contain the entire 45° straight line. Thus we can judge that $F \neq G$ at approximate 90% confidence.

In Example 3, the two samples came from the same survival distributions with different censoring distributions. The first sample came from a 40% uniform censoring distribution (i.e., $b = 2.2316$ for the first sample), and the second sample came from a 20% uniform censoring distribution (i.e., $b = 4.9651$ for the second sample). The consequent PL estimators are sketched in Figure 5. Once again, we cannot infer whether $F = G$ from Figure 5. The approximate confidence band of bootstrap Q-Q plots is exhibited in Figure 6. As the approximate simultaneous confidence bands contain the 45° straight line entirely, we can determine that $F = G$ at ap-

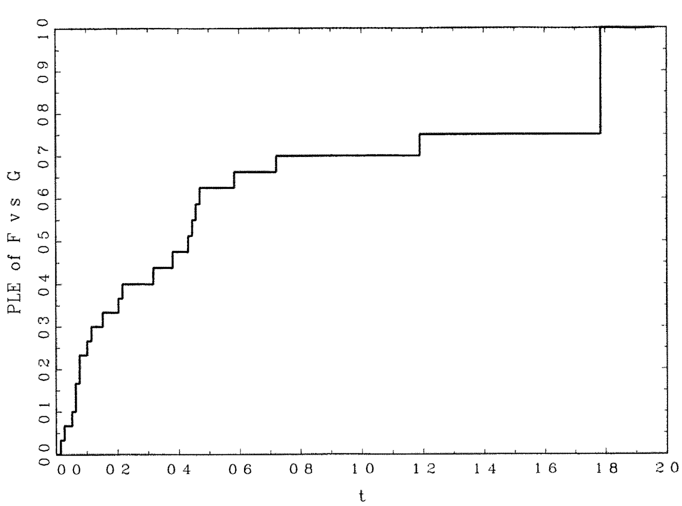

Figure 3. Estimated Distributions of the First Sample Exponential (1) With 40% Uniform Censoring Times (Solid Line) and the Second Sample from Exponential (2) With 40% Uniform Censoring Times (Dotted Line).

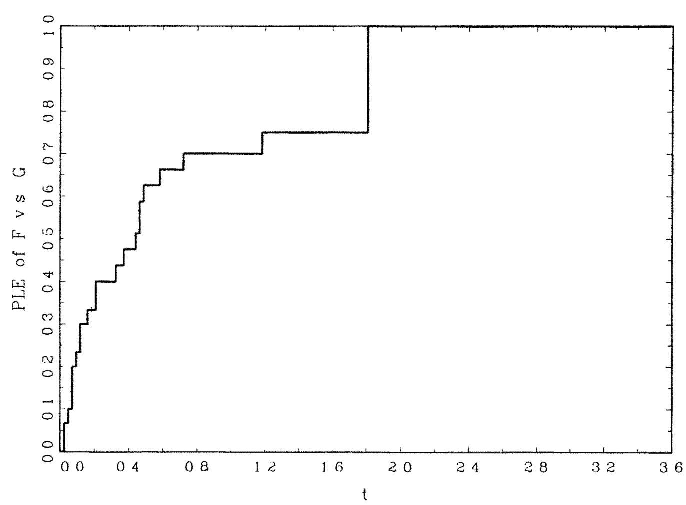

Figure 5. Estimated Distributions of the First Sample From Exponential (1) With 40% Uniform Censoring Times (Solid Line) and the Second Sample From Exponential (1) With 20% Uniform Censoring Times (Dotted Line).

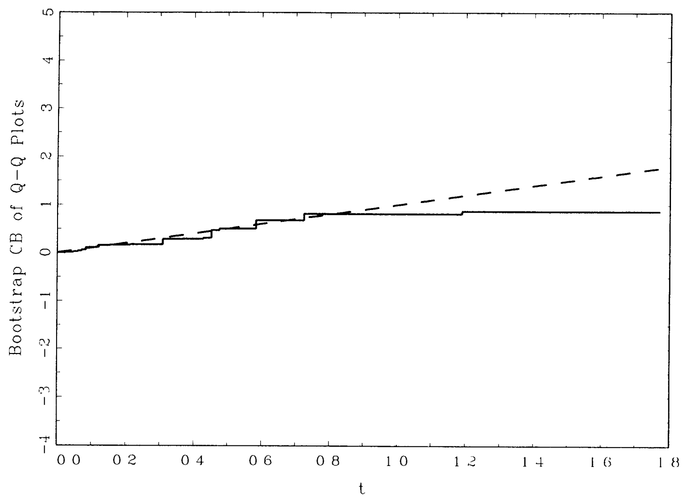

Figure 6. Estimated Q-Q Plot of Example 3 (Solid Line) and 90% Bootstrap Confidence Band (Dotted Line). The estimated confidence band includes the 45° dashed line.

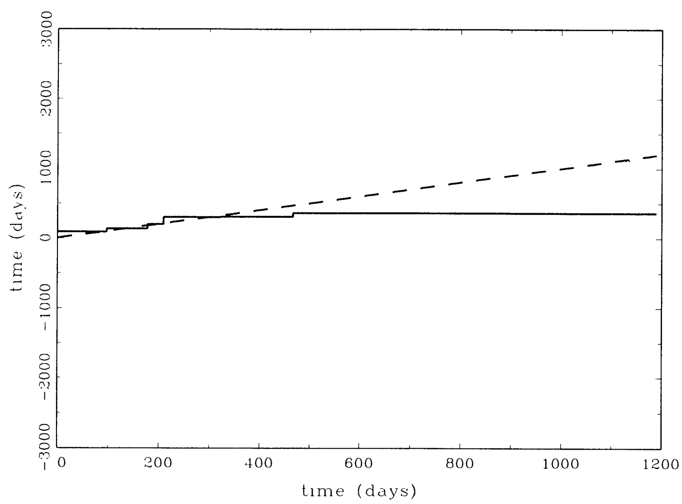

Figure 8. Estimated Q-Q Plot of Empiric Data From Fleming et al. (1980) (Solid Line) and 90% Bootstrap Confidence Band (Dotted Line). The estimated confidence band does not include the 45° dashed line.

proximate 90% confidence. The nonparametric inference for the Q-Q plots for the figures are consistent with the true states of nature even in the presence of nuisance censoring distributions (H_1 and H_2). These Monte Carlo studies confirm the theoretical results of the previous section.

As a real data example, we examine a study performed at the Mayo Clinic of patients with limited Stage II or IIIa ovarian carcinoma. One main goal was to determine whether or not grade of disease was associated with time to progression of disease. The data were taken from a study by Fleming, O'Fallon, O'Brien, and Harrington (1980). For the patients with low-grade or well-differentiated cancer, there were five uncensored and nine censored data points; for the high-grade or undifferentiated cancer patients, there were fifteen uncensored and four censored data points. The estimated distributions are plotted in Figure 7. The Q-Q plot and its associated confidence band are shown in Figure 8. The estimated confidence band does not include the 45° dashed line; hence the distributions in Figure 7 are different.

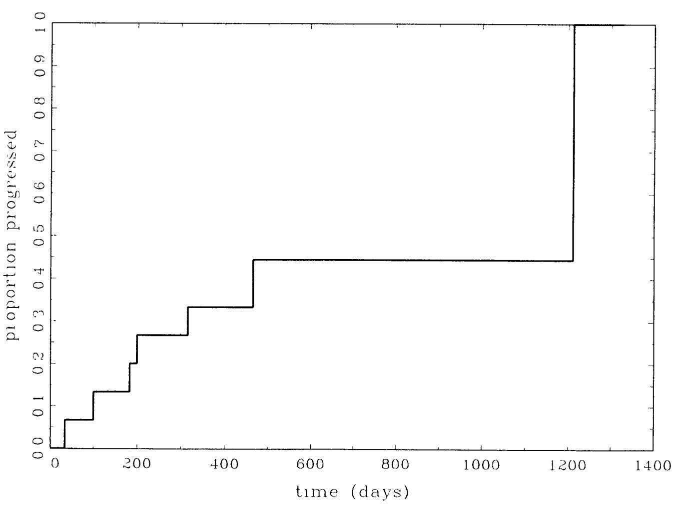

Figure 7. Estimated Distributions of Progressed Proportion of Patients With Low-Grade (Solid Line) and High-Grade (Dotted Line) Ovarian Carcinoma Using the Data of Fleming et al. (1980).

Monte Carlo simulation was conducted to examine the level and power of the Q-Q goodness-of-fit statistic. In Table 2 we present the level simulation. We consider the null hypothesis of equality of three distributions—the exponential (1), Weibull (1, .5), and Wiebull (1, 1.5)—where the distribution function of the Wiebull (λ, α) equals 1 $-\exp[-(\lambda t)^{\alpha}]$. We also vary the amount of censoring and the sample sizes. The sizes of the tests were estimated from 2,500 simulation samples. The results indicated that the nominal level of the test is close to the actual level, even for small sample sizes and heavier censoring levels. This is no doubt due to the fact that we are using bootstrap levels rather than the asymptotic ones.

Table 2. Simulated Level of the Q-Q Goodness-of-Fit Statistic

Survival distribution	Censoring % H_1/H_2	m	n	Level of test .01	.05	.1
Exp(1)	40/40	15	10	.008	.046	.092
		20	15	.008	.048	.095
		25	20	.009	.051	.097
	40/20	15	10	.008	.047	.093
		20	15	.012	.049	.104
		25	20	.011	.051	.102
Wiebull (1, .5)	40/40	15	10	.008	.045	.106
		20	15	.009	.047	.094
		25	20	.012	.051	.097
	40/20	15	10	.009	.046	.092
		20	15	.014	.053	.105
		25	20	.013	.051	.101
Wiebull (1, 1.5)	40/40	15	10	.007	.046	.091
		20	15	.008	.053	.104
		25	20	.013	.052	.103
	40/20	15	10	.007	.046	.093
		20	15	.008	.052	.103
		25	20	.012	.052	.102

Table 3. Power Study: Simulated Powers of the Q-Q, Gehan, and Logrank Tests Under Crossing Hazard Alternatives

			Q-Q		Gehan		Logrank	
Survival distribution	Censoring distributions	m = n	.01	.05	.01	.05	.05	.1
EARLY	U[0, 1]	20	.287	.562	.161	.416	.081	.177
		50	.723	.856	.463	.776	.125	.341
	U[0, 2]	20	.310	.551	.152	.336	.052	.162
		50	.792	.903	.371	.642	.068	.243
MIDDLE	U[0, 1]	20	.321	.612	.147	.394	.094	.205
		50	.784	.902	.302	.580	.231	.426
	U[0, 2]	20	.404	.613	.099	.311	.076	.184
		50	.799	.926	.297	.536	.142	.336
LATE 1	U[0, 1]	20	.323	.534	.021	.049	.096	.215
		50	.764	.901	.025	.051	.236	.531
	U[0, 2]	20	.692	.924	.082	.176	.301	.590
		50	.991	.999	.098	.261	.849	.981
LATE 2	U[0, 1]	20	.256	.346	.031	.108	.071	.246
		50	.376	.610	.056	.142	.221	.431
	U[0, 2]	20	.624	.791	.125	.294	.404	.643
		50	.982	.998	.274	.513	.821	.941

In Table 3 we give power comparisons of the Q-Q goodness-of-fit statistic to Gehan's (1965) extension of the Mann–Whitney test and the logrank test (Mantel 1966). Gehan's test and the logrank test are both known to perform poorly when the hazard rates of distributions under test cross. We simulate the power of these three statistics under several crossing hazard alternatives for variable sample sizes, levels, and censoring distributions. The design of this power study is quite similar to that of Fleming et al. (1980). The alternatives under study are an early hazard difference (EARLY),

$$\begin{array}{lll} \lambda_F = 3 & \lambda_G = .75 & t \in (0, .2) \\ \lambda_F = .75 & \lambda_G = 3 & t \in [.2, .4) \\ \lambda_F = 1 & \lambda_G = 1 & t \in [.4, \infty); \end{array}$$

a middle hazard difference (MIDDLE),

$$\begin{array}{lll} \lambda_F = 2 & \lambda_G = 2 & t \in (0, .1) \\ \lambda_F = 3 & \lambda_G = .75 & t = [.1, .4) \\ \lambda_F = .75 & \lambda_G = 3 & t \in [.4, .7) \\ \lambda_F = 1 & \lambda_G = 1 & t \in [.7, \infty); \end{array}$$

a late hazard difference (LATE 1),

$$\begin{array}{lll} \lambda_F = 1 & \lambda_G = 1 & t \in (0, .8) \\ \lambda_F = 2 & \lambda_G = .2 & t \in [.8, \infty); \end{array}$$

and another later hazard difference where F and G are the Wiebull (2, 2) and Wiebull (.5, .5) (LATE 2). The numbers speak for themselves: The Q-Q goodness-of-fit statistic is much more powerful than the Gehan and logrank competitors. These comparisons look roughly similar to the ones given in table 5 of Fleming et al. (1980).

4. CONCLUDING REMARKS

We studied the nonparametric bootstrap inference for censored data in one- and two-sample cases. In addition, a variety of applications of the bootstrap to various shift functional statistics of censored data can be analogously derived via the similar techniques given in this article. The Lorenz curves method (Lorenz 1905) is a good example. This method has been generalized to measure the concentration and inequality in distributions in many fields, such as economics, politics, and many other social sciences (Csörgö, Csörgö, and Horváth 1986). It also has known asymptotic convergence properties and is adapted to the censored case as well as to the bootstrap resampling. Cumulative hazard functions are another possible class of examples (see Dabrowska, Doksum, and Song 1989). One can study these functionals using the techniques developed previously.

APPENDIX: BOOTSTRAPPING FUNCTIONALS OF THE PL ESTIMATOR AND PROOFS

In this section we give a self-contained development of the necessary weak convergence results for bootstrap problem in the presence of randomly right-censored data. The results given here are more general than what are needed; however, the general results are no more difficult than the specific results.

Henceforth we consider the space of $D[0, \infty)$, the space of cadlag real-valued functions on infinite time scales $[0, \infty)$. These functions can have finite jump discontinuities, such as the edf's or the eqf's. Any stochastic process having all sample paths in $D[0, \infty)$ can be regarded as a random element in $D[0, \infty)$. We will use the supremum norm over all compact finite subintervals, $[0, T]$, to construct a normed vector space $B = \{D[0, \infty), \|\cdot\|\}$, where $\|x\| = \sup_{t \le T} |x(t)|$, for any $x \in D[0, \infty)$ and $0 < T < \infty$. (One can find a further discussion of this norm or the other metrics in Gill 1989, p. 99, Pollard 1984, p. 108, and Shorack and Wellner 1986, p. 26.)

We also use the compact (or Hadamard) differentiability. Assume that ϕ: $B_1 \to B_2$, where B_1 and B_2 are normed vector spaces and ϕ is compactly differentiable at x. That is,

$$[\phi(x_n + t_n h_n) - \phi(x_n)]/t_n \to d\phi(x)h$$

as $t_n \to 0$ in $R = (-\infty, \infty)$, $x_n \to_{\|\cdot\|} x$ and $h_n \to_{\|\cdot\|} h$ in all compact subsets of B_1.

To prove the bootstrap version of a known empirical process result, we start with the following fundamental result, the Skorohod–Dudley–Wichura almost sure representation theorem (see Shorack and Wellner 1986, p. 47).

Lemma A.1. If $X_n \to_d X$ and X takes values in a separable subset of a normed vector space B with a σ-algebra $\mathfrak{B}$, then there exists a sequence of X'_n with $X'_n =_d X_n$, for all n, and $X' =_d X$ such that $X'_n \to_{\text{a.s.}} X'$.

Thus, once we have the weak convergence of a stochastic process, we can construct another distributionally equivalent sequence of random elements that have the a.s. property, provided the conditions of the theorem are satisfied. Using this stronger property, we can investigate the functionals of X_n (or X_n) in any general space. Because the conditions of the Skorohod–Dudley–Wichura almost sure representation theorem hold for our problem settings, we can use this theorem to avoid such defects. Gill (1989) demonstrated various applications of this method and also proved the weak consistency of the bootstrap functionals for the complete data problem. We will restate theorem 5 of Gill (1989) as follows.

Lemma A.2. Assume that (a) F_n is the edf, F_n^* is the bootstrap edf, and $\sqrt{n}[F_n(t) - F(t)] \Rightarrow B^\circ \circ F$; and (b) ϕ: $B_1 \to B_2$ are compactly differentiable at F, ψ: $B_2 \to R$, and ψ are measurable and continuous in a subset of B_2, where $d\phi(F)B^\circ \circ F$ lies in B_2 with probability 1. Then $L^*(\psi\{\sqrt{n}[\phi(F_n^*) - \phi(F_n)]\}) \to_p L(\psi\{d\phi(F)B^\circ \circ F\})$ as $n \to \infty$, where $L(\cdot)$ denotes "the distribution of" and $L^*(\cdot)$ denotes the "bootstrap distribution of." Furthermore, if $\psi\{d\phi(F)B^\circ \circ F\}$ has a continuous distribution, then

$$\sup_t |P^*(\psi\{\sqrt{n}[\phi(F_n^*) - \phi(F_n)]\} \le t) - P(\psi\{d\phi(F)B^\circ \circ F\} \le t)| \to_p 0 \quad \text{as} \quad n \to \infty,$$

where $P(\cdot)$ denotes "the probability of" under L and P^* denotes the "probability of" under L^*.

Assumption (a) in Lemma A.2 is a consequence of Donsker's theorem (Pollard 1984, thm. V11). The result of this lemma reduces to the convergence of F_n^* if both ϕ and ψ are the identity functions. Most of all, the method of proof in the lemma is very illuminating. For our specific example, we will apply it to prove the weak convergence of the bootstrap Q-Q plots for censored data in the two-sample case. The following lemmas are from Gill (1980) and Akritas (1986).

Lemma A.3. Let T_1 be a finite constant such that $[1 - F(T_1)][1 - H_1(T_1)] > 0$. Then $\sqrt{n}[\hat{F}_n(t) - F(t)] \Rightarrow Z_1(t)$ on $D[0, T_1]$ as $n \to \infty$, where Z_1 is a mean 0 Gaussian process with covariance function

$$C_1(s, t) = [1 - F(s)][1 - F(t)] \int_0^{s\wedge t} \frac{dF(u)}{[1 - F(u)]^2[1 - H_1(u)]}, \quad \text{for} \quad s, t \in [0, T_1].$$

Lemma A.4. Let T_1 be defined as in Lemma A.3 and let $\hat{F}_n^*$ denote the bootstrap estimator of $\hat{F}_n$. Then $\sqrt{n}[\hat{F}_n^*(t) - \hat{F}_n(t)] \Rightarrow Z_1^*(t)$ on $D[0, T_1]$ as $n \to \infty$, where Z_1^* is a mean 0 Gaussian process with covariance function $C_1^*(s, t) = C_1(s, t)$, for $s, t \in [0, T_1]$.

We can draw the parallels between Lemmas A.3 and A.4 by noting that each process has the same limiting Gaussian process. Using Lemmas A.3 and A.4, we can generalize the convergence theorem for the uncensored case in Lemma A.2 to the censored case.

Theorem A.5. Under the assumptions on ϕ and ψ as in Lemma A.2,

$$\psi\{\sqrt{n}[\phi(\hat{F}_n^*) - \phi(\hat{F}_n)]\} \Rightarrow \psi\{d\phi(F)Z_1 \circ \phi \circ F\} \quad \text{as} \quad n \to \infty.$$

Proof. By Lemma A.3 and the Skorohod–Dudley–Wichura almost sure representation theorem in Lemma A.1, it is possible to construct $\hat{F}'_n =_d \hat{F}_n$ and $Z'_1 =_d Z_1$ such that $\sqrt{n}[\hat{F}'_n - F] \to Z'_1$ a.s. Let $\hat{F}'^*_n$ be the bootstrap estimator of $\hat{F}'_n$; then by Lemma A.4, $\sqrt{n}[\hat{F}'^*_n - \hat{F}'_n] \Rightarrow Z_1^*[1 - \hat{F}'_n] =_d Z_1^*[1 - \hat{F}_n]$. Again, by Lemma A.1, it is possible to construct $\hat{F}'^{*\prime}_n =_d \hat{F}'^*_n$ and $Z_1^{*\prime} =_d Z_1^*$ such that $\sqrt{n}[\hat{F}'^{*\prime}_n - \hat{F}'_n] \to Z_1^{*\prime}[1 - \hat{F}'_n] =_d Z_1^*[1 - \hat{F}_n]$ a.s. Hence $\sqrt{n}[\hat{F}'^{*\prime}_n - F] = \sqrt{n}[\hat{F}'^{*\prime}_n - \hat{F}'_n] + \sqrt{n}[\hat{F}'_n - F] \to Z_1^{*\prime}[1 - \hat{F}'_n] + Z'_1$ a.s. Hence $\sqrt{n}[\phi(\hat{F}'^{*\prime}_n) - \phi(\hat{F}'_n)] = \sqrt{n}[\phi(\hat{F}'^{*\prime}_n) - \phi(F)] - \sqrt{n}[\phi(\hat{F}'_n) - \phi(F)] \to d\phi(F)\{[Z_1^{*\prime} \circ \phi \circ F][1 - \hat{F}_n] + Z'_1 \circ \phi \circ F\} - d\phi(F)[Z_1^{*\prime} \circ \phi \circ F][1 - \hat{F}_n]$ a.s. By the continuity of ψ, $\psi\{\sqrt{n}[\phi(\hat{F}'^{*\prime}_n) - \phi(\hat{F}'_n)]\} \to \psi\{d\phi(F)[Z_1^{*\prime} \circ \phi \circ F][1 - \hat{F}_n]\}$ a.s. Because $\hat{F}'^{*\prime}_n =_d \hat{F}'^*_n$, $Z_1^{*\prime} =_d Z_1^*$, and $Z_1^*[1 - \hat{F}_n] =_d Z_1$, it follows that $\psi\{\sqrt{n}[\phi(\hat{F}'^*_n) - \phi(\hat{F}'_n)]\} \to_d \psi\{d\phi(F)Z_1 \circ \phi \circ F\}$. Finally, because the left side is a measurable function of $\hat{F}'_n =_d \hat{F}_n$, this completes the proof.

This theorem gives the necessary results to deduce the large-sample bootstrap distribution theory for a wide class of functionals of the PL estimator. These examples include L, M, and R estimators as well as a variety of functionals discussed by Gill (1989).

We now specialize these general results to the problem at hand—the weak convergence of the bootstrapped shift process with censored data. Theorem A.5 deals with $\hat{F}_n$ in the first sample. Because the two samples are totally symmetric, we can have a similar theorem for $\hat{G}_m$ in the second sample.

Corollary A.6. Let T_2 be a finite constant such that $(1 - G(T_2))(1 - H_2(T_2)) > 0$. Under the assumptions on ϕ and ψ as in Lemma A.2,

$$\psi\{\sqrt{m}[\phi(\hat{G}_m^*) - \phi(\hat{G}_m)]\} \Rightarrow \psi\{d\phi(G)Z_2 \circ \phi \circ G\} \quad \text{as} \quad m \to \infty,$$

where Z_2 is a mean 0 Gaussian process with covariance function,

$$C_2(s, t) \equiv [1 - G(s)][1 - G(t)] \int_0^{s\wedge t} \frac{dG(u)}{[1 - G(u)]^2[1 - H_2(u)]}, \quad \text{for} \quad s, t \in [0, T_2].$$

Proof. After the proper identification in the proof of Theorem A.5—$F \leftrightarrow G$, $n \leftrightarrow m$, $H_1 \leftrightarrow H_2$, $Z_1 \leftrightarrow Z_2$—the proof follows.

Choosing the functional $\phi(\cdot) = (\cdot)^{-1}$, the inverse function, we have the following results.

Corollary A.7. Assume that $g(G^{-1}(\cdot))$ is continuous and bounded away from 0 on $[0, T_2]$, where g is the pdf corresponding to G. Let $\hat{G}_m^{*-1}$ denote the bootstrap estimator of $\hat{G}_m^{-1}$. Then

$$\sqrt{m}[\hat{G}_m^{*-1}(t) - \hat{G}_m^{-1}(t)] \Rightarrow Z_2(G^{-1}(t))/g(G^{-1}(t)) \text{ on } D[0, T_2] \quad \text{as} \quad m \to \infty.$$

Proof. Let $\phi(\cdot) = (\cdot)^{-1}$, the inverse function, and $\psi(\cdot) = (\cdot)$, the identity function. Because $d\phi(G) = 1/g \circ G^{-1}$, the theorem follows directly from Corollary A.6.

Putting all of the foregoing results together yields the main result on the weak convergence of the bootstrapped shift process with censored data. The statement of the theorem is given in Theorem 2.1.

Proof of Theorem 2.1

First, we recall the following result of theorem 3.2 of Wells and Tiwari (1989a): Suppose that the assumptions in Lemma A.3 and Corollary A.7 hold. If $n/N \to \theta \in (0, 1)$, then $\sqrt{mn/N}[\hat{\Delta}_{mn}(t) - \Delta(t)] = \sqrt{mn/N}[\hat{G}_m^{-1} \circ \hat{F}_n(t) - G^{-1} \circ F(t)] \Rightarrow Z(t)/g \circ G^{-1} \circ F(t)$ on $D[0, T_1]$, as $m \wedge n \to \infty$, where $N = m + n$ and Z is a mean 0 Gaussian process with covariance kernel given by $C(s, t)$

$= (1-\theta)C_1(s,t) + \theta C_2(G^{-1}\circ F(s), G^{-1}\circ F(t))$ for $s, t \in [0, T_1]$. Thus $Z \equiv \sqrt{1-\theta}Z_1 + \sqrt{\theta}Z_2 \circ G^{-1} \circ F$. Now for the bootstrap analog, we need to study the process, $\sqrt{mn/N}[\hat{G}_m^{*-1}\circ \hat{F}_n^* - \hat{G}_m^{-1}\circ \hat{F}_n]$. This is equal to

$$\equiv \sqrt{\frac{mn}{N}}[\hat{G}_m^{*-1}\circ \hat{F}_n^* - \hat{G}_m^{-1}\circ \hat{F}_n^*] + \sqrt{\frac{mn}{N}}[\hat{G}_m^{-1}\circ \hat{F}_n^* - \hat{G}_m^{-1}\circ \hat{F}_n]$$

$$\equiv \mathbb{A}_N + \mathbb{B}_N, \text{ say,}$$

where

$$\mathbb{A}_N \equiv \sqrt{\frac{mn}{N}}[\hat{G}_m^{*-1}\circ \hat{F}_n^* - \hat{G}_m^{-1}\circ \hat{F}_n^*]$$
$$\Rightarrow \sqrt{\frac{n}{N}}Z_2 \circ G^{-1}\circ \hat{F}_n^*/g\circ G^{-1}\circ \hat{F}_n^*, \quad \text{by Corollary A.7}$$
$$\Rightarrow \sqrt{\theta}Z_2\circ G^{-1}\circ F/g\circ G^{-1}\circ F, \quad \text{by (17.7)–(17.9) of Billingsley (1968).}$$

In addition,

$$\mathbb{B}_N \equiv \sqrt{\frac{mn}{N}}[\hat{G}_m^{-1}\circ \hat{F}_n^* - \hat{G}_m^{-1}\circ \hat{F}_n]$$
$$= \sqrt{\frac{mn}{N}}[\hat{F}_n^* - \hat{F}_n]\{[\hat{G}_m^{-1}\circ \hat{F}_n^* - \hat{G}_m^{-1}\circ \hat{F}_n]/[\hat{F}_n^* - \hat{F}_n]\}$$
$$\Rightarrow \sqrt{\frac{m}{N}}Z_1\{[\hat{G}_m^{-1}\circ \hat{F}_n^* - \hat{G}_m^{-1}\circ \hat{F}_n]/[\hat{F}_n^* - \hat{F}_n]\},$$
$$= \sqrt{1-\theta}Z_1/g\circ G^{-1}\circ \hat{F}_n^{**}.$$

This follows from theorem 3.2 of Wells and Tiwari (1989a) and the mean value theorem, where $|\hat{F}_n^{**}(t) - \hat{F}_n(t)| \le |\hat{F}_n^*(t) - \hat{F}_n(t)|$. Hence, by (17.7)–(17.9) of Billingsley (1968), it follows that $\mathbb{B}_N \Rightarrow \sqrt{1-\theta}Z_1/g\circ G^{-1}\circ F$. Finally, we can use the a.s. representation theorem in Lemma A.1 or the method in the proof of theorem 3.2 of Wells and Tiwari (1989a) to conclude that $\mathbb{A}_N + \mathbb{B}_N \Rightarrow Z/g\circ G^{-1}\circ F$.

Proof of Theorem 2.2

From theorem 3.2 of Wells and Tiwari (1989a), Theorem 2.1, and consistency of Q_m, it follows that

$$\sqrt{\frac{mn}{N}}[\hat{\Delta}_{mn}(t) - \Delta(t)]/[dQ_m\circ \hat{F}_n(t)] \Rightarrow Z(t)$$

and

$$\sqrt{\frac{mn}{N}}[\hat{\Delta}_{mn}^*(t) - \hat{\Delta}_{mn}(t)]/[dQ_m\circ \hat{F}_n(t)] \Rightarrow Z(t).$$

Because $\sup_t|\cdot|$ is a continuous mapping in the Skorohod topology, invoking the continuous mapping theorem yields

$$\sup_t\left|\sqrt{\frac{mn}{N}}[\hat{\Delta}_{mn}^*(t) - \hat{\Delta}_{mn}(t)]\Big/[dQ_m\circ \hat{F}_n(t)]\right| \to_d \sup_t|Z(t)|$$

and

$$\sup_t\left|\sqrt{\frac{mn}{N}}[\hat{\Delta}_{mn}(t) - \Delta(t)]\Big/[dQ_m\circ \hat{F}_n(t)]\right| \to_d \sup_t|Z(t)|.$$

Hence $c_{mn}(\Delta)$ converges to the $(1-\alpha)$th point of the distribution of $\sup_t|Z(t)|$. Similarly,

$$\Pr\left\{\sup_t\left(\left|\sqrt{\frac{mn}{N}}[\hat{\Delta}_{mn}(t) - \Delta(t)]/[dQ_m\circ \hat{F}_n(t)]\right|\right) \le c_{mn}(\Delta)\right\} \to 1-\alpha.$$

Thus

$$\Pr\left\{\sup_t\left(\left|\sqrt{\frac{mn}{N}}[\hat{\Delta}_{mn}(t) - \Delta(t)]/[dQ_m\circ \hat{F}_n(t)]\right|\right) \le c_{mn}(\Delta)\right\}$$
$$= \Pr\left\{\left|\sqrt{\frac{mn}{N}}[\hat{\Delta}_{mn}(t) - \Delta(t)]/[dQ_m\circ \hat{F}_n(t)]\right| \le c_{mn}(\Delta), \forall t\in[0,T_1]\right\}$$
$$= \Pr\left\{\hat{\Delta}_{mn}(t) - \sqrt{\frac{N}{mn}}c_{mn}(\Delta)[d\dot{Q}_m\circ \hat{F}_n(t)] \le \Delta(t) \le \hat{\Delta}_{mn}(t) + \sqrt{\frac{N}{mn}}c_{mn}(\Delta)[dQ_m\circ \hat{F}_n(t)], \forall t\in[0,T_1]\right\} \to 1-\alpha.$$

[Received August 1992. Revised September 1993.]

REFERENCES

Akritas, M. (1986), "Bootstrapping the Kaplan-Meier Estimator," *Journal of the American Statistical Association,* 81, 1032–1039.

Bickel, P. J., and Freedman, D. A. (1981), "Some Asymptotic Theory for the Bootstrap," *The Annals of Statistics,* 9, 1196–1217.

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

Csörgö, M., Csörgö, S., and Horváth, L. (1986), *An Asymptotic Theory for Empirical Reliability and Concentration Processes,* New York: Springer-Verlag.

Dabrowska, D. M., Doksum, K. A., and Song, J. K. (1989), "Graphical Comparison of Cumulative Hazards for Two Populations," *Biometrika,* 76, 763–773.

Doksum, K. (1974), "Empirical Probability Plots and Statistical Inference for Nonlinear Models in the Two-Sample Case," *The Annals of Statistics,* 2, 267–434.

Doksum, K. A., and Sievers, G. L. (1976), "Plotting With Confidence: Graphical Comparisons of Two Populations," *Biometrika,* 63, 421–434.

Efron, B. (1981), "Censored Data and the Bootstrap," *Journal of the American Statistical Association,* 76, 312–319.

Fleming, T. R., O'Fallon, J. R., O'Brien, P. C., and Harrington, D. P. (1980), "Modified Kolmogorov-Smirnov Test Procedures With Application to Arbitrarily Right-Censored Data," *Biometrics,* 36, 607–625.

Gehan, E. A. (1965), "A Generalized Wilcoxian Test for Comparing Arbitrarily Singly Censored Samples," *Biometrika,* 52, 203–223.

Gill, R. D. (1980), *Censoring and Stochastic Integrals,* Mathematical Centre Tract Vol. 124, Amsterdam: Mathematisch Centrum.

——— (1989), "Non- and Semi-Parametric Maximum Likelihood Estimators and the Von Mises Method (Part 1)," *Scandinavian Journal of Statistics,* 16, 97–128.

Hawkins, D. L., and Kochar, S. C. (1991), "Inference for the Crossing of Two Continuous CDF's," *The Annals of Statistics,* 19, 1626–1638.

Hodges, J. L., and Lehmann, E. L. (1963), "Estimates of Location Based on Rank Tests," *Annals of Mathematical Statistics,* 34, 598–611.

Hollander, M., and Korwar, R. M. (1982), "Nonparametric Bayesian Estimation of the Horizontal Distance Between Two Populations," in *Nonparametric Statistical Inference I,* New York: North-Holland.

Kaplan, E. L., and Meier, P. (1958), "Nonparametric Estimation From Incomplete Observations," *Journal of the American Statistical Association,* 53, 457–481.

Lio, Y. L., and Padgett, W. J. (1992), "Asymptotically Optimal Bandwidths for a Smooth Nonparametric Quantile Estimator Under Censoring," *Journal of Nonparametric Statistics,* 1, 219–229.

Lorenz, M. O. (1905), "Methods of Measuring the Concentration of Wealth," *Publication of the American Statistical Association,* 9, 209–219.

Mantel, N. (1966), "Evaluation of Survival Data and Two New Rank Order Statistics Arising in Its Consideration," *Cancer Chemotherapy Reports,* 50, 163–170.

Meng, X. L., Bassiakos, Y., and Lo, S. H. (1991), "Large-Sample Properties for a General Estimator of the Treatment Effect in the Two-Sample Problem With Right Censoring," *The Annals of Statistics,* 19, 1786–1812.

Miller, R. G. (1981), *Survival Analysis,* New York: John Wiley.

Padgett, W. J. (1986), "A Kernel-Type Estimator of a Quantile Function From Right-Censored Data," *Journal of the American Statistical Association,* 81, 215–222.

Padgett, W. J., and Thombs, L. A. (1986), "Smooth Nonparametric Quantile Estimation Under Censoring: Simulations and Bootstrap Methods," *Communications in Statistics—Simulations,* 15, 1003–1025.

Padgett, W. J., and Wei, L. J. (1980), "Estimation of the Scale Parameter in the Two-Sample Problem with Arbitrary Censoring," *Biometrika,* 69, 252–256.

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

Reid, N. (1981), "Estimating the Median Survival Time," *Biometrika,* 68, 601–608.

Sheather, S. J., and Marron, J. S. (1990), "Kernel Quantile Estimators," *Journal of the American Statistical Association,* 85, 410–416.

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

Silverman, B. W. (1986), *Density Estimation,* New York: Chapman and Hall.

Wang, J. L., and Hettmansperger, T. P. (1990), "Two-Sample Inference for Median Survival Times Based on One-Sample Procedures for Censored Data," *Journal of the American Statistical Association,* 85, 529–536.

Wei, L. J., and Gail, M. (1983), "Nonparametric Estimation for a Scale-Change with Censored Observations," *Journal of the American Statistical Association,* 78, 382–388.

Wells, M. T., and Tiwari, R. C. (1989a), "Bayesian Quantile Plots and Statistical Inference for Nonlinear Models in the Two-Sample Case With Incomplete Data," *Communications in Statistics—Theory and Methods,* 18, 2955–2964.

——— (1989b), "On The Bayesian Bootstrap With Censored Data," preprint, Cornell University, Dept. of Economic and Social Statistics.

Wilk, M. B., and Gnanadesikan, R. (1968), "Probability Plotting Methods for the Analysis of Data," *Biometrika,* 55, 1–17.