Statistica Sinica **15**(2005), 969-980

BAYESIAN BOOTSTRAP ANALYSIS OF DOUBLY CENSORED DATA USING GIBBS SAMPLER

Yongdai Kim, Jaeyong Lee and Jinseog Kim

Seoul National University

Abstract: The Bayesian bootstrap for doubly censored data is constructed from the empirical likelihood perspective, and a Gibbs sampler algorithm is proposed for evaluating the Bayesian bootstrap posterior. The proposed Bayesian bootstrap posterior is shown to be the limit of the nonparametric posteriors with Dirichlet process priors as the prior information vanishes, and to be equivalent to the weighted bootstrap on the observables. A small simulation study shows that the proposed Bayesian bootstrap estimator compares favorably with the nonparametric maximum likelihood estimator; furthermore its asymptotic properties are studied.

Key words and phrases: Bayesian bootstrap, doubly censored data, empirical likelihood, survival model.

1. Introduction

In survival analysis, data are subject to censoring. The most common type of censoring is right censoring, in which the survival time is larger than the observed right censoring time. In some cases, however, data are subject to left, as well as, right censoring. When left censoring occurs, the only information available to a statistician is that the survival time is less than or equal to the observed left censoring time. Data with both right and left censored observations are known as *doubly censored data*. Examples of doubly censored data have been given by Gehan (1965), Mantel (1967), Peto (1973) and Turnbull (1974), among others.

Analysis of doubly censored data has been studied by many statisticians. Turnbull (1974) proposed the self-consistent estimator (SCE), the concept first introduced by Efron (1967), and showed that the SCE is the nonparametric maximum likelihood estimator (NPMLE) for grouped survival data. Tsai and Crowley (1985) studied the theoretical properties of the SCE and applied their results to doubly censored data. The asymptotic properties of the SCE were studied more rigorously by Chang and Yang (1987), Chang (1990) and Gu and Zhang (1993). Gu and Zhang (1993) showed that SCE and the NPMLE may not coincide, while Mykland and Ren (1996) provided a necessary and sufficient condition for an SCE to be the NPMLE, and argued by simulation that finite

sample performance of the NPMLE is better than that of an SCE. Wellner and Zhan (1997) proposed an efficient algorithm for computing the NPMLE.

This paper is concerned with Bayesian bootstrap (BB) analysis of doubly censored data. Rubin (1981) first introduced the Bayesian bootstrap as a Bayesian alternative to the bootstrap (Efron (1979)). It was further extended to the finite population model (Lo (1988)), right censored data (Hjort (1991), Lo (1993)), and proportional hazard model (Kim and Lee (2003)). There are three views for the BB — the weighted bootstrap, the limit of the full Bayesian posterior, and the Bayesian version of the empirical likelihood approach, all of which are summarized in Section 2.

In this paper, we devise a BB procedure for doubly censored data, taking the Bayesian version of the empirical likelihood approach as our reference point. This choice has two important implications. First, the proposed BB procedure is the same as the full Bayesian procedure of time discrete doubly censored data proposed by Kuo and Smith (1992) and based on a certain noninformative prior. Moreover, we show that the proposed BB posterior is the limit of the full Bayesian posterior of the time continuous model. Therefore, the BB analysis may be considered noninformative Bayesian inference. Also, we can use the efficient Gibbs sampler algorithm proposed by Kuo and Smith (1992) for evaluating BB posteriors. Second, the proposed BB procedure is equivalent to the weighted bootstrap of Wellner and Zhang (1996), which is the first of its kind for doubly censored data. In addition, the proposed BB algorithm is simpler and faster than the standard bootstrap algorithm.

This paper is organized as follows. Section 2 summarizes the three views on the BB. In Section 3, a BB procedure for doubly censored data is proposed. The components of the BB — the BB likelihood, prior, posterior — and a computational algorithm using Markov Chain Monte Carlo (MCMC) are discussed. Section 4 explores the connection of the proposed BB to the weighted bootstrap and the full Bayesian approach. Section 5 presents simulation results and a data set is analyzed using the proposed BB. In Section 6, the asymptotic properties of the proposed BB posterior are studied. Discussion follows in Section 7.

2. Three Views for Bayesian Bootstrap

In order to provide a background for the proposed BB procedure for doubly censored data, this section reviews the three views of the BB. For the discussion in this section, we consider the analysis of uncensored data.

The first view of the BB is that it extends the Efron's bootstrap (Efron (1979)). Let $\mathcal{X} = (X_1, \ldots, X_n)$ be a random sample from an unknown distribu-

tion F, and suppose a functional of F, $T(F)$, is of interest. In this situation, a typical bootstrap procedure consists of drawing many bootstrap samples $\mathcal{X}_1^*, \ldots, \mathcal{X}_B^*$, where each bootstrap sample is a random sample from the empirical distribution of the original sample $\mathcal{X}$, and inference about $T(F)$ is based on the $T(F_i^*)$'s, where F_i^* is the empirical distribution of $\mathcal{X}_i^*$. Noting that $F_i^* \stackrel{d}{=} \sum_{i=1}^n w_i \delta_{X_i}$, where $nw = n(w_1, \ldots, w_n) \sim Multinomial(n, 1/n, \ldots, 1/n)$, Rubin proposed a smoother alternative, $Dirichlet(1, \ldots, 1)$, for the distribution of w. For this reason, the BB is viewed as a weighted bootstrap. Theoretical properties of the BB and weighted bootstrap have been studied by many authors including Lo (1987), Weng (1989), Mason and Newton (1992), Praestgaard and Wellner (1993), Gasparini (1995), James (1997) and Choudhuri (1998).

The second view is that the BB posterior is the limit of the full Bayesian posterior as the amount of the prior information vanishes, and that the BB constitutes a noninformative (or default) Bayesian analysis of nonparametric problems. More formally, if the prior on F is the Dirichlet process with parameter α, a nonnull finite measure, then the posterior is the Dirichlet process with parameter $\alpha + \sum_{i=1}^n \delta_{X_i}$ (Ferguson (1973)). As the total mass of α (or the prior sample size) goes to 0, the posterior converges to Rubin's BB posterior. As Gasparini (1995) noted, this provides the basis for using the BB as a default nonparametric Bayesian analysis.

The third view is that the BB posterior is obtained from

$$\text{BB posterior} \propto \text{empirical likelihood} \times \text{prior}. \tag{1}$$

Assume that there are no ties in $\mathcal{X}$. Furthermore, suppose the true distribution belongs to $\mathcal{F}_n = \{\sum_{i=1}^n w_i \delta_{X_i} : \sum_{i=1}^n w_i = 1, w_i \geq 0, i = 1, \ldots, n\}$. Then the model is effectively parametric, and Bayesian analysis can be carried out. Since there is one observation at each X_i, the likelihood is given by

$$L(F) = \prod_{i=1}^n w_i. \tag{2}$$

Adopting a noninformative prior, $\prod_{i=1}^n w_i^{-1}$, we get, by the usual Bayesian computation, Rubin's BB posterior. This description of BB was noted by many authors (Owen (1990), Choudhuri (1998), Lazar (2000) and Kim and Lee (2003)). Note that the likelihood used in this derivation is the empirical likelihood of Owen (1990). This is the main idea we use to derive the BB posterior for doubly censored data. That is, the model is reduced to a data-dependent parametric model, $\mathcal{F}_n$, and the usual Bayesian analysis is performed with an appropriate prior. Here we call the likelihood on $\mathcal{F}_n$ the "BB likelihood". In what follows, we use the terms "empirical likelihood" and "BB likelihood" interchangeably.

3. BB for Doubly Censored Data

3.1. Model

Let X_i, $i = 1, \ldots, n$, be independent and identically distributed (i.i.d.) positive random variables with a common survival function S_X. Independent of the X_i's, the $Y_i \geq Z_i$ are i.i.d. pairs of right and left censoring times with possibly defective marginal distribution functions S_Z and S_Y. Under the doubly censoring mechanism, we observe only pairs of (W_i, δ_i) :

$$(W_i, \delta_i) = \begin{cases} (X_i, 1), & \text{if } Z_i < X_i \leq Y_i, \\ (Y_i, 2), & \text{if } X_i > Y_i, \\ (Z_i, 3), & \text{if } X_i \leq Z_i. \end{cases} \tag{3}$$

The censoring indicator δ_i takes on the values 1, 2 and 3, if the survival time X_i is observed, right censored, and left censored, respectively. This section develops a BB procedure for inferring S_X based on the doubly censored observations $(W_1, \delta_1), \ldots, (W_n, \delta_n)$.

3.2. BB Likelihood

Let $\mathcal{F}$ be the space of all distribution functions on $\mathbb{R}^+ = [0, \infty)$. The first step in constructing the BB likelihood is to reduce $\mathcal{F}$ to a parametric model $\mathcal{F}_n$. A choice of $\mathcal{F}_n$ is a family of distribution functions F of the form $F(t) = \sum_{i=1}^{l} p_i I(u_i \leq t)$, with some predefined data-dependent points, $\mathbf{U} = \{u_1 < \cdots < u_l\}$. For doubly censored data, we propose $\mathbf{U}$ as follows. Let $V_1 < \ldots < V_m$ be distinct points of $\{W_1, \ldots, W_n\}$ and define, for $k = 1, \ldots, m$, $\alpha_k^* = \sum_{i=1}^{n} I(W_i = V_k, \delta_i = 1), \beta_k^* = \sum_{i=1}^{n} I(W_i = V_k, \delta_i = 2)$ and $\gamma_k^* = \sum_{i=1}^{n} I(W_i = V_k, \delta_i = 3)$. Let $\beta_0^* = 0$.

- Case 1. If $\alpha_i^* > 0$, then $V_i \in \mathbf{U}$.
- Case 2. If $\alpha_i^* = 0, \gamma_i^* > 0$, then
  1. if $\alpha_j^* = 0$ and $\gamma_j^* = 0$ for $j < i$, then $0 \in \mathbf{U}$, or
  2. if $\beta_{i-1}^* > 0$ and $\beta_j^* = 0$ and $\alpha_j^* = 0$ for all $i \leq j \leq m$, then $(V_{i-1}+V_i)/2 \in \mathbf{U}$.
- Case 3. If $\alpha_m^* = 0$ and $\beta_m^* > 0$, then $\infty \in \mathbf{U}$.

Remark. If $\alpha_1^* > 0$ and $\alpha_m^* > 0$, then $\mathbf{U}$ consists of only V_is with $\alpha_i^* > 0$. That is, $\mathbf{U}$ contains all distinct uncensored observations. Cases 2 and 3 deal with situations in which the smallest or largest observation is censored, respectively.

The main motivation behind our choice of $\mathbf{U}$ is that the support of the limit of the full Bayesian posteriors with Dirichlet process priors is $\mathbf{U}$, provided $\alpha_1^* > 0$

and $\alpha_m^* > 0$. In addition, the maximum likelihood estimator of S_X on $\mathcal{F}_n$ is an SCE (Mykland and Ren (1996)).

With $\mathbf{U}$, the parameters to be estimated are $\mathbf{p} = (p_1, \ldots, p_l)$, and the likelihood of $\mathbf{p}$ is

$$L(\mathbf{p}) = \prod_{k=1}^{l} p_k^{\alpha_k} \left(1 - \sum_{i=1}^{k} p_i\right)^{\beta_k} \left(\sum_{i=1}^{k} p_i\right)^{\gamma_k}, \tag{4}$$

where $\alpha_k = \sum_{i=1}^n I(W_i = u_k, \delta_i = 1)$, $\beta_k = \sum_{i=1}^n I(u_{k-1} \le W_i < u_k, \delta_i = 2)$ and $\gamma_k = \sum_{i=1}^n I(u_{k-1} < W_i \le u_k, \delta_i = 3)$, for $k = 1, \ldots, l$.

3.3. Prior

For the prior on $\mathbf{p}$, we propose the improper

$$\pi(\mathbf{p}) = \prod_{k=1}^{l} \frac{1}{p_k}. \tag{5}$$

A motivation for (5) is that the BB posterior constructed via the product of the BB likelihood (4) and the prior (5) can be obtained as a limit of the full Bayesian posterior as the prior information vanishes. Consider the full Bayesian analysis with the Dirichlet process with base measure $\alpha F_0(t)$, where $F_0(t)$ is the prior mean and α governs the amount of prior information. If we let α go to 0 (the amount of prior information goes to 0), then the full Bayesian posterior converges to the BB posterior. See Section 4.1 for details.

Remark. Without censored observations, the BB posterior distribution derived from (4) and (5) is the same as Rubin's BB posterior (1981). Furthermore, the proposed BB posterior is equivalent to Lo's (1993) BB for right censored data.

3.4. MCMC

The BB posterior of $\mathbf{p}$ is

$$\pi(\mathbf{p}|\mathrm{D}) \propto \prod_{k=1}^{l} p_k^{\alpha_k} \left(1 - \sum_{i=1}^{k} p_i\right)^{\beta_k} \left(\sum_{i=1}^{k} p_i\right)^{\gamma_k} \prod_{k=1}^{l} \frac{1}{p_k}, \tag{6}$$

where $\mathrm{D} = \{(W_1, \delta_1), \ldots, (W_n, \delta_n)\}$. This posterior is by no means a well known distribution. However, Bayesian computation using the Gibbs sampler algorithm proposed by Kuo and Smith (1992) poses no difficulty in this case.

Let $\mathbf{X}_n^* = (X_1^*, \ldots, X_n^*)$ be the survival times, some of which are observed (when $\delta_i = 1$) while others are not (when $\delta_i = 2$ or 3). The main idea of the Gibbs sampler is to generate $\mathbf{X}_n^*$ from $\mathcal{L}(\mathbf{X}_n^*|\mathbf{p}, \mathrm{D})$, and then $\mathbf{p}$ from $\mathcal{L}(\mathbf{p}|\mathbf{X}_n^*, \mathrm{D})$, successively.

Sampling from $\mathcal{L}(\mathbf{X}_n^*|\mathbf{p}, \mathrm{D})$ can be easily done as follows. Once $\mathbf{p}$ and D are given, it suffices to generate X_i^* from $\mathcal{L}(X_i^*|\mathbf{p}, W_i, \delta_i)$, for $i = 1, \ldots, n$. When $\delta_i = 1$, X_i^* should be W_i with probability 1. When $\delta_i = 2$, we generate X_i^* from the distribution $F(t|t > W_i)$, where $F(t) = \sum_{i=1}^{l} p_i I(u_i \le t)$. Since F is finitely supported, a random number from $F(t|t > W_i)$ can be generated easily. Similarly, when $\delta_i = 3$, X_i^* is generated from the distribution $F(t|t \le W_i)$. Once $\mathbf{X}_n^*$ and D are given, $\mathcal{L}(\mathbf{p}|\mathbf{X}_n^*, \mathrm{D})$ is the same as Rubin's BB posterior: $\mathcal{L}(\mathbf{p}|\mathbf{X}_n^*, \mathrm{D}) \sim Dirichlet(r_1, \ldots, r_p)$, where $r_i = \sum_{k=1}^{n} I(X_k^* = u_i)$.

4. Some Properties of the Proposed BB

4.1. BB as the limit of full Bayesian posteriors

The proposed BB posterior is the same as the full Bayesian posterior of the time discrete model considered by Kuo and Smith (1992), if F is assumed in advance to have mass on $\mathbf{U}$. What seems a difference is that the BB uses the noninformative prior on $\mathbf{p}$ while Kuo and Smith (1992) use a proper Dirichlet distribution prior. However, since the prior (5) can be obtained as the limit of the Dirichlet distribution, the BB procedure can be considered a noninformative Bayesian analysis of Kuo and Smith's time discrete model. The explanation of the BB via the empirical likelihood approach also justifies using Kuo and Smith's model even for time continuous observations with appropriately chosen support points.

This analogy persists for time continuous models. That is, the proposed BB can be obtained as a limit of the full Bayesian posterior of the time continuous model. Suppose a priori that F is a Dirichlet process with mean F_0 and precision parameter $\alpha > 0$, $F \sim DP(\alpha F_0)$. Let $\mathcal{L}(F|\mathrm{D}, \alpha, F_0)$ be the posterior distribution of F when the prior distribution of F is $DP(\alpha F_0)$. Then,

$$\mathcal{L}(F|\mathrm{D}, \alpha, F_0) \xrightarrow{d} \mathcal{L}^{(B)}(F|\mathrm{D}) \tag{7}$$

on $D[0, \infty)$, provided $\alpha_1 > 0$ and $\alpha_l > 0$, where $D[0, \infty)$ is the space of right continuous functions with left limits defined on $[0, \infty)$ equipped with the uniform topology. The proof is provided in the Appendix.

Remark. If $\alpha_{(1)} = 0$ and $\gamma_{(1)} > 0$, the full Bayesian posteriors of F on $[0, W_{(1)}]$ do not vanish as $\alpha \to 0$. Hence, the limiting posterior differs from the BB posterior on $[0, W_{(1)}]$. As n gets larger, however, $W_{(1)}$ converges to 0 if $\mathbb{R}^+$ is the support of the distribution of W_i; therefore, the discrepancy between the limiting posterior and BB posterior becomes minimal for large n. Similar remarks apply when $\alpha_{(m)} = 0$ and $\beta_{(m)} > 0$.

4.2. BB as a weighted bootstrap

Let $Q(t) = \Pr(W > t)$ and $Q_j(t) = \Pr(W > t, \delta = j)$, for $j = 1, 2, 3$. Let $Q_j^{(n)} = \sum_{i=1}^n I(W_i > t, \delta = j)/n$ for $j = 1, 2, 3$, and let $Q^{(n)}(t) = Q_1^{(n)} + Q_2^{(n)} + Q_3^{(n)}$. Put $S_X(t) = \Pr(X > t), S_Y(t) = \Pr(Y > t)$ and $S_Z(t) = \Pr(Z > t)$. Then we have

$$S_X(t) = Q(t) - \int_{u \le t} \frac{S_X(t)}{S_X(u)} dQ_2(u) + \int_{t < u} \frac{1 - S_X(t)}{1 - S_X(u)} dQ_3(u), \tag{8}$$

$$S_Y(t) = 1 + \int_0^t \frac{dQ_2(u)}{S_X(u)}, \tag{9}$$

$$S_Z(t) = -\int_t^\infty \frac{dQ_3(u)}{1 - S_X(u)}. \tag{10}$$

See Chang and Yang (1987) for details. In fact, the SCE $S_X^{(n)}$ of S_X is defined as the solution of the empirical version of (8), where the empirical version means population quantities are replaced by their empirical estimators; for example, $Q_1(t)$ is replaced by $Q_1^{(n)}(t)$. The estimators $S_Y^{(n)}$ and $S_Z^{(n)}$ of S_Y and S_Z are given by the empirical version of (9) and (10).

Using (8), we can devise a weighted bootstrap for doubly censored data. The standard weighted bootstrap procedure for the observables is to assign the mass r_i to (W_i, δ_i), where $(r_1, \ldots, r_n) \sim Dirichelt(1, 1, \ldots, 1)$. Once the weighted bootstrap for the observables is defined, the weighted bootstrap distributions of Q, Q_j, for $j = 1, 2, 3$, are defined as $Q^{(w)}(t) = \sum_{i=1}^n r_i I(W_i \le t)$ and $Q_j^{(w)}(t) = \sum_{i=1}^n r_i I(W_i \le t, \delta_i = j)$, for $j = 1, 2, 3$. Then we can define the weighted bootstrap distributions of S_X, S_Y and S_Z, denoted by $S_X^{(w)}$, $S_Y^{(w)}$ and $S_Z^{(w)}$, as the solutions of (8), (9) and (10) with Q and Q_i replaced by $Q^{(w)}$ and $Q_j^{(w)}$ for $j = 1, 2, 3$. This weighted bootstrap procedure for doubly censored data was studied by Wellner and Zhan (1996).

In general, solutions of (8), (9) and (10) are not unique. However, if we assume that the support of $S_X^{(w)}$ is confined to $\mathbf{U}$, there exists a unique solution. See Mykland and Ren (1996). Now, since $S_X^{(w)}, S_Y^{(w)}$ and $S_Z^{(w)}$ are functions of $(r_1, \ldots, r_n)$, we can show by the standard change of variable technique, with a carefully calculated Jacobian matrix, that the law of $S_X^{(w)}$, whose support is $\mathbf{U}$, is the same as the BB posterior of S_X, provided

$$\{W_i : \delta_i = 1\} \cap \{W_i : \delta_i \ne 1\} = \emptyset. \tag{11}$$

The detailed proof of this assertion is available from the authors upon request. This equivalence allows frequentists to use the proposed BB for approximating the sampling distribution of $S_X^{(n)}$, since the weighted bootstrap is thought to be an alternative to Efron's bootstrap.

5. Illustrations

In this section we illustrate the application of the proposed BB to simulation models, as well as to a data set. The BB posteriors are calculated via the MCMC algorithm presented in Section 3.4 with 10,000 iterations, of which the first 1,000 iterations are discarded as burn-in.

First, we perform a small simulation to evaluate the true coverage probability of the probability interval based on the BB posterior. Survival time X is generated from $Exp(100)$ — the exponential distribution with mean 100. The left and right censoring variables (Z, Y) are generated by $(Z, Y) = (Z, Z + W)$, where $Z \sim Exp(10)$ and $W \sim Exp(140)$, Z and W independent. Under this model, the censoring probability is about 48%, of which 38% is due to right censoring and 10% is due to left censoring. Table 1 presents the coverage probabilities of the 95% probability intervals of the 25% and 50% (median) quantiles based on the proposed BB. The results show that the BB posterior can be used as an alternative frequentist's method for doubly censored data.

Table 1. The coverage probabilities of the 95% equal tail probability intervals of the 25% and 50% quantiles. These coverage probabilities are calculated based on 10000 independent samples.

Sample size n	20	50	100
25% quantile	94.1	94.5	95
50% quantile	93.2	93.8	94.9

Next, we compare the performance of the Bayes estimator and NPMLE of F. We use the following simulation model: $X \sim Exp(1)$, $Y \sim Exp(2)$ and $Z = (Y - 0.5)I(Y \geq 0.5)$. Table 2 presents the average distances (sup-norm) of the estimators to the true distribution. The results for the NPMLE are quoted from Mykland and Ren (1996). The Bayesian bootstrap estimator seems to provide results that are closer to the true model than those of the NPMLE.

Table 2. The average distance (sup-norm) of the estimators from the true model.

$n = 25$	$\|S - \text{NPMLE}\|$	0.3276 (0.0959)
	$\|S - \text{Bayes}\|$	0.3133 (0.0884)
$n = 50$	$\|S - \text{NPMLE}\|$	0.2427 (0.0648)
	$\|S - \text{Bayes}\|$	0.2343 (0.0594)

Finally, we analyze a data set consisting of times to the first use of marijuana collected from 191 California high school boys. The students were asked " When did you first use marijuana?" For this question, there are three possible answers – " I started at age 14." (uncensored) " I never used it." (right censoring) and " I have used it but don't remember" (left censoring). The full data set is

available in Turnbull and Weiss (1978). Figure 1 presents the Bayesian bootstrap estimate of the survival function, along with the 5% and 95% quantiles of the Bayesian bootstrap posterior. The survival curves in the figure are drawn by linearly interpolating the estimates of survival probabilities at $t \in \mathbf{U}$.

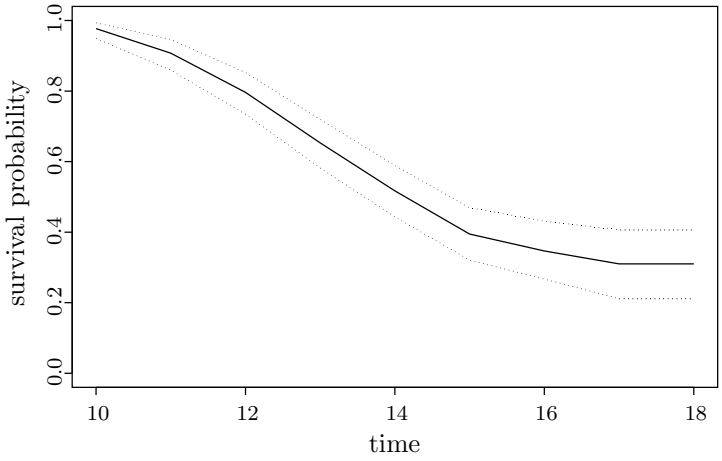

Figure 1. Posterior mean and the 5%, 95% quantiles of survival probabilities

6. Large Sample Properties

In this section, we prove that the BB posterior of S_X centered at the SCE $S_X^{(n)}$ is asymptotically equivalent to the distribution of the SCE $S_X^{(n)}$ centered at the true survival function S_X. We assume the following regularity conditions:

A1. $\Pr(Z \leq Y) = 1$;
A2. $S_Y(t) - S_Z(t) > 0$ on $(0, \infty)$;
A3. S_X, S_Y and S_Z are continuous functions of t and $0 < S_X(t) < 1$ for $t > 0$;
A4. $S_X(0) = S_Y(0) = 1$, $S_X(\infty) = S_Y(\infty) = S_Z(\infty) = 0$;
A5. there exists δ and T, $0 < \delta < T < \infty$ such that $S_Z(t)$ is constant with a value less than 1 on $[0, \delta]$ and $S_Z(T) = 0$, i.e., $\Pr(Z = 0) > 0, \Pr(Z \in (0, \delta)) = 0$ and $\Pr(Z \leq T) = 1$.

Remark. Chang (1990) used the above regularity conditions to prove the weak convergence of a self-consistent estimator.

Theorem 1. *$\mathcal{L}^{(B)}(\sqrt{n}(S_X - S_X^{(n)})|\mathrm{D})$ and $\mathcal{L}(\sqrt{n}(S_X^{(n)} - S_X))$ converge in distribution to the same Gaussian process on $D[0,T]$, where $D[0,T]$ is the space of right continuous functions with left limits defined on $[0,T]$ equipped with the uniform topology.*

Proof. Let $q^{(n)} = \sqrt{n}(Q_1^{(n)} - Q_1, Q_2^{(n)} - Q_2, Q_3^{(n)} - Q_3)$ and $q^{(w)} = \sqrt{n}(Q_1^{(w)} - Q_1^{(n)}, Q_2^{(w)} - Q_2^{(n)}, Q_3^{(w)} - Q_3^{(n)})$. Then it is not hard to prove that $\mathcal{L}^{(w)}(q^{(w)}|\mathrm{D})$

and $\mathcal{L}(q^{(n)})$ converge weakly to the same Gaussian process on $D[0,T]^3$ (see Lo (1987)).

Let $r^{(n)} = \sqrt{n}(S_X^{(n)} - S_X, S_Y^{(n)} - S_Y, S_Z^{(n)} - S_Z)$ and $r^{(w)} = \sqrt{n}(S_X^{(w)} - S_X^{(n)}, S_Y^{(w)} - S_Y^{(n)}, S_Z^{(w)} - S_Z^{(n)})$. Let Φ be the mapping from $D[0,T]^3$ to $D[0,T]^3$ such that $\Phi(q^{(n)}) = r^{(n)}$ defined by the system of equations (8), (9) and (10). Since Φ is an asymptotically linear continuous mapping (Chang, (1990)), $\mathcal{L}^{(w)}(r^{(w)}|\mathrm{D})$ and $\mathcal{L}(r^{(n)})$ have the same limiting distribution, which is also a Gaussian process on $D[0,T]^3$.

Finally, by **A3**, (11) holds with probability 1, and hence the law of $S_X^{(w)}$ is the same as the BB posterior.

7. Discussion

The computation of the BB posterior is much easier than that of the standard bootstrap method with the NPMLE. Wellner and Zhan (1997) proposed an algorithm for finding the NPMLE, which is a combination of EM and ICM (iterative convex minorant) algorithms. However, this does not provide a variance estimator. In practice, to estimate the variance, a bootstrap with the Wellner and Zhan algorithm is used. This is executed by first generating a bootstrap sample $(W_1^*, \delta^*), \ldots, (W_n^*, \delta_n^*)$ and finding the MLE based on the bootstrap sample. This procedure is repeated many times to have many bootstrap estimates of F, and the variance of the NPMLE is estimated based on the bootstrap estimates of F. Computation is extremely demanding. For comparison of the proposed BB and the bootstrap with NPMLE, we repeated the simulation of Wellner and Zhan (1997). With a sample of size 5,000 and censoring probability 0.6, the bootstrap with NPMLE needed 119 hours, or about 5 days, for 1,000 bootstrap samples − this is quoted from Wellner and Zhan (1997). In contrast, the BB proposed here took less than 40 minutes to generate $10,000$ BB posterior samples. We acknowledge that this is not a fair comparison, because the computations were executed on different machines with different programs. However, we believe that it gives a clear impression on the speed of the BB computation.

Appendix. The proof of (7)

Since F has increasing sample paths (distribution functions), it suffices to show convergence in distribution of the finite dimensional distributions of F. That is, for given sequence of numbers $0 = t_0 < t_1 < \ldots < t_K < t_{K+1} = \infty$, it suffices to prove that

$$\begin{aligned} &\mathcal{L}(F(t_1), F(t_2) - F(t_1), \ldots, F(t_{K+1}) - F(t_K)|D, \alpha, F_0) \\ &\xrightarrow{d} \mathcal{L}^{(B)}(F(t_1), F(t_2) - F(t_1), \ldots, F(t_{K+1}) - F(t_K)|D). \end{aligned} \tag{12}$$

For $i = 1, 2, 3$, let $\Delta_i = \{W_k : \delta_k = i, k = 1, \ldots, n\}$. Then, the posterior distribution of F given only the uncensored observations Δ_1 is a Dirichlet process with mean F_n and precision parameter α_n, where

$$F_n(t) = \frac{\sum_{i:W_i \in \Delta_1} I(W_i \le t) + \alpha F_0(t)}{n_1 + \alpha}$$

and $\alpha_n = n_1 + \alpha$. Here, n_1 is the number of observations in Δ_1. Hence, the posterior distribution of F is the same as the posterior distribution of F given Δ_2 and Δ_3 only, with the Dirichlet process prior having mean F_n and precision parameter α_n.

Next, let $\mathcal{S} = \{t_1, \ldots, t_K\} \cup \Delta_2 \cup \Delta_3$ and let $s_1 < \cdots < s_m$ be the ordered distinct numbers in $\mathcal{S}$. Then the likelihood of Δ_2 and Δ_2 depends on $(F(s_1), F(s_2) - F(s_1), \ldots, F(s_{m+1}) - F(s_m))$, where $s_{m+1} = \infty$. Hence we can obtain the posterior density of $(F(s_1), F(s_2) - F(s_1), \ldots, F(s_{m+1}) - F(s_m))$ given Δ_2 and Δ_3 by direct calculation. Finally, it is easy to show that the posterior density of $(F(s_1), F(s_2) - F(s_1), \ldots, F(s_{m+1}) - F(s_m))$ converges to the BB posterior density as $\alpha \to 0$, and the proof is done.

Acknowledgement

This research was supported (in part) by KOSEF through Statistical Research Center for Complex Systems at Seoul National University. Also, the research of the first author was supported in part by grant R01-2004-000-10284-0 from the Basic Research Program of the Korea Science & Engineering Foundation.

References

Chang, M. N. (1990). Weak convergence of a self-consistent estimator of the survival function with doubly censored data. *Ann. Statist.* **18**, 391-404.

Chang, M. N. and Yang, G. L. (1987). Strong consistency of a nonparametric estimator of the survival function with doubly censored data. *Ann. Statist.* **15**, 1536-1547.

Choudhuri, N. (1998). Bayesian bootstrap credible sets for multidimensional mean functional. *Ann. Statist.* **26**, 2104-2127.

Efron, B. (1979). Bootstrap methods: another look at the jackknife. *Ann. Statist.* **7**, 1-26.

Ferguson, T. S. (1973). A Bayesian analysis of some nonparametric problems. *Ann. Statist.* **1**, 209-230.

Gasparini, M. (1995). Exact multivariate Bayesian bootstrap distributions of moments. *Ann. Statist.* **23**, 762-768.

Gehan, E. A. (1965). A generalized two-sample Wilcoxon test for doubly censored data. *Biometrika* **52**, 650-653.

Gu, M. G. and Zhang, C. H. (1993). Asymptotic properties of self-consistent estimators based on doubly censored data. *Ann. Statist.* **21**, 611-624.

Hjort, N. L. (1991). Bayesian and empirical Bayesian bootstrapping. Technical Report, University of Oslo and Norwegian Computing Centre.

James, L. F. (1997). A study of a class of weighted bootstraps for censored data. *Ann. Statist.* **25**, 1595-1621.

Kim, Y. and Lee, J. (2003). Bayesian bootstrap for proportional hazards models. *Ann. Statist.* **31**, 1905-1922.

Kuo, L. and Smith, F. M. (1992). Bayesian computations in survival models via the Gibbs sampler. In *Survival Analysis: State of the Art* (Edited by J. P. Klein and P. K. Goel), 11-24. Kluwer.

Lazar, N. A. (2000). Bayesian empirical likelihood. Technical report 721, Statistics department, Carnegie Mellon University.

Lo, A. Y. (1987). A large-sample study of the Bayesian bootstrap. *Ann. Statist.* **15**, 360-375.

Lo, A. Y. (1988). A Bayesian bootstrap for a finite population. *Ann. Statist.* **16**, 1684-1695.

Lo, A. Y. (1993). A Bayesian bootstrap for censored data. *Ann. Statist.* **21**, 100-123.

Mantel, N. (1967). Ranking procedures for arbitrarily restricted observations. *Biometrika* **23**, 65-78.

Mason, D. M. and Newton, M. A. (1992). A rank statistics approach to the consistency of a general bootstrap. *Ann. Statist.* **20**, 1611-1624.

Mykland, P. A. and Ren, J. J. (1996). Algorithms for computing self-consistent and maximum likelihood estimators with doubly censored data. *Ann. Statist.* **24**, 1740-1764.

Owen, A. B. (1990). Empirical likelihood ratio confidence regions. *Ann. Statist.* **18**, 90-120.

Peto, R. (1973). Exponential survival curves for interval-censored data. *Appl. Statist.* **22**, 86-91.

Praestgaard, J. and Wellner, J. A. (1993). Exchangeably weighted bootstraps of the general empirical process. *Ann. Probab.* **21**, 2053-2086.

Rubin, D. B. (1981). The Bayesian bootstrap. *Ann. Statist.* **9**, 130-134.

Tsai, W. Y. and Crowley, J. (1985). A large sample study of generalized maximum likelihood estimators from incomplete data via self-consistency. *Ann. Statist.* **13**, 1317-1334.

Turnbull, B. W. (1974). Nonparametric estimation of a survivorship function with doubly censored data. *J. Amer. Statist. Assoc.* **69**, 169-173.

Turnbull, B. W. and Weiss, L. (1978). A likelihood ratio statistic for testing goodness of fit with randomly censored data. *Biometrics* **34**, 367-375.

Wellner, J. A. and Zhan, Y. (1996). Bootstrapping Z-estimators. Technical Report 308, Department of Statistics, University of Washington.

Wellner J. A. and Zhan, Y. (1997). A hybrid algorithm for computation of the nonparametric maximum likelihood estimator from censored data. *J. Amer. Statist. Assoc.* **92**, 945-959.

Weng, C. S. (1989). On a second-order asymptotic property of the Bayesian bootstrap mean. *Ann. Statist.* **17**, 705-710.

Department of Statistics, Seoul National University, Seoul 151-742, Korea.

E-mail: ydkim@stats.snu.ac.kr

Department of Statistics, Seoul National University, Seoul 151-742, Korea.

E-mail: leej@stats.snu.ac.kr

Department of Statistics, Seoul National University, Seoul 151-742, Korea.

E-mail: jskim@stats.snu.ac.kr

(Received November 2003; accepted October 2004)