BAYESIAN BOOTSTRAP FOR PROPORTIONAL HAZARDS MODELS

BY YONGDAI KIM[1] AND JAEYONG LEE

Ewha Womans University, Korea and Pennsylvania State University

We propose two Bayesian bootstrap extensions, the binomial and Poisson forms, for proportional hazards models. The binomial form Bayesian bootstrap is the limit of the posterior distribution with a beta process prior as the amount of the prior information vanishes, and thus can be considered as a default nonparametric Bayesian analysis. It is also the same as Lo's Bayesian bootstrap for censored data when covariates are absent. The Poisson form Bayesian bootstrap is equivalent to the Bayesian analysis with Cox's profile likelihood. When the baseline distribution is discrete, thus when the data set has many ties, simulation study suggests that the binomial form Bayesian bootstrap performs better than standard frequentist procedures in the frequentist sense. An advantage of the proposed Bayesian bootstrap procedures over the standard Bayesian analysis is conceptual and computational simplicity. Finally, it is shown that both Bayesian bootstrap posteriors are asymptotically equivalent to the sampling distribution of the maximum likelihood estimator.

1. Introduction. Bayesian analysis of the proportional hazards model has been studied for more than two decades by many statisticians, Kalbfleisch (1978), Hjort (1990) and Laud, Damien and Smith (1998), to name just a few, but it still remains a field only for experts. Difficulties of Bayesian analysis, both conceptual and computational, come from modelling the baseline cumulative hazard function (c.h.f.) with nonparametric priors, for example, beta process [Hjort (1990)] or gamma process [Kalbfleisch (1978)], whose theory and computation are often formidable to practitioners. To circumvent such complications of the full Bayesian analysis, we propose Bayesian bootstrap (BB) procedures which, we believe, are easily accessible to practitioners and at the same time are reliable inference procedures. The proposed BB procedures approximate the baseline c.h.f. with the family of c.h.f.'s having masses only at uncensored data points. Thus, the BB procedures are conceptually parametric and computationally simple but retain the flexibility of nonparametric models. Another advantage of the BB procedures is that it is unnecessary to elicit prior information on the baseline c.h.f. This is an important advantage, for elicitation of prior information on nonparametric objects is not an easy task.

From the frequentist's point of view, the proposed BB procedures can be

considered as new methods of approximating the sampling distribution of the maximum likelihood estimator (MLE), because the BB posterior is proved to be asymptotically equivalent to the sampling distribution of the MLE. Moreover, simulation study suggests that the binomial form BB posterior approximates the sampling distribution of the MLE better than standard frequentist methods for grouped survival data.

BB was first introduced by Rubin (1981) and extended to finite population model [Lo (1988)] and right censored data [Lo (1993)]. There are three views of BB. We will briefly review these to give background for the proposed BB procedures.

The first view is that BB extends Efron's bootstrap [Efron (1979)]. Consider a situation where $\mathcal{X} = (X_1, X_2, \ldots, X_n)$ is a random sample from an unknown distribution F and a functional of F, $T(F)$, is of interest. In this situation, a typical bootstrap procedure consists of sampling 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 on $T(F)$ is based on $T(F_i^*)$s where F_i^* is the empirical distribution of $\mathcal{X}_i^*$. Noting that $F_i^* \stackrel{d}{=} \sum_{j=1}^n w_j \delta_{X_j}$, where $nw = n(w_1, w_2, \ldots, w_n) \sim Multinomial(n, 1/n, \ldots, 1/n)$, Rubin proposed a smoother alternative, $Dirichlet(1, 1, \ldots, 1)$, for the distribution of w. For this reason, BB is viewed as a weighted bootstrap. The weighted bootstrap provides a unified framework for theoretical study of Efron's bootstrap and Rubin's BB. Theoretical properties of the BB and weighted bootstrap have been studied by many authors, for instance, Lo (1987), Weng (1989), Mason and Newton (1992), Præstgaard and Wellner (1993), Gasparini (1995), James (1997) and Choudhuri (1998). Lo (1993) [see also James (1997)] extended this idea to censored data. He observed that the Kaplan–Meier estimator of the survival function $S(t) = 1 - F(t)$ is

$$(1) \qquad \hat{S}(t) = \prod_{t \in \mathcal{T}_n} \left(\frac{1 - \sum_{j \in D(t)} 1}{\sum_{j \in R(t)} 1} \right),$$

where $\mathcal{T}_n$ is the set of distinct uncensored observations, $D(t)$ the set of uncensored observations at time t and $R(t)$ the risk set at time t. He proposed to replace the 1's in (1) by independent gamma random variables. Lo (1993) showed that this BB posterior can be used as an approximation to the posterior obtained from the *beta neutral process* prior, and James (1997) studied the second order properties of this approximation. Furthermore, Wellner and Zhan (1996) extended the weighted bootstrap to Z-estimators and showed their method can be applied to doubly censored data and a frailty model.

The second view is that the BB posterior is the limit of the full Bayesian posterior as the amount of prior information vanishes, and that BB is a noninformative (or default) Bayesian analysis of nonparametric problems. If the prior on F is a Dirichlet process with parameter α, a non-null finite measure, the posterior is again a 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 the use of the BB as a default nonparametric Bayesian analysis. At the same time, BB gives a good approximation to the posterior. The same property holds for Lo's BB and the binomial BB proposed in this paper for censored data and the proportional hazards model, respectively. See Section 2.4 for details. Even though this view provides a clear connection between the BB and full Bayesian analysis, it requires detailed information about the full Bayesian posterior for constructing the BB, which may be infeasible in most situations.

The third view is that the BB posterior is obtained by

$$
\text{(2)} \qquad \text{BB posterior} \propto \text{empirical likelihood} \times \text{prior.}
$$

For the following discussion, assume that there are no ties among $\mathcal{X}$. Suppose the true distribution belongs to $\mathcal{F}_n = \{\sum_{i=1}^{n} w_i \delta_{X_i}\}$. 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

$$
\text{(3)} \qquad L(F) = \prod_{i=1}^{n} w_i.
$$

Adopting a noninformative prior, $\prod_{i=1}^{n} w_i^{-1}$, by the usual Bayesian computation we get Rubin's BB posterior for w, $Dirichlet(1, 1, \ldots, 1)$. This description of BB was noted by many authors, including Rubin (1981), Owen (1990), Choudhuri (1998) and Lazar (2000). Note that the likelihood used in the above derivation is the empirical likelihood of Owen (1990). Thus, this view can be effectively summarized by (2). This is the main idea we will use to derive the BB posterior for the proportional hazards model. The above discussion provides an idea for constructing the empirical likelihood which will be used in this paper. That is, first the model is reduced to a data-dependent parametric model, $\mathcal{F}_n$, and the likelihood on $\mathcal{F}_n$ is the empirical likelihood.

This paper is organized as follows. In Section 2, BB procedures for the proportional hazards model are proposed and components of the BB—BB likelihoods, priors, posteriors—and computational algorithms using Markov chain Monte Carlo (MCMC) are discussed. In Sections 3 and 4, simulation results are presented and a real data set is analyzed by the proposed BB methods. Discussion of issues around the BB is presented in Section 5. Finally, the asymptotic theory is given in Section 6.

2. BB for Cox's proportional hazards model.

2.1. *The model.* Let $X_1, \ldots, X_n$ be survival times with covariates $Z_1, \ldots, Z_n \in \mathcal{R}^p$. Suppose the distribution F_i of X_i with covariate Z_i is given by $1 - F_i(t) = (1 - F(t))^{\exp(\beta^T Z_i)}$ for an unknown regression parameter $\beta \in R^p$, where F is the baseline distribution function. The c.h.f. A of the cumulative

distribution function (c.d.f.) F is defined by $dA(t) = dF(t)/(1 - F(t-))$ and the c.h.f. A_i of F_i is given by $dA_i(t) = 1 - (1 - dA(t))^{\exp(\beta^T Z_i)}$. If A is absolutely continuous with respect to Lebesgue measure, there exists a hazard function a such that $A(t) = \int_0^t a(s)\, ds$ and the hazard function of A_i is $a(t) \exp(\beta^T Z_i)$. The survival times are subject to right censoring, and only $(T_1, \delta_1, Z_1), \ldots, (T_n, \delta_n, Z_n)$ are observed, where $T_i = \min(C_i, X_i)$, $\delta_i = I(X_i \le C_i)$ and $C_1, \ldots, C_n$ are independent censoring variables with a common c.d.f. G. For $i = 1, 2, \ldots, n$, define counting processes $N_i(t) = I(T_i \le t, \delta_i = 1)$ and $Y_i(t) = I(T_i \ge t)$. Also, let $N(t) = \sum_{i=1}^n N_i(t)$, $\Delta N(t) = N(t) - N(t-)$, and $Y(t) = \sum_{i=1}^n Y_i(t)$.

2.2. *BB likelihoods.* Two forms of the likelihood using product-integration, the binomial and Poisson forms, are available for the proportional hazards model. They are denoted by $L^B(\beta, A)$ and $L^P(\beta, A)$, respectively, and are

$$(4) \qquad L^B(\beta, A) = \prod_{i=1}^n \prod_{t \in [0, \tau]} dA_i(t)^{dN_i(t)} (1 - dA_i(t))^{Y_i(t) - dN_i(t)},$$

$$(5) \qquad L^P(\beta, A) = \prod_{i=1}^n \prod_{t \in [0, \tau]} (Y_i(t)\, dA_i(t))^{dN_i(t)} \exp\left(-\int_0^\tau Y_i(t)\, dA_i(t)\right),$$

where $[0, \tau]$ is the time period of interest. Both of them are the real likelihoods when the c.d.f. is absolutely continuous, while only the former is when the c.d.f. is discrete. Nevertheless, both are used in the derivation of BB likelihoods since each has its own merits. For the interpretation of the two likelihoods (4) and (5), see Andersen, Borgan, Gill and Keiding (1993).

To construct the BB likelihoods for the proportional hazards model, an appropriate reduced parametric model $\mathcal{F}_n$ needs to be chosen first. Lo (1993) made a clever choice of $\mathcal{F}_n$ for censored data, namely, $\mathcal{F}_n = \{\sum p_i \delta_{y_i} : p_i \ge 0, \sum p_i = 1\}$, where the y_i's are distinct uncensored observations. We employ the same $\mathcal{F}_n$ for the proportional hazards models. For each t, define $D(t) = \{i : \Delta N_i(t) = 1\}$ and $R(t) = \{i : Y_i(t) = 1\}$. The set $D(t)$ is the set of observations which fail at time t and $R(t)$ is the set of observations still at risk at time t. Let $\mathcal{T}_n = \{t : \Delta N(t) \ge 1\}$. Then $\mathcal{F}_n$ becomes the set of all probability measures with support $\mathcal{T}_n$. The BB likelihoods can be derived from Poisson as well as binomial form likelihoods. Replacing $1 - dA_i(t)$ with $(1 - \Delta A(t))^{\exp(\beta^T Z_i)}$ at each $t \in \mathcal{T}_n$ in the binomial form likelihood (4), we obtain the binomial form BB (BFBB) likelihood

$$L_n^B(\beta, A) = \prod_{t \in \mathcal{T}_n} \left(\prod_{i \in D(t)} \left(1 - (1 - \Delta A(t))^{\exp(\beta^T Z_i)}\right)^{\Delta N_i(t)} \right)$$
$$\times (1 - \Delta A(t))^{\sum_{i \in R(t) \setminus D(t)} \exp(\beta^T Z_i)}.$$

Similarly, replacing $dA_i(t)$ with $\exp(\beta^T Z_i)\Delta A(t)$ for $t \in \mathcal{T}_n$ in (5), we get the Poisson form BB (PFBB) likelihood

$$L_n^P(\beta, A) = \prod_{t \in \mathcal{T}_n} \exp\left(\sum_{i \in D(t)} \beta^T Z_i\right) \Delta A(t)^{\Delta N(t)} \exp\left(-\Delta A(t) \sum_{i \in R(t)} \exp(\beta^T Z_i)\right).$$

REMARK. Prentice and Gloeckler (1978) proposed to maximize the BFBB likelihood to get estimators of β and the baseline survival probability for grouped survival data. They argued that the BFBB likelihood was computationally simpler than the real and exact partial likelihoods [Peto (1972) and Cox (1972)] and yielded reasonable answers. However, they reported instability of numerical maximization and the possibility of several local maxima. The Bayesian bootstrap avoids these problems by exploring the whole surface of the BFBB likelihood with efficient MCMC algorithms. The simulation results in Section 3 show that the BFBB posterior yields reliable results for both continuous and discrete survival models.

REMARK. The PFBB likelihood $L_n^P(\beta, A)$ is also used by frequentists to derive the partial likelihood. See Andersen, Borgan, Gill and Keiding (1993) for details.

2.3. *Priors.* Note that the prior on ΔA used for Rubin's BB is not proper. Interestingly, no proper priors produce consistent posteriors. We propose the following two priors for the BFBB and PFBB, respectively:

$$(6) \qquad \pi(A) = \prod_{t \in \mathcal{T}_n} \Delta A(t)^{-1}(1 - \Delta A(t))^{-1},$$

$$(7) \qquad \pi(A) = \prod_{t \in \mathcal{T}_n} \Delta A(t)^{-1}.$$

A motivation for (6) is that the BFBB posterior with prior (6) can be obtained by the limit of the full Bayesian posterior as the prior information vanishes. Consider the full Bayesian analysis with the beta process with parameters $c(t)$ and $A_0(t)$ on the c.h.f., where $A_0(t)$ is the prior mean, and $c(t)$ governs the amount of prior information. See Hjort (1990) for a description of beta processes. If we let $c(t)$ go to 0 (the amount of prior information goes to 0), the full Bayesian posterior converges to the BFBB posterior.

The motivation of prior (7) is more involved. Popular priors of c.h.f. A for full Bayesian analysis—beta processes (including Dirichlet processes) and gamma processes—are Lévy processes. Interestingly, not all Lévy process priors for A have consistent posteriors for the full Bayesian analysis [Kim and Lee (2001)]. Suppose A is a Lévy process with Lévy measure $\nu([0, t] \times D) = \int_0^t \int_D f_s(x)\,dx\,ds$, for $D \subset [0, 1]$. Roughly, a sufficient condition for the posterior consistency is that $f_s(x) \approx c/x$ near $x = 0$, $c > 0$. Since $f_s(x)$ can be considered as the density of the

jump size at s of the Lévy process, this result suggests that the prior of $\Delta A(t)$ needs to behave like $1/\Delta A(t)$ for $\Delta A(t)$ near 0. Among such priors, we choose (7) since the marginal posterior distribution of the regression coefficients is proportional to the product of the partial likelihood and the prior. Also, in Section 6, we justify the use of the priors (6) and (7) by proving that the two BB posteriors are asymptotically equivalent to the sampling distribution of the MLE.

We recommend a subjective proper prior for β when prior information is available for β. The constant prior for β can be used if prior information is not available, but propriety of the posterior needs to be checked. For the PFBB, the posterior is proper if $\text{coni}(\mathcal{A}) = \mathcal{R}^p$, where $\mathcal{A} = \{z_i - z_j : j \in R(t) \setminus D(t), i \in D(t), t \in \mathcal{T}_n\}$ and the conical hull of $\mathcal{A}$, $\text{coni}(\mathcal{A})$, is the set of all conical (non-negative linear) combinations of the vectors in $\mathcal{A}$. This is because the partial likelihood is log-concave and has a unique maximizer [Jacobsen (1989)].

More caution needs to be taken for the propriety of the BFBB, because even with a proper prior on β the BFBB posterior can be improper. Suppose that there are no tied uncensored observations. Then the marginal BFBB posterior of β with constant prior is

$$\pi(\beta | Data) = c \prod_{t \in \mathcal{T}_n} \left[\psi\left(\sum_{i \in R(t)} \exp(\beta^T Z_i) \right) - \psi\left(\sum_{i \in R(t) \setminus D(t)} \exp(\beta^T Z_i) \right) \right],$$

for some positive constant c, where $\psi(x) = \int_0^1 (1 - (1 - y)^{x-1})/y \, dy$. Since $d\psi(x)/dx = \int_0^\infty (y e^{-yx})/(1 - e^{-y}) \, dy \geq 1/x^2$, the mean value theorem yields that

$$\pi(\beta | Data) \geq c \prod_{t \in \mathcal{T}_n} \frac{\exp(\beta^T Z_{d(t)})}{(\sum_{i \in R(t)} \exp(\beta^T Z_i))^2},$$

where $d(t) \in D(t)$. Suppose the data observed are $(1, 1, -1)$, $(2, 1, -1.9)$, and $(3, 1, -1.5)$ with coordinates being T_i, δ_i and Z_i, respectively. Then the BFBB posterior is improper even with some proper priors on β, whereas the PFBB posterior is proper with constant prior on β.

2.4. *BB posteriors.* The BFBB posterior with prior (6) for $\Delta A(t)$ and $\pi(\beta)$ for β is given by

$$(8) \qquad \pi(\beta, \Delta A | Data) \propto L_n^B(\beta, A) \prod_{t \in \mathcal{T}_n} \Delta A(t)^{-1} (1 - \Delta A(t))^{-1} \pi(\beta).$$

Due to the form of the BFBB likelihood, neither of the full conditional distributions of β and $\Delta A(t)$ is a well-known distribution. But Bayesian computation with a Metropolis–Hastings algorithm does not pose any difficulty.

With prior (7) for $\Delta A(t)$ and $\pi(\beta)$ for β, the PFBB posterior of β and $\Delta A(t)$ is given by

$$\pi(\beta, \Delta A | Data) \propto \pi(\beta) \prod_{t \in \mathcal{T}_n} \exp\left(\sum_{i \in D(t)} \beta^T Z_i \right) \Delta A(t)^{\Delta N(t) - 1}$$

$$\times \exp\left(-\Delta A(t) \sum_{i \in R(t)} \exp(\beta^T Z_i) \right).$$

Thus, $\Delta A(t)$ for $t \in \mathcal{T}_n$, given β, are independent and follow

$$\Delta A(t) | \beta, Data \sim Gamma\left(\Delta N(t), \sum_{i \in R(t)} \exp(\beta^T Z_i) \right).$$

Integrating out $\Delta A(t)$, we get the marginal posterior of β,

$$\pi(\beta | Data) \propto \pi(\beta) \prod_{t \in \mathcal{T}_n} \frac{\exp(\sum_{i \in D(t)} \beta^T Z_i)}{(\sum_{i \in R(t)} \exp(\beta^T Z_i))^{\Delta N(t)}}$$

$$= \pi(\beta) \times \text{Cox's partial likelihood.}$$

Since the partial likelihood for β is log-concave, so is the posterior of β with a log-concave prior. In this case, adaptive rejection sampling [Gilks and Wild (1992)] can be used coordinatewise.

REMARK.   The MCMC schemes for the BBs do not include the sample path generation of a Lévy process and, hence, are much simpler than that of the full Bayesian computation.

REMARK.   Consider the survival model without covariates. Then, the BFBB posterior reduces to Lo's BB posterior for censored data: $\Delta A(t)$, $t \in \mathcal{T}_n$, are independent and follow $Beta(\Delta N(t), Y(t) - \Delta N(t))$. In the PFBB case, $\Delta A(t)$ are independent and follow $Gamma(\Delta N(t), r(t))$, where $r(t)$ is the cardinality of $R(t)$. This provides an alternative to Lo's BB for censored data without covariates.

REMARK.   When the regression coefficient β is the only parameter of interest, the PFBB gives a justification for the Bayesian analysis with the partial likelihood and a prior on β, as was done in Volinsky, Madigan, Raftery and Kronmal (1996).

REMARK.   Since the BFBB and PFBB were derived under implicit assumptions of discrete and continuous c.h.f., respectively, the survival probabilities have slightly different forms. For the estimation of $1 - F(t)$, we recommend $\prod_{T_i \leq t, \delta_i = 1} (1 - \Delta A(T_i))$ for the BFBB and $\exp(-\sum_{T_i \leq t, \delta_i = 1} \Delta A(T_i))$ for the PFBB. In fact, the former cannot be used for the PFBB because the PFBB distribution of $\Delta A(t)$ is not restricted to $[0, 1]$.

3. A simulation study. In this section, we conduct a simulation experiment to compare frequentist coverage probability of the proposed BB posterior of β with that of standard frequentist procedures, Breslow and Efron partial likelihoods. Breslow's and Efron's partial likelihoods coincide for data sets without ties, but differ otherwise. For a detailed description of Breslow's and Efron's likelihoods, see Kalbfleisch and Prentice (1980).

We consider two baseline c.d.f.'s: a continuous distribution $F_0 = Exponential(1)$ and a discrete distribution F_1 with support $\{0.3, 0.5, 1\}$ and probability $1/3$ at each support point. The covariate Z is one-dimensional and generated from $Bernoulli(1/2)$. The exponential distribution with mean 2 is used for the censoring distribution.

We consider 4 sample sizes, 30, 50, 100 and 200. For the BB posteriors, Markov chains for the BB computation were run for 10,000 iterations with the first 500 iterations discarded as burn-in. For each method considered, empirical frequentist coverage probability of the 95% interval estimate of β is calculated based on 2000 independent samples. Nominal coverage probability is 95%. The baseline distributions F_0 and F_1 are $Exponential(1)$ and the discrete distribution with support $\{0.3, 0.5, 1\}$ and probability $1/3$ at each support point.

The result of the simulation study is shown in Table 1. The simulation study reveals the striking feature that when the baseline c.d.f. is discrete (F_1 in Table 1), the performance of all methods except the BFBB gets worse as the sample size increases. The strong performance of the BFBB seems to be due to the fact that only the BFBB likelihood is the true likelihood when the baseline c.d.f. is discrete. The BFBB works consistently well for moderately large samples ($n = 100, 200$) under both the continuous and discrete models, although it shows smaller coverage probabilities than the nominal one for small sample sizes ($n = 30$). Thus, when the data has many ties and the sample size is moderately large, the BFBB credible interval is recommended even for the frequentists. For the continuous c.d.f., both Breslow's and Efron's methods work very well. Both BB methods perform well

TABLE 1
Empirical frequentist coverage probabilities of the BFBB, PFBB and
Breslow's and Efron's partial likelihoods

Baseline c.d.f.	Sample size	BFBB	PFBB	Breslow	Efron
F_0	30	91.40	93.45	95.15	95.15
	50	94.10	95.25	95.55	95.55
	100	94.10	95.05	95.70	95.70
	200	93.95	94.55	95.40	95.40
F_1	30	91.95	95.70	93.85	94.80
	50	93.40	62.15	82.70	92.75
	100	94.00	78.35	63.90	90.55
	200	94.85	47.35	24.50	85.35

for the continuous case, although their empirical coverage probabilities are slightly smaller than the nominal level for small sample sizes.

4. Analysis of leukemia patient remission data.   In this section, we analyze the leukemia patient remission time data set analyzed by Laud, Damien and Smith (1998). A total 42 patients is divided into control and treatment groups with 21 patients each. The treatment and control groups are given 6-mercaptopurine (or 6-MP) and placebo, respectively. Of the 42 observations, 12 are censored and 16 are tied.

In Table 2, we present the point estimates and 90% interval estimates of β from various methods. The first block of rows of Table 2 is quoted from Laud, Damien and Smith (1998). Using a Bayesian approach, they analyzed the data with the beta process prior with mean $A^*(t) = 0.05t$ and $c(t) = ke^{-0.05t}$. The constant k governs the amount of the prior information. Thus as k tends to 0, it is expected that the result of the BFBB is close to that of the full Bayesian analysis with small k. The BB analyses are in the second block, and the frequentist analyses in the third block. As also expected, the PFBB produces results similar to Breslow's and Efron's methods.

The PFBB and BFBB generate slightly different results which seem to be due to small sample size and many ties. However, Figure 1, depicting the histograms of samples from the PFBB and BFBB posteriors, shows that the difference of the two BB posterior distributions of β is not significant.

Figure 2 shows the Bayes estimate of the survival function with 5% and 95% quantiles and sample paths of the survival function from the PFBB posterior. The BFBB gives similar results and the figures are not included here. In these two figures, survival curves are drawn by linearly interpolating the estimates of survival probabilities at $t \in \mathcal{T}_n$. An advantage of the full Bayesian analysis is that it enables one to see the variation of each sample c.d.f. The BB retains this advantage. Thus, these figures show that with the BB posterior the same analysis can be done as with the full Bayesian analysis.

TABLE 2
Point and interval estimates of β of various methods

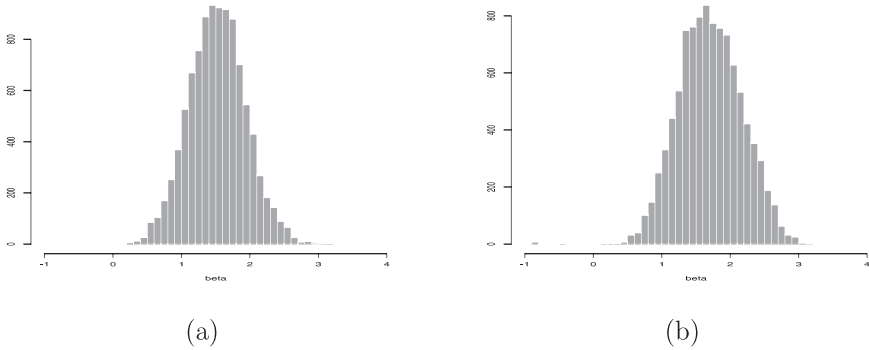| Method	k	Point estimate	90% C.I.
Full Bayesian	0.10	1.71	(1.06, 2.44)
	1.0	1.71	(1.02, 2.42)
	10.0	1.67	(0.98, 2.36)
PFBB		1.54	(0.86, 2.24)
BFBB		1.71	(1.00, 2.44)
Breslow		1.51	(0.83, 2.18)
Efron		1.56	(0.88, 2.24)

FIG. 1.   *Posterior samples of β*: *Bayesian bootstrap using* (a) *Poisson form and* (b) *binomial form.*
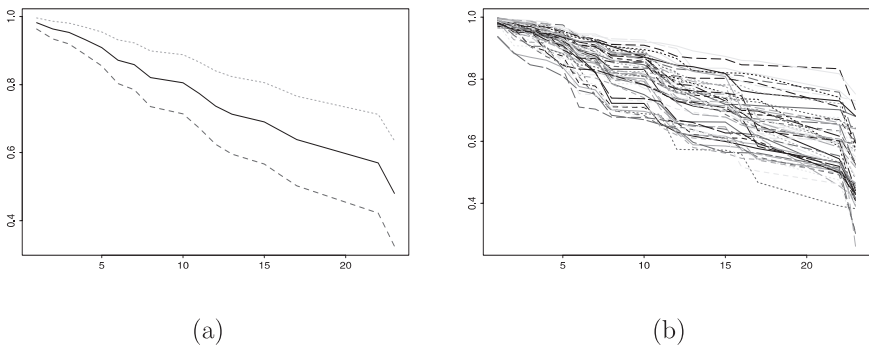

FIG. 2.   (a) *Posterior mean*, *and* 5%, 95% *quantiles of survival probabilities*; (b) *posterior samples of survival curves*. *The y-axis is survival probability and the x-axis is in months.*

5. Discussion. We proposed two BB procedures based on the Poisson and binomial form likelihoods. Both BB procedures can be considered as noninformative Bayesian analyses, but the BFBB approximates the posterior distribution better as we have seen in Table 2. The BFBB also performs better for discrete distributions than the PFBB and standard frequentist methods. Thus, it is recommended to frequentists for data sets with moderate sample size and many ties. The PFBB has nice mathematical equivalence with the Bayesian analysis of Cox profile likelihood and does not have propriety problems as in the BFBB.

In Section 2, we chose the prior $\pi(x)$ for $\Delta A(t)$ as $\pi(x) = 1/x$ for the PFBB and $\pi(x) = 1/(x(1-x))$ for the BFBB. These priors have many desirable properties. First, the BFBB posterior can be obtained as the limit of the full Bayesian posterior. Second, the marginal posterior of β for the PFBB is proportional to the product of the partial likelihood and the prior. Third, the BFBB and PFBB posteriors are asymptotically equivalent to the sampling distribution of the maximum likelihood estimator. However, this does not mean that we exclude other alternatives. For example, with priors $\pi(x) = x^{-1}(1-x)^{1-\alpha}, \alpha > 0$, a simple

propriety condition for the BFBB can be established. However, the choice of α is not obvious. More investigation into this matter needs to be done.

There are many possible applications of the BB for more complicated problems such as general censored data (doubly censoring and interval censoring) and random effect models. In such problems, the estimation of the variance of the MLE is computationally demanding. In contrast, BB may be implemented easily for those models with MCMC, and the computed BB posterior variance may be used as the variance estimate of the MLE.

6. Asymptotic results. In this section, we prove that the BB posterior distributions of β and A are asymptotically equivalent to the sampling distribution of the maximum likelihood estimators, $\hat{\beta}$ and $\hat{A}$. Denote by A_0 and β_0 the true values of A and β. We only consider the PFBB. The proof for the BFBB can be done similarly. We assume the following regularity conditions:

1. The true c.h.f. A_0 is absolutely continuous.
2. There is a positive constant τ such that $G(\tau) = 1$, $G(\tau-) < 1$ and $A_0(\tau) < \infty$, where G is the c.d.f. of the censoring variables.
3. Covariates $Z_i \in R^p$ are i.i.d. with $|Z_i| < M$ for a constant $M > 0$, where $|Z_i|$ is defined by

$$|Z_i| = \sum_{k=1}^{p} |Z_{ik}|.$$

4. The support of the distribution of Z_i contains an open set.
5. The prior density of β, $\pi(\beta)$, is continuous at β_0 with $\pi(\beta_0) > 0$.

The above conditions are assumed throughout this section. Condition (1) prevents ties. Condition (2) implies that the follow-up ends before all observations fail; this is satisfied in most practical problems. Condition (3) is for technical purposes and can be relaxed to the finiteness of certain moments of Z. For example, see Assumption 3.1 in Tsiatis (1981). Condition (4) avoids collinearity among the covariates and ensures the uniqueness of the MLE $\hat{\beta}$. Condition (5) is a standard assumption in Bayesian asymptotic theory. We consider only proper priors of β here.

Before we state theorems, we introduce the necessary notation. Let

$$S^0(t:\beta) = E\left(\sum_{i \in R(t)} \exp(\beta^T Z_i) \right),$$

$$S^1(t:\beta) = E\left(\sum_{i \in R(t)} Z_i \exp(\beta^T Z_i) \right),$$

$$S^2(t:\beta) = E\left(\sum_{i \in R(t)} Z_i Z_i^T \exp(\beta^T Z_i) \right)$$

and

$$V(t:\beta) = S^2(t:\beta)/S^0(t:\beta) - e_0(t)e_0(t)^T,$$

where

$$e_0(t) = \int_0^t S^1(s:\beta_0)/S^0(s:\beta_0)\,dA_0(s).$$

Define $\Lambda_0(t) = \int_0^t 1/S^0(s:\beta_0)\,dA_0(s)$ and $I(\beta) = \int_0^\tau V(t:\beta)S^0(t:\beta)\,dA_0(t)$. The next theorem is the main result of this section.

THEOREM 1. *Let $\mathcal{L}^B(\cdot|Data)$ denote the PFBB posterior distribution given the data. Then, as $n \to \infty$,*

$$\mathcal{L}^B\big(\sqrt{n}(A(\cdot) - \hat{A}(\cdot), \beta - \hat{\beta})|Data\big) \overset{d}{\to} (W(\Lambda_0(\cdot)) - Xe_0(\cdot), X)$$

on $D[0, \tau] \times R^p$ with probability 1, where X follows the multivariate normal distribution with mean 0 and variance $I^{-1}(\beta_0)$ and W is a standard Brownian motion. Here $D[0, \tau]$ is the space of right continuous functions on $[0, \tau]$ with left limits equipped with the uniform topology.

We divide the proof of Theorem 1 into two theorems.

THEOREM 2. *As $n \to \infty$,*

$$(9) \qquad\qquad \lim_{n\to\infty} \|f_n - \phi\| = 0 \qquad a.s.,$$

where f_n is the posterior density of $\sqrt{n}(\beta - \hat{\beta})$ from the PFBB, ϕ is the normal density with mean 0 and variance $I(\beta_0)^{-1}$ and $\|\cdot\|$ is the L_1 norm.

THEOREM 3.

$$\mathcal{L}^B\big(\sqrt{n}(A(\cdot) - \hat{A}(\cdot))|\sqrt{n}(\beta - \hat{\beta}) = x, Data\big) \overset{d}{\to} W(\Lambda_0(\cdot)) - xe_0(\cdot)$$

on $D[0, \tau]$ with probability 1.

Throughout the proofs, we write $o(n^\delta)$ for a sequence of random variables Z_n such that $Z_n/n^\delta \to 0$ with probability 1.

PROOF OF THEOREM 1. Theorems 2 and 3 prove the convergence of the marginal distribution of β and the conditional distribution of A given β. To prove the convergence of the joint PFBB posterior distribution of β and A, note that Theorem 2 implies the strong convergence of the marginal PFBB posterior distribution of $\sqrt{n}(\beta - \hat{\beta})$ to the distribution of X. Applying Theorem 2 of Sethuraman (1961) completes the proof. □

PROOF OF THEOREM 2. Let

$$L_n(\beta) = \prod_{i \in \mathcal{T}_n} \frac{\exp(\beta^T Z_i)}{\sum_{j \in R(T_i)} n^{-1} \exp(\beta^T Z_j)}$$

and

$$(10) \qquad g_n(h) = \frac{L_n(\hat{\beta} + h/\sqrt{n})}{L_n(\hat{\beta})} \pi(\hat{\beta} + h/\sqrt{n}),$$

where with a slight notational abuse $\mathcal{T}_n$ denotes the index set of uncensored observations. Since $f_n(h) = g_n(h)/\int g_n(h)\,dh$, it suffices to prove that

$$(11) \qquad \int |g_n(h) - \varphi(h)\pi(\beta_0)|\,dh \to 0$$

with probability 1, where $\varphi(h) = \exp[-h^T I(\beta_0)h/2]$. The left-hand side of (11) can be decomposed into four pieces as follows:

$$\int |g_n(h) - \varphi(h)\pi(\beta_0)|\,dh$$

$$\leq \int_{|h| \leq B} |g_n(h) - \varphi(h)\pi(\beta_0)|\,dh + \int_{B < |h| < \sqrt{n}\delta} g_n(h)\,dh$$

$$+ \int_{|h| \geq \sqrt{n}\delta} g_n(h)\,dh + \int_{B < |h|} \varphi(h)\pi(\beta_0)\,dh,$$

where $\delta, B > 0$. Denote the four terms on the right-hand side of the above inequality by I_1, I_2, I_3 and I_4, respectively. We will prove that for any given $\varepsilon > 0$ there exist δ and B such that $I_i \leq \varepsilon$, $i = 1, 2, 3, 4$, for all but finitely many n with probability 1.

For I_1, let $l_n(\beta) = \log L_n(\beta)$. Taylor expansion yields that

$$(12) \qquad l_n(\hat{\beta} + h/\sqrt{n}) - l_n(\hat{\beta}) = \frac{1}{2}h^T \frac{l_n^{(2)}(\hat{\beta})}{n}h + R_n(h),$$

where

$$(13) \qquad R_n(h) = \frac{1}{\sqrt{n}} \sum_{i,j,k=1}^{p} \frac{h_i h_j h_k}{6} l_{nijk}^{(3)}(\tilde{\beta})/n$$

for $|\tilde{\beta} - \hat{\beta}| \leq |h|/\sqrt{n}$. Here, $l_n^{(2)}(\beta)$ is the Hessian matrix of $l_n(\beta)$ and $l_{nijk}^{(3)}(\beta) = \partial^3 l_n(\beta)/\partial\beta_i\,\partial\beta_j\,\partial\beta_k$. Using (12), we can write

$$\int_{|h| \leq B} |g_n(h) - \psi(h)\pi(\beta_0)|\,dh \leq r_n \pi(\beta_0) \int_{|h| \leq B} \psi(h)\,dh,$$

where

$$r_n = \sup_{|h| \leq B} \left| \exp\left(-\frac{1}{2}h^T\left(\frac{-l_n^{(2)}(\hat{\beta})}{n} - I(\beta_0) \right)h + R_n(h) \right)\pi(\hat{\beta} + h/\sqrt{n})/\pi(\beta_0) - 1 \right|.$$

Since $\{l_{nijk}^{(3)}(\tilde{\beta})/n\}$ are bounded sequences in n with probability 1 for all combinations of (i, j, k), we have $\sup_{|h| \leq B} R_n(h) \to 0$ with probability 1. The continuity of π at β_0 together with $\hat{\beta} \to \beta_0$ [Tsiatis (1981)] implies $\sup_{|h| \leq B} |\pi(\hat{\beta} + h/\sqrt{n})/\pi(\beta_0) - 1| \to 0$ with probability 1. Finally, since $-l_n^{(2)}(\hat{\beta})/n \to I(\beta_0)$ with probability 1 [Tsiatis (1981)], we have $r_n \to 0$ with probability 1. Hence, I_1 converges to 0 with probability 1 for any given $B > 0$ since $\int_{|h| \leq B} \psi(h)\, dh < \infty$.

For I_2, let M be a positive constant such that $\max_{i,j,k}\{l_{nijk}^{(3)}(\tilde{\beta})/n\} < M$ for all but finitely many n with probability 1. Existence of the constant M can be proved easily by the law of large numbers. Then for $|h| \leq \sqrt{n}\delta$ we have $|R_n(h)| \leq p^2 \delta M h^T h/6$, where $R_n(h)$ is defined in (13). Let η be the smallest eigenvalue of $I(\beta_0)$ and let $C_n = -l_n^{(2)}(\hat{\beta})/n - I(\beta_0)$. Then

$$l_n(\hat{\beta} + h/\sqrt{n}) - l_n(\hat{\beta}) = \frac{1}{2}h^T \frac{l_n^{(2)}(\hat{\beta})}{n}h + R_n(h)$$

$$\leq -\frac{1}{2}h^T(I(\beta_0) + C_n)h + p^2 \delta M h^T h/6$$

$$\leq -\frac{1}{2}(\eta - a_n - p^2 \delta M/6)h^T h,$$

where $a_n = p \sum_{i,j=1}^{p} |C_{nij}|$ and C_{nij} is the (i, j)th element of C_n. Note that $a_n \to 0$ with probability 1. Now, choose δ such that $\eta - a_n - p^2 \delta M/6 > \tau$ for some positive constant τ for all but finitely many n with probability 1. Let

$$\rho = \sup_{h:\, |h| \leq 2\delta} \left| \frac{\pi(\beta_0 + h)}{\pi(\beta_0)} \right|.$$

Since $\hat{\beta} \to \beta_0$,

$$\sup_{h:\, |h| \leq \sqrt{n}\delta} \left| \frac{\pi(\hat{\beta} + h/\sqrt{n})}{\pi(\beta_0)} \right| \leq \rho$$

for all but finitely many n with probability 1. For given $\varepsilon > 0$, choose B such that

$$\int_{|h|>B} \exp\left(-\frac{1}{2}\tau h^T h\right) dh < \frac{\varepsilon}{\rho\pi(\beta_0)}.$$

Then

$$\int_{B<|h|<\sqrt{n}\delta} g_n(h)\, dh \leq \pi(\beta_0) \int_{|h|>B} \exp\left(-\frac{1}{2}\tau h^T h\right) \frac{\pi(\hat{\beta} + h/\sqrt{n})}{\pi(\beta_0)}\, dh \leq \varepsilon$$

for all but finitely many n with probability 1.

For I_3, let

$$l(\beta) = \beta^T \mathrm{E}(Z_1 I(\delta = 1))$$

$$- \int_0^\tau \log\big(\mathrm{E}(\exp(\beta^T Z_1) I(T_1 \geq t))\big) \mathrm{E}(\exp(\beta_0^T Z_1) I(T_1 \geq t)) \, dA_0(t).$$

It is not hard to see that $l(\beta)$ is strictly concave and attains its maximum at β_0. Since $l_n(\beta)$ are also concave and $\sup_{|\beta - \beta_0| \leq \delta} |l_n(\beta)/n - l(\beta)| \to 0$ for any $\delta > 0$, $l_n(\hat\beta)/n \to l(\beta_0)$ with probability 1. Let $\varrho = \sup_{\{\beta : |\beta - \beta_0| = \delta/2\}} l(\beta) - l(\beta_0)$. Note that $\varrho < 0$. Let

$$r_n = \sup_{\{\beta : |\beta - \beta_0| = \delta/2\}} |l_n(\beta)/n - l(\beta)| + |l_n(\hat\beta)/n - l(\beta_0)|.$$

For sufficiently large n such that $r_n \leq -\varrho/2$, $\{\beta : |\beta - \hat\beta| \geq \delta\} \subset \{\beta : |\beta - \beta_0| \geq \delta/2\}$ and $n^{p/2} \exp(n\varrho/2) \leq \varepsilon$, we have

$$\int_{|h| \geq \sqrt{n}\delta} g_n(h) \, dh$$

$$= n^{p/2} \int_{|\beta - \hat\beta| \geq \delta} \exp(l_n(\beta) - l_n(\hat\beta)) \pi(\beta) \, d\beta$$

$$\leq n^{p/2} \sup_{\{\beta : |\beta - \beta_0| \geq \delta/2\}} \exp(l_n(\beta) - l_n(\hat\beta)) \int_{R^p} \pi(\beta) \, d\beta$$

$$\leq n^{p/2} \exp\bigg[n\bigg(\sup_{\{\beta : |\beta - \beta_0| = \delta/2\}} (l(\beta) - l(\beta_0)) + r_n\bigg)\bigg] \int_{\beta \in R^p} \pi(\beta) \, d\beta$$

$$\leq n^{p/2} \exp(n\varrho/2) < \varepsilon$$

with probability 1, where the second inequality is due to the concavity of $l(\beta)$.

For I_4, it is obvious that we can choose a large B such that $I_4 \leq \varepsilon$, and the proof is complete. $\square$

PROOF OF THEOREM 3. Let $\mathcal{T}_n(t) = \{T_i \leq t : \delta_i = 1\}$ and $\theta_n = (\sqrt{n}(\beta - \hat\beta) = x, Data)$. Write

$$(14) \qquad \sqrt{n}(A - \hat A) = \sqrt{n}(A(\cdot) - \mathrm{E}(A(\cdot)|\theta_n))$$

$$(15) \qquad\qquad + \sqrt{n}(\mathrm{E}(A(\cdot)|\theta_n) - \hat A(\cdot)).$$

We will prove that (14) converges weakly to $W(\Lambda_0(\cdot))$ and (15) converges to $x e_0(\cdot)$ with probability 1.

For (14), since it is a Lévy process, we utilize Theorem 19 of Section V.4 in Pollard (1984). Let $Y_n(t) = \sqrt{n}(A(t) - \mathrm{E}(A(t)|\theta_n))$. First, we show the convergence of finite-dimensional distributions by verifying Lyapounov's condition. Suppose $0 \leq s < t \leq \tau$ are given. Note that

$$Y_n(t) - Y_n(s) = \sum_{s < T_i \leq t, \delta_i = 1} \sqrt{n}(\Delta A(T_i) - \mathrm{E}(\Delta A(T_i)|\theta_n)).$$

Using the moments of gamma distributions and Lemma A.2 in Tsiatis (1981), we have

$$\sum_{s < T_i \leq t, \delta_i = 1} E\{\sqrt{n}(\Delta A(T_i) - E(\Delta A(T_i)|\theta_n))\}^4$$

$$= \int_{(s,t]} \frac{9n^2}{(\sum_{j \in R(u)} \exp((\hat{\beta} + x/\sqrt{n})^T Z_j))^4} \, dN(u) \to 0 \qquad \text{a.s.}$$

Again, Lemma A.2 in Tsiatis (1981) implies

(16)
$$\text{Var}(Y_n(t) - Y_n(s)|\theta_n)$$
$$= n \int_{(s,t]} \frac{dN(u)}{(\sum_{i \in R(u)} \exp((\hat{\beta} + x/\sqrt{n})^T Z_i))^2} \to \Lambda_0(t) - \Lambda_0(s) \qquad \text{a.s.}$$

These two facts together show that the finite-dimensional distributions of Y_n converge to those of $W(\Lambda_0(\cdot))$. For tightness, note that

$$P\{|Y_n(t) - Y_n(s)| \geq \varepsilon|\theta_n\} \leq \frac{1}{\varepsilon^2} \text{Var}(Y_n(t) - Y_n(s)|\theta_n).$$

By (16), we have

$$\text{Var}(Y_n(t) - Y_n(s)|\theta_n) = \Lambda_0(t) - \Lambda_0(s) + o(1)$$

with probability 1. Since $\Lambda_0(t)$ is continuous, we can make $P\{|Y_n(t) - Y_n(s)| \geq \varepsilon|\theta_n\}$ as small as possible for all sufficiently large n with probability 1 by choosing t and s sufficiently close. Hence, by Theorem 19 of Section V.4 in Pollard (1984), we conclude that Y_n converges weakly to $W(\Lambda_0)$ on $D[0, \tau]$ with probability 1.

For (15), the Taylor expansion implies that

$$\sqrt{n}(E(A(t)|\theta_n) - \hat{A}(t)) = x \int_0^t E(s : \beta_n(s)^*) \frac{dN(s)}{\sum_{i \in R(s)} \exp(\beta_n(s)^{*T} Z_i)},$$

where

$$E(s : \beta) = \frac{\sum_{i \in R(s)} Z_i \exp(\beta^T Z_i)}{\sum_{i \in R(s)} \exp(\beta^T Z_i)}$$

and $\beta_n(s)^*$ lies between $\hat{\beta}$ and $\hat{\beta} + x/\sqrt{n}$. Since $\sup_{t \in [0,\tau]} |\beta_n(t)^* - \beta_0| \leq x/\sqrt{n} \to 0$, using Lemma A.2 of Tsiatis (1981), it is not hard to see that

$$\sup_{t \in [0,\tau]} \left| E(t : \beta_n(t)^*) - \frac{S^1(t : \beta_0)}{S^0(t : \beta_0)} \right| \to 0$$

and

$$\sup_{t \in [0,\tau]} \left| \int_0^t \frac{dN(s)}{\sum_{i \in R(s)} \exp(\beta_n(s)^{*T} Z_i)} - A_0(t) \right| \to 0.$$

Hence, (15) converges to $x e_0(\cdot)$. $\square$

REFERENCES

ANDERSEN, P. K., BORGAN, Ø., GILL, R. D. and KEIDING, N. (1993). *Statistical Models Based on Counting Processes*. Springer, New York.

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

COX, D. R. (1972). Regression models and life tables (with discussion). *J. Roy. Statist. Soc. Ser. B* **34** 187–220.

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.

GILKS, W. R. and WILD, P. (1992). Adaptive rejection sampling for Gibbs sampling. *Appl. Statist.* **41** 337–348.

HJORT, N. L. (1990). Nonparametric Bayes estimators based on beta processes in models for life history data. *Ann. Statist.* **18** 1259–1294.

JACOBSEN, M. (1989). Existence and unicity of MLEs in discrete exponential family distributions. *Scand. J. Statist.* **16** 335–349.

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

KALBFLEISCH, J. D. (1978). Nonparametric Bayesian analysis of survival time data. *J. Roy. Statist. Soc. Ser. B* **40** 214–221.

KALBFLEISCH, J. D. and PRENTICE, R. L. (1980). *The Statistical Analysis of Failure Time Data*. Wiley, New York.

KIM, Y. and LEE, J. (2001). On posterior consistency of survival models. *Ann. Statist.* **29** 666–686.

LAUD, P. W., DAMIEN, P. and SMITH, A. F. M. (1998). Bayesian nonparametric and covariate analysis of failure time data. In *Practical Nonparametric and Semiparametric Bayesian Statistics. Lecture Notes in Statist.* **133** 213–225. Springer, New York.

LAZAR, N. A. (2000). Bayesian empirical likelihood. Technical Report 721, Dept. Statistics, Carnegie Mellon Univ.

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.

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

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

PETO, R. (1972). Contribution to the discussion of "Regression models and life tables," by D. R. Cox. *J. Roy. Statist. Soc. Ser. B* **34** 205–207.

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

PRÆSTGAARD, J. and WELLNER, J. A. (1993). Exchangeably weighted bootstraps of the general empirical process. *Ann. Probab.* **21** 2053–2086.

PRENTICE, R. L. and GLOECKLER, L. A. (1978). Regression analysis of grouped survival data with application to breast cancer data. *Biometrics* **34** 57–67.

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

SETHURAMAN, J. (1961). Some limit theorems for joint distributions. *Sankyhā Ser. A* **23** 379–386.

TSIATIS, A. A. (1981). A large sample study of Cox's regression model. *Ann. Statist.* **9** 93–108.

VOLINSKY, T. C., MADIGAN, D., RAFTERY, A. E. and KRONMAL, R. A. (1996). Bayesian model averaging in proportional hazard models: Assessing stroke risk. Technical Report 302, Dept. Statistics, Univ. Washington.

WELLNER, J. A. and ZHAN, Y. (1996). Bootstrapping Z-estimators. Technical Report 308, Dept. Statistics, Univ. Washington.

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

DEPARTMENT OF STATISTICS
EWHA WOMANS UNIVERSITY
11-1 DAEHYUN-DONG, SEODAEMUN-GU
SEOUL, 120-750
SOUTH KOREA
E-MAIL: ydkim@mm.ewha.ac.kr

DEPARTMENT OF STATISTICS
PENNSYLVANIA STATE UNIVERSITY
326 THOMAS BUILDING
UNIVERSITY PARK, PENNSYLVANIA 16802
E-MAIL: leej@stat.psu.edu