The Annals of Statistics
1986, Vol. 14, No. 1, 238–249

ESTIMATION OF SURVIVAL CURVES FROM DEPENDENT CENSORSHIP MODELS VIA A GENERALIZED SELF-CONSISTENT PROPERTY WITH NONPARAMETRIC BAYESIAN ESTIMATION APPLICATION[1]

BY WEI-YANN TSAI

Brookhaven National Laboratory

This article presents a family of estimators of the survival function based on right-censored observations which admit the possibility that the censoring variables may not be independent of the true failure variables. This family is obtained by generalizing the self-consistent property (Efron, 1967) of the product limit estimator (Kaplan and Meier, 1958). By assuming a Dirichlet process prior distribution of the observable random vectors, nonparametric Bayesian estimators of the survival curve—which is also a member of this family—are derived under a special loss function. These nonparametric Bayesian estimators generalize results of Susarla and Van Ryzin (1976), who impose a Dirichlet process prior on the failure survival function without considering any prior distribution of the censoring variables. Large sample properties of this family of nonparametric Bayesian estimators are also derived.

1. Introduction. Let $X_1^0, \dots, X_n^0$ be independent random variables, each sharing the same survival function $S^0(t) = P(X^0 > t)$ with the random variable X^0. The variables X_i^0, $i = 1, \dots, n$ present the true failure times on n individuals subject to right censoring. The observations consist only of independent random vectors (X_i, δ_i), $i = 1, \dots, n$, with the same distribution as (X, δ) where X is an observable random variable and

$$\delta = \begin{cases} 1 & \text{if } X = X^0, \\ 0 & \text{if } X < X^0. \end{cases} \tag{1.1}$$

The aim is to estimate S^0 from the data $(X_1, \delta_1), \dots, (X_n, \delta_n)$. This type of problem arises in many practical situations—such as cancer research, biomedical studies of survival, and life testing—and has been treated by a number of authors. [See Kalbfleisch and Prentice (1980) for a recent list of references.]

Kaplan and Meier (1958) suggested a product limit (PL) estimator for $S^0(\cdot)$ and showed that this estimator is in fact a maximum likelihood estimator. Their

Received September 1983; revised August 1985.

[1] The submitted manuscript has been authored under contract DE-AC02-76CH00016 with the U.S. Department of Energy. Accordingly, the U.S. Government retains a nonexclusive, royalty-free license to publish or reproduce the published form of this contribution, or allow others to do so, for U.S. Government purposes.

AMS 1980 *subject classifications*. Primary 62E20; secondary 62G05.

Key words and phrases. Generalized self-consistency, survival function, censored data, Dirichlet process, nonparametric Bayesian estimation, dependent censorship models.

estimator $\hat{S}^0_{\mathrm{PL}}(t)$ is given by

$$(1.2)\quad \hat{S}^0_{\mathrm{PL}}(t) = \begin{cases} 1 & t \le X^*_{(1)}, \\ \prod_{j=1}^{k} (1 - d_j/n_j) & X^*_{(k)} \le t < X^*_{(k+1)}, \quad k = 1, \dots, m-1, \\ 0 & X^*_{(m)} \le t, \end{cases}$$

where $X^*_{(1)} < X^*_{(2)} < \cdots < X^*_{(m)}$ represent m distinct observations among $X_1, \dots, X_n$, with

$$n_j = \sum_{i=1}^{n} I(X_i \ge X^*_{(j)}), \qquad j = 1, 2, \dots, m,$$

$$d_j = \sum_{i=1}^{n} I(X_i = X^*_{(j)}, \delta_i = 1), \qquad j = 1, 2, \dots, m,$$

and $I(\cdot)$ denotes the indicator function.

This product limit (PL) estimator has received a great amount of attention in recent years. Breslow and Crowley (1974), Földes and Rejtö (1981), and Meier (1975) based study of the properties of PL estimators on some continuity assumptions and special structure of the censoring mechanism. Langberg, Proschan, and Quinzi (1981) show that the PL estimator is strongly consistent in dependent random censorship models.

Efron (1967) established a property of the PL estimator, which he named the self-consistency property. An estimator $\hat{S}^0$ is said to be self-consistent if

$$(1.3)\qquad \hat{S}^0(t) = \sum_{i=1}^{n} \frac{I(X_i > t)}{n} + \sum_{X_i \le t} \frac{1 - \delta_i}{n} \frac{\hat{S}^0(t)}{\hat{S}^0(X_i)}.$$

That is the proportion estimated to survive past t is equal to the proportion of the subjects observed to survive past t plus the sum for all individuals censored before t, of the estimated conditional probability of surviving past t given survival to the censoring time.

In Section 2 we extend the definition (1.3) to a more general situation. Based on this extension, we derive a family of estimators. Large sample properties of these estimators are also derived, under weaker conditions that the conditions imposed by Breslow and Crowley (1974).

Susarla and Van Ryzin (1976) [hereafter referred to as SV (1976)] applied Dirichlet process priors of Ferguson (1973) to S^0 and obtained a nonparametric Bayesian estimator under a squared error loss function. They found that their Bayesian estimator reduces to the Kaplan–Meier PL estimator as the "prior sample size" tends to zero.

In Section 3 we derive a Bayesian type estimator from the generalized self-consistent property. We also show this Bayesian type estimator can be derived formally for a special loss function by use of Dirichlet process priors for the distribution of the random vector (X_i, δ_i). If one puts Dirichlet process prior only on the uncensored observation, then the above nonparametric Bayesian estimators reduces to the estimators of SV (1976).

We conclude with some extensions to more general cases.

2. A generalized self-consistency property. In order to unify the notation for discrete, continuous, and mixed cases of random variable X^0, we follow the notation of Kalbfleisch and Prentice (1980, pages 8–9) and define

$$\Lambda^0(t) = -\int_0^{t^+} \frac{DS^0(u)}{S^0(u^-)} \tag{2.1}$$

and

$$\gamma(\Lambda^0)(t) = \lim_{r\to\infty} \prod_{i=1}^{r} \left\{1 - \left[\Lambda^0(u_i) - \Lambda^0(u_{i-1})\right]\right\}, \tag{2.2}$$

where $0 = u_0 < u_1 < \cdots < u_r = t$, the limit $r \to \infty$ is taken as $u_k - u_{k-1} \to 0$; and the integral and operator "D" in (2.1) are Riemann–Stieltjes integral and differential operator, respectively. From definitions (2.1) and (2.2), we obtain

$$S^0(t) = \gamma(\Lambda^0)(t) = \exp\left(\oint_0^t \frac{DS^0(u)}{S^0(u)}\right) \prod_{s\le t}\left(1 - \frac{\Delta S^0(s)}{S^0(s^-)}\right),$$

where the integral $\oint_0^t DF/H$ means integration over the intervals of points less than t for which $F(\cdot)$ is continuous, and $\Delta F(s) = F(s^-) - F(s^+)$.

In estimating the survival function $S^0(t)$ in the presence of censoring, various authors such as Breslow and Crowley (1974) and Meier (1975) typically adopt, for mathematical simplicity, one of the following two censorship models:

(M.1) Independent random censorship models: There exist independent, identically distributed censoring random variables $C_1, \dots, C_n$ such that $X_i = \min(X_i^0, C_i)$, $\delta_i = I(X_i^0 \le C_i)$, and C_i and X_i^0 are independent.

(M.2) Fixed censorship models: There exist n constants $c_1, \dots, c_n$ such that $X_i = \min(X_i^0, c_i)$ and $\delta_i = I(X_i^0 \le c_i)$.

Furthermore, all authors make one of the following continuity assumptions for S^0 and S_c^0 (where S_c^0 is the survival function of C_i):

(C.1) The functions S^0 and S_c^0 have no common discontinuities.

(C.2) The function S^0 is absolutely continuous and/or S_c^0 is absolutely continuous.

Assumption (C.1) is obviously weaker than (C.2), but it need not hold in many practical situations of interest. As a matter of fact, the PL estimator $\hat{S}_{\mathrm{PL}}^0$ defined in (1.2) is adjusted for ties, and to the best of my knowledge, no rigorous proof of the consistency of $\hat{S}_{\mathrm{PL}}^0$ exists in the literature that omits assumption (C.1). In this section we show how a generalized self-consistent property can be used to obtain a family of consistent estimators of S^0 without making any of the assumptions (M.1), (M.2), (C.1), or (C.2).

Let

$$S_u(t) = P(X > t, \delta = 1), \tag{2.3}$$

$$S_c(t) = P(X > t, \delta = 0), \tag{2.4}$$

and

$$S(t) = P(X > t) = S_u(t) + S_c(t). \tag{2.5}$$

We now assume that the censoring mechanisms should satisfy

$$\text{(A.1)} \qquad \int_0^t P(X^0 > t | X = x, \delta = 0)\, DS_c(x^-) = \int_0^t S^0(t)/S^0(x)\, DS_c(x^-)$$

for $0 \leq x \leq t < \infty$.

From Lemma 2.2 and Definition 2.4 below, it is natural to require (A.1) in order to obtain consistent estimators of S^0 by using self-consistency approach. There are other conditions suggested by other authors from different approaches. The detailed comparison is postponed to the Remark 2.10.

It can be readily seen that (A.1) holds under the models (M.1) and (M.2). Moreover, even if the censoring random variable C_i under model (M.1) is not independent of the failure random variable X_i^0, in certain cases (A.1) may still hold. We illustrate with an example.

EXAMPLE 2.1. Let (X_i^0, C_i), $i = 1, \ldots, n$ be independent, identically distributed random vectors having the bivariate exponential distribution of Marshall and Olkin (1967) with the survival function

$$\begin{aligned} S(t_1, t_2) &= P(X_i^0 > t_1, C_i > t_2) \\ &= \exp(-\lambda_1 t_1 - \lambda_2 t_2 - \lambda_{12} \max(t_1, t_2)) \end{aligned} \tag{2.6}$$

for $t_1, t_2 \geq 0$ and $\lambda_1, \lambda_2, \lambda_{12} > 0$. Straightforward computations show that

$$\begin{aligned} P(X_i^0 > t | X_i = x, \delta_i = 0) &= P(X_i^0 > t | X_i^0 > C_i = x) \\ &= \exp(-(\lambda_1 + \lambda_{12})(t - x)) \\ &= S(t,0)/S(x,0); \end{aligned}$$

hence (A.1) holds.

LEMMA 2.2. *Let S_u and S_c be as defined above. Then*

$$S^0(t) = S_u(t) + S_c(t^*) - \int_0^{t^*} P(X^0 > t | X = x, \delta = 0)\, DS_c(x^-) \qquad \textit{for } t \geq 0, \tag{2.7}$$

where $t^* = t^+$ or t^-.

PROOF.

$$\begin{aligned} S^0(t) &= E(I(X^0 > t)) = E\big(E\big(I(X^0 > t)|(X,\delta)\big)\big) \\ &= \int_0^\infty E\big(I(X^0 > t)|X = x, \delta = 1\big)D(1 - S_u(x^-)) \\ &\quad + \int_{t^*}^\infty E\big(I(X^0 > t)|X = x, \delta = 0\big)D(1 - S_c(x^-)) \\ &\quad + \int_0^{t^*} E\big(I(X^0 > t)|X = x, \delta = 0\big)D(1 - S_c(x^-)) \\ &= S_u(t) + S_c(t^*) - \int_0^{t^*} P(X^0 > t|X = x, \delta = 0)\, DS_c(x^-). \square \end{aligned}$$

THEOREM 2.3. *If and only if* (A.1) *holds, then*

$$S^0(t) = S_u(t) + S_c(t^*) - \int_0^{t^*} \frac{S^0(t)}{S^0(x)} DS_c(x^-) \quad \textit{for } t \geq 0, \tag{2.8}$$

where $t^* = t^+$ *or* t^-.

PROOF. Lemma 2.2 implies the result. □

Theorem 2.3 has been derived by Tsai and Crowley (1985) under models (M.1), and served as an important equation for studying large sample properties of PL estimators.

Now we define a generalized self-consistency property of an estimator $\hat{S}$ of S^0.

DEFINITION 2.4. An estimator $\hat{S}^0$ of S^0 is said to have a generalized self-consistency property if and only if there exist consistent estimators $\hat{S}_u$ and $\hat{S}_c$, respectively, of S_u and S_c, such that

$$\hat{S}^0(t) = \hat{S}_u(t) + \hat{S}_c(t) - \int_0^{t^+} \frac{\hat{S}^0(t)}{\hat{S}^0(x)} D\hat{S}_c(x^-) \quad \text{for } t \geq 0. \tag{2.9}$$

Let $\hat{S}_u(t)$ and $\hat{S}_c(t)$ be the empirical subsurvival functions of S_u and S_c, respectively, so that $\hat{S}_u(t) = S_u^e(t) = n^{-1}\sum_{i=1}^n I(X_i > t, \delta_i = 1)$ and $\hat{S}_c(t) = S_c^e(t) = n^{-1}\sum_{i=1}^n I(X_i > t, \delta_i = 0)$. Then (2.9) reduces to (1.3), which is the definition of the self-consistency property of $\hat{S}^0$ given by Efron (1967).

THEOREM 2.5. *Let* $T = \sup\{t|S(t) > 0\}$. *Then the unique solution of* (2.7) *or* (2.8) *for* $t < T$ *has the following explicit expression*:

$$\begin{aligned} S^0(t) &= \gamma\left(-\int_0^{T^+} \frac{DS_u(x)}{S(x^-)}\right)(t) \\ &= \exp\left(\oint_0^t \frac{DS_u(x)}{S(x)}\right) \prod_{x \leq t}\left(1 - \frac{\Delta S_u(x)}{S(x^-)}\right). \end{aligned} \tag{2.10}$$

PROOF. From (2.8), we obtain

$$\begin{aligned}0 &= -DS^0(t) + DS_u(t) + DS_c(t^-) - \left(DS^0(t)\right)\int_0^t \frac{1}{S^0(x)}\,DS_c(x^-)\\ &\quad - \frac{S^0(t)}{S^0(t)}\,DS_c(t^-)\\ &= -DS^0(t)\left(1 + \int_0^t \frac{1}{S^0(x)}\,DS_c(x^-)\right) + DS_u(t).\end{aligned}$$

For $t < T$, $S(t^-) \neq 0$ implies $S^0(t^-) \neq 0$; therefore we have

$$\frac{DS^0(t)}{S^0(t^-)}\left(S^0(t^-) + \int_0^{t^-} \frac{S^0(t^-)}{S^0(x)}\,DS_c(x^-)\right) = DS_u(t).$$

Thus, by using (2.8),

$$\frac{DS^0(t)}{S^0(t^-)} = \frac{DS_u(t)}{S_u(t^-) + S_c(t^-)}.$$

Hence

$$\Lambda^0(t) = -\int_0^{t^+} \frac{DS^0(x)}{S^0(x^-)} = -\int_0^{t^+} \frac{DS_u(x)}{S(x^-)} \tag{2.11}$$

or

$$\begin{aligned}S^0(t) &= \gamma(\Lambda^0)(t) = \gamma\left(-\int_0^{\tau^+} \frac{DS_u(x)}{S(x^-)}\right)(t)\\ &= \exp\left(\oint_0^t \frac{DS_u(x)}{S(x)}\right)\prod_{x\le t}\left(1 - \frac{\Delta S(x)}{S(x^-)}\right). \square\end{aligned}$$

The following two corollaries are direct results of Theorems 2.3 and 2.5.

COROLLARY 2.6. *If and only if* (A.1) *holds, then the subsurvival functions* $S_u(\cdot)$, $S_c(\cdot)$ *determine the survival function* $S^0(t)$ *according to expression* (2.10).

COROLLARY 2.7. $\hat{S}^0(t)$ *is a generalized self-consistent estimator if and only if there exist consistent estimators* $\hat{S}_u(\cdot)$ *and* $\hat{S}_c(\cdot)$, *respectively, of* $S_u(\cdot)$ *and* $S_c(\cdot)$ *such that*

$$\hat{S}^0(t) = \gamma\left(-\int_0^{\tau^+} \frac{D\hat{S}_u(x)}{\hat{S}_u(x^-) + \hat{S}_c(x^-)}\right)(t).$$

Let ψ be a family of estimators $(\hat{S}_u(\cdot), \hat{S}_c(\cdot))$ of $(S_u(\cdot), S_c(\cdot))$ such that

$$\begin{aligned}&\sup_{0\le t<T} |\hat{S}_u(t) - S_u(t)| = O\left(\sqrt{\frac{\log\log n}{n}}\right) \quad \text{a.e.,}\\ &\sup_{0\le t<T} |\hat{S}_c(t) - S_c(t)| = O\left(\sqrt{\frac{\log\log n}{n}}\right) \quad \text{a.e.}\end{aligned} \tag{A.2}$$

and

$\sqrt{n}((\hat{S}_u, \hat{S}_c) - (S_u, S_c))$ converges weakly to a bivariate Gaussian process (X, Y) which has mean $(0, 0)$ and a covariance structure given for $s \le t < T$ by

$$\begin{aligned} \operatorname{Cov}(X(s), X(t)) &= [1 - S_u(s)] S_u(t), \\ \operatorname{Cov}(X(s), Y(t)) &= -S_u(s) S_c(t), \\ \operatorname{Cov}(Y(s), X(t)) &= -S_c(s) S_u(t), \\ \operatorname{Cov}(Y(s), Y(t)) &= [1 - S_c(s)] S_c(t). \end{aligned} \tag{A.3}$$

Define a family ψ^0 of estimators of S^0 by

$$\psi^0 = \left\{ \hat{S}^0 | \hat{S}^0(t) = \gamma\left(-\int_0^{T^+} \frac{D\hat{S}_u(x)}{\hat{S}_u(x^-) + \hat{S}_c(x^-)} \right)(t), (\hat{S}_u, \hat{S}_c) \in \psi \right\}.$$

By slightly modifying Theorem 4.4 of Tsai and Crowely (1985), the following large sample properties of $\hat{S}^0 \in \psi^0$ are established.

THEOREM 2.8. *If* (A.1) *holds, then the following two properties hold for every* $\hat{S}^0 \in \psi^0$:

(i) $$\sup_{0 \le t \le T^* < T} |\hat{S}^0(t) - S^0(t)| = O\left(\sqrt{\frac{\log \log n}{n}} \right) \quad a.e. \text{ as } n \to \infty,$$

(ii) $\sqrt{n}(\hat{S}^0 - S^0)$ *converges weakly to a Gaussian process* Z *with mean* 0 *and*

$$\operatorname{Cov}(Z(s), Z(t)) = S^0(s) S^0(t) \int_0^s S^{-2}(x-)\, DS_n(x) \quad \text{for } s \le t \le T^* < T.$$

PROOF. Since the only two properties of S_u^e and S_c^e used in Theorem 4.4 of Tsai and Crowley (1985) are properties (A.2) and (A.3), therefore in replacing (S_u^e, S_c^e) by $(\hat{S}_u, \hat{S}_c) \in \psi$, the whole proof here may be carried through the same way. □

REMARK 2.9. Under (M.1) and (C.1), Peterson (1977) derived a unique expression for S^0 in terms of S_u and S_c which is equivalent to formula (2.10). In addition, formula (2.11) is a well known result under (M.1) and (C.2) (Breslow and Crowley, 1974).

REMARK 2.10. For the special case in which $S_u = S_u^e$ and $S_c = S_c^e$, Efron (1967) derives the result of Theorem 2.5 by mathematical induction.

REMARK 2.11. In the literature, there are quite a few mathematical formulations of what restrictions should be placed on the censoring mechanism so that the standard methods of analysis are appropriate. Williams and Lagakos (1977) derived constant-sum models from the likelihood function. A model for right censored survival data is of the constant-sum type if and only if

$$\text{(A.4)} \qquad \frac{dS_u(t)}{dS(t)} - \int_0^t \frac{dS_c(x)}{S(x)} = 1.$$

A special model of the constant-sum type is a survival independent censoring model introduced by Williams and Lagakos (1977) which satisfies

$$\text{(A.5)} \qquad P\big(X_i^0 > t | X_i = x, \delta_i = 0\big) = \frac{S^0(t)}{S^0(x)} \quad \text{for } 0 \le x \le t < \infty;$$

that is, censoring at time t carries the same information as survival beyond time x. Since (A.5) implies (A.1), (A5) is therefore also a sufficient condition for Theorem 2.3.

Another formulation, outlined by Cox (1975) and more formally defined by Andersen and Gill (1982), is the model which satisfies

$$\text{(A.6)} \qquad \int_0^t \frac{dS_u(x)}{S(x^-)} = \int_0^t \frac{dS^0(x)}{S^0(x^-)} \quad \text{for } 0 \le t < \infty,$$

that is, the failure rate of an item on test at time t should be unaltered by the censoring that has taken place. Kalbfleisch and MacKay (1979) proved that the constant-sum model (A.4) is equivalent to (A.6). It can be readily proved that (A.1) is also equivalent to (A.6) by following a similar line of proof as for Theorem 2.5.

Another formulation was due to Langberg, Proschan, and Quinzi (1981), who under assumption (C.1), derived a necessary and sufficient condition of Corollary 2.6 as follows:

$$\text{(A.7)} \qquad \frac{S(t)}{S(t^-)} = \begin{cases} \dfrac{S^0(t)}{S^0(t^-)} & \text{for every jump point } t \text{ of } S_u, \\ 1 & \text{otherwise,} \end{cases}$$

and

$$P\big(C \ge t | X^0 = t\big) = P\big(C > t | X^0 > t\big).$$

They also proved the strong consistency of the PL estimator, under (A.7) and (C.1). It is conjectured that (A.7) is equivalent to (A.6) under assumption (C.1). In certain realistic situations, assumption (C.1) may not always hold; then (A.1) or (A.6) is preferable, and the results of Theorem 2.6 are stronger.

REMARK 2.12. Recently, Robertson and Uppuluri (1984) [hereafter referred to as RU (1984)] generalized the PL estimator by using the idea of redistribution of mass to the right which was first considered by Efron (1967). Their generalization has strong connection with (2.7). Let $P(X^0 > X_{(j)} | X = X_{(i)}, \delta = \delta_{(i)}) = W_{ij}$, where $X_{(1)} < \cdots < X_{(n)}$ are the order statistics of $X_1, \ldots, X_n$, and $(W_{ij})_{n \times n}$ be the RR matrix defined by RU (1984, page 368). Then the estimator $\hat{S}^0$ obtained from the procedure 1 of RU (1984, page 369) is a solution of

$$\hat{S}^0(t) = S_u^e(t) + S_c^e(t) - \int_0^{t+} P\big(X^0 > t | X = x, \delta = 0\big)\, DS_c^e(x),$$

although their generalization is restricted to discrete estimators and cannot guarantee to obtain a consistent estimator of S^0.

3. Nonparametric Bayesian estimator of S^0. Let P be a random probability measure on (Ω, Π), where $\Omega = R^+ \times \{0,1\}$, $\Pi = B \times C$, B is the σ-field of Borel sets restricted to R^+, and $C = \{\phi, \{0\}, \{1\}, \{0,1\}\}$. Let α^* be a nonnull finite measure on (Ω, Π). Furthermore, assume the random measure P to be a Dirichlet process on (Ω, Π) with parameter α^* and $(X_1, \delta_1), \dots, (X_n, \delta_n)$ to be a random sample of size n from this Dirichlet process P. Our purpose is to estimate the survival function S^0 from a Bayesian point of view. [For the definition of the Dirichlet process and some basic results, see Ferguson (1973).]

The nonparametric Bayes estimators of $S_u(t)$ and $S_c(t)$ are, respectively,

$$\tilde{S}_u(t) = \frac{\alpha^*((t,\infty),\{1\}) + \sum_{i=1}^n I(X_i > t, \delta_i = 1)}{\alpha^*(\Omega) + n},$$

and

$$\tilde{S}_c(t) = \frac{\alpha^*((t,\infty),\{0\}) + \sum_{i=1}^n I(X_i > t, \delta_i = 0)}{\alpha^*(\Omega) + n},$$

under the squared error loss function

$$L(\tilde{S}_*, S_*) = \int_0^\infty [\tilde{S}_*(u) - S_*(u)]^2 \, dw(u),$$

where $w(\cdot)$ is a weight function, $\hat{S}_*(u)$ is an estimator of $S_*(u)$, and $* = u$ or c.

Therefore, we may derive a self-consistent estimator $\tilde{S}^0(\cdot)$ of $S^0(\cdot)$ from $\tilde{S}_u(\cdot)$ and $\tilde{S}_c(\cdot)$ by

$$\begin{aligned} \tilde{S}^0(t) &= \gamma\left(-\int_0^{\tau^+} \frac{D\tilde{S}_u(x)}{\tilde{S}_u(x^-) + \tilde{S}_c(x^-)} \right)(t) \\ &= \gamma\left(-\int_0^{\tau^+} \frac{D(\alpha^*((x,\infty),\{1\}) + \sum_{i=1}^n I(X_i > x, \delta_i =))}{\alpha^*([x,\infty),\{0,1\}) + \sum_{i=1}^n I(X_i \geq x)} \right)(t). \end{aligned} \tag{3.1}$$

It is easy to show $\tilde{S}_u(\cdot)$ and $\tilde{S}_c(\cdot)$ satisfy conditions (A.2) and (A.3); therefore, the following results follow.

THEOREM 3.1. *If assumption* (A.1) *holds, then*

(*i*) $\sup_{0 \leq t \leq T^* < T} |\tilde{S}^0(t) - S^0(t)| = O\left(\sqrt{\frac{\log\log n}{n}}\right)$ *a.s.*,

(*ii*) $\sqrt{n}(\tilde{S}^0 - S^0)$ *converges weakly to a Gaussian process Z, as defined in Theorem* 2.8.

REMARK 3.2. It can be shown that $\tilde{S}^0$ is the Bayesian estimator of S^0 under the loss function $L(\tilde{S}^0, S^0) = \int_0^\infty [\gamma^{-1}(\tilde{S}^0)(t) - \gamma^{-1}(S^0)(t)]^2 \, Dw(t)$, where γ^{-1} denotes the inverse operator of γ. [For a proof, see Tsai (1983).]

REMARK 3.3. When $\alpha^*(R^+, \{0\}) = 0$ and $\alpha^*((t, \infty), \{1\}) = \alpha(t, \infty)$, then $\tilde{S}^0(t)$ defined in (3.1) reduces to

$$\tilde{S}^0(t) = \frac{\alpha(t, \infty) + \sum_{j=1}^n I(x_j > t)}{\alpha(R^+) + n} \prod_{i | X^*_{(i)} \le t} \frac{\alpha(X^*_{(i)}, \infty) + n_i - d_i}{\alpha(X^*_{(i)}, \infty) + n_{i+1}}, \tag{3.2}$$

where $X^*_{(i)}$, n_i, and d_i, $i = 1, \ldots, m$, are defined in (1.2). If $\alpha(t, \infty)$ is a continuous function in R^+ and $S^e_u(\cdot)$ and $S^e_c(\cdot)$ do not have any common discontinuities, (3.2) will be a version of the formula derived by Susarla and Van Ryzin (1976) with the following main differences:

(i) If $S^e_u(\cdot)$ and $S^e_c(\cdot)$ have common jump points, then formula (3.1), as well as formula (3.2), will reduce to the $\hat{S}_{\text{PL}}$ defined in (2.1) as $\alpha^*(\Omega) \to 0$, but the formula derived by SV (1976) will not.

(ii) It should be clear that our censoring scheme is more general than the censoring models considered by SV (1976).

(iii) In the present paper, it is assumed that the Dirichlet process prior is given to the probability measure of the random vector (X_i, δ_i), whereas in SV (1976) the prior is incorporated in the survival function S^0 of X^0 and they do not consider any prior in the distribution of the censoring variable.

(iv) The estimator $\tilde{S}^0$ is derived with respect to squared error loss on $\gamma^{-1}(S^0)$, while the result of SV (1976) is derived with respect to squared error loss on S^0.

4. Discussion. In this paper we have presented a unified approach to estimating the survival function of right censored data which combine the results of Efron (1967), Breslow and Crowley (1974), Meier (1975), Susarla and Van Ryzin (1976), Peterson (1977), and Langberg, Proschan, and Quinzi (1981). We only consider "exclusive censoring," where the censoring observation is of the type $X_i^0 > X_i$. There are no conceptual difficulties in extending this method to "inclusive censoring" problems (where the censoring observations are of the type $X_i^0 \ge X_i$), to "doubly censoring" problems (where the censoring observations are of the type $X_i^0 > X_i$ or $X_i^0 < X_i$), to the competing risk problem, or to other incomplete observation problems.

In Section 1 we assume that $(X_1, \delta_1), \ldots, (X_n, \delta_n)$ are independent, identically distributed random vectors. The results of Sections 2 and 3 still hold even if this assumption is weakened so that $(X_1, \delta_1), \ldots, (X_n, \delta_n)$ are independent random vectors, not necessarily identically distributed.

The proofs of these results proceed along the lines of the proofs given in Sections 2 and 3, so we will omit most of the details. The main chore remaining to complete the proof is to establish (2.8) with some modification.

Let

$$S_u^i(t) = P(X_i > t, \delta_i = 1) \quad \text{for } i = 1, 2, \ldots, n,$$

$$S_c^i(t) = P(X_i > t, \delta_i = 0) \quad \text{for } i = 1, 2, \ldots, n,$$

$$\bar{S}_u(t) = \frac{1}{n} \sum_{i=1}^n S_u^i(t),$$

and

$$\bar{S}_c(t) = \frac{1}{n} \sum_{i=1}^{n} S_c^i(t).$$

If

$$\int_0^t P(X_i^0 > t | X_i = x, S_i = 1)\, dS_c^i(x) = \int_0^t \frac{S^0(t)}{S^0(x)}\, DS_c^i(x^-)$$

for $i = 1, \dots, n$, then we have

$$\begin{aligned} S^0(t) &= \frac{1}{n} E \sum_{i=1}^{n} I(X_i^0 > t) \\ &= \frac{1}{n} \sum_{i=1}^{n} \left\{ S_u^i(t) + S_c^i(t) - \int_0^t \frac{S^0(t)}{S^0(x)}\, DS_c^i(x) \right\} \\ &= \bar{S}_u(t) + \bar{S}_c(t) - \int_0^t \frac{S^0(t)}{S^0(x)}\, D\bar{S}_c(x). \end{aligned}$$

Therefore, Theorems 2.7 and 3.1 still hold as long as conditions (A.2) and (A.3) are satisfied when S_u and S_c are replaced by $\bar{S}_u$ and $\bar{S}_c$, respectively.

Acknowledgment. The author wishes to thank John Van Ryzin for helpful discussions and the referee and associate editor for insightful comments and suggestions.

REFERENCES

ANDERSEN, P. K. and GILL, R. D. (1982). Cox's regression model for counting processes: a large sample study. *Ann. Statist.* **10** 1100–1120.

BRESLOW, N. E. and CROWLEY, J. (1974). A large sample study of the life table and product limit estimates under random censorship. *Ann. Statist.* **13** 437–453.

COX, D. R. (1975). Partial likelihood. *Biometrika* **62** 269–276.

EFRON, B. (1967). The two sample problem with censored data. *Proc. Fifth Berkeley Symp. Math. Statist. Prob.* 4 831–853.

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

FÖLDES, A. and REJTŐ, L. (1981). Strong uniform consistency for nonparametric survival curve estimators from randomly censored data. *Ann. Statist.* **9** 122–129.

KALBFLEISCH, J. D. and MACKAY, R. J. (1979). On constant-sum models for censored survival data. *Biometrika* **66** 87–90.

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

KAPLAN, E. L. and MEIER, P. (1958). Nonparametric estimation from incomplete observations. *J. Amer. Statist. Assoc.* **53** 87–90, 457–481.

LANGBERG, N., PROSCHAN, F. and QUINZI, A. J. (1981). Estimating dependent life lengths, with applications to the theory of competing risks. *Ann. Statist.* **9** 157–167.

MARSHALL, A. W. and OLKIN, I. (1967). A multivariate exponential distribution. *J. Amer. Statist. Assoc.* **62** 30–44.

MEIER, P. (1975). Estimation of a distribution function from incomplete observations. In *Perspectives in Probability and Statistics* (J. Gani, ed.) 67–87. Academic, New York.

PETERSON, A. V. (1977). Expressing the Kaplan–Meier estimator as a function of empirical sub-survival functions. *J. Amer. Statist. Assoc.* **72** 854–858.

ROBERTSON, J. B. and UPPULURI, V. R. R. (1984). A generalized Kaplan–Meier estimator. *Ann. Statist.* **12** 366–371.

SUSARLA, V. and VAN RYZIN, J. (1976). Nonparametric Bayesian estimation of survival curves from incomplete observations. *J. Amer. Statist. Assoc.* **71** 897–902.

TSAI, W. Y. (1983). Estimation of survival curves from dependent censorship models via a generalized self-consistent property with nonparametric bayesian estimation application. Technical Report 33642, Brookhaven National Laboratory.

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.

WILLIAMS, J. A. and LAGAKOS, S. W. (1977). Models for censored survival analysis: constant sum and variable-sum models. *Biometrika* **64** 215–224.

APPLIED MATH DEPARTMENT
BROOKHAVEN NATIONAL LABORATORY
UPTON, NEW YORK 11973