Bayesian Nonparametric Estimation in a Series System or a Competing-Risks Model

Victor H. Salinas–Torres[1], Carlos A.B. Pereira[2] and Ram C. Tiwari[3]

U. of Santiago of Chile, U. of São Paulo, U. of North Carolina at Charlotte and National Cancer Institute

Abstract

This article presents a Bayesian nonparametric approach to the estimation of a system and its components' survival functions arising from observing the failure of a series system or a competing risk model. A Dirichlet multivariate process is used as a prior for the vector of the components' random subsurvival function to derive Bayes estimator of the survival function when the cause of failure belongs to a certain risk subset. This is done as follows. First, Peterson's formula is evaluated using the Bayes estimators of the subsurvival functions corresponding to the risk subset, to obtain a plugged-in nonparametric estimator of the survival function associated with the risk subset. Then, using the product-integration approach, it is proved that this nonparametric estimator is in fact the Bayes estimator of the survival function corresponding to the risk subset under quadratic loss function and the Dirichlet multivariate process. The weak convergence and the strong consistency of the estimator is established. The special case when the system has only two components corresponds to well studied randomly censored model.

Key Words and Phrases. Dirichlet processes, Dirichlet multivariate processes, randomly censored data, cumulative hazard rate, product-integration.

AMS (1991) subject classifications. Primary 62G05 – 62A15; Secondary 60K99.

Revised Version. November, 2001

---

[1]Department of Mathematics, U. of Santiago of Chile. P.O. Box 307/2, Santiago, Chile.

[2]Department of Statistics, U. of São Paulo. P.O. Box 66281, 05315-970, São Paulo-SP, Brasil.

[3]Department of Mathematics, U. of North Carolina at Charlotte, NC28223, and National Cancer Institute, DCCPS/SRAB, 6116 Executive Blvd., Room 5039, Bethseda, Maryland, USA.

1 Introduction

Consider a series system with r $(r > 1)$ components or a competing-risks model with r sources of failure. Let X_j $(j = 1, \ldots, r)$ denote the failure time of the j-th component or source of failure, where $X_1, \ldots, X_r$ are independent random variables. When the system (or organism) fails , the observed random vector is (Z, δ), where $Z = \min(X_1, \ldots, X_r)$, and $\delta = j$ if $Z = X_j$, $j = 1, \ldots, r$.

1.1 Survival and Subsurvival Functions

The survival function of the j-th component is denoted by

$$S_j(t) = \Pr(X_j > t) \quad j = 1, \ldots, r .$$

Let

$$S_j^*(t) = \Pr(Z > t, \delta = j) ,$$

be the subsurvival function, called the cumulative incidence function (CIF), of the j-th component $(j = 1, \ldots, r)$. Then, the system survival function is given by

$$S(t) = \Pr(Z > t) = \sum_{j=1}^{r} S_j^*(t) .$$

Let Δ be a nonempty subset of $\{1, \ldots, r\}$, called a risk subset, and denote by Δ^c its complement. Corresponding to Δ, we define the subsurvival and survival functions

$$S_\Delta^*(t) = \Pr(Z > t, \delta \in \Delta) ,$$

and

$$S_\Delta(t) = \Pr\left(\min_{j \in \Delta} X_j > t\right) ,$$

respectively.

1

1.2 Cumulative Hazard Rate and Cause Specific Hazard Rate

Some notions related to hazard failure rate are explained in the context of risk subsets; however, we obtain their usual notions when the risk set consists of a singleton.

The cumulative hazard rate Λ_Δ for the survival function S_Δ or random variable $Y_\Delta = \min_{j \in \Delta}(X_j)$ is a nonnegative, nondecreasing, right-continuous function on $[0, \infty)$ satisfying

$$
\begin{aligned}
d\Lambda_\Delta(s) &= \Lambda_\Delta[s, s+ds) = Pr\{Y_\Delta \in [s, s+ds)|Y_\Delta \geq s\} \\
&= -dS_\Delta(s)/S_\Delta(s^-).
\end{aligned}
$$

So that

$$
\Lambda_\Delta(t) = \int_0^t \frac{-dS_\Delta(s)}{S_\Delta(s^-)}, \tag{1.1}
$$

and

$$
\begin{aligned}
S_\Delta(t) &= \int_0^t -S_\Delta(s^-)d\Lambda_\Delta(s) \\
&= \prod_{[0,t]} (1 - d\Lambda_\Delta(s)), \ t \geq 0, \tag{1.2}
\end{aligned}
$$

where $\prod_{[0,t]}$ is the product-integral. If $\Lambda_\Delta(\cdot)$ is continuous, $\prod_{[0,t]}(1-d\Lambda_\Delta(s)) = exp(-\Lambda_\Delta(t))$; otherwise

$$
\prod_{[0,t]} (1 - d\Lambda_\Delta(s)) = \exp(-\Lambda_\Delta^c(t)) \prod_{i:t_i \leq t} (1 - \Lambda_\Delta\{t_i\}) ,
$$

where $t_1, t_2, \ldots$, are the discontinuity points of Λ_Δ and Λ_Δ^c is the continuous part of Λ_Δ, i.e., $\Lambda_\Delta^c(t) = \Lambda_\Delta(t) - \sum_{i:t_i \leq t} \Lambda_\Delta\{t_i\}$. For a complete review of product-integration in survival analysis see, for example, Gill and Johansen (1990).

If S_Δ and S_Δ^c have no common discontinuities, (1.1) becomes

$$
\Lambda_\Delta(t) = \int_0^t \frac{-dS_\Delta^*(s)}{S(s^-)}. \tag{1.3}
$$

The function $(-dS_\Delta^*/S)$ is called the cause-specific hazard rate of the risk subset Δ.

1.3   Peterson's Formula

Peterson's formula (Peterson, 1977) expresses the survival function associated with the risk subset Δ as a functional of the subsurvival functions associated with Δ and Δ^c. The formula is given by

$$S_\Delta(t) = \varphi(S_\Delta^*(\cdot), S_{\Delta^c}^*(\cdot); t) , \quad \text{for } t \le t^* = \min(t_{S_\Delta}, t_{S_{\Delta^c}}) , \tag{1.4}$$

where

$$\varphi(F(\cdot), G(\cdot); t) = \exp\left\{\oint_0^t \frac{dF(s)}{F(s) + G(s)}\right\} \prod_t \frac{F(s_+) + G(s_+)}{F(s_-) + G(s_-)} ,$$

$t_{S_\Delta} = \sup\{t : S_\Delta(t) > 0\}$, and $\oint_0^t$ is the integral over the union of intervals of points less than t for which $F(\cdot)$ is continuous. $\prod_t$ indicates the product over the set $\{s \le t : s$ is a jump point of $F\}$. For (1.4) to be well defined we assume that S_Δ and S_{Δ^c} have no common discontinuities. If S_Δ and S_{Δ^c} are continuous, then (1.4) is analogous to equation (7.5) in Breslow and Crowley (1974):

$$S_\Delta(t) = \exp\left[\int_0^t \frac{dS_\Delta^*(s)}{S_\Delta^*(s) + S_{\Delta^c}^*(s)}\right] .$$

1.4   Objectives

Let $(X_{i1}, \ldots, X_{ir})$, $i = 1, \ldots, n$, be n independent latent or imaginary observations on $(X_1, \ldots, X_r)$. The actual observations consist of pairs (Z_i, δ_i), $i = 1, \ldots, n$, which forms a sample on (Z, δ).

The main objective of this paper is to derive Bayes estimator of S_Δ from the data $(Z_1, \delta_1), \ldots, (Z_n, \delta_n)$.

3

When $r = 2$, the risk subset is the singleton subset $\Delta = \{1\}$ and $\Delta^c = \{2\}$, the data $(Z_1, \delta_1), \ldots, (Z_n, \delta_n)$ correspond to a random (right) censored model and the estimation of $S_1(t) = \Pr(X_1 > t)$ has been considered by a number of authors. For example, Kaplan and Meier (1958) derived the product-limit (PL) estimator for S_1 and showed that it is in fact a maximum likelihood estimator. Breslow and Crowley (1974) studied several properties of the PL estimator.

Susarla and Van Ryzin (1976) considered the Dirichlet process prior (Ferguson, 1973) for S_1 and obtained a nonparametric Bayes estimator (denoted hereafter as SV estimator). The SV estimator reduces to PL estimator as the "prior sample size" tends to zero. Susarla and Van Ryzin (1978) studied the large sample properties of the SV estimator. For a survey of work on the Bayesian estimation of survival function using Dirichlet processes, see Ferguson, Phadia and Tiwari (1992).

Other works consider different approaches. Hjort (1990) studied the problem of finding Bayes estimators for cumulative hazard rates and related parameters considering a class of Beta processes as a prior distribution. Bayesian inference for a weigthed distribution model is considered by Lo (1993) using Dirichlet processes defined in terms of Gamma processes.

This paper has a special interest in the general case $r \geq 2$, i.e, a series system or a competing-risks model.

In Section 2, we use Dirichlet multivariate processes, introduced by Salinas-Torres, Pereira and Tiwari (1997), to derive Bayes estimators of the system survival function corresponding to a risk subset, Δ. This is done as follows. We first substitute the Bayes estimator of $\mathbf{S}_\Delta^* = (S_\Delta^*, S_{\Delta^c}^*)$ in Peterson's formula (1.4), and obtain a nonparametric estimator of S_Δ. Using the product-integration approach described in Subsection 1.2, we then show the nonparametric estimator of the S_Δ is in fact a Bayes estimator under the quadratic loss function, i.e., the posterior mean of S_Δ given the data. The case $r = 2$,

4

which corresponds to the censored data problem is a consequence. In Section 3, the large sample properties such as strong consistency, law of iterated logarithm, and weak convergence of our estimator of S_Δ are studied.

2 Bayesian Analysis

The objective of this section is to consider the Bayesian approach to a competing-risks model or a series system and to derive the Bayes estimator of the survival function S_Δ associated with a risk set Δ.

2.1 Dirichlet Multivariate Processes

For definitions of the Dirichlet process and singular Dirichlet distributions we refer to Ferguson (1973) and Wilks (1962), Section 7.7, respectively. The Dirichlet multivariate process is defined as follows.

Let $(\mathcal{X}, \mathcal{A})$ be a measurable space, and $\alpha_1, \ldots, \alpha_r$ be finite, nonnull, and nonnegative measures defined on $(\mathcal{X}, \mathcal{A})$. Let $\boldsymbol{\rho} = (\rho_1, \ldots, \rho_r); P_1, \ldots, P_r$ be mutually independent random elements defined on a probability space $(\Omega, \mathcal{F}, Q)$. Suppose that $\boldsymbol{\rho}$ has a singular Dirichlet distribution $D_s(\alpha_1(\mathcal{X}), \ldots, \alpha_r(\mathcal{X}))$, and P_j is the Dirichlet process with parameter α_j, that is, $P_j \sim D(\alpha_j)$, $j = 1, \ldots, r$. Define $\mathbf{P}^* = (P_1^*, P_2^*, \ldots, P_r^*) = (\rho_1 P_1, \ldots, \rho_r P_r)$. Then $\mathbf{P}^*$ is a Dirichlet multivariate (r-variate) process with parameter $\boldsymbol{\alpha} = (\alpha_1, \alpha_2, \ldots, \alpha_r)$, that is, $\mathbf{P}^* \sim DM_r(\boldsymbol{\alpha})$.

In a context of a series system or a competing risks model, the definition of the Dirichlet multivariate process is established as follows.

Let $\alpha_1, \ldots, \alpha_r$ be finite, nonnull, and nonnegative measures on $((0, \infty), \mathcal{B}_{(0,\infty)})$. Let $\boldsymbol{\rho} = (\rho_1, \ldots, \rho_r) = (\Pr(\delta = 1), \ldots, \Pr(\delta = r)) \sim D_s(\alpha_1(0, \infty), \ldots, \alpha_r(0, \infty))$, $T_j^*(t) = \Pr(Z > t | \delta = j)$, $T^* \sim D(\alpha_j)$; $j = 1, \ldots, r$. Suposse that $\boldsymbol{\rho}, T_1^*, \ldots, T_r^*$

are mutually independent. Then the prior for $\mathbf{S}^*$ is given by $\mathbf{S}^* = (\rho_1 T_1^*, \ldots, \rho_r T_r^*) \sim DM_r(\alpha_1, \ldots, \alpha_r)$. (See Salinas-Torres, Pereira and Tiwari, 1997).

The induced prior for S_Δ^* is given by

$$S_\Delta^*(t) \sim Beta\left(c_r S_{\Delta,0}^*(t), \; c_r(1 - S_{\Delta,0}^*(t))\right), \; t > 0, \tag{2.1}$$

where $c_r = \sum_{j=1}^r \alpha_j(0, \infty)$ and $S_{\Delta,0}^*(t) = \sum_{j \in \Delta} \alpha_j(t, \infty)/c_r$ is the prior mean of S_Δ^*. Also, $S_0(t) = S_{\Delta,0} + S_{\Delta^c,0}$ is the prior mean of S.

Note that $S_\Delta^*(\cdot)$ is a Beta process on $[0, \infty)$ with independent increments of the type (see Hjort,1990)

$$dS_\Delta^*(t) \sim Beta\left(c_r dS_{\Delta,0}^*(t), \; c_r(1 - dS_{\Delta,0}^*(t))\right). \tag{2.2}$$

The following result gives the prior mean of the survival function S_Δ in terms of the prior mean of its associated cumulative hazard rate Λ_Δ.

Lemma 1. *Suppose that S_Δ and S_{Δ^c} have no common discontinuities. Under the prior (2.1) for $S_\Delta^*(\cdot)$, the prior mean of the survival function S_Δ is given by, for each $t > 0$,*

$$S_{\Delta,0}(t) := \mathbb{E}\left[S_\Delta(t)\right] = \prod_{[0,t]}(1 - d\Lambda_{\Delta,0}(s)),$$

where $\Lambda_{\Delta,0}(s) := \mathbb{E}\left[\Lambda_\Delta(s)\right]$ is the prior mean of Λ_Δ associated with the survival function S_Δ.

Proof.

Define a sequence of partitions of $[0, \infty)$ by $\{[(k-1)/l, k/l) : k, l = 1, 2, \ldots,\}$ and let

$$\widetilde{S}_\Delta^*(k/l) = S_\Delta^*((k-1)/l) - S_\Delta^*(k/l).$$

So that

$$\widetilde{S}(k/l) = \widetilde{S}_\Delta^*(k/l) + \widetilde{S}_{\Delta^c} * (k/l).$$

Also, let

$$\widetilde{S}_{\Delta,0}^*(k/l) = S_{\Delta,0}^*((k-1)/l) - S_{\Delta,0}^*(k/l),$$

$$\widetilde{S}_0(k/l) = \widetilde{S}_{\Delta,0}^*(k/l) + \widetilde{S}_{\Delta^c,0}^*(k/l).$$

Under the prior (2.1) for $S_\Delta^*(t)$, we have

$$\widetilde{S}_\Delta^*(k/l) \sim Beta(c_r \widetilde{S}_{\Delta,0}^*(k/l) \,,\, c_r(1 - \widetilde{S}_{\Delta,0}^*(k/l))),$$

and

$$\frac{\widetilde{S}_\Delta^*(k/l)}{\sum_{i \geq k} \widetilde{S}(i/l)} \sim Beta(c_r \widetilde{S}_{\Delta,0}^*(k/l) \,,\, c_r(\sum_{i \geq k} \widetilde{S}_0(i/l) - \widetilde{S}_{\Delta,0}^*(k/l)))$$

$$\sim Beta(c_r \widetilde{S}_{\Delta,0}^*(k/l) \,,\, c_r \sum_{i > k} \widetilde{S}_0(i/l)).$$

Also,

$$S_\Delta(t) = \lim_{\substack{l \to \infty \\ \frac{k-1}{l} \to t}} \prod_{i:i \leq k} \left(1 - \frac{\widetilde{S}_\Delta^*(i/l)}{\sum_{r \geq i} \widetilde{S}(r/l)}\right)$$

and hence

$$S_{\Delta,0}(t) = \mathbb{E}[S_\Delta(t)] = \lim_{\substack{l \to \infty \\ \frac{k-1}{l} \to t}} \prod_{i:i \leq k} \left(1 - \frac{\widetilde{S}_{\Delta,0}^*(i/l)}{\sum_{r > i} \widetilde{S}_0((r/l))}\right)$$

$$= \prod_{[0,t]} \left(1 - \frac{dS_{\Delta,0}^*(s)}{S_0(s^-)}\right) = \prod_{[0,t]} (1 - d\Lambda_{\Delta,0}(s)),$$

where the last equality follows from (1.3). □

Let

$$S_{jn}^*(t) = \frac{1}{n} \sum_{i=1}^n I(Z_i > t, \delta = j), j = 1, \ldots, r,$$

be the empirical subsurvival function of the j-th component, $j = 1, \ldots, r$, where $I(\cdot)$ is the indicator function. Let $S_n(t) = \sum_{j=1}^r S_{jn}^*(t)$ be the empirical survival function of the system.

7

The posterior distribution of $\mathbf{S}^*(t)$ is again an updated Dirichlet multivariate process given by $\mathbf{S}^*(t)|n\mathbf{S}_n^*(t) \sim DM_r(\alpha_1 + nS_{1n}^*, \ldots, \alpha_r + nS_{2n}^*)$; see Salinas-Torres, Pereira and Tiwari (1997).

Let $S_{\Delta n}^*(t) = \sum_{j \in \Delta} S_{jn}^*(t) = \frac{1}{n} \sum_{i=1}^{n} I(Z_i > t, \delta \in \Delta)$ be the empirical subsurvival associated with the risk subset Δ.

Consider the quadratic loss function

$$L(\mathbf{S}^*, \widehat{\mathbf{S}}^*) = \int_0^\infty \| \mathbf{S}^*(t) - \widehat{\mathbf{S}}^*(t) \|^2 \, dW(t), \tag{2.3}$$

where $\| \cdot \|$ is the usual R_r norm, $\widehat{\mathbf{S}}^* = (\widehat{\mathbf{S}}_1^*, \ldots, \widehat{\mathbf{S}}_r^*)$ is an estimator of $\mathbf{S}^* = (S_1^*, \ldots, S_r^*)$ and $W(\cdot)$ is a weight function.

Let $p_n = c_r/(c_r + n)$. Then the Bayes estimators of $S_\Delta^*(\cdot)$ and $S(\cdot)$ are given by

$$\widehat{S}_\Delta^*(t) = p_n S_{\Delta,0}^*(t) + (1 - p_n)S_{\Delta,n}^*(t),$$

and

$$\widehat{S}(t) = \sum_{j=1}^{r} \widehat{S}_j^*(t) = \widehat{S}_\Delta^*(t) + \widehat{S}_{\Delta^c}^*(t) \ ,$$

respectively.

As a remark we would like to point out that the Bayes estimators $\widehat{S}_\Delta^*$, $\widehat{S}$ are strongly consistent. For instance, using Glivenko Cantelli Theorem and the fact that $p_n \downarrow 0$, it can be shown that $\widehat{S}_\Delta^*$ converges to S_Δ^* uniformly w.p. 1.

2.2   Main Result

Let the $m(\leq n)$ distinct order statistics of Z be $Z_{(1)} < \ldots < Z_{(m)}$. Set $n_j = \sum_{i=1}^{n} I(Z_i \geq Z_{(j)})$ and $d_j = \sum_{i=1}^{n} I(Z_i = Z_{(j)}, \delta_i = 1)$, $j = 1, \ldots, m$. Define

$$i_\Delta(t) = exp \left\{ \frac{-1}{c_r + n} \sum_{j \in \Delta^c} \int_0^t \frac{d\alpha_j(s, \infty)}{\widehat{S}(s)} \right\}$$

8

and

$$\pi_r(t) = \prod_{i:Z_{(i)} \leq t} \frac{\sum_{j=1}^{r} \alpha_j(Z_{(i)}, \infty) + n_i - d_i}{\sum_{j=1}^{r} \alpha_j(Z_{(i)}, \infty) + n_i},$$

where Δ^c is the complement of Δ. The main result is given below.

Theorem 1. *Suppose that the function* $\boldsymbol{f}(s) = (\alpha_1(s, \infty), \ldots, \alpha_r(s, \infty))$ *is continuous on* $(0, t)$, *for each* $t > 0$, *and* S_Δ *and* S_{Δ^c} *have no common discontinuities then, for* $t \leq Z_{(m)}$,

$$\widehat{S}_\Delta(t) = \varphi(\widehat{S}_\Delta^*, \widehat{S}_{\Delta^c}^*; t) = \widehat{S}(t) i_\Delta(t) \pi_\Delta(t) ,$$

is the Bayes estimator of $S_\Delta(t)$ *under the quadratic loss function* (2.3).

Proof. Substituting the Bayes estimates S_Δ^* and $S_{\Delta^c}^*$ in Peterson's formula (1.4) , we have

$$\widehat{S}_1(t) = \exp\left\{\int_0^t \frac{d\widehat{S}_\Delta^*}{\widehat{S}(s)}\right\} \prod \frac{\sum_{j=1}^{r} \widehat{S}_j^*(s+)}{\sum_{j=1}^{r} \widehat{S}_j^*(s-)} , \quad t \leq t^*. \tag{2.4}$$

Note that $d\widehat{S}_\Delta^*(s) = d\widehat{S}(s) - \dfrac{\sum_{j \in \Delta^c} d\alpha_j(s, \infty)}{c_r + n}$, and the first term in (**??**) becomes $\widehat{S}(t) i_\Delta(t)$. Since for each fixed $t > 0$, $f_j(\cdot, \infty)$ is monotone decreasing and continuous on $(0, t)$, and $1/\widehat{S}$ is monotone increasing on $(0, t)$ and both are of bounded variation on $(0, t)$, f_j can be decomposed uniquely as the difference of two monotone continuous functions (cf. Rudin, 1964, Corollary 1 of Theorem 6.27) and $1/\widehat{S}$ as the difference of two monotone functions (not necessarily unique). Thus, the integral $\displaystyle\int_0^t \frac{\sum_{j \in \Delta^c} d\alpha_j(s, \infty)}{\widehat{S}(s)}$ is well defined. Moreover, the second factor in (**??**) is

$$\prod_{s \leq t} \frac{\sum_{j=1}^{r} \alpha_j(s, \infty) + \sum_{i=1}^{n} I(Z_i > s+)}{\alpha_1(s, \infty) + \alpha_2(s, \infty) + \sum_{i=1}^{n} I(Z_i > s-)} = \pi_r(t).$$

On the other hand, proceeding as in Lemma 1,

$$\widehat{S}_\Delta(t) = \mathbb{E}[S_\Delta(t) \,|\, data\,] = \prod_{[0,t]} (1 - d\widehat{\Lambda}_\Delta(s)), \tag{2.5}$$

9

where $d\widehat{\Lambda}_{\Delta}(s) = \dfrac{c_r dS^*_{\Delta,0}(s) + nd\widehat{S}^*_{\Delta}(s)}{cS_0(s^-) + n\widehat{S}(s^-)}$. Simplifying (2.5) yields (2.4). □

Note that (2.5) is the posterior mean of $S_{\Delta}(t)$ given the data; that is the Bayes estimator of S_{Δ} under the loss function (2.3).

When $r = 2$, $\Delta = \{1\}$ and $\Delta^c = \{2\}$, the series system corresponds to the randomly censored data model. In this case, letting $\alpha_1(t, \infty) + \alpha_2(t, \infty) = \alpha(t, \infty)$, for each t, the product $\widehat{S}(t)\pi_r(t)$ is analogous to the equation (3.2) of Tsai (1986). Also, $\widehat{S}(t)\pi_r(t)$ is similar to the SV estimator. If $\alpha_j(0, \infty)$ tend to zero, for all $j = 1, 2$ the estimator $\widehat{S}_1$ reduces to the PL estimator.

3   Large sample properties of $\widehat{S}_{\Delta}$

Let $D[0, T]$ be the space of *cadlag* functions on $[0, T]$ equipped with the supremum norm $||F||_T = \sup_{0 \leq t \leq T} |F(t)|$. Denote by $D[0, T] \times D[0, T]$ the product space, where $||(F, G)||_T = \max(||F||_T, ||G||_T)$.

The strong consistency of $\widehat{S}_{\Delta}$ follows from the strong consistency of $\widehat{\mathbf{S}}^* = (\widehat{S}^*_1, \ldots, \widehat{S}^*_r)$ in conjunction with the following result.

Lemma 2. *The functional φ from $D[0, T] \times D[0, T]$ into $D[0, T]$ defined in Theorem 1 is continuous.* □

Lemma 2 can be easily shown by considering a sequence (F_n, G_n) in $D[0, T] \times D[0, T]$ such that $||(F_n, G_n) - (F, G)||_T \to 0$ for some $(F, G) \in D[0, T] \times D[0, T]$ and showing that for $H_n(\cdot) = \varphi(F_n, G_n; \cdot)$, $||H_n - H||_T \to 0$, where $H(\cdot) = \varphi(F, G; \cdot)$.

To establish the weak convergence of $\widehat{S}_{\Delta}$, note that

$$n^{1/2}[(\widehat{S}^*_\Delta, \widehat{S}^*_{\Delta^c}) - (S^*_\Delta, S^*_{\Delta^c})] = n^{1/2}[(S^*_{\Delta,n}, S^*_{\Delta^c,n}) - (S^*_\Delta, S^*_{\Delta^c})] + o_p(1) \ , \qquad (3.1)$$

since $p_n \downarrow 0$. The bivariate empirical process $n^{1/2}[(S^*_{\Delta,n}, S^*_{\Delta^c,n})(t) - (S^*_\Delta, S^*_{\Delta^c})(t)]$ converges weakly to a Gaussian process (U, V) with mean $(0,0)$ and a covariance kernel given by, for $0 \le s \le t \le t^*$, (see Tsai and Crowley, 1985)

$$
\begin{aligned}
\text{Cov}\,(U(s), U(t)) &= (1 - S^*_\Delta(s))S^*_\Delta(t) \\
\text{Cov}\,(U(s), V(t)) &= -S^*_\Delta(s)S^*_{\Delta^c}(t) \\
\text{Cov}\,(V(s), U(t)) &= -S^*_{\Delta^c}(s)S^*_\Delta(t) \\
\text{Cov}\,(V(s), V(t)) &= (1 - S^*_{\Delta^c}(s))S^*_{\Delta^c}(t) \ .
\end{aligned}
\qquad (3.2)
$$

From (**??**) it follows that $n^{1/2}[(\widehat{S}^*_\Delta, \widehat{S}^*_{\Delta^c})(t) - (S^*_\Delta, S^*_{\Delta^c})(t)]$ converges weakly to the Gaussian process (U, V) in (3.2).

Using the von Mises' (1947) representation of statistics and applying results from Breslow and Crowley (1974), and Tsai and Crowley (1985), the following theorem gives the weak convergence and the law of iterated logarithm for $\widehat{S}_\Delta$.

Theorem 2.

a) *Suppose that the survival functions S_Δ and S_{Δ^c} have no common discontinuities. Then $n^{1/2}(\widehat{S}_\Delta - S_\Delta)$ converges weakly to a Gaussian process $W(\cdot)$ with mean zero and a covariance kernel given by, for $s \le t < t^* = \min(t_{S_\Delta}, t_{S_{\Delta^c}})$,*

$$\text{Cov}\,(W(s), W(t)) = S_\Delta(s)S_\Delta(t) \int_0^s \frac{dS^*_\Delta(u)}{[S(u-)]^2} \ .$$

b) *Suppose that, in addition, the survival functions S_Δ and S_{Δ^c} are continuous. Then*

$$\sup_{0 \le t \le t^*} |\widehat{S}_\Delta(t) - S_\Delta(t)| = O\left(\sqrt{\frac{\log \log n}{n}}\right) \quad a.s. \ as \ n \to \infty \ .$$

11

Proof. The arguments being fairly routine, we only sketch the proofs.

a) $n^{1/2}(\widehat{S}_\Delta - S_\Delta) = \varphi'((S^*_\Delta, S^*_{\Delta^c}); n^{1/2}((\widehat{S}^*_\Delta, \widehat{S}^*_{\Delta^c}) - (S^*_\Delta, S^*_{\Delta^c})) + o(n^{1/2}||(\widehat{S}^*_\Delta, \widehat{S}^*_{\Delta^c}) - (S^*_\Delta, S^*_{\Delta^c})||_{t^*}$, where $\varphi'((S^*_\Delta, S^*_{\Delta^c}); \cdot)$ is the Frechet differential of φ at point $(S^*_\Delta, S^*_{\Delta^c})$. Applying, Lemma 4.3 of Tsai and Crowley(1985), and Kiefer's inequality to the second term, we obtain

$$
n^{1/2}(\widehat{S}_\Delta - S_\Delta)(t) = S_\Delta(t)\left\{ n^{1/2}\frac{(\widehat{S}^*_\Delta - S^*_\Delta)(t)}{S(t)} + \int_0^t n^{1/2}\frac{(\widehat{S}^*_\Delta(u) - S^*_\Delta(u))}{[S(u-)]}dS(u)\right.
$$
$$
\left. - \int_0^t n^{1/2}\frac{(\widehat{S}(u) - S(u))}{[S(u-)]^2}dS_\Delta(u)\right\} + o_p(1).
$$

We conclude the proof using Theorem 4.4 of Tsai and Crowley (1985) and Breslow and Crowley (1974).

b) Note that the $\varphi'((F, G); \cdot)$ is a linear and continuous transformation, therefore

$$
||\widehat{S}_\Delta - S_\Delta||_{t^*} \le ||\varphi'|| ||(\widehat{S}^*_\Delta, \widehat{S}^*_{\Delta^c}) - (S^*_\Delta, S^*_{\Delta^c})||_{t^*} + O(||(\widehat{S}^*_\Delta, \widehat{S}^*_{\Delta^c}) - (S^*_\Delta, S^*_{\Delta^c})||_{t^*}),
$$

where $||\varphi'||$ is the norm on the space of linear transformations. The strong consistency of the estimator $(\widehat{S}^*_\Delta, \widehat{S}^*_{\Delta^c})$ and the law of iterated logarithm for $\widehat{S}^*_\Delta$ and $\widehat{S}^*_{\Delta^c}$ along with the continuity assumption, imply the conclusion. $\square$

We note that $\widehat{S}_\Delta$ is a generalized self consistent estimator (Tsai, 1986), since it is obtained by using Peterson's formula and the asymptotic properties of $\mathbf{S}^*$.

Acknowledgements

References

Breslow, N.E. and Crowley, J. (1974). A large sample study of the life table and product limit estimates under random censorship. *The Annals of Statistics*, **13**, 437–453.

Ferguson, T.S. (1973). A Bayesian analysis of some nonparametric problems. *The Annals of Statistics*, **1**, 209–230.

Ferguson, T.S., Phadia, E.G., and Tiwari, R.C. (1992). Bayesian nonparametric inference. In *Current Issues in Statistical Inference: Essays in Honor of D. Basu* (M. Ghosh and P.K. Pathak, eds.), 127-150. IMS Lecture Notes & Monograph Series, Vol. 17.

Gill, R.D. and Johansen, S. (1990). A survey of product-integration with a view towards application in survival analysis. *The Annals of Statistics*, **18**, 1501-1555.

Hjort, N.L. (1990). Nonparametric Bayes estimators based on Beta processes in models for life history data. *The Annals of Statistics*, **18**, 1259-1294.

Kaplan, E.L. and Meier, P. (1958). Nonparametric estimation from incomplete observations. *Journal of the American Statistical Association*, **53**, 87–90, 457–481.

Lo, A.Y. (1993). A Bayesian method for weigthed sampling. *The Annals of Statistics*, **21**, 2138-2148.

Peterson, A.V. (1977). Expressing the Kaplan-Meier estimator as a function of empirical subsurvival functions. *Journal of the American Statistical Association*, **72**, 854-858.

Rudin, W. (1964). *Principles of Mathematical Analysis*. 2nd. ed. McGraw-Hill, New York.

Salinas-Torres, V.H.; Pereira, C.A.B. and Tiwari, R.C. (1997). Convergence of Dirichlet measures arising in context of Bayesian analysis of competing risks models. *Journal of Multivariate Analysis*, **62**, 24-35.

Susarla, V. and Van Ryzin, J. (1976). Nonparametric Bayesian estimation of survival curves from incomplete observations. *Journal of the American Statistical Association*, **71**, 897-902.

Susarla, V. and Van Ryzin, J. (1978). Large sample theory for a Bayesian nonparametric survival curve estimator based on censored data. *The Annals of Statistics*, **6**, 755–768.

Tsai, W.Y. (1986). Estimation of survival curves from dependent censored models via a generalized self-consistent property with nonparametric Bayesian estimation application. *The Annals of Statistics*, **14**, 238–249.

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

Von Mises, R. (1947). On the asymptotic distribution of differentiable statistical functions. *The Annals of Mathematical Statistics*, **18**, 309–348.

Wilks, S.S. (1962). *Mathematical Statistics*. Wiley, New York.