The Korean Communications
in Statistics Vol. 12 No. 3, 2005
pp. 545-556

Weak Convergence for Nonparametric Bayes Estimators Based on Beta Processes in the Random Censorship Model

Jee Chang Hong[1)]

Abstract

Hjort(1990) obtained the nonparametric Bayes estimator $\hat{F}_{c,\alpha}$ of F_0 with respect to beta processes in the random censorship model. Let $X_1, \cdots, X_n$ be i.i.d. F_0 and let $C_1, \cdots, C_n$ be i.i.d. G. Assume that F_0 and G are continuous. This paper shows that $\{\hat{F}_{c,\alpha}(u) | 0 < u < \tau\}$ converges weakly to a Gaussian process whenever $\tau < \infty$ and $\tilde{F}_0(\tau) < 1$.

Keywords : Nonparametric Bayes estimator, Compact differentiability, Delta method

1. Introduction and Summary

Let $X_1, \cdots, X_n$ be independent and identically distributed(i.i.d.) random variables from a distribution F_0 on $[0, \infty)$ having $F_0(0) = 0$ and let $C_1, \cdots, C_n$ be i.i.d. random variables with cumulative distribution function(cdf) G on $[0, \infty)$. Assume that the X_i are independent of the C_i. Let $T_i = min\{X_i, C_i\}$, $\delta_i = 1\{X_i \leq C_i\}$ for each $i = 1, \cdots, n$, and let $T_1, \cdots, T_n$ be i.i.d. random variables with cdf H_0. Then $1 - H_0 = (1 - F_0)(1 - G)$. In the usual random censorship model one observes only $(T_1, \delta_1), \cdots, (T_n, \delta_n)$.

The problem of constructing nonparametric Bayes estimators(NPBE) for F_0 based on the censored data $(T_1, \delta_1), \cdots, (T_n, \delta_n)$ has been considered by many authors by placing a prior distribution for F_0 on the space $\mathfrak{F}$ of all cdf's on $[0, \infty)$. Using the Dirichlet process introduced by Ferguson(1973), an NPBE for F_0 based on the censored data has been considered by Susarla and Van Ryzin(1976). Ferguson and Phadia(1979) obtained an NPBE for F_0 with respect to the prior process neutral to the right introduced by Doksum(1974).

1) Assistant Professor, Department of Liberal Arts, Hanzhong University, Donghae 240-713, Korea.
E-mail : jchong@donghae.ac.kr

Ferguson and Phadia(1979) extend the result of Susarla and Van Ryzin(1978) that a Dirichlet process is a process neutral to the right.

Given a cdf F_0 on $[0,\infty)$, the cumulative hazard function(chf) A_0 is defined by

$$A_0(t)=\int_{[0,t]}\frac{dF_0(s)}{1-F_0(s-)},\qquad t\geq 0. \tag{1.1}$$

The formula (1.1) yields

$$F_0(t)=\int_{[0,t]}[1-F_0(s-)]dA_0(s) \tag{1.2}$$

which is well known as the Volterra integral equation. The unique solution of the equation (1.2) is given in terms of the product integral by

$$F_0(t)=1-\prod_{[0,t]}(1-dA_0),\quad t\geq 0. \tag{1.3}$$

See Gill and Johansen(1990). Here $\prod$ is the product integral.

Let N be the process counting observed failures and Y be the process giving the number at risk defined by

$$\begin{aligned} N(t)&=\sum_{i=1}^{n}1\{T_i\leq t,\delta_i=1\},\\ Y(t)&=\sum_{i=1}^{n}1\{T_i\geq t\}. \end{aligned} \tag{1.4}$$

The Nelson-Aalen estimator $\hat{A}_{NA}$ of A_0 based on $(T_1,\delta_1),\cdots,(T_n,\delta_n)$ is given by

$$\hat{A}_{NA}(t)=\int_{[0,t]}\frac{dN}{Y}. \tag{1.5}$$

By a substitution of (1.5) into the right-hand side(rhs) of (1.3) we obtain the Kaplan-Meier estimator $\hat{F}_{KM}$ of F_0 which is given by

$$\hat{F}_{KM}(t)=1-\prod_{[0,t]}(1-d\hat{A}_{NA}),\quad t\geq 0. \tag{1.6}$$

For investigation of the survival phenomena the chf of A_0 is as a basic object as the survival function F_0. Let $\mathbb{A}$ be the space of all chf's. Hjort(1990) introduced a beta process for A_0 with parameter functions $c(\cdot)$ and $\alpha(\cdot)$, denoted by $A_0\sim\text{beta}(c,\alpha)$, where $c(\cdot)$ is a piecewise continuous and nonnegative function on $[0,\infty)$ and $\alpha(\cdot)$ is a cumulative hazard function. A beta process is an $\mathbb{A}$-valued Lèvy process with independent increments.(See the definition of a beta process in Hjort(1990).)

The NPBE $\hat{A}_{c,\alpha}$ of A_0 with respect to the beta process $A_0\sim\text{beta}(c,\alpha)$ based on

$(T_1, \delta_1), \cdots, (T_n, \delta_n)$ obtained by Hjort(1990) is given by

$$\hat{A}_{c,\alpha}(t) = \int_{[0,t]} \frac{cd\alpha + dN}{c+Y}. \tag{1.7}$$

If A is a beta process, then the random distribution F given by (1.3) is a process neutral to the right. By substitution of (1.7) into rhs of (1.3) we obtain an estimator $\hat{F}_{c,\alpha}$ of F given by

$$\hat{F}_{c,\alpha}(t) = 1 - \prod_{[0,t]} (1 - d\hat{A}_{c,\alpha}), \quad t \geq 0. \tag{1.8}$$

Using the fact that the posterior of a beta process given data is also a beta process and a beta process has independent increments, one can easily see that $\hat{F}_{c,\alpha}$ is a conditional expectation of F given data. Therefore we see that the estimator $\hat{F}_{c,\alpha}$ given in (1.8) is an NPBE of F_0 with respect to a process neutral to the right under a squared error loss function. Recently, Gill and Johansen(1990) extended the usual delta method to a class of compactly differentiable functionals and then used it to prove the weak convergence of the Kaplan-Meier estimator. This new approach fits beautifully with the differentiability and now becomes one of the most powerful tools in proving weak convergence of many important statistical functionals. This technique which is referred to as the functional delta method is used extensively in our discussion of the weak convergence of the NPBE $\hat{F}_{c,\alpha}$.

Let $(\Omega, \Im, P)$ be the underlying probability space for this model and take filtration as

$$\Im_t = \sigma\{1\{T_i \leq s, \delta_i = 1\}, 1\{T_i \geq s\} : 0 \leq s \leq t, i = 1, \cdots, n\}, \quad t \geq 0. \tag{1.9}$$

Now, $(\Omega, \Im, \{\Im_t : t \geq 0\}, P)$ is the stochastic basis for this model. Martingale technique is used efficiently for obtaining the covariance structure of the weak limits of our estimators. Consider the process M on $[0, \infty)$ given by

$$M(t) = N(t) - \int_0^t Y dA_0, \tag{1.10}$$

where A_0 is given by (1.1) and the processes N and Y are given by (1.4). It is well-known that the process M is a square-integrable zero-mean martingale with respect to the filtration in (1.9) and it has the predictable variation process $\langle M, M \rangle$ given by

$$\langle M, M \rangle (t) = \int_0^t Y(1 - \Delta A_0) dA_0. \tag{1.11}$$

This is the unique, nondecreasing, predictable process such that $M^2 - \langle M, M \rangle$ is again a martingale.

Now we outline the idea how the functional delta method works combined with differentiability and Skorohod almost sure representation. Let U_n be random elements of a normed vector space such that

$$Z_n = a_n(U_n - x) \xrightarrow{D} Z, \tag{1.12}$$

where $a_n \to \infty$ is a sequence of real constants. Let ϕ be a function from this space into another normed vector space, compactly differentiable at x in the sense that for all $t_n \to 0$, $t_n \in \mathbb{R}$, and all $h_n \to h$

$$t_n^{-1}(\phi(x + t_n h_n) - \phi(x)) \to d\phi(x) \cdot h,$$

where $d\phi(x)$ is a continuous linear map between the two spaces.

By the Skorohod almost sure representation we may pretend that Z_n converges almost surely to Z. Now apply the definition of differentiability with x as given, $t_n = a_n^{-1}$, $h_n = Z_n$, $h = Z$, so that $x + t_n h_n = x + a_n^{-1} Z_n = U_n$. We obtain that $a_n(\phi(U_n) - \phi(x)) \xrightarrow{a.s.} d\phi(x) \cdot Z$, which implies that $a_n(\phi(U_n) - \phi(x)) \xrightarrow{D} d\phi(x) \cdot Z$ as desired. In most statistical applications the random elements U_n and Z_n in (1.12) are the empirical distribution functions of data and the empirical processes with $a_n = n^{\frac{1}{2}}$ in which (1.12) is guaranteed by Donsker's theorem(Van der Vaart and Wellner(1996)).

Define $F_n^1, \tilde{F}_n, F_0^1, \tilde{F}_0$ as (3.1) and (3.2). Here $\phi^{(n)}$ is given by (3.5). In our application, the functional delta method with $U_n = (F_n^1, \tilde{F}_n)$, $x = (F_0^1, \tilde{F}_0)$, $a_n = \sqrt{n}$

$$Z_n = (Z_n^1, \tilde{Z}_n) = (\sqrt{n}(F_n^1 - F_0^1), \sqrt{n}(\tilde{F}_n - \tilde{F}_0)) \xrightarrow{D} (Z^1, \tilde{Z}) = Z \tag{1.13}$$

for the weak convergence of the process $\hat{F}_{c,\alpha}$ is not directly applicable since the functional $\phi^{(n)}$ varies over $n \geq 1$ and it is not clear that $\phi^{(n)}$ is compactly differentiable.

Susarla and Van Ryzin(1978) verified consistency of $\hat{F}_\alpha$, the NPBE of F with respect to the Dirichlet process. Perhaps the most important small sample advantage of NPBE $\hat{F}_{c,\alpha}$ over the Kaplan-Meier estimator is admissibility. As noted by Susarla and Van Ryzin(1978), NPBE is probably admissible, although it may not be easy to establish. Some other small sample advantages of NPBE are discussed in Susarla and Van Ryzin(1978).

Section 2 extends functional delta method for our purposes. Using the extended functional delta method and an analytic property of the product integral operator, in Section 3, we verify the weak convergence of $\hat{F}_{c,\alpha}$. We introduce the compact differentiability of the mappings ϕ_1, ϕ_2 and ϕ_3 and give their derivatives in Appendix.

2. Generalized functional delta method

Let B_1 and B_2 be normed vector spaces and let $\phi : B_1 \to B_2$. We briefly give a definition of differentiability, see Gill(1989) for further background.

Definition 2.1 (Gill(1989)) We say that $\phi : B_1 \to B_2$ is compactly differentiable at $x \in B_1$ if there exists a continuous linear operator $d\phi(x) : B_1 \to B_2$ such that for all $t_n \to 0 (t_n \in \mathbb{R})$ and $h_n \to h \in B_1$,

$$\left\| \frac{1}{t_n} (\phi(x + t_n h_n) - \phi(x)) - d\phi(x) \cdot h \right\|_\infty \to 0 \quad \text{as } n \to \infty.$$

Here $\| \cdot \|_\infty$ is the supremum norm on the space of B_2.

The following Gill's(1989) theorem, finally present definitive version of delta method.

Theorem 2.2 (functional delta method) Let U_n be a sequence of random elements of B_1, $a_n \to \infty$ a real sequence, such that $a_n(U_n - x) \xrightarrow{D} Z$ for some fixed point $x \in B_1$ and a random element Z of B_1. Suppose ϕ is compactly differentiable at $x \in B_1$. Then,

$$a_n(\phi(U_n) - \phi(x)) \xrightarrow{D} d\phi(x) \cdot Z,$$

$$a_n(\phi(U_n) - \phi(x)) - d\phi(x) \cdot Z \xrightarrow{P} 0.$$

For our purposes, we extend Theorem 2.2 to the case where the functional $\phi^{(n)}$ may vary over $n \geq 1$ and may not be differentiable. The following theorem is the generalized functional delta method for weak convergence of the process.

Theorem 2.3 Let U_n be a sequence of random elements of B_1, $a_n \to \infty$ a real sequence, such that $a_n(U_n - x) \xrightarrow{D} Z$ for some fixed point $x \in B_1$ and a random element Z of B_1. Suppose ϕ is compactly differentiable at x. If the condition

$$a_n \| \phi^{(n)} - \phi \|_\infty \xrightarrow{a.s.} 0 \quad \text{as } n \to \infty \tag{2.1}$$

is satisfied. Then

$$a_n(\phi^{(n)}(U_n) - \phi(x)) \xrightarrow{D} d\phi(x) \cdot Z$$

and moreover, $a_n(\phi^{(n)}(U_n) - \phi(x))$ and $d\phi(x) \cdot a_n(U_n - x)$ are asymptotically equivalent.

Proof. We show that

$$\left\| a_n(\phi^{(n)}(U_n) - \phi(x)) - d\phi(x) \cdot Z \right\|_\infty \xrightarrow{a.s.} 0.$$

But we have

$$\begin{aligned} &\left\| a_n(\phi^{(n)}(U_n)-\phi(x))-d\phi(x)\cdot Z\right\|_\infty \\ &\le \left\| a_n(\phi^{(n)}(U_n)-\phi(U_n))+a_n(\phi(U_n)-\phi(x))-d\phi(x)\cdot Z\right\|_\infty . \end{aligned} \tag{2.2}$$

By condition (2.1) and Theorem 2.2, the first and second terms of (2.2) converge to 0 a.s., respectively. □

3. Weak convergence of $\hat{F}_{c,\alpha}$

Let F_n^1 be the empirical (sub)distribution function of the T_i with $\delta_i=1$ and let $\tilde{F}_n$ be the empirical distribution function of all the T_i. With the processes N and Y in (1.4) we may write

$$\begin{aligned} F_n^1(t) &= \frac{N(t)}{n}, \\ 1-\tilde{F}_n(t-) &= \frac{Y(t)}{n}. \end{aligned} \tag{3.1}$$

Let

$$\begin{aligned} F_0^1(t) &= EF_n^1(t), \\ \tilde{F}_0(t) &= E\tilde{F}_n(t). \end{aligned} \tag{3.2}$$

Then it can be easily seen that

$$\begin{aligned} F_0^1(t) &= \int_0^t (1-G^-)dF_0, \\ \tilde{F}_0(t) &= 1-(1-F_0(t))(1-G(t)). \end{aligned} \tag{3.3}$$

For $\tilde{F}_0(\tau)<1$, $\tau<\infty$, it follows from (3.3) that

$$\begin{aligned} A_0(t) &= \int_0^t \frac{dF_0}{1-F_0^-} = \int_0^t \frac{dF_0^1}{1-\tilde{F}_0^-}, \\ F_0(t) &= 1-\prod_{[0,t)}\left(1-\frac{dF_0}{1-F_0^-}\right) = 1-\prod_{[0,t)}\left(1-\frac{dF_0^1}{1-\tilde{F}_0^-}\right). \end{aligned} \tag{3.4}$$

For fixed $\tau>0$, define the following mappings $\phi_1,\ \phi_2 : D[0,\tau]^2 \to D[0,\tau]$, $\phi_3 : D[0,\tau] \to D[0,\tau]$ by

$$\phi_1(x,y) = \left(x, \frac{1}{1-y^-}\right),$$

$$\phi_2(x,y) = \int y\,dx,$$

$$\phi_3(x) = 1 - \prod(1 - dx)$$

where $D[0,\tau]$ is the space of all real functions on $[0,\tau]$ which are right-continuous on $[0,\tau)$ and have left-limits on $(0,\tau]$.

It turns out that

$$A_0 = (\phi_2 \circ \phi_1)(F_0^1, \tilde{F}_0),$$
$$F_0 = (\phi_3 \circ \phi_2 \circ \phi_1)(F_0^1, \tilde{F}_0) = \phi(F_0^1, \tilde{F}_0).$$

Then (1.5) and (1.6) can be rewritten

$$\hat{A}_{NA} = (\phi_2 \circ \phi_1)(F_n^1, \tilde{F}_n)$$

and

$$\hat{F}_{KM} = (\phi_3 \circ \phi_2 \circ \phi_1)(F_n^1, \tilde{F}_n) = \phi(F_n^1, \tilde{F}_n), \tag{3.5}$$

respectively.

We prove the compact differentiability of the mappings ϕ_1, ϕ_2 and ϕ_3 in Appendix with their derivatives.

For each $n \geq 1$, let

$$\phi_1^{(n)}(x, y) = \left(\frac{c}{n}\alpha + x, \frac{1}{\frac{c}{n} + (1 - y^-)}\right),$$

where c, α are the parameter functions of the beta process. Then we may write

$$\hat{A}_{c,\alpha} = (\phi_2 \circ \phi_1^{(n)})(F_n^1, \tilde{F}_n),$$
$$\hat{F}_{c,\alpha} = (\phi_3 \circ \phi_2 \circ \phi_1^{(n)})(F_n^1, \tilde{F}_n) = \phi^{(n)}(F_n^1, \tilde{F}_n). \tag{3.6}$$

We prove weak convergence of $\hat{F}_{c,\alpha}$ by assuming further that the parameter function $c(\cdot)$ of the beta process beta$\{c, \alpha\}$ is bounded by a positive constant $K > 0$ so that

(A1) $0 \leq c(t) \leq K,\ t \geq 0$

Theorem3.1 Let $\tau < \infty$ and $\tilde{F}_0(\tau) < 1$. Let F_0 and G be continuous. Then

$$\sqrt{n}(\hat{F}_{c,\alpha} - F_0) \xrightarrow{D} R \quad \text{in } D[0,\tau],$$

where $R = (1 - F_0)\int \frac{1}{1 - \tilde{F}_0^-}(dZ^1 + \tilde{Z}dA_0)$ is a zero-mean Gaussian process with the asymptotic covariance structure given for $0 \leq s, t \leq \tau$ by

$$(1 - F_0(s))(1 - F_0(t))\int_0^{s \wedge t} \frac{dF_0}{(1 - F_0)^2(1 - G)}. \tag{3.7}$$

Proof. Applying Lemma A4 to $(h, k) = (Z^1, \tilde{Z})$ at the point $(x, y) = (F_0^1, \tilde{F}_0)$ yields

$$R = d\phi(F_0^1, \widetilde{F}) \cdot (Z^1, \widetilde{Z})$$
$$= \int \prod (1 - d(\int \frac{dF_0^1}{1-\widetilde{F}_0^-}))d(\int \frac{\widetilde{Z}}{(1-\widetilde{F}_0^-)^2} dF_0^1 + \int \frac{dZ^1}{1-\widetilde{F}_0^-}) \prod (1 - d(\int \frac{dF_0^1}{1-\widetilde{F}_0^-}))$$
$$= \int \prod (1 - dA_0)(\frac{\widetilde{Z}}{(1-\widetilde{F}_0^-)^2} dF_0^1 + \frac{dZ^1}{1-\widetilde{F}_0^-})$$
$$= (1-F_0)\int \frac{1}{1-\widetilde{F}_0^-}(dZ^1 + \widetilde{Z}dA_0).$$

By Skorohod-Dudley almost-sure representation theorem(Billingsley(1986)) and (1.13), we will show that

$$\left\| \sqrt{n}(\hat{F}_{c,\alpha} - F_0) - R \right\|_\infty \xrightarrow{a.s.} 0.$$

Using (3.5), (3.6) we have

$$\begin{aligned} &\left\| \sqrt{n}(\hat{F}_{c,\alpha} - F_0) - R \right\|_\infty \\ &= \left\| \sqrt{n}[\hat{F}_{c,\alpha} - \hat{F}_{KM} + \hat{F}_{KM} - F_0] - R \right\|_\infty \\ &\le \left\| \sqrt{n}[\phi^{(n)}(F_n^1, \widetilde{F}_n) - \phi(F_n^1, \widetilde{F}_n)] \right\|_\infty \\ &+ \left\| \sqrt{n}[\phi(F_n^1, \widetilde{F}_n) - \phi(F_0^1, \widetilde{F}_0)] - d\phi(F_0^1, \widetilde{F}_0) \cdot (Z^1, \widetilde{Z}) \right\|_\infty. \end{aligned} \tag{3.8}$$

The first term of (3.8) shows by integration by parts that this difference is bounded by a constant C times $\|\hat{A}_{c,\alpha} - \hat{A}_{NA}\|_\infty$. We have

$$\begin{aligned} &\left\| \sqrt{n}[\phi^{(n)}(F_n^1, \widetilde{F}_n) - \phi(F_n^1, \widetilde{F}_n)] \right\|_\infty \\ &= \left\| \sqrt{n}[(1 - \prod_{[0,t]}(1 - d\hat{A}_{c,\alpha})) - (1 - \prod_{[0,t]}(1 - d\hat{A}_{NA}))] \right\|_\infty \\ &= \left\| \sqrt{n}[\prod_{[0,t]}(1 - d\hat{A}_{NA}) - \prod_{[0,t]}(1 - d\hat{A}_{c,\alpha})] \right\|_\infty \\ &= \left\| \sqrt{n}\int_0^t \prod_{[0,u)}(1 - d\hat{A}_{NA})(\hat{A}_{c,\alpha} - \hat{A}_{NA})(du)\prod_{(u,t]}(1 - d\hat{A}_{c,\alpha}) \right\|_\infty \\ &\le C\left\| \sqrt{n}(\hat{A}_{c,\alpha}(t) - \hat{A}_{NA}(t)) \right\|_\infty \\ &\le \frac{C}{\sqrt{n}} sup_{s \le t}\left| \int_0^s \frac{cd\alpha}{c/n + (1-\widetilde{F}_n^-)} \right| + \frac{C}{\sqrt{n}} sup_{s \le t}\left| \int_0^s \frac{cdF_n^1}{(1-\widetilde{F}_n^-)(c/n + (1-\widetilde{F}_n^-))} \right| \\ &\le \frac{C}{\sqrt{n}} \frac{\|c\|_\infty \alpha(t)}{1-\widetilde{F}_n^-(t)} + \frac{C}{\sqrt{n}} \frac{\|c\|_\infty F_n^1(t)}{(1-\widetilde{F}_n^-(t))^2} \end{aligned}$$

$$\leq \frac{1}{\sqrt{n}} \frac{CK\alpha(t)}{1-\widetilde{F}_n^-(t)} + \frac{1}{\sqrt{n}} \frac{CKF_n^1(t)}{(1-\widetilde{F}_n^-(t))^2} \to 0$$

as $n \to \infty$, since $1-\widetilde{F}_n^-(t) \to 1-\widetilde{F}_0^-(t)$ and $F_n^1(t) \to F_0^1(t)$.

By Gill and Johanson(1990), the second term of (3.8) shows that

$$\left\| \sqrt{n}[\phi(F_n^1, \widetilde{F}_n) - \phi(F_0^1, \widetilde{F}_0)] - d\phi(F_0^1, \widetilde{F}_0) \cdot (Z^1, \widetilde{Z}) \right\|_\infty \xrightarrow{a.s.} 0.$$

Thus

$$\left\| \sqrt{n}(\hat{F}_{c,\alpha} - F_0) - R \right\|_\infty \xrightarrow{a.s.} 0.$$

By Theorem 2.3, $\sqrt{n}(\hat{F}_{c,\alpha} - F_0) = \sqrt{n}(\phi^{(n)}(F_n^1, \widetilde{F}_n) - \phi(F_0, \widetilde{F}_0))$ and

$$\begin{aligned} R_n &= d\phi(F_0^1, \widetilde{F}_0) \cdot (Z_n^1, \widetilde{Z}_n) \\ &= (1-F_0)\int \frac{1}{1-\widetilde{F}_0^-}(dZ_n^1 + \widetilde{Z}_n^- dA_0) \end{aligned} \tag{3.9}$$

are asymptotically equivalent. The covariance structure of $\sqrt{n}(\hat{F}_{c,\alpha} - F_0)$ is a asymptotically equivalent to the covariance structure of R_n. The asymptotic covariance structure calculations involved in (3.7) are given in Appendix. We notice here that (3.7) coincides with (6.5) of Gill(1994). This completes the proof. □

Appendix

In the following three lemmas, we compute derivatives of ϕ_1, ϕ_2 and ϕ_3(see Gill(1989), Gill and Johansen(1990)).

Lemma A1 Let $E_1 = \{(x,y) \in D[0,\tau] \times D[0,\tau] : 0 < \|y\|_\infty \leq M_1, M_1 \in (0,1)\}$ and let $\phi_1 : E_1 \subset D[0,\tau] \times D[0,\tau] \to D[0,\tau] \times D[0,\tau]$ be defined by $\phi_1(x,y) = (x, \frac{1}{1-y})$. Then ϕ_1 is compactly differentiable at $(x,y) \in E_1$ with derivative

$$d\phi_1(x,y) \cdot (h,k) = (h, \frac{k}{(1-y)^2}).$$

Lemma A2 Let $E_2 = \{(x,y) \in D[0,\tau] \times D[0,\tau] : \int_0^\tau |dx| \leq M_2\}$ and let $\phi_2 : E_2 \subset D[0,\tau] \times D[0,\tau] \to D[0,\tau]$ be defined by $\phi_2(x,y) = \int y dx$. Let (x,y) now be a fixed point of E_2 such that $\int_0^\tau |dy|$ is finite. Then ϕ_2 is compactly differentiable at $(x,y) \in E_2$ with derivative

$$d\phi_2(x,y) \cdot (h,k) = \int kdx + \int ydh,$$

where the integral with respect to h is defined by the integration by parts formula if h is not of bounded variation.

Lemma A3 Let $E_3 = \{v \in D[0,\tau] : \int_0^\tau |dv| \leq M_3\}$ and let $\phi_3 : D[0,\tau] \to D[0,\tau]$ be defined by $\phi_3(v) = 1 - \prod(1 - dv)$. Then ϕ_3 is compactly differentiable at $v \in E_3$ with derivative

$$d\phi_3(v) \cdot l = \int \prod(1 - dv) dl \prod(1 - dv),$$

where a continuous linear operator is l all of which need not be of bounded variation.

One of the most important properties of compact differentiation is that it satisfies the chain rule : if $\psi : B_1 \to B_2$ and $\varphi : B_2 \to B_3$ are compact differentiable at $x \in B_1$ and $\psi(x) \in B_2$ respectively, then $\varphi \circ \psi : B_1 \to B_3$ is compact differentiable at x with derivative $d\varphi\{\psi(x)\} \cdot d\psi(x)$(a continuous linear map from B_1 into B_3).

Lemma A4 The composite $\phi = \phi_3 \circ \phi_2 \circ \phi_1$ is compactly differentiable with derivative

$$d\phi(x,y) \cdot (h,k)$$
$$= \int \prod(1 - d(\int \frac{dx}{1-y}))d(\int \frac{k}{(1-y)^2}dx + \int \frac{dh}{1-y})\prod(1 - d(\frac{dx}{1-y}))$$

Proof. ϕ is compactly differentiable with derivative

$$d\phi(x,y) \cdot (h,k)$$
$$= d(\phi_3 \circ \phi_2 \circ \phi_1)(x,y) \cdot (h,k)$$
$$= d\phi_3(\phi_2 \circ \phi_1(x,y)) \circ d\phi_2(\phi_1(x,y)) \circ d\phi_1(x,y) \cdot (h,k)$$
$$= d\phi_3(\phi_2 \circ \phi_1(x,y)) \circ d\phi_2(\phi_1(x,y)) \cdot (h, \frac{k}{(1-y)^2})$$
$$= d\phi_3(\int \frac{dx}{1-y}) \cdot d\phi_2(x, \frac{1}{1-y}) \cdot (h, \frac{k}{(1-y)^2})$$

$$= d\phi_3\left(\int \frac{dx}{1-y}\right) \cdot \left(\int \frac{k}{(1-y)^2}dx + \int \frac{dh}{1-y}\right)$$

$$= \int \prod(1-d(\int \frac{dx}{1-y}))d(\int \frac{k}{(1-y)^2}dx + \int \frac{dh}{1-y})\prod(1-d(\int \frac{dx}{1-y})).$$

Lemma A5 Let $\tau < \infty$ and $\tilde{F}_0(\tau) < 1$. Let F_0 and G be continuous. Then

$$Cov(R_n(s), R_n(t)) = (1-F_0(s))(1-F_0(t))\int_0^{s\wedge t} \frac{dF_0}{(1-F_0)^2(1-G)}.$$

Proof.

$$\begin{aligned}\sqrt{n}\,(dZ_n^1 + \tilde{Z}_n^- dA_0) &= n\,(dF_n^1 - dF^1 + (\tilde{F}_n^- - \widetilde{F_0^-})dA_0) \\ &= dN - YdA_0.\end{aligned} \quad (1)$$

By substitution of (1) into rhs of (3.9) we have

$$R_n = \frac{1}{\sqrt{n}}(1-F_0)\int \frac{1}{1-\tilde{F}_0}(dN - YdA_0) = \frac{1}{\sqrt{n}}(1-F_0)\int \frac{1}{1-\tilde{F}_0}dM.$$

Using (1.10), (1.11) we see that

$$\begin{aligned}
&Cov(R_n(s), R_n(t)) \\
&= Cov\left(\frac{1}{\sqrt{n}}(1-F_0(s))\int_0^s \frac{1}{1-\tilde{F}_0^-}dM, \frac{1}{\sqrt{n}}(1-F_0(t))\int_0^t \frac{1}{1-\tilde{F}_0^-}dM\right) \\
&= \frac{1}{n}(1-F_0(s))(1-F_0(t))E\left(\int_0^s \frac{1}{1-\tilde{F}_0^-}dM\int_0^t \frac{1}{1-\tilde{F}_0^-}dM\right) \\
&= \frac{1}{n}(1-F_0(s))(1-F_0(t))E\int_0^{s\wedge t} \frac{YdA_0}{(1-\tilde{F}_0^-)^2} \\
&= \frac{1}{n}(1-F_0(s))(1-F_0(t))\int_0^{s\wedge t} \frac{n(1-\tilde{F}_0^-)}{(1-\tilde{F}_0^-)^2}\frac{dF_0^1}{1-\tilde{F}_0^-} \\
&= (1-F_0(s))(1-F_0(t))\int_0^{s\wedge t} \frac{(1-G^-)dF_0}{(1-F_0^-)^2(1-G^-)^2} \\
&= (1-F_0(s))(1-F_0(t))\int_0^{s\wedge t} \frac{dF_0}{(1-F_0)^2(1-G)}.
\end{aligned}$$

References

[1] Billingsley, P. (1986). *Probability and Measure.* John Wiley & Son. New York.

[2] Doksum, K. (1974). Tailfree and neutral random probabilities and their posterior distributions. *Ann. Probability* **2** 183-201.

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

[4] Ferguson, T. S. and Phadia, E. G. (1979). Bayesian nonparametric estimation based one censored data. *Ann. Statist.* **7** 163-186.

[5] Gill, R. D. (1989). Non- and Semi-parametric maximum likelihood estimators and the von Mises method (Part I). *Scand. J. Statist.* **16** 97-128.

[6] Gill, R. D. (1994). *Lectures on probability theory.* Springer-Verlag. New York.

[7] Gill, R. D. and Johansen, S. (1990). A survey of product-integration with a view toward application in survival analysis. *Ann. Statist.* **18** 1501-1555.

[8] Hjort, N. L. (1990). Nonparametric Bayes estimators based on beta processes in models for life history data. *Ann. Statist.* **18** 1259-1294.

[9] Susarla, V. and Van Ryzin, J. (1976). Nonparametric Bayesian estimation of survival curve from incomplete observations. *J. Amer. Statist. Assoc.* **61** 897-902.

[10] Susarla, V. and Van Ryzin, J. (1978). Large sample theory for a Bayesian nonparametric survival curve estimator based on censored samples. *Ann. Statist.* **6** 755-768.

[11] Van der Vaart, A. W. and Wellner, J. A. (1996). *Weak convergence and Empirical processes.* Springer-Verlag. New York.

[Received April 2005, Accepted July 2005]