A NONPARAMETRIC ESTIMATOR OF THE SURVIVAL FUNCTION UNDER PROGRESSIVE CENSORING

Joseph C. Gardiner

Michigan State University, East Lansing, Michigan

and

V. Susarla

State University of New York, Binghamton, New York

1. Introduction

The subject of nonparametric estimation of the survival function from incomplete or censored observations has received much attention for more than two decades. We may cite here the celebrated work of Kaplan and Meier (1958) where a product-limit (PL-) estimator of the survival curve is obtained from a sample in which each lifetime may be truncated (fixed censorship) due to limits on observation. In Breslow and Crowley (1974) the properties of this estimator are considered in the case of random censorship, where each lifetime has its own censoring random variable, and the lifetimes and censoring times being each independent and identically distributed (i.i.d.) sequences and also independent of each other. By utilizing the notion of Dirichlet process priors introduced by Ferguson (1973), Susarla and Van Ryzin (1976) obtain a nonparametric Bayesian estimator of the survival function which generalizes the PL-estimator of Kaplan and Meier.

The basic formulation in these works involves consideration of a random sample of lifetimes $X_1,\ldots,X_n$ which may not be completely observable due to the existence of corresponding censoring variables $Y_1,\ldots,Y_n$. The recorded data for the sample is therefore $(Z_1,\delta_1),\ldots,(Z_n,\delta_n)$ where $Z_i=\min(X_i,Y_i)$ and $\delta_i=0$ or 1 according as $X_i>Y_i$ or $X_i\leq Y_i$. In several longitudinal investigations the variables $\{Z_i: 1\leq i\leq n\}$ are time-ordered: $Z_{(1)}\leq Z_{(2)}\leq\cdots\leq Z_{(n)}$. The first observation $Z_{(1)}$ is the smallest one followed by the next smallest $Z_{(2)}$ and so on until the largest observation $Z_{(n)}$ is recorded last. In these circumstances cost and time limitations often preclude prolonged experimentation until the complete set of data $\{(Z_i,\delta_i): 1\leq i\leq n\}$ has been recorded. Furthermore, cogent ethical reasons make it imperative that observation be ceased at the earliest possible stage if the current accumulated data warrants a clear statistical decision. Thus a progressively censored scheme may be advocated in which observation is curtailed at an intermediate stage determined by the cumulative statistical information. If the experimentation is terminated at the k_n^{th} stage, where $k_n\in\{1,\ldots,n\}$ may be a stopping time, then the recorded data are $\{(Z_{(i)},\delta_i^*): 1\leq i\leq k_n\}$, with $\delta_i^*=0$ or 1 according as $Z_{(i)}$ is a censoring time or a true lifetime. The only information available on the remaining $n-k_n$ units is that both their censoring and survival times exceed $Z_{(k_n)}$; that is $Z_{(j)}>Z_{(k_n)}$, $j=k_n+1,\ldots,n$.

In this paper we construct a nonparametric Bayesian estimator, under squared error loss, for the survival function F from the data $\{(Z_{(i)},\delta_i^*): 1\leq i\leq k_n;\ Z_{(j)}>Z_{(k_n)},\ j=k_n+1,\ldots,n\}$ when F follows a Dirichlet process prior. Our estimator thus generalizes, to the progressively censored case, the estimator of Susarla and Van Ryzin (1976) and encompasses both fixed and random censorship. It includes, of course, the cases in which the complete sample is observed ($k_n=n$), an extension of an estimator of Ferguson (1973) when no censoring is present and the Kaplan-Meier estimator. It should be noted that in a progressively censored scheme as described here the observed duration variables $\{Z_{(i)}: 1\leq i\leq k_n\}$ and their corresponding identifiers $\{\delta_i^*: 1\leq i\leq k_n\}$ are neither independent nor identically distributed. The absence of this important technical facility in the case of progressive censoring (which is available when the complete data set is observed as in the works cited earlier) introduces additional complications and subtleties in the analysis of progressively censored schemes. For some applications of progressive censoring see Sen, et al. (1973, 1978, 1981).

The substantive material in this paper is distributed in the following three sections. Section 2 introduces the basic assumptions, notation and preliminary notions and provides a brief genesis of our estimator. Various special cases are also dealt with here. We have placed the laborious technical manipulations of construction in Section 4, while Section 3 provides a numerical example.

2. Preliminaries

We are concerned with longitudinal studies in which n specimens under test are followed from the onset with either the time to decrement (survival time) X or its competing censoring time Y recorded for each unit up to the time of the k^{th} response, $k \in \{1,\ldots,n\}$. We suppose the survival distribution F of X is a Dirichlet process (for the definition of a Dirichlet process and other terms, See Ferguson (1973)) and given F, the survival times $X_1,\ldots,X_n$ of the sample are independent and identically distributed (with distribution 1-F). Furthermore, we consider the corresponding censoring times $Y_1,\ldots,Y_n$ to be independent of F, $X_1,\ldots,X_n$, but make no further assumptions on the distribution of the Y_i's themselves.

The objective is to estimate the survival curve

$$F(t) = P[X > t | F], \quad t \geq 0 \quad . \tag{1}$$

We do not have at our disposal the complete set of data $\{(Z_i,\delta_i): 1 \leq i \leq n\}$ where

$$Z_i = \min(X_i, Y_i)$$

$$\delta_i = 0 \text{ or } 1 \quad \text{according as } X_i > Y_i \quad \text{or} \quad X_i \leq Y_i \quad ,$$

but rather the first k order statistics $\{Z_{(1)},\ldots,Z_{(k)}\}$ from $\{Z_1,\ldots,Z_n\}$ and their corresponding identifiers $\{\delta_1^*,\ldots,\delta_k^*\}$, where $\delta_i^* = 1$ or 0 according as $Z_{(i)}$ is a survival time or censoring time, as well as the information that $Z_{(j)} > Z_{(k)}$, $j=k+1,\ldots,n$. On the basis of these recorded data we seek the Bayes estimator $\hat{F}(t)$ of $F(t)$ under the loss ℓ

$$\ell(\hat{F},F) = \int_0^\infty (\hat{F}(x) - F(x))^2 \, dw(x) \quad ,$$

where w is a weight function. Thus $\hat{F}(t)$ is simply the posterior conditional expectation of $F(t)$ given the data; that is, we need to evaluate $E(F(t)|(Z_{(i)},\delta_i^*):$ $1 \le i \le k$, $Z_{(j)} > Z_{(k)}$, $j = k+1,\ldots,n)$, where E denotes expectation with respect to the Dirichlet process with parameter α. As argued in Susarla and Van Ryzin (1976) this may be accomplished in two stages. First relabel the data $\{(Z_{(i)},\delta_i^*): 1 \le i \le k\}$ as follows: let $Z^*_{(1)},\ldots,Z^*_{(\ell)}$ and $Z^*_{(\ell+1)},\ldots,Z^*_{(k)}$ denote, respectively, the ordered survival times and ordered censoring times recorded among $Z_{(1)},\ldots,Z_{(k)}$. Now consider a random sample of size ℓ, say $\eta_1,\ldots,\eta_\ell$ from a Dirichlet process ξ with parameter α and then a random sample size $(k-\ell)$, say $\eta_{\ell+1},\ldots,\eta_k$ from the conditional process of ξ given $\eta_1,\ldots,\eta_\ell$. Then this conditional process is itself a Dirichlet process with parameter β, with β given by

$$\beta(\cdot) = \alpha(\cdot) + \sum_{i=1}^{\ell} I_{\{\eta_i\}}(\cdot) \quad .$$

Therefore if the conditional process of F given $(Z^*_{(1)},1),\ldots,(Z^*_{(\ell)},1)$ is a Dirichlet process with parameter

$$\beta = \alpha + \sum_{i=1}^{\ell} I_{\{Z^*_{(i)}\}} \quad , \tag{2}$$

then the construction of our estimator $\hat{F}(t)$ reduces to the evaluation of $E(F(t)|(Z^*_{(i)},0): \ell < i \le k$, $Z_{(j)} > Z_{(k)}$, $j = k+1,\ldots,n)$, where E now denotes expectation with respect to the distribution of the Dirichlet process with parameter β of (2). This will be shown to reduce to

$$\hat{F}(t) = \frac{E(F(t)\ \{\prod_{i=\ell+1}^{k} F(Z^*_{(i)})\}\ \{F(Z_{(k)})\}^{n-k})}{E(\{\prod_{i=\ell+1}^{k} F(Z^*_{(i)})\}\ \{F(Z_{(k)})\}^{n-k})} \quad . \tag{3}$$

We shall defer the details of the evaluation of (3) to Section 4. The final form of $\hat{F}(t)$ can be written as

$$\hat{F}(t) = B(t)W(t) \tag{4}$$

where

$$B(t) = \frac{\alpha(t,\infty) + N^+(t) + (n-k)\ [t < Z_{(k)}]}{\alpha(R^+) + n}$$

$$W(t) = \prod_{j=1}^{k} \left\{ \frac{\alpha(Z_{(j)},\infty) + N^+(Z_{(j)}) + (n-k) + \lambda_j}{\alpha(Z_{(j)},\infty) + N^+(Z_{(j)}) + (n-k)} \right\}^{[Z_{(j)} \le t,\ \delta^*_j = 0]/\lambda_j}$$

$$\cdot \left\{ \frac{\alpha(Z_{(k)},\infty) + (n-k)}{\alpha(Z_{(k)},\infty)} \right\}^{[Z_{(k)} \le t]} \quad ,$$

and $[A]$ denotes the indicator of a set A. Also,

$$N^+(t) = \sum_{j=1}^{k} [Z_{(j)} > t] \quad ,$$

$$\alpha(R^+) = \alpha(0,\infty) \quad ,$$

λ_j = number of censored observations tied at $Z_{(j)}$, and $\frac{0}{0}$ in an exponent is interpreted as unity. It is easy to see that $\hat{F}$ is left continuous at censored observations provided the measure α has no atoms at these points.

Several special cases follow from (4):

(a) Suppose the entire data set $\{(Z_i,\delta_i): 1 \le i \le n\}$ is available. Then setting $k = n$ throughout we obtain

$$(5) \qquad \hat{F}(t) = \left\{\frac{\alpha(t,\infty) + N^+(t)}{\alpha(R^+) + n}\right\} \prod_{j=1}^{n} \left\{\frac{\alpha(Z_j,\infty) + N^+(Z_j) + \lambda_j}{\alpha(Z_j,\infty) + N^+(Z_j)}\right\}^{[Z_j \le t, \delta_j = 0]/\lambda_j}$$

which is the estimator given by Susarla and Van Ryzin (1976). It is also shown there that in the limit $\alpha(R^+) \to 0$, (5) reduces to the Kaplan-Meier product-limit estimator. If, however, we have only the partial data set $\{(Z_{(i)},\delta_i^*): 1 \le i \le k,\ Z_{(j)} > Z_{(k)},\ j = k+1,\ldots,n\}$ of a progressively censored sample, then for $t < Z_{(k)}$, the limit of (4) as $\alpha(R^+) \to 0$ is

$$(6) \qquad \frac{N^+(t) + (n-k)}{n} \prod_{j=1}^{k} \left\{\frac{N^+(Z_{(j)}) + (n-k) + \lambda_j}{N^+(Z_{(j)}) + (n-k)}\right\}^{[Z_{(j)} \le t, \delta_j^* = 0]/\lambda_j} .$$

Now writing $\dfrac{N^+(t) + (n-k)}{n} = \prod_{\{j: Z_{(j)} \le t\}} \left\{\dfrac{N^+(Z_{(j)}) + (n-k)}{N^+(Z_{(j)}) + (n-k) + \lambda_j^*}\right\}$ where λ_j^* is the multiplicity of $Z_{(j)}$, we find that (6) reduces to

$$(7) \qquad \prod_{j=1}^{k} \left\{\frac{N^+(Z_{(j)}) + (n-k)}{N^+(Z_{(j)}) + (n-k) + \lambda_j^*}\right\}^{[Z_{(j)} \le t, \delta_j^* = 1]}$$

If there are no ties among the uncensored observations this is precisely the product-limit estimator for $t < Z_{(k)}$. When $t \geq Z_{(k)}$ the behavior of $\hat{F}$ depends on α even in the limit and we cannot recover $\hat{F}$ since for any $M > 0$ one can choose measures α_1, α_2 which agree on $(0,M]$ but differ on (M,∞).

(b) Suppose that in addition to the entire set $\{(Z_i, \delta_i) : 1 \leq i \leq n\}$ being available there is no censoring present. Then setting $k = n$ and $[Z_{(j)} \leq t, \delta_i^* = 0] = 0$ in the terms following (4) we obtain

$$\hat{F}(t) = \frac{\alpha(t,\infty) + N^+(t)}{\alpha(R^+) + n} \quad , \tag{8}$$

which is the estimator of $F(t)$ proposed by Ferguson (1973). Again in the limit as $\alpha(R^+) \to 0$, (8) reduces to

$$\hat{F}(t) = \frac{N^+(t)}{n} = n^{-1} \sum_{j=1}^{n} [X_j > t] \quad , \tag{9}$$

which is the empirical survival function of $X_1, \ldots, X_n$.

On the other hand, if under a progressively censored scheme the only available data are $\{(Z_{(i)}, \delta_i^*) : 1 \leq i \leq k,\ Z_{(j)} > Z_{(k)},\ j = k+1, \ldots, n\}$, then the absence of observed censoring times among $Z_{(1)}, \ldots, Z_{(k)}$ reduces (4) to

$$\hat{F}(t) = \left\{ \frac{\alpha(t,\infty) + N^+(t) + (n-k)[t < Z_{(k)}]}{\alpha(R^+) + n} \right\} \left\{ \frac{\alpha(Z_{(k)},\infty) + (n-k)}{\alpha(Z_{(k)},\infty)} \right\}^{[Z_{(k)} \leq t]} . \tag{10}$$

We may thus regard (10) as the appropriate generalization of the Ferguson estimator (8) to the progressively censored case. Observe that if $t < Z_{(k)}$, the limit of (10) as $\alpha(R^+) \to 0$ is again the empirical survival function (9). For $t \geq Z_{(k)}$ this limit will depend on α and our previous remark in (a) applies.

With the restrictions noted here we depict the interrelation among the various estimators of F(t) diagrammatically in Figure 1:

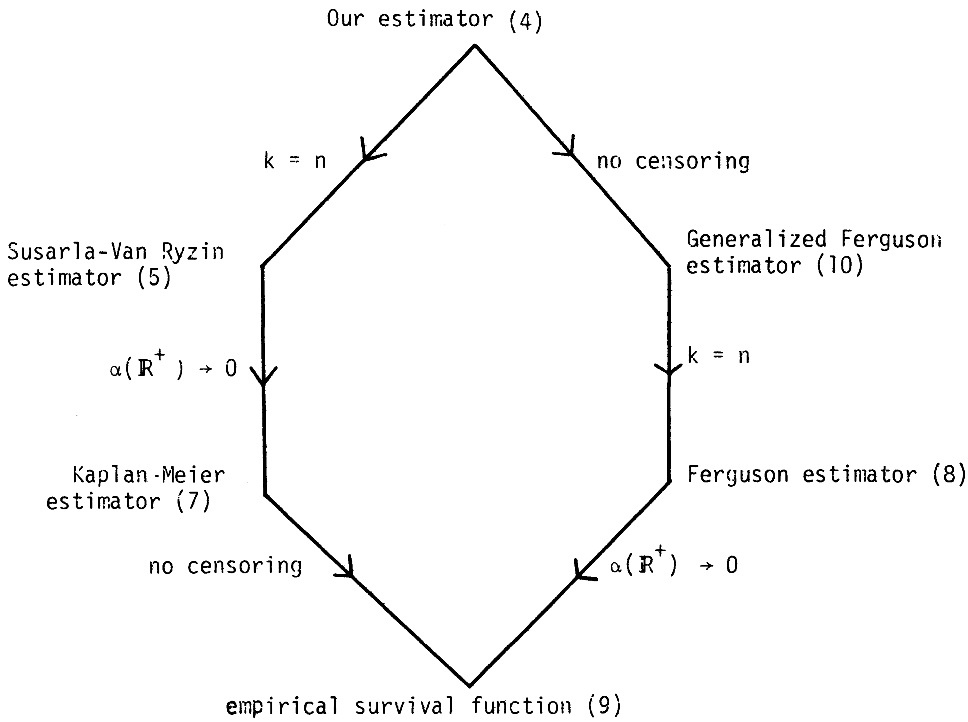

FIGURE 1: Various estimators of F(t)

3. A Numerical Example

We illustrate here the power of a progressively censored scheme with the partial data set $\{(Z_{(i)},\delta_i^*): 1 \leq i \leq k,\ Z_{(j)} > Z_{(k)},\ j = k+1,\ldots,n\}$ $(k < n)$ to yield results that are almost in agreement with those obtained when the complete survival profiles $\{(Z_i,\delta_i): 1 \leq i \leq n\}$ of the sample have been recorded. The data, taken from Johnson and Elandt-Johnson (1980) (p. 179) represent the survival times in weeks of 81 patients in a melanoma study conducted through the Central Oncology Group at the University of Wisconsin, Madison.

The censored survival times are indicated by a + sign.

136,	58,	55+,	181+,	21,	23,	190+,	65,	234,
194+,	14,	90,	20,	130,	213+,	215+,	124,	108+,
54,	98,	193+,	138,	141,	110,	67+,	50,	26,
103,	59,	134+,	147+,	152+,	65,	40,	34,	57,
81+,	152+,	125+,	151+,	34,	158+,	27,	148+,	27,
132+,	140+,	32,	130+,	38,	85,	129+,	100+,	19,
118,	53,	120+,	66,	46,	37,	50+,	114+,	124+,
26,	102,	93+,	80+,	60,	86+,	21+,	44+,	23,
70,	73+,	19,	38,	31,	25,	76+,	13,	16+,

We choose for our parameter α the measure generated through $\alpha(t,\infty) = \exp(-\theta t)$, $t \geq 0$ where $\theta > 0$ is a real parameter. Since from (1) $E(F(t)) = \alpha(t,\infty)/\alpha(R^+) = e^{-\theta t}$ (expectation with respect to the Dirichlet process F), we estimate θ by

$$\hat{\theta}_k = \sum_{i=1}^{k} \delta_i^* / \sum_{i=1}^{k} Z_{(i)} \quad .$$

A reason for this is that when the censoring times $\{Y_i : 1 \leq i \leq n\}$ are i.i.d. and the survival times are i.i.d. with survival distribution $F_0(t) = \exp(-\theta t)$ then $\{\hat{\theta}_{k_n} : n \geq 1\}$ converges in probability to θ when $n^{-1}k_n \to 1$ as $n \to \infty$. We have computed $\hat{F}(t)$ in three cases: 1) $k = n = 81$; 2) $k = 73$ and 3) $k = 65$. The agreement between the three curves is very good for time points $< Z_{(k)}$.

t =	25	44	54	65	76	100	148	190
1)	.89779	.74449	.69184	.61091	.59691	.54368	.35401	.31245
2)	.89736	.74382	.69107	.61005	.59597	.53361	.35259	.24209
3)	.89691	.74317	.69034	.60925	.59512	.53267	.34307	.22655

4. Proofs

Recall the notation introduced in Section 2. The proof that the conditional process of F given $\{(Z^*_{(i)},1): 1 \le i \le \ell\}$ is a Dirichlet process with parameter β as specified in (2), follows along exactly the same lines as that of Theorem 4 in Susarla and Van Ryzin (1976). In order to demonstrate (3) write

$$(11) \qquad E(F(t) \mid (Z^*_{(i)},0): \ell < i \le k) = \int_0^1 P[F(t) > a \mid (Z^*_{(i)},0): \ell < i \le k]da \quad ,$$

where E denotes (and in the sequel) expectation with respect to the Dirichlet process with parameter β. Now for $\ell < i \le k$, $\delta^*_i = 0$, $Z^*_{(i)} = X^*_i \wedge Y^*_i = Y^*_i$ so that $X^*_i > Z^*_{(i)}$ and for $k < j \le n$ of course X^*_j, $Y^*_j > Z_{(k)}$ where $Z_{(j)} = X^*_j \wedge Y^*_j$. Therefore the integrand in (11) may be written as the ratio I_1/I_2 where

$$(12) \qquad I_1 = E(P[F(t) > a,\ X^*_i \varepsilon (Z^*_{(i)},\infty),\ \ell < i \le k,\ X^*_j,\ Y^*_j \varepsilon (Z_{(k)},\infty),\ k < j \le n \mid F(t),\ F(Z_{(k)}),\ F(Z^*_{(i)}),\ \ell < i \le k]) \quad ,$$

and I_2 is the resulting expectation obtained by suppressing both F(t) and "$F(t) > a$" in (12). Since $(Y_1,\ldots,Y_n)$ is independent of $(F, X_1,\ldots,X_n)$ and including the terms $Z^*_{(i)} \equiv X^*_i \varepsilon (0,\infty)$ (when $\delta^*_i = 1$), $1 \le i \le \ell$ we get on simplification

$$(13) \qquad I_1 = E([F(t) > a] \cdot P[X^*_i \varepsilon (0,\infty),\ 1 \le i \le \ell,\ X^*_i \varepsilon (Z^*_{(i)},\infty),\ \ell < i \le k\ ,\ X^*_j \varepsilon (Z_{(k)},\infty),\ k < j \le n \mid F(t),\ F(Z_{(k)}),\ F(Z^*_{(i)}),\ \ell < i \le k]) \cdot P[Y^*_j \varepsilon (Z_{(k)},\infty),\ k < j \le n] \quad .$$

Likewise I_2 is obtained from (13) by suppressing $[F(t) > a]$ and $F(t)$. Since $\{X_1,\ldots,X_n\}$ is a random sample from F, the inner conditional probability in (13) is almost surely

$$\text{(14)} \qquad \{\prod_{i=\ell+1}^{k} F(Z^*_{(i)})\}\,\{F(Z_{(k)})\}^{n-k} \quad .$$

Finally, using (14) in I_1/I_2 and carrying out the integration in (11) we obtain (3).

We are now left with the tedious task of carrying out the integrations in (3). Several cases must be considered depending on the position of the time point t among the observed points $Z^*_{(\ell+1)},\ldots,Z^*_{(k)}$ and $Z_{(k)}$ $(=\max\{Z_{(i)}: 1 \leq i \leq k\})$.

Suppose $Z^+_{(\ell+1)},\ldots,Z^+_{(m)}$ denote the distinct ordered values among $Z^*_{(\ell+1)},\ldots$., $Z^*_{(k)}$ with corresponding multiplicities $\lambda_{\ell+1},\ldots,\lambda_m$. Thus

$$\lambda_i \geq 1,\ \ell < i \leq m; \quad 1 \leq \ell \leq k \leq n \quad \text{and} \quad \ell < m \leq k,\ \sum_{i=\ell+1}^{n} \lambda_i = k-\ell \quad .$$

The largest recorded observable $Z_{(k)}$ may be either a survival time or a censoring time. Suppose we are in the latter case so that $Z_{(k)} = Z^+_{(m)}$. Consider the case $t > Z^+_{(m)}$. Select the partition of $R^+ = (0,\infty)$ given by the points $\{Z^+_{(i)}: \ell \leq i \leq m+2\}$ where $Z^+_{(\ell)} = 0, Z^+_{(m+1)} = t$ and $Z^+_{(m+2)} = \infty$. Then defining $U_i = F(Z^+_{(i)}) - F(Z^+_{(i+1)}), \ell < i \leq m+1$, the random vector $(U_\ell,\ldots,U_{m+1})$ has the Dirichlet distribution with parameter $(\beta_\ell,\ldots,\beta_{m+1})$ where

$$\text{(15)} \qquad \beta_i = \beta(Z^+_{(i)},\ Z^+_{(i+1)}]$$

and β as given in (2).

Now $F(Z^+_{(i)}) = (1 - \sum_{j=\ell}^{i-1} U_j)$, $\ell < i \leq m+1$. Therefore the integrand in the numerator of (3), $F(t)\ \{\prod_{i=\ell+1}^{m} (F(Z^+_{(i)}))^{\lambda_i}\}\ \{F(Z^+_{(m)})^{n-k}\}$, can be written

$$(16) \qquad \prod_{i=\ell+1}^{m+1} \left(1 - \sum_{j=\ell}^{i-1} U_j\right)^{\lambda_i} \left(1 - \sum_{j=\ell}^{m-1} U_j\right)^{n-k} ,$$

with $\lambda_{m+1} = 1$. For the denominator in (3) the integrand is the same as (16) except that $\lambda_{m+1} = 0$.

Now $(U_\ell,.,,,U_m)$ has probability density

$$\left\{ \frac{\Gamma\left(\sum_{i=\ell}^{m+1} \beta_i\right)}{\prod_{i=\ell}^{m+1} \Gamma(\beta_i)} \right\} \left\{ \prod_{j=\ell}^{m} u_j^{\beta_j - 1} \right\} \left(1 - \sum_{j=\ell}^{m} u_j\right)^{\beta_{m+1}-1} ,$$

$$0 < \text{all } u_j < 1$$

$$\text{and} \quad 0 < \sum_{j=\ell}^{m} u_j < 1 ,$$

where Γ denotes the Gamma function. The expectation of (16) involves integration over the variables $u_\ell,\ldots,u_m$. Suppressing all terms not involving u_m, the integral over u_m is

$$\int_0^{(1 - \sum_{j=\ell}^{m-1} u_j)} u_m^{\beta_m - 1} \left(1 - \sum_{j=\ell}^{m-1} u_j - u_m\right)^{(\beta_{m+1} + \lambda_{m+1}) - 1} du_m$$

$$= \frac{\Gamma(\beta_m)\, \Gamma(\beta_{m+1} + \lambda_{m+1})}{\Gamma(\beta_m + \beta_{m+1} + \lambda_{m+1}} \left(1 - \sum_{j=\ell}^{m-1} u_j\right)^{\beta_m + \beta_{m+1} + \lambda_{m+1} - 1} .$$

Now proceeding with the successive integrations over $u_{m-1},\ldots,u_\ell$ we finally obtain for the numerator of (3)

$$(17)\quad \left\{\frac{\Gamma\left(\sum_{i=\ell}^{m+1}\beta_i\right)}{\prod_{i=\ell}^{m+1}\Gamma(\beta_i)}\right\}\left\{\frac{\Gamma(\beta_m)\,\Gamma(\beta_{m+1}+\lambda_{m+1})}{\Gamma(\beta_m+\beta_{m+1}+\lambda_{m+1})}\right\}\prod_{i=\ell}^{m-1}\left\{\frac{\Gamma(\beta_i)\,\Gamma\left(\sum_{j=i+1}^{m+1}(\beta_j+\lambda_j)+(n-k)\right)}{\Gamma\left(\beta_i+\sum_{j=i+1}^{m+1}(\beta_j+\lambda_j)+(n-k)\right)}\right\}.$$

Note that in (17) the value of λ_{m+1} is 1. For the denominator of (3) we have the same expression except that $\lambda_{m+1} = 0$.

Proceeding with cancellations of the common factors in the numerator and denominator of (3) yields our estimator

$$(18)\qquad \hat{F}(t) = \left(\frac{\beta_{m+1}}{\beta_m+\beta_{m+1}}\right)\prod_{j=\ell+1}^{m}\left\{\frac{\sum_{i=j}^{m}(\beta_i+\lambda_i)+\beta_{m+1}+(n-k)}{\sum_{i=j}^{m}(\beta_i+\lambda_i)+\beta_{m+1}+(n-k)+\beta_{j-1}}\right\}.$$

Recall (15) and (2). A trite computation shows

$$\beta_{m+1} = \alpha(t,\infty) + \#\text{ (observed lifetimes} > t)$$

$$\beta_m + \beta_{m+1} = \alpha(Z^+_{(m)},\infty) + \#\text{ (observed lifetimes} > Z^+_{(m)})$$

$$(19)\quad \sum_{i=j}^{m}(\beta_i+\lambda_i)+\beta_{m+1} = \alpha(Z^+_{(j)},\infty) + \#\text{ (observed lifetimes} > Z^+_{(j)})$$

$$+ \#\text{ (observed censoring times} \geq Z^+_{(j)})$$

$$= \alpha(Z^+_{(j)},\infty) + N^+(Z^+_{(j)}) + \lambda_j$$

$$\sum_{i=j}^{m}(\beta_i+\lambda_i)+\beta_{m+1}+\beta_{j-1} = \alpha(Z^+_{(j-1)},\infty) + N^+(Z^+_{(j-1)}) \quad .$$

Substituting in (18) and again cancelling out common factors we get for the product term in (18)

$$(20)\qquad \frac{\alpha(Z^+_{(m)},\infty) + N^+(Z^+_{(m)}) + (n-k) + \lambda_m}{\alpha(Z^+_{(\ell)},\infty) + N^+(Z^+_{(\ell)}) + (n-k)} \prod_{j=\ell+1}^{m} \left\{ \frac{\alpha(Z^+_{(j)},\infty) + N^+(Z^+_{(j)}) + (n-k) + \lambda_j}{\alpha(Z^+_{(j)},\infty) + N^+(Z^+_{(j)}) + (n-k)} \right\}.$$

Now $Z^+_{(m)} = Z_{(k)}$ and $Z^+_{(\ell)} = 0$. Also in the case considered here $N^+(Z^+_{(m)}) = 0$, $N^+(t) = 0$. Using these in (19) and (20) yields

$$\hat{F}(t) = \frac{\alpha(t,\infty)}{\alpha(R^+) + n} \prod_{j=\ell+1}^{m} \left\{ \frac{\alpha(Z^+_{(j)},\infty) + N^+(Z^+_{(j)}) + (n-k) + \lambda_j}{\alpha(Z^+_{(j)},\infty) + N^+(Z^+_{(j)}) + (n-k)} \right\} \cdot \left\{ \frac{\alpha(Z_{(k)},\infty) + (n-k)}{\alpha(Z_{(k)},\infty)} \right\}$$

which is the form of (4) for this case.

All other cases are handled in exactly the same manner and lead to the general form of $\hat{F}(t)$ given in (4).

5. Concluding Remarks

It can be shown that when the censoring times $\{Y_i: i \geq 1\}$ are i.i.d. with continuous right distribution function G on $(0,\infty)$, the survival times $\{X_i: i \geq 1\}$ are i.i.d. with continuous right distribution F, and $n^{-1}k_n \to \gamma \varepsilon (0,1]$, then for any $T > 0$ with $F(T)G(T) > 1-\gamma$, the process $\{n^{\frac{1}{2}}(\hat{F}(t) - F(t)): t \varepsilon [0,T\}$ converges weakly to a Gaussian process. Furthermore under appropriate conditions strong convergence and consistency can be demonstrated.

ACKNOWLEDGEMENTS

The authors wish to thank Professor P.K. Wong for his programming assistance. Research sponsored in part by the National Institutes of Health under Grants NIGMS-1R01-GM28030 and NIGMS-1R01-GM28405, and in part by the Office of Naval Research under contract N00014-79-C0522.

REFERENCES

Breslow, N. and Crowley, J. (1974). A large sample study of the life table and product limit estimates under random censorship. Annals of Statistics 2, 437-453.

Chatterjee, S.K. and Sen, P.K. (1973). Nonparametric testing under progressive censoring. Calcutta Statistical Association Bulletin 22, 13-50.

Delong, E. and Sen, P.K. (1981). Estimation of $P[X > Y]$ based on progressively truncated versions of the Wilcoxon-Mann-Whitney statistics. Communications in Statistics A10 (10), 963-981.

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

Gardiner, J. and Sen, P.K. (1978). Asymptotic normality of a class of time-sequential statistics and applications. Communications in Statistics A7 4, 373-388.

Johnson, N. and Elandt-Johnson, R. (1980). Survival Models and Data Analysis. Wiley & Sons, New York.

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

Susarla, V. and Van Ryzin, J. (1976). Nonparametric Bayesian estimation of survival curves from incomplete observations. Journal of the American Statistical Association 61, 897-902.