NONPARAMETRIC BIVARIATE ESTIMATION WITH RANDOMLY CENSORED DATA

By

Gregory Campbell
Purdue University

Department of Statistics
Division of Mathematical Sciences
Mimeoseries #79-25

September 1979

NONPARAMETRIC BIVARIATE ESTIMATION WITH RANDOMLY CENSORED DATA

by

Gregory Campbell
Purdue University

ABSTRACT

The problem of estimation of a bivariate distribution function with randomly censored data is considered. It is assumed that the censoring occurs independently of the lifetimes, and that deaths and losses which occur simultaneously can be separated. Two estimators are developed: a reduced-sample (RS) estimator and a self-consistent (SC) one. It is shown that the SC estimator satisfies a nonparametric "likelihood" function and is unique up to the final censored values in any dimension; it jumps at the points of double deaths in both dimensions. The two estimators are compared. An example is presented illustrating the estimates in a reliability setting.

AMS 1970 subject classification. Primary 62G05.

Key words and phrases. Kaplan-Meier estimator, product-limit estimator, censored data, survival estimation, life tables.

NONPARAMETRIC BIVARIATE ESTIMATION WITH RANDOMLY CENSORED DATA

by

Gregory Campbell
Purdue University

1. Introduction and summary. The nonparametric estimation of a bivariate distribution function using randomly censored data is treated. For each pair of items, one attempts to observe the lifelength of each item. However, there is a random mechanism which can censor the data before the lifelengths have been observed. The problem is to estimate the bivariate distribution function, or, equivalently, the bivariate survival function in the presence of censored data.

There are numerous situations where bivariate estimation is important. Times to death or times to initial contraction of a disease may be of interest for littermate pairs of rats or for twin studies in humans. The time to a deterioration level or the time to reaction of a treatment may be of interest in pairs of lungs, kidneys, eyes, or ears of humans. In demography, there are the lifelengths of married couples. In reliability applications, the distribution of the lifelengths of a particular pair of components in a system may be of interest. The censoring may arise for a number of reasons. The items may withdraw from the study, they may be withdrawn due to a change of health status or contamination. They may be censored by death from a cause unrelated to the study. The censoring may also be due to random entry into the experiment and then truncation of the experiment at a fixed time.

This paper reviews in Section 2 the one-dimensional survival estimation problem, focussing on the nonparametric approach of Kaplan and Meier (1958) and the related self-consistency technique of Efron (1967). In Section 3, two bivariate estimators are proposed: a reduced-sample (RS) estimator and a self-consistent (SC) one. Conditions for the existence and uniqueness of the latter estimator are presented. The relative merits of the estimators are discussed. The final section illustrates the calculation of the estimators for a reliability example.

2. <u>One-dimensional survival estimation</u>. Let X_i^0 $(i=1,\ldots,n)$ be independent, identically distributed (i.i.d.) random variables from a continuous survival distribution function $F(x) = P(X^0 > x)$. Let C_i $(i=1,\ldots,n)$ be i.i.d. censoring variables (observation limits) from the survival distribution $H(x) = P(C > x)$. The pair (X_i, ε_i) $(i=1,\ldots,n)$ are observable, where

$$X_i = \min(X_i^0, C_i); \qquad \varepsilon_i = \begin{cases} 1 & \text{if } X_i = X_i^0 \quad \text{(uncensored)} \\ 0 & \text{if } X_i < X_i^0 \quad \text{(censored)}. \end{cases}$$

Let the survival distribution of the X's be denoted by $G(x) = P(X > x)$. It is assumed that the C_i's are independent of the X_i's; i.e.,

$$G(x) = F(x)\,H(x). \tag{1}$$

Kaplan and Meier (1958) develop two estimators of $F(x)$. The first, called the reduced-sample (RS) estimator, requires additionally that all the C_i's are observable. It is given by:

$$\tilde{F}(t) = N(t)/N^0(t), \tag{2}$$

where $N(t)$ is the number of items observed and surviving at time t where deaths but not losses at time t are excluded, and where $N^0(t)$ is the number of items for which $C_i \geq t$.

For the second estimator of Kaplan and Meier, assume initially that the times s_i of deaths are fixed and discrete $(i=1,\ldots,I)$. Then the estimator of F(t) proposed by Kaplan and Meier is the product of estimated conditional probabilities:

$$\hat{F}(t) = \prod_{i \leq j} \hat{p}_i \quad \text{for } t \in [s_j, s_{j+1}) \quad (j=1,\ldots,I-1),$$

where $p_i = N_i/M_i$ for M_i the number of $X_j \geq t$ and N_i the number of $X_j \geq s_i$ excluding the deaths at s_i. The limiting estimator as I tends to infinity such that $\max|s_{i+1}-s_i|$ tends to zero is called the Kaplan-Meier product-limit estimator. They show that it is the nonparametric maximum likelihood estimator (i.e., it maximizes the probability or "likelihood" function). This estimator is not defined for any losses which follow the final death. Efron (1967) shows that the product-limit (PL) estimate is the unique solution of the following self-consistency equation

$$n\hat{F}(s_j) = N_j + \sum_{i<j} (1-\varepsilon_i) \frac{\hat{F}(s_j)}{\hat{F}(s_i)}, \tag{3}$$

where N_j is the number of $X_i \geq s_j$ excluding deaths at s_j.

The estimator which satisfies (3) has a closed-form solution, but this will not be the case for the bivariate situation. Consequently, the following result which describes an iterative solution will prove useful in the bivariate treatment of Section 3.

<u>Theorem 1</u>. Let $F_j^{(0)} = N_j/n$ and define $F_j^{(k)}$ iteratively as

$$nF_j^{(k)} = nF_j^{(0)} + \sum_{i<j} (1-\varepsilon_i) \frac{F_j^{(k-1)}}{F_i^{(k-1)}} \qquad (k=1,2,\ldots).$$

Then the sequence $\{F_j^{(k)}\}_{k=1}^{\infty}$ is non-decreasing sequence which converges to the Kaplan-Meier solution $\hat{F}(s_j)$ of equation (3).

Proof. To show that $F_j^{(k)}$ is increasing in k, it suffices to show that $F_j^{(k)}/F_i^{(k)}$ $(i < j)$ is non-decreasing in k. This follows by induction. Note that, for k = 0 and 1, $\frac{F_j^{(0)}}{F_i^{(0)}} \leq \frac{F_j^{(1)}}{F_i^{(1)}}$ since $\sum_{k<i} (1-\varepsilon_k)\frac{1}{N_k} \leq \sum_{k<j} (1-\varepsilon_k)\frac{1}{N_k}$ for $i < j$. Assume that $F_j^{(k)}/F_i^{(k)}$ is increasing up to and including $k=\ell$. Then it suffices to show that $F_j^{(\ell)}/F_i^{(\ell)} \leq F_j^{(\ell+1)}/F_i^{(\ell+1)}$. Now

$$F_j^{(\ell)} = F_j^{(0)} + \sum_{i<j} (1-\varepsilon_i)\frac{F_j^{(\ell-1)}}{F_i^{(\ell-1)}}$$

and

$$F_j^{(\ell+1)} = F_j^{(0)} + \sum_{i<j} (1-\varepsilon_i)\frac{F_j^{(\ell)}}{F_i^{(\ell)}} = F_j^{(\ell)} + \sum_{i<j} (1-\varepsilon_i)\left[\frac{F_j^{(\ell)}}{F_i^{(\ell)}} - \frac{F_j^{(\ell-1)}}{F_i^{(\ell-1)}}\right].$$

Then

$$\frac{F_j^{(\ell+1)}}{F_i^{(\ell+1)}} - \frac{F_j^{(\ell)}}{F_i^{(\ell)}} = \{F_i^{(\ell)}(F_j^{(\ell)} + \sum_{k<j} (1-\varepsilon_k)\left[\frac{F_j^{(\ell)}}{F_k^{(\ell)}} - \frac{F_j^{(\ell-1)}}{F_k^{(\ell-1)}}\right])$$

$$- F_j^{(\ell)}(F_i^{(\ell)} + \sum_{k<i} (1-\varepsilon_k)\left[\frac{F_i^{(\ell)}}{F_k^{(\ell)}} - \frac{F_i^{(\ell-1)}}{F_k^{(\ell-1)}}\right])\}/F_i^{(\ell)}F_j^{(\ell)}$$

$$= \{F_i^{(\ell)}F_j^{(\ell)} \sum_{i\leq k<j} \frac{1-\varepsilon_k}{F_k^{(\ell)}} + (F_j^{(\ell)}F_i^{(\ell-1)} - F_i^{(\ell)}F_j^{(\ell-1)}) \cdot$$

$$\sum_{k<i} \frac{1-\varepsilon_k}{F_k^{(\ell-1)}} - F_i^{(\ell)}F_j^{(\ell-1)} \sum_{i\leq k<j} \frac{1-\varepsilon_k}{F_k^{(\ell-1)}}\}/F_i^{(\ell)}F_j^{(\ell)}$$

$$= \{F_i^{(\ell)} \sum_{i\le k<j} (1-\varepsilon_k)\left[\frac{F_j^{(\ell)}}{F_k^{(\ell)}} - \frac{F_j^{(\ell-1)}}{F_k^{(\ell-1)}}\right] + \left[\frac{F_j^{(\ell)}}{F_i^{(\ell)}} - \frac{F_j^{(\ell-1)}}{F_i^{(\ell-1)}}\right] F_i^{(\ell)} F_i^{(\ell-1)} .$$

$$\sum_{k<i} \frac{1-\varepsilon_k}{F_k^{(\ell-1)}} \}/F_i^{(\ell)} F_j^{(\ell)} .$$

Now apply the induction hypothesis twice to the right-hand side to conclude that it is nonnegative. Therefore, the sequence $F_j^{(k)}$ is nondecreasing. Since it is also bounded, the limit of $F_j^{(k)}$ (as $k\to\infty$) exists; call it $\hat{F}_j$. Then it is clear, since $\hat{F}_j$ satisfies (3), that $\hat{F}_j = \hat{F}(s_j)$. □

Note that the proof does not hinge on the initial selection $(F_j^{(0)} = N_j/n)$ but works equally well for any $F_j^{(0)}$ and that $F_j^{(0)} \le \hat{F}(s_j)$.

Numerous authors have continued work in this area of one-dimensional survival estimation. Breslow and Crowley (1974) supplied the details for the asymptotic behavior of these estimates of Kaplan and Meier. Efron (1967) applied the product-limit, self-consistent estimator to the two-sample testing problem. J. Sander Chmiel (1975) has studied asymptotic behavior of functions of the Kaplan-Meier PL estimate. Turnbull (1974) looked at the doubly censored data problem, and Meier (1975) extended the theory to the more general censoring situation than the random censoring mechanism.

3. Bivariate survival estimators. Let $\{(X_i^0, Y_i^0)\}_{i=1}^n$ be i.i.d. pairs of random (lifelength) variables from the continuous joint survival distribution $F(s,t) = P(X^0 > s, Y^0 > t)$. Let $\{(C_i, D_i)\}_{i=1}^n$ be i.i.d. pairs of censoring variables from continuous joint distribution $H(s,t) = P(C > s, D > t)$. The variables X_i, Y_i, ε_i, ε_i' are observed, for $i=1,\dots,n$:

$$X_i = \min(X_i^0, C_i); \qquad \varepsilon_i = \begin{cases} 1 & \text{if } X_i = X_i^0 \\ 0 & \text{if } X_i < X_i^0 ; \end{cases}$$

$$Y_i = \min\,(Y_i^0, D_i); \qquad \varepsilon_i' = \begin{cases} 1 & \text{if } Y_i = Y_i^0 \\ 0 & \text{if } Y_i < Y_i^0 . \end{cases}$$

Let the distribution function of (X,Y) be denoted $G(s,t) = P(X > s,\ Y > t)$. It is assumed that (C_i, D_i) is independent of (X_i^0, Y_i^0) for all $i=1,\ldots,n$; that is, the censoring takes place independently of the lifetimes. Thus,

$$G(s,t) = F(s,t)\ H(s,t). \tag{4}$$

For the bivariate reduced-sample estimate, assume that not only are X_i, Y_i, ε_i, ε_i' available for each pair, but also C_i and D_i (even if both members of the pair may have failed prior to censoring). The natural extension of the RS estimate of Kaplan and Meier is based on solving equation (4) in terms of F(s,t) and then estimating G(s,t) and H(s,t). In particular, the RS estimate is:

$$\tilde{F}(s,t) = G_n(s,t)/H_n^-(0,t) \tag{5}$$

where $nH_n^-(s,t)$ is the number of pairs such that $C_i \geq s$ and $D_i \geq t$ and $nG_n(s,t)$ is the number of pairs for which $X_i^0 > s$, $Y_i^0 > t$ with $C_i \geq s$ and $D_i \geq t$ (i.e., $nG_n(s,t)$ is the number alive at (s,t), excluding deaths at s in the first or at t in the second).

There are two main disadvantages of this RS estimator. First, there are relatively few situations in which one enjoys the luxury of observing the censoring variables when the lifetime precede them. Secondly, there is no guarantee that the point estimate in (5) based on a binomial model with random number of trials is indeed a bivariate distribution function as s and t range over all values. In particular, it is possible to construct examples for which the probability of a rectangle is negative (see Section 4).

The second estimator is a bivariate self-consistent estimate. Assume initially that the possible times for losses and deaths occur at discrete points. Let $s_1,\ldots,s_I$ be the distinct times for the first item of the pair, and $t_1,\ldots,t_J$ be the distinct times for the second item. Define δ_{ij}, α_{ij}, β_{ij}, and λ_{ij} as follows:

δ_{ij} = number of pairs for which $X_k=s_i$, $Y_k=t_j$, $\varepsilon_k=1$, $\varepsilon'_k=1$ (double death);

α_{ij} = number of pairs for which $X_k=s_i$, $Y_k=t_j$, $\varepsilon_k=0$, $\varepsilon'_k=1$ (censored in first coordinate, death in second);

β_{ij} = number of pairs for which $X_k=s_i$, $Y_k=t_j$, $\varepsilon_k=1$, $\varepsilon'_k=0$ (death in first, censored in second);

λ_{ij} = number of pairs for which $X_k=s_i$, $Y_k=t_j$, $\varepsilon_k=0$, $\varepsilon'_k=0$ (double loss).

Just as in the one-dimensional case, it is necessary to separate deaths and losses which occur at any point (s_i,t_j). The convention adopted here is that deaths precede losses in each dimension. Therefore, in the first coordinate, the pairs δ_{ij} and β_{ij} precede α_{ij} and λ_{ij}; in the second coordinate δ_{ij} and α_{ij} precede β_{ij} and λ_{ij}. Pictorially, this can be represented by Figure 1. There is intuitive appeal for this convention in that if a loss occurs at time t, the item can be assumed to be surviving at the time of the censoring.

The problem is to estimate $F_{ij} = F(s_i,t_j)$ for $i=1,\ldots,I$, $j=1,\ldots,J$. Consider the maximum likelihood criterion extended to the bivariate case. The nonparametric "likelihood" or, more correctly, probability function is

$$L = \prod_{i=1}^{I} \prod_{j=1}^{J} \Delta_{ij}^{\delta_{ij}} F_{ij}^{\lambda_{ij}} Q_{ij}^{\alpha_{ij}} R_{ij}^{\beta_{ij}} \tag{6}$$

where $\Delta_{ij} = F_{ij} + F_{i-1,j-1} - F_{i,j-1} - F_{i-1,j}$ is the probability of death in the rectangle $(s_{i-1},s_i] \times (t_{j-1},t_j]$, $Q_{ij} = F_{i,j-1} - F_{ij}$ is the probability of death in $(s_i,\infty) \times (t_{j-1},t_j]$, and $R_{ij} = F_{i=1,j} - F_{ij}$ is the probability of death in $(s_{i-1},s_i] \times (t_j,\infty)$. In order to maximize L by choice of F_{ij} for fixed values of δ_{ij}, λ_{ij}, α_{ij}, and β_{ij}, differentiate $\ell n L$ with respect to F_{ij} and set equal to zero to get the "likelihood" equation

$$\frac{\delta_{ij}}{\Delta_{ij}} + \frac{\delta_{i+1,j+1}}{\Delta_{i+1,j+1}} - \frac{\delta_{i,j+1}}{\Delta_{i,j+1}} - \frac{\delta_{i+1,j}}{\Delta_{i+1,j}} + \frac{\lambda_{ij}}{F_{ij}} + \frac{\beta_{i+1,j}}{R_{i+1,j}}$$
$$- \frac{\beta_{ij}}{R_{ij}} + \frac{\alpha_{i,j+1}}{Q_{i,j+1}} - \frac{\alpha_{ij}}{Q_{ij}} = 0 \tag{7}$$

Then solutions to (7) are possible maximum "likelihood" estimates.

Consider a self-consistent approach. At time (s_i,t_j), to estimate F_{ij}, one would surely count all N_{ij} pairs that are known to be alive at (s_i,t_j) (so exclude deaths at s_i or at t_j but not losses). Now for the $\alpha_{k\ell}$ pairs that were censored in the first coordinate at s_k but died in the second coordinate at time t_ℓ, for $k < i$ and $\ell > j$, the expected number to survive to (s_i,t_ℓ) is $\alpha_{k\ell}\,\frac{Q_{i\ell}}{Q_{k\ell}}$. Similarly for the $\beta_{k\ell}$ pairs censored in the second coordinate for $k > i$, $\ell < j$, the expected number to survive to (s_k,t_j) is $\beta_{k\ell}\,\frac{R_{kj}}{R_{k\ell}}$. Finally, of the $\lambda_{k\ell}$ doubly censored pairs for $k < i$ <u>or</u> $\ell < j$, the expected number to survive to (s_i,t_j) is $\lambda_{k\ell}\,\frac{F_{\max(i,k),\max(j,\ell)}}{F_{k\ell}}$. This argument leads to the bivariate self-consistency equation:

$$n\hat{F}_{ij} = N_{ij} + \sum_{\substack{\ell>j \\ k<i}} \alpha_{k\ell} \frac{\hat{Q}_{i\ell}}{\hat{Q}_{k\ell}}$$

$$+ \sum_{\substack{k>i \\ \ell<j}} \beta_{k\ell} \frac{\hat{R}_{kj}}{\hat{R}_{k\ell}} + \sum_{\substack{k<i \\ \text{OR} \\ \ell<j}} \lambda_{k\ell} \frac{\hat{F}_{\max(i,k),\max(j,\ell)}}{\hat{F}_{k\ell}}, \tag{8}$$

where $N_{ij} = \sum_{\substack{k>i \\ \ell>j}} \delta_{k\ell} + \sum_{\substack{k\geq i \\ \ell>j}} \alpha_{k\ell} + \sum_{\substack{k>i \\ \ell\geq j}} \beta_{k\ell} + \sum_{\substack{k\geq i \\ \ell\geq j}} \lambda_{k\ell}$, $\hat{Q}_{ij} = \hat{F}_{i,j-1} - \hat{F}_{ij}$, and

$\hat{R}_{ij} = \hat{F}_{i-1,j} - \hat{F}_{ij}$.

Theorem 2. An estimate F_{ij} satisfying (8) also is a solution of equation (7) with F_{ij} replaced by $\hat{F}_{ij}$ and Δ_{ij} by $\hat{\Delta}_{ij} = \hat{F}_{ij} + \hat{F}_{i-1,j-1} - \hat{F}_{i-1,j} - \hat{F}_{i,j-1}$.

Proof. From equation (8)

$$n\hat{R}_{ij} = n(\hat{F}_{i-1,j}-\hat{F}_{ij}) = \sum_{\ell>j} \delta_{i\ell} + \sum_{\ell>j} \alpha_{i-1,\ell} + \sum_{\ell\geq j} \beta_{i\ell} + \sum_{\ell\geq j} \lambda_{i-1,\ell}$$

$$- \sum_{\ell>j} \alpha_{i-1,\ell} \frac{\hat{Q}_{i\ell}}{\hat{Q}_{i-1,\ell}} + \sum_{\substack{\ell>j \\ k<i-1}} \alpha_{k\ell} \frac{\hat{Q}_{i-1,\ell}-\hat{Q}_{i\ell}}{\hat{Q}_{k\ell}} + \sum_{\ell<j} \beta_{i\ell} \frac{\hat{R}_{ij}}{\hat{R}_{i\ell}} \tag{9}$$

$$+ \sum_{\substack{k\leq i-1 \\ \ell\leq j-1 \\ \text{exc.}(i-1,j-1)}} \lambda_{k\ell} \frac{\hat{R}_{ij}}{\hat{F}_{k\ell}} + \sum_{\substack{k<i-1 \\ \ell>j}} \lambda_{k\ell} \frac{\hat{R}_{i\ell}}{\hat{F}_{k\ell}} - \sum_{\ell\geq j} \lambda_{i-1,\ell} \frac{\hat{F}_{i\ell}}{\hat{F}_{i-1,\ell}}.$$

Form an equation for $n\hat{R}_{i,j-1}$ in a similar fashion and subtract equation (9) from it to obtain:

$$n(\hat{R}_{i,j-1}-\hat{R}_{ij}) = n\hat{\Delta}_{ij} = \delta_{ij} + \sum_{k\leq i-1} \alpha_{kj} \frac{\hat{Q}_{i-1,j}-\hat{Q}_{ij}}{\hat{Q}_{kj}}$$

$$+ \sum_{\ell\leq j-1} \beta_{i\ell} \frac{\hat{R}_{i,j-1}-\hat{R}_{ij}}{\hat{R}_{i\ell}} + \sum_{\substack{k\leq i-1\\ \ell\leq j-1}} \lambda_{k\ell} \frac{\hat{R}_{i,j-1}-\hat{R}_{ij}}{\hat{F}_{k\ell}} . \tag{10}$$

Divide equation (9) by $\hat{\Delta}_{ij} = \hat{R}_{i,j-1}-\hat{R}_{ij} = \hat{Q}_{i-1,j}-\hat{Q}_{ij}$:

$$n = \frac{\delta_{ij}}{\hat{\Delta}_{ij}} + \sum_{k\leq i-1} \frac{\alpha_{kj}}{\hat{Q}_{kj}} + \sum_{\ell\leq j-1} \frac{\beta_{i\ell}}{\hat{R}_{i\ell}} + \sum_{\substack{k\leq i-1\\ \ell\leq j-1}} \frac{\lambda_{k\ell}}{\hat{F}_{k\ell}} \tag{11}$$

Form three other equations similar to (11) based on $\hat{\Delta}_{i+1,j}$, $\hat{\Delta}_{i,j+1}$, $\hat{\Delta}_{i+1,j+1}$ instead of the $\hat{\Delta}_{ij}$ of (11). Add equations based on $\hat{\Delta}_{ij}$ and $\hat{\Delta}_{i+1,j+1}$ and subtract those based on $\hat{\Delta}_{i+1,j}$ and $\hat{\Delta}_{i,j+1}$:

$$0 = \frac{\delta_{ij}}{\hat{\Delta}_{ij}} + \frac{\delta_{i+1,j+1}}{\hat{\Delta}_{i+1,j+1}} - \frac{\delta_{i,j+1}}{\hat{\Delta}_{i,j+1}} - \frac{\delta_{i+1,j}}{\hat{\Delta}_{i+1,j}} + \frac{\lambda_{ij}}{\hat{F}_{ij}}$$

$$+ \frac{\beta_{i+1,j}}{\hat{R}_{i+1,j}} - \frac{\beta_{ij}}{\hat{R}_{ij}} + \frac{\alpha_{i,j+1}}{\hat{Q}_{i,j+1}} - \frac{\alpha_{ij}}{\hat{Q}_{ij}} . \tag{12}$$

The proof is complete in that (12) is just (7) with $F_{k\ell}$ replaced by $\hat{F}_{k\ell}$. □

The above theorem proves that if there is a self-consistent (SC) estimator, it is also "maximum likelihood". It is necessary to show the SC estimator exists and to find conditions for its uniqueness.

Assume initially that all the β's and λ's are zero so that only α-censoring occurs. Then the self-consistent equation (8) reduces to

$$F_{ij} = \frac{N_{ij}}{n} + \sum_{\substack{k<i \\ \ell>j}} \frac{\alpha_{k\ell}}{n} \frac{\hat{Q}_{i\ell}}{\hat{Q}_{k\ell}} .$$

Calculating $\hat{Q}_{ij} = \hat{F}_{i,j-1} - \hat{F}_{ij}$

$$\hat{Q}_{ij} = \frac{(N_{i,j-1}-N_{ij})}{n} + \sum_{k<i} \frac{\alpha_{kj}}{n} \frac{\hat{Q}_{ij}}{\hat{Q}_{kj}} .$$

Note that this equation is similar to one-dimensional self-consistency equation (3). Therefore, letting $Q_{ij}^{(0)} = \frac{N_{i,j-1}-N_{ij}}{n}$ and defining $Q_{ij}^{(m)}$ iteratively, by Theorem 1 the sequence $\{Q_{ij}^{(m)}\}$ is increasing. Consequently, the $F_{ij}^{(m)}$ so derived would also be increasing. In a similar manner, assume that α's and λ's are zero so that only β-censoring occurs. Then the sequence $\{R_{ij}^{(m)}\}$ is nondecreasing and hence so is $F_{ij}^{(m)}$. It is also possible to prove that if the α's and β's are zero so that only double censoring occurs that $\{F_{ij}^{(m)}\}$ is nondecreasing sequence in m. These results suggest the conjecture that in general the iterative procedure $\{F_{ij}^{(m)}\}$ beginning with $F_{ij}^{(0)} = N_{ij}/n$ is non-decreasing in m. This then implies, since $F_{ij}^{(m)} \leq 1$ for all i, j, and m, that the sequence $F_{ij}^{(m)}$ converges to $\hat{F}_{ij}$, the SC estimate, for i=1,...,I; j=1,...,J.

Given an estimate which satisfies the SC equation (8), the next issue is whether the estimate is unique. Let J denote the IJ x IJ matrix whose $((ij),(k,\ell))$ entry is

$$J_{(ij)(ij)} = -\frac{\partial^2 \log L}{\partial F_{ij}^2}$$

$$= d_{ij} + d_{i+1,j+1} + d_{i,j+1} + d_{i+1,j} + \ell_{ij} + a_{ij} + a_{i,j+1}$$

$$+ b_{ij} + b_{i+1,j},$$

where $d_{ij} = \delta_{ij}/(\Delta_{ij})^2$, $\ell_{ij} = \lambda_{ij}/F_{ij}^2$, $a_{ij} = \alpha_{ij}/Q_{ij}^2$,

and $b_{ij} = \beta_{ij}/R_{ij}^2$. Further,

$$J_{(ij)(i+1,j+1)} = d_{i+1,j+1}$$

$$J_{(ij)(i+1,j-1)} = d_{i+1,j}$$

$$J_{(ij)(i-1,j+1)} = d_{i,j+1}$$

$$J_{(ij)(i-1,j)} = -d_{ij} - d_{i,j+1} - b_{ij}$$

$$J_{(ij)(i,j-1)} = -d_{ij} - d_{i+1,j} - a_{ij}$$

$$J_{(ij)(k,\ell)} = 0 \quad \text{if } |i-k| \geq 2 \text{ or } |j-\ell| \geq 2.$$

Theorem. The matrix J is non-negative definite.

Proof. Multiply the matrix J by $Q = \{q_{(ij)(k,\ell)}\}$ where $q_{(ij)(k,\ell)} = (-1)^{i+j}(-1)^{k+\ell}$. Then the resulting matrix (call it K) has all non-negative entries. To prove J is non-negative definite it suffices to prove K is. Decompose K in the following manner:

Let L be diagonal IJ x IJ matrix with entries ℓ_{ij} at the entry (ij),(ij) and zeros elsewhere. Let A_{ij} denote the IJ x IJ matrix with ones at (ij)(ij), (ij)(i,j-1), (i,j-1)(ij), and (i,j-1)(i,j-1) with zeros elsewhere. Let B_{ij} denote the IJ x IJ matrix with ones at (ij)(ij), (i-1,j)(ij), (ij)(i-1,j) and (i-1,j)(i-1,j) with zeros elsewhere. Lastly, let D_{ij} denote the IJ x IJ with zeros except for the ones at (ij)(ij), (ij)(i,j-1), (ij)(i-1,j) (ij)(i-1,j-1), (i,j-1)(ij), (i,j-1)(i,j-1), (i,j-1)(i-1,j), (i,j-1)(i-1,j-1), (i-1,j)(ij), (i-1,j)(i,j-1), (i-1,j)(i-1,j), (i-1,j)(i-1,j-1), (i-1,j-1)(ij), (i-1,j-1)(i,j-1), (i-1,j-1)(i-1,j) and (i-1,j-1)(i-1,j-1). Then, it is readily seen that

$$Q = L + \sum_{i,j} a_{ij} A_{ij} + \sum_{i,j} b_{ij} B_{ij} + \sum_{i,j} d_{ij} D_{ij}.$$

Since L, A_{ij}, B_{ij}, and D_{ij} are all non-negative definite, K and hence J are also. □

Because J is non-negative definite, the likelihood function L of equation (6) is convex in the F_{ij}'s. This implies that the maximum L is unique up to possible flat spots by shifts in adjacent F_{ij}'s. In particular, the same uniqueness problem in one-dimension carries over to the bivariate case; namely, there is no uniqueness possible to the right of last death for α-censoring, above the last death for β-censoring and both above and to the right for λ-censoring. That is, deaths must follow losses in the appropriate way for the estimate to be unique. In the event that that is not the case, one can achieve specificity or uniqueness by arbitrarily converting the final losses (with no deaths thereafter) in any appropriate dimension to deaths.

Both the reduced-sample and the self-consistent estimators facilitate estimation with random censorship. The RS estimator requires the additional information of (C_i, D_i) even if X_i^0 precedes C_i or Y_i^0 precedes D_i. Its advantages are that it is easy to compute and is unbiased and consistent. In the event that all C_i and D_i exceed the largest X_i^0 and the largest Y_i^0, respectively, the estimator reduces to the empirical bivariate estimator with no censorship. The main disadvantage is that the resultant estimator need not be a distribution function. The self-consistent estimator, while more difficult to compute, does not require the complete censoring information concerning the C_i's and D_i's. Further, it is always a distribution function; in the event of no censoring it reduces to the empirical distribution function. The estimator jumps only at the points of double-deaths or final censored values in any dimension.

4. An example. Consider a number of identical systems of two (relevant) components A and B. Each system can function if either component is functioning, even although the components lifetimes may be dependent. In the first year of a two-year study, 100 such systems are placed on tests. Of these, ten fail in both A and B at the end of year 1, 8 fail in both components in the second year, 20 have A fail in year 1, and B in year 2, 15 have B fail in year 1, and A in year 2. Of the systems still functioning, 10 have failed in A in the first year, 5 in B in year 1, 16 in A in year 2, 12 in B in year 2, and 4 systems have no component failures after two years. The 90 systems still functioning at the end of year 1 are joined in the second year by an additional 1000 identical new systems. Of these 1000, 110 fail in both A and B by the end of year 2, 325 fail only in A, 275 in B and 290 do not fail after 2 years. (See figure 1). Let T denote the time beyond which no component functions in the figure.

The bivariate RS estimates for the jumps at years (ij) are:

$$\tilde{f}_{11} = .11 \qquad \tilde{f}_{21} = .22 \qquad \tilde{f}_{T1} = .05$$

$$\tilde{f}_{12} = .22 \qquad \tilde{f}_{22} = -.02 \qquad \tilde{f}_{T2} = .12$$

$$\tilde{f}_{1T} = .10 \qquad \tilde{f}_{2T} = .16 \qquad \tilde{f}_{TT} = .04$$

Note that as remarked earlier that the resulting estimator need not be a bivariate distribution function; in this case, it is not, due to the negative mass at (2,2).

The self-consistent estimates are:

$$\hat{f}_{11} = .11 \qquad \hat{f}_{21} = .20 \qquad \hat{f}_{T1} = .11$$

$$\hat{f}_{12} = .21 \qquad \hat{f}_{22} = .06 \qquad \hat{f}_{T2} = .09$$

$$\hat{f}_{1T} = .07 \qquad \hat{f}_{2T} = .12 \qquad \hat{f}_{T3} = .03$$

Acknowledgements. The author would like to acknowledge several helpful discussions with H. Rubin and C. Srinivasan.

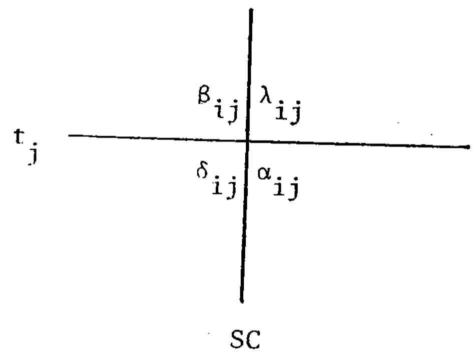

FIGURE 1

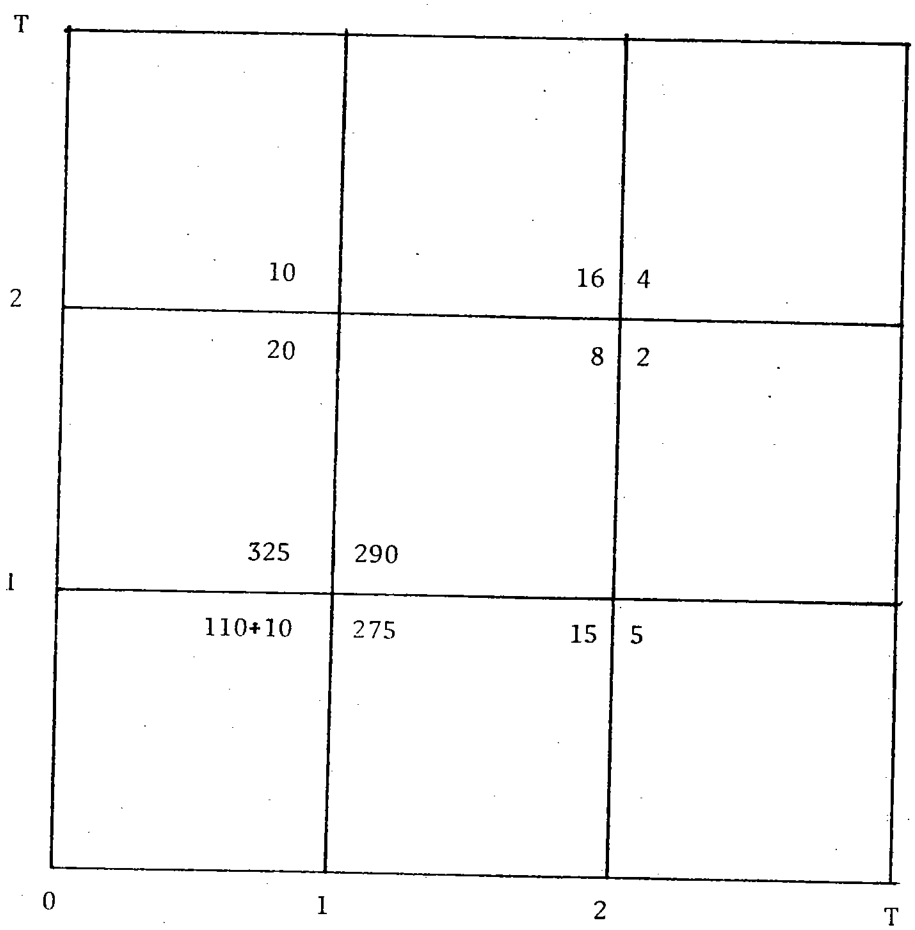

FIGURE 2

REFERENCES

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

Chmiel, J. Sander (1975a). The weak convergence of quantiles of the product-limit estimator. Stanford University Department of Statistics Public Health Service Technical Report No. 5.

Chmiel, J. Sander (1975b). Asymptotic properties for estimators derived from the Kaplan-Meier estimator. Stanford University Department of Statistics Public Health Service Technical Report No. 8.

Chmiel, J. Sander (1975c). A comparison of three estimators derived from the Kaplan-meier estimator. Stanford University Department of Statistics Public Health Service Technical Report No. 11.

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

Kaplan, E. L. and Meier, P. (1958). Nonparametric estimation from incomplete observations. J. Amer. Statist. Assoc. 53, 457-481.

Meier, Paul (1975). Estimation of a distribution function from incomplete observations. Perspectives in Probability and Statistics. Applied Probability Trust, Shelfield. 67-87.

Turnbull, Bruce W. (1974). Nonparametric estimation of a survivorship function with doubly-censored data. J. Amer. Statist. Assoc. 69, 169-173.