The Annals of Statistics
1988, Vol. 16, No. 4, 1475–1489

KAPLAN–MEIER ESTIMATE ON THE PLANE[1]

BY DOROTA M. DABROWSKA

Carnegie Mellon University and University of California, Berkeley

Estimation of the bivariate survival function from censored data is considered. The product integral representation of univariate survival functions is generalized to the bivariate case and used to determine identifiability of the survival function of the partially observed data. A bivariate analogue of the Kaplan–Meier estimate is introduced and its almost sure consistency is studied. Extensions to the general multivariate case are sketched.

1. Introduction. Survival and reliability studies often involve observations on paired individuals subject to censoring. Let $T = (T_1, T_2)$ be a pair of nonnegative random variables (rv). The variables T_1 and T_2 are thought of as survival or failure times and may represent lifetimes of married couples, times from initiation of a treatment until first response in two successive courses of a treatment in the same patient, etc. Under bivariate right censoring, the observable variables are given by $Y = (Y_1, Y_2)$ and $\delta = (\delta_1, \delta_2)$, where $Y_i = \min(T_i, Z_i)$ and $\delta_i = I(T_i = Y_i)$. Here $Z = (Z_1, Z_2)$ is a pair of fixed or random censoring times thought to represent times to withdrawals from the study. We refer to Clayton (1978), Hanley and Parnes (1983), Campbell (1981) and Clayton and Cuzick (1985) for examples of this type of censoring mechanism.

Two problems are addressed in this paper. First, we discuss conditions which ensure identifiability of the underlying joint survival function of the partially observable failure times. In the univariate case Aalen and Johansen (1978) and Gill and Johansen (1987) show that the survival function can be expressed as a product integral of the cumulative hazard function. A similar representation is available in the bivariate case for a suitably defined bivariate cumulative hazard function. The latter is a vector function representing cumulative hazards corresponding to "single" and "double" failures. Under the assumption of independence of the failure and censoring times, the bivariate cumulative hazard and the associated bivariate survival function can be easily expressed in terms of the joint distribution function of the observable variables.

Further, we consider estimation of the survival function of the censored failure times. Our estimation procedure rests on the natural "substitution principle," i.e., the estimate is based on the sample counterpart of the product integral. We refer to it as a bivariate Kaplan–Meier estimate. The name seems to be justified since apart from its product integral form, the marginals are given by the univariate Kaplan–Meier (1958) estimates and in the absence of censoring, the estimator reduces to the usual empirical survival function. The almost sure consistency of the bivariate Kaplan–Meier estimate is established.

Received April 1987; revised February 1988.
[1]This research was supported by a University of California Presidential Fellowship and National Institute of General Medical Sciences Grant GM-35416-02.
AMS 1980 *subject classifications.* Primary 62G05, 62N05; secondary 26B30.
Key words and phrases. Product integrals, identifiability, censored data.

The relevance of identifiability questions in inference problems related to competing risks models has been raised by many authors; see Tsiatis (1975) and Peterson (1975) for instance. In the context of estimation of the bivariate survival function from censored data, this problem was considered by Langberg and Shaked (1982). The authors suggested looking at $P(T_1 > s,\ T_2 > t) = A(s, t)B(t)$, where $A(s, t) = P(T_1 > s | T_2 > t)$ and $B(t) = P(T_2 > t)$, and applying the product integral representation of univariate survival functions to the terms A and B separately. Properties of the corresponding estimator of the survival function were developed by Campbell and Földes (1982), Campbell (1982), Horváth (1983), Burke (1984), Lo and Wang (1986) and Horváth and Yandell (1986), among others. The estimator suffers from various drawbacks, in particular it is not a proper survival function since it is not monotone in each of its coordinates, it does not reduce to the usual empirical survival function in the case of uncensored data and is dependent on the selected path and ordering of the components. Ruymgaart (1987) considered estimation of the related cumulative hazard function.

Tsai, Leurgans and Crowley (1986) suggested an estimation procedure which involves estimation of conditional survival functions using Beran's (1981) nonparametric regression methods for censored data. The estimator rests on smoothing techniques appropriate for nonparametric density and regression estimation and, although it is consistent, its rate of almost sure consistency is very slow, as compared to the Campbell–Földes estimate or ours.

Campbell (1981), Hanley and Parnes (1983) and Muñoz (1980) studied nonparametric MLE estimation using Efron's (1967) self-consistency algorithm and the EM algorithm of Dempster, Laird and Rubin (1977). The nonparametric MLE in this model does not have closed form expression and is not unique.

In connection with testing for independence, Pons (1986) derived a weak convergence result for the estimate of the bivariate cumulative hazard function corresponding to double failures (estimator $\hat{\Lambda}_{11}$ of Section 3). Bickel (personal communication) suggested the use of this estimator to construct an estimate of the bivariate survival function. His approach boils down to solving the integral equation

$$F(s,t) = F(s,0) + F(0,t) - 1 + \int_0^s \int_0^t F(u-, v-)\Lambda_{11}(du, dv)$$

subject to the initial condition

$$F(s,0) = \prod_{u \le s} (1 - \Lambda_{10}(du, 0)),$$

$$F(0,t) = \prod_{v \le t} (1 - \Lambda_{01}(0, dv)),$$

where Λ_{10}, Λ_{01} and Λ_{11} are cumulative hazard functions corresponding to single and double failures (Section 2). This is an inhomogeneous Volterra equation and has a solution in terms of the Peano series

$$1 + \sum_{n=1}^{\infty} \int \cdots \int_{\substack{0 \le u_1 < \cdots < u_n \le s \\ 0 \le v_1 < \cdots < v_n \le t}} \Lambda_{11}(du_1, dv_1) \cdots \Lambda_{11}(du_n, dv_n).$$

The corresponding estimator obtained by plugging in estimates $\hat{\Lambda}_{01}$, $\hat{\Lambda}_{01}$ and $\hat{\Lambda}_{11}$ (Section 3) has several of the same properties as the estimate considered in this paper. The definition is symmetric, the marginals are given by the Kaplan–Meier estimators, in the absence of censoring the estimate reduces to the empirical survival function. It has an extra nice property, namely it is always a survival function. However, an undesirable property is that it throws an important part of the data away.

2. Survival and cumulative hazard functions. In this section we show the correspondence between the bivariate survival and cumulative hazard functions. For the sake of completeness, we briefly consider the univariate case first. Next we define a bivariate hazard function and show that in analogy to the univariate case, it determines the bivariate survival function. Extensions to higher dimensions are outlined.

2.1. *Univariate survival times.* Let T be a univariate nonnegative rv defined on some probability space $(\Omega, \mathscr{F}, P)$ and let $F(t) = P(T > t)$ be its survival function. Furthermore, let $\Lambda(t)$, $\Lambda(dt) = -F(dt)/F(t-)$, $\Lambda(0) = 0$ be the associated cumulative hazard function. Then for $t \in [0, \tau]$ such that $F(\tau) > 0$, we have

$$F(t) = \exp\{-\Lambda^c(t)\} \prod_{u \le t} \{1 - \Lambda(\Delta u)\}, \tag{2.1}$$

where Λ^c is the continuous component of Λ, the product is taken over the discontinuity points of F and $\Lambda(\Delta u) = \Lambda(u) - \Lambda(u-)$ denotes the size of the jump at time u. This is the well-known representation of univariate survival functions. We refer to Peterson (1977), Gill (1980), Beran (1981) and Wellner (1985) for its derivations. Aalen and Johansen (1978) and Gill and Johansen (1987) show that (2.1) can be written as a product integral

$$F(t) = \prod_{s \le t} (1 - \Lambda(ds)) = \lim_{\max|s_i - s_{i-1}| \to 0} \prod_i (1 - \Lambda((s_{i-1}, s_i])),$$

where $0 = s_0 < s_1 < \cdots < s_n = t$ is a partition of $(0, t]$ and $\Lambda((s_{i-1}, s_i]) = \Lambda(s_i) - \Lambda(s_{i-1})$.

While various proofs of (2.1) seem to be available, the following simple argument will be useful in the sequel. For $t \in [0, \tau]$ such that $F(\tau) > 0$, we have $F(t) = \exp\{A(t)\}$, where $A(t) = \log F(t)$. By the Jordan decomposition of functions of bounded variation,

$$A(t) = \int_0^t A(du) = \int_0^t A^c(du) + \int_0^t A^d(du),$$

where A^c and A^d are the continuous and discrete components of A. Specifically,

$$\begin{aligned} A^d(t) &= \sum_{u \le t} A^d(\Delta u) = \sum_{u \le t} \log\{F(u)/F(u-)\} \\ &= \sum_{u \le t} \log(1 - \Lambda(\Delta u)), \end{aligned}$$

where the sum is taken over the discontinuity points of F, and

$$A^c(t) = A(t) - A^d(t) = \int_0^t F(u-)^{-1} F^c(du) = -\Lambda^c(t).$$

2.2. *Bivariate survival times.* Let $T = (T_1, T_2)$ be a pair of nonnegative rv's defined on a probability space $(\Omega, \mathscr{F}, P)$ and let $F(s,t) = P(T_1 > s, T_2 > t)$ be the corresponding joint survival function. By a bivariate cumulative hazard function, we mean a vector function $\Lambda(s,t) = (\Lambda_{10}(s,t), \Lambda_{01}(s,t), \Lambda_{11}(s,t))$, where

$$\Lambda_{11}(ds, dt) = \frac{P(T_1 \in ds, T_2 \in dt)}{P(T_1 \geq s, T_2 \geq t)} = \frac{F(ds, dt)}{F(s-, t-)},$$

$$\Lambda_{10}(ds, t) = \frac{P(T_1 \in ds, T_2 > t)}{P(T_1 \geq s, T_2 > t)} = \frac{-F(ds, t)}{F(s-, t-)},$$

$$\Lambda_{01}(s, dt) = \frac{P(T_1 > s, T_2 \in dt)}{P(T_1 > s, T_2 \geq t)} = \frac{-F(s, dt)}{F(s, t-)}$$

and

$$\Lambda_{10}(0, t) = \Lambda_{01}(s, 0) = \Lambda_{11}(0, 0) = 0.$$

If F has a density $f(s,t)$, we have $\Lambda_{11}(ds, dt) = \lambda_{11}(s,t)\,ds\,dt$, $\Lambda_{10}(ds,t) = \lambda_{10}(s,t)\,dt$ and $\Lambda_{01}(s, dt) = \lambda_{01}(s,t)\,dt$, where

$$\begin{aligned}
\lambda_{11}(s,t) &= \lim_{(h_1, h_2) \to 0} \frac{1}{h_1 h_2} P(T_1 \in [s, s+h_1], T_2 \in [t, t+h_2] | T_1 \geq s, T_2 \geq t) \\
&= \frac{f(s,t)}{F(s-, t-)},
\end{aligned}$$

$$\begin{aligned}
\lambda_{10}(s,t) &= \lim_{h \to 0} \frac{1}{h} P(T_1 \in [s, s+h] | T_1 \geq s, T_2 > t) \\
&= \int_t^\infty f(s,u)\,dv / F(s-, t),
\end{aligned}$$

$$\begin{aligned}
\lambda_{01}(s,t) &= \lim_{h \to 0} \frac{1}{h} P(T_2 \in [t, t+h] | T_1 > s, T_2 \geq t) \\
&= \int_s^\infty f(u,t)\,du / F(s, t-).
\end{aligned}$$

Thus $\lambda_{11}(s,t)$ represents the instantaneous rate of a "double failure" at point (s,t), given that the individuals were alive at times $T_1 = s-$ and $T_2 = t-$. Further, $\lambda_{10}(s,t)$ represents the rate of a "single failure" at time s given that the first individual was alive at time $T_1 = s-$ and the second survived beyond time $T_2 = t$. The meaning of $\lambda_{01}(s,t)$ is analogous.

We next give the representation of the bivariate survival function $F(s,t)$ in terms of the bivariate hazard function $\Lambda(s,t)$. Set $A(s,t) = \log F(s,t)$. Then for $(s,t) \in [0,\tau_1] \times [0,\tau_2]$, $F(\tau_1,\tau_2) > 0$, we have

$$\begin{aligned} F(s,t) &= \exp\{A(s,t)\} \\ &= \exp\left\{\int_0^s \int_0^t A(du,dv) + A(s,0) + A(0,t)\right\} \\ &= F(s,0)F(0,t)\exp\left\{\int_0^s \int_0^t A(du,dv)\right\}. \end{aligned}$$

We exploit the Jordan decomposition of $A(s,t)$. For $(s,t) \in [0,\tau_1] \times [0,\tau_2]$, $F(\tau_1,\tau_2) > 0$, the function $A(s,t) = \log F(s,t)$ is a function of bounded variation in the sense of Vitali and Hardy and Krause. We refer to Hildebrandt (1963), Chapter 3, and Clarkson and Adams (1933) for a survey of basic results on functions of bounded variation on the plane. By Theorem 5.4 in Hildebrandt (1963), page 110, $A(s,t)$ has a finite or countable number of discontinuities and they lie on a denumerable set of lines orthogonal to the coordinate axes.

In what follows, for any bivariate function $\phi(s,t)$ we write

$$\begin{aligned} \phi(\Delta s,t) &= \phi(s,t) - \phi(s-,t), \\ \phi(s,\Delta t) &= \phi(s,t) - \phi(s,t-) \end{aligned}$$

and

$$\phi(\Delta s,\Delta t) = \phi(s,t) - \phi(s,t-) - \phi(s-,t) + \phi(s-,t-).$$

Introduce sets

$$\begin{aligned} E_1 &= \{(s,t)\colon A(s,t) < 0,\ A(\Delta s,t) = A(s,\Delta t) = 0\}, \\ E_2 &= \{(s,t)\colon A(s,t) < 0,\ A(\Delta s,t) < 0,\ A(\Delta s,\Delta t) = 0\}, \\ E_3 &= \{(s,t)\colon A(s,t) < 0,\ A(s,\Delta t) < 0,\ A(\Delta s,\Delta t) = 0\}, \\ E_4 &= \{(s,t)\colon A(s,t) < 0,\ A(\Delta s,\Delta t) > 0\}. \end{aligned}$$

By the right-continuity and monotonicity of F, the set E_1 corresponds to the support of the purely continuous component of A, while E_4 is the support of the purely discrete component. Further, E_2 and E_3 are supports of components of A that have discontinuities lying along lines orthogonal to the coordinate axes. By the Jordan decomposition of functions of bounded variation on the plane, we have

$$\int_0^s \int_0^t A(du,dv) = \sum_{i=1}^4 A_i(s,t),$$

where

$$\begin{aligned} A_1(s,t) &= \int_0^s \int_0^t I[(u,v) \in E_1]\, A(du,dv), \\ A_2(s,t) &= \sum_{u \le s} \int_0^t I[(u,v) \in E_2][A(u,dv) - A(u-,dv)], \\ A_3(s,t) &= \sum_{v \le t} \int_0^s I[(u,v) \in E_2][A(du,v) - A(du,v-)], \\ A_4(s,t) &= \sum_{u \le s} \sum_{v \le t} I[(u,v) \in E_4]\, A(\Delta u,\Delta v). \end{aligned}$$

We compute now the explicit form of this decomposition. The following identities will be useful:

$$
\begin{aligned}
F(u,v)/F(u-,v-) &= 1-\Lambda_{10}(\Delta u, v-) \\
&\qquad -\Lambda_{01}(u-,\Delta v)+\Lambda_{11}(\Delta u,\Delta v), \\
F(u-,v)/F(u-,v-) &= 1-\Lambda_{01}(u-,\Delta v), \\
F(u,v-)/F(u-,v-) &= 1-\Lambda_{10}(\Delta u, v-).
\end{aligned}
\tag{2.2}
$$

Consider the purely discrete part first. By (2.2), we have

$$
\begin{aligned}
A_4(s,t) &= \sum_{\substack{u\le s,\ v\le t\\ (u,v)\in E_4}}\sum \left[\log\left(\frac{F(u,v)}{F(u-,v-)}\right)-\log\left(\frac{F(u-,v)}{F(u-,v-)}\right)\right. \\
&\qquad\qquad \left. -\log\left(\frac{F(u,v-)}{F(u-,v-)}\right)\right] \\
&= \sum_{\substack{u\le s,\ v\le t\\ (u,v)\in E_4}}\sum \log\left[1-\frac{\Lambda_{10}(\Delta u,v-)\Lambda_{01}(u-,\Delta v)-\Lambda_{11}(\Delta u,\Delta v)}{(1-\Lambda_{10}(\Delta u,v-))(1-\Lambda_{01}(u-,\Delta v))}\right].
\end{aligned}
\tag{2.3}
$$

Further, using (2.2) again,

$$
\begin{aligned}
A_3(s,t) &= \sum_{v\le t}\int_0^s I[(u,v)\in E_3]\left[\frac{F(du,v)}{F(u-,v)}-\frac{F(du,v-)}{F(u-,v-)}\right] \\
&= \sum_{v\le t}\int_0^s I[(u,v)\in E_3]\left[\frac{F(du,\Delta v)}{F(u-,v)}-\frac{F(du,v-)F(u-,\Delta v)}{F(u-,v)F(u-,v-)}\right] \\
&= \sum_{v\le t}\int_0^s I[(u,v)\in E_3]\frac{F(u-,v-)}{F(u-,v)} \\
&\qquad \times\left[\frac{F(du,\Delta v)}{F(u-,v-)}-\frac{F(du,v-)}{F(u-,v-)}\frac{F(u-,\Delta v)}{F(u-,v-)}\right] \\
&= \sum_{v\le t}\int_0^s \frac{I[(u,v)\in E_3]}{[1-\Lambda_{01}(u-,\Delta v)]} \\
&\qquad \times[\Lambda_{11}(du,\Delta v)-\Lambda_{10}(du,v-)\Lambda_{01}(u-,\Delta v)].
\end{aligned}
\tag{2.4}
$$

Similarly,

$$
\begin{aligned}
A_2(s,t) &= \sum_{u\le s}\int_0^t \frac{I[(u,v)\in E_2]}{[1-\Lambda_{10}(\Delta u,v-)]} \\
&\qquad \times[\Lambda_{11}(\Delta u,dv)-\Lambda_{10}(\Delta u,v-)\Lambda_{01}(u-,\Delta v)].
\end{aligned}
\tag{2.5}
$$

Finally,

$$(2.6)\qquad \begin{aligned} A_1(s,t) &= \int_0^s\int_0^t I[(u,v)\in E_1]\left[\frac{F(du,dv)}{F(u,v)} - \frac{F(du,v)F(u,dv)}{F(u,v)F(u,v)}\right] \\ &= \int_0^s\int_0^t I[(u,v)\in E_1][\Lambda_{11}(du,dv) - \Lambda_{10}(du,v)\Lambda_{01}(u,dv)]. \end{aligned}$$

Note that formulas (2.4)–(2.6) follow heuristically from (2.3) through a Taylor expansion by noting that the quotient in (2.3) is infinitesimal on E_i, $i = 1,2,3$. Define function $L(s,t)$ by

$$L(du,dv) = \frac{\Lambda_{10}(du,v-)\Lambda_{01}(u-,dv) - \Lambda_{11}(du,dv)}{\{1-\Lambda_{10}(\Delta u, v-)\}\{1-\Lambda_{01}(u-,\Delta v)\}}.$$

By the definition of the sets E_i, we have $\Lambda_{10}(\Delta u, v-) = 0$ for $(u,v)\in E_3\cup E_1$, and $\Lambda_{01}(u-,\Delta v) = 0$ for $(u,v)\in E_2\cup E_1$. Therefore, combining (2.3)–(2.6) with the product integral representation of the univariate survival functions, we arrive at Proposition 2.1.

PROPOSITION 2.1. *For* (s,t) *such that* $F(s,t) > 0$, *we have* $F(s,t) = F(s,0)F(0,t)\prod_{i=1}^4 B_i(s,t)$, *where*

$$\begin{aligned} B_i(s,t) &= \exp\{A_i(s,t)\} \\ &= \exp\left\{-\int_0^s\int_0^t I[(u,v)\in E_i]L(du,dv)\right\}, \qquad i = 1,2,3, \\ B_4(s,t) &= \prod_{\substack{u\le s\ v\le t \\ (u,v)\in E_4}}\prod [1 - L(\Delta u,\Delta v)], \end{aligned}$$

and

$$\begin{aligned} F(s,0) &= \exp\{-\Lambda_{10}^c(s,0)\}\prod_{u\le s}\{1-\Lambda_{10}(\Delta u,0)\}, \\ F(0,t) &= \exp\{-\Lambda_{01}^c(0,t)\}\prod_{v\le t}\{1-\Lambda_{01}(0,\Delta v)\}. \end{aligned}$$

Gill pointed out to me that the representation of Proposition 2.1 can be rewritten as a product integral

$$(2.7)\qquad \begin{aligned} F(s,t) = {}& \prod_{u\le s}(1-\Lambda_{10}(du,0))\prod_{v\le t}(1-\Lambda_{01}(0,dv)) \\ & \times \prod_{\substack{u\le s \\ v\le t}}(1-L(du,dv)), \end{aligned}$$

where the last factor on the right-hand side is defined by

$$\prod_{\substack{u\le s \\ v\le t}}(1-L(du,dv)) = \lim_{\substack{\max|u_i-u_{i-1}|\to 0 \\ \max|v_j-v_{j-1}|\to 0}} \prod_{i,j}\Big(1 - L\big((u_{i-1},u_i]\times(v_{j-1},v_j]\big)\Big),$$

where $0 = u_0 < \cdots < u_m = s$, $0 = v_0 < \cdots < v_n = t$ is a partition and $L((u_{i-1}, u_i] \times (v_{j-1}, v_j]) = L(u_i, v_j) - L(u_{i-1}, v_j) - L(u_i, v_{j-1}) + L(u_{i-1}, v_{j-1})$. While a rigorous argument requires extension of Gill and Johansen's (1987) results on product integration, heuristically note that

$$\begin{aligned}
1 - \Lambda_{10}(du, v-) &= P(T_1 > u | T_1 \geq u, T_2 \geq v) = F(u, v-)/F(u-, v-),\\
1 - \Lambda_{01}(u-, dv) &= P(T_2 > v | T_1 \geq u, T_2 \geq v) = F(u-, v)/F(u-, v-),\\
1 - \Lambda_{11}(du, dv) &= (1 - \Lambda_{10}(du, v-)) + (1 - \Lambda_{01}(u-, dv))\\
&\quad - P(T_1 > u, T_2 > v | T_1 \geq u, T_2 \geq v)\\
&= \{F(u, v-) + F(u-, v) - F(u, v)\}/F(u-, v-)
\end{aligned}$$

and after some algebra

$$\begin{aligned}
1 - L(du, dv) &= \frac{P(T_1 > u, T_2 > v | T_1 \geq u, T_2 \geq v)}{P(T_1 > u | T_1 \geq u, T_2 \geq v) P(T_2 > v | T_1 \geq u, T_2 \geq v)}\\
&= \frac{F(u, v) F(u-, v-)}{F(u, v-) F(u-, v)}.
\end{aligned}$$

Substitution of these expressions into the right-hand side of (2.7) and a little algebra gives an alternative proof of Proposition 2.1.

2.3. *Extensions to the multivariate case.* We consider now the general multivariate case. Since the notation is cumbersome, we merely sketch the main points.

Let $T = (T_1, \ldots, T_k)$ be a nonnegative rv defined on some probability space $(\Omega, \mathscr{F}, P)$ and let $F(t_1, \ldots, t_k) = P(T_1 > t_1, \ldots, T_k > t_k)$ be the corresponding survival function. We define first the k-variate cumulative hazard function. Roughly speaking, this is the collection of $2^k - 1$ functions representing the instantaneous risk of all possible "q-tuple failures," $q = 1, \ldots, k$, of components $m_1, \ldots, m_q$ at times $t_{m_1}, \ldots, t_{m_q}$ given that they were alive at times $t_{m_1}-, \ldots, t_{m_q}-$ and that the remaining components survived beyond times t_m, $m \in \{1, \ldots, k\} - \{m_1, \ldots, m_q\}$. More precisely, let J be the collection of all zero–one sequences $(j) = \{j_1, \ldots, j_k\}$ such that $\sum j_i \neq 0$. By a k-variate cumulative hazard function we mean a triangular array

$$\Lambda(t_1, \ldots, t_k) = \left\{\Lambda_{q(j)}(t_1, \ldots, t_k) \colon q = 1, \ldots, k, (j) \in J, \sum j_m = q\right\},$$

consisting of k rows with $\binom{k}{q}$ elements in the qth row, $q = 1, \ldots, k$. For $(j) \in J$ such that $\sum j_m = q$ and $j_{i_1} = \cdots = j_{i_q} = 1$, the function $\Lambda_{q(j)}$ is defined by

$$\begin{aligned}
&\Lambda_{q(j)}\left(t_1, \ldots, t_{j_{i_1}-1}, dt_{j_{i_1}}, t_{j_{i_1}+1}, \ldots, t_{j_{i_q}-1}, dt_{j_{i_q}}, t_{j_{i_q}+1}, \ldots, t_k\right)\\
&\quad = (-1)^q \frac{F\left(t_1, \ldots, t_{j_{i_1}-1}\, dt_{j_{i_1}}, t_{j_{i_1}+1}, \ldots, t_{j_{i_q}-1}, dt_{j_{i_q}}, t_{j_{i_q}+1}, \ldots, t_k\right)}{F\left(t_1, \ldots, t_{j_{i_1}-1}, t_{j_{i_1}}-, t_{j_{i_1}+1}, \ldots, t_{j_{i_q}-1}, t_{j_{i_q}}-, t_{j_{i_q}+1}, \ldots, t_k\right)}.
\end{aligned}$$

Thus for instance if $k = 3$, then the first row of $\Lambda(t_1, t_2, t_3)$ consists of three cumulative hazard functions corresponding to "single failures," the second row consists of three cumulative hazard functions corresponding to "double failures" and the third consists of one function representing the cumulative hazard function of a "triple failure."

The derivation of the product integral representation of $F(t_1,\dots,t_k)$ is inductive and follows along the same lines as in the bivariate case. Set $A(t_1,\dots,t_k) = \log F(t_1,\dots,t_k)$. Then

$$F(t_1,\dots,t_k) = \exp\left\{\int_0^{t_1}\cdots\int_0^{t_k} A(du_1,\dots,du_k)\right\} \times \prod_{i=1}^{k-1}\left[\prod_{(m)} F(m_1t_1,\dots,m_kt_k)\right]^{(-1)^{k-i}}, \tag{2.8}$$

where the inner product in (2.8) is taken over all zero–one sequences $(m) = (m_1,\dots,m_k)$ such that $\Sigma m_j = i$. The double product in (2.8) represents, the ratio of products of i-dimensional marginal survival functions, $i = 1,\dots,k-1$. For instance if $k = 3$, then this factor reduces to

$$F(t_1,t_2,0)F(t_1,0,t_3)F(0,t_2,t_3)/\{F(t_1,0,0)F(0,t_2,0)F(0,0,t_3)\}.$$

To obtain the product integral representation of $F(t_1,\dots,t_k)$, it remains to use the product integral representation of i-dimensional marginal survival functions, $i = 1,\dots,k-1$, and next apply the Jordan decomposition to the integral appearing in the exponent of (2.6). In the last step, note that the discontinuities lie on hyperplanes orthogonal to the axes of the coordinate system. The form of the purely continuous component can be deduced from the derivative of $A(t_1,\dots,t_k)$ with respect to $t_1,\dots,t_k$. The form of the purely discrete component follows from a little algebra applied to $\Sigma_{u_1\le t_1}\cdots\Sigma_{u_k\le t_k} A(\Delta u_1,\dots,\Delta u_k)$.

3. Estimation of the bivariate survival function from censored data. We assume now that the data are censored and consider the identifiability question first. The failure times $T = (T_1, T_2)$ and censoring times $Z = (Z_1, Z_2)$ are defined on a common probability space $(\Omega, \mathscr{F}, P)$ and the respective joint survival functions are denoted by $F(s,t)$ and $G(s,t)$. The observable rv's are given by $Y = (Y_1, Y_2)$ and $\delta = (\delta_1, \delta_2)$, where $Y_i = \min(T_i, Z_i)$ and $\delta_i = I(T_i = Y_i)$, $i = 1, 2$.

By Proposition 2.1, the identifiability of $F(s,t)$ will follow if we can show that the bivariate hazard function $\Lambda(s,t)$ can be expressed in terms of the joint distribution function of Y and δ.

Set $H(s,t) = P(Y_1 > s,\ Y_2 > t)$, $K_1(s,t) = P(Y_1 > s,\ Y_2 > t,\ \delta_1 = 1,\ \delta_2 = 1)$, $K_2(s,t) = P(Y_1 > s,\ Y_2 > t,\ \delta_1 = 1)$ and $K_3(s,t) = P(Y_1 > s,\ Y_2 > t,\ \delta_2 = 1)$. Assume

(A) $T = (T_1, T_2)$ and $Z = (Z_1, Z_2)$ are independent.

Assumption (A) is sufficient to ensure identifiability of F on the support of H. Indeed, for (s,t) such that $H(s,t) > 0$, we have $H(s,t) = G(s,t)F(s,t)$,

$K_1(ds, dt) = G(s-, t-)F(ds, dt)$, $K_2(ds, t) = G(s-, t)F(ds, t)$ and $K_3(s, dt) = G(s, t-)F(s, dt)$ so that

$$\Lambda_{11}(s,t) = \int_0^s \int_0^t K_1(du, dv)/H(u-, v-),$$

$$\Lambda_{10}(s,t) = -\int_0^s K_2(du, t)/H(u-, t),$$

$$\Lambda_{01}(s,t) = -\int_0^t K_3(s, dv)/H(s, v-).$$

Suppose now that $Y_i = (Y_{1i}, Y_{2i})$, $\delta_i = (\delta_{1i}, \delta_{2i})$, $i = 1, \dots, n$, is an iid sample, each (Y_i, δ_i) having the same distribution as (Y, δ). To estimate the survival function F of the partially observable survival times, define

$$\hat{H}(s,t) = n^{-1} \sum I(Y_{1i} > s, Y_{2i} > t),$$

$$\hat{K}_1(s,t) = n^{-1} \sum I(Y_{1i} > s, Y_{2i} > t, \delta_{1i} = 1, \delta_{2i} = 1),$$

$$\hat{K}_2(s,t) = n^{-1} \sum I(Y_{1i} > s, Y_{2i} > t, \delta_{1i} = 1),$$

$$\hat{K}_3(s,t) = n^{-1} \sum I(Y_{1i} > s, Y_{2i} > t, \delta_{2i} = 1).$$

Further, let $\hat{\Lambda}(s,t) = (\hat{\Lambda}_{10}(s,t), \hat{\Lambda}_{01}(s,t), \hat{\Lambda}_{11}(s,t))$ be an estimator of the bivariate cumulative hazard function given by

$$\hat{\Lambda}_{11}(s,t) = \int_0^s \int_0^t \hat{K}_1(du, dv)/\hat{H}(u-, v-),$$

$$\hat{\Lambda}_{10}(s,t) = -\int_0^s \hat{K}_2(du, t)/\hat{H}(u-, t),$$

$$\hat{\Lambda}_{01}(s,t) = -\int_0^t \hat{K}_3(s, dv)/\hat{H}(s, v-).$$

A natural candidate for an estimator of $F(s,t)$ is provided by

$$\hat{F}(s,t) = \hat{F}(s,0)\hat{F}(0,t) \prod_{\substack{0<u\le s\\0<v\le t}} \left[1 - \hat{L}(\Delta u, \Delta v)\right],$$

where

$$\hat{L}(\Delta u, \Delta v) = \frac{\hat{\Lambda}_{10}(\Delta u, v-)\hat{\Lambda}_{01}(u-, \Delta v) - \hat{\Lambda}_{11}(\Delta u, \Delta v)}{\{1 - \hat{\Lambda}_{10}(\Delta u, v-)\}\{1 - \hat{\Lambda}_{01}(u-, \Delta v)\}}$$

and $\hat{F}(s,0)$ and $\hat{F}(0,t)$ are the usual Kaplan–Meier estimates, i.e.,

$$\hat{F}(s,0) = \prod_{u \le s} \left[1 - \hat{\Lambda}_{10}(\Delta u, 0)\right],$$

$$\hat{F}(0,t) = \prod_{v \le t} \left[1 - \hat{\Lambda}_{01}(0, \Delta v)\right].$$

The marginals of $\hat{F}(s,t)$ are given by the univariate Kaplan–Meier estimates. In the absence of censoring $\hat{F}(s,t)$ reduces to the usual empirical survival function. This can be verified by noting that the empirical survival function is purely discrete and by carrying the same calculation as in (2.3). A referee pointed out that in the presence of censoring $\hat{F}(s,t)$ may fail to be monotone. As an example consider points $(Y_{1i}, Y_{2i}, \delta_{1i}, \delta_{2i})$, $i = 1, \dots, 4$, given by $(0.51, 0.02, 1, 1)$, $(0.68, 0.68, 1, 1)$, $(0.11, 0.62, 1, 0)$ and $(0.24, 0.24, 0, 0)$. Then $\hat{F}(s, 0) = 1,\ 0.75,\ 0.75, 0.375, 0$ for $s = 0, 0.11, 0.24, 0.51, 0.68$, $\hat{F}(0, t) = 1, 0.75, 0.75, 0.75, 0$ for $t = 0, 0.02, 0.24, 0.62, 0.68$, $\hat{F}(s,t) = 0.5$ for $(s,t) \in [0.11, 0.68) \times [0.02, 68)$ and $\hat{F}(s,t) = 0$ if $s \geq 0.68$ or $t \geq 0.68$. Note that $\hat{F}(0.51, 0.02) = 0.5 > 0.375 = \hat{F}(0.51, 0)$. Roughly speaking, this anomaly can be explained by the fact that three different portions of the data set are involved in estimation of Λ_{10}, Λ_{01} and Λ_{11}. It can be easily verified that Λ_{10}, Λ_{01} and Λ_{11} satisfy

$$\Lambda_{10}(ds, dt) = \big(1 - \Lambda_{10}(\Delta s, t-)\big) L(ds, dt),$$
$$\Lambda_{01}(ds, dt) = \big(1 - \Lambda_{01}(s-, \Delta t)\big) L(ds, dt).$$

However, unless the data are uncensored, $\hat{\Lambda}_{10}$, $\hat{\Lambda}_{01}$ and $\hat{\Lambda}_{11}$ no longer satisfy this constraint.

Extension to the general multivariate case is inductive. In Section 2.3, we have outlined the extension of the product integral representation of survival functions to the case of k-dimensional failure times. Consider an iid sample $Y_i = (Y_{1i}, \dots, Y_{ki})$ and $\delta_i = (\delta_{1i}, \dots, \delta_{ki})$, where $Y_{ji} = \min(T_{ji}, Z_{ji})$ and $\delta_{ji} = I(T_{ji} = Y_{ji})$, $j = 1, \dots, k$ and $i = 1, \dots, n$. If the censoring variables $Z_i = (Z_{1i}, \dots, Z_{ki})$ are independent of the failure times $T_i = (T_{1i}, \dots, T_{ki})$, then the multivariate hazard function $\Lambda(t_1, \dots, t_k)$ of Section 2.3 can be expressed in terms of the joint distribution function of Y's and δ's. The sample counterpart of $\Lambda(t_1, \dots, t_k)$ coupled with the product integral representation of the survival function $F(t_1, \dots, t_k)$ yields the multivariate Kaplan–Meier estimate.

4. Consistency of the bivariate Kaplan–Meier estimate.

In this section we consider consistency of the estimator $\hat{F}(s,t)$. For $\tau = (\tau_1, \tau_2)$ let $\|\cdot\|_\tau$ denote the supremum norm on $[0, \tau_1] \times [0, \tau_2]$.

PROPOSITION 4.1. *Suppose that condition* (A) *holds and* $\tau = (\tau_1, \tau_2)$ *satisfies* $H(\tau_1, \tau_2) > 0$. *Then* $\|\hat{F} - F\|_\tau \to 0$ *almost surely.*

We have $\hat{F}(s,t) = \hat{F}(s,0)\hat{F}(0,t)\prod_{i=1}^4 C_i(s,t)$, where

$$C_i(s,t) = \prod_{\substack{0 < u \leq s \\ 0 < v \leq t \\ (u,v) \in E_i}} \{1 - \hat{L}(\Delta u, \Delta v)\}.$$

By Proposition 2.1 and the uniform consistency of the univariate Kaplan–Meier estimates [Földes and Rejtő (1981), Csörgő and Horváth (1983) and Shorack and Wellner (1986), page 305], it is enough to show $\|B_i - C_i\|_\tau \to 0$ a.s. for $i = 1, \dots, 4$. This will be established in a sequence of lemmas.

LEMMA 4.1. *Under the assumptions of Proposition* 3.1, $\|\hat{\Lambda}_{11} - \Lambda_{11}\|_\tau \to 0$, $\|\hat{\Lambda}_{10} - \Lambda_{10}\|_\tau \to 0$ *and* $\|\hat{\Lambda}_{01} - \Lambda_{01}\|_\tau \to 0$ *almost surely.*

PROOF. This follows from the Glivenko–Cantelli theorem and simple algebra. We omit the details. □

LEMMA 4.2. *Under the assumptions of Proposition* 3.1, $\|B_i - C_i\|_\tau \to 0$ *almost surely, for* $i = 1, 2, 3$.

PROOF. We consider the case of $i = 3$, the proof in the remaining two cases is analogous. By the inequality $|x - y| \le |\log x - \log y|$ for $0 < x, y < 1$, we have

$$\begin{aligned} &\sup|B_3(s,t) - C_3(s,t)| \\ (4.1)\qquad &\le \sup\left|\sum\sum_{\substack{u \le s,\, v \le t \\ (u,v) \in E_3}} \left[\log(1 - \hat{L}(\Delta u, \Delta v)) - \hat{L}(\Delta u, \Delta v)\right]\right| \\ &\quad + \sup\left|\int_0^s \int_0^t I[(u,v) \in E_3](\hat{L} - L)(du, dv)\right|, \end{aligned}$$

where the sup is taken over $(s, t) \in [0, \tau_1] \times [0, \tau_2]$. Consider the first sum on the right-hand side of (4.1). By the elementary inequality $-\log[1 - (1 + x)^{-1}] - (1 + x)^{-1} < [x(1 + x)]^{-1}$ for $x > 0$ and $x < -1$ we have

$$\begin{aligned} &\sum\sum_{\substack{u \le s,\, v \le t \\ (u,v) \in E_3}} |\log(1 - \hat{L}(\Delta u, \Delta v)) - \hat{L}(\Delta u, \Delta v)| \\ &\le \sum\sum_{\substack{u \le s,\, v \le t \\ (u,v) \in E_3}} \hat{L}^2(\Delta u, \Delta v)\{1 - \hat{L}(\Delta u, \Delta v)\}^{-1} \\ (4.2)\qquad &\le \sup\{\hat{L}(\Delta u, \Delta v) \colon (u,v) \in E_3 \cap [0, \tau_1] \times [0, \tau_2]\} \\ &\quad \times \sum\sum_{\substack{u \le \tau_1,\, v \le \tau_2 \\ (u,v) \in E_3}} \left|\frac{\hat{\Lambda}_{10}(\Delta u, v-)\hat{\Lambda}_{01}(u-, \Delta v) - \hat{\Lambda}_{11}(\Delta u, \Delta v)}{1 - \hat{\Lambda}_{10}(\Delta u, v-) - \hat{\Lambda}_{01}(u-, \Delta v) + \hat{\Lambda}_{11}(\Delta u, \Delta v)}\right|. \end{aligned}$$

For $(u, v) \in E_3$, we have $L(\Delta u, \Delta v) = 0$. Therefore

$$\begin{aligned} &\sup\{\hat{L}(\Delta u, \Delta v) \colon (u,v) \in E_3 \cap [0, \tau_1] \times [0, \tau_2]\} \\ &\quad \le \sup\{|\hat{L}(\Delta u, \Delta v) - L(\Delta u, \Delta v)| \colon 0 \le u \le \tau_1, 0 \le v \le \tau_2\}, \end{aligned}$$

which converges almost surely to 0 by Lemma 4.1. Furthermore, the sum on the right-hand side of (4.2) stays bounded by the consistency of $\hat{\Lambda}_{10}$, $\hat{\Lambda}_{01}$ and $\hat{\Lambda}_{11}$ and a little algebra. Finally, the second term on the right-hand side of (4.1) converges almost surely to 0 by Lemma 4.1 and a simple calculation. □

LEMMA 4.3. *Under the assumptions of Proposition* 3.1

$$\sum\sum_{\substack{u \le \tau_1,\, v \le \tau_1 \\ (u,v) \in E_4}} |\hat{L}(\Delta u, \Delta v) - L(\Delta u, \Delta v)| \to 0$$

almost surely.

Proof. We have

$$\begin{aligned}
&\sum\sum\left|\hat{L}(\Delta u, \Delta v) - L(\Delta u, \Delta v)\right| \\
&\quad\le \sum\sum\left|\hat{L}_1(\Delta u, \Delta v) - L_1(\Delta u, \Delta v)\right| \qquad (4.3)\\
&\qquad + \sum\sum\left|\hat{L}_2(\Delta u, \Delta v) - L_2(\Delta u, \Delta v)\right|.
\end{aligned}$$

Here the sums extend over $u \le \tau_1$, $v \le \tau_2$ such that $(u, v) \in E_4$. Furthermore,

$$L_1(\Delta u, \Delta v) = \frac{\Lambda_{11}(\Delta u, \Delta v)}{\{(1 - \Lambda_{10}(\Delta u, v-))(1 - \Lambda_{01}(u-, \Delta v))\}},$$

$$L_2(\Delta u, \Delta v) = \frac{\Lambda_{10}(\Delta u, v-)\Lambda_{01}(u-, \Delta v)}{\{(1 - \Lambda_{10}(\Delta u, v-))(1 - \Lambda_{01}(u-, \Delta v))\}}.$$

The terms L_1 and L_2 are defined by replacing Λ_{11}, Λ_{10} and Λ_{01} by their sample counterparts in L_1 and L_2.

Lemma 4.1 and a little algebra imply $\sup|\hat{L}_1(\Delta u, \Delta v) - L_1(\Delta u, \Delta v)| \to 0$ a.s. with sup taken over $0 \le u \le \tau_1$ and $0 \le v \le \tau_2$, and $\|\hat{L}_1 - L_1\|_\tau \to 0$ a.s. Since almost sure convergence implies weak convergence and the set E_4 is closed, we have

$$\limsup \sum_{\substack{u \le \tau_1,\, v \le \tau_2 \\ (u, v) \in E_4}}\sum \hat{L}_1(\Delta u, \Delta v) \le \sum_{\substack{u \le \tau_1,\, v \le \tau_2 \\ (u, v) \in E_4}}\sum L_1(\Delta u, \Delta v) < \infty.$$

Scheffé's theorem [see, e.g., Shorack and Wellner, (1986), page 862] implies

$$\sum_{\substack{u \le \tau_1,\, v \le \tau_2 \\ (u, v) \in E_4}}\sum \left|\hat{L}_1(\Delta u, \Delta v) - L_1(\Delta u, \Delta v)\right| \to 0$$

almost surely. The second sum in (4.3) can be treated in a similar way. □

Lemma 4.4. *Under the assumptions of Proposition* 3.1, $\|B_4 - C_4\|_\tau \to 0$ *a.s.*

Proof. Since $H(\tau_1, \tau_2) > 0$, we can find $\eta < 1$ such that $\sup\{L(\Delta u, \Delta v)\colon 0 \le u \le \tau_1,\ 0 \le v \le \tau_2\} < \eta$. Fix $\varepsilon > 0$. By Lemma 4.1 and a little algebra, $\sup\{|\hat{L}(\Delta u, \Delta v) - L(\Delta u, \Delta v)|/(1 - L(\Delta u, \Delta v))\colon 0 \le u \le \tau,\ 0 \le v \le \tau_2\} < \varepsilon$ for n sufficiently large. Further, by the mean value theorem, we have $|\log(1 - x)| \le |x|(1 - \varepsilon)^{-1}$ for $|x| \le \varepsilon$. Therefore, for n sufficiently large

$$\begin{aligned}
\sup|B_4(s, t) - C_4(s, t)| &\le \sum\sum\left|\log(1 - \hat{L}(\Delta u, \Delta v)) - \log(1 - L(\Delta u, \Delta v))\right| \\
&= \sum\sum\left|\log\left(1 - \frac{\hat{L}(\Delta u, \Delta v) - L(\Delta u, \Delta v)}{1 - L(\Delta u, \Delta v)}\right)\right| \\
&\le (1 - \varepsilon)^{-1}\sum\sum\frac{|\hat{L}(\Delta u, \Delta v) - L(\Delta u, \Delta v)|}{(1 - L(\Delta u, \Delta v))} \\
&\le (1 - \varepsilon)^{-1}(1 - \eta)^{-1}\sum\sum\left|\hat{L}(\Delta u, \Delta v) - L(\Delta u, \Delta v)\right|.
\end{aligned}$$

Here all the sums are taken over $0 \le u \le \tau_1$, $0 \le v \le \tau_2$ such that $(u, v) \in E_4$. Lemma 4.3 implies that this bound converges almost surely to 0. □

Acknowledgments. I thank Richard Gill and the referees for their comments.

REFERENCES

AALEN, O. O. and JOHANSEN, S. (1978). An empirical transition matrix for non-homogeneous Markov chains based on censored observations. *Scand. J. Statist.* **5** 141–150.

BERAN, R. (1981). Nonparametric regression with randomly censored survival data. Technical Report, Dept. Statistics, Univ. California, Berkeley.

BURKE, M. D. (1984). An almost-sure approximation of a multivariate product-limit estimator under random censorship. Unpublished.

CAMPBELL, G. (1981). Nonparametric bivariate estimation with randomly censored data. *Biometrika* **68** 417–422.

CAMPBELL, G. (1982). Asymptotic properties of several nonparametric multivariate distribution function estimators under random censorship. In *Survival Analysis* (J. Crowley and R. A. Johnson, eds.) 243–256. IMS, Hayward, Calif.

CAMPBELL, G. and FÖLDES, A. (1982). Large-sample properties of nonparametric bivariate estimators with censored data. In *Nonparametric Statistical Inference* (B. V. Gnedenko, M. L. Puri and I. Vincze, eds.) 103–122. North-Holland, Amsterdam.

CLARKSON, J. A. and ADAMS, C. R. (1933). On the definitions of bounded variation for functions of two variables. *Trans. Amer. Math. Soc.* **35** 824–854.

CLAYTON, D. G. (1978). A model of association in bivariate life tables and its application in epidemiological studies of familial tendency in chronic disease incidence. *Biometrika* **65** 141–151.

CLAYTON, D. G. and CUZICK, J. (1985). Multivariate generalizations of the proportional hazards model (with discussion). *J. Roy. Statist. Soc. Ser. A* **148** 82–117.

CSÖRGŐ, S. and HORVÁTH, L. (1983). The rate of strong uniform consistency for the product-limit estimator. *Z. Wahrsch. verw. Gebiete* **62** 411–426.

DEMPSTER, A. P., LAIRD, N. M. and RUBIN, D. B. (1977). Maximum likelihood estimation for incomplete data via the EM algorithm (with discussion). *J. Roy. Statist. Soc. Ser. B* **39** 1–38.

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

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

GILL, R. D. (1980). *Censoring and Stochastic Integrals*. Mathematical Centre Tract 124, Mathematisch Centrum, Amsterdam.

GILL, R. D. and JOHANSEN, S. (1987). Product integrals and counting processes. Technical Report MS-R8707, Mathematisch Centrum, Amsterdam.

HANLEY, J. A. and PARNES, M. N. (1983). Nonparametric estimation of a multivariate distribution in the presence of censoring. *Biometrics* **39** 129–139.

HILDEBRANDT, T. K. (1963). *Introduction to the Theory of Integration*. Academic, New York.

HORVÁTH, L. (1983). The rate of strong uniform consistency for the multivariate product-limit estimator. *J. Multivariate Anal.* **13** 202–209.

HORVÁTH, L. and YANDELL, B. S. (1986). Bootstrapped multidimensional product limit process. Technical Report, Dept. Statistics, Univ. Wisconsin, Madison.

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

LANGBERG, N. A. and SHAKED, M. (1982). On the identifiability of multivariate life distribution functions. *Ann. Probab.* **10** 773–779.

LO, S.-H. and WANG, J. L. (1986). Iid representations for the bivariate product limit estimators and the bootstrap versions. Technical Report, Div. Statistics, Univ. California, Davis.

MUÑOZ, A. (1980). Nonparametric estimation from censored bivariate observations. Technical Report, Dept. Statistics, Stanford Univ.

PETERSON, A. V. (1975). Nonparametric estimation in the competing risk problem. Technical Report, Dept. Statistics, Stanford Univ.

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

PONS, O. (1986). A test of independence between two censored survival times. *Scand. J. Statist.* **13** 173–185.

RUYMGAART, F. (1987). Some properties of bivariate empirical hazard processes under random censoring. Technical Report 8718, Catholic Univ. of Nijmegen, Netherlands.

SHORACK, G. R. and WELLNER, J. A. (1986). *Empirical Processes with Applications to Statistics.* Wiley, New York.

TSAI, W. Y., LEURGANS, S. and CROWLEY, J. (1986). Nonparametric estimation of a bivariate survival function in the presence of censoring. *Ann. Statist.* **14** 1351–1365.

TSIATIS, A. (1975). A non-identifiability aspect of the problem of competing risks. *Proc. Nat. Acad. Sci. U.S.A.* **72** 20–22.

WELLNER, J. A. (1985). A heavy censoring limit theorem for the product limit estimator. *Ann. Statist.* **13** 150–162.

DIVISION OF BIOSTATISTICS
SCHOOL OF PUBLIC HEALTH
UNIVERSITY OF CALIFORNIA
LOS ANGELES, CALIFORNIA 90024