LARGE-SAMPLE PROPERTIES OF NONPARAMETRIC BIVARIATE ESTIMATORS WITH CENSORED DATA

by

G. Campbell and A. Földes*

Department of Statistics
Division of Mathematical Sciences
Mimeo Series #80-10

May 1980

*Mathematical Institute of the Hungarian Academy of Sciences. This research was done while the author was visiting the Departments of Statistics and Mathematics at Purdue University. The author would like to acknowledge the financial assistance of the two departments.

LARGE-SAMPLE PROPERTIES OF NONPARAMETRIC

BIVARIATE ESTIMATORS WITH CENSORED DATA

by

G. Campbell

and

A. Földes

Let $\{(X_i^0, Y_i^0)\}_{i=1}^n$ be independent from the continuous, bivariate distribution function F and let $\{(C_i, D_i)\}_{i=1}^n$ denote censoring variables (independent of (X_i^0, Y_i^0)) from the continuous bivariate distribution G. The pair $X_i = \min(X_i^0, C_i)$ and $Y_i = \min(Y_i^0, D_i)$ is observable along with $(\epsilon_{1i}, \epsilon_{2i})$, where $\epsilon_{1i}(\epsilon_{2i}) = 1$ or 0 according as $X_i = X_i^0 (Y_i = Y_i^0)$ or not. Nonparametric estimators $\tilde{F}$ of F are developed based on the hazard gradient of Marshall (1975). The hazard function estimate is path-dependent. Consequently two piecewise linear paths are considered: one linear from (0,0) to (s,0) to (s,t) and the other linear from (0,0) to (0,t) to (s,t). Two estimators $\tilde{F}_{1n}$ and $\tilde{F}_{2n}$ result. For each path there is another easily-computed estimator $\hat{F}_{in}$ expressible as the product of conditional probabilities and hence the product of two Kaplan-Meier-type estimators. It is proved that for S,T such that $\bar{H}(S,T) = P\{X > S, Y > T\} > 0$,

$$\sup_{\substack{0<s\leq S\\ 0<t\leq T}} |\hat{F}_{in}(s,t) - \tilde{F}_{in}(s,t)| = \mathcal{O}(\tfrac{1}{n}) \text{ a.s.}$$

Further, each estimator $\tilde{F}_{in}$ (and hence $\hat{F}_{in}$) is almost surely uniformly consistent for F; in particular, for $i = 1,2$, $\sup_{\substack{0<s\leq S\\ 0<t\leq T}} |\tilde{F}_{in}(s,t) - F(s,t)| = \mathcal{O}(\sqrt{\log\log n}/\sqrt{n})$ a.s., provided $\bar{H}(S,T) > 0$.

LARGE-SAMPLE PROPERTIES OF NONPARAMETRIC BIVARIATE ESTIMATORS WITH CENSORED DATA

by

G. Campbell and A. Földes

1. Introduction and Summary. The estimation of a bivariate distribution function under random censoring is considered. The problem is to estimate the distribution of the life times under random censoring in which one knows whether the observations are losses (censored) or deaths (uncensored). There are numerous examples to demonstrate the importance of this bivariate problem. In some experiments the data are naturally paired such as observations on eyes, lungs, twins, married couples, or matched pairs. Oftentimes there are two sequential observations on the same individual (pre-test, post-test). In a reliability setting, a pair of components in a system can be observed.

Examples of (possibly bivariate) censoring mechanisms are plentiful. There is, of course, the censoring due to patient drop-out or non-compliance. The competing risk framework can be thought of as censoring. The censoring may be essentially univariate, such as the random entry of subjects into the study with fixed cutoff time for evaluation.

The one-dimensional random censoring model has been treated in great detail in the recent literature, beginning with the landmark papers of Kaplan and Meier (1958) and Efron (1967). The asymptotic normality and weak convergence of the product-limit estimator of Kaplan and Meier was treated by Breslow and Crowley (1974). Strong uniform consistency was treated by Winter, Földes and Rejtő (1978), Földes, Rejtő and Winter (1979), and Földes and Rejtő (1979).

The bivariate estimation problem with discrete times of deaths or losses has been considered by Campbell (1979) using an extension of the self-consistent approach of Efron (1967). Hanley and Parnes (1980) have treated maximum likelihood approaches to bivariate estimation. In contrast to the iterative estimators of these researchers this paper considers several new closed-form estimators for the bivariate model and proves strong uniform consistency to the true bivariate distribution of the lifetimes.

Two path-dependent estimators are introduced in Section 2. Each estimator is the product of two one-dimensional Kaplan-Meier product limit estimators.

A hazard function approach is employed in Section 3 to estimate $-\ln F(s,t)$ and hence $F(s,t)$. Two path-dependent estimators of $-\ln F(s,t)$ are proposed and these lead to estimators of the bivariate distribution function.

Section 4 explores the relationship of the estimators of Sections 2 and 3. In Section 5 the pointwise consistency of the estimators that are products of Kaplan-Meier estimators is considered under mild conditions on F. Under stronger conditions, all the estimators of Sections 2 and 3 are proved to be uniformly almost sure consistent for F with rate $O(\sqrt{\frac{\log\log n}{n}})$ on the rectangle $[0,S] \times [0,T]$. The final section presents an example and some discussion.

2. Two Path-Dependent Product-Limit Estimators.

Let $\{X_i^\circ, Y_i^\circ\}_{i=1}^\infty$ be independent identically distributed pairs of nonnegative random variables with continuous bivariate survival function $F(s,t) = P(X^\circ>s, Y^\circ>t)$. Let $\{C_i,D_i\}_{i=1}^\infty$ denote another sequence of nonnegative i.i.d. pairs of random (censoring) variables with continuous survival function $G(s,t) = P(C>s, D>t)$. Define

$$X_i = \min\{X_i^\circ, C_i\}, \qquad Y_i = \min\{Y_i^\circ, D_i\} \qquad i=1, 2, \ldots, n$$

$$\varepsilon_{1i} = \begin{cases} 1 & X_i = X_i^\circ \text{ (uncensored)} \\ 0 & X_i < X_i^\circ \text{ (censored)} \end{cases}$$

$$\varepsilon_{2i} = \begin{cases} 1 & Y_i = Y_i^\circ \text{ (uncensored)} \\ 0 & Y_i < Y_i^\circ \text{ (censored)} \end{cases}$$

It is assumed that the two sequences $\{X_i^\circ, Y_i^\circ\}_{i=1}^\infty$ and $\{C_i, D_i\}_{i=1}^\infty$ are mutually independent. Let $H(s,t) = P(X_i>s, Y_i>t)$ denote the survival function of (X_i, Y_i). By independence,

$$H(s,t) = F(s,t)\, G(s,t). \tag{2.1}$$

Based on the elementary observation

$$F(s,t) = F(s,0)\, F_1(t|s), \tag{2.2}$$

where $F_1(t|s) = P(Y^\circ>t|X^\circ>s)$, the survival function $F(s,t)$ is estimated by separately estimating each of the two terms on the right of (2.2). This leads to an estimator $\hat{F}_1(s,t)$ based on the path from (0,0) to (s,t) which is linear from (0,0) to (s,0) and linear from (s,0) to (s,t).

The following notation is established: Let

$$N_n(s,t) = N(s,t) = \sum_{i=1}^{n} I_{\{X_i>s,\ Y_i>t\}}; \tag{2.3}$$

$$\alpha_i(s,t) = I_{\{X_i\le s,\ Y_i>t,\ \varepsilon_{1i}=1\}} \qquad i=1, 2, \ldots, n; \tag{2.4}$$

$$\beta_j(s,t) = I_{\{X_j>s,\ Y_j\le t,\ \varepsilon_{2j}=1\}} \qquad j=1, 2, \ldots, n. \tag{2.5}$$

To estimate $F(s,0)$, project all points vertically onto the line $y=0$, and, ignoring the (Y_i, ε_{2i}) values, calculate the Kaplan-Meier product-limit estimator of $F(s,0)$ using the one dimensional censored sample $\{X_i, \varepsilon_{1i}\}_{i=1}^n$. This produces the estimator

$$\hat{F}_{1n}(s,0) = \begin{cases} \prod_{i=1}^{n} \left(\frac{N(X_i,0)}{N(X_i,0)+1} \right)^{\alpha_i(s,0)} & \text{if } s \leq \tau_{1n} \\ 0 & \text{otherwise,} \end{cases}$$

where $\tau_{1n} = \max_{1 \leq i \leq n} \{X_i\}$.

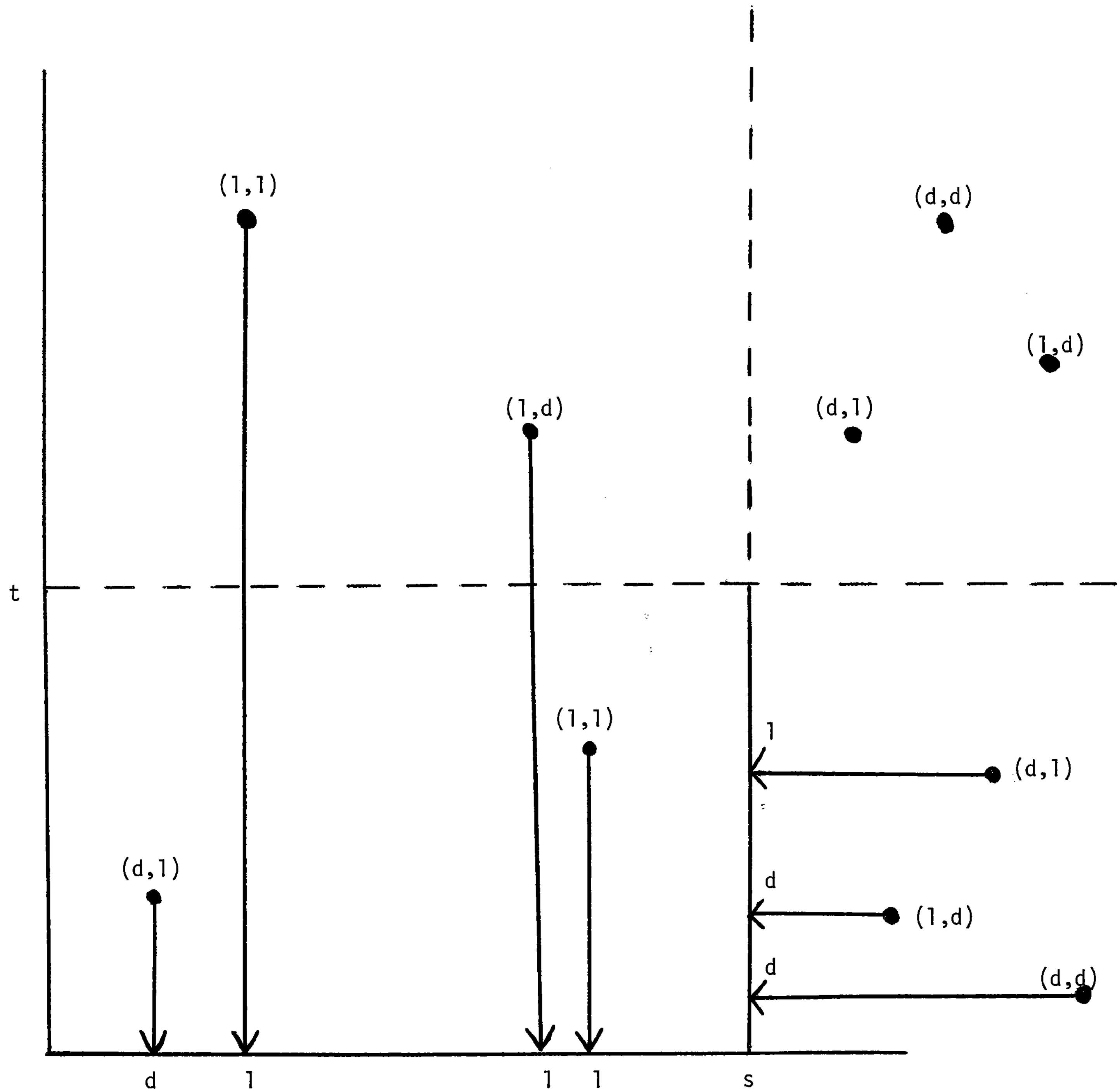
(The one-dimensional convention that the last observation is converted to a death (if it is censored) is adhered to here.) To estimate $F_1(t|s)$, the second term of (2.2), project all points for which $X_i > s$ horizontally to the line $X = s$, and ignoring the (X_i, ε_{1i}) values calculate the Kaplan-Meier product-limit estimator based on the data $\{Y_j, \varepsilon_{2j}\}_{j=1}^n$ for which $X_j > s$ (see Figure 1). Observe that this method estimates the probability $P(Y^\circ > t | X > s)$ but

$$(2.6) \qquad P(Y^\circ>t|X>s) = P(Y^\circ>t|X^\circ>s, C>s) = P(Y^\circ>|X^\circ>s)$$

since C is independent of the pair (X°, Y°). Thus the following estimator of $F_1(t|s)$ is obtained:

$$(2.7) \qquad \hat{F}_{1n}(t|s) = \begin{cases} \prod_{j=1}^{n} \left(\frac{N(s,Y_j)}{N(s,Y_j)+1} \right)^{\beta_j(s,t)} & \text{if } t \leq \tau_{2n}(s) \\ 0 & \text{otherwise,} \end{cases}$$

where $\tau_{2n}(s) = \max_{1 \leq i \leq n} \{Y_i : X_i > s\}$.



d = death
l = loss

FIGURE 1

Consequently the estimator for the F(s,t) is

$$\text{(2.8)}\qquad \hat{F}_{1n}(s,t) = \begin{cases} \prod_{i=1}^{n} \left(\frac{N(X_i,0)}{N(X_i,0)+1}\right)^{\alpha_i(s,0)} \prod_{j=1}^{n} \left(\frac{N(s,Y_j)}{N(s,Y_j)+1}\right)^{\beta_j(s,t)} & \text{if } N(s,t) > 0 \\ 0 & \text{otherwise.} \end{cases}$$

Remark 2.1. In the event of no censoring in either coordinate this estimator reduces to the ordinary empirical survival function $\frac{1}{n} \sum I_{\{X_i^\circ > s, Y_i^\circ > t\}}$.

Remark 2.2. By changing the role of s and t it is possible to develop our estimator $\hat{F}_{2n}(s,t)$ (based on the relation $F(s,t) = F(t,0)F_2(s|t) = F(t,0)P(X^\circ > s | Y^\circ > t)$) using the linear path from (0,0) to (0,t) and to (s,t). The corresponding estimator is

$$\text{(2.10)}\qquad \hat{F}_{2n}(s,t) = \begin{cases} \prod_{j=1}^{n} \left(\frac{N(0,Y_j)}{N(0,Y_j)+1}\right)^{\beta_j(0,t)} \prod_{j=1}^{n} \left(\frac{N(X_i,t)}{N(X_i,t)+1}\right)^{\alpha_i(s,t)} & \text{if } N(s,t) > 0 \\ 0 & \text{otherwise.} \end{cases}$$

All of the results which are true for $\hat{F}_{1n}(s,t)$ hold (using the same type of arguments) for $\hat{F}_{2n}(s,t)$. Therefore in what follows only $\hat{F}_{1n}(s,t)$ is treated.

Remark 2.3. One can easily give an example which shows that $\hat{F}_{1n}(s,t)$ is not necessarily a distribution function.

3. Estimators Based on the Bivariate Hazard Function.

The multivariate hazard gradient approach of Marshall (1975) is employed to develop bivariate survival function estimators based on the hazard function. Define the hazard function R(s,t) as

(3.1) $$R(s,t) = -\log F(s,t).$$

Assume that R is absolutely continuous with partial derivatives that exist almost everywhere. Let $\underline{r}(\underline{z})$ denote the gradient of $R(\underline{z})$ for $\underline{z} = (z_1, z_2)$; i.e.,

(3.2) $$\underline{r}(\underline{z}) = (r_1(\underline{z}), r_2(\underline{z})), \text{ where}$$

(3.3) $$r_1(\underline{z}) = \frac{\partial R(\underline{z})}{\partial z_1} \text{ and } r_2(\underline{z}) = \frac{\partial R(\underline{z})}{\partial z_2}.$$

Then R(s,t) can be reconstructed as the path integral of $\underline{r}(\underline{z})$ from (0,0) to (s,t). By path independence one can write

(3.4) $$R(s,t) = \int_{(0,0)}^{(s,t)} \underline{r}(\underline{z})\, d\underline{z}.$$

In particular, consider the path linear from (0,0) to (0,s) and linear from (s,0) to (s,t). Then

(3.5) $$R(s,t) = \int_0^s r_1(u,0)\, du + \int_0^t r_2(s,v)\, dv$$

By (3.1) and (3.3) from (3.5)

(3.6) $$-\log F(s,t) = -\int_0^s \frac{1}{F(u,0)}\, d_u F(u,0) - \int_0^t \frac{1}{F(s,v)}\, d_v F(s,v)$$

where $d_u F(u,t)$ denotes Lebésgue-Stieltjes integration over u for t fixed. Using (2.1),

(3.7) $$-\log F(s,t) = -\int_0^s \frac{G(u,0)}{H(u,0)} \frac{\partial F(u,0)}{\partial u}\, du - \int_0^t \frac{G(s,v)}{H(s,v)} \frac{\partial F(s,v)}{\partial v}\, dv.$$

Introduce the following functions:

(3.8) $$\tilde{K}(s,t) = \int_0^s G(u,t) \frac{\partial P(X^\circ \le u, Y^\circ > t)}{\partial u}\, du;$$

$$(3.9) \qquad \tilde{L}(s,t) = \int_0^t G(s,v) \frac{\partial P(X^\circ>s, Y^\circ\leq v)}{\partial v} \, dv.$$

Applying the trivial

$$(3.10) \qquad \frac{\partial P(X^\circ\leq u, Y^\circ>t)}{\partial u} = -\frac{\partial P(X^\circ>u, Y^\circ>t)}{\partial u}$$

and

$$(3.11) \qquad \frac{\partial P(X^\circ>s, Y^\circ\leq v)}{\partial v} = -\frac{\partial P(X^\circ>s, Y^\circ>v)}{\partial v}$$

relations yields

$$(3.12) \qquad -\log F(s,t) = \int_0^s \frac{1}{H(u,0)} \, d_u\tilde{K}(u,0) + \int_0^t \frac{1}{H(s,t)} \, d_v\tilde{L}(s,v).$$

Equation (3.12) suggests that H, $\tilde{K}$ and $\tilde{L}$ be estimated first. The natural estimator of H(u,v) is the empirical survival function:

$$(3.13) \qquad H_n(s,t) = \frac{1}{n} \sum I_{\{X_i>s, Y_i>t\}} = \frac{N(s,t)}{n}.$$

The basic idea of estimating $\tilde{K}(u,v)$ and $\tilde{L}(u,v)$ is the following observation:

$$(3.14) \qquad \tilde{K}(s,t) = \int_0^s G(u,t) \, d_u \, P(X^\circ\leq u, Y^\circ>t) = \int_0^s P(C>u, D>t) \, d_u \, P(X^\circ\leq u, Y^\circ>t)$$

$$= \int_0^s P(C>u, D>t | X^\circ = u, Y^\circ>t) \, d_u \, P(X^\circ\leq u, Y^\circ>t)$$

$$= \int_0^s P(C>u, Y>t | X^\circ = u, Y^\circ>t) \, d_u \, P(X^\circ\leq u, Y^\circ>t) = P(C\geq X^\circ, Y>t, X^\circ<s)$$

$$= P(X^\circ\leq C, Y>t, X\leq s) = E(\alpha_i(s,t)),$$

the last equality following from (2.4).

And similarly

$$(3.15) \qquad \tilde{L}(s,t) = E(\beta_j(s,t)).$$

Hence the natural estimators of $\tilde{K}(s,t)$ and $\tilde{L}(s,t)$ are

(3.16) $$\tilde{K}_n(s,t) = \frac{1}{n}\sum_{i=1}^{n} \alpha_i(s,t)$$

and

(3.17) $$\tilde{L}_n(s,t) = \frac{1}{n}\sum_{j=1}^{n} \beta_j(s,t).$$

Consequently by (3.12) - (3.17) estimate $R(s,t) = -\log F(s,t)$ by

(3.18) $$\tilde{R}_{1n}(s,t) = \int_0^s \frac{1}{H_n(u,0)}\, d_u\, \tilde{K}_n(u,0) + \int_0^t \frac{1}{H_n(s,v)}\, d_v\, \tilde{L}_n(s,v)$$
$$= \frac{1}{n}\sum_{i=1}^{n} \frac{\alpha_i(s,0)}{H_n(X_i,0)} + \frac{1}{n}\sum_{j=1}^{n} \frac{\beta_j(s,t)}{H_n(s,Y_j)}$$

if $N(s,t) > 0$ and let $\tilde{R}_1(s,t) = +\infty$ otherwise. Moreover, let

(3.19) $$\tilde{F}_{1n}(s,t) = \exp(-\tilde{R}_{1n}(s,t)).$$

4. Relationship of the Product-Limit and the Hazard Function Estimators.

Lemma 4.1.

(4.1) $$\sup_{\substack{0\le s<\infty \\ 0\le t<\infty}} |H_n(s,t) - H(s,t)| = 0\left(\sqrt{\frac{\log\log n}{n}}\right) \quad \text{a.s.}$$

(4.2) $$\sup_{\substack{0\le s<\infty \\ 0\le t<\infty}} |\tilde{K}_n(s,t) - \tilde{K}(s,t)| = 0\left(\sqrt{\frac{\log\log n}{n}}\right) \quad \text{a.s.}$$

(4.3) $$\sup_{\substack{0\le s<\infty \\ 0\le t<\infty}} |\tilde{L}_n(s,t) - \tilde{L}(s,t)| = 0\left(\sqrt{\frac{\log\log n}{n}}\right) \quad \text{a.s.}$$

Proof. Result (4.1) simply follows from the multi-dimensional law of the iterated logirithm for empirical distributions of Kiefer (1961). To prove (4.2) it is enough to observe that

(4.4) $\tilde{K}(s,t) = P(X^\circ \leq s, X^\circ - C \leq 0) - P(X^\circ \leq s, X^\circ - C \leq 0, Y \leq t).$

Therefore $\tilde{K}_n(s,t)$ can be considered as the difference of two empirical distributions:

(4.5) $\tilde{K}_n(s,t) = \frac{1}{n}\sum_{i=1}^{n} I_{\{X_i^\circ \leq s, X_i^\circ - C_i \leq 0\}} - \frac{1}{n}\sum_{j=1}^{n} I_{\{X_j^\circ \leq s, X_j^\circ - C_j \leq 0, Y_j \leq t\}}.$

That means that applying again Kiefer's result (once in two-, once in three-dimensions) (4.2) is obtained. A similar argument proves (4.3). □

Remark 4.2. Suppose that $H(S,T) > 0$ for $S,T < \infty$. Then from the above-mentioned Kiefer Theorem for almost all ω there exist an $n_0(\omega)$ such that if $n > n_0(\omega)$ then

(4.6) $H_n(s,t) > \frac{1}{2} H(s,t)$ for all $0 \leq s \leq S,\ 0 \leq t \leq T.$

For technical reasons introduce $\check{F}_{1n}$, the modified Kaplan-Meier-type estimator (a similar idea was used in Földes and Rejtő (1979)):

(4.7) $$\check{F}_{1n}(s,t) = \begin{cases} \prod_{i=1}^{n}\left(\frac{N(X_i,0)+1}{N(X_i,0)+2}\right)^{\alpha_i(s,0)} \Big/ \left(\prod_{j=1}^{n}\frac{N(s,Y_j)+1}{N(s,Y_j)+2}\right)^{\beta_j(s,t)} & \text{if } N(s,t) > 0 \\ 0 & \text{otherwise.} \end{cases}$$

Lemma 4.3. If $H(S,T) > 0$ $(0<S,T<\infty)$ then

(4.8) $$\sup_{\substack{0\leq s\leq S\\ 0\leq t\leq T}} |\hat{F}_{1n}(s,t) - \check{F}_{1n}(s,t)| = O\left(\frac{1}{n}\right) \quad \text{a.s.}$$

Proof. By the well-known inequality

(4.9) $$\left|\prod_{k=1}^{m} a_k - \prod_{k=1}^{m} b_k\right| \leq \sum_{k=1}^{m} |a_k-b_k| \quad \text{if } |a_k| \leq 1,\ |b_k| \leq 1,\ k=1,2,\ldots,m,$$

estimate the difference of the original and the modified estimators as follows:

(4.10) $$|\hat{F}_{1n}(s,t) - \check{F}_{1n}(s,t)| \leq \sum_{i=1}^{n} \frac{\alpha_i(s,0)}{(N(X_i,0)+1)^2} + \sum_{j=1}^{n} \frac{\beta_j(s,t)}{(N(s,Y_j)+1)^2}.$$

Hence applying Remark 4.2

(4.11) $$\sup_{\substack{0\leq s\leq S\\ 0\leq t\leq T}} |\hat{F}_{1n}(s,t) - \check{F}_{1n}(s,t)| \leq \frac{2n}{N^2(S,T)} \leq \frac{2n}{n^2H_n^2(S,T)}$$

$$\leq \frac{8}{nH^2(S,T)} = O(\tfrac{1}{n}) \quad \text{a.s.} \quad \square$$

Lemma 4.4. If $H(S,T) > 0$ $(0<S,T<\infty)$ then

(4.12) $$\sup_{\substack{0\leq s\leq S\\ 0\leq t\leq T}} |\hat{F}_{1n}(s,t) - \tilde{F}_{1n}(s,t)| = O(\tfrac{1}{n}) \quad \text{a.s.}$$

Proof. First observe that by the elementary inequality

(4.13) $$|x-y| < |\log x - \log y| \qquad (\text{for } 0 < x,y \leq 1)$$

and by Lemma 4.3:

$$(4.14)\ \sup_{\substack{0\le s\le S\\0\le t\le T}} |\hat{F}_{1n}(s,t) - \tilde{F}_{1n}(s,t)| \le \sup_{\substack{0\le s\le S\\0\le t\le T}} |\hat{F}_{1n}(s,t) - \check{F}_{1n}(s,t)| +$$

$$\sup_{\substack{0\le s\le S\\0\le t\le T}} |\check{F}_{1n}(s,t) - \tilde{F}_{1n}(s,t)|$$

$$= O(\tfrac{1}{n}) + \sup_{\substack{0\le s\le S\\0\le t\le T}} |\log \check{F}_{1n}(s,t) + \tilde{R}_{1n}(s,t)| \quad \text{a.s.}$$

Using logarithmic expansion

$$\log \check{F}_{1n}(s,t) = -\sum_{i=1}^{n} \frac{\alpha_i(s,0)}{N(X_i,0)+2} - \sum_{i=1}^{n} \alpha_i(s,0) \sum_{k=2}^{\infty} \frac{1}{(N(X_i,0)+2)^k k}$$

(4.15)

$$- \sum_{j=1}^{n} \frac{\beta_j(s,t)}{N(s,Y_j)+2} - \sum_{j=1}^{n} \beta_j(s,t) \sum_{k=2}^{\infty} \frac{1}{(N(s,Y_j)+2)^k k} =$$

$$- \sum_{i=1}^{n} \frac{\alpha_i(s,0)}{N(X_i,0)+2} - \sum_{j=1}^{n} \frac{\beta_j(s,t)}{N(s,Y_j)+2} + D_n(s,t)$$

where $D_n(s,t)$ denotes the remainder term. Now

$$(4.16)\ \sup_{\substack{0\le s\le S\\0\le t\le T}} |D_n(s,t)| \le \frac{2n}{(N(S,T)+2)^2} \le \frac{8n}{n^2H^2(S,T)} = O(\tfrac{1}{n}) \quad \text{a.s.}$$

by estimating the infinite sums by geometric series, and using Remark 4.2. Hence from (4.13) - (4.16) and by (3.18)

$$(4.17)\quad \sup_{\substack{0\le s\le S\\0\le t\le T}} |\hat{F}_{1n}(s,t) - \tilde{F}_{1n}(s,t)| \le \left|\frac{1}{n}\sum_{i=1}^{n} \alpha_i(s,0)\left(\frac{1}{H_n(X_i,0)} - \frac{1}{H_n(X_i,0)+\frac{2}{n}}\right)\right| +$$

$$\left|\frac{1}{n}\sum_{j=1}^{n} \beta_j(s,t)\left(\frac{1}{H_n(s,Y_j)} - \frac{1}{H_n(s,Y_j)+\frac{2}{n}}\right)\right| + 0(\tfrac{1}{n}) \le$$

$$\frac{2\,\frac{2}{n}}{(H_n(S,T))(H_n(S,T)+\frac{2}{n})} + 0(\tfrac{1}{n}) \le \frac{16}{nH^2(S,T)} + 0(\tfrac{1}{n}) = 0(\tfrac{1}{n}) \quad \text{a.s.}$$

again using Remark 4.2. □

Remark 4.5. Lemma 4.4 gives a large-sample result for the proximity of $\hat{F}_{1n}$ and $\tilde{F}_{1n}$. In fact, an absolute bound can be obtained by repeated application of a one-dimensional result of Breslow and Crowley (1974); namely,

$$0 < -\ln \hat{F}_{1n}(s,t) + \ln \tilde{F}_{1n}(s,t) \le (n-N_n(s,0))/nN_n(s,0)$$

$$+ (N_n(s,0)-N_n(s,t))/N_n(s,0)N_n(s,t) = (n-N_n(s,t))/nN_n(s,t).$$

In particular, $\tilde{F}_{1n}(s,t) \ge \hat{F}_{1n}(s,t)$.

5. Consistency.

The pointwise consistency of the estimator $\hat{F}_{1n}$ follows from the corresponding one-dimensional results, in that the estimator was constructed as a product of two one-dimensional Kaplan-Meier estimators. The pointwise consistency remains true in case of not necessarily continuous functions F and G, as one can develop, using the same projecting argument the corresponding bivariate estimator as the product of two one-dimensional Kaplan-Meier estimators. Observe that for this pointwise consistency neither the continuity condition on F nor G is required in that the Kaplan-Meier estimator is consistent (see Winter, Földes and Rejtö (1978) and Földes, Rejtö and Winter (1979)). Under some smoothness conditions the following much stronger theorem is now proved.

Theorem 5.1. If F and G are continuous and if F is such that -ln F is absolutely continuous with partial derivatives that exist almost everywhere and if, for $0 < S,T < \infty$, $H(S,T) > 0$, then

$$(4.18)\quad \sup_{\substack{0\le s\le S\\ 0\le t\le T}} |\tilde{F}_{1n}(s,t) - F(s,t)| = O\left(\sqrt{\frac{\log\log n}{n}}\right) \text{ a.s.}$$

Proof. Applying again (4.13), (3.12), (3.18) and (3.19) it is possible to estimate the left-hand side of (4.18) as follows:

$$(4.19)\quad |\tilde{F}_{1n}(s,t) - F(s,t)| \le \left|\int_0^s \frac{1}{H_n(u,0)}\, d_u\tilde{K}_n(u,0) - \int_0^s \frac{1}{H(u,0)}\, d_u\tilde{K}(u,0)\right| +$$

$$\left|\int_0^t \frac{1}{H_n(s,v)}\, d_v\tilde{L}_n(s,v) - \int_0^t \frac{1}{H(s,v)}\, d_vL(s,v)\right|$$

Both of the terms of (4.19) can be estimated using Lemma 4.1, Remark 4.2 and partial integration as follows (we perform only their estimation for the first terms). (Observe that $\tilde{K}_n(u,0)$ and $\tilde{K}(u,0)$ are nondecreasing in u, $\tilde{L}_n(s,t)$ and $\tilde{L}(s,t)$ are non-decreasing in t for fixed s, and $H_n(s,t)$ and $H(s,t)$ non-increasing in both arguments.

$$\sup_{\substack{0<s<S\\0\le t\le T}} \left|\int_0^s \frac{1}{H_n(u,0)}\, d_u\tilde{K}_u(u,0) - \int_0^s \frac{1}{H(u,0)}\, d_u\tilde{K}_n(u,0)\right| \le$$

$$\sup_{\substack{0<s<S\\0\le t\le T}} \int_0^s \left|\frac{1}{H_u(u,0)} - \frac{1}{H(u,0)}\right| d_u\tilde{K}_n(u,0) + \sup_{\substack{0<s<S\\0\le t\le T}} \left|\int \frac{1}{H(u,0)}\, d_u(\tilde{K}_n(u,0)-\tilde{K}(u,0))\right|$$

$$\le \sup_{\substack{0\le s\le S\\0\le t\le T}} |H_n(s,t) - H(s,t)| \int_0^s \frac{2}{H^2(u,0)}\, d_u\tilde{K}_n(u,0) + \frac{1}{H(S,T)}\, 2 \sup_{\substack{0<s\le S\\0\le t\le T}} |\tilde{K}_n(s,t)-\tilde{K}(s,t)| \le$$

$$= 0\left(\sqrt{\frac{\log\log n}{n}}\right)\left(\frac{2}{H^2(S,T)} + \frac{2}{H(S,T)}\right) = 0\left(\sqrt{\frac{\log\log n}{n}}\right) \quad \text{a.s.} \quad \square$$

Corollary 5.2. Under the conditions of Theorem 5.1,

$$(4.20) \quad \sup_{\substack{0<s<S\\0\le t\le T}} |\hat{F}_{1n}(s,t) - F(s,t)| = 0\left(\sqrt{\frac{\log\log n}{n}}\right) \quad \text{a.s.}$$

Proof. Apply Lemmas 4.3 and 4.4 in conjunction with Theorem 5.1.

6. An Example.

Consider the example of Figure 2 consisting of four points. Here l denotes loss, and d death, so that the (d,l) at (x_3,y_3) denotes a point which is a death in the first coordinate and censored in the second. At each point in the rectangle $[0,x_4] \times [0,y_4]$ the estimator $\hat{F}_{14}$ can be calculated. Note that in the calculation of $\hat{F}_{14}(x,0)$ the final loss in the first coordinate is converted to a death (as is the convention in one-dimensional Kaplan-Meier estimation). Suppose we wish to compute $\hat{F}_1(s,t)$ where $s \in (x_2,x_3)$ and $t \in (y_3,y_2)$. Then

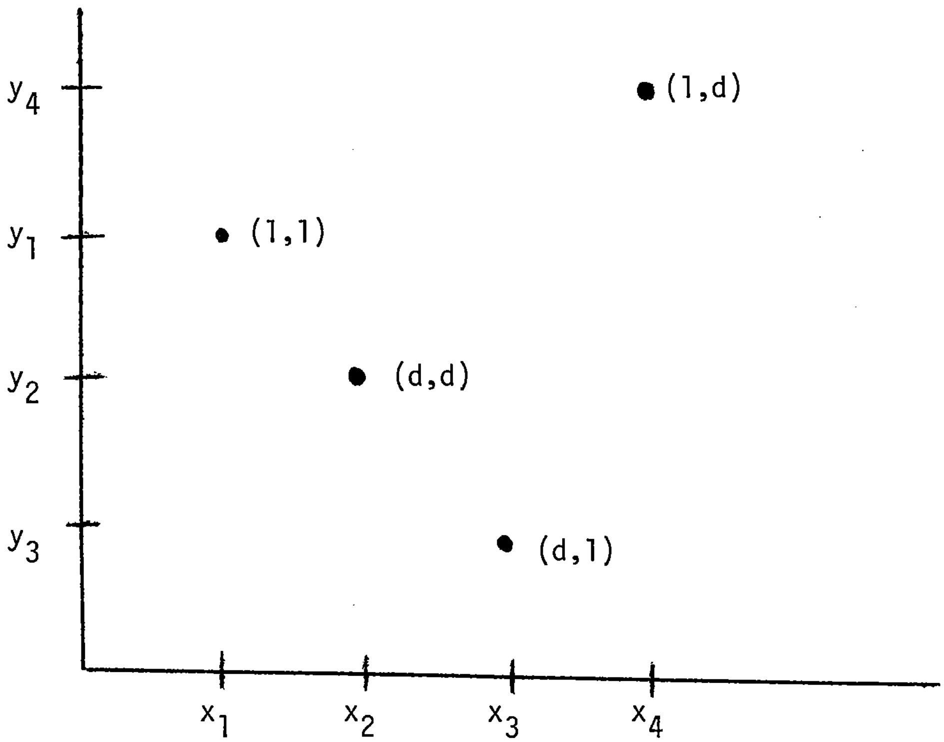
$$\hat{F}_{14}(s,0) = \left(\frac{3}{4}\right)^0\left(\frac{2}{3}\right)^1$$

To compute $\hat{F}_{14}(t|s)$, note that there are two points such that $x_i > s$, the point (x_3,y_3) projected back to the line $x = s$ is a loss, the point (x_4,y_4) projected to $x = s$ is a death. Thus $\hat{F}_{14}(t|s) = \left(\frac{1}{2}\right)^0 = 1$. Therefore $\hat{F}_{14}(s,t) = \hat{F}_{14}(s,0)\ \hat{F}_{14}(t|s) = \frac{2}{3} \cdot 1 = \frac{2}{3}$. Another way to arrive to the same conclusion is to follow (2.8) and then we get (see Figure 3)

$$\hat{F}_{14}(s,t) = \left(\frac{N(x_1,0)}{N(x_1,0)+1}\right)^{\varepsilon_{11}} \cdot \left(\frac{N(x_2,0)}{N(x_2,0)+1}\right)^{\varepsilon_{12}} \cdot \frac{N(s,y_3)}{N(s,y_3)+1}^{\varepsilon_{23}} = \left(\frac{3}{4}\right)^0\left(\frac{2}{3}\right)^1\left(\frac{1}{2}\right)^0 = \frac{2}{3}\ .$$

In this way the entire estimator $\hat{F}_{14}$ can be calculated. It is displayed in Figure 4 where the function is constant on the smaller rectangles. It was remarked earlier that $\hat{F}_{14}$ need not be a bivariate survival function and it is not one for this example. Figure 5 presents the estimator $\hat{F}_{24}$ based on the alternate path.

While the estimator in $\hat{F}_{14}$ is not guaranteed to be a bivariate survival function, Section 5 nonetheless proves under suitable conditions that as the sample size tends to infinity that the estimator is uniformly almost surely consistent for the true survival distribution function F (with rate $0\left(\sqrt{\frac{\log\log n}{n}}\right)$). This technique can be generalized from two dimensions to higher dimensions. The difference is that the number of possible paths (and hence the estimators) increases from 2 to 2^{k-1} for the k-dimensional analog, but the uniformly almost sure consistency (with the same rate) of each of these 2^{k-1} estimators can be shown.



$\varepsilon_{11} = 0$ $\quad \varepsilon_{21} = 0$

$\varepsilon_{12} = 1$ $\quad \varepsilon_{22} = 1$

$\varepsilon_{13} = 1$ $\quad \varepsilon_{23} = 0$

$\varepsilon_{14} = 0$ $\quad \varepsilon_{24} = 1$

FIGURE 2
A FOUR-POINT EXAMPLE

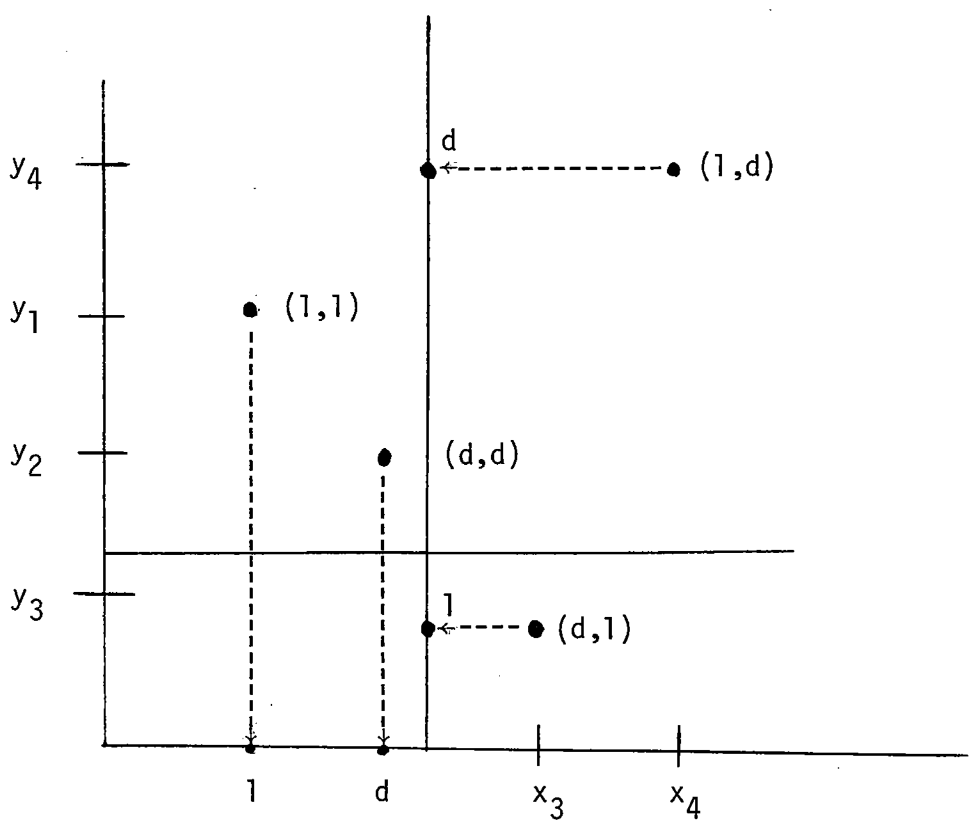


FIGURE 3
THE PROJECTION RULE

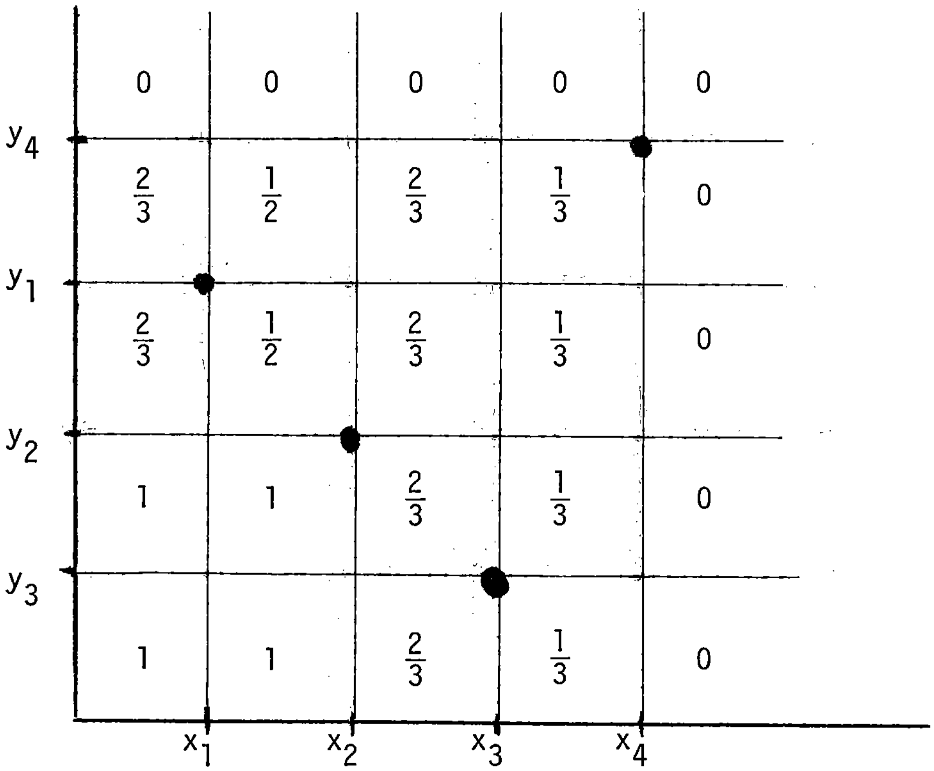


FIGURE 4
THE SURVIVAL ESTIMATE $\hat{F}_1$

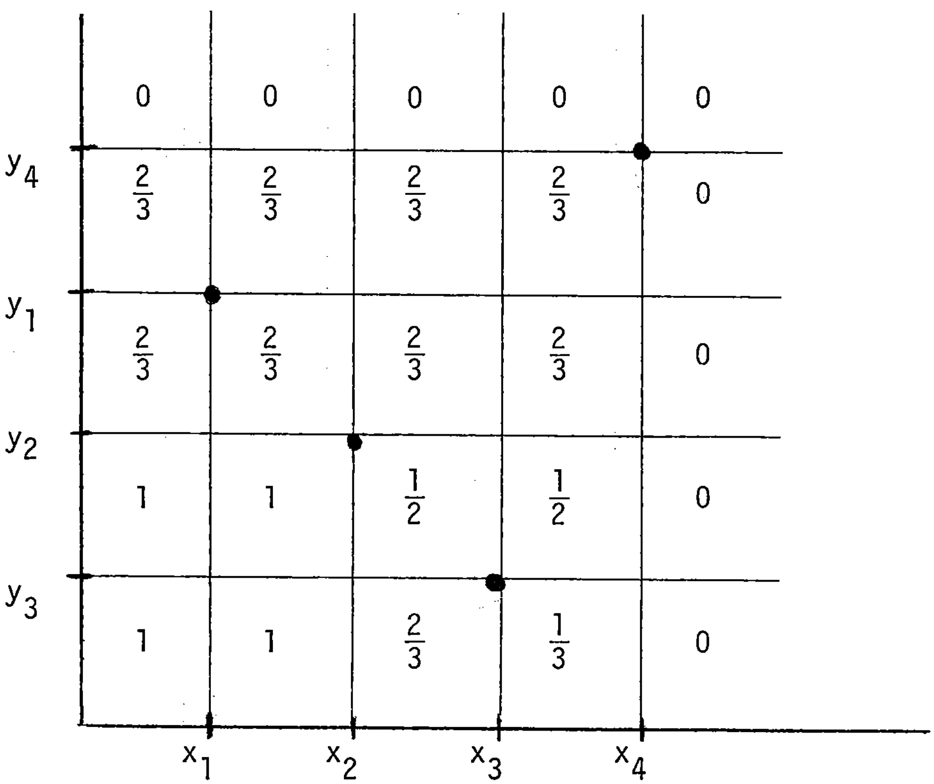


FIGURE 5
THE SURVIVAL ESTIMATE $\hat{F}_2$

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.

Campbell, G. (1979). Nonparametric bivariate estimation with randomly censored data. Purdue University Department of Statistics Mimeoseries #79-25.

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

Földes, A. and Rejtő, L. (1979). A LIL type result for the product limit estimator on the whole line. Mathematical Institute of the Hungarian Academy of Sciences Technical Report No. 50.

Földes, A., Rejtő, L., and Winter, B. B. (1979). Strong consistency properties of nonparametric estimators for randomly censored data. to appear in Periodica Math Hung.

Hanley, J. A. and Parnes, M. N. (1980). Estimation of a multivariate distribution in the presence of censoring. Biometrics (to appear).

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

Kiefer, J. (1961). On large deviations of the empiric d.f. of vector chance variables and a law of the iterated logarithm. Pacific J. of Math. 11, 649-660.

Marshall, A. W. (1975). Some comments on the hazard gradient. Stochastic Processes and their Applications 3, 293-300.

Winter, B. B., Földes, A., and Rejtő, L. (1978). Glivenko-Cantelli theorems for the PL estimate. Problems of Control and Information Theory 7, 213-225.