ASYMPTOTIC PROPERTIES OF SEVERAL NONPARAMETRIC MULTIVARIATE DISTRIBUTION FUNCTION ESTIMATORS UNDER RANDOM CENSORING

by

Gregory Campbell
Purdue University

Mimeograph Series #81-54

Department of Statistics
November 1981

1. INTRODUCTION

The problem of nonparametric estimation of a multivariate distribution function in the presence of random censoring is considered. The multivariate lifetimes could represent the times to death of animals in fixed-sized litters, the failure times of components in a multicomponent system, the observations of participants of a matched triples study, or the onset times to stages of a disease in a patient. In the special bivariate case, there are the numerous examples of paired data on eyes, lungs, kidneys, twins or married couples. It is possible that the censoring is univariate or multivariate. Whereas the censoring of times to death of animals in litters born at random times yet truncated at a fixed time is an example of univariate censoring, the truncation at a fixed time of measures on the participants in a matched triple study would provide trivariate independent censoring. The study of lifelengths of twins and married couples would provide an example of bivariate censoring with possible dependence between the two censoring variables.

The estimation of one-dimensional distribution function estimators with randomly censored data has been extensively developed. The product-limit estimator was proposed by Kaplan and Meier (1958). Under suitable conditions, asymptotic normality and weak convergence of this estimator was established by Breslow and Crowley (1974) and strong uniform almost sure convergence was proved by Földes and Rejtő (1981).

The bivariate problem has merited some attention recently. Campbell (1981) estimated the bivariate distribution function under bivariate censoring for discrete or grouped data via the EM algorithm. Leurgans, Tsai, and Crowley (1982) have proposed an estimator for univariate censoring that utilizes Freund's bivariate exponential distribution. Campbell and

Földes (1982) have proposed several estimators based on hazard gradient estimators and on products of one-dimensional product-limit estimators. It is the weak convergence of the latter estimators which is the purpose of this paper.

A path-dependent distribution function estimator based on the hazard gradient is introduced in §2 after some notational development. The result of strong uniform almost sure convergence of the estimator which was proved in Campbell and Földes (1982) is presented.

A topological discussion in §3 precedes an important lemma on empirical processes in two-dimensional time. The main theorems of §4 prove the weak convergence of the suitably normalized estimator. The discussion in the final section considers estimators with different paths as well as estimators which are products of product-limit estimators. The extension from two to k dimensions is noted.

2. NOTATION AND THE ESTIMATOR

For simplicity of exposition, bivariate observations are considered. Let $\{\underset{\sim}{X}_i\}_{i=1}^{\infty}$ denote a sequence of independent random variables, $\underset{\sim}{X}_i = (X_{i1}, X_{i2})$, from the continuous bivariate distribution function F. Each $\underset{\sim}{X}_i$ represents the lifetimes of a pair of (possibly dependent) items. Let $\{\underset{\sim}{C}_i\}_{i=1}^{\infty}$ denote a sequence of independent random variables, $\underset{\sim}{C}_i = (C_{i1}, C_{i2})$, from the continuous bivariate distribution function G. It is assumed that $\{\underset{\sim}{X}_i\}_{i=1}^{\infty}$ and $\{\underset{\sim}{C}_i\}_{i=1}^{\infty}$ are mutually independent.

In general $\underset{\sim}{X}_i$ and $\underset{\sim}{C}_i$ are not both observable. Define

$$Z_{ji} = \min(X_{ji}, C_{ji});$$

$$\varepsilon_{ji} = I_{\{X_{ji} \leq C_{ji}\}}, \qquad j=1,2;\ i=1,2,\ldots,$$

where $I_A(x)$ is 1(0) if $x \in (\notin) A$. Note that ε corresponds to whether Z is an uncensored value (ε=1) or censored value (ε=0). It is assumed that $\underset{\sim}{Z}_i = (Z_{1i}, Z_{2i})$ and $\underset{\sim}{\varepsilon}_i = (\varepsilon_{1i}, \varepsilon_{2i})$ are observable. Let H denote the distribution function of $\underset{\sim}{Z}_i$. Define the bivariate survival function

$$\bar{F}(\underset{\sim}{t}) \equiv \bar{F}(t_1, t_2) \equiv P(X_1 > t_1, X_2 > t_2)$$

and with abuse of notation for $\underset{\sim}{t} = (t_1, t_2)$

$$F(\bar{t}_1, t_2) = P(X_1 > t_1, X_2 \le t_2); \; F(t_1, \bar{t}_2) = P(X_1 \le t_1, X_2 > t_2).$$

Similar functions can be defined for G and H. By independence of $\underset{\sim}{X}$ and $\underset{\sim}{C}$

$$\bar{H}(t_1, t_2) = \bar{F}(t_1, t_2)\bar{G}(t_1, t_2) \qquad (1)$$

for all t_1, t_2.

The hazard gradient approach of Marshall (1975) was employed by Campbell and Földes (1982) to estimate the distribution function as indicated below. The cumulative hazard function is given by

$$R(\underset{\sim}{t}) = -\ln \bar{F}(\underset{\sim}{t}). \qquad (2)$$

Assuming R is absolutely continuous with partial derivatives that exist almost everywhere, let $\underset{\sim}{r}(\underset{\sim}{t})$ denote the gradient of $R(\underset{\sim}{t})$. Then $R(\underset{\sim}{t})$ can be represented as the path-independent integral of $r(\underset{\sim}{z})$ from (0,0) to $\underset{\sim}{t}$. In particular, for the linear path (0,0) to $(t_1, 0)$ to (t_1, t_2),

$$R(\underset{\sim}{t}) = \int_0^{t_1} r_1(u,0)du + \int_0^{t_2} r_2(t_1,v)dv$$

where $r_1(\underset{\sim}{s}) = \frac{\partial R(\underset{\sim}{s})}{\partial s_1}$, $r_2(\underset{\sim}{s}) = \frac{\partial R(\underset{\sim}{s})}{\partial s_2}$ for $\underset{\sim}{s} = (s_1, s_2)$; i.e.,

$$R(\underset{\sim}{t}) = \int_0^{t_1} (\bar{F}(u,0))^{-1} d_u F(u,\bar{0}) + \int_0^{t_2} (\bar{F}(t_1,v))^{-1} d_v F(\bar{t}_1, v)$$

where $d_u F(u,\bar{s})$ and $d_v F(\bar{s},v)$ denote Lebesgue-Stieljes integration over u and v, respectively, with s fixed.

Define

$$K_1(\underset{\sim}{t}) = P(Z_1 \le t_1, Z_2 > t_2, \varepsilon_1 = 1) = \int_0^{t_1} \bar{G}(u,t_2)d_uF(u,\bar{t}_2);$$
$$K_2(\underset{\sim}{t}) = P(Z_1 > t_1, Z_2 \le t_2, \varepsilon_2 = 1) = \int_0^{t_2} \bar{G}(t_1,v)d_vF(\bar{t}_1,v). \tag{3}$$

Then

$$R(\underset{\sim}{t}) = \int_0^{t_1} (\bar{H}(u,0))^{-1}d_uK_1(u,0) + \int_0^{t_2} (\bar{H}(t_1,v))^{-1}d_vK_2(t_1,v) \tag{4}$$

Estimate H, K_1, and K_2 by the empiricals

$$H_n(\underset{\sim}{t}) = \frac{1}{n}\sum_{i=1}^{n} I_{(Z_{1i}\le t_1, Z_{2i}\le t_2)};$$

$$K_{1n}(\underset{\sim}{t}) = \frac{1}{n}\sum_{i=1}^{n} \alpha_{1i}(\underset{\sim}{t});$$

$$K_{2n}(\underset{\sim}{t}) = \frac{1}{n}\sum_{i=1}^{n} \alpha_{2i}(\underset{\sim}{t});$$

where $\alpha_{1i}(\underset{\sim}{t}) = I_{(Z_{1i}\le t_1, Z_{2i}>t_2, \varepsilon_{1i}=1)}$ and $\alpha_{2i}(\underset{\sim}{t}) = I_{(Z_{1i}>t_1, Z_{2i}\le t_2, \varepsilon_{2i}=1)}$. Then R(t) is estimated from $\{\underset{\sim}{Z}_i\}_{i=1}^n$ and $\{\underset{\sim}{\varepsilon}_i\}_{i=1}^n$ by

$$R_n(\underset{\sim}{t}) = \int_0^{t_1} (\bar{H}_n(u,0))^{-1}d_uK_{1n}(u,0) + \int_0^{t_2} (\bar{H}_n(t_1,v))^{-1}d_vK_{2n}(t_1,v) \tag{5}$$

and $\bar{F}(\underset{\sim}{t})$ by

$$\bar{F}_n(\underset{\sim}{t}) = \exp\{-R_n(t)\}. \tag{6}$$

If F and G are continuous and if T_1 and T_2 are such that $\bar{H}(T_1,T_2) > 0$, Campbell and Földes (1982) proved

$$\sup_{\substack{0<t_1\le T_1 \\ 0<t_2\le T_2}} |\bar{F}_n(\underset{\sim}{t}) - \bar{F}(\underset{\sim}{t})| = \mathcal{O}\left(\sqrt{\frac{\ln\ln n}{n}}\right) \quad \text{a.s.}$$

3. WEAK CONVERGENCE OF EMPIRICALS IN TWO-DIMENSIONAL TIME

The study of the weak convergence of empirical processes in multidimensional time culminated in articles by Neuhaus (1971) and Straf (1971). The approach of Neuhaus (1971) is the reference for the topological discussion below.

For simplicity one can reduce the domain of the bivariate distribution function F from $[0,\infty)\times[0,\infty)$ to the unit square, $[0,1]\times[0,1]$ by the transformation $u_1 = F(t_1,\bar{0})$ and $u_2 = F(t_2|t_1) = P(X_2 \leq t_2|X_1 \leq t_1)$, as suggested in Durbin (1970). The approach of Neuhaus (1971) is to restrict the real-valued functions from the unit square. For the point $\underset{\sim}{t} = (t_1,t_2)$ inside the unit square, let Q_1,Q_2,Q_3,Q_4 denote the four open quadrants in the square determined by $\underset{\sim}{t}$, where Q_1 is the upper right quadrant. The space D_2 is the set of all real functions from the unit square such that if $\{\underset{\sim}{t}_n\}$ denotes a sequence in Q_i such that $\lim_{n\to\infty} \underset{\sim}{t}_n = \underset{\sim}{t}$ then $\lim_{n\to\infty} f(\underset{\sim}{t}_n)$ exists for $i=1,2,3,4$ and for $i=1$ its limiting value is $f(\underset{\sim}{t})$. Let Λ denote the class of all continuous functions from [0,1] onto itself. Let $\underset{\sim}{\lambda} = (\lambda_1,\lambda_2)\in\Lambda\times\Lambda$ and $|\underset{\sim}{t}|$ denote Euclidean distance in the plane. Define the metric d (which can be thought of as an extension of the one-dimensional Skorohod metric) for f,g in D_2 as

$$d(f,g) = \inf_{\varepsilon>0}\{\varepsilon\text{: there exists } \underset{\sim}{\lambda} = \underset{\sim}{\lambda}_\varepsilon \text{ with } \sup_{\underset{\sim}{t}}|\underset{\sim}{\lambda}(\underset{\sim}{t})-\underset{\sim}{t}| \leq \varepsilon$$

$$\text{and } \sup_{\underset{\sim}{t}}|f(\underset{\sim}{t})-g(\underset{\sim}{\lambda}(\underset{\sim}{t}))| \leq \varepsilon\}\ .$$

Then (D_2,d) is a separable complete metric space, unlike the space of discontinuous functions in the unit square with the metric d or the sup metric. Therefore, the Prohorov development of weak convergence is applicable.

Let

$$U_{1n}(\underset{\sim}{t}) = \sqrt{n}\ (H_n(t_1,\bar{t}_2) - H(t_1,\bar{t}_2);\ U_{2n}(\underset{\sim}{t}) = \sqrt{n}\ (H_n(\bar{t}_1,t_2) - H(\bar{t}_1,t_2));$$
$$V_{jn}(\underset{\sim}{t}) = \sqrt{n}\ (K_{jn}(\underset{\sim}{t})-K_j(\underset{\sim}{t})),\quad j=1,2. \tag{7}$$

<u>Lemma.</u> Let T_1,T_2 be such that $\bar{H}(T_1,T_2) > 0$, then as $n \to \infty$ $(U_{1n}(\underset{\sim}{t}),U_{2n}(\underset{\sim}{t}),V_{1n}(\underset{\sim}{t}),V_{2n}(\underset{\sim}{t}))$ converges weakly to a quadrivariate, two-dimensional-time Gaussian process $(U_1(\underset{\sim}{t}),U_2(\underset{\sim}{t}),V_1(\underset{\sim}{t}),V_2(\underset{\sim}{t}))$ with mean $(0,0,0,0)$ and covariance structure given below, where $a \wedge b = \min(a,b)$, $a \vee b = \max(a,b)$, and $\underset{\sim}{s} = (s_1,s_2)$

$$\mathrm{Cov}(U_1(\underset{\sim}{s}),U_1(\underset{\sim}{t})) = H(s_1 \wedge t_1,\overline{s_2 \vee t_2}) - H(s_1,\bar{s}_2)H(t_1,\bar{t}_2);$$

$$\mathrm{Cov}(U_2(\underset{\sim}{s}),U_2(\underset{\sim}{t})) = H(\overline{s_1 \vee t_1},s_2 \wedge t_2) - H(\bar{s}_1,s_2)H(\bar{t}_1,t_2);$$

$$\mathrm{Cov}(V_1(\underset{\sim}{s}),V_1(\underset{\sim}{t})) = K_1(s_1 \wedge t_1,s_2 \vee t_2) - K_1(s_1,s_2)K_1(t_1,t_2);$$

$$\mathrm{Cov}(V_2(\underset{\sim}{s}),V_2(\underset{\sim}{t})) = K_2(s_1 \vee t_1,s_2 \wedge t_2) - K_2(s_1,s_2)K_2(t_1,t_2);$$

$$\mathrm{Cov}(U_1(\underset{\sim}{s}),U_2(\underset{\sim}{t}) = \begin{cases} -H(s_1,\bar{s}_2)H(\bar{t}_1,t_2) & \text{if } s_1 < t_1 \text{ or } s_2 > t_2 \\ H(s_1,t_2) + H(s_2,t_1) - H(s_1,s_2) - H(t_1,t_2) & \\ \qquad - H(s_1,\bar{s}_2)H(\bar{t}_1,t_2) & \text{if } s_1 \geq t_1,\ s_2 \leq t_2; \end{cases}$$

$$\mathrm{Cov}(U_1(\underset{\sim}{s}),V_1(\underset{\sim}{t})) = K_1(s_1 \wedge t_1,s_2 \vee t_2) - H(s_1,\bar{s}_2)K_1(t_1,t_2);$$

$$\mathrm{Cov}(U_1(\underset{\sim}{s}),V_2(\underset{\sim}{t})) = \begin{cases} -H(s_1,\bar{s}_2)K_2(t_1,t_2) & \text{if } s_1 < t_1 \text{ or } s_2 > t_2 \\ K_2(s_1,t_2) + K_2(s_2,t_1) - K_2(s_1,s_2) - K_2(t_1,t_2) & \\ \qquad - H(s_1,\bar{s}_2)K_2(t_1,t_2) & \text{if } s_1 \geq t_1,\ s_2 \leq t_2; \end{cases}$$

$$\mathrm{Cov}(U_2(\underset{\sim}{s}),V_1(\underset{\sim}{t})) = \begin{cases} -H(\bar{s}_1,s_2)K_1(t_1,t_2) & \text{if } s_1 > t_1 \text{ or } s_2 < t_2 \\ K_1(s_1,t_2) + K_1(t_1,s_2) - K_1(s_1,s_2) - K_1(t_1,t_2) & \\ \qquad - H(\bar{s}_1,s_2)K_1(t_1,t_2) & \text{if } s_1 \leq t_1,\ s_2 \geq t_2; \end{cases}$$

$$\mathrm{Cov}(U_2(\underset{\sim}{s}),V_2(\underset{\sim}{t})) = K_2(s_1 \vee t_1, s_2 \wedge t_2) - H(\bar{s}_1,s_2)K_2(t_1,t_2);$$

$$\mathrm{Cov}(V_1(\underset{\sim}{s}),V_2(\underset{\sim}{t})) = \begin{cases} -K_1(s_1,s_2)K_2(t_1,t_2) & \text{if } s_1 < t_1 \text{ or } s_2 > t_2 \\ J(s_1,t_2) + J(t_1,s_2) - J(s_1,s_2) - J(t_1,t_2) & \\ \qquad - K_1(s_1,s_2)K_2(t_1,t_2) & \text{if } s_1 \geq t_1,\ s_2 \leq t_2, \end{cases}$$

where $J(t_1,t_2) = P(Z_1 \leq t_1, Z_2 \leq t_2, \varepsilon_1=1, \varepsilon_2=1)$.

Proof. The finite dimensional distributions converge to multivariate normal distributions by an application of the multivariate central limit theorem to the four-dimensional variables $(I_{(Z_{1i} \leq t_1, Z_{2i} > t_2)} - H(t_1,\bar{t}_2)$, $I_{(Z_{1i} > t_1, Z_{2i} \leq t_2)} - H(\bar{t}_1,t_2)$, $I_{(Z_{1i} \leq t_1, Z_{2i} > t_2, \varepsilon_{1i}=1)} - K_1(\underset{\sim}{t})$, $I_{(Z_{1i} > t_1, Z_{2i} \leq t_2, \varepsilon_{2i}=1)} - K_2(\underset{\sim}{t}))$. A simple calculation on these indicator variables yields the covariance structure of the lemma. In order to prove tightness in four dimensions, the tightness result of Neuhaus (1971, pp. 1292-5) for a one-dimensional empirical function of a multidimensional time parameter is applied for U_{1n}, U_{2n}, V_{1n}, and V_{2n} separately. Then tightness of the distributions of $(U_{1n}, U_{2n}, V_{1n}, V_{2n})$ follows immediately.

4. MAIN THEOREMS

The theorems in this section proceed in a fashion similar to the proof of weak convergence in one dimension in Breslow and Crowley (1974). The random variables $(U_{1n}, U_{2n}, V_{1n}, V_{2n})$ and (U_1, U_2, V_1, V_2) can be replaced by random variables with the same finite dimensional distribution but which also satisfy the condition that $d((U_{1n}, U_{2n}, V_{1n}, V_{2n}), (U_1, U_2, V_1, V_2))$ converges to zero almost surely, where d also represents the extension to $D_2 \times D_2 \times D_2 \times D_2$ of the metric d on D_2.

THEOREM 1. If F and G are continuous bivariate distribution functions and if T_1, $T_2 < \infty$ are such that $\bar{H}(T_1,T_2) > 0$, then for $\underset{\sim}{t} = (t_1,t_2)$ with $0 < t_1 < T_1$, $0 < t_2 < T_2$,

$$\sqrt{n}\ (R_n(\underset{\sim}{t})-R(\underset{\sim}{t}))$$

converges weakly to a two-dimensional-time Gaussian process $W(\underset{\sim}{t})$ given by:

$$W(\underset{\sim}{t}) = A_1(\underset{\sim}{t}) + B_1(\underset{\sim}{t}) + A_2(\underset{\sim}{t}) + B_2(\underset{\sim}{t}),$$

where

$$\begin{aligned}
A_1(\underset{\sim}{t}) &= \int_0^{t_1} (\bar{H}(u,0))^{-2}U_1(u,0)d_uK_1(u,0) \\
B_1(\underset{\sim}{t}) &= (\bar{H}(t_1,0))^{-1}V_1(t_1,0) - \int_0^{t_1} (\bar{H}(u,0))^{-2}V_1(u,0)d_uH(u,\bar{0}) \\
A_2(\underset{\sim}{t}) &= \int_0^{t_2} (\bar{H}(t_1,v))^{-2}U_2(t_1,v)d_vK_2(t_1,v) \\
B_2(\underset{\sim}{t}) &= (\bar{H}(t_1,t_2))^{-2}V_2(t_1,t_2) - \int_0^{t_2} (\bar{H}(t_1,v))^{-2}V_2(t_1,v)d_vH(\bar{t}_1,v).
\end{aligned} \tag{8}$$

<u>Proof</u>. From equations (4) and (5)

$$\sqrt{n}\ (R_n(\underset{\sim}{t})-R(\underset{\sim}{t})) = \sqrt{n}\ [\int_0^{t_1} (\bar{H}_n(u,0))^{-1}d_uK_{1n}(u,0) - \int_0^{t_1} (\bar{H}(u,0))^{-1}d_uK_1(u,0)$$

$$+ \int_0^{t_2} (\bar{H}_n(t_1,v))^{-1}d_vK_{2n}(t_1,v) - \int_0^{t_2} (\bar{H}(t_1,v))^{-1}d_vK_2(t_1,v)].$$

Now

$$\sqrt{n}\,[\int_0^{t_1}(\bar{H}_n(u,0))^{-1}d_uK_{1n}(u,0) - \int_0^{t_1}(\bar{H}(u,0))^{-1}d_uK_1(u,0)]$$

$$= \sqrt{n}\left\{\int_0^{t_1}(\bar{H}(u,0))^{-2}[\bar{H}(u,0) - \bar{H}_n(u,0)]d_uK_1(u,0)\right.$$

$$+ \int_0^{t_1}(\bar{H}(u,0))^{-1}\, d_u(K_{1n}(u,0) - K_1(u,0))$$

$$+ \int_0^{t_1}[\bar{H}(u,0) - \bar{H}_n(u,0)][\frac{1}{\bar{H}(u,0)\,\bar{H}_n(u,0)} - \frac{1}{\bar{H}^2(u,0)}]d_uK_1(u,0)$$

$$\left.+ \int_0^{t_1}[\frac{1}{\bar{H}_n(u,0)} - \frac{1}{\bar{H}(u,0)}]d_u(K_{1n}(u,0) - K_1(u,0))\right\}$$

$$= A_{1n}(\underset{\sim}{t}) + B_{1n}(\underset{\sim}{t}) + E_{1n}(\underset{\sim}{t}) + E^*_{1n}(\underset{\sim}{t})$$

where

$$A_{1n}(\underset{\sim}{t}) = \int_0^{t_1}(\bar{H}(u,0))^{-2}U_{1n}(u,0)d_uK_1(u,0)$$

$$B_{1n}(\underset{\sim}{t}) = (\bar{H}(t_1,0))^{-1}V_{1n}(t_1,0) - \int_0^{t_1}(\bar{H}(u,0))^{-2}V_{1n}(u,0)d_uH(u,\bar{0})$$

$$E_{1n}(\underset{\sim}{t}) = \frac{1}{\sqrt{n}}\int_0^{t_1}\frac{U^2_{1n}(t_1,0)}{\bar{H}_n(u,0)\bar{H}^2(u,0)}\, d_uK_1(u,0)$$

$$E^*_{1n}(t) = \int_0^{t_1}\frac{U_{1n}(u,0)}{\bar{H}_n(u,0)\bar{H}(u,0)}\, d_u(K_{1n}(u,0) - K_1(u,0)).$$

In a similar manner,

$$\sqrt{n}\left[\int_0^{t_2}(\bar{H}_n(t_1,v))^{-1}d_vK_{2n}(t_1,v) - \int_0^{t_2}(\bar{H}(t_1,v))^{-1}d_vK_2(t_1,v)\right]$$

$$= A_{2n}(\underset{\sim}{t}) + B_{2n}(\underset{\sim}{t}) + E_{2n}(\underset{\sim}{t}) + E^*_{2n}(\underset{\sim}{t}),$$

where

$$A_{2n}(\underset{\sim}{t}) = \int_0^{t_2}(\bar{H}(t_1,v))^{-2}\, U_{2n}(t_1,v)d_vK_2(t_1,v)$$

$$B_{2n}(\underset{\sim}{t}) = (\bar{H}(t_1,t_2))^{-1}V_{2n}(t_1,t_2) - \int_0^{t_2}(\bar{H}(t_1,v))^{-2}V_{2n}(t_1,v)dH(\bar{t}_1,v)$$

$$E_{2n}(\underset{\sim}{t}) = \frac{1}{\sqrt{n}}\int_0^{t_2}\frac{U^2_{2n}(t_1,v)}{\bar{H}_n(t_1,v)\bar{H}^2(t_1,v)}d_vK_2(t_1,v)$$

$$E^*_{2n}(\underset{\sim}{t}) = \int_0^{t_2}\frac{U_{2n}(t_1,v)}{\bar{H}_n(t_1,v)\bar{H}(t_1,v)}d_v(K_{2n}(t_1,v) - K_2(t_1,v)).$$

Now as n tends to infinity, $E_{jn}(\underset{\sim}{t})$ and $E^*_{jn}(\underset{\sim}{t})$ converge in probability to zero in the supremum metric by an argument similar to that of Breslow and Crowley (1974) and hence converge in probability to zero in the metric d, for j = 1,2. Further, $A_{jn}(\underset{\sim}{t})$ converges almost surely to $A_j(\underset{\sim}{t})$ and $B_{jn}(\underset{\sim}{t})$ to $B_j(\underset{\sim}{t})$ in the sup metric and hence in d, for j = 1, 2. Therefore, $\sqrt{n}(R_n(\underset{\sim}{t}) - R(\underset{\sim}{t}))$ converges weakly to $W(\underset{\sim}{t})$. That $W(\underset{\sim}{t})$ is a Gaussian process with mean 0 follows immediately from the Lemma. The covariance structure of $W(\underset{\sim}{t})$ can be calculated from (8) and from the covariance structure of (U_1, U_2, V_1, V_2).

THEOREM 2. If F and G are continuous bivariate distribution functions and if T_1, $T_2 < \infty$ are such that $\bar{H}(T_1, T_2) < \infty$, then for $0 < t_1 < T_1$, $0 < t_2 < T_2$

$$\sqrt{n}(\bar{F}_n(\underset{\sim}{t}) - \bar{F}(\underset{\sim}{t}))$$

converges weakly to a two-dimensional-time Gaussian process $W^*(\underset{\sim}{t})$ which has mean 0 and covariance

$$\mathrm{Cov}(W^*(s_1,s_2), W^*(t_1,t_2)) = F(s_1,s_2)F(t_1,t_2)\mathrm{Cov}(W(s_1,s_2), W(t_1,t_2)).$$

<u>Proof</u>. By (6)

$$\sqrt{n}(\bar{F}_n(\underset{\sim}{t}) - \bar{F}(\underset{\sim}{t})) = \sqrt{n}(e^{-R_n(\underset{\sim}{t})} - e^{-R(\underset{\sim}{t})}).$$

A Taylor's series expansion yields

$$\sqrt{n}(\bar{F}_n(\underset{\sim}{t}) - \bar{F}(\underset{\sim}{t})) = -e^{-R(\underset{\sim}{t})}\sqrt{n}(R_n(\underset{\sim}{t}) - R(\underset{\sim}{t})) - e^{-R_n^*(\underset{\sim}{t})}\sqrt{n}(R_n(\underset{\sim}{t}) - R(\underset{\sim}{t}))^2 \quad (9)$$

where $\sup |R_n^* - R| \leq \sup |R_n - R|$. The second term on the right of (9) converges to zero in probability in the sup norm and hence in d. Thus, $\sqrt{n}(\bar{F}_n(\underset{\sim}{t}) - \bar{F}(\underset{\sim}{t}))$ converges weakly to $-\bar{F}(\underset{\sim}{t})W(\underset{\sim}{t})$ which is a Gaussian process with mean 0 and desired covariance.

5. OTHER ESTIMATORS

Campbell and Földes (1982) introduced another path-dependent estimator. For $N_n(t) = n\,\bar{H}_n(t)$, define the estimator $S_n(\underset{\sim}{t})$ of $\bar{F}(\underset{\sim}{t})$:

$$S_n(t) = \prod_{i=1}^{n}\left(\frac{N(Z_{1i},0)}{N(Z_{1i},0)+1}\right)^{\alpha_{1i}(t_1,0)} \prod_{i=1}^{n}\left(\frac{N(t_1, Z_{2i})}{N(t_1, Z_{2i})+1}\right)^{\alpha_{2i}(t_1,t_2)},$$

provided $N(\underset{\sim}{t}) > 0$. This is the product of two one-dimensional product-limit estimators, one a marginal estimator for the first coordinate, and the second a conditional one on the second coordinate given $Z_1 > t_1$. It has been proved that for $T_1, T_2 < \infty$ such that $\bar{H}(T_1, T_2) > 0$

$$\sup_{\substack{0<t_1\leq T_1 \\ 0<t_2\leq T_2}} |\bar{F}_n(\underset{\sim}{t}) - S_n(\underset{\sim}{t})| = \mathcal{O}(\tfrac{1}{n}) \quad \text{a.s.}$$

Therefore $S_n(t)$ inherits strong uniform almost sure consistency as well as weak convergence from $\bar{F}_n(\underset{\sim}{t})$.

The estimators $\bar{F}_n(\underset{\sim}{t})$ and $S_n(\underset{\sim}{t})$ depend on the path from (0,0) to $(t_1,0)$ to (t_1,t_2). Since $R(\underset{\sim}{t})$ and $F(\underset{\sim}{t})$ are path independent, it is possible to have also developed estimators for the linear path from (0,0) to $(0,t_2)$ to (t_1,t_2). In general these estimators differ from $\bar{F}_n(\underset{\sim}{t})$ and $S_n(\underset{\sim}{t})$. However, the strong consistency and weak convergence results follow in the same way. The covariance structure of the limiting Gaussian process does depend on the path.

The extension of these estimators and their asymptotic properties from two-dimensional time to k(>2)-dimensional time is straightforward. In k-dimensions there are k! piecewise linear paths similar to the two mentioned above. Strong uniform almost consistency follows readily. It is convenient to use the definition of D_k in Neuhaus (1971) to develop the weak convergence results.

The drawback of the estimators $\bar{F}_n$ and S_n is that the estimators are not necessarily survival distribution functions. Although monotonicity is assured along the path of definition, it is not guaranteed along other ever-increasing paths nor is it the case that non-negative mass is assigned to rectangles. The fact that the estimator is uniformly strongly consistent for the properly behaved function $\bar{F}$ minimizes this problem for large samples.

In that S_n can be thought of as a generalized maximum likelihood estimator along the designated path, one could obtain such a generalized maximum likelihood estimator subject to the constraint that the estimator is a distribution function. It is conjectured that these asymptotic results will not be changed for such an estimator.

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*, 437-53.

Campbell, G. (1981). Nonparametric bivariate estimation with randomly censored data. *Biometrika* *68*, 417-22.

Campbell, G. and Földes, A. (1982). Large-sample properties of nonparametric bivariate estimators with censored data. *Proceedings of the International Colloquium on Nonparametric Statistical Inference*, 1980. North Holland (to appear).

Durbin, J. (1970). Asymptotic distributions of some statistics based on the bivariate sample distribution function. In *Nonparametric Techniques in Statistical Inference*, Ed. M.L. Puri, pp. 435-49. Cambridge University Press.

Földes, A. and Rejtö, L. (1981). Strong uniform consistency for nonparametric survival curve estimators from randomly censored data. *Ann. Statist* *9*, 122-29.

Kaplan, E.L. and Meier, P. (1958). Nonparametric estimation from incomplete observations. *J. Am. Statist. Assoc.* *53*, 457-81.

Leurgans, S., Tsai, W., and Crowley, J. (1982). Freund's bivariate exponential distribution and censoring. *Proceedings of the Institute of Mathematical Statistics, Special Topics Meeting on Survival Analysis*, 1981.

Marshall, A.W. (1975). Some comments on the hazard gradient. *Stoch. Proc. and their Applic.* *3*, 293-300.

Neuhaus, Georg (1971). On weak convergence of stochastic processes with multidimensional time parameter. *Ann. Math. Statist* *42*, 1285-95.

Straf, M.L. (1972). Weak convergence of stochastic processes with several parameters. *Proceedings of the Sixth Berk. Symp. on Math. Statist and Probab.* *II*. pp. 187-221.