The Annals of Statistics
1981, Vol. 9, No. 1, 157–167

ESTIMATING DEPENDENT LIFE LENGTHS, WITH APPLICATIONS TO THE THEORY OF COMPETING RISKS

BY N. LANGBERG[1], F. PROSCHAN[1] AND A. J. QUINZI[2]

Florida State University and Temple University

In the classical theory of competing risks (as well as in many reliability models and incomplete data problems) it is assumed that (A1) the risks (i.e., the random variables of interest) are independent and that (A2) death does not result from simultaneous causes. Employing our probabilistic solution to a related problem in probability modelling, we obtain strongly consistent estimators for the unobservable marginal distributions of interest. These estimators are analogous to those of Kaplan and Meier [*J. Amer. Statist. Assoc.* (1958) *63*] and are appropriate when the assumptions of independence and no simultaneous causes of death [(A1) and (A2), above] fail to hold. We show how our methods can be used to unify and simplify the nonparametric approach toward estimation in the competing risks model. As a consequence we obtain an elementary proof of the strong consistency of the Kaplan-Meier estimator. Our results extend and simplify the work of Peterson [*J. Amer. Statist. Assoc.* (1977) **72**] and Desu and Narula [*The Theory and Applications of Reliability, I* (I. Shimi and C. P. Tsokos, eds.) (1977)].

Introduction and summary. Langberg, Proschan, and Quinzi (1978) [hereafter referred to as LPQ (1978)] show that under certain mild conditions it is possible to establish a particular equivalence between an arbitrary system of *dependent* components and a system of independent components. The two systems are equivalent in the sense that they possess (1) the same distribution for the time to system failure and (2) the same probabilities of occurrence of each failure pattern. In the case of a series system, (more generally, any coherent system), a particular failure pattern occurs when the failure of that set of components coincides with the failure of the system. The LPQ (1978) result, as well as the methods developed in this paper, apply in at least three contexts of interest—(1) engineering or reliability models, (2) competing risks models and (3) various models involving incomplete data.

In the classical theory of competing risks it is assumed that the causes act independently and that death does not result from simultaneous causes. In Section 3 we describe the competing risks model. We examine the classical assumptions and introduce the Kaplan-Meier (1958) (K-M) estimator. Employing the LPQ (1978) result, we obtain in Section 4 strongly consistent estimators for the unobservable marginal distributions of interest in the competing risks model. These estimators are analogous to those of Kaplan and Meier (1958) and are appropriate when the assumptions of independence and no simultaneous causes of death fail to hold. We show how our methods can be used to unify and simplify the nonparametric approach toward estimation in the competing risks model. As a consequence we obtain an elementary proof of the strong consistency of the Kaplan-Meier estimator. Our results extend and simplify the work of Desu and Narula (1977) and Peterson (1975, 1977). Section 5 consists of proofs.

2. From a dependent model to an independent one. In this section we state for future reference, the equivalence result of LPQ (1978).

Received February 1978; revised August 1979.

[1] Research supported by the Air Force Office of Scientific Research under AFOSR Grants 74-2581C and 78-3678.

[2] Research supported by the Air Force Office of Scientific Research under AFOSR Grant 78-3673.

AMS 1970 *subject classification*. 62N05, 62G05.

Key words and phrases. Competing risks, Kaplan-Meier estimator, reliability, life testing, series system, estimation.

Let $\mathscr{I}$ denote the collection of nonempty subsets of $\{1, \cdots, r\}$. Let $\mathbf{T}$ represent the vector of component life lengths of an r-component system ($r \geq 2$) with life length τ. Let $\mathbf{H}$ be the vector of component life lengths in a series system of $(2^r - 1)$ components with life length H, where the coordinates of $\mathbf{H}$ are indexed lexicographically by $I \in \mathscr{I}$. For $I \in \mathscr{I}$, *failure pattern I occurs* in the system with life length τ (alternatively, H) if $\xi(\mathbf{T})$ [alternatively, $\xi^*(\mathbf{H})$] is I, where

$$\xi(\mathbf{T}) = \begin{cases} I, & \text{if } \tau = T_i \text{ for each } i \in I \text{ and } \tau \neq T_i \text{ for each } i \notin I \\ \emptyset, & \text{otherwise;} \end{cases}$$

and

$$\xi^*(\mathbf{H}) = \begin{cases} I, & \text{if } H_I < H_J \text{ for each } J \neq I \\ \emptyset, & \text{otherwise.} \end{cases}$$

The two systems are *equivalent in life length and failure patterns* [$\mathbf{H} =_{\mathrm{LP}} \mathbf{T}$] if

$$P(H > t, \xi^*(\mathbf{H}) = I) = P(\tau > t, \xi(\mathbf{T}) = I) \tag{2.1}$$

for each $t \geq 0$ and every $I \in \mathscr{I}$. In general, for a life length S with distribution function K let $\bar{K}(s) = P(S > s)$ denote the corresponding survival probability and let $\alpha(K) = \sup\{s\colon \bar{K}(s) > 0\}$. LPQ (1978) prove the following:

THEOREM 2.1. *Let $\tau = \min(T_i, 1 \leq i \leq r)$ denote the life length of an r-component series system, where T_i represents the life length of component i, $i = 1, \cdots, r$. Define $\bar{F}(t, I) = P(\tau > t, \xi(\mathbf{T}) = I)$ $(P(\tau \leq t, \xi(\mathbf{T}) = I)$, $I \in \mathscr{I}$, $F(t) = 1 - \sum_{I \in \mathscr{I}} F(t, I)$, and $\bar{F}(t) = 1 - F(t)$.*

Then the following statements hold:

(i) *A necessary and sufficient condition for the existence of a set of independent random variables $\{H_I, I \in \mathscr{I}\}$ which satisfy $\mathbf{H} =_{\mathrm{LP}} \mathbf{T}$, where $H = \min(H_I, I \in \mathscr{I})$, is that the functions $F(\cdot, I)$, $I \in \mathscr{I}$, have no common discontinuities in the interval $[0, \alpha(F))$.*

(ii) *The random variables $\{H_I, I \in \mathscr{I}\}$ in (i) have corresponding survival probabilities $\{\bar{G}_I(\cdot), I \in \mathscr{I}\}$ which are uniquely determined on the interval $[0, \alpha(F)]$ as follows:*

$$\bar{G}_I(t) = \prod\nolimits_{a \leq t} [\bar{F}(a)/\bar{F}(a^-)] \exp\left[- \int_0^t (dF^C(\cdot, I)/\bar{F}) \right], \qquad 0 \leq t \leq a(F), \tag{2.2}$$

where $F^C(\cdot, I)$ is the continuous part of $F(\cdot, I)$, the product is over the set of discontinuities $\{a\}$ of $F(\cdot, I)$, $I \in \mathscr{I}$, and the product over an empty set is defined as unity.

REMARK 2.2. Formula (2.2) is defined in LPQ (1978) for t in the half-open interval $[0, \alpha(F))$. The present formulation is, however, equivalent.

REMARK 2.3. LPQ (1978) show that if the original system is such that $P(T_i = T_j) = 0$ for $i \neq j$ and $P(\tau = T_i) > 0$, $1 \leq i \leq r$, then the original system is equivalent in life length and failure patterns to a system involving the same number r of independent random variables $\{H_i, 1 \leq i \leq r\}$.

REMARK 2.4. Suppose that the original vector $\mathbf{T}$ in Theorem 2.1 is *itself* a vector of independent random variables and the functions $F(\cdot, I)$, $I \in \mathscr{I}$, have no common discontinuities. Then it is not difficult to show that $\mathbf{T}$ itself is a unique solution to the equation $\mathbf{H} =_{\mathrm{LP}} \mathbf{T}$.

3. The independent competing risks model: The Kaplan-Meier estimator. Let there be a finite number of causes of death labelled $1, \cdots, r$. We associate with each cause j a nonnegative random variable T_j, $j = 1, \cdots, r$. The random variable T_j represents the age at death if cause j were the *only* cause present in the environment. In a reliability setting T_j denotes the life length of component j in a series system of r components. In an incomplete or censored data problem, one of the random variables T_j represents the time at which an individual becomes "unobservable" for a reason other than death, while the remaining variables typically represent various causes of death. The complete collection of random variables $T_1, \cdots, T_r$ is not observed. Instead, only two quantities are observed: the *age at death* given by $\tau = \min(T_1, \cdots, T_r)$ and *the cause of death*, labeled ξ, given by $I \in \mathscr{I}$ such that $\xi(T) = I$. When death results from exactly one of the r possible causes, as is usually assumed, then ξ is the index i for which $\tau = T_i$. The biomedical researcher is interested in making inferences about *unobservable* quantities (viz., the random variables $T_1, \cdots, T_r$) by using data from *observable* quantities—in this case, the lifetime τ and the cause of death ξ. In particular, he seeks to estimate the *marginal* survival probability corresponding to a given cause (or combination of causes) operating alone without competition from the other causes. That is, he wishes to estimate the $2^r - 1$ survival probabilities

$$\bar{M}_J(t) = P[\min(T_j, j \in J) > t],\ J \in \mathscr{I}.$$

In analyzing competing risk data, various authors typically assume one or more of the following:

(A1) $T_1, \cdots, T_r$ are mutually independent.

(A2) Death does not result from simultaneous causes. [Consequently, $P(T_i = T_j) = 0$ for $i \neq j$.]

(A3) The distributions of $T_1, \cdots, T_r$ have no common discontinuities.

(A4) The joint distribution of $T_1, \cdots, T_r$ is absolutely continuous.

For many years there have been several approaches to problems of estimation in competing risk theory which employ, in varying degrees, the above assumptions. For example, the assumption of independence (A1) was until recently almost universally made even though it is obviously inappropriate in many problems. Moreover, assumptions (A2) through (A4) need not hold in certain situations of interest. For example, in engineering systems where system failure can occur as a result of the simultaneous failure of two (or more) components, assumptions (A2) and (A4) do not hold. Assuming (A1), (A2) and (A3), Peterson (1975, 1977) shows how the Kaplan-Meier estimator may be expressed as a function of the empirical counterparts of the functions $F(\cdot, I)$, $I \in \mathscr{I}$, in Theorem 2.1. He thus indicates a way to obtain strong consistency of the estimator when $\mathbf{T}$ is a vector of random variables. In this paper we show how Theorem 2.1 can be used to estimate the marginal distributions of interest in a unified way without making any of the assumptions (A2), (A3), or (A4). Moreover, we are able to drop the assumption of independence (A1) and find necessary and sufficient conditions for the existence of consistent estimators for the marginal distributions of interest. Peterson (1975) also considers the case of dependent risks. In Section 4 we show how our methods extend and simplify those of Peterson.

Let $T_i = (T_{1i}, \cdots, T_{ri})$, $i = 1, \cdots, n$, represent a random sample from the joint distribution of the nonnegative random variables $T_1, \cdots, T_r$. Denote the marginal distributions (survival probabilities) of $T_1, \cdots, T_r$ by $M_i(\bar{M}_i)$, $i = 1, \cdots, r$. For each $I \in \mathscr{I}$, let $M_I(t) = P(\tau_I \leq t)$, where $\tau_I = \min(T_i, i \in I)$. Assume (A1), (A2), and (A3). Then the cause of death $\xi(T) = i$ if and only if $T_i < T_j$ for each $j \neq i$, $1 \leq i, j \leq r$. For each $i = 1, \cdots, n$ only τ_i and ξ_i are observed, where $\tau_i = \min(T_{1i}, \cdots, T_{ri})$ and $\xi_i = j$ whenever $\tau_i = T_{ji}$. Consider the case $r = 2$ and suppose we seek to estimate the marginal survival probability $\bar{M}_1(t) = P(T_1 > t)$. Let $0 \equiv \tau_{(0)} \leq \tau_{(1)} \leq \cdots \leq \tau_{(n)}$ denote the ordered values of the observations $\tau_1, \cdots, \tau_n$. Then the Kaplan-Meier product-limit estimator (hereafter referred to simply as the K-M estimator) of $\bar{M}_1(t)$ is

$$\hat{\bar{M}}_1(t) = \textstyle\prod_i [(n - i)/(n - i + 1)], \tag{3.1}$$

where the product is over the ranks i of those ordered observations $\tau_{(i)}$, $1 \leq i \leq n$, such that $\tau_{(i)} \leq t < \tau_{(n)}$ and $\tau_{(i)}$ corresponds to a death from cause 1 [$\tau_{(i)} = T_{1j}$ for some j]. If $\tau_{(n)}$ corresponds to a death from cause 1, then $\hat{\bar{M}}_1(t)$ is defined to be zero for $t > \tau_{(n)}$. Otherwise, $\hat{\bar{M}}_1(t)$ is undefined for $t > \tau_{(n)}$. [In the original formulation by Kaplan and Meier (1958), T_1 corresponded to the time until death, while T_2 represented the time at which a loss occurred.]

Peterson (1975) considers the following straightforward extension of (3.1) for the survival probabilities $\bar{M}_J(t)$, $J \in \mathscr{I}$. The (extended) K-M estimator $\hat{\bar{M}}_J$ is given by

$$\hat{\bar{M}}_J(t) = \textstyle\prod_i [(n - i)/(n - i + 1)], \tag{3.2}$$

where now the product is over the ranks i of those ordered observations $\tau_{(i)}$ such that $\tau_{(i)} \le t < \tau_{(n)}$ and $\tau_{(i)}$ corresponds to a death from *at least one* cause $j \in J$. Conventions analogous to those used in defining (3.1) also hold in (3.2) when $t > \tau_{(n)}$.

Assuming independence (A1), no simultaneous causes of death (A2), and disjoint sets of discontinuities for the marginal distributions (A3), Peterson (1977) indicates a way to obtain strong consistency of the K-M estimator (3.1). Breslow and Crowley (1974) and Meier (1975) show that the estimator is asymptotically normal and, as a process in t, converges to a normal process. More recently, Aalen (1976) shows that the bivariate vector of K-M estimators $(1 - \hat{\bar{M}}_1(t_1),\ 1 - \hat{\bar{M}}_2(t_2))$ is asymptotically bivariate normal, and that, regarded as a bivariate process in t_1 and t_2, converges to a normal process. Estimators analogous to (3.1) and (3.2) are proposed in the next section by using formula (2.2) of Theorem 2.1. Such extensions will apply in situations *where the assumptions of independence* (A1) *and in simultaneous causes of death* (A2) *fail to hold.*

4. The dependent case. In this section, unless otherwise indicated, we drop the assumption (A1) of independent risks. By replacing $\bar{F}(t)$ and $\bar{F}(t, i)$ in (2.2) with their empirical counterparts, we can *estimate* the distributions G_I, associated with the unobservable variables H_I, $I \in \mathscr{I}$. However, the distributions G_I *differ*, in general, from the marginal distributions M_I which we seek to estimate. But suppose for a moment that $T_1, \cdots, T_r$ are, in fact, *independent* and that

(A3)′ The functions $F(\cdot, I)$, $I \in \mathscr{I}$, in Theorem 2.1 have no common discontinuities. In this case a simple relationship holds between the functions M_I and the survival functions $\bar{G}_I$ in (2.2). For by Remark 2.4, $G_i = M_i$, $i = 1, \cdots, r$. Consequently,

$$\bar{M}_I(t) = \textstyle\prod_{i \in I} \bar{M}_i(t) = \prod_{i \in I} \bar{G}_i(t) \tag{4.1}$$

for every $t \in [0, \alpha(F)]$. In the case $r = 2$, if we replace the functions $\bar{F}(t)$ and $\bar{F}(t, 1)$ on the right in (2.2) by their empirical counterparts, then the resulting statistic is the K-M estimator (3.1) of $P(T_1 > t)$. In view of (4.1), if r is an arbitrary integer greater than 2, a reasonable estimator for $\bar{M}_I(t)$ ought to be $\prod \hat{\bar{G}}_i(t)$, where the product is over $i \in I$ and $\hat{\bar{G}}_i$ is the function resulting from (2.2) by replacing the functions $\bar{F}$ and $\bar{F}(\cdot, i)$ with their empirical counterparts, $i = 1, \cdots, r$. Again, it is easy to show that the resulting statistic is, in fact, the (generalized) K-M estimator (3.2) for $\bar{M}_I(t)$. Thus, in the case of independent risks, (2.2) leads directly to well-known estimators possessing several optimal properties. It is therefore reasonable to expect that (2.2) also plays a role in the estimation problem when the risks are mutually *dependent*. This is, indeed the case.

THEOREM 4.1. [Peterson (1975)]. *Let* $T_1, \cdots, T_r$ *be nonnegative random variables satisfying* (A2) *and* (A3) [*but not necessarily* (A1)]. *Let* $\mathscr{P}$ *be a partition of* $\{1, \cdots, r\}$ *and define* $\bar{G}_i$, $1 \le i \le r$, *as in* (2.2). *Then for each* $t \in [0, \alpha(F)]$,

$$\bar{M}_I(t) = \textstyle\prod_{i \in I} \bar{G}_i(t) \qquad \textit{for each } I \in \mathscr{P} \tag{4.2a}$$

if and only if

$$P(\tau > t, \xi(T) \subseteq I) = \int_t^\infty \bar{M}_{I'}(x)\, dM_I(x) \qquad \textit{for each } I \in \mathscr{P}, \tag{4.2b}$$

where I' *denotes the complement of* $I \in \mathscr{I}$.

Peterson (1975) uses an operator defined on a space of distribution functions to prove an equivalent version of Theorem 4.1. Employing Theorem 2.1, we give a proof which is considerably simpler. The following notation and lemma are useful in interpreting (4.2a, b).

Let $\{T_i, 1 \le i \le r\}$ and $\{T_i^*, 1 \le i \le r\}$ be two collections of random variables. For each function f of the random variables $T_1, \cdots, T_r$, let f^* denote the value of the same function of $T_1^*, \cdots, T_r^*$. For each set I belonging to a partition $\mathscr{P}$ of $\{1, \cdots, r\}$, define $\bar{F}_{\mathscr{P}}(t, I) = P(\tau > t, \xi(T) \in I)$. On the right in (2.2), replace $F(t, I)$ by $F_{\mathscr{P}}(t, I)$ and call the resulting expression $\bar{G}_{I,\mathscr{P}}(t)$.

The following simple lemma is useful in proving Theorem 4.1:

LEMMA 4.2. *Assume the hypothesis of Theorem* 4.1. *Let* $T_1^*, \cdots, T_r^*$ *be independent random variables such that* τ_I^* and τ_I *have the same distribution* (i.e., $M_I^* = M_I$) *for each* $I \in \mathscr{P}$. *Then* (4.2a) *is equivalent to*

(4.3a) $$\bar{G}^*_{I,\mathscr{P}} = \bar{G}_{I,\mathscr{P}} \quad \textit{for each} \quad I \in \mathscr{P},$$

and (4.2b) *is equivalent to*

(4.3b) $$\bar{F}_{\mathscr{P}}(t, I) = \bar{F}^*_{\mathscr{P}}(t, I) \qquad \text{for each} \quad I \in \mathscr{P}.$$

PROOF. By (A2) and (A3), we have that M_I and M_J have no common discontinuities and $P(\tau_I = \tau_J) = 0$ for each $I, J \in \mathscr{P}, I \neq J$. Since $M_I = M_I^*$, we deduce that $P(\tau_I^* = \tau_J^*) = 0$ for each $I, J \in \mathscr{P}, I \neq J$. It follows from Remark 2.3 that if $\mathscr{P}$ has exactly k members, $2 \leq k \leq r$, then each of the collections $\{\tau_I, I \in \mathscr{P}\}$ and $\{\tau_I^*, I \in \mathscr{P}\}$ has at most k occurring failure patterns. By Remark 2.4,

(4.4) $$\bar{M}_I^* = \bar{G}_{I,\mathscr{P}}(= \bar{M}_I).$$

By (2.2) and the fact that $\bar{F}_{\mathscr{P}}(t, I) = \sum_{i \in I} \bar{F}(t, i)$,

(4.5) $$\bar{G}_{I,\mathscr{P}} = \prod_{\iota \in I} \bar{G}_\iota.$$

A simple calculation shows that

(4.6) $$\bar{F}^*_{\mathscr{P}}(t, I) = \int_t^\infty \bar{M}_I^*(x)\, dM_I(x).$$

The conclusion follows from (4.4), (4.5), and (4.6). □

We now give the following elementary proof of Theorem 4.1.

PROOF OF THEOREM 4.1. By Lemma 4.2, it is enough to show that (4.3a) holds if and only if (4.3b) holds.

Suppose first that (4.3b) holds. Since

$$\bar{F}(t) = \textstyle\sum_{I \in \mathscr{P}} \bar{F}_{\mathscr{P}}(t, I) \qquad \text{and} \qquad \bar{F}^*(t) = \sum_{I \in \mathscr{P}} \bar{F}^*(t, I),$$

it follows from (2.2) that $\bar{G}_{I,\mathscr{P}} = \bar{G}^*_{I,\mathscr{P}}$.

Conversely, suppose that (4.3a) holds. It follows from (2.2) that $\bar{F} = \prod \bar{G}_{I,\mathscr{P}}$ and $\bar{F}^* = \prod \bar{G}^*_{I,\mathscr{P}}$, where each product is over $I \in \mathscr{P}$. Since

$$\bar{F}_{\mathscr{P}}(t, I) = \int_t^\infty (\bar{F}/\bar{G}_{I,\mathscr{P}})\, dG_{I,\mathscr{P}}$$

and

$$\bar{F}^*_{\mathscr{P}}(t, I) = \int_t^\infty (\bar{F}^*/\bar{G}^*_{I,\mathscr{P}})\, dG^*_{I,\mathscr{P}}$$

for each $I \in \mathscr{P}$, relation (4.3b) holds. □

One drawback of Peterson's formulation is that the assumption of no simultaneous causes of death (A2) does not hold, e.g., in an engineering system where system failure can occur as a result of the simultaneous failure of two or more components. In Theorem 4.4 below we give necessary and sufficient conditions for a functional relation to exist between the functions $\bar{G}_I$ in (2.2) and the functions M_I which we seek to estimate *without assuming* (A2). Our only assumption is that the functions $F(\cdot, I), I \in \mathscr{I}$, have no common discontinuities in $[0, \alpha(F)]$ [Assumption (A3)′]. Assumption (A3) implies (A3)′, but the converse does not hold.

Another disadvantage in Peterson's (1975) approach is that in order to establish a relation between the function M_J for an *individual* subset J of $\{1, \cdots, r\}$ and the functions $\bar{G}_I$, $I \in$

$\mathscr{I}$, in (2.2), he requires that (4.2b) hold simultaneously for each set I in a partition $\mathscr{P}$ of $\{1, \cdots, r\}$ which contains J. In general, it is not difficult to construct joint distributions which satisfy *none* of the conditions in (4.2b) yet for which *at least one* relationship exists between the function M_J and the survival probabilities $\bar{G}_I$, $I \in \mathscr{I}$.

EXAMPLE 4.3. Let the discrete random vector (T_1, T_2) have a joint probability distribution as given in Table 4.1. A simple calculation verifies that $\bar{G}_1$ given by (2.2) equals $\bar{M}_1$, but $\bar{G}_2 \neq \bar{M}_2$. Furthermore, *neither* of the conditions (4.2b) is satisfied. Thus, Theorem 4.1 above is not applicable in this simple example. We will show, however, that our generalization of Theorem 4.1 (Theorem 4.4 below) does apply here. Before we state Theorem 4.4, we introduce the following notation. Let $\mathscr{I}_I = \{J \in \mathscr{I}: J \cap I \neq \phi\}$ and $F(t, \mathscr{I}_I) = P(\tau > t, \xi(T) \in \mathscr{I}_I)$. Thus, $F(t, \mathscr{I}_I) = \sum P(\tau > t, \xi(\mathbf{T}) = J) = \sum F(t, J)$, where each sum is over $J \in \mathscr{I}_I$. For each function G, let $D(G)$ $[C(G)]$ denote the set of discontinuities (continuities) of G. For simplicity of notation let $D[F(\cdot, I)] \equiv D(I)$ and let $C[F(\cdot, I)] \equiv C(I)$ for $I \in \mathscr{I}$.

TABLE 4.1
Distribution of (T_1, T_2)

$T_1 \backslash T_2$	2	4	6
1	1/24	1/8	1/12
3	1/12	1/16	0
5	1/6	1/6	1/6

Theorem 4.4 below resembles Theorem 4.1 in that we find necessary and sufficient conditions for a relationship to exist between the functions M_I, $I \in \mathscr{I}$, and the survival probabilities $\bar{G}_I$, $I \in \mathscr{I}$, given by (2.2). It *generalizes* Theorem 4.1 in the following ways. First, the assumption (A2) of no simultaneous causes of death is dropped. Secondly, we assume a weaker version of assumption (A3), namely the assumption (A3)′ that the functions $F(\cdot, I)$, $I \in \mathscr{I}$, have no common discontinuities.

THEOREM 4.4. *Let $T_1, \cdots, T_r$ be nonnegative random variables satisfying* (A3)′. *Let $I \in \mathscr{I}$. Then for each $t \in [0, \alpha(F)]$,*

$$\bar{M}_I(t) = \textstyle\prod_{J \in \mathscr{I}_I} \bar{G}_J(t) \tag{4.7}$$

if and only if the following two conditions hold:

$$M_I(a)/M_I(a^-) = \bar{F}(a)/\bar{F}(a^-), \qquad a \in D(F(\cdot, \mathscr{I}_I)) \tag{4.8a}$$
$$= 1, \qquad \textit{otherwise}$$

and

$$P(\tau_{I'} \geq t \mid \tau_I = t) = P(\tau_{I'} > t \mid \tau_I > t), \tag{4.8b}$$

where $\bar{G}_J$ is given by (2.2).

The proof of Theorem 4.4 is given in Section 5.

REMARK 4.5. Suppose that the random variables $\tau_I = \min(T_i, i \in I)$, $I \in \mathscr{I}$, have absolutely continuous distributions. Let $m_1(t)[M_I(t)]$ and $m_{I|I}'(t)[M_{I|I'}(t)]$ denote respectively the density (distribution) function and conditional density (distribution) function of τ_I and of τ_I given $\tau_{I'} > t$. Then condition (4.8b) is equivalent to

$$m_{I|I'}(t)/M_{I|I'}(t) = m_I(t)/M_I(t).$$

In other words, the conditional failure rate function of τ_I given $\tau_{I'} > t$ is equal to the (unconditional) failure rate function of τ_I. Stated differently, the random variables τ_I and $\tau_{I'}$ are independent "along the diagonal $\tau_I = \tau_{I'}$". Desu and Narula (1977) arrive at a condition similar to (4.8b) in the special case when the assumption of absolute continuity (A4) [and hence also (A2)] holds.

Note that conditions (4.8a, b), in contrast to (4.2b), apply to only *one* subset $I \in \mathscr{I}$ at a time. Consequently, we can proceed in Example 4.3 as follows. It is easy to verify that conditions (4.8a, b) hold. Since $\bar{G}_{\{1,2\}} \equiv 1$, it follows from Theorem 4.4 that $\bar{M}_1 = \bar{G}_1$.

We have previously remarked that the assumption (A2) of no simultaneous causes of death is unrealistic in certain models of interest. An important family of multivariate distributions for which assumption (A2) fails is the family of multivariate exponential (MVE) distributions of Marshall and Olkin (1967). We illustrate with an example.

EXAMPLE 4.6. For simplicity, suppose that the random vector (T_1, T_2) has the Marshall-Olkin bivariate exponential distribution with survival probability:

$$P(T_1 > t_1, T_2 > t_2) = \exp[-\lambda_1 t_1 - \lambda_2 t_2 - \lambda_{12} \max(t_1, t_2)],$$

for $t_1, t_2 \geq 0$ and $\lambda_1, \lambda_2, \lambda_{12} > 0$. Since the marginal distributions M_1 and M_2 are continuous, condition (4.8a) trivially holds. Condition (4.8b) with $I = \{1\}$ states that

$$P(T_2 \geq t \mid T_1 = t) = P(T_2 > t \mid T_1 > t).$$

An easy computation shows that these conditional probabilities are each equal to $\exp(-\lambda_2 t)$. Thus, Theorem 4.4 may be applied when the joint distribution belongs to the family of Marshall-Olkin MVE distributions, whereas Theorem 4.1 cannot be applied here since (A2) fails.

In Section 3 we assumed that the risks were mutually independent (A1) and that the functions $F(\cdot, I)$, $I \in \mathscr{I}$, had no common discontinuities (A3)′. Under these assumptions the basic formula (2.2) yielded the K-M estimators for the marginal distribution M_1 ($r = 2$) and the functions M_I, $I \in \mathscr{I}$, ($r \geq 2$). In a similar fashion formula (2.2) (via Theorem 4.4) can be used to determine strongly consistent estimators for the functions M_I, $I \in \mathscr{I}$, in the important practical cases when independence fails to hold and simultaneous causes of death are allowed.

The key tool we shall use in establishing consistent estimators for the marginal distributions of interest is given in Theorem 4.7 below. First we introduce some notation. As above, $\mathbf{T} = (T_1, \cdots, T_r)$ is a vector of nonnegative random variables, $\tau = \min(T_j, 1 \leq j \leq r)$ represents the life length of an individual exposed to r risks of death and ξ represents the cause of death, where $\xi = J$ if and only if $\tau = T_j$ *for each* $j \in J$ and $\tau \neq T_j$ for each $j \notin J$, $J \in \mathscr{I}$. For each $I \in \mathscr{I}$ and Borel set A, $F(A, I) = P(\tau \in A, \xi = I)$ taking liberties with the notation for $F(\cdot, I)$. There exists, by Theorem 2.1, a collection $\{H_I, I \in \mathscr{I}\}$ of independent random variables such that $\mathbf{H} =_{\mathrm{LP}} \mathbf{T}$, where $H = \min(H_I, I \in \mathscr{I})$. Moreover, the probability $\bar{G}_I(t^+) = P(H_I \geq t)$ may be obtained from (2.2). Now let $\mathbf{T}_i = (T_{1i}, \cdots, T_{ri})$ $i = 1, 2, \cdots$, be a sequence of nonnegative random vectors (representing a sequence of individuals) and let τ_i, ξ_i, $F_i(\cdot, I)$, F_i, $D(I, i)$, $C(I, i)$, $H_{I,i}$, H_i, and $\bar{G}_{I,i}$ be the analogues of τ, ξ, $F(\cdot, I)$, F, $D(I)$, $C(I)$, H_I, H, and $\bar{G}_I$ above. Let $F^C(F^D)$ denote the continuous (discontinuous) part of F.

THEOREM 4.7. *Suppose the following conditions hold*:

(i) *For $I \neq J$, the pair $\{F(\cdot, I), F(\cdot, J)\}$ as well as each pair $\{F_i(\cdot, I), F_i(\cdot, J)\}$, $i = 1, 2, \cdots$, have no common discontinuities.*

(ii) *For $I \in \mathscr{I}$ and $0 \leq t \leq \alpha(F)$,*

$$\lim_{n\to\infty} F_n(\{[0, t] \cap C(I)\}, I) = F^C([0, t], I).$$

(iii) *For $I \in \mathscr{I}$,*

$$\lim_{n\to\infty} \sup_{0\leq t\leq \alpha(F)} | F_n(([0, t] \cap D(I)), I) - F^D([0, t], I) | = 0.$$

(iv)
$$\lim_{n\to\infty} \sup_{0\leq t\leq \alpha(F)} | \bar{F}_n(t) - \bar{F}(t) | = 0.$$

Then for $I \in \mathscr{I}$ and $0 \leq t \leq \alpha(F)$,

(4.9)
$$\lim_{n\to\infty} G_{I,n}(t^+) = G_I(t^+).$$

The proof of Theorem 4.7 is given in Section 5.

Consider now the estimation problem posed above. Let $\mathbf{T}_i = (T_{1i}, \cdots, T_{ri})$, $i = 1, \cdots, n$, be independent and identically distributed as $\mathbf{T} = (T_1, \cdots, T_r)$. Replace $F(\cdot, I)$ and F by their empirical counterparts, $\hat{F}_n(\cdot, I)$ and $\hat{F}_n$, respectively, on the right in (2.2), where

$$\hat{F}_n(t, I) = n^{-1} \textstyle\sum_{i=1}^{n} \chi\{\tau_i \le t, \xi_i = I\},$$

$$\hat{F}_n(t) = n^{-1} \textstyle\sum_{i=1}^{n} \chi\{\tau_i \le t\},$$

and $\chi(A)$ is the indicator function of the set A. The resulting expression, call it $\hat{\bar{G}}_{I,n}$, is an estimator for $\bar{G}_I$. Assume that $\mathbf{T}$ satisfies (4.8a, b). Then (4.7) holds. A natural estimator for $M_I(t)$ in this case, then, is

$$\hat{\bar{M}}_{I,n}(t) = \textstyle\prod_{J \in \mathscr{I}_I} \bar{G}_{J,n}(t). \tag{4.10}$$

Such an estimator will be strongly consistent if for every $J \in \mathscr{I}_I$,

$$\bar{G}_{J,n}(t^+) \to \bar{G}_J(t^+) \quad \text{a.s.} \tag{4.11}$$

To show that this is a simple consequence of Theorem 4.7, it suffices to verify conditions (i) through (iv). Condition (i) holds trivially. Condition (ii) follows from the Strong Law of Large Numbers. Condition (iv) is a consequence of the Glivenko-Cantelli theorem. Condition (iii) follows from (iv) and the continuity of $F^C([0, t], I)$.

REMARK 4.8. Let $0 \equiv \tau_{(0)} \le \tau_{(1)} \le \cdots \le \tau_{(n)}$ denote the ordered values of $\tau_1, \cdots, \tau_n$. In analogy with the Kaplan-Meier estimator (3.2), the estimator (4.10) may be expressed as follows:

$$\hat{\bar{M}}_I(t) = \textstyle\prod_i [(n-i)/(n-i+1)], \tag{4.12}$$

where the product is over the ranks i of those ordered observations $\tau_{(i)}$, $1 \le i \le n$, such that $\tau_{(i)} \le t < \tau_{(n)}$ and $\tau_{(i)}$ corresponds to a death from the simultaneous causes $j \in J$, $J \in \mathscr{I}_I$. If for some i, $\tau_{(n)} = T_{ji}$ for each $j \in J$, $J \in \mathscr{I}_I$, then (4.12) is defined to be zero for $t > \tau_{(n)}$. Otherwise, (4.12) is undefined for $t > \tau_{(n)}$.

In view of Remark 4.8 and the preceding argument, we have proven:

THEOREM 4.9. *In the competing risks model of Section* 3, *assume only that the functions* $F(\cdot, 1)$, $I \in \mathscr{I}$, *have no common discontinuities, and the joint distribution of* $(T_1, \cdots, T_r)$ *satisfies* (4.8a, b). *Then the estimator* (4.12) *is strongly consistent for* $\bar{M}_I$.

REMARK 4.10. To show consistency of the K-M estimator in the independent case, Peterson (1977) must rely on a property of an operator defined on a space of discrete distribution functions (which he states without proof). If we assume that the risks are independent and that assumption (A3)′ holds, then by applying Theorem 4.7 above, as we did in the dependent case, we have a proof of the consistency of the K-M estimator which is considerably more elementary.

5. Proofs. Before we give a proof of Theorem 4.4, we state two lemmas which are proven in LPQ (1978).

LEMMA 5.1. *For every probability measure* Q (*with a possible atom at* ∞) *such that* $\bar{Q}(0^-) = 1$, *and every* $t \ge 0$, *the following holds*:

$$\bar{Q}(t) = \exp\left[-\int_0^t (dQ^C/\bar{Q}) \cdot \textstyle\prod_{a \le t} [\bar{Q}'(a)/\bar{Q}(a^-)]\right], \tag{5.1}$$

where the product is over the set $\{a\}$ *of discontinuities of* Q, *and the product over an empty set is defined to be* 1.

LEMMA 5.2. *Let $\{\bar{G}_I, I \in \mathscr{I}\}$ be a collection of survival probabilities satisfying* (2.2). *Then for each $I \in \mathscr{I}$ and $t \in [0, \alpha(F))$,*

$$\begin{aligned} \bar{G}_I(t)/\bar{G}_I(t^-) &= \bar{F}(t)/\bar{F}(t^-), \qquad t \in D(F(\cdot, I)) \\ &= 1 \qquad \textit{otherwise.} \end{aligned}$$

PROOF OF THEOREM 4.4. Suppose (4.8a, b) holds. By (4.8a),

$$D(F(\cdot, \mathscr{I}_I)) = D(M_I),\ I \in \mathscr{I}. \tag{5.2}$$

For every Borel set $B \subset [0, \alpha(F))$, $F(B, \mathscr{I}_I) \equiv \mathrm{P}(\tau \in \mathrm{B}, \xi(\mathbf{T}) \in \mathscr{I}_I)$

$$\begin{aligned} &= P(\tau_I \in B, \tau_I \le \tau_{I'}) \\ &= \int_B P(\tau_{I'} \ge u \mid \tau_I = u)\, dM_I(u) \\ &= \int_B P(\tau_{I'} > u \mid \tau_I > u)\, dM_I(u) \qquad \text{[by 4.8 b]} \\ &= \int_B [\bar{F}(u)/\bar{M}_I(u)]\, dM_I(u). \end{aligned}$$

Thus, for every Borel set $B \subseteq [0, \alpha(F))$,

$$F(B, \mathscr{I}_I) = \int_B \{\bar{F}(u)/\bar{M}_I(u)\}\, dM_I(u). \tag{5.3}$$

Relations (5.2) and (5.3) together imply that

$$dF^C(u, \mathscr{I}_I)/dM_I^C = \bar{F}(u)/\bar{M}_I(u).$$

It follows by (4.8a) and by Lemma 5.1, that

$$\begin{aligned} \textstyle\prod_{J \in \mathscr{I}_I} \bar{G}_J(t) &= \textstyle\prod_a \chi\{a \in D[F(\cdot, \mathscr{I}_I)] \cap [0, t]\}[\bar{F}(a)/\bar{F}(a^-)] \\ &\quad \cdot \exp\left\{-\int_0^t dF^C(u, \mathscr{I}_I)/\bar{F}(u)\right\} \\ &= \textstyle\prod_a \chi\{a \in D(M_I) \cap [0, t]\}[\bar{M}_I(a)/\bar{M}_I(a^-)] \\ &\quad \cdot \exp\left\{-\int_0^t dM_I^C(u)/\bar{M}_I(u)\right\} = \bar{M}_I(t). \end{aligned}$$

Conversely, suppose (4.7) holds. By (2.2) and by Lemma 5.1,

$$\begin{aligned} &\textstyle\prod_a \chi\{a \in D[F(\cdot, \mathscr{I}_I)] \cap [0, t]\}[\bar{F}(a)/\bar{F}(a^-)]\exp\left\{-\int_0^t dF^C(u, \mathscr{I}_I)/\bar{F}(u)\right\} \\ &\quad = \textstyle\prod_a \chi[a \in D(M_I) \cap [0, t]\}[\bar{M}_I(a)/M_I(a^-)]\exp\left\{-\int_0^t dM_I^C(u)/\bar{M}_I(u)\right\}. \end{aligned} \tag{5.4}$$

Letting $\prod$ denote the product over sets $J \in \mathscr{I}_I$, we have

$$\begin{aligned} \bar{M}_I(a)/\bar{M}_I(a^-) &= \textstyle\prod [\bar{G}_J(a)]/\prod[\bar{G}_J(a^-)] \\ &= \textstyle\prod[\bar{F}(a)/\bar{F}(a^-)], \qquad a \in D(F(\cdot, J)) \\ &= 1 \qquad \text{otherwise} \\ &= \bar{F}(a)/\bar{F}(a^-), \qquad a \in D(F(\cdot, \mathscr{I}_I)) \\ &= 1 \qquad \text{otherwise.} \end{aligned}$$

Thus, (4.8a) holds. Equation (4.8b) follows from (5.4) by cancellation. □

Before we prove Theorem 4.7, we present two explanatory remarks and introduce some notation.

REMARK 5.3. Assumptions (i) and (ii) of Theorem 4.7 together imply that for $I \in \mathscr{I}$,

$$\lim_{n\to\infty} \sup_{0\le t\le\alpha(F)} | F_n(\{t\}, I) - F(\{t\}, I) | = 0.$$

Remark 5.4. Remark 5.3 implies that

$$D(I) \subseteq \cup_{n=1}^{\infty} \cap_{k\ge n} D(I, k) \qquad \text{for all } I \in \mathscr{I}.$$

Let

$$B_{1,n}^{I}(t) \equiv \textstyle\sum_{a\in D(I)} \chi\{a \in D(I, n) \cap [0, \alpha(F_n)) \cap [0, t)\} \cdot \ln[\bar{F}_n(a)/\bar{F}_n(a^-)],$$

$$B_1^I(t) \equiv \textstyle\sum_{a\in D(I)} \chi\{a \in [0, \alpha(F)) \cap [0, t)\} \cdot \ln[\bar{F}(a)/\bar{F}(a^-)],$$

$$B_{2,n}^{I}(t) \equiv \textstyle\sum_{a\in D(I,n)} \chi\{a \in C(I) \cap [0, \alpha(F_n)) \cap [0, t)\} \cdot \ln[\bar{F}_n(a)/\bar{F}_n(a^-)] - \int_{[0,\, t)} \chi\{u \in C(I, n) \cap [0, \alpha(F_n))\}\, dF_n(u, \mathscr{I})/\bar{F}_n(u),$$

and

$$B_2^I(t) \equiv -\int_{[0,\, t)} \chi\{u \in C(I) \cap [0, \alpha(F))\}\, dF(u, I)/\bar{F}(u).$$

Note that for $0 \le t \le \alpha(F)$, $I \in \mathscr{I}$, and $n = 1, 2, \cdots$:

$$\bar{G}_{I,n}(t^+) = \exp\{B_{1,n}^I(t) + B_{2,n}^I(t)\},$$

$$\bar{G}_I(t^+) = \exp\{B_1^I(t) + B_2^I(t)\},$$

$$\textstyle\prod_{I\in\mathscr{I}} \bar{G}_{I,n}(t^+) = \bar{F}_n(t^+),$$

and that

$$\textstyle\prod_{I\in\mathscr{I}} \bar{G}_I(t^+) = \bar{F}(t^+).$$

Since $\bar{F}(t^+) = \lim_{n\to\infty} \bar{F}_n(t^+) = \lim_{n\to\infty} \prod_{I\in\mathscr{I}} \bar{G}_{I,n}(t^+) \le \prod_{I\in\mathscr{I}} \limsup_{n\to\infty} \bar{G}_{1,n}(t^+)$, to prove Theorem 4.7 it suffices to show that for an arbitrary subsequence $\{m\}$ of $\{1, 2, \cdots\}$, $I \in \mathscr{I}$, and $0 \le t < \alpha(F)$,

$$\limsup_{m\to\infty} B_{j,m}^I(t) \le B_j^I(t), \qquad j = 1, 2. \tag{5.5}$$

In the following three lemmas we prove (5.5).

LEMMA 5.5. *Under assumptions* (i)–(iv) *of Theorem* 4.7,

$$\alpha \equiv \liminf_{n\to\infty} \alpha(F_n) \ge \alpha(F). \tag{5.6}$$

PROOF. It suffices to consider the case $\alpha < \infty$. Let $\{m\}$ be a subsequence of $\{1, 2, \cdots\}$ such that $\lim_{m\to\infty} \alpha(F_m) = \alpha$. By definition, $\bar{F}_m[\alpha(F_m)] = 0$. By (iv), $\lim_{m\to\infty} \bar{F}[\alpha(F_m)] = 0$. Hence $\bar{F}(\alpha) = 0$. Consequently relation (5.6) follows from the definition of $\alpha(F)$.

In Lemmas 5.6 and 5.7 below, let $\{m\}$ denote an arbitrary infinite subsequence of $\{1, 2, \cdots\}$.

LEMMA 5.6. *Let us assume that assumptions* (i)–(iv) *of Theorem* 4.7 *hold. Then for all* $I \in \mathscr{I}$ *and* $0 \le t \le \alpha(F)$,

$$\limsup_{m\to\infty} B_{1,m}^I(t) \le B_1^I(t).$$

PROOF. By Fatou's Lemma,

$\limsup_{m\to\infty} B^I_{1,m}(t) \le \sum_{a\in D(I)} \limsup_{m\to\infty} [\chi\{a \in D(I, m) \cap$

$$[0, \alpha(F_m)) \cap [0, t)\}\cdot\ln\{\bar F_n(a)/\bar F_n(a^-)\}].$$

Consequently the result of the lemma follows by Remark 5.3, Lemma 5.5, and by Assumption (iv) of Theorem 4.7. □

LEMMA 5.7. *Let us assume that assumptions* (i)–(iv) *of Theorem* 4.7 *hold. Then for all* $I \in \mathscr{I}$ *and* $0 \le t \le \alpha(F)$,

$$\limsup_{m\to\infty} B^I_{2,m}(t) \le B^I_2(t).$$

PROOF. By Assumption (i) of Theorem 4.7,

$$\sum_{a\in D(I,m)} \chi\{a \in C(I) \cap [0, \alpha(F_m)) \cap [0, t)\}\cdot\ln[\bar F_m(a)/\bar F_m(a^-)]$$

$$= -\int_{[0,t)} \chi\{u \in D(I, m) \cap C(I)$$

$$\cap [0, \alpha(F_m))\}\cdot\ln[1 + F_m(\{u\}, I)/\bar F_m(w)]\, dF_m(u, I)/F_m(\{u\}, I).$$

By Lemma 5.5,

$$B^I_{2,m}(t) \le -\int_{[0,\, t)} \chi\{u \in C(I)\}\, dF_m(u, I)/\bar F_m(u)$$

$$-\int_{[0,t)} \chi\{u \in D(I, m) \cap C(I)\}$$

$$\cdot\{[\ln[1 + F_m(\{u\}, I)/\bar F_m(u)]/F_m(\{u\}, I)] - 1/\bar F_m(u)\}\cdot dF_m(u, I).$$

Consequently the result of the lemma follows by Assumptions (ii)–(iv) of Theorem 4.7 and the Helly-Bray theorem. □

REFERENCES

[1] AALEN, O. (1976). Nonparametric inference in connection with multiple decrement models. *Scand. J. of Statist.* **3** 15–27.

[2] 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.

[3] DESU, M. M. and NARULA, S. C. (1977). Reliability estimation under competing causes of failure. In *The Theory and Applications of Reliability I*, (I. Shimi and C. P. Tsokos Eds.) Academic, New York.

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

[5] LANGBERG, N., PROSCHAN, F. and QUINZI, A. J. (1978). Converting dependent models into independent ones, preserving essential features. *Ann. Probability* **6** 174–181.

[6] MARSHALL, A. W. and OLKIN, I. (1967). A multivariate exponential distribution. *J. Amer. Statist. Assoc.* **62** 30–44.

[7] MEIER, P. (1976). Estimation of a distribution function from incomplete observations. *Perspectives in Probability and Statistics.* Applied Probability Trust, Sheffield.

[8] MILLER, D. R. (1977). A note on independence of multivariate lifetimes in competing risks models. *Ann. Statist.* **5** 576–579.

[9] PETERSON, A. V., JR. (1975). Nonparametric estimation in the competing risks problem. Stanford Univ. Tech. Report No. 13.

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

[11] TSIATIS, A. (1975). A nonidentifiability aspect of the problem of competing risks. *Proc. Nat. Acad. Sci. USA* **72** 20–22.

N. LANGBERG AND F. PROSCHAN
DEPARTMENT OF STATISTICS
FLORIDA STATE UNIVERSITY
TALLAHASSEE, FLORIDA 32306

A. J. QUINZI
DEPARTMENT OF STATISTICS
TEMPLE UNIVERSITY
PHILADELPHIA, PENNSYLVANIA 19122