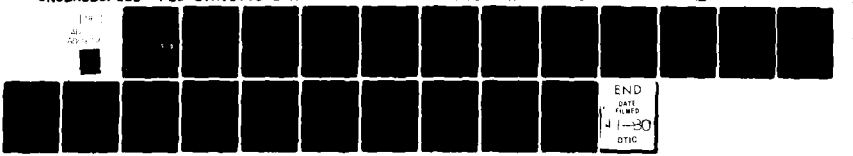
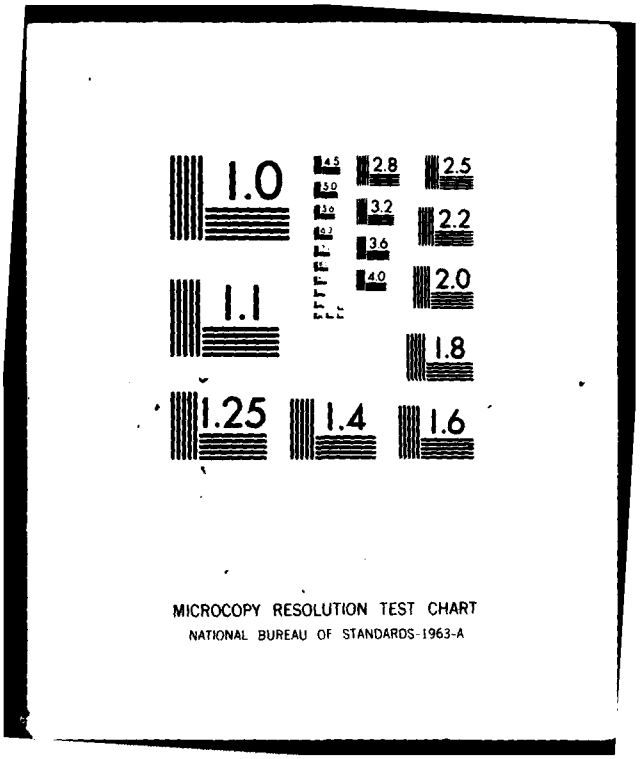
AD-A089 799 FLORIDA STATE UNIV TALLAHASSEE DEPT OF STATISTICS F/G 12/1
ON THE IDENTIFIABILITY OF MULTIVARIATE LIFE DISTRIBUTION FUNCTI--ETC(U)
JUL 80 N A LANGBERG, M SHAKED AFOSR-78-3678
UNCLASSIFIED FSU-STATISTICS-M553 AFOSR-TR-80-0940 NL



END
DATE FILMED
J 1-80
DTIC


1.0
1.1
1.25
1.4
1.6
1.8
2.0
2.2
2.5
2.8
3.2
3.6
4.0
4.5
5.0
5.6
6.3
7.1
MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

(4) LEVEL II

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER (18) (19) AFOSR-TR-80-0940	2. GOVT ACCESSION NO. AD-A089799
3. RECIPIENT'S CATALOG NUMBER	
4. TITLE (and Subtitle) (2) ON THE IDENTIFIABILITY OF MULTIVARIATE LIFE DISTRIBUTION FUNCTIONS.	5. TYPE OF REPORT & PERIOD COVERED (9) Technical rept. Interim
	6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) (10) Naftali A. Langberg, Moshe Shaked	8. CONTRACT OR GRANT NUMBER(s) (15) AFOSR-78-3678, NSF-MCS 79-27150
9. PERFORMING ORGANIZATION NAME AND ADDRESS Florida State University, Department of Statistics, Tallahassee, FL 32306 (11) Jul 80	10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS (16) 61102F (17) 2304/A5
11. CONTROLLING OFFICE NAME AND ADDRESS Air Force Office of Scientific Research/NM, Bolling AFB, Washington, DC 20032	12. REPORT DATE July 1980
	13. NUMBER OF PAGES 19 (12) 22p.
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office) (14) FSU-STATISTICS-M353, TR-78-114-AFOSR	15. SECURITY CLASS. (of this report) UNCLASSIFIED
	15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approve for public release; distribution unlimited.

DTIC ELECTE S OCT 2 1980 D B

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Identifiability of distributions, Product limit estimators, censored data, multivariage distributions

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

Please see back for abstract.

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

UNCLASSIFIED 400277 dk

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

AD A089799

DDC FILE COPY

Let (T_1, T_2) and (L_1, L_2) be two independent bivariate random vectors with distributions F and H. Let $\tau_1 = \min(T_1, L_1)$, $\tau_2 = \min(T_2, L_2)$ and let $G_{0,0}(s, t) = P\{\tau_1 \leq s, \tau_2 \leq t, T_1 \leq L_1, T_2 \leq L_2\}$, $G_{0,1}(s, t) = P\{\tau_1 \leq s, \tau_2 \leq t, T_1 \leq L_1, L_2 < T_2\}$, $G_{1,0}(s, t) = P\{\tau_1 \leq s, \tau_2 \leq T, L_1 < T_1, T_2 \leq L_2\}$ and $G_{1,1}(s, t) = P\{\tau_1 \leq s, \tau_2 \leq t, L_1 < T_1, L_2 < T_2\}$. Under mild conditions the distributions F and H are expressed explicitly as functionals of $G_{0,0}$, $G_{0,1}$, $G_{1,0}$ and $G_{1,1}$. Necessary and sufficient conditions for the formulas to hold even when (T_1, T_2) and (L_1, L_2) are not independent are derived. Numerous applications are indicated. Extension of the results to p-dimensional distributions $(p > 2)$ is given.

Accession For
NTIS GRA&I ☑
DTIC TAB ☐
Unannounced ☐
Justification
By
Distribution/
Availability Codes
Dist | Avail and/or Special
A

AFOSR-TR- 80-0940

On the Identifiability of Multivariate Life Distribution Functions

by

Naftali A. Langberg[1] and Moshe Shaked[1,2]
Florida State University Indiana University

FSU Statistics Report No. M553
AFOSR Technical Report No. 78-114

July, 1980
The Florida State University
Department of Statistics
Tallahassee, Florida 32306

[1]Research supported by the Air Force Office of Scientific Research, AFSC, USAF, under Grant AFOSR-78-3678.

[2]Research supported by N.S.F. grant MCS-79-27150.

Approved for public release; distribution unlimited.

80 9 25 118

On the Identifiability of Multivariate Life Distribution Functions

by

Naftali A. Langberg and Moshe Shaked

Abstract.

Let (T_1, T_2) and (L_1, L_2) be two independent bivariate random vectors with distributions F and H. Let $\tau_1 = \min(T_1, L_1)$, $\tau_2 = \min(T_2, L_2)$ and let $G_{0,0}(s, t) = P\{\tau_1 \leq s, \tau_2 \leq t, T_1 \leq L_1, T_2 \leq L_2\}$, $G_{0,1}(s, t) = P\{\tau_1 \leq s, \tau_2 \leq t, T_1 \leq L_1, L_2 < T_2\}$, $G_{1,0}(s, t) = P\{\tau_1 \leq s, \tau_2 \leq T, L_1 < T_1, T_2 \leq L_2\}$ and $G_{1,1}(s, t) = P\{\tau_1 \leq s, \tau_2 \leq t, L_1 < T_1, L_2 < T_2\}$. Under mild conditions the distributions F and H are expressed explicitly as functionals of $G_{0,0}$, $G_{0,1}$, $G_{1,0}$ and $G_{1,1}$. Necessary and sufficient conditions for the formulas to hold even when (T_1, T_2) and (L_1, L_2) are not independent are derived. Numerous applications are indicated. Extension of the results to p-dimensional distributions $(p > 2)$ is given.

Key words: Identifiability of distributions, Product limit estimators, censored data, multivariate distributions.

AMS subject classification. Primary 62N05; Secondary 62E10.

AIR FORCE OFFICE OF SCIENTIFIC RESEARCH (AFSC)
NOTICE OF TRANSMITTAL TO DDC
This technical report has been reviewed and is approved for public release IAW AFR 190-12 (7b).
Distribution is unlimited.
A. D. BLOSE
Technical Information Officer

1. Introduction and Summary.

By observing a series system of d components (d = 2, 3, ...) we can only determine its lifelength and the components that cause the system to fail. In particular the lifelength of a series system that consists of a nonempty subset of the original d components is unobservable. We refer to the distribution function (d.f.) of the lifelength of this series system as an unobservable d.f. For example, some of the original d components, such as wires and switches, may be placed in the system to support the operation of the main components. The d.f. of the lifelength of the series system that consists only of the main components is of great importance to the theory of engineering reliability but is unobservable.

Langberg, Proschan, and Quinzi (LPQ) (1978), although addressing themselves to a different problem present, in particular, a way to relate the unobservable d.f.'s to observable quantities (see Th. 4.1). First, LPQ (1978) define $2^d - 1$ observable subdistribution functions (s.d.f.'s). The value of an unobservable s.d.f. at the point $s (s \geq 0)$ is the probability that the original series system fails at a time less than or equal to s as a result of the simultaneous failure of a particular nonempty subset of the d components and of no other components. Then LPQ (1978) express the various unobservable d.f.'s as functionals of the observed s.d.f.'s, whenever the lifelengths of the d components are independent and satisfy a mild regularity condition. Thus, they provide the researchers in the theory of engineering reliability with a theoretical tool to determine the unobservable d.f.'s. The problem of relating unobservale d.f.'s to observable s.d.f.'s has been considered by several other researchers such as Peterson (1975), Tsiatis (1975), and Miller (1977).

The theoretical relationship discussed in the previous paragraph suggest a "naive" method of estimating the unobservable d.f.'s. LPQ (1980) consider the problem of estimating the various unobservable d.f.'s based on data collected from n independent and identical series systems each consisting of d components. First, LPQ (1980) estimate the observable s.d.f.'s by their empirical counterparts. Then they replace the observable s.d.f's by their empirical estimators in the respective functionals. Using this method LPQ (1980) obtain the Product Limit Estimators (P.L.E.'s) for the observable d.f.'s, first introduced by Kaplan and Meier (1958).

An important statistical problem is to test whether the life distribution of a component, F, is exponential versus various wearout classes of alternatives on the basis of a random sample from the d.f. F in the absence of the other d-1 nuisance components. Tests for these hypotheses have been proposed, for example, by Proschan and Pyke (1967), Barlow (1968), Bickel (1969), Bickel and Doksum (1969), Barlow and Doksum (1972), Hollander and Proschan (1972), (1975), (1979), and Koul (1977), (1978a), (1978b). There seems to be a growing interest among statisticians to test these hypotheses in a more realistic situation: Where additional "nuisance" components are present. Koul and Susarla (1978) and Chen, Hollander, and Langberg (1980a), (1980b) suggest tests for some of the hypotheses described above based on the P.L.E.

In this paper we present a formula that relates unobservable multivariate d.f.'s to observable multivariate s.d.f.'s. This formula suggest further development of statistical estimation and testing procedures concerning multivariate d.f.'s in the presence of "nuisance" components.

Assume that a pair of individuals, a wife and a husband for example, are under study. The observation of each of the two individuals is terminated in the event of death or in case of a withdrawal from the study. The joint lifelengths of the two individuals is of great importance to the theory of biostatistics but is unobservable.

In Section 2 we relate the unobservable joint life distribution of the two individuals to observable quantities. First we define 4 observable joint s.d.f.'s. The value of an observable joint s.d.f. at times t, s($t, s \geq 0$) is the probability that the two individuals are removed from the study at times less than or equal to t and s, respectively, as a result of a specific combination of deaths and withdrawals. Then we express the unobservable joint d.f. as a functional of the observable s.d.f's, whenever some mild conditions hold. Thus we provide the researchers in biostatistics with a *theoretical tool to determine the unobservable d.f.*

Langberg and Shaked (1980a) utilize the theoretical relationship established in Section 2 to estimate the unobservable joint d.f. from a sample of n independent and identical pairs of individuals. Following LPQ (1980)s idea they obtain a bivariate P.L.E. for the unobservable bivarite d.f. and establish the strong consistency, weak convergence and some other desirable properties of the bivariate P.L.E. [For more details see Langberg and Shaked (1980a), (1980b)].

In Section 3 we consider a group of p individuals, $p = 1, 2, \ldots$, a family for example, and assume that each individual is exposed to d risks, $d = 2, 3, \ldots$. Using the results obtained in Section 2 we relate the joint d.f. of the times to deaths of the p individuals from risk j, $j = 1, \ldots, d$, to the observable s.d.f's. The value of an observable s.d.f. at times

$t_1, \ldots, t_p (t_1, \ldots, t_p \geq 0)$ is the probability that the p individuals die at times less than or equal to $t_1, \ldots, t_p$, respectively, as a result of the occurrence of a specific combination of risks.

Finally, we note that although we have chosen to employ the languages of reliability theory, biostatistics and the theory of competing risks (series system, components, death, withdrawal, risks, etc.) the results presented here apply to any model where observations include (1) The time at which a particular event occurs and (2) The cause(s) of the occurrence of the event. As is the case, for example, in mortality studies (Hoel(1972)) and some mathematical epidemiology models (see Billard, Lacayo and Langberg (1979), Lacayo and Langberg (1980a), (1980b) and Langberg (1980).)

Throughout we define a product over an empty set of indicies as 1 and an integral over an empty set as zero.

2. The Bivariate case.

Consider a pair of individuals, a wife and a husband for example, with lifelengths T_1 and T_2. The observation of each of the two individuals is terminated in the event of death or in case of a withdrawal from the study. Let L_1 and L_2 denote the withdrawal times of the two individuals respectively. We assume that the random variables (r.v.'s) T_1, T_2, L_1 and L_2 are defined on a common probability space $(\Omega, \mathcal{B}, P)$.

In this section we present the joint d.f. of the lifelengths (T_1, T_2) as a functional of observable quantities. To be more specific we introduce two definitions and some notation.

Definition 2.1. We say that individual i is removed from the study due to death (withdrawal) and write $\xi_i = 1 (\xi_i = 2)$ if $T_i \leq L_i (L_i < T_i)$, $i = 1, 2$.

Definition 2.2. Let K be a function defined on $(-\infty, \infty)$. We say that K is a subdistribution function (s.d.f.) if K is nondecreasing, right-continuous, and assumes values in $[0, 1]$.

Let $\tau_i = \min\{T_i, L_i\}$ be the removal time of individual i, and let G_i be the d.f. of τ_i, $i = 1, 2$. Further, let $F(t, s, I_1, I_2) = P\{\tau_1 \leq t, \tau_2 \leq s, \xi_1 = I_1, \xi_2 = I_2\}$, $I_1, I_2 \in \{1, 2\}$ be the observable joint s.d.f.'s and let $H(t, s) = P\{T_1 \leq t, T_2 \leq s\}$ be the unobservable joint d.f. of the lifelengths of the two individuals under consideration. For a s.d.f. K, let $\overline{K}(t) = \lim_{s \to \infty} K(s) - K(t)$ be the subsurvivalfunction corresponding to K, let $\overline{K}(t-) = \lim_{s \to t-} \overline{K}(s)$, and let $\alpha(K) = \sup\{t: \overline{K}(t) > 0\}$. Further, let K_c be the continuous s.d.f. corresponding to K and let C(K)(D(K)) be the set of all continuity (discontinuity) points of K. In this section we present the d.f. $H(\cdot, \cdot)$ as a functional of the s.d.f.'s: $F(\cdot, \cdot, I_1, I_2)$, $I_1, I_2 = 1, 2$, provided that:

(A.1) The lifelengths (T_1, T_2) and the withdrawal times (L_1, L_2) are independent random vectors,

and that

(A.2) For $i = 1, 2$, the sets $D(P\{\tau_i \leq \cdot, \xi_i = 1\})$ and $D(P\{\tau_i \leq \cdot, \xi_i = 2\})$ are disjoint.

First, we reduce the problem of obtaining the desired functional that relates the observable joint s.d.f.'s to the unobservable joint d.f. concerning two individuals to a similar, but simpler problem, concerning only one individual. Note that by Assumption (A.1)

$$(2.1) \qquad P\{T_1 > t, T_2 > s\} = P\{T_1 > t | T_2 > s\}P\{T_2 > s\}$$
$$= P\{T_1 > t | \tau_2 > s\}P\{T_2 > s\} \text{ for } t\in(-\infty,\infty),\ s\in(-\infty,\alpha(G_2)),$$

(2.2) The r.v.'s T_2 and L_2 are independent,

and that

(2.3) For $s\in(-\infty,\alpha(G_2))$ the conditional r.v.'s $\{T_1|\tau_2 > s\}$ and $\{L_1|\tau_2 > s\}$ are independent.

Further, by (A.2):

(2.4) For $s\in(-\infty,\alpha(G_2))$ the sets $D(P\{\tau_1 \leq \cdot, \xi_1 = 1|\tau_2 > s\})$ and $D(P\{\tau_1 \leq \cdot, \xi_1 = 2|\tau_2 > s\})$ are disjoint.

Observe that to express $H(\cdot, \cdot)$ as a functional of the s.d.f.'s: $F(\cdot, \cdot, I_1, I_2)$, $I_1, I_2 = 1, 2$, it suffices (a) to express $P\{T_2 > \cdot\}$ as a functional of the s.d.f.'s: $P\{\tau_2 \leq \cdot, T_2 \leq L_2\}$, $P\{\tau_2 \leq \cdot, L_2 < T_2\}$, (b) to express $P\{T_1 > \cdot|\tau_2 > s\}$ as a functional of the s.d.f.'s: $P\{\tau_1 \leq \cdot, T_1 \leq L_1|\tau_2 > s\}$, $P\{\tau_1 \leq \cdot, L_1 < T_1|\tau_2 > s\}$ for every $s\in(-\infty,\alpha(G_2))$ and

(c) to insert these functionals in Equation (2.1). This is done in Theorem 2.5, but first we need some preliminaries.

Let T and L be two independent r.v.'s representing, respectively, the lifelength and withdrawal time of a single individual. Further, let $\tau = \min\{T, L\}$ be the removal time of the individual from the study and let G denote the d.f. of τ. We express now the unobservable survival function: $P\{T > \cdot\}$ as a functional of the observed s.d.f.'s: $P\{\tau \leq \cdot, T \leq L\}$, $P\{\tau \leq \cdot, L < T\}$. The following lemma is useful.

Lemma 2.3. Let $Q(\cdot)$ be a s.d.f. Then for $t \in (-\infty, \alpha(Q))$:

$$\bar{Q}(t) = \lim_{u \to -\infty} \bar{Q}(u) \Big\{ \prod_{\substack{a \leq t \\ a \in \bar{D}(Q)}} [\bar{Q}(a)][\bar{Q}(a-)]^{-1} \Big\}$$

$$\cdot \exp\Big\{ -\int_{-\infty}^{t} [\bar{Q}(a)]^{-1} dQ_C(a) \Big\}. \tag{2.5}$$

For a proof of Lemma 2.1 see Lee and Thompson (1976), Th. 2.2.

We present now the desired functional for a single individual.

Lemma 2.4. Let C and D be, respectively, the sets of continuity and discontinuity points of the s.d.f. $P\{\tau \leq \cdot, T \leq L\}$. Assume that:

(A.3) The r.v.'s T and L are independent,

and that

(A.4) The sets D and $D(P\{\tau \leq \cdot, L < T\})$ are disjoint.

Then for $t \in (-\infty, \alpha(G))$:

$$P\{T > t\} = \Big\{ \prod_{\substack{a \leq t \\ a \in D}} [\bar{G}(a)][\bar{G}(a-)]^{-1} \Big\}$$

$$\cdot \exp\Big\{ -\int_{-\infty}^{t} [\bar{G}(a)]^{-1} dP\{\tau \in \{(-\infty, a]\} \cap C,\ T \leq L\} \Big\}. \tag{2.6}$$

Proof. By Assumptions (A.3) and (A.4)

$$(2.7)\qquad \bar{G}(a)[\bar{G}(a-)]^{-1} = P\{T > a\}[P\{T \geq a\}]^{-1},\ a\in(-\infty,\ \alpha(G)),$$

$$(2.8)\qquad P\{\tau \leq a,\ T \leq L\} = \int_{-\infty}^{a} P\{L \geq v\}dP\{T \leq v\},\ a\in(-\infty,\ \infty),$$

$$P\{\tau \leq a,\ L < T\} = \int_{-\infty}^{a} P\{T > v\}dP\{L \leq v\},\ a\in(-\infty,\ \infty),$$

and

$$(2.9)\qquad P\{\tau > a\} = P\{T > a\}P\{L > a\},\ a\in(-\infty,\ \infty).$$

Thus, by Assumption (A.4)

(2.10) The sets $D(P\{T \leq \cdot\})\cap\{(-\infty,\ \alpha(G))\}$ and $D(P\{L \leq \cdot\})\cap\{(-\infty,\ \alpha(G))\}$ are disjoint.

Let χ denote the indicator function. Then by Assumption (A.3) and Equations (2.8), (2.9) and (2.10):

$$(2.11)\qquad \int_{-\infty}^{t} \chi(C)(a)[\bar{G}(a)]^{-1}dP\{\tau \leq a,\ T \leq L\} =$$

$$= \int_{-\infty}^{t} \chi(C)(a)[P\{T > a\}]^{-1}dP\{T \leq a\},\ t\in(-\infty,\alpha(G)).$$

Finally, note that $\alpha(G) \leq \sup\{t\colon P\{T > t\} > 0\}$ and that $C\cap\{(-\infty,\alpha(G))\} = C(P\{T \leq \cdot\})\cap\{(-\infty,\alpha(G))\}$. Consequently the desired result follows from Equations (2.5), (2.7), and (2.11). ||

We return now to our original problem, and present the functional that relates the joint d.f. of the lifelengths of the two individuals considered to their observable s.d.f.'s.

Theorem 2.5. Assume (T_1, T_2, L_1, L_2) satisfies Conditions (A.1) and (A.2). Let $C_2(D_2)$ be the sets of continuity (discontinuity) points of the

s.d.f. $P\{\tau_2 \le \cdot, T_2 \le L_2\}$, let $C_{1,s}(D_{1,s})$ be the sets of continuity (discontinuity) points of the conditional s.d.f. $P\{\tau_1 \le \cdot, T_1 \le L_1 | T_2 > s\}$, and let $\alpha_{1,s} = \sup\{t: P\{\tau_1 > t | \tau_2 > s\} > 0\}$, $s \in (-\infty, \alpha(G_2))$. Then for $(t, s) \in \{(u,v): v \in (-\infty, \alpha(G_2)), u \in (\infty, \alpha_{1,v})\}$:

$$P\{T_1 > t, T_2 > s\} = \{ \prod_{\substack{u \le s \\ u \in D_2}} [\bar{G}_2(u)][\bar{G}_2(u-)]^{-1} \}$$

$$\{ \prod_{\substack{v \le t \\ v \in D_{1,s}}} [P\{\tau_1 > v | \tau_2 > s\}][P\{\tau_1 \ge v | \tau_2 > s\}]^{-1} \} \tag{2.12}$$

$$\cdot \exp\{-\int_{-\infty}^{s} [\bar{G}_2(u)]^{-1} \chi(C_2)(u) dP\{\tau_2 \le u, T_2 \le L_2\}\}$$

$$\cdot \exp\{-\int_{-\infty}^{t} [P\tau_1 > v | \tau_2 > s]^{-1} \chi(C_{1,s})(v) dP\{\tau_1 \le v, T_1 \le L_1 | \tau_2 > s\}\}.$$

Proof. First note that by Assumptions (A.1) and (A.2) Conditions (A.3) and (A.4) hold for $(T, L) = (T_2, L_2)$.

Thus, by Lemma 2.4 for $s \in (-\infty, \alpha(G_2))$:

$$P\{T_2 > s\} = \{ \prod_{\substack{a \le s \\ a \in D_2}} [\bar{G}_2(a)][\bar{G}_2(a-)]^{-1} \}$$

$$\cdot \exp\{-\int_{-\infty}^{s} [\bar{G}_2(u)]^{-1} \chi(C_2)(u) dP\{\tau_2 \le u, T_2 \le L_2\}. \tag{2.13}$$

Let s $(-\infty, \alpha(G_2))$. Note that by Assumptions (A.1) and (A.2) Conditions (A.3) and (A.4) hold for the random pair (T, L) that is stochastically equal to the conditional random pair $[(T_1, L_1) | \tau_2 > s]$. Thus, by Lemma 2.4 for $t \in (-\infty, \alpha_{1,s})$:

$$P\{T_1 > t | \tau_2 > s\} = \prod_{\substack{v \le t \\ v \in D_{1,s}}} [P\{\tau_1 > v | \tau_2 > s\}][P\{\tau_1 \ge v | \tau_2 > s\}]^{-1}$$

$$\cdot \exp\{-\int_{-\infty}^{t} [P\{\tau_1 > v | \tau_2 > s\}]^{-1} \chi(C_{1,s})(v) dP\{\tau_1 \le v, T_1 \le L_1 | \tau_2 > s\}\}. \tag{2.14}$$

Consequently the desired result follows by Equations (2.1), (2.13) and (2.14) by insertion. ||

In many life testing situations the d.f.'s of the lifelengths and censoring times of the two individuals considered are anticipated to be strictly increasing smoothly on $[0, \infty)$. For these cases we obtain now from Theorem 2.5 a corollary of practical importance.

Corollary 2.6. Let τ_1, τ_2 be continuous r.v.'s such that for t, $s\epsilon(-\infty,\infty)$ $P\{T_1 > t, T_2 > s\}$, $P\{L_1 > t, L_2 > s\}$ are strictly positive. Assume Condition (A.1) holds. Then for t, $s\epsilon(-\infty,\infty)$:

$$P\{T_1 > t, T_2 > s\} = \exp\{-\int_{-\infty}^{s} [\overline{G}_2(u)]^{-1} dP\{\tau_2 \leq u, T_2 \leq L_2\}$$

(2.15

$$-\int_{-\infty}^{t} [P\{\tau_1 > v | \tau_2 > s\}]^{-1} dP\{\tau_1 \leq v, T_1 \leq L_1 | \tau_2 > s\}\}.$$

Proof. Apply Theorem 2.5 and note that $\alpha(G_2) = \alpha_{1,s} = \infty$, $s\epsilon(-\infty,\infty)$, and that D_2, $D_{1,s}s\epsilon(-\infty,\infty)$ are empty sets. ||

Finally, we note that the relationship between the unobservable d.f. and the observable s.d.f's given by Equation (2.12) holds for some cases where the independence assumption, given by (A.1), does not hold. First, we present a necessary and sufficient condition for the relationship given by (2.12) to hold. Then we give an example for which Assumption (A.1) does not hold, however the desired relationship, given by Equation (2.12), holds. We need the following lemma.

Lemma 2.7. (LPQ (1980), Th. 4.4) Let (L, T) be a pair of r.v.'s defined on a common probability space. Let C, D be as in Lemma 2.4, but

now do not assume Conditions (A.3) or (A.4). Then equation (2.6) holds if the following conditions are satisfied.

$$(A.5)\quad [P\{T > t\}][P\{T \geq t\}]^{-1} = \begin{cases} [\overline{G}(t)][\overline{G}(t-)]^{-1}, & t \in D \cap \{(-\infty, \alpha(G))\} \\ 1, & t \in C \cap \{(-\infty, \alpha(G))\}, \end{cases}$$

and

(A.6) For $t \in C \cap \{(-\infty, \alpha(G))\}$, $P\{L \geq t | T = t\} =$

$P\{L > t | T > t\}$ a.s. with respect to $P\{T \leq \cdot\}$.

We are ready to present the necessary and sufficient conditions.

Theorem 2.8. Let T_1, L_1, T_2, and L_2 be r.v.'s defined on a common probability space. Let C_2, D_2, $C_{1,s}$, $D_{1,s}$, $\alpha_{1,s}$, $s \in (-\infty, \alpha(G_2))$ be as in Theorem 2.5. Assume Condition (2.1) holds. Then Equation (2.12) holds iff the following conditions hold.

$$(A.7)\quad [P\{T_2 > s\}][P\{T_2 \geq s\}]^{-1} = \begin{cases} [\overline{G}_2(t)][\overline{G}_2(t-)]^{-1}, & t \in D_2 \cap \{(-\infty, \alpha(G_2))\} \\ 1, & t \in C_2 \cap \{(-\infty, \alpha(G_2))\} \end{cases}$$

$$(A.8)\quad [P\{T_1 > t,\ \tau_2 > s\}][P\{T_1 \geq t,\ \tau_2 > s\}]^{-1} =$$

$$= \begin{cases} [P\{\tau_1 > t,\ \tau_2 > s\}][P\{\tau_1 \geq t,\ \tau_2 > s\}]^{-1}, & t \in D_{1,s} \cap \{(-\infty, \alpha_{1,s})\} \\ 1, & t \in C_{1,s} \cap \{(-\infty, \alpha_{1,s})\}, \end{cases}$$

for $t \in (-\infty, \alpha_{1,s})$, $s \in (-\infty, \alpha(G_2))$.

(A.9) For $s \in C_2 \cap \{(-\infty, \alpha(G_2))\}$, $P\{L_2 \geq s | T_2 = s\} =$

$P\{L_2 > s | T_2 > s\}$ a.s. with respect to $P\{T_2 \leq \cdot\}$,

and

(A.10) For $t \in C_{1,s} \cap \{(-\infty, \alpha_{1,s})\}$, $s \in (-\infty, \alpha(G_2))$

$P\{L_1 \geq t | \tau_2 > s,\ T_1 = t\} = P\{L_1 > t | \tau_2 > s,\ T_1 > t\}$

a.s. with respect to $P\{T_1 \leq \cdot \mid \tau_2 > s\}$.

Proof. Note that

$$(2.16) \qquad P\{T_2 \epsilon B, \epsilon T_2 \leq L_2\} = \int_B P\{L_2 \geq u \mid T_2 = u\} dP\{T_2 \leq u\}$$

for all Borel sets in $(-\infty, \alpha(G_2))$ and that

$$(2.17) \qquad P\{T_1 \epsilon B,\ T_1 \leq L_1 \mid \tau_2 > s\} =$$

$$= \int_B P\{L_1 \geq u \mid \tau_2 > s,\ T_1 = u\} dP\{T_1 \leq u \mid \tau_2 > s\}$$

for all Borel sets in $(-\infty, \alpha(G_1))$.

First, assume Conditions (A.7) through (A.10) hold. Then Equation (2.12) follows by substitution and by Equations (2.6), (2.16) and (2.17).

Now, assume that Equation (2.12) holds. Then in particular the equation holds for $t = 0$, $s \epsilon (-\infty, \alpha(G_2))$. Conditions (A.7) and (A.9) follow now by Lemma 2.7. Since from (A.7) and (A.9) the pair (T_2, L_2) satisfies Equation (2.6), it follows from Equations (2.1) and (2.12) that the conditional pair $(\{T_1 \mid \tau_2 > s\}, \{L_1 \mid \tau_2 > s\})$, $s \epsilon (-\infty, \alpha(G_2))$ satisfies Equation (2.6). Thus, Conditions (A.8) and (A.10) follow by Lemma 2.7. ||

Finally we present an example where the independence assumption, given by (A.1), does not hold, but the desired relationship given by Equation (2.12) is satisfied.

The following definition is needed.

Definition 2.9. (Marshall and Olkin (1967).) Let λ_I, $\phi \neq I \subseteq \{1, \ldots, n\}$ be nonnegative real numbers, $\sum_{\phi \neq I \subseteq \{1, \ldots n,\}} \lambda_I > 0$, and let V_I, $\phi \neq I \subseteq \{1, \ldots, n\}$ be independent exponential r.v.'s with rates λ_I, $\phi \neq I \subseteq \{1, \ldots, n\}$, respectively. The random vector $< U_1, \ldots, U_n >$ has a Marshall-Olkin Multivariate Exponential Distribution (M.O.M.E.D.) with parameters

λ_I, $\phi \neq I \subseteq \{1, \ldots, n\}$ if the two random vectors $< U_1, \ldots, U_n >$ and $< \min\{U_I: I, 1 \in I\}, \ldots, \min\{u_I: I, n \in I\} >$ are stochastically equal.

Example 2.10. Let $< U_1, \ldots, U_4 >$ have a M.O.M.E.D. with parameters λ_I, $\phi \neq I \subseteq \{1, \ldots, 4\}$, and let us assume that $< T_1, T_2, L_1, L_2 >$ and $< U_1, \ldots, U_4 >$ are stochastically equal. Let D_2, $D_{1,s}$ and $\alpha_{1,s}$, $s \in (-\infty, \alpha(G_2))$ be as in Theorem 2.5. Note that $\alpha(G_2) = \alpha_{1,s} = \infty$, that D_2, $D_{1,s}$ are empty sets, and that the two random pairs (T_1, T_2) and (L_1, L_2) are not necessarily independent. We want to show that formula (2.12) holds. To obtain the desired result it suffices, by Theorem 2.8, to verify conditions (A.9) and (A.10). These two conditions follow by some simple calculation with independent exponential r.v.'s. Equation (2.1) does not hold in general, however if $\lambda_{\{1,3,4\}} = \lambda_{\{1,4\}} = 0$, then Equation (2.1) holds.

Note that by Assumption (A.1)

$$P\{T_1 > t, T_2 > s\} = P\{T_2 > s | \tau_1 > t\} P\{T_1 > t\} \text{ for}$$
$$s \in (-\infty, \infty), t \in (-\infty, \alpha(G_1)).$$

Thus, Theorems 2.5, 2.8, and Corollary 2.6 hold when we exchange T_1 and T_2 and L_1 and L_2.

Finally, note that the if we exchange T_1 and L_1 and T_2 and L_2 Theorems 2.5, 2.8, and Corollary 2.6 hold for the random pair (L_1, L_2).

3. The Multivariate Case.

Consider a group of p individuals, a family for example, p = 1, 2, ..., and assume that each individual is exposed to d risks, d = 2, 3, ... Let denote the time to death of individual i from risk j by a nonnegative r.v. $T_{i,j}$, i = 1, ..., p, and let $T_{0,j} = \infty$, j = 1, ..., d. Assume that all these r.v.'s are defined on a common probability space $(\Omega,\mathcal{B},P)$.

In this section we present the joint d.f. of the times to death of the p individual from risk j: $P\{T_{1,j} \leq \cdot, \ldots, T_{d,j} \leq \cdot\}$ as functionals of observable quantities. To be more specific we introduce some notation. Throughout we fix the risk considered and denote it by j. Let $L_{ij} = \min\{T_{iq}, q = 1, \ldots, d, q \neq j\}$, let $\tau_i = \min\{L_{i,j}, T_{i,j}\}$ be the lifelength of individual i and let G_i be the d.f. of τ_i, i = 1, ..., p. Further, let $\xi_{i,j} = 1$ on the set $\{T_{i,j} \leq L_{i,j}\}$, and = 2 on the set $\{L_{i,j} < T_{i,j}\}$, i = 1, ..., p, let $H_j(t_1, \ldots, t_p) = P\{T_{i,j} \leq t_i, i = 1, \ldots, p\}$ be the probability that the p individuals die at times less than or equal to $t_1, \ldots, t_p$, respectively, from risk j in the absence of all the other risks, and let $F_j(t_1, \ldots, t_p, I_1, \ldots, I_p) = P\{\tau_i \leq t, \xi_{i,j} = I_i, i = 1, \ldots, p\}$, $I_i = 1, 2$, i = 1, ..., p, be the observed s.d.f.'s: the probability that the p individuals die at times less than or equal to $t_1, \ldots, t_p$, respectively, as a result of a specific combination of risks. In this section we present each of the unobservable d.f.'s $H_j(\cdot, \ldots, \cdot)$ as a functional of the observed s.d.f.'s: $F_j(\cdot, \ldots, \cdot, I_1, \ldots, I_p)$, provided that:

(B.1) The random vectors $< T_{1,q}, \ldots T_{p,q} >$, q = 1, ..., d are independent, and that

(B.2) For r = 1, ..., p the sets:

$D(P\{\tau_r \le \cdot,\ \xi_{r,q} = 1\})$ and $D(P\{\tau_r \le \cdot,\ \xi_{r,q} = 2\})$ are disjoint.

Theorem 3.1. Let $C^j_{r,t_0,\ldots t_{r-1}}$ ($D^j_{r,t_0,\ldots,t_{r-1}}$) be the sets of continuity (discontinuity) points of $P\{\tau_r \le \cdot,\ \xi_{r,j} = 1,\ \tau_i > t_i,\ i = 0, \ldots, r-1\}$, $r = 1, \ldots, p$, $t_1, \ldots, t_p \in (-\infty,\infty)$, $t_0 = -\infty$. Further, let $\alpha_{1,t_0} = \sup\{t:\ P\{\tau_1 > t\} > 0\}$, $\alpha_{r,t_0,\ldots,t_{r-1}} = \sup\{t:\ P\{\tau_r > t | \tau_1 > t_1, \ldots, \tau_r > t_{r-1}\}\}$, $t_1 < \alpha_{1,t_0}$, $t_2 < \alpha_{2,t_0,t_1}$, $\ldots$, $t_r < \alpha_{r,t_0,\ldots,t_{r-1}}$, $r = 2, \ldots, p$, $t_0 = -\infty$. Assume Conditions (B.1) and (B.2) hold. Then for $(t_0, \ldots, t_p) \in \{(u_0, \ldots, u_p):\ u_0 = -\infty,\ u_r < \alpha_{r,u_0,\ldots,u_{r-1}},\ r = 2, \ldots, p\}$,

$$(3.1) \qquad P\{T_{1,j} > t_1, \ldots, T_{p,j} > t_p\} =$$

$$= \Big[\prod_{r=1}^{p} \prod_{\substack{a \le t \\ a \in D^j_{r,t_0,\ldots,t_{r-1}}}} [P\{\tau_r > a,\ \tau_i > t_i,\ i = 1, \ldots r-1\}]$$

$$\cdot [P\{\tau_r \ge a,\ \tau_i < t_i,\ i = 1, \ldots, r-1\}]^{-1}\Big]$$

$$\cdot \exp\{-\Sigma_{r=1}^{p} \int_{-\infty}^{t_p} [P\{\tau_r > a,\ \tau_i > t_i,\ i = 0, \ldots, r-1\}]^{-1} \chi(C^j_{r,t_0,\ldots,t_{r-1}}) \ (a)$$

$$dP\{\tau_r \le a,\ \xi_{r,j} = 1,\ \tau_i > t_i,\ i = 0, \ldots, r-1\}\}.$$

Proof. First, note that by Condition (B.1)

$$P\{T_{1,j} > t_1, \ldots, T_{p,j} > t_p\} =$$

$$= \prod_{r=1}^{p} P\{T_{r,j} > t_r | T_{i,j} > t_i,\ i = 0, \ldots, r-1\} =$$

$$= \prod_{r=1}^{p} P\{T_{r,j} > t_r | \tau_i > t_i,\ i = 0, \ldots, r-1\}.$$

Consequently the desired result follows now by Condition (B.2) and Theorem 2.5 applied to the conditional pairs $[\{T_{r,j}, L_{r,j}\} | \tau_i > t_i, i = 0, \ldots, r-1]$ $r = 1, \ldots, p$. ||

References

(1) Barlow, R. E. (1968). Likelihood ratio tests for restricted families. Ann. Math. Statist. 39, 547-560.

(2) Barlow, R. E. and Doksum, K. A. (1972), Isotonic tests for convex ordering. Proc. 6th Berk. Symp. 1, 293-323.

(3) Bickel, P. J. (1969). Tests for monotone failure rate II. Ann. Math. Statist. 40, 1250-1260.

(4) Bickel, P. J. and Doksum, K. A. (1969). Tests for monotone failure rate based on normalized spacings. Ann. Math. Statist. 40, 1216-1235..

(5) Billard, L. Lacayo, H. and Langberg, N. A. (1979). The symmetric m-dimensional simple epidemic process. J. R. Statist. Sco., series B, Vol. 41, 196-202.

(6) Chen, Y. Y., Hollander, M., and Langberg, N. A. (1980a). Testing whether new is better than used with randomly censored data. Florida State University Technical Report M-503.

(7) Chen, Y. Y., Hollander, M., and Langberg, N. A. (1980b). Testing for monotone mean residual life distribution with randomly censored data. Florida State University, Technical Report M-550.

(8) Hoel, David G. (1972). A representation of mortality data by competing risks. Biometrics 28, 475-488.

(9) Hollander, M. and Proschan, F. (1972). Testing whether new is better than used. Ann. Math. Statist. 43, 1136-1146.

(10) Hollander, M., and Proschan, F. (1975). Testing for the mean residual life. Biometrika 62, 586-593.

(11) Hollander, M. and Proschan, F. (1979). Testing to determine the underlying distribution using randomly censored data. Biometrics 35, 393-401.

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

(13) Koul, H. L. (1977). A test for new better than used. Commun. Statist-Theo. Meth. A6, 563-573.

(14) Koul, H. L. (1978a). A class of tests for testing new is better than used. Canadian Jour. Statist. V6 No. 2. 240-271.

(15) Koul, H. L. (1978b). Testing for new better than used in expectation. Commun. Statist.-Theor. Meth. A7, 685-701.

(16) Koul, H. L. and Susarla, V. (1978). Testing for new better than used In expectation with incomplete data. Michigan State University, Technical Report No. RM-386.

(17) Lacayo, H. and Langberg, N. A. (1980a). The exact and asymptotic formulas for the state probabilities in simple epidemics with m kinds of susceptibles. Florida State University, Technical Report M-530.

(18) Lacayo, H. and Langberg, N. A. (1980b). The convergence of the state probabilities in a class of m-dimensional simple epidemic models. Florida State University, Technical Report M(R)-464

(19) Langberg, N. A. (1980). The construction and asymptotic behavior of some m-dimensional simple epidemic models. Florida State University, Technical Report M-485.

(20) Langberg, N. A. Proschan, F. and Quinzi, A. J. (1978). Converting dependent models into independent ones preserving essential features. Ann. Prob. Vol. 6, No. 1, 174-181.

(21) Langberg, N. A., Proschan, F. and Quinzi, A. J. (1980). Estimation of dependent lifelengths with application to the theory of competing risks. Ann. Statist. Vol. 8, No. 6.

(22) Langberg, N. A., and Shaked, M. (1980a). Estimation of bivariate life distributions in the presence of censoring.

(23) Langberg, N. A. and Shaked, M. (1980b). The weak convergence of a bivariate Product Limit Estimator.

(24) Lee, L. and Thompson, W. A. Jr. (1976). Failure rate - a unified approach. J. Appl. Prob. 13, 176-182.

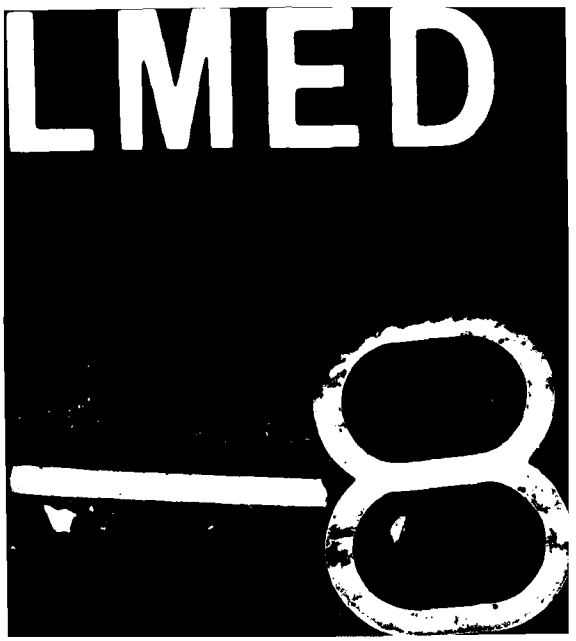
(25) Marshall, A. W. and Olkin, I. (1967). A multivariate exponential distribution. J. Amer Statist. Assoc., 62 30-44.

(26) Miller, D. R. (1977). A note on independence of multivariate lifetimes in competing risks models. Ann. Statist. 5, 576-579.

(27) Peterson, A. V. Jr. (1975). Nonparametric estimation in the competing risks problem. Stanford University Technical Report No. 13.

(28) Proschan, F., and Pyke, R. (1967). Tests for monotone failure rate. Proc. 5th Berk. Symp. 3, 293-312.

(29) Tsiatis, A. (1975). A nonidentifiability aspect of the problem of competing risks. Proc. Nat. Acad. Sci. USA. 72.


LMED