	1.0		2.8	2.5
			3.2	2.2
			3.6	
			4.0	2.0
	1.1			
				1.8
	1.25	1.4		1.6

MICROCOPY RESOLUTION TEST CHART

Unclassified

REPORT DOCUMENTATION PAGE	READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER AFOSR-TR. 87-1094	2 GOVT ACCESSION NO	3. RECIPIENT'S CATALOG NUMBER

AD-A186 038

4. TITLE (and Subtitle) Some new approaches to multivariate probability distributions	5. TYPE OF REPORT & PERIOD COVERED Journal Technical - December 1986
	6. PERFORMING ORG. REPORT NUMBER 86-44

7. AUTHOR(s) D.N. Shanbhag and S. Kotz	8. CONTRACT OR GRANT NUMBER(s) F49620-85-C-0008

9. PERFORMING ORGANIZATION NAME AND ADDRESS Center for Multivariate Analysis Fifth Floor Thackeray Hall University of Pittsburgh, Pittsburgh, PA 15260	10. PROGRAM ELEMENT. PROJECT, TASK AREA & WORK UNIT NUMBERS 6.1102F 2304 A5

11. CONTROLLING OFFICE NAME AND ADDRESS Air Force Office of Scientific Research Department of the Air Force Bolling Air Force Base, DC 20332	12. REPORT DATE December 1986
	13. NUMBER OF PAGES 35

14. MONITORING AGENCY NAME & ADDRESS(If different from Controlling Office)	15. SECURITY CLASS. (of this report) Unclassified
	15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

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

18. SUPPLEMENTARY NOTES

19. Key words and phrases: Expected remaining life function, hazard measure, reliability, representation and stability theorems, identifiability, multivariate probability distributions.

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

In this paper we extend and generalize to the multivariate set-up our earlier investigations related to expected remaining life functions and general hazard measures including representations and stability theorems for arbitrary probability distributions in terms of these con-

DD FORM 1473
1 JAN 73

Unclassified

20

cepts.   (The univariate case is discussed in detail in Kotz and Shanbhag,

Advances in Applied Probability 12 (1980).)

SOME NEW APPROACHES TO
MULTIVARIATE PROBABILITY DISTRIBUTIONS

D. N. Shanbhag*
University of Sheffield, U.K.

and

S. Kotz
University of Maryland, U.S.A.

Center for Multivariate Analysis

University of Pittsburgh

SOME NEW APPROACHES TO

MULTIVARIATE PROBABILITY DISTRIBUTIONS

D. N. Shanbhag*
University of Sheffield, U.K.

and

S. Kotz
University of Maryland, U.S.A.

December 1986

Technical Report No. 86-44

Accession For	
NTIS GRA&I	☒
DTIC TAB	☐
Unannounced	☐
Justification	
By	
Distribution/	
Availability Codes	
Dist	Avail and/or Special
A-1	

Center for Multivariate Analysis
Fifth Floor Thackeray Hall
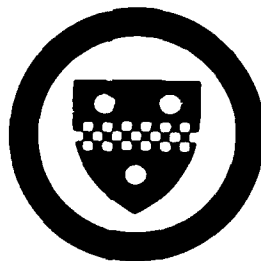University of Pittsburgh
Pittsburgh, PA 15260

SOME NEW APPROACHES TO

MULTIVARIATE PROBABILITY DISTRIBUTIONS

D. N. Shanbhag*
University of Sheffield, U.K.

and

S. Kotz
University of Maryland, U.S.A.

ABSTRACT

In this paper we extend and generalize to the multivariate set-up
our earlier investigations related to expected remaining life functions
and general hazard measures including representations and stability
theorems for arbitrary probability distributions in terms of these con-
cepts. (The univariate case is discussed in detail in Kotz and Shanbhag,
Advances in Applied Probability 12 (1980).)

AMS 1979 subject classification: 62H05, 62N05.

Key words and phrases: Expected remaining life function, hazard measure,
reliability, representation and stability theorems, identifiability,
multivariate probability distributions.

1. INTRODUCTION

In the earlier paper by the authors, Kotz and Shanbhag (1980) (to be referred to as KSh for brevity) presented a detailed discussion of new approaches to univariate probability distributions. We concentrated on representations and characterizations of probability distribution functions in terms of conditional expectations (specifically in terms of the expected remaining life function - e.r.l. function) and in terms of hazard measures.

In the course of our investigations, we succeeded in extending, generalizing and simplifying a number of results dealing with e.r.l. functions and hazard measures which have appeared in the literature of the last two decades. We also presented some convergence theorems which shed light on the structure of e.r.l. functions, hazard measures and distribution functions in both the continuous and discrete cases (but not restricted to these cases only).

In many instances of practical applications, requiring model building, there are indications of such results being of special potential importance.

The present paper is structured along the lines of KSh (1980) but is an initial attempt towards studying more subtle and difficult problems of multivariate distributions. In this paper, we shall attempt to unify, extend, generalize and simplify results scattered in the literature related to structures of multivariate distributions (in particular but not exclusively of a non-absolutely continuous nature), of various definitions of hazard measures. (Unlike the univariate case there is no unique definition of this concept in the multivariate case in the literature.) Among other results, an over-compassing generalization of the scalar multivariate hazard measure is given and an overall

structure as well as certain convexity properties and their implications related to this measure are revealed. In addition, we define and investigate multivariate analogues and extensions of e.r.l. functions and trace their relations, first to the multivariate probability distribution functions and then to the corresponding univariate concept on the one hand, as well as to (various generalizations of) multivariate hazard measures on the other. Following the approach adopted in KSh (1980) for the univariate case, we do not restrict ourselves necessarily to non-negative random variables. (The notions of the hazard measure as well as that of the e.r.l. functions in the literature are usually limited to the non-negative case.)

Most of the groundwork as far as the convergence and representation theorems is concerned has been laid in KSh (1980). However, in the present paper we clarify, using examples of specific distributions, some ambiguities and certain inconsistencies related to the structure of various characteristics of multivariate distributions in our search for the most meaningful and practically attractive expressions and representations of these distributions which would expose the hidden dependencies among jointly distributed random variables. These findings could prove to be of some significance in future developments at least in areas such as reliability and pattern recognition.

2. A GENERALIZED MULTIVARIATE HAZARD GRADIENT AND A MULTIVARIATE GENERALIZATION OF THE e.r.l. FUNCTION

In this section, we shall give, among other things, two theorems that follow as direct corollaries of KSh (1980). These concern respectively a generalized multivariate hazard gradient and an analogous

multivariate generalization of the e.r.l. function.

For multivariate distributions, there exist in the literature basically two approaches to defining hazard functions, both confined predominantly to absolutely continuous distributions on Euclidean spaces.

The first definition, adopted and analyzed by, among others, Basu (1971) and Puri and Rubin (1973), is a straightforward extension of the univariate concept. (A purely discrete case was also considered by Puri and Rubin (1973).) The hazard function of a random vector $\underset{\sim}{X} = (X_1,\ldots,X_p)$ is defined in this case to be a real-valued function r on $\{x: \bar{F}(\underset{\sim}{x}) > 0\}$ with values

$$r(\underset{\sim}{x}) = f(\underset{\sim}{x})/\bar{F}(\underset{\sim}{x}),$$

where $\underset{\sim}{x} = (x_1,\ldots,x_p) \in R^p$, $f(\underset{\sim}{x})$ is the probability density function, and $\bar{F}(\underset{\sim}{x})$ is the survivor function given by

$$\bar{F}(\underset{\sim}{x}) = P(\underset{\sim}{X} \geq \underset{\sim}{x}).$$

(Here as well as in what follows the inequalities for vectors are to be understood componentwise.) This concept was further discussed by Block (1977) where additional closely related variants were proposed, and treated in a somewhat more unified manner in Galambos and Kotz (1978). We intend to generalize this definition and examine it in greater detail. However, since our contribution in this case is to be rather substantial without relying very heavily on KSh (1980), we shall deal with it separately in the next section (i.e., Section 3 of the paper).

The second approach, due to Johnson and Kotz (1975a) and Marshall (1975), defines a multivariate hazard gradient (in an absolutely continuous case) as the vector-valued function h on $\{x: \bar{F}(\underset{\sim}{x}) > 0\}$ with values

$$h(\underset{\sim}{x}) = (-\frac{\partial}{\partial x_1},\ldots,-\frac{\partial}{\partial x_p})\log \bar{F}(\underset{\sim}{x})$$
$$= -\text{grad} \log \bar{F}(\underset{\sim}{x})$$

(except for a set of Lebesgue measure zero). As was shown by Marshall
(1975) in the absolutely continuous case, the vector-valued h uniquely
determines the probability distribution function (d.f.) or equivalently
the survivor function. Note that each one of the components of $h(x)$
depends in general on all the variables $x_i(i = 1,2,...,p)$. In the first
part of this section (i.e. in part a) we shall generalize the gradient
h to the case of arbitrary d.f.'s and at the same time reduce some re-
dundancy existing in the structure of the components of this gradient.
The main result involving a representation given in this part subsumes
Marshall's (1975) result and is essentially a corollary of Propositions
5 and 8 of KSh (1980).

In KSh (1980) - motivated by the remark contained in Shanbhag
(1970) and the results of Hamdan (1972), Kotlarski (1972), Shanbhag
and Bhaskara Rao (1975) and Gupta (1975) - we also extended the con-
cept of the e.r.l. function of a positive random variable to an arbi-
trary random variable and have given a representation for a probability
distribution in terms of this function. Some possibilities of the
applicability of the concept in practice have been indicated in KSh
(1980) and the references cited above. (Also, see Hall and Wellner
(1981), Hollander and Proschan (1984) and Glänzel et al (1984) for
further information and references on the e.r.l. function.) A variety
of multivariate generalizations of this function can of course be con-
structed. However, we intend in this case to deal only with a certain
construction that has features closely resembling those of the multi-
variate hazard function of the present section. The representation
theorem in this latter case follows as a corollary of KSh (1980). In
view of the prevailing analogy, we shall devote the second part of this

section (i.e. part b) to discussing this particular version of e.r.l.
functions and revealing some of its properties including the aforementioned
theorem. For a related but independently carried out investigation of
multivariate analogues of e.r.l. functions, the reader may wish to
consult Zahedi (1985). This work is however along different lines.

a. <u>A generalized hazard gradient and some of its basic properties</u>.

Let $p \geq 2$, F be a d.f. on R^p and $\underset{\sim}{X} = (X_1,X_2,\ldots,X_p)$ be a random vector
distributed according to this d.f. Let $\nu_F^{(i)}(\cdot|\underset{\sim}{x}_{(i+1)})$ with $\underset{\sim}{x}_{(i)} =$
$(x_i,x_{i+1},\ldots,x_p)$, $\underset{\sim}{x}_{(1)} = \underset{\sim}{x}$ denote the hazard measure on R^1 for the
conditional distribution of X_i given that $X_{i+1} \geq x_{i+1},\ldots,X_p \geq x_p$
(as stipulated in Section 4 of KSh (1980)) for every $\underset{\sim}{x}_{(i+1)} \in R^{p-i}$ and
$i = 1,2,\ldots,p-1$. (We define the conditional distribution to be arbitrary
for any conditioning set of measure zero.) Also, let $\nu_F^{(p)}(\cdot)$ denote
the corresponding hazard measure on R^1 for the marginal distribution
of X_p. Extending and modifying the definition of Johnson and Kotz (1975a)
and Marshall (1975), we call the family

$$\{\nu_F^{(i)}(\cdot|\underset{\sim}{x}_{(i+1)}):\underset{\sim}{x}_{(i+1)} \in R^{p-i}, \ i = 1,2,\ldots,p-1\}, \ \nu_F^{(p)}(\cdot)$$

<u>the hazard gradient relative to the d.f.F</u>. We have the following theorem
which is essentially a corollary of Propositions 5 and 8 of KSh (1980)
(see, also, Cox (1972)):

THEOREM 1. The survivor function corresponding to F is represented by

$$\overline{F}(\underset{\sim}{x}) = P(\underset{\sim}{X} \geq \underset{\sim}{x}) = \prod_{i=1}^{p} \{[\prod_{y_i \in D_i(\underset{\sim}{x}_{(i)})} \{1-\nu_F^{(i)}(\{y_i\}|\underset{\sim}{x}_{(i+1)})\}]$$

$$\times \exp[-\nu_F^{(c,i)}((-\infty,x_i]|\underset{\sim}{x}_{(i+1)})]\}, \ \underset{\sim}{x} \in R^p \qquad (2.1)$$

and for a continuous F the representation is

$$\overline{F}(x) = \exp\{-\sum_{i=1}^{p} \nu_F^{(i)}((-\infty,x_i]|\underset{\sim}{x}_{(i+1)})\}, \; \underset{\sim}{x} \in R^p, \tag{2.2}$$

where the notation $\nu_F^{(p)}(\cdot|\underset{\sim}{x}_{(p+1)})$ is used for convenience to denote $\nu_F^{(p)}(\cdot)$, $e^{-\infty}$ is defined to be zero, $D_i(\underset{\sim}{x}_{(i)})$ is the set of real points $y_i < x_i$ at which $\nu_F^{(i)}(\{y_i\}|\underset{\sim}{x}_{(i+1)})$ is positive, and $\nu_F^{(c,i)}(\cdot|\underset{\sim}{x}_{(i+1)})$ the continuous (non-atomic) part of $\nu_F^{(i)}(\cdot|\underset{\sim}{x}_{(i+1)})$. Furthermore, if F is continuous and $\{F_n: n=1,2,\dots, \}$ is a sequence of d.f.'s on R^p, then using the same notation

$$\nu_{F_n}^{(i)}((-\infty,x_i]|\underset{\sim}{x}_{(i+1)}) \rightarrow \nu_F^{(i)}((-\infty,x_i]|\underset{\sim}{x}_{(i+1)}) \tag{2.3}$$

for each $\underset{\sim}{x}$ such that $\overline{F}(\underset{\sim}{x}) > 0$ and $i = 1,2,\dots,p$ if and only if $\{F_n\}$ converges to F.

Proof. (2.1) and, if F is continuous, (2.2) follow immediately from Proposition 5 of KSn (1980) in view of the relation

$$P(\underset{\sim}{X} \geq \underset{\sim}{x}) = P(X_p \geq x_p)\prod_{i=1}^{p-1} P(X_i \geq x_i|X_{i+1} \geq x_{i+1},\dots,X_p \geq x_p),$$

$$\underset{\sim}{x} \in R^p. \tag{2.4}$$

If F is continuous, then the marginal distribution function of X_p is continuous and for every $\underset{\sim}{x}$ such that $\overline{F}(\underset{\sim}{x}) > 0$ and $i = 1,2,\dots,p-1$, the conditional distribution of X_i given $X_{i+1} \geq x_{i+1},\dots,X_p \geq x_p$ is continuous. Also, if $\underset{\sim}{x}^{(n)} = (X_1^{(n)},\dots,X_p^{(n)})$ for each $n \geq 1$ is a random vector distributed according to F_n, then for each $n \geq 1$

$$P(\underset{\sim}{x}^{(n)} \geq \underset{\sim}{x}) = P(X_p^{(n)} \geq x_p)\prod_{i=1}^{p-1} P(X_i^{(n)} \geq x_i|X_{i+1}^{(n)} \geq x_{i+1},\dots,X_p^{(n)} \geq x_p),$$

$$\underset{\sim}{x} \in R^p. \tag{2.5}$$

Applying Proposition 8 of KSh (1980) to the survivor functions on the
r.h.s. of (2.5), it can be easily verified that the convergence part
of the theorem is valid.

Remark 1.

For _absolutely_ continuous distributions, representation (2.2)
reduces to that of Marshall (1975). Both (2.2) and (2.1) are thus
extensions of Marshall's hazard gradient representation. Moreover,
the general representation for purely discrete distributions follows
from (2.1) in the obvious manner.

Remark 2.

The "convergence" part of Theorem 1 fails to be valid if the
assumption of continuity of F is omitted. Examples 1-3 presented in
KSh (1980) following Proposition 8 in Section 4 are sufficient to
illustrate this situation.

Remark 3.

The hazard gradient obviously has other versions when the ordering
of the variables is altered. Under a specific situation, one may find
a particular version to be the most natural and easiest to handle. In
that case, we shall of course consider the corresponding ordering to be
the one implied in our Theorem 1. A similar remark applies to the result
of Theorem 2 below.

Remark 4.

The following observation related to univariate hazard measures
may be appropriate at this point. (See also the beginning of Section 4
of KSh (1980).) If G is a d.f. on R^1, then according to representation

(4.1) in KSh (1980) either

$$\prod_{x_r \in D}(1-\nu_G(\{x_r\})) = 0 \text{ or } H_c(\infty) = \infty,$$

where ν_G is the hazard measure corresponding to G, D is the set of dis-
continuities of ν_G and $H_c(x) = \nu_G^{(c)}((-\infty,x])$, $\nu_G^{(c)}$ being the continuous

part of ν_G. Whenever the right extremity of G is not one of its dis-
continuity points, we have $\nu_G(\{x_r\}) < 1$ for all $x_r \in D$. Now the Borel

zero-one law and relation (16) given in Burrill (1972), p.245, imply

that $\prod_{x_r \in D}(1-\nu_G(\{x_r\}) = 0$ if and only if $\sum_{x_r \in D}\nu_G(\{x_r\}) = \infty$ provided

$\nu_G(\{x_r\}) < 1$, $x_r \in D$. This leads us to the relation

$$\nu_G((-\infty,\infty)) = \sum_{x_r \in D}\nu_G(\{x_r\}) + H_c(\infty) = \infty \tag{2.6}$$

whenever the right extremity of G is not one of its discontinuity

points. (This result was obtained earlier by Shanbhag (1979) using

a somewhat different argument.)

Remark 5.

As a corollary of Theorem 1, it follows that the components of $\underset{\sim}{X}$

are independent if and only if there exists a version of the hazard

gradient of F such that $\nu_F^{(i)}(\cdot|x_{\sim(i+1)})$ is independent of $x_{\sim(i+1)}$ for

each i = 1,2,...,p-1. The theorem also yields several other interesting

corollaries. In particular, since the theorem also implies that every

distribution on R^p is characterized by its hazard gradient, one could

obviously use it to give further characterizations of distributions,

such as the Marshall-Olkin bivariate distribution or Fréchet's multi-

variate distribution with continuous marginals or a multivariate Pareto

distribution, for which the hazard gradients are of a particularly

appealing form.

b. <u>The generalized e.r.l. function and some relevant comments.</u>

In view of Proposition 3 of KSh (1980), (2.4) in the proof of Theorem 1 above implies that under some mild assumptions there exists a representation for the survivor function of every p-component random vector $X = (X_1,\ldots,X_p)$ in terms of the conditional expectations $E\{h_i(X_i)|X_i \geq x_i,\ldots,X_p \geq x_p\}$, $(x_i,\ldots,x_p) \in R^{p-i+1}$ of monotone transforms h_i, $i = 1,2,\ldots,p$. This is given by the following Theorem 2. The theorem yields, among other things, that if X is a random vector with $E\{X_i^+\} < \infty$ for all $i = 1,2,\ldots,p$ (where $X_i^+ = \max\{0,X_i\}$), then the conditional expectations $E\{X_i-x_i|X_i \geq x_i,\ldots,X_p \geq x_p\}$, $i = 1,2,\ldots,p$, $x(= (x_1,\ldots,x_p)) \in R^p$ (and hence $E\{X-x|X \geq x\}$, $x \in R^p$) characterize the distribution of X; the representation in this latter case is also obvious now. Since the family of expectations $\{E\{X_i-x_i|X_i \geq x_i,\ldots,X_p \geq x_p\}$: $i = 1,2,\ldots,p$, $x \in R^p\}$ avoids some of the redundancies existing in the function $E\{X-x|X \geq x\}$, $x \in R^p$ and has all the obvious requirements of an e.r.l. function, it would be reasonable to adopt it to be the e.r.l. function of a multivariate probability distribution on R^p.

THEOREM 2. (A representation theorem). Let $X = (X_1,\ldots,X_p)$ be a random vector with p components and h_i, $i = 1,2,\ldots,p$ be real-valued non-decreasing functions on the real line such that $E\{h_i^+(X_i)\} < \infty$ for all $i = 1,2,\ldots,p$ (where $h_i^+(X_i) = \max\{0,h_i(X_i)\}$). If h_i, $i = 1,2,\ldots,p$, are such that $h_i(x_i) < E\{h_i(X_i)|X_i > x_i, X_{i+1} \geq x_{i+1},\ldots,X_p \geq x_p\}$ whenever $P\{X_i > x_i, X_{i+1} \geq x_{i+1},\ldots,X_p \geq x_p\} > 0$, then the survivor function corresponding to X is given by

$$P(X \geq x) = G(x), \qquad x(= (x_1,\ldots,x_p)) \in R^p, \qquad (2.7)$$

where G is the left continuous function satisfying

$$
G(\underset{\sim}{x}) = \begin{cases}
0 \text{ if } x_j > b_j^* \text{ for some } j \geq 1 \ \& \ \leq p \\
\\
\prod_{i=1}^{p} \lim_{y_i \to -\infty} \frac{g_i(y_i, \underset{\sim}{x}_{(i+1)})}{g_i(\underset{\sim}{x}_{(i)})} \{ \prod_{z \in D_{y_i,x_i}^{(i)}} g_i^*(z, \underset{\sim}{x}_{(i+1)}) \} \exp\{-\int_{y_i}^{x_i} \frac{dh_i^{(c)}(z)}{g_i(z, \underset{\sim}{x}_{(i+1)})}\} \\
\\
\hspace{4cm} \text{if } x_j < b_j^* \text{ for all } j \geq 1 \ \& \ \leq p \hspace{2cm} (2.8)
\end{cases}
$$

in which $D_{y_i,x_i}^{(i)}$ denotes the set of discontinuity points of h_i in $[y_i, x_i)$, $h_i^{(c)}$ denotes the continuous part of h_i (i.e. of its right contin-uous version), $x_{(i)} = (x_i, \ldots, x_p)$,

$$
g_i(\underset{\sim}{x}_{(i)}) = E\{h_i(X_i) | \underset{\sim}{X}_{(i)} \geq \underset{\sim}{x}_{(i)}\} - h_i(x_i-),
$$

$$
g_i^*(z, \underset{\sim}{x}_{(i+1)}) = \left\{ \frac{g_i(z, \underset{\sim}{x}_{(i+1)}) - (h_i(z) - h_i(z-))}{g_i(z+, \underset{\sim}{x}_{(i+1)}) + (h_i(z+) - h_i(z))} \right\} \frac{g_i(z+, \underset{\sim}{x}_{(i+1)})}{g_i(z, \underset{\sim}{x}_{(i+1)})} , \hspace{1cm} (2.9)
$$

and

$$
b_i^* = \begin{cases}
\infty \text{ if } \{y: \lim_{x_i \uparrow y} E\{h_i(X_i)|\underset{\sim}{X}_{(i)} \geq \underset{\sim}{x}_{(i)}\} \text{ exists and } \leq h_i(y)\} \text{ is empty} \\
\\
\inf\{y: \lim_{x_i \uparrow y} E\{h_i(X_i)|\underset{\sim}{X}_{(i)} \geq \underset{\sim}{x}_{(i)}\} \text{ exists and } \leq h_i(y)\} \text{ otherwise}
\end{cases}
$$

with $\underset{\sim}{X}_{(i)} = (X_i, \ldots, X_p)$.

(The conditional expectations are defined arbitrarily when the conditioning sets are of measure zero; also (2.8) and (2.9) in the statement above are to be read without $x_{(i+1)}$ in the case of $i = p$.)

Remark 6.

In view of Theorem 3 and the information given in the Remarks in Section 3 of KSh (1980), it is possible to present several extensions and variants of Theorem 2 given above.

Remark 7.

If h_i's in Theorem 2 are assumed additionally to be continuous, then the representation (2.7) with G_i's given by (2.8) without the term $\{ \prod_{z \in D^{(i)}_{y_i, x_i}} g_i^*(z, x_{(i+1)}) \}$ and with $h_i^{(c)}$'s replaced by h_i's is valid.

Remark 8.

If $h_i (i = 1, 2, \ldots, p)$ of Theorem 2 are taken as strictly increasing, the representation (2.7) for a survivor function is obviously valid in the case of every distribution satisfying the integrability condition of the theorem. One may be interested in seeing whether there exists a representation for the survivor function for X in terms of the conditional expectations corresponding to a fewer number of functions, which are appealing in some sense, at least when the domains of the definition of h_i are taken as Euclidean spaces with $h_i(X_i)$ considered above replaced by $h_i(X^{(i)})$, $X^{(i)}$ being a subvector of X. However, it is not difficult to see that in general merely with the integrability condition such a representation does not exist. This could be verified by noting, for example, that if h_i, $i = 1, 2, \ldots, p-1$ are given to be real-valued Borel

measurable functions on R^p, then there exist random vectors $\underset{\sim}{X}$ and $\underset{\sim}{Y}$ with distinct distributions having a common support (such as $\{(0,\ldots,0),$ $(1,0,\ldots,0), \ldots, (0,\ldots,0,1)\}$) such that

$$E\{h_i(\underset{\sim}{X})|\underset{\sim}{X} \geq \underset{\sim}{x}\} = E\{h_i(\underset{\sim}{Y})|\underset{\sim}{Y} \geq \underset{\sim}{x}\} \text{ for all } \underset{\sim}{x} \in R^p \text{ and } i = 1,2,\ldots,p-1.$$

Remark 9.

Prakasa Rao (1974) has essentially attempted to solve under some constraints the problem mentioned in Remark 8. He has given in this context a uniqueness theorem in the bivariate case under certain assumptions. The following example shows that the theorem is not valid.

EXAMPLE 1. Define h to be a real-valued function on R^2 such that

$$h(x,y) = (1-e^{-x^2})\xi(y), \quad x,y \in R^1,$$

where

$$\xi(y) = \begin{cases} c & \text{if } y \leq 1 \\ c + \dfrac{(y-1)^3}{6} & \text{if } 1 < y \leq 2 \\ c + (y-2) + \dfrac{(3-y)^3}{6} & \text{if } 2 < y \leq 4 \\ c + 2 + \dfrac{(y-5)^3}{6} & \text{if } 4 < y \leq 5 \\ c + 2 & \text{if } y > 5, \end{cases}$$

where c is a positive number. Alternatively, one could consider the h with a slightly more trivial situation of $\xi \equiv c$ for $c \neq 0$. Let (X,Y) and (Z,W) be random vectors with absolutely continuous independent nonnegative components such that X and Z are identically distributed but the distributions of Y and W are not identical. Also assume the random vectors to be such that their marginal distributions have all left extremities to be equal to zero and

$$P(Y \leq 1) = P(W \leq 1), \quad P(Y \leq y|Y \geq 1) = P(W \leq y|W \geq 1) \text{ for all } y \geq 1.$$

Observe that all the assumptions in Theorem 2.1 of Prakasa Rao (1974) are satisfied with $x_0 = y_0 = 0$. Moreover (X,Y) and (Z,W) satisfy Prakasa Rao's stipulation (2.0). However, in this case, the conclusions of the theorem are not valid. (It is obviously possible to illustrate this point by other examples of a similar nature.)

Remark 10.

In view of Theorem 2, characterizations based on e.r.l. functions are now obvious for the well known distributions such as the Marshall-Olkin bivariate distribution, the Farlie-Gumbel-Morgenstern distribution discussed in Johnson and Kotz (1975b), Gumbel's bivariate exponential distribution, the multivariate Pareto distribution and several other multivariate distributions appearing in Johnson and Kotz (1972). One could also apply the theorem to arrive at further characterizations based on conditional expectations for distributions such as Fréchet's and those discussed by Krishnaiah (1977). The following example may serve as an illustration of this point.

EXAMPLE 2. (Fréchet's bivariate continuous distribution).

Consider F to be the continuous d.f. on R^2 such that the corresponding survivor function is given by

$$\overline{F}(x_1,x_2) = \min\{1-F_1(x_1), 1-F_2(x_2)\}, \quad (x_1,x_2) \in R^2$$

with F_1 and F_2 as univariate d.f.'s. Clearly, since F is assumed to be continuous, we require F_1 and F_2 to be continuous here also. Define

$$h_i(x_i) = (F_i(x_i))^{\alpha_i}, \quad x_i \in R^1, \quad i = 1,2,$$

where $0 < \alpha_i < \infty$ and fixed. Then it follows that if $\underset{\sim}{X} = (X_1,X_2)$ is a random vector with d.f. F, we have for every $\underset{\sim}{x}(=(x_i,x_2)) \in R^2$ and $i = 1,2$

$$E \{h_i(X_i) | \underset{\sim}{X} \geq \underset{\sim}{x}\}$$

$$= \begin{cases} \dfrac{1}{1+\alpha_i} \{1-(G(\underset{\sim}{x}))^{1+\alpha_i}\}\{1-G(\underset{\sim}{x})\}^{-1} & \text{if } G(\underset{\sim}{x}) < 1 \\ \\ 1 & \text{if } G(\underset{\sim}{x}) = 1, \end{cases}$$

where $G(\underset{\sim}{x}) = \max\{F_1(x_1), F_2(x_2)\}$. (On the set $\{G(\underset{\sim}{x}) = 1\}$, one could also define $E\{h_i(X_i)|\underset{\sim}{X} \geq \underset{\sim}{x}\}$ differently.) Obviously, given α_i and F_i, $\{E(h_1(X_1)|X_1 \geq x_1, X_2 \geq x_2), E(h_2(X_2)|X_2 \geq x_2): (x_1,x_2) = R^2\}$ characterizes the distribution considered above among all bivariate distributions. (This distribution has several other interesting characterization properties also, the recent characterization based on discretized Shannon entropy given in Bertoluzza and Forte (1985) being one of these.)

3. EXTENDED VERSIONS OF THE RESULTS OF BASU AND PURI AND RUBIN DEALING WITH THE HAZARD FUNCTION

We shall now discuss a rather substantial generalization of what is known in the literature as the "scalar" multivariate hazard function.

Let, as in the previous section, F be a d.f. on R^p, $\underset{\sim}{X}$ be a p-component random vector with this distribution and $\bar{F}$ be the corresponding survivor function. Denote by P_F the measure determined by F on (the Borel σ-field $\mathcal{B}_p$ of) R^p. Since, in the multivariate case, we can have an F such that $P_F\{\underset{\sim}{x}: \bar{F}(\underset{\sim}{x}) = 0\} > 0$, (e.g., if we take F to be continuous such that $P_F\{\underset{\sim}{x}(=(x_1,\ldots,x_n)): x_1 = -x_2\} = 1$, we obtain $P_F\{\underset{\sim}{x}: \bar{F}(\underset{\sim}{x}) = 0\} = 1$), the definition of a hazard measure in KSh (1980) is not extendable as it stands. However, if we restrict ourselves only to the set $\mathbf{C}$ (say) of distributions F for which $\bar{F}(\cdot) > 0$ almost surely $[P_F]$, the definition

in KSh (1980) of a hazard measure admits an obvious extension. Suppose then that $F \in C$ and define ν_F to be the scalar hazard measure on R^p given by

$$\nu_F(B) = \int_B \frac{1}{\overline{F}(\underset{\sim}{x})} \, dP_F(x) \text{ for all } B \in B_p. \tag{3.1}$$

The integral on the r.h.s. of the equation can be written following the accepted convention in the literature as $\int_B \frac{1}{\overline{F}(\underset{\sim}{x})} \, dF(\underset{\sim}{x})$.

In the case when F is an absolutely continuous d.f. with respect to the Lebesgue measure on R^p, ν_F also possesses this property and thus the Radon-Nikodym derivative becomes the hazard function, studied by earlier authors, a.e. on $\{\underset{\sim}{x}: \overline{F}(\underset{\sim}{x}) > 0\}$. It follows from the investigations of Basu (1971) and Puri and Rubin (1973) (see also, Galambos and Kotz (1978)) that the measure ν_F does not in general determine uniquely the distribution F. Consider then $\mathcal{D}_F$ to be the set of d.f.'s on R^p that are members of C having the same scalar hazard measure as F. Clearly the set $\mathcal{D}_F$ defined herein is convex although not necessarily closed relative to weak convergence.

Consider now the set of all d.f.'s on the compactified Euclidean space $[-\infty,\infty]^p$. There exists a normed linear space of which this is a compact subset with the corresponding relative metric as a metric of weak convergence. Then, as a further subset of this compact set, the closure $\overline{\mathcal{D}}_F$ of the set $\mathcal{D}_F$ is also compact. (For simplicity we abuse the notation slightly here and elsewhere in this section by denoting the set of all d.f.'s on $[-\infty,\infty]^p$ which are extensions of members of $\mathcal{D}_F$ also by $\mathcal{D}_F$.) Since $\overline{\mathcal{D}}_F$ is also convex, Choquet's theorem

(cf. Phelps (1965) page 19 and also Kendall (1963)) implies that each $F^* \in \mathcal{P}_F$ can be represented as the centroid or barycenter of a probability measure on the Borel σ-field of the linear space which is concentrated on the set of extreme points of $\bar{\mathcal{P}}_F$. In general, the problem of obtaining the extreme points of $\bar{\mathcal{P}}_F$ or merely of $\mathcal{P}_F$ seems to be a difficult one and we have not as yet obtained any positive information in this connection. However, through a theorem and two corollaries to follow, we shall provide some valuable information concerning the problem of characterizing F on the basis of ν_F. This gives, among other things, the Poisson-Martin representation for F in terms of ν_F when F is continuous and a more natural extension of the univariate hazard measure to the multivariate case than the hazard gradient of the last section, possessing the uniqueness and stability requirements.

Before discussing our main results of this section, the following instructive examples making some specific points are worth revealing.

EXAMPLE 3. Let

$$F(\underset{\sim}{x}) = \prod_{i=1}^{p} F_i(x_i), \quad \underset{\sim}{x} = (x_1,\ldots,x_p) \in R^p,$$

where F_i are continuous d.f.'s on R^1. Then, appealing to the result of Puri and Rubin (1973) or our observation above concerning a representation for the members of $\bar{\mathcal{D}}_F$, we can easily see that each member $F^* \in \mathcal{D}_F$ has the following form:

$$F^*(\underset{\sim}{x}) = \int_{R^p} \prod_{i=1}^{p} \{1-(1-F_i(x_i))^{\lambda_i}\} dG(\underset{\sim}{\lambda}), \quad \underset{\sim}{x} \in R^p, \tag{3.2}$$

where G is a d.f. on R^p such that the corresponding measure is concen-

trated on the set $\{\underset{\sim}{\lambda}: \lambda_i > 0, \ i = 1,2,\ldots,p, \ \prod_{i=1}^{p} \lambda_i = 1\}$. Also, this

can be seen via the Poisson-Martin integral representation given for

the members of $\boldsymbol{\mathcal{D}}_F$ in Corollary 1 below. Incidentally, in the present

case, the extreme points of $\boldsymbol{\mathcal{D}}_F$ are given precisely by the d.f.'s F^* of

the form

$$F^*(\underset{\sim}{x}) = \prod_{i=1}^{p} \{1 - (F_i(x_i))^{\lambda_i}\}, \quad \underset{\sim}{x} \in R^p$$

with $\lambda_i > 0$, $i = 1,2,\ldots,p$ and $\prod_{i=1}^{p} \lambda_i = 1$ and any extreme point of $\overline{\boldsymbol{\mathcal{D}}}_F$

is either an extreme point of $\boldsymbol{\mathcal{D}}_F$ or a d.f. on $[-\infty,\infty]^p$ which is the

weak limit of a sequence of extreme points of $\boldsymbol{\mathcal{D}}_F$. Looking at an arbi-

trary member F^* given by (3.2) in the case of $p \geq 2$ for $\boldsymbol{\mathcal{D}}_F$, we observe

a curious property of $\boldsymbol{\mathcal{D}}_F$ that if $F^* \in \boldsymbol{\mathcal{D}}_F$ and any p-1 of the p univariate

marginals of F^* agree with those corresponding to F, then $F^* = F$. In

other words, we have in this case that if a d.f. on R^p has p-1 of its

univariate marginals precisely the same as those corresponding to F and

its scalar hazard measure on R^p is defined and is given by ν_F, then

this d.f. has to be F. Since every univariate d.f. is uniquely determin-

ed by its hazard measure, we could also restate this property using

only hazard measures. (For some recent advances connected with the re-

sults discussed herein, see Lau and Rao (1982), Rao and Shanbhag (1986)

and Davies and Shanbhag (1987).)

EXAMPLE 4. Let $p \geq 2$, k be a real number and S be a countable subset

of R^{p-1}. Also let $\boldsymbol{G}$ denote the set of d.f.'s on R^{p-1} that are concentrated

on S giving a positive probability mass to each point of S. For each

$G \in \boldsymbol{G}$, let F_G denote the d.f. on R^p which is concentrated on

$\{\underset{\sim}{x}: \underset{\sim}{x} \in R^p, \ \sum_{1}^{p} x_i = k\}$ with

$$F_G(x_1,\ldots,x_{p-1},\infty) = G(x_1,\ldots,x_{p-1}), \quad (x_1,\ldots,x_{p-1}) \in R^{p-1}$$

(in the usual notation). It is easily seen that here ν_F are all (well defined and) identical. If we now consider $p \geq 4$ and any of the F_G's to be F, then it is clearly seen that the condition that $F^* \in \mathcal{D}_F$ does not imply $F^* = F$ even if it is given that F^* has all of its univariate marginals or bivariate marginals to be the same as those of F. However, for the F in this example, the condition that $F^* \in \mathcal{D}_F$ together with

$$F^*(x_1,\ldots,x_{p-1},\infty) = F(x_1,\ldots,x_{p-1},\infty), \quad (x_1,\ldots,x_{p-1}) \in R^{p-1}$$

implies that $F^* = F$. Note also that here we have the set of extreme points of $\mathcal{D}_F$ to be empty and the set of extreme points of $\bar{\mathcal{D}}_F$ to be the closure (relative to weak convergence) of the set of the degenerate d.f.'s on $[-\infty,\infty]^p$ that are concentrated on $\{x: x \in R^p, \sum_1^p x_i = k\}$; clearly now the situation of *the last example that each* $F^* \in \mathcal{D}_F$ *has an integral* representation in terms of the extreme points of $\mathcal{D}_F$ is not valid.

In spite of certain isolated cases, such as that of Fréchet's distribution of Example 2 or of a d.f. F that satisfies for some $b \in R^p$ the conditions $F(b) = 1$ and $\bar{F}(b) = P_F(\{b\}) > 0$, in which the F is characterized by ν_F, it now follows that, in general, unless at least one of the (p-1)-variate marginals of the distribution (or something equivalent to it) is given, ν_F does not characterize F. One might then be interested to know whether F is characterized by ν_F given any one of the (p-1)-variate marginals. Our attempt to answer this question has been only partially successful so far and the findings of this investigation are presented, among other things, in the following results.

We are now ready to give our main theorem of the section together with the two of its interesting corollaries. (The reader can find some analogy between the proof of the theorem given here and Seneta's (1981)

proof of the Poisson-Martin integral representation theorem for a super regular vector corresponding to a non-negative matrix.)

THEOREM 3. If $F^* \in \mathcal{D}_F$ and, for each $i = 1,2,\ldots,p$, we have (in the standard notation)

$$F^*(x_1,\ldots,x_{i-1},\infty, x_{i+1},\ldots,x_p) = F(x_1,\ldots,x_{i-1},\infty, x_{i+1},\ldots,x_p)$$

$$\text{for all } x_i \in R^1, \; i = 1,2,\ldots,i-1,i+1,\ldots,p, \qquad (3.3)$$

then $F^* = F$. Furthermore, given an $F^* \in \overline{\mathcal{D}}_F$, there exists a probability measure μ^* on the set of all d.f.'s, G on $[-\infty,\infty]^p$, such that

$$F^*(\underset{\sim}{x}) = \int_{\mathcal{K}} G(\underset{\sim}{x})d\mu^*(G), \; \underset{\sim}{x} \in [-\infty,\infty]^p, \qquad (3.4)$$

where $\mu^*(\mathcal{K}) = 1$ and $\mathcal{K}$ is the closure (relative weak convergence) of the set of the d.f.'s $K_t(\cdot)$ for $\underset{\sim}{t}$ such that $F(\underset{\sim}{t})$, $\overline{F}(\underset{\sim}{t}) > 0$ ($\overline{F}$ being the survivor function of F as in the last section), where each of the $K_t(\cdot)$ is defined to be a d.f. on $[-\infty,\infty]^p$ such that it is the degenerate d.f. at $\underset{\sim}{t}$ if $P_F(\{\underset{\sim}{t}\}) = \overline{F}(\underset{\sim}{t})$ and the d.f. satisfying the following otherwise:

$$K_t(\underset{\sim}{x}) = \frac{k(\underset{\sim}{x},\underset{\sim}{t})}{k(\underset{\sim}{t},\underset{\sim}{t})} \; , \quad \underset{\sim}{x} \in (-\infty,\underset{\sim}{t}] \qquad (3.5)$$

with

$$k(\underset{\sim}{y},\underset{\sim}{t}) = \triangle_t(\underset{\sim}{y}) + \sum_{n=1}^{\infty} \int_{(-\infty,\underset{\sim}{y}]} \int_{[\underset{\sim}{y}_1,\underset{\sim}{t}]} \cdots \int_{[\underset{\sim}{y}_{n-1},\underset{\sim}{t}]} d\nu_F(\underset{\sim}{y}_n)\ldots d\nu_F(\underset{\sim}{y}_1), \; \underset{\sim}{y} \in (-\infty,\underset{\sim}{t}]$$

$\triangle_t(\cdot)$ being the d.f. degenerate at $\underset{\sim}{t}$. (The proof of the theorem asserts that $K_t(\cdot)$ is well defined.)

Proof. In view of Fubini's Theorem and relation $P_F(B) = \int_B \overline{F}(\underset{\sim}{x})d\nu_F(\underset{\sim}{x})$

for every Borel subset B of R^p, we have for each $\underset{\sim}{t}$ such that $\overline{F}(\underset{\sim}{t}) > P_F(\{\underset{\sim}{t}\})$, $n \geq 1$,

$$\int_{(-\infty,\underset{\sim}{t}]}\int_{(-\infty,\underset{\sim}{y}_n]}\cdots\int_{(-\infty,\underset{\sim}{y}_2]} d\nu_F(\underset{\sim}{y}_1)\ldots d\nu_F(\underset{\sim}{y}_n)$$

$$= \int_{(-\infty,\underset{\sim}{t}]}\int_{[\underset{\sim}{y}_1,\underset{\sim}{t}]}\cdots\int_{[\underset{\sim}{y}_{n-1},\underset{\sim}{t}]} d\nu_F(\underset{\sim}{y}_n)\ldots d\nu_F(\underset{\sim}{y}_1)$$

$$\leq \frac{1}{\overline{F}(\underset{\sim}{t})} \int_{(-\infty,\underset{\sim}{t}]}\int_{[\underset{\sim}{y}_1,\underset{\sim}{t}]}\cdots\int_{[\underset{\sim}{y}_{n-1},\underset{\sim}{t}]} \overline{F}(\underset{\sim}{y}_n)d\nu_F(\underset{\sim}{y}_n)\ldots d\nu_F(\underset{\sim}{y}_1)$$

$$\leq \frac{\alpha_{\underset{\sim}{t}}}{\overline{F}(\underset{\sim}{t})} \int_{(-\infty,\underset{\sim}{t}]}\int_{[\underset{\sim}{y}_1,\underset{\sim}{t}]}\cdots\int_{[\underset{\sim}{y}_{n-2},\underset{\sim}{t}]} \overline{F}(\underset{\sim}{y}_{n-1})d\nu_F(\underset{\sim}{y}_{n-1})\ldots d\nu_F(\underset{\sim}{y}_1)$$

$$\leq \frac{\alpha_{\underset{\sim}{t}}^{n-1}}{\overline{F}(\underset{\sim}{t})} \int_{(-\infty,\underset{\sim}{t}]} \overline{F}(\underset{\sim}{y}_1)d\nu_F(\underset{\sim}{y}_1) = \frac{\alpha_{\underset{\sim}{t}}^{n-1}}{\overline{F}(\underset{\sim}{t})} F(\underset{\sim}{t}), \tag{3.6}$$

where $\alpha_{\underset{\sim}{t}} = F(\underset{\sim}{t})/\{F(\underset{\sim}{t}) + \overline{F}(\underset{\sim}{t}) - P_F(\{\underset{\sim}{t}\})\} < 1$, since $\alpha_{\underset{\sim}{t}} > P_F([\underset{\sim}{y},\underset{\sim}{t}])/\overline{F}(\underset{\sim}{y})$ for $\underset{\sim}{y} \leq \underset{\sim}{t}$. (3.6) establishes, among other things, that $K_t(\cdot)$ in the statement of the theorem is well defined. Now, for each d.f. $\overline{F}^*$ on R^p such that $F^* \in \mathcal{D}_F$ and $\underset{\sim}{t}$ as in (3.6), we have, in view of relation $P_{F^*}(B) = \int_B \overline{F}^*(\underset{\sim}{x})d\nu_F(\underset{\sim}{x})$ with B as an arbitrary Borel set and $\overline{F}^*$ as the survivor function corresponding to F^*,

$$\overline{F}^*(\underset{\sim}{t}) = \xi_0(\underset{\sim}{t}) + (-1)^p P_{F^*}((-\infty,\underset{\sim}{t}))$$

$$= \xi_0(\underset{\sim}{t}) + (-1)^p \int_{(-\infty,\underset{\sim}{t})} \overline{F}^*(\underset{\sim}{x})d\nu_F(\underset{\sim}{x})$$

$$= \xi_1(\underset{\sim}{t}) + (-1)^{2p} \int_{(-\infty,\underset{\sim}{t})} P_{F^*}((-\infty,\underset{\sim}{x}))d\nu_F(\underset{\sim}{x})$$

$$= \xi_n(\underset{\sim}{t}) + (-1)^{(n+1)p} \int_{(-\infty,\underset{\sim}{t})} \int_{(-\infty,\underset{\sim}{x}_n)} \cdots \int_{(-\infty,\underset{\sim}{x}_2)} P_{F^\star}((-\infty,\underset{\sim}{x}_1))$$

$$d\nu_F(\underset{\sim}{x}_1)\ldots d\nu_F(\underset{\sim}{x}_n), \quad n \geq 1, \qquad (3.7)$$

where the sequence $\{\xi_m(\underset{\sim}{t}): m = 0,1,\ldots\}$ (for each given $\underset{\sim}{t}$) is such that it depends only on ν_F and d.f.'s $F^\star(x_1,\ldots,x_{i-1},\infty,x_{i+1},\ldots,x_p)$, $(x_1,\ldots,x_{i-1},x_{i+1},\ldots,x_p) \in R^{p-1}$, $i = 1,\ldots,p$. It follows trivially from (3.6) that the multiple integral on the r.h.s. of (3.7) tends to zero and $n \to \infty$. This in turn implies that the sequence $\{\xi_n(\underset{\sim}{t}): n = 1,2,\ldots\}$ in (3.7) converges to $\overline{F}^\star(\underset{\sim}{t})$ and hence we have that if (3.3) is valid, then

$$\overline{F}^\star(\underset{\sim}{t}) = \overline{F}(\underset{\sim}{t}) \text{ for each } \underset{\sim}{t} \text{ such that } \overline{F}(\underset{\sim}{t}) > P_F(\{\underset{\sim}{t}\}). \qquad (3.8)$$

In view of the left continuity of $\overline{F}$ and $\overline{F}^\star$ and the fact that $\{x: \underset{\sim}{x} \in R^p, \overline{F}(\underset{\sim}{x}) = 0\} = \{x: \underset{\sim}{x} \in R^p, \overline{F}^\star(\underset{\sim}{x}) = 0\}$, we can conclude that if (3.8) is valid, then we have $\overline{F}^\star = \overline{F}$ or equivalently $F^\star = F$. This establishes the first part of the theorem.

To establish the second part of the theorem, define

$$B = \{\underset{\sim}{t}: \underset{\sim}{t} \in R^p, F(\underset{\sim}{t}), \overline{F}(\underset{\sim}{t}) > 0\},$$

$$B_0 = \{\underset{\sim}{t}: \underset{\sim}{t} \in R^p, \overline{F}(\underset{\sim}{t}) = P_F(\{\underset{\sim}{t}\}) > 0\},$$

and $\quad B_m = \{\underset{\sim}{t}: \underset{\sim}{t} \in B, \overline{F}(\underset{\sim}{t}) \geq P_F(\{\underset{\sim}{t}\}) + \frac{1}{m}\}, \quad m = 1,2,\ldots .$

If $F^\star \in \mathcal{D}_F$, then by the monotone convergence theorem, we get

$$\overline{F}^\star(\underset{\sim}{x}) = \int_{[\underset{\sim}{x},\infty)} \overline{F}^\star(\underset{\sim}{y}) d\nu_F(\underset{\sim}{y})$$

$$= \int_{B_0} \overline{K}_t(\underset{\sim}{x}) dP_{F^\star}(\underset{\sim}{t}) + \lim_{m\to\infty} \int_{B_m} \overline{\Delta}_t(\underset{\sim}{x}) \overline{F}^\star(\underset{\sim}{t}) d\nu_F(\underset{\sim}{t}), \quad \underset{\sim}{x} \in R^p. \qquad (3.9)$$

Now, for every $m \geq 1$ and $\underset{\sim}{t} \in B_m$, $\alpha_{\underset{\sim}{t}}$ of (3.6) is bounded by $m/(m+1)$

and hence it follows from (3.6) that $k(\underset{\sim}{t},\underset{\sim}{t})$ is bounded on B_m for each

$m \geq 1$. Also, if we define $\overline{k}(\underset{\sim}{x},\underset{\sim}{t}) = k(\underset{\sim}{t},\underset{\sim}{t}) - k(\underset{\sim}{x}-,\underset{\sim}{t})$ for each $\underset{\sim}{t} \in B \cap B_0^c$

and $\underset{\sim}{x} \in R^p$ (with this to be zero if $\underset{\sim}{x} \notin (-\infty,\underset{\sim}{t}]$). Fubini's Theorem

implies that for each $m \geq 1$ and $\underset{\sim}{x} \in R^p$

$$\int_{B_m} \overline{k}(\underset{\sim}{x},\underset{\sim}{t})\overline{F}^*(\underset{\sim}{t})d\nu_F(\underset{\sim}{t})$$

$$= \int_{B_m} \overline{\Delta}_{\underset{\sim}{t}}(\underset{\sim}{x})\overline{F}^*(\underset{\sim}{t})d\nu_F(\underset{\sim}{t}) + \int_{B_m} \overline{k}(\underset{\sim}{x},\underset{\sim}{t})P_{F^*}(B_m(\underset{\sim}{t}))d\nu_F(\underset{\sim}{t}), \qquad (3.10)$$

where $B_m(\underset{\sim}{t}) = [\underset{\sim}{t},\infty) \cap B_m$. Observe that (3.10) follows easily from the

relations:

$$\overline{k}(\underset{\sim}{x},\underset{\sim}{t}) = \overline{\Delta}_{\underset{\sim}{t}}(\underset{\sim}{x}) + \sum_{n=1}^{\infty} \int_{[\underset{\sim}{x},\underset{\sim}{t}]}\int_{[\underset{\sim}{x},\underset{\sim}{y}_n]} \cdots \int_{[\underset{\sim}{x},\underset{\sim}{y}_2]} d\nu_F(\underset{\sim}{y}_1)\ldots d\nu_F(\underset{\sim}{y}_n),$$

$$\underset{\sim}{x} \leq \underset{\sim}{t}, \quad \underset{\sim}{t} \in B \cap B_0^c,$$

and

$$\int_{B_m(\underset{\sim}{y}_n)} \overline{F}^*(\underset{\sim}{t})d\nu_F(\underset{\sim}{t}) = P_{F^*}(B_m(\underset{\sim}{y}_n)), \quad \underset{\sim}{y}_n \in B \cap B_0^c, \quad m,n \geq 1.$$

From (3.9) and (3.10), it consequently follows that there exists a

sequence $\{\mu_m: m = 1,2,\ldots\}$ of measures on R^p such that $\mu_m(R^p) \leq 1$ for

all m and

$$\overline{F}^*(\underset{\sim}{x}) = \lim_{m \to \infty} \int_B \overline{k}(\underset{\sim}{x},\underset{\sim}{t})d\mu_m(\underset{\sim}{t}), \quad \underset{\sim}{x} \in R^p, \qquad (3.11)$$

which in turn implies that $\{\mu_m(B)\}$ converges to 1 and hence that there

exists a sequence $\{\mu_m\}$ of probability measures on R^p for which (3.11) is

valid. Since K is compact, using Parthasarathy's (1967) Theorem 6.4,

it can then be easily seen that there exists a probability measure μ^* on K such that

$$\overline{F}^*(\underset{\sim}{x}) = \int_K \overline{G}(\underset{\sim}{x})d\mu^*(G), \quad \underset{\sim}{x} \in R^p.$$

Since $\overline{\mathcal{D}}_F$ is the closure of $\mathcal{D}_F$, a further application of Parthasarathy's Theorem yields the validity of the second part of our theorem.

The following two corollaries of Theorem 3 are easy to prove:

COROLLARY 1. (The Poisson-Martin representation): If F is continuous, then we have a d.f. F^* on R^p to be a member of $\mathcal{D}_F$ if and only if it has a representation

$$F^*(\underset{\sim}{x}) = \int_{K \cap \mathcal{D}_F} G(\underset{\sim}{x})d\mu(G), \quad \underset{\sim}{x} \in R^p,$$

for some probability measure μ on $K \cap \mathcal{D}_F$. (In the present case, we also have $K \cap \mathcal{D}_F$ to be a G_δ set of the space of all d.f.'s on $[-\infty,\infty]^p$.)

COROLLARY 2. The hazard measure ν_F jointly with the hazard measures relative to all the univariate and multivariate marginals of F determines F uniquely. (This corollary can be verified by induction.)

Remark 11.

It can be noted that the result of Corollary 1 does not remain valid if the assumption that F is continuous is dropped. Also, in view of what we have observed, it can be concluded that if F is continuous, then we have the set of extreme points of $\mathcal{D}_F$ to be a nonempty subset of K and each $F^* \in \mathcal{D}_F$ to be the barycenter of a probability measure that is carried by the set of extreme points of $\mathcal{D}_F$.

Remark 12.

The finite collection of hazard measures given in Corollary 2
appears, in spite of the restriction that $F \in \mathbf{C}$, to be a more natural
multivariate analogue of the univariate hazard measure than the hazard
gradient of the last section. A stability theorem for the collection
is valid when F is continuous, as is shown by Corollary 3 of the next
section.

4. A STABILITY THEOREM

We conclude the paper by proving and commenting in this section on
a general stability theorem for probability measures on metric spaces,
which yields, among other things, the two stability propositions in
KSh (1980) as simple corollaries. The proof of the present theorem uses
Prohorov's (1956) and related theorems in Billingsley (1968) dealing
with the convergence of probability measures. It might be instructive
to compare this with the proofs of earlier stability propositions in KSh
(1980). The techniques used for proving the theorem here are indeed of
a more global nature than those which are sufficient in the case of
probability measures on the real line.

Now, let S be a metric space, T an index set, $\mathbf{S}$ the Borel σ-field
on S, $\mathbf{P}$, $\mathbf{P}_1$, $\mathbf{P}_2$, $\mathbf{P}_3$ families of probability measures on $(S,\mathbf{S})$, $\{A_t: t \in T\}$
a family of collections of sets with $\mathbf{A}_t \subseteq \mathbf{S}$ for every $t \in T$, and
$\{h(\cdot|t,A_t,P): A_t \in \mathbf{A}_t, P \in \mathbf{P}, t \in T\}$ a family of real-valued Borel
measurable functions on $(S,\mathbf{S})$ satisfying the following conditions in
which the notation $D(t,A_t,P)$ stands for the set of discontinuity points
of $h(\cdot|t,A_t,P)$.

(i) P_1, P_2, $P_3 \subset P$, also P_2 is closed (under weak convergence),

(ii) $P_1^{(1)}$, $P_2^{(1)}$, $P_3^{(1)}$, ... $\in \dot{P}_1$ and $\{P_n^{(1)}: n \geq 1\}$ converges weakly to $P^* \in P \Rightarrow h(\cdot|t,A_t,P_n^{(1)}) \to h(\cdot|t,A_t,P^*)$ as $n \to \infty$ uniformly almost surely $[P^*]$ on $A_t \cap D^c(t,A_t,P^*)$ and

$$\sup_{n \geq 1} E_{P_n^{(1)}}\{|h(\cdot|t,A_t,P_n^{(1)})|I_{A_t \cap \{|h(\cdot|t,A_t,P_n^{(1)}| \geq \alpha\}}\} \to 0$$

as $\alpha \to \infty$ for each $t \in T$ and P^*-continuity set A_t in A_t with $P^*(A_t)$ positive,

(iii) $P_1^{(2)}$, $P_2^{(2)} \in P_2$ and are distinct $\Rightarrow$ there exist $t \in T$ and $A_t \in A_t$ such that $P_1^{(2)}(A_t)$, $P_2^{(2)}(A_t)$ are both positive, A_t is both $P_1^{(2)}$-continuity set and $P_2^{(2)}$-continuity set and

$$E_{P_1^{(2)}}\{h(\cdot|t,A_t,P_1^{(2)})|A_t\} \neq E_{P_2^{(2)}}\{h(\cdot|t,A_t,P_2^{(2)})|A_t\},$$

and

(iv) $P^{(3)} \in P_3 \Rightarrow D(t,A_t,P^{(3)})$ has zero $P^{(3)}$-measure for every t in T and $P^{(3)}$-continuity set A_t in A_t.

Further, let $P \in P$ and $\{P_n: n \geq 1\}$ be a sequence of members of P_1 such that $\{P_n: n = 1,2,...\}$ is relatively compact. Then we have the following stability theorem:

THEOREM 4. (a) The condition that

$$P \in P_3, \quad \{P_n: n \geq 1\} \text{ converges weakly to } P \tag{4.1}$$

implies that

$$E_{P_n}\{h(\cdot|t,A_t,P_n)|A_t\} \to E_P\{h(\cdot|t,A_t,P)|A_t\} \tag{4.2}$$

as $n \to \infty$ for every $t \in T$ and P-continuity set $A_t \in A_t$ with $P(A_t) > 0$. Moreover, (b) if additionally $P, P_1, P_2, ... \in P_2$ and the set of cluster points of $\{P_n: n = 1,2,...\}$ (relative to weak convergence) is a subset

of P_3, then the converse assertion is valid.

<u>Proof.</u> Assume first that (4.1) is valid. Since $P \in P_3$, it is obvious that the set of discontinuity points of $h(\cdot|t,A_t,P)I_{A_t}$ has zero P-measure for every $t \in T$ and P-continuity set $A_t \in A_t$. Now, let $t \in T$ and P-continuity set $A_t \in A_t$ be arbitrarily fixed. Since $P_n \in P_1$, $n \geq 1$, the requirements of Billingsley's (1968) Theorem 5.5 are clearly met with $h(\cdot|t,A_t,P)I_{A_t}$ as h and $h(\cdot|t,A_t,P_n)I_{A_t}$ as h_n. This theorem implies that $\{P_n h_n^{-1}, n = 1,2,\ldots\}$ converges weakly to Ph^{-1}. If we now consider X_n, $n \geq 1$ and X to be some random variables having distributions $P_n h_n^{-1}$, $n \geq 1$ and Ph^{-1} respectively, we have $\{X_n: n = 1,2,\ldots\}$ converging to X in distribution. Also, the fact that $P_n \in P_1$, $n \geq 1$ implies that $\{X_n: n = 1,2,\ldots\}$ considered here is uniformly integrable. Since Billingsley's (1968) Theorem 5.4 yields that $E\{X_n\} \to E\{X\}$ as $n \to \infty$ in such a situation, we can conclude that

$$E_{P_n}\{h(\cdot|t,A_t,P_n)I_{A_t}\} \to E_P\{h(\cdot|t,A_t,P)I_{A_t}\} \text{ as } n \to \infty. \qquad (4.3)$$

In view of the assumptions that $\{P_n\}$ converges weakly to P and A_t is a P-continuity set, it follows that $P_n(A_t) \to P(A_t)$ as $n \to \infty$. If $P(A_t) > 0$, we have (4.2) then as an obvious consequence of (4.3). Hence we have the first part of the stability theorem to be valid.

To establish that the second part of the theorem holds, assume that $P, P_1, P_2, \ldots \in P_2$ and the set of cluster points of $\{P_n: n = 1,2,\ldots\}$ is a subset of P_3 and also that (4.2) is valid. Since each cluster point of $\{P_n: n = 1,2,\ldots\}$ is an element of P_3 and $\{P_n: n = 1,2,\ldots\}$ is relatively compact, we should have a subsequence $\{P_{n_r}: r = 1,2,\ldots\}$ of $\{P_n: n = 1,2,\ldots\}$ converging weakly to $Q \in P_3$ with $Q \neq P$ unless (4.1) is

valid. If Q^* denotes the (weak) limit of a subsequence of $\{P_n\}$, then clearly we have $Q^* \in \boldsymbol{P}_3$ and hence the first part of the theorem and the validity of (4.2) lead us to

$$E_P\{h(\cdot|t,A_t,P)|A_t\} = E_{Q^*}\{h(\cdot|t,A_t,Q^*)|A_t\} \qquad (4.4)$$

for every $t \in T$ and $A_t \in \boldsymbol{A}_t$ such that A_t is a P-continuity set with $P(A_t) > 0$ as well as a Q^*-continuity set with $Q^*(A_t) > 0$. We have assumed that $P \in \boldsymbol{P}_2$ and for each $n \geq 1$, $P_n \in \boldsymbol{P}_2$ and also we have $\boldsymbol{P}_2$ to be closed. In that case, we have $P, Q^* \in \boldsymbol{P}_2$ and hence, in view of (4.4), $Q^* = P$. It is therefore impossible that (4.1) will not be valid. Hence we have the second part of the theorem.

Remark 13.

In the case of $h(\cdot|t,A_t,P)$ being independent of P, obviously the part of condition (ii) that $h(\cdot|t,A_t,P_n^{(1)}) \to h(\cdot|t,A_t,P^*)$ uniformly almost surely $[P^*]$ on $A_t \cap D^c(t,A_t,P^*)$ for every $t \in T$ and P^*-continuity set A_t with $P^*(A_t) > 0$ is trivially met. Also, if $h(\cdot|t,A_t,P)$ are all continuous, then the condition (iv) above is obviously satisfied with $\boldsymbol{P}_3 = \boldsymbol{P}$. If S is a Polish space or in particular, if it is a Euclidean space, we have a sequence $\{P_n: n = 1,2,\ldots\}$ of members of $\boldsymbol{P}$ to be relatively compact if and only if it is tight in the sense of Billingsley (1968: p.37) (cf. Theorems 6.1 and 6.2 in Billingsley (1968)). Thus, it is evident that in various specialized situations, the theorem given above has simplified and perhaps more appealing versions.

Remark 14.

If the stipulation "the set of cluster points of $\{P_n: n = 1,2,\ldots\}$ is a subset of $\boldsymbol{P}_3$" is replaced by the weaker stipulation "the set of

cluster points of the range of $\{P_n: n = 1,2,\ldots\}$ is a subset of $\mathcal{P}_3$",
Theorem 4 still remains valid provided we also replace "the converse
assertion is valid" by "(4.2) implies that $\{P_n: n = 1,2,\ldots\}$ converges
weakly to P".

Remark 15.

To illustrate that the stability theorem just proved does not remain
valid if the assumptions $P \in \mathcal{P}_3$ and the set of cluster points of
$\{P_n: n = 1,2,\ldots\}$ is a subset of $\mathcal{P}_3$ respectively appearing in the two
parts of the theorem are omitted, it is sufficient to consider the
following example:

EXAMPLE 5. Let $\{x_n: n = 1,2,\ldots\}$ be a sequence of strictly increasing,
real numbers converging to a real number x'. Let x'' be a real number
greater than x'. Define P, P', $\{P_n: n = 1,2,\ldots\}$ to be a sequence of
probability measures on the Borel σ-field of R^1 such that for some
$0 < \alpha < 1$

$$P_n(\{x\}) = \begin{cases} \alpha & \text{if } x = x_n \\ 1 - \alpha & \text{if } x = x'', \end{cases}$$

$$P(\{x\}) = \begin{cases} \alpha & \text{if } x = x' \\ 1 - \alpha & \text{if } x = x'', \end{cases}$$

and

$$P'(\{x\}) = \begin{cases} \alpha + \dfrac{\alpha(d-c)}{x''-x'} & \text{if } x = x' \\ 1 - \alpha - \dfrac{\alpha(d-c)}{x''-x'} & \text{if } x = x'', \end{cases}$$

where c and d are given real numbers such that $c < d$ and $\{\alpha(d-c)/(x''-x')\}$
$< 1 - \alpha$. Also, define h on R such that

$$h(x) = \begin{cases} c & \text{if } x < x' \\ d + (x-x') & \text{if } x \geq x''. \end{cases}$$

If we take T = the singleton $\{1\}$, $A_1 = \{(-\infty,x): -\infty < x < x''\}$,
$P = \{P,P',P_1,P_2,\ldots\}$ and $h(\cdot|1,A,P^*) = h(\cdot)$ for every member A of A_1
and $P^* \in P$, then it follows that P itself satisfies the requirement
of P_1 and P_2 mentioned above. However, in this case we cannot have
a nonempty subset P_3 of P satisfying the condition (iv) as required.
Consequently, it follows that in this example neither the requirement
of $P \in P_3$ nor the requirement of the set of cluster points of
$\{P_n: n = 1,2,\ldots\}$ being a subset of P_3 is met. Observe that here
$\{P_n: n = 1,2,\ldots\}$ converges to P weakly, $P \neq P'$ and (4.2) is not valid
(since $E_{P_n}\{h(\cdot)|A\} \neq E_p\{h(\cdot)|A\}$ whenever $A = (-\infty,x)$ with $x < x'$) but
(4.2) with P replaced by P' is valid. This implies that with the de-
letions mentioned above neither the first part of the theorem nor the
second part remains valid.

Theorem 4 has several interesting corollaries. In particular it
yields that if a characteristic property exists, based on conditional
expectations of the type $E_p\{h(\cdot|t)|A_t\}$ for probability measures P within
a certain class, then, under certain mild conditions, one can produce
a stability version of the property. It is easily seen that Proposi-
tion 4 of KSh (1980) is an obvious corollary of Theorem 4 and also it
is not difficult now to state a stability version of our Theorem 2 of
Section 2 based on Theorem 4. (Note that in view of what was revealed
in Remark 13, the statement of Theorem 4 simplifies under the situation
in Theorem 2.) It is also worth pointing out in this place that (in view
of Proposition 5 of KSh (1980)) the "only if" part of Proposition 8 of
KSh (1980) follows as a corollary of the first part of Theorem 4 by

letting $S = R^1$, $T = (-\infty, b)$, $A_t = \{R^1\}$ for every $t \in (-\infty, b)$, $\boldsymbol{P} = \boldsymbol{P}_1 =$ the set of measures in the sequence

$$\{P_{F_n} : n \geq 0,\ F_0 = F\},\ (\boldsymbol{P}_2 =)\boldsymbol{P}_3 = \{P_F\}$$

and for each t in T and P^* in P

$$h(x|t, A_t, P^*) = \begin{cases} (P^*([x, \infty)))^{-1} I_{(-\infty, t]} & \text{if } P^*([x, \infty)) > 0 \\\\ 0 & \text{otherwise;} \end{cases}$$

moreover, if some simple initial observations are made and $\boldsymbol{P}$, $\boldsymbol{P}_1$, $\boldsymbol{P}_2$ and $\boldsymbol{P}_3$ are appropriately redefined, the "if" part of Proposition 8 of KSh (1980) follows from the second part of Theorem 4. Essentially the same argument leads to the following stability version of the characterization result in our Corollary 2 of Section 3. This result clearly subsumes Proposition 8 of KSh (1980).

COROLLARY 3. Let $p \geq 1$ and $\{F_n : n = 1, 2, \ldots\}$ be a sequence of d.f.'s on R^p and F be a continuous d.f. on R^p. Assume that F and for each n, F_n are members of the set $\boldsymbol{C}$ defined in the last section. Then

$$F_n(\underset{\sim}{x}) \to F(\underset{\sim}{x}) \text{ for all } \underset{\sim}{x} \in R^p$$

if and only if

$$\underset{\sim}{\nu}^*_{F_n}(\underset{\sim}{x}) \to \underset{\sim}{\nu}^*_F(\underset{\sim}{x}) \text{ for all } \underset{\sim}{x} \text{ with } \overline{F}(\underset{\sim}{x}) > 0,$$

where the notation $\underset{\sim}{\nu}^*_G(\underset{\sim}{x})$ stands for the vector whose elements (given in some specified order) are $\nu_G((-\infty, \underset{\sim}{x}])$ and its counterparts relative to all the univariate and multivariate marginals of G, with appropriate subvectors in place of $\underset{\sim}{x}$ and appropriate number of components in ∞, and $\overline{F}$ stands for the survivor function corresponding to F as in the earlier sections.

ACKNOWLEDGMENTS

REFERENCES

1. Basu, A.P. (1971). Bivariate failure rate, J. Amer. Statist. Assoc., 66, 103-104.

2. Bertoluzza, C. and Forte, B. (1985). Mutual dependence of random variables and maximum discretized entropy, Ann. Prob., 13, 630-637.

3. Billingsley, P. (1968). Convergence of Probability Measures, Wiley, New York.

4. Burrill, C.W. (1972). Measure, Integration and Probability, McGraw Hill, New York.

5. Cox, D.R. (1972). Regression models and life tables, J. Roy, Statist. Soc., B, 34, 187-220.

6. Davies, P.L. and Shanbhag, D.N. (1987). A generalization of a theorem of Deny with applications in characterization theory, Quart. J. Math. (Oxford) (to appear).

7. Galambos, J. and Kotz, S. (1978). Characterization of Probability Distributions, Lecture Notes in Mathematics, No. 675, Springer-Verlag, Heidelberg, New York.

8. Glänzel, W., Telcs, A. and Schubert, A. (1984). Characterization by truncated moments and its application to Pearson-type distributions, Z. Wahrsch. Verw. Geb. 66, No. 2, 173-183.

9. Gupta, R.C. (1975). On the characterization of distributions by conditional expectations, Commun. in Statist., 4, 99-103.

10. Hall, W.J. and Wellner, J.A. (1981). Mean residual life, Statistics and Related Topics, North Holland, Amsterdam, 169-184 (eds. M. Csörgö, D.A. Dawson, J.N.K. Rao, and A.K. Md. E. Saleh).

11. Hamdan, M.A. (1972). On a characterization by conditional expectations, Technometrics, 14, 497-499.

12. Hollander, M. and Proschan, F. (1984). Nonparametric Concepts and Methods in Reliability, Handbook of Statistics, 4, 613-655, (eds. P.R. Krishnaiah and P.K. Sen).

13. Johnson, N.L. and Kotz, S. (1972). Distributions in Statistics: Continuous Multivariate Distributions, Wiley, New York.

14. Johnson, N.L. and Kotz, S. (1975a). A vector multivariate hazard rate, J. Multiv. Analysis, 5, 53-66, 498.

15. Johnson, N.L. and Kotz, S. (1975b). On some generalized Farlie-Gumbel-Morgenstern distributions, Commun. in Statistics, 4, 415-428.

16. Kendall, D.G. (1963). Extreme points methods in stochastic analysis. Z. Wahrsch. Verw. Geb. 1, 295-300.

17. Kotlarski, I.I. (1972). On a characterization of some probability distributions by conditional expectation, Sankhyā, Ser. A, 34, 461-467.

18. Kotz, S. and Shanbnag, D.N. (1980). Some new approaches to probability distributions, Advances Appl. Probability, 12, 903-921.

19. Krishnaiah, P.R. (1977). On generalized gamma type distributions and their applications in reliability, Theory and Applications in Reliability, Vol. 1, 47, Academic Press, New York (eds. C.P. Tsokos and I.N. Shimi).

20. Lau, Ka-Sing and Rao, C.R. (1982). Integrated Cauchy functional equation and characterizations of the exponential law, Sankhyā A. 44, 72-90.

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

22. Parthasarathy, K.R. (1967). Probability Measures on Metric Spaces, Academic Press, New York.

23. Phelps, R.R. (1966). Lectures on Choquet's Theorem, Van Nostrand Co., Princeton.

24. Prakasa Rao, B.L.S. (1974). On a property of bivariate distributions, J. Multiv. Analysis, 4, 106-113.

25. Prohorov, Yu. V. (1956). Convergence of random processes and limit theorems in probability theory, Theor. Probl. and Applic., 1, 177-238.

26. Puri, P.S. and Rubin, H. (1974). On a characterization of the family of distributions with constant multivariate failure rates, Ann. Prob., 2, 738-740.

27. Rao, C.R. and Shanbhag, D.N. (1986). Recent results on characterization of probability distributions: a unified approach through extensions of Deny's theorem, Adv. Appl. Prob., 18, 660-678.

28. Seneta, E. (1981). Non-negative Matrices and Markov Chains, 2nd ed. Springer, New York.

29. Shanbhag, D.N. (1970). Characterizations for exponential and geometric distributions, J. Amer. Statist. Assoc., 65, 1256-1259.

30. Shanbhag, D.N. (1979). Some refinements in distribution theory, Sankhya, Ser. A, 41, 252-262.

31. Shanbhag, D.N. and Bhaskara Rao, M. (1975). A note on characterizations of probability distributions based on conditional expected values, Sankhya, Ser. A, 37, 297-300.

32. Zahedi, H. (1985). Some new classes of multivariate survival distribution functions, Journal of Statist. Planning and Inference, 11, 171-188.

END

DATE FILMED

DEC. 1987