JOURNAL OF MULTIVARIATE ANALYSIS 5, 53-66 (1975)

A Vector Multivariate Hazard Rate

N. L. JOHNSON*

University of North Carolina at Chapel Hill

AND

SAMUEL KOTZ

Temple University, Philadelphia, Pennsylvania

Communicated by P. K. Sen

A vector definition of multivariate hazard rate, and associated definitions of increasing and decreasing multivariate hazard rate distributions are presented. Consequences of these definitions are worked out in a number of special cases. Relationships between hazard rate and orthant dependence are established.

1. INTRODUCTION

The *hazard rate* $h_X(x)$ of the distribution of an absolutely continuous random variable X with cumulative distribution $F_X(x) = \Pr[X \leqslant x]$, and density function $f_X(x) = F_X'(x)$ is defined as

$$h_X(x) = -(d/dx)\log(1 - F_X(x)) = f_X(x)/\{1 - F_X(x)\} = f_X(x)/G_X(x), \quad (1)$$

with $G_X(x) = 1 - F_X(x)$, in the interval $0 < G_X(x) < 1$ and is undefined otherwise.

The hazard rate has been used (see, for example [1]) as a basis for certain kinds of classification of univariate distributions. If $h_X(x)$ is an increasing (decreasing) function of x (for those values for which it is defined), the distribution is termed *increasing* (*decreasing*) *hazard rate*, denoted by IHR (DHR).

Received July, 1974.

AMS classification number: 62H05.

Key words: Hazard rate, multivariate distributions, Morgenstern–Gumbel–Farlie system, orthant dependence.

* This work was supported by the U. S. Army Research Office, Durham, under Contract No. DAHC04-71-C-0042.

Of course, most distributions are neither IHR nor DHR — $h_X(x)$ may increase over certain ranges of x, and decrease over others.

2. Multivariate Hazard Rate

Some attempts have been made to extend the definition of IHR and DHR to multivariate distributions [2, 4, 6, 7, 10, 11], in which multivariate hazard rate has been defined as a single scalar quantity, or there has been no explicit definition of a multivariate hazard rate.

In [8] we took the point of view that, for a concept such as "hazard rate," it is unreasonable to expect a single value to represent this aspect of a multivariate distribution. The basic idea underlying the univariate definition is that of rate of decrease in "survivors" with increase in value (x) of X (as in a life table where the hazard rate is in fact the force of mortality). When there are two or more variates this rate depends on which variate is changed (or more generally, the proportions in which different variates are changed) and we need a different "rate" for each variate.

We defined the (joint) ***multivariate hazard rate of m*** absolutely continuous random variables $X_1, ..., X_m$ as the vector

$$h_{\mathbf{X}}(\mathbf{x}) = (-(\partial/\partial x_1), ..., -(\partial/\partial x_m)) \log G_{\mathbf{X}}(\mathbf{x}) = -\text{grad} \log G_{\mathbf{X}}(\mathbf{x}), \qquad (2)$$

where $G_{\mathbf{X}}(\mathbf{x}) = P\{X_i > x_i, i = 1, ..., m\}$. For convenience we will write $-(\partial/\partial x_j) \log G_{\mathbf{X}}(\mathbf{x}) = h_{\mathbf{X}}(\mathbf{x})_j, j = 1, ..., m$.

If for *all* values of $\mathbf{x}$, *all* components of $h_{\mathbf{X}}(\mathbf{x})$ are increasing (decreasing) functions of the corresponding variable, i.e., $h_{\mathbf{X}}(\mathbf{x})_j$ is an increasing (decreasing) function of x_j for $j = 1, 2, ..., m$, then the distribution is called ***(multivariate) IHR (DHR)***.

If we wish to emphasize this particular definition we will call it ***vector*** (or ***gradient***) ***multivariate IHR (DHR)***.

If $h_{\mathbf{X}}(\mathbf{x})_j$ is an increasing (decreasing) function of x_j ($j = 1, ..., m$) at $\mathbf{x}' = (x_1, x_2, ..., x_m)$ we say that the distribution is vector (or gradient) multivariate IHR (DHR) *at the point* $(x_1, ..., x_m)$.

Marshall [10] and Block [3] also regard $h_{\mathbf{X}}(\mathbf{x})$ as being relevant to the concept of multivariate hazard rate. Marshall's definition of multivariate IHR appears to require that $h_{\mathbf{X}}(\mathbf{x})_j$ is an increasing function of each of $x_1, x_2, ..., x_m$; Block [3] requires (i) $h_{\mathbf{X}}(\mathbf{x} + \Delta \mathbf{1})_j$ to be a decreasing function of Δ (for any $\mathbf{x}$ and all j), and (ii) $h_{\mathbf{X}}(\mathbf{x})_j$ is an increasing function of x_i for all $i \neq j$ (Note that both of these definitions place conditions on variation of $h_{\mathbf{X}}(\mathbf{x})_j$ with respect to variables other than x_j.)

3. Some Properties

(i) If $X_1, \ldots, X_m$ are mutually independent then $h_{\mathbf{X}}(\mathbf{x})_j = h_{X_j}(x_j)$ where, of course, the left-hand side is a component of a multivariate hazard rate and the right-hand side is a univariate hazard rate.

So, if $X_1, \ldots, X_m$ are independent, their joint distribution is IHR (DHR) if and only if the distribution of each $X_1, \ldots, X_m$ is IHR (DHR).

(ii) Suppose that $X_1, \ldots, X_m$ *are exchangeable*, i.e., $G_{\mathbf{X}}(\mathbf{x})$ is unchanged if $X_1, \ldots, X_m$ are permuted in any order. This implies that if the joint distribution is IHR (DHR) for some particular set of values $x_1, \ldots, x_m$ it will also be IHR (DHR) for any permutation of these values.

(iii) If $Z_1 = -X_1$ then (remembering the variables are continuous)

$$h_{Z_1, X_2, \ldots, X_m}(z_1, x_2, \ldots, x_m)_1 = h_{\mathbf{X}}(-z_1, x_2, \ldots, x_m)_1, \tag{3.1}$$

while for $j \geqslant 2$

$$h_{Z_1, X_2, \ldots, X_m}(z_1, x_2, \ldots, x_m)_j = h_{X_2, \ldots, X_m}(x_2, \ldots, x_m)_j - h_{\mathbf{X}}(-z_1, x_2, \ldots, x_m)_j. \tag{3.2}$$

Similar (though more complicated) results can be obtained when several of the X_j's are changed in sign.

If Y_j is a continuous increasing monotonic function of X_j for $j = 1, 2, \ldots, m$ and $X_j = \xi_j(Y_j)$ $(j = 1, \ldots, m)$, then

$$h_{\mathbf{Y}}(\mathbf{y})_j = h_{\mathbf{X}}(\boldsymbol{\xi}(\mathbf{y}))_j(\partial \xi_j / \partial y_j). \tag{4}$$

If $\mathbf{X}$ is IHR (DHR) and $\partial \xi_j / \partial y_j$ is a nondecreasing (nonincreasing) function of y_j for $j = 1, 2, \ldots, m$ then $\mathbf{Y}$ also is IHR (DHR). In particular, if $\mathbf{X}$ is IHR (DHR) then so is $(b_1 X_1 + a_1, \ldots, b_m X_m + a_m)$ if $\mathbf{b} > \mathbf{0}$.

(iv) If the multivariate hazard rate is constant, (i.e., does not vary with *any* of $x_1, x_2, \ldots, x_m$), so that $h_{\mathbf{X}}(\mathbf{x}) = \mathbf{c}$, this means that (whenever the hazard rate exists) $(\partial \log G_{\mathbf{X}}(\mathbf{x}) / \partial x_j) = -c_j$ $(j = 1, \ldots, m)$.

Hence $G_{\mathbf{X}}(\mathbf{x}) = \exp(-c_j x_j)\, g_j(x_1, \ldots, x_{j-1}, x_{j+1}, \ldots, x_m)$ $(j = 1, \ldots, m)$, whence

$$G_{\mathbf{X}}(\mathbf{x}) \propto \exp\left(-\sum_{j=1}^{m} c_j x_j\right). \tag{5}$$

Thus, the X's are mutually independent exponential variables if and only if the multivariate hazard rate is constant.

We may distinguish between *strictly* constant vector hazard rates ($h_{\mathbf{X}}(\mathbf{x}) = \mathbf{c}$, as in this Section) and *locally* constant rates, for which $h_{\mathbf{X}}(\mathbf{x})_j$ does not depend on x_j, though it may depend on the other x's. We shall see (in Sect. 5.4) that bivariate distributions with locally constant vector multivariate hazard rates belong to the family of Gumbel's bivariate exponential distributions.

(v) Noting that $G_{\mathbf{X}}(\mathbf{x}) = G_{X_2,\ldots,X_m}(x_2,\ldots,x_m)\,\tilde{G}_{X_1|X_2,\ldots,X_m}(x_1 \mid x_2,\ldots,x_m)$ where $\tilde{G}_{X_1|X_2,\ldots,X_m}(x_1 \mid x_2,\ldots,x_m) = \Pr[X_1 > x_1 \mid \bigcap_{j=2}^{m}(X_j > x_j)]$, we see that

$$\begin{aligned} h_{\mathbf{X}}(\mathbf{x})_1 &= -(\partial/\partial x_1)\log \tilde{G}_{X_1|X_2,\ldots,X_m}(x_1 \mid x_2,\ldots,x_m), \\ &= \tilde{h}_{X_1|X_2,\ldots,X_m}(x_1 \mid x_2,\ldots,x_m), \end{aligned} \tag{6}$$

in an obvious notation. Generally

$$h_{\mathbf{X}}(\mathbf{x})_j = \tilde{h}_{X_j|(X_1,\ldots,X_m)_j}(x_j \mid (x_1,\ldots,x_m)_j), \tag{7}$$

where $(X_1,\ldots,X_m)_j$ denotes $X_1,\ldots,X_{j-1}$, $X_{j+1},\ldots,X_m$, and similarly for $(x_1,\ldots,x_m)_j$.

Thus the components of the vector multivariate hazard rate are in fact univariate hazard rates of conditional distributions of each variate, given certain inequalities on the remainder.

A direct consequence of these results is that if the joint distribution of $X_1,\ldots,X_m$ is IHR (DHR) so is the conditional joint distribution of any subset of the X's, given inequalities of form $(X_j > x_j)$ on the remainder.

Since $\tilde{h}_{X_1|X_2,\ldots,X_m}(x_1 \mid x_2,\ldots,x_m)$ is the expected value, with respect to variation of $X_2,\ldots,X_m$ of the hazard rate $h_{X_1|X_2,\ldots,X_m}(x_1 \mid X_2,\ldots,X_m)$ of the distribution of X_1 given $X_2, X_3,\ldots,X_m$, we might expect the joint distribution of $X_1,\ldots,X_m$ to be IHR (DHR) if each of the conditional joint distributions of X_j given the other $(m-1)$ X's $(j = 1,\ldots,m)$ is IHR (DHR). This would certainly be so if the appropriate distribution of the other variables did not depend on the value of x_j. However, in the equation

$$\tilde{h}(x_1 \mid x_2,\ldots,x_m) = E[h(x_1 \mid X_2,\ldots,X_m)] \tag{8}$$

(subscripts omitted for conveneince) the relevant joint distribution of $X_2,\ldots,X_m$ is that *conditioned on* $X_1 > x_1$ and truncated by $\bigcap_{j=2}^{m}(X_j > x_j)$.

It is known [1, p. 37] that mixtures of univariate DHR distributions are also DHR. (A similar property is not valid for IHR distributions.) Since $h_{\mathbf{X}}(\mathbf{x})_1$ is the hazard rate of the conditional (univariate) distribution of X_1 given $\bigcap_{j=2}^{m}(X_j > x_j)$, (see (6)), it follows that mixtures of multivariate DHR distributions are also DHR.

(vi) If $X_1, X_2,\ldots,X_m$ are mutually independent and each is IHR (DHR) then the least of them, $L = \min(X_1, X_2,\ldots,X_m)$, is also IHR (DHR).

In the general case, however,

$$h_L(l) = \sum_{j=1}^{m} h_{\mathbf{X}}(l, l,\ldots, l)_j\,. \tag{9}$$

If $h_{\mathbf{X}}(l, l,..., l)_j$ is an increasing (decreasing) function of l for all j, then L is IHR (DHR), but this is not ensured simply by the condition that $\mathbf{X}$ is vector multivariate IHR (DHR). If $h_{\mathbf{X}}(l_1, l_2,..., l_m)_j$ is an increasing (decreasing) function of *each* of $l_1, l_2,..., l_m$ (not only of l_j) for all j, this does ensure that L is IHR (DHR).

On the other hand, it appears that L might be IHR (DHR), even though $\mathbf{X}$ is not vector multivariate IHR (DHR).

We note that if $X_1, X_2,..., X_m$ are exchangeable, then $h_{\mathbf{X}}(l, l,..., l)_j$ does not depend on j (equal to $H_{\mathbf{X}}(l)$, say) and

$$h_L(l) = mH_{\mathbf{X}}(l). \tag{10}$$

4. Hazard Rates and Orthant Dependence

Lehmann [9] has defined *positive* (*negative*) *quadrant dependence* between two variables X_1, X_2 by

$$R_{F(X_1,X_2)} = F_{\mathbf{X}}(\mathbf{x}) \Big/ \left\{\prod_{j=1}^{2} F_{X_j}(x_j)\right\} \geqslant (\leqslant) 1 \tag{11}$$

for all x_1, x_2. He has shown that these definitions are equivalent to those obtained by replacing $R_{F(X_1,X_2)}(x_1, x_2)$ by

$$R_{G(X_1,X_2)}(x_1, x_2) = G_{\mathbf{X}}(\mathbf{x}) \Big/ \left\{\prod_{j=1}^{2} G_{X_j}(x_j)\right\}. \tag{12}$$

Dykstra *et al.* [5] have extended these definitions to define *positive* (*negative*) *orthant dependence* among $m(\geqslant 2)$ variables $X_1, X_2,..., X_m$ to correspond to

$$R_{F(\mathbf{X})}(\mathbf{x}) = F_{\mathbf{X}}(\mathbf{x}) \Big/ \left\{\prod_{j=1}^{m} F_{X_j}(x_j)\right\} \geqslant (\leqslant) 1 \tag{13}$$

for all $\mathbf{x}$.

We, however, will work with the equally natural extended definition

$$R_{G(\mathbf{X})}(\mathbf{x}) = G_{\mathbf{X}}(\mathbf{x}) \Big/ \left\{\prod_{j=1}^{m} G_{X_j}(x_j)\right\} \geqslant (\leqslant) 1 \tag{14}$$

for all $\mathbf{x}$, which we will call *G-positive* (*negative*) *orthant dependence.*

For $m = 2$, the two definitions are equivalent; for $m > 2$ this is not so. (A counterexample is easily constructed (for $m = 3$) by noticing that (13) is satisfied if the conditional joint distribution of X_1 and X_2 given $(X_3 \leqslant x_3)$ has positive (negative) quadrant dependence, while (14) is satisfied if the joint distribution given $(X_3 > x_3)$ has positive (negative) quadrant dependence.)

We will now develop some relations among $R_{G(\mathbf{x})}(\mathbf{x})$ and various hazard rates. We first note that

$$\lim_{x_1 \to -\infty} R_{G(X_1,\ldots,X_m)}(x_1, \ldots, x_m) = R_{G(X_2, \ldots X_m)}(x_2, \ldots, x_m). \tag{15}$$

In particular, for $m = 2$

$$\lim_{x_1 \to -\infty} R_{G(X_1,X_2)}(x_1, x_2) = R_{G(X_2)}(x_2) = G_{X_2}(x_2)/G_{X_2}(x_2) = 1. \tag{16}$$

Since

$$\begin{aligned} \partial \log R_{G(\mathbf{X})}(\mathbf{x})/\partial x_j &= \partial \log G_{\mathbf{X}}(\mathbf{x})/\partial x_j - \partial \log G_{X_j}(x_j)/\partial x_j, \\ &= h_{X_j}(x_j) - h_{\mathbf{X}}(\mathbf{x})_j, \end{aligned} \tag{17}$$

we see that for bivariate distributions ($m = 2$), if $h_{X_j}(x_j) > (<) h_{\mathbf{X}}(\mathbf{x})_j$ for all $\mathbf{x}$ and $j = 1, 2$ then (remembering (16)), the distribution has *positive* (*negative*) *quadrant dependence*. The converse is not necessarily true.

We cannot immediately extend this result to cases $m > 2$. In view of (16), we do have results like:

If $X_1, \ldots, X_{m-1}$ have G-positive (negative) orthant dependence and $h_{X_m}(x_m) > (<) h_{\mathbf{X}}(\mathbf{x})_m$ for all $\mathbf{x}$, then $X_1, \ldots, X_m$ have G-positive (negative) orthant dependence.

Recently Yanagimoto [12] refined the concept of positive dependence for bivariate distributions and proposed three closely related definitions. Relations between his definitions and vector bivariate hazard rates will be discussed in another place.

5. Examples

5.1. *Bivariate Normal Distributions.* It is known that all normal distributions are IHR. For a multinormal distribution, the conditional distributions of any variable, given the remainder, is normal, and so IHR. However, this does not mean that all multivariate normal distributions are IHR, according to the vector multivariate hazard rate defined in this paper.

In fact, bivariate normal distributions with positive (or zero) correlation are vector bivariate IHR. A proof of this result, by J. Galambos, is given in the Appendix. If X_1, X_2 have a joint standard bivariate normal distribution with correlation coefficient ρ, then

$$h_{\mathbf{X}}(\mathbf{x})_1 = \frac{\left\{1 - \Phi\left(\dfrac{x_2 - \rho x_1}{\sqrt{(1-\rho^2)}}\right)\right\} \phi(x_1)}{L_\rho(x_1, x_2)}, \tag{18}$$

where

$$L_\rho(x_1, x_2) = [2\pi\sqrt{(1-\rho^2)}]^{-1} \int_{x_2}^{\infty}\int_{x_1}^{\infty} \exp[-\tfrac{1}{2}(1-\rho^2)^{-1}(t_1^2 - 2\rho t_1 t_2 + t_2^2)]\, dt_1\, dt_2,$$

$$\phi(x) = (\sqrt{2\pi})^{-1}\exp(-\tfrac{1}{2}u^2), \qquad \Phi(x) = \int_{-\infty}^{x} \phi(u)\, du.$$

The ratio of this component of the multivariate hazard rate to the univariate hazard rate, $h_{X_1}(x_1)$, is

$$\frac{h_{\mathbf{X}}(\mathbf{x})_1}{h_{X_1}(x_1)} = \frac{\left\{1 - \Phi\left(\frac{x_2 - \rho x_1}{\sqrt{(1-\rho^2)}}\right)\right\}\{1 - \Phi(x_1)\}}{L_\rho(x_1, x_2)}. \tag{19}$$

Since $\lim_{x_2\to-\infty} G_{X_1,X_2}(x_1, x_2) = G_{X_1}(x_1)$, we might expect that

$$\lim_{x_2\to-\infty} h_{\mathbf{X}}(\mathbf{x})_1 = h_{X_1}(x_1). \tag{20}$$

From Table I it can be seen that this limit is approached more rapidly when (i) ρ is large and positive, or (ii) x_1 is large and positive.

More extensive tables are given in [8].

TABLE I

Values of $h_{\mathbf{X}}(\mathbf{x})_1/h_{X_1}(x_1)$

	$\rho = 0.4$					$\rho = 0.8$				
x_2 \ x_1	−2.0	−1.0	0	1.0	2.0	−2.0	−1.0	0	1.0	2.0
−2.0	0.9235	0.9733	0.9924	0.9981	0.9996	0.7575	0.9806	0.9996	1.0000	1.0000
−1.0	0.6908	0.8498	0.9389	0.9778	0.9927	0.1847	0.6799	0.9629	0.9990	1.0000
0	0.3766	0.6069	0.7924	0.9000	0.9544	0.0075	0.1552	0.6288	0.9418	0.9976
1.0	0.1529	0.3505	0.5837	0.7593	0.8656	–[a]	0.0072	0.1561	0.6003	0.9176
2.0	0.0484	0.1662	0.3772	0.5900	0.7416	–[a]	0.0001	0.0094	0.1730	0.5847

[a] – denotes "<0.00005."

5.2. *Multivariate Pareto Distributions.* Consider the m-dimensional multivariate Pareto distribution, with density function

$$P_{\mathbf{X}}(\mathbf{x}) = a(a+1)\cdots(a+m-1)\left(\prod_{j=1}^{m}\theta_j\right)^{-1}\left(\sum_{j=1}^{m}\theta_j^{-1}x_j - m + 1\right)^{-(a+m)}$$

$$(a > 0;\ x_j > \theta_j > 0). \tag{21}$$

For this distribution $G_{\mathbf{X}}(\mathbf{x}) = (\sum_{j=1}^{m} \theta_j^{-1}x_j - m + 1)^{-a}$. Hence

$$-\partial \log G_{\mathbf{X}}(\mathbf{x})/\partial x_r = a\theta_r^{-1}\left(\sum_{j=1}^{m} \theta_j^{-1}x_j - m + 1\right)^{-1}. \tag{22}$$

This is a decreasing function of x_r for all r $(=1,..., m)$. Hence this is a decreasing vector multivariate hazard rate distribution.

5.3. *Morgenstern–Gumbel–Farlie Distributions.* Consider the family of bivariate distributions for which

$$G_{X_1,X_2}(x_1, x_2) = G_{X_1}(x_1)\, G_{X_2}(x_2)[1 + \alpha(1 - G_{X_1}(x_1))(1 - G_{X_2}(x_2))] \quad \text{with} \quad |\alpha| < 1. \tag{23}$$

(These were originally defined in terms of a similar equivalent relation among the cumulative distribution functions.) For this family

$$\begin{aligned} h_{X_1,X_2}(\mathbf{x})_j &= h_{X_j}(x_j) - \alpha\{1 - G_{X_{3-j}}(x_{3-j})\} f_{X_j}(x_j)/[1 + \alpha\{1 - G_{X_1}(x_1)\}\{1 - G_{X_2}(x_2)\}], \\ &= [1 - \{\beta(G_{X_j}(x_j))^{-1} - 1\}^{-1}]\, h_{X_j}(x_j), \end{aligned} \tag{24}$$

where $\beta = 1 + [\alpha\{1 - G_{X_{3-j}}(x_{3-j})\}]^{-1}$.

Note that β has the same sign as α, since $|\alpha|^{-1} > 1$ and $|1 - G_{X_{3-j}}(x_{3-j})|^{-1} > 1$.

If α is positive (negative) then (noting that $G_{X_j}(x_j)$ is a decreasing function of x_j) $h_{\mathbf{X}}(\mathbf{x})_j$ is an increasing (decreasing) function of x_j.

It follows that if both X_1 and X_2 are IHR (DHR) then $\mathbf{X}$ is multivariate IHR (DHR) if α is positive (negative).

We have

$$R_{G(X_1,X_2)}(x_1, x_2) = 1 + \alpha\{1 - G_{X_1}(x_1)\}\{1 - G_{X_2}(x_2)\}. \tag{25}$$

($R_G(\cdot)$ is defined in Section 4, Eq. 12.) In this case $R_{G(X_1,X_2)}(x_1, x_2) \geqslant (\leqslant)\, 1$ according as $\alpha \geqslant (\leqslant)\, 0$. Also, from Eq. (24), $h_{X_j}(x_j) - h_{\mathbf{X}}(\mathbf{x})_j$ has the same sign as α, so in this case positive (negative) quadrant dependence does imply $h_{X_j}(x_j) > (<)\, h_{\mathbf{X}}(\mathbf{x})_j$).

In the special case when both X_1 and X_2 have exponential distributions, so that $h_1(x_1) = h_1$ and $h_2(x_2) = h_2$ are constants, we have

$$h_{\mathbf{X}}(\mathbf{x})_j = [1 - \{\beta \exp(h_j x_j) - 1\}^{-1}]\, h_1, \tag{26}$$

with $\beta = 1 + [\alpha\{1 - \exp(-h_{3-j}x_{3-j})\}]^{-1}$. Hence the joint distribution is multivariate IHR if α is positive, DHR if α is negative.

If, in (24), X_1 and X_2 each have Weibull distributions, with

$$G_{X_j}(x_j) = \exp(-x_j^{c_j}) \qquad (c_j > 0, x_j > 0; j = 1, 2),$$

then

$$h_{\mathbf{X}}(\mathbf{x})_j = [1 - \{\beta \exp(x_j^{c_j}) - 1\}^{-1}]\, c_j x_j^{c_j - 1} \qquad (j = 1, 2), \tag{27}$$

where $\beta = 1 + [\alpha\{1 - \exp(-x_{3-j}^{c_{3-j}})\}]^{-1}$. Further detailed analysis gives:

(a) If $\alpha > 0$ (<0) and $c_j > 1$ (<1) for $j = 1, 2$, then the joint distribution of X_1 and X_2 is multivariate IHR (DHR).

(b) The joint distribution is vector bivariate DHR if $0 < \alpha < 1$ and $c_1, c_2 \leqslant 1/2$.

(c) If $1/2 < c_j < 1$ $(j = 1, 2)$ then the distribution is DHR provided α (>0) is less than

$$(2b_j)^{-1}\{1 + 8b_j^2)^{1/2}\} \exp[-1 + 3b_j - 8b_j^2\{1 + (1 + 8b_j^2)^{1/2}\}^{-1}],$$

where $b_j = c_j^{-1} - 1$, for $j = 1, 2$. Some numerical values are:

c_j	0.6	0.7	0.8	0.9,
Upper limit for α	0.948	0.797	0.566	0.285.

(d) If $\alpha < 0$ and $c_j > 1$ the distribution is bivariate IHR provided α exceeds certain lower limits (see [8]).

5.4 *Gumbel's Bivariate Exponential Distributions.* The standard joint cumulative distribution function for these distributions

$$R_{X_1,X_2}(x_1, x_2) = 1 - e^{-x_1} - e^{-x_2} + e^{-(x_1 + x_2 + \theta x_1 x_2)} \qquad (x_1 > 0, x_2 > 0; \theta > 0), \tag{28}$$

so that

$$G_{X_1,X_2}(x_1, x_2) = \exp(-x_1 - x_2 - \theta x_1 x_2), \tag{29}$$

whence $h_{\mathbf{X}}(\mathbf{x}) = (1 + \theta x_2, 1 + \theta x_1)$.

These components are constant with respect to variation in the corresponding variable (i.e., $h_{\mathbf{X}}(\mathbf{x})_1$ does not depend on x_1, nor $h_{\mathbf{X}}(\mathbf{x})_2$ on x_2) but not with respect to variation in the other variable. Using the terminology introduced in Section 3(iv), the distribution of $\mathbf{X}$ has *locally,* but not *strictly* constant bivariate hazard rate. We now show that under fairly general conditions a locally constant bivariate hazard rate implies that the distribution is of Gumbel's bivariate exponential form (possibly with location and scale parameters not equal to 0 and 1, respectively).

If we assume $h_{\mathbf{X}}(\mathbf{x}) = (f_1(x_2), f_2(x_1))$, then $\log G_{\mathbf{X}}(\mathbf{x}) = -x_1 f_1(x_2) + A_1(x_1) = -x_2 f_2(x_1) + A_2(x_2)$, where $A_1(\cdot)$, $A_2(\cdot)$ are arbitrary functions subject to the standard conditions on $G_{\mathbf{X}}(\mathbf{x})$.

Putting $x_1 = 0$ gives $A_1(0) = -x_2 f_2(0) + A_2(x_2)$, so $A_2(x_2)$ must be a linear function of x_2. Similarly $A_1(x_1)$ must be a linear function of x_1. Now putting $x_1 = a \neq 0$, we obtain

$$-af_1(x_2) = -x_2 f_2(a) + (\text{linear function of } x_2),$$

so that $f_1(x_2)$ must be a linear function of x_2. Similarly $f_2(x_1)$ must be a linear function of x_1. It follows that the joint distribution of appropriate linear transforms of X_1, X_2 is of form (29).

To summarize, Gumbel's family of bivariate exponential distributions includes all distributions with locally constant bivariate hazard rates.

5.5 *Bivariate Exponential Distributions of Marshall and Olkin.* These have joint survival functions of form

$$G_{X_1,X_2}(x_1, x_2) = \exp\{-\lambda_1 x_1 - \lambda_2 x_2 - \lambda_{12} \max(x_1, x_2)\}$$
$$(\lambda_1, \lambda_2, \lambda_{12} > 0; x_1, x_2 > 0). \quad (30)$$

For these distributions

$$h_{\mathbf{X}}(\mathbf{x}) = \begin{cases} (\lambda_1, \lambda_2 + \lambda_{12}) & (x_1 < x_2), \\ (\lambda_1 + \lambda_{12}, \lambda_2) & (x_1 > x_2). \end{cases} \quad (31)$$

They are not strictly IHR, but for $x_2(x_1)$ fixed the first (second) component is a nondecreasing function of $x_1(x_2)$.

The analysis for multivariate ($m > 2$) Marshall–Olkin distributions follows similar lines.

5.6 *Freund's Bivariate Exponential Distribution.* For this distribution, the joint density is

$$f_{X_1,X_2}(x_1, x_2) = \begin{cases} \alpha\beta' \exp\{-\beta' x_2 - (\alpha + \beta - \beta')x_1\} & (0 \leqslant x_1 < x_2), \\ \alpha'\beta \exp\{-\alpha' x_1 - (\alpha + \beta - \alpha')x_2\} & (0 \leqslant x_2 < x_1), \end{cases} \quad (32)$$

with $\alpha, \beta, \alpha', \beta' > 0$.

We will assume $\alpha + \beta \neq \alpha', \beta'$, and $\alpha' > \alpha$, $\beta' > \beta$. We find

$$G_{X_1,X_2}(x_1, x_2) = \begin{cases} (\alpha + \beta - \beta')^{-1}[\alpha \exp\{-\beta' x_2 - (\alpha + \beta - \beta')x_1\} \\ \quad + (\beta - \beta') \exp\{-(\alpha + \beta)x_2\}] & (0 \leqslant x_1 < x_2), \\ (\alpha + \beta - \beta')^{-1}[\beta \exp\{-\alpha' x_1 - (\alpha + \beta - \alpha')x_2\} \\ \quad + (\alpha - \alpha') \exp\{-(\alpha + \beta)x_1\}] & (0 \leqslant x_2 < x_1), \end{cases} \quad (33)$$

and

$$h_X(x)_1 = \begin{cases} \dfrac{\alpha(\alpha + \beta - \beta')}{\alpha - (\beta' - \beta) \exp\{-(\alpha + \beta - \beta')(x_2 - x_1)\}} & (0 \leqslant x_1 < x_2) \\ \dfrac{\beta\alpha' - (\alpha' - \alpha)(\alpha + \beta) \exp\{-(\alpha + \beta - \alpha')(x_1 - x_2)\}}{\beta - (\alpha' - \alpha) \exp\{-(\alpha + \beta - \alpha')(x_1 - x_2)\}} & (0 \leqslant x_2 < x_1). \end{cases} \quad (34)$$

Taking the case $\alpha + \beta > \alpha' > \alpha$; $\alpha + \beta > \beta' > \beta$ we find that for $x_1 < x_2$; $h_{\mathbf{X}}(\mathbf{x})_1$ is of form $c_1\{c_2 - c_3 \exp(c_4 x_1)\}^{-1}$ where $c_j > 0$ ($j = 1, 2, 3, 4$), and so is an increasing function of x_1.

For $x_1 > x_2$; $h_{\mathbf{X}}(\mathbf{x})_1$ is of form $(c_5 \exp(c_6 x_1) - c_7)/(c_8 \exp(c_6 x_1) - c_9)$, where $c_j > 0$, and is an increasing function of x_1 if $c_7 c_8 - c_5 c_9 > 0$. Now $c_7 c_8 - c_5 c_9 = [\beta(\alpha' - \alpha) \exp\{(\alpha + \beta - \alpha') x_2\}][\alpha + \beta - \alpha'] > 0$. So $h_{\mathbf{X}}(\mathbf{x})_1$ is also an increasing function of x_1 for $x_1 > x_2$. Similarly $h_{\mathbf{X}}(\mathbf{x})_2$ is an increasing function of x_2 for all x_2. The joint distribution is IHR.

Similar results are obtained for other inequalities between $\alpha + \beta$, α' and β' (but always keeping $\alpha' > \alpha$, $\beta' > \beta$).

5.7 *Bivariate Logistic.* We consider the joint cumulative distribution function:

$$F_{X_1,X_2}(x_1, x_2) = (1 + e^{-x_1} + e^{-x_2})^{-1}, \tag{35}$$

for which

$$\begin{aligned} G_{X_1,X_2}(x_1, x_2) &= 1 - (1 + e^{-x_1})^{-1} - (1 + e^{-x_2})^{-1} + (1 + e^{-x_1} + e^{-x_2})^{-1}, \\ &= e^{-x_1-x_2}(1 + e^{-x_1})^{-1}(1 + e^{-x_2})^{-1}(1 + e^{-x_1} + e^{-x_2})^{-1}(2 + e^{-x_1} + e^{-x_2}). \end{aligned} \tag{36}$$

We find

$$h_{\mathbf{X}}(\mathbf{x})_1 = (1 + e^{-x_2})(2 + e^{-x_1} + e^{-x_2})^{-1}\{(1 + e^{-x_1} + e^{-x_2})^{-1} + (1 + e^{-x_1})^{-1}\} \tag{37}$$

which is an increasing function of x_1. The joint distribution is IHR.

6. Concluding Remarks

We feel that enough results have been presented to provide evidence of the usefulness of our proposed vector definition of multivariate hazard rate.

Using the definitions of IHR and DHR based on this concept we are able to classify a number of important multivariate distributions, comparable with those so classified by other definitions of IHR and DHR.

We feel however that the criteria of overall IHR and DHR are too sweeping to be of general usefulness. Even for univariate distributions, the DHR property implies that (if continuous) the density must be a decreasing function of x over its support.

Rather, the hazard rate function itself is of value as a description of the distribution—in particular of the intervals in which hazard rate is increasing or decreasing.

From this point of view, the desirability of some form of vector hazard rate for multivariate distributions is evident. Whether the particular one we have chosen is, in any sense, optimal, is an open question. We believe that we have shown it to be useful.

APPENDIX I: Proof that Bivariate Normal Distributions with Nonnegative Correlation Coefficients are Multivariate IHR (J. Galambos)

LEMMA.

$$\phi(x) - x\{1 - \Phi(x)\} > 0.$$

Proof.

$$(d/dx)[\phi(x) - x\{1 - \Phi(x)\}] = -\{1 - \Phi(x)\} < 0,$$

and

$$\lim_{x\to\infty}[\phi(x) - x\{1 - \Phi(x)\}] = 0.$$

THEOREM. *If* $1 > \rho \geqslant 0$ *then*

$$h_{\mathbf{X}}(\mathbf{x})_1 = \{1 - \Phi[(x_2 - \rho x_1)/\sqrt{(1 - \rho^2)}]\}\,\phi(x_1)/L_\rho(x_1, x_2)$$

is an increasing function of x_1 *for all* x_2.

Proof. The result is clearly true for $\rho = 0$, so we suppose $\rho > 0$.

$$\begin{aligned}\frac{\partial h}{\partial x_1} &= \frac{1}{\{L_\rho(x_1, x_2)\}^2}\left[\frac{\rho}{\sqrt{(1 - \rho^2)}} L_\rho(x_1, x_2)\phi\left(\frac{x_2 - \rho x_1}{\sqrt{(1 - \rho^2)}}\right)\phi(x_1)\right.\\ &\quad - x_1\phi(x_1)\{1 - \Phi((x_2 - \rho x_1)/\sqrt{(1 - \rho^2)})\}\, L_\rho(x_1, x_2)\\ &\quad \left.+ [\phi(x_1)\{1 - \Phi((x_2 - \rho x_1)/\sqrt{(1 - \rho^2)})\}]^2\right]\\ &> \frac{\phi(x_1)\{1 - \Phi((x_2 - \rho x_1)/\sqrt{(1 - \rho^2)})\}}{L_\rho(x_1, x_2)^2}\\ &\quad \times [\phi(x_1)\{1 - \Phi((x_2 - \rho x_1)/\sqrt{(1 - \rho^2)})\} - x_1 L_\rho(x_1, x_2)] \qquad (\text{since } \rho > 0).\end{aligned}$$

So $\partial h/\partial x_1 > 0$ if $x_1 \leqslant 0$ or if (cf. definition $(L_\rho(x_1, x_2))$

$$\phi(x_1)\left\{1 - \Phi\left(\frac{x_2 - \rho x_1}{\sqrt{(1 - \rho^2)}}\right)\right\} - x_1\int_{x_1}^{\infty}\phi(t)\left\{1 - \Phi\left(\frac{x_2 - \rho t}{\sqrt{(1 - \rho^2)}}\right)\right\} dt > 0. \quad (\text{A.1})$$

Denote the left-hand side of (A.1) by $\beta(x_1, x_2)$. Now

$$\lim_{x_2\to\infty}\beta(x_1, x_2) = 0. \qquad (\text{A.2})$$

We also have

$$(1-\rho^2)^{1/2}\frac{\partial\beta}{\partial x_2} = -\phi(x_1)\phi\left(\frac{x_2-\rho x_1}{\sqrt{(1-\rho^2)}}\right) + x_1\int_{x_1}^{\infty}\phi(t)\phi\left(\frac{x_2-\rho t}{\sqrt{(1-\rho^2)}}\right)dt.$$

Again using the lemma, we have

$$(1-\rho^2)^{1/2}\frac{\partial\beta}{\partial x_2} < x_1\left[\int_{x_1}^{\infty}\phi(t)\phi\left(\frac{x_2-\rho t}{\sqrt{(1-\rho^2)}}\right)dt - \{1-\Phi(x_1)\}\phi\left(\frac{x_2-\rho x_1}{\sqrt{(1-\rho^2)}}\right)\right]$$
$$= x_1\gamma(x_1, x_2; \rho).$$

When $x_1 > 0$, $(\partial\beta/\partial x_2) < 0$ if $\gamma(x_1, x_2; \rho) < 0$. Now

$$\lim_{x_1\to-\infty}\gamma(x_1, x_2; \rho) = \lim_{x_1\to\infty}\gamma(x_1, x_2; \rho) = 0, \tag{A.3}$$

and

$$\frac{\partial\gamma}{\partial x_1} = \frac{\rho}{1-\rho^2}\{1-\Phi(x_1)\}(x_2-\rho x_1)\phi\left(\frac{x_2-\rho x_1}{\sqrt{(1-\rho^2)}}\right).$$

This changes sign only at $x_1 = \rho^{-1}x_2$, and for this value of x_1,

$$\gamma(\rho^{-1}x_2, x_2; \rho) = \int_{\rho^{-1}x_2}^{\infty}\phi(t)\phi\left(\frac{x_2-\rho t}{\sqrt{(1-\rho^2)}}\right)dt - \{1-\Phi(\rho^{-1}x_2)\}(\sqrt{2\pi})^{-1}$$
$$< \int_{\rho^{-1}x_2}^{\infty}\phi(t)\cdot(\sqrt{2\pi})^{-1}\,dt - \{1-\Phi(\rho^{-1}x_2)\}(\sqrt{2\pi})^{-1} = 0.$$

Hence from (A.3), it follows that $\gamma(x_1, x_2; \rho) < 0$ for all x_1, x_2 and so

$$\partial\beta/\partial x_2 < 0 \qquad \text{for all} \quad x_2 \text{ and all } x_1 > 0.$$

Combining this with (A.2) we see that $\beta(x_1, x_2) > 0$ for all x_2 and all $x_1 > 0$. Hence

$$\partial h/\partial x_1 > 0 \qquad \text{for all} \quad x_2 \text{ and all } x_1.$$

(We have already noted that $\partial h/\partial x_1 > 0$ for $x_1 \leqslant 0$.)

References

[1] Barlow, R. E., Proschan, F., and Hunter, L. C. (1965). *Mathematical Theory of Reliability*. Wiley, New York.

[2] Basu, A. P. (1971). Bivariate failure rate. *J. Amer. Statist. Assoc.* **66**, 103–104.

[3] Block, H. W. (1973). Monotone hazard and failure rates for absolutely continuous multivariate distributions. *Research Report* #73-20, University of Pittsburgh.

[4] Brindley, E. C. and Thompson, W. A. (1972). Dependence and aging aspects of multivariate survival. *J. Amer. Statist. Assoc.* **67** 822–830.
[5] Dykstra, R. L., Hewett, J. E., and Thompson, W. A. (1973). Events which are almost independent. *Ann. Statist.* **1** 674–681.
[6] Goodman, I. R. (1972). Ph.D. Thesis, Temple University, Philadelphia.
[7] Harris, R. (1970). A multivariate definition for increasing hazard rate distribution functions. *Ann. Math. Statist.* **41** 713–717.
[8] Johnson, N. L. and Kotz, S. (1973). A vector-valued multivariate hazard rate. Institute of Statistics, University of North Carolina, Mimeo Series No. 873. (Part of this was also presented at the ISI meeting, Vienna, 1973.)
[9] Lehmann, E. L. (1966). Some concepts of dependence. *Ann. Math. Statist.* **37** 1137–1153.
[10] Marshall, A. W. (1973). Some comments on the hazard gradient. *Research Report.* Dept. of Statistics, University of Rochester, N. Y.
[11] Puri, P. and Rubin, H. (1974). On a characterization of the family of distributions with constant multivariate failure rates. *Annals of Probability* **2** 738–740.
[12] Yanagimoto, T. (1972). Families of positively dependent random variables. *Ann. Inst. Statist. Math., Tokyo* **24** 559–573.