Ann. Inst. Statist. Math.
Vol. 56, No. 2, 361–367 (2004)

A CHARACTERIZATION OF THE MULTIVARIATE NORMAL DISTRIBUTION BY USING THE HAZARD GRADIENT

JORGE NAVARRO AND JOSE M. RUIZ*

Facultad de Matemáticas, Universidad de Murcia, 30100 Murcia, Spain,
e-mail: jorgenav@um.es; jmruizgo@um.es

(Received July 15, 2002; revised February 17, 2003)

Abstract. We give a general result to characterize a multivariate distribution from a relationship between the left truncated mean function and the hazard gradient function. This result allows us to obtain new characterizations of multivariate distributions. In particular, we show that, for the multivariate normal distribution, the simple relationship, obtained in standardized form by McGill (1992, *Communications in Statistics. Theory Methods*, **21**(11), 3053–3060), actually characterizes the multivariate normal distribution.

Key words and phrases: Hazard gradient function, failure rate, mean residual life, left truncated mean function, multivariate normal.

1. Introduction

Let $X = (X_1, \dots, X_n)'$ be a random vector with density $f(x)$ and multivariate reliability (survival) function

$$R(x) = \Pr(X \geq x) = \Pr(X_1 \geq x_1, \dots, X_n \geq x_n)$$

for $x = (x_1, \dots, x_n)' \in \mathbb{R}^n$.

The multivariate left truncated mean is defined by

$$m(x) = E(X \mid X \geq x) = (m_1(x), \dots, m_n(x))'$$
$$m_i(x) = \frac{1}{R(x)} \int_{x_n}^{\infty} \cdots \int_{x_1}^{\infty} x_i f(x_1, \dots, x_n) dx_1 \cdots dx_n$$

and it is equivalent to the multivariate mean residual life $e(x) = m(x) - x$ introduced by Jupp and Mardia (1982) (see also Arnold and Zahedi (1988) or Shaked and Shanthikumar (1991) for another dynamic definition). It is well known that $m(x)$ (or $e(x)$) uniquely determines $R(x)$ (see Johnson and Kotz (1975), Shanbhag and Kotz (1987) and Ruiz *et al.* (1993)).

The multivariate hazard (failure) gradient was defined by Johnson and Kotz (1975) and Barlow and Proschan (1975) from

$$h(x) = -\nabla \log R(x) = (h_1(x), \dots, h_n(x))'$$

*Supported by Ministerio de Ciencia y Tecnologia under grant BFM2000-0362.

$$h_i(x) = -\frac{1}{R(x)} \frac{\partial}{\partial x_i} R(x)$$
$$\nabla = \left(\frac{\partial}{\partial x_1}, \ldots, \frac{\partial}{\partial x_n} \right)'$$

as the multivariate extension of the hazard (failure) rate function. The meaning is similar to the univariate case

$$\begin{aligned} h_i(x) &= \lim_{h \to 0} \frac{\Pr(X_1 \geq x_1, \ldots, x_i \leq X_i \leq x_i + h, \ldots, X_n \geq x_n)}{h \Pr(X_1 \geq x_1, \ldots, X_i \geq x_i, \ldots, X_n \geq x_n)} \\ &= \lim_{h \to 0} \Pr(X_i \leq x_i + h \mid X_1 \geq x_1, \ldots, X_i \geq x_i, \ldots, X_n \geq x_n)/h \end{aligned}$$

that is, the punctual failure probability for the i-th component when all the components are working and have age x_i, $i = 1, \ldots, n$. Note that $g(t) = h_i(x_1, \ldots, x_{i-1}, t, x_{i+1}, \ldots, x_n)$ is the failure rate for the conditional random variable

$$(X_i \mid X_1 \geq x_1, \ldots, X_{i-1} \geq x_{i-1}, X_{i+1} \geq x_{i+1}, \ldots, X_n \geq x_n).$$

Marshall and Olkin (1979) showed that $h(x)$ also uniquely determines $R(x)$ and Block (1977) used this function to obtain the multivariate increasing/decreasing failure (hazard) rate classes (MIFR and MDFR). Basu (1971) and Puri and Rubin (1974) defined the multivariate hazard function by $f(x)/R(x)$ while Shaked and Shanthikumar (1987) preferred a dynamic definition.

Kotz and Shanbhag (1980) characterized the univariate normal distribution by

$$m(x) = \mu + \sigma^2 h(x). \tag{1.1}$$

In Ruiz and Navarro (1994), a general way to obtain $f(x)$ from

$$m(x) = k + q(x)h(x)$$

is given. This result allows us to obtain some useful characterizations for distributions without an explicit expression for $h(x)$ or $m(x)$ (as in the case of the univariate normal distribution).

In the multivariate case, Ma (1998) and Asadi (1998, 1999) (see also the references given there) characterized the distribution function from different relationships between reliability functions. McGill (1992), Gupta and Gupta (1997) and Ma (2000) studied the hazard gradient for the multivariate normal and Roy and Mukherjee (1998) studied the extensions of univariate life distributions from reliability properties.

The purpose of this paper is to extend the results of Kotz and Shanbhag (1980) and Ruiz and Navarro (1994) to the multivariate case.

2. The main result

First, we obtain the characterization for the multivariate normal distribution.

THEOREM 2.1. *If X is a multivariate random vector with density $f(x)$, left truncated mean $m(x)$ and hazard gradient $h(x)$, then X has a (non degenerate) multivariate* *normal distribution with mean* μ *and variance-covariance matrix* $V = (\sigma_{ij})$ $(N_n(\mu, V))$ *if, and only if*

$$m(x) = \mu + Vh(x) \quad \textit{for} \quad x \in \mathbb{R}^n. \tag{2.1}$$

PROOF. Let us suppose that $X \equiv N_n(\mu, V)$, then

$$(X_1 \mid X_2 = x_2, \ldots, X_n = x_n) \equiv N_1(\mu_1 + V_{12}V_{22}^{-1}(y - \mu_Y), \sigma_{11} - V_{12}V_{22}^{-1}V_{21})$$

where $y = (x_2, \ldots, x_n)'$, $\mu_Y = (\mu_2, \ldots, \mu_n)'$ and

$$V = \begin{pmatrix} \sigma_{11} & V_{12} \\ V_{21} & V_{22} \end{pmatrix}.$$

If $n = 1$ then (2.1) holds (see Kotz and Shanbhag (1980)). By induction on n, let us suppose that (2.1) holds for $k < n$. Thus

$$\begin{aligned} R_1(x_1 \mid y)m_1(x_1 \mid y) = R_1(x_1 \mid y)(\mu_1 + V_{12}V_{22}^{-1}(y - \mu_Y)) \\ + (\sigma_{11} - V_{12}V_{22}^{-1}V_{21})f_1(x_1 \mid y) \end{aligned}$$

and

$$\begin{aligned} R_{2\ldots n}(y \mid x_1)m_{2\ldots n}(y \mid x_1) = R_{2\ldots n}(y \mid x_1)(\mu_Y + V_{21}\sigma_{11}^{-1}(x_1 - \mu_1)) \\ - (V_{22} - V_{21}\sigma_{11}^{-1}V_{12})\nabla R_{2\ldots n}(y \mid x_1), \end{aligned}$$

where

$$R_1(x_1 \mid y) = \int_{x_1}^{\infty} f_1(x_1 \mid y)dx_1$$

and

$$R_{2\ldots n}(y \mid x_1) = \int_{y}^{\infty} f_{2\ldots n}(y \mid x_1)dy.$$

Hence, if $m(x) = (m_1(x), \ldots, m_n(x))'$, then

$$\begin{aligned} R(x)m_1(x) &= \int_x^{\infty} x_1 f(x)dx \\ &= \int_y^{\infty} f_{2\ldots n}(y) \int_{x_1}^{\infty} x_1 f_1(x_1 \mid y)dx_1 dy \\ &= \int_y^{\infty} f_{2\ldots n}(y)[R_1(x_1 \mid y)(\mu_1 + V_{12}V_{22}^{-1}(y - \mu_Y)) \\ &\qquad + (\sigma_{11} - V_{12}V_{22}^{-1}V_{21})f_1(x_1 \mid y)]dy \\ &= \int_x^{\infty} (\mu_1 + V_{12}V_{22}^{-1}(y - \mu_Y))f(x)dx + (\sigma_{11} - V_{12}V_{22}^{-1}V_{21}) \int_y^{\infty} f(x)dy \\ &= R(x)\mu_1 + V_{12}V_{22}^{-1}I + (\sigma_{11} - V_{12}V_{22}^{-1}V_{21})\frac{-\partial R(x)}{\partial x_1}, \end{aligned}$$

where

$$\begin{aligned}
I &= \int_{x_1}^{\infty} f_1(x_1) \int_{y}^{\infty} (y - \mu_Y) f_{2\ldots n}(y \mid x_1) dy dx_1 \\
&= \int_{x_1}^{\infty} f_1(x_1)[R_{2\ldots n}(y \mid x_1) V_{21}\sigma_{11}^{-1}(x_1 - \mu_1) - (V_{22} - V_{21}\sigma_{11}^{-1}V_{12})\nabla R_{2\ldots n}(y \mid x_1)]dx_1 \\
&= \int_{x}^{\infty} V_{21}\sigma_{11}^{-1}(x_1 - \mu_1) f(x) dx \\
&\quad - (V_{22} - V_{21}\sigma_{11}^{-1}V_{12}) \int_{x_1}^{\infty} f(x) \left(\nabla \int_{y}^{\infty} f_{2\ldots n}(y \mid x_1) dy \right) dx_1 \\
&= V_{21}\sigma_{11}^{-1} R(x)(m_1(x) - \mu_1) - (V_{22} - V_{21}\sigma_{11}^{-1}V_{12}) \left(\frac{\partial R(x)}{\partial x_2}, \ldots, \frac{\partial R(x)}{\partial x_n} \right)'
\end{aligned}$$

and, finally, we obtain

$$\begin{aligned}
&(1 - V_{12}V_{22}^{-1}V_{21}\sigma_{11}^{-1})R(x)(m_1(x) - \mu_1) \\
&\quad = (\sigma_{11} - V_{12}V_{22}^{-1}V_{21}) \frac{-\partial R(x)}{\partial x_1} \\
&\qquad - V_{12}V_{22}^{-1}(V_{22} - V_{21}\sigma_{11}^{-1}V_{12}) \left(\frac{\partial R(x)}{\partial x_2}, \ldots, \frac{\partial R(x)}{\partial x_n} \right)'
\end{aligned}$$

and

$$\begin{aligned}
m_1(x) &= \mu_1 + \sigma_{11}h_1(x) + V_{12}(h_2(x), \ldots, h_n(x))' \\
&= \mu_1 + \sum_{j=1}^{n} \sigma_{1j} h_j(x).
\end{aligned}$$

Analogously, we can obtain the expressions for $m_i(x)$ when $i = 2, \ldots, n$. Conversely, if (2.1) holds, then $m_i(x) = \mu_i + \sum_{j=1}^{n} \sigma_{ij} h_j(x)$,

$$\int_{x}^{\infty} (x_i - \mu_i) f(x) dx = -\sum_{j=1}^{n} \sigma_{ij} \frac{\partial R(x)}{\partial x_j}$$

and, by differentiating,

$$(-1)^n (x_i - \mu_i) f(x) = -\sum_{j=1}^{n} \sigma_{ij} \left(\frac{\partial^n}{\partial x_1 \ldots \partial x_n} \frac{\partial R(x)}{\partial x_j} \right)$$

holds, which jointly with

$$\frac{\partial^n R(x)}{\partial x_1 \ldots \partial x_n} = (-1)^n f(x)$$

imply

$$(x_i - \mu_i) f(x) = -\sum_{j=1}^{n} \sigma_{ij} \frac{\partial f(x)}{\partial x_j}.$$

Thus,

$$(x - \mu) f(x) = -V(\nabla f(x))$$

holds, which implies $X \equiv N_n(\mu, V)$. □

Remark 2.1. McGill (1992) calculated $m(x)$ by using the hazard gradient for the standardized variables $(Z_i = (X_i - \mu)/\sigma_i)$. Hence he did not obtain expression (2.1). Moreover, he did not give the reverse implication (he did not characterize the multivariate normal distribution). Note that the multivariate normal distribution cannot be characterized from inversion formulas for the mean residual life $(e(x))$ or the multivariate hazard gradient $(h(x))$ because, in this case, we do not have an explicit expression for $e(x)$ $(m(x))$ nor $h(x)$.

Remark 2.2. We have obtained the following general result for absolutely continuous random vectors.

If X is a random vector with density $f(x)$, support S, left truncated mean $m(x)$ and hazard gradient $h(x)$, then the following conditions are equivalent:

1. $m(x) = k + Q(x)h(x)$, for all $x \in S$
2. $\nabla \log f(x) = Q^{-1}(x)(k - x - Q(x)\nabla)$, for all $x \in S$

where $k = (k_1, \ldots, k_n)' \in \mathbb{R}^n$, $x = (x_1, \ldots, x_n)' \in \mathbb{R}^n$,

$$Q(x) = \begin{pmatrix} q_{11}(x_1) & \cdots & q_{1n}(x_n) \\ \cdots & \cdots & \cdots \\ q_{n1}(x_1) & \cdots & q_{nn}(x_n) \end{pmatrix},$$

$$Q(x)\nabla = \begin{pmatrix} q_{11}(x_1) & \cdots & q_{1n}(x_n) \\ \cdots & \cdots & \cdots \\ q_{n1}(x_1) & \cdots & q_{nn}(x_n) \end{pmatrix} \begin{pmatrix} \dfrac{\partial}{\partial x_1} \\ \cdots \\ \dfrac{\partial}{\partial x_n} \end{pmatrix} = \begin{pmatrix} \displaystyle\sum_{j=1}^{n} \frac{\partial}{\partial x_j} q_{1j}(x_j) \\ \cdots \\ \displaystyle\sum_{j=1}^{n} \frac{\partial}{\partial x_j} q_{nj}(x_j) \end{pmatrix},$$

and there exists $Q^{-1}(x)$ for all x and $\lim_{x_j \to \infty} q_{ij}(x_j) \frac{\partial}{\partial x_j} R(x) = 0$, $i, j = 1, \ldots, n$.

The proof can be obtained from the authors. We have also obtained a similar result in the discrete case.

Remark 2.3. If $\lim_{x \to -\infty} Q(x)h(x) = 0$, then $k = E(X)$. In particular, if X is a positive random vector and $Q(\mathbf{0})h(\mathbf{0}) = 0$ then $k = E(X)$ (where $\mathbf{0} = (0, \ldots, 0)' \in \mathbb{R}^n$). If we take $k = \mathbf{0}$, then the preceding theorem provides a criterion to characterize the distribution function from this relationship $m(x) = Q(x)h(x)$. Note also that we can replace the left truncated mean function $m(x) = E(X \mid X \geq x)$ by the mean residual life function $e(x) = E(X - x \mid X \geq x) = m(x) - x$.

Remark 2.4. In the univariate case, the characterization theorem given in Ruiz and Navarro (1994) is a general result because all the distributions can be characterized from a relationship as

$$m(x) = k + q(x)h(x)$$

since we can define $q(x)$ by

$$q(x) = (m(x) - k)/h(x).$$

However, in the multivariate case, this result cannot be applied to all the multivariate distributions. That is, there exist distributions which do not satisfy the relation $m(x) = k+Q(x)h(x)$, for a matrix-function $Q(x) = (q_{ij}(x_j))$ (i.e. we cannot obtain $Q(x)$ from k, $m(x)$ and $h(x)$). For example, the theorem cannot be applied to the bivariate Gumbel's exponential distribution.

Note also that if $Q(x) = (q_{ij}(x))$, then there exist infinite matrix-functions $Q(x)$ satisfying $m(x) = k + Q(x)h(x)$. For example, we can make $q_{ij}(x) = 0$ for $j \neq i$ and $q_{ii}(x) = (m_i(x)-k)/h_i(x)$. The extension of the preceding theorem to all the multivariate distributions is an open question.

Remark 2.5. The characterization for the multivariate normal distribution can be easily obtained from the general result, without a direct calculation for $m(x)$, as follows

$$\log f(x) = \log c - \frac{1}{2}(x-\mu)'V^{-1}(x-\mu)$$
$$\nabla \log f(x) = -V^{-1}(x-\mu) = V^{-1}(\mu - x - V\nabla).$$

Acknowledgements

The authors gratefully acknowledge the many helpful suggestions made by Professor Samuel Kotz and the anonymous referees during the preparation of the paper.

References

Arnold, B. and Zahedi, H. (1988). A note on the characterization of bivariate densities by conditional densities, *Communications in Statistics*, **13**(3), 395–409.

Asadi, M. (1998). Some general characterizations of the bivariate Gumbel distribution and the bivariate Lomax distribution based on truncated expectations, *Journal of Multivariate Analysis*, **67**, 190–202.

Asadi, M. (1999). Multivariate distributions characterized by a relationship between mean residual life and hazard rate, *Metrika*, **49**, 121–126.

Barlow, R. E. and Proschan, F. (1975). *Statistical Theory of Reliability and Life Testing*, Holt, Rinehart and Winston, New York.

Basu, A. P. (1971). Bivariate failure rate, *Journal of the American Statistical Association*, **66**, 103–104.

Block, H. W. (1977). Monotone failure rate for multivariate distributions, *Naval Research Logistics*, **24**, 627–637.

Gupta, P. L. and Gupta, R. C. (1997). On the multivariate normal hazard, *Journal of Multivariate Analysis*, **62**, 64–73.

Johnson, N. L. and Kotz, S. (1975). A vector of multivariate hazard rate, *Journal of Multivariate Analysis*, **5**, 53–66.

Jupp, P. E. and Mardia, K. V. (1982). A characterization of the multivariate Pareto distribution, *The Annals of Statistics*, **10**, 1021–1024.

Kotz, S. and Shanbhag, D. N. (1980). Some new approaches to probability distributions, *Advances in Applied Probability*, **12**, 903–921.

Ma, C. (1998). Characteristic properties of multivariate survival functions in terms of residual life distribution, *Metrika*, **47**, 227–240.

Ma, C. (2000). A note on the multivariate normal hazard, *Journal of Multivariate Analysis*, **73**, 282–283.

Marshall, A. W. and Olkin, I. (1979). *Inequalities: Theory of Majorization and Its Applications*, Academic Press, San Diego, California.

McGill, J. I. (1992). The multivariate hazard gradient and moments of the truncated multinormal distribution, *Communications in Statistics. Theory and Methods*, **21**(11), 3053–3060.

Puri, P. S. and Rubin, H. (1974). On a characterization of the family of distributions with constant multivariate failure rates, *The Annals of Probability*, **2**, 738–740.

Roy, D. and Mukherjee, S. P. (1998). Multivariate extensions of univariate distributions, *Journal of Multivariate Analysis*, **67**, 72–79.

Ruiz, J. M. and Navarro, J. (1994). Characterization of distributions by relationships between failure rate and mean residual life, *IEEE Transactions on Reliability*, **43**, 640–644.

Ruiz, J. M., Marín, J. and Zoroa, P. (1993). A characterization of continuous multivariate distributions by conditional expectations, *Journal of Statistical Planning and Inference*, **37**, 13–21.

Shaked, M. and Shanthikumar, J. G. (1987). Multivariate hazard rates and stochastic ordering, *Advances in Applied Probability*, **19**, 123–137.

Shaked, M. and Shanthikumar, J. G. (1991). Dynamic multivariate mean residual life functions, *Journal of Applied Probability*, **28**, 613–629.

Shanbhag, D. N. and Kotz, S. (1987). Some new approaches to multivariate probability distributions, *Journal of Multivariate Analysis*, **22**, 189–211.