〈Research Paper〉

Journal of the Korean
Statistical Society
Vol. 18. No1, 1989

Characterization of Some Multivariate Distributions

N. Unnikrishnan Nair*

ABSTRACT

In this article the problem of characterizing multivariate distributions, possessing certain conditional distributions that have the same form as the parent model, are considered. It is shown that the forms of such conditional distributions characterize some well known distributions like the multivariate exponential, multivariate Burr, multivariate Lomax etc.

1. Introduction

It is well known that a multivariate distribution is not always uniquely determined by its marginal densities. However, methods of constructing bivariate distributions with specified marginals have been considered by many authors like Morgenstern (1956), Farlie (1960) and Mardia (1970). Kotz (1975) has generalised some of these ideas to the multivariate case, in which he has also strssed the use of survival functions in the process. A some what similar problem is the determination of a multivariate distribution when the conditional distributions are known. Abrahams and Thomas (1984) answered this question by obtaining a necessary and sufficient condition for the conditional densities $f_1(x \mid y)$ and $f_2(y \mid x)$ to determine a bivariate distribution uniquely, as

$$f_1(x \mid y) / f_2(y \mid x) = g(x) / h(y) \tag{1}$$

* Department of Mathematics and Statistics Cochin University of Science and Technology Cochin, 682022 INDIA

where g and h are non-negative integrable functions with equal integrals. In the present article a similar result in the more general multivariate case is attempted based on a property of the ratio of two conditional survival functions. The shift in the focus of attention from densities to survival functions enables us to draw conclusions about the forms of all marginal and conditional distributions involved, a fact which could not be accomplished in the earlier work. We give several examples of multivariate distributions such as Burr, exponential, Pareto etc. that can be characterized by the property of identical forms for certain conditional distributions.

2. Basic result

Consider a random vector $\underline{X}=(X_1, X_2,\cdots\cdots, X_n)$ possessing absolutely continuous distribution with respect to Lebesgue measure in the support of $R_n^+=\{(x_1, x_2,\cdots\cdots,x_n) \mid x_i>0, i=1,2,\cdots\cdots, n\}$. Joint survival function of $\underline{X}$ is

$$S(\underline{t}) = P[\underline{X}>\underline{t}]$$

where $\underline{t}=(t_1,t_2,\cdots\cdots, t_n)$ is a vector of non-negative reals and $\underline{X}>\underline{t}$ stands for the event $\bigcap_1^n (X_i>t_i)$. Let E_r denote the set of r-tuples formed out of the elements of $\underline{X}$ and $\underline{Z}$ be an element of E_r. Defining $\underline{Z}^*$ as the compliment of $\underline{Z}$ with respect to $\underline{X}$, the conditional survival function of $\underline{Z}$ given $\underline{Z}^*$ is

$$S(t_{(r)} \mid t^*_{(r)}) = P[\underline{Z}>t_{(r)} \mid \underline{Z}^* > t^*_{(r)}] \qquad (2)$$

whth $t_{(r)}$ and $t^*_{(r)}$ as the realizations of Z and Z^* corresponding to $\underline{t}$ of $\underline{X}$. It is to be noted that the survival function(2) determines the conditional distribution of $\underline{Z}$ given $\underline{Z}^*> t^*_{(r)}$ uniquely. By conditional distribution in this paper, we mean distributions of the type just mentioned in contrast to the usual notion of $\underline{Z}^*=t^*_{(r)}$.

Theorem 1. The density function of X is uniquely determined by the surival functions $S_1(t_{(r)} \mid t^*_{(r)})$ and $S_2(t^*_{(r)} \mid t_{(r)})$ at those points for which these functions are non-zero if and only if

$$S_1/S_2 = g(t_{(r)})/h(t^*_{(r)}) \qquad (3)$$

for an arbitrary Z in E_r and any one $r=1,2,\cdots, n$, provided $g(\cdot)$ and $h(\cdot)$ are non-negative real functions with continuous partial derivatives in the subspaces R_r^+ and R_{n-r}^+ respectively, satisfying $g(0+) = h(0+)$

Proof. To prove the sufficiency of the conditions we note that the left side exists finitely as the ratio of two absolutely continuous functions. Further, there exist continuous function u in R_r^+ and V in R_{n-r}^+ such that

$$g(t_{(r)}) = \int_{t_{(r)}}^{\infty} u(\underline{y})\,d\underline{y}$$

and

$$h(t^*_{(r)}) = \int_{t^*_{(r)}}^{\infty} v(\underline{y})\,d\underline{y}$$

where the single integrals stand for multiple integrals in relevant dimensions. Equation (3) is equivalent to

$$\frac{S_1}{S_2} = \frac{\int_{t_{(r)}}^{\infty} u(\underline{y})\,d\underline{y} \Big/ \int_0^{\infty} u(\underline{y})\,d\underline{y}}{\int_{t^*_{(r)}}^{\infty} v(\underline{y})\,d\underline{y} \Big/ \int_0^{\infty} v(\underline{y})\,d\underline{y}} \tag{4}$$

Denoting by $S(t_{(r)})$, the expressions on the numerator of the right side in the last equation, we see that S is non-increasing in $t_{(r)}$, $S(+\infty)=0$ and $S(0)=1$.
Further

$$S(t_{(r)}+\underline{h}) - S(t_{(r)}) = -\int_{t_{(r)}}^{t_{(r)}+\underline{h}} u(\underline{y})\,d\underline{y} \Big/ \int_0^{\infty} u(\underline{y})\,d\underline{y}$$

so that $S(t_{(r)}+\underline{0}) = S(t_{(r)})$, proving the right continuity of S. Consequently S is the survival function of Z. Similarly the denominator on the right of (4) is the survival function $S(t^*_{(r)})$ of $\underline{Z}^*$. From this the distribution of $\underline{X}$ is uniquely specified through

$$S(\underline{t}) = S_1(t_{(r)} \mid t^*_{(r)}) \cdot S(t^*_{(r)})$$

The necessary part follows by equating the two representations of $S(\underline{t})$ in terms of the conditional and marginal survival functions. This completes the proof. ■

It is obsered that (3) gives a class of necessary and sufficient conditions corresponding to the various $\underline{Z}$'s in $E_1, E_2, \cdots\cdots, E_n$. However, representation (3) for an arbitrary $\underline{Z}$ in E_r does not suffice to guarantee a unique joint distribution whose form is identical with that of all marginal and conditional distributions. The relevant modifications are prescrebed in the next section.

3. Characterizations

We now proceed to characterize several multivariate distributions by assuming well known forms for the conditional survival functions in addition to the conditions of Theorem and some modifications thereof. Apart from the forms assumed for the conditionals, other general conditions remain the same for all distributions.

(1) Multivariate exponential distribution.

The model illustrated here is the general case of the bivariate exponential distribution of Gumbel (1960) with survival function

$$S(\underline{t}) = \exp\left[-\Sigma a_i t_i - \sum_{i<j} a_{ij} t_i t_j - \cdots - a_{12\cdots n} t_1 t_2 \cdots t_n\right] \qquad (5)$$

where the a's are all non-negative. We characterize (5) by stating

Theorem 2. A necessary and sufficient condition for the distribution of $\underline{X}$ and all its marginals and conditionals to be of multivariate exponential form is that for every $\underline{Z}$ in E_r and any one $r= 1,2,\cdots, n$ the survival functions S_1 correspond to exponential form in appropriate dimensions for all $\underline{t}>\underline{o}$.

Proof. We specialize to the case $r=1$, as the proof is similarly obtained for all other values of r. In this case, all X_i's have univariate exponential distributions for $X_j>t_j$, $i\neq j$.
That is

$$S(t_{(1)} \mid t^*_{(1)}) = \exp\left[-a(t^*_{(1)})\, t_{(1)}\right] \qquad (6)$$

where $t_{(i)}= t_i$ for $i= 1,2\cdots\cdots, n$. Allowing t_i to tend to zero for $i>3$, when $t_{(i)}=t_1, t_2$

$$S(t_1 \mid t_2) = \exp\left[-a_1(t_2)\, t_1\right]$$

$$S(t_2 \mid t_1) = \exp\left[-a_2(t_1)\, t_2\right]$$

Applying Theorem 1 to these survival functions by taking $g(t_1)= \exp[-a_1t_1]$ and $h(t_2) = \exp[-a_2t_2]$ for some $a_1, a_2> 0$, it is seen that the distribution of (X_1,X_2) is uniquely determined. We now show that this bivariate model is of form (5) with $n=2$. From representation (3)

$$a_1(t_2)\, t_1 + a_2 t_2 = a_2(t_1)\, t_2 + a_1 t_1 \qquad (7)$$

or

$$(a_1(t_2)-a_1)t_2^{-1}=(a_2(t_1)-a_2)t_1^{-1}$$

The left side is a function of t_2 only and the right side contains only t_1. That this holds for t_1, $t_2>0$ implies that either is a constant, say a_{12}. Thus the only solution of (7) is

$$a_i(t_j)=a_i+a_{12}t_j, i, j=1,2, i\neq j \tag{8}$$

Using (9)

$$S(t_1,t_2)=\exp[-a_1t_1-a_2t_2-a_{12}t_1t_2]$$

The non-negativity of S for all t_1 and t_2 ensures $a_{12}\geq 0$. Considering other pairs of t values, it is easy to see that for every pair (X_i,X_j), $i\neq j=1,2,\cdots,n$

$$S(t_i,t_j)=\exp[-a_it_i-a_jt_j-a_{ij}t_it_j] \tag{9}$$

In the second stage we let t_j tend to zero for $j>3$ in (5) and write

$$S(t_i|t_jt_l)=\exp[-a_i(t_j,t_l)t_i] \tag{10}$$

for $i\neq j\neq l$ and $i,j,l=1,2,3$. By virtue of (9) and (10) three different expressions for $S(t_1,t_2,t_3)$ are obtained. When equated in pairs, leaves functional equations in $a_l(t_j,t_l)$ of the form

$$\begin{aligned}&a_1(t_2,t_3)t_1+a_2t_2+a_{23}t_2t_3+a_3t_3\\&=a_2(t_1,t_3)t_2+a_1t_1+a_3t_3+a_{13}t_1t_3\\&=a_3(t_1,t_2)t_3+a_1t_1+a_2t_2+a_{12}t_1t_2\end{aligned} \tag{11}$$

The first pair of equations gives

$$a_1(t_2,t_3)t_1-a_1t_1-a_{13}t_1t_3=a_2(t_1,t_3)t_2-a_2t_2-a_{23}t_2t_3$$

This being an identity holding good for all $t_1,t_2,t_3>0$, the linearity in t_1 of the left side implies that $a_2(t_1,t_2)$ contains only first degree terms in t_1. The same arguments extends to the other functions in (11). This leaves solutions of the form

$$a_1(t_2,t_3)=a+bt_2+ct_3+a_{123}t_2t_3$$

$$a_2(t_1,t_3)=a'+b't_1+c't_3+a'_{123}t_1t_3$$

Substituting in (11) and equating coefficients of like terms lead to

$$a_1(t_2,t_3)=a_1+a_{12}t_2+a_{13}t_3+a_{123}t_2t_3$$

and two other similar expressions. Introducing these in (10) and multiplying the resulting expression by $S(t_j,t_l)$ obtained form (9), the survival function of (X_1,X_2,X_3) is obtained as

$$S(t_1,t_2,t_3)=\exp\left[-\sum_1^3 a_i t_i-\sum_{i<j} a_{ij}t_it_j-a_{123}t_1t_2t_3\right]$$

which conforms to (5) with n=3. The proof for higher dimensions follows by induction on n. From the course of the proof it is obvious that all the marginal distributions and the conditional distributions not encountered here are of identical form in relevant dimensions. When X follows, survival function (5), direct calculations by setting appropriate combinations of thet's to zero establishes the required form for the marginals. This completes the proof. ■

(2) **Multivariate Burr Distribution.**

Consider the multivariate Burr distribution in R_n^+ (obtained by generalising the Durling's (1975) bivariate form) specified by

$$S(t)=\left[1+\sum_1^n a_it_i^{c_i}+\sum_{i<j} a_{ij}t_i^{c_i}t_j^{c_j}+\cdots+a_{12\cdots n}\ t_1^{c_1}\cdots t_n^{c_n}\right]^{-k}$$

where, all a, c_i, k>0. The statement of Theorem 2 holds here also with the change that in the present case S_1 should have multivariate Burr distributions in R_r^+. Since the arguments are identical to that in Theorem 2, we give here only the important steps in the proof.

In the special case r=1,

$$S(t_i\,|\,t_j)=\left[1+a_i(t_j)t_i^{c_i}\right]^{-k},\ i,j=1,2,\ i\neq j \tag{12}$$

Taking $g(t_1)=(1+a_1t_1^{c_1})^{-k}$ and $h(t_2)=(1+a_2t_2^{c_2})^{-k}$ for some $a_1,a_2,k>0$ by Therorem 1, the distribution of (X_1,X_2) is uniquely determined by the functions in (12). Equating the two different expressions for $S(t_1,t_2)$ and simplifying

$$(1/t_2^{c_2})(a_1(t_2)-a_1)+a_1(t_2)a_2=(1/t_1^{c_1})(a_2(t_1)-a_2)+a_2(t_1)a_1 \qquad (13)$$

for all $t_1, t_2, > O$. The solution of (13) as in (7) is,

$$a_j(t_i)=(a_j+a_{12}t_i^{c_i})/(1+a_i t_i^{c_i}) \quad j\neq i,\ \ i,j=1,2$$

Thus

$$S(t_1,t_2)=(1+a_1t_1^{c_1}+a_2t_2^{c_2}+a_{12}t_1^{c_1}\ t_2^{c_2}),$$

proving the result for n=2. The proof to higher to dimensions and conclusions are omitted as they have identical pattern with example 1. We note that multivariate Burr distribution of Takahasi (1965) is realised when the a's with more than one suffix tend to zero, so that the relevant properties hold for that model as well.

(3) Multivariate Lomax distribution

This distribution introduced by Lindley and Singpurwalla (1986) and investigated further by Nayak(1987) has survival function,

$$S(\underline{t})=(1+\sum_{1}^{n} a_i t_i)^{-k}, \quad a_i, t_i,\ k>0 \qquad (13)$$

Being a special case of the multivariate Burr distribution, all the required results follow from example (2). Interestingly, the multivariate Pareto distribution of Mardia (1962), bivariate model of Hutchison (1979) share common properties with model (13). The Gumbel (1961) bivariate logistics, its extension by Satterthwaite and Hutchison (1978) and also the logistic model of Malik and Abraham (1973), can be obtained by transformations (see Nayak(1987) for details) of the Lomax model. Even though the distributions in examples (2) and (3) are related, it does not appear that the multivariate model in (1) is related to the others inspite of the common feature all of them possess.

In conclusion, we emphasize the distinction between the above results with that of Abrahams and Thomas (1984), on account of using different types of conditional distributions used in each. For example by assuming conditional densities $f_1(x \mid y)$ and $f_2(y \mid x)$ to be of exponential form, the latter obtains a bivariate distribution different form the Gumbel's variety. On the other hand the bivariate exponential we have considered does not have exponential conditionals in the sense of Abrahams and Thomas (1984).

References

(1) Abrahams, J and Thomas, J.B. (1984). A note on characterization of bivariate densities by conditional densities, *Communications in Statistics* Vol. 13(3), 395-400.

(2) Durling, F. C. (1975). The bivariate Burr distribution, *Statistical Distributions in scientific work,* D. Reidel, Holland, 329-335.

(3) Farlie, D.J.G.(1960). The performance of some correlation coefficients for a general bivariate distribution, *Biometrika,* Vol. 47, 307-323.

(4) Gumbel, E.J.(1960). Bivariate exponential distributions, *Journal of the American Statistical Assciation,* Vol.55. 698-707.

(5) Gumbel, E.J.(1961). Bivariate logistic distributions, *Journal of the American Statistical Association,* Vol.56. 335-349.

(6) Hutchison, T.P. (1979). Four applications of a bivariate Pareto distribution, *Bionetrical Journal.* Vol. 21. 553-563.

(7) Kotz, S. (1975). Multivariate distributions at a cross road, *Statistical Distributions in scientific work,* Vol. I, 247-270 D. Reidel, Holland.

(8) Lindley D.V and Singpurwalla, N.D.(1986). Multivariate distribution for the life lengths of a system sharing a common environment, *Journal of Applied Probability,* Vol. 23 418-431.

(9) Malik, H.J and Abraham, B.(1973). Multivariate Logistic distributions, *Annals of Statistics,* Vol 3. 588-590.

(10) Mardia, K.V.(1962). Multivariate Pareto distribution, *The Annals of Mathematical Statistices.,* Vol. 35, 170-177.

(11) Mardia, K.V. (1970). *Families of bivariate distributions,* Griffin, London.

(12) Morgenstern, D.(1956). Einf ache Beispiele Zwei dimensionaler Ver teilungen, *Mitt. Math. Statist.* Vol. 8 234-235.

(13) Nayak, T.K.(1987). Multivariate Lomax distribution Properites and use fulness in reliability theory, *Jornal of Applied Probabiliay* Vol. 24 170-177.

(14) Satterth Waite, S.P. and Hutchison, T.P.(1978). A generalization of Gumbel's bivariate logist distribution, *Metrika,* Vol. 25 163-170.

(15) Takahasi, K.(1965). Note on the multivariate Burr's distribution, *Annals of the Insitute of statistical Mathematics* Vol.17, 257-260.