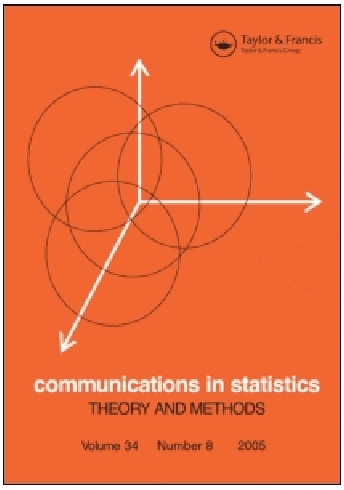

Communications in Statistics - Theory and Methods
Publication details, including instructions for authors and subscription information:
http://www.informaworld.com/smpp/title~content=t713597238

Conditional specifications of multivariate pareto and student distributions

Jacek Wesolowski [a]; Mohammad Ahsanullah [b]
[a] Mathematical Institute, Warsaw University of Technology, Warsaw, Poland [b] Department of Management Sciences, Rider University, Lawrenceville, NJ, U.S.A

To cite this Article Wesolowski, Jacek and Ahsanullah, Mohammad(1995) 'Conditional specifications of multivariate pareto and student distributions', Communications in Statistics - Theory and Methods, 24: 4, 1023 — 1031
To link to this Article: DOI: 10.1080/03610929508831537
URL: http://dx.doi.org/10.1080/03610929508831537

COMMUN. STATIST.—THEORY METH., 24(4), 1023–1031 (1995)

CONDITIONAL SPECIFICATIONS OF MULTIVARIATE PARETO AND STUDENT DISTRIBUTIONS

Jacek Wesolowski
Mathematical Institute, Warsaw University of Technology
Plac Politechniki 1, 00-661 Warsaw, Poland
e-mail: wesolo@plwatu21.bitnet
and
Mohammad Ahsanullah
Department of Management Sciences, Rider University
Lawrenceville, NJ, 08648-3099, U.S.A.
e-mail: ahsan@rider.bitnet

Key words and phrases: multivariate Pareto distribution, multivariate Student distribution, multivariate Cauchy distribution, Student conditional distribution, Pareto conditional distribution, conditional specification, characterization of probability distributions.

AMS (1980) subject classification: 62H05.

ABSTRACT

A random vector has a multivariate Pareto distribution if one of its univariate conditional distribution is Pareto and some of its marginals are identi-

cally distributed. A general method developed in the course of the proof of this result is applied also to characterize the multivariate Student (Cauchy) measure by one univariate Student conditional distribution.

1. Introduction

In this paper we are interested in a specification of some multivariate distributions by assuming that one of univariate conditional distributions is of a special form and that some marginal distributions are identical. The first result of such a type was obtained by Ahsanullah (1985), where it was proved that in the bivariate case one normal conditional distribution and equidistribution of marginals guarantee the bivariate normality. In that paper also the following conjecture was formulated: Let $X = (X_1, \cdots, X_n)$ be a random vector fulfilling the conditions:

(i) The conditional distribution of X_n given $X_1, \cdots, X_{n-1}$ is normal $\mathcal{N}(\alpha_0 + \sum_{j=1}^{n-1} \alpha_j X_j, \sigma^2)$, where $\alpha_0, \alpha_1, \cdots, \alpha_{n-1}, \sigma$ are some real constants and $\sigma > 0$.

(ii) The r.v's $X_1, \cdots, X_n$ are identically distributed.

Is it true that X has a n-variate Gaussian distribution?

That question was the beginning of the general discussion on possibility of characterization of distribution of any n-variate random vector X by specifying the conditional distribution $X_n|X_1, \cdots, X_{n-1}$ and assuming identity of distribution for some marginals. Now we recall main results of that discussion to give a reasonable background for the main results of this paper.

Ahsanullah and Sinha (1986) provided a counter example showing that the conjecture is false. They proved that if instead of (ii) it is assumed that X is exchangeable then the conjecture holds. In Arnold and Pourahmadi (1988) (we refer to it by AP in the sequel) it was shown that excheangability is much too strong. Applying a Markovian scheme they obtained a characterization of multivariate normality by (i) and the following equidistribution condition:

$$(X_1, \cdots, X_{n-1}) \stackrel{d}{=} (X_2, \cdots, X_n) \tag{1}$$

($\stackrel{d}{=}$ stands for equation in distribution). In a recent paper by Ahsanullah and Wesolowski (1994) (we refer to it by AW in the sequel) (*ii*) was replaced by

$$(X_0, X_1, \cdots, X_k) \stackrel{d}{=} (X_0, X_1, \cdots, X_{k-1}, X_{k+1}), \quad k = 1, \cdots, n-1, \qquad (2)$$

$X_0 = 0$ a.s. and the n-variate normal distribution was characterized by (*i*) and (2).

In AP also other multivariate distributions were considered. However its main result concerns unique determination of the bivariate measure by one conditional distribution and identically distributed marginals provided some indecomposability condition is fulfilled.

Our aim is to develop a quite general method of characterizing multivariate probability measures by (2) and a conditional distribution of X_n given $X_1, \cdots, X_{n-1}$. It is done while treating the multivariate Pareto distribution in Section 2. The approach is applied to the multivariate Student distribution (which has Student conditionals) in Section 3.

2. Characterization of the multivariate Pareto distribution

As an example of application of the main result of AP a characterization of the bivariate Pareto distribution was obtained there. Let us state it now in a form of the solution of some homogeneous Fredholm integral equation of the second kind we will use later.

Lemma 1 *If for all $x > 0$*

$$f(x) = \int_0^\infty \frac{(\beta+1)(\gamma+y)^{\beta+1}}{(\gamma+y+x)^{\beta+2}} f(y)\, dy \quad ,$$

$\beta > 0$, $\gamma > 0$, *where f is a probability density function (pdf), then*

$$f(x) = \frac{\beta\gamma^\beta}{(\gamma+x)^{\beta+1}}, \quad x > 0,$$

i.e. it is a pdf of the Pareto $\mathcal{P}(\beta, \gamma)$ distribution.

In this section we investigate multivariate measures generated by r.v's $X = (X_1, \cdots, X_n)$ with the following property: the conditional distribution of X_n given $X_1, \cdots, X_{n-1}$ is Pareto with parameters $\alpha+n-1$ and $\sum_{i=1}^{n-1} X_i+1$, where α is a positive constant, i.e.

$$\mu_{X_n|X_1,\cdots,X_{n-1}} = \mathcal{P}(\alpha + n - 1, \sum_{i=1}^{n-1} X_i + 1) \ . \tag{3}$$

In AP it was pointed out that along the lines of the Markovian approach used in the proof of characterization of the multivariate Gaussian measure the following result may be obtained.

Theorem 1 *Let $X = (X_1, \cdots, X_n)$ be a random vector fulfilling* (1) *and* (3). *Then it has a Mardia multivariate Pareto distribution with the pdf of the form*

$$f(x_1, \cdots, x_n) = \frac{(\alpha + n - 1)_n}{(\sum_{i=1}^n x_i + 1)^{\alpha+n}} \ , \tag{4}$$

$x_i > 0$, $i = 1, \cdots, n$, *where* $(\alpha + n - 1)_n = \alpha(\alpha + 1) \cdots (\alpha + n - 1)$.

This kind of a multivariate Pareto distribution was introduced in Mardia (1962) and proved to be useful in many applications; for more details see Arnold (1983).

This section is devoted to a parallel problem of characterizing the Mardia multivariate Pareto distribution by (2) and (3). A method developed in the proof of the main result works also for other conditional distributions (see Section 3).

Before presenting the main result let us give a brief survey of other possible conditional specifications of multivariate Pareto measures. In Arnold (1987) a characterization of a bivariate Pareto distribution through both Pareto conditionals without giving the precise forms of parameters was given. Then it was slightly extended in Castillo and Sarabia (1990). These results are reproduced in a recent monograph on conditional specifications by Arnold et al. (1992), where also a multivariate case involving all univariate Pareto conditionals was treated. Another interesting result contained in this book is a characterization of the Mardia bivariate Pareto distribution by Pareto conditionals and linearity of regression functions. In Arnold et al.

(1993a) multivariate measures with all bivariate Pareto conditionals were investigated while in Arnold et al. (1993b) multivariate distributions with generalized Pareto conditionals are studied.

Now we complement Theorem 1 by replacing the equidistribution condition (1) by (2). In this sense the result is parallel to AW, where similarly a new version of the multivariate normal characterization complementary to one (mentioned earlier) from AP, was obtained. However we can not adopt simply a method used for the normal case in AW (as it was done in AP) since it was based on considering conditional characteristic functions which does not seem to be convenient in the Paretian case. Instead we develop new and more universal approach which can be applied for other conditional distributions (a Student conditional is considered in Section 3) too and allows also to reprove the normality characterization from AW.

Theorem 2 *Let $X = (X_1, \cdots, X_n)$ be a random vector fulfilling* (2) *and* (3). *Then X has a Mardia multivariate Pareto distribution with the pdf given in* (4).

Let us recall that a detailed discussion of relations between conditions (1) and (2) was given in AW. It was pointed out there that they are essentially different and of the comparable strength but under (2) the Markovian approach is not possible.

Proof. Observe that if the conditional distribution of X_k given X_0, X_1, $\cdots$, X_{k-1} has the form

$$\mu_{X_k|X_0,X_1,\cdots,X_{k-1}} = \mathcal{P}(\alpha + k - 1, \sum_{i=0}^{k-1} X_i + 1) \tag{5}$$

for any $k = 1, 2, \cdots, n$, then one can reconstruct the joint density f of X by the formula

$$f(x_1, \cdots, x_n) = \prod_{k=1}^{n} f_{X_k|X_0=x_0,X_1=x_1,\cdots,X_{k-1}=x_{k-1}}(x_i)$$

$$= \prod_{k=1}^{n} \frac{(\alpha + k - 1)(\sum_{j=0}^{i-1} x_j + 1)^{\alpha+k-1}}{(\sum_{j=1}^{k} x_j + 1)^{\alpha+k}} = \frac{(\alpha + n - 1)_n}{(\sum_{i=1}^{n} x_i + 1)^{\alpha+n}}$$

which is of the form (4). By $f_{X_k|X_0,X_1,\cdots,X_{k-1}}$ we denoted the pdf of the conditional distribution of X_k given $X_0, X_1, \cdots, X_{k-1}, k = 1, \cdots, n$.

Now it suffices to show that (5) holds. We aply backward induction with respect to k. For $k = n$ it follows from (3). We will prove (5) for $k = m - 1$ provided it holds for $k = m$, $m \geq 2$.

To this end observe that since $\mu_{X_m|X_0,X_1,\cdots,X_{m-1}}$ has a density then also $\mu_{X_m|X_0,X_1,\cdots,X_{m-2}}$ is absolutely continuous and by (2) it also holds for the conditional distribution $\mu_{X_{m-1}|X_0,X_1,\cdots,X_{m-2}}$.

Consider now the following elementary formula for conditional densities:

$$f_{X_m|X_0=x_0,X_1=x_1,\cdots,X_{m-2}=x_{m-2}}(x) = \\ = \int f_{X_m|X_0=x_0,X_1=x_1,\cdots,X_{m-2}=x_{m-2},X_{m-1}=y}(x) \cdot \\ \cdot f_{X_{m-1}|X_0=x_0,X_1=x_1,\cdots,X_{m-2}=x_{m-2}}(y)\, dy\,. \tag{6}$$

Applying the induction assumption to (6) we have

$$f_{X_m|X_0=x_0,X_1=x_1,\cdots,X_{m-2}=x_{m-2}}(x) = \\ \int_0^\infty \frac{(\alpha+m-1)(\sum_{i=0}^{m-2} x_i + y + 1)^{\alpha+m-1}}{(\sum_{i=0} x_i + y + x + 1)^{\alpha+m}} \cdot \\ \cdot f_{X_{m-1}|X_0=x_0,X_1=x_1,\cdots,X_{m-2}=x_{m-2}}(y)\, dy$$

for any positive $x, x_1, \cdots, x_{m-2}$, $x_0 = 0$. Fix any $x_1, \cdots, x_{m-2}$. Then by (2) the above equation yields

$$g(x) = \int_0^\infty \frac{(\alpha+m-1)(\theta+y+1)^{\alpha+m-1}}{(\theta+y+x+1)^{\alpha+m}} g(y)\, dy\,,$$

where $\theta = x_0 + x_1 + \cdots + x_{m-2}$ and

$$g(x) = f_{X_m|X_0=x_0,X_1=x_1,\cdots,X_{m-2}=x_{m-2}}(x) = \\ = f_{X_{m-1}|X_0=x_0,X_1=x_1,\cdots,X_{m-2}=x_{m-2}}(x)\,, \quad x > 0.$$

Hence by Lemma 1

$$g(x) = \frac{(\alpha+m-2)(\theta+1)^{\alpha+m-2}}{(\theta+1+x)^{\alpha+m-1}}\,, \quad x > 0\,,$$

and consequently $\mu_{X_{m-1}|X_0,X_1,\cdots,X_{m-2}} = \mathcal{P}(\alpha+m-2, \sum_{i=0}^{m-2} X_i + 1)$. □

3. Characterization of the multivariate Student distribution

The n-variate Student $Mt_n(m, a)$ distribution with m degrees of freedom is defined by its pdf

$$f(x_1, \cdots, x_n) = \frac{a^m \Gamma(\frac{m+n}{2})}{\pi^{\frac{n}{2}} \Gamma(\frac{m}{2})(a^2 + x_1^2 + \cdots + x_n^2)^{\frac{m+n}{2}}}, \tag{7}$$

$x_i \in \mathbf{R}$, $i = 1, \cdots, n$, for any $n \geq 1$. The paramter a is a positive constant (a scale). By a simple computation it can be easily checked that for a random vector $X = (X_1, \cdots, X_n)$ distributed according to the Student $Mt_n(m, a)$ distribution the conditional distribution of X_n given $X_1, \cdots, X_{n-1}$ is a scaled univariate Student distribution with $m + n - 1$ degrees of freedom; more precisely

$$\mu_{X_n | X_1, \cdots, X_{n-1}} = t_{m+n-1}\left(\sqrt{\frac{a^2 + X_1^2 + \cdots + X_{n-1}^2}{m + n - 1}}\right), \tag{8}$$

where $t_k(\alpha)$ denotes the univariate Student distribution with k degrees of freedom. To obtain its density put $m = k$ and $n = 1$ in the formula (7). More details on multivariate Student distribution can be found for instance in a recent monograph on symmetric multivariate measures by Fang et al. (1990).

Now we can give another example where the main result of AP (its Theorem 2.1) is applicable. It is a characterization of the bivariate Student distribution. We formulate it as a lemma analogous to our Lemma 1.

Lemma 2 *If for all real x's*

$$f(x) = \int_{\mathbf{R}} \frac{\Gamma(\frac{k+2}{2})(\alpha^2 + y^2)^{\frac{k+1}{2}}}{\sqrt{\pi}\Gamma(\frac{k+1}{2})(\alpha^2 + y^2 + x^2)^{\frac{k+2}{2}}} f(y)\, dy,$$

where α is positive real, k is integer and f is a pdf then

$$f(x) = \frac{\Gamma(\frac{k+1}{2})\alpha^k}{\sqrt{\pi}\Gamma(\frac{k}{2})(\alpha^2 + x^2)^{\frac{k+1}{2}}}, \quad x \in \mathbf{R},$$

i.e. it is a pdf of the $t_k(\alpha)$.

Applying the Markovian approach from AP we easily obtain the following characterization in the multivariate case.

Theorem 3 *Let* $X = (X_1, \cdots, X_n)$ *be a random vector fulfilling* (8) *and* (1). *Then it has a multivariate Student distribution with the pdf given in* (7).

Now making use of the method developed in Section 2 we can state an analogue with (1) replaced by (2).

Theorem 4 *Let* $X = (X_1, \cdots, X_n)$ *be a random vector fulfilling* (8) *and* (2). *Then it has a multivariate Student distribution with the pdf given in* (7).

Proof. (A sketch.) The proof has the same structure as that of Theorem 2. It suffices to prove by backward induction that

$$\mu_{X_j|X_0,X_1,\cdots,X_{j-1}} = t_{m+j-1}\left(\sqrt{\frac{a^2 + X_1^2 + \cdots + X_{j-1}^2}{m+j-1}}\right)$$

for any $j = 1, 2, \cdots, n$. By (6) and the induction assumption we have for any fixed $x_1, \cdots, x_{j-2}$, $j > 1$,

$$g(x) = \int_{\mathbf{R}} \frac{\Gamma(\frac{m+j}{2})(\theta^2 + y^2)^{\frac{m+j-1}{2}}}{\sqrt{\pi}\Gamma(\frac{m+j-1}{2})(\theta^2 + y^2 + x^2)^{\frac{m+j}{2}}} g(y)\, dy\ ,$$

where $\theta = a^2 + x_0^2 + x_1^2 + \cdots + x_{j-2}^2$. Hence by Lemma 2 we have

$$g(x) = \frac{\Gamma(\frac{m+j-1}{2})\theta^{m+j-2}}{\sqrt{\pi}\Gamma(\frac{m+j-2}{2})(\theta^2 + x^2)^{\frac{m+j-1}{2}}}$$

for any real x. □

Remark. Since $Mt_n(1, a)$ is the multivariate Cauchy distribution then by putting in all the above formulas and results $m = 1$ characterizations of this distribution by a univariate Student conditional and (1) or (2) can be easily obtained.

BIBLIOGRAPHY

[1.] AHSANULLAH, M. (1985). Some characterizations of the bivariate normal distribution. *Metrika* **32**, 215-218.

[2.] AHSANULLAH, M., SINHA, B.K. (1986). On normality via conditional normality. *Calcutta Statist. Assoc. Bull.* **35**, 199-202.

[3.] AHSANULLAH, M., WESOLOWSKI, J. (1994). Multivariate normality via conditional normality. *Statist. Probab. Lett.* **20**, 235-238.

[4.] ARNOLD, B.C. (1983). *Pareto Distributions.* International Cooperative Publishing House, Fairland, Maryland.

[5.] ARNOLD, B.C., CASTILLO, E., SARABIA, J.M. (1992). *Conditionally Specified Distributions*, Lect. Notes in Statist. **73**, Springer, Berlin.

[6.] ARNOLD, B.C., CASTILLO, E., SARABIA, J.M. (1993a). A variation on the conditional specification theme. In *Bull. Inter. Statist. Inst., Contributed Papers, 49-th Session, Firenze, Book 1*, 51-52.

[7.] ARNOLD, B.C., CASTILLO, E., SARABIA, J.M. (1993b). Multivariate distributions with generalized Pareto conditionals. *Statist. Probab. Lett.* **17**, 361-368.

[8.] ARNOLD, B.C., POURAHMADI, M. (1988). Conditional characterizations of multivariate distributions. *Metrika* **35**, 99-108.

[9.] CASTILLO, E., SARABIA, J.M. (1990). Bivariate distributions with second kind beta conditionals. *Comm. Statist., Theory and Meth.* **19(9)**, 3433-3445.

[10.] FANG, K.-T., KOTZ, S., NG K.W. (1990). *Symmetric Multivariate and Related Distributions.* Chapman and Hall, London.

[11.] MARDIA, K.V. (1962). Multivariate Pareto distributions. *Ann. Math. Statist.* **33**, 1008-1015.

Received September, 1994; Revised November, 1994.