Ann. Inst. Statist. Math.
Vol. 47, No. 1, 177–183 (1995)

BIVARIATE DISTRIBUTIONS VIA A PARETO CONDITIONAL DISTRIBUTION AND A REGRESSION FUNCTION

JACEK WESOLOWSKI

Mathematical Institute, Warsaw University of Technology, Plac Politechniki 1, 00-661 Warsaw, Poland

(Received February 22, 1994; revised June 17, 1994)

Abstract. Uniqueness of specification of a bivariate distribution by a Pareto conditional and a consistent regression function is investigated. New characterizations of the Mardia bivariate Pareto distribution and the bivariate Pareto conditionals distribution are obtained.

Key words and phrases: Conditional specification, conditional distribution, regression function, bivariate probability distribution, Pareto conditional distribution, Mardia bivariate Pareto distribution, bivariate Pareto conditionals distribution.

1. Introduction

Specifications of multivariate probability distributions by conditional characteristics of Paretian type have been intensively studied in recent years. First result is due to Arnold (1987), where a characterization of a bivariate distribution with the both Pareto conditional distributions, without specifying precise forms of parameters, was obtained. Then it was extended while Castillo and Sarabia (1990) investigated bivariate distributions with the second kind beta conditionals. We recall this result following the recent monograph on conditional specifications of this kind by Arnold *et al.* ((1992), Section 5.2):

Denote by $\mathcal{P}(\sigma, \beta)$ the Pareto distribution with the density function

$$f(x) = \begin{cases} \dfrac{\beta\sigma^\beta}{(\sigma + x)^{\beta+1}} & x > 0 \\ 0 & x \leq 0 \end{cases},$$

where β and σ are positive numbers. For a random vector (X, Y) denote by $\mu_{X|Y}$ and $\mu_{Y|X}$ conditional distributions of X given Y and of Y given X, respectively. Consider a random vector (X, Y) with

$$\mu_{X|Y} = \mathcal{P}(\sigma_1(Y), \beta), \qquad \mu_{Y|X} = \mathcal{P}(\sigma_2(X), \beta), \tag{1.1}$$

where σ_1 and σ_2 are some positive functions.

Then

$$\sigma_1(Y) = \frac{a+Y}{b+cY}, \qquad \sigma_2(X) = \frac{a+bX}{1+cX}$$

and the joint distribution, named the bivariate Pareto conditionals distribution, has a density of the form

$$f(x,y) = \frac{K}{(a+bx+y+cxy)^{\beta+1}}, \qquad x>0, \quad y>0, \tag{1.2}$$

$f(x,y)=0$ otherwise, K is a normalizing constant, $a \geq 0$, $b>0$, $c \geq 0$; if $a=0$ then $\beta \in (0,1)$, if $c=0$ then $\beta \in (1,\infty)$—observe that in the case $\beta=1$ we can not have $a=0$ as it was allowed in Arnold *et al.* (1992). A special case of this measure for $\beta>1$ and $c=0$ is the Mardia bivariate Pareto distribution introduced in Mardia (1962). This distribution was also characterized in the Arnold *et al.* (1992) by (1.1) and linearity of regressions $E(X \mid Y)$ and $E(Y \mid X)$—see Section 7.2.

Other conditional specification of the similar nature are the following: In Arnold *et al.* ((1992), Section 8.5), a multivariate distribution with all univariate Pareto conditionals was considered while multivariate measures with all bivariate Pareto conditionals has been recently treated in Arnold *et al.* (1993*a*). Univariate generalized Pareto conditionals are studied in Arnold *et al.* (1993*b*). Another kind of conditional specification featuring one univariate Pareto conditional and equidistribution of some marginals was given in Arnold and Pourahmadi (1988), developed recently by Wesolowski and Ahsanullah (1994).

In this paper we consider the Pareto conditional distribution in the general scheme of specification of the distribution of a random vector (X,Y) by the conditional distribution $\mu_{Y|X}$ and the conditional mean $E(X \mid Y)$. Such a method of specification of bivariate measures goes back to Korwar (1975), where binomial and Pascal conditional distribution together with a regression function $m(y) = E(X \mid Y=y) = ay+b$, for some real constants a and b, were considered. Then Cacoullos and Papageorgiou (1983) investigated the uniqueness of determination of the bivariate distribution for binomial, Pascal and Poisson type conditionals and an arbitrary consistent function m. The same question for hypergeometric and negative hypergeometric conditionals was treated by Papageorgiou (1985). Kyriakoussis (1988) worked on uniqueness of specification of the bivariate distribution assuming that $\mu_{Y|X}$ is an X-fold convolution of a discrete measure and m is a polynomial. The discussion given in Johnson and Kotz (1992) revealed some important limitations of such a model. Logarithmic series distributions are characterized by the binomial conditional distribution and $m(0)=b$, $m(y)=ay$, $y=1,2,\dots$ or by the Pascal conditional distribution and $m(y) = \frac{\theta p y}{q} \frac{(1+\theta p/q)^{y-1}}{[(1+\theta p/q)^y - 1]}$, $y=1,2,\dots$, where $p, \theta \in (0,1)$, $q=1-p$ and $[\cdot]$ denotes the integer part of $\cdot$, in Kyriakoussis and Papageorgiou (1991). More recent contributions are the following: Arnold *et al.* (1993) uniqueness result for the shifted exponential conditional distribution and a consistent m; Ahsanullah and Wesolowski (1993) characterization of the bivariate normality by the normal conditional law and the linearity of m; Wesolowski (1993*a*) uniqueness theorems for the power series conditional distribution and a consistent m; Wesolowski (1993b) specification of the bivariate Poisson conditionals distribution by the Poisson conditional and a power function m.

2. Uniqueness and characterization

Let (X, Y) be a random vector with non-negative components. As it was pointed out in Introduction in this paper we consider possible distributions of (X, Y) with $\mu_{Y|X}$ as in the case of the bivariate Pareto conditionals distribution, i.e. we assume that the conditional density of the measure $\mu_{Y|X}$ has the form

$$f_{Y|X=x}(y) = \frac{\beta(a+bx)^{\beta}(1+cx)}{(a+bx+y+cxy)^{\beta+1}}, \qquad y > 0, \quad x \in S_X \tag{2.1}$$

or equivalently

$$\mu_{Y|X} = \mathcal{P}\left(\frac{a+bX}{1+cX}, \beta\right), \tag{2.2}$$

where $a \geq 0$, $b > 0$, $c \geq 0$ (if $a = 0$ then $\beta \in (0,1)$, if $c = 0$ then $\beta \in (1, \infty)$) and $S_X = [0, \infty)$.

Additionally we assume $ac \neq b$. Observe that in the case $ac = b$ by (2.1)

$$f_{Y|X=x}(y) = \frac{\beta a^{\beta}}{(a+y)^{\beta+1}}$$

for any $y > 0$, $x \in S_X$. Consequently Y has the Pareto $\mathcal{P}(a, \beta)$ distribution and X, Y are independent. Hence $E(X \mid Y) = E(X)$ and it is the only restriction imposed on the distribution of X.

Our aim in this section is to prove that under the given above assumptions the conditional expectation $E(X \mid Y)$ or $E[(a + bX + Y + cXY)^{-1} \mid Y]$ uniquely determine the joint distribution.

THEOREM 2.1. *Let (X, Y) be a random vector fulfilling (2.2). Then its distribution is uniquely determined by $E(X \mid Y)$.*

PROOF. Denote by m the regression function of X given Y:

$$m(y) = E(X \mid Y = y) = \int_{S_X} x dF_{X|Y=y}(x), \qquad y > 0,$$

where $F_{X|Y=y}$ is the conditional distribution function. From (2.1) it follows that Y has a density, say f_Y. Then the obvious identity

$$f_Y(y) dF_{X|Y=y}(x) = f_{Y|X=x}(y) dF_X(x),$$

where F_X denotes the distribution function of X, yields for any $y > 0$

$$m(y) \int_{S_X} f_{Y|X=x}(y) dF_X(x) = \int_{S_X} x f_{Y|X=x}(y) dF_X(x). \tag{2.3}$$

Denote a generalized distribution function H by the formula

$$dH(x) = (a+bx)^{\beta}(1+cx)dF_X(x), \qquad x \geq 0.$$

Then (2.1) and (2.3) imply

$$m(y)\phi_{\beta+1}(y) = \int_{S_X} x(a+bx+y+cxy)^{-(\beta+1)}dH(x), \tag{2.4}$$

where

$$\phi_{\gamma}(y) = \int_{S_X} (a+bx+y+cxy)^{-\gamma}dH(x)$$

for $\gamma \in \{\beta, \beta+1\}$. Observe that from the definition of ϕ_β and $\phi_{\beta+1}$ it follows that for any $y > 0$

$$(b+cy)\int_{S_X} x(a+bx+y+cxy)^{-(\beta+1)}dH(x) = \phi_\beta(y) - (a+y)\phi_{\beta+1}(y). \tag{2.5}$$

Now we join (2.4) and (2.5) to obtain

$$[(b+cy)m(y)+a+y]\phi_{\beta+1}(y) = \phi_\beta(y), \qquad y > 0. \tag{2.6}$$

Observe that ϕ_β is differentiable and its derivative can be expressed in terms of ϕ_β and $\phi_{\beta+1}$:

$$\phi'_\beta(y) = \frac{\beta}{b+cy}[(ac-b)\phi_{\beta+1}(y) - c\phi_\beta(y)], \qquad y > 0. \tag{2.7}$$

Hence by the assumption $ac \neq b$ after some easy algebra (2.6) and (2.7) yield

$$[(b+cy)m(y)+a+y]\phi'_\beta(y) = -\beta(cm(y)+1)\phi_\beta(y), \qquad y > 0.$$

Consequently

$$\phi_\beta(y) = K\exp\left(-\beta\int\frac{cm(y)+1}{(b+cy)m(y)+a+y}dy\right), \qquad y > 0,$$

where K is a positive constant.

Hence by (2.6) $\phi_{\beta+1}(y) = KG(y)$, $y > 0$, where G is a function uniquely determined by a, b, c. On the other hand observe that for any $y > 0$

$$\phi_{\beta+1}(y) = E\left(\frac{(a+bX)^{\beta}(1+cX)}{[a+bX+y(1+cX)]^{\beta+1}}\right).$$

The function $\phi_{\beta+1}$ is k-differentiable for any $k = 1, 2, \ldots$, in each point $y \geq 0$ (in $y = 0$ we consider the right-hand side derivatives). Consider now a r.v. $Z = \frac{1+cX}{a+bX}$.

Since $Z \le \max(1/a, c/b)$ a.s. then $E(Z^k)$ exists for any $k = 1, 2, \dots$ and the distribution of Z is uniquely determined by the sequence of its moments. Since

$$KG^{(k)}(0) = \phi_{\beta+1}^{(k)}(0) = (-1)^k \frac{\Gamma(\beta+k+1)}{\Gamma(\beta+1)} E(Z^{k+1}), \qquad k = 0, 1, 2, \dots$$

then a, b, c, K characterize uniquely the distribution of Z. To prove that the distribution of X is uniquely determined by a, b, c, consider $\hat{X}$ fulfilling the assumptions of the theorem. Then, similarly as above, for $\hat{Z} = \frac{1+c\hat{X}}{a+b\hat{X}}$ we have $E(\hat{Z}^k) = \hat{K}G^{(k)}(0)$, $k = 1, 2, \dots$, with the same function G. Hence $\hat{K}E(Z^k) = KE(\hat{Z}^k)$, $k = 1, 2, \dots$. Consequently for $L = K/\hat{K}$

$$\tau(s) = L\hat{\tau}(s) + 1 - L, \qquad s \ge 0,$$

where τ and $\hat{\tau}$ are the Laplace-Stjeltjes transforms of Z and $\hat{Z}$, respectively. Since Z and $\hat{Z}$ are positive a.s. then both the transforms vanish as $s \to \infty$. Thus $L = 1$ and $K = \hat{K}$. This yields unique determination of the distribution of Z and finally by the definition of Z the distribution of X is also uniquely characterized by a, b, c. □

By Theorem 2.1 a bivariate distribution can be specified by the conditional Pareto distribution and a regression function. Now we use this result to characterize the bivariate Pareto conditionals and the Mardia bivariate Pareto distributions. Observe that the mean of the Pareto $\mathcal{P}(\sigma, \beta)$ distribution exists only for $\beta > 1$ and is equal $\sigma/(\beta - 1)$. Hence for (X, Y) with the density (1.2) and $\beta > 1$

$$E(X \mid Y) = \frac{a+Y}{(\beta-1)(b+cY)}. \tag{2.8}$$

COROLLARY 2.1. *Let (X, Y) be random vector fulfilling* (2.2) *and* (2.8). *Then (X, Y) has the bivariate Pareto conditionals distribution with the density* (1.2). *If $E(X \mid Y)$ is linear, i.e. $c = 0$ then the joint distribution is the Mardia bivariate Pareto distribution.*

Now we are going to replace the conditional mean $E(X \mid Y)$ by $E[(a + bX + Y + cXY)^{-1} \mid Y)$. It appears that in this case we can also obtain a uniqueness result like Theorem 2.1 using similar methods.

THEOREM 2.2. *Let (X, Y) be a random vector fulfilling* (2.2). *Then the distribution of (X, Y) is uniquely determined by $E[(a + bX + Y + cXY)^{-1} \mid Y)$.*

PROOF. Adopting the notations from the proof of Theorem 2.1 and additionally denoting

$$\zeta(y) = E[(a + bX + Y + cXY)^{-1} \mid Y = y), \qquad y > 0$$

we have

$$\zeta(y)\phi_{\beta+1}(y) = \phi_{\beta+2}(y), \qquad y > 0.$$

Now (2.7) with β changed into $\beta+1$ yields

$$(ac-b)\zeta(y)\phi_{\beta+1}(y) = \frac{b+cy}{\beta+1}\phi'_{\beta+1}(y) + c\phi_{\beta+1}(y), \qquad y>0.$$

Hence

$$\phi_{\beta+1}(y) = \frac{K}{(b+cy)^{\beta+1}} \exp\left[(\beta+1)(ac-b)\int \frac{\zeta(y)}{b+cy}dy\right], \qquad y>0.$$

Consequently $\phi_{\beta+1}(y) = KG(y)$, $y>0$, where G is a function uniquely determined by a, b, c. Now it suffices to follow the final steps of the proof of Theorem 2.1. □

Theorem 2.2 can be also used to characterize the bivariate Pareto conditionals distribution. Observe that for a r.v. X with the Pareto $\mathcal{P}(\sigma,\beta)$ distribution $E(\frac{\sigma}{\sigma+X}) = \frac{\beta}{\beta+1}$. Consequently for (X,Y) with the density (1.2) (and conditional distributions given in (1.1))

$$E[(a+bX+Y+cXY)^{-1} \mid Y) = \frac{\beta}{(\beta+1)(a+Y)}. \tag{2.9}$$

COROLLARY 2.2. *Let (X,Y) be a random vector fulfilling* (2.2) *and* (2.9). *Then it has the bivariate Pareto conditionals distribution with the density* (1.2). *If $c=0$ then it is the Mardia bivariate Pareto distribution.*

Acknowledgements

Thanks to the referees' suggestions the exposition of the material has been considerably improved.

REFERENCES

Ahsanullah, M. and Wesolowski, J. (1993). Bivariate normality via Gaussian conditional structure (preprint).

Arnold, B. C. (1987). Bivariate distributions with Pareto conditionals, *Statist. Probab. Lett.*, **5**, 263–266.

Arnold, B. C. and Pourahmadi, M. (1988). Conditional characterizations of multivariate distributions, *Metrika*, **35**, 99–108.

Arnold, B. C., Castillo, E. and Sarabia, J. M. (1992). Conditionally specified distributions, *Lecture Notes in Statist.*, **73**, Springer, Berlin.

Arnold, B. C., Castillo, E. and Sarabia, J. M. (1993*a*). Conditionally specified models: structure and inference, *Multivariate Analysis: Future Directions 2* (eds. C. M. Cuadras and C. R. Rao), 441–450, North Holland Series in Statistics and Applied Probability, **7**, Elsevier Science B. V., Amsterdam.

Arnold, B. C., Castillo, E. and Sarabia, J. M. (1993*b*). Multivariate distributions with generalized Pareto conditionals, *Statist. Probab. Lett.*, **17**, 361–368.

Cacoullos, T. and Papageorgiou, H. (1983). Characterizations of discrete distributions by a conditional distribution and a regression function, *Ann. Inst. Statist. Math.*, **35**, 95–103.

Castillo, E. and Sarabia, J. M. (1990). Bivariate distributions with second kind beta conditionals, *Comm. Statist. Theory Methods*, **19**(9), 3433–3445.

Johnson, N. L. and Kotz, S. (1992). Non-existence of certain polynomial regression in random sum theory, *Sankhyā*, *Ser. A*, **54**, 455–459.

Korwar, R. M. (1975). On characterizing some discrete distributions by linear regression, *Comm. Statist.*, **4**, 1133–1147.

Kyriakoussis, A. (1988). Characterization of bivariate discrete distributions, *Sankhyā*, *Ser. A*, **50**, 286–287.

Kyriakoussis, A. and Papageorgiou, H. (1991). Characterizations of logarithmic series distributions, *Statist. Neerlandica*, **45**, 1–8.

Mardia, K. V. (1962). Multivariate Pareto distributions, *Ann. Math. Statist.*, **33**, 1008–1015.

Papageorgiou, H. (1985). On characterizing some discrete distributions by a conditional distributions and a regression function, *Biometrical J.*, **27**, 473–479.

Wesolowski, J. (1993*a*). Bivariate discrete measures via a power series conditional distribution and a regression function (preprint).

Wesolowski, J. (1993*b*). A new conditional specification of the bivariate Poisson conditionals distribution, *Statist. Neerlandica* (to appear).

Wesolowski, J. and Ahsanullah, M. (1994). Conditional specifications of multivariate Pareto and Student distributions (preprint).