

Characterizations of Bivariate Pareto and Yule Distributions

Evdokia Xekalaki* and Caterina Dimaki

Department of Statistics, Athens University of Economics and
Business, Athens, Greece

ABSTRACT

This article provides two characterizations of Mardia's Type I bivariate Pareto distribution. In particular, it is shown that the distribution of a random vector (X, Y) is uniquely determined as a Mardia's Type I bivariate Pareto distribution if its weighted form follows a Mardia's Type I bivariate Pareto distribution. Furthermore, this distribution is characterized by a condition on its tail probabilities. Analogous results hold for the bivariate extension of the Yule distribution which can be considered as the discrete analogue of the distribution under study.

*Correspondence: Evdokia Xekalaki, Department of Statistics, Athens University of Economics and Business, 76 Patission St., 104 34 Athens, Greece; E-mail: exek@aub.gr.

Key Words: Bivariate weighted distributions; Form invariance; Partial-sums distributions; Mardia's Type I bivariate Pareto; Bivariate generalized Waring distribution; Bivariate Waring distribution; Bivariate Yule distribution; Characterization.

Mathematics Subject Classification: Primary 62E10; Secondary 60E05.

1. INTRODUCTION

The concept of weighted distributions introduced some 40 years ago by Rao (1963) has been proved to be a very useful tool in describing a plethora of observed data by providing the means of interpreting certain practical complications that have been noticed since over half a century ago (e.g., Cox, 1962; Fisher, 1934) and has laid the way for much interesting research (e.g., Dimaki and Xekalaki, 1996; Hart, 1978; Patil, 1991; Patil and Ord, 1976; Patil and Rao, 1978). Weighted distributions arise very often in practice and represent distorted versions of the original distributions of data producing mechanisms, which may not be reproduced due to the method of data collection. This type of distortion induced by the adopted sampling procedure gives the units of the original distribution unequal probabilities of inclusion in the sample, often directly or inversely proportional to their size (size biased sampling). As a result, the value x of an original random variable X is observed with a frequency, which is different from that anticipated under the hypothesized distribution. It represents in fact the value of a random variable X^w whose frequency distribution is obtained by adjusting the frequencies of the anticipated distribution through weighting them with a weight $w(\cdot)$ analogous to the probabilities of inclusion of the values of X .

The wide applicability potential of weighted models motivated the study of the concept of weighting in bivariate contexts, thus generating equally interesting results for the interpretation of data arranged in a bivariate frequency distribution (see for example, Arnold and Nagaraja, 1991; Kocherlacota, 1995; Mahfoud and Patil, 1982; Patil et al, 1986). Identifying the original distribution from the form of its observed (weighted) version has naturally attracted much of the interest in such models, particularly in size biased sampling or income underreporting contexts and largely in the case where the distorting weighting mechanism leaves the original distribution form invariant (e.g., Dimaki and Xekalaki, 1996; Krishnaji, 1970a,b; Xekalaki, 1986). This article focuses on the identifiability problem in the bivariate case. In particular, it provides characterizations of some bivariate versions of distributions

that frequently arise in contexts of income distribution analysis and inventory decision making. These are Mardia's Type I bivariate Pareto distribution and its discrete analogue the bivariate Yule distribution. In particular, following a brief introduction of notation and terminology a characterization of Mardia's Type I bivariate Pareto distribution is provided by the property of form invariance under weight functions leading to size biased versions of it (Sec. 2). In Sec. 3, uniqueness results for the discrete analogue of the Mardia's Type I bivariate Pareto distribution, the bivariate Yule distribution, are presented. Furthermore, a more general form of distribution having the Yule as a special case is shown to be form invariant under a multiplicative weight function. Finally, in Sec. 4 a characterization of the bivariate Yule distribution, is given based on a bivariate extension of partial-sum distributions.

Before we proceed with the main results, we provide some notation and terminology.

Let us consider a random vector (X, Y) with nonnegative components and let its probability (density) function be denoted by $f_{X,Y}(x, y)$. Let $w(x, y)$ be a nonnegative weight function and assume that $E[w(X, Y)]$ exists. Denote a new probability (density) function (e.g., Kocherlacota, 1995) by,

$$f_{X,Y}^w(x, y) = \frac{w(x, y)f_{X,Y}(x, y)}{E[w(X, Y)]} \quad (1.1)$$

and let $(X, Y)^w = (X^w, Y^w)$ be the random vector whose probability (density) function is $f_{X,Y}^w(x, y)$. In the sequel $f_{X,Y}(x, y)$ is referred to as the original probability density function and $f_{X,Y}^w(x, y)$ as the observed weighted probability density function.

Cases that have attracted much interest for one dimensional variables are those where the weight function w is directly or inversely proportional to the size of X (size biased sampling) and those where the weighting allows for the functional form of the probability (density) function of X to be retained under the transition from the unobserved (original) to the observed (weighted) distribution, while permitting the possibility of a change in the vector of parameters (form invariance). Patil and Ord (1976) studied extensively the property of form invariance of univariate distributions under size biased sampling. In what follows, adopting a natural extension of their idea in the two-dimensional case we focus on the form invariance property for certain bivariate distributions. Proofs of the results are omitted as they follow directly along lines similar to the proofs of existing one dimensional cases or they naturally follow by utilizing discrete (respectively continuous) analogues of concepts on which

characterizations existing in the literature are based. Our results can be extended so as to provide characteristic properties for the multivariate extensions of the studied distributions.

2. CHARACTERIZATIONS OF MARDIA'S TYPE I BIVARIATE PARETO DISTRIBUTION

The bivariate Pareto distribution of Mardia's type I form is defined by the probability density function (pdf)

$$f_{X,Y}(x, y) = \frac{\lambda(\lambda + 1)}{\theta_1 \theta_2} \left(\frac{x}{\theta_1} + \frac{y}{\theta_2} - 1 \right)^{-(\lambda+2)},$$

for $x > \theta_1$, $y > \theta_2$, and $\lambda > 0$ (2.1)

and has Pareto marginal distributions with parameters θ_1, λ and θ_2, λ , respectively (Mardia, 1965). As can be easily verified, the standard form of (2.1) ($\theta_1 = \theta_2 = 1$), like its univariate version, is characterized by form invariance in the context of size biased sampling. The result can be formulated by the theorem that follows.

Proposition 2.1. *Let (X, Y) be a random vector with continuous components defined on $(1, +\infty)$ having finite means and joint pdf $f_{X,Y}(x, y)$. Let $f_{X,Y}^w(x, y)$ denote its observed sampled version as defined by (1.1) with $w(x, y) = x + y - 1$. Then, the original distribution (X, Y) has a bivariate Pareto distribution with pdf given by*

$$f_{X,Y}(x, y) = (\lambda + 1)(\lambda + 2)(x + y - 1)^{-(\lambda+3)}$$

if and only if its observed distribution is a bivariate Pareto with pdf $f_{X,Y}^w(x, y) = \lambda(\lambda + 1)(x + y - 1)^{-(\lambda+2)}$.

Thus, the form of the original distribution is retained if the sampling bias that induces the observed sampled distribution can be represented by $w(x, y) = x + y - 1$. This leads to size-biased forms of bivariate distributions, being a bivariate version of $w(x) = x$, widely used in size biased sampling contexts (e.g., Patil and Ord, 1976; Patil et al., 1986; Rao, 1963).

Let us now consider a random vector (X, Y) with nonnegative components and let its probability (density) function be denoted by $f_{X,Y}(x, y)$. From this distribution, a new bivariate distribution can be

defined by the formula

$$f_{X^*, Y^*}(u, v) = C \int_v^\infty \int_u^\infty \frac{f_{X, Y}(x, y)}{x + y - 1} dx dy, \quad u > 0, v > 0. \quad (2.2)$$

This is the continuous version of the bivariate STER model considered by Xekalaki (1986). In fact, the distribution of (X^*, Y^*) can be thought of as a weighted form of the distribution of (X, Y) with a weight, which is an integrated version of $w(x, y) = 1/(x + y - 1)$, a bivariate analog of the integrated version of $w(x) = 1/x$ used in the context of income underreporting by Krishnaji (1970b) for the Pareto distribution, and which can be viewed as a continuous analog of the summed version of $w(x) = 1/x$ used by Krishnaji (1970a) and Xekalaki (1983) in the context of income underreporting and inventory decision making. Dimaki and Xekalaki (1996) have also considered integrated and summed versions of $w(x) = 1/x$ for Pareto and Yule type distributions in the context of multiplicative distortion of observations. In a more general context, it is interesting to observe that the one-dimensional analogue of (2.2) arises as the distribution of the forward recurrence time in a renewal process.

The proposition that follows formulates a uniqueness result concerning bivariate Pareto distributions of Mardia's Type I form, which stems from (4.3).

Proposition 2.2. *Let the random vectors (X, Y) and (X^*, Y^*) with continuous components be defined on $(1, +\infty)$. Assume also that the pdf's of (X, Y) and (X^*, Y^*) satisfy the relationship*

$$f_{X^*, Y^*}(u, v) = C \int_v^\infty \int_u^\infty \frac{f_{X, Y}(x, y)}{x + y - 1} dx dy, \quad u > 1, v > 1.$$

Then, the pdf of (X^, Y^*) is Mardia's Type I bivariate Pareto with parameters $\theta_1 = \theta_2 = 1$ and λ if and only if the pdf of (X, Y) is Mardia's Type I bivariate Pareto with parameters $\theta_1 = \theta_2 = 1$ and $\lambda + 1$.*

3. CHARACTERIZING THE BIVARIATE YULE DISTRIBUTION

The distribution that can be regarded as the discrete analogue of Pareto distribution is known in the literature as the Yule distribution

(Kekalaki and Panaretos, 1988). In the univariate case, it is defined by

$$p_x = P(X = x) = \frac{\rho x!}{(\rho + 1)_{(x+1)}}, \quad x = 0, 1, 2, \dots, \rho > 0 \quad (3.1)$$

where $\alpha_{(\beta)} = \Gamma(\alpha + \beta)/\Gamma(\alpha)$, $\alpha > 0$, $\beta \in R$.

This and the univariate Pareto distribution exhibit a duality in their properties similar to that exhibited by the geometric and the exponential distributions (see, e.g., Kekalaki and Panaretos, 1988). As demonstrated in the sequel, a similar duality is revealed to exist in the bivariate case as, like the bivariate Pareto, the bivariate Yule is shown to be characterised by form invariance in the case of weighting leading to a size biased version of it.

Definition 3.1. (Kekalaki, 1986). A random vector (X, Y) is said to have the bivariate Yule distribution with parameter ρ (BYD(ρ)) if its probability function is given by

$$p_{x,y} = P(X = x, Y = y) = \frac{\rho_{(2)}(x+y)!}{(\rho + 1)_{(x+y+2)}}, \quad x, y = 0, 1, 2, \dots, \rho > 0.$$

This distribution can be thought of as a natural two-dimensional extension of the Yule distribution with pf given by (3.1) since both of its marginals are univariate Yule distributions with parameter ρ . It can easily be shown that the bivariate Yule distribution, like the bivariate Pareto distribution, is uniquely determined by the form of its weighted version when $w(X, Y) = X + Y + 1$. The next proposition presents the result.

Proposition 3.1. Let (X, Y) , (X^w, Y^w) be random vectors with nonnegative integer valued components. Assume that the probability functions $p_{r,m}$ and $p_{r,m}^w$ of (X, Y) and (X^w, Y^w) , respectively, satisfy the condition

$$p_{r,m}^w \equiv P(X^w = r, Y^w = m) = \frac{w(r, m)p_{r,m}}{E[w(X, Y)]}.$$

Then, the distribution of (X^w, Y^w) is the weighted bivariate Yule $(\rho + 1)$, $\rho > 0$ with weight function $w(X, Y) = X + Y + 1$ if and only if the distribution of (X, Y) is the bivariate Yule $(\rho + 1)$, $\rho > 0$.

The above result characterizes the Waring distribution, a more general distribution which contains the Yule as a special case. Both of these distributions are special cases of a bivariate distribution known as

the bivariate generalized Waring distribution (BGWD) introduced by (Xekalaki, 1984a,b, 1985, 1986) and applied in the study of the accident experience of a population in two successive time periods in the context of a proneness-liability model.

Definition 3.2. (Xekalaki, 1984a). A random vector (X, Y) is said to have the bivariate generalized Waring distribution, with parameters α, k, m and ρ (BGWD($\alpha, k, m; \rho$)) if its probability function (pf) is given by,

$$f_{X,Y}(x, y) = \frac{\rho_{(k+m)} \alpha_{(x+y)} k_{(x)} m_{(y)}}{(\alpha + \rho)_{(k+m)} (\alpha + k + m + \rho)_{(x+y)}} \frac{1}{x!} \frac{1}{y!},$$

$x, y = 0, 1, \dots; \alpha, k, m, \rho > 0$, where

$$\beta = \Gamma(\beta + \gamma) / \Gamma(\beta), \quad \beta > 0, \quad \gamma \in R.$$

By the next proposition, it is demonstrated that the BGWD is form invariant under the multiplicative weight function $w(x, y) = x^{(i)} y^{(j)}$, where $x^{(i)}$ stands for $\frac{x!}{(x-i)!}$. This is a bivariate version of $w(x) = x^{(i)}$ considered by Kocherlacota (1995) that has been used in economics contexts to model income inequality and leads to what economists term moment-distributions (e.g. Hart, 1978).

Proposition 3.2. Let the random vector (X, Y) have nonnegative, integer-valued components and assume that the distribution of (X, Y) is the BGWD ($\alpha, k, m; \rho$). Then, under the multiplicative weight function $w(x, y) = x^{(i)} y^{(j)}$, the distribution of the random vector $(X, Y)^w$ defined through (1.1) is a shifted BGWD with support $\{i, i+1, \dots\} \times \{j, j+1, \dots\}$ and parameters $\alpha + i + j, k + i, m + j, \rho - i - j$ where $\alpha, k, m > 0, \rho > i + j$.

The proof is tedious but straightforward.

Definition 3.3 (Xekalaki, 1986). A random vector (X, Y) is said to have the bivariate Waring distribution with parameters α and ρ (BWD($\alpha; \rho$)) if its probability function is given by

$$p_{x,y} = P(X = x, Y = y) = \frac{\rho_{(2)}}{(\alpha + \rho)_{(2)}} \frac{\alpha_{(x+y)}}{(\alpha + \rho + 2)_{(x+y)}},$$

where $\alpha > 0$ and $\rho > 0$; $x = 0, 1, 2, \dots$ $y = 0, 1, 2, \dots$

It is obvious that the BGWD is a generalization of the bivariate Waring distribution with parameters α and ρ ($\text{BWD}(\alpha; \rho)$) which, in turn, is a generalization of the bivariate Yule distribution, that is

$$(X, Y) \sim \text{BWD}(\alpha; \rho) \Leftrightarrow (X, Y) \sim \text{BGWD}(\alpha; 1, 1; \rho).$$

$$(X, Y) \sim \text{BYD}(1; \rho) \Leftrightarrow (X, Y) \sim \text{BGWD}(1; 1, 1; \rho).$$

Remark 3.1. As a result of the above interrelationships, and in particular since $(X, Y) \sim \text{BYD}(\rho) \Leftrightarrow (X, Y) \sim \text{BWD}(1, \rho)$, we can conclude that the bivariate Waring distribution is form invariant under the weight function $X + Y + 1$.

4. UNIQUENESS RESULTS BASED ON BIVARIATE EXTENSIONS OF DISTRIBUTIONS OF PARTIAL SUMS

Based on a given two-dimensional distribution with pf $p_{x,y} \equiv P(X = x, Y = y)$, new two-dimensional distributions can be constructed through the models:

$$p_{r,l}^w = \frac{w(r, l)p_{r,l}}{E[w(X, Y)]} \quad (4.1)$$

$$q_{r,l}^* = C \sum_{x=r+1}^{\infty} \sum_{y=l+1}^{\infty} \frac{p_{x,y}}{x+y} \quad (4.2)$$

$$q_{r,l}^+ = \frac{\sum_{x=r+1}^{\infty} \sum_{y=l+1}^{\infty} p_{x,y}}{E(XY)} \quad (4.3)$$

The first of these models was demonstrated in the previous section to characterize the bivariate Yule distribution. A continuous version of it was shown in Sec. 2 to lead to a uniqueness property of Mardia's Type I bivariate Pareto distribution. The second model given by (4.2), known as the bivariate STER model, was introduced by Xekalaki (1986) and was shown to lead to a characterization of the bivariate Yule distribution. The third model given by (4.3) is a bivariate version of $\sum_{n=r+1}^{\infty} P(X = x)/E(X)$ considered by Dimaki and Xekalaki (2004) and shown to be characteristic of the Yule distribution in the univariate case. Analogously, a characterization of the Yule distribution can be shown to

stem from (4.3) in the bivariate case. The result is formulated by the proposition that follows.

Proposition 4.1. *Let (X, Y) and (W_1, W_2) be non-negative integer valued random vectors with probability functions $p_{x,y}$ and $q_{r,m}^+$, respectively, satisfying the condition*

$$q_{r,m}^+ \equiv P(W_1 = r, W_2 = m) \\ = \frac{\sum_{x=r+1}^{\infty} \sum_{y=m+1}^{\infty} p_{x,y}}{E(XY)}, \quad r = 0, 1, \dots; \quad m = 0, 1, \dots$$

Then, the distribution of the random vector (W_1, W_2) is the bivariate Waring $(3, \rho)$, $\rho > 0$ if and only if the distribution of (X, Y) is the bivariate Yule $(\rho + 2, \rho > 0)$.

REFERENCES

- Arnold, B. C., Nagaraja, H. N. (1991). On some properties of bivariate weighted distributions. *Commun. Statist. Theory Meth.* 20(5&6):1853-1860.
- Cox, D. R. (1962). *Renewal Theory*. New York: Barnes & Noble.
- Dimaki, C., Xekalaki, E. (1996). Additive and multiplicative distortion of observations: some characteristic properties. *J. Appl. Statist. Sci.* 5(2):113-127.
- Dimaki, C., Xekalaki, E. (2004). Identifying the Yule distribution by properties of certain reliability measures. *J. Statist. Planning Infer.* To appear.
- Fisher, R. A. (1934). The effects of methods of ascertainment upon the estimation of frequencies. *Ann. Eugenics* 6:13-25.
- Hart, P. E. (1978). Moment distributions in economics: an exposition. *J. Roy. Statist. Soc. A* 138:423-434.
- Kocherlacota, S. (1995). Discrete bivariate weighted distributions under multiplicative weight function. *Commun. Statist. Theory Meth.* 24(2):533-551.
- Krishnaji, N. A. (1970a). Characteristic property of the Yule distribution. *Sankyā A* 32:343-346.
- Krishnaji, N. (1970b). Characterization of the pareto distribution through a model of under-reported incomes. *Econometrica* 38: 251-255.

- Mahfoud, M., Patil, G. P. (1982). On weighted distributions. In: Kallianpur, G., Krishnaji, P. R., Ghosh, J. K., eds. *Statistics and Probability: Essays in Honor of C.R. Rao*, Amsterdam: North Holland. pp. 479–492.
- Mardia, K. V. (1965). *Families of Bivariate Distributions*. London: Griffin.
- Patil, G. P. (1991). Encountered data, statistical ecology, environmental statistics, and weighted distribution methods. *Environmetrics* 24:377–423.
- Patil, G. P., Ord, J. K. (1976). On size-biased sampling and related form-invariant weighted distributions. *Sankyā, Series B* 38:48–61.
- Patil, G. P., Rao, C. R. (1978). Weighted distributions and size biased sampling with applications to wildlife populations and human families. *Biometrics* 34:179–184.
- Patil, G. P., Rao, C. R., Ratnaparkhi, M. V. (1986). On discrete weighted distributions and their use in model choice for observed data. *Commun. Statist. Theory Meth.* 15(3):907–918.
- Rao, C. R. (1963). On discrete distributions arising out of methods of ascertainment. *Sankyā A* 25:311–324.
- Xekalaki, E. (1983). A property of the Yule distribution and its applications. *Commun. Statist., Part A (Theory Methods)* A12(10): 1181–1189.
- Xekalaki, E. (1984a). The bivariate, generalized Waring distribution and its application to accident theory. *J. Royal Statist. Soc. Series A (General)* 147(3):488–498.
- Xekalaki, E. (1984b). Models leading to the bivariate generalized Waring distribution. *Utilitas Mathematica* 25:263–290.
- Xekalaki, E. (1985). Some bivariate extensions of the generalized Waring distribution. *Studia Scientiarum Mathematicarum Hungarica* 20:173–181.
- Xekalaki, E. (1986). The bivariate Yule distribution and some of its properties. *Statistics* 17(2):311–317.
- Xekalaki, E., Panaretos, J. (1988). On the association of the Pareto and the Yule distributions. *Teoriya Veroyatnostei i ee Primeneniya* 33(1):206–210.

Copyright of Communications in Statistics: Theory & Methods is the property of Marcel Dekker Inc. and its content may not be copied or emailed to multiple sites or posted to a listserv without the copyright holder's express written permission. However, users may print, download, or email articles for individual use.