MULTIVARIATE PARETO DISTRIBUTIONS

BY K. V. MARDIA

University of Rajasthan, Jaipur

1. Introduction and summary. It is well known that the family of Pareto distributions with densities

$$(1.1) \quad \begin{aligned} f(x; a, p) &= pa^p/x^{p+1}, & x > a > 0, \\ &= 0, & x \leqq a, p > 0, \end{aligned}$$

provides reasonably good fits to many empirical distributions, e.g., to distributions of income and of property values. In most of these cases, ancillary information is present, which could be utilized if an appropriate multivariate Pareto distribution were available.

The objects of this note are (i) to suggest two families of bivariate Pareto distributions with the property that both marginal distributions are of univariate Pareto form; (ii) to extend these to multivariate forms; and (iii) to discuss estimation of the parameters in the bivariate distributions.

2. The bivariate Pareto distribution of type 1. A simple bivariate density function satisfying the marginal property is

$$(2.1) \quad \begin{aligned} f_1(x, y; a, b, p) &= [p(p+1)(ab)^{p+1}]/(bx + ay - ab)^{p+2}, \\ & \qquad x > a > 0, y > b > 0, \\ &= 0, \qquad x \leqq a, y \leqq b, p > 0. \end{aligned}$$

We shall call it a bivariate Pareto distribution of type 1.

The density for this distribution is constant on every line $a^{-1}x + b^{-1}y = c$. The marginal density functions of x and y are $f(x; a, p)$ and $f(y; b, p)$ respectively. Also, the conditional distribution of x given y is

$$(2.2) \quad \begin{aligned} f_1(x \mid y) &= [b(p+1)(ay)^{p+1}]/(bx + ay - ab)^{p+2}, \quad x > a > 0, y > b > 0, \\ &= 0 \text{ otherwise}, \end{aligned}$$

which is again of Pareto form but with displaced origin. Further, we have

$$(2.3) \quad \begin{aligned} E(x) &= ap/(p-1), \quad p > 1; \\ V(x) &= a^2p/\{(p-1)^2(p-2)\}, \quad p > 2, \end{aligned}$$

$$(2.4) \quad \text{Cor}(x, y) = 1/p, \quad p > 2,$$

$$(2.5) \quad E(x \mid y) = a + \{(ay)/(bp)\},$$

$$(2.6) \quad V(x \mid y) = a^2y^2(p+1)/\{b^2(p-1)p^2\}.$$

Similarly $E(y)$, $V(y)$, $f_1(y \mid x)$ and $V(y \mid x)$ can be immediately obtained.

Received December 14, 1959; revised March 10, 1962.

3. A bivariate distribution of type 2. The density function of $u = p\log(x/a)$, where x has the density function $f(x; a, p)$ is

$$(3.1) \qquad \begin{array}{ll} e^{-u}, & u > 0, \\ 0, & u \leqq 0. \end{array}$$

Therefore, the density function of u is a gamma function with index parameter equal to 1, i.e., $\phi(u, 1)$ in the notation of Kibble [3]. Similarly, $v = q\log(y/b)$ has the density function $\phi(v, 1)$, where y is a random variate with the density function $f(y; b, q)$. Hence, a bivariate distribution of u and v may be taken as the joint distribution of two dependent gamma type variables in the sense of Kibble [3], with the index parameter equal to unity. On writing $p = 1$ and $\rho^2 = \alpha$ in Kibble [3], pp. 141, the joint density function of u and v is given by

$$(3.2) \qquad \begin{array}{ll} \dfrac{1}{1-\alpha}\exp\left\{-\dfrac{u+v}{1-\alpha}\right\} I_0\left\{\dfrac{2(\alpha uv)^{\frac{1}{2}}}{1-\alpha}\right\}, & u > 0, v > 0, \\ 0, & u \leqq 0, v \leqq 0, 0 \leqq \alpha < 1, \end{array}$$

where

$$I_k(z) = \sum_{r=0}^{\infty} [\Gamma(r+1)\Gamma(k+r+1)]^{-1}(\tfrac{1}{2}z)^{k+2r}$$

is a modified Bessel function of the first kind and of the kth order, and

$$\alpha = \rho^2 = \mathrm{Cor}(u, v).$$

Therefore, the joint density function of x and y is

$$\begin{aligned} f_2(x, y; a, b, p, q, \alpha) &= [pq/(1-\alpha)xy][(a/x)^p(b/y)^q]^{1/(1-\alpha)} \\ &\quad \cdot \sum_{r=0}^{\infty}\left[\frac{\alpha pq\,\{\log(x/a)\}\{\log(y/b)\}}{(1-\alpha)^2}\right]^r \frac{1}{(r!)^2}, \\ (3.3) \qquad &= \frac{pq}{(1-\alpha)xy}\left[\left(\frac{a}{x}\right)^p\left(\frac{b}{y}\right)^q\right]^{1/(1-\alpha)} \\ &\quad \cdot I_0\left[\frac{2[\alpha pq\,\{\log(x/a)\}\{\log(y/b)\}]^{\frac{1}{2}}}{1-\alpha}\right]; \\ &\quad x > a > 0, y > b > 0, p > 0, q > 0, 0 \leqq \alpha < 1. \end{aligned}$$

For $x \leqq a$, $y \leqq b$, the density function is zero. The marginal density functions of x and y are $f(x; a, p)$ and $f(y; b, q)$ respectively and (3.3) is termed the Pareto distribution of type 2.

$E(x)$ and $V(x)$ are the same as in (2.3), and $E(y)$ and $V(y)$ are immediate by interchanging (a, p) and (b, q). Further,

$$(3.4) \qquad \mathrm{Cor}(x, y) = \alpha\{pq(p-2)(q-2)\}^{\frac{1}{2}}\{(p-1)(q-1)-\alpha\}^{-1}, \qquad p > 2, q > 2,$$

$$f_2(x \mid y) = [p/(1-\alpha)x][(a/x)^p(b/y)^q]^{1/(1-\alpha)} \cdot I_0[2\{\alpha pq[\log(x/a)][\log(y/b)]\}^{\frac{1}{2}}/(1-\alpha)], \tag{3.5}$$

$$E(x \mid y) = [pa/(p-1+\alpha)][y/b]^{q\alpha/(q-1+\alpha)}, \tag{3.6}$$

$$V(x \mid y) = a^2 p\left[\frac{(b^{-1}y)^{2\alpha q/(2\alpha-2+p)}}{2\alpha-2+p} - \frac{(b^{-1}y)^{2\alpha q/(\alpha-1+p)}}{\alpha-1+p}\right]. \tag{3.7}$$

Similarly, $f_2(y \mid x)$, $E(y \mid x)$ can be written by symmetry. Further, the moment generating function of u and v derived by Kibble [3], gives the characteristic function $\phi(t_1, t_2)$ of $\log(x/a)$ and $\log(y/b)$ as

$$\phi(t_1, t_2) = [1 - i(p^{-1}t_1 + q^{-1}t_2) - p^{-1}q^{-1}t_1t_2(1-\alpha)]^{-1}. \tag{3.8}$$

4. Multivariate Pareto Type 1. Writing (2.1) in the form

$$f_1(x, y; a, b, p) = p(p+1)/\{ab(a^{-1}x + b^{-1}y - 1)^{p+2}\},$$
$$x > a > 0, y > b > 0, p > 0,$$

a generalization to the k-variate case becomes obvious and the joint density function can be written as

$$f_1(x_1, x_2, \cdots, x_k) = \frac{p(p+1)\cdots(p+k-1)}{\left(\prod_{i=1}^{k} a_i\right)\left\{\left(\sum_{i=1}^{k} a_i^{-1}x_i\right) - k + 1\right\}^{p+k}}, \tag{4.1}$$
$$x_i > a_i > 0, i = 1, \cdots, k; p > 0.$$

The density of the mass for this distribution is constant on every hyperplane $\sum_{i=1}^{k} a_i^{-1}x_i = c$. The marginal density function of x_i is $f(x_i; a_i, p)$ and the conditional distributions are again of the Pareto form but with displaced origin. The regressions of $x_1, \cdots, x_r$, $(r < k)$, on $x_{r+1}, \cdots, x_k$, are linear. All the correlations of zero order are equal to $1/p$. Every partial correlation coefficient of the sth order is $1/(p+s)$. The multiple correlation of a variate with the other $(k-1)$ variate is $[(k-1)/\{p(p+k-2)\}]^{\frac{1}{2}}$.

5. Multivariate Pareto type 2. If the dependent random variables x_i have the marginal density functions $f(x_i; a_i, p_i)$, then the joint distribution of random variables $u_i = p_i \log(x_i/a)$, $i = 1, \cdots, k$, might be taken as a multivariate gamma distribution with parameter $p = 1$ in the sense of Krishnamoorthy and Parthasarthy [4]. Hence in the notation of this reference, the required joint density function of $x_1, \cdots, x_k$ is

$$\prod_{i=1}^{k} f(x_i; a_i, p_i) \sum_{r=0}^{\infty} [\sum_{i<j} C_{ij}L(p_i \log(x_i/a_i), 1)L(p_j \log(x_j/a_j), 1) + \cdots + C_{12\cdots k}L(p_1 \log(x_1/a_1), 1) \cdots L(p_k \log(x_k/a_k), 1)]^r, \tag{5.1}$$

where the marginal density function of x_i is $f(x_i; a_i, p_i)$. The regressions of $\log x_1$ on $\log x_2, \cdots, \log x_k$, $k \neq 2$, and of x_1 on $x_2, \cdots, x_k$, are non-linear.

6. Comparison of the two bivariate distributions. The two distribution functions are distinct. In type 2, unlike type 1 but like the normal distribution, both parameters in the marginal distribution of x may be different from the parameters in the marginal distribution of y, and $\mathrm{Cor}(x, y) = 0$ implies that x and y are independent Paretos. The correlation between x and y is positive in type 1 while non-negative in type 2. The regressions of x on y in type 1 and type 2 are linear and exponential respectively, but $V(x \mid y)$ depends on y. In type 2, the regression of $\log x$ on $\log y$ is linear.

7. Estimation of the parameters in bivariate Pareto type 2. Let (x_i, y_i), $i = 1, \cdots, n$, be a random sample of size n from bivariate Pareto type 2 population.

THEOREM 7.1. *Consistent estimates of a and b are X and Y respectively, where X is minimum $(x_1, \cdots, x_n)$ and Y is minimum $(y_1, \cdots, y_n)$.*

PROOF. The sample extremes are consistent estimates of the population extremes when the latter are finite.

THEOREM 7.2. *Consistent estimates of p and q are $\{\log(G_x/X)\}^{-1}$ and $\{\log(G_y/Y)\}^{-1}$ respectively, where G_x and G_y are the geometric means of $(x_1, \cdots, x_n)$ and $(y_1, \cdots, y_n)$ respectively.*

PROOF. It is shown by Muniruzzaman [8], Section 9, that

$$E[\{\log(G_x/X)\}^{-1}] = np/(n-2),$$

and

$$V[\{\log(G_x/X)\}^{-1}] = p^2/n.$$

THEOREM 7.3. *A consistent estimate of α is*

$$\bar{r} = \frac{\sum_{i=1}^{n} [\log(x_i/G_x)] [\log(y_i/G_y)]}{\left\{\sum_{i=1}^{n} [\log(x_i/G_x)]^2 \sum_{i=1}^{n} [\log(y_i/G_y)]^2\right\}^{\frac{1}{2}}}.$$

PROOF. From (3.8), the central moments of $\log(x/a)$ and $\log(y/b)$ may be derived as

$$\mu_{20} = 1/p^2, \qquad \mu_{40} = 9/p^4,$$
$$\mu_{11} = \alpha/(pq), \qquad \mu_{22} = \{(2\alpha+1)/(pq)\}^2,$$
$$\mu_{21} = 2\alpha/(p^2q), \qquad \mu_{31} = 9\alpha/(p^3q).$$

Interchanging p and q, μ_{02}, μ_{04}, μ_{12}, and μ_{13} can be obtained immediately.

For large values of n, by Cramér [2], Section 27.8, and the above moments, we have

$$\begin{aligned} E(\bar{r}) &= E[\text{Sample Cor}\{\log(x/a), \log(y/b)\}] \\ &\doteq \text{Population Cor}\{\log(x/a), \log(y/b)\} \\ &= \alpha, \end{aligned}$$

and

$$V(\bar{r}) \doteq \frac{\alpha^2}{4n}\left[\frac{\mu_{40}}{\mu_{20}^2}+\frac{\mu_{04}}{\mu_{02}^2}+\frac{2\mu_{22}}{\mu_{20}\mu_{02}}-\frac{4\mu_{31}}{\mu_{11}\mu_{20}}-\frac{4\mu_{13}}{\mu_{11}\mu_{02}}\right]$$
$$= (1-\alpha^2)(2\alpha^2+6\alpha+1)/n.$$

Therefore, $\bar{r}$ is a consistent estimate of α.

The above theorems provide consistent estimates of a, b, p, q and α.

It is interesting to note that, for given a and b, the maximum likelihood estimates of p and q on the basis of the sample (x_i, y_i), $i = 1, \cdots, n$, from the bivariate population are the same as that on the basis of the samples (x_i) and (y_i), $i = 1, \cdots, n$, from the univariate Pareto populations $f(x; a, p)$ and $f(y; b, q)$ respectively, while the maximum likelihood estimate of α is given by

$$\alpha^{\frac{1}{2}} = n^{-1}\sum_{i=1}^{n} c_i\{I_1[2\alpha^{\frac{1}{2}}c_i/(1-\alpha)]/I_0[2\alpha^{\frac{1}{2}}c_i/(1-\alpha)]\}$$

where

$$c_i = [\{\log(x_i/a)\}\{\log(y_i/b)\}/\{\log(G_x/a)\}\{\log(G_y/b)\}]^{\frac{1}{2}}.$$

8. Estimates of parameters in bivariate Pareto type 1. The maximum likelihood estimates from a random sample (x_i, y_i), $i = 1, \cdots, n$, can be shown easily to be the following:

$$\hat{a} = \text{minimum } (x_1, \cdots, x_n) = X,$$
$$\hat{b} = \text{minimum } (y_1, \cdots, y_n) = Y,$$

and

$$\hat{p} = (C^{-1} - \tfrac{1}{2}) + (C^{-2} + \tfrac{1}{4})^{\frac{1}{2}},$$

where

$$C = n^{-1}\sum_{i=1}^{n}\log(X^{-1}x_i + Y^{-1}y_i - 1).$$

9. Sampling problem for the bivariate Pareto type 2. The general problem in the succeeding sections is to obtain the appropriate ratio and regression estimates of the measure of location of y under a few important sampling techniques, when x is an ancillary variable and a, b are given. The measure of central tendency is taken to be the geometric mean. Its appropriateness and importance in the univariate Pareto population is exhibited by Muniruzzaman [8]. The log-scale is convenient in the subsequent discussion.

10. Ratio estimate. Let (x_i, y_i), $i = 1, \cdots, n$, be a random sample from Pareto type 2. Suppose ν and δ are the population geometric means of x and y respectively, i.e.,

(10.1) $$\nu = ae^{1/p}, \qquad \delta = be^{1/q}.$$

We define a ratio estimate of δ (parallel to the standard ratio estimate) for given ν, as

$$R = (G_y/G_x)\nu. \tag{10.2}$$

LEMMA 10.1. *Log G_y is an unbiased and consistent estimate of log δ.*

PROOF. We have

$$E(\log G_y) = \log \delta, \qquad V(\log G_y) = 1/(nq^2). \tag{10.3}$$

THEOREM 10.1. *Log R is an unbiased, consistent estimate of log δ and has smaller variance than log G_y if $\alpha > \frac{1}{2}qp^{-1}$.*

PROOF. Define

$$z = \log\{aR/(b\nu)\} = \sum_{i=1}^{n} \log(y_i/b)(a/x_i). \tag{10.4}$$

Its characteristic function is

$$\{1 - it(q^{-1} - p^{-1})n^{-1} + t^2(1 - \alpha)(n^2pq)^{-1}\}^{-n}. \tag{10.5}$$

We obtain $E(z)$ and $V(z)$, and use (10.1) and (10.4), to get $E(\log R) = \log \delta$ and $V(\log R) = (1/n)\{p^{-2} + q^{-2} - 2\alpha(pq)^{-1}\} < 1/(nq^2) = V(\log G_y)$ if $\alpha > \frac{1}{2}qp^{-1}$.

Incidentally we can obtain the probability density function of R. On application of the inverse theorem to the characteristic function (10.5), we find that the density function of z is

$$\begin{aligned} f(z) &= \frac{1}{\pi^{\frac{1}{2}}} \frac{c^n d^n}{(n-1)!} e^{-\frac{1}{2}z(c-d)}[z/(c+d)]^{n-\frac{1}{2}} K_{n-\frac{1}{2}}\left[\frac{1}{2}z(c+d)\right], \quad z \geqq 0, \\ &= \frac{1}{\pi^{\frac{1}{2}}} \frac{c^n d^n}{(n-1)!} e^{-\frac{1}{2}z(c-d)}[-z/(c+d)]^{n-\frac{1}{2}} K_{n-\frac{1}{2}}\left[-\frac{1}{2}z(c+d)\right], \quad z \leqq 0, \end{aligned} \tag{10.6}$$

where

$$c = \{(q - p) + [(p + q)^2 - 4pq\alpha]^{\frac{1}{2}}\}/\{2(1 - \alpha)\},$$

$$d = \{(p - q) + [(p + q)^2 - 4pq\alpha]^{\frac{1}{2}}\}/\{2(1 - \alpha)\},$$

and

$$K_{n+\frac{1}{2}}(w) = (\pi/2w)^{\frac{1}{2}}e^{-w} \sum_{r=0}^{n} (n + r)!/\{r!(n - r)!(2w)^r\},$$

is a modified Bessel function of the second kind of order $n + \frac{1}{2}$. (See, for integration, Section 2 in Mardia [12].) Hence the probability density function of R is known.

11. Regression estimate (simple sampling). A random sample (x_i, y_i), $i = 1, \cdots, n$, is drawn from the Pareto type 2. We define

$$\hat{y} = \log G_y + \hat{\beta}(\log \nu - \log G_x),$$

where

$$\hat{\beta} = \frac{\sum_{i=1}^{n} [\log (y_i/b)] [\log (x_i/G_x)]}{\sum_{i=1}^{n} [\log (x_i/G_x)]^2}.$$

THEOREM 11.1. *$\hat{y}$ is an unbiased estimate of log δ.*

PROOF. By (3.2), we have

$$E[\log(y/b) \mid x] = (1 - \alpha)q^{-1} + p\alpha q^{-1}\{\log(x/a)\},$$

so that

$$E(\hat{y} \mid x) = (1 - \alpha) + \log b + p\alpha q^{-1} \log G_x .$$

On taking expectations with respect to x, the theorem follows. Therefore $\hat{y}$ can be termed the simple regression estimate of log δ.

LEMMA 11.1. *The joint distribution of log G_x and log G_y is asymptotically bivariate normal with parameters (log ν, log δ, $1/(np)$, $1/(nq)$, α) where (m_1, m_2, σ_1, σ_2, ρ) denote the standard parameters of the bivariate normal population.*

PROOF. The central limit theorem of equal components applies to the joint distribution of log G_x and log G_y. Hence if we know a solution to a sampling problem for a random sample (x_i, y_i), $i = 1, \cdots, n$, from a bivariate normal population with parameters (m_1, m_2, σ_1, σ_2, ρ) then the same problem, and its solution for large n, can be written for a random sample from Pareto type 2, on replacing x_i, y_i, $\bar{x}$, $\bar{y}$, m_1, m_2, σ_1, σ_2, and ρ by the respective quantities log x_i, log y_i, log G_x, log G_y, log ν, log δ, $1/p$, $1/q$ and α.

THEOREM 11.2. *$\hat{y}$ is a more efficient estimate than log G_y, at least for large values of n.*

PROOF. Application of Lemma 11.1 to the variance of the corresponding regression estimate for $(\bar{x}, \bar{y})$ of bivariate normal population (Sukhatme [10], pp. 203) gives for large values of n

$$V(\hat{y}) \doteq (1 - \alpha^2)/(nq^2) < 1/(nq^2) = V(\log G_y).$$

12. Regression estimate (double sampling). Let $(x_1', \cdots, x_m')$ be a random sample on variate x from Pareto type 2, $(x_1, \cdots, x_n)$ be a sub-sample of it $(m < n)$, and $(y_1, \cdots, y_n)$ are corresponding observed values of variate y. Suppose G_x' is the geometric mean of $(x_1', \cdots, x_m')$. We define

$$\hat{y}_1 = \log G_y + \hat{\beta}(\log G_x' - \log G_x).$$

THEOREM 12.1. *$\hat{y}_1$ is an unbiased estimate and, at least for large n, more efficient estimate of log δ than log G_y.*

PROOF. Unbiasedness may be established as in Theorem 11.1. Further, application of Lemma 11.1 in the solution to a similar problem for $(\bar{x}, \bar{y})$ of the bivariate normal population (Cochran [9], pp. 278) gives for large n,

$$V(\hat{y}_1) \doteq \{\alpha^2/(mq^2)\} + \{(1 - \alpha^2)/(nq^2)\}; \qquad n < m,$$

$$< 1/(nq^2) = V(\log G_y).$$

Therefore $\hat{\theta}_1$ can be termed the regression estimate of log δ in double sampling.

13. Remarks. Consistent estimates of the parameters in the multivariate distributions can be obtained on the line of Sections 7 and 8. The generalization of Lemma 11.1 to the multivariate case is obvious. Its application provides, for large n, solutions to problems in multi-stage sampling from Pareto type 2.

14. Acknowledgments. The author is very grateful to B. V. Shah, William Kruskal and the referees for their valuable suggestions and comments, and to C. Y. Raosahib, K. S. Rao, M. C. Chakrabarti, D. D. Kosambi and G. C. Patni for their interest and encouragement.

REFERENCES

[1] HAGSTROEM, K. G. (1925). La loi de Pareto et la réassurance. *Skand. Aktuarietidskr.* **25** 65–105.

[2] CRAMÉR, HARALD (1946). *Mathematical Methods of Statistics.* Princeton Univ. Press.

[3] KIBBLE, W. F. (1941). A two-variate gamma type distribution. *Sankhyā* **5** 137–150.

[4] KRISHNAMOORTHY, A. S. and PARTHASARTHY, M. (1951). A multivariate gamma type distribution. *Ann. Math. Statist.* **22** 549–557.

[5] KRISHNAIAH, P. R. and RAO, M. M. (1960). A correction to "a multivariate gamma type distribution." *Ann. Math. Statist.* **31** 229.

[6] KRISHNAIAH, P. R. and RAO, M. M. (1961). Remarks on a multivariate gamma distribution. *Amer. Math. Monthly* **68** 342–346.

[7] RAO, C. RADHAKRISHNA (1952). *Advanced Statistical Methods in Biometric Research.* Wiley, New York.

[8] MUNIRUZZAMAN, A. N. M. (1957). On measures of location and dispersion and tests of hypotheses in a Pareto population. *Calcutta Statist. Assn. Bull.* **7** 115–123.

[9] COCHRAN, W. G. (1953). *Sampling Techniques.* Wiley, New York.

[10] SUKHATME, P. V. (1953). *Sampling Theory of Surveys with Applications.* Iowa State College Press, Ames.

[11] WATSON, G. N. (1948). *A Treatise on the Theory of Bessel Functions.* Macmillan, New York.

[12] MARDIA, K. V. (1961). An important integral and application in partial fractions. *Math. Student* **29** 15–20.