BULLETIN OF THE
INSTITUTE OF MATHEMATICS
ACADEMIA SINICA
Volume 28, Number 2, June 2000

# TWO MULTIVARIATE PARETO DISTRIBUTIONS AND THEIR RELATED INFERENCES

BY

HSIAW-CHAN YEH (葉小蓁)

**Abstract.** Two different multivariate Pareto distributions and their related inferences are studied in this paper. Both of Marida (1962) and Arnold (1983) studied mainly for the bivariate case. Some of their results can be analogously extended to the multivariate case.

**1. Introduction.** Two different multivariate Pareto distributions and their related inferences are considered in this paper. The first one is the homogeneous multivariate Pareto $(IV)$, $MP^{(m)}(IV)(\underline{0},\underline{1},\gamma,\alpha)$, some of its properties has been studied in Yeh (1994). The inferences for the parameters $\gamma$ and $\alpha$ will be continued to study in this paper. According to Arnold (1983)'s Chapter 6, this distribution is generated by m's i.i.d. exponential random variables, $\{U_i\}^m$, and $Z$-Gamma $(\alpha, 1)$ with $\{U_i\}$ and $Z$ being independent. Define $X_i = (U_i/Z)^\gamma$, $i = 1, 2, \ldots, m$, then the random vector

$$\underline{X} = (X_1, X_2, \ldots, X_m) \sim MP^{(m)}(IV)(\underline{0},\underline{1},\gamma,\alpha)$$

is the so-called homogeneous multivariate Pareto $(IV)$ distribution with joint survival function

$$\bar{F}_{\underline{x}}(\underline{x}) = \left\{1 + \sum_{i=1}^{m} x_i^{1/\gamma}\right\}^{-\alpha} \tag{1.1}$$

Received by the editors March 25, 1998 and in revised form March 25, 1999.

Key words and phrases: Homogeneous Multivariate Pareto (IV) Distribution; Marida's Multivariate Pareto Distribution; Wicksell-Kibble Multivariate Exponential; Feller-Pareto Distribution; Joint Characteristic Function; Inferences; Pivotal Confidence Interval; Asymptotic Distribution; Consistent.

and the joint pdf

$$f_{\underline{x}}(\underline{x}) = \frac{1}{\gamma^m}\frac{\Gamma(\alpha+m)}{\Gamma(\alpha)}\left\{1+\sum_{i=1}^{m} x_i^{1/\gamma}\right\}^{-\alpha-m}\left(\prod_{i=1}^{m} x_i\right)^{1/\gamma-1} \tag{1.2}$$

The second one is the Marida's multivariate Pareto distribution. This $m$-variate ($m \geq 3$) Pareto distribution is the extension of Marida (1962)'s second type of bivariate Pareto distribution. It is constructed through the equation

$$X_i = \sigma_i e^{U_i/\alpha_i},\ i = 1, 2, \ldots, m$$

where the $m$-variate random vector

$$\underline{U} = (U_1, U_2, \ldots, U_m)$$

is the Wicksell-Kibble (1993, 1941) exponential variable, then the random vector

$$\underline{X} = (X_1, X_2, \ldots, X_m)$$

is the Marida's multivariate Pareto distribution and is denoted by $WKP^{(m)}$ $(I)(\underline{\sigma}, \underline{\alpha}, p)$. This notation emphasizes its genesis from the Wicksell-Kibble distribution and the fact that it has the classical Pareto, $P(I)$, marginal distribution. The parameters in this distribution are respectively

$$\underline{\sigma} = (\sigma_1, \sigma_2, \ldots, \sigma_m), \sigma_i > 0 \text{ for scale,}$$

$$\underline{\alpha} = (\alpha_1, \alpha_2, \ldots, \alpha_m), \alpha_i > 0 \text{ for shape,}$$

and $p, 0 < p < 1$. The estimators of $\underline{\sigma}, \underline{\alpha}$ and $p$ and their sampling distributions will be discussed in section 3.

**2. Inferences for the homogeneous $MP^{(m)}(IV)(\underline{0}, \underline{1}, \gamma, \alpha)$ distributions.** Yeh (1994) studied only the probability properties of the homogeneous $MP^{(m)}$ $(IV)(\underline{0}, \underline{1}, \gamma, \alpha)$ with tow parameters $\alpha$ and $\gamma$. The inferences for $\alpha$ and $\gamma$ will be studied in this section.

**2.1 The numerical solutions for the estimates of $\alpha$ and $\gamma$.** There are two commonly used methods to estimate $\gamma$ and $\alpha$ numerically, one is the maximum likeihood estimate ($MLE$), the other one is the method of moment estimate ($MM$).

Let $\{\underline{X}^{(i)} = (X_1^{(i)}, X_2^{(i)}, X_m^{(i)})\}_{i=1}^{n}$ be a random sample from a $MP^{(m)}\ (IV)(\underline{0}, \underline{1}, \gamma, \alpha)$ population. From Yeh (1994) equation (2.2), knowing that the likelihood function of $(\gamma, \alpha)$ is

(2.1)

$$L = L(\gamma,\alpha) = \frac{1}{\gamma^{mn}}\left\{\frac{\Gamma(\alpha+m)}{\Gamma(\alpha)}\right\}^{n}\prod_{i=1}^{n}\left\{1+\sum_{j=1}^{m}(X_j^{(i)})^{1/\gamma}\right\}^{-(\alpha+m)}\prod_{i=1}^{n}\left(\prod_{j=1}^{m}X_j^{(i)}\right)^{1/\gamma-1}$$

Taking natural logarithm,

$$\begin{aligned}(2.2)\quad &\ln(L)\\ =&(-mn)\ln(\gamma)+n\{\ln\Gamma(\alpha+m)-\ln\Gamma(\alpha)\}\\ &-(\alpha+m)\sum_{i=1}^{n}\ln\left\{1+\sum_{j=1}^{m}(X_j^{(i)})^{1/\gamma}\right\}+(1/\gamma-1)\sum_{i=1}^{n}\sum_{j=1}^{m}\ln(X_j^{(i)}).\end{aligned}$$

The MLE $(\hat{\gamma}, \hat{\alpha})$, of $(\gamma, \alpha)$ exist, it must satisfy $\frac{\partial \ln(L)}{\partial \gamma} = \frac{\partial \ln(L)}{\partial \alpha} = 0$, where

$$\begin{aligned}(2.3)\quad \frac{\partial \ln(L)}{\partial\gamma} =& -\frac{mn}{\gamma}+\frac{(\alpha+m)}{\gamma^2}\sum_{i=1}^{n}\left\{\frac{\sum_{j=1}^{m}(X_j^{(i)})^{1/\gamma}\ln(X_j^{(i)})}{1+\sum_{j=1}^{m}(X_j^{(i)})^{1/\gamma}}\right\}\\ &-\frac{1}{\gamma^2}\sum_{i=1}^{n}\sum_{j=1}^{m}\ln(X_j^{(i)})=0\end{aligned}$$

and

$$(2.4)\quad \frac{\partial \ln(L)}{\partial\alpha} = n\left\{\frac{\Gamma'(\alpha+m)}{\Gamma(\alpha+m)}-\frac{\Gamma'(\alpha)}{\Gamma(\alpha)}\right\}-\sum_{i=1}^{n}\ln\left\{1+\sum_{j=1}^{m}(X_j^{(i)})^{1/\gamma}\right\}=0.$$

Clearly, there is no closed form for the MLE $(\hat{\gamma}, \hat{\alpha})$ of $(\gamma, \alpha)$, $(\hat{\gamma}, \hat{\alpha})$ can be solved theoretically only. The numerical values of $(\hat{\gamma}, \hat{\alpha})$ can be obtained via an iterative procedure; in practice, a standard computer program such as Fletcher and Powell (1963) is often used to maximize the likelihood function.

An alternative way to estimate $\gamma$ and $\alpha$ is the method of moment estimate. According to Yeh (1994) equations (1.1) and (1.4), it is easily discerned that each $X_j, j=1,2,\ldots,m$, in $\underline{X}=(X_1,X_2,\ldots,X_m)\sim MP^{(m)}(IV)$ $(\underline{0},\underline{1},\gamma,\alpha)$ is an univariate Pareto, $P(IV)$ $(0,1,\gamma,\alpha)$, variable, and hence from Arnold (1983)'s Chapter 3, equation (3.3.11), the $\delta$'th moments of each $X_j$ can be obtained as

$$E(X_j^\delta)=\Gamma(\alpha-\gamma\delta)\Gamma(1+\gamma\delta)/\Gamma(\alpha), -\frac{1}{\gamma}<\delta<\frac{\alpha}{\gamma} \text{ for } j=1,2,\ldots,m.$$

For each $j\in\{1,2,\ldots,m\}$, since $\{X_j^{(1)},X_j^{(2)},\ldots,X_j^{(n)}\}$ are i.i.d. $P(IV)$ $(0,1,\gamma,\alpha)$ variables, the method of moment estimates of $\gamma$ and $\alpha$ are obtained by equating first and second sample and theoretical moments. They are thus

$$M_1=\frac{1}{mn}\sum_{j=1}^{m}\sum_{i=1}^{n}X_j^{(i)}=\frac{\Gamma(\alpha-\gamma)\Gamma(1+\gamma)}{\Gamma(\alpha)}, \tag{2.5}$$

$$M_2=\frac{1}{mn}\sum_{j=1}^{m}\sum_{i=1}^{n}(X_j^{(i)})^2=\frac{\Gamma(\alpha-2\gamma)\Gamma(1+2\gamma)}{\Gamma(\alpha)}, \tag{2.6}$$

It is possible for the system of equations (2.5) and (2.6) not to have a solution, but this problem can be overcomed by taking large sample sizes.

**2.2 Interval estimates of $\alpha$ and $\gamma$.** The universal pivotal technique is utilized to construct a confidence interval of $\alpha$.

Let $\{\underline{X}^{(i)}=(X_1^{(i)},X_2^{(i)},\ldots,X_m^{(i)})\}_{i=1}^m$ be a random sample of size $n$ from a $MP^{(m)}(IV)(\underline{0},\underline{1},\gamma,\alpha)$ population, consider the transformation

$$Y_j^{(i)}=\frac{X_j^{(i)\,1/\gamma}}{1+(X_1^{(i)})^{1/\gamma}+(X_2^{(i)})^{1/\gamma}+\ldots,+(X_{j-1}^{(i)})^{1/\gamma}} \tag{2.7}$$

for each sample unit $i=1,2,\ldots,n$ and each $j^{th}$ coordinate, $j=1,2,\ldots,m$.

From Property 2.2 of Yeh (1994), knowing that for fixed $j\in\{1,2,\ldots,m\}$, the $n$'s random variables

$$\{Y_j^{(i)}\}_{i=1}^n \text{ i.i.d. } FP(0,1,1,\alpha+j-1,1)$$

and for fixed $i \in \{1, 2, \ldots, n\}$, the $m$'s random variables $\{Y_j^{(i)}\}_{j=1}^m$ are statistically independent, where $FP$ denotes the Feller-Pareto distribution and its relation to the univariate Pareto $(IV)$ distribution is identified as $FP$ $(0, 1, 1, \alpha + j - 1) = P(IV)$ $(0, 1, 1, \alpha + j - 1)$ with survival function

$$\bar{F}_{Y_j^{(i)}}(y) = (1+y)^{-(\alpha+j-1)}, \ y \geq 0, \ \text{(Arnold (1983). Ch. 3)}.$$

The following two theorems had been proved by Targhetta (1979):

**Theorem 2.1.** *If $X_1, X_2, \ldots, X_n$ are independently identically distributed (i.i.d.) with common univariate cdf $F(x;\theta)$ for every parameter $\theta$, then* (i) *the random variable*

$$-\sum_{i=1}^{n} \ln(F(X_i;\theta)) \sim \ Gamma(n, 1)$$

*and* (ii) $\{\theta : -\sum_{i=1}^n \ln(F(X_i;\theta)) \in (C_1, C_2)\}$ *is a* $100(1-\xi)\%$ *confidence interval of $\theta$, where $C_1, C_2$ represents respectively the lower $(\xi/2)^{\text{th}}$ and the upper $(\xi/2)^{\text{th}}$ percentile of a Gamma $(n, 1)$ distribution.*

Targhetta also generalized his result to an $m$-dimensional random vector with conditional distribution being well-defined. It is also stated as a Theorem

**Theorem 2.2.** *Let $\underline{X}^{(i)} = (X_1^{(i)}, X_2^{(i)}, \ldots, X_m^{(i)})$, for $i = 1., 2, \ldots, n$ be an $m$-dim random vector, let*

$$U_{i1}(\theta) = F_{X_1^{(i)}}(X_1^{(i)};\theta), \ \text{for } i = 1, 2, \ldots, n$$

*and* $$U_{ij}(\theta) = F_{X_j^{(j)}|X_1^{(i)}, X_2^{(i)}, \ldots, X_{j-1}^{(i)}}, (X_j^{(i)}|(X_1^{(i)}, X_2^{(i)}, \ldots, X_{j-1}^{(i)};\theta),$$

*for $i = 1, 2, \ldots, n$, $j = 2, \ldots, m$. Then* (i) *the random variable*

$$Z_n(\theta) = -\sum_{i=1}^{n}\sum_{j=1}^{m} \ln(U_{ij}(\theta)) \sim \ Gamma(mn, 1)$$

*and* (ii) *the region* $\{\theta : -\sum_{i=1}^n \sum_{j=1}^m \ln(U_{ij}(\theta)) \in (C_1, C_2)\}$ *is a* $100(1-\xi)\%$ *confidence interval of $\theta$, where $C_1$, is the lower $(\xi/2)^{th}$ and $C_2$ is the upper $(\xi/2)^{th}$ percentile of a Gamma $(mn, 1)$ distribution.*

Applying these two theorems to the random variable $\{Y_1^{(i)}, Y_2^{(i)}, \ldots, Y_m^{(i)}\}$ defined in equation (2.7) for each sample unit $i \in \{1, 2, \ldots, n\}$, take

$$
\begin{aligned}
U_{i1}(\alpha) &= \bar{F}_{Y_1^{(i)}}(Y_1^{(i)}; \alpha) = (1 + Y_1^{(i)})^{-\alpha} \sim U(0,1), \\
U_{i2}(\alpha) &= \bar{F}_{Y_2^{(i)}|Y_1^{(i)}}(Y_2^{(i)}|Y_1^{(i)}; \alpha) \\
&= \bar{F}_{Y_2^{(i)}}(Y_2^{(i)}; \alpha) = (1 + Y_2^{(i)})^{(-\alpha+1)} \sim U(0,1), \\
&\vdots \\
U_{ij}(\alpha) &= F_{Y_j^{(i)}|Y_1^{(i)}, Y_2^{(i)}, \ldots, Y_{j-1}^{(i)}}(Y_j^{(i)}|Y_1^{(i)}, Y_2^{(i)}, \ldots, Y_{j-1}^{(i)}; \alpha) \\
&= \bar{F}_{Y_j^{(i)}}(Y_j^{(i)}) = (1 + Y_j^{(i)})^{-(\alpha+j+1)} \sim U(0,1)
\end{aligned} \tag{2.8}
$$

for $j = 1, 2, \ldots, m$.

Since $\{Y_1^{(i)}, Y_2^{(i)}, \ldots, Y_m^{(i)}\}$ are independent, so

$$\{U_{i1}(\alpha), U_{i2}(\alpha), \ldots, U_{im}(\alpha)\} \text{ i.i.d. } U(0,1)$$

and for each $i = 1, 2, \ldots, n$,

$$\left\{ -\sum_{j-1}^{m} \ln(U_{ij}(\alpha)) \right\}_{i=1}^{n} \text{ i.i.d. } Gamma(m, 1)$$

so $-\sum_{i=1}^{n} \sum_{j=1}^{m} \ln(U_{ij}(\alpha)) \sim Gamma(mn, 1)$, hence, $-\sum_{i=1}^{n} \sum_{j=1}^{m} \ln(U_{ij}(\alpha))$ is a pivotal quantity for the shape parameter, $\alpha$. By the definition of $U_{ij}(\alpha)$ in equation (2.8), then $\ln(U_{ij}(\alpha)) = -(\alpha + j - 1)\ln(1 + Y_j^{(i)})$. Let

$$
\begin{aligned}
Z_n(\alpha) &= -\sum_{i=1}^{n} \sum_{j=1}^{m} \ln(U_{ij}(\alpha)) \\
&= \sum_{i=1}^{n} \sum_{j=1}^{m} (\alpha + j - 1)\ln(1 + Y_j^{(i)}) \sim Gamma(mn, 1).
\end{aligned}
$$

Expanding $Z_n(\alpha)$ by substituting $j = 1, 2, \ldots, m$ in the above expression, then

$$
\begin{aligned}
Z_n(\alpha) =& \alpha \sum_{i=1}^{n} \ln(1 + Y_1^{(i)}) + (\alpha + 1) \sum_{i=1}^{n} \ln(1 + Y_2^{(i)}) \\
&+ (\alpha + 2) \sum_{i=1}^{n} \ln(1 + Y_3^{(i)}) + \ldots + (\alpha + m - 1) \sum_{i=1}^{n} \ln(1 + Y_m^{(i)}).
\end{aligned}
$$

Let $T_j = \sum_{i=1}^{n} \ln(1 + Y_j^{(i)})$ for each $j = 1, 2, \ldots, m$, so

$$Z_n(\alpha) = \alpha(\sum_{j=1}^{m} T_j) + \sum_{j=2}^{m} (j-1)T_j.$$

Since $Z_n(\alpha)$ is a Gamma $(mn, 1)$ variable, so if we let $C_1, C_2$ represent the $(\xi/2)^{th}$ and the $(1-\xi/2)^{th}$ percentiles of a Gamma $(mn, 1)$ distribution, then

$$P\{C_1 < Z_n(\alpha) < C_2\} = 1 - \xi,$$

$$\text{i.e.,} \quad P\left\{C_1 < \alpha(\sum_{j=1}^{m} T_j) + \sum_{j=2}^{m}(j-1)T_j < C_1\right\} = 1 - \xi$$

$$\text{(2.9)} \quad \begin{aligned} &= P\left\{C_1 - \sum_{j=2}^{m}(j-1)T_j < \alpha(\sum_{j=1}^{m} T_j) < C_2 - \sum_{j=2}^{m}(j-1)T_j\right\} \\ &= P\left\{\frac{C_1 - \sum_{j=2}^{m}(j-1)T_j}{\sum_{j=1}^{m} T_j} < \alpha < \frac{C_2 - \sum_{j=2}^{m}(j-1)T_j}{\sum_{j=1}^{m} T_j}\right\} \end{aligned}$$

$$\text{Thus} \left(\frac{C_1 - \sum_{j=2}^{m}(j-1)T_j}{\sum_{j=1}^{m} T_j}, \frac{C_2 - \sum_{j=2}^{m}(j-1)T_j}{\sum_{j=1}^{m} T_j}\right)$$

is a $100(1-\xi)$ % confidence interval for $\alpha$. The two limits of this confidence interval are statistic if the parameter $\gamma$ is known. However, if $\gamma$ is unknown, then it should be estimated first, get $\hat{\gamma}$ and plug in the definition of $Y_j^{(i)} = \frac{(X_j^{(i)})^{1/\hat{\gamma}}}{1+\sum_{l=1}^{j-1}(X_l^{(i)})^{1/\hat{\gamma}}}$, and hence the confidence interval in equation (2.9) can then be evaluated.

**3. Marida's multivariate Pareto families.** Marida (1962) concentrated on the analysis of his second type of bivariate Pareto by applying the transformation $X_i = \sigma_i e^{U_i/\alpha_i}$, $i = 1, 2$, with $(U_1, U_2)$ having a bivariate Wicksell-Kibble Exponential distribution. Arnold (1983) did some inferences for this bivariate Pareto distribution. The more general $m$-variate Pareto distribution, denoted by $WKP^{(m)}(I)$ $(\underline{\sigma}, \underline{\alpha}, p), m \geq 3$, is discussed in this section. It is obtained by an $m$-variate Wicksell-Kibble exponential random vector $\underline{U} = (U_1, U_2, \ldots, U_m)$, each $U_i$ is a geometric sum of i.i.d.

exponential random variables with mean $p$, ie.

$$(3.1)\qquad \begin{aligned} &U_i = \sum_{j=1}^{N} U_{ij}, i = 1, 2, \ldots, m \\ \text{and}\quad &N \text{ is independent of each } U_{ij} \text{ with pmf} \\ &P_N(n) = p(1-p)^{n-1},\ n = 1, 2, \ldots. \end{aligned}$$

The random variable $U_i$ defined in this manner is the standard exponential, $Exp(1)$, random variable.

Let $\underline{X} = (X_1, X_2, \ldots, X_m)$ be an $m$-variate random vector and $\underline{\sigma} = (\sigma_1, \sigma_2, \ldots, \sigma_m)$, $\underline{\alpha} = (\alpha_1, \alpha_2, \ldots, \alpha_m)$, $\underline{\sigma} > 0$ and $\underline{\alpha} > 0$ are two fixed parameter vectors. If each coordinated $X_j$ is defined by $X_j = \sigma_j e^{U_j/\alpha_j}$, $j = 1, 2, \ldots, m$, then the random vector $\underline{X}$ is distributed as $WKP^{(m)}(I)$ $(\underline{\sigma}, \underline{\alpha}, p)$.

From the construction of each $U_i$, $i = 1, 2, \ldots, m$, knowing that given $N = n$, the conditional pdf of $U_i|_{N=n}$ is distributed as a Gamma $(n, p)$ variable. Because the survival function of Gamma variable has no closed form, so the joint survival functions of $\underline{U}$ and hence $\underline{X}$ are not tractable. Although Marida (1962) gave the joint pdf of $\underline{X}$, yet its form is too complicate to be useful. However, the joint characteristic function of $\underline{U}$ can be easily derived as

$$(3.2)\qquad \begin{aligned} &\varphi_{\underline{U}}(t_1, t_2, \ldots, t_m) \\ &= \sum_{n=1}^{\infty} E(e^{i(t_1U_1 + t_2U_2 + \ldots, + t_mU_m)}|_{N=n}) P_N(n) \\ &= \sum_{n=1}^{\infty} (1 - ipt_1)^{-n} \cdot (1 - ipt_2)^{-n} \ldots (1 - ipt_m)^{-n} \cdot p(1-p)^{n-1} \\ &= \Big\{ 1 - i(\sum_{i=1}^{m} t_i) - p(\sum_{1=i}^{m-1} \sum_{<j=2}^{m} t_i t_j) + ip^2 \sum_{1=i}^{m-2} \sum_{<j<k}^{m-1} \sum_{=3}^{m} (t_i t_j t_k) \\ &\quad - \ldots + (-i)^m p^{m-1} \prod_{i=1}^{m} t_i \Big\}^{-1} \end{aligned}$$

As for the two parameter vectors $\underline{\sigma} = (\sigma_1, \sigma_2, \ldots, \sigma_3)$ and $\underline{\alpha} = (\alpha_1, \alpha_2, \ldots, \alpha_m)$ in the $WKP^{(m)}(I)$ $(\underline{\sigma}, \underline{\alpha}, p)$ families, they can be treated by applying the transformation ln on $X_i = \sigma_i e^{U_i/\alpha_i}, i = 1, 2, \ldots, m$, i.e., define

$$Y_i = \ln(X_i) = \ln \sigma_i + U_i/\alpha_i, \; i = 1, 2, \ldots, m. \tag{3.3}$$

Since each $Y_i$ is a linear transformation of the standard exponential variable $U_i$, so it is obvious that the random vector $\underline{Y} = (Y_1, Y_2, \ldots, Y_m)$ has an $m$-variate Wicksell-Kibble exponential distribution with location parameter $(\ln \sigma_1, \ln \sigma_2, \ldots, \ln \sigma_m)$ and scale parameter $(1/\alpha_1, 1/\alpha_2, \ldots, 1/\alpha_m)$. The joint characteristic function of $\underline{Y}$ can be obtained from equation (3.2) as

$$\begin{aligned} &\varphi_{\underline{Y}}(t_1, t_2, \ldots, t_m) \\ =&e^{i(t_1 \ln \sigma_1 + t_2 \ln \sigma_2 + \ldots + t_m \ln \sigma_m)} \\ &\cdot \left\{ 1 - i \sum_{j=1}^{m} \frac{t_j}{\alpha_j} - p \sum_{1=j}^{m-1} \sum_{<k=2}^{m} \frac{t_j}{\alpha_j} \frac{t_k}{\alpha_k} + ip^2 \sum_{1=j}^{m-2} \sum_{<k<}^{m} \sum_{l=3}^{m} \left( \frac{t_j t_k t_l}{\alpha_j \alpha_k \alpha_l} \right) \right. \\ &\left. + \ldots + (-i)^m p^{m-1} \prod_{j=1}^{m} \frac{t_j}{\alpha_j} \right\}^{-1}. \end{aligned} \tag{3.4}$$

The joint characteristic function of $\underline{X} = (X_1, X_2, \ldots, X_m)$ is not tractable, yet $\underline{X}$ has the following property:

**Property 3.1.** *Each $X_i = \sigma_i e^{U_i/\alpha_i}$ in $\underline{X}$ is a univariate classical Pareto $P(I)$ $(\sigma_i, \alpha_i)$.*

*Proof.* Consider the marginal survival function of each $X_i$, given $x \geq \sigma_i$,

$$\begin{aligned} \bar{F}_{X_i}(x) &= P(X_i \geq x) = P(\sigma_i e^{U_i/\alpha_i} \geq x) \\ &= P(U_i \geq \alpha_i \ln(\frac{x}{\sigma_i})) \\ &= e^{-\alpha_i \ln(\frac{x}{\sigma_i})} = (\frac{x}{\sigma_i})^{-\alpha_i} \sim P(I)(\sigma_i, \alpha_i). \end{aligned} \tag{3.5}$$

The parameter $p$ which appears in the joint characteristic function of $\underline{U}$ is equation (3.2) is actually related to the pairwise correlation between $U_i$ and $U_j, i \neq j, i, j \in \{1, 2, \ldots, m\}$.

**Property 3.2.** *Let $\underline{U}$ be an $m$-variate Wicksell-Kibble standard exponential random vector, then each pair $(U_i, U_j)$ in $\underline{U}$ is correlated, for $i \neq$*

*$j, j \in \{1, 2, \ldots, m\}$, and all their correlations are equal to*

$$\rho = \rho(U_i, U_j) = 1 - p$$

*Proof.* Let $\rho_{ij}$ be the correlation between $U_i$, and $U_j$, there are totally $\binom{m}{2} = \frac{m(m-1)}{2}$ such correlations. By definition

$$\rho_{ij} = \frac{\mathrm{Cov}(U_i, U_j)}{\sqrt{\mathrm{Var}U_i \cdot \mathrm{Var}\ U_j}} = \frac{E(U_iU_j) - E(U_i)E(U_j)}{\sqrt{[E(U_i^2) - (E(U_i))^2][E(U_j^2) - (E(U_j))^2]}}.$$

Recall in equation (3.1), each $U_i$ is a standard exponential variablle, so the mean of $U_i$ is $E(U_i) = 1$, and the variance of $U_i$ is $\mathrm{Var}(U_i) = 1$. As for $E(U_iU_j)$, it is calculated by the conditional expectation

$$E(U_1U_2) = E_N(E(U_1U_2)|_N),$$

given the random variable $N = n$, $U_1|_{N=n}$ and $U_2|_{N=n}$ are independently distributed as a Gamma $(n, p)$ variable, hence

$$\begin{aligned} E(U_1U_2) &= \sum_{n=1}^{\infty} E(\sum_{j=1}^{n} U_{1j})E(\sum_{j=2}^{n} U_{2j})p(1-p)^{n-1} \\ &= \sum_{n=1}^{\infty} (np)^2 p(1-p)^{n-1} = p^3 \sum_{n=1}^{\infty} n^2(1-p)^{n-1} \\ &= p^3 \sum_{n=1}^{\infty} \{n(n-1) + n\}(1-p)^{n-1} \\ &= p^3 \left\{ \sum_{n=1}^{\infty} n(n-1)(1-p)^{n-1} + \sum_{n=1}^{\infty} n(1-p)^{n-1} \right\} \\ &= p^3 \left\{ (1-p)\frac{2}{p^3} + \frac{1}{p^2} \right\} = 2(1-p) + p = 2 - p, \end{aligned}$$

hence $\rho_{12} = \frac{(2-p)-1}{\sqrt{1}} = 1 - p$.

By symmetry, all the correlations between any pair $(U_i, U_j)$ in $\underline{U}$ are equal, and $\rho_{ij} = \rho = 1 - p$ and thus the role of the parameter $p$ in the joint characteristic function (equation (3.2)) is just $1 - \rho$, where $\rho$ is the common correlation of any pair $(U_i, U_j)$ in $\underline{U}$.

The parameter $p$ which appears in the joint characteristic function $\varphi_{\underline{U}}$ $(t_1, t_2, \ldots, t_m)$ is actually related to the correlation between any pair $(Y_i, Y_j)$, $i \neq j$, in $\underline{Y} = (Y_1, Y_2, \ldots, Y_m)$.

Since the relation between $Y_j$ and $U_j$ is a linear transformation, i.e., $Y_j = \ln(X_j) = \ln \sigma_j + U_j/\alpha_j$, so the correlation between $Y_j$ and $Y_k$ is the same as $\rho_{jk} = \mathrm{Corr}(U_j, U_k) = \rho = 1 - p$.

Let $\underline{X} = (X_1, X_2, \ldots, X_m)$, $\sim WKP^{(m)}(I)(\underline{\sigma}, \underline{\alpha}, p)$, the correlation of any pair $(X_i, X_j)$ in $\underline{X}$

$$\rho(X_i, X_j) = \frac{E(X_i, X_j) - E(X_i)E(X_j)}{\sqrt{\mathrm{Var}(X_i) \cdot \mathrm{Var}(X_j)}}, i \neq j$$

can be calculated from the fact of property 3.1 that each $X_i$ in $\underline{X}$, $i = 1, 2, \ldots, m$ is marginally distributed as a univariate classical Pareto $P(I)(\sigma_i, \alpha_i)$ variable, hence

$$E(X_i) = \frac{\alpha_i \sigma_i}{\alpha_i - 1}, \ \mathrm{Var}(X_i) = \frac{\alpha_i \sigma_i^2}{(\alpha_i - 1)^2(\alpha_i - 2)}.$$

As for $E(X_i X_j)$, it is easily obtained from the relation $X_i = \sigma_i e^{U_i/\alpha_i}$ and the joint mgf of $(U_i, U_j)$,

$$M_{(U_i, U_j)}(t_i, t_j) = E(e^{t_i U_i + t_j U_j}).$$

$$\begin{aligned} E(X_i X_j) &= E(\sigma_i \sigma_j e^{U_i/\alpha_i + U_j/\alpha_j}) = \sigma_i \sigma_j E(e^{1/\alpha_i U_i + 1/\alpha_j U_j}) \\ &= \sigma_i \sigma_j M_{(U_i, U_j)}(\frac{1}{\alpha_i}, \frac{1}{\alpha_j}), \end{aligned}$$

where the mgf of $(U_i, U_j)$ can be obtained from the joint characteristic function of $\underline{U} = (U_1, U_2, \ldots, U_m)$ in equation (3.2) by letting all the other $t_k = 0$, for those $k \in \{1, 2, \ldots, m\} - \{i, j\}$, and letting $t_i = \frac{-i}{\alpha_i}, t_j = \frac{-i}{\alpha_j}$ in equation (3.2), then

$$E(X_i X_j) = \frac{\sigma_i \sigma_j}{1 - (\frac{1}{\alpha_i} + \frac{1}{\alpha_j}) + p(\frac{1}{\alpha_i \alpha_j})},$$

after some algebraic simplifications, the correlation between $X_i, X_j$ is

$$\rho(X_i, X_j) = \frac{(1-p)\sqrt{\alpha_i\alpha_j(\alpha_i-2)(\alpha_j-2)}}{(\alpha_i-1)(\alpha_j-1)-(1-p)}$$

for all $X_i, X_j$ in $\underline{X} = (X_1, X_2, \ldots, X_m) \sim WKP^{(m)}(I)(\underline{\sigma}, \underline{\alpha}, p)$, $\forall i \neq j$.

If $p = 1$, then $\rho(U_i, U_j) = \rho(Y_i, Y_j) = \rho(X_i, X_j) = 0$, i.e. $U_i$ and $U_j$ and thus $Y_i$ and $Y_j$, $X_i$ and $X_j$ are uncorrelated. Moreover, from equation (3.2), the joint characteristic function of $\underline{U}$, know that if $p = 1$, then all the $U_i, i = 1, 2, \ldots, m$, in $\underline{U}$ are independently distributed and hence all the $X_i, i = 1, 2, \ldots, m$ in $\underline{X}$ are also independently distributed, so the multivariate $WKP^{(m)}(I)(\underline{\sigma}, \underline{\alpha}, p)$ distribution is analogous to the multivariate normal distribution. It is stated as a theorem below.

**Theorem 3.1.** *Let $\underline{X} = (X_1, X_2, \ldots, X_m) \sim WKP^{(m)}(I)(\underline{\sigma}, \underline{\alpha}, p)$, then for any pair $X_i, X_j$ in $\underline{X}$, $i \neq j$, $\rho(X_i, X_j) = 0$ if and only if all the $X_i, i = 1, 2, \ldots, m$ are independently distributed.*

**3.2 Estimation for the multivariate pareto $WKP^{(m)}(I)(\underline{\sigma}, \underline{\alpha}, p)$ distributions.** Let $\underline{X}^{(i)} = (X_1^{(i)}, X_2^{(i)}, \ldots, X_m^{(i)})$, $i = 1, 2, \ldots, n$, be a random sample of size $n$ from a $WKP^{(m)}(I)(\underline{\sigma}, \underline{\alpha}, p)$ population.

For each $i = 1, 2, \ldots, n, j = 1, 2, \ldots, m$,

$$X_j^{(i)} = \sigma_j e^{U_j^{(i)}/\alpha_j}, \tag{3.6}$$

where $U_j^{(i)}$ is given as in equation (3.1), consider the transformation $Y_j^{(i)} = \ln X_j^{(i)}$, then the random vector

$$\underline{Y}^{(i)} = (Y_1^{(i)}, Y_2^{(i)}, \ldots, Y_m^{(i)}),\ i = 1, 2, \ldots, n$$

is also Wicksell-Kibble $m$-variate exponentially distributed, hence the inferences for the parameters $\underline{\sigma} = (\sigma_1, \ldots, \sigma_m), \underline{\alpha} = (\alpha_1, \ldots, \alpha_m)$, and $0 < p < 1$ in the $WKP(\underline{\sigma}, \underline{\alpha}, p)$ population reduce to that of analyzing a random sample from an $m$-variate exponential population.

**Theorem 3.2.** *For each $j^{th}$ coordinate, $j = 1, 2, \ldots, m$, the MLE of $\sigma_j$ and $\alpha_j$ are respectively*

$$\hat{\sigma}_j = \min X_j^{(i)}, \hat{\alpha}_j = (\bar{Y}_j - \ln \hat{\sigma}_j)^{-1}, \ \textit{where } \bar{Y}_j = \frac{\sum_{i=1}^{n} Y_j^{(i)}}{n}.$$

*Proof.* For fixed $j \in \{1, 2, \ldots, m\}$, the random variables

$$Y_j^{(i)} = \ln\ \sigma_j + \frac{U_j^{(i)}}{\alpha_j},\ i = 1, 2, \ldots, m$$

are i.i.d. distributed as two-parameters exponential $Exp(\alpha_j, \ln \sigma_j)$ variable with pdf

$$f_{Y_j^{(i)}}(y_j^{(i)}) = \begin{cases} \alpha_j e^{-\alpha_j (y_j^{(i)} - \ln \sigma_j)} & \text{if } y_j^{(i)} \geq \ln \sigma_j, \\ 0 & \text{o.w.} \end{cases}$$

The log-likelihood function of $\alpha_j$ is

$$\ln(L) = n \ln \alpha_j - \alpha_j \sum_{i=1}^{n} y_j^{(i)} + n\alpha_j (\ln \sigma_j). \tag{3.7}$$

The MLE of $\sigma_j$ is discerned by $y_j^{(i)} = \ln X_j^{(i)} \geq \ln \sigma_j$ for all $i = 1, 2, \ldots, n$ and also from equation (3.6), knowing that $X_j^{(i)} \geq \sigma_j$, $\forall i = 1, 2, \ldots, n$, hence it is clearly that the MLE of $\sigma_j$ is $\min\limits_{1 \leq i \leq n} X_j^{(i)}$. The MLE of $\alpha_j$ is obtained by solving

$$\frac{\partial \ln(L)}{\partial \alpha_j} = -\sum_{i=1}^{n} y_j^{(i)} + n \ln \sigma_j + \frac{n}{\alpha_j} = 0, \tag{3.8}$$

so

$$\hat{\alpha}_j = \frac{n}{\sum_{i=1}^{n} y_j^{(i)} - n \ln \hat{\sigma}_j} = \frac{1}{\bar{y}_j - \ln \hat{\sigma}_j}, \tag{3.9}$$

moreover,

$$\left.\frac{\partial^2 \ln(L)}{\partial^2 \alpha_j}\right|_{\hat{\alpha}_j} = \left.-\frac{n}{\alpha_j^2}\right|_{\hat{\alpha}_j} < 0. \tag{3.10}$$

Thus,

$$\hat{\alpha}_j = (\bar{Y}_j - \ln \hat{\sigma}_j)^{-1} = \left(\bar{Y}_j - \ln\left(\min_{1 \leq i \leq n} X_j^{(i)}\right)\right)^{-1}. \tag{3.11}$$

is the MLE of $\alpha_j$ for each $j = 1, 2, \ldots, m$, whre $\bar{Y}_j = \frac{\sum_{i=1}^{n} Y_j^{(i)}}{n}$.

**Theorem 3.3.** *For each fixed* $j \in \{1,2,\ldots,m\}$, (1) *the exact sampling distribution of* $\hat{\sigma}_j$ *is classical Pareto*

$$P(I)(\sigma_j, n\alpha_j) \text{ with pdf } f_{\hat{\sigma}_j}(x) = \begin{cases} n\alpha_j(\sigma_j)^{n\alpha_j} x^{-(n\alpha_j+1)}, & \text{if } x > \sigma_j \\ 0, & \text{o.w.} \end{cases}$$

*and also* $2n\alpha_j \ln(\hat{\sigma}_j/\sigma_j) \sim \chi^2_{(2)}$.

(2) *The exact sampling distribution of* $\hat{\alpha}_j$ *is* $\hat{\alpha}_j \sim Gamma(n-1, (n\alpha_j)^{-1})$, *or, equivalently* $2n\alpha_j/\hat{\alpha}_j \sim \chi^2_{(2n-2)}$.

(3) $\hat{\sigma}_j$ *and* $\hat{\alpha}_j$ *are independently distributed.*

*Proof.* Refer to Arnold (1983) chapter 5 and Baxter (1980).

It is easy to determine the means and variances of the MLE $\hat{\sigma}_j$ and $\hat{\alpha}_j$ from the above theorem.

**Property 3.3.** *For each* $j \in \{1,2,\ldots,m\}$

(i) $E(\hat{\sigma}_j) = \sigma_j(1 - \frac{1}{n\alpha_j})^{-1} > \sigma_j$, *ie.,* $\hat{\sigma}_j$ *always overestimates* $\sigma_j$; $Var\ (\hat{\sigma}_j) = \sigma_j^2 n\alpha_j(n\alpha_j - 1)^{-2}(n\alpha_j - 2)^{-1}$.

(ii) $E(\hat{\alpha}_j) = \alpha_j n(n-2)^{-1} \neq \alpha_j$, *i.e.,* $\hat{\alpha}_j$ *is a biased estimator of* $\alpha_j$ *and* $Var\ (\hat{\alpha}_j) = (\alpha_j n)^2(n-2)^{-2}(n-3)^{-1}$.

The following two properties deal the asymptotic behaviors of $\hat{\sigma}_j$ and $\hat{\alpha}_j$ as the sample size $n$ is very large.

**Property 3.4.** *For each* $j \in \{1,2,\ldots,m\}$, *the MLE of* $\sigma_j, \hat{\sigma}_j = \min_{1\le i\le n} X_j^{(i)}$ *is a consistent estimator of* $\sigma_j$, *and the asymptotic distribution of* $\hat{\sigma}_j$ *is an exponential distribution, more specifically, for* $z > 0$

$$\lim_{n\to\infty} P\{n\alpha_j(\hat{\sigma}_j/\sigma_j - 1) > z\} = e^{-z}.$$

*Proof.* From Property 3.1, for $j \in \{1,2,\ldots,m\}$, each $X_j^{(i)}$ is a classic Pareto $P(I)(\sigma_j, \alpha_j)$ variable for all $i = 1,2,\ldots,n$, so it satisfies the "regularity smooth conditions" in distribution theory, hence its sample extreme $\hat{\sigma}_j = \min_{1\le i\le n} X_j^{(i)}$ is a strong consistent estimator of $\sigma_j$, also, the survival function of $\hat{\sigma}_j \sim P(I)(\sigma_j, n\alpha_j)$ is $P(\hat{\sigma}_j > t) = (t/\sigma_j)^{-n\alpha_j}$, $t \ge \sigma_j$, hence

$$P\Big\{(n\alpha_j)(\hat{\sigma}_j/\sigma_j - 1) > z\Big\} = P\Big\{\hat{\sigma}_j > \sigma_j(1 + \frac{z}{n\alpha_j})\Big\}$$
$$= (1 + \frac{z}{n\alpha_j})^{-n\alpha_j} \to e^{-z}, \text{ as } n \to \infty.$$

**Property 3.5.** *Let the sample size $n$ be large, $n > 30$, then for each $j \in \{1, 2, \ldots, m\}$, the MLE of $\alpha_j$, $\hat{\alpha}_j = (\bar{Y}_j - \ln(\min_{1 \le i \le n} X_j^{(i)}))^{-1}$ is a consistent estimator of $\alpha_j$ and the asymptotic distribution of $\hat{\alpha}_j$ is a normal, $N(\alpha_j, \alpha_j^2/n)$, variable.*

*Proof.* It is easily discerned that the regularity conditions of the log-likelihood function of $\alpha_j$ equation (3.7) are satisfied, hence by the optimum properties of the MLE, $\hat{\alpha}_j$, is consistent and asymptotically distributed as a normal $N(\alpha_j, \frac{-1}{nE\left(\frac{\partial^2 \ln(L)}{\partial^2 \alpha_j}\right)})$ variable, (Kulldorf (1957)), where $E(\frac{\partial^2 \ln(L)}{\partial^2 \alpha_j})$ is obtained from equation (3.10), $\frac{\partial^2 \ln(L)}{\partial^2 \alpha_j} = -\frac{n}{\alpha_j^2}$, therefore, when $n$ is large ($n > 30$), the asymptotic distribution of $\hat{\alpha}_j$ is normally $N(\alpha_j, \frac{\alpha_j^2}{n})$, distributed.

For the parameter $p, 0 < p < 1$, it is derived in Property 3.2, know that $p = 1 - \rho$, where $\rho$ is the common correlation of any pair $(U_i, U_j)$ in $\underline{U}$.

In practice, use sample correlation $r_{jk}$ between $Y_j$ and $Y_k$ to estimate $\rho_{jk}$, i.e.,

$$r_{jk} = \frac{\sum_{i=1}^n (Y_j^{(i)} - \bar{Y}_j)(Y_k^{(i)} - \bar{Y}_k)}{\sqrt{\sum_{i=1}^n (Y_j^{(i)} - \bar{Y}_j)^2 \cdot \sum_{i=1}^n (Y_k^{(i)} - \bar{Y}_k)^2}}.$$

There are totally $\binom{m}{2} = \frac{m(m-1)}{2}$ such $r_{jk}$'s to estimate the common $\rho$ by pooling $\{r_{jk}\}$ together and get the moment estimator, $\hat{\rho}$, of $\rho$, i.e.,

$$\hat{\rho} = \frac{\sum\limits_{j=k}^{m-1} \sum\limits_{k=2}^{m} r_{jk}}{\binom{m}{2}} \tag{3.12}$$

Thus, the parameter $p$ is estimated by

$$\hat{p} = 1 - \hat{\rho} \tag{3.13}$$

**Property 3.6.** *If the sample size $n$ is large, $n > 30$, then the moment estimator of $p$ is $\hat{p} = 1 - \hat{\rho}$, where $\hat{\rho}$ is defined in equation* (3.12), *$\hat{p}$ is consistent, and $\hat{p}$ is asymptotically normally distributed.*

*Proof.* From equation (3.13), know that $\hat{\rho}$ is the moment estimator of $\rho = \mathrm{Corr}(Y_j, Y_k)$ where $Y_j = \ln(X_j)$ and $(Y_j, Y_k)$ is any pair in the random vector $\underline{Y} = (Y_1, Y_2, \ldots, Y_m)$, so $\hat{\rho}$ is consistent and the asymptotic normality of $\hat{\rho}$ is followed by the central limit theorem on the sample moment estimator. Since $\hat{p} = 1 - \hat{\rho}$, is a linear transformation, so $\hat{p}$ is also consistent and asymptotically normally distributed.

## References

1. Arnold, *Pareto Distributions*, International Cooperative Publ. House, Fairland, Maryland, 1983.
2. Baxter, *Minimum variance unbiased estimation of the parameters of the Pareto distribution*, Metrika, **27** (1980), 133-138.
3. Fletcher and Powell, *A rapid descent method for minimization*, Computer Journal, **6:2** (1963), 163-168.
4. Kibble, *A two variable Gamma type distribution*, Sankhya, **5** (1941), 137-150.
5. Kulldorf, *On the conditions for consistency and asymptotic efficiency of maximum likelihood estimates*, Skand. Aktuarietidskr, **40** (1957), 129-144.
6. Marida, *Multivariate Pareto distributions*, Annals of Math. Stat., **33** (1962), 1008-1015.
7. Targhetta, *Confidence intervalus for a one-parameter family of multivariate distributions*, Biometrika, **66** (1979), 687-688.
8. Wicksell, *On correlation functions of type III*, Biometrika, **25** (1933), 121-133.
9. Yeh, *Some properties of the homogeneous multivariate Paretor (IV) distributions*, Journal of Multivariate Analysis, **51:1** (1994), 46-53.

Department of Finance, National Taiwan University, Republic of China