

Sri Lankan Journal of Applied Statistics, Vol (15-3)

Modeling and Estimation of a Bivariate Pareto Distribution using the Principle of Maximum Entropy

Jagathnath Krishna K.M.*

Economics Research Division, CSIR-Central Leather Research Institute,
Adyar, Chennai – 600 020, India

*Corresponding Author: jagath.krishna@gmail.com

Received: 24, November 2013 / Revised: 1, July 2014 / Accepted: 9, October 2014

© IAppStat-SL 2013

ABSTRACT

In this paper we modeled a bivariate Pareto I distribution using the method of principle of maximum entropy probability distribution. Properties of the model are discussed. Further the estimation of the parameters involved in the model is done in two stages using two different methods namely, principle of maximum entropy estimation (POME) and maximum likelihood estimation. From the simulation study conducted to compare the performance of the estimates obtained by the above two methods, we conclude that POME method is performing better than MLE and the two methods are comparable.

Keywords: Bivariate Pareto I, Maximum Entropy probability Distribution, POME, MLE, Bivariate Failure Rate.

1. Introduction

Every probability distribution has some uncertainty associated with it. Accordingly, for some given partial information about some characteristics of the distribution or the random variate, we wish to derive a model that best approximate the distribution which is consistent with the given information. An approach to produce a model for the data generating distribution is the well-known maximum entropy method.

Let X be a non-negative random variable representing the lifetime of a component with distribution function $F(x) = P[X \leq x]$, survival function $\bar{F}(x) = 1 - F(x)$ and probability density function $f(x)$. Then the entropy function (Shannon (1948)) which provides a quantitative measure for the uncertainty of the random variable X is given by

$$H(f) = -\int_0^{\infty} f(x) \log f(x) dx. \tag{1.1}$$

The maximum entropy distribution is the distribution whose probability density function $f(\cdot)$ maximizes $H(f)$ in a set of distributions with given constraints. In maximum entropy principle, we begin with the fact that the distribution function is unknown. It is well known that among the set of distribution, the most uncertain distribution is the uniform. Thus in maximum entropy procedure, we have access to some information regarding the characteristic properties of the life time of the component or system. The additional information or constraints help us to obtain an appropriate model that maximizes the entropy function $H(f)$ (see Kapur (1989)).

In reliability studies, the problem of interest is the residual life time of a component which has survived beyond an age t . Ebrahimi (1996) has defined the residual entropy function applicable to such situations. The maximum entropy distribution for the univariate residual life time distributions has been discussed in Ebrahimi (2000) and Asadi et. al. (2004).

Now let (X_1, X_2) represents the life time of a two component parallel system with survival function $\bar{F}(x_1, x_2) = P[X_1 > x_1, X_2 > x_2]$, the bivariate density function $f(x_1, x_2)$ and total failure rate $\lambda(t)$ Cox's (1972). Then the bivariate residual entropy function (Asha et. al (2009)) $H(f, t, t_1, t_2)$ is given by

$$H(f, t, t_1, t_2) = (H(f_X, t), H_{12}(f, t_1, t_2), H_{21}(f, t_1, t_2)) \tag{1.2}$$

where

$$H(f_X, t) = 1 - \frac{1}{\bar{F}_X(t)} \int_t^{\infty} \lambda(x) \bar{F}_X(x) \log \lambda(x) dx; \quad t > 0 \tag{1.3}$$

$$H_{12}(f, t_1, t_2) = 1 - \frac{1}{\left[\frac{\partial \bar{F}(t_1, u)}{\partial u} \right]_{u=t_2}} \int_{t_1}^{\infty} \lambda_{12}(x_1 | t_2) \left[\frac{\partial \bar{F}(x_1, u)}{\partial u} \right]_{u=t_2} \log \lambda_{12}(x_1 | t_2) dx_1; t_1 > t_2 \tag{1.4}$$

$$H_{21}(f, t_1, t_2) = 1 - \frac{1}{\left[\frac{\partial \bar{F}(u, t_2)}{\partial u} \right]_{u=t_1}} \int_{t_2}^{\infty} \lambda_{21}(x_2 | t_1) \left[\frac{\partial \bar{F}(u, x_2)}{\partial u} \right]_{u=t_1} \log \lambda_{21}(x_2 | t_1) dx_2; t_1 < t_2 \tag{1.5}$$

The paper is organized into seven sections. In section 2, we derived a bivariate Pareto model and discussed its properties. In section 3, we focused on the estimation of only the scale parameter by using the principle of maximum entropy estimation (POME) and maximum likelihood estimation (MLE) from the density of minimum. The estimation of the rest of the parameters are obtained by substituting the estimate of the scale parameter obtained from the univariate minimum density in the MLE for the bivariate Pareto I distribution. Also an asymptotic property of the model is also discussed in section 4. In section 5, simulation studies for the estimates of the parameters by using the two methods are done and root mean square errors of the estimates have also been discussed. In section 6, a discussion based on the simulation is given and finally concluding remarks is given in section 7.

2. Maximum Entropy Probability Distribution

In this section, we formalized a bivariate maximum entropy distribution. The model is obtained by maximizing the bivariate residual entropy function $H(f, t, t_1, t_2)$ given in (1.2) subject to a series of constraints and is discussed as follows.

Result 1: Maximize $H(f, t, t_1, t_2)$ overall probability density function $f(x_1, x_2)$ with support of $(\sigma, \infty) \times (\sigma, \infty)$ satisfying

(i) $f(x_1, x_2) > 0$ for all $x_1, x_2 > \sigma$

(ii) $\int_{\sigma}^{\infty} \int_{\sigma}^{\infty} f(x_1, x_2) dx_1 dx_2 = 1$ and $p_i = P[X_i < X_j], i, j = 1, 2, i \neq j.$

(iii)
$$\left(\frac{\left(\frac{d\lambda(t)}{dt} \right)}{\lambda(t)}, \frac{\left(\frac{\partial \lambda_{12}(t_1 | t_2)}{\partial t_1} \right)}{\lambda_{12}(t_1 | t_2)}, \frac{\left(\frac{\partial \lambda_{21}(t_2 | t_1)}{\partial t_2} \right)}{\lambda_{21}(t_2 | t_1)} \right) \geq (-k\lambda(t), -k_1\lambda_{12}(t_1 | t_2), -k_2\lambda_{21}(t_2 | t_1))$$

with $\lambda(\sigma) = \frac{1}{k\sigma}$, $\lambda_{12}(\sigma | t_2) = \frac{1}{k_1\sigma}$ and $\lambda_{21}(\sigma | t_1) = \frac{1}{k_2\sigma}$ is the bivariate

Pareto I distribution.

Proof:

To obtain the maximum entropy function differentiate (1.3), (1.4) and (1.5) with respect to t, t_1 and t_2 and equating to zero. Thus we have

$$\begin{aligned} \frac{d}{dt} H(f_x, t) &= \lambda(t) [\log \lambda(t) - 1 + H(f_x, t)] = 0 \\ \Rightarrow H(f_x, t) &= 1 - \log \lambda(t) \end{aligned} \tag{2.1}$$

$$\begin{aligned} \frac{\partial}{\partial t_1} H_{12}(f, t_1, t_2) &= \lambda_{12}(t_1 | t_2) [\log \lambda_{12}(t_1 | t_2) - 1 + H_{12}(f, t_1, t_2)] = 0 \\ \Rightarrow H_{12}(f, t_1, t_2) &= 1 - \log \lambda_{12}(t_1 | t_2) \end{aligned} \tag{2.2}$$

$$\begin{aligned} \frac{\partial}{\partial t_2} H_{21}(f, t_1, t_2) &= \lambda_{21}(t_2 | t_1) [\log \lambda_{21}(t_2 | t_1) - 1 + H_{21}(f, t_1, t_2)] = 0 \\ \Rightarrow H_{21}(f, t_1, t_2) &= 1 - \log \lambda_{21}(t_2 | t_1) \end{aligned} \tag{2.3}$$

Now we have to obtain the optimum values for $\lambda(t)$, $\lambda_{12}(t_1 | t_2)$ and $\lambda_{21}(t_2 | t_1)$ that maximizes (2.1), (2.2) and (2.3). Hence we will maximize the entropy functions (2.1), (2.2) and (2.3) subject to the constraints (i), (ii) and (iii).

From (iii), we have,

$$\frac{\left(\frac{d\lambda(t)}{dt} \right)}{\lambda^2(t)} \geq -k \tag{2.4}$$

$$\frac{\left(\frac{\partial \lambda_{12}(t_1 | t_2)}{\partial t_1} \right)}{\lambda_{12}^2(t_1 | t_2)} \geq -k_1 \tag{2.5}$$

$$\frac{\left(\frac{\partial \lambda_{21}(t_2 | t_1)}{\partial t_2} \right)}{\lambda_{21}^2(t_2 | t_1)} \geq -k_2 \tag{2.6}$$

Integrating (2.4) with respect to t , we get

$$\lambda(t) \geq \frac{1}{kt}$$

Using the condition $\lambda(\sigma) = \frac{1}{k\sigma}$ in constrain (iii), the inequality becomes

$$\lambda(t) = \frac{1}{kt}$$

Integrating (2.5) with respect to t_1 , we get

$$\lambda_{12}(t_1 | t_2) \geq \frac{1}{k_1 t_1}$$

Using the condition $\lambda_{12}(\sigma | t_2) = \frac{1}{k_1 \sigma}$ in constrain (iii), the inequality becomes

$$\lambda_{12}(t_1 | t_2) = \frac{1}{k_1 t_1}$$

Integrating (2.6) with respect to t_2 , we get

$$\lambda_{21}(t_2 | t_1) \geq \frac{1}{k_2 t_2}$$

Using the condition $\lambda_{21}(\sigma | t_1) = \frac{1}{k_2 \sigma}$, in constrain (iii) the inequality becomes

$$\lambda_{21}(t_2 | t_1) = \frac{1}{k_2 t_2}$$

Hence $\underline{\lambda}(t) = \left(\frac{1}{kt}, \frac{1}{k_1 t_1}, \frac{1}{k_2 t_2} \right)$ and $\lambda_{10}(t) = p_1 \lambda(t), \lambda_{20}(t) = p_2 \lambda(t)$.

The Cox's TFR uniquely determine the distribution through the expression.

$$f(x_1, x_2) = \exp \left\{ - \int_{\sigma}^{x_1} \left(\frac{1}{ku} \right) du - \int_{x_1}^{x_2} \left(\frac{1}{k_2 u} \right) du \right\} \left(\frac{p_1}{kx_1} \right) \left(\frac{1}{k_2 x_2} \right)$$

$$f(x_1, x_2) = \frac{p_1 \sigma^{1/k}}{kk_2} x_1^{-(1/k_1 + 1)} x_2^{-(1/k - 1/k_1 + 1)}; \sigma < x_1 < x_2$$

Similarly for $x_1 > x_2$. Thus

$$f(x_1, x_2) = \begin{cases} \frac{p_1 \sigma^{1/k}}{kk_2} x_2^{-(1/k_2 + 1)} x_1^{-(1/k - 1/k_2 + 1)}; \sigma < x_1 < x_2 \\ \frac{p_2 \sigma^{1/k}}{kk_1} x_1^{-(1/k_1 + 1)} x_2^{-(1/k - 1/k_1 + 1)}; \sigma < x_2 < x_1 \end{cases} \quad (2.7)$$

Writing $\frac{1}{k} = \alpha_1 + \alpha_2, \frac{1}{k_1} = \alpha'_1, \frac{1}{k_2} = \alpha'_2, \frac{p_1}{k} = \alpha_1$ and $\frac{p_2}{k} = \alpha_2$, we get

$$f(x_1, x_2) = \begin{cases} \alpha_1 \alpha'_2 \sigma^{(\alpha_1 + \alpha_2)} x_1^{-(\alpha_1 + \alpha_2 - \alpha'_2 + 1)} x_2^{-(\alpha'_2 + 1)}; \sigma < x_1 < x_2 \\ \alpha_2 \alpha'_1 \sigma^{(\alpha_1 + \alpha_2)} x_2^{-(\alpha_1 + \alpha_2 - \alpha'_1 + 1)} x_1^{-(\alpha'_1 + 1)}; \sigma < x_2 < x_1 \end{cases} \quad (2.8)$$

The model (2.8) has a dependency similar to the Freund's (1961) bivariate exponential distribution. This model is different from all the models

discussed in literature in the sense that, it do not have Pareto marginal but have mixture Pareto marginal. However, this distribution enjoys the bivariate extension of several properties of the univariate Pareto distribution. This model is applicable to a two component parallel system. When one of the components fails, the other bear an extra load and works with a renewed parameter. The survival function of the model (2.8) is given by

$$\bar{F}(x_1, x_2) = \begin{cases} \frac{\alpha_1 \sigma^{(\alpha_1 + \alpha_2)}}{(\alpha_1 + \alpha_2 - \alpha_2')} (x_1)^{-(\alpha_1 + \alpha_2 - \alpha_2')} (x_2)^{-\alpha_2'} + \frac{(\alpha_2 - \alpha_2') \sigma^{(\alpha_1 + \alpha_2)}}{(\alpha_1 + \alpha_2 - \alpha_2')} (x_2)^{-(\alpha_1 + \alpha_2)} ; & \sigma < x_1 < x_2 \\ \frac{\alpha_2 \sigma^{(\alpha_1 + \alpha_2)}}{(\alpha_1 + \alpha_2 - \alpha_1')} (x_2)^{-(\alpha_1 + \alpha_2 - \alpha_1')} (x_1)^{-\alpha_1'} + \frac{(\alpha_1 - \alpha_1') \sigma^{(\alpha_1 + \alpha_2)}}{(\alpha_1 + \alpha_2 - \alpha_1')} (x_1)^{-(\alpha_1 + \alpha_2)} ; & \sigma < x_2 < x_1 \end{cases} \quad (2.9)$$

Properties of the model:

1. Let $\underline{X} = (X_1, X_2)$ be a random vector in the support of $(\sigma, \infty) \times (\sigma, \infty)$ with survival function $\bar{F}(x_1, x_2)$ specified in (2.9), then \underline{X} follows bivariate Pareto I distribution (2.8), if and only if \underline{X} is totally dull at the point $\underline{t} = (t, t)$ (Asha and Jagathnath (2008)).

A distribution of $\underline{X} = (X_1, X_2)$ is called dull at the point $\underline{t} = (t, t)$ whenever

$$P[X_1 \geq s_1 t, X_2 \geq s_2 t \mid X_1 \geq t, X_2 \geq t] = P[X_1 \geq s_1, X_2 \geq s_2] \quad (2.10)$$

for all $s_1, s_2 \geq 1$.

2. A bivariate random variable \underline{X} has density function specified by (2.8) if and only if $\underline{\lambda}(\underline{x})$ is reciprocal linear given by $\underline{\lambda}(\underline{x}) = \left(\frac{(\alpha_1 + \alpha_2)}{x}, \frac{\alpha_1'}{x_1}, \frac{\alpha_2'}{x_2} \right)$

with $\lambda_{10}(x) = \frac{\alpha_1}{x}$ and $\lambda_{20}(x) = \frac{\alpha_2}{x}$ (Asha and Jagathnath (2008)).

The important task after modeling is the estimation of the parameters involved in the model. Now we look into the problem of estimation.

3. Estimation of Parameters from the Density of the Minimum

In reliability and survival analysis, the distribution of minimum of a set of components is of interest to many of the researchers. In this section we considered two methods of estimating parameters from the density of the

minimum. The methods are Principle of Maximum Entropy estimation (POME) and Maximum Likelihood Estimation (MLE).

The probability density function corresponding to the minimum is given by

$$f_x(x) = \frac{\alpha}{\sigma} \left(\frac{x}{\sigma}\right)^{-(\alpha+1)}, x > \sigma, \alpha = \alpha_1 + \alpha_2 \tag{3.1}$$

3.1. Principle of Maximum Entropy Estimation (POME)

In this method, using the given constraints (information) we maximize the entropy function. To obtain an estimate of the parameters, we proceed as follows.

The entropy function for the density function $f_x(x)$ given in (3.1) is obtained by inserting (3.1) in the definition of entropy function given in (1.1).

That is

$$H(f) = -\log\left(\frac{\alpha}{\sigma}\right) \int_{\sigma}^{\infty} f(x) dx + (\alpha + 1) \int_{\sigma}^{\infty} \log\left(\frac{x}{\sigma}\right) f(x) dx \tag{3.2}$$

From (3.2), the constraints are obtained as

$$\int_{\sigma}^{\infty} f(x) dx = 1 \tag{3.3}$$

and

$$\int_{\sigma}^{\infty} \log\left(\frac{x}{\sigma}\right) f(x) dx = E\left[\log\left(\frac{X}{\sigma}\right)\right] \tag{3.4}$$

The first constraint specifies the total probability and the second constraint represents the geometric mean. These constraints are unique and sufficient to explain the model. Using the general expression for the probability density function given in Kapur (1989), we have

$$f(x) = \exp\left(-\lambda_0 - \lambda_1 \log\left(\frac{x}{\sigma}\right)\right) \tag{3.5}$$

where λ_0 and λ_1 are lagrangian multipliers.

By applying conditions (3.3) to (3.5), we obtain

$$\lambda_0 = \log\left(\frac{\sigma}{\lambda_1 - 1}\right) \tag{3.6}$$

Substituting (3.6) in (3.5) yields,

$$f(x) = \frac{\lambda_1 - 1}{\sigma} \left(\frac{x}{\sigma} \right)^{-\lambda_1} \tag{3.7}$$

Comparing (3.7) with (3.1), we get $\lambda_1 = \alpha + 1$.

Taking logarithm of equation (3.7),

$$\log f(x) = \log(\lambda_1 - 1) - \log \sigma - \lambda_1 \log \left(\frac{x}{\sigma} \right) \tag{3.8}$$

The negative expected value of equation (3.8) gives the entropy function

$$H(f) = -\log(\lambda_1 - 1) + \log \sigma + \lambda_1 E \left(\log \left(\frac{X}{\sigma} \right) \right) \tag{3.9}$$

Now the lagrangian multipliers are obtained by taking the partial derivative of (3.9) with respect to λ_1 and equating to zero. Thus we have the equation,

$$E \left[\log \left(\frac{X}{\sigma} \right) \right] = \frac{1}{\lambda_1 - 1} \tag{3.10}$$

Differentiating (3.6) with respect to λ_1 and using (3.10), we get

$$-\frac{\partial \lambda_0}{\partial \lambda_1} = E \left[\log \left(\frac{X}{\sigma} \right) \right] = \frac{1}{\lambda_1 - 1} \tag{3.11}$$

From Tribus (1969), we have

$$\frac{\partial^2 \lambda_0}{\partial \lambda_1^2} = V \left[\log \left(\frac{X}{\sigma} \right) \right] = \frac{1}{(\lambda_1 - 1)^2} \tag{3.12}$$

Hence the parametric estimation of POME consists of two equations (3.11) and (3.12). Inserting $\lambda_1 = \alpha + 1$, (3.11) and (3.12) becomes,

$$E \left[\log \left(\frac{X}{\sigma} \right) \right] = \frac{1}{\alpha} \tag{3.13}$$

$$V \left[\log \left(\frac{X}{\sigma} \right) \right] = \frac{1}{\alpha^2} \tag{3.14}$$

The estimates of the parameters σ and α can be obtained by solving the equations (3.13) and (3.14) numerically.

3.2. Maximum Likelihood Estimation (MLE)

To obtain the maximum likelihood estimate for the parameters σ and α , considering the log-likelihood function.

$$\log L = n \log \alpha + n \alpha \log \sigma - (\alpha + 1) \sum_{i=1}^n \log(x_i)$$

According to Arnold (1983) for a fixed α , the likelihood is maximized when σ is set equal to $Min(X_1, X_2)$. Differentiating with respect to α can then be used to obtain the maximum value of α . Thus we have,

$$\hat{\sigma} = Min(X_1, X_2) \quad \text{and} \quad \hat{\alpha} = \frac{n}{\sum_{i=1}^n \log\left(\frac{X_i}{\hat{\sigma}}\right)}$$

4. Maximum Likelihood Estimation for the Bivariate Pareto I Model.

Estimation of the parameters $\alpha_1, \alpha_2, \alpha'_1, \alpha'_2, \sigma$ by the method of MLE is a tedious job. By considering $\hat{\sigma} = Min(X_1, X_2)$ for some fixed α 's, the problem of estimation becomes easier. In this section we obtain the estimates for the shape parameters by considering the estimate of σ obtained in section 3.

Consider a random sample of size n from a population having bivariate density function given in (2.8). Let n_1 and n_2 be the sample sizes corresponding to $x_1 < x_2$ and $x_1 > x_2$, where X_1 and X_2 are life times of the components. Now the likelihood function of the sample is given by

$$L = \alpha_1^{n_1} (\alpha'_1)^{n_2} \alpha_2^{n_2} (\alpha'_2)^{n_1} \sigma^{n(\alpha_1 + \alpha_2)} \prod_{i=1}^{n_1} (x_{1i})^{-(\alpha_1 + \alpha_2 - \alpha'_2 + 1)} (x_{2i})^{-(\alpha'_2 + 1)} \prod_{i=1}^{n_2} (x_{2i})^{-(\alpha_1 + \alpha_2 - \alpha'_1 + 1)} (x_{1i})^{-(\alpha'_1 + 1)} \tag{4.1}$$

Considering the log-likelihood,

$$\begin{aligned} \log L = & n_1 \log \alpha_1 + n_2 \log(\alpha'_1) + n_2 \log \alpha_2 + n_1 \log(\alpha'_2) + n(\alpha_1 + \alpha_2) \log \sigma \\ & - (\alpha_1 + \alpha_2 - \alpha'_2 + 1) \sum_{i=1}^{n_1} \log(x_{1i}) - (\alpha'_2 + 1) \sum_{i=1}^{n_1} \log(x_{2i}) \\ & - (\alpha_1 + \alpha_2 - \alpha'_1 + 1) \sum_{i=1}^{n_2} \log(x_{2i}) - (\alpha'_1 + 1) \sum_{i=1}^{n_2} \log(x_{1i}) \end{aligned} \tag{4.2}$$

The estimates are obtained by solving the five simultaneous normal equations. Unfortunately, the solutions shows convergence problem. Hence we have to estimate the parameters either by considering some of the parameters to be known or to be estimated from some other methods. The estimate of the scale parameter estimated from the density of the minimum by the method of POME and MLE discussed in section 3 are used to estimate the rest of the parameters. Thus the estimates of the shape parameters are given by

$$\hat{\alpha}_1 = \frac{n_1}{\sum_{i=1}^{n_1} \log(x_{1i}) + \sum_{i=1}^{n_2} \log(x_{2i}) - n \log \hat{\sigma}} \tag{4.3}$$

$$\hat{\alpha}_2 = \frac{n_2}{\sum_{i=1}^{n_1} \log(x_{1i}) + \sum_{i=1}^{n_2} \log(x_{2i}) - n \log \hat{\sigma}} \tag{4.4}$$

$$\hat{\alpha}'_1 = \frac{n_2}{\sum_{i=1}^{n_2} \log(x_{1i}) - \sum_{i=1}^{n_2} \log(x_{2i})} \tag{4.5}$$

$$\hat{\alpha}'_2 = \frac{n_1}{\sum_{i=1}^{n_2} \log(x_{2i}) - \sum_{i=1}^{n_1} \log(x_{1i})} \tag{4.6}$$

Where $\hat{\sigma}$ is the estimate of σ which is obtained by two methods discussed in section 3.

5. Simulation Study

Bivariate Pareto I random samples are generated from Freund (1961) bivariate exponential distribution by using the transformation given in the following theorem.

Theorem 5.1. Let \underline{X} be a random vector with $\frac{X_i}{\sigma} = e^{-rT_i}$ for $\sigma, r > 0$. Then

\underline{X} is distributed as in (2.8) if and only if (T_1, T_2) is distributed as Freund

(1961) bivariate exponential distribution with $\alpha_i = \frac{\beta_i}{r}, \alpha'_i = \frac{\beta'_i}{r}, i = 1, 2$.

For each population, 10000 random samples of sizes 25, 50 and 100 were generated and the parameters are estimated in two phases as discussed in section 3 and section 4. Several combinations of parameters were considered for the simulation study. From the simulation study, it is observed that when

the parameters are having higher values, i.e. for σ greater than 3 and rest of the parameters above 5 is giving comparatively large bias or root mean square error values compared to that of parameters with smaller values. But still the estimates obtained using POME is performing better than that of MLE. For illustration, root mean square error and variance of the estimates for two different sets of parameters are obtained and are given in Table 6.1 and Table 6.2.

6. Results and Discussion

The simulation study has been conducted for different set of parameters. Simulation studies for two sets of parameters are given in Table 6.1 and table 6.2.

Table 6.1: Comparison of Estimation Methods - Simulation Study 1

Sample Size	Parameters	Measures	POME	MLE
n=25	σ (1.8)	RMSE	0.1034	0.2532
		Variance	0.0107	0.0641
	α_1 (1.73)	RMSE	0.2008	0.4411
		Variance	0.0403	0.1946
	α_2 (1.04)	RMSE	0.1887	0.3756
		Variance	0.0356	0.1411
	α_1' (1.38)	RMSE	1.0825	1.0825
		Variance	1.1718	1.1718
	α_2' (0.69)	RMSE	0.3484	0.3484
		Variance	0.1214	0.1214
n=50	σ (1.8)	RMSE	0.0743	0.2253
		Variance	0.0055	0.0508
	α_1 (1.73)	RMSE	0.1135	0.3017
		Variance	0.0129	0.0910
	α_2 (1.04)	RMSE	0.1085	0.2567
		Variance	0.0118	0.0659
	α_1' (1.38)	RMSE	0.6239	0.6239
		Variance	0.3893	0.3893
	α_2' (0.69)	RMSE	0.2259	0.2259
		Variance	0.0510	0.0510
n=100	σ (1.8)	RMSE	0.0535	0.2006
		Variance	0.0029	0.0402
	α_1 (1.73)	RMSE	0.0671	0.2089
		Variance	0.0045	0.0436
	α_2 (1.04)	RMSE	0.0647	0.1774
		Variance	0.0042	0.0315
	α_1' (1.38)	RMSE	0.4032	0.4032
		Variance	0.1626	0.1626
	α_2' (0.69)	RMSE	0.1532	0.1532
		Variance	0.0235	0.0235

Table 6.2: Comparison of Estimation Methods - Simulation Study 2

Sample Size	Parameters	Measures	POME	MLE	
n=25	σ (2)	RMSE	0.2340	0.3364	
		Variance	0.0548	0.1132	
	α_1 (4)	RMSE	0.1484	0.3299	
		Variance	0.0220	0.1088	
	α_2 (3.3)	RMSE	0.1401	0.2829	
		Variance	0.0196	0.0801	
	α_1' (4.6)	RMSE	0.7488	0.7488	
		Variance	0.5608	0.5608	
	α_2' (2.7)	RMSE	0.2464	0.2464	
		Variance	0.0607	0.0607	
	n=50	σ (2)	RMSE	0.1220	0.2940
			Variance	0.0149	0.0865
α_1 (4)		RMSE	0.0831	0.2180	
		Variance	0.0069	0.0475	
α_2 (3.3)		RMSE	0.0784	0.1820	
		Variance	0.0061	0.0331	
α_1' (4.6)		RMSE	0.4441	0.4441	
		Variance	0.1972	0.1972	
α_2' (2.7)		RMSE	0.1612	0.1612	
		Variance	0.0260	0.0260	
n=100		σ (2)	RMSE	0.0800	0.2471
			Variance	0.0064	0.0611
	α_1 (4)	RMSE	0.0486	0.1510	
		Variance	0.0024	0.0228	
	α_2 (3.3)	RMSE	0.0466	0.1269	
		Variance	0.0022	0.0161	
	α_1' (4.6)	RMSE	0.2925	0.2925	
		Variance	0.0855	0.0855	
	α_2' (2.7)	RMSE	0.1104	0.1104	
		Variance	0.0122	0.0122	

From the Table 6.1 and Table 6.2, it is evident that irrespective of the values of the parameters and the sample sizes, the estimates obtained by method of POME for the minimum probability density applied in MLE for a bivariate Pareto distribution performs better than the method of MLE for all the parameter. The RMSE seems to be comparatively smaller for POME than MLE.

From Table 6.1 it can be observed that, for smaller parametric values, as the sample size increases the estimates obtained for scale parameters using POME as a plug in estimator is giving least RMSE, whereas the RMSE for the estimates obtained from MLE is higher. Simulation study is also conducted for larger parameter values. From Table 6.2, we can observe that the estimates obtained for all the parameters are giving least RMSE values as the sample size increases for both the method of estimation. This shows that whatever be the choice of parameter values, higher the sample values the least is the RMSE. From the simulation study carried out, we recommend POME method of estimation for estimating parameters to MLE.

7. Conclusion

In this paper, a bivariate Pareto I distribution has been modeled using the principle of maximum entropy distribution method. The parameters involved in the model are estimated using two phases, the first phase corresponding to the density of the minimum and the second phase by MLE corresponding to bivariate random variable. A simulation study has been conducted to study the performance of these methods. From the simulation study, it can be seen that POME performs better than that of MLE. Here in this study, we considered two method of estimation of the parameters namely, POME and MLE. POME method is applied only for estimating σ and is being used as a plug in estimate in MLE. Instead of that if a bivariate conditional POME method is derived, it may give more efficient estimates than those obtained from the methods discussed in this article.

Acknowledgements

The author is thankful to the CSIR-Central Leather research Institute, Government of India, for the support rendered to carry out the work and acknowledge the CSIR-CLRI Communication No.1028. The author is indebted to the unknown referee for the critical reviews and suggestions which helped in improving the article.

References

1. Arnold B.C. (1983). Pareto distributions, International Co-operative Publishing House, Fairland, MD.
2. Asadi M., Ebrahimi N., Hamedani G.G. and Soofi E.S. (2004). Maximum dynamic entropy models. *Journal of Applied Probability*, 41, 379-390. DOI:10.1239/jap/1082999073.

3. Asha G. and Jagathnath Krishna K.M. (2008). Modeling and characterizations of a bivariate Pareto distribution. *Journal of Statistical Theory and Applications*, 7(4), 435-452.
4. Asha G., Jagathnath Krishna K.M. and Muraleedharan Nair K.R. (2009). Characterization of bivariate distributions using residual measure of uncertainty. *Journal of Probability and Statistical Science*, 7(2), 139-152.
5. Cox D.R. (1972). Regression models and life tables. *Journal of Royal Statistical Society. B*, 34, 187-220.
6. Ebrahimi N. (1996). How to measure uncertainty in residual life time distributions. *Sankhya A*, 58, 48-56.
7. Ebrahimi N. (2000). The maximum entropy method for lifetime distributions. *Sankhya A*, 62, 236-243.
8. Freund J. (1961). A bivariate extension of exponential distribution. *Journal of American Statistical Association*, 56, 971-977. DOI:10.1080/01621459.1961.10482138
9. Kapur J.N. (1989). Maximum entropy models in science and engineering. Wiley Eastern Limited, New Delhi.
10. Shannon C.E. (1948). A mathematical theory of communication. *Bell System Technical Journal*, 27, 379 – 423. DOI:10.1002/j.1538-7305.1948.tb01338.x, DOI:10.1002/j.1538-7305.1948.tb00917.x
11. Tribus M. (1969). Rational Descriptions, Decisions and Designs. Pergamon, New York, USA.