Efficient Estimation of the Parameters of the Pareto Distribution in the Presence of Outliers

Dixit, U. J.^{1,a}, Jabbari Nooghabi, M.^a

^aDepartment of Statistics, University of Mumbai

Abstract

The moment(MM) and least squares(LS) estimations of the parameters are derived for the Pareto distribution in the presence of outliers. Further, we have derived a mixture method(MIX) of estimations with MM and LS that shows that the MIX is more efficient. In the final section we have given an example of actual data from a medical insurance company.

Keywords: Pareto distribution, maximum likelihood estimator, moment estimator, least squares estimation, mixture method, outliers, insurance.

1. Introduction

The Pareto distribution is used as a model for insurance, business, economics, engineering, reliability, hydrology, and mineralogy. These models have been studied by Quandt (1996), Malik (1970), Asrabadi (1990), Hossain and Zimmer (2000), and Nadeau and Teorey (2003). The Pareto distribution was originally used to describe the allocation of wealth among individuals since it seemed to adequately show the way that a larger portion of the wealth in society is owned by a smaller percentage of the people in that society. It can be shown that from a probability density function(pdf), graph of the population f(x), the probability or fraction of f(x) that own a small amount of wealth per person is high. The probability then steadily decreases as wealth increases.

Another application of this distribution is for On-Line Analytical Processing(OLAP) view size estimation. Nadeau and Teorey (2003) used Pareto distribution for OLAP provides useful information quickly from large amounts of data residing in a data warehouse. To improve the quickness of response to queries, pre-aggregation is a useful strategy. However, it is usually impossible to pre-aggregate along all combinations of the dimensions. The multi-dimensional aspects of the data lead to combinatorial explosion in the number and potential storage size of the aggregates. Nadeau and Teorey (2003) suggested to selectively pre-aggregate. Cost/benefit analysis involves estimating the storage requirements of the aggregates in question. They (Nadeau and Teorey, 2003) presented an original algorithm to estimate the number of rows in an aggregate based on the Pareto distribution model. They also tested the Pareto model algorithm empirically against four published algorithms, and concluded that the Pareto model algorithm is consistently the best of these algorithms to estimate view size.

Hossain and Zimmer (2000) compared methods of estimation for the parameters of the Pareto distribution to determine which method provides the better estimates when the observations are censored.

¹ Corresponding author: Professor, Department of Statistics, University of Mumbai, Mumbai, India.

E-mail: Jabbarinm@yahoo.com; Jabbarinm@um.ac.ir

They used unweighted least squares(LS), the maximum likelihood estimate(MLE) and modified likelihood estimate(MML) for censored and uncensored data. They proposed that the LS method be generally preferred over the ML and MML methods to estimate the parameters of the Pareto distribution for complete samples.

Dixit and Jabbari Nooghabi (2011) have considered the Pareto distribution in the presence of outliers when the parameter α is unknown and the parameters β and θ are known. ML and uniformly minimum variance unbiased estimator(UMVUE) of α , the pdf, cumulative distribution function(cdf) and the *r*th moment are derived. These estimators are compared empirically for their mean square error(MSE) and are investigated with the help of numerical technique. They have shown that MLE of pdf and cdf are better than the UMVUEs. In addition, it is shown that the expectation of the MLE of *r*th moment does not exist. Finally, they have illustrated these methods with the help of real data from an insurance company.

Let a set of random variables $(X_1, X_2, ..., X_n)$ represent the claim amounts of a medical cure insurance company. It is assumed that the claims of some of passengers are β times higher than claims of the passengers who have normal injuries.

Hence, we assume that the random variables $(X_1, X_2, ..., X_n)$ are such that any k of them are distributed with pdf

$$f_2(x;\alpha,\beta,\theta) = \frac{\alpha(\beta\theta)^{\alpha}}{x^{\alpha+1}}, \quad 0 < \beta\theta \le x, \ \alpha > 0, \ \beta > 1, \ \theta > 0, \tag{1.1}$$

and the remaining (n - k) random variables are distributed as

$$f_1(x;\alpha,\theta) = \frac{\alpha\theta^{\alpha}}{x^{\alpha+1}}, \quad 0 < \theta \le x, \ \alpha > 0.$$
(1.2)

In this paper, we assume that all three parameters are unknown and we have derived the moment(MM), LS and mixture method(MIX) of MM and LS estimators of the parameters of the Pareto distribution in the presence of outliers. We have shown that the MLE of these parameters does not exist. In addition, it is shown that the MIX estimator of the parameters are more efficient than their MM. Finally, we give an example of claims for a medical insurance company.

2. Joint Distribution of (X_1, X_2, \ldots, X_n) with k Outliers

The joint pdf of $(X_1, X_2, ..., X_n)$ in the presence of k outliers is given by

$$f(x_1, x_2, \dots, x_n; \alpha, \beta, \theta) = \frac{\alpha^n \theta^{n\alpha} \beta^{k\alpha}}{C(n, k)} \left(\prod_{i=1}^n x_i \right)^{-(\alpha+1)} \\ \times \sum_{A_1=1}^{n-k+1} \sum_{A_2=A_1+1}^{n-k+2} \cdots \sum_{A_k=A_{k-1}+1}^n \prod_{j=1}^k \mathbf{I}(x_{A_j} - \beta \theta) \mathbf{I}(x_{A_j} - \theta),$$
(2.1)

where $C(n,k) = n!/\{k!(n-k)!\}$ and **I** is the indicator function defined as

$$\mathbf{I}(y) = \begin{cases} 1, & y > 0, \\ 0, & \text{otherwise.} \end{cases}$$

Note that marginal distribution of X_i is

$$f(x_i;\alpha,\beta,\theta) = b \frac{\alpha(\beta\theta)^{\alpha}}{x_i^{\alpha+1}} \mathbf{I}(x_i - \beta\theta) + \bar{b} \frac{\alpha\theta^{\alpha}}{x_i^{\alpha+1}} \mathbf{I}(x_i - \theta), \quad \alpha > 0, \ \beta > 1, \ \theta > 0,$$
(2.2)

where b = k/n, $\bar{b} = 1 - b$ and (X_1, X_2, \dots, X_n) are not independent (Dixit, 1987, 1989; Dixit and Nasiri, 2001; Dixit and Jabbari Nooghabi, 2011).

3. Method of Moment

Let $X_1, X_2, ..., X_n$ be a random sample of size *n* from the Pareto distribution in the presence of outliers. If we define

$$\begin{pmatrix}
m_1 = E(X), \\
m_2 = E(X)^2, \\
m_3 = E\left(\frac{1}{X}\right), \\
m_4 = E\left(\frac{1}{X}\right)^2.
\end{cases}$$
(3.1)

By using the pdf of X in Equation (2.2), we obtain

$$\begin{cases} m_1 = \left(\frac{\alpha}{\alpha - 1}\right) \theta \left(b\beta + \bar{b}\right), \\ m_2 = \left(\frac{\alpha}{\alpha - 2}\right) \theta^2 \left(b\beta^2 + \bar{b}\right), \\ m_3 = \left(\frac{\alpha}{\alpha + 1}\right) \theta^{-1} \left(b\beta^{-1} + \bar{b}\right), \\ m_4 = \left(\frac{\alpha}{\alpha + 2}\right) \theta^{-2} \left(b\beta^{-2} + \bar{b}\right). \end{cases}$$

$$(3.2)$$

Therefore

$$m_1 m_3 = \frac{\alpha^2}{\alpha^2 - 1} \left(b\beta + \bar{b} \right) \left(b\beta^{-1} + \bar{b} \right), \tag{3.3}$$

and

$$m_2 m_4 = \frac{\alpha^2}{\alpha^2 - 4} \left(b\beta^2 + \bar{b} \right) \left(b\beta^{-2} + \bar{b} \right).$$
(3.4)

Now, if we find α from Equation (3.3) and substitute it in Equation (3.4) and using some elementary algebra, then

$$A\beta^4 + B\beta^3 + C\beta^2 + D\beta + E = 0, (3.5)$$

$$\begin{cases}
A = E = m_1 m_3 b \bar{b}, \\
B = D = -4 m_2 m_4 b \bar{b}, \\
C = 3 m_2 m_4 m_3 m_1 + (m_1 m_3 - 4 m_2 m_4) (b^2 + \bar{b}^2).
\end{cases}$$
(3.6)

According to Abramowitz and Stegun (1970, p.17) and Pachner (1984, p.6.1), we obtain the roots of the quartic equation as

$$\hat{\beta}_{1mm} = \frac{z_1^{\frac{1}{2}} + z_2^{\frac{1}{2}} + z_3^{\frac{1}{2}}}{2} - \frac{B}{4A},$$
(3.7)

$$\hat{\beta}_{2mm} = \frac{z_1^{\frac{2}{2}} - z_2^{\frac{2}{2}} - z_3^{\frac{2}{3}}}{2} - \frac{B}{4A},$$
(3.8)

$$\hat{\beta}_{3mm} = \frac{-z_1^{\frac{1}{2}} + z_2^{\frac{1}{2}} - z_3^{\frac{1}{2}}}{2} - \frac{B}{4A},\tag{3.9}$$

and

$$\hat{\beta}_{4mm} = \frac{-z_1^{\frac{1}{2}} - z_2^{\frac{1}{2}} + z_3^{\frac{1}{2}}}{2} - \frac{B}{4A},$$
(3.10)

where z_1 , z_2 and z_3 are the roots of cubic equation $z^3 + 2pz^2 + (p^2 - 4r)z - q^2 = 0$,

$$p = \frac{C}{A} - 6\left(\frac{B}{4A}\right)^2,$$
$$q = \frac{D}{A} + 2\frac{B}{4A}\left[4\left(\frac{B}{4A}\right)^2 - \frac{C}{A}\right],$$

and

$$r = \frac{E}{A} + \frac{B}{4A} \left\{ \frac{B}{4A} \left[\frac{C}{A} - 3 \left(\frac{B}{4A} \right)^2 \right] - \frac{D}{A} \right\}.$$

Character of the roots of the quartic equation:

Let

$$F = \frac{3AB - 3B^2 + C^2}{12},$$
$$G = \begin{vmatrix} A & \frac{B}{4} & \frac{C}{6} \\ \frac{B}{4} & \frac{C}{6} & \frac{B}{4} \\ \frac{C}{6} & \frac{B}{4} & A \end{vmatrix},$$

and

$$H = \frac{AC}{6} - \frac{B^2}{16}.$$

Then the discriminant is

$$\Delta^* = F^3 - 27G^2.$$

If the two roots are equal, then $\Delta^* = 0$. Excluding this case, any complex roots occur in pairs and the possible cases are:

- (1) If $\Delta^* < 0$, two roots are real and two are complex.
- (2) If $\Delta^* > 0$ and both *H* and 2HF 3AG are negative, all the roots are real.
- (3) If $\Delta^* > 0$ and at least one of two *H* and 2*HF* 3*AG* is positive, then all the roots are complex.

Note: One should note that according to Pachner (1984, p.6.1), if all three roots z_1 , z_2 and z_3 are real and positive, then all four roots $\hat{\beta}_{1mm}$, $\hat{\beta}_{2mm}$, $\hat{\beta}_{3mm}$ and $\hat{\beta}_{4mm}$ are real and we may obtain more than one feasible solution in either case. In such a situation, estimates can be selected by evaluating the likelihood for each feasible solution and choosing the one that maximizes likelihood.

Then, for finding the moment estimator of α , we replace the moment estimator of β in Equation (3.3). So

$$\hat{\alpha}_{mm} = \left[\frac{m_1 m_3}{m_1 m_3 - \left(b\hat{\beta}_{mm} + \bar{b}\right) \left(b\hat{\beta}_{mm}^{-1} + \bar{b}\right)}\right]^{\frac{1}{2}}.$$
(3.11)

Finally, by using the first equation in Equation (3.2), we obtain the moment estimator of θ as

$$\hat{\theta}_{mm} = \frac{m_1 \left(\hat{\alpha}_{mm} - 1\right)}{\hat{\alpha}_{mm} \left(b\hat{\beta}_{mm} + \bar{b}\right)}.$$
(3.12)

4. Maximum Likelihood Estimator

Let $X_1, X_2, ..., X_n$ be a random sample of size *n* from the Pareto distribution in the presence of outliers. So from Equation (2.1)

$$\begin{cases} \tilde{\theta} = \text{MLE}(\theta) = X_{(1)}, & \theta < x_{(1)} < \beta \theta, \\ \tilde{\beta \theta} = \text{MLE}(\beta \theta) = X_{(1)}, & \beta \theta < x_{(1)}, \end{cases}$$
(4.1)

where $X_{(1)}$ is the first order statistic.

Then, with substituting these equations in the likelihood function we find the likelihood equation respect to α as

$$\begin{cases} \frac{n}{\alpha} + n \ln (x_{(1)}) + k \ln(\beta) - \ln(t) = 0, \qquad \theta < x_{(1)} < \beta \theta, \\ \frac{n}{\alpha} + n \ln (x_{(1)}) - (n - k) \ln(\beta) - \ln(t) = 0, \qquad \beta \theta < x_{(1)}, \end{cases}$$
(4.2)

where $T = \prod_{i=1}^{n} X_i$. So we can find the MLE of α and it is given as $\tilde{\alpha}$.

$$\tilde{\alpha} = \begin{cases} \frac{n}{\ln(t) - n \ln(x_{(1)}) - k \ln(\beta)}, & \theta < x_{(1)} < \beta \theta, \\ \frac{n}{\ln(t) - n \ln(x_{(1)}) + (n - k) \ln(\beta)}, & \beta \theta < x_{(1)}. \end{cases}$$
(4.3)

Finally, to find the MLE of β , we replace the above equations in the likelihood function of β . If $\tilde{\theta} = x_{(1)}$, then the likelihood function of β is

$$\ln\left(L\left(\underline{x},\beta\right)\right) \simeq n\ln\left(\frac{n}{\ln(t) - n\ln\left(x_{(1)}\right) - k\ln(\beta)}\right) - n$$

Thus, $\ln(L(\underline{x},\beta))$ is increasing as β increasing.

Also, if $\beta \theta = x_{(1)}$, then the likelihood function of β is

$$\ln\left(L\left(\underline{x},\beta\right)\right) \simeq n \ln\left(\frac{n}{\ln(t) - n \ln(x_{(1)}) + (n-k) \ln(\beta)}\right) - n$$

and $\ln(L(\underline{x},\beta))$ is decreasing as β increasing.

Moreover, if β tends to infinity, then the $\ln(L(\underline{x},\beta))$ tends to $-\infty$. Therefore, the MLE of β does not exist.

Note: One should note that we can obtain the MLE on the boundary of parameter space. When $\tilde{\beta}\theta = x_{(1)}$, the MLE of β is actually 1. As we assume that β is greater than 1, this estimator is not meaningful. When $\tilde{\theta} = x_{(1)}$, the MLE of β is actually $(t/x_{(1)}^n)^{1/k}$. If β is larger than this, the likelihood function cannot be defined.

For example, the plots of $\ln(L(\underline{x},\beta))$ respect to β are given in Figure 1 for n = 5, k = 1, $x_{(1)} = 15$ and t = 95454000.

Hence, for all parameters α , β and θ , the maximum likelihood estimator does not exist.

5. Least Squares Estimator

For the Pareto distribution in the presence of outliers, the reliability function is

$$R(x;\alpha,\beta,\theta) = 1 - F(x;\alpha,\beta,\theta) = \left(\frac{\theta}{x}\right)^{\alpha} \left(b\beta^{\alpha} + \bar{b}\right).$$
(5.1)

So we write

$$y = -\ln(R(x)) = \alpha \ln(x) - \alpha \ln(\theta) - \ln\left(b\beta^{\alpha} + \bar{b}\right).$$
(5.2)

By using an estimate of R(x), we obtain the LS estimator of α as

$$\hat{\alpha}_{ls} = \frac{\sum_{i=1}^{n} y_i \ln(x_i) - n\bar{y}\overline{\ln(x)}}{\sum_{i=1}^{n} \left[\ln(x_i) - \overline{\ln(x)}\right]^2},$$
(5.3)

where $\overline{\ln(x)} = \sum_{i=1}^{n} \ln(x_i)/n$ and $\overline{y} = \sum_{i=1}^{n} y_i/n$.

Here, we can use this method to find the value of $R(x_i)$

$$\hat{F}(X_{(i)}) = 1 - \hat{R}(X_{(i)}), \quad i = 1, 2, \dots, n,$$
(5.4)

$$\hat{R}(X_{(i)}) = \frac{r_i}{r_i + 1} \hat{R}(X_{(i-1)}), \quad \hat{R}(X_{(0)}) = 1,$$
(5.5)



 $\ln(L(\beta | \underline{x}))$ when $\tilde{\theta} = x_{(1)}$ for $n = 5, k = 1, x_{(1)} = 15, t = 95454000$



Figure 1

and r_i is the reverse rank of the *i*th value of x and

$$\hat{F}(X_{(i)}) = \frac{i}{n+1}.$$
 (5.6)

This estimate has the property that $E(\hat{F}(X_{(i)})) = F(X_{(i)})$. For more details, see Hossain and Zimmer (2000).

6. Mixture Method of Moment and Least Squares Estimators

Read (1981) proposed the method that avoids the difficulty of complicated equations. According to Read (1981), the fourth replacement of some (but not all) of the equations in the system may make it

more manageable. From Equation (5.3), we can obtain the estimator of α as

$$\hat{\alpha}_{mix} = \frac{\sum_{i=1}^{n} y_i \ln(x_i) - n\bar{y} \overline{\ln(x)}}{\sum_{i=1}^{n} \left[\ln(x_i) - \overline{\ln(x)} \right]^2}.$$
(6.1)

In addition, from the first and third equations in Equation (3.2), we have

$$\beta = \frac{-\left(\alpha^2 \left(b^2 + \bar{b}^2\right) - m_1 m_3 \left(\alpha^2 - 1\right)\right) + \sqrt{\Delta}}{2b\bar{b}\alpha^2},\tag{6.2}$$

where

$$\Delta = \alpha^4 \left(b^4 + \bar{b}^4 \right) + m_1^2 m_3^2 \left(\alpha^2 - 1 \right)^2 - 2\alpha^2 m_1 m_3 \left(\alpha^2 - 1 \right) \left(b^2 + \bar{b}^2 \right) - 2b^2 \bar{b}^2 \alpha^4.$$

Now, if we substitute Equation (6.1) in Equation (6.2), we can get the mixture estimator of β .

$$\hat{\beta}_{mix} = \frac{-\hat{\alpha}_{mix}^2 \left(b^2 + \bar{b}^2 \right) + m_1 m_3 \left(\hat{\alpha}_{mix}^2 - 1 \right) + \sqrt{\Delta_1}}{2 b \bar{b} \hat{\alpha}_{mix}^2},\tag{6.3}$$

where

$$\Delta_1 = \hat{\alpha}_{mix}^4 \left(b^4 + \bar{b}^4 \right) + m_1^2 m_3^2 \left(\hat{\alpha}_{mix}^2 - 1 \right)^2 - 2\hat{\alpha}_{mix}^2 m_1 m_3 \left(\hat{\alpha}_{mix}^2 - 1 \right) \left(b^2 + \bar{b}^2 \right) - 2b^2 \bar{b}^2 \hat{\alpha}_{mix}^4.$$
(6.4)

Finally, by using the first equation in Equation (3.2), we find the mixture estimator of θ .

$$\hat{\theta}_{mix} = \frac{m_1 \left(\hat{\alpha}_{mix} - 1\right)}{\hat{\alpha}_{mix} \left(b\hat{\beta}_{mix} + \bar{b}\right)}.$$
(6.5)

6.1. Asymptotically unbiasedness of the estimators

Now, we shall show that $\hat{\alpha}_{mix}$, $\hat{\beta}_{mix}$ and $\hat{\theta}_{mix}$ are asymptotically unbiased estimators. We have

 $\hat{\alpha}_{mix} = \hat{\alpha}_{ls}.$

Therefore, according to the least squares estimation

$$E\left(\hat{\alpha}_{mix}\right) = \alpha. \tag{6.6}$$

Let $T_1 = 1/n \sum_{i=1}^n X_i$, $T_2 = \hat{\alpha}_{mix}$ and $T_3 = 1/n \sum_{i=1}^n 1/X_i$. For the mixture estimator of β , we can write $\hat{\beta}_{mix}$ as a function of T_1 , T_2 and T_3 . Hence,

$$\hat{\beta}_{mix} = f(T_1, T_2, T_3) = \frac{-T_2^2 \left(b^2 + \bar{b}^2\right) + T_1 T_3 \left(T_2^2 - 1\right) + \sqrt{\Delta_1}}{2b\bar{b}T_2^2},$$
(6.7)

$$\Delta_1 = T_2^4 \left(b^4 + \bar{b}^4 \right) + T_1^2 T_3^2 \left(T_2^2 - 1 \right)^2 - 2T_2^2 T_1 T_3 \left(T_2^2 - 1 \right) \left(b^2 + \bar{b}^2 \right) - 2b^2 \bar{b}^2 T_2^4.$$

Let $\lambda_1 = E(T_1)$, $\lambda_2 = E(T_2)$ and $\lambda_3 = E(T_3)$. Expand the function $f(T_1, T_2, T_3)$ around $(\lambda_1, \lambda_2, \lambda_3)$ by Taylor series as

$$\hat{\beta}_{mix} = f(\lambda_1, \lambda_2, \lambda_3) + (T_1 - \lambda_1) \frac{df}{dT_1} \Big|_{T_1 = \lambda_1, T_2 = \lambda_2, T_3 = \lambda_3} + (T_2 - \lambda_2) \frac{df}{dT_2} \Big|_{T_1 = \lambda_1, T_2 = \lambda_2, T_3 = \lambda_3} + (T_3 - \lambda_3) \frac{df}{dT_3} \Big|_{T_1 = \lambda_1, T_2 = \lambda_2, T_3 = \lambda_3} + O(n^{-1}).$$

$$(6.8)$$

Hence, from Equations (6.7) and (6.8)

$$E\left(\hat{\beta}_{mix}\right) \simeq f\left(\lambda_{1}, \lambda_{2}, \lambda_{3}\right)$$

$$= \frac{-\lambda_{2}^{2}\left(b^{2} + \bar{b}^{2}\right) + \lambda_{1}\lambda_{3}\left(\lambda_{2}^{2} - 1\right) + \sqrt{\Delta_{2}}}{2b\bar{b}\lambda_{2}^{2}},$$
(6.9)

where

$$\Delta_2 = \Delta_1|_{T_1 = \lambda_1, T_2 = \lambda_2, T_3 = \lambda_3} = \lambda_2^4 \left(b^4 + \bar{b}^4 \right) + \lambda_1^2 \lambda_3^2 \left(\lambda_2^2 - 1 \right)^2 - 2\lambda_2^2 \lambda_1 \lambda_3 \left(\lambda_2^2 - 1 \right) \left(b^2 + \bar{b}^2 \right) - 2b^2 \bar{b}^2 \lambda_2^4.$$

From Equations (3.2) and (6.6), We have $\lambda_1 = \alpha \theta (b\beta + \bar{b})/(\alpha - 1)$, $\lambda_2 = \alpha$ and $\lambda_3 = \alpha \theta^{-1} (b\beta^{-1} + \bar{b})/(\alpha + 1)$. Therefore

$$\begin{split} \Delta_{2} &= \alpha^{4} \left(b^{4} + \bar{b}^{4} \right) + \left(\frac{\alpha \theta \left(b\beta + \bar{b} \right)}{\alpha - 1} \right)^{2} \left(\frac{\alpha \theta^{-1} \left(b\beta^{-1} + \bar{b} \right)}{\alpha + 1} \right)^{2} \left(\alpha^{2} - 1 \right)^{2} \\ &- 2\alpha^{2} \left(\frac{\alpha \theta \left(b\beta + \bar{b} \right)}{\alpha - 1} \right) \left(\frac{\alpha \theta^{-1} \left(b\beta^{-1} + \bar{b} \right)}{\alpha + 1} \right) \left(\alpha^{2} - 1 \right) \left(b^{2} + \bar{b}^{2} \right) - 2b^{2}\bar{b}^{2}\alpha^{4} \\ &= \frac{\alpha^{4}}{\beta^{2}} \left\{ \beta^{2} \left(b^{4} - 2b^{2}\bar{b}^{2} + \bar{b}^{4} \right) + \left(b^{2}\beta + b\bar{b}\beta^{2} + b\bar{b} + \bar{b}^{2}\beta \right) \left[b^{2}\beta + b\bar{b}\beta^{2} + b\bar{b} + \bar{b}^{2}\beta - 2 \left(b^{2} + \bar{b}^{2} \right) \beta \right] \right\}. \end{split}$$

So by using some elementary algebra

$$\Delta_2 = \frac{\alpha^4 b^2 \bar{b}^2}{\beta^2} \left(\beta^2 - 1\right)^2.$$

Then

$$\begin{split} \mathbf{E}\left(\hat{\beta}_{mix}\right) &\simeq \frac{-\alpha^2 \left(b^2 + \bar{b}^2\right) + \left(\frac{\alpha\theta(b\beta + \bar{b})}{\alpha - 1}\right) \left(\frac{\alpha \theta^{-1}(b\beta^{-1} + \bar{b})}{\alpha + 1}\right) \left(\alpha^2 - 1\right) + \sqrt{\frac{\alpha^4 b^2 \bar{b}^2}{\beta^2}} \left(\beta^2 - 1\right)^2}{2b \bar{b} \alpha^2} \\ &= \frac{\alpha^2 b \bar{b} \left(\beta + \beta^{-1}\right) + \frac{\alpha^2 b \bar{b}}{\beta} \left(\beta^2 - 1\right)}{2b \bar{b} \alpha^2}, \end{split}$$

and

$$E(\hat{\beta}_{mix}) \simeq \beta, \quad \text{as } n \to \infty \& k \to \infty.$$
 (6.10)

Now, for mixture estimator of θ , we have

$$\hat{\theta}_{mix} = g\left(T_1, T_2, T_4\right) = \frac{T_1\left(T_2 - 1\right)}{T_2\left(bT_4 + \bar{b}\right)},\tag{6.11}$$

where $T_1 = 1/n \sum_{i=1}^n X_i$, $T_2 = \hat{\alpha}_{mix}$ and $T_4 = \hat{\beta}_{mix}$. By using Taylor series

$$\hat{\theta}_{mix} = g(\lambda_1, \lambda_2, \lambda_4) + (T_1 - \lambda_1) \frac{dg}{dT_1} \Big|_{T_1 = \lambda_1, T_2 = \lambda_2, T_4 = \lambda_4} + (T_2 - \lambda_2) \frac{dg}{dT_2} \Big|_{T_1 = \lambda_1, T_2 = \lambda_2, T_4 = \lambda_4} + (T_4 - \lambda_4) \frac{dg}{dT_4} \Big|_{T_1 = \lambda_1, T_2 = \lambda_2, T_4 = \lambda_4} + O(n^{-1}),$$
(6.12)

where $\lambda_1 = E(T_1)$, $\lambda_2 = E(T_2)$ and $\lambda_4 = E(T_4)$. Then, according to the previous procedure

$$E\left(\hat{\theta}_{mix}\right) \simeq g\left(\lambda_{1}, \lambda_{2}, \lambda_{4}\right)$$
$$= \lim_{n,k \to \infty} \frac{\frac{\alpha \theta}{\alpha - 1} \left(b\beta + \bar{b}\right)(\alpha - 1)}{\alpha \left(b\beta + \bar{b}\right)},$$

and

$$\mathrm{E}(\hat{\theta}_{mix}) \simeq \theta$$
, as $n \to \infty \& k \to \infty$. (6.13)

6.2. Asymptotic distribution of the estimators

In this subsection, we wish to show that the estimators are consistent. From the strong law of large numbers, it is known that $T_1 = 1/n \sum_{i=1}^n X_i$, $T_2 = \hat{\alpha}_{mix}$ and $T_3 = 1/n \sum_{i=1}^n 1/X_i$ converge almost surely (as) to E(X), α and E(1/X), respectively. Also, from the central limit theorem and Lehmann and Casella (1998, Theorem 8.21 in p.61), we readily observe that T_1 , T_2 and T_3 are asymptotically normally distributed.

So

$$\sqrt{n} \begin{pmatrix} T_1 - \mathcal{E}(X) \\ T_2 - \mathcal{E}(\hat{\alpha}_{mix}) \\ T_3 - \mathcal{E}\left(\frac{1}{X}\right) \end{pmatrix} \sim N\left(\underline{0}, \Sigma\right), \tag{6.14}$$

where

$$\underline{0} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, \qquad \Sigma = \begin{pmatrix} \sigma_{11} & \sigma_{12} & \sigma_{13} \\ \sigma_{21} & \sigma_{22} & \sigma_{23} \\ \sigma_{31} & \sigma_{32} & \sigma_{33} \end{pmatrix},$$

and

$$\sigma_{11} = \operatorname{Var}(T_1), \qquad \sigma_{12} = \sigma_{21} = \operatorname{Cov}(T_1, T_2), \qquad \sigma_{13} = \sigma_{31} = \operatorname{Cov}(T_1, T_3), \\ \sigma_{22} = \operatorname{Var}(T_2), \qquad \sigma_{23} = \sigma_{32} = \operatorname{Cov}(T_2, T_3), \qquad \sigma_{33} = \operatorname{Var}(T_3).$$

Now, we should find the covariance matrix. We have

$$\sigma_{11} = V(T_1) = E(T_1)^2 - (E(T_1))^2$$

826

Then

$$E(T_{1})^{2} = \frac{1}{n^{2}} \left[\sum_{i=1}^{n} E(X_{i}^{2}) + \sum_{i\neq j}^{n} E(X_{i}X_{j}) \right]$$
$$= \frac{1}{n} E(X^{2}) + \frac{n-1}{n} E(X_{i}X_{j}), \quad i \neq j.$$
(6.15)

In addition, we have

$$\mathrm{E}\left(X^{2}\right) = \frac{\alpha}{\alpha - 2}\theta^{2}\left(b\beta^{2} + \bar{b}\right),$$

and we know that from Equations (1.1) and (1.2)

$$h(x_i, x_j) = \frac{1}{C(n,k)} \Big[C(n-2, k-2) f_2(x_i) f_2(x_j) + C(n-2, k-1) \Big\{ f_2(x_i) f_1(x_j) + f_1(x_i) f_2(x_j) \Big\} \\ + C(n-2, k) f_1(x_i) f_1(x_j) \Big], \quad i \neq j.$$
(6.16)

So

$$\mathbf{E}(X_i X_j) = \frac{\alpha^2 \theta^2}{(\alpha - 1)^2 n(n-1)} \left[k(k-1)\beta^2 + 2k(n-k)\beta + (n-k)(n-k-1) \right], \quad i \neq j,$$
(6.17)

and by using some elementary algebra

$$\sigma_{11} = \frac{\alpha \theta^2 \left(b\beta^2 + \bar{b} \right)}{n(\alpha - 1)^2 (\alpha - 2)}.$$
(6.18)

By using Taylor series

$$\begin{split} \sigma_{12} &= \mathrm{E}(T_1T_2) - \mathrm{E}(T_1)\mathrm{E}(T_2) \\ &\simeq \alpha \mathrm{E}(T_1) - \alpha \mathrm{E}(T_1), \end{split}$$

and

$$\sigma_{12} \simeq 0. \tag{6.19}$$

Similarly

$$\sigma_{23} \simeq 0. \tag{6.20}$$

Now, we use the same method as σ_{11} to obtain σ_{13} .

$$\sigma_{13} = \text{Cov}(T_1, T_3) = \text{E}(T_1T_3) - \text{E}(T_1)\text{E}(T_3),$$

and

$$\mathcal{E}(T_1T_3) = \frac{1}{n} + \frac{n-1}{n} \mathcal{E}\left(X_i \frac{1}{X_j}\right), \quad i \neq j.$$

Therefore, by using Equation (6.16)

$$E\left(X_{i}\frac{1}{X_{j}}\right) = \frac{\alpha^{2}}{n(n-1)(\alpha^{2}-1)}\left[k(k-1) + k(n-k)\left(\beta + \beta^{-1}\right) + (n-k)(n-k-1)\right], \quad i \neq j, \quad (6.21)$$

and after simplification

$$\sigma_{13} = -\frac{1}{n(\alpha^2 - 1)}.$$
(6.22)

From the least squares estimation, we have

$$\sigma_{22} = \operatorname{Var}\left(\hat{\alpha}_{mix}\right) = \frac{\sigma^2}{\sum_{i=1}^n \left[\ln(x_i) - \overline{\ln(x)}\right]^2},\tag{6.23}$$

and an unbiased estimator of σ^2 is

$$\tilde{\sigma}^2 = \frac{\sum_{i=1}^n (y_i - \bar{y})^2}{n-2} \left(1 - r^2\right),\tag{6.24}$$

where *r* is Pearson linear correlation coefficient between *y* and ln(x).

To obtain σ_{33} similarly as the previous, we have

$$\sigma_{33} = V(T_3) = \mathrm{E}(T_3)^2 - (\mathrm{E}(T_3))^2.$$

So

$$E(T_3)^2 = \frac{1}{n}E\left(\frac{1}{X^2}\right) + \frac{n-1}{n}E\left(\frac{1}{X_i}\frac{1}{X_j}\right), \quad i \neq j,$$
(6.25)

and

$$\mathbf{E}(X_i X_j)^{-1} = \frac{\alpha^2 \theta^{-2}}{n(n-1)(\alpha+1)^2} \left[k(k-1)\beta^{-2} + 2k(n-k)\beta^{-1} + (n-k)(n-k-1) \right], \quad i \neq j.$$
(6.26)

Therefore, by using some algebra

$$\sigma_{33} = \frac{\alpha \left(b + \bar{b}\beta^2\right)}{n\beta^2\theta^2(\alpha+1)^2(\alpha+2)}.$$
(6.27)

Now, we have

$$\hat{\alpha}_{mix} = f_1(T_1, T_2, T_3) = T_2, \tag{6.28}$$

$$\hat{\beta}_{mix} = f_2(T_1, T_2, T_3) = \frac{-T_2^2 \left(b^2 + \bar{b}^2\right) + T_1 T_3 \left(T_2^2 - 1\right) + \sqrt{\Delta_1}}{2b\bar{b}T_2^2},$$
(6.29)

$$\Delta_1 = T_2^4 \left(b^4 + \bar{b}^4 \right) + T_1^2 T_3^2 \left(T_2^2 - 1 \right)^2 - 2T_2^2 T_1 T_3 \left(T_2^2 - 1 \right) \left(b^2 + \bar{b}^2 \right) - 2b^2 \bar{b}^2 T_2^4,$$

and

$$\hat{\theta}_{mix} = f_3(T_1, T_2, T_3) = \frac{T_1(T_2 - 1)}{T_2 \left(b f_2(T_1, T_2, T_3) + \bar{b} \right)}.$$
(6.30)

So by using Taylor series expansion and Lehmann and Casella (1998, Theorem 8.22 in p.61), we obtain

$$\sqrt{n} \begin{pmatrix} \hat{\alpha}_{mix} - \alpha \\ \hat{\beta}_{mix} - \beta \\ \hat{\theta}_{mix} - \theta \end{pmatrix} \sim N(\underline{0}, \Sigma_1), \qquad (6.31)$$

where $\Sigma_1 = B\Sigma B^T (B^T \text{ is transpose of } B)$ and

$$B = \begin{pmatrix} \frac{df_1}{dT_1} & \frac{df_1}{dT_2} & \frac{df_1}{dT_3} \\ \frac{df_2}{dT_1} & \frac{df_2}{dT_2} & \frac{df_2}{dT_3} \\ \frac{df_3}{dT_1} & \frac{df_3}{dT_2} & \frac{df_3}{dT_3} \end{pmatrix}_{T_1 = \mathbb{E}(X), T_2 = \mathbb{E}(\hat{\alpha}_{mix}), T_3 = \mathbb{E}(\frac{1}{X})}$$
(6.32)

From Equation (6.28), we have

$$\left. \frac{df_1}{dT_1} \right|_{T_1 = \mathsf{E}(X), T_2 = \mathsf{E}(\hat{\alpha}_{mix}), T_3 = \mathsf{E}(\frac{1}{X})} = 0, \tag{6.33}$$

$$\left. \frac{df_1}{dT_2} \right|_{T_1 = E(X), T_2 = E(\hat{\alpha}_{mix}), T_3 = E\left(\frac{1}{X}\right)} = 1, \tag{6.34}$$

and

$$\frac{df_1}{dT_3}\Big|_{T_1 = \mathbb{E}(X), T_2 = \mathbb{E}(\hat{\alpha}_{mix}), T_3 = \mathbb{E}(\frac{1}{X})} = 0.$$
(6.35)

By using Equation (6.29), we obtain

$$\frac{df_2}{dT_1} = \frac{1}{2b\bar{b}T_2^2} \left[T_3 \left(T_2^2 - 1 \right) + \frac{1}{2} \Delta_1^{-\frac{1}{2}} \frac{d\Delta_1}{dT_1} \right],$$

where

$$\frac{d\Delta_1}{dT_1} = 2T_1T_3^2 \left(T_2^2 - 1\right)^2 - 2T_2^2T_3 \left(T_2^2 - 1\right) \left(b^2 + \bar{b}^2\right).$$

Thus, with substituting $T_1 = E(X)$, $T_2 = \alpha$ and $T_3 = E(1/X)$ and using some elementary algebra, we find

$$\frac{df_2}{dT_1}\Big|_{T_1 = \mathrm{E}(X), T_2 = \mathrm{E}(\hat{\alpha}_{mix}), T_3 = \mathrm{E}(\frac{1}{X})} = \frac{\beta(\alpha - 1)\left(b + \bar{b}\beta\right)}{b\bar{b}\alpha\theta\left(\beta^2 - 1\right)}.$$
(6.36)

In addition, we have

$$\frac{df_2}{dT_2} = \frac{T_1 T_3 T_2^{-3}}{b\bar{b}} - \frac{T_2^{-3}}{b\bar{b}} \Delta_1^{\frac{1}{2}} + \frac{T_2^{-2}}{4b\bar{b}} \Delta_1^{-\frac{1}{2}} \frac{d\Delta_1}{dT_2},$$
(6.37)

where

$$\frac{d\Delta_1}{dT_2} = 4T_2^3 \left(b^4 + \bar{b}^4\right) + 4T_1^2 T_3^2 \left(T_2^2 - T_2\right) - 4T_1 T_3 T_2 \left(2T_2^2 - 1\right) \left(b^2 + \bar{b}^2\right) - 8b^2 \bar{b}^2 T_2^3.$$

Therefore, after simplification, we obtain

$$\frac{df_2}{dT_2}\Big|_{T_1 = \mathbb{E}(X), T_2 = \mathbb{E}(\hat{\alpha}_{mix}), T_3 = \mathbb{E}(\frac{1}{X})} = \frac{2\beta \left(b\beta + \bar{b}\right) \left(b + \bar{b}\beta\right)}{b\bar{b}\alpha \left(\alpha^2 - 1\right) \left(\beta^2 - 1\right)}.$$
(6.38)

Further,

$$\frac{df_2}{dT_3} = \frac{T_1 T_2^{-2} \left(T_2^2 - 1\right)}{2b\bar{b}} + \frac{T_2^{-2}}{4b\bar{b}} \Delta_1^{-\frac{1}{2}} \frac{d\Delta_1}{dT_3},\tag{6.39}$$

where

$$\frac{d\Delta_1}{dT_3} = 2T_1^2 \left(T_2^2 - 1\right)^2 T_3 - 2T_1 T_2^2 \left(T_2^2 - 1\right) \left(b^2 + \bar{b}^2\right).$$

Then, by using some elementary algebra

$$\frac{df_2}{dT_3}\Big|_{T_1=\mathrm{E}(X),T_2=\mathrm{E}(\hat{\alpha}_{mix}),T_3=\mathrm{E}(\frac{1}{X})} = \frac{(\alpha+1)\left(b\beta+\bar{b}\right)\beta^2\theta}{b\bar{b}\alpha\left(\beta^2-1\right)}.$$
(6.40)

From Equation (6.30), we find

$$\frac{df_3}{dT_1} = \frac{T_2 \left(bf_2(T_1, T_2, T_3) + \bar{b} \right) \frac{d}{dT_1} \left[T_1(T_2 - 1) \right] - T_1(T_2 - 1) \frac{d}{dT_1} \left[T_2 \left(bf_2(T_1, T_2, T_3) + \bar{b} \right) \right]}{T_2^2 \left(bf_2(T_1, T_2, T_3) + \bar{b} \right)^2},$$

and

$$\frac{d}{dT_1} \left[T_2 \left(b f_2(T_1, T_2, T_3) + \bar{b} \right) \right] = b T_2 \frac{d}{dT_1} f_2(T_1, T_2, T_3).$$

Therefore, with replacing $T_1 = E(X)$, $T_2 = \alpha$ and $T_3 = E(1/X)$ and using some algebra, we get

$$\frac{df_3}{dT_1}\Big|_{T_1 = E(X), T_2 = E(\hat{\alpha}_{mix}), T_3 = E(\frac{1}{X})} = -\frac{\alpha - 1}{\bar{b}\alpha \left(\beta^2 - 1\right)}.$$
(6.41)

In addition, we have

$$\frac{df_3}{dT_2} = \frac{T_2 \left(bf_2(T_1, T_2, T_3) + \bar{b} \right) \frac{d}{dT_2} \left[T_1(T_2 - 1) \right] - T_1(T_2 - 1) \frac{d}{dT_2} \left[T_2 \left(bf_2(T_1, T_2, T_3) + \bar{b} \right) \right]}{T_2^2 \left(bf_2(T_1, T_2, T_3) + \bar{b} \right)^2}, \quad (6.42)$$

830

and

$$\frac{d}{dT_2} \left[T_2 \left(b f_2(T_1, T_2, T_3) + \bar{b} \right) \right] = b f_2(T_1, T_2, T_3) + \bar{b} + b T_2 \frac{d}{dT_2} f_2(T_1, T_2, T_3).$$

Therefore, after substitution and simplification, we obtain

$$\frac{df_3}{dT_2}\Big|_{T_1=\mathrm{E}(X),T_2=\mathrm{E}(\hat{\alpha}_{mix}),T_3=\mathrm{E}(\frac{1}{X})} = \frac{\left[\bar{b}\alpha\left(\beta^2-1\right)-\bar{b}\left(\beta^2+1\right)-2b\beta\right]\theta}{\bar{b}\alpha\left(\alpha^2-1\right)\left(\beta^2-1\right)}.$$
(6.43)

Finally, we have

$$\frac{df_3}{dT_3} = \frac{T_2 \left(bf_2(T_1, T_2, T_3) + \bar{b} \right) \frac{d}{dT_3} [T_1(T_2 - 1)] - T_1(T_2 - 1) \frac{d}{dT_3} \left[T_2 \left(bf_2(T_1, T_2, T_3) + \bar{b} \right) \right]}{T_2^2 \left(bf_2(T_1, T_2, T_3) + \bar{b} \right)^2}, \quad (6.44)$$

and

$$\frac{d}{dT_3} \left[T_2 \left(b f_2(T_1, T_2, T_3) + \bar{b} \right) \right] = b T_2 \frac{d}{dT_3} f_2(T_1, T_2, T_3).$$

Then, by using some elementary algebra, we find

$$\frac{df_3}{dT_3}\Big|_{T_1 = \mathcal{E}(X), T_2 = \mathcal{E}(\hat{\alpha}_{mix}), T_3 = \mathcal{E}(\frac{1}{\bar{X}})} = -\frac{(\alpha+1)\beta^2\theta^2}{\bar{b}\alpha(\beta^2-1)}.$$
(6.45)

So by substituting Equations (6.33), (6.34), (6.35), (6.36), (6.38), (6.40), (6.41), (6.43) and (6.45) into Equation (6.32), we obtain the matrix *B* and then the matrix Σ_1 . In addition, the determinant of matrix Σ_1 , after using some elementary algebra and simplification is

$$\det(\Sigma_1) = \frac{\left[b\bar{b}\alpha^2\left(\beta^4 + 1\right) + \alpha^2\beta^2\left(b^2 + \bar{b}^2 - 1\right) + 4\beta^2\right]\beta^2\theta^2\sigma^2}{n^2b^2\bar{b}^2\alpha^4\left(\alpha^2 - 4\right)\left(\beta^2 - 1\right)^2\left[\sum_{i=1}^n\left(\ln(x_i)\right)^2 - n\left(\overline{\ln(x)}\right)^2\right]}.$$
(6.46)

It is appear that all elements of the matrix *B* are fixed when $n \to \infty$ and $k \to \infty$. In addition, we can see that all elements of the matrix Σ tend to zero as $n \to \infty$ and $k \to \infty$. Then, the variance and covariance of the mixture estimators of the parameters tend to zero when $n \to \infty$ and $k \to \infty$.

Further, determinant of the matrix Σ_1 tends to zero as *n* and *k* tend to ∞ .

Therefore, the mixture estimator of the parameters are asymptotically consistent and normally distributed.

7. Comparison of Moment and Mixture Methods and an Example

In order to have some idea about efficiency, we perform sampling experiments using R software. Our model has three parameters and we have calculated covariance matrix of the estimate α , β and θ . Determinant of the covariance matrix is

$$det = \operatorname{Var}\left(\hat{\alpha}\right) \left[\operatorname{Var}\left(\hat{\beta}\right) \operatorname{Var}\left(\hat{\theta}\right) - \left(\operatorname{Cov}\left(\hat{\beta},\hat{\theta}\right)\right)^{2}\right] + \operatorname{Var}\left(\hat{\beta}\right) \left[\operatorname{Var}\left(\hat{\alpha}\right) \operatorname{Var}\left(\hat{\theta}\right) - \left(\operatorname{Cov}\left(\hat{\alpha},\hat{\theta}\right)\right)^{2}\right] + \operatorname{Var}\left(\hat{\theta}\right) \left[\operatorname{Var}\left(\hat{\alpha}\right) \operatorname{Var}\left(\hat{\beta}\right) - \left(\operatorname{Cov}\left(\hat{\alpha},\hat{\beta}\right)\right)^{2}\right].$$
(7.1)



Figure 2

We have generated a sample of size n = 6, 7, ..., 30 from the Pareto distribution with k = 1, 2, 3, $\alpha = 5, 10, \beta = 1.5, 2, 3$ and $\theta = 0.5, 5, 10$. We have given graphs based on one thousand independent replications of each experiment (Figure 2, 3).

The graphs show that the mixture estimation it is more efficient than the moment estimation.





Determinant when k=2, α =10, β =2 and θ =0.5

Determinant when k=2, α =10, β =3 and θ =10



Figure 3

Example 1. In an insurance company, one of services is medical insurance. Any claim may be made by a passenger for compensating medical problems received in an accident. This compensation is to be at least θ Rials since less than θ is not reasonable for them to claim for the problem. Most of the passenger claims are almost near to θ Rials that represents a modal value of their claims; however,

k	$\hat{\alpha}_{mix}$	\hat{eta}_{mix}	$\hat{\theta}_{mix}$
1	3.1278	2.7138	84337.2
2	3.1278	2.8271	78625.8
3	3.1278	2.7924	74166.3

Table 1: The mixture estimators of the parameters for k = 1, 2, 3

Table 2: The likelihood function corresponding to	b k, for $k = 1, 2, 3$
k	$\ln L(\underline{x}; \hat{\alpha}_{mix}, \hat{\beta}_{mix}, \hat{\theta}_{mix})$
1	-294.6088
2	-296.7135
3	-298.1447

the claims depend on factors such as type and nature of injury. The claim amounts vary according to these factors. Therefore, it is observed that the passenger claims are β times higher than passengers whose claims are near to the modal value. It is clear that the claims follow the Pareto distribution in the presence of outliers with parameters α , β and θ , where α , β and θ are unknown and the number of outliers (*k*) is unknown.

The claims data from the Iran Insurance company records for the year 2009 are given by drawing a random sample of size 25 of the claim amounts.

280870,	110147,	100483,	108729,	142800
102108,	107852,	163073,	118722,	108948
117307,	180237,	115422,	123086,	113936
221617,	112211,	106790,	178104,	101561
104325,	110343,	112843,	131537,	138744

In the example, we know that the number of outliers k is unknown; subsequently, we use the following method to estimate it. We calculate the mixture estimator for the parameters α , β and θ for different values of k. Then, k can be selected by evaluating the likelihood function respect to k for different values of k choosing the one that maximizes the corresponding likelihood.

The mixture estimators of the parameters for k = 1, 2, 3 are shown in Table 1. In addition, from the likelihood function corresponding to k, $\ln L(\underline{x}; \hat{\alpha}_{mix}, \hat{\beta}_{mix}, \hat{\theta}_{mix})$ for k = 1, 2, 3 are shown in Table 2. Finally, we see that the likelihood function respect to k is maximized for k = 1.

Acknowledgments

We are thankful to the two referees for their valuable comments.

References

Abramowitz, M. and Stegun, A. I. (1970). *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*, Dover Publications, New York.

Asrabadi, B. R. (1990). Estimation in the Pareto distribution, Metrika, 37, 199–205.

- Dixit, U. J. (1987). Characterization of the gamma distribution in the presence of *k* outliers, *Bulletin Bombay Mathematical Colloquium*, **4**, 54–59.
- Dixit, U. J. (1989). Estimation of parameters of the Gamma Distribution in the presence of Outliers, *Communications in Statistics Theory and Methods*, **18**, 3071–3085.

- Dixit, U. J. and Jabbari Nooghabi, M. (2011). Efficient estimation in the Pareto distribution with the presence of outliers, *Statistical Methodology*, **8**, 340–355.
- Dixit, U. J. and Nasiri, F. P. (2001). Estimation of parameters of the exponential distribution in the presence of outliers generated from uniform distribution, *Metron* LIX(3–4), 187–198.
- Hossain, A. M. and Zimmer, W. J. (2000). Comparisons of methods of estimation for a Pareto distribution of the first kind, *Communications in Statistics - Theory and Methods*, 29, 859–878.
- Lehmann, E. L. and Casella, G. (1998). *Theory of Point Estimation*, John Wiley & Sons, Second edition, New York.
- Malik, H. J. (1970). Estimation of the parameter of the Pareto distribution, *Metrika*, 15, 126–132.
- Nadeau, T. P. and Teorey, T. J. (2003). A Pareto model for OLAP view size estimation, *Information Systems Frontiers*, 5, 137–147.
- Pachner, J. (1984). Handbook of Numerical Analysis Applications with Programs for Engineers and Scientists, McGraw-Hill Inc., New York.
- Quandt, R. E. (1996). Old and new methods of estimation and the Pareto distribution, *Metrika*, **10**, 55–82.
- Read, R. R. (1981). Representation of certain covariance matrices with application to asymptotic efficiency, *Journal of the American Statistical Association*, **76**, 148–154.

Received March 2011; Accepted September 2011