Quality & Quantity (2007) 41:913–926
DOI 10.1007/s11135-007-9100-8

The Estimation of Pareto Distribution by a Weighted Least Square Method

HAI-LIN LU[1] and SHIN-HWA TAO[2,*]
[1]*Department of Management Information Science, Chia-Nan University of Pharmacy and Science, Tainan, Taiwan, ROC;* [2]*Department of Hospital and Health Care Administration, Chia-Nan University of Pharmacy and Science, Tainan, Taiwan, ROC*

Abstract. The two-parameter Pareto distribution provides reasonably good fit to the distributions of income and property value, and explains many empirical phenomena. For the censored data, the two parameters are regularly estimated by the maximum likelihood estimator, which is complicated in computation process. This investigation proposes a weighted least square estimator to estimate the parameters. Such a method is comparatively concise and easy to perceive, and could be applied to either complete or truncated data. Simulation studies are conducted in this investigation to show the feasibility of the proposed method. This report will demonstrate that the weighted least square estimator gives better performance than unweighted least square estimators with simulation cases. We also illustrate that the weighted least square estimator is very close to maximum likelihood estimator with simulation studies.

Key words: Pareto distribution, weighted least square method, type II censored data.

1. Introduction

Suppose that a random variable X has two-parameter Pareto distribution, then its probability density function (pdf) and cumulative distribution function (cdf) are respectively as follows:

$$f_X(x) = \beta\alpha^{\beta}x^{-(\beta+1)}, \quad \alpha > 0,\ \ \beta > 0,\ \ x \geq \alpha \tag{1}$$

and

$$F_X(x) = 1 - \left(\frac{\alpha}{x}\right)^{\beta}.$$

Also the survivor function $S_X(x) = 1 - F_X(x) = P(X > x) = (\alpha/x)^{\beta}$ is the probability of income greater than x in economics.

* Author for correspondence: 60, Erh-Jen RD., Sec.1, Jen-Te Town, Tainan County, Taiwan 717, ROC. Tel: (886)-6-2664911 Ext 125; Fax: (886)-6-2667308;
E-mail: lintao@mail.chna.edu.tw

Cohen and Whitten (1988) proposed a moment estimator (ME) to estimate the parameters of Pareto distribution. They provided tables to evaluate α and β because the expected values of the parameters α and β do not always exist. In case the data are complete, then the maximum likelihood estimators (MLE) of α and β are easy to calculate. It was also proved by Quandt (1966) that the MLE $\hat{\alpha}$ and $\hat{\beta}$ are consistent with the parameters α and β, respectively. But it is complicated to compute MLE for censored data. Furthermore, since $\hat{\alpha} = x_{(1)}$, where $x_{(1)}$ is the minimum observation of a data set, then if $x_{(1)}$ is not observed in the censored schemes, the MLE of α will have a large bias. Geisser (1984, 1985) has provided extensive analysis of the application of Bayesian methods in predicting the future values of this distribution of random variables from observed values in a complete random sample. Nigm and Hamdy (1987) have considered a similar problem under censored data. The least square method is another measure often used to estimate the parameters. This investigation suggests a weighted least squares method for the estimation of the parameters, which is suitable for both complete and censored data. Section 2 will provide the methodology of the proposed method in detail. Section 3 will evaluate the new approach by simulation cases. A brief discussion is included in the last section.

2. Methodology and Results

2.1. THE MOMENT ESTIMATOR AND THE MAXIMUM LIKELIHOOD ESTIMATOR

First, we give a brief introduction of the ME and MLE. Suppose that random variables $X_1, X_2, \ldots, X_n$ are independent and identically distributed as (1), then Quandt (1966) proposed the following ME for both of the parameters in (1):

$$\widehat{\beta} = \frac{n\bar{x} - x_{(1)}}{n(\bar{x} - x_{(1)})}, \quad \widehat{\alpha} = \frac{(n\widehat{\beta} - 1)x_{(1)}}{n\widehat{\beta}},$$

where $\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$, $x_{(1)}$ is the minimum order statistics and n is the sample size. The MLE for α and β, according to Johnson et al. (1994), are respectively as follows:

$$\hat{\alpha} = \min_{1 \le i \le n} x_i = x_{(1)}, \quad \hat{\beta} = n\left[\sum_{j=1}^{n} \log\left(\frac{x_j}{\hat{\alpha}}\right)\right]^{-1}.$$

Next, we are going to introduce the least square methods in linear regression analysis.

2.2. UNWEIGHTED LEAST SQUARE METHODS

By taking natural log on both sides of (1), we have

$$\log(1 - F_X(x)) = \beta \log x - \beta \log \alpha. \tag{2}$$

It is easy to estimate α and β by unweighted least square method if we use the cdf of sample, $\hat{F}_X(x_{(i)}) = (i/n)$, as the estimator of $F_X(x)$ and $\log(1 - \hat{F}_X(x_{(i)})$ as the dependent variable. In this investigation, we choose

$$\hat{F}_X(x_{(i)}) = \frac{i - c}{n - 2c + 1}, \quad 0 \leq c \leq 1,$$

as the estimator of $F_X(x)$ where c takes the value 0.3 by following David (1981). In this investigation, we denote the estimators derived by this method as OLSE for simplification.

Next, we introduce another unweighted least square method. By multiplying −1 to both sides of (2), we obtain

$$-\log(1 - F_X(x)) = -\beta \log x + \beta \log \alpha. \tag{3}$$

From the property of Pareto distribution, we learn that $-\log(1 - F_X(x))$ is standard exponential distribution. Let

$$Z = -\log(1 - F_X(X)) \text{ and } Z = -\beta \log X + \beta \log \alpha. \tag{4}$$

It turns out that $Z_{(1)} < Z_{(2)} < \cdots < Z_{(n)}$ are the corresponding order statistics of $X_{(1)} < X_{(2)} < \cdots < X_{(n)}$. Therefore, according to Balakrishnan and Cohen (1991), the expected value of $Z_{(i)}$ is given as:

$$EZ_{(i)} = \sum_{j=1}^{i} \frac{1}{n - j + 1}. \tag{5}$$

If we choose $EZ_{(i)}$ as the estimate of $Z_{(i)}$ and the dependent variable of (4), then the estimates of α and β can be derived by applying least square method. The estimators derived here are denoted by ELSE for the purpose of distinction.

2.3. WEIGHTED LEAST SQUARE METHOD

In Section 2.2, we assume that every location has the same weight while applying least square method no matter what value does the dependent variable take. Bergman (1986) emphasized that it is not proper to have the same weight on each point while using regression analysis to estimate the parameters of Weibull distribution. He suggested adding a weighting function while applying regression analysis. Meanwhile, if the regression model is in the form of

$$Y_i = a + bV_i + \varepsilon_i$$

then we have $Y_i = -\log(1 - F_X(X_{(i)}))$, $a = \beta \log \alpha$, $b = -\beta$, $V_i = \log X_{(i)}$ from (3), where ε_i is the random error with expected value $E(\varepsilon_i) = 0$. Apparently, $-\log(1 - F(X_{(i)}))$ in (3) is standard exponential distribution and the variance of the order statistics does not satisfy the condition of being constant. That is, the weight of each point is not identical. We suggest considering a weighting function in executing the regression analysis of (3). The weighting function we suggest is different from that of Bergman (1986), and the method we propose is based on the following reasoning. Since $Z = -\log(1 - F_X(X))$ is standard exponential distribution and its expected value is as stated in (5), and the variance of Z is

$$\mathrm{Var}(Z_{(i)}) = \sum_{j=1}^{i} \frac{1}{(n-j+1)^2}.$$

Bickel and Doksum (2001) proposed an approach of variance stabilizing transformation. We adopt such an idea here that for a transformation $h(Z)$ its variance is

$$\mathrm{Var}(h(Z)) = \mathrm{Var}(Z)[h'(Z)]^2.$$

In this investigation, $h(Z) = Z$, that is, $\mathrm{Var}(h(Z)) = \mathrm{Var}(Z)$and $h'(Z) = 1$. We suggest taking the reciprocal of $\mathrm{Var}(h(Z))$ as the weighting factor, that is,

$$w_i = \frac{1}{\mathrm{Var}(h(Z_{(i)}))} = \frac{1}{\mathrm{Var}(Z_{(i)})}.$$

Therefore, $\mathrm{Var}(\varepsilon_i)$ is constant at this time. By adding up the weighted square terms, we have

$$Q = \sum_{i=1}^{n} w_i [Z_{(i)} - Z_i(X_{(i)})]^2,$$

where

$$w_i = \frac{1}{\sum_{j=1}^{i} \left(\frac{1}{n-j+1}\right)^2}, \quad Z_{(i)} \approx E(Z_{(i)}) = \sum_{j=1}^{i} \frac{1}{n-j+1},$$

$$Z_i(X_{(i)}) = \beta \log \alpha \; - \beta V_i \quad \text{and} \quad V_i = \log X_{(i)}, \;\; i = 1, \ldots, n.$$

Table I. The estimation of β by different estimators at true values $\alpha = 1$ and $\beta = 0.5$ (note: the values of each entry are mean $\pm$ SD and (MSE))

n	OLSE	WLSE	ELSE	ME	MLE
10	0.5109 ± 0.2340 (0.2343)	0.5078 ± 0.1965 (0.1967)	0.5436 ± 0.2390 (0.2430)	1.0816 ± 0.1323 (0.5965)	0.6305 ± 0.2448 (0.2774)
12	0.5010 ± 0.2020 (0.2021)	0.5015 ± 0.1693 (0.1693)	0.5345 ± 0.2137 (0.2165)	1.0637 ± 0.0935 (0.5714)	0.6022 ± 0.1991 (0.2238)
14	0.4979 ± 0.1853 (0.1853)	0.5029 ± 0.1545 (0.1545)	0.5289 ± 0.1976 (0.1998)	1.0534 ± 0.0806 (0.5593)	0.5871 ± 0.1766 (0.1969)
16	0.4935 ± 0.1714 (0.1715)	0.4988 ± 0.1395 (0.1395)	0.5189 ± 0.1729 (0.1739)	1.0451 ± 0.0678 (0.5493)	0.5714 ± 0.1569 (0.1724)
18	0.4912 ± 0.1590 (0.1593)	0.4989 ± 0.1302 (0.1302)	0.5168 ± 0.1654 (0.1663)	1.0391 ± 0.0596 (0.5424)	0.5611 ± 0.1445 (0.1569)
20	0.4896 ± 0.1497 (0.1500)	0.4986 ± 0.1231 (0.1231)	0.5156 ± 0.1576 (0.1584)	1.0356 ± 0.0524 (0.5382)	0.5571 ± 0.1346 (0.1463)
30	0.4846 ± 0.1200 (0.1210)	0.4986 ± 0.0998 (0.0998)	0.5081 ± 0.1248 (0.1250)	1.0224 ± 0.0326 (0.5234)	0.5358 ± 0.1019 (0.1080)
40	0.4874 ± 0.1059 (0.1066)	0.4978 ± 0.0852 (0.0852)	0.5040 ± 0.1081 (0.1082)	1.0169 ± 0.0244 (0.5175)	0.5272 ± 0.0859 (0.0901)
50	0.4848 ± 0.0930 (0.0943)	0.4997 ± 0.0764 (0.0764)	0.5008 ± 0.0950 (0.0950)	1.0128 ± 0.0184 (0.5132)	0.5202 ± 0.0745 (0.0772)
60	0.4880 ± 0.0868 (0.0876)	0.4993 ± 0.0691 (0.0691)	0.5000 ± 0.0882 (0.0882)	1.0111 ± 0.0155 (0.5113)	0.5183 ± 0.0687 (0.0711)
80	0.4872 ± 0.0738 (0.0749)	0.4983 ± 0.0597 (0.0597)	0.5003 ± 0.0762 (0.0762)	1.0082 ± 0.0117 (0.5084)	0.5126 ± 0.0580 (0.0593)
100	0.4883 ± 0.0668 (0.0678)	0.4994 ± 0.0537 (0.0537)	0.4988 ± 0.0680 (0.0680)	1.0064 ± 0.0089 (0.5065)	0.5099 ± 0.0515 (0.0524)

After minimizing Q, we obtain the weighted least square estimator (WLSE) of β and α which are respectively as follows:

$$\hat{\beta} = -\frac{\sum_{i=1}^{n} w_i Z_i \sum_{i=1}^{n} w_i V_i - \sum_{i=1}^{n} w_i \sum_{i=1}^{n} w_i Z_i V_i}{\sum_{i=1}^{n} w_i \sum_{i=1}^{n} w_i V_i^2 - \left(\sum_{i=1}^{n} w_i V_i\right)^2} \text{ and } \hat{\alpha} = \exp\left[\frac{\sum_{i=1}^{n} w_i Z_i - \hat{\beta} \sum_{i=1}^{n} w_i V_i}{\hat{\beta} \sum_{i=1}^{n} w_i}\right] \tag{6}$$

3. Simulation Studies

In this section we are going to evaluate the feasibility of the proposed WLSE method. In order to achieve an effective evaluation, Monte Carlo method is

Table II. The estimation of β by different estimators at true values $\alpha = 1$ and $\beta = 1$ (note: the values of each entry are mean $\pm$ SD and (MSE))

n	OLSE	WLSE	ELSE	ME	MLE
10	1.0174 ± 0.4496 (0.4499)	1.0127 ± 0.3814 (0.3817)	1.0891 ± 0.4752 (0.4836)	1.4356 ± 0.3716 (0.5725)	1.2502 ± 0.4670 (0.5299)
12	1.0116 ± 0.3965 (0.3967)	1.0042 ± 0.3449 (0.3450)	1.0689 ± 0.4323 (0.4378)	1.3941 ± 0.3230 (0.5095)	1.1989 ± 0.4028 (0.4492)
14	0.9894 ± 0.3606 (0.3608)	0.9986 ± 0.3068 (0.3069)	1.0536 ± 0.3838 (0.3838)	1.3616 ± 0.2770 (0.4556)	1.1636 ± 0.3447 (0.3815)
16	0.9867 ± 0.3400 (0.3403)	0.9987 ± 0.2808 (0.2808)	1.0403 ± 0.3519 (0.3542)	1.3428 ± 0.2536 (0.4265)	1.1402 ± 0.3118 (0.3419)
18	0.9861 ± 0.3183 (0.3186)	0.9937 ± 0.2627 (0.2627)	1.0314 ± 0.3372 (0.3387)	1.3281 ± 0.2359 (0.4042)	1.1234 ± 0.2874 (0.3128)
20	0.9793 ± 0.2956 (0.2963)	0.9943 ± 0.2473 (0.2474)	1.0265 ± 0.3128 (0.3139)	1.3186 ± 0.2223 (0.3885)	1.1145 ± 0.2693 (0.2926)
30	0.9709 ± 0.2398 (0.2416)	0.9938 ± 0.1939 (0.1940)	1.0107 ± 0.2506 (0.2509)	1.2709 ± 0.1761 (0.3231)	1.0715 ± 0.2088 (0.2207)
40	0.9713 ± 0.2091 (0.2111)	0.9959 ± 0.1714 (0.1715)	1.0039 ± 0.2158 (0.2159)	1.2512 ± 0.1518 (0.2935)	1.0580 ± 0.1753 (0.1847)
50	0.9745 ± 0.1879 (0.1896)	0.9951 ± 0.1510 (0.1511)	1.0025 ± 0.1922 (0.1922)	1.2313 ± 0.1343 (0.2675)	1.0416 ± 0.1524 (0.1580)
60	0.9712 ± 0.1713 (0.1737)	0.9994 ± 0.1393 (0.1393)	0.9999 ± 0.1760 (0.1760)	1.2185 ± 0.1222 (0.2504)	1.0320 ± 0.1355 (0.1393)
80	0.9791 ± 0.1504 (0.1519)	0.9965 ± 0.1201 (0.1202)	1.0012 ± 0.1524 (0.1524)	1.2022 ± 0.1080 (0.2293)	1.0232 ± 0.1154 (0.1177)
100	0.9759 ± 0.1339 (0.1361)	0.9969 ± 0.1057 (0.1057)	1.0005 ± 0.1363 (0.1363)	1.1938 ± 0.0992 (0.2178)	1.0210 ± 0.1030 (0.1051)

used to generate the simulation data, and all data are of Pareto distribution. The sample sizes chosen in our simulation range from 10 to 100. In each sample size, 1000 random data sets are generated for simulation studies. For comparison purpose, we calculate the estimates of the parameters by the methods mentioned in this investigation, i.e. ME, MLE, OLSE, ELSE and WLSE. All the estimated results are listed in Tables I– VI, in which we choose $\alpha = 1$ and $\beta = 0.5, 1, 4$ separately as true values and calculate the mean square error (MSE) and standard deviation (SD) for each random sample and each mentioned method. It can be observed from Tables I to III that for the estimator $\hat{\beta}$ the MSE, SD and bias of the proposed WLSE are much smaller than those of OLSE and ELSE which are both unweighted least square estimators in regression analysis. Meanwhile, the MSE of ME is comparatively larger than WLSE method. And there shows no significant difference between WLSE and MLE about MSE and SD values. It could also be observed from the tables that the $\hat{\beta}$ of ME has a large bias at $\beta \leq 1$. From Tables IV to VI, we

Table III. The estimation of β by different estimators at true values $\alpha = 1$ and $\beta = 4$ (note: the values of each entry are mean $\pm$ SD and (MSE))

n	OLSE	WLSE	ELSE	ME	MLE
10	4.0580 ± 1.7948 (1.7959)	4.0971 ± 1.5706 (1.5737)	4.3751 ± 1.9666 (2.0021)	4.6165 ± 1.6818 (1.7914)	4.9782 ± 1.8706 (2.1110)
12	3.9909 ± 1.5991 (1.5992)	4.0506 ± 1.3831 (1.3841)	4.2492 ± 1.7262 (1.7442)	4.5363 ± 1.4907 (1.5843)	4.8226 ± 1.6306 (1.8264)
14	3.9613 ± 1.4773 (1.4779)	4.0104 ± 1.2442 (1.2443)	4.2008 ± 1.5457 (1.5588)	4.4321 ± 1.2930 (1.3633)	4.6629 ± 1.3923 (1.5422)
16	3.9497 ± 1.3533 (1.3543)	3.9628 ± 1.1147 (1.1154)	4.1440 ± 1.4183 (1.4257)	4.3710 ± 1.1688 (1.2263)	4.5641 ± 1.2413 (1.3635)
18	3.9285 ± 1.2396 (1.2417)	3.9787 ± 1.0388 (1.0391)	4.1361 ± 1.3101 (1.3172)	4.3645 ± 1.1090 (1.1674)	4.5300 ± 1.1679 (1.2826)
20	3.9143 ± 1.1933 (1.1964)	3.9774 ± 0.9965 (0.9968)	4.1161 ± 1.2514 (1.2568)	4.3002 ± 1.0268 (1.0698)	4.4426 ± 1.0726 (1.1604)
30	3.8721 ± 0.9679 (0.9764)	3.9817 ± 0.7748 (0.7750)	4.0418 ± 1.0008 (1.0017)	4.2038 ± 0.8138 (0.8389)	4.2893 ± 0.8290 (0.8781)
40	3.8895 ± 0.8468 (0.8540)	3.9792 ± 0.6770 (0.6774)	4.0276 ± 0.8665 (0.8669)	4.1599 ± 0.6802 (0.6988)	4.2194 ± 0.6870 (0.7212)
50	3.8739 ± 0.7424 (0.7530)	3.9894 ± 0.6050 (0.6051)	4.0209 ± 0.7776 (0.7779)	4.1349 ± 0.6113 (0.6260)	4.1789 ± 0.6151 (0.6407)
60	3.9039 ± 0.6928 (0.6995)	3.9820 ± 0.5525 (0.5528)	4.0067 ± 0.7121 (0.7122)	4.0977 ± 0.5425 (0.5512)	4.1358 ± 0.5413 (0.5581)
80	3.8939 ± 0.5842 (0.59379)	3.9896 ± 0.4759 (0.4760)	3.9949 ± 0.6093 (0.6094)	4.0772 ± 0.4732 (0.4795)	4.1013 ± 0.4672 (0.4781)
100	3.9071 ± 0.5297 (0.5378)	4.0021 ± 0.4327 (0.4327)	3.9994 ± 0.5437 (0.5437)	4.0602 ± 0.4160 (0.4204)	4.0788 ± 0.4113 (0.4188)

can see that the MSE, SD and bias of WLSE are a lot smaller than those of OLSE and ELSE for the estimator $\hat{\alpha}$. The bias of WLSE is less than that of MLE if $n < 50$, and less than that of ME if $n < 20$. In general, WLSE is not significantly different from MLE and is better than ME. Figures 1—6 display the ratio of MSE values between various least square estimators and MLE. It is obvious from the figures that unweighted least square estimators, OLSE and ELSE, are less efficient than WLSE. Therefore, the proposed weighted least square method is a reasonable and suitable approach. In the meantime, it is clear from the figures that the MSE of WLSE and MLE are almost equivalent, i.e. the ratio is close to 1 for a large n.

4. An Illustrated Example

In this section, we use the example proposed by BalaKrishnan and Cohen (1991) to illustrate the estimation of the parameters of Pareto

Table IV. The estimation of α by different estimators at true values $\alpha = 1$ and $\beta = 0.5$ (note: the values of each entry are mean $\pm$ SD and (MSE))

n	OLSE	WLSE	ELSE	ME	MLE
10	0.9866 ± 0.5038 (0.5040)	0.9637 ± 0.2766 (0.2790)	0.9765 ± 0.5083 (0.5088)	1.1350 ± 0.2793 (0.3102)	1.2520 ± 0.3077 (0.3978)
12	0.9782 ± 0.4707 (0.4713)	0.9658 ± 0.2253 (0.2278)	0.9622 ± 0.4609 (0.4625)	1.1041 ± 0.2167 (0.2404)	1.1986 ± 0.2351 (0.3078)
14	0.9623 ± 0.4240 (0.4257)	0.9684 ± 0.1868 (0.1895)	0.9633 ± 0.4154 (0.4170)	1.0864 ± 0.1821 (0.2016)	1.1659 ± 0.1953 (0.2562)
16	0.9583 ± 0.4024 (0.4046)	0.9740 ± 0.1674 (0.1694)	0.9614 ± 0.3970 (0.3989)	1.0736 ± 0.1526 (0.1694)	1.1422 ± 0.1622 (0.2157)
18	0.9540 ± 0.3831 (0.3859)	0.9765 ± 0.1535 (0.1553)	0.9600 ± 0.3755 (0.3776)	1.0657 ± 0.1381 (0.1530)	1.1261 ± 0.1459 (0.1929)
20	0.9497 ± 0.3642 (0.3676)	0.9778 ± 0.1335 (0.1353)	0.9599 ± 0.3648 (0.3670)	1.0564 ± 0.1166 (0.1296)	1.1101 ± 0.1226 (0.1648)
30	0.9405 ± 0.3026 (0.3084)	0.9870 ± 0.0937 (0.0946)	0.9662 ± 0.3081 (0.3099)	1.0357 ± 0.0730 (0.0812)	1.0706 ± 0.0754 (0.1033)
40	0.9480 ± 0.2751 (0.2800)	0.9913 ± 0.0757 (0.0762)	0.9705 ± 0.2708 (0.2724)	1.0271 ± 0.0537 (0.0601)	1.0530 ± 0.0550 (0.0764)
50	0.9437 ± 0.2466 (0.2530)	0.9919 ± 0.0598 (0.0604)	0.9672 ± 0.2471 (0.2493)	1.0213 ± 0.0429 (0.0479)	1.0418 ± 0.0438 (0.0606)
60	0.9522 ± 0.2289 (0.2339)	0.9943 ± 0.0522 (0.0525)	0.9684 ± 0.2284 (0.2306)	1.0178 ± 0.0351 (0.0394)	1.0349 ± 0.0357 (0.0499)
80	0.9529 ± 0.2027 (0.2081)	0.9949 ± 0.0421 (0.0424)	0.9749 ± 0.2010 (0.2026)	1.0128 ± 0.0259 (0.0289)	1.0255 ± 0.0262 (0.0366)
100	0.9562 ± 0.1819 (0.1872)	0.9965 ± 0.0361 (0.0363)	0.9770 ± 0.1806 (0.1820)	1.0107 ± 0.0206 (0.0232)	1.0208 ± 0.0208 (0.0294)

distribution. The following data are a set of simple right truncated samples of size 20.

$$1010, 1035, 1049, 1062, 1076, 1098, 1109, 1130, 1180, 1210$$

$$1270, 1315, 1332, 1450, 1480, 1505, 1600, 1652, 1813, 1951$$

We use the iterative numerical method to calculate the MLE that gives $\hat{\alpha}_m = 1010$ and $\hat{\beta}_m = 3.662$, and the asymptotic variance of $\hat{\beta}_m, V(\hat{\beta}_m) = 1.0713$. In this example, the WLSE proposed in this investigation produces $\hat{\alpha}_\omega = 999.42$ and $\hat{\beta}_\omega = 3.5828$, and the asymptotic variance of $\hat{\beta}_\omega, V(\hat{\beta}_\omega) = 0.8583$. These values can be derived by the covariance matrix in weighted least square method. Now using the goodness of fit test of Cramer–Von Mises to examine if the distribution of the parameter is appropriate. That is,

Table V. The estimation of α by different estimators at true values $\alpha = 1$ and $\beta = 1$ (note: the values of each entry are mean $\pm$ SD and (MSE))

n	OLSE	WLSE	ELSE	ME	MLE
10	0.9581 ± 0.2564 (0.2598)	0.9743 ± 0.1301 (0.1326)	0.9514 ± 0.2503 (0.2550)	1.0297 ± 0.1147 (0.1185)	1.1112 ± 0.1228 (0.1657)
12	0.9633 ± 0.2314 (0.2343)	0.9771 ± 0.1054 (0.1079)	0.9534 ± 0.2296 (0.2343)	1.0240 ± 0.0946 (0.0976)	1.0921 ± 0.0997 (0.1358)
14	0.9534 ± 0.2204 (0.2253)	0.9800 ± 0.0897 (0.0920)	0.9574 ± 0.2165 (0.2207)	1.0179 ± 0.0787 (0.0807)	1.0763 ± 0.0823 (0.1123)
16	0.9570 ± 0.2074 (0.2118)	0.9836 ± 0.0807 (0.0823)	0.9586 ± 0.2053 (0.2094)	1.0161 ± 0.0690 (0.0708)	1.0673 ± 0.0720 (0.0985)
18	0.9562 ± 0.2010 (0.2057)	0.9847 ± 0.0717 (0.0733)	0.9567 ± 0.2007 (0.2053)	1.0138 ± 0.0597 (0.0612)	1.0593 ± 0.0619 (0.0858)
20	0.9568 ± 0.1935 (0.1983)	0.9872 ± 0.0667 (0.0680)	0.9618 ± 0.1905 (0.1943)	1.0113 ± 0.0528 (0.0540)	1.0522 ± 0.0546 (0.0755)
30	0.9579 ± 0.1613 (0.1667)	0.9920 ± 0.0468 (0.0475)	0.9645 ± 0.1619 (0.1658)	1.0074 ± 0.0353 (0.0361)	1.0350 ± 0.0361 (0.0503)
40	0.9626 ± 0.1435 (0.1483)	0.9942 ± 0.0362 (0.0366)	0.9705 ± 0.1438 (0.1468)	1.0051 ± 0.0264 (0.0269)	1.0259 ± 0.0269 (0.0374)
50	0.9671 ± 0.1299 (0.1340)	0.9954 ± 0.0302 (0.0306)	0.9751 ± 0.1289 (0.1313)	1.0038 ± 0.0206 (0.0210)	1.0206 ± 0.0209 (0.0293)
60	0.9676 ± 0.1195 (0.1238)	0.9965 ± 0.0262 (0.0264)	0.9758 ± 0.1198 (0.1222)	1.0032 ± 0.0171 (0.0174)	1.0172 ± 0.0173 (0.0244)
80	0.9720 ± 0.1046 (0.1083)	0.9976 ± 0.0208 (0.0209)	0.9813 ± 0.1034 (0.1051)	1.0009 ± 0.0125 (0.0127)	1.0125 ± 0.0126 (0.0177)
100	0.9730 ± 0.0959 (0.0996)	0.9979 ± 0.0175 (0.0176)	0.9842 ± 0.0932 (0.0946)	1.0016 ± 0.0100 (0.0102)	1.0101 ± 0.0101 (0.0143)

$$\begin{cases} H_0 : F(t) = \hat{F}(t) \\ H_1 : F(t) \neq \hat{F}(t) \end{cases} \quad \text{for all } t$$

where $\hat{F}(t) = \hat{F}(t, \hat{\alpha}_m, \hat{\beta}_m)$ or $\hat{F}(t) = \hat{F}(t, \hat{\alpha}_\omega, \hat{\beta}_\omega)$. The test statistics of Cramer–Von Mises is

$$W^2 = \sum_{i=1}^{n} \left(\hat{F}(x_{(i)}) - \frac{i - 0.5}{n} \right)^2 + \frac{1}{12n}.$$

If we substitute $\hat{F}(t)$ with $\hat{F}(t, \hat{\alpha}_m, \hat{\beta}_m)$, then we have $W^2 = 0.1064 < W_{\mathrm{cr}}^{2*} = 0.461$ [i.e. $P(W_{\mathrm{cr}}^{2*} > 0.461) = 0.05$ according to Kececioglu (1993, p. 742, Table 21.1)]. Hence we cannot reject H_0. If we substitute $\hat{F}(t)$ with $\hat{F}(t, \hat{\alpha}_\omega, \hat{\beta}_\omega)$, then we have $W^2 = 0.113 < W_{\mathrm{cr}}^{2*} = 0.461$. Apparently, the data in this example fit the proposed distribution and there is no significant difference between the value of W^2 of WLSE and that of MLE.

Table VI. The estimation of α by different estimators at true values $\alpha = 1$ and $\beta = 4$ (note: the values of each entry are mean $\pm$ SD and (MSE))

n	OLSE	WLSE	ELSE	ME	MLE
10	0.9835 ± 0.0697 (0.0716)	0.9923 ± 0.0312 (0.0322)	0.9809 ± 0.0707 (0.0732)	1.0009 ± 0.0271 (0.0271)	1.0257 ± 0.0265 (0.0369)
12	0.9827 ± 0.0659 (0.0681)	0.9939 ± 0.0262 (0.0269)	0.9809 ± 0.0658 (0.0685)	1.0006 ± 0.0221 (0.0221)	1.0212 ± 0.0217 (0.0303)
14	0.9825 ± 0.0617 (0.0641)	0.9945 ± 0.0223 (0.0230)	0.9832 ± 0.0607 (0.0630)	1.0003 ± 0.0184 (0.0184)	1.0181 ± 0.0181 (0.0256)
16	0.9842 ± 0.0578 (0.0599)	0.9951 ± 0.0198 (0.0204)	0.9833 ± 0.0577 (0.0601)	1.0004 ± 0.0160 (0.0160)	1.0160 ± 0.0159 (0.0225)
18	0.9846 ± 0.0551 (0.0572)	0.9959 ± 0.0180 (0.0184)	0.9864 ± 0.0537 (0.0553)	1.0003 ± 0.0145 (0.0145)	1.0140 ± 0.0143 (0.0200)
20	0.9854 ± 0.0528 (0.0548)	0.9963 ± 0.0163 (0.0167)	0.9853 ± 0.0527 (0.0547)	1.0004 ± 0.0131 (0.0131)	1.0128 ± 0.0129 (0.0182)
30	0.9858 ± 0.0441 (0.0464)	0.9978 ± 0.0114 (0.0116)	0.9884 ± 0.0434 (0.0449)	1.0001 ± 0.0086 (0.0086)	1.0084 ± 0.0085 (0.0120)
40	0.9879 ± 0.0390 (0.0409)	0.9984 ± 0.0090 (0.0091)	0.9907 ± 0.0380 (0.0391)	1.0001 ± 0.0063 (0.0063)	1.0063 ± 0.0063 (0.0089)
50	0.9889 ± 0.0348 (0.0365)	0.9990 ± 0.0076 (0.0076)	0.9922 ± 0.0342 (0.0350)	1.0000 ± 0.0052 (0.0052)	1.0050 ± 0.0051 (0.0071)
60	0.9906 ± 0.0316 (0.0330)	0.9991 ± 0.0066 (0.0066)	0.9927 ± 0.0314 (0.0322)	1.0001 ± 0.0043 (0.0043)	1.0042 ± 0.0042 (0.0060)
80	0.9918 ± 0.0274 (0.0286)	0.9994 ± 0.0053 (0.0053)	0.9943 ± 0.0266 (0.0272)	1.0000 ± 0.0031 (0.0031)	1.0031 ± 0.0031 (0.0043)
100	0.9927 ± 0.0246 (0.0256)	0.9995 ± 0.0044 (0.0044)	0.9955 ± 0.0239 (0.0243)	1.0000 ± 0.0025 (0.0026)	1.0025 ± 0.0025 (0.0036)

5. Discussion

We propose an alternative weighted least square method to estimate the parameters of two-parameter Pareto distribution. For Pareto distribution, the bias is generally large when we adopt $X_{(1)}$, the minimum observation, to calculate the MLE of α. But $X_{(1)}$ does not necessarily exist when the data is not complete, such as Type II double censored data and multiple censored data. Besides, we generally have to apply Newton–Raphson method or EM method in an iterative manner to solve for MLE under any type of data. Such iterative processes demand a careful selection of initial values otherwise it might cause wandering and sometime even divergent as one can see in solving likelihood equation. Such problems are removed from the proposed WLSE, which can apply to either complete or censored data and has a solution of the closed form as (6). From the results of simulation, one can see that adding weighting factor gives superior estimation in the regression analysis of linear

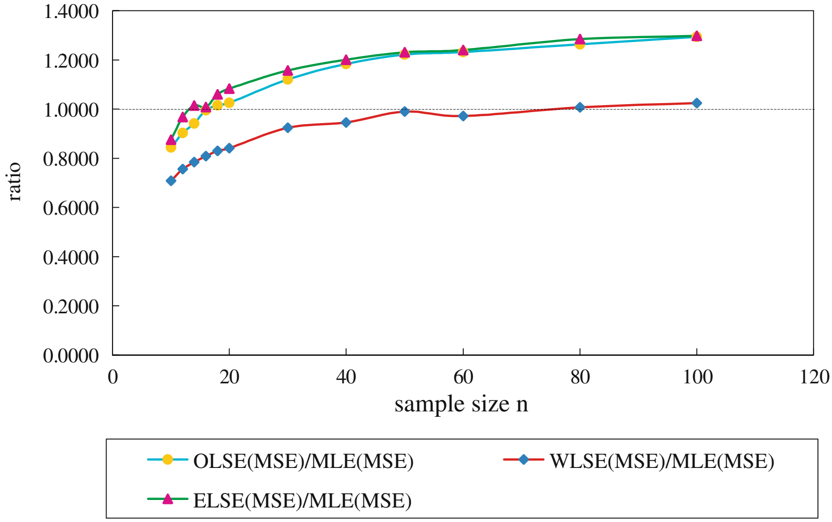

Figure 1. The ratio of MSE of $\hat{\beta}$ between various least square estimators and MLE at $\alpha = 1, \beta = 0.5$.

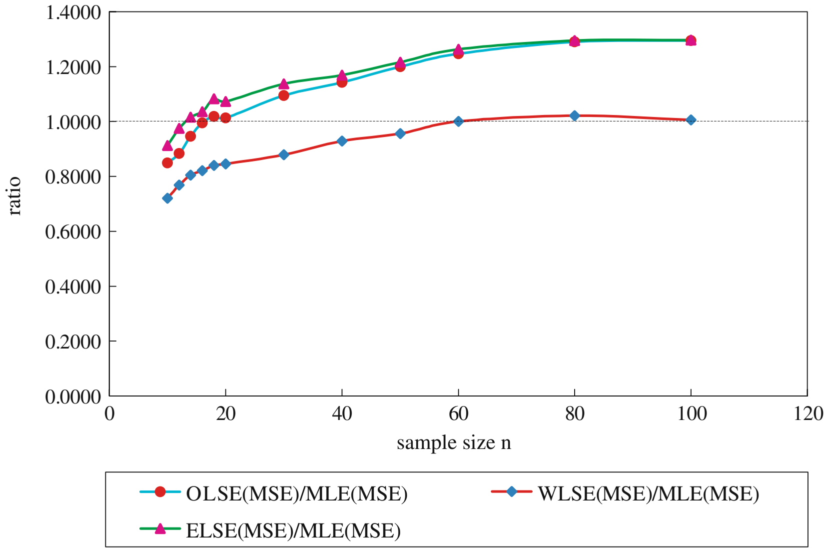

Figure 2. The ratio of MSE of $\hat{\beta}$ between various least square estimators and MLE at $\alpha = 1, \beta = 1$.

models. The MSE and bias of WLSE are comparatively smaller than those of OLSE and ELSE. The results of WLSE and MLE under complete data are of no significant difference. The MSE of WLSE and MLE are approximately

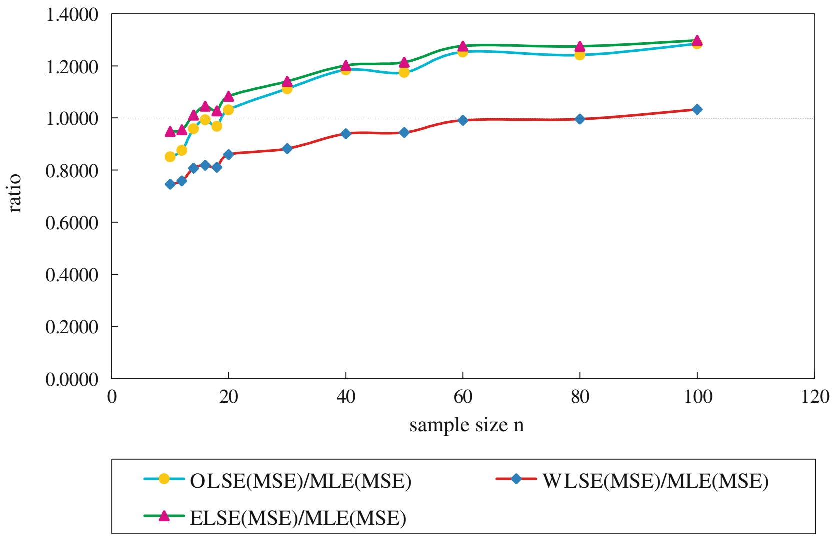

Figure 3. The ratio of MSE of $\hat{\beta}$ between various least square estimators and MLE at $\alpha = 1, \beta = 4$.

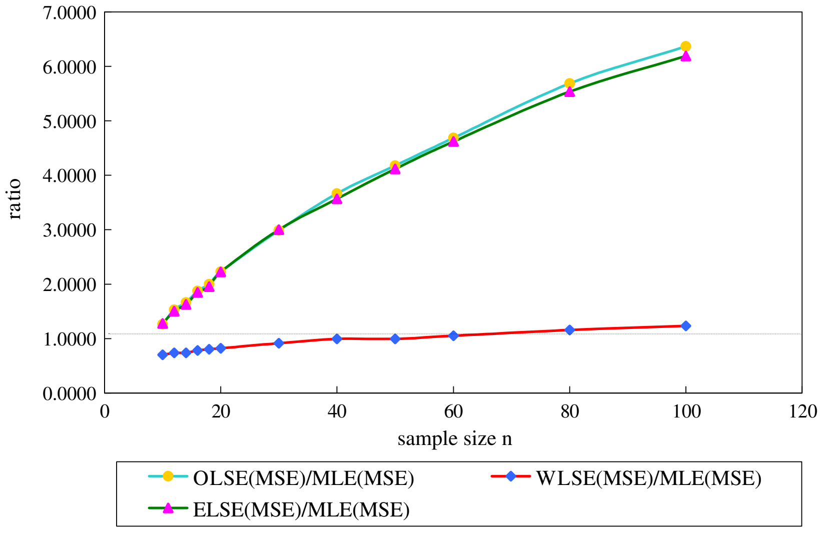

Figure 4. The ratio of MSE of $\hat{\alpha}$ between various least square estimators and MLE at $\alpha = 1, \beta = 0.5$.

equivalent when the sample size n is large. All the above results show the importance and essential of the weighting factor. We conclude that WLSE is an easy, feasible and effective method to estimate the parameters of two-parameter Pareto distribution.

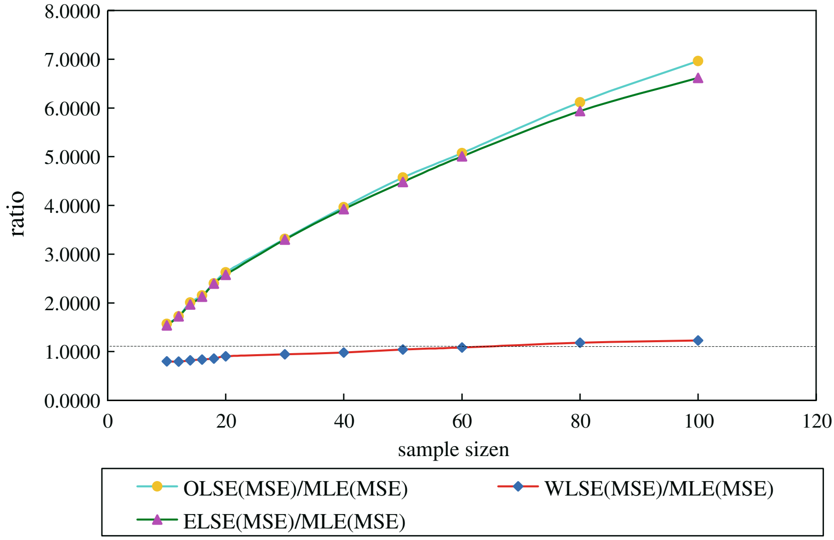

Figure 5. The ratio of MSE of $\hat{\alpha}$ between various least square estimators and MLE at $\alpha = 1, \beta = 1$.

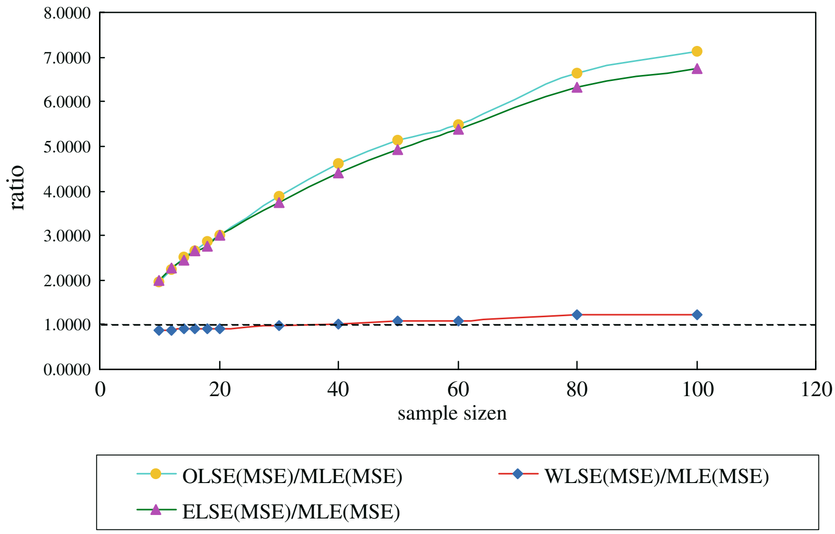

Figure 6. The ratio of MSE of $\hat{\alpha}$ between various least square estimators and MLE at $\alpha = 1, \beta = 4$.

References

Balakrishnan, N. & Cohen, A. C. (1991). *Order Statistics and Inference*. San Diego: Academic Press, Inc.

Bergman, B. (1986). Estimation of weibull parameters using a weight function. *Journal of Materials Science Letters* 5: 611–619.

Bickel, P. J. & Doksum, K. A. (2001). *Mathematical Statistics-Basic Ideas and Selected Topics*, Vol I. New Jersey: Prentice-Hall, Inc.

Cohen, A. C. & Whitten, B. J. (1988). *Parameter Estimation in Reliability and Life Span Models*. New York: Marcel Dekker.

David, H. A. (1981). *Order Statistics*, 2nd edn. New York: Wiley.

Geisser, S. (1984). Predicting Pareto and Exponential Observables. *Canadian Journal of Statistics* 12: 143–152.

Geisser, S. (1985). Interval prediction for Pareto and Exponential Observables. *Journal of Econometrics* 29: 173–185.

Johnson, N. L., Kotz, S. & Balakrishnan, N. (1994). *Continuous Univariate Distributions*. New York: John Wiley & Sons, Inc.

Kececioglu, D. (1993). *Reliability and Life Testing Handbook*, Vol. 1. New Jersey: PTR Prentice-Hall Inc.

Nigm, A. M. & Hamdy, H. I. (1987). Bayesian prediction bounds for the Pareto lifetime Model. *Communication in Statistics – Theory & Method* 16: 1761–1772.

Quandt, R. E. (1966). Old and new methods of estimation and the Pareto distribution. *Metrika* 10: 55–82.