*University of Thi-Qar Journal Vol. 13 No.3 SEP 2018*
*Web Site:* https://jutq.utq.edu.iq/index.php/main
*Email: journal@jutq.utq.edu.iq*

# Generating random samples from Pareto Distribution to computing the methods to estimating the parameters by simulation

**Ali . L . Arif**
**College of Medicine / Thi-qar University**

## Abstract

In this paper, we carry out a comparison between some estimating methods for the two parameters of Paerto distribution like least squares, moment, ridge regression, relative least squares and maximum likelihood by using the Mean squared errors and total deviation to determine the best method by using different values for the parameters and different sample size. Were generated with use programming simulation by Quick basic.

**Keywords:** Pareto distribution, Methods of estimation**,** numerical imulation.

**الخلاصة:**

في هذا البحث تمت المقارنة بين بعض طرائق التقدير لمعلمتي توزيع باريتو مثل طريقة المربعات الصغرى والعزوم وانحدار العتبة والمربعات الصغرى النسبية والإمكان الأعظم.وذلك باستخدام طريـــــــقة متوسط مربعات الخطأ وطريقة الانحراف الكلي للوصول إلى أفضل الطرائق في التقدير باستخدام قيــــــــم افتراضية مختلفة للمعلمات وبأحجام مختلفة للعينات والتي تم توليدها باستخدام المحاكاة بواسطة برنــــــامج ( q-basic ) .

**الكلمات المفتاحية :** توزيع باريتو , طرق التقدير , محاكاة عددية .

## 1. Introduction

The Pareto distribution is a probability model for continuous variable..The distribution is commonly used in many situations incomes, economic, and city population within a given area .The density function has two primary parameters, shape and scale as given below.

$$f(x;\alpha,\beta)=\frac{\alpha}{x}\left(\frac{\beta}{x}\right)^{\alpha} \qquad \alpha\ ,\ \beta > 0\ ,\quad x > \beta$$

(1)

Where α is shape parameter , β is scale parameter

The scale is the position of the left edge of the probability density and the shape determines the steepness of ski slope.[3]

## 2. Least Squares Method

The least square method is extensively used in reliability engineering mathematics problems and the estimation of probability distribution parameters.

The cumulative distribution function and reliability function are given respectively by

$$F=(x;\alpha,\beta)=1-\left(\frac{\beta}{x}\right)^{\alpha} \tag{2}$$

$$R(x)=\left(\frac{x}{\beta}\right)^{-\alpha} \tag{3}$$

Taking logarithm of both sides of (3).

$$\log R(x_i)=-\alpha\log x_i+\alpha\log\beta \qquad i=1, 2\ldots n$$

then

$$\log(x_i)=\log\beta-\frac{1}{\alpha}\log R(x_i) \tag{4}$$

Equation (4) can be written in the form

$$Y_i=A+BX_i\ , \tag{5}$$

where A=log $\beta$ , $B=-\frac{1}{\alpha}$ , $Y_i=\log(x_i)$, $X_i=\log R(x_i)$

Let $$S = \sum_{i=1}^{n} (Y_i - A - BX_i)^2 \qquad (6)$$

Differentiation $S$ w.r.t A and B and equating to zero, we have the least square (LS) estimates of A and B as .[1]

$$\hat{A} = \frac{\sum \log R(x_i) \sum \log R(x_i) \log(x_i) - \sum (\log R(x_i))^2 \sum \log(x_i)}{(\sum \log R(x_i))^2 - n \sum (\log R(x_i))^2}$$

(7)

$$\hat{B} = \frac{\sum \log R(x_i) \sum \log(x_i) - n \sum \log R(x_i) \log(x_i)}{(\sum \log R(x_i))^2 - n \sum (\log R(x_i))^2} \qquad (8)$$

**3. Relative Least Squares Method.**

The relative least square estimators of A and B can be obtained by minimizing the sum of squares of the relative residuals, [2]

$$S = \sum_{i=1}^{n} \left( \frac{y_i - A - Bx_i}{x_i} \right)^2 \qquad (9)$$

$$S = \sum_{i=1}^{n} (1 - Aw_i - Bz_i)^2 \qquad (10)$$

Where $$w_i = \frac{1}{y_i} \qquad z_i = \frac{x_i}{y_i} \qquad (11)$$

Differentiating w.r.t A and B and equating to zero

$$\sum_{i=1}^{n} w_i = A \sum_{i=1}^{n} w_i^2 + B \sum_{i=1}^{n} w_i z_i \qquad (12)$$

$$\sum_{i=1}^{n} z_i = A\sum_{i=1}^{n} w_i z_i + B\sum_{i=1}^{n} z_i^{\,2} \qquad (13)$$

After simplification, we get

$$\hat{A} = \frac{\sum\left(\frac{1}{\log(x_i)}\right)\left(\frac{\log(R(x_i))}{\log((x_i))}\right)\sum\left(\frac{\log(R(x_i))}{\log((x_i))}\right) - \sum\left(\frac{1}{\log((x_i))}\right)\sum\left(\frac{\log(R(x_i))}{\log((x_i))}\right)^2}{\left(\sum\left(\frac{1}{\log((x_i))}\right)\left(\frac{\log(R(x_i))}{\log((x_i))}\right)\right)^2 - \sum\left(\frac{\log(R(x_i))}{\log((x_i))}\right)^2 \sum\left(\frac{1}{\log((x_i))}\right)^2}$$

(14)

$$\hat{B} = \frac{\sum\left(\frac{1}{\log(x_i)}\right)\left(\frac{\log(R(x_i))}{\log((x_i))}\right)\sum\left(\frac{1}{\log(x_i)}\right) - \sum\left(\frac{1}{\log((x_i))}\right)^2 \sum\left(\frac{\log(R(x_i))}{\log((x_i))}\right)}{\left(\sum\left(\frac{1}{\log((x_i))}\right)\left(\frac{\log(R(x_i))}{\log((x_i))}\right)\right)^2 - \sum\left(\frac{\log(R(x_i))}{\log((x_i))}\right)^2 \sum\left(\frac{1}{\log((x_i))}\right)^2}$$

(15)

Where $A = \log(\beta)$ and $B = -\dfrac{1}{\alpha}$

(16)

## 4. Ridge Regression

The ridge regression estimates of A and B can be obtained by minimizing the error sum of squares for the model (5) subject to a single constraint that $A^2 + B^2 = \rho$ where $\rho$ is a finite positive constant . The method of Lagrange multipliers requires the differentiation of

$$L = \sum_{i=1}^{n} (Y_i - A - BX_i)^2 + \lambda(A^2 + B^2 - \rho) \qquad (17)$$

with respect to A and B . When these derivatives are equated to zero , we obtain the following two equations [6]

$$\sum_{i=1}^{n} Y_i = (n+\lambda)A + B\sum_{i=1}^{n} X_i \qquad (18)$$

$$\sum_{i=1}^{n} Y_i X_i = A\sum_{i=1}^{n} X_i + B(\lambda + \sum X_i^{\ 2})$$

$$\hat{A} = \frac{\sum \log R(x_i)\sum \log R(x_i)\log(x_i) - \sum(\log R(x_i))^2 \sum(\lambda + \log(x_i))}{(\sum \log R(x_i))^2 - (n+\lambda)(\lambda + \sum(\log R(x_i))^2)}$$

(19)

$$\hat{B} = \frac{\sum \log R(x_i)\sum \log(x_i) - (n+\lambda)\sum \log R(x_i)\log(x_i)}{(\sum \log R(x_i))^2 - (n+\lambda)(\lambda + \sum(\log R(x_i))^2)} \qquad (20)$$

Where $0 < \lambda < 1$ is the ridge coefficient. If $\lambda = 0$ ,we obtain the least square estimates.

## 5. Maximum likelihood Estimation

The likelihood function of two –parameter model given in (1) can be written as

$$L = (\alpha, \beta \mid x) = \prod_{i=1}^{n} \frac{\alpha\beta^{\alpha}}{x_i^{\ \alpha+1}} \qquad (21)$$

Recall that the likelihood function considered as a function of parameters for fixed values of $x_1 .........x_n$. The maximum likelihood estimates for $\alpha$ and $\beta$ are the values of $\alpha$ and $\beta$ make L as large as possible given the data we have. Recall that $\beta$ can be no larger than the smallest value of x in our data, so the best we can do in maximizing L by adjusting $\beta$ is as follows:[5]

$$\hat{\beta} = \min\{x_i\} \qquad (22)$$

In order to find the maximum likelihood estimate for $\alpha$ , calculus is appropriate. So that the maximam likelihood estimator of $\alpha$ is given ,

$$\log L = (\alpha, \beta \mid x) = \prod_{i=1}^{n} \log\left(\frac{\alpha\beta^{\alpha}}{x_i^{\alpha+1}}\right) \qquad (23)$$

$$= n\log(\alpha) + \alpha n\log(\beta) - (\alpha - 1)\sum_{i=1}^{n}\log(x_i) \qquad (24)$$

$$\frac{d\log L}{d\alpha} = \frac{n}{\alpha} + n\log(\beta) - \sum_{i=1}^{n}\log(x_i) \qquad (25)$$

Setting the derivative equal to zero, a little algebra and an omitted second derivative check to confirm we are maximizing L rather than minimizing L yields:

$$\hat{\alpha} = \frac{n}{\sum_{i=1}^{n}\log(\frac{x_i}{\hat{\beta}})} \qquad (26)$$

**6. The Method of Moment**

The mean of the Pareto distribution exists and given by

$$E(x) = \frac{\alpha\beta}{\alpha - 1} \qquad (27)$$

provided $\alpha > 1$ .We estimate $\alpha$ by equating (27) to the sample mean $\bar{x}$ , yielding

$$\hat{\alpha} = \frac{\bar{x}}{\bar{x} - \hat{\beta}} \ , \qquad (28)$$

where $\hat{\beta}$ is some estimate of $\beta$ [4]

The estimation of $\beta$ from samples of n observations is accomplished as follows.

Equating the lowest sample value, $x_{0}$ , we obtain

$$\hat{\beta} = \frac{(n\alpha - 1)x_0}{n\alpha} \quad (29)$$

And
$$\hat{\alpha} = \frac{n\bar{x} - x_0}{n(\bar{x} - x_0)} \quad (30)$$

## 7. Goodness of Fit Analysis

Some methods of goodness of fit are employed. Mean square error MSE

$$MSE = \sum_{i=1}^{n} \frac{(\hat{R}(t(i) - R(t(i)))^2}{n} \quad (31)$$

And total deviation TD is two measurements that give an indication of accuracy of parameter estimation [7]

$$TD = \left|\frac{\hat{\alpha} - \alpha}{\alpha}\right| + \left|\frac{\hat{\beta} - \beta}{\beta}\right| \quad (32)$$

## 8. Application

A simulation study is carry out to compare the performance of the proposed estimation methods, We generate a data set of certain values of $\alpha$ and $\beta$ for samples of sizes of 10 , 20 , 30 ,100, 500 .The true value pairs are ($\alpha$ , $\beta$) : (2,1), (2,2), (3,2) and $\lambda$ :( 0, 0.1, 0.5 ) are the parameter of ridge regression .

The generation of random sample by observing that if U is uniform (0, 1), then $t = \beta(1-U)^{-1/\alpha}$

All results are based on 10000 replications.

**Table (1) Estimation of parameters with sample size 10 .**

| Method | $\alpha,\beta$ | $\lambda$ | $\hat{\alpha}$ | $\hat{\beta}$ | T D | MSE |
|---|---|---|---|---|---|---|
| L S M | 2 , 1 | 0 | 2 | 0.999999 | 0.00000035 | 0.001559 |
| | 2 , 2 | .1 | 1.99999 | 2 | 0.00000047 | 0.0012944 |
| | 3 , 2 | .5 | 2.99999 | 2 | 0.0000047 | 0.007650 |
| R R M | 2 , 1 | 0 | 2.232829 | 1.105914 | 1.550682 | 0.1765685 |
| | 2 , 2 | .1 | 2.0067 | 1.965146 | 0.699577 | 0.288801 |
| | 3 , 2 | .5 | 2.78333 | 2.03624 | 1.158939 | 0.1037725 |
| R L S | 2 , 1 | 0 | 2 | 0.999999 | 1.714711 | 0.0155992 |
| | 2 , 2 | .1 | 2 | 2 | 0.0000035 | 0.0012947 |
| | 3 , 2 | .5 | 3 | 2 | 0.8333335 | 0.007659 |
| M L E | 2 , 1 | 0 | 1.028253 | 1.00543 | 1.025012 | 0.3861737 |
| | 2 , 2 | .1 | 2.027394 | 1.368726 | 0.3293341 | 0.1856301 |
| | 3 , 2 | .5 | 2.055356 | 2.4264 | 0.5280811 | 0.095082 |
| M O M | 2 , 1 | 0 | 1.952334 | 0.9768406 | 1.550682 | 0.0016221 |
| | 2 , 2 | .1 | 1.069106 | 2.231544 | 0.581219 | 0.0096877 |
| | 3 , 2 | .5 | 7.416289 | 2.543262 | 1.743727 | 1.606918 |

**Table (2)** **Estimation of parameters with sample size 20**

| Method | $\alpha, \beta$ | $\lambda$ | $\hat{\alpha}$ | $\hat{\beta}$ | T D | MSE |
|---|---|---|---|---|---|---|
| L S M | 2 , 1 | 0 | 2 | 1 | 0.00000011 | 0.041987 |
| | 2 , 2 | .1 | 2.00001 | 2 | 0.00000476 | 0.00215 |
| | 3 , 2 | .5 | 3 | 2 | 0.00000198 | 0.00112 |
| R R M | 2 , 1 | 0 | 2 | 1 | 1.696735 | 0.2745048 |
| | 2 , 2 | .1 | 1.997 | 1.994 | 0.69819 | 0.1820 |
| | 3 , 2 | .5 | 2.907922 | 2..012921 | 1.127625 | 0.1957148 |
| R L S | 2 , 1 | 0 | 2 | 1 | 1.5 | 0.041989 |
| | 2 , 2 | .1 | 1.999 | 2 | 0.00000957 | 0.00217 |
| | 3 , 2 | .5 | 3 | 2 | 0.8333334 | 0.00114 |
| M L E | 2 , 1 | 0 | 1.0075 | 0.6329 | 0.8633025 | 0.523236 |
| | 2 , 2 | .1 | 2.0099 | 1.5239 | 0.243009 | 0.12307 |
| | 3 , 2 | .5 | 2.065457 | 2.199818 | 0.4114235 | 0.1662209 |
| M O M | 2 , 1 | 0 | 2.6162 | 1.2281 | 0.5362163 | 0.1847517 |
| | 2 , 2 | .1 | 1.0946 | 5.4332 | 3.263937 | 0.24389 |
| | 3 , 2 | .5 | 3.466332 | 2.129676 | 0.220282 | 0.0318 |

**Table (3)** **Estimation of parameters with sample size 50**

| Method | $\alpha, \beta$ | $\lambda$ | $\hat{\alpha}$ | $\hat{\beta}$ | T D | MSE |
|---|---|---|---|---|---|---|
| L S M | 2 , 1 | 0 | 2 | 1 | 0.00000238 | 0.003801 |
| | 2 , 2 | .1 | 2 | 2 | 0.00000476 | 0.00162 |
| | 3 , 2 | .5 | 3.00002 | 2 | 0.00000695 | 0.004055 |
| R R M | 2 , 1 | 0 | 2 | 1 | 1.696735 | 0.228759 |
| | 2 , 2 | .1 | 1.999165 | 1.99998 | 0.6972154 | 0.375217 |
| | 3 , 2 | .5 | 2.984304 | 1.997209 | 1.253727 | 0.2314573 |
| R L S | 2 , 1 | 0 | 2 | 1 | 1.5 | 0.003804 |
| | 2 , 2 | .1 | 2 | 1.99999 | 0.00000655 | 0.016266 |
| | 3 , 2 | .5 | 3.00003 | 2.00001 | 0.8333346 | 0.004056 |
| M L E | 2 , 1 | 0 | 1.000715 | 1.099609 | 0.5992509 | 0.4620868 |
| | 2 , 2 | .1 | 2.010422 | 1.757314 | 0.1265538 | 0.327123 |
| | 3 , 2 | .5 | 2.00888 | 1.57955 | 0.5405955 | 0.1398175 |
| M O M | 2 , 1 | 0 | 2.035579 | 1.017389 | 0.003517 | 0.003584 |
| | 2 , 2 | .1 | 4.13403 | 4.450605 | 4.292317 | 1.72858 |
| | 3 , 2 | .5 | 12.85723 | 2.759592 | 3.665539 | 460.662 |

**Table (4)** **Estimation of parameters with sample size 100**

| Method | $\alpha , \beta$ | $\lambda$ | $\hat{\alpha}$ | $\hat{\beta}$ | T D | MSE |
|---|---|---|---|---|---|---|
| L S M | 2 , 1 | 0 | 2 | 1 | 0.00000238 | 0.00161 |
| | 2 , 2 | .1 | 2.000001 | 2.000001 | 0.00000953 | 0.00144 |
| | 3 , 2 | .5 | 2.99998 | 2 | 0.00000874 | 0.00242 |
| R R M | 2 , 1 | 0 | 2 | 1 | 1.696735 | 0.162212 |
| | 2 , 2 | .1 | 1.99914 | 1.99864 | 0.69723 | 0.2016497 |
| | 3 , 2 | .5 | 2.963253 | 1.997448 | 1.243768 | 0.1984202 |
| R L S | 2 , 1 | 0 | 2 | 1 | 1.5 | 0.00165 |
| | 2 , 2 | .1 | 2.000001 | 2 | 0.00000834 | 0.0046 |
| | 3 , 2 | .5 | 3.000002 | 2 | 0.8333342 | 0.00242 |
| M L E | 2 , 1 | 0 | 1.01063 | 1.122689 | 0.6173692 | 0.3514912 |
| | 2 , 2 | .1 | 2.045797 | 1.554014 | 0.2458915 | 0.1393401 |
| | 3 , 2 | .5 | 2.01462 | 1.720789 | 0.4680655 | 0.1374648 |
| M O M | 2 , 1 | 0 | 2.481063 | 1.19156 | 0.4320916 | 0.00604 |
| | 2 , 2 | .1 | 2.615239 | 3.227582 | 0.9214106 | 1.99699 |
| | 3 , 2 | .5 | 3.003621 | 2.001189 | 0.000180 | 0.00243 |

**Table (5)** **Estimation of parameters with sample size 500**

| Method | $\alpha , \beta$ | $\lambda$ | $\hat{\alpha}$ | $\hat{\beta}$ | T D | MSE |
|---|---|---|---|---|---|---|
| L S M | 2 , 1 | 0 | 2 | 1 | 0 | 0.02376 |
| | 2 , 2 | .1 | 1.99996 | 1.99995 | 0.00000476 | 0.00133 |
| | 3 , 2 | .5 | 3.00008 | 2.00003 | 0.00000405 | 0.0016765 |
| R R M | 2 , 1 | 0 | 2 | 1 | 1.696735 | 0.158178 |
| | 2 , 2 | .1 | 1.999838 | 2.000007 | 0.6968281 | 0.1883615 |
| | 3 , 2 | .5 | 2.996747 | 1.999882 | 1.127625 | 0.258769 |
| R L S | 2 , 1 | 0 | 2 | 1 | 1.5 | 0.02388 |
| | 2 , 2 | .1 | 2 | 2 | 0.00000357<br>6 | 0.001338 |
| | 3 , 2 | .5 | 3.000014 | 2.00002 | 0.8333339 | 0.001678 |
| M L E | 2 , 1 | 0 | 1.0001 | 0.96678 | 0.5331373 | 0.3628124 |
| | 2 , 2 | .1 | 2.000275 | 1.766382 | 0.1169465 | 0.156668 |
| | 3 , 2 | .5 | 2.065457 | 2.199818 | 0.475487 | 0.1799932 |
| M O M | 2 , 1 | 0 | 34.06602 | 1.939463 | 16.97248 | 1.51934 |
| | 2 , 2 | .1 | 1.232339 | 2.378553 | .5731071 | 0.11246 |
| | 3 , 2 | .5 | 8.934164 | 2.66303 | 2.30957 | 21.09681 |

## 6. Conclusions

The results of simulation are listed in tables (1), (2), (3), (4) and (5). From these tables, we see that the L S estimates of parameters are too close to the true values and the values of MSE and TD are very small .The parameter estimates from RR ,RLS ,MLE and MOM methods are close to the true values but not as close as L S estimates, because the values of MSE and TD are greater than the corresponding values from L S . Is shown from the computational results that the estimators are much closer to the true parameter values when no outliers were present L S method are best for estimation of scale and shape parameters of the Pareto distribution.

## 7. References

1. Das, I. (2000). Applicability of existing continuous methods in determining the Pareto
set for nonlinear, mixed-integer multicriteria optimization problems. In Proceedings
of the 8th AIAA/USAF/NASA/ISSMO symposium on multidisciplinary analysis &
Optimization.

2. Hosking, J.R.M. and J.R. Wallis. 1987. Parameter and quantile estimation for the
generalized Pareto distribution. Technometrics 29: 339-349.

3. Mierlus-Mazilu (2010), “On Generalized Pareto Distributions”, Romanian Journal of
Economic Forecasting, 8: 107-117

4. Norman L. Johnson and Samuel Kotz, Distributions in Statistics Continuous Univariate
distributions - 1. , Houghton Mifflin Company, Boston,1970

5. Ranneby, Bo (1984) The maximum spacing method. An estimation method related to the
maximum likelihood method. Scand. J. Statist., 11, 93-112.

6. Ronald, E.W. and Raymond, H.M. (1978) Probability and statistics for engineers and

scientists. Second Edition. Macmillan publishing Co., Inc, NewYork.

7. Wingo, D. R. (1982) Unimodality of the Pareto distribution likelihood function for multi censored samples and implications for estimation.Commun. Statistics, Theory and methods. 11 (10), 1129- 1138.