

Estimating the Parameters of a Pareto Distribution

Introducing a Quantile Regression Method

Joseph Lee Petersen

Introduction. A broad approach to using correlation coefficients for parameter estimation and not merely as descriptive statistics has been developed. [1,2,3,4] It was a goal of this project to extend these ideas specifically to estimating the parameters of the Pareto distribution. In this paper we will recall the definition of the Pareto distribution, some basic properties, and some previously developed methods of estimating the parameters of the Pareto distribution from which a random sample comes. We will introduce a new parameter estimation scheme based on correlation coefficients. Finally, we will study and compare the performance of each of the parameter estimation schemes.

1 The Pareto Distribution

The Pareto Distribution was first proposed as a model for the distribution of incomes. It is also used as a model for the distribution of city populations within a given area. The Pareto distribution is defined by the following functions:

$$\text{CDF: } F(x|\alpha, k) = 1 - \left(\frac{k}{x}\right)^\alpha; \quad k \leq x < \infty; \alpha, k > 0$$

$$\text{PDF: } f(x|\alpha, k) = \frac{\alpha k^\alpha}{x^{\alpha+1}}; \quad k \leq x < \infty; \alpha, k > 0$$

The first parameter marks a lower bound on the possible values that a Pareto distributed random variable can take on. To illustrate we can see in figure 1 a plot of the density of a Pareto(1, 1) random variable.

A few well known properties follow:

$$E(X) = \alpha k / (\alpha - 1), \alpha > 1$$

$$\text{Var}(X) = \alpha k^2 / [(\alpha - 1)^2 (\alpha - 2)], \alpha > 2$$

2 Parameter Estimation

We are interested in estimating the parameters of the Pareto distribution from which a random sample comes. We will outline a few parameter estimation schemes.

2.1 Method of Moments

We actually modify the usual method of moments scheme according to a method laid out in Johnson and Kotz[5]. If we set the sample mean equal to the distribution's theoretical expected value mentioned above and if we set the sample minimum, x_1 , equal to the theoretical expected value of the minimum of a size n sample of Pareto(k, α) random variables, we obtain two equations and two unknowns:

$$\bar{x} = \frac{\tilde{\alpha} \tilde{k}}{(\tilde{\alpha} - 1)}$$

$$x_1 = \frac{n \tilde{\alpha} \tilde{k}}{n \tilde{\alpha} - 1}$$

Solving these equations yields the following estimators:

$$\tilde{\alpha} = \frac{n \bar{x} - x_1}{n (\bar{x} - x_1)}$$

$$\tilde{k} = \frac{(n \tilde{\alpha} - 1) x_1}{n \tilde{\alpha}}$$

2.2 Median Estimator

As far as the author knows, this is a new estimator. The idea is that in method of moments we set the sample mean equal to the theoretical mean, so here we will set the sample median equal to the theoretical median. Many of the estimation schemes discussed in this paper were first studied in the

case where k was known to be equal to one, so that it was only α that needed to be estimated. In this case we can see from the CDF above that if $x = m_s$, the median of the sample, we can estimate α as follows:

$$1 - m_s^{-\tilde{\alpha}} = 0.5$$

$$\implies \tilde{\alpha} = \frac{\ln 2}{\ln m_s}$$

It is likely though that we will be interested in many cases in which k is not equal to one, so if we already have an estimate for k , call it k_{est} , we will make the following adjustment to this estimate for α :

$$\tilde{\alpha} = \frac{\ln 2}{\ln \left(\frac{m_s}{k_{est}} \right)}$$

For the purposes of analyzing the performance of this estimator, we will use the minimum sample value as the estimate for k .

2.3 Maximum Likelihood

The likelihood function, L , for the Pareto distribution has the following form:

$$L(k, \alpha | x) = \prod_{i=1}^n \frac{\alpha k^\alpha}{x_i^{\alpha+1}}; 0 < k \leq \min\{x_i\}, \alpha > 0$$

Recall that the likelihood function tells as a function of the distribution parameters how likely it is to have observed the data that we did in fact observe. The maximum likelihood estimates for k and α are the values of k and α that make L as large as possible given the data we have. The most familiar method of maximizing functions involves calculus. However, we need no calculus to see that L gets large beyond bound for increases in k . It is key then to recall that k can be no larger than the smallest value of x in our data, so the best we can do in maximizing L by adjusting k is as follows:

$$\hat{k} = \min\{x_i\}$$

In order to find the maximum likelihood estimate for α , calculus is appropriate. Since L is nonnegative, we can take its logarithm. We do this because

it is easier to differentiate $\log L$ than L itself. Logarithms are bijective functions, so the value of α that maximizes L also maximizes $\log L$. The process in brief looks like this:

$$\begin{aligned} \log L(k, \alpha | x) &= \sum_{i=1}^n \log \left(\frac{\alpha k^\alpha}{x_i^{\alpha+1}} \right) \\ &= n \log(\alpha) + \alpha n \log(k) - (\alpha + 1) \sum_{i=1}^n \log(x_i) \\ \implies \frac{d}{d\alpha} &= n/\alpha + n \log(k) - \sum_{i=1}^n \log(x_i) \end{aligned}$$

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:

$$\hat{\alpha} = n / \sum_{i=1}^n \log \left(\frac{x_i}{\hat{k}} \right)$$

2.4 Correlation Coefficients

Gideon[4] has shown using a correlation based interpretation of linear regression that the mean and standard deviation for a normally distributed data set of size n can be estimated by regression of the sorted data on the 1st through n th $(n + 1)$ -tiles of the standard normal. In such a regression, the intercept of the fitted linear model serves as an unbiased estimate of the mean of the distribution from which the data came, and the slope of the fitted linear model serves as an unbiased estimate of the standard deviation. This quantile regression estimation scheme not only is appropriate for normally distributed data though, but rather it works for distributions from any scale regular family.¹

The connection between regression and correlation is laid out explicitly in Gideon[4]. We will describe the connection here in brief. The value of s

¹A scale regular family is a family of distributions such that for any member in the family there is a scalar that can multiply the member to yield another member in the same family with unit variance. A common example of the use of the scale regular property of a scale regular family is standardizing a Normally distributed random variable.

that satisfies the following equation (See Gideon[1,2,3,4]) estimates standard deviation:

$$r(q, u^o - sq) = 0$$

where q is a vector of the 1st through n th $(n + 1)$ -tiles of the standard distribution, n is the sample size of the data, u^o is a vector of the ordered data, and r is any correlation coefficient. When r is Pearson's correlation, the solution is exactly the least squares estimate of the slope of a linear model. As for an estimate of the model intercept, if r is Pearson's, the estimate is $\bar{u}^o - s\bar{q}$. If r is Kendall's r or the Greatest Deviation Correlation Coefficient (See Gideon[2]), the estimate of intercept is $\text{median}(u^o) - s \text{median}(q)$. Again if r is Pearson's correlation coefficient, the solution is exactly the least squares line. However, it is important to note that r need not be Pearson's r , but rather other correlation statistics will yield estimates, with perhaps more desirable properties, of the slope and intercept of the line.

We will now describe how these ideas can be applied to the Pareto distribution. Let X be a Pareto(α, k) distributed random variable. Then $U = \ln X$ is a two-parameter Exponentially distributed random variable with parameters λ and θ . That is U has probability density function $f(u) = \left(\frac{1}{\lambda}\right) e^{-(u-\theta)/\lambda}$, where $\lambda = 1/\alpha$ and $\theta = \ln k$. Its expected value is $\lambda + \theta$ or in terms of the original Pareto random variable, $1/\alpha + \ln k$, and its variance is λ^2 or $1/\alpha^2$. (This means the standard deviation is $1/\alpha$.) It should be noted now that the Exponential distributions are a scale regular family. Now the random variable $Z = \alpha U$ will be rather a "standard" Exponential with probability density function $f(z) = e^{-(z-\theta)}$ and so cumulative density function $F(z) = 1 - e^{-(z-\theta)}$.

These facts suggest that if we log transform data which we hypothesize to be Pareto distributed, we can solve $r(q, u^o - sq) = 0$, q being the quantiles of the standard Exponential, to get an estimate of the scale parameter for the Exponential distribution to which the log transformed Pareto data corresponds, and the appropriate center measure, i , of the uncentered residuals will give an estimate of the location parameter of the Exponential distribution to which the log transformed Pareto data corresponds. In order to turn these estimates back in terms of the Pareto parameters, we estimate k by e^i and α by $1/s$. In our studies we used Pearson's r and the Greatest Deviation Correlation Coefficient (See Gideon[1,2]) to evaluate the performance of this estimation scheme.

3 An Example

We used the probability integral transformation to write a function in SPlus that would generate random samples of Pareto distributed data. For this example we generated a size 100 sample from a Pareto(1, 0.5) distribution. We can see a histogram of this data in figure 2.

We've truncated the histogram in order to compare it to the density plot in Figure 1, and because there are some extreme outliers that would have made the histogram not useful had they been included. (The max was 5715 and the next highest was 1581.) The next plot shows the sorted logarithm transformed data plotted as a function of the first through 100th 101-tiles of an Exponential(1) random variable.

The dotted line is a least squares regression line. The solid line is a Greatest Deviation based regression line. The intercept and slope of the least squares line are 0.126 and 2.006 which imply estimates for k and α of 1.134 and 0.498. The intercept and slope of the GD line are 0.023 and 2.139 which imply estimates for k and α of 1.023 and 0.467. Method of moments estimates $k = 1.025$ and $\alpha = 1.009$. It is important to note how horribly method of moments does in estimating α . The reason for this is that the integral defining variance for a Pareto distribution does not converge if α is less than or equal to two, and similarly the integral defining the distribution's mean is infinite if α is less than or equal to one. This means that the distribution is prone to extreme outliers. The limit of $\tilde{\alpha}$ as \bar{x} goes to infinity is one, so one extreme outlier can yield an estimate of one for α . The maximum likelihood estimates are $k = 1.035$ and $\alpha = 0.487$. The median estimate for α assuming a k estimate of k_{min} (1.035) is 0.504. Finally to illustrate the equation $r(q, u^o - sq) = 0$ we can see a plot of $r(q, u^o - sq)$ as a function of s . Remember u^o is actually the logarithm of the sorted presumed Pareto data and hence is distributed Exponentially.

The value of s where the graph crosses the s axis is the correlation estimate of the standard deviation of the Exponential distribution. Our estimate of α is 0.498 which was obtained more precisely than can be done looking at a graph using a midpoint algorithm. Recall that solving the correlation function is embedded in the least squares regression. This is why the correlation estimate is the same as that from the LS regression.

4 Estimator Performance

It was interesting to see one example, but to truly evaluate the performance of estimators, we need to do some simulations. Our simulations were done by setting a value for k and α . SPlus generated a size 100 sample of Pareto(k, α) data and estimated k using one of the parameter estimation schemes we discussed earlier. It repeated 1000 times. The mean, median, standard error (SE) of those 1000 estimates were computed. The estimated bias was calculated as the mean minus the true value of the parameter. The mean squared error (MSE) was calculated as the bias squared plus the SE squared. This process was repeated for α using the same estimation scheme and the same values of k and α to generate the random samples. When all the estimation schemes were tested, the entire process was repeated using different values of k and α to generate the data. We tested for values of k at 1, 1,000, and 1,000,000, and for each value of k , we tested for values of α at 0.1, 1, 10, 100, and 1,000. The estimation schemes were GDCC based quantile regression for k (GDqk), GDCC based quantile regression for α (GDq α), Least squares based quantile regression for k (LSqk), Least squares based quantile regression for α (LSq α), method of moments for k (momk), method of moments for α (mom α), maximum likelihood for k (MLk), maximum likelihood for α (ML α), median estimator for α assuming $k = x_{min}$ (med α), and a Pearson's correlation based estimate for α . For the purposes of reporting in this paper, SPlus rounded all results to 3 digits as per its significant digits function. The results of one simulation can be found here. The output for the other fourteen simulations can be found at the end of the paper.

Simulation Statistics for Pareto(1,0.1)

(1, 0.1)	Mean	Median	SE	Bias	MSE
GDqk	1.1200	0.9600	6.44e-01	0.12100	0.429000
GDq α	0.1010	0.1000	1.19e-02	0.00133	0.000144
LSqk	1.1100	0.7780	1.08e+00	0.10500	1.190000
LSq α	0.0971	0.0964	1.29e-02	-0.00293	0.000175
momk	1.1000	1.0600	1.24e-01	0.10400	0.026100
mom α	1.0000	1.0000	1.42e-12	0.90000	0.810000
MLk	1.1100	1.0700	1.23e-01	0.10700	0.026500
ML α	0.1020	0.1010	1.02e-02	0.00183	0.000108
med α	0.1040	0.1020	1.55e-02	0.00354	0.000254

5 Results

Something very fascinating is revealed in looking at all the output. The method of moments approach, despite being totally inappropriate for estimating α , performed consistently best as an estimator for k . There were only two instances in which the GD based quantile regression method was less biased ($k = 1000$ and $\alpha = 100$, $k = 1,000,000$ and $\alpha = 1$) than method of moments and one instance ($k = 1,000,000$ and $\alpha = 0.1$) in which maximum likelihood yielded a smaller MSE.

As for estimating α , the trend seemed to be that the GD based quantile regression estimator was in all cases the least biased estimator and maximum likelihood yielded the least MSE for low values of α while method of moments yielded the least MSE for high values of α .

6 Conclusion

Among the estimators studied, method of moments clearly performed the best for estimating k . While this is interesting theoretically, it could be less interesting practically. On page 242 Johnson and Kotz[5] say that the Pareto distribution serves well as a model for incomes at the extremities, but not as well over the entire range of income levels. We are guessing then that if one were to try to fit a Pareto model to a set of data, he would first through other means determine at what value k the Pareto model takes over and then only be interested in the power parameter α that models the data greater than k . However if one would need to estimate k , we recommend the method of moments estimator.

As for estimating α , the Greatest Deviation based quantile regression showed to be the least biased estimator, but it generally had a 20% greater standard error than maximum likelihood. Maximum likelihood yielded smaller mean squared error than the Greatest Deviation based quantile regression over the parameter values we studied. If one needs an unbiased estimator, we recommend the Greatest Deviation based quantile regression estimator. If small MSE is desired, we recommend maximum likelihood estimation.

Part of the motivation for this project was to develop a parameter estimation scheme for the Pareto distribution that fits into the broader subject of using correlation statistics for parameter estimation and not merely as

descriptive statistics. So far the quantile regression approach has only been applied to symmetric distribution such as Normal, Gosset's t , and Cauchy. We see that correlations, specifically in the case of the Greatest Deviation Correlation Coefficient, can produce good estimates for skewed distributions. The GD quantile regression approach produced estimates for α that were relatively unbiased.

7 Complete Simulation Output

Simulation Statistics for Pareto(1, 1)

(1, 1)	Mean	Median	SE	Bias	MSE
GDqk	1.000	0.999	0.05020	0.00366	0.002530
GDq α	1.000	0.998	0.12200	0.00390	0.014900
LSqk	0.971	0.969	0.08990	-0.02870	0.008900
LSq α	0.962	0.958	0.13100	-0.03790	0.018600
momk	1.000	0.999	0.01030	0.00164	0.000109
mom α	1.190	1.180	0.10100	0.18900	0.045800
MLk	1.010	1.010	0.00949	0.00983	0.000187
ML α	1.020	1.010	0.10400	0.02030	0.011200
med α	1.020	1.010	0.15400	0.02440	0.024400

Simulation Statistics for Pareto(1, 10)

(1, 10)	Mean	Median	SE	Bias	MSE
GDqk	1.000	1.000	0.005000	.000282	2.51e-05
GDq α	10.000	9.960	1.220000	0.033800	1.49e+00
LSqk	0.997	0.997	0.009230	-0.003360	9.64e-05
LSq α	9.670	9.550	1.340000	-0.334000	1.92e+00
momk	1.000	1.000	0.001010	0.000007	1.02e-06
mom α	10.100	10.000	1.040000	0.085700	1.08e+00
MLk	1.000	1.000	0.000914	0.000978	1.79e-06
ML α	10.200	10.100	1.020000	0.202000	1.07e+00
med α	10.300	10.100	1.500000	0.253000	2.31e+00

Simulation Statistics for Pareto(1, 100)

(1, 100)	Mean	Median	SE	Bias	MSE
GDqk	1.0	1.0	5.03e-04	1.80e-05	2.53e-07
GDq α	100.0	99.1	1.23e+01	3.71e-01	1.51e+02
LSqk	1.0	1.0	9.32e-04	-2.85e-04	9.50e-07
LSq α	96.0	94.9	1.27e+01	-4.00e+00	1.78e+02
momk	1.0	1.0	9.33e-05	-9.00e-07	8.71e-09
mom α	101.0	99.7	1.02e+01	6.77e-01	1.05e+02
MLk	1.0	1.0	9.84e-05	1.03e-04	2.03e-08
ML α	102.0	101.0	1.01e+01	2.30e+00	1.08e+02
med α	103.0	102.0	1.48e+01	3.03e+00	2.28e+02

Simulation Statistics for Pareto(1, 1000)

(1, 1000)	Mean	Median	SE	Bias	MSE
GDqk	1	1	4.50e-05	3.00e-06	2.04e-09
GDq α	1010	1000	1.24e+02	1.03e+01	1.56e+04
LSqk	1	1	9.59e-05	-3.17e-05	1.02e-08
LSq α	970	965	1.30e+02	-3.00e+01	1.79e+04
momk	1	1	1.02e-05	0.00e+00	1.03e-10
mom α	1010	1010	1.03e+02	1.27e+01	1.08e+04
MLk	1	1	9.33e-06	9.00e-06	1.68e-10
ML α	1020	1010	1.07e+02	1.72e+01	1.17e+04
med α	1030	1020	1.51e+02	2.65e+01	2.36e+04

Simulation Statistics for Pareto(1000, 0.1)

(1000, 0.1)	Mean	Median	SE	Bias	MSE
GDqk	1.14e+03	9.66e+02	6.72e+02	1.38e+02	4.71e+05
GDq α	1.01e-01	1.00e-01	1.19e-02	8.77e-04	1.43e-04
LSqk	1.13e+03	7.61e+02	1.19e+03	1.34e+02	1.44e+06
LSq α	9.66e-02	9.56e-02	1.32e-02	-3.36e-03	1.87e-04
momk	1.10e+03	1.06e+03	1.13e+02	9.63e+01	2.20e+04
mom α	1.00e+00	1.00e+00	1.42e-12	9.00e-01	8.10e-01
MLk	1.11e+03	1.07e+03	1.33e+02	1.12e+02	3.03e+04
ML α	1.02e-01	1.01e-01	1.02e-02	1.88e-03	1.08e-04
med α	1.03e-01	1.02e-01	1.54e-02	3.29e-03	2.49e-04

Simulation Statistics for Pareto(1000, 1)

(1000, 1)	Mean	Median	SE	Bias	MSE
GDqk	1000.000	997.000	49.500	1.9500	2.46e+03
GDq α	1.000	0.992	0.120	0.0017	1.43e-02
LSqk	974.000	975.000	91.800	-25.6000	9.08e+03
LSq α	0.961	0.950	0.130	-0.0394	1.84e-02
momk	1000.000	998.000	11.100	1.8600	1.27e+02
mom α	1.190	1.180	0.102	0.1900	4.63e-02
MLk	1010.000	1010.000	9.760	9.8200	1.92e+02
ML α	1.020	1.010	0.106	0.0200	1.17e-02
med α	1.030	1.010	0.152	0.0295	2.40e-02

Simulation Statistics for Pareto(1000, 10)

(1000, 10)	Mean	Median	SE	Bias	MSE
GDqk	1000.00	1000.00	4.89	0.0960	23.90
GDq α	10.10	9.96	1.24	0.0937	1.54
LSqk	997.00	998.00	9.33	-2.6300	93.90
LSq α	9.67	9.57	1.36	-0.3310	1.95
momk	1000.00	1000.00	1.04	0.0330	1.07
mom α	10.20	10.10	1.05	0.1810	1.13
MLk	1000.00	1000.00	1.03	0.9630	1.98
ML α	10.20	10.20	1.02	0.2290	1.10
med α	10.30	10.10	1.54	0.3070	2.45

Simulation Statistics for Pareto(1000, 100)

(1000, 100)	Mean	Median	SE	Bias	MSE
GDqk	1000.0	1000.0	0.4870	0.0020	2.38e-01
GDq α	101.0	100.0	12.0000	0.9540	1.46e+02
LSqk	1000.0	1000.0	0.8930	-0.2640	8.66e-01
LSq α	96.2	95.1	12.7000	-3.7900	1.76e+02
momk	1000.0	1000.0	0.0996	-0.0028	9.93e-03
mom α	102.0	101.0	10.1000	1.6300	1.04e+02
MLk	1000.0	1000.0	0.0976	0.1020	1.99e-02
ML α	102.0	102.0	10.0000	2.2800	1.06e+02
med α	103.0	101.0	15.3000	2.6900	2.42e+02

Simulation Statistics for Pareto(1000, 1000)

(1000, 1000)	Mean	Median	SE	Bias	MSE
GDqk	1000	1000	0.0490	0.002	2.41e-03
GDq α	1000	994	126.0000	3.770	1.59e+04
LSqk	1000	1000	0.0951	-0.029	9.88e-03
LSq α	963	955	128.0000	-36.800	1.78e+04
momk	1000	1000	0.0098	0.000	9.61e-05
mom α	1010	999	107.0000	7.490	1.15e+04
MLk	1000	1000	0.0107	0.010	2.14e-04
ML α	1020	1010	102.0000	15.200	1.07e+04
med α	1020	1010	146.0000	19.500	2.17e+04

Simulation Statistics for Pareto(1mil, 0.1)

(1mil, 0.1)	Mean	Median	SE	Bias	MSE
GDqk	1.23e+06	1.00e+06	8.63e+05	2.26e+05	7.96e+11
GDq α	1.00e-01	9.94e-02	1.20e-02	1.67e-04	1.43e-04
LSqk	1.12e+06	7.92e+05	1.11e+06	1.25e+05	1.24e+12
LSq α	9.64e-02	9.56e-02	1.24e-02	-3.56e-03	1.67e-04
momk	1.10e+06	1.06e+06	1.23e+05	1.01e+05	2.53e+10
mom α	1.00e+00	1.00e+00	1.31e-12	9.00e-01	8.10e-01
MLk	1.11e+06	1.07e+06	1.16e+05	1.05e+05	2.44e+10
ML α	1.02e-01	1.02e-01	9.98e-03	2.45e-03	1.06e-04
med α	1.03e-01	1.02e-01	1.55e-02	2.96e-03	2.50e-04

Simulation Statistics for Pareto(1mil, 1)

(1mil, 1)	Mean	Median	SE	Bias	MSE
GDqk	1.00e+06	9.96e+05	4.90e+04	-4.58e+02	2.40e+09
GDq α	1.01e+00	9.98e-01	1.23e-01	9.45e-03	1.52e-02
LSqk	9.75e+05	9.76e+05	8.85e+04	-2.52e+04	8.47e+09
LSq α	9.64e-01	9.53e-01	1.27e-01	-3.62e-02	1.75e-02
momk	1.00e+06	9.98e+05	1.02e+04	1.34e+03	1.06e+08
mom α	1.20e+00	1.19e+00	9.97e-02	1.95e-01	4.81e-02
MLk	1.01e+06	1.01e+06	1.01e+04	1.02e+04	2.07e+08
ML α	1.02e+00	1.01e+00	1.03e-01	1.96e-02	1.09e-02
med α	1.02e+00	1.01e+00	1.48e-01	-1.00e+06	1.00e+12

Simulation Statistics for Pareto(1mil, 10)

(1mil, 10)	Mean	Median	SE	Bias	MSE
GDqk	1.00e+06	1.00e+06	4670.00	2.23e+02	2.18e+07
GDq α	1.00e+01	9.93e+00	1.18	1.79e-02	1.40e+00
LSqk	9.97e+05	9.97e+05	9090.00	-3.31e+03	9.35e+07
LSq α	9.62e+00	9.48e+00	1.32	-3.82e-01	1.89e+00
momk	1.00e+06	1.00e+06	1040.00	-1.29e+01	1.08e+06
mom α	1.01e+01	1.01e+01	1.01	1.21e-01	1.04e+00
MLk	1.00e+06	1.00e+06	942.00	9.82e+02	1.85e+06
ML α	1.02e+01	1.01e+01	1.06	1.85e-01	1.15e+00
med α	1.03e+01	1.01e+01	1.52	2.80e-01	2.40e+00

Simulation Statistics for Pareto(1mil, 100)

(1mil, 100)	Mean	Median	SE	Bias	MSE
GDqk	1.00e+06	1.00e+06	478.0	-12.400	228000
GDq α	1.01e+02	1.00e+02	12.4	0.512	153
LSqk	1.00e+06	1.00e+06	910.0	-281.000	906000
LSq α	9.65e+01	9.57e+01	13.7	-3.470	200
momk	1.00e+06	1.00e+06	99.3	-0.100	9870
mom α	1.01e+02	1.00e+02	10.3	1.020	107
MLk	1.00e+06	1.00e+06	104.0	104.000	21600
ML α	1.03e+02	1.02e+02	10.9	2.510	124
med α	1.03e+02	1.01e+02	15.4	2.680	245

Simulation Statistics for Pareto(1mil, 1000)

(1mil, 1000)	Mean	Median	SE	Bias	MSE
GDqk	1000000	1000000	47.90	-0.30	2290.0
GDq α	1010	997	122.00	6.03	15000.0
LSqk	1000000	1000000	97.50	-28.50	10300.0
LSq α	974	962	125.00	-26.50	16400.0
momk	1000000	1000000	9.60	-0.20	92.1
mom α	1010	1000	103.00	11.00	10800.0
MLk	1000000	1000000	9.42	10.00	189.0
ML α	1020	1020	105.00	21.80	11500.0
med α	1030	1010	146.00	25.00	22100.0

References

- [1] Rudy A. Gideon, *Generalized Interpretation of Pearson's r* , unpublished paper (URL: <http://www.math.umt.edu/gideon/Generalized%20CC.pdf>)
- [2] Rudy A. Gideon, *Random Variables, Regression, and the Greatest Deviation Correlation Coefficient*, unpublished paper (URL: <http://www.math.umt.edu/gideon/SLR%20theory%20revision.pdf>)
- [3] Rudy A. Gideon, *Correlation in Simple Linear Regression*, unpublished paper (URL: <http://www.math.umt.edu/gideon/CORR-N-SPACE-REG.pdf>)
- [4] Rudy A. Gideon, *Location and Scale Estimation with Correlation Coefficients*, unpublished paper (URL: <http://www.math.umt.edu/gideon/locscale.pdf>)
- [5] Norman L. Johnson and Samuel Kotz, *Distributions in Statistics Continuous Univariate Distributions - 1*, Houghton Mifflin Company, Boston, 1970