Adamawa State University Journal of Scientific Research
Volume 7 Number 1, April, 2019; Article no. ADSUJSR 0701013
ISSN:2705-1900(Online); ISSN: 2251-0702 (Print)
http://www.adsujsr.com

Some Statistical Inference for Pareto Distribution under Simple Random Sampling in the presence of Outliers

Paul Inuwa Dalatu[1,*] and Osama. A. A. Alsattari[2]
[1]Department of Mathematics, Faculty of Science, Adamawa State University, Mubi-Nigeria
[2]University of Palestine, Al-Zahra City-Gaza Strip-Palestine
Contact: dalatup@gmail.com,: +2348132490488

Abstract

Pareto is a continuous distribution which has several applications in real life fields including the extreme value events. Hence, in this paper, the maximum likelihood (ML) and method of moment (MOM) estimations are proposed to estimate the shape parameter estimates for Pareto Type I distribution. The BIAS, Mean Square Error (MSE), and the Standard Error (SE) of the shape parameter estimates for both the two methods were obtained based on different simulated sample sizes. The scale parameter of Pareto distribution is treated as a fixed value. In this paper, the data is contaminated with 5% and 10% of outliers to investigate the effect of the outliers on the parameter estimates. The results show that when the sample size was large, the parameter estimates were more reliable and accurate. However, in the presence of outliers, the method of moment is better than the maximum likelihood estimation evident by having the smallest standard errors.

Keywords: Pareto distribution; Maximum likelihood; Outliers; Contamination; Estimation.

Introduction

Vilfredo Pareto, (born July 15, 1848, Paris, France-died August 19, 1923, Geneva, Switzerland) (Arnold, 2015). Vilfredo Pareto was an Italian economist and sociologist who was known for his theory on mass and elite interaction as well as for his application of Mathematics to economic analysis (Chhetri et al., 2017). In 1906, Vilfredo Pareto (Dunford et al., (2014); Bhatti et al., (2019)), introduced the concept of Pareto Distribution when he observed that 20% of the pea pods were responsible for 80% of the peas planted in his garden. He related this phenomenon to the nature of wealth distribution in Italy, and found that 80% of the country`s wealth was owned by about 20% of its population. However, this was described in terms of land ownership. The Italian observed that 80% of the land was owned by a handful of wealthy citizens, who comprised about 20% of the population.

The practical applications of the Pareto Distributions are enormous in many fields like; Business Management, Company`s Revenues, Employee Evaluation, Insurance, Survival Analysis, Computer Science, Economics, Better Decisions, Biomedical, Financial Risk Management, etc., (see details in Mohamed et al., (2018); Alzaatreh et al., (2012); Thiago et al., (2017)).

The probability density function of Pareto distribution according to (Ihtisham et al., 2019) can be written as follows:

$$f(x;\alpha,\beta)=\frac{\beta\alpha^{\beta}}{x^{\beta+1}} \qquad (1)$$

Where, $0<\alpha<x$ and $\beta>0$ are the scale and the shape parameters respectively, X is the random variable such that $x>0$. In this study, the scale parameter is constant and equals 1. As a consequence, the Model (1) above becomes

$$f(x;1,\beta)=\frac{\beta}{x^{\beta+1}} \qquad (2)$$

The mean of the random variable comes from Pareto distribution with fixed scale parameter can be calculated as:

$$E(X)=\int_0^{\infty} x\, f(x;1,\beta)dx=\int_0^{\infty} x\frac{\beta}{x^{\beta+1}}dx=\beta\int_0^{\infty}\frac{x}{x^{\beta+1}}dx=\beta\int_0^{\infty}x^{-\beta}\,dx.$$

$$E(X) = \beta\left(\frac{x^{-\beta+1}}{-\beta+1}\right)\Big|_0^\infty = \frac{\beta}{\beta-1}, \beta > 1.$$

As well as, the variance of the Model (2) above is given as follows:

$$var(X) = \frac{\beta}{(\beta-1)^2(\beta-2)}, \beta > 2.$$

The cumulative density function (CDF) of Model (2) above is provided as follows:

$$F(x;\alpha,\beta) = 1-(x)^{-\beta} \qquad (3)$$

The behavior of Pareto distribution is determined by the shape parameter β. Hence, it is considered as the most important quantity in Pareto distribution. This distribution is usually known as the Pareto distribution, and we will soon relate it to the Pareto principle. However, this term also known and refers to as related truncated distribution (Huang et al., 2013).

The social sciences have found that the Pareto distribution embodies a useful power law. The Pareto Distribution is most often presented in terms of its survival function, which gives the probability of seeing larger values than x. (This is often known as the complementary CDF, since it is 1-CDF. It is sometimes called the reliability function or the tail function). The survival functions of a Pareto distribution.

In the simulation of the random variable experiment, select the Pareto distribution. Vary the shape parameter and note the shape and location of the density function. For selected values of the parameter, run the simulation 1000 times with an update frequency of 10 and note the apparent convergence of the empirical density to the true density.

The aim of this paper is to propose maximum likelihood and method of moment estimators to estimates for Pareto Type I Distribution. However, the BIAS, Mean Square Error and Standard Error of the shape parameter estimates for both methods based on different simulated sample sizes were obtained, where the results were visualized in Tables (1-3) and Figures (1-3).

Pareto distributions in firm size and income occur under very limited assumptions on the distribution of underlying primitives. Unlike in previous theories, large firms or incomes can appear instantaneously and result from an arbitrarily small level of ex ante heterogeneity. In contrast, economists' current understanding of why Pareto distributions emerge falls into two categories. The first theory works through a "transfer of power law". One Pareto distribution can be explained by assuming that some other variable is distributed according to a Pareto distribution; for example, entrepreneurial skills (Charles, 2015), firm productivities (Gómez-Déniz and Calderin-Ojeda (2014); Perla and Tonetti (2014)), or firm size by Benhabib and Bisin (2018). A functional form for the production function also needs to be assumed, which preserves the Pareto functional form, such as a power function. The second theory holds that Pareto distributions result from a dynamic, proportional, "random growth" process, following Luckstead and Devadoss (2014) law. In this theory, many firms or incomes are large because they have been hit by a long and unlikely continued sequence of good idiosynchratic shocks (Ihtisham et al., (2019); Atkinson and Voitchovsky (2010); Arnold (2014); Akinsete et al., (2016); or Luttmer (2007)).

The Beta-exponentiated Pareto distribution (BEP) was investigated by Zea et al. (2012), using Mahdavi and Kundu (2017), method providing different shapes for the density and hazard functions. The BEP distribution has several sub-models. The MLE method was used to estimate the parameters and to derive the observed information matrix. A bladder cancer data set was utilized to illustrate the flexibility of the proposed model. The simplicity of these two models, and the fact that Hassan et al., (2018) model was not purposefully developed to generate Pareto distributions, in fact suggest that there might be something more general about joint production with complementarities that leads to Pareto generating production functions.

Materials and Methods

Maximum Likelihood Estimation

The likelihood estimation method is one of the common methods which is used for estimating the parameters. In this particular Pareto distribution, the likelihood function is (Ihtisham et al., 2019):

$$L(\beta; x_1, \dots, x_n) = \prod_{i=1}^{n} \frac{\beta}{x_i^{\beta+1}} = \beta^n \prod_{i=1}^{n} x_i^{-(\beta+1)} \qquad (4)$$

As a result, the log-likelihood is obtained as:

$$\log(L) = n\log(\beta) - (\beta+1)\sum_{i=1}^{n} \log(x_i) \qquad (5)$$

By using the partial derivative with respect to β, it gives

$$\frac{n}{\beta} - \sum_{i=1}^{n} \log(x_i) \qquad (6)$$

Set the Model (6) above to be zero and solve it for β to obtain the MLE of β

$$\frac{n}{\hat{\beta}} - \sum_{i=1}^{n} \log(x_i) = 0 \Leftrightarrow \hat{\beta} = \frac{n}{\sum_{i=1}^{n} \log(x_i)} = \frac{1}{\left(\frac{\sum_{i=1}^{n} \log(x_i)}{n}\right)} = \frac{1}{mean(\log(x_i))} \qquad (7)$$

Method of Moment

Let $X_1, X_2, \dots, X_n$ be a random sample from a Pareto distribution with probability density function (Gómez-Déniz and Calderin-Ojeda (2014), $f(x; 1, \beta) = \frac{\beta}{x^{\beta+1}}$.

It can be observed that the mean of Pareto distribution with scale parameter equals 1 is $\frac{\beta}{\beta-1}$. Hence, by setting the mean of Pareto distribution to be equal to the sample mean $\bar{x}$ we obtained

$$\bar{x} = \frac{\hat{\beta}}{\hat{\beta}-1} \Leftrightarrow \bar{x}(\hat{\beta}-1) = \hat{\beta} \Leftrightarrow \bar{x}\hat{\beta} - \bar{x} = \hat{\beta} \Leftrightarrow \hat{\beta} = \frac{\bar{x}}{\bar{x}-1} \qquad (8)$$

In this paper, different sample sizes are generated from Pareto distribution with parameters $\alpha = 1,\ \beta = 3$ and the shape parameter estimates are obtained and compared.

The BIAS, Mean Square Error and Standard Error

BIAS is an estimator in finite samples. In general, BIAS is computed as in (Kosmidis, (2014); Zavershynski, (1017)):

$$\text{BIAS} = E(\hat{\theta}) - \theta, \qquad (9)$$

where θ is population mean and $\hat{\theta}$ is its estimator. Note: We say that $\hat{\theta}$ is unBIASed estimator if $Bias(\hat{\theta}) = 0.$ While, the Mean Square Error is defined as in (Kosmidis, (2014); Zavershynski, (2017)):

$$MSE(\hat{\theta}) = E\left[(\hat{\theta}-\theta)^2\right] \qquad (10)$$

$$= Bias^2(\hat{\theta}, \theta) + Var(\hat{\theta}).$$

The sample variance is one estimator of σ^2. It is defined as in (Kosmidis, (2014); Chernick, M. R., (2012)):

$$E[S^2] = E\left[\frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2\right] = \sigma^2 \qquad (11)$$

$$= \frac{n-1}{n}\sigma^2,$$

where $(x_1, x_2, x_3, \dots, x_n)$ are the samples and $\bar{x}$ is the sample mean. Therefore, the square root of the variance is called the Standard Error, denoted $SE(\hat{\theta})$.

Results and Discussion

The R Studio version 3.4.3 (2018) was used for simulation study to compare the performances of both MLE and MOM methods for data without outliers and in the presence of outliers with both percentages of 5% and 10%. The parameters α, β are set equal to 1 and 3 as the true parameters of Pareto model. In the presence of outliers, the simulated data is contaminated from uniform distribution $unif(0,70)$. The simulation for each sample size involves a total of 1000 replications. The two estimation methods such as MLE and MOM were then applied to the data. The outcomes of simulation study are summarized in Tables (1-3). It can be observed that with the increase in the percentage of outliers, the standard errors of the preceding methods decrease for various sample sizes. It can be seen that the BIAS of both MLE and MOM increases with the increase in the percentage of outliers and MOM is less affected in the presence of outliers. The effect of outliers on the standard errors of the parameter estimates is displayed in Figures 1-3. It can be observed that MLE is less affected for clean data. However, in the presence of outliers, MOM is better than MLE evident by having the smallest standard errors.

Table 1: BIAS, MSE, and standard errors of the parameter estimates for clean data, without outliers

SAMPLE SIZE	METHOD	ESTIMATES	BIAS	MSE	SE
50	MLE	3.0607	0.0607	0.2096	0.4537
	MOM	3.1011	0.1011	0.2387	0.4780
100	MLE	3.0222	0.0222	0.0971	0.3109
	MOM	3.0445	0.0445	0.1109	0.3301
200	MLE	3.0161	0.0161	0.0467	0.2155
	MOM	3.0293	0.0293	0.0551	0.2330
400	MLE	3.0106	0.0106	0.0240	0.1545
	MOM	3.0213	0.0213	0.0294	0.1701
500	MLE	3.0104	0.0104	0.0170	0.1298
	MOM	3.0194	0.0194	0.0209	0.1432
1000	MLE	3.0093	0.0093	0.0091	0.0949
	MOM	3.0134	0.0134	0.0123	0.1102

Table 2: BIAS, MSE and standard errors of the parameter estimates for data with 5% outliers

SAMPLE SIZE	METHOD	ESTIMATES	BIAS	MSE	SE
50	MLE	2.1691	0.8309	0.7502	0.2446
	MOM	1.4635	1.5365	2.4068	0.2141
100	MLE	2.0055	0.9945	1.0133	0.1557
	MOM	1.3526	1.6474	2.7222	0.0915
200	MLE	2.0053	0.9947	1.0011	0.1084
	MOM	1.3421	1.6579	2.7516	0.0551
400	MLE	1.9995	1.0005	1.0067	0.0761
	MOM	1.3363	1.6637	2.7693	0.0375
500	MLE	2.0002	0.9998	1.0043	0.0686
	MOM	1.3367	1.6633	2.7678	0.0343
1000	MLE	2.0009	0.9991	1.0007	0.0487
	MOM	1.3356	1.6644	2.7710	0.0233

Table 3: BIAS, MSE and standard errors of the parameter estimates or data with 10% outliers

SAMPLE SIZE	METHOD	ESTIMATES	BIAS	MSE	SE
50	MLE	1.5132	1.4868	2.2298	0.1392
	MOM	1.1943	1.8057	3.2639	0.0576
100	MLE	1.5035	1.4965	2.2486	0.0949
	MOM	1.1879	1.8121	3.2848	0.0340
200	MLE	1.4985	1.5015	2.2587	0.0649
	MOM	1.1851	1.8149	3.2942	0.0232
400	MLE	1.5010	1.4990	2.2493	0.0471
	MOM	1.1835	1.8165	3.2999	0.0156
500	MLE	1.4981	1.5019	2.2574	0.0424
	MOM	1.1830	1.8170	3.3016	0.0142
1000	MLE	1.4977	1.5023	2.2577	0.0287
	MOM	1.1822	1.8178	3.3045	0.0098

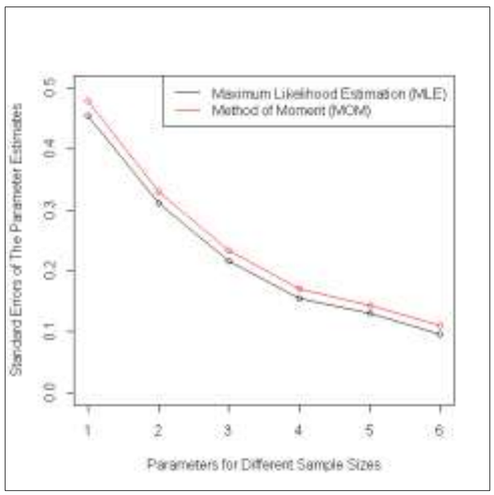

Figure 1: The effect of 0% *outliers* on the *standard errors* of the *parameter estimates* for different sample sizes.

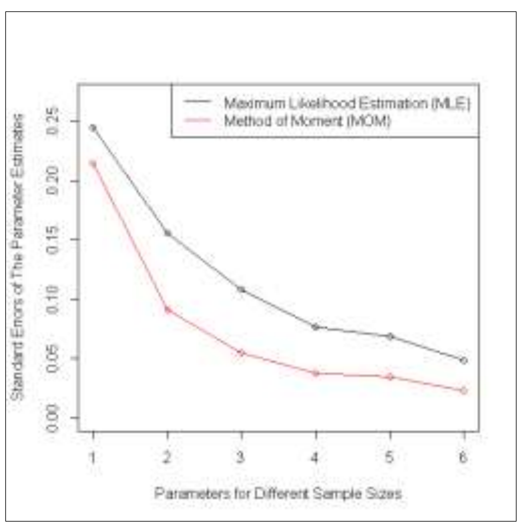

Figure 2: The effect of 5% *outliers* on the *standard errors* of the *parameter estimates* for different sample sizes

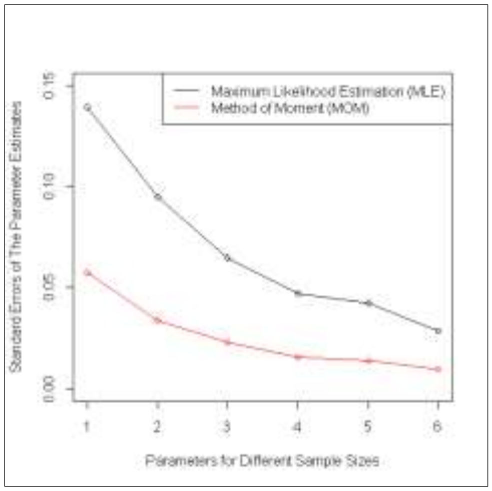

Figure 3: The effect of 10% *outliers* on the *standard errors* of the *parameter estimates* for different sample sizes.

Conclusion

The main objective of this paper is to compare the estimates of the shape parameter for both MLE and MOM in the presence of leverage points in the simulated Pareto distribution's data set. The data is divided into two types namely, data without outliers and data with outliers. It can be observed that for clean data, the MLE performance is better than the MOM. Nevertheless, for data with outliers, the MOM outperforms the MLE evident by having the smallest standard error estimates for all sample sizes.

References

Akinsete, A., Famoye, F. and Lee, C. (2016). The Beta-Pareto Distribution. *Taylor and Francis, (6): 547—563.*

Arnold, B. C. (2014). Statistical Distributions and Applications. *Wiley Statistical Reference, 1:1-16.*

Arnold, B. C. (2015). Pareto Distributions. *Monographs on Statistics and Applied Probability, 2nd Edition, Chapman and Hall/CRC Press, Taylor and Francis*.

Atkinson, A. B., and Voitchovsky, S. (2010). The distribution of top earnings in the UK since the Second World War. *Economica, Volume 78: 440-459.*

Ayman Alzaatreh, Felix Famoye, and Carl Lee. (2012). Gamma-Pareto Distribution and Its Applications. *Journal of Modern Applied Statistics Methods, 11(1): 1-18.*

Charles I. Jones. (2015). Pareto and Piketty: The Macroeconomics of Top Incomes and Wealth Inequality. *Journal of Economic Perspective, 29(1): 29-46.*

Chernick, M. R. (2012). The Jackknife connections to the bootstrap. *Wiley Interdicip Rev Comput Stat, 4: 224-226.*

Emilio Gómez-Déniz and Enrique Calderin-Ojeda. (2014). A suitable Alternative to the Pareto Distribution. *Hacettepe Journal of Mathematics and Statistics, 43(5): 843-860.*

Fiaz Ahmad Bhatti, G. G. Hamedian, Wenhui Sheng and Munir Ahmad. (2019). Cubic Rank Transmuted Modified Burr III Pareto Distribution: Development, Properties, Characterizations and Applications. *International Journal of Statistics and Probability, 8(1): 1-19.*

Hassan, A. S., Mohamd, R. E., Elgarhy, M. and Fayomi, A. (2018). Alpha power transformed extended exponential distribution: properties and applications. *Journal of Nonlinear Sciences and Applications, 12(4): 62-67.*

Ihtisham, S., Khalil, A., Manzoor, S., Khan, S.A., and Ali, A. (2019). Alpha-Power Pareto distribution: Its properties and applications. *PLoS ONE, 14(6), E0218027:1-11.*

Jesse Benhabib and Alberto Bisin. (2018). Skewed Wealth Distribution: Theory and Empirics. *Journal of Economic Literature, 56(4): 1261-1291.*

Loannis Kosmidis. (2014). BIAS in parametric estimation: reduction and useful side-effects.

WIREs Computational Statistics, 2014, 6:185-196.

Luckstead, J. and Devadoss, S. (2014). A nonparametric analysis of the growth process of Indian Cities. *Econom. Lett. , 124(3): 516-519.*

Luttmer, Erzo G.J. (2007). Selection, Growth, and the Size Distribution of Firms. *Quarterly Journal of Economics, 122 (3): 1103–1144.*

Mahdavi, A. and Kundu, D. (2017). A new method for generating distributions with an application exponential distribution. *Communications in Statistics-Theory and Methods, 46(13):6543-6557.*

Mohamed, E. Ghitany; Emilio, Gómez-Déniz and Saralees, Nadarajah. (2018). A New Generalization of the Pareto Distribution and Its Application to Insurance Data. *Journal of Risk and Financial Management, 11, 10:1-14.*

Maksym Zavershynski. (2017). MSE and BIAS-Variance Decomposition. *Towards Data Science, 2017.7:1-8.*

Mei Ling Huang, Vincenzo Coia, and Percy Brill. (2013). A Cluster Truncated Pareto Distribution and Its Applications. *Hindawi Publishing Corporation, ISRN Probability and Statistics, Volume 2013: 10 pages.*

Perla, Jesse and Christopher Tonetti, (2014). Equilibrium Imitation and Growth. *Journal of Political Economy, 122 (1): 1–21.*

Rosie Dunford, Quanrong Su, Ekraj Tamang and Abigail Wintour. (2014). The Pareto Principle. *The Plymouth Student Scientist, 7(1): 140-148.*

R Studio version 3.4.3. (2018). http://www.rstudio.org/

Sher B. Chhetri, Hongwei Long, and Gokarna Aryal. (2017). The Beta Transmuted Pareto Distribution: Theory and Applications. *Journal of Statistics Applications and Probability, 6(2): 243-258.*

Thiago A. N. de Andrade, Luz Milena Zea Fernande, Frank Gomes-Silva and Gauss M. Cordiero. (2017).The Gamma Generalized Pareto Distribution with Applications in Survival Analysis. *International Journal of Statistics and Probability, 6(3): 1-16.*

Zea, L. M., Silva, R. B., Bourguignon, M., Santos, A. M., and Cordeiro, G. M. (2012). The Beta Exponentiated Pareto Distribution with Application to Bladder Cancer Susceptibility. *International Journal of Statistics and Probability, 21(2): 3-8.*