Parameter Estimation for the Double Pareto Distribution

Faris M. Al-Athari
Department of Mathematics,
Faculty of Science and Information Technology,
Zarqa University, Zarqa 13132, Jordan

Abstract: Problem statement: The double Pareto distribution appeared most often as model for variety of fields, including archaeology, biology, economics, environmental science, finance and physics. The distribution exhibits Paretian power-law behavior in both tails. The family of double Pareto distributions has recently been proposed for modeling growth rates such as annual gross domestic product, stock prices, foreign currency exchange rates and company sizes. In this study, I develop parameter estimates for the double Pareto distribution that are easy to compute. I compare the performance of the maximum likelihood estimate with Bayesian and the method of moments estimates. **Approach:** This study contracted with maximum likelihood, the method of moments and Bayesian using Jeffrey's prior and the extension of Jeffrey's prior information. The comparisons are made on the performance of these estimators with respect to the Mean Squared Error (MSE) for small, moderate and large samples and for some values of the scale and the extension of Jeffrey's prior parameters using the simulation techniques. **Results:** It turns out that the maximum likelihood method and Bayesian method with Jeffrey's prior result in smaller MSE compared to others in all cases. **Conclusion:** Based on the results of the simulation, the maximum likelihood method and Bayesian method with Jeffrey's prior are found to be the best with respect to MSE.

Key words: Bayesian method, modified Jeffrey's prior, maximum likelihood, moments estimators

INTRODUCTION

The probability density function and the cumulative distribution function of the symmetric double Pareto distribution are Eq. 1 and 2:

$$f(x;\theta,\beta) = \frac{\theta}{2\beta}\begin{cases} \left(\dfrac{x}{\beta}\right)^{\theta-1} & \text{for } 0 < x < \beta \\ \left(\dfrac{\beta}{x}\right)^{\theta+1} & \text{for } x \geq \beta \end{cases} \qquad (1)$$

And:

$$F(x;\theta,\beta) = \begin{cases} \dfrac{1}{2}\left(\dfrac{x}{\beta}\right)^{\theta}, & \text{for } 0 \leq x < \beta \\ 1 - \dfrac{1}{2}\left(\dfrac{\beta}{x}\right)^{\theta} & \text{for } x \geq \beta \end{cases} \qquad (2)$$

For some $\theta > 0$ and $\beta > 0$ respectively (Reed, 2001).

In practice, the double Pareto distribution has been widely used as a model for growth rates and size distributions of various phenomena arising in different areas of inquiry such as economics, physical Sciences, geology, geography and finance (Lee *et al.*, 1998; Madan *et al.*, 1998; Kozubowiski and Podgoriski, 2001; Amaral *et al.*, 1998; Buldyrev *et al.*, 1997; Stanley *et al.*, 1996; Takayasu and Okuyama, 1998; Reed, 2001).

The double Pareto distribution arises as exponential function of double exponential distribution and can be derived by combining the Pareto distribution and the distribution of the reciprocal of a Pareto random variable and has power tail behavior at zero and infinity (Reed, 2001; Kotz *et al.*, 2001).

The double Pareto distribution has been proposed as a model for heavy-tailed phenomena. Heavy-tailed distributions are very important in modeling finance (Embrechts *et al.*, 1997; Fama, 1965; Jansen and Varies, 1991; Loretan nand Phillips, 1994; Mandelbrot, 1963; Madfal and Raw, 1996; Rachev, 2003; Rachev and Mittnik, 2000), Physics (Barkai *et al.*, 2000), engineering (Nikias and Shao, 1995; Resnick, 1997; Resnick and Starica, 1995; Uchaikin and Zolotarev, 1999). Burroughs and Tebbens (2001a; 2001b) estimated parameters of the truncated distribution by least squares fitting on a probability plot and by minimizing mean squared error fit on a plot of the tail

distribution function. Ramirez *et al.* (2008) used a mixture of K Pareto distributions to model Ethernet packets interarrival times. Reed and Jorgensen (2004) considered the double Pareto lognormal distribution as model for heavy-tailed data and introduced various approaches to inference this distribution. Cobo *et al.* (2010) developed an algorithm to implement Bayesian inference for the double Pareto lognormal distribution. Al-Athari (2011) considered the double translated Pareto distribution as model for infinite excess kurtosis data and as model for non-kurtosis data and proposed various robust confidence intervals to inference this distribution. Ahmed *et al.* (2010), the maximum likelihood estimation, Bayesian using Jeffrey's prior and the extension of Jeffrey's prior for estimating the parameters of Weibull distribution are compared using the mean squared error and the mean percentage error. Ahmed and Ibrahim (2011) compared the performance of maximum likelihood and Bayesian using Jeffrey's prior and the extension of Jeffrey's prior for estimating the survival function of Weibull distribution with right censored data.

In this study, I develop parameter estimates for the double Pareto distribution that are easy to compute. The Maximum Likelihood Estimation (MLE), the method of Moments Estimation (MM) and Bayesian using Jeffrey's prior (BAJ) and extension of Jeffrey's Prior Information (BAEJ) are developed. The comparisons are made on the performance of these estimators with respect to the Mean Squared Error (MSE) for varying sample sizes and for some values of the scale parameter and the extension of Jeffrey's prior using the simulation techniques.

These methods should be useful for the practitioners in area of science and engineering where power law probability tails are prevalent.

MATERIALS AND METHODS

Maximum likelihood estimation: Let $X_1, X_2,\ldots X_n$ be a random sample of size n taken from the double Pareto distribution given in Eq. 1. Then the likelihood and the log-likelihood functions for the double Pareto density can be written respectively as Eq. 3 and 4:

$$L(\theta) = 2^{-n}\theta^n (\prod_{i=1}^{n} x_i)^{-1} \exp(-\theta \sum_{i=1}^{n} |\ln x_i - \ln\beta|) \qquad (3)$$

And:

$$\ln L(\theta) = -n\ln 2 + n\ln\theta - \ln\prod_{i=1}^{n} x_i \\ -\theta\sum_{i=1}^{n} |\ln x_i - \ln\beta| \qquad (4)$$

where, the scale parameter β is taken to be constant. Maximizing Eq. 3 or Eq. 4 we get the maximum likelihood estimator $\hat{\theta}_m$ for θ. It follows that

$$\partial\ln L(\theta)/\partial\theta = n/\theta - \sum_{i=1}^{n} |\ln x_i - \ln\beta| \ .$$

Setting $\partial\ln L(\theta)/\partial\theta = 0$ we get the equation $n/\hat{\theta}_m - \sum_{i=1}^{n} |\ln x_i - \ln\beta| = 0$ and hence the MLE of $\theta, \hat{\theta}_m$ is given by Eq. 5:

$$\hat{\theta}_m = \frac{n}{\sum_{i=1}^{n} |\ln x_i - \ln\beta|} \qquad (5)$$

Under the regularity conditions (Johnson, 1996; Verbeek, 2008; Zacks, 1971), this estimator possesses the major properties of the maximum likelihood estimator, that is $\hat{\theta}_m$ is consistent, asymptotic efficient and best asymptotically normal estimator with mean θ and asymptotic variance, $\mathrm{avar}(\hat{\theta}_m)$ attains the Cramer Rao lower bound.

Method of moments estimator: Given data assumed to be from the double Pareto distribution given in Eq. 1 one could obtain method of moments estimates of θ and β using the first two moments of the log-transformed data distribution. Use of the double Pareto moments (with untransformed data) is not recommended, since the population moments of order θ or greater do not exist. So, we need to prove the following Theorems

Theorem 1: If X has double Pareto distribution given by Eq. 1, then the density of Y= ln X is a shifted double exponential with mean in β and scale $\frac{1}{\theta}$.

Proof: The p.d.f. of Y is easily found from (1). It can be expressed as:

$$g(y;\theta,\beta) = e^y f(e^y;\theta,\beta)$$

With some algebra, the p.d.f. of Y can be shown to be:

$$g(y;\theta,\beta) = \frac{\theta}{2} \begin{cases} e^{\theta(y-\ln\beta)}, & \text{for } y < \ln\beta \\ e^{-\theta(y-\ln\beta)}, & \text{for } y \geq \ln\beta \end{cases}$$

which is equal to double exponential with mean in β and scale $\frac{1}{\theta}$.

It is easy to show that the moment generating function (mgf) of the shifted double exponential distribution is:

$$M_Y(t) = \frac{\beta^t}{1 - t^2/\theta^2} \text{ for } |t| < \theta$$

And the variance of Y is:

$$Var(Y) = \frac{2}{\theta^2} \tag{6}$$

To find the method of moment estimator using the shifted double exponential distribution, one needs only solve Eq. 6, with Var (Y) set to its sample equivalent and gets:

$$\hat{\theta}_{mm} = \sqrt{\frac{2n}{\sum_{i=1}^{n}(Y_i - \overline{Y})^2}} \tag{7}$$

where, Y=ln X. The estimator in Eq. 7 is independent of the parameter β.

Bayesian estimation: Given a random sample X_1, X_2,….,X_n from the double Pareto distribution given in Eq. 1, a Bayes' estimator is a decision function $\hat{\theta}$ that minimizes the conditional expectation of the loss function $\ell(\theta, \hat{\theta})$ given that $X_1 = x_1, X_2 = x_2,..., X_n = x_n$ and, accordingly, it minimizes the expected risk function. By using a quadratic loss function $\ell(\theta, \hat{\theta}) = (\theta - \hat{\theta})^2$, then the Bayes' estimator is given by $\hat{\theta}_B = E(\theta | x_1, x_2,..., x_n)$, the mean of the conditional distribution of θ, given $X_1 = x_1, X_2 = x_2,..., X_n = x_n$. This conditional distribution $K(\theta | x_1, x_2,..., x_n)$ is called a posterior distribution of θ.

The goal is to compute the posterior distribution which depends on the likelihood function, the prior Π (θ) and the marginal probability density function.

Jeffrey's prior information: Consider the likelihood function L (θ) given by Eq. 3 and its Fisher information $I(\theta) = E(\frac{\partial \ln L(\theta)}{\partial \theta})^2 = \frac{n}{\theta^2}$.

Jeffrey's (Jeffreys, 1946) suggested $\Pi(\theta) \propto \sqrt{I(\theta)} = \frac{\sqrt{n}}{\theta}$ as a prior distribution for the parameter θ. Then from standard Bayesian theory (Box and Tiao, 1992; Hogg and Craig, 1994), the posterior distribution is proportional to the joint distribution L (θ)Π(θ) that is $K(\theta | x_1, x_2,..., x_n) \propto L(\theta).\Pi(\theta)$ with the possibility of dropping all functions involving constants and X_1, X_2,….,X_n alone (not θ) from the right-hand member of this expression.

This posterior distribution is proper provided that θ>0 and β>0 with the constant of proportionality turning out to be:

$$[\Gamma(n)]^{-1} . (\sum_{i=1}^{n}|\ln X_i - \ln\beta|)^n$$

That is:

$$K(\theta | x_1, x_2,..., x_n) = \frac{\theta^{n-1}(\sum_{i=1}^{n}|\ln X_i - \ln\beta|)^n}{\Gamma(n)} \text{ x}$$

$$\left(e^{-\theta \sum_{i=1}^{n}|\ln X_i - \ln\beta|} \right)$$

Which is gamma p.d.f. with parameters a = n and $\beta^* = \frac{1}{\sum_{i=1}^{n}|\ln X_i - \ln\beta|}$.

Assuming the quadratic loss function, the Bayes' estimation $\hat{\theta}_{Bj}$ is the mean of the posterior distribution. Thus:

$$\hat{\theta}_{Bj} = \frac{n}{\sum_{i=1}^{n}|\ln X_i - \ln\beta|}$$

This estimator is the same as the maximum likelihood estimator $\hat{\theta}_m$ given in Eq. 5 and is independent of the parameter β.

Extension of Jeffrey's prior information: The extension of Jeffrey's prior is given by:

$$\Pi(\theta) \propto [I(\theta)]^c = \frac{n^c}{\theta^{2c}}$$

where, c is positive constant (Al-Kutubi and Ibrahim, 2009). Therefore, the posterior distribution is proper provided that θ>0, β>0 and n>2c-1 with the constant of proportionality turning out to be:

$$[\Gamma(n - 2c + 1)]^{-1} . (\sum_{i=1}^{n}|\ln X_i - \ln\beta|)^{n-2c+1}$$

And hence:

$$K(\theta \mid x_1, x_2,...,x_n) = \frac{\theta^{n-2c} \left(\sum_{i=1}^{n} \mid \ln X_i - \ln \beta \mid\right)^{n-2c+1}}{\Gamma(n-2c+1)} \; x$$

$$X\left(e^{-\theta \sum_{i=1}^{n} \mid \ln X_i - \ln \beta \mid}\right)$$

Which is gamma p.d.f. with parameters $\alpha = n-2c+1$ and $\beta^* = \dfrac{1}{\sum_{i=1}^{n} \mid \ln X_i - \ln \beta \mid}$.

Assuming the quadratic loss function, the Bayes' estimation $\hat{\theta}_{BE}$ is the mean of the posterior distribution. Thus:

$$\hat{\theta}_{BE} = \frac{n-2c+1}{\sum_{i=1}^{n} \mid \ln X_i - \ln \beta \mid}$$

RESULTS

A computer simulation experiment was done to compare four methods of estimation of the parameter of the double Pareto distribution. Simulations were performed for sample sizes n= 10, 25, 50, 100 with values of the parameter $\theta = 0.5 (1.5)5$, $\beta = 1.0$ and values of Jefferys extension c = 0.4, 3.0. Without losing the generality I assumed β equals to 1.0. For each combination of values of n,θ, c and β 40000

samples were simulated from the uniform (0, 1) and then used to generate random samples from the double Pareto distribution by using the probability transform. For each sample, the parameter was estimated by the proposed four methods of estimation and then the mean- squared errors and the estimated means of the parameter were calculated to compare the estimation methods with the help of MATLAB, the language of technical computing version 6.5 (Part-Enander, 1996).

The transformation to the double Pareto distributed variable is given by:

$$X_i = F^{-1}(U_i) = \begin{cases} \beta(2U_i)^{1/\theta} & \text{for } 0 < U_i < 1/2 \\ \beta / [2(1-U_i)]^{1/\theta} & \text{for } U_i > 1/2 \end{cases}$$

Where:
F (.) = The distribution function given in Eq. 2
U_i = Uniformly distributed random variable on (0,1)

The simulation results for the mean-squared errors and the estimated means of the parameter for each estimation of (a) the Maximum Likelihood method (ML) (b) the Method of Moments (MM) (c) Bayesian with Jeffrey's prior (BAJ) and (d) Bayesian with Extension Jeffrey's prior (BAEJ) are summarized and tabulated in Table 1.

Table 1: The estimated mean and the mean-squared error of the estimators

n	θ	c	MLE	Men	MM	Mean	BAEJ	Mean
10	0.5	0.4	0.0413	0.5553	0.0614	0.5977	0.0442	0.5664
		3.0	0.0413	0.5553	0.0614	0.5977	0.0590	0.2776
	2.0	0.4	0.6614	2.2212	0.9827	2.3907	0.7078	2.2656
		3.0	0.6614	2.2212	0.9827	2.3907	0.9442	1.1106
	3.5	0.4	2.0255	3.8870	3.0097	4.1837	2.1675	3.9648
		3.0	2.0255	3.8870	3.0097	4.1837	2.8916	1.9435
	5.0	0.4	4.1338	5.5529	6.1422	5.9768	4.4235	5.6639
		3.0	4.1338	5.5529	6.1422	5.9768	5.9012	2.7764
25	0.5	0.4	0.0125	0.5216	0.0164	0.5381	0.0129	0.5258
		3.0	0.0125	0.5216	0.0164	0.5381	0.0145	0.4173
	2.0	0.4	0.1997	2.0864	0.2624	2.1523	0.2060	2.1031
		3.0	0.1997	2.0864	0.2624	2.1523	0.2325	1.6692
	3.5	0.4	0.6116	3.6513	0.8035	3.7666	0.6308	3.6805
		3.0	0.6116	3.6513	0.8035	3.7666	0.7120	2.9210
	5.0	0.4	1.2482	5.2161	1.6398	5.3808	1.2873	5.2578
		3.0	1.2482	5.2161	1.6398	5.3808	1.4531	4.1729
50	0.5	0.4	0.0056	0.5107	0.0071	0.5190	0.0057	0.5127
		3.0	0.0056	0.5107	0.0071	0.5190	0.0061	0.4596
	2.0	0.4	0.0897	2.0428	0.1138	2.0758	0.0911	2.0509
		3.0	0.0897	2.0428	0.1138	2.0758	0.0972	1.8385
	3.5	0.4	0.2746	3.5748	0.3484	3.6327	0.2791	3.5891
		3.0	0.2746	3.5748	0.3484	3.6327	0.2978	3.2173
	5.0	0.4	0.5603	5.1069	0.7110	5.1895	0.5695	5.1273
		3.0	0.5603	5.1069	0.7110	5.1895	0.6077	4.5962
100	0.5	0.4	0.0027	0.5053	0.0034	0.5095	0.0027	0.5063
		3.0	0.0027	0.5053	0.0034	0.5095	0.0028	0.4801
	2.0	0.4	0.0431	2.0213	0.0539	2.0381	0.0434	2.0254
		3.0	0.0431	2.0213	0.0539	2.0381	0.0448	1.9203
	3.5	0.4	0.1320	3.5373	0.1651	3.5666	0.1331	3.5444
		3.0	0.1320	3.5373	0.1651	3.5666	0.1373	3.3605
	5.0	0.4	0.2693	5.0533	0.3370	5.0952	0.2715	5.0634
		3.0	0.2693	5.0533	0.3370	5.0952	0.2802	4.8007

DISCUSSION

Table 1 show that the Maximum Likelihood Method (MLE) and Bayesian method with Jeffrey's prior (BAJ) give, for all values of sample size and parameters, the closest estimated mean estimator for θ. In other words the maximum likelihood method and Bayesian method with Jeffrey's prior have the lowest absolute bias. The estimated mean value of Bayesian method with extension Jeffrey's prior (BAEJ) is closer to the parameter than that of the Method of Moments (MM) when the parameter c is small otherwise it is farther. Table 1 show also that, for all values of sample size and parameters, the mean-squared errors of the maximum likelihood and Bayesian with Jeffrey's prior followed by the method of moments estimators are less than that of Bayesian with extension Jeffrey's prior estimator. This suggests that the maximum likelihood method and Bayesian method with Jeffrey's prior are the best followed by the method of moments a result which was supported partially by (Ahmed and Ibrahim, 2011; Ahmed *et al.*, 2010). When the sample size increases the mean-squared error decreases for all cases but for any fixed value of n, the mean-squared error increases as the parameter θ increases for all cases. Finally, the mean-squared error for the Bayesian method with extension Jeffrey's prior increases as the value of the extension parameter c increases.

CONCLUSION

Estimations of the shape parameter of the symmetric double Pareto distribution have been studied in this research. It turns out that the maximum likelihood estimator and Bayesian estimator with Jeffrey's prior are equivalent and the best compared to the moments and Bayesian with extension Jeffrey's prior estimators. The method of moment's estimator gives better results than Bayesian estimator with extension Jeffrey's prior.

The mean-squared error decreases as long as the sample size increases.

The Maximum Likelihood Estimator (MLE) is like Bayesian estimator with Jeffrey's prior has several advantages. It is always exist, fast and straightforward to compute and yields feasible values and under regularity conditions it is consistent, asymptotic efficient and best asymptotically normal with mean θ and asymptotic variance attains the Cramer Rao lower bound.

ACKNOWLEDGEMENT

REFERENCES

Ahmed, A.O.M. and N.A. Ibrahim, 2011. Bayesian Survival Estimator for weibull Distribution with censored Data. J. Applied Sci., 11: 393- 396.

Ahmed, A.O.M.Al Omari, M.A., H.S. Al-Kutubi and N.A. Ibrahim, 2010. Comparison of the bayesian and maximum likelihood estimation for weibull distribution. J. Math. Statis., 6: 100-104. DOI: 10.3844/jmssp.2010.100.104

Al-Athari, F.M., 2011. Confidence interval for locations of non-kurtosis and large kurtosis leptokurtic symmetric distributions. J. Applied Sci., 11: 528-534.

Al-Kutubi, H.S. and N.A. Ibrahim, 2009. Bayes estimator for exponential distribution with extension of Jeffery prior information. Malaysian J. Math. Sci., 3: 297-313.

Amaral, L.A.N., S.V. Buldyrev, S. Havlin, M.A. Salinger and H.E. Stanley, 1998. Power law scaling for a system of interacting units with complex internal structure. Physical Rev. Lett., 80: 1385-1388.

Barkai, E., R. Metzler and J. Klafter, 2000. From continuous time random walks to the fractional fokker-planck equation. Phys. Rev. E Stat. Phys. Plasmas Fluids Relat Interdiscip Topics, 61: 132-138. PMID: 11046248

Box, G.E.P. and G.C. Tiao, 1992. Bayesian Inference in statistical Analysis. 1st Edn., Wiley, New York, ISBN: 0471574287, pp: 588.

Buldyrev, S.V., L.A.N. Amaral, S. Havlin, H. Leschorn and P. Mass *et al.*, 1997. Scaling behavior in economics: II. Modeling of company growth. J. Phys. I (France), 7: 635-650.

Burroughs, S.M. and S.F. Tebbens, 2001a. Upper-truncated power law distributions. Fractals, 9: 209-222.

Burroughs, S.M. and S.F. Tebbens, 2001b. Upper-truncated power laws in natural systems. J. Pure Applied Geophysics, 158: 741-758.

Cobo, P.R., R.E. Lillo, S. Wilson and M.P. Wiper, 2010. Bayesian inference for double pareto lognormal queues. Ann. Applied Stat., 4: 1533-1557. DOI: 10.1214/10-AOAS336

Embrechts, P., C. Kluppelberg and T. Mikosch, 1997. Modelling Extremal Events for Insurance and Finance. 1st Edn., Sipringer-Verlag, Berlin, ISBN: 3540609318, pp: 645.

Fama, E.F., 1965. The behavior of stock market prices. J. Bus., 38: 34-105.

Hogg, R.V. and A. Craig, 1994. Introduction to Mathematical Statistics. 5th Edn., Prentice-Hall Inc., New Jersey, ISBN-10: 0023557222, pp: 576.

Jansen, D.W. and C.G.D. Varies, 1991. On the frequency of large stock market returns: Putting booms and busts into perspective. Rev. Econ. Stat., 73: 18-24.

Jeffreys, H., 1946. An invariant form for the prior probability in estimation problems. Proc. R. Soc. Lond. A., 186: 453-461. DOI: 10.1098/rspa.1946.0056

Johnson, J., 1996. Econometric Methods. 4th Edn., McGraw-Hill, New York, ISBN: 0070327203, pp: 480.

Kotz, S., T.J. Kozubowski and K. Podgorski, 2001. The Laplace Distribution and Generalizations: A Revisit with Applications to Communications, Economics, Engineering, and Finance. 1st Edn., Birkhauser, Boston, ISBN: 0817641661, pp: 349.

Kozubowiski, T.J. and K. Podgoriski, 2001. Asymmetric Laplace laws and modeling financial data. Math. Comput. Model., 34: 1003-1021. DOI: 10.1016/S0895-7177(01)00114-5

Lee, Y., L.A.N. Amaral, D. Canning, M. Meyer and H.E. Stanley, 1998. Universal features in the growth dynamics of complex organizations. Phys. Rev. Lett., 81: 3275- 3278.

Loretan, M. and P.C.B. Phillips, 1994. Testing the covariance stationarity of heavy-tailed time series: An overview of the theory with applications to several financial datasets. J. Empirical Finance, 1: 211-248. DOI: 10.1016/0927-5398(94)90004-3

Madan, D.B., P.P. Carr and E.C. Chang, 1998. The variance gamma process and option pricing. Eur. Finance Rev., 2: 79-105. DOI: 10.1023/A:1009703431535

Madfal, G.S. and C.R. Raw, 1996. Statistical Methods in Finance. 1st Edn., Elsevier, Amsterdam, ISBN: 0444819649, pp: 733.

Mandelbrot, B., 1963. The variation of certain speculative prices. J. Bus., 36: 394-419.

Nikias, C.L. and M. Shao, 1995. Signal Processing with Alpha-Stable Distributions and Applications. 1st Edn., Wiley, New York, ISBN: 047110647X, pp: 168.

Part-Enander, 1996. The MATLAB Handbook. 1st Edn., Addison Wesley, Longman, ISBN: 0201877570, pp: 423.

Rachev, S.T. and S. Mittnik, 2000. Stable Paretian Models in Finance. 1st Edn., Wiley, Chichester, UK., ISBN: 0471953148, pp: 855.

Rachev, S.T., 2003. Handbook of Heavy Tailed Distributions in Finance. 1st Edn., Elsevier, New York, ISBN: 0444508961, pp: 680.

Ramirez, P., R.E. Lillo and M. Wiper, 2008. Bayesian analysis of a queueing system with a Long-tailed arrival process. Comm. Stat. Simulat. Comput., 37: 697-712.

Reed, W.J. and M. Jorgensen, 2004. The double Pareto-lognormal distribution-a new parametric model for size distributions. Comm. Statist. Theory Meth., 33: 1733-1754.

Reed, W.J., 2001. The Pareto, Zipf and other power laws. Econ. Lett., 74: 15-19. DOI: 10.1016/S0165-1765(01)00524-9

Resnick, S. and C. Starica, 1995. Consistency of hill's estimator for dependent data. J. Applied Probability, 32: 139-167.

Resnick, S.I., 1997. Heavy tail modeling and teletrafic data. Ann. Stat., 25: 1805-1869.

Stanley, M.H.R., L.A.N. Amaral, S.V. Buldyrev, S. Havlin and H. Leschorn *et al.*, 1996. Scaling behaviour in the growth of companies. Nature, 397: 804-806.

Takayasu, H. and K. Okuyama, 1998. Country dependence on company size distributions and a numerical model based on competition and cooperation. Fractals, 6: 67-69. DOI: 10.1142/s0218348x98000080

Uchaikin, V.V. and V.M. Zolotarev, 1999. Chance and Stability: Stable Distributions and Their Applications. 1st Edn., VSP, Utrecht, ISBN: 9067643017, pp: 570.

Verbeek, M., 2008. Guide to Modern Econometrics. 3rd Edn., John Wiley and Sons, New York, ISBN: 9780470517697, pp: 472.

Zacks, S., 1971. The theory of Statistical Inference. 1st Edn., John Wiley and Sons, New York, ISBN: 0471981036, pp: 609.