EUROPEAN ACADEMIC RESEARCH
Vol. VI, Issue 12/ March 2019

ISSN 2286-4822
www.euacademic.org

Impact Factor: 3.4546 (UIF)
DRJI Value: 5.9 (B+)

Lorenz Curve and Gini Coefficient in Right Truncated Pareto's Income Distribution

EHTESHAM HUSSAIN
Department of Statistics, University of Karachi, Pakistan
MUHAMMAD AHSANUDDIN
Department of Economics, University of Karachi, Pakistan
MASOOD UL HAQ
Usman Institue Technology, Karachi, Pakistan

Abstract

Gini's coefficient of concentration has been studied extensively for the study of dispersion of the variable. Geometrically, it is related to the area enclosed by the Lorenz Curve and the Distribution function of the uniform distribution. The study of the Lorenz Curve for some symmetric and some skewed distributions has been done algebraically and graphically. The distribution which has been selected is Right Truncated Pareto's Income Distribution. The Gini's Index has been obtained by calculating the relevant area for different parameters.

Keywords: Lorenz Curve, Gini Coefficient, Right Truncated Pareto's Income Distribution

1. INTRODUCTION

The concept of spread, variability or inequality has been discussed in the economics and statistical literature in various forms (Atkinson, 1970; Allison, 1978; Acemoglu, 2003; Yitzhaki & Schechtman, 2005; Williams & Doessel, 2006). Analytically to measure the spread, variability or inequality in random variables of interest (e.g. income) several measures have been suggested, namely, standard deviation, mean deviation, Gini's mean difference etc. (Fellman, 2012). To measure relative variability their coefficients e.g. coefficient of variation, Gini coefficient etc. are also used and they are in terms of the parameter(s) of the underlying distribution. In order to measure variability in the variable of interest geometrically Lorenz curves are used (Arnold, 2001; Fellman, 2018).

In this paper we will discuss Lorenz curve, coefficient of variation and Gini coefficient, (which is related to Lorenz curve) associated with the truncated two parameter Pareto distribution (Arnold, 2015). Usually left truncated Pareto distribution is used as income distribution (Chernobai et al., 2015), in this form the Lorenz curve and Gini coefficient are independent of location parameter, but coefficient of variation depends on both parameters. While in right truncated Pareto distribution, these three measures depend on both parameters.

It is suggested that, to measure the variability, right truncation is important in Pareto distribution because it is related to real situations in the formulation of meaningful indices.

2. DEFINITIONS

Consider a non-negative random variable X with density function $f(x)$, distribution function $F(x)$ and finite mean μ. The moment distribution function of X, $F_1(x)$ is defined by Stuart and Ord (1987).

$$p = F(x) = \int_o^x f(x)dx \qquad 0<p<1 \tag{1}$$

$$L\ (p) = F_1(x) = \mu^{-1}\int_o^x x\ \mathrm{f}(x)\ dx \qquad \mu\neq 0 \tag{2}$$

$$\mu = E(X) = \int_0^\infty f(x)dx \tag{3}$$

The graph of $L(p)$ as a function of p which we call as the Lorenz function, is the population Lorenz curve on the unit interval [0,1],(see Fig.1). If $L(p)=p$, then there is complete equality and if $L(p)>p$ then disparity appears i.e. variability in the distribution of X exists.

Twice the area between the Lorenz curve and the line $L(p) = p$ is called the Gini coefficient and define as:

$$G = 1 - 2\int_0^1 L(p)\,dp \qquad 0<G<1 \qquad (4)$$

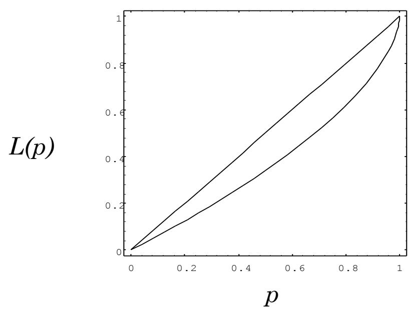


Fig-1. Lorenz curve

3. RIGHT TRUNCATION

Corresponding to $F(x)$, let $F_k(x)$ denote the distribution function of the truncated distribution, truncated on the right at $x=k$, where k is an element of the support {o $<F(x)$ <1}. More specifically we have right truncated distribution function

$$F_k(x) = \begin{cases} 0 & x \langle a \\ \dfrac{F(x)}{F(k)} & a \langle x \langle k \\ 1 & x \le k \end{cases}$$

where

$F_x(k)=Pr.\ (X \le k)$

4. LORENZ CURVE IN TRUNCATED DISTRIBUTION

The Lorenz curve $L_k(p)$ of the truncated distribution is given by Moothathu (1985a)

$$L_k(p) = \frac{L\,(p\,F(k))}{L\,(F\,(k))} \qquad (5)$$

For applications of Lorenz curve and Gini coefficient see Goldie (1977), Sarabia (2008), Moothathu (1985b) and references cited therein. Moothathu (1985a) also contains references to the applications of the Pareto distribution, Lorenz curve and Gini coefficient.

5. PARETO DISTRIBUTION

The Pareto distribution with its density function also known as income distribution is

$$f(x) = va^{v}x^{-v-1} \qquad x \geq a > 0, v > 0$$

where a is location parameter and v is shape or inequality parameter. Using (1), (2), (3), (4) and (5), the following results obtained

$$\text{Mean}: E(x) = \frac{av}{(v-1)} \quad , \quad v > 1$$

$$\text{Gini coefficient}: G(v) = \frac{1}{(2v-1)} \quad , \quad v > 1$$

$$\text{Standard deviation}: \sigma = \frac{a}{(v-1)}\sqrt{\frac{v}{(v-2)}} \quad , \quad v > 2$$

$$\text{Lorenz Function}: L(p) = \{1-(1-p)^{1-\frac{1}{v}}\} \quad , \quad v \geq 1$$

In Pareto distribution, measures of variability cannot be used for $1<v\leq 2$, since σ does not exists in this case.

From the above properties we see that mean, and standard deviation are functions of both parameters (*a, v*) but Lorenz curve and Gini coefficient are only the function of parameter *v* and no contribution of *a*.

6. RIGHT TRUNCATED PARETO DISTRIBUTION

The right truncated Pareto distribution and related measures are as follows:

$$f(x) = \frac{va^{v}}{F(k)} x^{-v-1}, \qquad a<x<k\,a,\ v>0$$

$$= 0 \qquad \text{elsewhere}$$

where

$$F(k) = \left\{1-\left(\frac{a}{k}\right)^{v}\right\}, \qquad k \geq a$$

Mean: $$E_k(X) = \frac{av}{F(k)(v-1)}\left(1-\left(\frac{a}{k}\right)^{v-1}\right)$$

Standard deviation: $$\sigma_k = \sqrt{\frac{a^2 v}{F(k)(v-2)}\left(1-\left(\frac{a}{k}\right)^{v-1}\right) - E(X)^2}$$

Gini coefficient:

$$G_k(a,v) = 1 - \frac{2}{\left(1-\left(\frac{a}{k}\right)^{v-1}\right)} + \frac{2v}{(2v-1)} \frac{\left(1-\left(\frac{a}{k}\right)^{2v-1}\right)}{\left(F(k)\left(1-\left(\frac{a}{k}\right)^{v-1}\right)\right)}$$

Lorenz curve: $$L_k(p;a,v) = \frac{\left\{1-\left(1-F(k)p\right)^{1-1/v}\right\}}{\left\{1-\left(\frac{a}{k}\right)^{v-1}\right\}}$$

From above properties we see that mean, standard deviation, Lorenz curve and Gini coefficient are function of both parameters *(a, v).*

7. NUMERICAL ILLUSTRATION

To see the effect of right truncation, a comparison is made by drawing Lorenz curves and computing mean, standard deviation and Gini coefficient. Many combinations of parameters *(a, v)* have been studied and it has been found that for comparing variability/ inequality in income distribution Right truncated distribution should be effective, since the location parameter *a* also has its contribution while in Pareto it is not. Due to lack of space, results for selected parameters *(a, v),* given in Table 1, are shown in following figures and tables

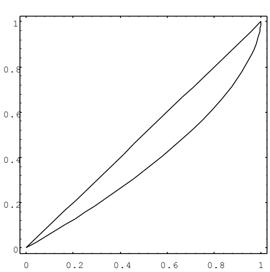
Table I: Parameters and Truncated Points

a	*v*	*k*
5	2.5	100
25	2.5	100

Table-I has set of parameters *(a, v)* and *k*. To draw Lorenz curves and required computations, Mathematica System is used (Wolfram, 1991).

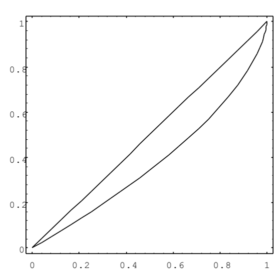
Lorenz Curves for Pareto's Income distribution and Truncated to right Pareto's Income distribution.

Lorenz Curve



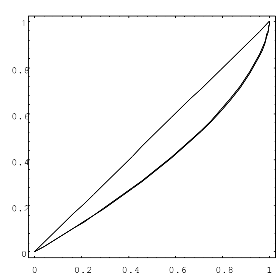
a=5, v=2.5
G=0.2500

Truncated Lorenz Curve

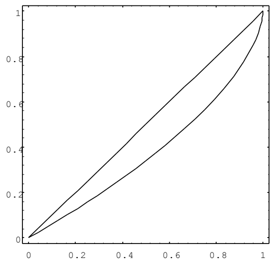


a=5, v=2.5, k=100
G=0.2422

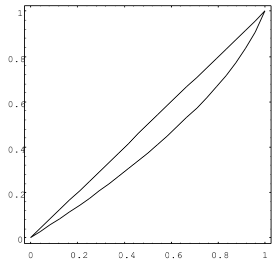
Combine Lorenz Curve



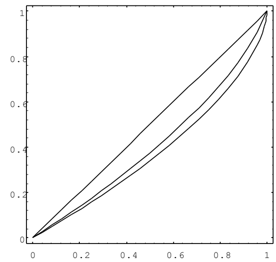
a=5, v=2.5, k=100
G=0.2500



a=25, v=2.5
G=0.2500



a=25, v=2.5, k=100
G=0.1832



a=25, v=2.5, k=100
$G>G_k$

Table II: Comparison μ, σ and G with a=5 and v =2.5

Parameters	Summary Results	Pareto distribution	Right Truncated Pareto distribution k=100
	μ	8.33333	8.24477
a=5, v=2.5	σ	7.45356	5.39696
	G	0.25000	0.242219

Table III: Comparison μ, σ and G with a=25 and v=2.5

Parameters	Summary Results	Pareto distribution	Right Truncated Pareto distribution k=100
	μ	41.66667	37.2678
a=25, v=2.5	σ	37.6344	14.0198
	G	0.25000	0.18318

Similar results are obtained by computation. Table-II contains results computed for a=5, v=2.5 verses a=25, v=2.5 and k=100 which show there is no much significance difference in mean, standard deviation and Gini coefficient. The results appear in Table-III show that if a is increased to a=25 and the remaining parameters are the same, there is a significant difference in mean, standard deviation and Gini coefficient, and reveals a reduction in variability.

8. CONCLUSION

The study of the Lorenz Curve for some symmetric and some skewed distributions are shown algebraically and graphically in this study. It is found that the distribution thus obtained by calculating the relevant area for different parameters is Right Truncated Pareto's Income Distribution which is true for real situations in the formulation of meaningful indices.

REFERENCES

1. Acemoglu, D. (2003). Cross-country inequality trends. *The Economic Journal*, *113*(485), 121-149.
2. Allison, P. D. (1978). Measures of inequality. *American sociological review*, 865-880.
3. Arnold, B. C. (2001). The Lorenz curve: evergreen after 100 years. In *Advances on income inequality and concentration measures* (pp. 34-46). Routledge.
4. Arnold, B. C. (2015). *Pareto distributions*. CRC press.
5.Atkinson, A. B. (1970). On the measurement of inequality. Journal of economic theory, 2(3), 244-263.
6. Chernobai, A., Rachev, S. T., &Fabozzi, F. J. (2015). Composite goodness-of-fit tests for left-truncated loss samples. Handbook of financial econometrics and statistics, 575-596.
7. Fellman, J. (2012) Estimation of Gini Coefficients Using Lorenz Curves. Journal of Statistical and Econometric Methods, 1, 31-38.
8. Fellman, J. (2018). Income Inequality Measures. Theoretical Economics Letters, 8, 557-574.
9. Goldie, C. M. (1977). Convergence Theorems for Empirical Lorenz Curves and their Inverses, Advances in Applied Probability. 9, 765-791.
10. Moothathu, T. S. K. (1985a). Sampling distributions of Lorenz curve and Gini index of the Pareto distribution. Sankhyā: The Indian Journal of Statistics, Series B, 247-258.

11. Moothathu, T. S. K. (1985b). Distribution of Maximum Likelihood Estimators of Lorenz Curve and Gini Index of Exponential Distribution. Ann. Inst. Statist. Math., 37, 473-479.

12. Sarabia, J. M. (2008). Parametric Lorenz curves: Models and applications. In Modeling income distributions and Lorenz curves (pp. 167-190). Springer, New York, NY.

13. Stuart, A. and Ord, J. K. (1987). Kendalls Advanced Theory of Statistics, Vol. I, fifth edition, CharlesGriffin, London.

14. Wolfram, S. (1991). Mathematica: A System for Doing Mathematics by Computer, 2nd ed.Redwood City: C. A. Addison-Wesley.

15. Williams, R. F., & Doessel, D. P. (2006). Measuring inequality: tools and an illustration. International journal for equity in health, 5(1), 5.

16. Yitzhaki, S., & Schechtman, E. (2005). The properties of the extended Gini measures of variability and inequality.