ISSN: 2319-5967
ISO 9001:2008 Certified
International Journal of Engineering Science and Innovative Technology (IJESIT)
Volume 1, Issue 2, November 2012

A Novel Concept of Partial Lorenz Curve and Partial Gini Index

Sudesh Pundir and Rajeswari Seshadri
Department of Statistics Pondicherry University, Puducherry – 605014, INDIA
Department of Mathematics, Pondicherry University, Puducherry – 605014, INDIA

***Abstract:** Lorenz curve and Gini index are the famous tools of income inequality. When one wants to study income inequality of a class of individuals, Partial Lorenz Curve is introduced and the corresponding measure of inequality called Partial Gini Index is defined. Partial Gini Index (PGI) is derived on Partial Lorenz Curve (PLC). Partial orderings are defined for comparison of partial Lorenz curves. For illustration purpose, real life data set of per capita Net State Domestic Product (NSDP) at current prices of States and Union territories of India are taken for the year 1998-99. Partial Gini Index is evaluated for three different classes of community and the results are discussed comparing with the Gini Index of the entire community. It is observed that Partial Gini Index gives more insight into the income inequality distribution between different classes of society.*

Keywords: **Lorenz curve, Gini index, Partial Lorenz Curve, Partial Gini Index.**

I. INTRODUCTION

Lorenz Curve is the most frequently used device to describe and compare income inequality in Economics. Lorenz curve is not only used in Economics field but is also used in reliability, survival analysis, medicine and many other fields [1]-[13]. Let income X of a unit be non negative random variable with probability distribution function (pdf) f(x) and the cumulative distribution function (cdf) of f(x) is given by

$F(x)=\int_0^x f(u)du$

Which can be interpreted as the proportion of units having income less than or equal to x. Let $F_1(x)$ be the proportion of income received by individual with income less than or equal to x assuming that $E(X) = \mu$ exists and is nonzero. The first moment distribution function of x is given as

$F_1(x) = \frac{1}{\mu}\int_0^x uf(u)du$

where $0 < F_1(x) < 1$. $F_1(x)$ is monotonically non decreasing function of x.
Generally Lorenz curve is expressed as
$L(p)= F_1(x)$
where $p = F(x)$, $0 \le p \le 1$.
Lorenz curve is the mathematical relationship between F(x) and $F_1(x)$ expressed graphically in general as in Fig. 1.

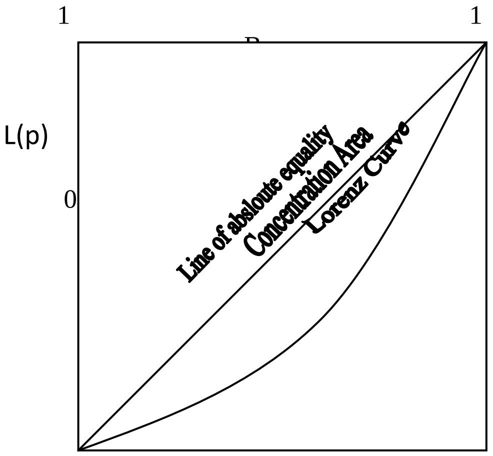


Fig. 1 Lorenz Curve

Lorenz Curve (LC) is defined as

ISSN: 2319-5967
ISO 9001:2008 Certified
International Journal of Engineering Science and Innovative Technology (IJESIT)
Volume 1, Issue 2, November 2012

$L(p) = \frac{1}{\mu}\int_0^p F^{-1}(t)dt$

where $\mu = \int_0^1 F^{-1}(t)\,dt$ is the mean of distribution

and $F^{-1}(p) = \inf\{x : F(x) \geq p\},\ 0 \leq p \leq 1.$

Gini index is defined as

$GI = 1 - 2\int_0^1 L(p)\,dp.$

Properties of Lorenz curve

1. The Lorenz curve L(p) is entirely contained into a square of side 1 units, because p is defined over [0, 1] and the the value of L(p) is also within [0, 1]. Both x- axis and the y- axis are percentages.
2. The Lorenz curve is not defined if the mean is either 0 or ∞.
3. The Lorenz curve is a continuous function and it lies always below the 45^0 line or equal to it.
4. Lorenz curve is strictly increasing convex function of p. Its first derivative

$\frac{d\,L(p)}{dp} = \frac{q(p)}{\mu} = \frac{x}{\mu}$ with $x = F^{-1}(p)$

is always positive as incomes are positive and so is its second order derivative. (Mathematically, a curve is convex when its second order derivative is positive).

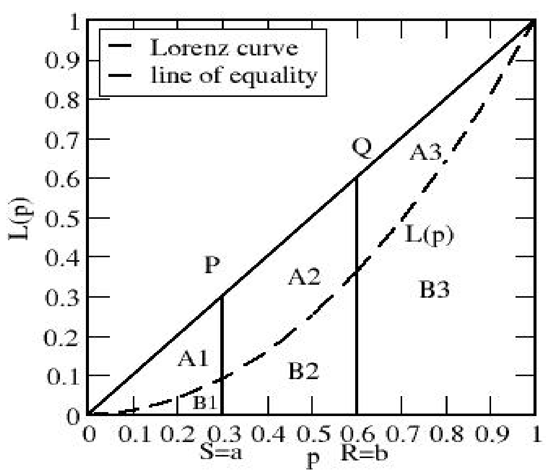
In section 2, partial Lorenz curve and partial Gini index are discussed. In section 3, partial Gini index is computed by using three methods of numerical integration viz. Trapezoidal rule, Simpson's 1/3 and 3/8 rule. In section 4, confidence intervals for partial Gini index are derived. In section 5, partial Lorenz ordering is defined so that the two partial Lorenz curves can be compared by partial Gini index. In section 6, a real life data set of per capita Net State Domestic Product (NSDP) at current prices of 31 States and Union territories of India for year 1998-99 is considered for illustration purposes.

II. PARTIAL LORENZ CURVE AND ITS GINI INDEX

When it is required to study income inequality among a portion or a class of society (say poor class, middle class, rich class), one needs to study only a portion of Lorenz curve and to measure the inequality of this portion, the measure of inequality called Partial Gini Index is defined. The advantage of defining partial Lorenz curve and Gini Index is that when we talk about social welfare of the society, it is more meaningful if we take care of that part of the society which needs more attention in terms of measuring income inequality.

Graphically, the partial area under Lorenz curve in interval [a,b] is shown in Fig 2.1.

Fig. 2.1 Lorenz Curve



The Partial area under the Lorenz curve (PAC) is defined as

$$PAC_{[a,b]} = \int_a^b L(p)\,dp \qquad (1)$$

and the Partial Gini Index (PGI) is defined as

ISSN: 2319-5967
ISO 9001:2008 Certified
International Journal of Engineering Science and Innovative Technology (IJESIT)
Volume 1, Issue 2, November 2012

$$PGI_{[a,b]} = \frac{A2}{A2+B2} = \text{Area of PQRS}\nabla\text{ Area under the LC in [a,b]}$$

$$= 1 - \frac{\text{PAC in the interval [a,b]}}{\text{Area of PQR}\nabla}$$

$$= 1 - \frac{2\int_a^b L(p)dp}{(a+b)(b-a)} \quad (2)$$

Note: Partial Gini Index on any [a,b] tends to Gini Index over [0,1] when a→ 0 and b →1.

III. NUMERICAL INTEGRATION TECHNIQUES FOR FINDING PARTIAL GINI INDEX

The expression for Partial Gini Index is derived by using trapezoidal rule, Simpson's 1/3 rule and 3/8 rule.

i) **Trapezoidal rule**

The area under the Lorenz curve in the interval [a,b] is given as

$$\int_a^b L(p)dp = \int_{a=p_0}^{p_1} L(p)dp + \int_{p_1}^{p_2} L(p)dp + \ldots\ldots\ldots\ldots + \int_{p_{n-1}}^{p_n} L(p)dp$$

$$= \frac{(b-a)}{n}\left[\frac{L(p_0)+L(p_n)}{2} + L(p_1)+\ldots\ldots+L(p_{n-1})\right] \quad (3)$$

On substituting (3) in (2), we get

$$PGI = 1 - \frac{1}{n(b+a)}[L(a)+L(b)+2(L(p_1)+L(p_2)+\ldots+L(p_{n-1}))]$$

$$= 1 - \frac{L(a)+L(b)}{n(b+a)} - \frac{2\sum_{i=1}^{n} L(pi)}{n(b+a)} \quad (4)$$

ii) **Simpson's 1/3 rule**

Using this rule, the area under the Lorenz curve is given as

$$\int_a^b L(p)dp = \frac{(b-a)}{3n}[(L(a)+L(b))+4(L(p_1)+\ldots\ldots+L(p_{n-1}))+2\{L(p_2)+\ldots\ldots\ldots+L(p_{n-2}\}] \quad (5)$$

Substituting (5) in (2), we get

$$PGI = 1 - \frac{2(L(a)+L(b))}{3n(b+a)} - \frac{8}{3n(b+a)}[L(p_1)+\ldots\ldots L(p_{n-1})] - \frac{4}{3n(b+a)}[L(p_2)+\ldots+L(p_{n-2})] \quad (6)$$

iii) **Simpson's 3/8 rule**

$$\int_a^b L(p)dp = \frac{3(b-a)}{8n}[(L(a)+L(b)) + 3(L(p_1)+\ldots+L(p_{n-1}))+2\{L(p_3)+\ldots+L(p_{n-3}\}] \quad (7)$$

Substituting (7) in (2), we get,

$$PGI = 1 - \frac{3(L(a)+L(b)}{4n(b+a)} - \frac{9}{4n(b+a)}[L(p_1)+\ldots L(p_{n-1})] - \frac{3}{2n(b+a)}[L(p_3)+\ldots\ldots L(p_{n-3})] \quad (8)$$

In the next Section Lower limit and upper limit for the Partial Gini Index are derived.

IV. LOWER LIMIT AND UPPER LIMIT FOR PARTIAL GINI INDEX

i) Lower limit of Partial Gini Index

To find the area under the Lorenz curve, instead of calculating the area as sum of trapeziums, if we approximate the area of each shape by inner rectangles, then

$$\int_a^b L(p)dp = \frac{(b-a)}{n}\{L(p_0)+\ldots\ldots+L(p_{n-1})\} \quad (9)$$

ii) Upper limit of Partial Gini Index

To find the area under the Lorenz curve, instead of calculating the area as sum of trapeziums, if we approximate the area of each shape by inner rectangles, then

$$\int_a^b L(p)dp = \frac{(b-a)}{n}\{L(p_1)+\ldots+L(p_n)\} \quad (10)$$

On substituting (9) and (10) in (2), PGI is given as

$$PGI_{LL} = 1 - \frac{2}{n(b+a)}\{L(p_0)+\ldots\ldots+L(p_{n-1})\} \quad (11)$$

ISSN: 2319-5967
ISO 9001:2008 Certified
International Journal of Engineering Science and Innovative Technology (IJESIT)
Volume 1, Issue 2, November 2012

$$PGI_{UL} = 1 - \frac{2}{n(b+a)}[L(p_1)+\ldots\ldots+L(p_n)] \quad (12)$$

Hence, one can easily see that expressions in (4), (6) and (8) lies between (11) and (12).

V. PARTIAL LORENZ ORDERING

The distribution F is said to have a less inequality than the distribution G in the partial Lorenz sense if the partial Lorenz Curve (PLC) for F dominates the partial Lorenz curve for G , i.e.

$PLC_F(p) \geq PLC_G(p)$

or $1 - \frac{2\int_a^b L_F(p)}{(b+a)(b-a)} \leq_{PL} 1 - \frac{2\int_a^b L_G(p)}{(b+a)(b-a)}$

or $PGI_F \leq PGI_G$

It is concluded that partial Lorenz curve for F dominates partial Lorenz curve for G if Partial Gini Index for F is less than Partial Gini Index of G.

VI. REAL LIFE ILLUSTRATION

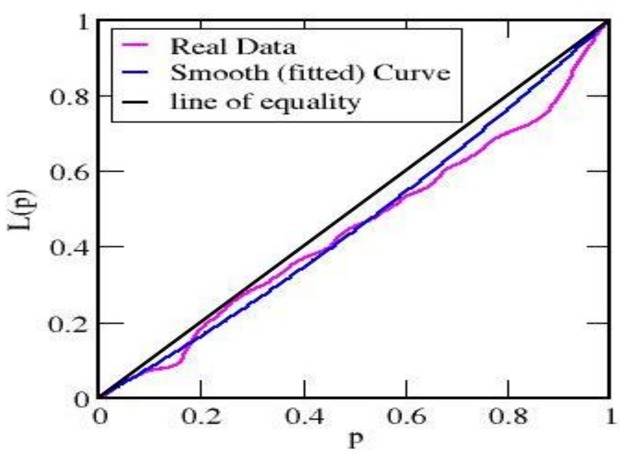
The proposed results are applied to the real life data taken from per capita Net state domestic product at current prices for 31 states and UTs of India for the year 1998-99. The data source is a publication of Directorate of Economics and Statistics, India. The Lorenz curve coordinates for the above data is given in the Table 1.

Table.1. Lorenz Curve Ordinates

Sl. No.	p	L(p) for 1998-99
0	0	0
1	0.1	0.070
2	0.2	0.180
3	0.3	0.287
4	0.4	0.394
5	0.5	0.473
6	0.6	0.548
7	0.7	0.631
8	0.8	0.713
9	0.9	0.798
10	1.0	1.000

To evaluate the Gini Index for the real data considered, first the Area under Lorenz Curve (AUC) is evaluated. Since the real data consists of only 31 data points and the data is not smooth when plotted, first a suitable interpolation is done to increase the plot points to 180 and 300 data points. Care has been taken to remove the error caused due to step size ‘h’. And a smooth curve fitting for the given data is given as shown in Fig. 6.1

Fig. 6.1 Lorenz Curve of Real Data



ISSN: 2319-5967
ISO 9001:2008 Certified
International Journal of Engineering Science and Innovative Technology (IJESIT)
Volume 1, Issue 2, November 2012

The area under the Lorenz curve for the real data for the year 1998-99 is presented in Table 2. The area was obtained using three different integration techniques mentioned earlier, viz. Trapizoidal rule, Simpson's 1/3rd rule and Simpson's 3/8th rule. Since the real data got was not very smooth, the area under the real data and the fitted curve is found have a 3% error.

Table.2 Area under Lorenz curve for Real Life Data (1998-99)

Area Under Lorenz Curve	Trapezoidal Rule	Simpson's 1/3 rule	Simpson's 3/8 rule
Real Data	0.4545992	0.4544867	0.4546349
Fitted Curve	0.4683867	0.4682509	0.4684798
% Error in AUC	2.945%	2.939%	2.955%

For finding partial Gini Index calculations, the real data is used since we want to illustrate the idea of partial Gini Index in real data. The Gini Index of the real data is found to be 0.090802. Partial Gini index is computed for 1998-1999 year in the interval [a,b] by using trapezoidal rule, Simpson's $^1/_3$ and $^3/_8$ rule. The values of PGI for different subintervals are given in Table 3.

Table 3 Partial Gini index for different class

Different Classes	PAUC	Partial Area of ∇	A	$PGI_{[a, b]}$
Poor class [0-0.3]	0.039994	0.045	0.005005	0.11124
Middle class [0.3-0.6]	0.121559	0.135	0.013441	0.099563
Rich class [0.6-1.0]	0.293046	0.320	0.026954	0.084231

It is clear from the Table 6.3 that the PAUC values of all the three sub intervals add up to total area under the curve in [0, 1]; the area of the trapeziums add up to the total area of the isosceles triangles with sides of unit length and the sum of the A's Columns add up to the value of A1 + A2 + A3 mentioned in the Figure 2.1. However, it is clear that we can't predict the value of Partial Gini index of any subinterval [a,b] having known the Gini index of the given data in [0, 1]. It is seen that the PGI of subintervals may be lower or higher than the Gini index themselves. It is observed in this real data that the PGI of poor class denoted by us in [0, 0.3] is higher by 18%, whereas the middle class PGI is higher by 8% and the rich class is lower by 7%.

VII. CONCLUSION

In this paper, the need for partial Lorenz curve is discussed and to measure the income inequality of a class, an inequality measure called as Partial Gini Index is proposed. The expression of partial Gini index is derived by using the three methods of numerical integration viz. trapezoidal method, Simpson's 1/3 and 3/8 method. Upper limit and lower limit for Partial Gini Index is derived by using the rectangular areas. The partial Lorenz ordering is useful in comparing the two partial Lorenz curves. The theoretical results are applied to a real life data. It is clear that the Partial Gini Index gives a better insight about the income inequality of a class compared to the Gini Index of the entire data. Using partial Lorenz curve and partial Gini Index one can check how much is the income inequality in a particular group and then economists can give some suggestion on how it can be reduced by making new policies, which significantly helps in improving the life style of deprived class. As Partial Gini Index measures the income inequality of a class, one can concentrate only on that class instead of judging the whole income inequality in the society. This will help us to work towards the betterment of the social welfare of the particular section of the society.

REFERENCES

[1] R Aaberge, "Characterization of Lorenz Curve and Income Distributions, Social Choice and Welfare," vol. 17, pp. 639-653, 2000.

[2] AB Atkinson, on the measurement of inequality," Journal of Economic Theory, vol. 2, pp. 244-263, 1970.

[3] M Chandra and ND Singpurwalla, "Relationships between some notions which are common to reliability theory and economics," Mathematics of Operations research, vol. 6, no.1, pp.113-121, 1981.

[4] MH Gail, "Applying the Lorenz curve to disease risk to optimize health benefits under cost constraints," Statistics and its interface, vol. 2, no. 2, pp.117-121, 2009.

[5] NC Kakwani, Income inequality and Poverty: Methods of Estimation and Policy, Oxford University Press, NY, 1980.

ISSN: 2319-5967
ISO 9001:2008 Certified
International Journal of Engineering Science and Innovative Technology (IJESIT)
Volume 1, Issue 2, November 2012

[6] B Klefsjo, "Reliability interpretations of some concepts of Economics", Naval Research Logistics Quarterly, vol. 31, 301-308, 1984.

[7] SC Kochar, "On extensions of DMRL and related partial orderings of life distributions," Communication in statistics-Stochastic Models, vol. 5, no. 2, pp. 235-245, 1989.

[8] WC Lee, "Analysis of Seasonal data using the Lorenz curve and the associated Gini index," International Journal of Epidemiology, vol. 25, no. 2, pp. 426-434, 1996.

[9] WC Lee, "Characterizing exposure –disease association in human population using the Lorenz curve and Gini index," Statistics in Medicine, vol. 16, no. 7, pp. 729-739, 1997.

[10] WC Lee, "Probabilistic analysis of Global performances of diagnostic tests: Interpreting the Lorenz curve based summary measures," Statistics in Medicine, vol.18, pp. 455-471, 1999.

[11] J Llorca and M Delqado-Rodriquez, "Visualizing exposure-disease association: the Lorenz curve and the Gini index," Medical Science Monitor: International medical Journal of experimental and clinical research, vol. 8, no.10, pp.193-197, 2002.

[12] S Pundir, Study of Inequality measures and related inference: A book for applied researchers and practitioners, LAP Lambert Academic Publishing Gmbh & Co. KG, Germany, 2012.

[13] A Ullah and David EA Giles, Handbook of Applied Economic Statistics, Marcel Dekker Inc., NY, 1998.

AUTHOR BIOGRAPHY

Dr. Sudesh Pundir is currently working as an Assistant Professor in the Department of Statistics, Pondicherry University, Puducherry. Her areas of research are Biostatistics, Applied Statistics, and Reliability. She has a no. of published research papers in reputed journals. She has also participated in many conferences in India as well as abroad. She has organized one International Conference and acted as an organizing committee member in many conferences. She has presented many research papers and gave invited talks, special invited talks and acted as a resource person in various National and International Conferences/Workshops. She is a life member of IISA and ISPS.

Dr. Sudesh Pundir
Assistant Professor,
Department of Statistics,
Pondicherry University,
Puducherry-605 014, INDIA

sudeshpundir19@gmail.com

Dr. Rajeswari Seshadri is currently working as a Reader in the Department of Mathematics, Pondicherry University, Puducherry. Her areas of research interests are Numerical solutions of differential equations, approximate analytic techniques, Boundary layer flows, and Laminar-Turbulent transition prediction. She has a no. of published research papers in reputed journals. She has organized many Conferences. She has presented many research papers, invited talks and acted as a resource person in various National and International Conferences/Workshops.

Dr. Rajeswari Seshadri
Reader,
Department of Mathematics,
Pondicherry University,
Puducherry-605 014, INDIA
oviaraji@yahoo.com