Economics Bulletin

Volume 37, Issue 2

An L^p-norm based approach to measuring inter-distributional inequality

Ritabrata Munshi
Indian Statistical Institute, Kolkata & TIFR, Mumbai

Soumyanetra Munshi
Economic Research Unit, Indian Statistical Institute, Kolkata

Abstract

This note proposes a class of indices of inter-distributional inequality (IDI) based on L^p-norm. We also modify the axiom of translation invariance and make it `level sensitive'. Level sensitivity, in the context of IDI measures, essentially captures the idea that an IDI is less affected by change in the distributional inequality at higher margins of the distributions than at the lower margins. Our proposed IDI measure, in addition to satisfying some standard properties of IDI measures, also satisfies the axiom of level sensitivity.

We are very grateful to Professor Richard McLean for his detailed comments on earlier drafts and impeccable attention to mathematical rigor. We also thank an anonymous referee for his careful reading and comments. The usual disclaimer applies.
Citation: Ritabrata Munshi and Soumyanetra Munshi, (2017) "An L^p-norm based approach to measuring inter-distributional inequality", *Economics Bulletin*, Volume 37, Issue 2, pages 1113-1121
Contact: Ritabrata Munshi - rmunshi@math.tifr.res.in, Soumyanetra Munshi - soumyanetra@isical.ac.in.
Submitted: October 27, 2016. **Published:** May 14, 2017.

1. Introduction

Income inequality is regarded by many as one of the most significant current public policy issues. Consequently, quantitative measures that compare income distributions have a significant role to play in the current discussions concerning income inequality. Various measures have been proposed for measurement of inequality (Cowell [2], Cowell and Kuga [3], Shorrocks [14], Lugo [13], Fields and Fei [6]). Recently there has also been an upsurge in the measurement of inter-distributional inequality, IDI (Yalontezky [18] and also [17]). Classically IDI indices have been considered by Gastwirth [9], Dagum [4], Ebert [5], Vinod [16], to name a few, and more recently by Breton et al. [1]. In this paper, we propose a measure of discrimination based on L^p-norm. We also introduce the idea of *level sensitivity* in the context of IDI measures which essentially captures the idea that an IDI is less affected by change in the distributional inequality at higher margins of the distributions than at the lower margins. This might be true of variables like income and literacy levels, for instance. Our proposed IDI measure, in addition to satisfying some standard properties of IDI measures, also satisfies the axiom of *level sensitivity*.

Consider the issue of comparing variations in distributions of an attribute. For example, consider two groups of population. These could be men and women within a country, or could be populations of two different countries. At any given point in time, each group may be described by its own distribution of a given attribute (like income, wealth, life expectancy, duration of unemployment). If attribute of one group is systematically different than the other, we say that the pair of distributions shows a 'discrimination pattern' or inter-distributional inequality (IDI). Measuring such IDI meaningfully through a new L^p-norm based index is the attempt of this paper.

Besides the usual axioms that a desirable IDI measure should satisfy (Yalonetzky [18]), we also mathematically formalize an axiom of *level sensitivity*, which is an adaptation of the axiom of translation invariance (see below). The basic idea of this can be found in thoughts of some earlier authors like Kakwani [10] in the context of poverty measures, Kolm [11], [12], and others, but, to the best of our knowledge, it has not been discussed in the context of IDI measures. Moreover our proposed measure also satisfies some other standard axioms of IDI indices.

The rest of the paper is organized as follows: In Section 2 we lay down some plausible axioms that an IDI index should ideally satisfy. In Section 3 we establish how our proposed L^p-norm based measure satisfies the axioms. Section 4 summarizes and concludes.

2. Axioms for inter-distributional inequality (IDI) indices

Let

$$\mathcal{F}_0 = \Big\{ F : \mathbb{R} \to [0,1] : \text{non-decreasing, right continuous,} \\ F((-\infty, 0)) = \{0\}, \lim_{x\to\infty} F(x) = 1 \Big\}$$

be the set of income distribution functions. The condition $F((-\infty, 0)) = \{0\}$ reflects our assumption that the income function is non-negative. However $F(0)$ is allowed to take any non-negative value. In particular we may have $F(0) > 0$ which means that a positive proportion of population has zero income. One may characterize the above set as the collection of all distribution functions of non-negative random variables. If the income distribution function of a population consisting of N individuals is $F \in \mathcal{F}_0$, then the income vector $(X_1, \dots, X_N)$ consists of N i.i.d. random variables $X_i \sim F$.

An inter-distributional inequality (IDI) index is a function

$$\Delta : \mathcal{F}_0 \times \mathcal{F}_0 \to \mathbb{R}_{\geq 0} \cup \{\infty\},$$

which assigns a non-negative number (including infinity) to every pair of distribution functions. The following are some of the desirable properties that such a measure should possess (Breton et al., 2011 [1], Yalonetzky [17], Yalonetzky [18]):

> **Axiom S: Symmetry** For any pair of distribution functions F_1 and F_2, we have $\Delta(F_1, F_2) = \Delta(F_2, F_1)$.

The symmetry axiom means that the measure of discrepancy between two distributions should remain the same, no matter in which order they are considered. In other words, an IDI index should not change when distributions are switched around[1] [2].

> **Axiom M: Monotonicity** Suppose F_1, F_2, F_3 are three distribution functions satisfying the following:
>
> $$\min\{F_1(t), F_2(t)\} \leq F_3(t) \leq \max\{F_1(t), F_2(t)\} \quad \forall t.$$
>
> Then
>
> $$\max\{\Delta(F_1, F_3), \Delta(F_2, F_3)\} \leq \Delta(F_1, F_2).$$

[1] Notice that the property of symmetry immediately rules out any kind of 'group-specific' disadvantage focus (GDF) as discussed in greater detail in Yalonetzky [18].

[2] See Shorrocks [15] for a critique of Dagum's measure (Dagum [4]) and other desirable properties of a 'distance function'. It was also adapted by Ebert [5]. Many discrimination measures are not symmetric and specifically measure discrepancy in the distribution of the 'comparison' population with respect to the distribution of the 'reference' population' (see Breton et. al [1] for example).

This means that if F_3 lies between F_1 and F_2 at every point then the measure of discrepancy between F_1 and F_3, and that between F_2 and F_3, must be smaller than that between F_1 and F_2. (Note that the standard convention for the extended non-negative real axis $\mathbb{R}_{\geq 0} \cup \{\infty\}$ is $x < \infty$ for all $x \in \mathbb{R}_{\geq 0}$.)

> **Axiom WSI+:** For $\lambda \in \mathbb{R}_+$, and $F \in \mathcal{F}_0$, let $F^{[\lambda]}(t) = F(t/\lambda)$. Then $F^{[\lambda]} \in \mathcal{F}_0$. The index is said to be SI+ if for all $F_1, F_2 \in \mathcal{F}_0$ the map $\lambda \mapsto \Delta(F_1^{[\lambda]}, F_2^{[\lambda]})$ is non-decreasing.

> **Axiom WSI–:** The index is said to be SI– if for all $F_1, F_2 \in \mathcal{F}_0$ the map $\lambda \mapsto \Delta(F_1^{[\lambda]}, F_2^{[\lambda]})$ is non-increasing.

We say that an IDI is **scale invariant (SI)** if it is both SI+ and SI–, i.e.

$$\Delta(F_1^{[\lambda]}, F_2^{[\lambda]}) = \Delta(F_1, F_2).$$

for all positive λ. This means that if the income function of both the populations are scaled up by the same factor then the index remains unchanged.

> **Axiom TI: Translation Invariance** For $L \in \mathbb{R}_+$, and $F \in \mathcal{F}_0$, let $F^{\{L\}}(t) = F(t - L)$. Then $F^{\{L\}} \in \mathcal{F}_0$. The index is said to be translation invariant if for all $F_1, F_2 \in \mathcal{F}_0$ and $L \in \mathbb{R}_+$ we have
>
> $$\Delta(F_1^{\{L\}}, F_2^{\{L\}}) = \Delta(F_1, F_2).$$

Under this axiom if two income functions are translated by the same factor then the index remains unchanged. This axiom is not compatible with scale invariance. One way to resolve this issue is to sacrifice SI in exchange of *Linear Homogeneity*, as was done, for example, by Ebert [5]. Since SI seems to be a more natural concept, we propose a variation of TI which is more intuitive. We call this Level Sensitivity.

> **Axiom LS: Level Sensitivity** An IDI index is said to be level sensitive if for any two distribution functions F_1 and F_2 we have $\Delta(F_1^{\{L\}}, F_2^{\{L\}})$ is non-increasing in L.

Level sensitivity w.r.t. a translation of incomes means that if everyone's income increased by a certain amount from an initial level, then the measure of discrepancy is smaller for the change at the translated incomes. For example, suppose we are interested in comparing the discrimination index between income distributions of men and women in India, with that between men and women in the US. Assume, as an example, it is known that income of all individuals in the US are L units (say Rs. 50,000) higher than those in India (and the men-women population structure in India and the US are the same). Then this axiom says that the discrimination index is higher for India than the US.

Axiom SSDE: Strong Sensitivity to Distributional Equality An IDI index is said to be strongly sensitive to distributional equality if:

$$F_1(t) = F_2(t)\ \forall t \iff \Delta(F_1, F_2) = 0.$$

Next, we propose an IDI index based on Minkowski's distance[3] [4].

3. An L^p-norm based measure of IDI

Let us define a natural metric of 'distance' between distributions as follows: For $p \in \mathbb{R}_+$,

$$\Delta_w^p(F_1, F_2) := \left(\int_{\mathbb{R}} |F_1(t) - F_2(t)|^p \, w(t) dt \right)^{\frac{1}{p}}, \tag{1}$$

where $w(t)$ is a non-negative weight function. Note that the integral on the right may not be finite, in which case we just put $\Delta_w^p(F_1, F_2) = \infty$. We will naturally be interested in the various properties of $\Delta_w^p(F_1, F_2)$ given various assumptions concerning w.

Theorem 1. *The above measure of discrimination Δ_w^p has the following properties.*

(i) Δ_w^p satisfies axioms S, M and SSDE if $w > 0$.

(ii) Δ_w^p satisfies LS if w is decreasing.

(iii) Suppose $w \in C^1(0,1)$, then Δ_w^p satisfies WSI+ if and only if $w(x) + w'(x)x \geq 0$, and it satisfies WSI– if and only if $w(x) + w'(x)x \leq 0$.

(iv) Suppose $w \in C^1(0,1)$, then Δ_w^p satisfies both WSI+ and WSI– (i.e. it satisfies SI) if and only if $w(x) + w'(x)x = 0$

(v) In particular if the weight function is taken to be $w(t) = t^{-1}$, then Δ_w^p satisfies S, M, SSDE, SI and LS.

The theorem relates the properties of the discrimination measure with the analytic properties of the weight function w. Before we present the formal proof of the theorem, let us motivate the analysis as follows. Fix $0 < \epsilon < \frac{1}{2}$ and consider the income distributions G_1

---

[3] Hence our proposed measure is very close to Ebert's measure [5] which takes the form

$$d^r(X, Y) = \left(\int_0^1 |F_X^{-1}(v) - F_Y^{-1}(v)|^r dv \right)^{1/r}, \quad r \geq 1,$$

where F_X^{-1} and F_Y^{-1} are the inverse distribution functions of the income distributions of X and Y. Notice that it satisfies, TI but not SI and LS.

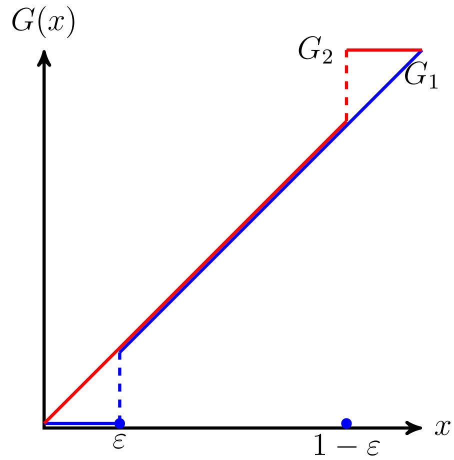
[4] We are formulating the IDI index in terms of the distribution function F. Usually IDI indices are defined in terms of the income vector $(X_1, \ldots, X_n)$ of the population with n members. Here X_i are random variables with cumulative distribution function F. Starting from a income data $\mathbf{X} \in \mathbb{R}^n_{\geq 0}$, we consider the sampling distribution function $F_{\mathbf{X}}$. Then we define $\Delta(\mathbf{X}, \mathbf{Y}) = \Delta(F_{\mathbf{X}}, F_{\mathbf{Y}})$.

and G_2 defined as follows:

$$G_1(x) = \begin{cases} 0, & x < \epsilon \\ x; & \epsilon \leq x < 1 \\ 1; & 1 \leq x. \end{cases}$$

$$G_2(x) = \begin{cases} 0, & x < 0 \\ x; & 0 \leq x < 1 - \epsilon \\ 1; & 1 - \epsilon \leq x. \end{cases}$$

The following figure gives a graph of these income distributions.



We wish to compare these two distributions with the "perfect equality" distribution $E(x) = x$ for all $x \in [0, 1]$. If one applies the simple unweighted L_1 inequality measure Δ, it follows that

$$\int_0^1 |x - G_1(x)|dx = \int_0^1 |x - G_2(x)|dx.$$

If $\int_0^1 |x - G_i(x)|dx$ is interpreted as a measure of the deviation of G_i from "perfect equality", i.e., the uniform distribution, then G_1 and G_2 would seem to exhibit the same degree of inequality. However one could argue that G_1 exhibits a greater degree of inequality than does G_2 if one feels that income disparities at low income levels correspond to greater inequality than income disparities at high income levels. The "problem" with the simple unweighted L_1 inequality measure stems from the fact that each income level $x \in [0; 1]$ is treated symmetrically by Δ. Hence a more general weighted version of Δ that places greater weight on income disparities at low income levels than at high income levels can result in a more refined inequality measure. If for example w is any strictly decreasing function on $[0; 1]$; then

$$\int_0^1 |x - G_1(x)|w(x)dx > \int_0^1 |x - G_2(x)|w(x)dx.$$

and we could conclude that F_1 exhibits a greater degree of inequality than F_2. This explains the statement (ii) of the theorem.

Proof. Axioms S and M are easily verified as follows:

Axiom S: Clearly $\Delta_w^p(F_1, F_2) = \Delta_w^p(F_2, F_1) = \left(\int_{\mathbb{R}} |F_1(t) - F_2(t)|^p\, w(t)dt\right)^{\frac{1}{p}}$.

Axiom M: Suppose $F_i \in \mathcal{F}_0$ for $i = 1, 2, 3$, satisfying

$$\min\{F_1(t), F_2(t)\} \leq F_3(t) \leq \max\{F_1(t), F_2(t)\} \quad \forall t.$$

It follows from the given condition that,

$$|F_1(t) - F_2(t)| \geq \max\{|F_2(t) - F_3(t)|, |F_1(t) - F_3(t)|\}$$

and hence

$$\begin{aligned}&\left(\int_{\mathbb{R}} |F_1(t) - F_2(t)|^p w(t)dt\right)^{\frac{1}{p}} \\ &\geq \max\left\{\left(\int_{\mathbb{R}} |F_2(t) - F_3(t)|^p w(t)dt\right)^{\frac{1}{p}}, \left(\int_{\mathbb{R}} |F_1(t) - F_3(t)|^p w(t)dt\right)^{\frac{1}{p}}\right\}.\end{aligned}$$

So we conclude that

$$\Delta_w^p(F_1, F_2) \geq \max\{\Delta_w^p(F_1, F_3), \Delta_w^p(F_2, F_3)\}.$$

Note that if the right hand side is ∞, so is the left hand side.

Axiom SSDE: Cleary $F_1(t) = F_2(t)\ \forall t \Rightarrow \Delta_w^p(F_1, F_2) = 0$. In the reverse direction we see that $\Delta_w^p(F_1, F_2) = 0$ implies that $F_1(t) = F_2(t)$ except possibly a measure zero set w.r.t the measure $w(t)dt$. So that SSDE is satisfied if $w(t) > 0$ for $t \geq 0$ [5].

Axiom LS: Suppose the weight function $w \in C^1(\mathbb{R})$ is decreasing. We have to show that $H(L) := \Delta_w^p(F_1^{\{L\}}, F_2^{\{L\}})$ is non-increasing in L. Now, by definition

$$\Delta_w^p(F_1^{\{L\}}, F_2^{\{L\}}) = \left(\int_{\mathbb{R}} |F_1(t - L) - F_2(t - L)|^p w(t)dt\right)^{\frac{1}{p}}.$$

Setting $t - L = z$, we get

$$\left(\int_{\mathbb{R}} |F_1(z) - F_2(z)|^p w(z + L)dz\right)^{\frac{1}{p}}.$$

[5] This is true for Ebert's measure as well and is called the reflexivity property in Ebert [5].

Suppose $L_1 < L_2$, then we have $w(z+L_1) \geq w(z+L_2)$, and consequently

$$\left(\int_{\mathbb{R}} |F_1(z) - F_2(z)|^p w(z+L_1) dz\right)^{\frac{1}{p}} \geq \left(\int_{\mathbb{R}} |F_1(z) - F_2(z)|^p w(z+L_2) dz\right)^{\frac{1}{p}},$$

i.e. $H(L_1) \geq H(L_2)$. So $\Delta_w^p(F_1^{\{L\}}, F_2^{\{L\}})$ is non-increasing in L.

Axiom WSI+ and WSI–: Let $\lambda > 0$ and consider

$$H(\lambda) = [\Delta_w^p(F_1^{[\lambda]}, F_2^{[\lambda]})]^p = \int_{\mathbb{R}} |F_1(t/\lambda) - F_2(t/\lambda)|^p w(t) dt.$$

Let $u = t/\lambda$, so that the right hand side reduces to

$$\int_{\mathbb{R}} |F_1(u) - F_2(u)|^p w(u\lambda) \lambda du.$$

Differentiating within the integral sign with respect to λ (which is allowed as $w \in C^1$) we get

$$\frac{\partial}{\partial \lambda} H(\lambda) = \int_{\mathbb{R}} |F_1(u) - F_2(u)|^p \{w(u\lambda) + uw'(u\lambda)\lambda\} du,$$

which is ≥ 0 (resp. ≤ 0) if $w(t) + tw'(t) \geq 0$ (resp. $w(t) + tw'(t) \leq 0$).

Axiom SI: It follows that if $w(t)+tw'(t) = 0$ for all t, then the derivative of $H(\lambda)$ vanishes and hence $H(\lambda)$ is constant, i.e. Δ_w^p is scale invariant. Conversely, scale invariance implies that $H'(\lambda) = 0$ for all possible choices for F_1 and F_2. If $w(t) + tw'(t)$ is not identically zero (almost sure), then one of the following two sets will have positive measure

$$W_{\pm} = \{t : \pm[w(t) + tw'(t)] > 0\}.$$

Without loss of generality suppose W_+ has positive measure. Since we are assuming that $w(t) + tw'(t)$ is continuous, this set should contain an interval, say $[a, b] \subset W_+$. We can further assume that $w(t) + tw'(t) > \delta > 0$ for all $t \in [a, b]$ and some small enough $\delta > 0$. Now consider

$$F_1(x) = \begin{cases} 0 & \text{for } x < b \\ 1 & \text{for } x \geq b \end{cases}$$

and

$$F_2(x) = \begin{cases} 0 & \text{for } x < a \\ 1/2 & \text{for } a \leq x < b \\ 1 & \text{for } x \geq b. \end{cases}$$

For this pair of F_1 and F_2, we have $H(\lambda) > 0$ for $\lambda = 1$, contradicting the fact that it is identically zero for all pairs of distributions.

Axiom TI: Let $L > 0$ and consider

$$[\Delta_w^p(F_1^{\{L\}}, F_2^{\{L\}})]^p = \int_{\mathbb{R}} |F_1(t-L) - F_2(t-L)|^p w(t)dt$$
$$= \int_{\mathbb{R}} |F_1(u) - F_2(u)|^p w(u+L)du.$$

The last integral equals $[\Delta_w^p(F_1, F_2)]^p$ if and only if $w(u+L) = w(u)$. So the weight function has to be a constant. Again observe that the argument holds even if ∞ is allowed as a possible value for the index. ■

4. Conclusion

In this paper, we propound an L^p-norm based index of inter-distributional inequality. We also introduce a variation of the standard translation invariance property of IDI indices (called level-sensitivity) which requires that IDI measures are decreasing and convex. This captures the idea that an IDI is less affected by change in the distributional inequality at higher margins of the distributions than at the lower margins which might be true in case of some variables like income. None of the usual IDI indices like those propounded by Gastwirth [9], Dagum [4], Ebert [5], Vinod [16], and Breton et al. [1], satisfies all the axioms.

References

Breton, Michel Le., Alessandra Michelangeli, Peluso, Eugenio (2012) "A stochastic dominance approach to the measurement of discrimination" *Journal of Economic Theory*, **147**, 1342–1350

Cowell, Frank. A. (2009) Measuring Inequality. Prentice Hall/Harvester Wheatsheaf

Cowell, Frank A. and Kuga, K. (1981) "Inequality Measurement: An Axiomatic Approach" *European Economic Review*, **15**, 287–305

Dagum, C. (1980) "Inequality Measures between Income Distributions with Applications" *Econometrica*, **48 (7)**, 1791–1803

Ebert, U. (1984) "Measures of Distance between Income Distributions" *Journal of Economic Theory*, **32**, 266–274

Fields, Gary S. and John C. H. Fei (1978) "On Inequality Comparisons" *Econometrica*, **46 (2)**, 303–316

Foster, J. E. (1984) "On economic poverty: a survey of aggregate measures" *Advances in Econometrics*, **3**, 215–251

Gastwirth, Joseph L. (1975) "Statistical Measures of Earnings Differentials", *The American Statistician*, **29 (1)**, 32–35

Gastwirth, Joseph L. (1985) Comment on "Measurement of Economic Distance between Blacks and Whites" by H. D. Vinod. *Journal of Business & Economic Statistics*, **3 (4)**, 405–407

Kakwani, N. (1980) On a Class of Poverty Measures. Econometrica. 48 (2), 437–446

Kolm, S-C. (1976) "Unequal Inequalities I", *Journal of Economic Theory*, **12**, 416–442.

Kolm, S-C. (1976) "Unequal Inequalities II", *Journal of Economic Theory*, **13**, 82–111.

Lugo, M.A. (2005) "Comparing Multidimensional Indices of Inequality: Methods and Application", ECINEQ WP 2005 - 14

Shorrocks, Anthony F. (1980) "The Class of Additively Decomposable Inequality Measures", *Econometrica*, **48 (3)**, 613–625

Shorrocks, Anthony F. (1982) "On the Distance between Income Distributions", *Econometrica*, **50 (5)**, 1337–1339

Vinod, H. D. (1985) "Measurement of Economic Distance between Blacks and Whites", *Journal of Business & Economic Statistics*, **3 (1)**, 78–88

Yalonetzky, G. (2010) Measuring group disadvantage with indices based on relative distributions. Oxford Poverty and Human Development Initiative, University of Oxford.

Yalonetzky, G. (2012) "Measuring group disadvantage with Inter-distributional Inequality Indices: A Critical Review and Some Amendments to Existing Indices", Economics: The Open-Access, Open-Assessment E- Journal, Vol. 6, 2012-9. http://dx.doi.org/10.5018/economics-ejournal.ja.2012-9.