Revista Colombiana de Estadística
Current Topics in Statistical Graphics
Diciembre 2014, volumen 37, no. 2, pp. 419 a 436
DOI: http://dx.doi.org/10.15446/rce.v37n2spe.47947

The Graphical Representation of Inequality

La representación gráfica de la desigualdad

Alberto Arcagni[a], Francesco Porro[b]

Dipartimento di Statistica e Metodi Quantitativi, Università degli Studi di Milano-Bicocca, Italy

Abstract

As of the past century, the analysis and the graphical representation of inequality play a very important role in economics. In the literature, several curves have been proposed and developed to simplify the description of inequality. The aim of this paper is a review and a comparison of the most known inequality curves, evaluating the features of each, with a particular focus on interpretation.

***Key words*:** Bonferroni Curve, Inequality Index, Income Distribution, Lorenz Curve, Zenga Inequality Curve.

Resumen

Desde el siglo pasado el análisis y representación gráfica de la desigualdad juega un papel importante en la economía. En la literatura varias curvas han sido propuestas y desarrolladas para simplificar la descripción de la desigualdad. El objetivo de este artículo es revisar y comparar las curvas de la desigualdad más conocidas evaluando sus características y enfocándose en su interpretación

***Palabras clave*:** curva de Bonferroni, curva de Lorenz, curva de Zenga, distribución del ingreso, índice de desigualdad.

1. Introduction

Inequality is an important characteristic of non-negative distributions. It is mainly analysed in socio-economics sciences, particularly in relation to income

[a]Professor. E-mail: alberto.arcagni@unimib.it
[b]Professor. E-mail: francesco.porro1@unimib.it

distributions. Inequality curves are graphical methods used to analyse this characteristic and are generally related to inequality indexes. The graphs of inequality functions can usually be drawn in the unitary square.

In this paper three curves are presented. The Lorenz curve (Lorenz 1905) is the oldest but also the most used nowadays despite its forced behaviour. The Bonferroni curve (Bonferroni 1930) is another classical curve. It is strictly related to the Lorenz curve and it has a forced behaviour, too. Finally the $I(p)$ curve (Zenga 2007) is the most recent although related to the other two curves, it can assume different shapes which allow to distinguish different situations in terms of inequality.

These three curves have the common characteristic that they can be defined using only the mean of the whole population and the means of particular subgroups. In the literature, other inequality curves have been introduced, studied and applied in different fields. One of the first proposals is the $\delta(p)$ of Gini which has the important feature of being uniform for the Pareto distribution, but it does not lie in the unitary square. Another which can be mentioned is the $Z(p)$ curve, proposed by Zenga (1984). Such curve is uniform for the Log-normal distribution. It originates from a different approach because it is based on a ratio of two quantiles, and therefore not included in this comparison.

In this paper curves are defined for continuous models, but they can be also applied to discrete distributions and to empirical distributions.

An important application of the inequality curves is that they can be used to define some orderings. Such orderings allow the comparison of distributions in terms of inequality. This kind of comparison within the same model allows to understand how the distribution parameters influence the inequality.

The article is structured as follows. First, all main definitions are introduced in Section 2. In this section, distribution models used to exemplify inequality curves are also described. Sections 3, 4, and 5 are devoted to the description of the Lorenz curve, the Bonferroni curve and the $I(p)$ curve, respectively. In Section 6 the orderings based on the considered curves are introduced and their relationship investigated. Section 7 provides a method to simplify the comparison of the curves: an application of this comparison is performed by using data from the 2012 Bank of Italy sample survey. Finally Section 8 is devoted to some final remarks.

2. Preliminary Definitions

In this section some definitions that will be useful hereunder, are introduced. The first one is as follows.

Definition 1 (Generalized inverse function)**.** Let F be a non-decreasing function defined from $\mathbb{R}$ to the interval $[0,1]$. The generalized inverse function of F is the function, denoted by F^{-1}, defined as:

$$F^{-1}(p) = \begin{cases} \inf\{y : F(y) \geq p\} & \text{if } p \in (0,1] \\ \inf\{y : F(y) > 0\} & \text{if } p = 0. \end{cases} \tag{1}$$

In the remainder of this paper, given a distribution function F, F^{-1} will denote the inverse function of F or, if needed, the generalized inverse function of F.

It is well recognized in the literature that the inequality does not change in case of scale-transformations, thus the inequality curves must be non dependent on the scale parameters of the distribution. The definition of scale parameter follows:

Definition 2 (Scale parameter)**.** Let $\{F_\alpha,\ \alpha > 0\}$ be a family of distribution functions. Then α is a scale parameter of such family if

$$F_\alpha(x) = F_1\left(\frac{x}{\alpha}\right), \qquad \forall x \in \mathbb{R}.$$

To simplify the explanation, the concepts of lower and upper groups are useful. Given a population and a statistic variable X evaluated on it, for each $p \in (0,1)$, the population can be split into two groups. The first one, called *lower group* consisting of the proportion p of people with the lowest values of X, and the second one called *upper group* composed by all the others. Once the population is split into the lower and the upper groups, the means of X in these two groups can be computed, obtaining the *lower* and the *upper mean*. The two following definitions refer to these two means.

Definition 3 (Lower mean)**.** Let X be a continuous random variable, with distribution function F, and support $[a,b]$, where $0 \leq a < b \leq +\infty$. For any $p \in [0,1]$, the lower mean $\overset{-}{M}(p)$ is defined as

$$\overset{-}{M}_{(p)} = \begin{cases} \frac{1}{p}\int_0^p F^{-1}(t)dt & \text{if } p \in (0,1] \\ a & \text{if } p = 0. \end{cases}$$

Definition 4 (Upper mean)**.** Let X be a continuous random variable, with distribution function F, and support $[a,b]$, where $0 \leq a < b \leq +\infty$. For any $p \in [0,1]$, the upper mean $\overset{+}{M}_{(p)}$ is defined as

$$\overset{+}{M}_{(p)} = \begin{cases} \frac{1}{1-p}\int_p^1 F^{-1}(t)dt & \text{if } p \in [0,1) \\ b & \text{if } p = 1. \end{cases}$$

***Note* 1.** In the Definitions 3 and 4 the lower mean and the upper mean have been extended by continuity in $p = 0$ and in $p = 1$, respectively. It is easy to verify that for a random variable X with expected value μ, the following formula holds true:

$$\mu = p \overset{-}{M}_{(p)} + (1-p) \overset{+}{M}_{(p)}, \qquad \forall p \in [0,1],$$

with the convention that whether the support of X is not finite:

$$(1-p) \overset{+}{M}_{(p)} = 0 \qquad \text{if } p = 1.$$

In the next sections, in order to calculate the inequality curves for some distribution models, the following is considered:

- the *(non-negative) uniform model* with distribution function

$$F(x) = \begin{cases} 0 & \text{if } x < \alpha(1-\theta) \\ \frac{x-(1-\theta)\alpha}{2\theta\alpha} & \text{if } \alpha(1-\theta) \leq x < \alpha(1+\theta) \\ 1 & \text{if } x \geq \alpha(1+\theta), \end{cases}$$

  where $0 \leq \theta \leq 1$ is a direct inequality indicator and $\alpha > 0$ is a scale parameter;

- the *exponential model* with distribution function

$$F(x) = \begin{cases} 1 - e^{-x/\alpha} & \text{if } x \geq 0 \\ 0 & \text{otherwise,} \end{cases}$$

  where $\alpha > 0$ is a scale parameter;

- the *Pareto model* with distribution function

$$F(x) = \begin{cases} 1 - \left(\frac{x_0}{x}\right)^\theta & \text{if } x \geq x_0 \\ 0 & \text{otherwise,} \end{cases}$$

  where $\theta > 1$ (to guarantee a finite expectation) is an inverse inequality indicator and $x_0 > 0$ is the lower bound of the support and a scale parameter;

- the *Log-normal model* with distribution function

$$F(x) = \begin{cases} \Phi\left(\frac{\ln(x)-\gamma}{\delta}\right) & \text{if } x > 0 \\ 0 & \text{otherwise,} \end{cases}$$

  where $\delta > 0$ is a direct inequality indicator, e^γ is a scale parameter and $\Phi(x)$ is the distribution function of the standard normal distribution;

- the *Dagum model* (see Dagum 1977) with distribution function

$$F(x) = \begin{cases} \left[1 + \left(\frac{x}{\alpha}\right)^{-\theta}\right]^{-\beta} & \text{if } x > 0 \\ 0 & \text{otherwise,} \end{cases}$$

  where $\beta > 0$ and $\theta > 1$ (to guarantee a finite expectation) are inverse inequality indicators where the other is fixed, and $\alpha > 0$ is a scale parameter.

3. The Lorenz Curve

The Lorenz curve introduced in the widely known paper by Lorenz (1905) is the most famous inequality curve used in the literature. Many equivalent definitions of it have been contributed, the following is from Pietra (1915) and it has been used also by Gastwirth (1972).

Definition 5. Let X be a non-negative continuous random variable, with positive and finite expected value μ, and distribution function F. The Lorenz curve of X is defined as

$$\begin{aligned} L(p) &= \frac{1}{\mu}\int_0^p F^{-1}(t)dt \\ &= \frac{p\,\bar{M}_{(p)}}{\mu}, \qquad p \in [0,1]. \end{aligned}$$

An inequality index that can be evaluated using the Lorenz curve is the Gini index G (Gini 1914). It is worth highlighting that the definition of such index does not require the Lorenz curve: only later was the relationship with this curve emphasized. From a graphical standpoint, the Gini index can be interpreted as the ratio of the concentration area and its theoretical maximum. The concentration area is the area between the bisector of the first quadrant and the Lorenz curve; its theoretical maximum corresponds to the area below such bisector.

Definition 6. Let X be a continuous random variable with Lorenz curve $L(p)$. The Gini index G is defined as

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

or equivalently as

$$G = \int_0^1 \frac{p - L(p)}{p}\cdot 2p\,dp = \int_0^1 \frac{\mu - \bar{M}(p)}{\mu}\cdot 2p\,dp. \tag{2}$$

The meaning of the Lorenz curve is not very immediate, since it compares the lower mean and the total mean, using the "weight" p, resulting in a less clear interpretation of such comparison. However, if the random variable X represents income, and $L(p)$ is the corresponding Lorenz curve, $L(p_0) = L_0$ means that the "bottom" proportion p_0 of the population has the proportion L_0 of total income.

It is easy to verify that the Lorenz curve is always zero at $p = 0$ and equals 1 at $p = 1$: such restrictions highlight that the behavior of $L(p)$ is a priori fixed. For this reason the explaining power of the Lorenz curve vanishes for values of p close to 0 or to 1. Moreover, it is well-known that the Lorenz curve is always convex. An interesting characteristic of the Lorenz curve is that the maximum length of the vertical segment between it and the bisector of the first quadrant is known as the Pietra index P and corresponds to the value $\tilde{p} = F(\mu)$:

$$P = \frac{E(|X-\mu|)}{2\mu} = F(\mu) - L[F(\mu)].$$

Moreover the derivative of the Lorenz curve at $\tilde{p} = F(\mu)$ is equal to 1.

Following the approach developed in Zenga (1984), using the Lorenz curve, it is possible to define a random variable which tends to the situation of maximal inequality as shown below. Let X be a random variable depending on the parameter θ. X is said to tend to the situation of maximal inequality as θ tends to θ_0 where

$$\lim_{\theta \to \theta_0} L_X(p;\theta) = L_M(p) = \begin{cases} 0 & \text{if } p \in [0,1) \\ 1 & \text{if } p = 1. \end{cases}$$

In such case, G is equal to 1. Analogously, X is said to tend to the situation of minimal inequality as θ tends to θ_0 if

$$\lim_{\theta \to \theta_0} L_X(p;\theta) = L_m(p) = p, \qquad \forall\, p \in [0,1],$$

meaning that the Lorenz curve tends to the bisector of the first quadrant, and therefore G tends to 0.

Table 1 shows the Lorenz curve and the Gini index for the distribution models described in Section 2, where

$$B(x;a,b) = \frac{\int_0^x t^{a-1}(1-t)^{b-1}dt}{\int_0^1 t^{a-1}(1-t)^{b-1}dt} \qquad x \in [0,1], a > 0, b > 0$$

is the incomplete Beta function ratio, and

$$\Gamma(x) = \int_0^\infty t^{x-1}e^{-t}dt$$

is the Gamma function. Figure 1 shows some examples of the Lorenz curves from Table 1.

Table 1: Lorenz curves and Gini indices for the considered models.

Model	**Lorenz curve**	**Gini Index**
Uniform	$L(p) = p(1-\theta+\theta p)$	$G = \theta/3$
Exponential	$L(p) = p + (1-p)\ln(1-p)$	$G = 0.5$
Pareto	$L(p) = 1-(1-p)^{(\theta-1)/\theta}$	$G = 1/(2\theta-1)$
Log-normal	$L(p) = \Phi[\Phi^{-1}(p) - \delta]$	$G = 2\Phi(\delta/\sqrt{2}) - 1$
Dagum	$L(p) = B\left(p^{1/\beta}; \beta+1/\theta; 1-1/\theta\right)$	$G = \frac{\Gamma(\beta)\Gamma(2\beta+1/\theta)}{\Gamma(2\beta)\Gamma(\beta+1/\theta)} - 1$

4. The Bonferroni Curve

The curve was introduced by Bonferroni (1930) and, to the present, has been analysed and studied by various authors: see for instance De Vergottini (1940), Tarsitano (1990), Giorgi & Crescenzi (2001) and Zenga (2013).

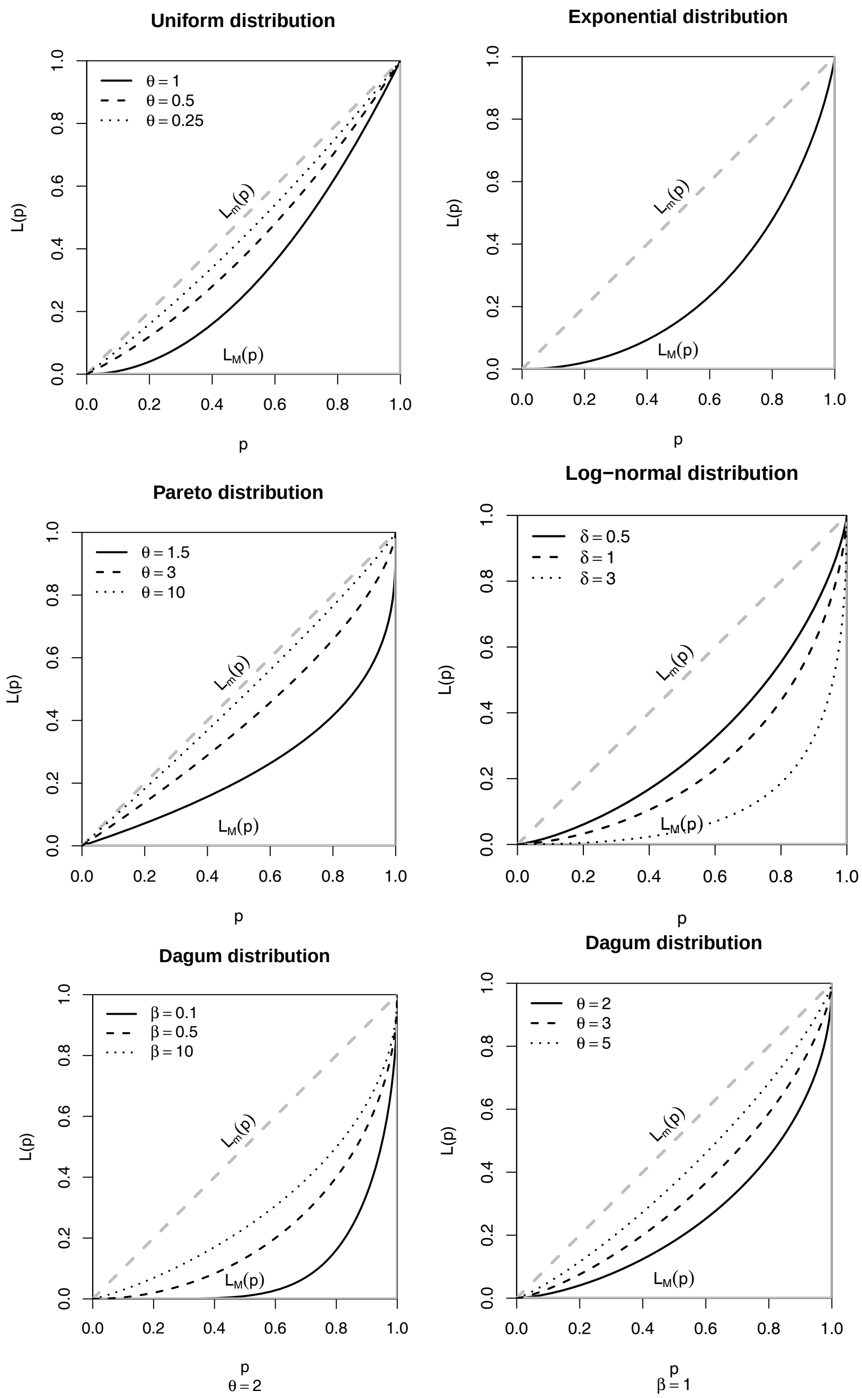

Figure 1: Graphs of different Lorenz curves for the considered models.

The Bonferroni curve is defined as follows:

Definition 7. Let X be a non-negative continuous random variable with a positive and finite expected value μ, and distribution function F. The Bonferroni curve of X is defined as

$$\begin{aligned} B(p) &= \frac{1}{p\mu}\int_0^p F^{-1}(t)dt \\ &= \frac{\bar{M}_{(p)}}{\mu} \qquad p \in (0,1]. \end{aligned}$$

The Bonferroni inequality index B represents the area above the Bonferroni curve in the unitary square, namely, the complement to 1 of the Bonferroni curve mean value.

Definition 8. Let X be a non-negative continuous random variable with Bonferroni curve $B(p)$. The Bonferroni index is defined as

$$B = 1 - \int_0^1 B(p)dp.$$

The Bonferroni curve compares the mean of the lower group with the total mean. Differently from the Lorenz curve no "weight" is applied. In other words, if the random variable X represents income, and $B(p)$ is the corresponding Bonferroni curve, $B(p_0) = B_0$ means that the average income of the "bottom" proportion p_0 of the population is B_0 times the average income of the entire population.

Using the Definitions 3 and 4 it is easy to see that

$$\lim_{p \to 0^+} B(p) = \frac{a}{\mu} \qquad \text{and} \qquad B(1) = 1,$$

where a denotes the lower bound of the support of the random variable originating the Bonferroni curve. Moreover, differently from the Lorenz one, the Bonferroni curve is not necessarily convex.

If the random variable X tends to the situation of maximal inequality, the $B(p)$ curve tends to the function $B_M(p)$, defined as:

$$B_M(p) = \begin{cases} 0 & \text{if } p \in (0,1) \\ 1 & \text{if } p = 1, \end{cases}$$

and consequently the inequality index is $B = 1$.

If the random variable X tends to the situation of minimal inequality, the corresponding Bonferroni curve tends to

$$B_m(p) = 1 \qquad \forall\, p \in (0,1],$$

and consequently the corresponding inequality index is $B = 0$.

A particular shape of the Bonferoni curve is obtained when the random variable X has a uniform distribution, as in this case, it is a linear function. Hence, if X has a uniform distribution with support $[\alpha(1-\theta), \alpha(1+\theta)]$ (see Section 2), then the corresponding Bonferroni curve is given by

$$B(p) = (1-\theta) + \theta p,$$

and the inequality index is $B = \theta/2$.

The Bonferroni curve is related with the Lorenz curve: if $L(p)$ is the Lorenz curve of X, then the Bonferroni curve can be obtained throught the simple transformation

$$B(p) = \frac{L(p)}{p}, \qquad \forall\, p \in (0,1]$$

Table 2 shows the Bonferroni curve and the Bonferroni index for the distribution models presented in Section 2, where

$$\Psi(x) = \frac{d}{dx} \ln \Gamma(x) = \frac{\Gamma'(x)}{\Gamma(x)}$$

denotes the Digamma function, that is, the logarithmic derivative of the Gamma function. Figure 2 shows some examples of the Bonferroni curves from Table 2.

TABLE 2: Bonferroni curves and Bonferroni indices for the considered models.

Model	**Bonferroni curve**	**Bonferroni Index**
Uniform	$B(p) = (1-\theta) + \theta p$	$B = \theta/2$
Exponential	$B(p) = 1 + \frac{(1-p)}{p} \ln(1-p)$	$B = 0.644934$
Pareto	$B(p) = \frac{1-(1-p)^{(\theta-1)/\theta}}{p}$	$B = 1 - \Psi(2-1/\theta) + \Psi(1)$
Log-normal	$B(p) = \frac{\Phi\left[\Phi^{-1}(p)-\delta\right]}{p}$	$B = 1 - \int_0^1 \frac{\Phi\left[\Phi^{-1}(p)-\delta\right]}{p} dp$
Dagum	$B(p) = \frac{B\left(p^{1/\beta}; \beta+1/\theta; 1-1/\theta\right)}{p}$	$B = \beta\left[\Psi\left(\beta + \frac{1}{\theta}\right) - \Psi(\beta)\right]$

5. The $I(p)$ Curve

The $I(p)$ curve was introduced in Zenga (2007). It is the most recent inequality curve of the three considered in this paper. Nevertheless, the number of papers discussing it and the related index I is increasing: see for instance Greselin & Pasquazzi (2009), Radaelli (2010), Langel & Tillé (2012) and Greselin, Pasquazzi & Zitikis (2013). This curve is defined below:

Definition 9. Let X be a non-negative continuous random variable, with a positive and finite expected value μ, and distribution function F. The $I(p)$ curve of X is defined as

$$\begin{aligned} I(p) &= 1 - \frac{(1-p)\int_0^p F^{-1}(t)dt}{p\int_p^1 F^{-1}(t)dt} \\ &= 1 - \frac{\overset{-}{M}_{(p)}}{\overset{+}{M}_{(p)}}, \qquad p \in (0,1). \end{aligned}$$

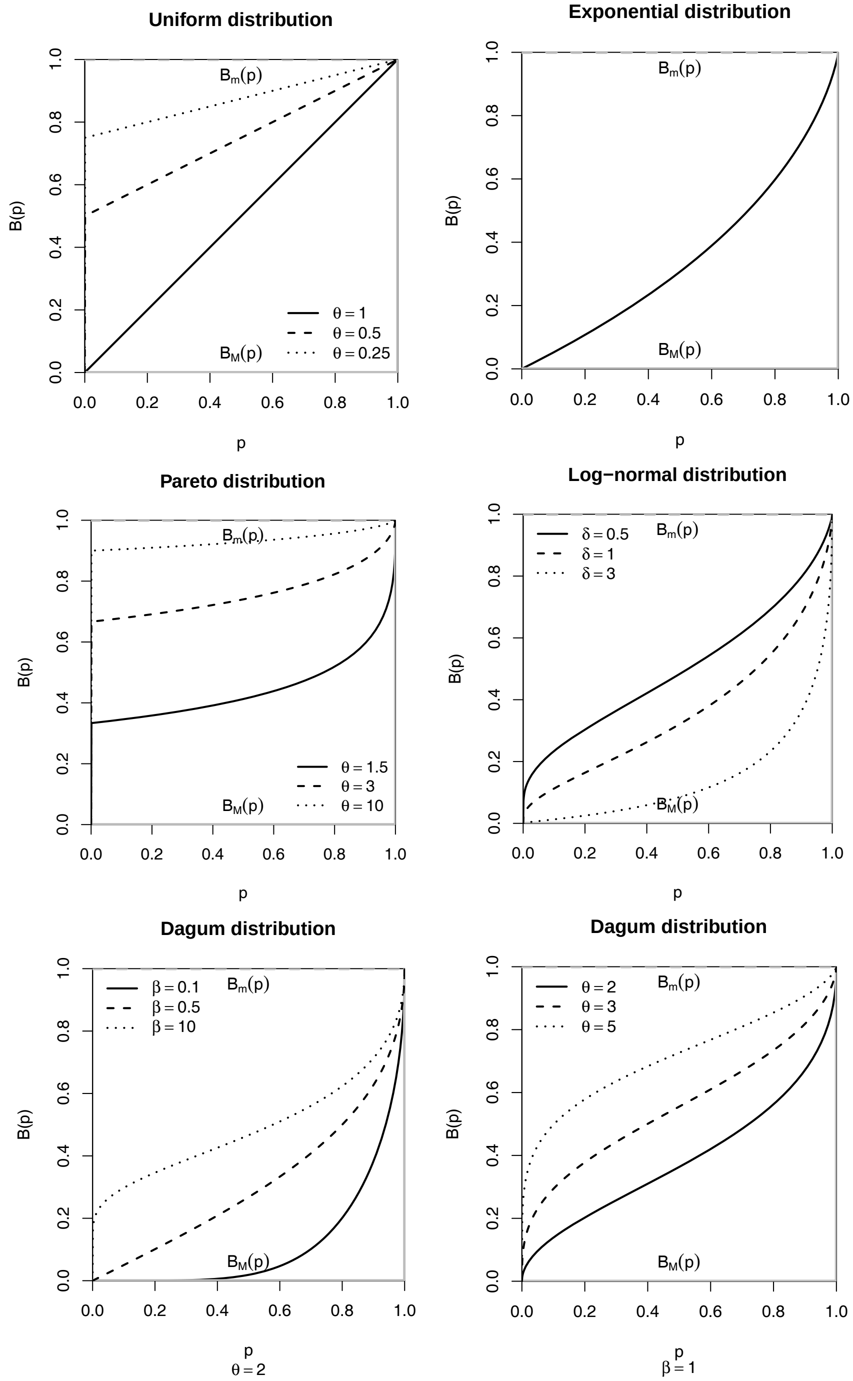

FIGURE 2: Graphs of different Bonferroni curves for the considered models.

Similar to the Bonferroni index, the inequality index I can be obtained from the mean value of the $I(p)$ curve but it represents the area below the $I(p)$ curve.

Definition 10. Let X be a continuous random variable and let $I(p)$ denotes its inequality $I(p)$ curve. The inequality index I is defined as

$$I = \int_0^1 I(p)dp.$$

The $I(p)$ curve can be easily interpreted, and its information is immediate and intuitive. If the random variable X models income distribution, it follows, by definition, that if the $I(p)$ curve is equal to I_0 at $p = p_0$, this means that the average income of the "bottom" proportion p_0 of the population is $(1 - I_0)$-times the average income of the remaining population.

As previously mentioned, the Lorenz curve assumes prefixed values for $p = 0$ and $p = 1$, whereas the Bonferroni curve is always equal to 1 for $p = 1$. The $I(p)$ curve is more flexible, since the values it assumes for the extreme values of p depend on the distribution function that originated the curve. Polisicchio (2008) proved that if X is a random variable with support $[a, b]$, where $0 \leq a < b \leq +\infty$ and with finite and positive expected value μ, then

$$\lim_{p \to 0^+} I(p) = 1 - \frac{a}{\mu} \qquad \text{and} \qquad \lim_{p \to 1^-} I(p) = 1 - \frac{\mu}{b},$$

with the convention that $\mu/b = 0$ if b is not finite. Moreover, also the $I(p)$ curve is not necessarily convex.

If the random variable X tends to the situation of maximal inequality, then the $I(p)$ curve tends to the function $I_M(p)$, defined as

$$I_M(p) = 1, \qquad \forall\, p \in (0, 1),$$

while, if the random variable X tends to the situation of minimal inequality, the $I(p)$ curve tends to zero for all $p \in (0, 1)$, that is

$$I_m(p) = 0, \qquad \forall\, p \in (0, 1).$$

Polisicchio (2008) proved that if the $I(p)$ curve of the random variable X is uniform and equal to $1-k$, then X has a truncated Pareto distribution with parameters $\theta = 0.5$, $x_0 = \mu k$, and μ/k as truncation point. Therefore, the distribution function of X is

$$F(x) = \begin{cases} 0 & \text{if } x \leq \mu k \\ \frac{1}{1-k}\left[1 - \sqrt{\frac{\mu k}{x}}\right] & \text{if } \mu k < x < \mu/k \\ 1 & \text{if } x \geq \mu/k. \end{cases}$$

The above truncated Pareto distribution has been analysed and from this model, a new distribution model, which seems very promising for modelling income distributions has been defined, see Zenga (2010), Arcagni & Porro (2013) and Arcagni & Zenga (2013).

As the Bonferroni curve, the $I(p)$ curve is also related to the Lorenz curve, and therefore to the Bonferroni curve itself. The relationships are (see Zenga 2007)

$$I(p) = \frac{p - L(p)}{p[1 - L(p)]} \qquad \forall\, p \in (0,1)$$

$$I(p) = \frac{1 - B(p)}{1 - pB(p)} \qquad \forall\, p \in (0,1).$$

Table 3 shows the $I(p)$ curve and the index I for the distribution models described in Section 2. Figure 3 includes some examples thereof.

TABLE 3: $I(p)$ curves and I indices for the considered models.

Model	$I(p)$ **curve**	I **index**
Uniform	$I(p) = \theta(1+\theta p)^{-1}$	$I = \ln(\theta+1)$
Exponential	$I(p) = \frac{\ln(1-p)}{p[\ln(1-p)-1]}$	$I = 0.843302$
Pareto	$I(p) = \frac{1-(1-p)^{1/\theta}}{p}$	$I = \Psi(1/\theta+1) - \Psi(1)$
Log-normal	$I(p) = \frac{p-\Phi[\Phi^{-1}(p)-\delta]}{p[1-\Phi(\Phi^{-1}(p)-\delta)]}$	$I = \int_0^1 \frac{p-\Phi[\Phi^{-1}(p)-\delta]}{p[1-\Phi(\Phi^{-1}(p)-\delta)]}\,dp$
Dagum	$I(p) = \frac{p-B\left(p^{1/\beta};\beta+1/\theta;1-1/\theta\right)}{p[1-B\left(p^{1/\beta};\beta+1/\theta;1-1/\theta\right)]}$	$I = \int_0^1 \frac{p-B\left(p^{1/\beta};\beta+1/\theta;1-1/\theta\right)}{p[1-B\left(p^{1/\beta};\beta+1/\theta;1-1/\theta\right)]}\,dp$

6. The Partial Orders

In the literature, an important application related to inequality curves, is the possibility to rank the distributions. Such ranking is obtained by a partial order which can be defined from an inequality curve. Below we include the definition of the well-known ordering based on the Lorenz curve.

Definition 11 (Partial order based on the Lorenz curve)**.** Let X and Y be two continuous non-negative random variables, both with finite and positive expected value. Let L_X and L_Y denote their Lorenz curves. X is said to be larger (or more unequal) than Y in the order based on the Lorenz curve (and it is denoted by $X \geq_L Y$), if

$$L_X(p) \leq L_Y(p) \qquad \forall\, p \in (0,1).$$

From the graphical point of view, the random variable X is larger than Y in this order, if its Lorenz curve lies below the Lorenz curve of Y for all $p \in (0,1)$. Analogously, to the ordering based on the Lorenz curve, the following can be defined.

Definition 12 (Partial order based on the Bonferroni curve)**.** Let X and Y be two continuous non-negative random variables, both with finite and positive expected value. Let B_X and B_Y denote their Bonferroni curves. X is said to be larger (or more unequal) than Y in the order based on the Bonferroni curve (and it is denoted by $X \geq_B Y$), if

$$B_X(p) \leq B_Y(p) \qquad \forall\, p \in (0,1).$$

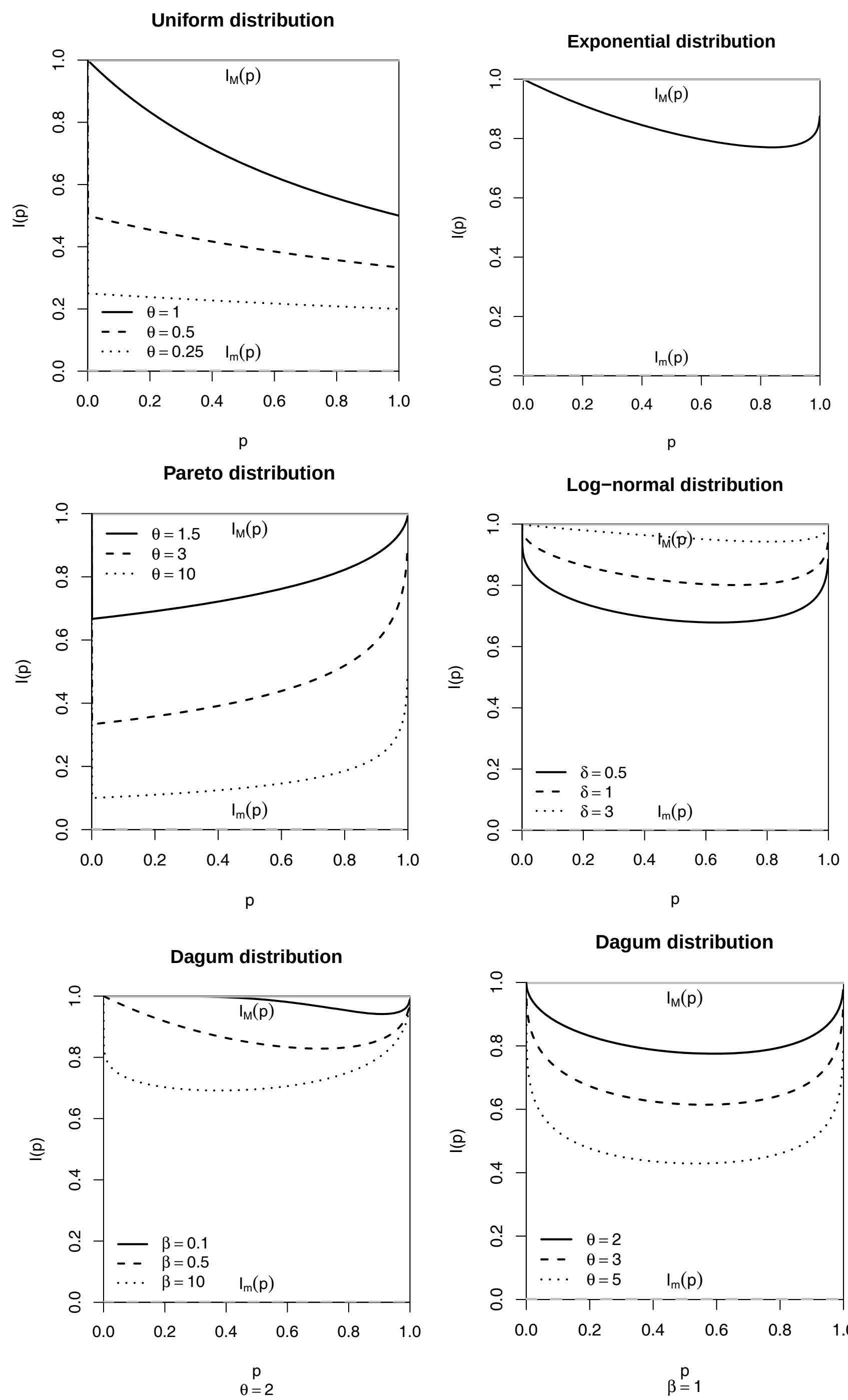

FIGURE 3: Graphs of different $I(p)$ curves for the considered models.

Even if less used, such ordering is well-known and is studied in the literature, see for example Tarsitano (1990), Giorgi & Crescenzi (2001), Pundir, Arora & Jain (2005).

The third partial order considered was introduced in Porro (2008).

Definition 13 (Partial order based on the $I(p)$ curve)**.** Let X and Y be two continuous non-negative random variables, both with finite and positive expected value. Let I_X and I_Y denote their inequality $I(p)$ curves. X is said to be larger (or more unequal) than Y in the ordering based on $I(p)$ curve (and it is denoted by $X \geq_I Y$), if

$$I_X(p) \geq I_Y(p) \qquad \forall\, p \in (0,1)$$

The relationship amongst these three orderings is summarized in the following result (for partial proof, see Polisicchio & Porro 2011).

Lemma 1 (Lemma of equivalence)**.** *Let X and Y be two continuous non-negative random variables X and Y, both with finite and positive expected value. Then:*

$$X \geq_L Y \Leftrightarrow X \geq_B Y \Leftrightarrow X \geq_I Y.$$

This lemma makes the coherence of the three curves evident; in fact, two distributions are ordered for one ordering if, and only if, they are ordered for the other two. It is important to mention that all these orderings are only partial orders, as there are some distributions with crossing $L(p)$ curves and therefore with crossing $B(p)$ and $I(p)$ curves, that can not be ordered for all $p \in (0,1)$. But, if the distributions belong to the same parametric model, these partial orders may allow to explain how the parameters influence them in terms of inequality. This is the case of the models defined in Section 2. Their parameters are classified in scale parameter or in direct and indirect inequality indicators. As defined in the same section, the scale parameters do not influence inequality. How other parameters influence the inequality curves is shown in Figures 1, 2 and 3, further showing that the curves do not intersect.

7. A Unified Point of View

All the curves presented in the previous sections are defined as introduced in the literature. As the partial orders described in the previous section show, it does not always happen that, given two inequality curves, the one related to the situation of more inequality lies above the other. From the graphical point of view, inequality curves can be more intuitive if they satisfy such restriction, meaning that for a fixed $p \in (0,1)$, the curve related to the situation of more inequality takes on a greater value than the curve related to the situation of less inequality.

Following the same approach used by Zenga (1984), such "increasing ranking" can be achieved by performing a suitable transformation on the inequality curves. In Zenga (1984) through a simple transformation on the $\delta(p)$ of Gini, the new $\lambda(p)$ curve is obtained: such new curve lies in the unitary square and satisfies the aforementioned "increasing ranking".

Then, from the Lorenz curve, the curve $G(p)$ can be obtained as:

$$G(p) = 2[p - L(p)] \qquad p \in (0, 1),$$

which coincides with the function in the first integral in formula (2). Analogously from the Bonferroni curve, the $V(p)$ curve can be otained as:

$$V(p) = 1 - B(p) \qquad p \in (0, 1).$$

From the definition of the Bonferroni curve and formula (2) it follows that

$$G(p) = V(p) \cdot 2p \qquad p \in (0, 1).$$

The inequality $I(p)$ curve needs no transformation, since it already satisfies the "increasing ranking".

Another interesting result of these transformations is that the new curves have the following feature: the related inequality indices are the areas below the curves.

As mentioned in the introduction, the curves presented for continuous models can be applied to empirical distributions. Replacing the distribution function F of the model by the empirical cumulative distribution function (ECDF) is good enough. The empirical quantile function is the generalized inverse function F^{-1} of the ECDF as defined in formula (1). The result is a step-function with integral between 0 and 1 clearly equal to the empirical mean.

As an example, the formulae presented in this section can be applied to the data provided by the Bank of Italy (2012). The 2012 sample survey has been analysed with the `R` software (R Core Team 2013). The considered dataset consists of 8114 non-negative household incomes with mean equal to EUR 30481.01. In Figure 4 empirical curves $G(p)$, $V(p)$ and $I(p)$, that satisfy the "increasing ranking", are drawn. The three curves are drawn together in the unitary square. The values of the related indexes corresponding to the areas below the curves are included.

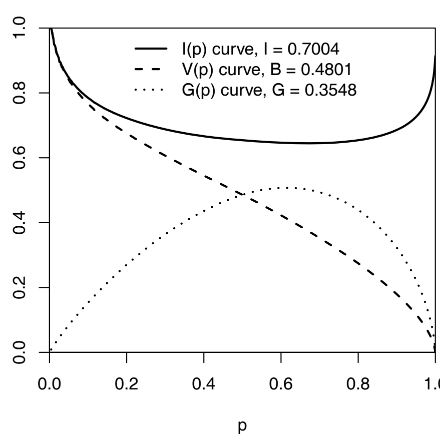

Figure 4: Unified representation of the inequality curves.

By using this unified representation it is easy to understand why the three indices assume such different values. In fact, index I is sensitive to the inequality in both the tails, index B is sensitive to the inequality in the poorest units but does not bring in the inequality from the richest ones, whereas the index G does not capture the inequality of both the tails.

8. Final Remarks

This paper is a review of the most known inequality curves. The considered curves are the Lorenz curve, the Bonferroni curve and the $I(p)$ curve. The main features of each are described with particular emphasis on their interpretation. Such curves are graphical methods used to analyse and compare inequality of non-negative distributions. For instance, inequality curves are used to rank the distributions through partial orders. The aforementioned curves are exemplified through five well-known non-negative distribution models, some of which can be used to describe income distributions. In the last section, a transformation of the Lorenz curve and a transformation of the Bonferroni curve allow an easier and more intuitive representation of such graphical tools.

[Recibido: mayo de 2014 — Aceptado: septiembre de 2014]

References

Arcagni, A. & Porro, F. (2013), 'On the parameters of Zenga distribution', *Statistical Methods & Applications* **22**(3), 285–303.

Arcagni, A. & Zenga, M. (2013), 'Application of Zenga's distribution to a panel survey on household incomes of European Member States', *Statistica & Applicazioni* **11**(1), 79–102.

Bank of Italy (2012), 'Indagine sui bilanci delle famiglie italiane'.
*http://www.bancaditalia.it/statistiche/storiche

Bonferroni, C. (1930), *Elementi di Statistica Generale*, Seeber, Firenze.

Dagum, C. (1977), *A New Model of Personal Income Distribution : Specification and Estimation*, Cahier de recherche, University of Ottawa, Faculty of Social Sciences, Department of Economics.
*http://books.google.com.co/books?id=9eckNAEACAAJ

De Vergottini, M. (1940), 'Sul significato di alcuni indici di concentrazione', *Annali di Economia Nuova Serie* **2**(5/6), 317–347.

Gastwirth, J. (1972), 'The estimation of the Lorenz curve and Gini index', *The Review of Economics and Statistics* **54**(3), 306–316.

Gini, C. (1914), 'Sulla misura della concentrazione e della variabilità dei caratteri', *Atti del Reale Istituto Veneto di Scienze, Lettere ed Arti* **73**, 1203–1248.

Giorgi, G. & Crescenzi, M. (2001), 'A look at Bonferroni inequality measure in a reliability framework', *Statistica* **41**, 571–583.

Greselin, F. & Pasquazzi, L. (2009), 'Asymptotic confidence intervals for a new inequality measure', *Communications in Statistics-Simulation and Computation* **38**(8), 1742–1756.

Greselin, F., Pasquazzi, L. & Zitikis, R. (2013), 'Contrasting the Gini and Zenga indices of economic inequality', *Journal of Applied Statistics* **40**(2), 282–297.

Langel, M. & Tillé, Y. (2012), 'Inference by linearization for Zenga's new inequality index: A comparison with the Gini index', *Metrika* **75**(8), 1093–1110.

Lorenz, M. (1905), 'Methods of measuring the concentration of wealth', *Publications of the American Statistical Association* **9**(70), 209–219.

Pietra, G. (1915), 'Delle relazioni fra indici di variabilità note I e II', *Atti del Reale Istituto Veneto di Scienze, Lettere ed Arti* **74**(2), 775–804.

Polisicchio, M. (2008), 'The continuous random variable with uniform point inequality measure I(p)', *Statistica & Applicazioni* **6**(2), 137–151.

Polisicchio, M. & Porro, F. (2011), 'A comparison between Lorenz L(p) curve and Zenga I(p) curve', *Statistica Applicata* **21**(3-4), 289–301.

Porro, F. (2008), Equivalence between partial order based on curve L(p) and partial order based on curve I(p), *in* 'Proceedings of SIS 2008', Padova.

Pundir, S., Arora, S. & Jain, K. (2005), 'Bonferroni curve and the related statistical inference', *Statistics & Probability Letters* **75**(2), 140–150.

R Core Team (2013), *R: A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing, Vienna, Austria.
*http://www.R-project.org/

Radaelli, P. (2010), 'On the decomposition by subgroups of the Gini index and Zenga's uniformity and inequality indexes', *International Statistical Review* **78**(1), 81–101.

Tarsitano, A. (1990), The Bonferroni index of income inequality, *in* 'Income and Wealth Distribution, Inequality and Poverty', C. Dagum and M. Zenga, Berlin, pp. 228–242.

Zenga, M. (1984), 'Tendenza alla massima ed alla minima concentrazione per variabili casuali continue', *Statistica* **44**(4), 619–640.

Zenga, M. (2007), 'Inequality curve and inequality index based on the ratios between lower and upper arithmetic means', *Statistica & Applicazioni* **5**(1), 3–27.

Zenga, M. (2010), 'Mixture of Polisicchio's truncated Pareto distributions with beta weights', *Statistica & Applicazioni* **8**(1), 3–25.

Zenga, M. (2013), 'Decomposition by sources of the Gini, Bonferroni and Zenga inequality indexes', *Statistica & Applicazioni* **11**(2), 133–161.