METRON - *International Journal of Statistics*
2004, vol. LXII, n. 1, pp. 137-160

AGOSTINO TARSITANO

A new class of inequality measures based on a ratio of L-statistics

Summary - The choice of a suitable weight function is a feasible and desirable direction for inequality comparison. This paper proposes a new class of indices based on a ratio of two L-statistics in which the weight function in the numerator involves the expected values of order statistics from a random variable describing the normative attitude to the measurement of inequality. Also, the article describes and discusses a graphical technique, based on the QQ plot, which is useful to select the most appropriate theoretical frequency distribution. A large-scale Monte Carlo study has been conducted to assess the sampling and asymptotic properties of some of a number of inequality indices. The simulations show that the new class is able to provide indices that compete well with the traditional measures of income inequality and industrial concentration.

Key Words - Gini family; Order statistics; QQ-plot; Economic inequality; Simulations.

1. Introduction

Linear functions of order statistics (L-statistics) have been widely applied in the study of income distribution to investigate asymptotic behavior of inequality indices and analyze their sensitivity to transfers at different points of the distributions. In fact, Stigler (1974) states that his principal motivation for carrying out research into L-statistics has been that "*...these estimators exhibit desirable robustness, particularly to heavy-tailed distributions or outlying distributions*". There has been an increased interest in measuring economic inequality over the recent decades. A number of monographs have been published in this field, including those by Sen (1997), Kakwani (1980), Nygard and Sandström (1981), and in particular for a survey on sampling aspects (from a parametric and distribution-free point of view) of some inequality indices investigated in the paper see Giorgi (1999).

A wide range of income inequality indices has been proposed and various attempts to unify the measures have been made (*e.g.* Zitikis 2002a, 2000b).

Received October 2003 and revised February 2004.

Here we propose a class of indices constructed as a ratio of L-statistics. What is new is that the weights are the expected values of a probability distribution describing the normative attitude to the measurement of income inequality. The new class has great flexibility, optimal sampling properties, and provides a fresh interpretation of some classical measures of income inequality. The contents of the various sections are as follows. Section 2 reviews the general formulation of L-statistics highlighting their more salient features. In the third section, the results of Section 2 are applied to show that a ratio of L-statistics yields valid measures of economic inequality. Section 4 reports on the results of a simulation experiment designed to compare the finite sampling properties of alternative measures of income inequality and to compare them with the asymptotic results. Finally, there is a conclusion in which the major merits of the new class are reviewed.

2. L-Statistics and inequality measures

Let $\{X_1, X_2, \ldots, X_n\}$ be a random sample of size n from a distribution function F and let $\{X_{1:n}, X_{2:n}, \ldots, X_{n:n}\}$ denote the associated order statistics. For a fixed sequence of weight functions J_n our interest will center on statistics of the type

$$\frac{L(J_n)}{\mu_n} = \frac{n^{-1}\sum_{i=1}^{n} J_n\left(\frac{i}{n+1}\right) X_{i:n}}{\mu_n} \quad \text{with} \quad \sum_{i=1}^{n} J_n\left(\frac{i}{n+1}\right) = 0\,; \tag{1}$$

where μ_n is the sample mean. Throughout the paper we assume that $\mu_n > 0$. Class (1) is particularly valuable in studies connected with income distributions because of its homogeneity of degree zero; moreover, it is computationally simple, exhibits desirable robustness when estimated from a sample and it is asymptotically efficient given a proper choice of the weight functions. The present paper deals with a class of L-statistics in which, following Chernoff *et al.* (1967), J_n allows weight to be put on all observations. This choice is particularly suited to income studies because every observation is relevant for an accurate measurement of inequality. In this sense, see Friedlin *et al.* (2003) for an illuminating application on very heavy-tailed distributions of the Chernoff *et al.* (1967) results on L-statistics. Moreover, the score function converges suitably to a real valued function $J(t)$ defined on $(0, 1)$ in such a way as to allow one to replace J_n with the limiting weight function $J(t)$. Table 1 presents the score function J_n of some indices falling into class (1). The scores are obtained by setting $t = i/(n+1)$ in $J_n(t)$. $F_n(.)$ is the empirical distribution function based on $\{X_1, X_2, \ldots, X_n\}$; the symbol $[x]$ is the largest integer less than or equal to x; $\{.\}$ denotes an indicator function which is one if the argument is true and zero otherwise. The Pietra-Ricci index is defined as one-half of the relative mean deviation; the Gini/2 index is given by the absolute mean deviation from the median divided by the arithmetic mean. The Eltetö-Frigyes index is defined as $(\mu_n - M_n)/\mu_n$ where M_n is the mean income of individuals with an income smaller then μ_n (Eltetö and Frigyes, 1968). The Gini, Mehran, and Piesch indices belong to the class of linear indices proposed by Mehran (1976). The Piesch index was also suggested by Giaccardi (1950). According to Sen and Singer (1996, 193-194), the Piesch-Giaccardi (Mehran) index measures the total inequality by looking at the propensity to fail of a system consisting of three units in parallel (series). Consequently, $P \leq M$. The Bonferroni index is obtained as the unweighted average of the relative differences between the mean and the partial means of the poorest units. The index proposed by De Vergottini is obtained as the weighted average of the relative differences between the mean and the partial means of the richest units. All but the Amato index and De Vergottini index lie in the interval from zero to unity. It must be noted that the inequality indices proposed by Amato, Bonferroni, and the De Vergottini are special cases of a general formula discussed by Amato (1948) which includes the linear indices as a subclass.

TABLE 1: *Score function of various measures belonging to class (1).*

Index	$J_n(t)$	$J(t)$
Gini (R)	$\left(\frac{n+1}{n-1}\right)(2t-1)$	$2t-1$
Pietra Ricci (D)	$F_n(\mu_n) - I\{t < F_n(\mu_n)\}$	$F(\mu) - I\{t < F(\mu)\}$
Eltetö Frygies (D_2)	$1 - I\{t < F_n(\mu_n)\}$	$1 - I\{t < F(\mu)\}$
Gini/2 (G_2)	$\mathrm{sgn}(t-0.5)$	$\mathrm{sgn}(t-0.5)$
Piesch (P)	$\frac{3[\lfloor(n+1)t\rfloor - 1][\lfloor(n+1)t\rfloor - 2]}{2(n-1)(n-2)} - \frac{1}{2}$	$(3t^2-1)/2$
Mehran (M)	$\frac{3[n - \lfloor(n+1)\rfloor][(n-1) - \lfloor(n+1)t\rfloor]}{(n-1)(n-2)}$	$1-3(1-t)^2$
Amato (A)	$1 - \frac{\sum_{j=\lfloor(n+1)t\rfloor}^{n} j^{-0.5}(n+1-j)^{-0.5}}{n^{-1}\sum_{j=1}^{n} j^{0.5}(n+1-j)^{-0.5}}$	$\frac{4}{\pi}\arcsin(t) - 1$
Bonferroni (B)	$\frac{n}{n-1}\left(1 - \sum_{j=\lfloor(n+1)t\rfloor}^{n} j^{-1}\right)$	$1 - Ln(t)$
De Vergottini (dV)	$\sum_{j=\lfloor(n+1)(1-t)\rfloor}^{t} j^{-1} - 1$	$-[1 + Ln(1-t)]$

The large sample results concerning the indices of Table 1 are contained in two fundamentals theorems.

Theorem S1. [Stigler (1974), Mason (1981)].
Assumption 1)

i) $J_n(t)$ *is uniformly bounded and continuous on* $[0, 1]$.
ii) $J(t)$ *is continuous except possibly at a finite number of points at which* $F^{-1}(t)$ *is continuous.*
iii) $J_n(t) \to J(t)$ *uniformly in an open neighborhood of* t *as* $n \to \infty$.
iv) $J(t)$ *satisfies the Hölder condition of order* $\alpha > 0.5$.
v) $\int J(t)dt = 0$.

Assumption 2)

$$\int \sqrt{F(x)[1-F(x)]}dx < \infty\,.$$

If assumptions 1) *and* 2) *are satisfied then*

a) $\operatorname*{Lim}_{n\to\infty} E[L(J_n)] = L(J) = \int_0^1 J(t)F^{-1}(t)dt\,;$

b) $$\operatorname*{Lim}_{n\to\infty} n\sigma^2[L(J_n)] = \sigma^2(J) = 2\iint_{0<t<s<1} J(t)J(s)t(1-s)dF^{-1}(t)dF^{-1}(s) \quad (2)$$

c) $\sqrt{n}\left[\dfrac{L(J_n)-L(J)}{\sigma[L(J_n)]}\right] \xrightarrow{d} N(0,1) \quad as \quad n\to\infty\,; \;\; if \quad \sigma > 0$

Theorem S2 [Shorack -Wellner (1986, 662-665)].
Assumption 1)

For some $M > 0$ *and* $B(t) = Mt^{-b_1}(1-t)^{-b_2}$ *for* $0 < t < 1$ *with* $\max(b_1, b_2) < 1$.

i) $| J_n(t) | \leq B(t)$.
ii) $J(t)$ *is continuous except possibly at a finite number of points at which* $F^{-1}(t)$ *is continuous.*
iii) $J_n(t) \to J(t)$ *uniformly in some small neighborhood of* t *as* $n \to \infty$.
iv)
   1) $| J(t) | \leq B(t)$*;*
   2) $J'(t)$ *exists on* $(0, 1)$ *and* $| J'(t) | \leq B(t)[t(1-t)]^{-1}$ *for* $0 < t < 1$.
v) $\int J(t)dt = 0$.

Assumption 2)

$$|F^{-1}(t)| < Mt^{-d_1}(1-t)^{-d_2} \;\; for \;\; 0 < t < 1\,, \max\{(b_1+b_2), (d_1+d_2)\} < 0.5\,. \quad (3)$$

If assumptions 1) *and* 2) *are satisfied then the conclusions of S1 still hold.*

Assumption 2 is cryptic in both theorems although the more familiar requirement of $E(|X|^{2+d}) < \infty$, $d > 0$ suffices in order that F should verify this assumption both in S1 (Singh, 1981) and in S2 (Shorack, 1972). We should consider, however, that the importance of these theorems relies on the fact that the variance of $L(J_n)$ converges even though the underlying distribution may not have a finite variance.

The asymptotic properties of (1) can be established by using a two-step procedure. First, (2) is ascertained for $L(J_n)$ via S1 or via S2, and then the Slutsky's theorem is applied to the ratio $L(J_n)/\mu_n$ (see *e.g.* Randles and Wolfe, 1991, 424-426). The asymptotic variance of the ratio (1) can be approximated by the methods outlined in Palmitesta *et al.* (1999) and Zitikis (2002a).

The conditions of Theorem S1 are fulfilled by the indices of Amato, Gini, Mehran and Piesch-Giaccardi. Theorem S1 is also applicable to the indices of Pietra-Ricci, Gini/2, and Eltetö-Frigyes provided that the corresponding $J(t)$ and $F^{-1}(t)$ possess no common discontinuities. Cf. Gastwirth (1974) for the asymptotic results concerning the Pietra-Ricci index. Theorem S2 extends (2) to the indices of Bonferroni and De Vergottini whose weight functions are not bounded but verify condition iv) of S2.

Helmers (1981) showed that under the conditions of theorem S1 (in particular, assumption iv.1 with $\alpha = 1$), the existence of a finite absolute third moment allows us to establish a Berry-Esséen type bound of order $n^{-0.5}$ for the indices belonging to class (1). Friedrich (1989) obtained a Berry-Esséen bound for an unbounded weight function, but at the cost of requiring the existence of a finite fourth moment.

3. A new class of inequality measures

Inequality comparisons are always subject to arbitrary specifications of weights. For the purpose of large-sample theory, the weight function J of an income inequality measure should have a reasonably smooth behavior. The purpose of this section is to show that the weight function can be conveniently specified by means of the order statistics of a given random variable.

A quantile plot method

The measurement of income inequality may receive a valuable stimulus from a graphical technique, similar in spirit to a QQ plot (see Shapiro and Brain, 1980). Let Z be a nondegenerate random variable with distribution function $G(z)$ such that $E(Z) < \infty$. We further assume that G does not contain unknown parameters. The model G represents the hypothesized distribution of income and may be chosen to direct attention to certain levels of the income variable.

Suppose that one individual has income X and that X is included in a random sample of size $n : \{Z_1, Z_2, \dots, Z_n\}$ drawn from G. If these observations are ranked in ascending order: $Z_{i:n}$ for $i = 1, 2, \dots, n$ with $E(Z_{i:n}) = m_{i:n}$, then our individual is prepared to occupy the position r whose expected income, under the model G, is nearest to X. More formally, $\|X - m_{r:n}\| \leq \|X - m_{i:n}\|$ for $i = 1, 2, \dots, n$ where $\|\,.\,\|$ is a fixed metric. Let $\{X_1, X_2, \dots, X_n\}$ be a random sample of incomes from F, the "true", but unknown distribution of incomes and consider the scatterplot with coordinates

$$\left\{\left(\frac{m_{i:n} - m_n}{\tau_n}\right), \left(\frac{X_{i:n} - \mu_n}{\mu_n}\right)\right\}; \quad i = 1, 2, \dots, n\,;$$
$$\tau_n = \frac{\sum_{i=1}^{n}(m_{i:n} - m_n)^2}{n(m_{n:n} - m_n)}\,; \quad m_n = n^{-1}\sum_{i=1}^{n} m_{i:n} \tag{4}$$

where μ_n is the arithmetic mean of the $\{X_{i:n}\}$. The rationale behind the plot is that if the plotting points lie near a straight line through the origin this provides an informal validation of G to represent the observed data. Departure from G would be indicated by increased curvature. If the linearity of the plot is satisfactory, then the least squares estimate of the slope

$$V(J_n) = \frac{n^{-1}\sum_{i=1}^{n} J_n\left(\frac{i}{n+1}\right) X_{i:n}}{n^{-1}\sum_{i=1}^{n} X_{i:n}} \tag{5}$$

with

$$J_n(t) = \frac{m_{\lfloor (n+1)t \rfloor:n} - m_n}{m_{n:n} - m_n} \qquad 0 \leq t \leq 1 \tag{6}$$

can measure the inequality in the sample distribution of incomes. First, $V(J_n)$ is scale invariant because it does not change if all $\{X_{i:n}\}$ are multiplied by a positive factor. Second, $V(J_n)$ obeys the Pigou-Dalton principle which implies that any transfer from a richer to a poorer individual which does not alter the ranking (neutral or order-preserving transfer) reduces inequality. In fact, the effect of a neutral transfer of $d > 0$ from the individual in the r-th position to that in the s-th position, with $s < r$, is $d(m_{s:n} - m_{r:n})/(m_{n:n} - m_n) < 0$. Third, from the fact that both the $\{X_{i:n}\}$ and the $\{m_{i:n}\}$ are nondecreasing it can be seen that $V(J_n) \geq 0$. Also, since $m_{i:n} \leq m_{n:n}$ then $V(J_n) \leq 1$. The lower limit $V(J_n) = 0$ can generally be interpreted as a tendency to the situation in which the $\{X_{i:n}\}$ are close together. The upper limit $V(J_n) = 1$ is achieved as n diverges while only one of the individuals gets all the income. Finally, the weight function J_n is a linear transformation of the income expected under

G which reveals the normative preferences to the measurement of income inequality implicit in any $V(J_n)$. Giaccardi (1950) defined a class of inequality measures similar to (5)-(6) in which $\{m_{i:n}\}$ is a bounded monotone increasing sequence of non negative constants.

In light of the high correlation and heteroschedasticity of the order statistics, the usual distributional results for the slope of a regression line do not apply. However, recalling that $J_n(t)$ is continuous and monotone nondecreasing for $0 < t < 1$, the asymptotic behavior of $V(J_n)$ can be assessed following the approach outlined in Section 2.

Of course, (5) does not directly assess the linear nature of the plot. The judgement of what can and cannot be considered a good approximation to a straight line is a personal matter and should depend on the magnitude of the departure from linearity and on the size of the sample. In case of doubt, a more formal decision rule can be used *e.g.* the correlation coefficient or its square.

Example 1. Let $G(z) = z$ be the uniform distribution over the unit interval. Then

$$m_{i:n} = \frac{i}{n+1}\,; \qquad m_n = \frac{1}{2}\,; \qquad m_{n:n} - m_n = \frac{n-1}{2(n+1)} \tag{7}$$

yields the Gini index which can now be regarded as a measure of inequality appropriate in a situation in which the *ex ante* proportion of individuals receiving an income less than or equal to z is directly proportional to z. One of the referees has called to my attention a paper by Cohen and Keppler (1996) where this model has been applied to analyze the relationship concerning firm size and research and development effort.

Example 2. Let the probability mass for Z be equally distributed at two points: $\{-1, 1\}$

$$m_{i:n} = \begin{cases} -1 & \text{if} \quad i \le n/2 \\ 1 & \text{if} \quad i > n/2 \end{cases}\,; \qquad m_n = 0\,, \quad m_{n:n} - m_n = 1 \tag{8}$$

This scheme leads to the Gini/2 index. The Pietra-Ricci index has an analogous interpretation. Note that when J_n is sectionally constant over $(0, 1)$ then $V(J_n)$ fails the Pigou-Dalton principle because (5) becomes insensitive to transfers between individuals associated with weights of equal sign and equal magnitude.

Example 3. If Z is the unit exponential distribution $G(z) = 1-e^{-z}$, $z > 0$ then

$$m_{i:n} = \sum_{j=n-i+1}^{n} j^{-1}\,; \qquad m_n = 1\,; \qquad m_{n:n} - m_n = \left(\sum_{j=2}^{n} j^{-1}\right) \tag{9}$$

define the De Vergottini index (more specifically, the version ranging between zero and one, denoted by dV*). The scores appearing in (9) show interesting aspects of the index. For example, if we assume that each unit possesses one of a set of n incomes and that these incomes are randomly distributed among infinite population, the minimum proportion of sampling units that must be included in order to obtain i different incomes coincides with $m_{i:n}$ in the De Vergottini index.

Example 4. Let G be the reflected unit exponential distribution $G(z) = e^z$, $z < 0$. In this case z represents a loss and not an income. It follows that

$$m_{i:n} = -\sum_{j=i}^{n} j^{-1}\,; \qquad m_n = -1\,; \qquad m_{n:n} - m_n = \frac{n-1}{n}\,; \tag{10}$$

corresponding to the Bonferroni index. The scores (10), termed log-rank scores, were introduced by Savage (1956). From another point of view, the i-th weight of the Bonferroni index represents the selection differential (in terms of mean units) obtained by comparing the mean of the bottom i ordered incomes with the total mean.

Example 5. Kadane (1971) showed that $m_{i:n} = fs^{i-1}$, $i = 1, 2, \ldots, n$ where s and f are parameters such that $s \geq 1$, $f > 0$ are the expectations of order statistics for a sample from some positive random variable with $E(Z) < \infty$. Let $s = \alpha^{1/(n+1)}$, $\alpha > 1$, $f = s$. Then

$$k_n(\alpha) = \frac{\sum_{i=1}^{n} \left[\frac{\alpha^{\frac{i}{n+1}} - m_n}{\alpha^{\frac{n}{n+1}} - m_n}\right] X_{i:n}}{\sum_{i=1}^{n} X_{i:n}} \quad \text{with} \quad m_n = n^{-1}\sum_{i=1}^{n} \alpha^{\frac{i}{n+1}} \to \frac{\alpha - 1}{Ln(\alpha)} = m \tag{11}$$

Since the limit weight function $J(t) = (\alpha^t t - m)/(\alpha - m)$ is convex, Theorem S1 applies to (11) for each $\alpha > 1$. $k_n(\alpha)$ is "rightist" in the sense that transfers at the upper end of the distribution are weighted more heavily than transfers at the lower end. The larger is α the greater is concern for those on higher incomes. In practical applications, one has to pick a certain value of $\alpha > 1$. Any such choice implies a further specification of the weight function. It can be shown (see the section on large sample approximation) that (11) is related to the logarithmic model $F(z) = Ln[(\alpha - m)z + m]/Ln(\alpha)$ for $(\alpha - m)^{-1} \leq z \leq 1$.

Numerical application

We illustrate the method for a sample of moderate size. Mehran (1981) reports $n = 200$ incomes data obtained from the results of a household survey where $\mu_n = 9438.3$, $\sigma_n = 9019.6$, $\gamma_1 = 2.8$, $\gamma_2 = 10.9$, median$= 6350.5$, range$= (600, 67403)$. Figure 1a-1d provide the presentation of the quantities in (4) for the following indices: Gini (R), Bonferroni (B), De Vergottini (dV) and $k_n(1.1)$. The marked curvature in the first two graphs contradicts the assumed model. A close (and surprising) resemblance between R and $k_n(1.1)$ is also apparent. The QQ plot against the scores of the De Vergottini index is highly linear except, perhaps, for two observations in the upper tail (the coefficient of determination is $R^2 = 0.97$). This suggests a linear transformation of the exponential distribution to be reasonable for the Mehran data. The corresponding level of the inequality is dV*$= 0.1892$.

The message that these examples convey is that one may select the best G distribution out of a finite family of admissible distributions which yields the most linear QQ-plot. In other words, if the hypothesized model and the true model of income are a linear transformation of each other, then an accurate examination of several QQ-plots can help to select the model of the size distribution of income. However, it is not our purpose to claim that a simple plot can capture simultaneously these two fundamental features of the distribution of income. Our modest proposal is to introduce a useful tool which can lead to the identification of stable structures in the income inequality of empirical distributions.

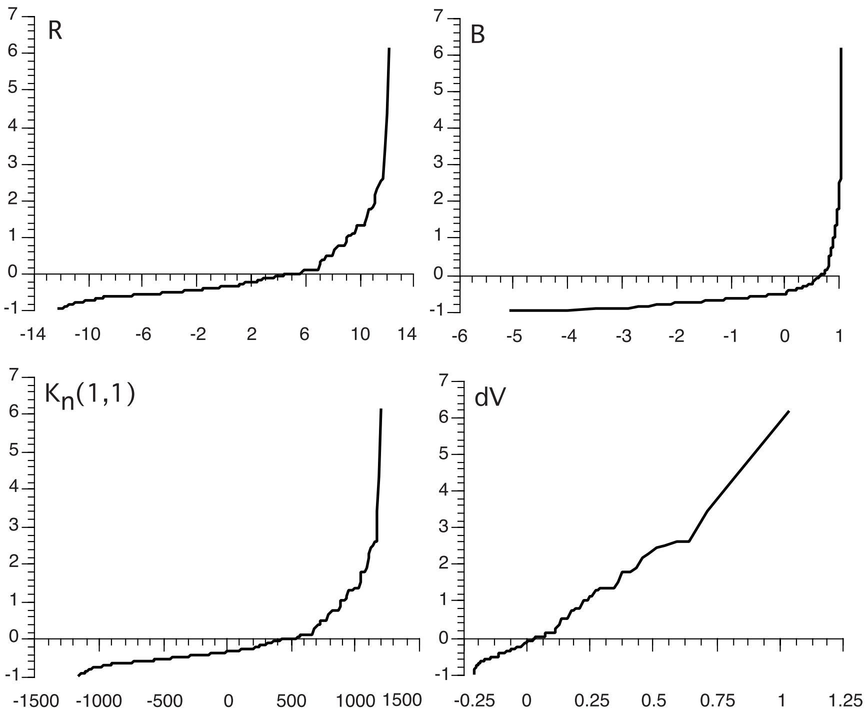

Figure 1. QQ-plot for four inequality measures.

Approximation of the expected value of order statistics

The proposed methodology may be difficult to apply because explicit algebraic expressions for $\{m_{i:n}\}$ and τ_n are rarely available. To alleviate this problem somewhat, we derive the weight function directly from the quantile function of the presumed model $G(z)$. To this end we further assume that Z is an absolutely continuous random variable with associated density function which is supposed to be continuous and strictly positive on the support of Z. This allows us to exploit the well known convergence result $m_{(n+1)t:n} \rightarrow G^{-1}(t)$. Since the size of the samples encountered in income distribution analysis is fairly large, this approximation should be satisfactory. To simplify the search for the weight function we restrict ourselves to those random variables Z with a limited range. Also, we impose the normalizing constraint $G^{-1}(1) \leq b$. Thus, the restricted class of indices is defined as follows

$$V(J_n) = \frac{n^{-1}\sum_{i=1}^{n} J\left(\frac{i}{n+1}\right) X_{i:n}}{\mu_n}$$
$$\text{with} \quad J\left(\frac{i}{n+1}\right) = \frac{G^{-1}\left(\frac{i}{n+1}\right) - m_n}{b - m_n}, \quad m_n = n^{-1}\sum_{i=1}^{n} G^{-1}\left(\frac{i}{n+1}\right) \tag{12}$$

The asymptotic properties of (11) follow from Theorem S1.

Example 6. If $G(z) = \sin(z)$, $0 \leq z \leq \pi/2$ then (12) produces the Amato index in Table 1.

$$J(t) = \frac{4}{\pi}\sin^{-1}(t) - 1 \tag{13}$$

It is easily seen that (13) is monotone increasing and satisfies the principle of diminishing transfers " ... *One values more such transfer between persons with given income difference if these incomes are lower than if they are higher*". Kolm (1976, p.87). This is equivalent to $J'(t) > 0$, $J''(t) < 0$ for all $0 \leq t \leq 1$ if the derivatives exist (Mehran, 1976). However, because of high concentration on lower incomes, the Amato index may assume negative values. Moreover, the value $A = 0$ does not necessarily represent perfect income equality since $A = 0$ for a uniform distribution too.

Example 7. Generalized beta of the first type: $G(z) = \left\{1 - [(1-z)]^{1/\alpha_2}\right\}^{1/\alpha_1}$, $0 \leq z \leq 1$ with $J(t) = 1 - c(1 - t^{\alpha_1})^{\alpha_2}$, $c = \alpha_1 / B(\alpha_1^{-1}, \alpha_2 + 1)$ $\alpha_1, \alpha_2 > 0$ where $B(x, y)$ is the complete beta function

$$T_n(\alpha_1, \alpha_2) = \frac{n^{-1}\sum_{i=1}^{n}\left\{1 - c_n^{-1}\left[1 - \left(\frac{i}{n+1}\right)^{\alpha_1}\right]^{\alpha_2}\right\} X_{i:n}}{\mu_n},$$
$$c_n = n^{-1}\sum_{i=1}^{n}\left[1 - \left(\frac{i}{n+1}\right)^{\alpha_1}\right]^{\alpha_2} \tag{14}$$

For $0 \leq \alpha_1 \leq 1$, $\alpha_2 > 1 T_n$ fulfills the Kolm principle. If $\alpha_1 = 1$ then T_n produces the generalization of the Gini index proposed by Donaldson and Weymark (1980) and analyzed by Zitikis and Gastwirth (2002). The parameter combination $(\alpha_1 = 1, \alpha_2 = 2)$ determines the Mehran index with $J(t) = 1 - 3(1-t)^2$ and $(\alpha_1 = 2, \alpha_2 = 1)$ gives the Piesch-Giaccardi index P with $J(t) = (3t^2 - 1)/2$. Note that P violates the Kolm principle because $J''(t) > 0$.

Example 8. For the generalized lambda distribution $G^{-1}(t) = t^{\alpha_1} - (1-t)^{\alpha_2}$, $0 \leq t \leq 1$,

$$\lambda_n(\alpha_1, \alpha_2) = \frac{\frac{1}{n(1-m_n)} \sum_{i=1}^{n} \left[\frac{B(\alpha_1, n+1)}{B(\alpha_1, i)} - \frac{B(\alpha_2, n+1)}{B(\alpha_2, n+1-i)} - m_n \right] X_{i:n}}{\mu_n}, \tag{15}$$
$$m_n = \frac{\Gamma(n+1)}{n} \sum_{i=1}^{n} \left[\frac{\Gamma(\alpha_1 + i)}{\Gamma(n+1+\alpha_1)\Gamma(i)} - \frac{\Gamma(\alpha_2 + n + 1 - i)}{\Gamma(n+1+\alpha_2)\Gamma(n+1-i)} \right]$$

The parameters α_1 and α_2 determine the shape of the score function. In particular, if $\alpha_1 = \alpha_2$ then $J(t)$ is symmetric about zero which means that a similar weight (of opposite sign) is attached to ordered incomes at equal distances from the extremes. In particular, the distribution with $\alpha_1 = \alpha_2 = 0.1349$ is practically indistinguishable from $N(0, \sigma)$ with $\sigma = 1.46357$ (Mudholkar *et al.*, 1991). For $0 \leq \alpha_1 \leq 1$, $1 \leq \alpha_2 < \infty$, (15) satisfies both the conditions of Theorem S1 and the Kolm principle.

4. A Monte Carlo experiment

In this section we attempt to gather the following information. The sample size beyond which the asymptotic results of class $V(J_n)$ become applicable to finite samples. The sensitivity of the indices when a change in the underlying distribution of income takes place. The peculiarities of the score function J_n that generate lower bias, smaller relative variability and faster rate of convergence to the normal distribution.

Models of income distribution

Giorgi and Pallini (1987, 1990), Palmitesta *et al.* (1999) carried out various experiments to assess the asymptotic behavior of Gini, Mehran and Piesch-Giaccardi indices in samples drawn from a Burr/3 distribution. See also Giorgi and Provasi (1995), Palmitesta *et al.* (2000). To the best of our knowledge, no Monte Carlo study has been performed on the Bonferroni and Gini/2.

Our Monte Carlo experiment involves two special cases of the generalized beta distribution of the second type introduced by McDonald and Xu (1985)

$$GB(x; a, b, c, d, e) = \frac{|a| x^{ad-1}\left[1-(1-c)(\frac{x}{b})^a\right]^{e-1}}{b^{ad} B(d, e)\left[1+c(\frac{x}{b})^a\right]^{d+e}} \tag{16}$$

$$\text{for} \quad 0 < x^a < \frac{b^a}{1-c}; \; 0 \leq c < 1, \; b, d, e \geq 0$$

which has a high flexibility of shape and provides realistic frameworks to test the inequality indices. The first model is the well known Burr/3 (or Dagum type 1) with $a = -\gamma^1$, $b = 1$, $c \rightarrow 1$, $d = 1$, $e = \delta^{-1}$ (if $\delta = 1$ then γ coincides with the Gini index). The other model is the beta density corresponding to (16) with $a = 1$, $b = 1$, $c = 0$, $d = \delta$, $e = \gamma$ which has a compact support. In this case the income X should be interpreted as a normalized value: $X = (y - y_{\min})/(y_{\max} - y_{\min})$ where $y_{\min}$ and $y_{\max}$ denote the lower and upper limit (known in advance) over the whole population considered. The closer the values of δ and γ are, the more symmetrical the density. The condition $\delta < \gamma$ suffices to ensure that the Beta density is positively skewed. The parameter combinations are reported in the following table:

TABLE 2: *values of* (γ, δ) *chosen for the simulations*

Models		C1	C2	C3	C4	C5	C6	C7
Dagum/1	δ	1.00000	1.00000	2.00000	2.00000	3.00000	3.00000	4.00000
	γ	0.15000	0.20000	0.20000	0.25000	0.25000	0.30000	0.30000
Beta	δ	6.94600	4.07300	2.58600	1.71900	1.17900	0.85100	0.63300
	γ	7.43300	5.10200	3.74600	2.84300	2.06900	1.71100	1.35400
	Gini	0.15000	0.20000	0.25373	0.31103	0.36603	0.42418	0.47357

It can be easily checked that for the combinations of parameters considered in Table 2, the corresponding distribution functions possess a finite moment of order 3 so that assumption 2 is satisfied both for Theorem S1 and for Theorem S2. Moreover, the models have the same value of the Gini index that is, the Gini index varies across the curves, but not across the models.

Simulation plan

The design consisted of generating $N = 25000$ samples of various sizes: $(n = 200, 500, 1000, 2000, 3000, 4000, 8000)$ for the seven parameter combinations with a total of 98 distinct experiments. The exponential spacings method proposed by Lurie and Mason (1973) was first used to generate uniform order statistics $U_{i:n}$, $i = 1, 2, \ldots, n$ and then the inversion method determined $\{X_{i:n}, i = 1, \ldots, n\}$ in a random sample from a Burr/3 distribution:

$X_{i:n} = [(U_{i:n})^{-\delta} - 1]^{-\gamma}$. To simulate the order statistics in a random sample from a beta distribution the independent variates $(X_1, X_2, \ldots, X_n)$ were generated according to the algorithms described in Rubinstein (1981, procedures Be-3 and Be-4) and sorted into natural order as $\{X_{i:n}, i = 1, \ldots, n\}$ by using Hoare's Quicksort. Since the simulation plan analyzes sequences longer than 2×10^8, the fast multiple recursive generator (FMRG) proposed by Deng and Lin (2000) was applied to produce uniform pseudorandom number on the unit interval (all the sequences were initialized from the same seeds). The FMRG has a period of $(2^{62} - 1)$ but yields values lower than 2^{31} so that repetitions in the same sample are probable. This is in contrast to the hypothesis of order statistics from continuous distributions in which the probability of ties among the observations is zero. However, because of the large discrepancy between the number of potential calls and the effective number of values used in a single sample, this drawback should not interfere constructively with the reliability of the simulations.

All the software has been written in Future Basic 3 running on a G4 (one processor, 400MHz) computer using MacOS 9.2 operating system. Program codes as well as numerical results are available from the author on request.

Assessing performance

Two simple coefficients of performance were considered for comparison

$$Mb(\theta) = \frac{1}{N}\left(\sum_{i=1}^{N}\left|\frac{\theta_i - \theta_0}{\theta_0}\right|\right) \;;\; \text{with } \theta_0 = \frac{\int_0^1 J(t)GB^{-1}(t; a, b, c, d, e)dt}{\int_0^1 GB^{-1}(t; a, b, c, d, e)dt} \tag{17}$$

$$Cv(\theta) = \sqrt{\frac{\sum_{i=1}^{N}\left(\frac{\theta_i - \theta_n}{\theta_n}\right)^2}{N}} \quad \text{where} \quad \theta_n = \frac{\sum_{i=1}^{N}\theta_i}{N} \tag{18}$$

The relative mean bias $Mb(\theta)$ quantifies the average magnitude of the estimator's accuracy and $Cv(\theta)$ reflects the estimator's variation from sample to sample. The computation of θ_0, *i.e.* the value of the given inequality index in the parent distribution, was carried out by standard numerical integration routines.

Eight indices were selected for inclusion in the study, that is $\theta \in \Theta = \{R, D, G_2, P, M, B, \lambda(0.8, 4), T(0.75, 2)\}$. The Amato index and De Vergottini index were not considered because they do not lie in the unit interval. The Eltetö-Frygies index was not included because all the essentials contained in the indices proposed by Eltetö and Frygies (1968) can be condensed into the Pietra-Ricci index. In fact, $1/(1 - D_2) = \mu_n/M_n$ (Kondor, 1971). The two indices $\lambda(0.8, 4)$, $T(0.75, 2)$ were chosen after a preliminary study (not reported here) involving 117 combinations of the parameters (α_1, α_2): $[0.1(0.1)0.9, 1(0.25)4]$

on the basis of the best values of (17) and (18) for the proposed sample sizes. Both $\lambda(0.8, 4)$ and $T(0.75, 2)$ are “leftist” giving more weight to transfers affecting lower income earners. This choice has the effort of giving less importance to observations in the upper tail than do other measures counteracting the fact that extreme incomes, though rare, have considerable effects on the sampling distributions. Moreover, since the top order statistics of the income random variable do not contribute too much to $\lambda(0.8, 4)$ and $T(0.75, 2)$, their convergence to the normal distribution should be faster than that of the other indices.

Results

For reason of space only a selection of the results can be presented here. Table 3, in two parts, shows the mean value (Ev), the mean bias (17) and the coefficient of variation (18) for the eight measures of inequality over the $N = 25000$ replications.

The general behavior of all the indices follows foreseeable lines. Consistent with asymptotic results, the absolute bias becomes smaller and the coefficient of variation diminishes as sample size increases for each $\theta \in \Theta$ although, the gain in accuracy and in low variance decays as the number of sampling units grows up. Table 3 reveals that the higher the level of inequality expressed by θ is, the lower $Mb(\theta)$ and $Cv(\theta)$ are for each θ (the estimated values of the population mean exhibit the reverse of this condition). To some extent this would be expected since (17) and (18) depend on the size of θ. To obtain a mean-independent comparison the quantities

$$\text{Relative Bias} = \frac{\sum_{j=1}^{7} Mb_j[\theta(\gamma, \delta)]}{7}, \qquad \text{Relative variation} = \frac{\sum_{j=1}^{7} Cv_j[\theta(\gamma, \delta)]}{7} \tag{19}$$

were computed over the seven curves for each model and reported in Table 4. The findings of this table confirm the tendencies already shown in Table 3. In fact, both the relative bias and the relative variation of all θ’s declines as the sample size tends to infinity and we could name several measures having similar optimum performance at least from a sample size $n \geq 500$. Also, we can note that the parent distribution has scarce or no influence on the behavior of the indices as far as mean bias and relative variation are concerned except that $Mb(\theta)$ and $Cv(\theta)$ associated with the beta model are slightly smaller than those obtained from the Burr/3 model.

TABLE 3a: *results fo the Burr/3 model.*

δ	γ	n		R	D	G_2	P	M	B	λ	T
2	0.2	200	EV	0.2536	0.1766	0.3493	0.2003	0.3595	0.3689	0.3293	0.3899
		500		0.2537	0.1769	0.3502	0.2002	0.3606	0.3707	0.3314	0.3918
		1000		0.2537	0.1769	0.3503	0.2001	0.3609	0.3712	0.3320	0.3924
		2000		0.2537	0.1769	0.3504	0.2000	0.3610	0.3714	0.3323	0.3926
		3000		0.2537	0.1770	0.3505	0.2000	0.3612	0.3716	0.3325	0.3928
		4000		0.2538	0.1770	0.3506	0.2001	0.3612	0.3717	0.3326	0.3929
		8000		0.2537	0.1770	0.3506	0.2000	0.3612	0.3717	0.3326	0.3929
		200	MB	0.0152	0.0106	0.0201	0.0138	0.0189	0.0182	0.0174	0.0198
		500		0.0096	0.0067	0.0126	0.0087	0.0119	0.0114	0.0110	0.0125
		1000		0.0068	0.0047	0.0090	0.0062	0.0085	0.0081	0.0078	0.0089
		2000		0.0048	0.0034	0.0064	0.0044	0.0060	0.0057	0.0055	0.0063
		3000		0.0039	0.0028	0.0058	0.0038	0.0055	0.0048	0.0044	0.0056
		4000		0.0034	0.0024	0.0045	0.0031	0.0042	0.0040	0.0039	0.0044
		8000		0.0024	0.0017	0.0032	0.0022	0.0030	0.0029	0.0028	0.0031
		200	CV	0.0600	0.0601	0.0574	0.0691	0.0527	0.0494	0.0528	0.0508
		500		0.0378	0.0377	0.0360	0.0435	0.0331	0.0309	0.0331	0.0318
		1000		0.0269	0.0268	0.0257	0.0310	0.0235	0.0219	0.0235	0.0226
		2000		0.0190	0.0189	0.0181	0.0219	0.0166	0.0154	0.0166	0.0160
		3000		0.0155	0.0155	0.0148	0.0179	0.0136	0.0126	0.0136	0.0130
		4000		0.0133	0.0133	0.0127	0.0154	0.0117	0.0109	0.0117	0.0112
		8000		0.0095	0.0094	0.0090	0.0109	0.0083	0.0077	0.0083	0.0080
4	0.3	200	EV	0.4731	0.3403	0.6531	0.3912	0.6357	0.6207	0.5790	0.6722
		500		0.4734	0.3408	0.6548	0.3912	0.6374	0.6220	0.5822	0.6748
		1000		0.4735	0.3409	0.6552	0.3911	0.6379	0.6224	0.5832	0.6757
		2000		0.4734	0.3409	0.6554	0.3910	0.6381	0.6226	0.5837	0.6760
		3000		0.4736	0.3410	0.6556	0.3911	0.6385	0.6228	0.5839	0.6762
		4000		0.4736	0.3411	0.6557	0.3912	0.6384	0.6228	0.5841	0.6764
		8000		0.4736	0.3411	0.6557	0.3911	0.6384	0.6228	0.5842	0.6764
		200	MB	0.0253	0.0188	0.0296	0.0262	0.0248	0.0220	0.0234	0.0237
		500		0.0161	0.0118	0.0187	0.0167	0.0157	0.0139	0.0148	0.0150
		1000		0.0116	0.0085	0.0134	0.0120	0.0112	0.0099	0.0106	0.0107
		2000		0.0082	0.0060	0.0095	0.0085	0.0079	0.0070	0.0075	0.0076
		3000		0.0071	0.0053	0.0088	0.0073	0.0064	0.0062	0.0068	0.0062
		4000		0.0058	0.0042	0.0066	0.0060	0.0056	0.0049	0.0053	0.0053
		8000		0.0041	0.0030	0.0047	0.0043	0.0040	0.0035	0.0038	0.0038
		200	CV	0.0535	0.0552	0.0454	0.0670	0.0390	0.0354	0.0403	0.0353
		500		0.0340	0.0347	0.0285	0.0428	0.0246	0.0223	0.0255	0.0222
		1000		0.0244	0.0248	0.0204	0.0308	0.0176	0.0160	0.0182	0.0158
		2000		0.0173	0.0175	0.0144	0.0218	0.0124	0.0113	0.0129	0.0112
		3000		0.0140	0.0142	0.0117	0.0177	0.0101	0.0092	0.0105	0.0091
		4000		0.0122	0.0123	0.0101	0.0154	0.0087	0.0079	0.0091	0.0079
		8000		0.0087	0.0087	0.0072	0.0109	0.0062	0.0056	0.0065	0.0056

TABLE 3b: *results fo the Beta model.*

δ	γ	n		R	D	G_2	P	M	B	λ	T
2.586	3.746	200	EV	0.2540	0.1828	0.3645	0.1942	0.3726	0.3807	0.3388	0.4062
		500		0.2538	0.1830	0.3652	0.1938	0.3736	0.3822	0.3407	0.4081
		1000		0.2538	0.1830	0.3655	0.1936	0.3739	0.3827	0.3414	0.4087
		2000		0.2538	0.1831	0.3656	0.1936	0.3740	0.3829	0.3417	0.4090
		3000		0.2538	0.1831	0.3657	0.1936	0.3741	0.3830	0.3418	0.4091
		4000		0.2537	0.1831	0.3657	0.1935	0.3741	0.3831	0.3419	0.4091
		8000		0.2537	0.1831	0.3657	0.1935	0.3741	0.3831	0.3419	0.4092
		200	MB	0.0383	0.0422	0.0419	0.0411	0.0367	0.0339	0.0364	0.0360
		500		0.0243	0.0268	0.0266	0.0260	0.0233	0.0213	0.0229	0.0227
		1000		0.0171	0.0188	0.0187	0.0183	0.0164	0.0150	0.0161	0.0159
		2000		0.0121	0.0134	0.0133	0.0130	0.0116	0.0106	0.0113	0.0113
		3000		0.0098	0.0109	0.0108	0.0105	0.0094	0.0086	0.0092	0.0091
		4000		0.0085	0.0094	0.0094	0.0091	0.0082	0.0075	0.0080	0.0080
		8000		0.0060	0.0067	0.0067	0.0064	0.0058	0.0053	0.0056	0.0056
		200	CV	0.0480	0.0531	0.0527	0.0513	0.0461	0.0423	0.0451	0.0448
		500		0.0304	0.0335	0.0333	0.0325	0.0291	0.0266	0.0285	0.0283
		1000		0.0214	0.0236	0.0234	0.0229	0.0205	0.0187	0.0201	0.0199
		2000		0.0152	0.0167	0.0166	0.0162	0.0145	0.0132	0.0142	0.0141
		3000		0.0123	0.0136	0.0135	0.0132	0.0118	0.0108	0.0115	0.0115
		4000		0.0107	0.0118	0.0117	0.0115	0.0103	0.0094	0.0100	0.0100
		8000		0.0076	0.0084	0.0083	0.0081	0.0072	0.0066	0.0071	0.0070
0.633	1.354	200	EV	0.4743	0.3581	0.6989	0.3796	0.6625	0.6375	0.5948	0.7021
		500		0.4739	0.3585	0.7004	0.3787	0.6638	0.6382	0.5975	0.7044
		1000		0.4737	0.3586	0.7008	0.3784	0.6642	0.6384	0.5984	0.7052
		2000		0.4736	0.3586	0.7011	0.3782	0.6644	0.6385	0.5988	0.7055
		3000		0.4736	0.3586	0.7011	0.3781	0.6645	0.6386	0.5990	0.7056
		4000		0.4736	0.3587	0.7012	0.3781	0.6645	0.6386	0.5990	0.7057
		8000		0.4736	0.3587	0.7013	0.3781	0.6646	0.6386	0.5992	0.7058
		200	MB	0.0329	0.0386	0.0337	0.0395	0.0262	0.0223	0.0259	0.0238
		500		0.0211	0.0248	0.0216	0.0253	0.0168	0.0143	0.0164	0.0151
		1000		0.0149	0.0175	0.0152	0.0178	0.0118	0.0100	0.0115	0.0106
		2000		0.0104	0.0123	0.0107	0.0125	0.0083	0.0070	0.0081	0.0075
		3000		0.0086	0.0101	0.0088	0.0103	0.0068	0.0058	0.0066	0.0061
		4000		0.0074	0.0088	0.0076	0.0089	0.0059	0.0050	0.0058	0.0053
		8000		0.0052	0.0062	0.0054	0.0062	0.0042	0.0035	0.0040	0.0037
		200	CV	0.0411	0.0484	0.0422	0.0492	0.0329	0.0280	0.0320	0.0296
		500		0.0263	0.0309	0.0269	0.0314	0.0210	0.0178	0.0203	0.0188
		1000		0.0186	0.0219	0.0191	0.0223	0.0148	0.0126	0.0144	0.0133
		2000		0.0131	0.0154	0.0134	0.0156	0.0104	0.0088	0.0101	0.0093
		3000		0.0107	0.0126	0.0110	0.0128	0.0086	0.0073	0.0083	0.0077
		4000		0.0093	0.0110	0.0096	0.0111	0.0074	0.0063	0.0072	0.0067
		8000		0.0065	0.0077	0.0067	0.0078	0.0052	0.0044	0.0051	0.0047

TABLE 4: *Overall measure of bias and relarive variation.*

		n	R	D	G_2	P	M	B	λ	T
Burr/3	Bias	200	0.0472	0.0478	0.0435	0.0557	0.0394	0.0369	0.0404	0.0376
		500	0.0300	0.0302	0.0274	0.0356	0.0248	0.0232	0.0253	0.0235
		1000	0.0215	0.0215	0.0195	0.0255	0.0177	0.0165	0.0180	0.0167
		2000	0.0152	0.0152	0.0138	0.0181	0.0125	0.0116	0.0127	0.0118
		3000	0.0132	0.0128	0.0124	0.0147	0.0111	0.0092	0.0108	0.0095
		4000	0.0107	0.0107	0.0097	0.0127	0.0088	0.0082	0.0089	0.0083
		8000	0.0076	0.0076	0.0069	0.0091	0.0062	0.0058	0.0064	0.0059
	Var.	200	0.0600	0.0604	0.0547	0.0711	0.0495	0.0463	0.0502	0.0469
		500	0.0379	0.0379	0.0343	0.0450	0.0311	0.0291	0.0316	0.0294
		1000	0.0270	0.0270	0.0245	0.0321	0.0222	0.0206	0.0225	0.0209
		2000	0.0191	0.0191	0.0173	0.0227	0.0156	0.0146	0.0159	0.0148
		3000	0.0157	0.0162	0.0163	0.0191	0.0134	0.0122	0.0127	0.0134
		4000	0.0134	0.0134	0.0121	0.0160	0.0110	0.0103	0.0112	0.0104
		8000	0.0096	0.0095	0.0086	0.0114	0.0078	0.0073	0.0079	0.0074
Beta	Bias	200	0.0368	0.0413	0.0397	0.0407	0.0338	0.0308	0.0334	0.0326
		500	0.0233	0.0263	0.0253	0.0258	0.0214	0.0195	0.0210	0.0206
		1000	0.0164	0.0187	0.0179	0.0182	0.0151	0.0138	0.0148	0.0145
		2000	0.0117	0.0134	0.0128	0.0129	0.0107	0.0099	0.0105	0.0103
		3000	0.0095	0.0111	0.0106	0.0105	0.0088	0.0081	0.0085	0.0084
		4000	0.0082	0.0097	0.0092	0.0091	0.0076	0.0071	0.0074	0.0073
		8000	0.0058	0.0072	0.0067	0.0064	0.0054	0.0052	0.0052	0.0052
	Var.	200	0.0460	0.0517	0.0497	0.0507	0.0424	0.0384	0.0414	0.0405
		500	0.0292	0.0328	0.0315	0.0322	0.0268	0.0243	0.0262	0.0256
		1000	0.0206	0.0231	0.0223	0.0227	0.0189	0.0171	0.0185	0.0181
		2000	0.0146	0.0164	0.0158	0.0161	0.0134	0.0121	0.0131	0.0128
		3000	0.0119	0.0134	0.0129	0.0131	0.0109	0.0099	0.0107	0.0105
		4000	0.0103	0.0116	0.0111	0.0114	0.0095	0.0085	0.0092	0.0091
		8000	0.0073	0.0082	0.0079	0.0080	0.0067	0.0060	0.0065	0.0064

As we have seen in Section 3, all the indices considered for the Monte Carlo experiment were shown to be asymptotically normally distributed. The asymptotic results however, are more readily applicable if we can establish how fast the convergence is taking place. Since the population distribution function is fixed, the rate of convergence depends only on the limit weight function J which can be used to investigate the asymptotic properties of (5). Table 5a and 5b show the results of testing

$$\begin{cases} H_0 : \psi \left[\dfrac{\theta - E(\Theta)}{\sigma(\theta)} \right] = N(0,1) \\[2ex] H_1 : \psi \left[\dfrac{\theta - E(\theta)}{\sigma(\theta)} \right] \neq N(0,1) \end{cases} \tag{20}$$

with the χ^2 goodness-of-fit test where $\psi(\,.\,)$ represents the distribution of the inequality index. More specifically, the 25000 repetitions of each $\theta \in \Theta$ were grouped into 40 non overlapping intervals each having an expected number of frequencies of 1250. The class limits of the empirical distributions of the inequality indices were chosen such that all theoretical probabilities are equal to 1/40 which gives equal importance to all partition of the Gaussian model. The observed values of the Chi-square greater than the threshold value $\chi^2_{0.05}(37) = 52.1923$, were replaced by "*" meaning that H_0 cannot be accepted at 5% level of significance. The Shapiro-Wilk test W would have been more akin to the issues discussed in this section than the χ^2. Nevertheless, our simulation plan considers $N = 25000$ replications of the quantity to be tested for normality and the computation of W becomes very imprecise as the number of cases increases. Royston (1992) gave an acceptable approximation for N in the interval [3-5000]. It is generally observed that, with sporadic exceptions, the tested empirical distributions are acceptably normal when the model underlying the sample values is the beta distribution and sample sizes of $n \geq 1000$ are available. The leftist measures B, M, $\lambda(0.8, 4)$ and $T(0.75, 2)$ perform well throughout. The normal model is also a plausible representation of the empirical distribution of the Pietra- Ricci index and Gini/2 index. The insufficient agreement between $\psi(\theta)$ and $N(0, 1)$ for smaller sized samples and low or moderate levels of inequality (curve C1-C3 in Table 2) should be ascribed to the slow convergence of the standard error of (5).

Under the Burr/3 model the performance of most indices deteriorates appreciably and a great variation is observed in the values of χ^2_c across levels of inequality and across sample sizes. Specifically, the $N(0, 1)$ was found to be appropriate for the empirical distribution of B, M and $\lambda(0.8, 4)$ for $n \geq 2000$. Also, D and G_2 have a sampling distribution which shows an encouraging degree of fitting to the normal whereas the empirical distribution for R and P contradict the postulated model. For high levels of inequality (curves C4-C7) the convergence to the normal is barely evident. In particular, R is very unreliable (except for the largest samples) and P exhibits a χ^2_c systematically greater than the others. On the other hand, the sampling behavior of B, M, $\lambda(0.8, 4)$ and $T(0.75, 2)$ appears to be quite close to what asymptotic theory suggests.

TABLE 5a: *asymptotic normality under the Burr/3 model.*

	n	R	D	G_2	P	M	B	λ	T
C1	200	*	*	*	*	*	47.38	*	*
C2	200	*	*	*	*	*	*	*	*
C3	200	*	*	*	*	45.44	*	*	40.01
C4	200	*	*	*	*	*	46.45	*	*
C5	200	*	*	*	*	34.00	28.93	*	*
C6	200	*	*	*	*	*	*	*	44.40
C7	200	*	*	48.47	*	*	*	*	45.60
C1	500	*	*	48.36	*	49.98	*	*	39.97
C2	500	*	*	*	*	*	*	*	*
C3	500	*	*	48.72	*	*	43.14	45.25	32.89
C4	500	*	*	*	*	50.29	45.57	*	39.09
C5	500	*	*	*	*	35.77	50.42	*	37.87
C6	500	*	*	*	*	*	*	*	45.36
C7	500	*	*	*	*	44.54	*	*	35.40
C1	1000	*	*	*	*	*	*	*	*
C2	1000	*	*	*	*	*	*	*	*
C3	1000	*	51.24	*	*	43.49	45.29	*	*
C4	1000	*	*	47.63	*	*	45.70	*	*
C5	1000	*	48.28	45.32	*	36.60	*	38.71	39.05
C6	1000	*	*	51.73	*	49.37	*	*	49.40
C7	1000	*	*	47.36	*	36.95	*	*	40.14
C1	2000	37.05	35.58	49.87	*	21.03	31.09	43.59	42.64
C2	2000	51.08	44.34	47.96	*	37.02	30.47	25.68	21.79
C3	2000	39.71	32.52	40.70	48.83	42.26	*	37.44	47.06
C4	2000	*	41.03	30.06	*	46.40	35.08	47.88	25.00
C5	2000	*	42.40	40.25	*	44.89	33.91	40.87	34.72
C6	2000	*	*	38.32	*	47.69	*	*	29.69
C7	2000	*	50.12	48.99	*	26.58	29.52	35.73	33.26
C1	3000	38.91	26.29	44.78	48.72	47.74	29.55	30.96	32.78
C2	3000	43.50	35.32	26.90	*	30.60	40.05	32.24	45.47
C3	3000	45.22	37.70	32.72	*	38.11	36.63	49.23	40.84
C4	3000	39.52	39.49	39.79	*	22.36	33.60	37.15	44.49
C5	3000	*	51.53	34.46	*	35.65	32.86	23.28	24.75
C6	3000	*	*	32.23	*	30.80	41.15	41.37	29.45
C7	3000	*	39.29	24.14	*	31.73	37.09	28.51	51.51
C1	4000	45.94	30.04	32.86	46.63	31.30	23.88	32.73	40.58
C2	4000	*	46.89	*	*	46.04	35.75	41.25	32.60
C3	4000	41.31	43.54	27.58	*	43.04	44.77	*	29.65
C4	4000	*	*	49.46	*	43.24	36.29	34.85	46.27
C5	4000	48.50	41.33	30.31	*	42.65	45.30	46.22	33.89
C6	4000	*	*	35.49	*	38.08	51.00	*	34.31
C7	4000	*	48.85	24.54	*	38.34	51.83	48.90	31.53
C1	8000	37.59	40.73	45.80	44.06	44.80	41.39	38.51	39.75
C2	8000	42.20	47.61	41.19	*	47.53	46.55	50.63 *	
C3	8000	46.06	47.25	39.87	*	39.19	33.71	46.64	33.89
C4	8000	39.81	42.82	*	*	35.40	42.13	37.72	28.09
C5	8000	36.67	28.84	36.09	39.81	34.90	31.12	51.18	28.02
C6	8000	*	50.36	36.04	*	44.40	48.03	*	37.18
C7	8000	*	34.33	34.06	*	36.38	32.08	40.94	*

TABLE 5b: *asymptotic normality under the beta model.*

	n	R	D	G_2	P	M	B	λ	T
C1	200	36.89	38.22	41.23	48.55	34.52	49.00	35.42	52.22
C2	200	40.65	41.17	24.25	40.00	*	51.86	44.38	46.43
C3	200	32.99	*	*	*	26.69	33.36	33.40	42.59
C4	200	37.15	39.97	35.74	47.13	41.00	28.07	39.03	38.95
C5	200	46.88	*	*	43.79	41.40	43.96	36.35	42.42
C6	200	43.83	37.34	36.56	43.01	*	*	*	*
C7	200	30.20	36.97	40.79	36.39	45.51	*	51.57	*
C1	500	43.36	43.77	44.61	*	49.04	50.55	33.09	48.74
C2	500	* v34.50	36.52	37.96	32.91	35.13	31.45	36.51	
C3	500	36.39	29.20	43.16	27.85	27.27	33.56	25.96	27.36
C4	500	40.03	37.02	40.01	29.59	44.63	49.16	*	41.23
C5	500	35.16	37.80	31.39	39.02	38.28	42.58	39.47	40.81
C6	500	25.88	31.07	46.44	28.92	*	48.19	*	44.03
C7	500	*	35.73	*	46.70	51.62	*	*	*
C1	1000	37.45	40.50	46.54	2.84	43.14	33.34	47.02	32.09
C2	1000	49.06	37.79	34.54	36.77	23.13	47.15	30.33	29.43
C3	1000	28.95	39.10	38.95	29.80	*	29.75	34.49	36.63
C4	1000	41.36	40.96	29.56	33.12	42.52	28.72	34.20	35.07
C5	1000	*	39.06	37.51	*	48.73	38.64	48.12	39.67
C6	1000	38.12	19.16	48.23	44.10	26.53	31.55	31.44	44.46
C7	1000	23.62	25.01	37.90	35.57	42.81	38.82	36.45	34.18
C1	2000	47.56	41.13	44.53	29.39	35.81	27.36	31.27	45.03
C2	2000	20.73	*	46.86	28.48	36.02	25.15	32.19	39.08
C3	2000	25.15	37.39	34.27	34.85	45.61	39.25	41.12	44.89
C4	2000	50.31	27.76	46.36	30.59	33.42	37.43	32.15	26.04
C5	2000	*	40.20	39.63	37.00	23.20	42.53	29.56	24.08
C6	2000	27.83	23.64	36.25	33.11	25.94	30.44	32.60	34.55
C7	2000	40.58	30.02	36.31	45.63	23.14	24.06	17.05	26.44
C1	3000	26.62	51.05	41.56	26.23	44.82	35.76	40.71	46.46
C2	3000	34.48	37.06	46.38	26.35	31.31	29.10	30.28	36.09
C3	3000	51.06	42.86	32.68	27.21	37.78	43.24	36.56	41.82
C4	3000	34.91	45.34	35.93	44.21	36.48	42.70	35.14	24.84
C5	3000	34.74	34.38	26.93	33.15	32.68	32.49	29.96	31.03
C6	3000	30.08	*	49.56	29.82	33.88	48.04	39.40	32.44
C7	3000	39.54	32.78	34.13	35.80	31.70	34.84	27.71	36.26
C1	4000	27.31	29.58	32.89	47.55	6.38	35.29	43.91	34.88
C2	4000	17.01	33.68	35.13	37.26	24.60	41.12	28.90	36.04
C3	4000	30.39	42.76	33.53	23.07	39.11	34.06	28.18	34.22
C4	4000	27.91	33.61	27.46	26.16	*	27.35	37.22	42.26
C5	4000	*	42.02	23.96	45.34	34.36	33.93	31.08	36.73
C6	4000	29.24	20.05	20.68	35.48	29.48	39.02	33.64	33.86
C7	4000	41.57	36.25	40.44	52.18	*	42.67	40.53	44.15
C1	8000	27.86	40.82	37.61	49.52	29.53	26.26	29.34	28.09
C2	8000	26.48	26.50	42.99	36.39	47.83	36.94	32.46	36.39
C3	8000	23.30	24.88	32.74	39.71	30.56	30.84	29.86	37.06
C4	8000	30.03	48.58	46.39	32.99	27.75	40.65	36.39	27.77
C5	8000	51.40	43.69	39.51	42.87	38.07	48.29	48.87	36.26
C6	8000	27.22	24.84	37.16	26.91	42.73	32.10	38.92	32.96
C7	8000	41.45	38.22	45.63	42.18	38.20	35.65	38.28	14.18

5. Concluding remarks

This article proposes the expected values of order statistics from a given random variable as weights of a new class of inequality measures. Also, a useful variant of the QQ-plot has been introduced. A Monte Carlo experiment has highlighted both the computational feasibility of the new indices and their strength and shortcomings.

The general conclusions that can be drawn, at least in the framework of the models considered in the simulations, are the following.

i) Although there appears to be a common thread among the indices of inequality, their sampling properties are not invariant. Specifically, the approach to normality for the Bonferroni index, Mehran index, $\lambda(0.8, 4)$, $T(0.75, 2)$ is faster than for the other indices.

ii) From a statistical point of view, the boundness of the weight function appears to be a factor of minor relevance for the choice of an inequality measure.

iii) The fashion of a “simple form” for the weight function of an income inequality index is not defensible since the relatively complex formulae for $\lambda(0.8, 4)$ and $T(0.75, 2)$ have often obtained better results than the more elegant and computationally simple expressions for the indices of Gini, Piesch-Giaccardi and Mehran.

iv) The importance given to inequality measures like the Pietra-Ricci index and Gini/2 index in empirical work is not questionable on the ground of their sampling accuracy which, on the contrary, seems to be very reasonable.

v) A question of both theoretical and practical interest is how the inequality measures would vary their performance as the parent population undergoes change. Our findings indicate that the performance of an index is unlikely to be affected seriously by the underlying model provided that the skewness of the potential distribution functions is at comparable level and the size of the sample is 8000 or larger.

vi) If the choice of an income-weighting scheme is based on a strictly statistical background, then the Bonferroni index, Mehran index, $\lambda(0.8, 4)$ and $T(0.75, 2)$ should be preferred because their asymptotic properties are better than those of the indices which lack the property of diminishing transfers. Thus considerations of both economic theory (the Kolm principle) and statistical theory point to the need to use leftist weight function to devise an income inequality measure.

ACKNOWLEDGMENTS

I wish to thank anonymous referees for their helpful comments that substantially improved this paper.

REFERENCES

AMATO, V. (1948) Sulla misura della concentrazione dei redditi, *Rivista Italiana di Demografia e Statistica*, 2, 504-529.

CHERNOFF, H., GASTWIRTH, J. L., AND JOHNS, M. V. (1967) Asymptotic distribution of linear combination of function of order statistics with applications to estimation, *The Annals of Mathematical Statistics*, 38, 52-62.

COHEN, W. M. AND KLEPPER, S. (1996) A reprise of size and R & D, *The Economic Journal*, 106, 925-951.

DENG, L.Y. AND LIN, D.K.J. (2000) Random number generation for the new century, *The American Statistician*, 54, 45-150.

DONALDSON, D. AND WEYMARK, J. A. (1980) A single-parameter generalization of the Gini indices of inequality, *Journal of Economic Theory*, 22, 67-86.

ELTETÖ, O. AND FRYGIES, E. (1968) New income inequality measures as efficient tools for causal analysis and planning, *Econometrica*, 36, 383-396.

FRIEDLIN, B., MIAO, W., AND GASTWIRTH J.L. (2000) On the use of the Shapiro-Wilk test in two-stage adaptive inference for paired data from moderate to very heavy tailed distributions, *Biometric Journal*, 7, 887-900.

FRIEDRICH, K. O. (1989) A Berry-Esséen bound for functions of independent random variables, *The Annals of Statistics*, 17, 170-183.

GASTWIRTH, J.L. (1974) Large sample theory of some measures of income inequality, *Econometrica*, 42, 191-196.

GIACCARDI, F. (1950) Indici di concentrazione, *Giornale degli Economisti ed Annali di Economia*, Anno IX (nuova serie), 282-290.

GIORGI, G. M. AND PALLINI, A. (1987) On the speed of convergence for the normal distribution of some inequality indices, *Metron*, 45, 93-105.

GIORGI, G. M. AND PALLINI, A. (1990) Inequality indices: theoretical and empirical aspects of their asymptotic behavior, *Statistical Papers-Statistiche Hefte*, 31, 65-76.

GIORGI, G. M. AND PROVASI, C. (1995) An experimental analysis to verify some properties of inequality indices of the Gini family, *In: Quantitative methods for applied sciences*, Nuova Immagine, Siena, 254-285.

GIORGI, G. M. (1999) Income inequality measurement: The statistical approach (with a comment by M.P. Victoria Feser), In: *Handbook on Income Inequality Measurement*, (J. Silber ed.), Recent Economic Thought series 71, Kluwer, Boston, 245-267.

HELMERS, R. (1981) A Berry-Esséen theorem for linear combinations of order statistics, *The Annals of Probability*, 5, 940-953.

KADANE, J. B. (1971) A moment problem for order statistics, *The Annals of Mathematical Statistics*, 42, 745-751.

KAKWANI, N. (1980) *Income inequality and poverty: Methods of estimation and policy applications*, (World Bank Research Publications), Oxford University Press, Oxford.

KOLM, S. C. (1976) Unequal inequalities, II, *Journal of Economic Theory*, 13, 82-111.

KONDOR, Y. (1971) An old-new measure of income inequality, *Econometrica*, 39, 1041-1042.

LURIE, D. AND MASON, L. (1973) Empirical investigation of several techniques for computer generation of order statistics, *Communication in Statistics*, 2, 363-371.

MASON, D. M. (1981) Asymptotic normality of linear combinations of order statistics with a smooth score function, *The Annals of Statistics*, 9, 899-908.

MCDONALD, J. B. AND XU, Y. J. (1995) A generalization of the beta distribution with applications, *Journal of Econometrics*, 66, 133-152. Erratum: 69,427-428.

MEHRAN, F. (1976) Linear measures of income inequality, *Econometrica*, 44, 805-809.

MEHRAN, F. (1981) Optimal grouping of income distribution data, *Journal of the American Statistical Association*, 76, 22-26.

MUDHOLKAR, G. S., KOLLIA, G. D., LIN, C. T., AND PATEL K. R. (1991) A graphical procedure for comparing goodness-of-fit tests, *Journal of the Royal Statistical Society*, B, 53, 221-232.

NYGARD, F. AND SANDSTRÖM, A. (1981) *Measuring Income Inequality*, Almqvist & Wiksell, Stockholm.

PALMITESTA, P., PROVASI C., AND SPERA, C. (1999) Approximated distributions of sampling inequality indices, *Computational Economics*, 13, 211-216.

PALMITESTA, P., PROVASI, C., AND SPERA, C. (2000) Confidence interval estimation for inequality indices of the Gini family, *Computational Economics*, 16, 137-147.

RANDLES, M. AND WOLFE, D.A. (1991) *Nonparametric statistical methods*, John Wiley & Sons, New York.

ROYSTON, P. (1992) Approximating the Shapiro-Wilk W-test for non-normality, *Statistics and Computing*, 2, 117-119.

RUBINSTEIN, R. Y. (1981) *Simulation and the Monte Carlo method*, John Wiley & Sons, New York.

SAVAGE, I. R. (1956) Contribution to the theory of rank order statistics. The two sample case, *The Annals of Mathematical Statistics*, 27, 509-516.

SEN, A. (1997) *On economic inequality*, (Expanded edition with a substantial annexe by J.E. Foster and A. Sen.), Clarendon Press, Oxford.

SEN, P. K. AND SINGER, J. M. (1993) *Large sample methods in statistics. An introduction with applications*, Chapman & Hall, London.

SHAPIRO, S. S. AND BRAIN, C. W. (1980) A review of distributional testing procedures and development of a censored sample distributional test, In: *Statistical Distributions in Scientific Work*, (C. Tailli *et al.* eds), vol. 5, D. Reidel Publishing Company, Dordrecht, Holland, 1-24.

SHORACK, G. R. (1972) Functions of order statistics, *The Annals of Mathematical Statistics*, 43, 412-427.

SHORACK, G. R. AND WELLNER, J. A. (1986) *Empirical processes with applications to statistics*, John Wiley & Sons, New York.

SINGH, K. (1981) An asymptotic representation and approximation to normality of L-Statistics, *Sankhya*, A, 43, 67-83.

STIGLER, S. M. (1974) Linear functions of order statistics with smooth weight functions, *The Annals of Statistics*, 2, 676-693.

ZITIKIS, R. (2002a) Large sample estimation of a family of economic inequality indices, *Pakistan Journal of Statistics,* (Special issue in honour of Dr. S. Ejaz Ahmad), 18, 225-248.

ZITIKIS, R. (2002b) Analysis of indices of economic inequality from a mathematical point of view, (Invited Plenary Lecture at the 11th Indonesian Mathematics Conference, State University of Malang, Indonesia.) *Matematika*, 8, 772-782.

ZITIKIS, R. AND GASTWIRTH, J. T. (2002) The asymptotic distribution of the S-Gini index, *Australian and New Zealand Journal of Statistics*, 44, 439-446.

AGOSTINO TARSITANO
Università degli studi della Calabria
Dipartimento di Economia e Statistica
87030 Arcavacata di Rende (Italy)
agotar@unical.it