THE GINI COEFFICIENT AND POVERTY INDICES: SOME RECONCILIATIONS*

by

Pranab Kumar Sen

Department of Biostatistics
University of North Carolina at Chapel Hill

Institute of Statistics Mimeo Series No. 1484

May 1985

THE GINI COEFFICIENT AND POVERTY INDICES : SOME RECONCILIATIONS *

By PRANAB KUMAR SEN

The Gini coefficient plays a vital role in the formulation of the poverty indices. TTT-transformations are adopted to provide a new interpretation of the Gini index, and the same is incorporated in the formulation of a robust version of the Sen poverty index.

1. INTRODUCTION

Poverty is usually defined as the extent to which individuals in a community or society fall below a minimal acceptable standard of living. An *index of poverty* is generally based on the distribution $F(y)$, $y \in R^+ = [0,\infty)$, of the *income variable* and a set *poverty line* ω (> 0), so that

(1.1) $\alpha = F(\omega)$ = proportion of people below the poverty line ω

may be taken as a crude index. The *income gap ratio* of the poor people may then be defined by

(1.2) $\beta = 1 - \omega^{-1}\{ \alpha^{-1} \int_0^\omega y dF(y)\}$.

Generally, both α and β are taken into consideration in the formulation of a meaningful poverty index. Sen (1976) considered the following :

(1.3) $\pi_S = \alpha\{\beta + (1-\beta)G_\alpha\}$,

where G_α ($0 \leq G_\alpha \leq 1$) is the *Gini coefficient* of the income distribution censored at ω (as will be defined later on). Though this index has been used extensively, there remains some issues relating to its normative content. Takayama (1979), incorporating a somewhat different set of axioms, has argued that a proper measure of the poverty is

(1.4) $\pi_T = G_\alpha$, the Gini coefficient for $\alpha^{-1}F(x)$, $0 \leq x \leq \omega$.

Whereas, Sen (1976) advocated that an index should be given by the weighted agrregate gap of the people below the poverty line, the Takayama (1979) index

* Work partially supported by the Office of Naval Research, Contract No. N00014-83-K-0387.

is based solely on the income inequality of the censored income distribution truncated from above by ω . Blackorby and Donaldson (1980) have suggested another index :

(1.5) $\pi_{BD} = \alpha\beta$,

which does not involve the Gini coefficient G_α explicitly. An obvious drawback of (1.4) is its insensitiveness to α , and this may lead to rather misleading interpretations in some cases. For example, if all the people below the poverty line have an equal income, say, y_o ($0 \le y_o \le \omega$), then $G_\alpha = 0$, so that $\pi_T = 0$, irrespective of the particular y_o and α . A situation of this type may arise due to heavy unemployment (where y_o is close to 0) or due to some social welfare system providing indirect income below the poverty line and thereby leading to some clusters in the censored income distribution; in both the cases, G_α (and hence, π_T) may be a meaningless entity. As we shall see in Section 2, the indices β and G_α are interrelated, and hence, all the three indices π_S, π_T and π_{BD} are dependent on the Gini coefficient G_α in some way. But, the role of G_α is not the same in all of these cases. Note that $\pi_S \ge \pi_{BD}$ but π_S and π_T (or π_T ans π_{BD}) may not be comparable in this way.

The object of the present study is to focus on the role of G_α in the formulation of poverty indices. This examination leads us to a more robust variant of the Sen poverty index , viz.,

(1.6) $\pi^* = \alpha(\beta^{1-G_\alpha})$.

In this context, the Gini coefficient is critically examined in Section 2. The main results are then considered in Section 3 and some general remarks are also appended there.

2. THE GINI COEFFICIENT : A FRESH LOOK

First, we consider the Gini coefficient G for the uncensored distribution F. We assume that

(2.1) $\mu = EY = \int_0^\infty y dF(y)$ is finite.

Thus, defining (the *survival distribution*) $\bar{F}(x) = 1 - F(x),\ x \in R^+$, we have

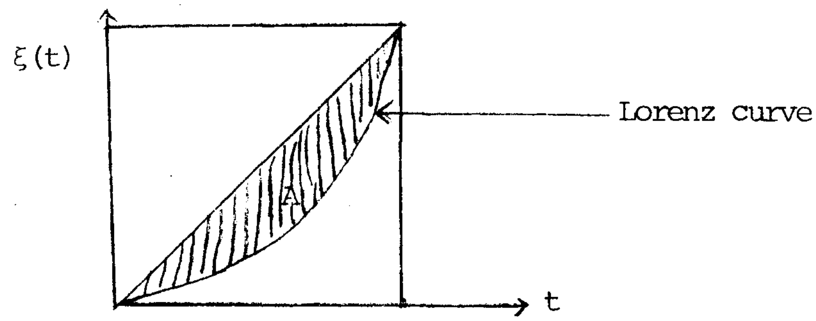
(2.2) $$\mu = \int_0^\infty \bar{F}(y)dy .$$

Let then

(2.3) $$F^{-1}(t) = \inf\{ x : F(x) \geq t \} ,\ 0 \leq t \leq 1,$$

(2.4) $$\xi(t) = \mu^{-1} \int_0^{F^{-1}(t)} y dF(y) ,\quad 0 \leq t \leq 1.$$

We may consider the *Lorenz curve* :



Then, in terms of the shaded area A , we have the Gini coefficient

(2.5) $$G = 2A .$$

Analytically, G may also be expressed as

(2.6) $$G = (2\mu)^{-1} E| Y_1 - Y_2 |,$$

where Y_1 and Y_2 are independent random variables , both having the same distribution F. Note that

(2.7) $$|Y_1 - Y_2| = (Y_1 \vee Y_2) - (Y_1 \wedge Y_2),$$

where $a \vee b$ and $a \wedge b$ stand respectively for the maximum and minimum of a and b. Therefore, by (2.6) and (2.7), we obtain that

(2.8) $$\begin{aligned} G &= (2\mu)^{-1}\{ E(Y_1 \vee Y_2) - E(Y_1 \wedge Y_2) \} \\ &= 1 - \mu^{-1} E(Y_1 \wedge Y_2) \\ &= 1 - 2\mu^{-1} \int_0^\infty y \bar{F}(y) dF(y). \end{aligned}$$

At this stage, we borrow the concept of *total time on test (TTT)* from *reliability theory* [viz., Klefsjö (1983)], and consider the *scaled-TTT transformation* :

(2.9) $$\phi(t) = \mu^{-1} \int_0^{F^{-1}(t)} \bar{F}(u)du = \int_0^{F^{-1}(t)} \bar{F}(u)du \Big/ \int_0^\infty \bar{F}(u)du ,\ 0 \leq t \leq 1 .$$

In this context, we may also note that

(2.10) $\quad TF(x) = \mu^{-1}\int_0^x \bar{F}(u)du\ , \quad x \in R^+$

is called the *equilibrium renewal distribution* corresponding to F. Further, by (2.4) and (2.9),

(2.11) $\quad \phi(t) = \xi(t) + \mu^{-1}(1-t)F^{-1}(t)\ , \ t \in [0,1].$

Let then

(2.12) $\quad \bar{\phi} = \int_0^1 \phi(t)dt = \int_0^\infty TF(x)dF(x) =$ *scaled-TTT mean,*

so that

(2.13)
$$\begin{aligned}\bar{\phi} &= \mu^{-1}\int_0^\infty \{\int_0^x \bar{F}(u)du\}dF(x)\\ &= \mu^{-1}\{\iint_{0\le u\le x<\infty} \bar{F}(u)dudF(x)\}\\ &= \mu^{-1}\{\int_0^\infty \bar{F}^2(u)du\ \}\\ &= -\ 2\ \mu^{-1}\{\int_0^\infty u\bar{F}(u)d\bar{F}(u)\}\\ &= 2\ \mu^{-1}\{\int_0^\infty u\ \bar{F}(u)dF(u)\ \}\\ &= 1 - G\ .\end{aligned}$$

Thus, the scaled-TTT mean $\bar{\phi}$ and the Gini coefficient G are complementary to each other. This representation not only provides additional insight into G, but also is quite useful in our subsequent analysis.

To extend the result to the censored case, we set

(2.14) $\quad F_\alpha(x) = \alpha^{-1}F(x)\ ,\ 0 \le x \le \omega \quad (\ F_\alpha(x) = 1,\ \forall\, x \ge \omega\ )\ ,$

(2.15) $\quad \bar{F}_\alpha(x) = 1 - F_\alpha(x) = 1 - \alpha^{-1}F(x)\ ,\ 0 \le x \le \omega,\quad \bar{F}_\alpha(x) = 0, \forall\, x \ge \omega\ ;$

(2.16) $\quad \mu_\alpha = \int_0^\infty xdF_\alpha(x) = \int_0^\omega xdF_\alpha(x) = \alpha^{-1}\int_0^\omega ydF(y).$

Then, parallel to (2.10), we have

(2.17) $\quad TF_\alpha(x) = \mu_\alpha^{-1}\int_0^x \bar{F}_\alpha(u)du = \alpha^{-1}\mu_\alpha^{-1}\int_0^x\{\alpha - F(y)\}dy\ ,\ 0 \le x \le \omega,$

so that

(2.18)
$$\begin{aligned}\bar{\phi}_\alpha &= 1 - G_\alpha = -\ \mu_\alpha^{-1}\int_0^\omega yd\bar{F}_\alpha^2(y)\\ &= 2\ \mu_\alpha^{-1}\int_0^\omega y\bar{F}_\alpha(y)dF_\alpha(y)\\ &= \mu_\alpha^{-1}\,E(\ Y_{\alpha 1}\wedge Y_{\alpha 2}\ ),\end{aligned}$$

where $Y_{\alpha 1}$ and $Y_{\alpha 2}$ are independent random variables , both having the common distribution F_α . Note that by (2.18),

$$(2.19)\qquad \mu_\alpha\bar{\phi}_\alpha = \mu_\alpha(1 - G_\alpha) = E(Y_{\alpha 1} \wedge Y_{\alpha 2}) ,$$

while, by (1.2) and (2.16),

$$(2.20)\qquad \beta = 1 - \omega^{-1}\mu_\alpha = 1 - (1-G_\alpha)^{-1}\,\omega^{-1}E(Y_{\alpha 1} \wedge Y_{\alpha 2}),$$

so that

$$(2.21)\qquad (1-\beta)(1 - G_\alpha) = (1-\beta)\cdot\bar{\phi}_\alpha = 2\omega^{-1}\int_0^\omega y\bar{F}_\alpha(y)dF_\alpha(y) = \omega^{-1}E(Y_{\alpha 1} \wedge Y_{\alpha 2}) = \gamma_\alpha, \text{ say.}$$

Obviously,

$$(2.22)\qquad 0 \le \gamma_\alpha \le (1-\beta) \vee (1-G_\alpha) = (\omega^{-1}\mu_\alpha) \vee \bar{\phi}_\alpha \le 1 .$$

Note that γ_α is a scale-free measure of the inequality of the censored income distribution F_α . Clearly, when $\beta = 0$ (i.e., $\mu_\alpha = \omega$), F_α has the unit mass at ω , so that $\bar{\phi}_\alpha = 1$ (i.e., $G_\alpha = 0$), and hence, $\gamma_\alpha = 1$. Similarly, $\gamma_\alpha = 0$ when $\beta = 1$ (i.e., $\mu_\alpha = 0$). In the other extreme case, where F_α has the unit mass at some intermediate point y_o $(0 < y_o < \omega)$, $G_\alpha = 0$ and hence, γ_α $(= 1 - \beta = \omega^{-1}y_o)$ depends solely on the ratio $\omega^{-1}y_o$. This reflects the desirability of γ_α as an index. We shall make more comments on it in the next section. In passing, we may remark that

$$(2.23)\qquad \gamma_\alpha = 2\alpha^{-1}\int_0^1 u\{1 - \alpha^{-1}F(u\omega)\}\,dF(u\omega),$$

so that if there are m people below the poverty line ω and if $y_1^* \le \dots \le y_m^*$ stand for the ordered values of their incomes, we have

$$(2.24)\qquad \gamma_\alpha = 2\,\omega^{-1}m^{-2}\,\Sigma_{i=1}^m\,(m-i+1)y_i^* ,$$

$$(2.25)\qquad \mu_\alpha = m^{-1}\Sigma_{i=1}^m\,y_i^* ,$$

$$(2.26)\qquad G_\alpha = (\Sigma_{i=1}^m\,\Sigma_{j=1}^m\,|y_i^* - y_j^*|)/(2m\,\Sigma_{i=1}^m\,y_i^*) .$$

We shall find these expressions useful in our subsequent analysis.

3. POVERTY INDICES : ROBUSTIFICATION

To start with, we may note that by (2.10) and (2.16),

$$(3.1)\qquad \mu_\alpha = -\alpha^{-1}\int_0^\omega x d\bar{F}(x) = -\alpha^{-1}\{[x\bar{F}(x)]_0^\omega - \int_0^\omega \bar{F}(x)dx\} = \mu\alpha^{-1}TF(\omega) + \omega - \omega\alpha^{-1} .$$

Therefore, by (1.2) and (3.1), we have

(3.2) $\beta = \alpha^{-1}\{ 1 - \mu\omega^{-1}TF(\omega)\}$,

so that by (1.2), (1.3), (2.21) and (3.2), we have

(3.3) $$\begin{aligned}\pi_S &= \alpha \{\beta + (1 - \beta)G_\alpha \} \\ &= \alpha \{1 - (1- \beta)(1- G_\alpha)\} \\ &= \alpha\{ 1 - \gamma_\alpha \} \\ &= \alpha\{ 1 - \omega^{-1}E(Y_{\alpha 1} \wedge Y_{\alpha 2})\} .\end{aligned}$$

This shows that in the Sen poverty index, both $\omega^{-1}\mu_\alpha$ (= $1-\beta$) and $\bar{\phi}_\alpha$ (= $1-G_\alpha$) play an equally important role. This feature is, however, not shared by the other two indices , π_T and π_{BD} . This also reveals that

(3.4) $$\begin{aligned}\alpha \geq \pi_S &\geq \{(1 - (1-\beta)) \vee (1-(1-G_\alpha)) \} \\ &= \alpha \{ \beta \vee G_\alpha \} \\ &= \pi_{BD} \vee (\alpha.\pi_T) \\ &= \pi_{BD} \vee \pi_T^* \text{ , say.}\end{aligned}$$

Note that γ_α is small if either β or G_α is close to 1 , while if both β and G_α are close to 1, then γ_α is of second order smallness. This may make π_S somewhat more inflated when both β and G_α are close to 1 , a case that may arise when there is a higher degree of poverty. Therefore, there seems to be a genuine need to curb the sensitivity of π_S when both β and G_α are close to 1. To illustrate this point, we consider a hypothetical case where $\beta = G_\alpha = 0.75$, so that $\gamma_\alpha = 0.0625$. Then $\pi_T^* = 0.75\alpha$, $\pi_{BD} = 0.75\alpha$ while $\pi_S = 0.9375\alpha$ and $\pi_T = 0.75$. The insensitivity of π_T to α makes it somewhat less appealing than π_S or π_{BD} , while π_S and π_{BD} differ by as much as 25% . We have noticed earlier that for $\mu_\alpha = 0$ (i.e., $\beta = 1$), γ_α is equal to 0 , so that π_S and π_{BD} both equal to α . Also, for $\beta = 0$, we have $G_\alpha = 0$, so that both π_S and π_{BD} are equal to 0. Consider a third case where the censored income distribution F_α has two clusters at 0 and ω with respective probability masses 1-p and p. In such a case, we have $\mu_\alpha = p$, $\bar{\phi}_\alpha = p$, $\beta = 1 - p$ and $\gamma_\alpha = p^2$, so that $\pi_S = \alpha(1-p^2)$ and $\pi_{BD} = \alpha(1-p) = \pi_T^*$. Thus $\pi_S > \pi_{BD} = \pi_T^*$ and the extent of divergence depends on p .

It also reveals that π_T^* is more appealing than the Takayama index π_T. Before we present a robust version, we consider the following:

Theorem 1. *For every* $\alpha : 0 < \alpha \leq 1$ *and every censored distribution* F_α,

(3.5) $$\alpha \geq \pi_S \geq \pi_{BD} \geq \pi_T^* \geq 0 ,$$

where all the three indices are equal to α *or 0 when* μ_α *is equal to 0 or* ω. *For* $0 < \mu_\alpha < \omega$, π_S *and* π_{BD} *are different from 0 while* π_T^* *may be equal to 0.*

Proof. By virtue of (3.4), to prove the set of inequalities in (3.5), it suffices to show that

(3.6) $$\beta \geq G_\alpha , \text{ for every } F_\alpha \text{ and } 0 < \alpha \leq 1.$$

Towards this, note that by (2.18) and (2.20),

(3.7) $$(1-\beta)/(1-G_\alpha) = -\mu_\alpha^2/\{\omega(\int_0^\omega y\bar{F}^2(y)dy)\}$$
$$= (\int_0^\omega \bar{F}_\alpha(y)dy)^2/\{(\int_0^\omega dy)(\int_0^\omega \bar{F}_\alpha^2(y)dy)\},$$

so that on using the Cauchy-Schwarz inequality on the numerator on the right hand side of (3.7), we immediately obtain that

(3.8) $(1-\beta)/(1-G_\alpha) \leq 1$, with the equlality sign holding only in the case of degenerate F_α,

and (3.8) ensures (3.6). When $\mu_\alpha = 0$, we have $\beta = 1$ and $\gamma_\alpha = 0$, so that $\pi_S = \pi_{BD} = \alpha$. Also, for $\mu_\alpha = 0$, there is a perfect equality of income of the people below the poverty line, but G_α is not properly defined. Conventionally, we may let G_α equal to 1, and this will make then $\pi_T^* = \alpha$. For $\mu_\alpha = \omega$, $\beta = 0$ and $G_\alpha = 0$, so that $\gamma_\alpha = 1$ and we have $\pi_S = \pi_{BD} = \pi_T^* = 0$. Note that for $0 < \mu_\alpha < \omega$, we have $0 < \beta \leq 1$, so that $\pi_{BD} > 0$. On the other hand, for $0 < \mu_\alpha < \omega$, we may have G_α arbitrarily close to 0, rendering π_T^* also close to 0. For example, if F_α has the unit mass at the point y_0 ($0 < y_0 < \omega$), we have $\beta = 1 - \omega^{-1}y_0$ (> 0), while G_α will be equal to 0, rendering that $\pi_T^* = 0$. Q.E.D.

Coming back to the inequality in (3.6), it is not difficult to construct suitable censored distributions F_α, for which β/G_α may be quite large. This is particularly true when F_α has one or few clusters which would make G_α quite small, while β need not be. As such, from the robustness considerations, we feel that π_T^* may not be a good competitor to π_S or π_{BD}. With this picture in mind, we would like to construct a robust variant of π_S. Note that by Theorem 1, π_S is bounded from above by α and from below by π_{BD}, where the two bounds agree when β is equal to 1. For β close to 0, π_S moves more towards π_{BD}, while, for β closer to 1, the upper bound is closer. In (3.6), we have further noticed that $\beta \geq G_\alpha$. Moreover, we may rewrite (1.3) as

(3.9) $$\pi_S = \alpha\{G_\alpha + (1-G_\alpha)\beta\} = \alpha G_\alpha + \pi_{BD}(1-G_\alpha).$$

Thus, π_S is an weighted arithmetic mean of the two bounds α and π_{BD}, where the relative weights are given by the Gini coefficient and its complement. While we do not want to change the role of the Gini coefficient in this context, we advocate the use of the geometric mean instead of the arithmatic mean, and thereby, propose the following index:

(3.10) $$\pi^* = \text{weighted geometric mean of } \alpha \text{ and } \pi_{BD}$$
$$= \alpha^{G_\alpha}\pi_{BD}^{1-G_\alpha} = \alpha(\beta^{1-G_\alpha}).$$

Note that as G_α goes to 0 (or 1), π^* converges to π_{BD} (or α). Also, using the well known inequality between the arithmetic and geometric means of nonnegative numbers, we readily obtain from (3.9) and (3.10) that

(3.11) $\pi_S \geq \pi^*$, where the equality sign holds when G_α is equal to 0 or 1, or when $\alpha = \pi_{BD}$.

Also, by virtue of the set of inequalities in (3.5),

(3.12) $$\pi_{BD} \leq \pi^* \leq \alpha,$$

where the equality signs hold when either G_α is 1 or 0, or $\alpha = \pi_{BD}$. Therefore, from (3.5), (3.11) and (3.12), we obtain that for every $\alpha : 0 < \alpha \leq 1$ and every censored distribution F_α,

(3.13) $$\alpha \geq \pi_S \geq \pi^* \geq \pi_{BD} \geq \pi_T^* \geq 0.$$

It is the median ordering of π^* that makes it more appealing as an index than π_S or π_{BD}. To illustrate this point, we consider some typical values of β and G_α and compare the three bounds.

TABLE 1

Table for the values of $\alpha^{-1}\pi_S$, $\alpha^{-1}\pi^*$ and $\alpha^{-1}\pi_{BD}$ for some typical (β, G_α).

β	G_α	$\alpha^{-1}\pi_S$	$\alpha^{-1}\pi^*$	$\alpha^{-1}\pi_{BD}$
0.1	0.05	0.145	0.112	0.100
0.1	0.10	0.190	0.126	0.100
0.2	0.10	0.280	0.235	0.200
0.2	0.20	0.360	0.276	0.200
0.3	0.20	0.440	0.382	0.300
0.3	0.25	0.475	0.405	0.300
0.4	0.20	0.520	0.481	0.400
0.4	0.30	0.580	0.527	0.400
0.5	0.30	0.650	0.616	0.500
0.5	0.40	0.700	0.660	0.500
0.6	0.40	0.760	0.736	0.600
0.6	0.50	0.800	0.793	0.600
0.7	0.50	0.850	0.837	0.700
0.7	0.60	0.880	0.859	0.700
0.8	0.50	0.900	0.895	0.800
0.8	0.60	0.920	0.915	0.800
0.9	0.60	0.960	0.959	0.900
0.9	0.70	0.970	0.969	0.900
0.95	0.70	0.985	0.985	0.950
0.95	0.80	0.990	0.990	0.950

A clear picture emerges for π_T^* as a middle runner between π_S and π_{BD} over the range of variation of β and G_α. Note that by virtue of (3.6), in Table 1, we have limited ourselves to values of G_α less than or equal to β. It is also clear that for small values of G_α (and β), π^* is closer to π_{BD}, while for G_α (and β) close to 1, π^* is closer to π_S than π_{BD}. To stress the robustness aspects of π^* a bit more, we consider again the situation where the censored income distribution F_α has two clusters at 0 and ω with respective probability masses 1-p and p, so that $\pi_S = (1-p^2)$ and $\pi_{BD} = (1-p)$. As here $1-G_\alpha = p$, we have $\pi^* = \alpha(1-p)^p \quad (\leq \alpha(1-p^2(1+p(1-p)/2))$. This shows that π^* is less affected by this clustering effect than π_S, though they are generally close to each other. This aspect merits consideration, particularly when a heavy unemployment dominates the poverty picture. In any case, β , the income gap ratio is not very sensitive to the income inequality , so that π_{BD} is not that sensitive to the income inequality , while G_α is not that sensitive to the income gap ratio, and this makes π_T^* rather insensitive to the income gap ratio. Looking from this perspective, both π_S and π^* take into account both β and G_α , and between the two, π^* is less fluctuative than π_S . On this ground, we recommend π^* as a proper poverty index.

University of North Carolina, Chapel Hill

REFERENCES

[1] Blackorby, C. and Donaldson, D. (1980). Ethical indices for the measurement of poverty. *Econometrica 48*, 4, 1053-1060.

[2] Klefsjö, B. (1983). Some tests against aging based on the total time on test transform. *Commun. Statist.- Theor. Meth. 12, 8*, 907-927.

[3] Sen, A.K. (1973). *On Economic Inequality*. Oxford University Press, London

[4] Sen, A.K. (1976). The measurement of poverty: An axiomatic approach. *Econometrica 44*, 1, 219-232.

[5] Takayama, N. (1979). Poverty, income inequality and their measures: Professor Sen's axiomatic approach reconsidered. *Econometrica, 47*, 3, 747-760.