SBrT 2019 1570553038

Contribuição ao estudo da medida de desigualdades – análise conjunta dos índices de Gini e Theil

Paulo Roberto Rosa Lopes Nunes

***Resumo*—Neste artigo os índices de Gini e Theil, utilizados para medida de desigualdade em uma população, são conjuntamente analisados. O índice de Theil é uma medida baseada na Teoria da Informação, podendo variar de 0 a ∞ se não houver limite ao tamanho da população. Já o índice de Gini é limitado ao intervalo $[0,1]$. Mostra-se de forma simples que o índice de Theil é uma entropia diferencial negativada, e, à luz da Teoria da Informação, um novo índice, S, aqui chamado índice de população escravizada equivalente, que varia no intervalo $[0,1]$, função do índice de Theil, é definido. Funções ideais de distribuição de riqueza e a conhecida distribuição de Pareto são analisadas no plano $S-G$. Os resultados sugerem a superioridade do índice S.**

Palavras-Chave — Desigualdade, Gini, Theil, Lorenz

***Abstract*— In this paper, the Gini, G, and the Theil, S, indices, which are used to measure inequality in a population, are jointly analyzed. The Theil index is a measure based on Information Theory, and can vary from 0 to ∞ if there is no limit on the size of the population. The Gini index is limited to the interval [0,1]. In a simple way, it is shown that the Theil index is the negative of a differential entropy, and, in the light of Information Theory, it is defined a new index, here called the equivalent slave population index, S, that is a function of the Theil index and varies in the interval $[0,1]$. Ideal wealth distribution functions and the well-known Pareto distribution are analyzed in the S-G plane. Results suggest superiority of the S index.**

Keywords — Inequality, Gini, Theil, Lorenz

I. INTRODUÇÃO

Neste trabalho será abordada a medida de desigualdade de riqueza em uma população, com ênfase no índice de Theil, que é uma medida oriunda da Teoria da Informação de Shannon. Os índices de Gini e de Theil são obtidos a partir de uma curva constante em intervalos, positiva e com área unitária, aqui denominada curva densidade de riqueza, função de uma variável contínua no intervalo $[0,1]$. Esta curva, em forma de escada, é essencialmente independente do tamanho da população, e para populações grandes, pode ser aproximada por uma curva contínua. A bem conhecida Curva de Lorenz é a integral desta densidade. Mostra-se que o índice de Gini é uma função linear simples do valor esperado da variável independente da função densidade de riqueza (generalizando-se a definição de valor esperado de variáveis aleatórias), e que o índice de Theil nada mais é do que o negativo da entropia diferencial da mesma.

À luz da Teoria da Informação de Shannon, para um dado valor do índice de Theil (T) são definidos dois índices, aqui denominados índice de população igualitária equivalente e índice de população escravizada equivalente ($S = 1 - I$). Estes dois índices variam no intervalo $[0,1]$. Para uma distribuição particular, aqui denominada "senhores e escravos", mostra-se que o índice de Gini (G) e o índice S são iguais.

No plano $S\ x\ G$, determina-se a região onde valores de S e G podem existir, e são estudadas curvas associadas a algumas distribuições de riqueza ideais e também à conhecida distribuição de Pareto. Observa-se que existem distribuições de riqueza que pelo índice de Gini mostram pouca desigualdade enquanto pelo índice S são bastante desiguais.

Assume-se aqui que riqueza é uma grandeza sempre positiva ou nula, que pode expressar desde posse de bens materiais até cultura ou formação acadêmica, por exemplo, passando por salário anual ou mensal recebido. De uma forma mais ampla, riqueza pode ser uma função positiva de vários destes bens materiais ou imateriais.

II. A FUNÇÃO DENSIDADE DE RIQUEZA

Neste trabalho, será definida inicialmente uma função de variável contínua que será denominada função densidade de riqueza. Na literatura, a abordagem mais usual é definir inicialmente a curva de Lorenz, que é a integral da densidade de riqueza. A escolha pela curva densidade de riqueza é mais conveniente quando o objetivo é o índice de Theil, que foi inicialmente definido por Theil [1] sobre uma distribuição de variável discreta. Com o uso da função densidade de riqueza, que é função de uma variável contínua, associa-se de imediato o indice de Theil com a entropia diferencial da mesma. Já o índice de Gini, é definido sobre a curva de Lorenz como uma relação de áreas. Neste trabalho mostra-se que índice de Gini está associado com o momento de 1ª ordem da densidade de riqueza. Em analogia com variáveis aleatórias, este momento será chamado de valor esperado da variável $r \in [0,1]$ para a qual é definida a função densidade de riqueza. Como se verá, o uso deste "valor esperado" como uma medida de desigualdade é bastante intuitivo. Tendo os dois índices de interesse referidos a uma única função, fica mais fácil compará-los e trabalhar com ambos. Embora a função densidade de riqueza seja definida inicialmente para uma população com N indivíduos, sua forma não se altera se N for substituído por qualquer múltiplo de N, o que permite estender a função para grandes populações, quando a forma descontinua, intrínseca à densidade, como será visto, pode ser negligenciada sem maiores consequências.

A constatação de que o índice de Theil está associado à integral da derivada da curva de Lorenz multiplicada pelo seu logaritmo foi feita por Rohde, em [2], a partir da expressão integral do índice em termos da função densidade de população por riqueza. A forma aqui desenvolvida é bem mais direta e simples. Em [2], a mencionada integral não é associada à entropia diferencial h.

A função densidade de riqueza é uma função não negativa cujo domínio é o intervalo $[0,1]$.

Assuma-se que a riqueza total de uma população de N indivíduos, $N \geq 1$, é igual a 1 e numere-se os indivíduos em ordem crescente da riqueza por cada um possuída, sendo 1 o individuo mais pobre e N o mais rico. Ao indivíduo n, cuja fração de riqueza é $W(n)$, associe-se o subintervalo $\left(\frac{n-1}{N}, \frac{n}{N}\right]$ do domínio da função densidade de riqueza. Neste intervalo, a função densidade de riqueza $w(r)$ assume o valor $NW(n)$, ou seja:

Paulo Roberto Rosa Lopes Nunes é Professor Titular aposentado pelo Instituto Militar de Engenharia. prnunes@alumni.stanford.edu

$$w(r) = NW(n) = NW(\lceil rN \rceil) \qquad 0 < r \leq 1\,, \qquad (1)$$

onde $n = \lceil rN \rceil$ é o valor do primeiro inteiro maior ou igual a rN. No ponto 0 a função densidade de riqueza é nula por definição. A partir desta função, a curva de Lorenz é obtida por integração. O coeficiente de Theil, como será mostrado, nada mais é que a entropia diferencial h associada à densidade de riqueza $w(r)$, que não é uma densidade de probabilidade, mas apresenta todas as características de uma. Como $w(r)$ é uma função não decrescente, sua integral, a curva de Lorenz associada, é uma função convexa. Na figura 1 é ilustrada uma densidade de riqueza *(a)* e a curva de Lorenz associada *(b)*. Estas curvas são validas para qualquer valor de N múltiplo de 8.

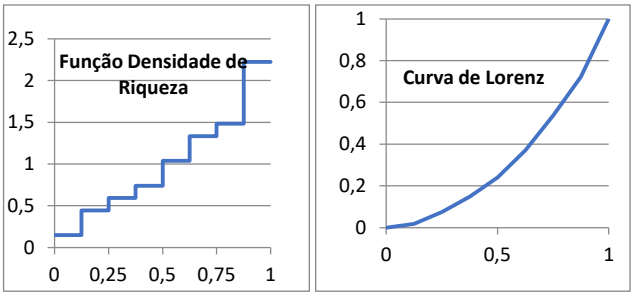

Fig 1. Exemplo de função densidade de riqueza (a) e a Curva de Lorenz Correspondente

III. O ÍNDICE DE GINI

O índice de Gini é certamente a mais popular medida de desigualdade utilizada no mundo. Ele é determinado com base na diferença de área entre a curva de Lorenz para uma perfeita distribuição de riqueza (todos os elementos são igualmente ricos) e a curva de Lorenz da população que está sendo medida.

A curva de Lorenz, $L(r)$, definida no intervalo [0,1], expressa qual fração da riqueza total da população pertence à fração r mais pobre da mesma. Como uma população sempre possui um número finito de indivíduos, a exata expressão da riqueza pertencente a fração r mais pobre da população seria uma função constante em intervalos. Para evitar este inconveniente, pode-se assumir que a população é infinita, ou, como é feito na abordagem aqui apresentada, que os indivíduos de uma população finita e suas riquezas podem ser fracionados. A função densidade de riqueza definida neste trabalho é construída assumindo-se a possibilidade deste fracionamento, sendo fácil constatar que a curva de Lorenz é a sua integral, sempre uma função contínua crescente e convexa.

Somente para uma população infinita pode-se admitir uma descontinuidade na curva de Lorenz, no ponto 1. Ocorreria se um número finito de indivíduos detivesse uma fração não infinitesimal da riqueza dessa população infinita. Na curva densidade de riqueza isto corresponderia a uma função impulso no ponto 1.

No caso de uma distribuição perfeitamente igualitária de riqueza, a função densidade de riqueza assume o valor 1 em todo o seu domínio, exceto no ponto zero, onde, por definição, vale zero, e a curva de Lorenz associada é uma reta com inclinação unitária. Qualquer outra distribuição de riqueza possui uma curva de Lorenz abaixo desta reta.

O índice de Gini associado a uma determinada distribuição de riqueza é definido como o valor da área da região entre a curva de Lorenz da distribuição igualitária e a curva de Lorenz da distribuição sendo medida, dividido pela área sob a curva da densidade igualitária, ou seja o índice de Gini vale duas vezes o valor da área A indicada na figura 2. Para o exemplo ilustrado na Figura 1, o coeficiente de Gini é igual a 0,347.

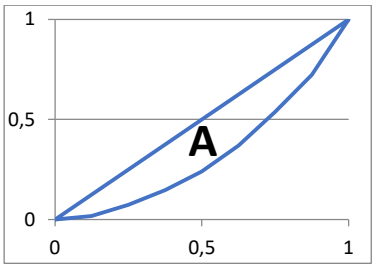

Fig. 2 - A Curva de Lorenz e o índice de Gini, que vale o dobro da área A

E interessante expressar o coeficiente de Gini em termos da curva $w(r)$. Observe-se que a curva de Lorenz é dada por:

$$L(r) = \int_0^r w(x)dx, \qquad (2)$$

e que a área A é dada por:

$$A = \frac{1}{2} - \int_0^1 L(r)dr = \frac{1}{2} - \int_{r=0}^1 \int_{x=0}^r w(x)dxdr \qquad (3)$$

A integral no 2º membro da equação acima é a integral de uma função $f(r,x) = w(x)$ na região do plano r,x dada pelo triangulo de vértices 0,0 1,0 e 1,1, e pode ser realizada variando x de 0 a 1 e r de x a 1, ou seja:

$$\begin{aligned}\int_{r=0}^1 \int_{x=0}^r w(x)dxdr &= \int_{x=0}^1 \int_{r=x}^1 w(x)dxdr \\ &= \int_{x=0}^1 w(x)dx \int_{r=x}^1 dr \\ &= \int_{x=0}^1 (1-x)w(x)dx = 1 - E_W\,, \qquad (4)\end{aligned}$$

onde

$$E_W = \int_{x=0}^1 xw(x)dx \qquad (5)$$

é o valor esperado da variável x considerando-se a densidade $w(x)$. Portanto, a área A é dada por

$$A = \frac{1}{2} - (1 - E_W) = E_W - \frac{1}{2} \qquad (6)$$

E o índice de Gini por

$$G = 2\left(E_W - \frac{1}{2}\right) \qquad (7)$$

Portanto, o coeficiente de Gini é duas vezes a diferença entre a média da função densidade de riqueza e a média da função densidade de riqueza ideal. A expressão acima será útil para determinar o limitante superior para o coeficiente de Gini

quando é dada a entropia diferencial da distribuição, ou seja, quando é dado o índice de Theil.

A expressão do índice de Gini na forma (7) não é comum, mas é bastante intuitiva: Quanto maior a concentração de riqueza mais próximo de 1 é o valor esperado E_W da variável r.

IV. O ÍNDICE DE THEIL

O índice de Theil é definido como a diferença entre a entropia da distribuição de riqueza se ela fosse igualitária (entropia máxima) e a entropia da real distribuição de riqueza entre os elementos da população [1]. A entropia da distribuição de riqueza é dada por:

$$H(W) = -\sum_{n=1}^{N} W(n) log_2 W(n) \ . \tag{8}$$

Se todos os indivíduos possuem frações idênticas de riqueza a entropia da distribuição é igual a $log_2 N$, pois $W(n) = 1/N$. O índice de Theil, T, portanto, é dado por:

$$T = log_2 N - H(W) \tag{9}$$

Calculando-se agora a entropia diferencial (vide e.g. [3]) associada a função densidade de riqueza previamente definida:

$$\begin{aligned} h(w) &= -\int_0^1 w(x) log_2 w(x) dx \\ &= -\int_0^1 NW([xN]) log_2 NW([xN])\, dx \\ &= -N\left\{\sum_{n=1}^{N} W(n)[log_2 W(n) + log_2 N]\frac{1}{N}\right\} \\ &= H(W) - \sum_{n=1}^{N} W(n) log_2 N = H(W) - log_2 N = -T \end{aligned} \tag{10}$$

Ou seja, a entropia diferencial da densidade de riqueza é igual a menos o índice de Theil associado a distribuição em questão. Quando $N \to \infty$ tanto $H(W)$ quanto $log_2 N$ tendem a infinito, mas a diferença entre os dois, ou seja, o índice de Theil, se mantém finita, a não ser que $\lim_{x\to 1} w(x) = \infty$, que é uma condição necessária, mas não suficiente, para que $T = -h(w) \to \infty$. Se N se mantém finito o máximo valor que T pode atingir é $log_2 N$, quando $H(W) = 0$, o que ocorre quando um único indivíduo concentra toda a riqueza. Observe-se que o índice de Theil é invariante à posição dos intervalos associados aos indivíduos no eixo horizontal, ou seja, não é necessário que a densidade de riqueza seja uma função crescente, o que permite que o índice de Theil de uma população seja facilmente obtido a partir de dados de subpopulações, o que geralmente não ocorre com o coeficiente de Gini, que requer o ordenamento dos indivíduos.

A definição do índice de Theil utilizada é a inicialmente proposta em [1]. Também se pode expressar o índice em termos da densidade de população por riqueza, $p(R)$, onde $0 \le R <\propto$ representa riqueza em valores absolutos e $p(R) = \frac{\mathrm{dP}}{\mathrm{dR}}$ é a razão entre a fração de indivíduos com riqueza no intervalo $[R, R + dR)$ e dR. Pode-se mostrar que a curva de Lorenz pode ser obtida de $p(R)$ pela seguinte equação:

$$L(r) = \frac{1}{\bar{R}} \int_0^{R(r)} xp(x) dx \tag{11}$$

Onde $R(r)$ soluciona

$$r = \int_0^{R(r)} p(x) dx \tag{12}$$

e

$$\bar{R} = \int_0^{\infty} xp(x) dx \tag{13}$$

é a riqueza média da população.

Para o caso de N finito, $p(R)$ será sempre composta por funções impulsos. Apenas quando se assume que $N \to \infty$ é que esta densidade pode ser representada por uma função contínua. No entanto, grandes populações são bem modeladas com $N \to \infty$.

O índice de Theil pode ser escrito como [1]:

$$T = \frac{1}{\bar{R}} \int_0^{\infty} xp(x) log_2 \left(\frac{x}{\bar{R}}\right) dx \tag{14}$$

Em geral, (14) é utilizada como definição do índice de Theil.

A. Os índices de população igualitária equivalente (I) e população escravizada equivalente (S)

Na Teoria da Informação de Shannon pode-se mostrar que a entropia de uma distribuição de probabilidades com N pontos é o menor número médio de bits por símbolo (ou de dígitos base k por símbolo se for utilizado logaritmo base k) necessário para codificar uma longa sequência de símbolos estatisticamente independentes gerada de acordo com a distribuição em questão (veja e.g. [3]). Parece claro então que a entropia é uma grandeza com característica descritiva, relacionada a uma distribuição.

No caso de distribuição de riqueza, é razoável supor que quanto mais rico for um elemento da população mais influente e atuante ele será no sistema econômico, e se alguém desejar reportar a atuação dos elementos da população no sistema ao longo do tempo, utilizando palavras-códigos formadas por bits, o menor número médio de bits para identificar cada elemento da população é igual à entropia da distribuição de riqueza.

A entropia, como já visto, pode assumir valores no intervalo $[0, \infty)$ se N não for limitado. Para comparar o índice de Theil com o coeficiente de Gini é interessante mapear esta grandeza no intervalo $[0,1)$. Para isso, é introduzido neste trabalho o conceito de População Igualitária Equivalente, à qual está associado o índice I. A População Igualitária Equivalente é aqui definida como uma população com a mesma entropia da população em estudo constituída por elementos igualmente ricos. O número de elementos nesta população é claramente $2^{H(W)}$, e o índice I é definido como a a fração da população correspondente a este número:

$$I = \frac{2^{H(W)}}{N} = 2^{H(W) - \log_2 N} = 2^{-T} = 2^{h(w)} \tag{15}$$

O índice I varia entre 0 e 1, e quanto maior, maior é a igualdade na população. Como se deseja uma medida de desigualdade pode-se definir um índice de População Escravizada Equivalente S. O nome decorre do fato desta população equivalente ser constituída por $N - 2^{H(W)}$

elementos completamente pobres, de modo que a combinação destas duas populações equivalentes possui o mesmo número de elementos e a mesma entropia da população original. O índice S é dado por:

$$S = 1 - I = 1 - 2^{h(w)} = 1 - 2^{-T} \quad (16)$$

Para a população utilizada para gerar os gráficos das figuras 1 e 2, são os seguintes os valores encontrados para T, I e S:

$$T = 0{,}289$$
$$I = 0{,}8185$$
$$S = 0{,}1815$$

A população extrema formada por elementos igualmente ricos e por elementos completamente pobres será chamada neste trabalho de uma população "senhores+escravos". A função densidade de riqueza para uma população "senhores+escravos" é nula para r no intervalo $[0, S]$ e igual a $1/(1-S)$ para r no intervalo $(S, 1]$, sendo fácil constatar que para estas populações $G = S$.

Vale observar que os valores de I e S não dependem da base utilizada para o cálculo da entropia, desde que a mesma base seja utilizada no cálculo dos índices. Para efeitos de simplificação matemática, pode-se então trabalhar com a base natural.

V. O PLANO S – G

De posse de dois índices para medir desigualdades que variam no mesmo intervalo, é interessante comparar o comportamento de ambos para algumas funções densidade de riqueza. A primeira questão que surge é quanto às limitações que um índice impõe ao outro.

A. Valores extremos de G dado S (ou, equivalentemente, dado T)

Dado que a densidade de riqueza possui um índice de Theil T, deseja-se determinar o intervalo possível para o valor de G. Trata-se de um problema de otimização funcional no intervalo $[0,1]$ com restrições, que pode ser resolvido com o uso de coeficientes de Laplace:

Como já foi visto, o coeficiente de Gini é proporcional à média E_W da densidade $w(r)$ (equações (5) e (7)). Basta, portanto, maximizar ou minimizar E_W com a restrição do valor de T, para encontrar os limites de G. Para simplificar as operações matemáticas será utilizado o logaritmo natural.

Seja o Laplaciano

$$\Lambda\big(w(r)\big) = \int rw(r)dr + \lambda_1 \int w(r)\ln\big(w(r)\big)dr + \lambda_2 \int w(r)dr \quad (17)$$

Derivando em relação a $w(R)$ e igualando a zero obtém-se:

$$\frac{\partial \Lambda\big(w(r)\big)}{\partial w(R)} = R + \lambda_1 \ln\big(w(R)\big) + \lambda_1 w(R)\left(\frac{1}{w(R)}\right) + \lambda_2 = 0 \quad (18)$$

Ou

$$\ln\big(w(R)\big) = \frac{-R + \lambda_1 + \lambda_2}{\lambda_1} \quad (19)$$

O que implica em

$$w(r) = Ke^{\alpha r} \qquad 0 \le r \le 1, \quad (20)$$

sendo as constantes K e α determinadas em função das restrições impostas por T_e e pela área unitária, além da condição de $w(r)$ ser uma função não decrescente.

Conclui-se então de $\alpha > 0$, $K = \alpha/(e^{\alpha} - 1)$ e α tal que

$$\int_0^1 Ke^{\alpha r} \ln(Ke^{\alpha r})\, dr = T_e. \quad (21)$$

Como a exponencial é crescente, o valor de E_w é um máximo. Um mínimo seria obtido com uma exponencial decrescente, mas que não atenderia a condição de $w(r)$ ser uma função não decrescente.

Integrando (21) conclui-se que o valor de α é determinado a partir de

$$T_e = -\left((e^{\alpha} - 1)\ln\left(\frac{\alpha}{e^{\alpha} - 1}\right) + (\alpha - 1)e^{\alpha} + 1\right)/(e^{\alpha} - 1) \quad (22)$$

E que o valor máximo que G pode atingir é dado por:

$$G_{max} = 2\frac{(\alpha - 1)e^{\alpha} + 1}{\alpha(e^{\alpha} - 1)} - 1 \quad (23)$$

Quanto ao valor mínimo que G pode atingir não há restrição, pois distribuições uniformes com um delta de Dirac no valor 1 possuem T infinito e G muito pequeno, a depender da área da função impulso. Trocando-se o impulso por um estreito pulso, e moldando-se a área de forma a atender o valor finito de T_e, pode-se obter um valor de G muito próximo de zero. Funções densidade de riqueza com estas características correspondem a populações essencialmente igualitárias mas com uma concentração significativa de riqueza nas mãos de um pequeno grupo. Estas populações quando medidas pelo índice de Gini mostram pouca desigualdade, mas quando medidas pelo índice S podem apresentar valores elevados. Neste trabalho estas populações serão denominadas "elite+povo".

Os valores máximos que o índice de Gini pode atingir, dado o índice S, correspondem à curva indicada na Fig. 3.

B. A distribuição ideal "senhores+escravos"

Como já visto na seção *IV-A*, a distribuição ideal *senhores+escravos* é constituída por uma parcela S da população que não possui bens e por uma parcela igualitária $I = 1 - S$ que possui todos os bens da população. Os valores *I* e *S* são, respectivamente, os índices de população igualitária equivalente e de população escravizada equivalente da população. Para esta distribuição, o índice de Theil é dado por $T = -\log_2(1 - S)$, e o índice de Gini é igual ao índice *S*.

No plano $S - G$ distribuições como estas mapeiam na diagonal que une os pontos *0,0* e *1,1* (vide Fig. 3).

C. A distribuição ideal "elite+povo"

A distribuição ideal elite+povo é constituída por uma fração igualitária t da população que possui uma fração $f > t$ da riqueza total. A fração restante da riqueza é possuída igualitariamente pela fração restante da população.

A função densidade de riqueza desta população ideal vale $(1 - f)/(1 - t)$ no intervalo $(0, t]$ e f/t no intervalo $(t, 1]$.

É fácil verificar que o índice de Gini é dado por $G = f - t$ e que o índice de Theil é dado por:

$$T = (1-f)\log_2\left(\frac{1-f}{1-t}\right) + f\log_2\left(\frac{f}{t}\right) \qquad (24)$$

Um pouco de manipulação matemática mostra que:

$$S = 1 - \left(\frac{1-t}{1-f}\right)\left(\frac{(1-t)f}{(1-f)t}\right)^{-f} \qquad (25)$$

Na Fig. 3 são ilustradas duas curvas para esta distribuição ideal, para os valores $t = 10^{-4}$ (ELITE 1E-4) e $t = 10^{-8}$ (ELITE 1E-8). Alguns valores do parâmetro f estão indicados ao longo das curvas.
Observa-se que para distribuições como estas o índice de Gini indica pouca desigualdade, ao contrário do índice S.

D. A distribuição de Pareto

Suponha-se que em uma população nenhum individuo possua menos (ou ganhe menos) que b, ou seja, o número relativo de indivíduos que possuem mais que b é 1. Suponha-se também que o número relativo de indivíduos que possuem mais que R, $R > b$, decresça seguindo uma lei de potência, ou seja, é igual a $\left(b/R\right)^a$, onde $a > 1$. Ou seja, se um indivíduo possui R, uma fração $\left(b/R\right)^a$ da população possui tanto ou mais que ele. A velocidade com que esta fração decresce depende do valor de a, mas não existe um indivíduo mais rico que todos os outros. Por mais que um indivíduo possua, sempre haverá alguém mais rico. Estas características se adaptam bem a várias populações reais existentes no mundo. A distribuição da população por riqueza de uma população com estas características é dada pela densidade de Pareto:

$$p_b(R) = \begin{cases} 0 & R < b \\ \frac{a}{b}\left(\frac{b}{R}\right)^{a+1} & R \geq b \end{cases} \qquad (26)$$

Para a densidade de Pareto pode-se mostrar que a curva de Lorenz, que independe do parâmetro b, é dada por:

$$L(x) = 1 - (1-x)(1-x)^{1-\frac{1}{a}}\ . \qquad (27)$$

Derivando a curva de Lorenz, encontra-se a curva densidade de riqueza:

$$w(x) = L'(x) = \frac{a-1}{a}(1-x)^{-\frac{1}{a}} \qquad (28)$$

Utilizando a curva de Lorenz ou a densidade de riqueza determina-se o índice de Gini:

$$G = \frac{1}{2a-1}\ , \qquad (29)$$

E o índice de Theil na base Neperiana:

$$T_e = log_e \frac{\mathrm{a}-1}{\mathrm{a}} + \frac{1}{a-1} \qquad (30)$$

Que permite calcular o valor de S:

$$S = 1 - \frac{a}{a-1}\mathrm{e}^{-\frac{1}{a-1}} \qquad (31)$$

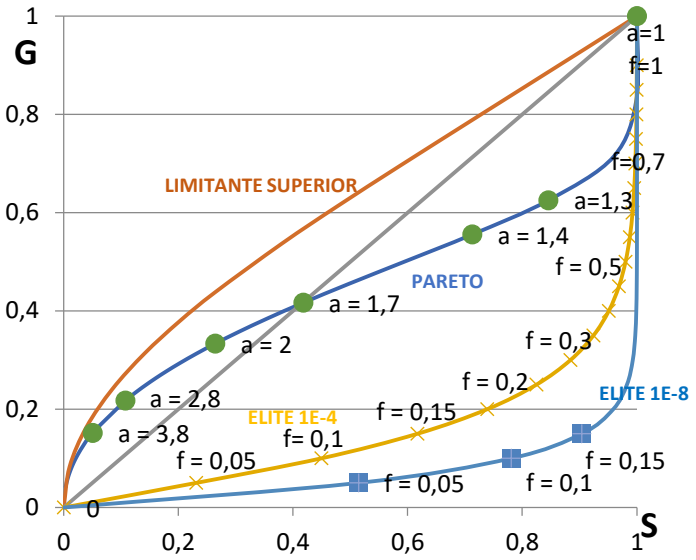

Fig. 3 - O plano S - G

VI. CONCLUSÕES

Neste trabalho os índices de Theil e Gini foram estudados de uma forma unificada. Mostrou-se que o índice de Theil está relacionado ao número médio de bits (ou dígitos de outra base) necessário para identificar, ao longo do tempo, qual elemento da população realiza atividade econômica. Mais exatamente, o índice de Theil é o decréscimo do número médio de bits necessário, caso todos os elementos da população fossem igualmente participantes.

Definiu-se então um índice S, função de T, que é a fração da população sem atividade econômica se apenas a parte restante da população atuasse de forma igualitária, mantendo o valor de T. T pode variar desde 0, para uma população igualitária, até $\log_2 N$ bits, para uma população onde toda a riqueza está concentrada em um único elemento. Já S varia entre 0 e $(N-1)/N$, em termos práticos entre 0 e 1.

Curvas relacionadas a distribuições ideais e a conhecida distribuição de Pareto foram analisada no plano S – G. Mostrou-se que, dado o valor de S, o valor do índice de Gini pode variar significativamente, desde 0 até um valor máximo função de S. Em especial, para populações com grande concentração de riqueza por uma fração muito pequena da população, verifica-se que o índice de Gini falha como indicador de desigualdade, o que não acontece com o índice S, indicando uma superioridade do índice S sobre o índice de Gini.

REFERÊNCIAS

[1] Theil, H., Economics and Information Theory, Amsterdam, North Holland, 1967

[2] Rohde, N., Derivation of Theil's Inequality Measure from Lorenz Curves Technical Report. University of Queensland, Brisbane, 2007

[3] Cover, T.M., Thomas, J.A. - Elements of Information Theory (2nd ed.), Wiley-Intercience, 2006