III.

Algumas Medidas de Concentração e Desigualdade e suas Aplicações

Ramonaval A. Costa(*)

1 INTRODUÇÃO

As medidas mais usadas no estudo da densidade da renda têm sido pouco utilizadas no caso de outros atributos que não sejam renda, apesar dos seus significados comportarem o emprego de uma gama de atributos. As medidas de desigualdade podem ser aproveitadas para fornecer informações sobre o grau de concentração na distribuição das terras, na produção industrial, na área plantada, ou em qualquer atributo que possamos imaginar, pelo menos teoricamente, como igualmente distribuído pelas pessoas ou espacialmente.

O fato destas medidas terem um significado estatístico bem definido abre a possibilidade de usá-las com outros atributos. O uso costumeiro delas relacionado com a renda, às vezes, é explicado por problemas de origem da própria medida, como é o caso da razão de concentração de Gini e da curva de Lorenz; outras vezes é apenas um problema de divulgação e concretização da utilidade das medidas e de seu significado.

Para o estudo de padrões de concentração é perfeitamente possível utilizarmos como instrumentos coadjuvantes aqueles que até agora têm sido largamente usados nos estudos da concentração e desigualdade da renda. Para evitar o uso indiscriminado das medidas de desigualdade como indicadores de padrões de

---

(*) O autor é professor da Universidade de Brasília e da FGV-IESAE. É chefe da Divisão de Pesquisas do Banco Central.

concentração é preciso que se entenda, pelo menos, a idéia, o sentido que está por trás destas medidas e assim como os problemas que surgem para uma aplicação sem ambigüidade nem erros de interpretação.

Com este objetivo discutimos algumas medidas muito usadas no estudo da concentração e desigualdade da renda, ora em função da simplicidade do seu cálculo, ora em função da clareza do seu significado. Entre elas destacamos a razão de concentração de Gini, a curva de Lorenz e suas propriedades, o Índice de Theil e algumas aplicações destes indicadores para o Brasil e algumas regiões.

2. MEDIDAS DE DESIGUALDADE E CONCENTRAÇÃO

2.1. Razão de Concentração de Gini

A razão de concentração de Gini é, sem dúvida, a medida de desigualdade mais conhecida usada e criticada nos estudos da distribuição da renda pessoal, tendo sido calculada para a maioria dos países existentes(1). Apesar de sua divulgação ter sido apresentada praticamente como uma medida vinculada à curva de Lorenz, a razão de concentração de Gini tem raízes bem estabelecidas do ponto de vista estatístico.

Do ponto de vista estatístico a razão de concentração de Gini é definida como uma medida de dispersão relativa; ou seja, a razão entre a média das diferenças e o dobro da média aritmética;

$$G = \frac{\Delta}{2\mu}$$

Consderando-se Δ como sendo a média das diferenças, isto é:

$$\Delta = \frac{\sum_{i=1} \sum_{j=1} | X_i - X_j | f_i f_j}{N^2}$$

(1) Vide Felix Paukert — «Income Distribution at Differente Levels of Development: a Survey of Evidence» — International Labour Review, 108, agosto/setembro/77, pp. 97/123.

Não há dúvida, portanto, que a razão de concentração é uma medida cujas bases estatísticas são relativamente sólidas, podendo ser comparadas às do coeficiente de variação. Trata-se, portanto, de um índice adimensional cujo limite inferior é zero e o limite superior é igual a um.

Recentemente, o reconhecimento destas características da razão de concentração levou **Joseph L. Gastwirth** a superar algumas de suas deficiências mais comumente apontadas. Primeiro foi a eliminação da crítica de subestimação da razão de concentração, apresentando um limite superior para tal medida[(2)]. Em seguida, permitindo a decomposição da desigualdade entre classes e intra-classes.

Em verdade, parece fora de dúvida de que a interpretação da razão de concentração como uma medida de dispersão não foi muito divulgada na literatura econômica anglo-saxônica. Mas, se observarmos as contribuições dos estatísticos italianos nas primeiras três décadas do século XX, principalmente em publicações na Revista METRON, internacionalmente conhecida, observaremos que o cálculo da média das diferenças (ou diferença média) foi motivo de muita discussão. Tais discussões giravam em torno do objetivo de simplificar sua obtenção, já que se tratava de uma medida de dispersão que ao invés de aferir as diferenças em relação a uma medida de tendência central, como a variância, pretendia aferir todas as diferenças possíveis entre os atributos, o que dificultava em demasia o seu cálculo[(3)] O significado estatístico da razão de concentração de Gini é, portanto, muito claro mesmo quando a medida está desvinculada da curva de Lorenz. Isto porque é possível demonstrar que o limite superior de $\triangle$ tende para 2μ, dando-nos a relação:

$$\frac{\triangle}{2\mu} = G$$

2.2. Curva de Lorenz

A razão de concentração de Gini tem sido muito divulgada como a razão entre a área de concentração real e área de concen-

(2) Vide Joseph L. Gastwirth — «The Estimation of the Lorenz Curve and Gini Index» — The Review of Economics and Statistics, LIV, agosto de 1972.

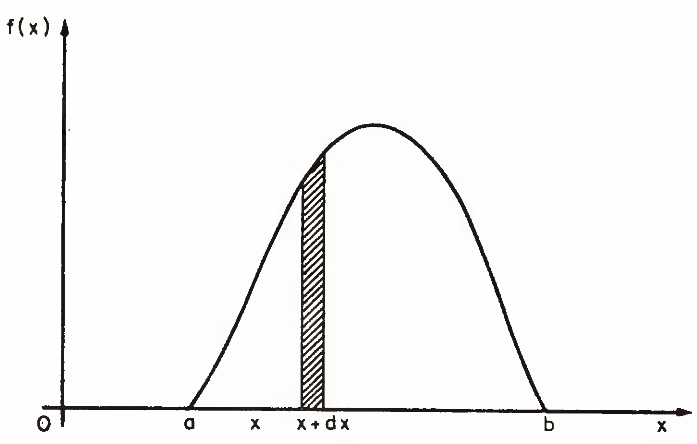
(3) Vide U. Paciello e B. De Finetti — «Calcolo della differenza media», Metron, Vol. VIII, n.º 3, fevereiro 1931.

tração máxima, em termos da curva de Lorenz. Tal fato obscurece um pouco o signficado estatístico da razão de concentração, apesar de facilitar enormemente a compreensão e, por conseguinte, a sua divulgação e utilização como medida de desigualdade.

Historicamente, o advento da curva de Lorenz como um instrumento adequado para observações de variações no grau de concentração de um atributo (renda), foi introduzido através dos trabalhos de **Lorenz, Chatelain, Seailles** e **Corrado Gini**[4]. Não há dúvida de que todos estes autores foram contemporâneos; no entanto, **M. O. Lorenz** é o mais conhecido, tendo a curva de concentração em questão recebido o seu nome.

Antes de apresentarmos de uma forma bem simples qual o significado da curva de Lorenz, é interessante que se dê um tratamento mais técnico ao seu verdadeiro significado. Para isto assumimos que temos um atributo quantitativo e variável entre **a** e **b**, tal que $a < b$. Em seguida, supomos que $f(x)$ seja a função densidade de freqüência e que $f(x)\,dx$ seja a freqüência com que o atributo assume valores no intervalo $x + dx$. Fazendo-se o gráfico de $f(x)$ temos:

FIGURA - 1



(4) Vide M.O. Lorenz «Methods of Measuring the Concentration of Wealth», ASA, Nova Série, junho 1905, 209; E. Chatelain «Les sucessions declarées in 1905», **Rev. Politique et Parlamentaire**, 1907; J. Seailles, «La repartition des fortunes en France», Alcan, Paris 1910; Corrado Gini, Indice di Concentrazione e di Dipendença, Roma, **Biblioteca Dell'Economista**, Editice Dorinese Milano, 1922, pp. 5/137.

Considerando **x** o atributo e **p** e **r** como coordenadas de um ponto qualquer **p** da curva de Lorenz temos, portanto, as respectivas equações:

$$p = \frac{\int_a^x f(x)\,dx}{\int_a^b f(x)\,dx} \quad (1) \quad e \quad r = \frac{\int_a^x x\,f(x)\,dx}{\int_a^b x\,f(x)\,dx} \quad (2)$$

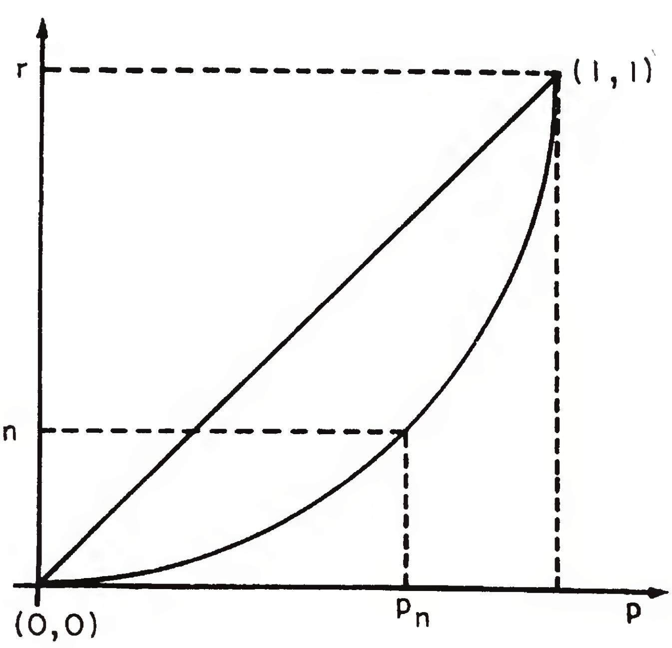
observamos que a equação (1) não é nada mais que a razão entre o momento incompleto e completo de ordem 0 da função distribuição F(x). Ao passo que a equação (2) representa a razão entre o momento incompleto e completo de ordem 1 da mesma função de distribuição. Logo podemos escrever **p** e **r** da seguinte forma:

$$p = \frac{m_0\,(x)}{m_0} = \frac{\int_a^x f(x)\,dx}{N} \quad e \quad r = \frac{m_1\,(x)}{m_1} = \frac{\int_a^x x\,f(x)\,dx}{N\overline{X}}$$

Onde N e $\overline{X}$ são respectivamente a freqüência total e a média aritmética do atributo x; evidentemente $N\overline{X}$ representa o total do atributo da distribuição em questão.

A curva de Lorenz, que é representada num diagrama cartesiano tendo como abcissa **p** e como ordenada **r**, será uma curva convexa em relação a **p**, cujos valores correspondentes aos limites **a** e **b** são respectivamente (0,0) e (1,1). Fazendo-se o gráfico correspondente obtemos:

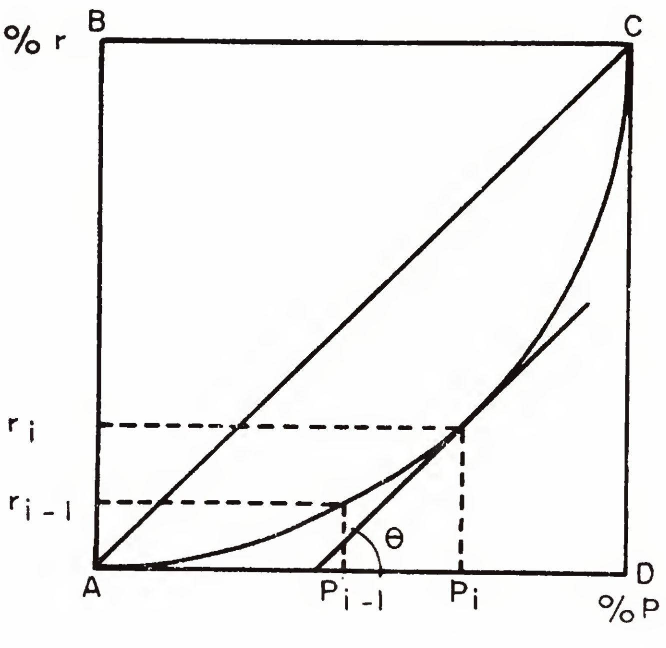
FIGURA - 2



Pela simples observação da Figura 2 podemos concluir que $p_i > r_i$ para todos os pontos, com exceção dos extremos, onde $p_i = r_i$[5]. Pelo exposto, não resta dúvida de que a curva de Lorenz existe para qualquer tipo de distribuição de um dado atributo. Não há necessidade de que x seja renda, apesar de **Lorenz** ter apresentado esta curva como a forma mais adequada para observações sobre as mudanças do grau de concentração da distribuição da renda[6].

Com o intuito de simplificar o entendimento da curva de Lorenz, através da eliminação das formalidades matemáticas e estatísticas, como foi feito no trabalho pioneiro de Lorenz, tentaremos expor rapidamente o significado da curva de Lorenz em termos mais simples. Para isto consideremos que temos numa distribuição de freqüência do atributo **x** com **n** classes, cujas freqüências são f_1, f_2 f_n e os respectivos elementos representativos de cada classe são x_1, x_2 x_n. Numa curva de Lorenz relacionamos os percentuais acumulados de rendas com percentuais acumulados das pessoas que recebem tais rendas a partir daquelas pessoas com níveis de renda mais baixos. Em geral, utiliza-se um quadrado cujo lado é igual à unidade, como na Figura 3 abaixo:

FIGURA – 3



(5) Vide L. Galvani — Sulle Curve di Concentrazione Relative a caratteri non Limitati e Limitati, Metron, Vol. X n.º 3, Roma, outubro de 1932.

(6) Vide M.O. Lorenz, op. cit.

Na Figura 3, AD representa o eixo das abcissas, no qual registramos $p_i = \frac{\sum\limits_{i<n} f_i}{N}$, o percentual acumulado das pessoas que recebem renda, consideradas em ordem crescente de renda. O lado **AB** do quadro considera-se como o eixo das ordenadas, onde registramos $r_i = \frac{\sum\limits_{i=1}^{n} f_i \, x_i}{\sum\limits_{i=1}^{n} f_i \, x_i}$, o correspondente percentual acumualado das rendas recebidas por P_i.

Os pares ordenados (r_i, p_i) representam o que chamamos de curva de Lorenz. No seu trabalho original, Lorenz coloca r_i nas abcissas e p_i nas ordenadas, mas isto não altera em nada as propriedades da curva. A única conseqüência é que a curva de Lorenz ficará acima da diagonal **AC.**

2.3. Algumas Propriedades da Curva de Lorenz

Esta curva tão utilizada quanto criticada desde o advento de sua divulgação, apresenta muitas propriedades que a tornam quase indispensável para quem estuda as características de um dado atributo, especialmente a renda. Vejamos algumas das principais características analíticas e práticas da curva de Lorenz:

1) É uma curva não decrescente e convexa em relação aos percentuais acumulados de pessoas que recebem renda;

2) Com exceção dos dois primeiros e últimos pares ordenados, r_i, será sempre menor que p_i, no caso do atributo x apresentar uma dispersão diferente de zero;

3) A inclinação da curva de Lorenz em qualquer ponto é igual ao limite da razão $\frac{r_i - r_{i-1}}{P_i - P_{i-1}}$ (7).

(7) Vide Edwards Ames, «A Method for Estimating the Size Distribution of a given Agregate Income», The Review of Economics and Statistics, julho de 1942.

4) Observando-se que a curva de Lorenz é convexa, percebe-se que existe um ponto em que a sua inclinação será paralela à diagonal **AC**. Justamente neste ponto a inclinação é igual à unidade, ou seja, a razão do percentual de renda e/o percentual de pessoas na classe são iguais. Esta propriedade nos permite identificar aproximadamente a classe, ou o nível de renda correspondente à renda média ($\bar{x}$). Daí a importância da renda média como uma medida resumo, já que representa o ponto da curva de Lorenz no qual a proporção de pessoas (ou famílias) e a proporção de renda são iguais.

5) A curva de Lorenz possibilita a visualização das modificações na distribuição da renda através do tempo ou as comparações no mesmo período, entre diferentes estados.

6) Quando a curva de Lorenz se reduz à diagonal **AC**, temos que $p_i = r_i$ para todo i. Estamos diante de um caso extremo de perfeita igualdade. Quando a curva de Lorenz se reduz ao triângulo **ACD**, onde $p_i = r_i = 0$, $p_j = r_j = 1$ para $j \neq i$, então estamos em outro extremo, o de máxima desigualdade, com uma pessoa (ou família) com toda a renda. Estes dois extremos são apenas pontos de referência, pois na realidade o que existe mesmo é uma curva convexa como **AEC**. Quanto mais próxima **AEC** estiver de **AC** menor a concentração, quanto mais afastada de **AC**, maior a concentração.

7) Apesar de suas propriedades e informações sobre o que se passa com a distribuição da renda, a curva de Lorenz apresenta situações de ambigüidade quando existe o cruzamento de duas curvas. A distinção será feita pela área entre a diagonal **AC** e a curva de Lorenz **AEC**, chamada área de concentração. Quando a área de concentração é igual, aí temos a situação de duas distribuições cuja dispersão relativa é a mesma[8].

A curva de Lorenz é um instrumento cujas características e propriedades ainda não foram completamente discutidas num só trabalho. Várias informações, sobre esta curva, estão espalhadas em diversos trabalhos relacionados com a distribuição de renda pessoal. Mas o que apresentamos, aqui, é o suficiente para compreendermos melhor o significado da razão de concentração de Gini quando derivado diretamente da curva de Lorenz.

---

de» — Revista Brasileira de Estatística — Rio de Janeiro, julho de 1974.

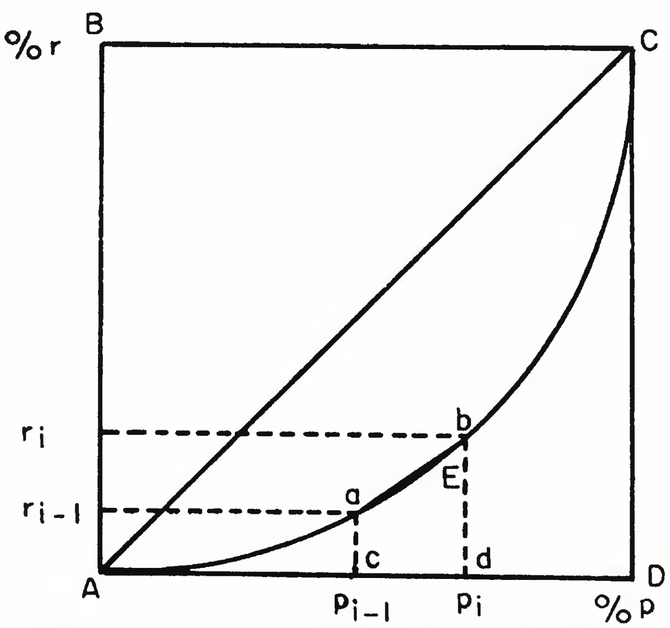
(8) Vide Ramonaval A. Costa, «Bem-Estar de Indicadores de Desigualda-

2.4. A Razão de Concentração de Gini e a Curva de Lorenz

Com as informações sobre a razão de concentração de Gini e sobre a curva de Lorenz podemos mostrar rapidamente como se deriva uma medida para a razão de concentração de Gini, no caso de uma distribuição de rendas. Esta derivação é uma forma bem simples de obtenção da razão de concentração a partir da curva de Lorenz. Existem várias maneiras de se obter esta medida. A que apresentamos utiliza as propriedades dos trapézios, cuja fórmula final se deve a **James Morgan**, um estudioso da distribuição de renda[9].

Consideremos como antes um quadrado **ABCD** com lados iguais à unidade, conforme a Figura 4:

FIGURA - 4



Onde **AC** representa a situação de perfeita igualdade. A área $ACD = \frac{1}{2}$ representa uma situação de máxima desigualdade. Ao passo que a área **ACE**, entre a diagonal **AC** e a curva de Lorenz **AEC**, representa a área de concentração ral, a qual pretendemos medir para construir um índice que representa a razão entre esta área de concentração real **ACE** e área máxima de concentração **ACD**, ou seja:

(9) Vide James Morgan «The Anatomy of Income Distribution». **Review of Economics and Statistics**, XLIV, agosto, 1962, p. 271.

$$G = \frac{\text{área ACE}}{\text{área ACD}} = \frac{\text{concentração real}}{\text{concentração máxima}}$$

Não há dúvida que calculando a área AECD e subtraindo da área **ACD** obteremos a área **ACE**, isto é:

$$G = \frac{\text{área ACD - área AECD}}{\text{área ACD}} = \frac{1/2 \text{ - área AECD}}{1/2} =$$

$$= \text{1-2 área AECD}$$

Portanto, a fórmula prática para a razão de concentração é obtida após o cálculo da área **AECD**. Mas esta área pode ser obtida dividindo-a em pequenos trapézios do tipo **abcd**. A soma total das áreas de todos estes pequenos trapézios resultará na área **AECD**, um pouco superestimada porque o segmento **ab** em verdade não é uma reta, mas sim uma curva. A conseqüência disto é a obtenção de um valor para **G** menor do que ele realmente é. Quando o número de classes é muito grande, esta diferença não chega a se constituir num grande obstáculo para aqueles que buscam apenas uma medida aproximada, sem maiores ambições.

Aplicando a definição da área de um trapézio temos que:

$$\text{área abcd} = \frac{bd + ac}{2} \quad cd$$

Mas $bd = r_i$, $ac = r_{i-1}$ e $cd = p_i - p_{i-1}$, ou seja:

$$\text{área abcd} = \frac{r_i + r_{i-1}}{2} \quad (p_i - p_{i-1})$$

Como temos k trapézios deste tipo, a área

$$\text{AECD} — \sum_{i=1}^{k} \frac{(r_i + r_{i-1})}{2} \quad (p_i — p_{i-1})$$

em termos das coordenadas da curva de Lorenz. Por conseguinte

$$G = 1 - 2\frac{1}{2} \sum_{i=1}^{k} (r_i + r_{i-1})\ (p_i - p_{i-1})$$

$$G = 1 - \sum_{i=1}^{k} (r_i + r_{i-1})\ (p_i - p_{i-1})$$

Esta é a fórmula que se utiliza na prática quando se quer obter o valor da razão de concentração de Gini. Os limites da razão de concentração de Gini ficam bem evidentes: **G=0** quando a área de concentração **ACE** for igual a zero, ou seja, a igualdade perfeita; **G = 1** quando a área de concentração **ACE** for igual a **ACD**, ou a máxima desigualdade.

Existem outras fórmulas práticas para a avaliação numérica de **G**, todas elas fornecem valores aproximados para **G**, mais ou menos equivalentes. A nossa preocupação nesta descrição foi mostrar como a razão de concentração de Gini tem raízes estatísticas bem fundamentadas, não sendo simplesmente um número que se obtém dividindo a área de concentração real pela área de concentração máxima numa curva de Lorenz. É uma medida de dispersão relativa, como o coeficiente de variação e outras medidas de desigualdade, não se podendo confundir o método aproximado de cálculo com o verdadeiro significado do índice.

2.5. Índice de Theil

O Índice de Theil, como a Razão de Concentração de Gini, é uma medida de desigualdade que independe da forma da distribuição estudada. Tem sua origem na Teoria da Informação[10]. Tem sido usado em Economia como uma medida do grau de desigualdade de renda.

Veremos que a idéia de desigualdade contida no Índice de Theil é relativamente simples. No entanto, o seu entendimento fica condicionado à compreensão de alguns conceitos básicos de Teoria da Informação. Como resultado desta dependência da Teoria de nformação, o Índice de Theil também exige certa familiaridade na manipulação de logaritmos. Portanto, antecipada-

(10) Vide Henry Theil — **Economics and Information Theory**, Amsterdam, North Holland Publishing Co., 1967, página 123.

mente percebe-se que não se trata de uma medida de fácil penetração populou; mesmo entre os profissionais em geral, já que a Teoria de Informação não faz parte do curriculum básico, e logaritmo, infelizmente, é esquecido e, às vezes, subestimada a sua utilização, pelo economista.

Os conceitos básicos de Teoria de Informação que se fazem necessários para chegarmos ao Índice de Theil são os seguintes: conteúdo de informação, entropia ou informação esperada e informação esperada de uma mensagem indireta[11]. Nos restringiremos aqui às idéias emitidas por **Henry Theil** no seu famoso livro **Economics and Information Theory.** Qualquer divergência de nomenclatura sobre os conceitos básicos de Teoria de Informação deve ser encarada com certa tolerância, por se tratar de área relativamente recente de conhecimento. Para ilustrar esta situação reproduziremos a observação feita por **Elnyn Edwards**[12]:

"Infelizmente a nomenclatura da Teoria de Informação não está padronizada. Como observaram Mc Gill e Quastler (1965): "Constitui uma ironia o fato de a Teoria de Informação colocar problemas de comuncação para os que a utilizam! O que chamamos de "incerteza" foi denominado "Entropia" por Shannon (1948), "

Agora tentaremos explicar o que se entende por conteúdo de informação. Dado um evento **E**, cuja probabilidade de acontecer é igual a **p**, chamamos de conteúdo de informação da mensagem afirmando que **E** aconteceu, à uma função inversamente proporcional a sua probabilidade. No caso de Teoria de Informação utilizamos uma função **log na base 2** e, em alguns casos **log na base e**. Logo, o conteúdo de informação de um evento **E** ter acontecido é dado por:

$$h\ (p) = \log \frac{1}{2} = -\log p, \text{ onde } 0 < p < 1$$

Quando a base dos logaritmos é 2, as unidades de conteúdo de informação denominam-se **bits**; no segundo caso, quando se

(11) Vide Ramonaval A. Costa, «Medida de Desigualdade de Renda», **Boletim Geográfico,** N.º 238, jan/fev. de 1974, Ano 33, pp. 45/72.

(12) Elwin Edwards — **Introdução à Teoria da Informação.** Editora Cultura, Editora da USP, 1971.

usa logaritmos neperianos, as unidades são chamadas de **nits.** Na área de comunicação, onde a Teoria de Informação é muito aplicada, as unidades mais comumente encontradas são **bits.** Quando a aplicação dos conceitos de Teoria de Informação em outras áreas, como economia, têm-se utilizado os logaritmos neperianos com mais freqüência[13].

O outro conceito básico da Teoria de Informação, relacionado com o Índice de Theil é o que se chama de entropia ou informação esperada de várias mensagens. Neste caso teríamos um conjunto de eventos E_1, E_2 ... E_n com as seguintes probabilidades p_1, p_2 ... p_n. A entropia não é nada mais nada menos que o valor esperado dos respectivos conteúdos de informação, ou seja:

$$H(p) = \sum_{i=1}^{n} p_i \log \frac{1}{p_i} = - \sum_{i=1}^{n} p_i \log p_i$$

No caso da entropia observa-se uma distribuição, cujos valores do atributo são os eventos E_i e as respectivas probabilidades são os valores de p_i. Justamente no caso deste conceito é que se tem algumas discordâncias de nomenclatura. Edwards em 1964 chamou-o de grau de incerteza, **Shannon** em 1948 deu o nome de entropia, ao passo que **Wiener** em 1948 preferiu chamá-lo de **"entropia negativa"** O que interessa, no caso em questão, é que se compreenda **H(p)** como uma valor médio do conteúdo de informação de vários eventos.

Finalmente, o último conceito que se faz necessário, para uma melhor compreensão do Índice de Theil, chama-se informação esperada de uma mensagem indireta. Considermos E_1, E_2 ... E_n um conjunto de eventos e seja p_1, p_2 ... p_n suas probabilidades **a priori,** ou a probabilidade deles acontecerem sem nenhuma outra informação. Seja t_1, t_2 ... t_n as respectivas probabilidades **a posteriori,** ou as probabilidades corrigidas depois dos eventos terem acontecido. Podemos calcular facilmente o conteúdo de informação de uma mensagem indireta para cada evento E_i levando-se em consideração as duas informações sobre as probabilidade p_i e t_i:

$$h(p_i) - h(t_i) = \log \frac{1}{p_i} - \log \frac{1}{t_i} = \log \frac{t_i}{p_i}$$

(13) Henry Theil, op. cit.

Agora podemos obter o conceito de informação esperada de uma mensagem indireta. Basta calcular, como anteriormente, o valor esperado de $\log \frac{t_i}{p_i}$ ou seja, do conteúdo de informação de uma mensagem indireta. A única diferença que surge é que as ponderações tanto podem ser as probabilidades a **priori,** como as probabilidades a **posteriori,** ou seja:

$$I\,(t : p) = \sum_{i=1}^{n} t_i \log \frac{t_i}{p_i} \quad \text{ou}$$

$$I\,(p, t) = \sum_{i=1}^{n} p_i \log \frac{t_i}{p_i}$$

Com a apresentação destes três conceitos podemos escrever com mais clareza o que chamamos de Índice de Theil.

2.6. Entropia, Índice de Theil ou Redundância

Os conceitos anteriormente descritos — quando aplicados ao estudo do grau de desigualdade de uma distribuição de rendas — nos fornecem uma medida de desigualdade de renda, para o caso em que as informações individuais de rendas estão disponíveis, e uma outra medida adequada para as situações em que os dados de renda estão grupados por classes de renda(14).

Consideremos o primeiro caso, e seja r_1, r_2 ... r_n as frações da renda individual entre **N** indivíduos. Tomando-se o conceito de entropia e r_i a fração da renda total referente ao indivíduo **i**, podemos facilmente obter uma medida do grau de igualdade entre as diversas frações de renda, ou seja:

$$H(r) = \sum_{i=1}^{N} r_i \log \frac{1}{r_i}$$

(14) Ião devemos esquecer que esta medida, apesar de estar sendo descrita em termos do atributo renda, é adequada no caso de outros atributos em que a idéia de desigualdade faça sentido.

Para mostrarmos como a entropia, neste caso, nos fornece uma medida de igualdade, basta examinarmos os dois casos extremos relacionados com a distribuição de rendas. O primeiro caso é o de perfeita igualdade, onde r_i seria igual para todos os indivíduos $i=1, 2 \ldots N$. Neste caso, o valor de r_i seria igual a $1/N$ e o valor da entropia seria igual a:

$$H(r) = \sum_{i-1}^{N} \frac{1}{N} \log \frac{1}{1/N} N \frac{1}{N} \log N = \log N$$

A máxima igualdade ou máxima entropia seria dada por **log N.** É necessário observar que a máxima igualdade depende do número de indivíduos envolvidos na distribuição.

O segundo caso é o de máxima desigualdade, quando um único indivíduo estaria de posse de toda a renda, ou seja $r_i = 0$ para todo $i \neq j$, ao passo que $r_j = 1$. Nestas circunstâncias a entropia se reduz a um único termo:

$$H(r) = r_j \log \frac{1}{r_j} = 1 . \log \frac{1}{1} = \log 1 = 0$$

A máxima desigualdade significa entropia zero. Na realidade, nunca lidamos com nenhum dos extremos, logo o valor da entropia, no caso da informação individual de renda, está sempre compreendido entre **0** e **log N**:

$$0 \leq H(r) \leq \log N$$

Uma medida de desigualdade de renda pode ser obtida, imediatamente, a partir da medida de entropia de uma dstribuição. Basta subtrairmos da entropia máxima, **log N,** o valor da verdadeira entropia e obteremos uma medida do grau de desigualdade das rendas, isto é:

$$T(r) = \log N - H(r) = \log N - \sum_{i=1}^{N} r_i \log \frac{1}{r_i} =$$

$$= \log N + \sum_{i=1}^{N} r_i \log r_i = \sum_{i=1}^{N} r_i \log N r_i$$

Tal medida é chamada de Redundância. Esta é a medida de desigualdade de renda que se pode obter utilizando as noções de Teoria de Informação. A redundância é o que chamamos de Índice de Theil, pois **Henry Theil** é o economista que tem divulgado muito a utilização da Teoria de Informação em Economia.

Não há dificuldade para se perceber que o Índice de Theil (ou redundância) possui o valor zero quando temos uma situação de máxima igualdade e um valor igual a **log N** numa situação de máxima desigualdade. Além disto, é uma medida que depende do valor de **N.** Apesar de o logaritmo reduzir substancialmente este aspecto indesejável do Índice de Theil, ainda assim tal dependência de **N** constitui uma deficiência, mormente quando se fazem necessárias comparações de distribuições com valores de **N** substancialmente diferentes. Às vezes, o uso do Índice de Theil (redundância ou da entropia) sem se dar conta da influência de N, pode levar a evidências obscuras, imprecisas e ambíguas.

Para evitar os problemas introduzidos por **N** basta dividirmos tanto a entropia como o Índice de Theil por **log N** e teremos dminuído este aspecto indesejável das medidas em questão:

$$h(r) = \frac{\sum_{i=1}^{N} r_i \log \frac{1}{r_i}}{\text{Log } N} \qquad \text{(Entropia relativa)}$$

$$t(r) = \frac{\sum_{i=1}^{N} r_i \log_2 N r_i}{\text{Log } N} \qquad \text{(Redundância relativa)}$$

Os limites para ambas as medidas ficarão reduzidas ao limite 0 e 1. O Índice de Theil será enão entendido como a medida de redundância relativa, que é igual a 1 (um) menos a entropia relativa:

$$t(r) = \frac{T(r)}{\log N} = 1 - \frac{H(r)}{\log N}$$

$$t(r) = 1 - \frac{H\ (r)}{\log N}$$

Fazendo tal modificação os valores para o Índice de Theil ficariam compreendidos entre **0** e **1**, como no caso da razão de concentração de Gini:

$$0 \leq t\,(r) \leq 1$$

A máxima igualdade corresponderia a zero, como antes; a máxima desigualdade corresponderia ao valor **1**, eliminando os problemas de dimensão, ou seja, não haveria necessidade de se preocupar com os **bits** ou **nits.**

2.7 Informação de Uma Mensagem Indireta

Até agora apresentamos o Índice de Theil com a hipótese de que os dados individuais estariam disponíveis. Como comumente daparamos com distribuições em classes de renda, é interessante mostrarmos neste caso como o Índice de Theil deve ser interpretado. Quando não temos os dados individuais, o conceito de Teoria de Informação utilizado é o valor esperado do conteúdo de informação de uma mensagem indireta.

Se temos uma distribuição de freqüência, cujo atributo é a renda **R**, observamos dois tipos de informações. O primeiro são as freqüências relativas da população correspondente a cada classe de renda, ou seja n_i. O segundo tipo de informação são as frações de renda, r_i, correspondentes a cada classe de renda, ou seja, a fração de renda r_i recebida pela fração na população n_i. Estes dois tipos de informações podem ser interpretados como as probabilidades **a priori** e **a posteriori.** Conforme consideremos as frações de renda ou as frações de população como a probabilidade **a priori** obteremos medidas de desigualdade de renda, cuja diferença consiste apenas nos pesos utilizados, como explicamos anteriormente. No caso do nosso experimento, optamos por ponderar pela fração de renda em cada classe:

$$I\,(R, N) = \sum_{i=1}^{n} r_i \log \frac{r_i}{n_i}$$

Este Índice de Theil, utilizado para o caso em que a renda é divulgada em classes de rendas ou extratos possui como limite inferior o valor **I (R, N) = 0** e como limite superior **I (R, N) =**

$= \log N$. O primeiro valor corresponde a uma situação de perfeita igualdade, aqui entendida como $r_i = n_i$.

$$I(R, N) = \sum_{i=1}^{n} r_i \log \frac{r_i}{n_i} = \sum_{i=1}^{N} r_i \log 1 = 0$$

O segundo valor representa uma situação de máxima desigualdade, ou seja, $r_i = 0$, $r_j = 1$ para $i \neq j$. Aqui, novamente, I (R,N) se reduz a um único termo, cujo valor é o seguinte:

$$I(R, N) = r_j \log \frac{r_j}{1/N} = 1 . \log \frac{1}{1/N} = \log N$$

2.8. Interpretação do Índice de Theil para Dados Grupados

Esta medida, utilizada no caso de dados grupados além da interpretação, como a informação esperada de uma mensagem indireta, na qual r_1, r_2 r_n seriam as probabilidades **a posteriori** e n_1, n_2, n_n as probabilidades **a priori**, possui uma outra interpretação relacionada com o conceito de média geométrica. Esta segunda interpretação é mais simples do que a anterior. Consideremos a fórmula de **I (R,N)**:

$$I(R,N) = \sum_{i=1}^{N} r_i \log \frac{r_i}{n_i}, \quad i = 1,2, \quad n$$

Lembrando que r_i/n_i é a renda **per capita** da classe de renda i deflacionada pela renda **per capita** total, é possível interpretar **I (R, N)** como o logarítmo da média geométrica ponderada das rendas **per capita** de cada classe deflacionada pela renda **per capita** total, isto é:

$$M_G = e^{I(R, N)} = \prod_{i=1}^{n} \left(\frac{r_i}{n_i}\right)^{r_i} \quad i = 1,2, \ldots n$$

Tal interpretação evidencia como o Índice de Theil diminui a influência das diferenças absolutas de renda, ora pelo uso do logaritmo ora porque a média geométrica é menos sensível aos valores extremos de uma distribuição.

3. APLICAÇÕES

3.1. Problemas Metodológicos

Para a aplicação das medidas anteriormente apresentadas é necessário que se observe alguns critérios de caráter metodológico a fim de se evitar arbitrariedades na intepretação dos resultados. Este tipo de preocupação se torna mais premente quando fazemos comparações das medidas no tempo e no espaço. É preciso comparar medidas obtidas através de uma mesma metodologia sob pena de se observar diferenças resultantes apenas dos métodos de aferição.

Os problemas mais importantes a que devemos estar atentos, quando do uso das medidas de desigualdade são o tamanho da amostra, a escolha do elemento representativo da classe e o número de classes. Numa mesma distribuição obtida duas vezes a partir de amostras de tamanhos diferentes as respectivas medidas de desigualdade podem diferir, principalmente quando as amostras diferem também pela estratificação ou por outras características. A diversidade de elementos representativos das classes, também introduz diferenças significantes nos valores das medidas de desigualdade, isto é, se usarmos o ponto médio em vez da média da classe, como elemento representativo, as medidas de desigualdades apresentarão valores diferentes. O número de classes é outro problema que devemos ter em conta quando comparamos duas distribuições no tempo e no espaço. É um fato bem reconhecido em estatística que a escolha do número de classes de uma distribuição observada deve atender alguns requisitos para evitar distorções nas características da distribuição. Especialmente, no caso do atributo renda, estes requisitos nem sempre podem ser obedecidos em função da falta de conhecimento adequado sobre a forma da distribuição.

As tabelas 1 e 2 contêm exemplos de variações nos valores da Razão de Concentração de Gini e do Índice de Theil segundo o tamanho da amostra. A distribuição utilizada foi obtida a partir do Censo Demográfico de 1970 segundo a amostra de 25% e sub-amostra de 1,3%. Tanto no caso da Razão de Concentração de Gini, como no caso do Índice de Theil, quando passamos de uma amostra de 25% para uma de 1,3%, os valores da desigualdade diminuem. Esta característica é a mesma para qualquer tipo de elemento representativo da classe. (Ponto Médio, Renda

Média, Limite Inferior e Limite Superior) No caso especial destes exemplos as variações são muito pequenas, não chegando a prejudicar a interpretação dos resultados. Quando a renda média é o elemento representativo, a variação para a Razão de Concentração de Gini é de 1,86% e para o Índice de Theil é de 4,59%.

TABELA 1

VARIAÇÃO DOS VALORES DA RAZÃO DE CONCENTRAÇÃO DE GINI EM FUNÇÃO DO TAMANHO DA AMOSTRA PARA OS DIFERENTES ELEMENTOS REPRESENTATIVOS

Amostra / Sub-Amostra → Elementos Representativos das Classes de Renda ↓	Amostra 25%	Sub-Amostra 1,3%	Variação
1. Ponto Médio	0.358070	0.361354	—0.003284
2. Renda Média	0.328930	0.335041	—0.006111
3. Limite Inferior.	0.371520	0.373858	—0.002338
4 Limite Superior	0.339270	0.343185	—0.003915

FONTE: Censo Demográfico de 1970 — FIBGE

As tabelas 3 e 4 mostram as variações nas medidas de desigualdades quando utilizamos diferentes elementos representativos das classes no cálculo das medidas. Para a Razão de Concentração, quando comparamos os valores obtidos usando a Renda Média e o Limite Inferior, a maior variação absoluta atinge o valor de 0.04259, o qual, em termos percentuais, corresponde mais ou menos a 12,95%. Para o Índice de Theil, quando comparamos os valores obtidos usando o Limite Inferior e o Limite Superior, a maior variação absoluta atinge cerca de 0.025947 que, em termos percentuais, significa aproximadamente 22,19%. Estas variações em função de diferentes elementos representativos, são muito significativos, pois introduzem diferenças acima de 5% nos valores das medidas utilizadas. Apesar de os elementos representativos mais comumente utilizados serem a Renda Mé-

TABELA 2

VARIAÇÃO DOS VALORES DO ÍNDICE DE THEIL EM FUNÇÃO DO TAMANHO DA AMOSTRA PARA OS DIFERENTES ELEMENTOS REPRESENTATIVOS

Amostra / Sub-Amostra → Elementos Representativos das Classes de Renda ↓	Amostra 25%	Sub-Amostra 1,3%	Variação
1 Ponto Médio	0.133270	0.134560	—0.00129
2. Renda Média	0.106115	0.110984	—0.00487
3. Limite Inferior	0.116956	0.118236	—0.00128
4. Limite Superior	0.142903	0.143776	—0.00087

FONTE: Censo Demográfico de 1970 — FIBGE

dia ou Ponto Médio, as tabelas 3 e 4 servem de testemunho, no caso da distribuição estudada, para o grande perigo de comparações das medidas sem observar como foram obtidas.

O outro tipo de preocupação metodológica que se deve ter, quando da utilização dos instrumentais anteriormente descritos, é o número das classes em que a distribuição é apresentada. Se apresentarmos uma mesma distribuição com diferente número de classe, tudo indica que as medidas de desigualdade obtidas deverão diferir.

A fórmula para avaliar numericamente a razão de concentração de Gini envolve algumas aproximações a fim de facilitar a sua computação. A possibilidade de mensuração exata é praticamente impossível e a Estatística tem como característica justamente obter estimativas, cujos valores não pretendem ser exatos. Portanto qualquer diferença que surja em função do método de cálculo, do tamanho da amostra e disposição ou número de classes da distribuição observada, deve ser evitada.

Consideremos as tabelas 5 e 6 como exemplos de uma mesma distribuição. Na tabela 5 a distribuição contém 10 classes de ren-

TABELA 3

VARIAÇÃO DA RAZÃO DE CONCENTRAÇÃO DE GINI SEGUNDO OS ELEMENTOS REPRESENTATIVOS DAS CIDADES PARA AS DUAS AMOSTRAS

VARIAÇÕES → AMOSTRA SUB-AMOSTRA ↓	[1,2]	[1,3]	[1,4]	[2,3]	[2,4]	[3,4]
Amostra	0.029140	0.01345	0.01880	0.04259	0.01034	0.03225
Sub-amostra	0.026313	0.012504	0.018169	0.038817	0.008144	0.030673

FONTE: Censo Demográfico de 1970 — IBGE (só valores absolutos)

1 — Ponto Médio 2 — Renda Média 3 — Limite Inferior 4 — Limite Superior

TABELA 4

VARIAÇÃO DO ÍNDICE DE THEIL SEGUNDO OS ELEMENTOS REPRESENTATIVOS DAS CLASSES PARA AS DUAS AMOSTRAS

VARIAÇÕES → AMOSTRA SUB-AMOSTRA ↓	[1,2]	[1,3]	[1,4]	[2,3]	[2,4]	[3,4]
Amostra	0.027155	0.016314	0.009633	0.010841	0.036788	0.025540
Sub-amostra	0.023576	0.016324	0.009216	0.007252	0.032792	0.025947

FONTE: Censo Demográfico de 1970 — IBGE (só valores absolutos)

1 — Ponto Médio 2 — Renda Média 3 — Limite Inferior 4 — Limite Superior

da e proporciona um valor para a Razão de Concentração de Gini em torno de 0.474688. Ao passo que, na tabela 6 a mesma distribuição é apresentada com apenas 7 classes de renda e o valor da Razão da Concentração de Gini diminui para 0.463927

Este pequeno exemplo, apesar de não estabelecer a direção da modificação da variação do valor da medida em questão, pois tudo depende de como as classes são escolhidas, serve apenas como uma ilustração de variações nas medidas em função do arranjo das distribuições em classes de atributos. No exemplo escolhido apesar da variação não ultrapassar a casa dos 2,5% é necessário que se tenha em mente mais esta possibilidade de variação, quando resultar ou de comparações no tempo, ou de comparações espaciais.

TABELA 5

Classes de Renda	População	Razão de Concentração de Gini
1. 0 — 67,20	106.664	
2. 67,20 — 134,40	324.342	
3. 134,40 — 268,80	1.070.870	
4. 268,80 — 537,60	1.935.874	
5. 537,60 — 806,40	810.886	0.474688
6. 806,40 — 1.344,00	474 472	
7 1.344,00 — 1.881,60	168.729	
8. 1.881,60 — 2.688,00	123.333	
9. 2.688,00 — 4.032,00	98.757	
10. 4.032,00 ou mais	52.275	

FONTE: PNAD/72 — Mão-de-Obra

3.2. Alguns Exemplos de Aplicação

A fim de ilustrarmos a utilização das medidas anteriormente discutidas apresentaremos alguns exemplos, bem simples, em que as medidas serão usadas com o intuito de evidenciar características regionais da desigualdade da renda que estariam escondidas se não tivesse havido a preocupação de desagregação espacial. Também mostraremos como estas medidas têm sido usadas em conexão com outros atributos diferentes da renda.

TABELA 6

	Classes de Renda		População	Razão de Concentração de Gini
1.	0 —	67,0	106.664	
2.	67,20 —	134,40	324.342	
3.	134,40 —	268,80	1.070.870	
4.	268,80 —	537,60	1.935.844	0.463927
5.	537,60 —	1.344,00	1.285.358	
6.	1.344,00 —	2.688,00	292.062	
7	2.688,00 ou mais			

FONTE: PNAD/72 — Mão-de-Obra

3.2.1. Aplicação com o Atributo Renda

A utilização das medidas de desigualdade com o atributo renda já se encontra muito divulgada, tendo até mesmo adquirido uma certa popularidade entre as pessoas, como por exemplo a Razão de Concentração de Gini. Qualquer exemplo será sempre uma repetição, já que existem vários livros e trabalhos documentando o uso deste instrumental.

Nas tabelas 7 e 8 destacamos dois tipos de utilização para a razão de concentração e para o coeficiente de Theil (Índice de Theil). Na primeira tabela, uma comparação internacional dramatizando as diferenças do grau de desigualdade de renda entre os países, com o Brasil apresentando um valor da medida de dispersão relativa de Gini, a razão de concentração muito elevada, em torno de 0.583 em oposição aos Estados Unidos e Tchecoslováquia cujos valores da razão de concentração são muito pequenos. Esta simples comparação põe em evidência quão desigual é a renda pessoal no Brasil, em 1970. Além disto mostra a factibilidade de um menor grau de desigualdade de renda em países desenvolvidos ou com outra organização política.

A informação sobre a distribuição da renda familiar de acordo com a situação urbana, suburbana e urbana pode ser vista na tabela 8 utilizando outra medida de desigualdade — o coeficiente de Theil. Parece muito claro que o grau de desigualdade da ren-

TABELA 7

COMPARAÇÃO INTERNACIONAL
DA DESIGUALDADE DE RENDA PESSOAL

Países	Razão de Concentração de Gini
Brasil (1970)	0.583
Estados Unidos (1969)	0.34
Tchecoslováquia (1970)	0.204

FONTE: Ramonaval A. Costa, **Size Income Distribution of Brazil in 1970. A cross-section Analysis of Income.** Dissertação de Doutorado, 1975.

da pessoal é maior nas áreas urbanas e menor nas áreas rurais; as áreas suburbanas teriam uma posição intermediária. Não há dúvida, novamente, que o coeficiente de Theil pode ser utilizado para ressaltar característieas interessantes que podem passar desapercebidas do pesquisador que não se conscientiza da possibilidade de utilização destas medidas de desigualdade como coadjuvante nas análises realizadas.

TABELA 8

DISTRIBUIÇÃO DA RENDA FAMILIAR
1970

Situação	Coeficiente de Theil
1. Rural	0.3782
2. Suburbana	0.3802
3. Urbana	0,5573

FONTE: Carlos Geraldo Langoni, **Distribuição da Renda e Desenvolvimento Econômico do Brasil**, Editora Expressão e Cultura, Rio de Janeiro, 1973.

Nos dois trabalhos que serviram de fonte para a elaboração das tabelas 7 e 8 podem ser encontrados inúmeros exemplos de utilização destes indicadores de desigualdade de renda pessoal ou familiar. Através do exame da utilização dessas medidas por vários técnicos, cuja imaginação trabalha, sem descanso, na procura de certas evidências a respeito da desigualdade das rendas, é que o iniciante vai apreendendo as múltiplas utilidades que estas medidas podem ter.

3.2.2. Aplicações com Outros Atributos

Como ressaltamos no começo, o uso destas medidas de desigualdade não se restringe a situações em que a renda é o atributo primordial. Elas todas podem ser utilizadas com atributos diferentes da renda; no entanto, esta utilização é ainda muito incipiente, ora porque os técnicos que teriam chances de usá-las em outros atributos, têm indicadores alternativos, ora por falta de melhor compreensão dos indicadores e suas propriedades. Não há dúvida que, em termos de renda, o seu significado se torna mais fácil de compreender em função da intensa utilização, ou porque o atributo renda permite interpretações mais interessantes e informativas. Mas creio que a grande inibição no uso destas medidas reside no fato de que o conhecimento estatístico das pessoas não permite o uso de indicadores que vão além do significado de média e da variância.

Em que pese as dificuldades para colher exemplos de utilização da razão de Concentração de Gini com atributos que não sejam a renda, podemos encontrar em alguns trabalhos a utilização de outras variáveis como: desigualdades dos anos de estudos, desigualdade na distribuição das terras, etc.

A desigualdade dos anos de estudos é uma variável que vem sendo utilizada na área de educação e sociologia como tentativa de medir variações na desigualdade dos anos de estudo de um país ou região. Na tabela 9 temos esta informação para os estados de São Paulo e Guanabara. A medida de desigualdade nos mostra a Guanabara apresentando um menor grau de desigualdade do que São Paulo. As diferenças na desigualdade educacional são mais difíceis de serem interpretadas em virtude da influência marcante da distribuição das idades impondo um certo grau de desigualdade do atributo.

TABELA 9

DESIGUALDADE DOS ANOS DE ESTUDOS

Estados	Coeficiente de Variação
São Paulo	0.676
Guanabara	0.619

FONTE: Ramonaval Augusto Costa, **Size Income Distribution of Brazil in 1970 — A cross-section Analysis of Income**, abril, 1975

Outra variável que pode ser observada sob o ângulo da desigualdade é a concentração da Propriedade da terra, de ter sido usada, mormente por geógrafos, nos estudos sobre áreas rurais. A tabela 10 apresenta esta variável associada a duas medidas de desigualdade, evidenciando que em São Paulo a diferença no grau desigualdade indica uma maior concentração de terras em São Paulo. Novamente a utilização das medidas de desigualdade oferecem a possibilidade de melhorar as características da análise, ressaltando, oportunamente, pontos que de outro modo não seriam devidamente descritos com grande chance de serem esquecidos no decorrer da análise.

TABELA 10

CONCENTRAÇÃO DE PROPRIEDADE DA TERRA, 1970

Estados	Coeficientes	
	Gini	Theil
São Paulo	0.610	0.088
Guanabara	0.459	0.560

FONTE: R. Hoffman — **"Contribuição à Análise da Distribuição da Renda e da Posse da Terra no Brasil"** Piracicaba, Mimeografada, E.S.A.L. Queiroz, 1971.

As informações das tabeas 7. 8, 9, e 10, apesar de serem interessantes pela informação que contêm, devem ser usadas em conjunto com outros atributos a fim de que a idéia de desigualdade seja apenas uma informação adicional entre outras.

4. CONCLUSÕES

O objetivo deste trabalho, como foi dito no início, era chamar atenção para o conjunto de medidas de Concentração e para sua possível utilização com outros atributos diferentes da renda, com a qual praticamente a maioria delas tiveram sua origem.

Quando a idéia de concentração, que se pretende medir, não se vincula a elementos nos quais a concentração se verifica, estas medidas de desigualdade são instrumentais adequados na tentativa de avaliar a posição relativa de unidades geográficas ou de pessoas. Desde que sua utilização seja cautelosa, procurando sempre que possível contornar os problemas metodológicos existentes, tais indicadores são certamente úteis e podem ser usados por qualquer técnico interessado, seja economista, geógrafo, agrônomo, sociólogo etc.

BIBLIOGRAFIA

(1) AMES, Edward, «A Method for Estimating the Size Distribution of a Given Aggregate Income» — **The Review of Economics and Statistics,** Julho de 1942.

(2) BOWMAN, Mary J. «A Graphical Analysis of Personal Income Distribution» **A.E.R.** XXXV, Setembro de 1945, pp. 607/628.

(3) CHENERY, Hollis, Ahluwalia, Montek S; Bel, G.L.; Duloay, John H. e Jolly, Richard, **Redistribution and Growth,** publicado pelo Banco Mundial, Oxford University Press, 1974.

(4) CLINE, William R. **Potential Effects of Income Redistribution on Economics Growth — Latin American Cases,** Praeger Publiser, 1972.

(5) COSTA, Ramonaval A. Size Income Distribution of Brazil in 1970 — A Cross-Section Analysis of Income Distribution bu Ocupations, Tese de Doutorado apresentada na Universidade de Vanderbilt, agosto de 1975, publicado pelo IBGE em 1977.

(6) ————————, «Bem-Estar e Indicadores de Desigualdade» **Revista Brasileira de Estatística,** Rio de Janeiro, Fundação IBGE, Julho de 1974.

(7) ————————, «Medidas de Desigualdade de Renda» **Boletim Geográfico** n.º 238, Jan/Fev. de 1974, Ano 33 pp. 45/72.

(8) DUARTE, J.C. «Aspectos da Distribuição de Renda no Brasil em 1970», Piracicaba, Tese de Mestrado da E.S.A.L. Queiroz, 1971.

(9) EDWARDS, Elwyn, **Introdução à Teoria da Informação,** Editora Cultura, Editora da USP, 1971.

(10) GALVANI, L. «Sulle Curve di Concentrazione Relativa a Caratteri non Limitati e Limitati», Metron, Vol. X, n.º 3, Roma, Otubro, 1932.

(11) GASTWIRTH, Joseph L., The Estimation of the Lorenz Curve and Gini Index, **The Review of Economics and Statistics,** LIV, agosto, 1972.

(12) GINI, Corrado «Indice di Concentrazione e di Dipendenza» **Biblioteca Dell'Economista,** Roma, Editice Torinese Milano, 1922 pp. 5.

(13) HOFFMAN, R. — «Contribuição à Análise da Distribuição da Renda e da Posse da Terra no Brasil» Piracicaba, Tese não publicada, E.S.A.L. Queiroz, 1971.

(14) KINGSTON, Jorge, Desigualdade na Distribuição das Rendas **Revista Brasileira de Economia,** Rio de Janeiro, FGV, março 1955, p. 5.

(15) LANGONI, C.G. Distribuição da Renda e Desenvolvimento Econômico do Brasil, Rio de Janeiro, Expressão e Cultura, 1973.

(16) LORENZ, M. O. — «Methods of Measuring the Concentration of Wealth» **A.S.A.** New Series, Junho de 1905, p. 209.

(17) MEIRELLES, Antonio C. «A Evolução da Estrutura do Sistema Bancário Brasileiro no período de 1966-1974» Sindicato dos Bancos do Estado da Guanabara, 1975, Ano III, n.º 8.

(18) DE FINETTI, B. e PACIELLO, U., «Calcolo Della Differenza Media» **Metron,** Vol. VIII n.º 3, Fevereiro de 1931.

(19) THEIL, Henry, **Economics and Information Theory,** Amsterdam, North Holland Publishing Co. 1967.