

# Um Novo Índice de Concentração: Aplicação ao Estudo da Evolução Recente da Distribuição de Renda no Brasil (\*)

João Luiz Maurity Saboia(\*\*)

## R E S U M O

Neste trabalho desenvolvemos um novo índice de concentração. Ele está baseado na hipótese de que a distribuição log-normal é adequada para representar o processo cuja concentração está sendo estudada. Utilizamos também a aproximação da distribuição normal pela logística, sugerindo, portanto, a distribuição log-logística para representar a distribuição de renda. Além de utilizada para a renda pessoal de indivíduos a log-normal tem sido também empregada em estudos de distribuição do capital das firmas. Desta forma, nosso índice pode ser utilizado tanto como medida de concentração de renda quanto para estudos de concentração industrial.

A grande vantagem do índice proposto está na sua facilidade de cálculo. Para tal, basta que se possua o número de observações (indivíduos, firmas, etc.) em cada faixa da variável estudada (renda, capital, etc.). A partir daí, através de simples operações matemáticas, obtemos uma linearização dos dados. De posse dos dados linearizados podemos obter nosso índice através de uma regressão linear simples. Uma das vantagens deste (...)

---

(\*) Trabalho baseado em duas versões anteriores. A primeira foi apresentada na XXIX Reunião Anual da SBPC, São Paulo, julho de 1977. A segunda foi apresentada no V Encontro da ANPEC, Rio de Janeiro, 1977. O autor agradece as críticas e sugestões apresentadas por Adriano Dias, Antonio Q. Cavalcante e Ricardo Varsano.

(\*\*) O autor é professor da Faculdade de Economia da Universidade do Rio de Janeiro.

---

## 1 INTRODUÇÃO

---

A divulgação dos resultados do Censo de 1970, com a revelação do crescente processo de concentração de renda verificado no Brasil, durante a década dos sessenta, deu origem a um intenso debate entre técnicos e economistas. De um lado se colocaram os críticos da política econômica governamental e de outro aqueles que se identificavam com a mesma<sup>(1)</sup>.

Como veremos no decorrer deste trabalho, o processo de concentração de renda parece ter continuado durante os primeiros anos da década dos setenta. Até 1972 o processo de concentração continuou acelerado. A partir de 1973 parece ter havido uma pequena recuperação. De qualquer forma ainda estaríamos em 1975 com uma distribuição de renda semelhante a do início da década.

As conclusões acima são o resultado da utilização de um novo índice de concentração proposto neste trabalho<sup>(2)</sup>. Nosso índice é bastante similar ao índice de Pareto. Diferentemente de Pareto, entretanto, o índice se aplicaria a toda a faixa de ren-

---

(...)

tipo de procedimento é que o índice de concentração não é um simples número, mas possui uma função de distribuição. Desta forma podemos utilizar testes de hipóteses para a verificação da melhora ou piora da distribuição da variável considerada.

Nosso índice foi utilizado inicialmente para comprovar o processo de concentração de renda ocorrido no país durante a década dos sessenta. Os resultados encontrados com nosso índice são coerentes com os resultados obtidos com outros índices (Gini, Theil e Variância dos Logs).

Em seguida o índice foi utilizado com os dados do Imposto de Renda e da Lei dos 2/3 para verificação da evolução da distribuição de renda a partir de 1968. Os resultados encontrados foram testados estatisticamente e nos sugerem que o processo de concentração de renda verificado no país durante a década passada continuou até 1972/73. A partir de então, parece ter havido uma pequena recuperação, de forma que a distribuição de renda em 1975 seria similar a de 1970.

(1) Veja Tolipan & Tinelli (1975) e Langoni (1973).

(2) Na realidade, trabalhando independentemente, e de forma bastante distinta, Dagum (1977) sugere o índice aqui proposto como sendo um «parâmetro de igualdade». Entretanto, Dagum não propõe que seu «parâmetro de igualdade» seja utilizado como índice de concentração.

da e não apenas às rendas dos extratos superiores. Existe também uma grande relação entre o índice que propomos, a variância dos logaritmos (variância dos logs) e o índice de Gini.

O índice aqui apresentado está baseado na hipótese de que a distribuição de renda pode ser representada adequadamente por uma distribuição log-normal<sup>(3)</sup>. Utilizamos também a aproximação da distribuição normal pela logística, sugerindo, portanto, a distribuição log-logística para representar a distribuição de renda. Além de utilizada para a renda pessoal de indivíduos a log-normal tem sido também empregada em estudos de distribuição de capital de firmas<sup>(4)</sup>. Portanto, é de se esperar que nosso índice possa ser também satisfatoriamente utilizado para estudos de concentração industrial.

A grande vantagem do índice ora proposto está na sua facilidade de cálculo. Para tal, basta que se possua o número de recipientes em cada faixa de renda. A partir daí, através de simples operações matemáticas, obtemos uma linearização dos dados. De posse dos dados linearizados podemos obter nosso índice através de uma regressão linear simples. Uma das vantagens deste tipo de procedimento é que o índice de concentração não é um simples número, mas possui uma função de distribuição. Desta forma podemos utilizar testes de hipóteses para a verificação da melhora ou piora da distribuição de renda considerada.

Além de dados dos Censos de 1960 e 1970, utilizamos neste trabalho dados do Imposto de Renda (IR) e da Lei dos 2/3. Os dados do IR correspondem ao período 1968-1975 e compreendem a população declarante do IR, a qual representava em 1975 cerca de metade da população economicamente ativa (PEA)<sup>(5)</sup>. Para não limitar o nosso estudo somente à população declarante do IR, levamos também em consideração a PEA não declarante do IR, incluindo-a na faixa mais baixa de rendimentos. Como veremos adiante as diferenças encontradas com nosso índice ao se

---

(3) A adequabilidade da distribuição log-normal para a descrição da variável renda já foi amplamente discutida por diversos autores, tais como, Aitkinson e Brown (1966).

(4) Veja Gibrat (1931), Kalecki (1945), Hart & Prais (1956) e Saboia (1977).

(5) Durante o período 1968-1975 a população declarante de IR aumentou em quase quatro vezes passando de 4.4 milhões para mais de 16 milhões.

trabalhar apenas com os declarantes do IR ou com toda a PEA não são significantes. Os dados da Lei dos 2/3 correspondem ao período 1968-1974 e compreendiam 8.8 milhões de trabalhadores em 1974. É interessante observar que os resultados encontrados com os dados do IR e da Lei dos 2/3 são similares, embora as populações consideradas nas duas amostras sejam distintas.

Na seção 2. é feita uma crítica à Lei de Pareto e ao índice de Pareto, o qual só se mostra adequado para as mais altas faixas de renda.

Na seção 3. apresentamos o índice de concentração a partir da hipótese de log-normalidade. Mostramos as transformações necessárias para a linearização dos dados e a forma de cálculo do índice após a linearização. O índice de concentração surge naturalmente ao aproximarmos a função de distribuição normal por uma distribuição logística<sup>(6)</sup>. Nesta seção mostramos também a relação entre o índice proposto, a variância dos logs e o índice de Gini.

Na seção 4. aplicamos o índice para a verificação da concentração de renda entre os anos 1960 e 1970. Os resultados encontrados confirmam o fato já bastante conhecido e discutido de que houve um intenso processo de concentração de renda durante aquele período.

Na seção 5. o índice é aplicado para a verificação da evolução da distribuição de renda no Brasil durante o período 1968-1975. O estudo é feito em 3 partes: considerando apenas a população declarante de IR, considerando uma PEA de 25 milhões e de 30 milhões de pessoas respectivamente<sup>(7)</sup>. Nos três casos as conclusões são semelhantes — o processo de concentração de renda foi intenso durante os anos do “milagre econômico” —, apresentando uma pequena recuperação a partir de 1973/74. Estes resultados são testados estatisticamente, o que vem a reforçar a tendência à piora de distribuição de renda, durante os anos do “milagre”

---

(6) A justificativa para esta aproximação pode ser encontrada em Berkson (1969). Alternativamente Dagum, *op. cit.*, desenvolve um modelo que sugere a distribuição log-logística para a distribuição de renda.

(7) Segundo dados do Censo de 1970 a PEA com algum tipo de remuneração compreendia cerca de 26 milhões de pessoas. Acreditamos que durante os anos considerados (1968-1975) a PEA com rendimentos tenha permanecido dentro da faixa entre 25 e 30 milhões de pessoas.

Na Seção 6. o índice é aplicado aos dados da Lei dos 2/3 para verificar a evolução da distribuição de renda no Brasil no período 1968-1974. Os resultados encontrados corroboram as conclusões tiradas com os dados do IR.

Finalmente, na seção 7., apresentamos as conclusões principais do trabalho.

O trabalho possui ainda 2 (dois) Apêndices. O Apêndice 1. apresenta alguns resultados sobre a função de distribuição logística, a qual, por fornecer uma boa aproximação à função de distribuição da normal, transforma-se em ponto de partida para a definição do índice. O Apêndice 2. apresenta o desenvolvimento estatístico para o teste de hipóteses utilizado.

---

## 2. UMA CRÍTICA À LEI DE PARETO

---

Pareto criou uma lei que, se aplicada à distribuição de renda, só se ajusta às rendas das classes privilegiadas. Trabalhando com dados de diversos países em diversos períodos (Inglaterra, Prússia e Saxônia no século XIX, Peru no século XVIII, Basileia medieval, etc.) chegou a uma curva de distribuição de renda que acreditava "natural". Além de não ter nada de natural, sendo uma característica do sistema social considerado, como já havia notado Oskar Lange<sup>(8)</sup>, gostaríamos de chamar a atenção para a inconveniência da utilização do índice de Pareto quando quisermos estudar a distribuição de renda de uma determinada população.

A lei de Pareto<sup>(9)</sup> pode ser representada pela seguinte equação

$$N = \frac{\alpha}{(R-R_0)^\beta} \quad (2.1)$$

---

(8) Veja Lange (1963).

(9) Veja Pareto (1897).

onde

$R_0$  — renda mínima

$R$  — renda

$N$  — número de pessoas com renda igual ou superior a  $R$

$\alpha, \beta$  — parâmetros positivos

Na prática é costume fazer-se  $R_0 = 0$  de forma que a lei de Pareto se transforma em

$$N = \frac{\alpha}{R^\beta} \quad (2.2)$$

A Equação (2.2) pode ser representada graficamente por uma hipérbole tangenciando os eixos  $R$  e  $N$ . (Veja Figura 2.1).

O coeficiente de Pareto é o parâmetro  $\beta$  da Equação (2.2). Quando  $\beta$  tende para infinito o gráfico da Figura 2.1. tende para sua assíntota vertical, o que significa que se tende para uma distribuição de renda perfeita. Quando  $\beta$  tende para zero, o gráfico da Figura 2.1. tende para uma reta horizontal, o que significa que a desigualdade de renda é máxima. Assim, à medida em que  $\beta$  cresce de zero para infinito a distribuição de renda melhora.

A Equação (2.2) pode ser facilmente linearizada tomando-se o logaritmo dos dois lados da equação. Assim obtemos:

$$\log N = \log \alpha - \beta \log R \quad (2.3)$$

Na prática os dados de renda não se localizam exatamente sobre a curva de Pareto de forma que um termo representando o erro deve ser incluído na Equação (2.3). Desta forma podemos estimar o coeficiente de Pareto  $\beta$  pelo método dos mínimos quadrados ordinários.

A imperfeição da Lei de Pareto pode ser facilmente observada notando-se que na Equação (2.2)  $N$  tende para infinito à medida em que  $R$  tende para zero. Na prática existe um número finito de pessoas, de forma que a situação real seria melhor representada pela Figura 2.2., onde  $R_1$  é a renda abaixo da qual a

FIGURA 2.1. - Curva de Pareto.

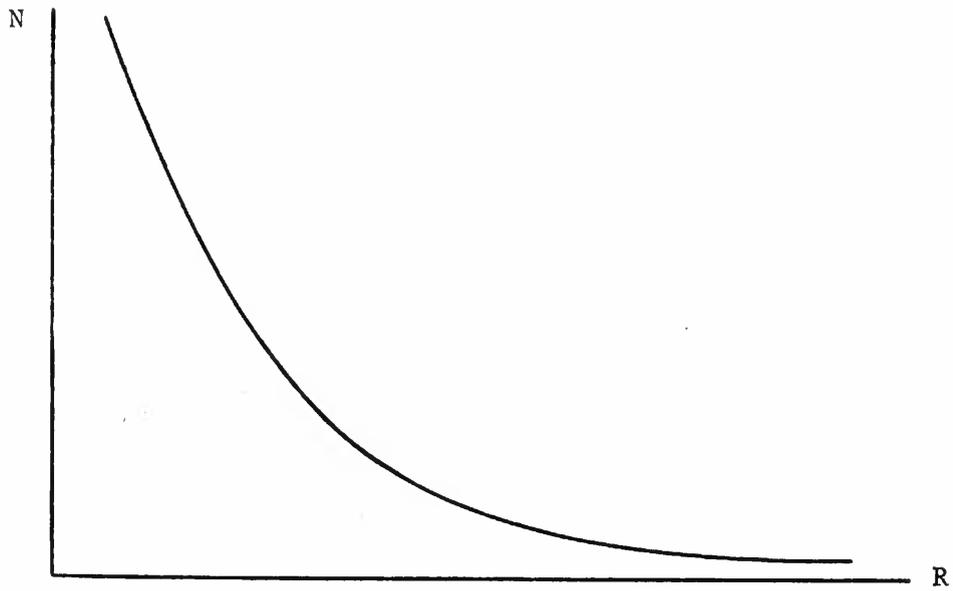


FIGURA 2.2. - A distribuição real de renda e a curva de Pareto

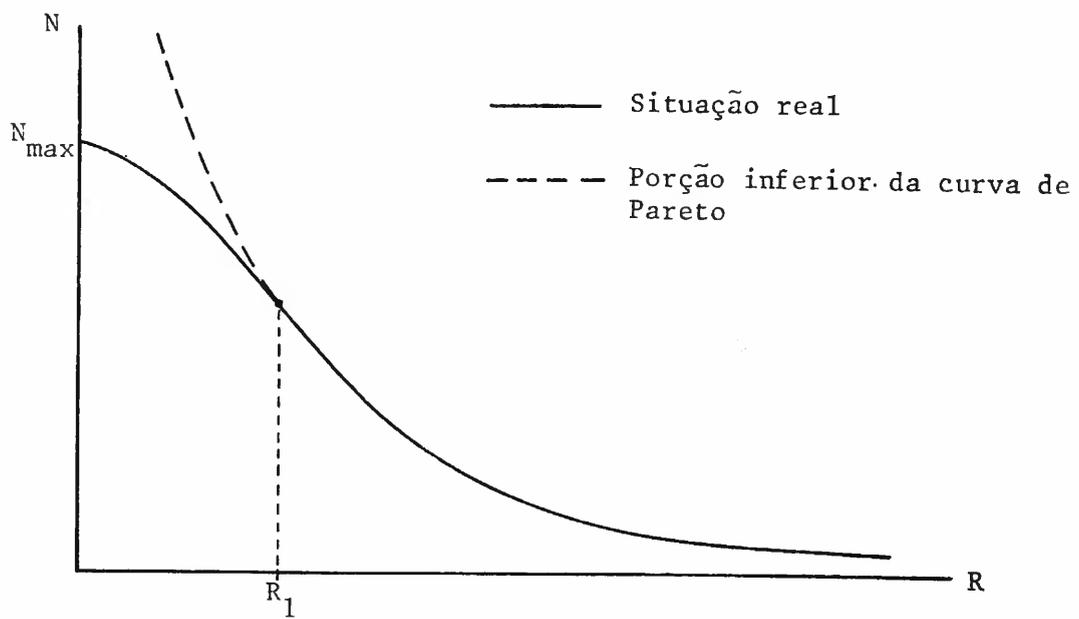
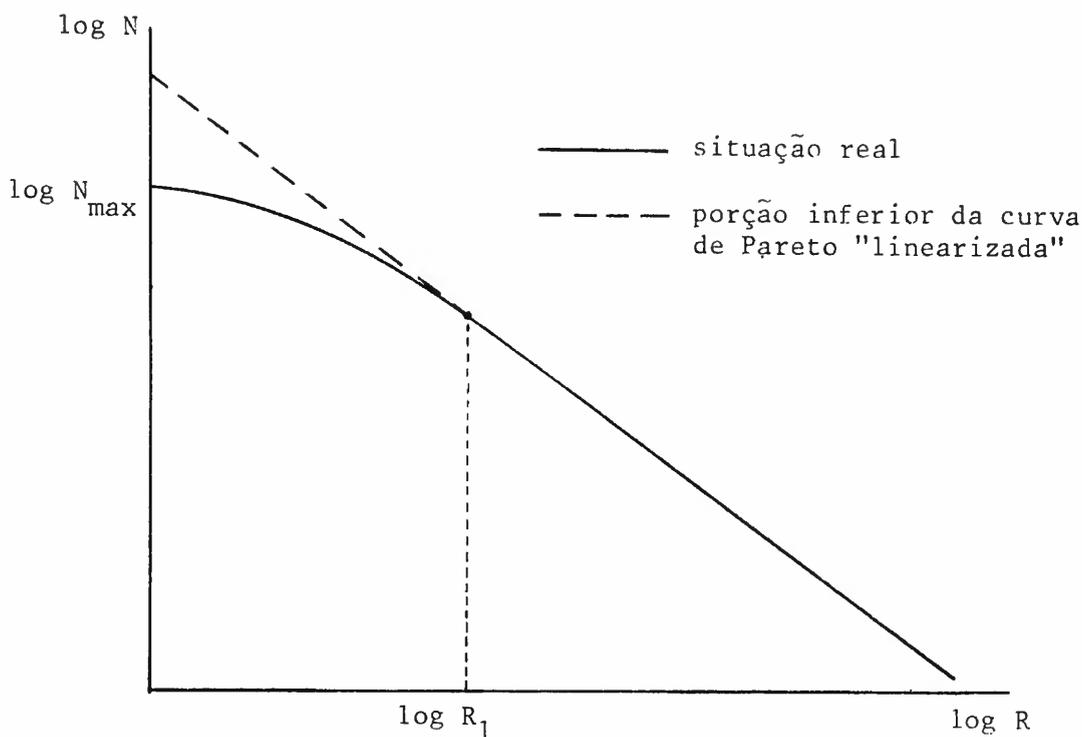


FIGURA 2.3. - "Linearização" da curva de Pareto.



Lei de Pareto deixa de ter validade e  $N_{\max}$  representa a população total. Desta forma, ao logaritmos as variáveis  $N$  e  $R$  obtemos a Figura 2.3., onde apenas a parte correspondente às rendas altas (a partir de  $R_1$ ) aparece efetivamente linearizada.

Na prática a renda  $R_1$  tende a ser elevada de forma que o coeficiente de Pareto só se aplica para as altas faixas de renda<sup>(10)</sup>

Um último ponto que gostaríamos de mencionar aqui é que a descoberta da Lei de Pareto e sua aplicabilidade à fenômenos tão distintos, tais como, distribuição de renda, firmas por capital, escravos entre seus senhores, artigos por autores, trabalhadores por sindicatos, etc.<sup>(11)</sup>, não tem nada de especial e não deve constituir uma surpresa. A grande maioria das funções de distribuição possuem a forma aproximada da Figura 2.2. (linha

(10) A adequabilidade da Lei de Pareto para altas rendas é contestada por alguns autores. Hayakawa (1951) afirma que a Lei de Pareto é adequada para rendas médias e não para rendas baixas ou altas.

(11) Veja Steindl (1965).

cheia). Portanto nada mais natural do que o fato da Lei de Pareto se aplicar satisfatoriamente à parte superior da maioria das distribuições. O que não deve ser esquecido entretanto é o fato de existir um valor mínimo da variável considerada ( $R_1$  em nosso caso) a partir do qual a distribuição em questão se aproxima da distribuição de Pareto.

---

### 3. O ÍNDICE DE CONCENTRAÇÃO PROPOSTO

---

As curvas de distribuição de renda se apresentam bastante assimétricas com longas caudas à direita (Veja Figura 3.1.). Uma das distribuições mais conhecidas que apresenta este tipo de comportamento é a distribuição log-normal. Sua aplicabilidade à distribuição de renda tem sido amplamente discutida, sendo que diversos modelos matemáticos foram desenvolvidos sugerindo a log-normal como adequada à descrição da distribuição de renda<sup>(12)</sup>

A distribuição log-normal aparece naturalmente quando estudamos variáveis que podem ser interpretadas como sendo um produto de variáveis aleatórias<sup>(13)</sup>. Este é o caso da variável renda que pode ser interpretada como resultante de uma renda inicial que sofre o efeito de taxas de crescimento a cada unidade de tempo. A grande aplicação da variável log-normal deve-se ao fato de se obter uma variável normal ao ser feita a transformação logarítmica (Veja Figura 3.2.).

Se ao invés de considerarmos a população nas diversas faixas, utilizarmos a distribuição cumulativa, i. e., considerarmos o número de pessoas com renda inferior a cada nível de renda, obtemos a Figura 3.3..

---

(12) A log-normal se ajusta tanto melhor quanto mais homogênea for a população e o tipo de rendimento considerado. Assim ao utilizarmos a log-normal para a distribuição de renda de uma determinada profissão, devemos obter um ajuste melhor do que se considerarmos toda a população. No segundo caso não apenas a população será mais heterogênea como também os tipos de rendimentos (salários, renda, etc.). Veja Aitkinson & Brown, op. cit..

(13) Veja Gibrat, Kalecki e Saboia, op. cit.

FIGURA 3.1. - Distribuição de uma população pelas diversas faixas de renda com respectivas curva ajustada.

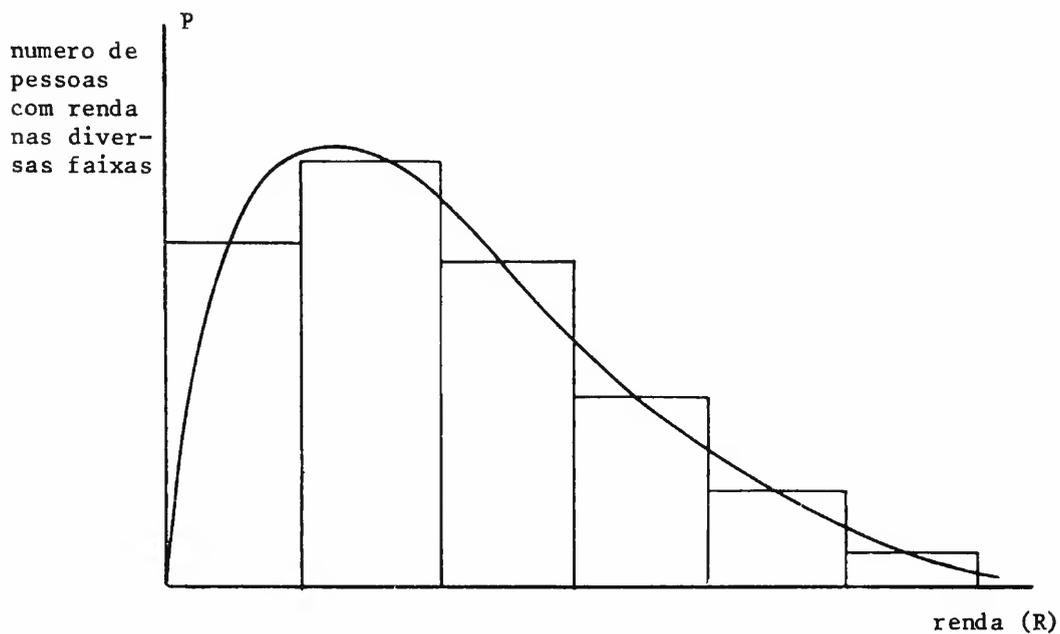
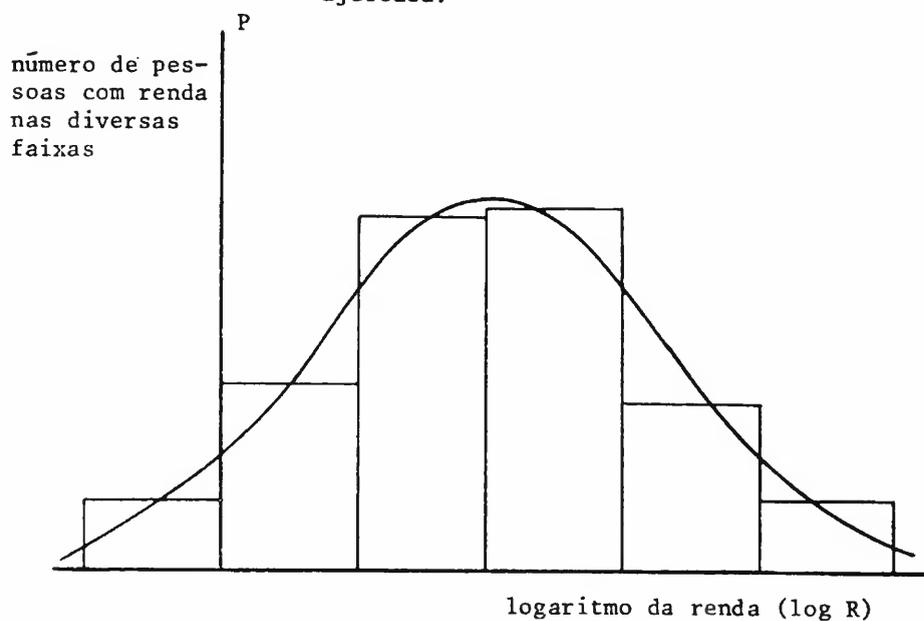


FIGURA 3.2. - Distribuição de uma população pelas diversas faixas de rendas logaritmadas com respectiva curva normal ajustada.



Para estudos comparativos da evolução da distribuição de renda torna-se vantajosa<sup>(14)</sup> a padronização da curva normal cumulativa considerando-se no eixo vertical, ao invés do número absoluto de pessoas, a porcentagem de pessoas com renda inferior à  $R$ . Assim obtemos a Figura 3.4..

Uma distribuição que se ajusta satisfatoriamente à distribuição acumulada da normal é a distribuição logística<sup>(15)</sup>. A distribuição logística é dada por

$$y = \frac{1}{1 + \alpha e^{-\beta x}} \quad (3.1)$$

onde

$x = \log R$  é o logaritmo da renda  $R$ ;

$y$  é a proporção de pessoas com renda inferior à  $R$ ;

$\alpha$  e  $\beta$  são parâmetros (positivos) da distribuição logística.

Os parâmetros  $\alpha$  e  $\beta$  determinam a forma da distribuição logística. O parâmetro  $\alpha$  é responsável pela localização da curva sobre o eixo  $x$ , i.e., dá uma idéia do padrão de vida médio da população ( $\alpha$  grande desloca a curva para a direita, o que significa uma população mais rica;  $\alpha$  pequeno desloca a curva para a esquerda, o que significa uma população mais pobre).

O parâmetro  $\beta$  é responsável pela inclinação da curva logística, i.e., dá uma idéia da distribuição da renda pela população ( $\beta$  grande implica numa curva mais inclinada, o que significa uma distribuição de renda mais igualitária;  $\beta$  pequeno implica numa curva menos inclinada, o que significa uma distribuição de renda mais desigual).

---

(14) Como a população varia ao longo do tempo a curva da Figura 3.3. se modificaria mesmo se a distribuição de renda não se modificasse. Com a curva padronizada da Figura 3.4. isto não ocorre.

(15) A vantagem de se aproximar a distribuição acumulada da normal pela distribuição logística é que esta última pode ser explicitada matematicamente enquanto o mesmo não ocorre com a distribuição normal. Veja nota (6).

FIGURA 3.4. - Distribuição cumulativa de uma população pelas diversas faixas de rendas logaritmadas.

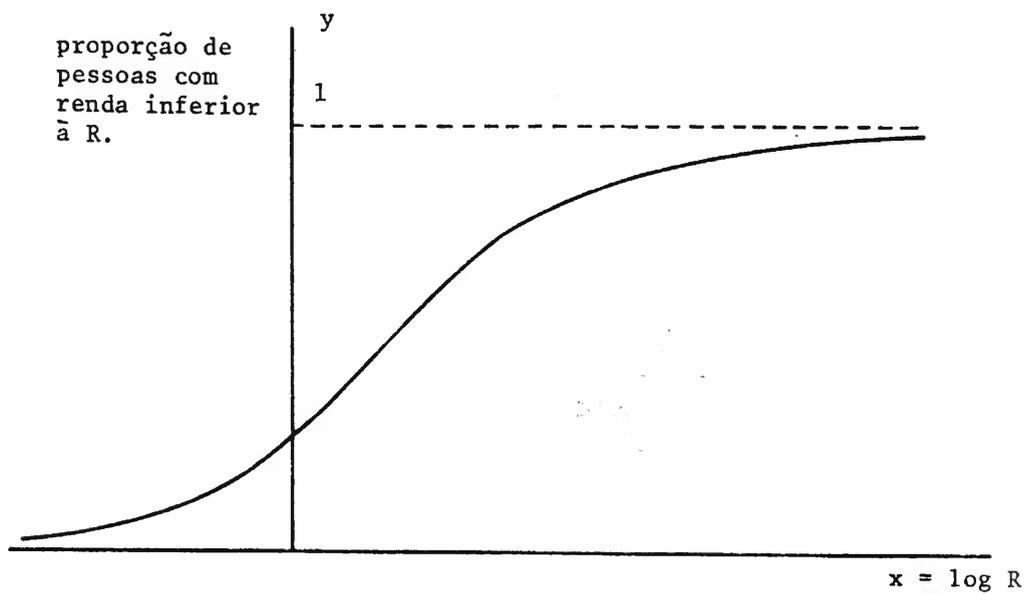
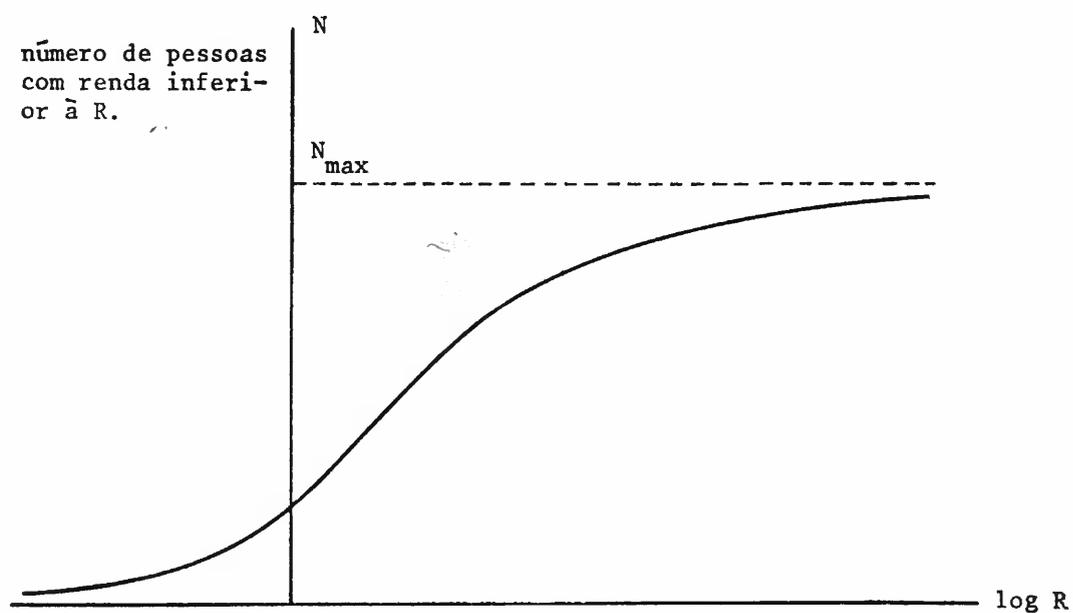


FIGURA 3.3. - Distribuição cumulativa de uma população pelas diversas faixas de rendas logaritmadas.



Podemos considerar os 4 (quatro) casos representados na Figura 3.5.. O caso 1 representa uma população pobre com uma má distribuição de renda. O caso 2 representa uma população pobre com uma boa distribuição de renda. O caso 3 representa uma população rica com uma má distribuição de renda. Finalmente o caso 4 representa uma população rica com uma boa distribuição de renda.

Até este momento estamos considerando a situação ideal em que os dados de renda se ajustam perfeitamente à curva logística. Na prática, entretanto, a situação é distinta, e o melhor que podemos esperar é que os dados se situem, próximos à curva logística. Assim se possuíssemos  $n$  pares de observações  $(x_i, y_i)$  poderíamos supor que, ao invés de Equação "ideal" (3.1) tivéssemos a seguinte situação:

$$y_i = \frac{1}{1 + \alpha e^{-\beta x_i - c_i}} \quad (3.2)$$

onde o índice  $i$  representa a  $i$ -ésima observação, e  $c_i$  é o resíduo responsável pela diferença entre os valores observados e a curva logística.

A estimação dos parâmetros da logística torna-se complicada se não procedermos inicialmente a uma linearização. Essa linearização pode ser obtida tomando-se o logaritmo da variável transformada  $y_i/(1 - y_i)$ . Assim obtemos

$$Y_i = \log\left(\frac{y_i}{1-y_i}\right) = -\log\alpha + \beta x_i + \epsilon_i \quad (3.3)$$

A Equação (3.3) é uma regressão linear simples, cujos parâmetros  $(\alpha$  e  $\beta)$  podem ser estimados através do método dos mínimos quadrados lineares ordinários.

Considerando-se a regressão linear da Equação (3.3) poderíamos ter diversas situações (Veja Figura 3.6.). O caso 1 ( $\beta \rightarrow 0$ ) representa uma população com perfeita desigualdade de distribuição de renda. O caso 2 ( $\beta$  pequeno) representa uma população com uma má distribuição de renda. O caso 3 ( $\beta$  grande) representa uma melhor distribuição de renda (a faixa total de ren-

das é bem menor do que no caso 2). Finalmente o caso 4, ( $\beta \rightarrow \infty$ ) representa uma população com perfeita igualdade de distribuição de renda (todas as pessoas teriam a mesma renda  $R_0$ ). Assim,  $\beta$  é o índice de concentração de renda que propomos neste trabalho. É um índice que varia entre zero e infinito ( $0 < \beta < \infty$ ). Quanto maior seu valor melhor a distribuição de renda da população considerada.

Neste ponto gostaríamos de chamar a atenção para a relação entre o índice  $\beta$  aqui proposto e um índice bastante conhecido como medida da distribuição de renda — a **variância dos logs**. Antes de mais nada observe que a variável renda é considerada, em nosso trabalho, logaritmada e que a variância da logística é dada por<sup>(16)</sup>

$$\sigma^2 = \frac{\pi^2}{3\beta^2} \quad (3.4)$$

Portanto, nosso índice  $\beta$  está intimamente ligado à variância dos logs<sup>(17)</sup>, sendo proporcional ao inverso do desvio padrão dos logs.

Pode-se também mostrar que a relação entre o índice  $\beta$  e o índice de Gini ( $G$ ) é dada por<sup>(18)</sup>

$$\beta = \frac{1}{G} \quad (3.5)$$

Outro ponto que deve ser mencionado é que tanto o índice de Pareto quanto o índice  $\beta$  são obtidos pela inclinação da curva de distribuição de renda “linearizada”. Desta forma, os dois índices são aproximadamente iguais. A diferença entre eles de-

(16) Veja Apêndice 1..

(17) Observe que a equação (3.4) não fornece exatamente a variância dos logs mas sim a variância da logística que se ajuste a nossos dados. A grande diferença entre os dois índices é a maneira de calculá-los. Para os dados de IR seria muito difícil calcular a variância dos logs devido às poucas faixas de renda existentes, pois grande parte da PEA está incluída numa única faixa.

(18) Dagum, op. cit., provou esta relação baseado na hipótese de que a distribuição de renda é log-logística.

FIGURA 3.5. - Quatro distribuições de renda possíveis segundo os Parâmetros  $\alpha$  e  $\beta$ .

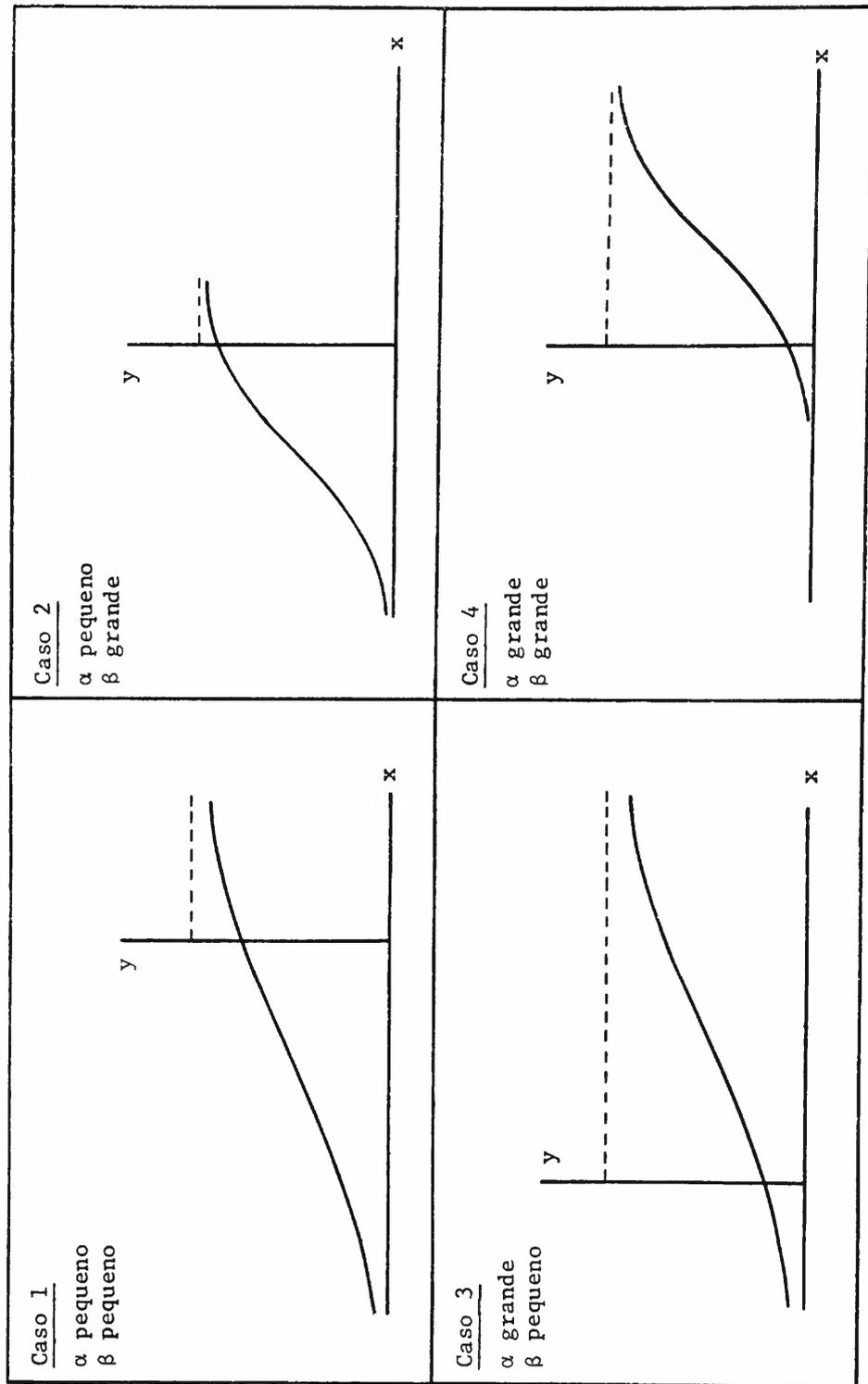
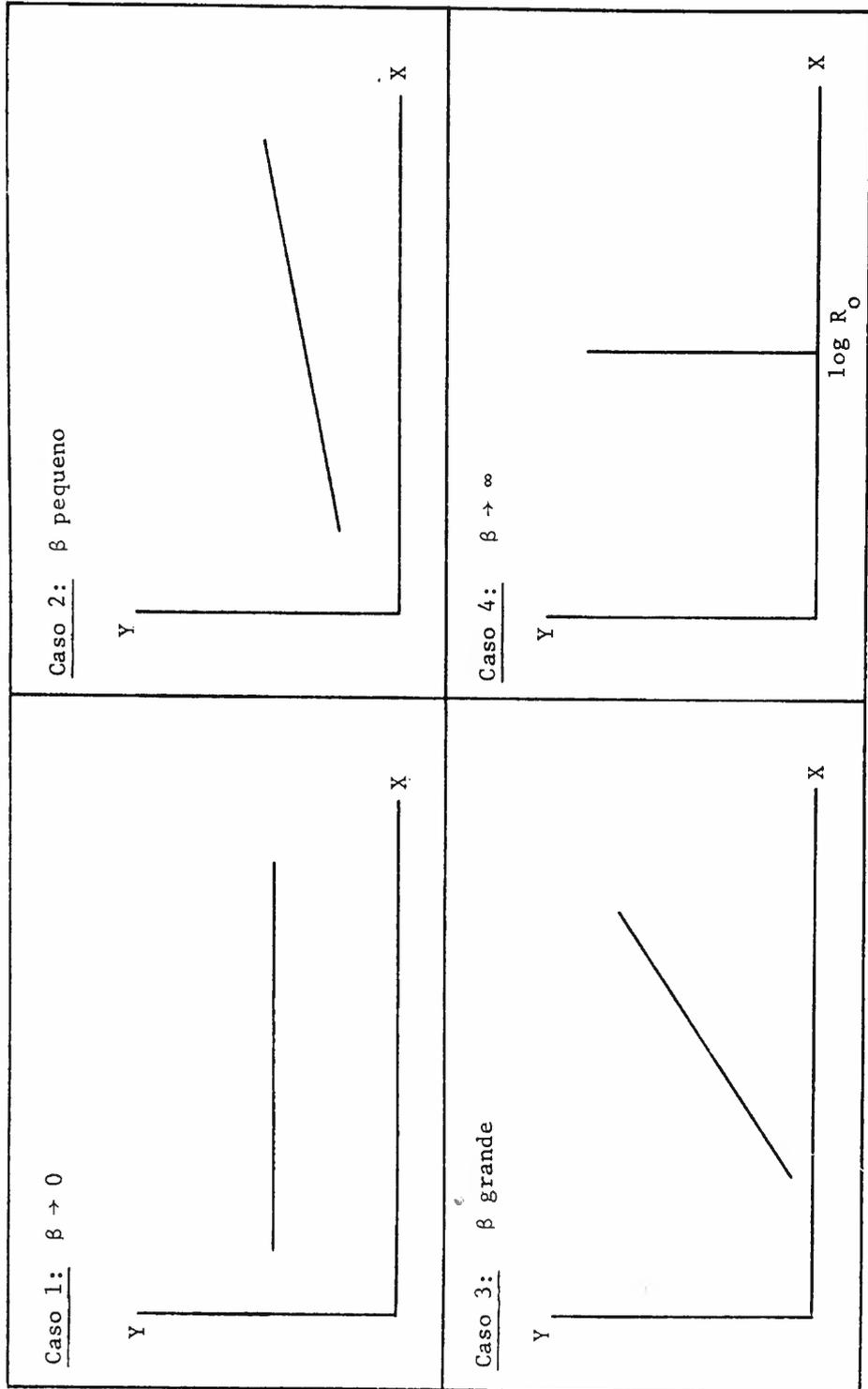


FIGURA 3.6. - Quatro casos de distribuição de renda segundo o parâmetro  $\beta$ .



ve-se ao fato de que as transformações efetuadas no eixo vertical para se obter a linearização nos dois casos são distintas.

Para finalizar a apresentação do índice, gostaríamos de considerar duas distribuições logísticas com médias idênticas e variâncias distintas, i.e.:

$$F_1(x) = \frac{1}{1 + \alpha_1 e^{-\beta_1 x}} \quad (3.6)$$

e

$$F_2(x) = \frac{1}{1 + \alpha_2 e^{-\beta_2 x}} \quad (3.7)$$

onde:

$$\mu = \frac{1}{\beta_1} \log \alpha_1 = \frac{1}{\beta_2} \log \alpha_2 \quad (3.8)$$

e

$$\beta_1 > \beta_2 \quad (3.9)$$

Inicialmente poderíamos considerar as duas possibilidades representadas pelas Figuras 3.7. e 3.8. para as duas distribuições. Entretanto, como veremos adiante, a possibilidade representada na Figura 3.8. não poderá ocorrer.

Da Equação (3.8) concluímos que

$$\alpha_1 = e^{\beta_1 \mu} \quad (3.10)$$

$$\alpha_2 = e^{\beta_2 \mu} \quad (3.11)$$

FIGURA 3.7. - Distribuições logísticas  $F_1(x)$  e  $F_2(x)$  com mesma média  $\mu$  e variância distintas (CASO I).

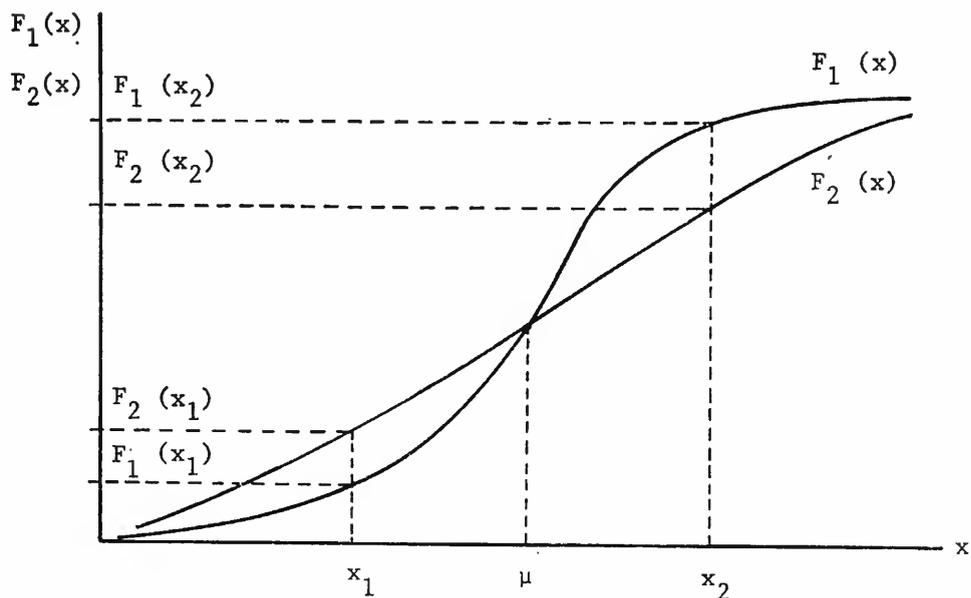
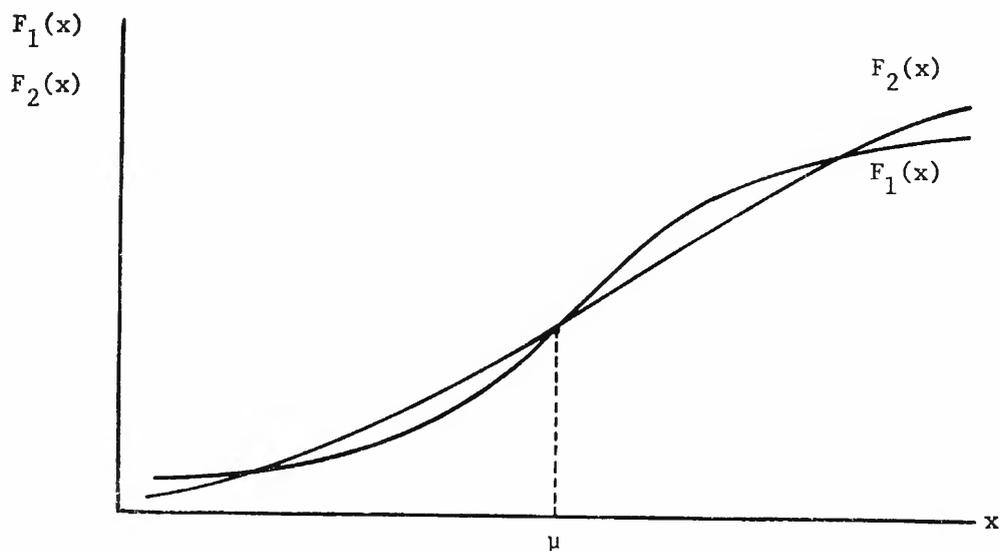


FIGURA 3.8. - Distribuições logísticas  $F_1(x)$  e  $F_2(x)$  com mesma média  $\mu$  e variância distintas (CASO II).



Substituindo nas Equações (3.6) e (3.7) obtemos

$$F_1(x) = \frac{1}{1 + e^{\beta_1(\mu-x)}} \quad (3.12)$$

$$F_2(x) = \frac{1}{1 + e^{\beta_2(\mu-x)}} \quad (3.13)$$

Portanto, para que  $F_1(x) = F_2(x)$  será necessário

que

$$e^{\beta_1(\mu-x)} = e^{\beta_2(\mu-x)} \quad (3.14)$$

cuja solução é dada por  $x = \mu$  em virtude da Equação (3.9). Assim, apenas a possibilidade representada pela Figura 3.7. poderá ocorrer.

O resultado acima é bastante importante pois ele implica em que, se considerarmos uma mesma faixa  $(x_2-x_1)$  de renda em torno da média, para as duas distribuições logísticas  $F_1(x)$  e  $F_2(x)$ , tal que  $\beta_1 > \beta_2$ , a primeira abrangerá uma maior porcentagem da população do que a segunda (Veja Figura 3.7., onde  $F_1(x_2) - F_1(x_1) > F_2(x_2) - F_2(x_1)$ ). Portanto, dentro deste critério,  $\beta_1 > \beta_2$  significa que  $F_1(x)$  fornece uma melhor distribuição de renda do que  $F_2(x)$ . Observe que se a situação descrita na Figura 3.8. ocorresse, a afirmativa acima não seria válida em geral<sup>(19)</sup>.

Discutimos acima apenas o caso em que as médias das duas distribuições  $F_1(x)$  e  $F_2(x)$  são iguais. O caso em que as médias são distintas pode ser reduzido a médias iguais deslocando-se umas das curvas sobre o eixo horizontal. Como estamos trabalhando com variáveis logaritmadas, isto seria equivalente a uma mudança de unidades, o que, obviamente, não muda em nada o estudo da distribuição de renda.

---

(19) No caso da Figura 3.8. a afirmativa seria válida apenas para faixas de rendimentos que não incorporassem os 2 pontos de interseção à direita e à esquerda da média.

---

#### 4. APLICAÇÃO I: COMPARAÇÃO DA DISTRIBUIÇÃO DE RENDA ENTRE 1960 E 1970

---

Nesta primeira aplicação utilizaremos nosso índice para verificar o processo de concentração de renda ocorrido no Brasil entre 1960 e 1970. Nosso intuito aqui é apenas testar nosso índice, pois este assunto já foi amplamente discutido e a verificação de tal processo de concentração já foi feita utilizando outros índices.

Apresentamos na Tabela 4.1. os resultados encontrados com o índice para os anos de 1960 e 1970<sup>(20)</sup>.

TABELA 4.1.  
RESULTADO DAS REGRESSÕES PARA OS ANOS DE  
1960 E 1970

Ano	$\log(\hat{\alpha})$ (a)	$\hat{\beta}$ (b)	$S_{\hat{\beta}}$ (c)	$R^2$ (d)	$S^2$ (e)
1960	2.725	2.033	.032	.999	.00737
1970	8.220	1.860	.033	.998	.00753

(a)  $\log \hat{\alpha}$  é o estimador de mínimos quadrados de  $\log a$ .

(b)  $\hat{\beta}$  é o estimador de mínimos quadrados de  $\beta$ .

(c)  $S_{\hat{\beta}}$  é o estimador para o desvio padrão de  $\hat{\beta}$ .

(d)  $R^2$  é o coeficiente de determinação.

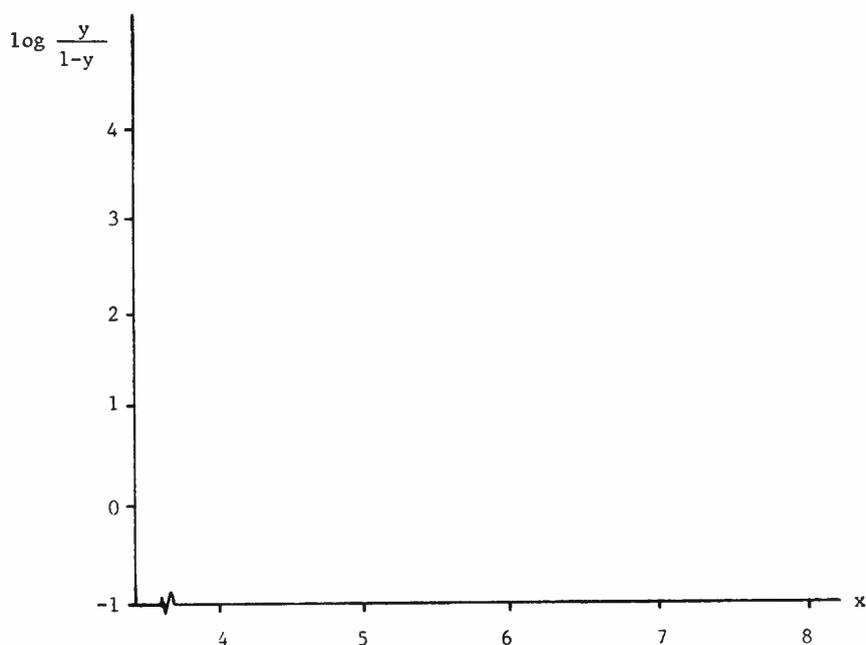
(e)  $S^2$  é o estimador para a variância dos resíduos.

---

(20) Foram utilizados dados dos Censos de 1960 e 1970 considerando apenas a PEA com rendimentos. Para o ano de 1960 utilizou-se as estimativas da Tabela 3.1 de Langoni, op. cit.

Observe que o ajuste de nossas regressões aos dados foi excelente ( $R^2 = .999$  e  $R^2 = .998$ ). Desta forma os coeficientes  $\beta$  foram estimados com bastante confiabilidade ( $S_{\hat{\beta}}$  muito pequeno comparativamente a  $\hat{\beta}$ ). (Veja Figura 4.1.).

FIGURA 4.1. - Linearização dos dados do Censo de 1970  
(x é o logaritmo da renda  $\underline{R}$  e y é a proporção de pessoas com renda inferior a  $\underline{R}$ )



Nossos resultados confirmam que houve uma piora na distribuição de renda entre 1960 e 1970 haja visto a diminuição sensível verificada no índice  $\beta$  durante aquele período. Como já dissemos anteriormente este resultado já é amplamente conhecido e já foi verificado com outros índices de concentração. Aqui, entretanto, podemos utilizar os testes de hipóteses desenvolvidos no Apêndice 2., para verificar com que confiabilidade poderíamos afirmar que houve uma piora real na situação da distribuição de renda entre 1960 e 1970. Assim teríamos:

$$H_0: \beta_{1960} = \beta_{1970}$$

$$H_A: \beta_{1960} > \beta_{1970}$$

Em nosso caso encontramos  $t = 8.184$  que é maior que o valor tabelado para a estatística  $t$  com 10 graus de liberdade no

nível de significância de 1%, i.e.,  $t_{10,.01} = 2.764$ . Podemos portanto afirmar que a piora indicada pelo nosso índice foi real e não se deveu a problemas de estimação<sup>(21)</sup>.

Apresentamos a seguir a Tabela 4.2. que mostra, além de nosso índice, outros índices de concentração para os anos de 1960 e 1970<sup>(22)</sup>. Esta tabela confirma a coerência de nosso índice quando comparado com outros índices — todos mostram claramente o processo de concentração de renda verificado em nosso país no período 1960-1970.

Como foi visto na seção anterior os índices de Gini e a variância dos logs podem ser calculados a partir do índice  $\beta$  utilizando-se as Equações (3.4) e (3.5). Através destas equações obtivemos para o índice de Gini os valores .492 em 1960 e .602 em 1970, e para a variância dos logs .796 em 1960 e 1.194 em 1970. Comparando-se estes resultados com a Tabela 4.2. verificamos uma razoável proximidade entre os índices, utilizando-se as duas formas de cálculo (diferenças relativas inferiores a 6% para o índice de Gini e 22% para a variância dos logs).

TABELA 4.2.

VARIAÇÃO DOS ÍNDICES DE CONCENTRAÇÃO  
ENTRE 1960 E 1970

ÍNDICES DE CONCENTRAÇÃO				
ANO	Gini	Theil	Variância dos logs	$\beta^{(a)}$
1960	.500	.470	.897	2.033
1970	.568	.644	.976	1.660

(a) Diferentemente dos índices de Gini, Theil e Variância dos Logs, o índice  $\beta$  diminui à medida em que a concentração aumenta.

(21) Foi utilizada a estatística  $t$  da Equação (A.21) do Apêndice 2. em virtude de não haver diferença significativa entre as variâncias dos resíduos quando testadas pela Equação (A.26) ( $1.02 < 5.05 = F_{(5,5),.05}$ ).

(22) Os índices de Gini, Theil e Variância dos Logs foram retirados da Tabela 3.6 de Langoni, op. cit., e excluem o pessoal com renda zero.

---

## 5. APLICAÇÃO II: A EVOLUÇÃO DA DISTRIBUIÇÃO DE RENDA NO BRASIL NO PERÍODO 1968-1975 (IMPOSTO DE RENDA)

---

O índice de concentração de renda foi também utilizada para a verificação da evolução da distribuição de renda no Brasil no período 1968-1975. Os dados foram obtidos de diversas edições do Anuário Estatístico Fiscal.

Para cada ano do período considerado (1968-1975) foram obtidas 3 (três) regressões lineares. A razão que nos levou a trabalhar com mais de uma regressão para cada ano deve-se ao fato de que os dados do Anuário Estatístico Fiscal levam em consideração apenas a população declarante de IR. Se desejarmos um índice que reflita a concentração de renda de toda a população devemos considerar toda PEA e não apenas aquela que declara IR. A população declarante de imposto de renda variou de um mínimo de cerca de 4.4 milhões de pessoas em 1968 até um máximo de cerca de 16 milhões de pessoas em 1975. Enquanto isso, sabemos pelo Censo de 1970, que naquele ano existiam cerca de 26 milhões de pessoas economicamente ativas com rendimentos no Brasil. Desta forma optamos por 3 (três) regressões para cada ano. A primeira leva em consideração apenas a população declarante do IR. As segunda e terceira regressões consideram PEA's de 25 e 30 milhões respectivamente, onde a população não declarante de IR está incluída na faixa de rendimentos mais baixos<sup>(23)</sup>. Como veremos posteriormente, as três regressões apresentam resultados semelhantes, embora as duas últimas, como era de se esperar<sup>(24)</sup>, indiquem um grau de concentração maior (valores de  $\beta$  menores que na primeira regressão). Nos três casos considerados obteve-se um ajuste excelente para as regressões ( $R^2$  consistentemente superior a .97). (Veja Tabela 5.1.).

---

(23) É claro que poderíamos utilizar apenas uma regressão, considerando a PEA de 1970 e utilizando a taxa de crescimento da PEA entre 1960 e 1970 para estimar a PEA no período 1968-1975. Entretanto, acreditamos que a utilização das 3 regressões é mais ilustrativa.

(24) Quanto mais pessoas forem incluídas na menor faixa de rendimentos maior será a desigualdade. Desta forma a regressão que só considera a população declarante de IR deverá fornecer um grau de concentração menor que a que utiliza uma PEA de 30 milhões de pessoas.

TABELA 5.1.

RESULTADO DAS REGRESSÕES PARA O PERÍODO 1968-1975 CONSIDERANDO-SE POPULAÇÕES CONSTITUÍDAS PELOS DECLARANTES DE IMPOSTO DE RENDA, 25 MILHÕES E 30 MILHÕES DE PESSOAS ECONOMICAMENTE ATIVAS

Ano	POPULAÇÃO ECONOMICAMENTE ATIVA (PEA)									
	DECLARANTES DO IR					30.000.000				
	$\log \hat{\alpha}^{(a)}$	$\hat{\beta}^{(b)}$	$S_{\hat{\beta}}^{(c)}$	$R^{2(d)}$	$S^2(e)$	$\log \hat{\alpha}$	$\hat{\beta}$	$S_{\hat{\beta}}$	$R^2$	$S^2$
1968	3.353	2.227	.045	.993	.1261	1.188	2.153	.048	.992	.1414
1969	4.044	2.230	.037	.997	.0576	2.248	2.155	.060	.991	1515
1970	4.023	2.148	.055	.992	.0948	2.384	2.044	.078	.983	.1879
1971	4.109	2.071	.058	.990	.1068	2.720	1.987	.076	.983	.1818
1972	4.226	1.980	.048	.993	.0700	2.973	1.906	.064	.987	1218
1973	4.779	2.009	.063	.989	.0992	3.692	1.949	.073	.983	1429
1974	5.610	2.057	.074	.985	.1457	4.401	1.987	.087	.977	.2059
1975	6.269	2.100	.066	.985	.1025	5.689	2.077	.070	.983	.1158

(c)  $S_{\hat{\beta}}$  é o estimador para o desvio padrão de  $\hat{\beta}$ .

(a)  $\log \hat{\alpha}$  é o estimador de mínimos quadrados de  $\log \alpha$ .

(d)  $R^2$  é o coeficiente de determinação.

(b)  $\hat{\beta}$  é o estimador de mínimos quadrados de  $\beta$ .

(e)  $S^2$  é o estimador para a variância dos resíduos.

NOTAS: Para os anos 1968 a 1973 foram utilizados dados publicados no Anuário Econômico Fiscal (1970-1975). Para os anos 1974 e 1975 foram utilizadas tabulações ainda não publicadas na época da pesquisa. Para 1975 os dados referentes aos declarantes isentos não estavam completos, razão pela qual os resultados correspondentes à regressão de declarantes para aquele ano estão um pouco distorcidos.

Os seguintes resultados podem ser tirados da Tabela 5.1.:

Nos 3 (três) casos observa-se uma diminuição acentuada do índice estimado  $\hat{\beta}$  a partir de 1968/69, atingindo seu ponto mínimo em 1972, aumentando lentamente a partir de 1973, atingindo em 1975 um valor próximo ao de 1970, porém ainda inferior ao valor em 1968/69 (Veja Figura 5.1.).

À medida em que consideramos uma população maior, o índice  $\hat{\beta}$  tende a diminuir (Veja Figura 5.1.).

O índice  $\hat{\beta}$  nos parece ter sido estimado com bastante confiabilidade haja visto seus pequenos desvios padrão  $S_{\hat{\beta}}$  proporcionalmente ao valor  $\hat{\beta}^{(25)}$ .

Os ajustes das regressões aos dados nos parecem excelentes. Os valores de  $R^2$  variam entre .976 e .997

O coeficiente  $\hat{\alpha}$  cresce à medida em que o tempo passa, possuindo valores maiores quando a população economicamente ativa é considerada menor.

Passemos agora a discutir os pontos abordados acima :

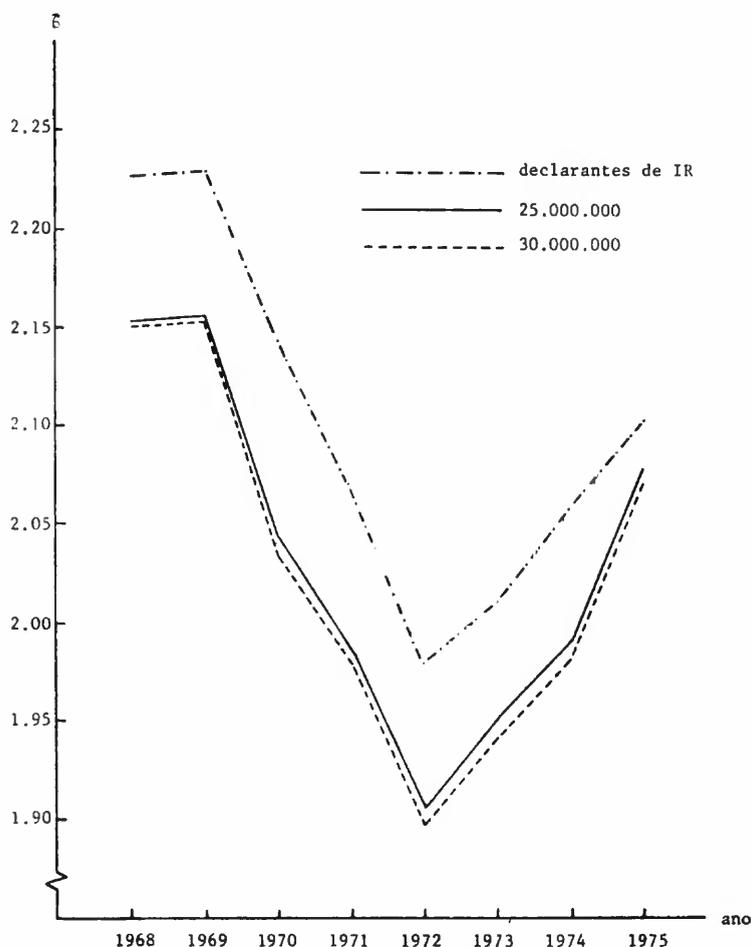
A diminuição de  $\hat{\beta}$  entre 1968 e 1972 seguida de um lento aumento a partir de 1973, nos indicaria uma piora da distribuição de renda até 1972 seguida de uma ligeira recuperação a partir de 1973. Nota-se, entretanto, que em 1975 a distribuição de renda ainda estaria em níveis inferiores aos de 1968/69, em níveis comparáveis aos de 1970<sup>(26)</sup>.

---

(25)  $\hat{\beta}/s_{\hat{\beta}}$  mede o grau de confiabilidade do estimador  $\hat{\beta}$ ; caso  $\beta = 0$  a relação  $\hat{\beta}/s_{\hat{\beta}}$  possui distribuição t e pode ser utilizada como teste de hipóteses (Veja Kmenta (1971)).

(26) Este ponto será discutido adiante através de testes de hipóteses estatísticas.

FIGURA 5.1. Evolução do índice de concentração estimado  $\beta$  durante o período 1968-1975 utilizando-se a população de declarantes do IR, 25 milhões e 30 milhões de pessoas economicamente ativas.



Sendo  $\beta$  um índice de concentração de renda, que cresce à medida em que a distribuição de renda melhora, verificamos que realmente ele deve diminuir quando consideramos uma PEA maior. Isto pode ser melhor compreendido, observando que a população correspondente à diferença entre os 25 ou 30 milhões da PEA e os declarantes de IR foram considerados como recipientes de rendimentos na faixa mais baixa de rendimentos. Assim, verificamos que o índice de concentração de renda é bastante coerente.

O crescimento do parâmetro  $\hat{\alpha}$  ao longo dos anos não deve ser tomado como uma melhora do nível geral de renda da popu-

lação como poderia parecer à primeira vista. Em nossas regressões trabalhamos com cruzeiros nominais sem preocupação com a inflação. Assim em termos nominais a população recebe mais cruzeiros ao longo dos anos, o que faz com que a curva se desloque para a direita sobre o eixo horizontal, como explicado na seção anterior. O fato de  $\hat{\alpha}$  diminuir quando a população considerada aumenta, pode ser explicado pelas mesmas razões da diminuição de  $\beta$  à medida em que a população considerada aumenta — as novas pessoas entram na faixa mais baixa, o que significa uma população mais pobre.

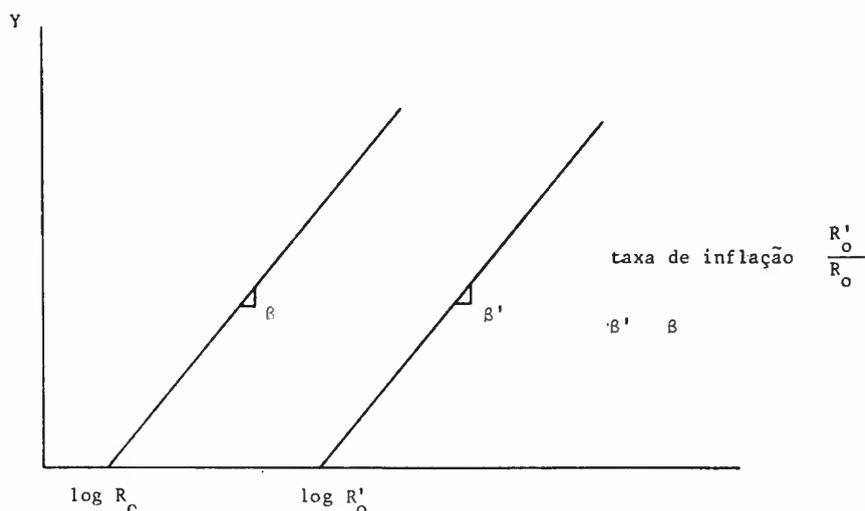
Um critério mais atento poderia levantar uma objeção ao fato de não se ter corrigido os cruzeiros de cada ano para se eliminar o efeito inflacionário. Isto não poderia distorcer os resultados encontrados? A resposta é **negativa** visto que se a distribuição de renda se mantivesse a mesma, havendo apenas inflação, isto significaria que as rendas (nominais) seriam multiplicadas por uma constante (taxa de inflação), de forma que ao serem logaritmadas, elas apareceriam igualmente defasadas ao longo do eixo horizontal. Desta forma, a inclinação das retas de regressão se manteriam constantes ( $\beta$  constante) havendo apenas variação no parâmetro  $\alpha$  como já foi dito anteriormente (Veja Figura 5.2.).

Como foi dito anteriormente, a Tabela 5.1. e a Figura 5.1. nos sugerem uma piora na distribuição de renda entre 1968/69 e 1972 seguida de uma lenta recuperação a partir de 1973. Seriam estas pioras e melhoras reais ou apenas devido a efeitos de estimação de parâmetros? Teria havido alguma mudança na distribuição de renda? Estas perguntas podem ser formalizadas da seguinte forma:

- T1 — Houve uma piora real entre 1968/69 e 1972?
- T2 — Houve uma melhora real entre 1972 e 1975?
- T3 — Como se apresentava a distribuição de renda em 1975 em relação ao início da década (1970)?

Recentemente o Governo reconheceu publicamente que os índices oficiais do custo de vida em 1973 foram subestimados. Além disso há outras indicações de que 1973 foi o ano em que os salários reais dos trabalhadores estiveram em seu nível mais

FIGURA 5.2. Situação em que a taxa de inflação é dada por  $R'_c/R'_o$  com distribuição de renda mantida constante ( $\beta=\beta'$ ).



baixo nos últimos anos<sup>(27)</sup>. Não seria, portanto, uma incoerência de nosso índice encontrar uma melhora já a partir de 1973? Para responder tal pergunta podemos testar a hipótese abaixo:

T4 — Houve uma melhora real entre 1972 e 1973?

Estas 4 (quatro) perguntas poderiam ser respondidas utilizando-se os testes de hipóteses descritos no Apêndice 2.. Desta forma obtemos a Tabela 5.2. com um resumo das estatísticas t encontradas para os 4 (quatro) testes de hipóteses abaixo, utilizando as regressões para 30 milhões de pessoas economicamente ativas<sup>(28)</sup>.

(27) Segundo cálculos do DIEESE, trabalhando com 81 categorias de trabalhadores de todo o Brasil, os salários reais dos trabalhadores diminuíram entre 1964 e 1973. Em 1974 houve uma pequena recuperação para as categorias cujos reajustes salariais eram feitos no segundo semestre (Veja DIEESE (1975)). Suplicy (1977) chega a conclusões semelhantes.

(28) Como os resultados encontrados para as três regressões foram semelhantes escolhemos a regressão com PEA de 30 milhões para os resultados de 1969 que são praticamente idênticos aos de 1968.

TABELA 5.2.

RESULTADOS DOS TESTES DE HIPÓTESES RELATIVOS À EVOLUÇÃO DA DISTRIBUIÇÃO DE RENDA NO PERÍODO 1968-1975 UTILIZANDO AS REGRESSÕES PARA UMA PEA DE 30 MILHÕES DE PESSOAS.

TESTE	F	Graus de liberdade	Conclusão <sup>(a)</sup>	t	Graus de liberdade	Conclusão <sup>(b)</sup>
T1 (1969x1972)	1.22	12,12	As variâncias dos resíduos de 1969 e 1972 são iguais	2.804	24	A distribuição de renda em 1972 é pior que em 1969
T2 (1972x1975)	1.07	12,15	As variâncias dos resíduos de 1972 e 1975 são desiguais	1.780	27	A distribuição de renda em 1975 é melhor que em 1972
T3 (1970x1975)	1.62	12,15	As variâncias dos resíduos de 1970 e 1975 são iguais	.309	27	A distribuição de renda em 1975 é semelhante à de 1970
T4 (1972x1973)	1.17	12,12	As variâncias dos resíduos de 1972 e 1973 são iguais	.434	24	A distribuição de renda em 1973 é semelhante à de 1972

a) conclusões baseadas no teste **F** da Equação (A.26) do Apêndice 2. no nível de significância de 5%.

b) conclusões baseadas no teste **t** da Equação (A.21) do Apêndice 2. no nível de significância de 5%.

**Teste 1 (T1)**

$$H_0: \beta_{1969} = \beta_{1972}$$

$$H_A: \beta_{1969} > \beta_{1972}$$

**Teste 2 (T2)**

$$H_0: \beta_{1975} = \beta_{1972}$$

$$H_A: \beta_{1975} > \beta_{1972}$$

**Teste 3 (T3)**

$$H_0: \beta_{1970} = \beta_{1975}$$

$$H_A: \beta_{1970} \neq \beta_{1975}$$

**Teste 4 (T4)**

$$H_0: \beta_{1973} = \beta_{1972}$$

$$H_A: \beta_{1973} > \beta_{1972}$$

Baseados na Tabela 5.2. podemos concluir que:

- Existe **bastante** evidência estatística de que houve uma piora na distribuição de renda entre 1968/69 e 1972.
- Existe **alguma** evidência estatística de que houve uma melhora na distribuição de renda entre 1972 e 1975.
- Existe **bastante** evidência estatística de que as distribuições de renda em 1970 e 1975 se apresentavam basicamente iguais.
- **Não** existe evidência estatística de que a distribuição de renda em 1973 era melhor que em 1972.

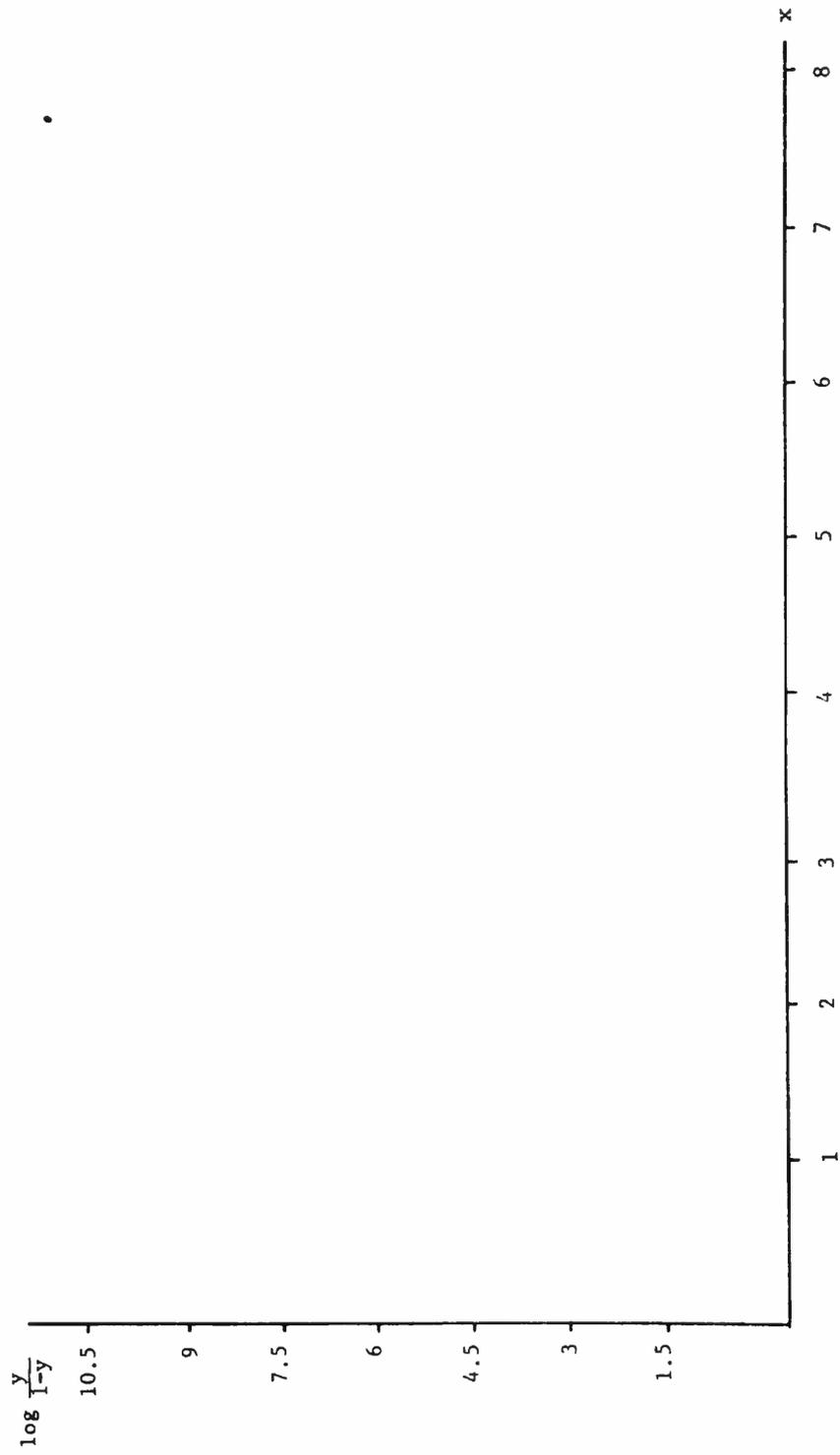
As conclusões acima estão baseadas na facilidade com que a hipótese nula foi rejeitada no teste T1, aceita nos testes T3 e T4, enquanto que no teste T2 ela foi rejeitada com dificuldade<sup>(29)</sup>.

Um ponto que gostaríamos de chamar atenção é que o fato de fazermos um estudo da evolução da distribuição de renda mantendo fixa a população (como nos casos em que consideramos 25 e 30 milhões de pessoas respectivamente) em nada prejudica nossas conclusões. Na verdade, o que ocorre é que mantendo fixa a população tendemos a subestimar a real concentração no final do período em relação ao início do período. Isto pode ser facilmente entendido observando-se que se considerássemos uma PEA crescente, esta diferença seria obrigatoriamente incorpora-

---

(29) Para se entender melhor estas conclusões deve-se observar que  $t_{24,05} = 1.711$ ,  $t_{24,01} = 2.492$ ,  $t_{27,05} = 1.703$  e  $t_{27,01} = 2.473$ . Desta forma as hipóteses nulas em T3 e T4 são aceitas com bastante facilidade, a hipótese nula em T1 é rejeitada a 5% e 1%, e a hipótese nula em T2 é rejeitada a 5% e aceita a 1%.

FIGURA 5.3. - Linearização dos dados do IR em 1975 utilizando PEA de 30 milhões de pessoas ( $x$  é o logaritmo R e  $y$  é a proporção de pessoas com renda inferior a R).



da à faixa de renda mais baixa, piorando a distribuição de renda (baixando  $\beta$ ). Desta forma, o fato de mantermos fixa nossa população ao longo do período é um fator para reforçar nossos argumentos em relação à tendência a uma piora na distribuição de renda no período 1968-1972 (Teste T1) ao mesmo tempo em que serve para dificultar a conclusão de que houve uma melhora na distribuição de renda no período 1972-1975. (Teste T2).

Outro ponto que gostaríamos de mencionar é que embora nossas regressões tenham fornecido ajustes excelentes, pode-se notar uma tendência a uma ligeira convexidade, ou seja, a linearização não foi perfeita<sup>(30)</sup>. (Veja Figura 5.3.).

Gostaríamos neste ponto de ressaltar o fato de que, devido a estarmos utilizando dados do IR não nos foi possível separar em diversas faixas de rendimentos uma grande massa de trabalhadores<sup>(31)</sup>. As consequências deste fato não foram graves em nossa aplicação haja visto a confiabilidade com que estimamos nossos parâmetros  $\beta$ . Caso tivéssemos uma maior diferenciação dos rendimentos das classes menos favorecidas da população brasileira teríamos mais pontos na nossa regressão, o que implicaria em melhores estimadores<sup>(32)</sup>.

Outra observação que gostaríamos de fazer é que, em virtude de não podermos separar grande parte da massa de trabalhadores, nossos índices foram estimados utilizando-se apenas os pontos superiores e à direita da reta de regressão. (Veja Figura 5.4.).

Entretanto, sendo aceita a hipótese de log-normalidade<sup>(33)</sup>, os pontos desconhecidos devem estar situados no prolongamen-

---

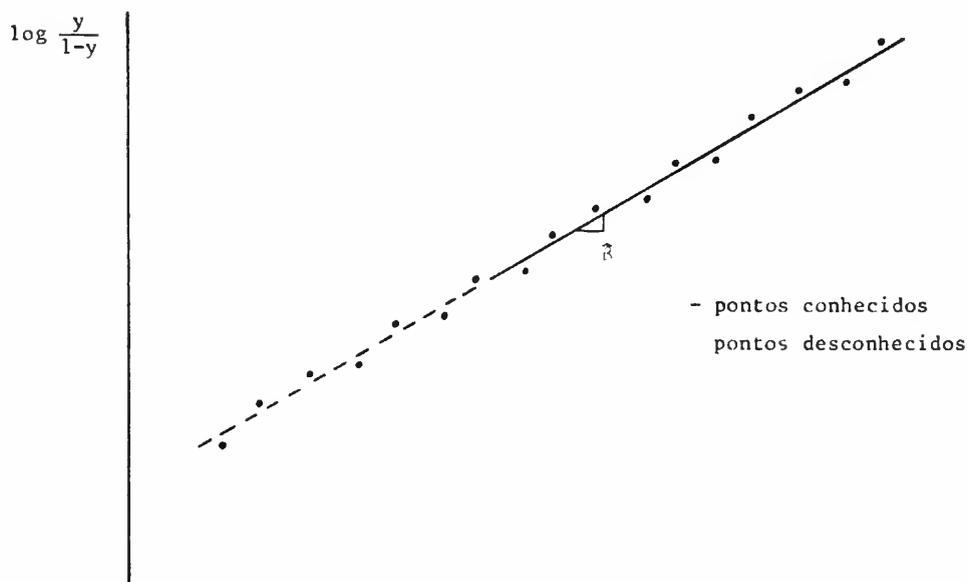
(30) Isto pode ser verificado com mais precisão através dos baixos valores das estatísticas de Durbin-Watson encontradas para as regressões (na faixa .2 — .5), significando existência de não-linearidade. (Veja Kmenta, op. cit.).

(31) Por estarmos utilizando dados do IR, a faixa mais baixa é bastante ampla. Em 1974 correspondia a todos aqueles que recebiam menos de 13900 cruzeiros anualmente. Desta forma grande parte da população trabalhadora está dentro desta faixa e não pode ser diferenciada. Este problema será contornado na próxima seção com dados da Lei dos 2/3.

(32) Veja Kmenta, op. cit.

(33) Ou sendo aceita a hipótese log-logística.

FIGURA 5.4. - Reta de Regressão ajustada utilizando apenas os pontos superiores e a direita ( $x$  é o logaritmo da renda  $R$  e  $y$  é a proporção de pessoas com renda inferior a  $R$ ).



to da reta de regressão. Desta forma não haveria distorção no cálculo de  $\beta$  ao utilizarmos apenas os pontos conhecidos superiores e à direita<sup>(34)</sup>.

Um ponto que gostaríamos de chamar a atenção é para o fato de que se compararmos os índices  $\beta$  para o ano de 1970 utilizando dados do Censo e do IR (Tabelas 4.1. e 5.1.) verificamos uma certa discrepância entre os valores encontrados, i.e., 1.660 (para o Censo de 1970) e 1.898 (para os dados do IR com uma PEA de 30 milhões de pessoas). Este resultado era de se esperar em virtude da ligeira tendência à convexidade mencionada acima e do fato de estarmos estimando nosso índice do IR com os pontos das faixas superiores de renda. Além disso a confiabilidade dos dois tipos de dados é bastante distinta, o que pode aumentar ainda mais a diferença entre os índices obtidos. Entretanto, acreditamos que se estivermos subestimando a desigualdade de renda com os dados do IR, a estaremos subestimando igualmente durante todo o período, e no nosso caso o que im-

(34) Na realidade deve estar havendo uma pequena distorção no cálculo de  $\beta$  em função da ligeira convexidade observada anteriormente. Por estarmos utilizando a cauda da distribuição, devemos estar superestimando os verdadeiros valores de  $\beta$ .

porta não é o valor absoluto do índice  $\beta$ , mas sim sua evolução que mostra um comportamento claro e inconfundível.

Para finalizar esta seção gostaríamos de comparar os resultados encontrados com o índice desenvolvido neste trabalho e um índice tradicional como, por exemplo, o índice de Gini. Os dados utilizados para o cálculo do índice de Gini também foram retirados dos Anuários Econômicos Fiscais, sendo que neste caso também há necessidade dos rendimentos brutos para cada faixa de rendimento. Foi feita a hipótese de que os rendimentos se distribuem uniformemente dentro de cada faixa<sup>(35)</sup> Os índices de Gini obtidos, assim como os nossos índices para o período 1968-1975, estão apresentados na Tabela 5.3..

TABELA 5.3.

ÍNDICE DE GINI E ÍNDICES  $\beta$  PARA O  
PERÍODO 1968-1975

ANO	GINI <sup>(a)</sup>	$\hat{\beta}$ (IR <sup>c</sup> )	$\hat{\beta}$ (25 mi- lhões)	$\hat{\beta}$ (30 mi- lhões)
1968	462	2.227	2.153	2.151
1969	457	2.230	2.155	2.152
1970	.542	2.148	2.044	2.037
1971	.612	2.071	1.987	1.980
1972	.626	1.980	1.906	1.898
1973	.709	2.009	1.949	1.941
1974	.709	2.057	1.987	1.979
1975	—	2.100	2.077	2.070

(a) O índice de Gini foi calculado para a população declarante do IR. No ano de 1975 não havia dados disponíveis de rendimentos para a faixa de rendimentos mais baixos.

(35) Esta é a hipótese mais simples, a qual não acarreta grandes diferenças no cálculo do índice de Gini comparativamente com outras hipóteses mais sofisticadas.

Antes de discutir os resultados comparativos dos dois índices acima gostaríamos de chamar a atenção para o fato de que o índice de Gini só foi calculado para a população declarante do IR. Até mesmo este cálculo foi dificultado, haja visto a grande parcela da população declarante de IR localizada na faixa de rendimentos mais baixos. Poderíamos tentar incluir o restante da PEA nesta faixa como foi visto para o cálculo do índice  $\beta$ , mas neste caso teríamos um grande problema: como estimar os rendimentos desta população com alguma confiabilidade? Desta forma optamos pela estimação de Gini apenas para a população declarante de IR.

Ao observarmos o comportamento do índice de Gini ao longo do período 1968-1974 notamos um comportamento bastante distinto do índice  $\beta^{(36)}$ , pois enquanto o índice  $\beta$  nos sugere uma piora até 1972/73 seguida de uma lenta recuperação, o índice de Gini nos sugere uma piora na distribuição de renda ao longo de todo o período considerado. Qual seria a razão para tal discrepância? A resposta a esta pergunta nos parece simples se recordarmos que a população declarante do IR sofreu grandes modificações ao longo do período, quase multiplicando seu número por 4 (quatro). Principalmente nos últimos anos, houve um grande número de pessoas que se tornaram contribuintes do IR, as quais na sua grande maioria localizam-se na faixa mais baixa de rendimentos. Para se ter uma idéia de como a composição desta primeira faixa se transformou ao longo do período, construímos a Tabela 5.4. que mostra o número de contribuintes e a renda média da população componente da faixa de rendimentos mais baixos durante o período 1968-1974.

Observe que ao longo do período houve um grande aumento no número de contribuintes na faixa de rendimentos inferiores e que a renda média destes contribuintes sofreu um decréscimo considerável se levarmos em consideração a inflação havida durante este período<sup>(37)</sup>. Este fato "per si" seria suficiente para acarretar um aumento no índice de Gini ao longo do período.

---

(36) Observe que o comportamento dos índices  $\hat{\beta}$  para as três regressões (declarantes de IR, PEA de 25 e 30 milhões respectivamente) é similar.

(37) Isto pode ser explicado pelo fato de que boa parte dos novos contribuintes devem ter passado a fazer declaração de IR com a finalidade de receber pequenos impostos retidos na fonte ou para possuir o CPF.

Com relação ao nosso índice  $\beta$ , apenas o aumento percentual do número de contribuintes na primeira faixa de rendimentos é capaz de distorcê-lo, e assim mesmo, apenas para o índice  $\beta$  que leva em consideração somente a população declarante de IR. Para corrigir este efeito é que utilizamos índices  $\beta$  considerando toda a PEA com rendimentos<sup>(38)</sup>. Portanto, os nossos índices  $\beta$  devem medir muito melhor a evolução da desigualdade de rendas ao longo do período que o índice de Gini quando aplicado aos dados de declarantes de IR.

TABELA 5.4.

NÚMERO DE CONTRIBUINTES DO IMPOSTO DE RENDA E RENDA MÉDIA PARA A FAIXA DE RENDIMENTOS INFERIORES NO PERÍODO 1968-1974

Ano	Número de contribuintes do IR na faixa inferior de rendimentos	% do total de contribuintes	Renda média <sup>(a)</sup> dos contribuintes na faixa inferior de rendimentos
1968	1 511 750	34.3	1 623
1969	2 523 469	40.1	2 390
1970	3 681 469	45.2	2 244
1971	4 857 196	50.5	1 680
1972	5 667 852	53.2	2 008
1973	6 992 066	59.6	1 099
1974	6 281 355	56.1	1 068

(a) A renda média é expressa em cruzeiros do ano considerado.

(38) Na realidade por considerarmos a PEA fixa, durante o período haverá uma ligeira tendência a subestimarmos a verdadeira desigualdade no final do período em relação ao início do período, pois o aumento da PEA existente durante o período deveria ser incorporado na faixa de rendimentos inferiores. Este efeito entretanto, é pequeno haja visto a proximidade dos índices  $\hat{\beta}$  para PEA's de 25 e 30 milhões respectivamente.

---

## 6. APLICAÇÃO III: A EVOLUÇÃO DA DISTRIBUIÇÃO DE RENDA NO BRASIL NO PERÍODO 1968-1974 (LEI DOS 2/3)

---

Nesta seção repetimos a aplicação da seção anterior utilizando dados da Lei dos 2/3. Estes dados foram retirados de diversos números do Boletim Técnico do Ministério do Trabalho<sup>(39)</sup>. Nossa intenção aqui foi testar os resultados encontrados anteriormente, utilizando uma outra amostra da população brasileira.

A Tabela 6.1. fornece os resultados das regressões obtidas com os dados da Lei dos 2/3 para o período 1968-1974.

Observe que os resultados encontrados na Tabela 6.1. são bastante similares aos resultados da Tabela 5.1.,  $\hat{\alpha}$  tende a crescer ao longo do tempo,  $\hat{\beta}$  se comporta de forma semelhante nas duas aplicações,  $S_{\hat{\beta}}$  é bastante pequeno comparativamente a  $\hat{\beta}$  e  $R^2$  é bastante alto<sup>(40)</sup>.

É bastante significativo o fato de que, não apenas  $\hat{\beta}$  se comporta da mesma forma para os dados do IR como para a Lei dos 2/3, como também os valores encontrados são bastante próximos. Na realidade os valores de  $\hat{\beta}$  encontrados com os dados da Lei dos 2/3 são ligeiramente inferiores aos encontrados com os dados do IR. Este resultado é razoável pois, como já foi dito anteriormente, as estimativas de  $\beta$  com os dados do IR devem estar superestimando ligeiramente o índice  $\beta$ .

Deve-se notar também que a linearização dos dados foi muito melhor com a utilização dos dados da Lei dos 2/3 que com os dados do IR<sup>(41)</sup>. (Veja Figura 6.1.).

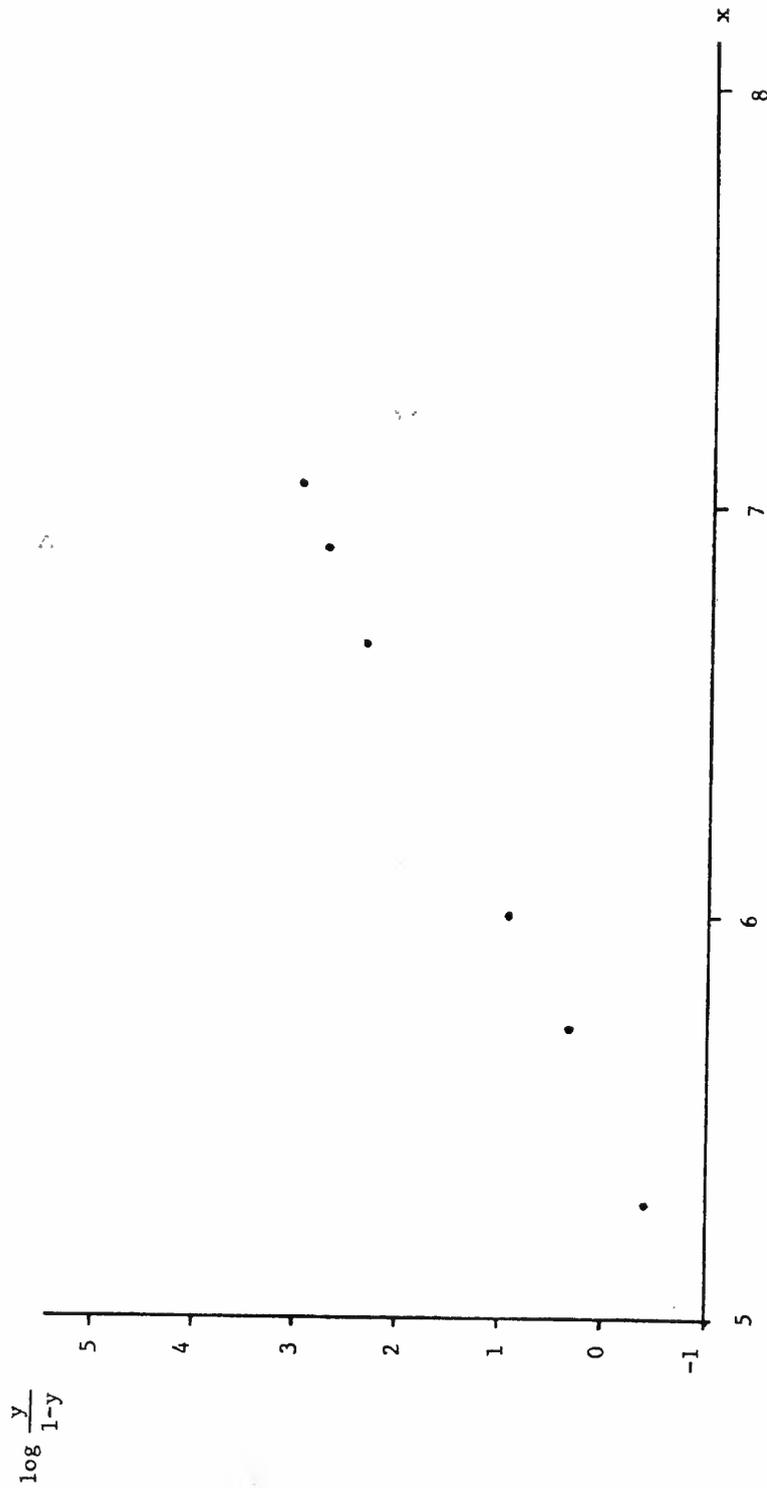
---

(39) Pela Lei dos 2/3 são cadastrados os empregados regidos pela CLT. Durante o período considerado, este número variou de 4.5 milhões em 1968 a 8.8 milhões em 1974.

(40) A grande diferença entre os valores calculados para o parâmetro  $\alpha$  deve-se ao fato de que trabalhamos com salários mensais para a Lei dos 2/3 e rendimentos anuais para os dados do IR.

(41) Isto pode ser observado não apenas pela Fig. 6.1. mas também pelos valores mais elevados encontrados para as estatísticas de Durbin-Watson (entre .3 e .9). (Veja nota (30)).

FIGURA 6.1. - Linearização dos dados da Lei dos 2/3 em 1970 (x é o logaritmo da renda R e y é a proporção de pessoas com renda inferior a R).



Obs.: Foram eliminadas as duas primeiras faixas de salários que não se ajustavam satisfatoriamente aos dados.

Desta forma, esta aplicação vem reforçar os resultados encontrados anteriormente, i.e., parece ter havido uma piora na distribuição de renda entre 1968 e 1972/73 seguida de alguma recuperação até meados da década dos setenta.

TABELA 6.1.

RESULTADO DAS REGRESSÕES PARA O PERÍODO  
1968-1974 UTILIZANDO DADOS DA LEI DOS 2/3

Ano	$\log \hat{\alpha}$ (a)	$\hat{\beta}$ (b)	$S_{\hat{\beta}}$ (c)	$R^2$ (d)	$S^2$ (e)
1969	11.048	2.149	.036	.997	.0088
1968	11.540	2.197	.036	.997	.0087
1970	10.710	1.951	.017	.999	.0026
1971	10.332	1.845	.011	.999	.0006
1972	10.445	1.795	.022	.998	.0023
1973	11.151	1.837	.054	.991	.0130
1974	13.186	2.063	.110	.972	.0548

(a)  $\log \hat{\alpha}$  é o estimador de mínimos quadrados de  $\log a$ .

(b)  $\hat{\beta}$  é o estimador de mínimos quadrados de  $\beta$ .

(c)  $S_{\hat{\beta}}$  é o estimador para o desvio padrão de  $\hat{\beta}$ .

(d)  $R^2$  é o coeficiente de determinação.

(e)  $S^2$  é o estimador para a variância dos resíduos.

**OBS.:** Para 1968 e 1969 foram eliminadas as três primeiras faixas de salários, para 1970 as duas primeiras faixas, para 1971 e 1972 as quatro primeiras faixas e para 1973 e 1974 a primeira faixa, as quais não se ajustavam satisfatoriamente aos dados.

## 7 CONCLUSÕES E COMENTÁRIOS FINAIS

Neste trabalho foi proposto um novo índice de concentração, o qual parte do princípio de que a distribuição log-normal é adequada para representar a distribuição de renda e que a distribui-

ção logística se aproxima da distribuição normal. Embora tenhamos restringido nosso trabalho a rendas pessoais, acreditamos que o índice aqui desenvolvido possa ser satisfatoriamente utilizado para estudos de concentração industrial.

Talvez a maior vantagem de nosso índice seja sua facilidade de cálculo. Ele pode ser determinado através de transformações logarítmicas das variáveis consideradas, seguida da estimação dos parâmetros de uma regressão linear simples.

Nosso índice foi utilizado para a comprovação do intenso processo de concentração de renda no período 1960-1970 e para a verificação da evolução da concentração de renda no Brasil no período 1968-1975. A primeira aplicação nos mostrou a coerência de nosso índice quando comparado com outros índices de concentração. A segunda aplicação nos sugere que o processo de concentração de renda verificado durante a década dos sessenta continuou durante os primeiros anos da década dos setenta. Até 1972/73 a distribuição de renda teria sofrido uma acentuada piora, recuperando-se parcialmente a partir de então, atingindo em 1975 níveis semelhantes aos de 1970. Portanto, poderíamos afirmar que a situação de distribuição de renda no Brasil em 1975 não diferia substancialmente da situação em 1970. Desta forma ainda nos encontraríamos no meio da década dos setenta com um distribuição de renda pior que a existente no país em 1960.

O período aqui estudado abrange os anos do conhecido "milagre econômico brasileiro" Assim verificamos que a distribuição de renda no Brasil parece ter piorado substancialmente durante grande parte daqueles anos. Apenas ao final do "milagre econômico" teria o Governo optado por uma política de melhoria dos salários reais dos trabalhadores<sup>(42)</sup>. Esta nova política salarial do Governo parece ter sido responsável por uma pequena melhoria na distribuição de renda no país após 1973.

Entretanto, também verificamos que a recuperação relativa dos salários dos trabalhadores foi muito pequena, haja visto que a situação de distribuição de renda em 1975, embora nos pareça melhor que em 1972/73 ainda era semelhante a de 1970. Portanto, a recuperação dos salários dos trabalhadores nos últimos anos foi apenas suficiente para colocá-los na mesma situação em que se encontravam em 1970 em relação ao resto da sociedade (em

---

(42) Veja DIEESE e Suplicy, *op. cit.*

situação ainda bastante inferior a de 1960). Como conclusão podemos afirmar que indiscutivelmente os mais prejudicados com o “milagre econômico” foram os trabalhadores que tiveram que suportar um violento arrocho salarial para que o país pudesse experimentar altas taxas de crescimento econômico.

Com a mudança da política salarial do Governo a partir de meados de 1976, mantendo os aumentos salariais por dissídios coletivos abaixo dos índices inflacionários, é de se esperar um reflexo imediato na situação de concentração de renda. A persistir tal política, a recuperação verificada entre 1972/73 e 1975 deverá ser eliminada, e podemos esperar uma tendência a uma piora na distribuição de renda no Brasil a partir de 1976/77.

#### (A1.) A DISTRIBUIÇÃO LOGÍSTICA<sup>(43)</sup>

A curva de Pearl, também conhecida como curva logística, foi descoberta pelo biólogo e demógrafo Raymond Pearl (1870-1940) ao estudar extensivamente o crescimento de organismos e de populações. Ele mostrou que sua curva logística descrevia bastante bem seus crescimentos. Desde então, a logística tem sido amplamente utilizada em diversos campos da ciência.

A curva logística possui uma versão probabilística que é a distribuição logística, cuja equação é dada por:

$$F(x) = \frac{1}{1 + ae^{-\beta x}} \quad -\infty < X < \infty \quad (A.1)$$

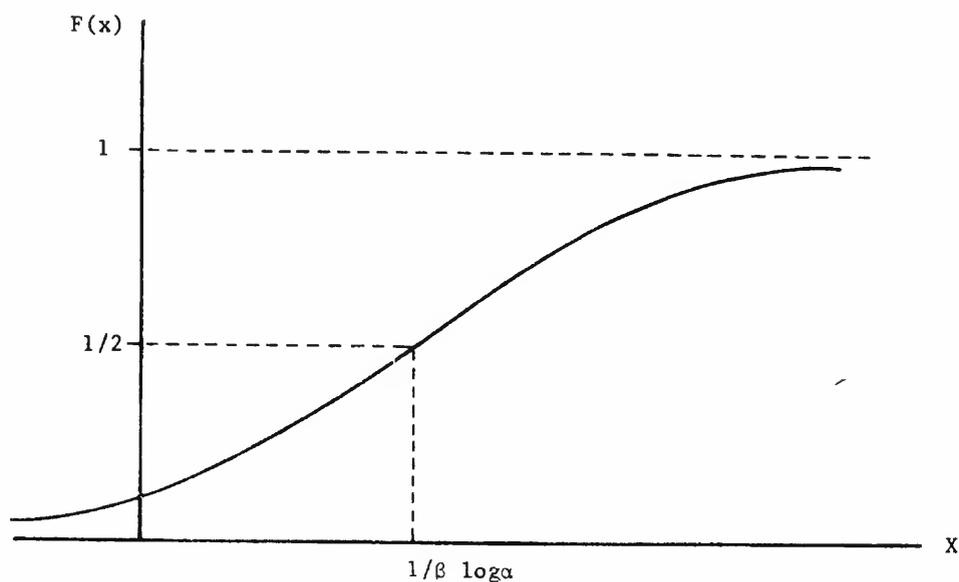
onde  $\alpha$  e  $\beta$  são parâmetros positivos. (Veja Figura A.1.).

A média e a variância da logística são dadas respectivamente por:

$$\mu = \frac{1}{\beta} \log \alpha \quad (A.2)$$

(43) Para uma melhor compreensão dos resultados discutidos no Apêndice 1. sugerimos a leitura de Hastings & Peacock (1975) e Berkson, op. cit.

FIGURA A.1. - Distribuição Logística



$$\sigma^2 = \frac{\pi^2}{3\beta^2} \quad (\text{A.3})$$

Pode-se provar que a logística é simétrica em relação à média  $\frac{1}{\beta} \log \alpha$  de forma que:

$$F(x) \Big|_{x = \frac{1}{\beta} \log \alpha} = \frac{1}{2} \quad (\text{A.4})$$

e

$$\frac{dF(x)}{dx} \Big|_{x = \frac{1}{\beta} \log \alpha} = \frac{\beta}{2} \quad (\text{A.5})$$

Observe que o parâmetro  $\alpha$  determina a localização da logística sobre o eixo horizontal, não influenciando em sua variância. Enquanto isso  $\beta$  se relaciona inversamente com o desvio padrão

$\sigma$  (Veja Equação A.3 ) ao mesmo tempo em que mede a inclinação da logística no ponto médio  $\frac{1}{\beta} \log \alpha$ . (Veja Equação A.5).

Do exposto acima verificamos que a distribuição logística possui várias propriedades similares à normal, sendo utilizada como aproximação da normal em aplicações onde se deseja uma forma explícita para a função de distribuição da variável considerada. Esta aproximação é excelente na parte central das curvas havendo alguma discrepância nas extremidades onde a normal se aproxima mais rapidamente dos valores assintóticos que a logística.

O ajuste da distribuição logística a um conjunto de  $N$  observações pode ser feito pelo método dos mínimos quadrados, ou seja,  $\alpha$  e  $\beta$ , podem ser estimados determinando-se os valores que minimizam a seguinte expressão:

$$\sum_{i=1}^N \left( F(x_i) - \frac{1}{1 + \alpha e^{-\beta x_i}} \right)^2 \quad (\text{A.6})$$

Este processo, entretanto, fornece equações transcendentais, que são difíceis de serem resolvidas. Uma forma de se evitar tal dificuldade consiste em se linearizar a logística. Uma vez obtida sua linearização podemos estimar os parâmetros da logística utilizando o método dos mínimos quadrados lineares o qual é bastante simples.

Da Equação (A.1) obtemos:

$$\frac{F(x)}{1 - F(x)} = \frac{1}{\alpha e^{-\beta x}} \quad (\text{A.7})$$

Tomando-se o logaritmo neperiano de (A.7) obtém-se finalmente

$$G(x) = \log \left( \frac{F(x)}{1 - F(x)} \right) = - \log \alpha + \beta x \quad (\text{A.8})$$

Podemos, portanto, estimar  $a$  e  $\beta$  fazendo-se a regressão da nova variável  $\tilde{G}(x)$  sobre  $x$ .

Deve-se notar que ao minimizarmos a expressão

$$\sum_{i=1}^N (G(x_i) + \log a - \beta x_i)^2 \quad (\text{A.9})$$

não estaremos simultaneamente minimizando a expressão (A.6). Na verdade, pode-se mostrar que este processo minimiza a soma das divisões entre os dados originais e os dados ajustados. Estas divisões são tomadas entre os dados originais e os ajustados, ou vice-versa, dependendo de qual das duas fornece um valor maior que a unidade. Apesar de todas as limitações este processo é simples e para efeitos práticos, pode ser utilizado satisfatoriamente.

Uma observação final deve ser feita aqui. O processo de linearização utilizado para a estimação dos parâmetros  $a$  e  $\beta$  pode ser eliminado se possuímos programas de computador que efetuem estimações não lineares. Neste caso, os parâmetros são obtidos pela minimização da Equação (A.6). Infelizmente, entretanto, estes programas estão implantados em poucos Centros de Computação, o que torna necessária a utilização do processo de linearização descrito neste apêndice.

## (A.2) — O TESTE DE HIPÓTESES UTILIZADO<sup>(44)</sup>

Consideremos uma regressão linear simples entre duas variáveis  $Y$  e  $X$

$$Y_i = \alpha + \beta X_i + \varepsilon_i \quad (\text{A.1})$$

onde o índice  $i$  representa a  $i$ -ésima observação de um total de  $n$  observações do par de variáveis  $X$ ,  $Y$ , e  $\sigma^2$  a variância do resíduo  $\varepsilon$ .

(44) Os resultados desenvolvidos neste Apêndice podem ser encontrados em diversos livros, tais como, Kmenta, *op. cit.*, e Wetherill (1972).

Seja  $\hat{\beta}$  o estimador de mínimos quadrados do parâmetro  $\beta$ .  
É sabido que

$$\hat{\beta} \sim N(\beta, \sigma_{\hat{\beta}}^2) \quad (\text{A.2})$$

onde  $\sigma_{\hat{\beta}}^2 = \text{Var}(\hat{\beta})$

Segue-se que

$$\frac{\hat{\beta} - \beta}{\sigma_{\hat{\beta}}} \sim N(0,1) \quad (\text{A.3})$$

e que

$$\frac{(n-2)s_{\hat{\beta}}^2}{\sigma_{\hat{\beta}}^2} \sim X_{n-2}^2 \quad (\text{A.4})$$

onde  $s_{\hat{\beta}}^2$  é um estimador de  $\sigma_{\hat{\beta}}^2$  dado por

$$s_{\hat{\beta}}^2 = \frac{s^2}{\sum_{i=1}^n (X_i - \bar{X})^2} \quad (\text{A.5})$$

onde

$$\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i \quad (\text{A.6})$$

e

$$s^2 = \frac{1}{n-2} \sum_{i=1}^n (Y_i - \hat{\alpha} - \hat{\beta}X_i)^2 \quad (\text{A.7})$$

Suponhamos que possuímos dois pares de variáveis,  $(Y^{(1)}, X^{(1)})$  e  $(Y^{(2)}, X^{(2)})$ , onde  $Y^{(1)}$  e  $Y^{(2)}$  são independentes entre si, e que existem  $n_1$  observações do par  $(Y^{(1)}, X^{(1)})$  e  $n_2$  observações do par  $(Y^{(2)}, X^{(2)})$ . Consideremos as seguintes regressões lineares simples

$$Y_i^{(1)} = \alpha_1 + \beta_1 X_i^{(1)} + \varepsilon_i^{(1)} \quad (\text{A.8})$$

e

$$Y_i^{(2)} = \alpha_2 + \beta_2 X_i^{(2)} + \varepsilon_i^{(2)} \quad (\text{A.8a})$$

onde  $\text{Var}(\varepsilon^{(1)}) = \sigma^2_1$  e  $\text{Var}(\varepsilon^{(2)}) = \sigma^2_2$

Como poderíamos realizar os testes de hipóteses abaixo?

$$H_0: \beta_1 = \beta_2 \quad (\text{A.9})$$

$$H_A: \beta_1 > \beta_2$$

ou

$$H_0: \beta_1 = \beta_2 \quad (\text{A.9a})$$

$$H_A: \beta_1 \neq \beta_2$$

Utilizando os resultados apresentados acima e devido à hipótese de independência entre  $Y^{(1)}$  e  $Y^{(2)}$  podemos escrever:

$$\hat{\beta}_1 - \hat{\beta}_2 \sim N\left(\beta_1 - \beta_2, \frac{\sigma^2_{\wedge}}{\beta_1} + \frac{\sigma^2_{\wedge}}{\beta_2}\right) \quad (\text{A.10})$$

$$\frac{(\hat{\beta}_1 - \hat{\beta}_2) - (\beta_1 - \beta_2)}{\sqrt{\frac{\sigma^2_{\wedge}}{\beta_1} + \frac{\sigma^2_{\wedge}}{\beta_2}}} \sim N(0, 1) \quad (\text{A.11})$$

$$\frac{(n_1 - 2) \frac{s^2_{\wedge}}{\beta_1}}{\sigma^2_{\wedge}} \sim X^2_{n_1 - 2} \quad (\text{A.12})$$

$$\frac{(n_2 - 2)s_{\beta_2}^2}{\sigma_{\beta_2}^2} \sim X^2_{n_2 - 2} \quad (\text{A.13})$$

$$\frac{(n_1 - 2)s_{\beta_1}^2}{\sigma_{\beta_1}^2} + \frac{(n_2 - 2)s_{\beta_2}^2}{\sigma_{\beta_2}^2} \sim X^2_{n_1 + n_2 - 4} \quad (\text{A.14})$$

Supondo que  $\sigma_1^2 = \sigma_2^2$ , o que pode ser facilmente verificado como veremos adiante, as Equações (A.11) e (A.14) transformam-se em

$$\frac{(\hat{\beta}_1 - \hat{\beta}_2) - (\beta_1 - \beta_2)}{\sqrt{\sigma_1^2(1/x^{(1)} + 1/x^{(2)})}} \quad (\text{A.15})$$

e

$$\frac{(n_1 - 2)S_{\beta_1}^2}{\sigma_1^2/x^{(1)}} + \frac{(n_2 - 2)S_{\beta_2}^2}{\sigma_1^2/x^{(2)}} \quad (\text{A.16})$$

onde

$$x^{(1)} = \sum_{i=1}^{n_1} (X_i^{(1)} - \bar{X}^{(1)})^2 \quad (\text{A.17})$$

$$x^{(2)} = \sum_{i=1}^{n_2} (X_i^{(2)} - \bar{X}^{(2)})^2 \quad (\text{A.18})$$

$$\bar{X}^{(1)} = \frac{1}{n_1} \sum_{i=1}^{n_1} X_i^{(1)} \quad (\text{A.19})$$

$$\bar{X}^{(2)} = \frac{1}{n_2} \sum_{i=1}^{n_2} X_i^{(2)} \quad (\text{A.20})$$

As variáveis aleatórias das Equações (A.15) e (A.16) são independentes entre si de forma que podemos obter uma variável  $t$

dividindo a Equação (A.15) pela raiz quadrada da Equação (A.16) dividida pelos respectivos graus de liberdade ( $n_1 + n_2 - 4$ ). Assim obtemos

$$\frac{(\hat{\beta}_1 - \hat{\beta}_2) - (\beta_1 - \beta_2)}{\sqrt{(n_1-2)S_1^2 + (n_2-2)S_2^2} \left( \frac{1}{x^{(1)}} + \frac{1}{x^{(2)}} \right)} \quad (\text{A. 21})$$

onde

$$S_1^2 = \frac{1}{n_1-2} \sum_{i=1}^{n_1} (Y_i^{(1)} - \hat{\alpha}_1 - \hat{\beta}_1 X_i^{(1)})^2 \quad (\text{A.22})$$

$$S_2^2 = \frac{1}{n_2-2} \sum_{i=1}^{n_2} (Y_i^{(2)} - \hat{\alpha}_2 - \hat{\beta}_2 X_i^{(2)})^2 \quad (\text{A.23})$$

Utilizando-se a estatística fornecida pela Equação (A.21), a qual possui distribuição  $t$  com  $(n_1 + n_2 - 4)$  graus de liberdade, podemos proceder aos testes de hipótese dados pelas Equações (A.9) e (A.9a) no nível de significância desejado.

Supusemos acima que  $\sigma_1^2 = \sigma_2^2$ . Esta hipótese pode ser testada facilmente utilizando-se os seguintes resultados:

$$\frac{(n_1 - 2)S_1^2}{\sigma_1^2} \sim X^2_{n_1 - 2} \quad (\text{A.24})$$

$$\frac{(n_2 - 2)S_2^2}{\sigma_2^2} \sim X^2_{n_2 - 2} \quad (\text{A.25})$$

Além disso, as variáveis aleatórias fornecidas pelas Equações (A.24) e (A.25) são independentes entre si devido à independência entre  $Y^{(1)}$  e  $Y^{(2)}$ . Desta forma, podemos obter uma estatística  $F$  dividindo as Equações (A.24) e (A.25) pelos respectivos graus de liberdade  $(n_1 - 2)$  e  $(n_2 - 2)$ , e dividindo os resultados obtidos.

Assim, se a hipótese nula  $\sigma_1^2 = \sigma_2^2$  for verdadeira, o seguinte resultado é verdadeiro:

$$\frac{S_1^2}{S_2^2} \sim F(n_1 - 2, n_2 - 2) \quad (\text{A.26})$$

No caso em que  $\sigma_1^2 \neq \sigma_2^2$  um outro teste deve ser utilizado. Neste caso não mais podemos obter uma estatística t como no caso anterior. Pode-se mostrar<sup>(45)</sup> que se  $\sigma_1^2 \neq \sigma_2^2$  então

$$\frac{(\hat{\beta}_1 - \hat{\beta}_2) - (\beta_1 - \beta_2)}{\frac{\sqrt{S_1^2}}{x^{(1)}} + \frac{S_2^2}{x^{(2)}}} \quad (\text{A.27})$$

possui aproximadamente uma distribuição t com

$$\frac{\left(\frac{S_1^2}{x^{(1)}} + \frac{S_2^2}{x^{(2)}}\right)^2}{\frac{\left(\frac{S_1^2}{x^{(1)}}\right)^2}{x^{(1)}-1} + \frac{\left(\frac{S_2^2}{x^{(2)}}\right)^2}{x^{(2)}-1}} \quad (\text{A.28})$$

graus de liberdade.

## REFERÊNCIAS BIBLIOGRÁFICAS

- [1] Aitkinson J. & J.A.C. Brown, The Lognormal Distribution, Cambridge, 1957.
- [2] Anuário Econômico Fiscal, CIEF-SERPRO, 1970-1975.
- [3] BERKSON, Mathematical Statistics by B.L. Van der Waerden, Springer-Verlag, 1969.

---

(45) Veja Wetherill, op. cit.

- [4] Boletim Técnico, Centro de Documentação e Informática, Ministério do Trabalho, 1969-1975.
- [5] DAGUM, C., «El Modelo Log-Logístico y la Distribucion del Ingreso en la Argentina», *El Trimestre Económico*, Vol. 44, n.º 176, out/dez., 1977.
- [6] DIEESE, Dez Anos de Política Salarial, *Estudos Sócio-Econômicos*, n.º 3, agosto de 1975.
- [7] GIBRAT, R., *Les Inégalités Économiques*, Paris, Sirey, 1931.
- [8] HART, P. E. & J.S. Prais, «The Analysis of Business Concentration», *Journal of Royal Statistical Society*, 1956, A. 119, pgs. 150-191.
- [9] HASTINGS, N.A.J. & J.B. Peacock, *Statistical Distributions*, Butterworths, 1975.
- [10] HAYAKAWA, M., «The Application of Pareto's Law of Income to Japanese Data», *Econometria*, Vol. 19, abril de 1951.
- [11] KALECKI, M., «On the Gibrat Distribution», *Econometrica*, Vol. 13, n.º 2, abril de 1945, pgs. 161-170.
- [12] KMENTA, J., *Elements of Econometrics*, Macmillan, 1971.
- [13] LANGE, O., *Introdução à Econometria*, Fundo de Cultura, 1963.
- [14] LANGONI, C. G., *Distribuição de Renda e Desenvolvimento Econômico do Brasil*, Expressão e Cultura, 1973.
- [15] PARETO, V., *Cours d'Economie Politique*, Lausanne, 1897.
- [16] SABOIA, J.L.M., «Uma Generalização da Lei de Gibrat para o Crescimento da Firma», *Pesquisa e Planejamento Econômico*, agosto de 1977.
- [17] STEINDL, J., *Random Process and the Growth of Firms — A Study of the Pareto Law*, Griffin, London, 1965.
- [18] SUPLICY, E. M., «A Política Salarial e os Índices de Preços», *Palestra proferida na Comissão de Economia do Senado Federal*, Brasília, setembro de 1977.
- [19] TOLIPAN, R. & A.C. Tinelli, *A Controvérsia sobre Distribuição de Renda e Desenvolvimento*, Zahar, 1975.
- [20] WETHERILL, G. B., *Elementary Statistical Methods*, Chapman & Hall, London, 1972.