I.C.M.S.C.

NOTAS DO INSTITUTO DE CIÊNCIAS MATEMÁTICAS DE SÃO CARLOS-USP

ASPECTOS DA DISTRIBUIÇÃO DE PARETO NO PROBLEMA DE AJUSTAMENTO DE DISTRIBUIÇÕES A DADOS ALTERADOS

E. C. Leme

Volume nº 13

UNIVERSIDADE DE SÃO PAULO

INSTITUTO DE CIÊNCIAS MATEMÁTICAS DE SÃO CARLOS

13560 - SÃO CARLOS - SÃO PAULO - BRASIL

ASPECTOS DA DISTRIBUIÇÃO DE PARETO NO PROBLEMA DE AJUSTAMENTO DE DISTRIBUIÇÕES A DADOS ALTERADOS

E. C. Leme

Volume nº 13

§ 0 - INTRODUÇÃO

Nosso estudo iniciou-se investigando a adequação do ajustamento da distribuição de Pareto a dados sobre rendimentos. Uma variedade de formas funcionais tem sido sugeridas como sendo convenientes para descrever distribuições de tais dados; algumas tem sido obtidas a partir de modelos que retratam a geração de uma distribuição de rendimentos enquanto outras se propõem apenas a ajustar razoavelmente bem as observações. Uma das que tem sido amplamente consideradas é a distribuição do tipo secante hiperbólica ao quadrado. Esta distribuição tem certas características que a tornam um instrumento de utilidade na análise e comparação de distribuições de rendimentos. A equação diferencial da qual a distribuição do tipo secante hiperbólica ao quadrado é obtida pode ser variada para permitir um grande número de distribuições diferentes para serem ajustadas. Existe uma semelhança entre essa distribuição e a distribuição de Pareto. Com o intuito de apresentar algumas dificuldades com as quais se depararam pesquisadores ao investigar o ajustamento de distribuições a dados sobre "rendimentos" e indicar alguns resultados mais significativos obtidos nesse contexto, são feitas as considerações apresentadas na secção 1.

Em se tratando de dados sobre rendimentos que estejam, de alguma forma, alterados ou modificados, foi de interesse a questão apresentada por Krishnaji (1970) em seu trabalho "Characterization of the Pareto Distribution through a model of underreported incomes", qual seja: assumindo Y como rendimento declarado e X como rendimento real, relacionados através do modelo $Y = RX$, com R independente de X e assumindo valores em (a,b), $(a,b) \subset [0,1]$, qual a classe das distribuições de X que satisfaz a propriedade $P(RX > y/RX > x_0) = P(X > y)$, para $y > x_0$ reais, (isto é, a distribuição da variável rendimento declarado, truncada à esquerda de um valor mínimo, coincide com a distribuição da variável rendimento real)? Krishnaji respondeu essa questão mostrando que, no caso em que a distribuição da variável R é uma Beta de parâmetros $p \in R^+$ e $q = 1$, a única família de soluções é a família das distribuições de Pareto.

Nas secções 2.1 e 2.2 abordamos a questão colocada, assumindo para a variável R uma distribuição Beta de parâmetros $p \in R^+$ e $q \in Z^{++}$ mostrando que, também neste caso, a única família de soluções satisfazendo a propriedade exigida é a família das distribuições de Pareto. Além disso,tornou-se de interesse a determinação de certas quantidades relacionadas às variáveis X e Y, em questão, tais como probabilidades, médias e variâncias condicionadas e o chamado índice de desonestidade; os resultados obtidos generalizam os apresentados por Hartley e Revankar [14].

Na secção 3 apresentamos uma tabela em que se pode obter valores do parâmetro λ da distribuição de Pareto correspondentes a valores de $V(x_0) = P(Y > x_0)$ e dos parâmetros da distribuição Beta, assim como valores do índice de desonestidade, correspondentes a cada caso.

Os comentários da secção 4 foram apresentados no sentido de situar os resultados aqui obtidos no contexto mais geral do problema.

Finalmente, desejo agradecer ao Prof. Barry C. Arnold por me haver sugerido este trabalho.

§ 1 - ALGUMAS CONSIDERAÇÕES SOBRE A DISTRIBUIÇÃO DE PARETO

Muito embora, em trabalhos teóricos, seja comum se desenvolver resultados usando-se distribuições de estatísticas contínuas, relativamente poucas dentre tais famílias de distribuições são comumente usadas como modelos para descrever fenômenos empíricos. A família das Normais, por exemplo, tem classicamente seu emprego como uma *lei de erros* para os dados empíricos. Para dados relacionados a variável tempo de vida, são usuais a Gama, a Lognormal, a Weibull e famílias com valores extremos modificados (Davis [8], Epstein e Sobel [9] e Barlow e Proschan [4]). Em Metereologia e Hidrologia foi também mostrada a importância da distribuição dos valores extremos (Gumbel [12]). Na descrição de dados empíricos, as famílias em geral usadas são a Lognormal e a Gama.

Uma classe de funções distribuição que aparece em uma grande variedade de dados empíricos, particularmente dados descrevendo fenômenos sociológicos, biológicos e econômicos, foi estudada por Simon [36]. As distribuições empíricas a que se referiu, especificamente, foram:

a) distribuições de palavras em amostras de frases, por sua frequência de ocorrência;
b) distribuições de cientistas, por número de trabalhos publicados;
c) distribuições de cidades, por população;
d) distribuições de rendimentos, por magnitude;

e) distribuições de gens biológicos, por nº de espécie.

Sua abordagem foi feita usando a terminologia de Processos Estocásticos, apresentando um tipo de processo que leva a uma classe de distribuições acentuadamente assimétricas (a distribuição de Yule) possuindo as seguintes características:

i) seu aspecto gráfico é o de uma curva fortemente assimétrica com longa cauda superior podendo esta, geralmente, ser aproximada por uma função da forma

$$f(x) = \left(\frac{a}{x}\right)^k b^x \tag{1}$$

onde a, b e k são constantes, com b tão próxima da unidade que, em uma primeira aproximação, o fator final tem um efeito significativo sobre f(x) apenas para grandes valores de x;

ii) o expoente k é maior do que 1 e, nos casos de distribuição de frequências de palavras, publicações e populações urbanas, é bastante próximo de 2 (Para numerosos exemplos de distribuições com essa propriedade, ver Zipf [39]).

iii) no caso de frequências de palavras, publicações e gens biológicos, a função (1) descreve a distribuição não meramente para valores caudais como também para pequenos valores de x.

A família das distribuições de Pareto, embora inicialmente formulada para caracterizar a graduação de rendimentos de uma população, tornou-se também usual em outros contextos tais como em análises de tamanhos de firmas, tamanhos de cidades, frequências de palavras, migração, bens patrimoniais, etc.

Particularmente na área de rendimentos (ou *recebimentos*, em geral) foi básico o trabalho de Wilfredo Pareto [29], professor de Economia que, investigando as condições para o bem-estar econômico de uma população e tentando descrever a configuração da desigualdade de rendimentos das pessoas, observou que, denotando por X o rendimento de uma pessoa qualquer da população, se

$F(x) = P(X \leq x)$ (= probabilidade do rendimento ser, no máximo, igual a um valor x)

então, ao menos para valores grandes de x,

$$\log[1-F(x)] \simeq \beta - \alpha \log x \qquad (2)$$

onde β e α são parâmetros, α representativo da forma da distribuição e conhecido como "constante de Pareto".

Na Europa, no século XIX, a adequação dessa lei aos dados de rendimentos foi tão perfeita que Pareto foi levado a considerar (2) como uma "lei natural" para esse problema concluindo, adicionalmente, que α deveria ser próxima de 1,5, a despeito de variações no tempo e espaço.

Em 1925, Levy [20] mostrou a existência de uma classe de distribuições que, além de seguir a forma assintótica da lei de Pareto, era caracterizada pelo fato de $0<\alpha<2$. Estas são conhecidas como as distribuições de Pareto estáveis, ou distribuições não-Gaussianas estáveis.

Como estabelecida originalmente por Pareto, a lei não considera a atuação nos dados de agentes externos (tais como impostos) e mesmo a ocupação ou condição política e social das pessoas envolvidas. Isto, e mais o fato de que, embora a lei se ajuste bem aos valores extremos, não é muito adequada para se ajustar a valores mais centrais da variável, levaram a posteriores estudos e novas sugestões para a resolução do problema. Dessa forma, várias outras distribuições foram propostas como modelos para descrever rendimentos devendo esses, além de fornecer uma aproximação razoavelmente boa para a distribuição dos valores reais, também possuir parâmetros simples de serem estimados e de serem interpretados em termos da Economia. O que foi observado no entanto na prática, é que esses dois critérios competem entre si.

Embora, com os novos modelos, um número considerável de evidências empíricas tenha sido inspecionado apresentando variações nos resultados, conjuntamente a estas, foi mostrado também que, muitas vezes, poder-se-ia aproximar a cauda superior de uma distribuição de rendimentos por (2). Dessa forma, tornou-se como que uma "imposição" às distribuições escolhidas como modelos para descrever rendimentos, terem suas caudas aproximadas pela lei de Pareto (2). Diversos autores tais como Champernowne [6]; [7], Mandelbrot [24], Fisk [11] e Ord [28], tentaram obter justificativas teóricas para tal imposição.

Champernowne [6], para ajustar a distribuição de rendimentos observados, sugeriu o modelo com função densidade dada por

$$f(x) \simeq \frac{u}{1+2\lambda u+u^2} \qquad (3)$$

com $u = \exp[-\alpha(y-y_0)]$ e $y = \log x$, que ele chamou de "*potência de rendimento*". Pelo fato dessa distribuição levar à cauda Paretiana em ambas as extremidades e ter sido apresentado para assegurar (3) um considerável conjunto de evidências, esse modelo tornou-se bastante atraente. No entanto, observou-se como desvantagem dessa distribuição que, apesar de existirem todos os momentos de $Y = \log X$, nem todos os momentos de X podem ser obtidos. As razões apresentadas para o estabelecimento de (3) levaram, em geral, a $\lambda < 1$. Fisk [11] desenvolveu seu trabalho para o caso especial de $\lambda = 1$ quando então obteve

$$f(y) = \operatorname{sech}^2[\alpha(y-y_0)] = \frac{u}{1+u^2} \qquad (4)$$

e $F(x) = 1 - (1+u)^{-1}$, observando também (4), o modelo logístico para Y, como um membro da família Burr (Ord [27]) e sugerindo outras convenientes transformações dessa família como modelos para distribuições de rendimentos. Posteriormente, Harrison [13] obteve resultados sensivelmente melhores com um tipo de abordagem por mínimos quadrados tendo, comparativamente, o único demérito de trabalhar com uma família de funções a quatro parâmetros, contra a de três introduzida por Champernowne.

Sob a terminologia estocástica de Cadeias de Markov, Champernowne [7] apresentou um modelo que levou a (1) e a generalizações dessa função, com hipóteses uma das quais a de ser a população em estudo, fechada (nenhuma ocorrência de nascimentos ou mortes). Por ser essa condição considerada irrealística, surgiram novas sugestões de modelos estocásticos tais como os sugeridos por Rutherford [33] com características de nascimento e morte, Sargan [34] incorporando além de nascimentos e mortes, também doações e "funções de segurança" e Wold e White [38], com características de nascimento e morte e a suposição da existência de um nível mínimo de rendimento. Por ser a abordagem estocástica não aceita universalmente surgiu, como um meio termo entre esta e a abordagem determinística, o modelo proposto por Lydall [21] que, assumindo a existência de um rendimento mínimo e uma estrutura tipo pirâmide, obteve uma lei geométrica para log X, implicando em uma cauda Paretiana.

No contexto das distribuições estáveis (ver Feller, 1966, 540-549), Mandelbrot [24] introduziu uma versão nova da lei de Pareto que ele denominou de Lei de Pareto-Levy que, parcialmente, também explica a configuração dos valores mais centrais da distribuição de rendimentos. A densidade dessa lei é determinada através de sua transformada bilateral de Laplace, (válida para $b > 0$),

$$G(b) = \int_{-\infty}^{\infty} \exp(-bx)\,|dF(x)| = \exp[(bx^*)^{\alpha} + Mb] \qquad (5)$$

onde x^* é um parâmetro escalar positivo, $M = E(X)$ e α, a constante de Pareto, é sujeita à condição $1 < \alpha < 2$.

De (5) obtem-se que, para $x \to \infty$,

$F(x) \simeq x^{-\alpha}[x^*\Gamma(1-\alpha)]^{\alpha}$

onde $\Gamma(1-\alpha)$ é função-gama de Euler; ou de outra forma,

$$\frac{F(x)}{x^{-\alpha}[x^*\Gamma(1-\alpha)]^{\alpha}} \to 1, \text{ quando } x \to \infty \qquad (6)$$

Um estudo comparativo, em termos dos valores de α, é feito com essa lei e a aqui denominada de Lei "forte" de Pareto, dada por

$$f(x) = \begin{cases} 1 - \left(\frac{x}{x_0}\right)^{-\alpha}, & x > x_0,\ \alpha > 0 \\ 0, & \text{caso contrário} \end{cases} \qquad (7)$$

com x_0 um fator escala e α um tipo de índice de desigualdade da distribuição, concluindo-se que as respectivas curvas se tornam rapidamente indestinguíveis quanto "mais distante" estiver α do valor 2.

Uma das principais reivindicações nesse trabalho foi a de mostrar que *é impossível explicar porque a lei de Pareto e não alguma outra lei é satisfeita por distribuições de rendimentos, sem conjuntamente estudar o problema provocado pelo fato que, essencialmente, a mesma lei continua a ser* *respeitada pela distribuição de rendimento, a despeito das mudanças na definição desse termo, uma vez que existem diversas maneiras de distinguir diferentes fontes de rendimento, podendo o mesmo ser uma soma de elementos tais como, A, rendimento agrícola, industrial ou comercial; B, rendimento em ações ou em dinheiro; C, rendimento usualmente tributado ou ganhos de capital; D, rendimento de diferentes membros de uma única fonte pagadora, etc.*

Sob esse tipo de argumentação e considerando não ser uma distribuição de rendimentos rigorosamente invariante sobre todo o campo de variação de x com respeito a uma mudança na definição do termo "rendimento", é colocado que: *Se a distribuição da soma de muitos componentes não for Gaussiana, for assimétrica e tal que* $E(X) < \infty$, *então a primeira afirmação mais razoável relativa à distribuição dessa soma é que ela segue a lei de Pareto-Levy.* Um outro resultado também apresentado é que a soma de um número finito de variáveis independentes e identicamente distribuidas de acordo com a lei "forte" de Pareto tem uma distribuição que tende assintoticamente a essa mesma lei, para valores caudais. A relação dada em (7) é mais propriamente conhecida como a *distribuição de Pareto do 1º tipo ou distribuição de Pareto clássica.*

É possível se definir famílias mais gerais de distribuições que incluem as distribuições de Pareto como casos especiais. Algumas famílias de distribuições de Pareto generalizadas, dadas através de sua função sobrevivência $P(x) = 1 - F(x) = P(X > x)$, são apresentadas por Arnold e La guna [3]. Mais especificamente,

a) Distribuição de Pareto do tipo I, ou distribuição de Pareto clássica, denotada por $P(I)(\sigma,\alpha)$, dada por

$$P(x) = \left(\frac{x}{\sigma}\right)^{-\alpha}, \quad x \geq \sigma,$$

onde $\sigma > 0$ é o parâmetro escala e $\alpha > 0$ é o índice de desigualdade, assumido, em geral, maior do que um.

b) Distribuição de Pareto do tipo II, $P(II)(\mu,\sigma,\alpha)$ dada por

$$P(x) = \left[1 + \frac{x-\mu}{\sigma}\right]^{-\alpha}, \quad x \geq \mu,$$

onde $\mu \in R$ é o parâmetro de locação, $\sigma > 0$ é o parâmetro escala e $\alpha > 0$ é chamado simplesmente de *o índice*, também, em geral, assumido maior do que um.

c) Distribuição de Pareto do tipo III, $P(III)(\mu,\sigma,\gamma)$, dada por

$$P(x) = \left[1 + \left(\frac{x-\mu}{\sigma}\right)^{1/\gamma}\right]^{-1}, \quad x \geq \mu,$$

onde $\mu \in R$ é o parâmetro de locação, $\sigma > 0$ é o parâmetro-escala e $\gamma > 0$ é o parâmetro de Gini, identificado no caso de $\mu = 0$ e $\gamma \leq 1$, com o índice de Gini de desigualdade de rendimentos.

d) Distribuição de Pareto do tipo IV, $P(IV)(\mu,\sigma,\gamma,\alpha)$, dada por

$$P(x) = \left[1+\left(\frac{x-\mu}{\sigma}\right)^{1/\gamma}\right]^{-\alpha}, \quad x \geq \mu,$$

onde $\mu \in R, \sigma > 0$, $\gamma > 0$ e $\alpha > 0$, $(\alpha < \gamma)$, podem ser identificados respectivamente como parâmetros de locação, escala, de Gini e o índice.

e) Distribuição de Feller-Pareto a cinco parâmetros, denotada por $FP(\mu,\sigma,\gamma,\lambda_1,\lambda_2)$, estabelecida através do resultado *se Y for uma Beta com parâmetros* λ_1 e λ_2 *e se para* $\mu \in R$, $\sigma > 0$ *e* $\gamma > 0$, *for definida a variável* $W = \mu+\sigma(Y^{-1}-1)^{\gamma}$, *então* W *terá uma distribuição de Feller-Pareto*". A função sobrevivência de W é obtida através de tabelas da função beta incompleta e através do resultado:

$$P(w) = P\{Y \leq [1 + (\frac{\omega-\mu}{\sigma})^{1/\gamma}]^{-1}\}$$

Para essas distribuições generalizadas são definidas e estudadas medidas de desigualdade tais como a curva de Lorenz e o índice de Gini, são apresentados resultados relacionados à construção de distribuições bivariadas e são sugeridas algumas técnicas sobre estimação e testes de hipóteses.

Kotz e Johnson [17] apresentam vários resultados, sobre a família das distribuições de Pareto relacionados à sua gênese, a certas estatísticas, parâmetros, estimação, caracterização, aplicações e distribuições relacionadas à mesma.

Na área de estimação, resultados foram apresentados por Malik [23] que desenvolveu uma estimação para os parâmetros da distribuição (7) através do método da máxima verossimilhança e do ponto de vista Bayesiano; Lwin [22] que apresentou uma estimação da probabilidade existente na cauda da distribuição (7) usando, além de estimadores não viciados de variância mínima, também estimadores Bayesianos. Também Mandelbrot [25], Quandt [30], Aigner [1], Aigner e Goldberger[2] e Kulldorf e Vannman [19] trataram com vários problemas de estimação em conexão com as distribuições de Pareto. Em todos esses trabalhos assume-se que os valores amostrais não contem erros oriundos do processo de observação. Em muitas aplicações entretanto, particularmente com respeito a variáveis tais como rendimentos e bens patrimoniais, pode-se esperar que os valores observados ou declarados estejam alterados ou modificados com relação aos valores reais da variável. Krishnaji [18] e Revankar, Hartley e Pagano [32] desenvolveram seus trabalhos nesse tipo de contexto obtendo caracterizações da distribuição (7) e Hartley e Revankar [14] consideraram problemas relacionados à estimação dos parâmetros dessa distribuição, a partir de amostra de observações alteradas da variável; posteriormente, Hinkley e Revankar [15] complementaram esses resultados com uma análise mais detalhada.

§2 - O MODELO DE KRISHNAJI PARA DISTRIBUIÇÃO DE DADOS ALTERADOS

§2.1 - No contexto de haver alteração nos valores reais da variável, foi de interesse o resultado obtido por Krishnaji [18] que, sob o modelo

$$Y = RX \qquad (8)$$

onde X representa a variável rendimento real, Y o rendimento declarado e R uma variável independente de X, distribuida de acordo com uma Beta de parâmetros p (p real positivo) e q = 1, mostrou que a distribuição de X coincide com a distribuição de Y, truncada à esquerda de um valor x_0, se e somente se X tiver uma distribuição de Pareto sobre (x_0,∞).

Uma questão aqui colocada foi se esse resultado permaneceria ainda válido independentemente da distribuição assumida para a variável R.

O problema que então consideramos foi o de investigar a classe das distribuições de X, relacionada com Y através do modelo (8), satisfazendo a condição

$$P(RX>y/RX>x_0) = P(X > y) \quad , \quad y > x_0, \tag{9}$$

com R uma variável independente de X, tendo uma distribuição Beta de parâmetros p (p real, positivo) e q ($q \in Z^{++}$) e ainda, analisar aspéctos relevantes das distribuições em pauta.

§2.2 - Denotando por F(.), G(.) e H(.), respectivamente, as funções distribuição de X, Y e R e por $U(x) = 1 - F(x)$, $V(y) = 1 - G(y)$, obtem-se de (8):

$$V(y) = \int_0^1 U(y/r)dH(r) \tag{10}$$

Pode-se observar que, com $y > x_0$,

$$P(RX > y/RX > x_0) = \frac{V(y)}{V(x_0)} = U(y) = P(X > y),$$ o que implica que, para $y \leq x_0$, $U(y) = 1$.

Sendo a densidade da distribuição assumida para R, dada por

$$h(r) = \begin{cases} \frac{1}{B(p,q)} r^{p-1}(1-r)^{q-1}, & 0 \leq r \leq 1,\ p>0,\ q\varepsilon Z^{++} \\ 0 & \text{, caso contrário} \end{cases} \qquad (11)$$

onde $B(p,q)$ representa a função Beta de parâmetros p,q, e considerando (9) e (10), tem-se:

$$V(y) = \int_0^1 U(y/r) \frac{1}{B(p,q)} r^{p-1}(1-r)^{q-1}\, dr.$$

Fazendo $x = y/r$ e denotando $c_q = V(x_0)B(p,q)$, obtem-se

$$c_q U(y) = y^p \int_0^1 U(x)x^{-(p+q)}(x-y)^{q-1}dx \qquad (E)$$

Assumindo a função distribuição U(.) absolutamente contínua, diferenciando (E) com relação a y, chega-se, para $q = 1$,

$$c_1[U(y)y^{-p}]^{(1)} = -U(y)y^{-(p+1)}, \quad y > x_0, \qquad (E_1)$$

cuja solução, sob a condição de ser $V(x_0) > 0$, é dada por

$$U(y) = k\ y^{-p[\frac{1}{V(x_0)} - 1]}$$

que, com $U(y) = 1$ para $y \leq x_0$, fornece

$$U(y) = \left(\frac{y}{x_0}\right)^{\lambda}, \quad y > x_0 > 0, \ \lambda = p\left[1 - \frac{1}{V(x_0)}\right] < 0,$$

que é a função sobrevivência da distribuição de Pareto clás sica.

Em (E), tomando $q = 2$ e derivando duas vezes, obtem-se

$$y^2 U^{(2)}(y) - 2pyU^{(1)}(y) + \frac{c_2 p(1+p)-1}{c_2} U(y) = 0 \qquad (E_2)$$

Com $y = e^t$, obtem-se a equação diferencial a coeficientes constantes

$$\left[D^2 - (2p+1)D + \frac{c_2 p(1+p)-1}{c_2}\right]U = 0, \text{ onde } D^n U = \frac{d^n U}{dt^n}.$$

As raizes da correspondente equação característica são reais e a solução geral de (E_2) é dada por:

$$U(y) = k_1 y^{\lambda_1} + k_2 y^{\lambda_2} \qquad (12)$$

onde $\begin{cases} \lambda_1 = \frac{1}{2}\left[2p + 1 - \left(\frac{4+c_2}{c_2}\right)^{1/2}\right] < 0 \\ \lambda_2 = \frac{1}{2}\left[2p + 1 + \left(\frac{4+c_2}{c_2}\right)^{1/2}\right] > 0 \end{cases}$

Como $F'(y) = -k_1\lambda_1 y^{\lambda_1 - 1} - k_2\lambda_2 y^{\lambda_2 - 1}$, e ainda, como para $\lambda_2 > 0$ tem-se $\int_{x_0}^{\infty} y^{\lambda_2 - 1} dy$ divergente, conclui-se que (12) não será uma solução de interesse a menos que $k_2=0$. Nesse caso, tem-se

$$U(y) = k_1 y^{\lambda_1} \quad , \quad y > x_0 > 0, \qquad \lambda_1 < 0,$$

que sob a condição $U(y) = 1$ para $y \leq x_0$, mostra que X tem uma distribuição de Pareto clássica.

Assim, sucessivamente, fazendo na equação (E), $q = n$, $n \in Z^+$, $n \geq 3$ e derivando n vezes, obtem-se:

$$y^n U^{(n)}(y) - \binom{n}{1} p y^{n-1} U^{(n-1)}(y) + \binom{n}{2} p(p+1) y^{n-2} U^{(n-2)}(y) +$$

$$- \binom{n}{3} p(p+1)(p+2) y^{n-3} U^{(n-3)}(y) + \ldots + (-1)^{n-1} \binom{n}{n-1} p(p+1) \ldots$$

$$\ldots (p+n-2) y U^{(1)}(y) + (-1)^n \frac{c_n p(p+1)\ldots(p+n-1) - (n-1)!}{c_n} U(y) = 0 \qquad (E_n)$$

(E_n) é uma equação $L(y) = 0$, de Euler, em que o operador L tem a forma

$$L(y) = A_0 x^n y^{(n)} + A_1 x^{n-1} y^{(n-1)} + \ldots + A_{n-1} x y^{(1)} + A_n y$$

Ao operador $L(U)$ associamos o polinômio indicial da equação, dado por

$$L(\lambda) = \lambda(\lambda-1)\ \ldots\ (\lambda-n+1) - \binom{n}{1}p\lambda(\lambda-1)\ \ldots\ (\lambda-n+2) +$$

$$\binom{n}{2}p(p+1)\lambda(\lambda-1)\ \ldots\ (\lambda-n+3) - \binom{n}{3}p(p+1)(p+2)\lambda(\lambda-1)\ldots$$

$$\ldots\ (\lambda-n+4)+\ldots+(-1)^{n-1}\binom{n}{n-1}p(p+1)\ldots(p+n-2) +$$

$$+(-1)^n \frac{C_n p(p+1)\ldots(p+n-1)-(n-1)!}{c_n}$$

Como $U(y) = P(X>y)$, as soluções de (E_n), de interesse, deverão ser funções monótonas, limitadas e ainda, deverão descrever probabilidade. Sob essas condições, as soluções deverão ser do tipo:

$$U(y) = \Sigma a_{ij}(\ell n y)^i y^{\lambda j}, \quad y > x_0 > 0, \quad \lambda j < 0 \tag{13}$$

$$\text{com } U(y) = 1, \text{ para } y \leq x_0 \tag{13'}$$

Uma análise mais detalhada do polinômio indicial pode ser feita no sentido de mostrar a não ocorrência de raízes múltiplas negativas em 13, compatíveis com 13'.

Sendo assim, teremos em (13), $i = 0$. (Shimizu[35] e Huang [16] estabelecem esse resultado num contexto bem mais geral para o problema em questão).

Pelas condições impostas a $U(x)$, concluimos que X tem uma distribuição de Pareto.

Também com o objetivo de determinar algumas quantidades de interesse relacionadas a X e Y, ainda sob o modelo (8) com R distribuida segundo (11) e X com uma distribuição clássica de Pareto cuja densidade é dada por

$$f(x) = \begin{cases} \dfrac{\lambda}{x_0} \left(\dfrac{x_0}{x}\right)^{\lambda+1}, & x > x_0 > 0,\ \lambda > 2, \\ 0 & , \text{caso contrário} \end{cases} \qquad (14)$$

consideramos, como em Hartley e Revankar [14], a proporção de X que é "alterada", isto é,

$$1 - R = \frac{X-Y}{X} = \frac{U}{X} \qquad (15)$$

Da independência entre X e R e em vista de (11), (14) e (15), obtem-se como densidade conjunta de X e U, a função

$$p_1(x,u) = f(x).h(1-\frac{u}{x}).\frac{1}{x}, \quad \text{para} \quad x \geq x_0,\ u \geq 0,$$

e como densidade conjunta de Y e U, a função

$$p_2(y,u) = f(y+u).h(\frac{y}{y+u})\frac{1}{y+u}, \quad \text{para } y + u \geq x_0,\ u \geq 0,$$

Para se determinar a densidade marginal de Y, deve-se considerá-la para valores y com $0 \leq y \leq x_0$ e com $y > x_0$.

Para $0 \leq y \leq x_0$,

$$g_1(y) = \int_{x_0-y}^{\infty} \frac{\lambda x_0^{\lambda+1}}{x_0 B(p,q)} \; y^{p-1} u^{q-1} \; \frac{1}{(y+u)^{\lambda+p+q}} \, du \qquad (16)$$

Adotando a transformação $x = y + u$ e após algum desenvolvimento chega-se a

$$g_1(y) = \frac{\lambda}{x_0 B(p,q)} \left(\frac{x_0}{y}\right)^{-(p+q-2)} \sum_{n=0}^{q-1} \binom{q-1}{n} \frac{(-1)^{q-1-n}}{\lambda+p+q-1-n} \left(\frac{x_0}{y}\right)^{n} \qquad (17a)$$

Para $y > x_0$, tem-se

$$g_2(y) = \int_{0}^{\infty} \frac{\lambda x_0^{\lambda+1}}{x_0 B(p,q)} \; y^{p-1} \; u^{q-1} \; \frac{1}{(y+u)^{\lambda+p+q}} \, du$$

que, com a mesma transformação adotada em (16) e após algum desenvolvimento, fornece

$$g_2(y) = \sum_{n=0}^{q-1} \binom{q-1}{n} \frac{(-1)^{q-1-n}}{\lambda+p+q-1-n} \cdot \frac{1}{B(p,q)} \frac{\lambda}{x_0} \left(\frac{x_0}{y}\right)^{\lambda+1} \qquad (17b)$$

A função densidade de Y é então dada por:

$$g(y) = \begin{cases} \dfrac{\lambda}{x_0 B(p,q)} \left(\dfrac{x_0}{y}\right)^{-(p+q-2)} \displaystyle\sum_{n=0}^{q-1} \binom{q-1}{n} \dfrac{(-1)^{q-1-n}}{\lambda+p+q-1-n} \left(\dfrac{x_0}{y}\right)^{n}, & 0 \leq y \leq x_0 \\[2ex] \displaystyle\sum_{n=0}^{q-1} \binom{q-1}{n} \dfrac{(-1)^{q-1-n}}{\lambda+p+q-1-n} \cdot \dfrac{1}{B(p,q)} \cdot \dfrac{\lambda}{x_0} \left(\dfrac{x_0}{y}\right)^{\lambda+1}, & y > x_0 \end{cases} \qquad (18)$$

Os resultados (13) e (18) generalizam o obtido por Krishnaji [18] em seu Teorema 1, como também o obtido por Hartley e Revankar [14]. Com isso, estabelecemos o seguinte resultado:

Seja X uma variável aleatória com função distribuição absolutamente contínua e seja R uma variável aleatória independente de X, com densidade dada por (11) e tal que $P(RX > x_0) > 0$ para algum $x_0 > 0$. Então a distribuição de RX, truncada à esquerda de x_0, coincide com a distribuição de X, se e somente se X tiver uma distribuição de Pareto sobre (x_0,∞).

De (18) obtem-se

$$P(0 \leq Y \leq x_0) = \frac{\lambda}{B(p,q)} \sum_{n=0}^{q-1} \binom{q-1}{n} \frac{(-1)^{q-1-n}}{(\lambda+p+q-1-n)(p+q-1-n)} \tag{19a}$$

e

$$P(Y > x_0) = \frac{1}{B(p,q)} \sum_{n=0}^{q-1} \binom{q-1}{n} \frac{(-1)^{q-1-n}}{\lambda+p+q-1-n} \tag{19b}$$

Consideremos agora o problema de predizer qual o valor de X que corresponde, em média, a um dado valor de Y.

Temos $Y = RX = X - U$; logo, alternativamente a $E(X/Y=y)$, podemos calcular $E(U/Y=y)$.

Para $0 \leq y \leq x$, a densidade condicional de U é dada por

$$\phi(u/y;\ 0 \leq y \leq x_0) = \frac{p_2(y,u)}{g_2(y)} =$$

$$= \frac{x_0^{\lambda+p+q-1}}{\sum_{n=0}^{q-1} \binom{q-1}{n} \frac{(-1)^{q-1-n}}{\lambda+p+q-1-n} \left(\frac{x_0}{y}\right)^n} \cdot \frac{1}{(y+u)^{\lambda+p+q}} \left(\frac{u}{y}\right)^{q-1},\quad x_0-y\leq u<\infty \tag{20a}$$

Para $y > x_0$, a densidade condicional de U é dada por

$$\phi(u/y;y>x_0) = \frac{p_2(y,u)}{g\ (y)}$$

Ou seja,

$$\phi(n/y;y>x_0) = \frac{1}{\sum_{n=0}^{q-1} \binom{q-1}{n} \frac{(-1)^{q-1-n}}{\lambda+p+q-1-n}} \cdot \frac{y^{\lambda+p} u^{q-1}}{(y+u)^{\lambda+p+q}},\quad u \geq 0 \tag{20b}$$

Portanto, com $\lambda + p > 1$,

$$E(U/Y=y;0\leq y\leq x_0) = \frac{x_0^{\lambda+p+q-1}\ y^{-(q-1)}}{\sum_{n=0}^{q-1} \binom{q-1}{n} \frac{(-1)^{q-1-n}}{\lambda+p+q-1-n} \left(\frac{x_0}{y}\right)^n} \int_{x_0-y}^{\infty} \frac{u^q}{(y+u)^{\lambda+p+q}}\, du$$

$$= \frac{\sum_{n=0}^{q} \binom{q}{n} \frac{(-1)^{q-n}}{\lambda+p+q-1-n} \left(\frac{x_0}{y}\right)^n}{\sum_{n=0}^{q-1} \binom{q-1}{n} \frac{(-1)^{q-1-n}}{\lambda+p+q-1-n} \left(\frac{x_0}{y}\right)^n} \cdot y \tag{21a}$$

$$E(U/Y=y;y>x_0) = \frac{y^{\lambda+p}}{\sum_{n=0}^{q-1}\binom{q-1}{n}\frac{(-1)^{q-1-n}}{\lambda+p+q-1-n}} \int_0^\infty \frac{u^q}{(y+u)^{\lambda+p+q}}\, du$$

$$= \frac{\sum_{n=0}^{q} \frac{\binom{q}{n}(-1)^{q-n}}{\lambda+p+q-1-n}}{\sum_{n=0}^{q-1}\binom{q-1}{n}\frac{(-1)^{q-1-n}}{\lambda+p+q-1-n}} \cdot y \tag{21b}$$

Portanto,

$$E(X/Y=y;\ 0\leq y\leq x_0) = \left[1 + \frac{\sum_{n=0}^{q}\binom{q}{n}\frac{(-1)^{q-n}}{\lambda+p+q-1-n}\left(\frac{x_0}{y}\right)^n}{\sum_{n=0}^{q-1}\binom{q-1}{n}\frac{(-1)^{q-1-n}}{\lambda+p+q-1-n}\left(\frac{x_0}{y}\right)^n} \right] \cdot y \tag{22a}$$

$$E(X/Y = y;\ y>x_0) = \left[1 + \frac{\sum_{n=0}^{q}\binom{q}{n}\frac{(-1)^{q-n}}{\lambda+p+q-1-n}}{\sum_{n=0}^{q-1}\binom{q-1}{n}\frac{(-1)^{q-1-n}}{\lambda+p+q-1-n}} \right] \cdot y \tag{22b}$$

que é proporcional a y.

Variâncias condicionais de X para um dado valor de Y também podem ser obtidas.

Tem-se $Var(X/Y = y) = Var(U/Y = y)$.

Logo, de (20a) e (20b) tem-se

$$\mathrm{Var}(X/Y=y;\ 0 \le y \le x_0) =$$

$$= \left(\frac{\sum_{n=0}^{q+1} \binom{q+1}{n} \frac{(-1)^{q+1-n}}{\lambda+p+q-1-n} \left(\frac{x_0}{y}\right)^n}{\sum_{n=0}^{q-1} \binom{q-1}{n} \frac{(-1)^{q-1-n}}{\lambda+p+q-1-n} \left(\frac{x_0}{y}\right)^n} - \left[\frac{\sum_{n=0}^{q} \binom{q}{n} \frac{(-1)^{q-n}}{\lambda+p+q-1-n} \left(\frac{x_0}{y}\right)^n}{\sum_{n=0}^{q-1} \binom{q-1}{n} \frac{(-1)^{q-1-n}}{\lambda+p+q-1-n} \left(\frac{x_0}{y}\right)^n} \right]^2 \right) y^2 \tag{23a}$$

$$\mathrm{Var}(X/Y = y,\ y > x_0) =$$

$$\left(\frac{\sum_{n=0}^{q+1} \binom{q+1}{n} \frac{(-1)^{n}}{\lambda+p+n-2}}{\sum_{n=0}^{q-1} \binom{q-1}{n} \frac{(-1)^{q-1-n}}{\lambda+p+q-1-n}} - \left[\frac{\sum_{n=0}^{q} \binom{q}{n} \frac{(-1)^{q-n}}{\lambda+p+q-1-n}}{\sum_{n=0}^{q-1} \binom{q-1}{n} \frac{(-1)^{q-1-n}}{\lambda+p+q-1-n}} \right]^2 \right) . y2 \tag{23b}$$

que é proporcional a y^2.

Calculando agora o quociente $\frac{E(Y)}{E(X)}$, observado como um *índice de desonestidade* em situações onde os dados são deliberadamente alterados obtem-se:

$$E(Y) = E(RX) = E(R).E(X) = \frac{p}{p+q} \cdot \frac{\lambda x_0}{\lambda - 1}$$

$$\text{Logo} \quad \delta = \frac{E(Y)}{E(X)} = \frac{p}{p+q} \tag{24}$$

É interessante notar que esse índice independe dos parâmetros (λ e x_0) da distribuição de X, dependendo exclusivamente dos parâmetros da distribuição Beta.

Com (19a) - (24), generalizamos os resultados apresentados por Hartley e Revankar [14].

§3 - Tabela de valores de λ em função de $V(x_0)$ e p, para $n = 1,2,3,4$.

Com o objetivo de fazer um estudo quantitativo sobre o parâmetro λ da distribuição de Pareto, solução de interesse da equação de Euler (E_n), foi considerado nessa equação $n = 1,2,3,4$ e para cada um desses valores foram fixados valores para a probabilidade $V(x_0)$ e para o parâmetro p da distribuição Beta (p,n).

Apresentamos o valor do parâmetro λ e do índice δ de desonestidade, em cada caso.

TABELA DE VALORES DE λ e δ, $\lambda = \lambda(V(x_0), p, n)$ e $\delta = \delta(p,n)$

	n \ $V(x_0)$	0,05	0,10	0,25	0,50	0,75	0,90	0,95	0,99	δ
p = 0,5	1	- 9,5	- 4,5	- 1,5	- 0,5	- 0,1667	- 0,0556	- 0,0263	- 0,0051	0,3333
	2	- 2,9050	- 1,7839	- 0,8028	- 0,3229	- 0,1180	- 0,0408	- 0,0196	- 0,0038	0,2
	3	- 1,9468	- 1,2820	- 0,6273	- 0,2670	- 0,1006	- 0,0353	- 0,0169	- 0,0033	0,1429
	4	- 1,5708	- 1,2386	- 0,6182	- 0,2538	- 0,0930	- 0,0323	- 0,0155	- 0,0030	0,1111
p = 1	1	- 19	- 9	- 3	- 1	- 0,3333	- 0,1111	- 0,0526	- 0,0101	0,5
	2	- 4,8443	- 3	- 1,3723	- 0,5616	- 0,2078	- 0,0723	- 0,0347	- 0,0067	0,3333
	3	- 2,9997	- 2,0001	- 1,0000	- 0,4348	- 0,1663	- 0,0587	- 0,0283	- 0,0055	0,25
	4	- 2,3147	- 1,7958	- 0,9140	- 0,3895	- 0,1466	- 0,0515	- 0,0248	- 0,0048	0,2
p = 1,5	1	- 28,5	- 13,5	- 4,5	- 1,5	- 0,5	- 0,1667	- 0,0790	- 0,0152	0,6
	2	- 6,6758	- 4,1437	- 1,9050	- 0,7839	- 0,2913	- 0,1016	- 0,0488	- 0,0095	0,4286
	3	- 3,9547	- 2,6474	- 1,3333	- 0,5840	- 0,2245	- 0,0795	- 0,0383	- 0,0074	0,3333
	4	- 2,9692	- 2,2854	- 1,1697	- 0,5050	- 0,1921	- 0,0679	- 0,0327	- 0,0064	0,2727
p = 2	1	- 38	- 18	- 6	- 2	- 0,6667	- 0,2222	- 0,1053	- 0,0202	0,6667
	2	- 8,4879	- 5,2613	- 2,4242	- 1,0001	- 0,3723	- 0,1300	- 0,0624	- 0,0121	0,5
	3	- 4,8702	- 3,2689	- 1,6519	- 0,7259	- 0,2798	- 0,0991	- 0,0478	- 0,0093	0,4
	4	- 3,5887	- 2,7500	- 1,4110	- 0,6132	- 0,2346	- 0,0831	- 0,0401	- 0,0078	0,3333
p = 2,5	1	- 47,5	- 22,5	- 7,5	- 2,5	- 0,8333	- 0,2778	- 0,1316	- 0,0253	0,7143
	2	- 10,2432	- 6,3425	- 2,9373	- 1,2132	- 0,4521	- 0,1578	- 0,0758	- 0,0147	0,5556
	3	- 5,7703	- 3,8651	- 1,9619	- 0,8639	- 0,3333	- 0,1183	- 0,0571	- 0,0110	0,4546
	4	- 4,1828	- 2,9139	- 1,6456	- 0,7394	- 0,2812	- 0,0985	- 0,0474	- 0,0092	0,3846
p = 3	1	- 57	- 27	- 9	- 3	- 0,1000	- 0,3333	- 0,1579	- 0,0303	0,75
	2	- 11,9938	- 7,4681	- 3,4468	- 1,4242	- 0,5311	- 0,1856	- 0,0890	- 0,0173	0,6
	3	- 6,6509	- 4,4766	- 2,2689	- 0,9997	- 0,3862	- 0,1370	- 0,0659	- 0,0129	0,5
	4	- 4,7886	- 3,6400	- 1,8729	- 0,8201	- 0,3153	- 0,1121	- 0,0544	- 0,0106	0,4286
p = 3,5	1	- 66,5	- 31,5	- 10,5	- 3,5	- 1,6667	- 0,3889	- 0,1842	- 0,0354	0,7778
	2	- 13,7682	- 8,5591	- 3,9537	- 1,6344	- 0,6093	- 0,2132	- 0,1022	- 0,0198	0,6364
	3	- 7,5534	- 5,0660	- 2,5720	- 1,1352	- 0,4381	- 0,1555	- 0,0753	- 0,0144	0,5385
	4	- 5,3756	- 4,0729	- 2,1041	- 0,9197	- 0,3549	- 0,1262	- 0,0613	- 0,0119	0,4667
p = 4	1	- 76	- 36	- 12	- 4	- 1,3333	- 0,4444	- 0,2105	- 0,0404	0,8
	2	- 15,5063	- 9,6510	- 4,4582	- 1,8443	- 0,6861	- 0,2405	- 0,1155	- 0,0224	0,6667
	3	- 8,4314	- 5,6509	- 2,8702	- 1,2689	- 0,4902	- 0,1740	- 0,0843	- 0,0163	0,5714
	4	- 5,9172	- 4,5404	- 2,3224	- 1,0209	- 0,3927	- 0,1398	- 0,0669	- 0,0135	0,5

§ 4 - COMENTÁRIOS FINAIS

Durante a execução do trabalho tive conhecimento de que Ramachandran (1977) em "On the strong Markov property of the Exponential laws" [31] obteve, com o auxílio da Análise Complexa, o seguinte teorema:

Sejam X e Y variáveis independentes e não negativas tais que $P(X>Y+x) = P(X>Y).P(X>x)$, $\forall\ x > 0$ *e* $P(Y=0)<c=P(X>Y)<1$. *Então a função distribuição de X é da forma* $F(x)=1-H(x)e^{-\lambda x}$, $x>0$, $\lambda>0$, *onde* $H(x+\mu)=H(x)$ *para todos os pontos de crescimento da distribuição* $G(x)$ *de* Y."

Convem observar que X tem distribuição exponencial se e somente se $Y = e^X$ tem distribuição de Pareto.

Claramente, portanto, o citado resultado de Ramachandran responde a questão colocada por Krishnaji [18], modelado na linguagem da distribuição de Pareto.

Outras provas desse resultado foram dadas por Shimizu (1978) [35] usando apenas técnicas da Análise Real e por Huang (1978) [16], via Análise Complexa, com técnicas diferentes das de Ramachandran.

Recentemente (1981), Ferman [10] apresentou uma versão forte multivariada do teorema em questão, na linguagem da distribuição exponencial, qual seja:

"Seja $X=(X_1,X_2,\ldots,X_n)$ um vetor aleatório. Então para qualquer $x = (x_1,\ldots, x_n)$ e $\Delta = (\delta_1,\ldots,\delta_n)$,

$P(X\geq x+\Delta/X\geq\Delta) = P(X\geq x)$, se e somente se $X_0 = \min\limits_{1\leq i\leq n} X_i$ tiver uma distribuição exponencial e for independente de $\{X_i-X_0:\ 1 \leq i \leq n\}$."

Em face desses resultados, desenvolvemos nossos estudos no sentido de investigar, especificamente para o caso em que a variável R tem uma distribuição Beta com parâmetros p e q (p real, positivo e $q \in Z^{++}$), quantidades rela cionadas às variáveis em questão e o comportamento quanti tativo dos parâmetros da solução do problema.

BIBLIOGRAFIA

[1] - AIGNER, D.J., - The estimation of moments for a Pareto distribution subject to both sampling and grouping errors - Review of International Statistical Institute, 38, 210-219, 1970.

[2] - AIGNER, D.J. and GOLDBERG, A.S. - Estimation of Pareto's law from grouped observations - Journal of the American Statistical Association, 65, 712-729 , 1970.

[3] - ARNOLD, B.C. and LAGUNA, L. - On generalized Pareto distributions with applications to income data. - International Studies in Economies, Monograph nº 10, 1977.

[4] - BARLOW, R.E. and PROSCHAN, F. - Mathematical Theory of Reliability - John Wiley and Sons, N.Y., 1965.

[5] - BRYSON, M.C. - Heavy tailed distributions: properties and tests - Technometrics, 16, 1, 61-68, 1974.

[6] - CHAMPERNOWNE. D.G. - The graduate of Income distributions - Econometrica, 20, 591-615, 1952.

[7] - CHAMPERNOWNE, D.G. - A model of income distribution- Economic Journal, 63, 318 - 351, 1953.

[8] - DAVIS, D.J. - An analysis of some failure data-Journal American Statistical Association, 47, 113-150, 1952.

[9] - EPSTEIN, B. and SOBEL, M. - Life Testing - Journal American Statistical Association 48, 485-502, 1953.

[10] - FERMAN, P. - Characterization of multidimensional exponential distribution- Vestinik Moskov Univ., Serv. 1, Mat. Mech. 1981, nº 4, 44-47,85 (Russian English Summary) English translation: Moscow Univ. Math. Bull, 36(1981), nº 4, 54-57, 1981.

[11] - Fisk, R. - The graduation of income distribution - -Econometrica, Vol 29, 2, 171-185, 1961.

[12] - GUMBEL, E.J. - Statistics of Extremes - Columbia University Press, N.Y., 1958.

[13] - HARRISON, A.J. - Inequality of income and the Champernowne distribution - Economics Department, University of Essex, 1974.

[14] - HARTLEY, M.J. and Revankar, N.S. - On the estimation of the Pareto law from under-reported data-Journal of Econometrics 2, 327-341, 1974.

[15] - HINKLEY, D.V, and REVANKAR, N.S. - Estimation of the Pareto law from under-reported data-Journal of Econometrics 5, 1-11, 1974.

[16] - HUANG, J.S. - On a lack of memory property-Statistical Series, 1978 - 84, University of Guelph, 1978.

[17] - KOTZ, S. and JOHNSON, N.L. - Distribution in Statistics: Continuous Univariate Distributions-1 , Houghton-Mifflin, Boston, 1970.

[18] - KRISHNAJI, N. - Characterization of the Pareto distribution through a model of under-reported incomes-Econometrica 38, 251-255, 1970.

[19] - KULLDORF, G. and VÄNNAMAN, K. - Estimation of the location and scale parameters of a Pareto distribution by linear functions of order statistics - Journal of the American Statistical Association 68, 341, 218-227, 1971.

[20] - LÉVY, P. - Calcul des Probabilités, Paris: Gauthier Villars, 1925.

[21] - LYDALL, H.L. - Econometrica 27, 110-115, 1959.

[22] - LWIN, T. - Estimation of the tail of the Paretian law - Skand Aktuar Tidskr, 1972.

[23] - MALIK, H.J. - Estimation of the parameters of the Pareto distribution - Metrika 15, 126-132, 1970.

[24] - MANDELBROT, B. - The Pareto Levy law and the distribution of income - International Economic Review, 1, 79-106, 1960.

[25] - MANDELBROT, B. - New methods in statistical economics- - Bulletin of the Internation Statistical Institute 40, Book 2, 699-721, 1964.

[26] - MARSAGLIA, G. and TUBILLA, A. - Annals of Probability- 3,2, 353-354, 1975.

[27] - ORD, J.K. - Families of frequency distributions - Griffin, London, p. 43, 1972.

[28] - ORD, J.K. - Statistical Models for Personal income distributions-Statistical Distributions in Scientific, 2, 151-158, 1975.

[29] - PARETO, V. - Cours d' Economie Politique (Lausanne), 1897.

[30] - QUANDT, R.E. - Old and new methods of estimation and the Pareto distribution-Metrika, 10, 55-82, 1966.

[31] - RAMACHANDRAN, B. - On the strong Markov property of the exponential laws - Colloquim on the Methods of Complex Analysis in the Theory of Probability and Statistics - Debrecen, Hungary, 1977.

[32] - REVANKAR, N.S., HARTLEY, M.J. and PAGANO, M. - A Characterization of the Pareto distribution-Annals of Statistics 2, 3, 599-602, 1974.

[33] - RUTHERFORD, R.S.G. - Econometrica, 23, 277-294, 1955.

[34] - SARGAN, J.D. - Econometrica 25, 568-590, 1957.

[35] - SHIMIZU, R. - Solution to a functional equations and its application to some characterization problems - Sankhyã: The Indian Journal of Statistics 40, séries A, 4, 319-332, 1978.

[36] - SIMON, H.A. - On a class of skew distributions functions - Biometrica 52, 425 - 440, 1955.

[37] - STIGHITZ, J.E. - Econometrica 37, 382-397, 1969.

[38] - WOLD, H.O.A. and WHITTLE, P. - Econometrica 25, 591-595, 1957.

[39] - ZIPF, G.K. - Human Behavior and the Principle of Least Effort - Addison - Wesley Press, 1949.