SPOLM 2008

ISSN 2175-6295 Rio de Janeiro- Brasil, 05 e 06 de agosto de 2008.

ESTIMATIVA BOOTSTRAP PARA O ENVIEZAMENTO, ERRO-PADRÃO E INTERVALO DE CONFIANÇA DO COEFICIENTE DE ELASTICIDADE DA CURVA DE PARETO

Giovani Glaucio de Oliveira Costa
Universidade Federal Rural do Rio de Janeiro
Instituto Multidisciplinar
Rua Professor Paris S/N. Centro. Nova Iguaçu. CEP: 26221-150. Rio de Janeiro. Brasil
giovaniglaucio@ufrrj.br

Resumo:

A Curva de Pareto ou “Curva ABC” tem sido aplicada como critério de tomadas de decisão em várias áreas do conhecimento, por empresas e instituições de vários ramos e em todo o mundo, isto é, como instrumento de modelagem de regularidade estatística para diversos fenômenos da natureza. Entretanto, apesar da Curva de Pareto ser utilizada com grande freqüência, o teste de sua significância ou aderência é pouco realizado, já que o desconhecimento da distribuição por amostragem do seu coeficiente angular(ou de sua elasticidade), torna inviável fazer acompanhar as estimativas do respectivo erro padrão, já para não falar na construção dos intervalos de confiança. Este trabalho tem a iniciativa preliminar de obter uma distribuição por amostragem empírica para a elasticidade da “Curva 80/20”, utilizando para esta finalidade a técnica de reamostragem bootstrap e com isso obter o viés e o erro padrão das estimativas da elasticidade da potência de pareto, possibilitando a construção de intervalos de confiança e testes de significância do ajustamento, minimizando assim a insuficiência de testes de significância para esta importante lei natural.

Palavras-chaves: curva de pareto, elasticidade, bootstrap, intervalos de confiança, teste de significância.

Abstract:

The Curve of Pareto or “ ABC Curve” has been applied as criterion of taking of decision in some areas of the knowledge, for companies and institutions of some branches and in the whole world that is, as instrument of regularity modeling statistics for many phenomena’s nature. However, despite the Curve of Pareto being used with great frequency, the test of its significance or tack little is carried through, since the unfamiliarity of the distribution for sampling of its regression coefficient (or of its elasticity), becomes impracticable to make to follow the estimates of the respective error standard, already not to mention in the construction of the confidence intervals. This work has the preliminary initiative to get a distribution for empirical sampling for the elasticity of “80/20 Curve”, using for this purpose the technique of bootstrap and with this to get the bias and the error standard of the estimates of the elasticity of the power of pareto, making possible construction confidence intervals and tests of significance of the adjustment, thus minimizing the insufficience of tests of significance for this important natural law.

Keywords: curve of pareto, elasticity, bootstrap, confidence interval, tests of significance.

1-Introdução

A Curva de Pareto ou a Curva ABC foi desenvolvida por Vilfredo Pareto, sociólogo e matemático italiano (1848-1923) para comparar duas curvas homogêneas e as conclusões que disso obteve resultaram em importantes contribuições ao estudo da economia e também da sociologia.

Especificamente no campo da distribuição da riqueza e a análise das eleições individuais, Pareto observou que 80% do território italiano pertenciam a 20% da sua população. Mais tarde,ele notou que essa regra podia ser utilizada para todos os aspectos da vida moderna.

Em economia, dá-se o nome de Princípio de Pareto ou Lei 80/20 a um fenômeno que afeta a distribuição do relacionamento entre produtos e fatores de um sistema qualquer.Vilfredo foi um dos primeiros a notar que 80% das riquezas do mundo estavam concentradas nas mãos de apenas uma elite de 20% ou menos da população. Daí o nome “Lei 80/20”.

O Princípio de Pareto que foi aparentemente descoberto em economia não se restringe à esse campo e tem sido aplicado aos mais diversos setores do conhecimento humano.

O economista Vilfredo Pareto sugeriu que a distribuição de renda de uma população seguia uma lei de potência simples, ou seja, que o número de pessoas que ganham uma certa quantia mensal, diminui conforme essa quantia vai aumentando. Essa lei é caracterizada por um índice, conhecido como o "Índice Pareto", que é pequeno se a renda é mal distribuída pela população, e grande se a distribuição de renda for mais uniforme. Entretanto, o Princípio de Pareto funciona apenas para aproximadamente 1% da população de melhor renda. Outro economista, chamado Gibrat, posteriormente sugeriu que os 99% restantes da população têm uma renda que segue uma curva de distribuição do tipo log-normal. Os ajustes das curvas de renda com o modelo de Gibrat levam ao "índice Gibrat", que também é pequeno para situações com rendas mal distribuídas.

Recentemente, um pesquisador japonês, Wataru Souma, do *ATR Human Information Science Laboratory,* buscou juntar esses dois modelos de modo empirico. Para isso, ele realizou uma análise detalhada de mais de 80% da população japonesa economicamente ativa. Ele analisou detalhadamente resultados de mais de 51 milhões de pessoas, entre 1887 e 1998. Para realizar esse estudo, ele empregou algumas técnicas freqüentemente usadas em Física da Matéria Condensada, usadas para modelar sistemas de "muitos-corpos". O resultado final é uma curva híbrida, que consiste em uma curva do tipo log-normal curve combinada com uma cauda que segue uma lei do tipo potência. Essa fórmula descreve adequadamente os dados dessa população, em todos os períodos investigados.

Desse modo, o pesquisador conseguiu verificar como a distribuição de renda evoluiu ao longo dos anos no Japão. Apesar de ainda não tentar explicar porque a distribuição de renda segue esse tipo de lei, o pesquisador acredita que essa fórmula empírica tem uma validade geral, e que poderá descrever a distribuição de renda em qualquer sociedade, em qualquer ponto da história. Se de fato isso ocorrer, esse resultado pode ser muito importante, pois uma descrição acurada da distribuição de salários é fundamental em várias áreas da economia, incluindo estudos de desigualdades sociais e pobreza.

A *IBM* descobriu que em 80% do tempo, os computadores estão executando apenas uma fração de 20% dos programas, qualquer que seja esse programa. Assim nasceu a idéia de criar processadores *RISC*, que são processadores projetados para serem rápidos e eficientes apenas na pequena porção de 20% de instruções mais usadas.

Edward Deming e Joseph Juran aplicaram a Princípio de Pareto para reduzir o número de defeitos nas linhas de montagens e foram grandes colaboradores para que o atual parque industrial japonês atingisse o nível atual de qualidade e produtividade. Nas palavras de Juran *"Uma pequena fração de 20% de práticas errôneas são responsáveis por cerca de 80% dos desperdícios e da baixa qualidade na produção e isso vale para qualquer empresa ou indústria"*.

Nasceu assim as bases para a Qualidade Total Japonesa. A idéia central é identificar os 20% de causas que geram 80% dos defeitos e focalizar energias nesse pequeno grupo de causas, ao invés de tentar reduzir 100% das causas, coisa que fatalmente tornaria os custos inevitavelmente inviáveis.

É patente que a Curva de Pareto tem sido aplicada como critério de tomadas de decisão em várias áreas do conhecimento, por empresas e instituições de vários ramos, isto é, como instrumento de modelagem de regularidade estatística para diversos de fenômenos da natureza. Portanto, a Curva de Pareto é utilizada com grande freqüência, mas sempre à nível descritivo. O desconhecimento da distribuição por amostragem (mesmo quando a distribuição da população é conhecida) do seu coeficiente angular (ou sua elasticidade) torna inviável fazer acompanhar as estimativas do respectivo erro padrão, para não falar na construção dos intervalos de confiança ou na realização de testes de significância. Este trabalho tem a iniciativa preliminar de obter uma distribuição por amostragem empírica para a elasticidade da Curva de Pareto e com isso obter o viés, o erro padrão e construir intervalos de confiança, minimizando assim a insuficiência de testes de significância para esta importante lei natural.

2-Objetivos da Pesquisa

Este estudo tem os seguintes objetivos:

- Gerar uma estimativa empírica, real , da distribuição por amostragem da variável aleatória 'elasticidade da Curva de Pareto';
- Aplicar metodologias *CIS*(*COMPUTER INTENSIVE STATISTICS*) para obtenção das distribuições por amostragem;
- Empregar o seguinte método *CIS*: bootstrap;
- Especificar o viés e o erro padrão da estatística 'elasticidade da Curva de Pareto'
- Construir intervalos de confiança e realizar testes de significância para aderência do ajustamento;
- Aplicar os procedimentos expostos a um estudo de caso real.

3-Justificativa da Utilização de Computação Estatística Intensiva

A abaixo se relaciona os motivos para aplicação de metodologias CIS:

- Não se conhecem com a precisão necessária os parâmetros característicos teóricos da distribuição por amostragem da variável aleatória ‘elasticidade da Curva de Pareto’;
- O estudo da distribuição por amostragem da ‘elasticidade da Curva de Pareto’ assume caráter instável, o que dificulta a especificação teórica do erro-padrão e do enviezamento desta estatística;
- A distribuição por amostragem fornece um modo direto para conhecer as “estimativas da elasticidade da Curva de Pareto”;
- O bootstrap permite obter, então, de forma experimental e empírica as distribuições por amostragem da estatística em foco;
- O método bootstrap permite ladear a insuficiência da teoria da amostragem que se faz sentir quando se aplica a Curva de Pareto em diversas situações práticas.

4-Viabilidade Operacional da Metodologia *CIS*

- O desenvolvimento recente da informática tem permitido que técnicas de reamostragem como o bootstrap possam ser operacionalizadas de maneira rápida e precisa;
- Recorrendo ‘a computação pesada’ consegue-se solucionar problemas para as quais a teoria da estatística tradicional não encontra solução.

5-Resumo Teórico de Reamostragem

O tipo de estatística não-paramétrica que foi ensinado no passado desempenhou um importante papel na análise de dados que não são variáveis contínuas, em escala nominal ou ordinal, e, portanto, não podem empregar a distribuição normal de probabilidade para fazer estimativas de parâmetros e de intervalo de confiança. Mas existe uma nova perspectiva sobre estimação não-paramétrica que também se relaciona com estimação de parâmetros e de intervalo de confiança para variáveis no mínimo em escala intervalar.

Com isso , não se tem que assumir que o intervalo de confiança para um parâmetro segue a distribuição normal. Pode-se até mesmo gerar intervalos de confiança para parâmetros como a mediana, o que geralmente é difícil de avaliar com as técnicas de inferência paramétrica tradicionais.

Essa abordagem não-paramétrica é conhecida como reamostragem e tem conquistado apoio como uma alternativa aos métodos clássicos de inferência paramétrica.

A reamostragem descarta a distribuição por amostragem assumida de uma estatística e calcula uma distribuição empírica – a real distribuição da estatística ao longo de centenas ou milhares de amostras.

Com a reamostragem, não se tem que confiar na distribuição assumida nem se tem que ser cuidadoso quanto à violação de uma das suposições inerentes. Pode-se calcular uma real distribuição de estatísticas da amostra e pode-se agora ver onde o 95 ou o 99 percentil estão realmente, acreditando-se que a amostra original seja confiável.

Mas de onde vêm as múltiplas amostras? É necessário reunir amostras separadas, aumentando sensivelmente o custo de coleta de dados? Ao longo dos anos estatísticos desenvolveram diversos procedimentos para criar as múltiplas amostras necessárias para a reamostragem *a partir da amostra original*.

Agora uma amostra pode gerar um grande número de outras amostras que podem ser empregadas para gerar a distribuição amostral empírica de uma estatística de interesse.

Reamostragem , contudo , não usa a distribuição de probabilidades assumida , mas ao invés disso ela calcula uma distribuição empírica de estatísticas estimadas. Criando múltiplas amostras da amostra original, a reamostragem agora precisa apenas do poder computacional para estimar um valor de uma estatística para cada amostra. Logo que eles estejam todos calculados, pode-se realizar o teste de normalidade dos valores e até mesmo construir intervalos de confiança e realizar testes de hipóteses.

A reamostragem engloba diversos métodos. Para este trabalho, se estudará e aplicará o b*ootstrap*.

Uma diferença chave entre os vários métodos de reamostragem é se as amostras são extraídas com ou sem reposição. A amostragem com reposição obtém uma observação a partir da amostra e então a coloca de volta na amostra para possivelmente ser usada novamente. A amostragem sem reposição obtém observações da amostra, mas uma vez obtidas eles não estão mais disponíveis.

O verdadeiro poder da reamostragem vem de amostragem **com reposição.** Pesquisas têm mostrado que esse método fornece estimativas diretas dos intervalos de confiança, apesar de ter havido avanços nos métodos simples para obtenção dos intervalos de confiança.

O método *bootstrap* obtém sua amostra via amostragem com reposição da amostra original. A chave é a substituição das observações após a amostragem, o que permite ao pesquisador criar tantas amostras quanto necessárias e jamais se preocupar quanto à duplicação de amostras, exceto quando isso acontecer ao acaso. Cada amostra pode ser analisada independentemente e os resultados compilados ao longo da amostra. Por exemplo, a melhor estimativa da média é exatamente a média de todas as médias estimadas ao longo das amostras.

O intervalo de confiança também pode ser diretamente calculado. As duas abordagens mais simples :

1) Calculam o erro padrão simplesmente como o desvio padrão das estimativas estimadas;
2) Literalmente ordenam as estimativas e definem os valores que contém os 5% extremos (ou 1%) dos valores estimados.

Matematicamente a obtenção da amostra bootstrap e suas estimativa do erro padrão é obtida da seguinte maneira:

Seja uma amostra original e a estatística de interesse abaixo:

$\boldsymbol{x = \{x_1, x_2, x_3, \ldots, x_{n-1}, x_n\}}$.

$$\hat{\theta} = F(x)$$

(1º) Geram-se as amostras *bootstrap* $x_{(1)}, x_{(2)}, x_{(3)}, \ldots, x_{(n^*)}$ com reposição de x.

(2º) Calculam-se as estimativas da estatística de interesse:

$$\hat{\theta}_{(b)} = F[x_{(b)}] \;, \; b=1, \ldots, B$$

(3º) Calcula-se o erro padrão *bootstrap*, $\hat{S}_{boot}$, dado por:

$$\left\{ \hat{S}_{boot} = \frac{1}{B-1} \cdot \sum_{b=1}^{B} [\hat{\theta}_b - \hat{\theta}_{(*)}]^2 \right\}^{1/2} \quad \text{, sendo}$$

$$\hat{\theta}_{(*)} = \frac{\sum_{b=1}^{B} \theta_{(b)}}{B}$$

Apesar de procedimentos de reamostragem não serem restritos por quaisquer suposições paramétricas, eles ainda têm certas limitações :

1) A amostra deve ser grande o bastante e obtida (a princípio aleatoriamente) de forma a ser representativa da população completa. Técnicas de reamostragem não podem conter quaisquer enviezamentos que traga como conseqüência uma amostra não representativa;
2) Métodos paramétricos são melhores em muitos casos para fazer estimativas pontuais. Os procedimentos de reamostragem podem completar as estimativas pontuais de métodos paramétricos fornecendo as estimativas de intervalos de confiança;

3) As técnicas de reamostragem não são adequadas para identificar parâmetros que têm um domínio amostral muito estreito, como os valores mínimos e máximos. A reamostragem funciona melhor quando a distribuição inteira é considerada para obter o parâmetro em análise.

6- Modelagem da Curva de Pareto

Após analisar várias distribuições de rendas de vários países com conjunturas iguais ou diferentes, o professor Vilfredo Pareto concluiu a existência de certa dependência funcional entre os níveis de renda de cada estrato de distribuição e o número de pessoas economicamente ativas, deduzindo o caráter de estabilidade associada à distribuição de renda.

A constatação permitiu a Pareto imaginar que seria uma aproximação razoável da realidade representar por uma hipérbole (ver expressão 1) o modelo analítico associado á distribuição de renda.

$$y= \frac{A}{(x - a)^2} \qquad (1)$$

Onde:

a= menor renda observada

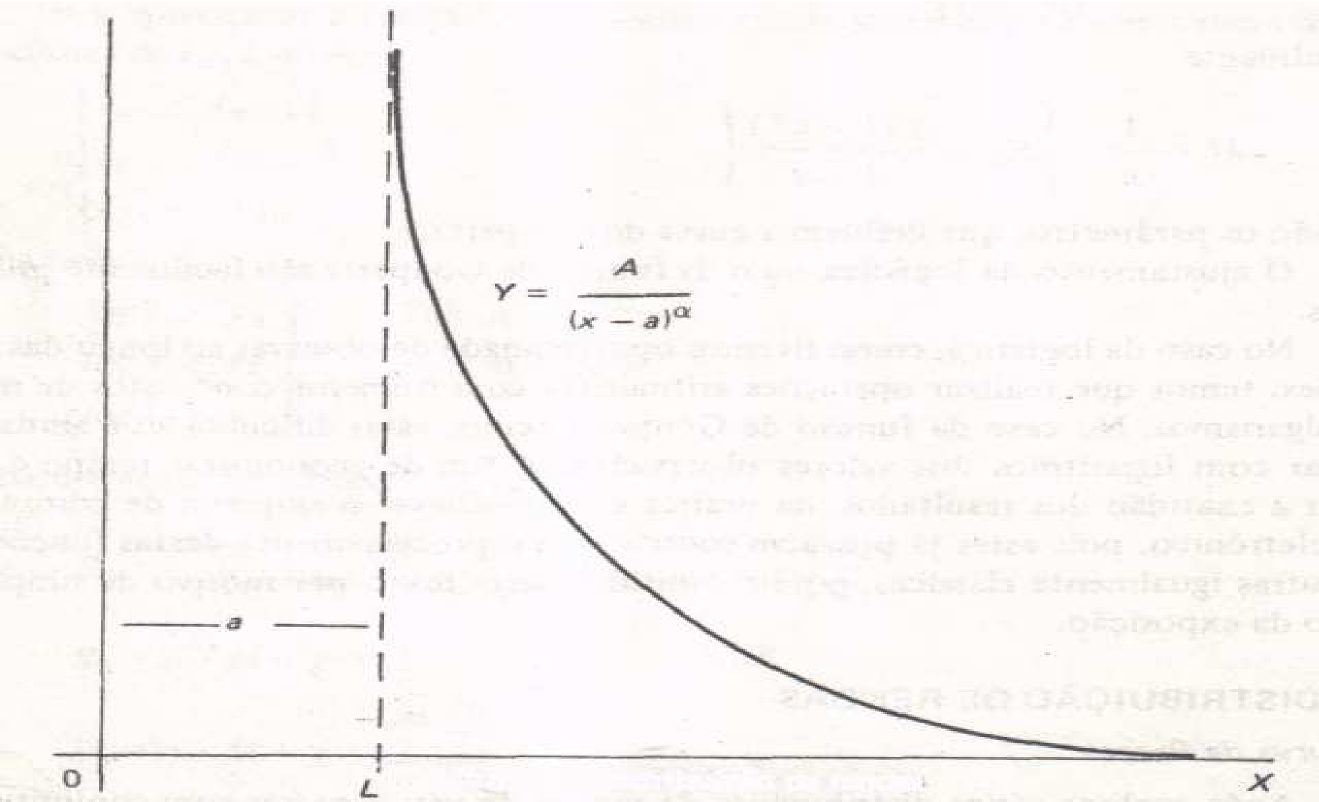
A=constante de proporcionalidade

x=um parâmetro indicativo do grau de concentração de renda

Chamando de x o eixo das abscissas e nele marcando as rendas e no eixo das ordenadas o mínimo de pessoas cujas rendas sejam iguais ou superiores a x, têm-se a Curva de Pareto, representada no gráfico 1.

Gráfico 1

Curva de Distribuição de Renda



Tem-se facilmente que:

lim y → α

x→a

lim y → 0

x→α

Então, a hipérbole considerada admite duas assíntotas: uma para x=a e outra para y=0.Por uma translação de eixo dos y´s, fazendo a origem deslocar-se até a(menor renda observada) a expressão 1 da Curva de Pareto reduz-se a potência simples:

$$y=A.x^{-\alpha} \qquad (2)$$

Então,

$$\log y = \log A - \alpha \cdot \log x \quad (3)$$

A expressão 3 traçada sobre um gráfico duplo-logaritmo é uma reta do tipo:

$z = w - \alpha r$, de declividade negativa e coeficiente angular($- \alpha$) (4)

Pareto, que concluiu estar em frente a uma lei natural afirma que α indica o grau de concavidade da curva.

O ajustamento da Curva de Pareto é imediato, utilizando o método de mínimos quadrados ordinários à expressão (4):

$$\begin{cases} \sum \log y = n \log A - \alpha \sum \log x \\ \sum \log x \cdot \log y = \sum \log A \log x - \alpha \cdot \sum (\log x)^2 \end{cases} \quad (5)$$

A solução do sistema da expressão 5 dá os valores de A e de α.

A expressão abaixo representa a variação proporcional de y, associada a uma variação proporcional de x, isto é, à elasticidade da curva de potência.

$$E = \frac{\frac{\partial y}{y}}{\frac{\partial x}{x}} = (\partial y / \partial x) \cdot (x/y) \quad (6)$$

Então,

$$-\alpha = \frac{\partial(\log y)}{\partial(\log x)} \qquad (7)$$

A expressão 7 mostra que $-\alpha$ corresponde à elasticidade da função de distribuição de rendas. É possível, então, interpretar o parâmetro α como a elasticidade do decréscimo do número de pessoas que se verifica quando se passa de uma para outra classe de renda mais elevada.

7-Estudo de Caso

O presente estudo de caso utiliza uma amostra de trabalhadores com idade entre 25 e 64 anos, residentes nas áreas urbanas do Brasil, extraída da PNAD (Pesquisa Nacional por Amostra de Domicílios) de 2002. Os dados correspondem a uma amostra aleatória de 10% das observações da PNAD.

A PNAD é uma pesquisa anual feita pelo IBGE representativa de toda a população brasileira. Em cada ano são entrevistados em torno de 330.000 indivíduos em cerca de 100.000 domicílios. A amostra utilizada envolveu 10.014 indivíduos da amostra original da PNAD.

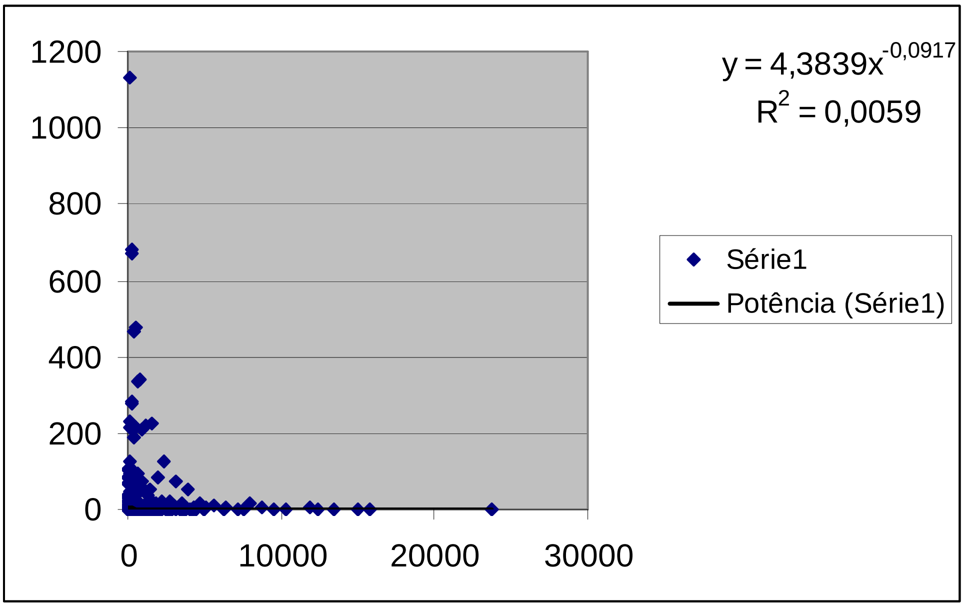
A pesquisa contém uma série de informações sobre educação, mercado de trabalho, indicadores sociais, etc. e o objetivo deste trabalho é ajustar segundo uma de Curva de Pareto a variável rendimento por pessoa (em R$ de 1999), obter a distribuição por amostragem empírica da elasticidade desta curva, seu viés , seu erro padrão, construir intervalos de confiança e realizar o teste de significância para a existência de aderência.

7.1-Ajustamento da Curva de Pareto aos Dados

Chamando de x o eixo das abscissas e nele marcando a variável rendimento por pessoa (em R$ de 1999) e no eixo das ordenadas o mínimo de pessoas cujas rendas sejam iguais ou superiores a x, têm-se a Curva de Pareto, representada no gráfico 2.

Gráfico 2

Curva de Pareto de Dados de Rendimentos de 10% de Pessoas da PNAD



A tabela 1 resume os parâmetros característicos associados à elasticidade da Curva de Pareto obtidos pelas expressões teóricas do coeficiente angular do modelo e do erro-padrão.

Tabela 1

Coeficiente Angular α Obtido pelo Método Tradicional e seus Parâmetros Característicos

Parâmetro	**Valor de *α***	**Erro-padrão *α***
Coeficiente angular α	-0,09179	0,04914

7.2-Ajustamento Bootstraps da Curva de Pareto aos Dados

A tabela 2 resume os parâmetros característicos associados à elasticidade da Curva de Pareto obtidos junto à distribuição por amostragem empírica bootstraps: de cada 1000 amostra gerada foi tirado o valor de α e desses milhares de valores de α foram calculados a sua esperança(valor médio de α) e o seu desvio-padrão(erro-padrão de α). O viés é a diferença entre o valor médio de α ***e*** o valor de α..

Tabela 2

Coeficiente Angular α Obtido pelo Método Bootstrap seus Parâmetros Característicos

Parâmetro	Valor Médio de α	Viés de α	Erro-padrão de α
Coeficiente angular α	-0,09074	0,00105	0,04824

Conhecidos o valor médio da estatística α e seu erro-padrão, pode-se, então, construir intervalos de confiança e realizar testes de hipóteses.

Tabela 3

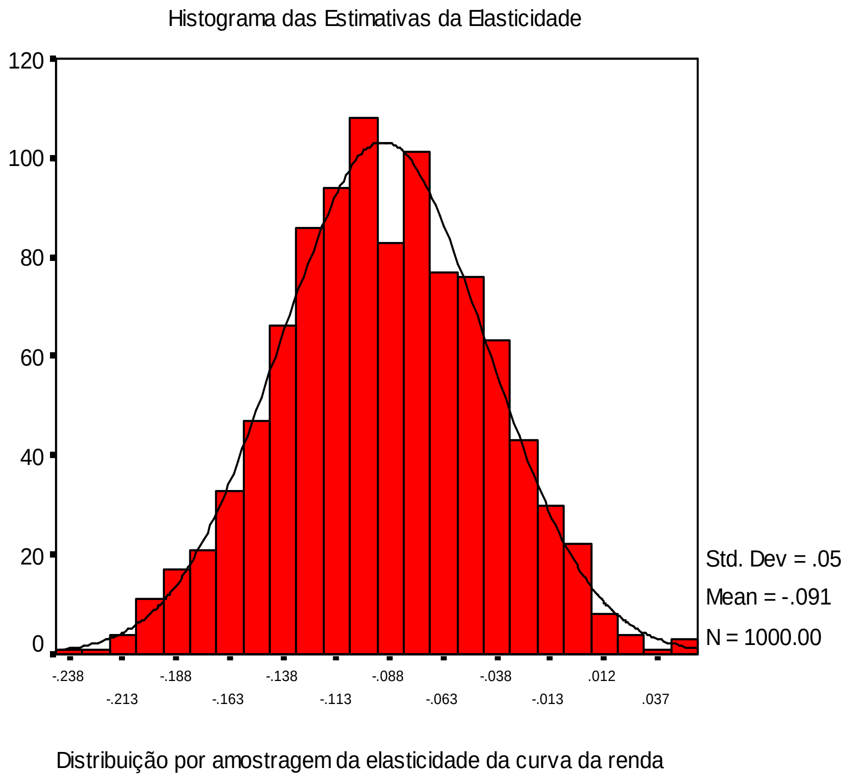
Intervalo de Confiança Bootstrap de 95% para o Coeficiente Angular α

Intervalos de Confiança	Limites de Confiança	
	Limite Inferior	Limite Superior
Percentílico	-0,18721	0,00149
Normal	-0,18646	0,00288

7.3-Teste de Normalidade das Estimativas da Distribuição por Amostragem da Elasticidade da Curva de Pareto

Abaixo se relaciona os gráficos: histograma, gráfico de ajustamento anormal e o gráfico de ajustamento dos resíduos para as estimativas da elasticidade obtidas pelo método bootstraps.

Gráfico 3



Percebe-se alto grau de simetria na distribuição amostral de α. O que sugere um modelo de estimativas normalmente distribuída.

Gráfico 4

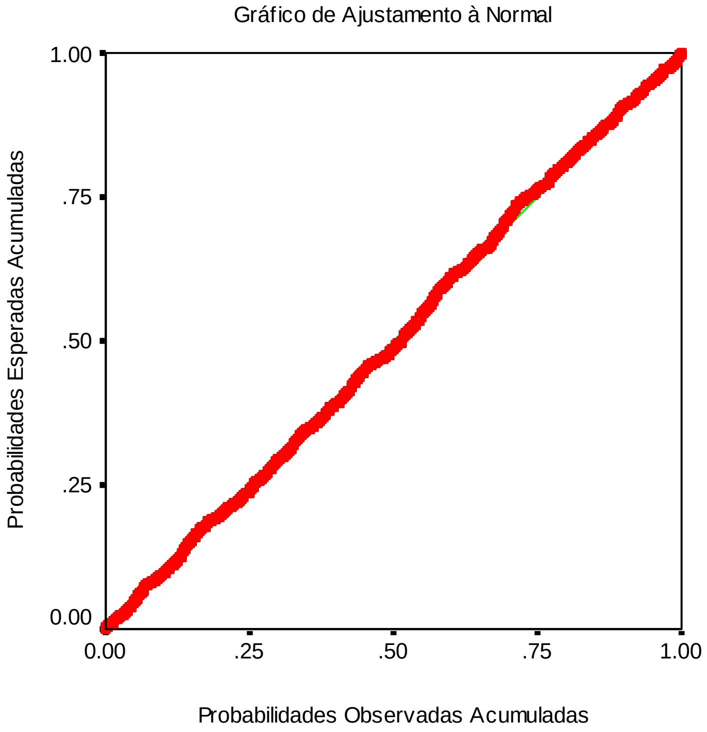
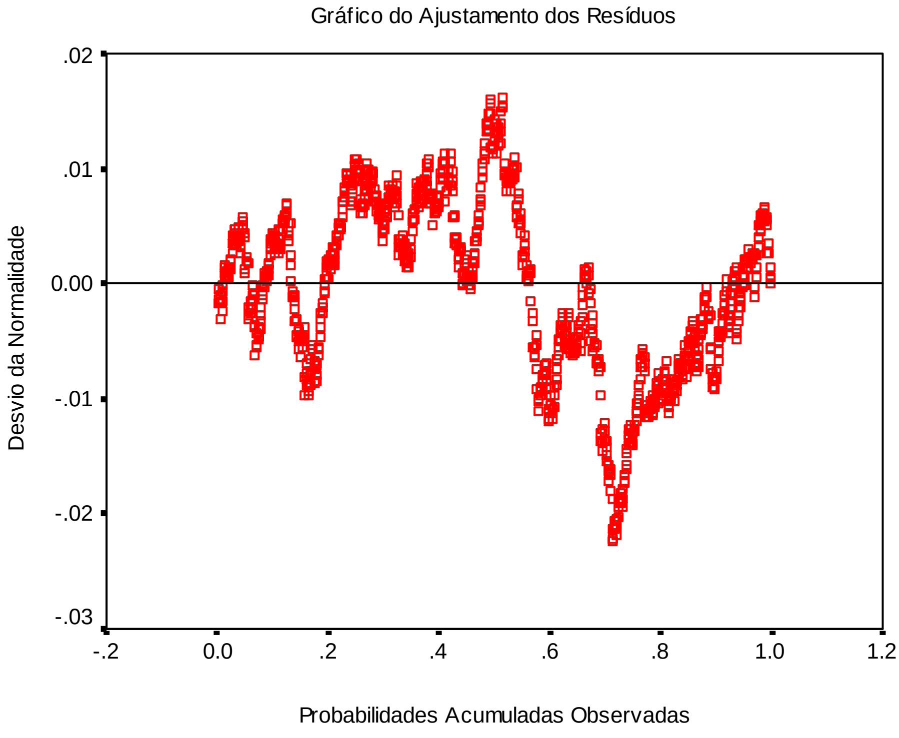


Gráfico 5



A tabela 1 contém a estimativa elasticidade e seu erro-padrão pelo ajustamento tradicional da Curva ABC. O erro padrão foi de 0,04914, relativamente pequeno, o que denota precisa da estimativa da elasticidade.A tabela 2 apresenta a estimativa da elasticidade por reamostragem bootstrap. Por este método, o valor médio estimado da elasticidade é de -0,09074 , bem próximo do valor que seria obtido pelo método tradicional o que resultou em um viés de somente 0,00105. O erro padrão da elasticidade bootstrap foi 0,04824 , bem próximo do obtido pelo método tradicional, porém levemente menor, o que denota que a esperança das estimativas da elasticidade bootstrap é um pouco mais precisa que o valor tradicional do coeficiente angular.

A tabela 3 mostra os intervalos de confiança bootstrap à nível de 5% de significância pelos métodos percentílico e pelo método baseado na normalidade das estimativas geradas.

Pelo gráfico 3, histograma das estimativas da elasticidade, observa-se que a estatística do coeficiente angular se assemelha bastante à Curva Normal:

O gráfico 4 apresenta o gráfico de ajustamento à normal, onde se observa um ajustamento perfeito da amostra à função de distribuição de probabilidade normal.O gráfico 5 inclui um gráfico de ajustamento dos resíduos. Se a amostra é perfeitamente normal, os resíduos distribuem-se segundo uma faixa horizontal em torno de zero, sem denotar qualquer padrão de distribuição. No estudo de caso considerado, é nítido o respeito a este pressuposto.

Para minimizar a probabilidade de erros de decisão nos testes de significância, seria mais prudente “conectar” os intervalos de confiança construídos a um modelo de distribuição de probabilidades das estimativas investigadas. A Curva Normal por ser a mais freqüente distribuição ajustável à descrição de variáveis aleatórias e por otimizar o poder de um teste de hipótese será utilizada como prova de aderência aos dados.

Diante de tudo que foi exposto em parágrafos acima, o intervalo de confiança bootstrap baseado na Curva Normal parece ser o mais indicado para estimar a elasticidade do rendimento de brasileiros: ***[-0,18646 ; 0,00288].*** Existe uma probabilidade de 95% deste intervalo conter a elasticidade da lei natural do comportamento dos rendimentos dos brasileiros.

7.3-Teste de Significância da Elasticidade da Curva de Pareto ou Teste da Existência de Aderência

Utilizando o intervalo de confiança bootstrap da elasticidade de Pareto, verifica-se que o valor zero está contido no intervalo de confiança, o que sugere, ao nível de 5% de significância, que o valor α para a elasticidade obtido para a curva pode não ser significante estatisticamente e que a amostra disponível do rendimento dos brasileiros pode não se aderir tolerantemente à Curva de Pareto, isto é, resulta em alto erro de amostragem.

8-Conclusão

O objetivo deste artigo foi gerar uma estimativa empírica, real, da distribuição por amostragem da variável aleatória 'elasticidade da Curva de Pareto', aplicando metodologias *CIS* (*COMPUTER INTENSIVE STATISTICS*), a técnica bootstrap. Com a distribuição por amostragem empírica pode-se especificar o viés e o erro padrão da estatística, possibilitando a construção de intervalos de confiança e realizando testes de significância para aderência do ajustamento.

Os resultados foram aplicados a um estudo de caso retirado da PNAD de 2000, onde se verificava a modelagem de pareto para a distribuição de renda de uma amostra de brasileiros.

A realização da pesquisa demonstrou que o ajustamento de dados de renda à curva de pareto utilizando métodos de reamostragem, bootstrap, evidencia resultados bem satisfatórios em comparação com o método tradicional e possibilita realizar testes de hipóteses aos coeficientes de regressão, como foi exemplificado com o da elasticidade, o que raramente é realizado atualmente quando se aplica a Curva ABC nas instituições e empresas.

Os resultados a que chegou deixam campo aberto para outras pesquisas na mesma área, como construção de intervalos de confiança e testes de significância para os coeficientes das curvas logística, gonpertz e de modelos de regressão múltipla.

9-Bibliografia

[1]Afifi, A. A. e **Clark**, V. (1984).Computer – Aided Multivariate Analysis. Lifetime Learning Publications. Belm. California.

[2]Anderson, T.W. (1984).An Introduction to Multivariate Statistical Analysis. 2ed.New York : John Wiley & Sons.

[3]Cazar, R. A. (2003).An Exercise on Chemometrics for a Quantitative Analysis Course.Madison: Journal of Chemical Education.

[4]Chatfield, C. e **Collins,** A. J. (1980).Introduction to Multivariate Analysis. Chapman and Hall. New York.

[5]Cliff, N., e **Hamburge**, C. D. (1967).The Study of Sampling Errors in Factor Analysis by Means of Artificial Experiments. Psychological Bulletin 68: 430-45.

[6] Costa, Giovani Glaucio de O. (2003). Busca de Fatores Associados à Prática de Atos Infracionais por Parte de Adolescentes no Estado do Rio de Janeiro: Um Estudo Preliminar, Estudo Orientado, PUC-RIO.

[7]David, A.Aaker , **Kumar**, V;**George**,S. Day.(1984).Marketing Research.

Dillon, W. R. e **Goldstein**, M.(1984).Multivariate Analysis : Methods and Applications . New York : John Wiley & Sons.

[8]Efron,B(1979).Bootstrap Methods:Another Look at the Jackknife, The Annals of Statistic,7,1-26.

[9]Efron,B.(1980).Computer Intensive Methods in Statistics" in Some Recent Advance in Statistic, Ed. J. Tiago de Oliveira e B.Epstein , Academia das Ciências de Lisboa, Lisboa.

[10]Efron,B.(1982).The Jackknife, the Bootstrap , and other Resampling Methods, CBNS 38,SIAM-NSF

[11] Ferreira, D. F. Análise Multivariada. Minas Gerais : Universidade Federal de Lavras.

[12]Hair, J. F. Jr. ; **Anderson,** R.E. ; **Tathan**, R. L. e **Black**, W. C.(2005). Trad. **Sant´Anna**,Adonai Schlup ; **Neto**, Anselmo Chaves. Análise Multivariada de Dados. 5. ed. Porto Alegre : Bookman.

[13]Hair, J. F. Jr. ; **Anderson,** R.E. ; **Tathan**, R. L. e **Black**, W. C.(1998). Multivariate Data Analysis. 5th ed. Upper Saddle River : Prentice Hall.

[14]Harman, Harry H.(1967).Modern Factor Analysis . 2 ed. Chicago : University of Chicago .

[15]Hawkins, D.M., Topics in Multivariate Analysis.Cambridge University Press: Cambridge.

[16]Johnson, D. E. (1998).Applied Multivariate Methods for Data Analysis. Pacific Grove: Duxbury Press.

[17]Johnson, R. A .e **Wichern** , D.W .(1998).Applied Multivariate Statistical Analysis. 4ed. Upper Saddle River: Prentice Hall.