

Ciência e Natura

ISSN: 0100-8307

cienciaenaturarevista@gmail.com

Universidade Federal de Santa Maria

Brasil

Mendonça Souza, Francisca; de Oliveira Silva, Pollyanna Kelly; Bourguignon Pereira, Marcelo; de Souza, Leandro Carlos

Avaliação do desempenho de estimadores dos parâmetros da distribuição Pareto corrigidos por Bootstrap

Ciência e Natura, vol. 34, núm. 1, 2012, pp. 7-22

Universidade Federal de Santa Maria

Santa Maria, Brasil

Disponível em: <http://www.redalyc.org/articulo.oa?id=467546170005>

- Como citar este artigo
- Número completo
- Mais artigos
- Home da revista no Redalyc

# AVALIAÇÃO DO DESEMPENHO DE ESTIMADORES DOS PARÂMETROS DA DISTRIBUIÇÃO PARETO CORRIGIDOS POR BOOTSTRAP

Francisca Mendonça Souza, Pollyanna Kelly de Oliveira Silva,  
Marcelo Bourguignon Pereira, Leandro Carlos de Souza

Universidade Federal de Pernambuco; Recife, PE  
*e-mail:* kikamatcom@gmail.com

## Resumo

Este trabalho avalia, via simulação de Monte Carlo, o desempenho de estimadores dos parâmetros  $\alpha$  e  $k$  que indexam a distribuição Pareto. A partir de amostras geradas por essa distribuição, foram calculadas estimativas pontuais dos parâmetros utilizando o método de máxima verossimilhança e método dos momentos, e suas respectivas versões corrigidas por *bootstrap*. As simulações foram consideradas em diversos cenários, destacando-se o tamanho amostral. Acrescenta-se que os estimadores foram avaliados no que se refere às suas variâncias, erros quadráticos médios e vieses. Os estimadores corrigidos por *bootstrap* apresentaram melhores resultados, em geral, quando comparados com os estimadores não corrigidos, quando o estimador é viesado.

**Palavras-chave:** *Bootstrap*, Correção de viés, Máxima Verossimilhança, Método dos Momentos, Simulação de Monte Carlo.

## Abstract

This paper evaluates empirically the performance of parameters estimators related to the Pareto distribution. From samples generated by this distribution, we calculated punctual perform estimations of the parameters using the maximum likelihood method, method of moments and their respective versions corrected by bootstrap. Many different scenarios were considered, specially the sample size used. We can add that the estimators were evaluated considering their variances, mean squared errors and bias. In general, the estimators corrected by bootstrap showed better results compared to the uncorrected estimators, i. e., when the estimator is biased.

**Keywords:** Bootstrap, Bias Correction, Maximum Likelihood, Method of Moments, Monte Carlo simulation.

## Introdução

Vilfredo Pareto foi um economista italiano que viveu entre 1848 e 1923. Ele constatou que apenas um número reduzido de pessoas detinha grande parte da riqueza existente. Motivado por este fato, Pareto desenvolveu uma distribuição que leva seu nome, distribuição Pareto (PEREIRA & REQUEIJO, 2008). A lei de Pareto, tal como formulado por Pareto em 1897, tratou da distribuição de renda em uma população e pode ser definida como  $N = Ax^\alpha$ , em que  $N$  é o número de pessoas com renda maior ou igual a  $X$ . Pareto percebeu que esta lei é universal e inevitável, independentemente da tributação e das condições sociais e políticas. Contestações da lei Pareto têm sido feitas por vários economistas nos últimos anos, por exemplo, Pigou (1932), Shirras (1935) e Hayakawa (1951). Recentemente, diversas tentativas foram realizadas para explicar muitos fenômenos empíricos usando a lei de Pareto. Atualmente, o princípio de Pareto está ligado à literatura industrial e uma de suas aplicações é no controle de qualidade.

Correções de viés têm sido bastante estudadas na literatura estatística. Os estimadores de máxima verossimilhança em modelos estatísticos paramétricos são, em geral, viesados para os valores verdadeiros dos parâmetros quando o tamanho da amostra é pequeno.

Diante desse fato, este trabalho tem como objetivo a implementação de um programa em C que realize simulações de Monte Carlo para avaliar, empiricamente, o desempenho dos estimadores pontuais e suas versões corrigidas por *bootstrap* dos parâmetros da distribuição Pareto em amostras de tamanho finito.

Para isso, são geradas amostras pseudo-aleatórias de variáveis que seguem a distribuição Pareto e, então, estimam-se os seus parâmetros utilizando os métodos de estimação de máxima verossimilhança, método dos momentos e suas respectivas versões corrigidas por *bootstrap*. São considerados diferentes tamanhos amostrais bem como diferentes valores para os parâmetros da distribuição Pareto.

## Referencial teórico

### A distribuição Pareto

A distribuição de Pareto é uma distribuição de probabilidade de lei de potência, que é utilizada para a modelagem de dados sociais, geofísicos, atuariais e muitos outros tipos de fenômenos observáveis. Fora do campo da economia, a distribuição Pareto, às vezes, é referenciada como distribuição de Bradford (JOHNSON et al. 1994).

Neste item, apresentamos a distribuição Pareto, incluindo a sua função densidade, função de distribuição acumulada e métodos para estimação dos parâmetros da distribuição.

### 2.2 Definições básicas

A função distribuição acumulada de uma variável aleatória  $X$  com distribuição Pareto é dada

$$F(x) = P(X \leq x) = 1 - \left( \frac{k}{x} \right)^\alpha, \quad k > 0, \alpha > 0, \quad x \geq k,$$

com a correspondente função densidade de probabilidade sendo

$$f(x; \alpha, k) = \frac{\alpha k^\alpha}{x^{\alpha+1}},$$

em que  $k$  é o parâmetro de escala e  $\alpha$  é o parâmetro de forma.

Os momentos da distribuição Pareto são dados por

$$E(X^n) = \int_k^{\infty} x^n f(x) dx = \int_k^{\infty} x^n \frac{\alpha k^\alpha}{x^{\alpha+1}} dx = \frac{\alpha k^\alpha}{\alpha - n}, \alpha > n.$$

Em especial, o valor esperado e a variância são, respectivamente

$$E(X) = \frac{\alpha k}{\alpha - 1}, \alpha > 1.$$

$$Var(X) = \frac{\alpha k^2}{(\alpha - 2)(\alpha - 1)^2}, \alpha > 2.$$

Na Figura 1, apresentamos o gráfico da densidade para alguns valores de  $\alpha$  considerando  $k = 1,0$ .

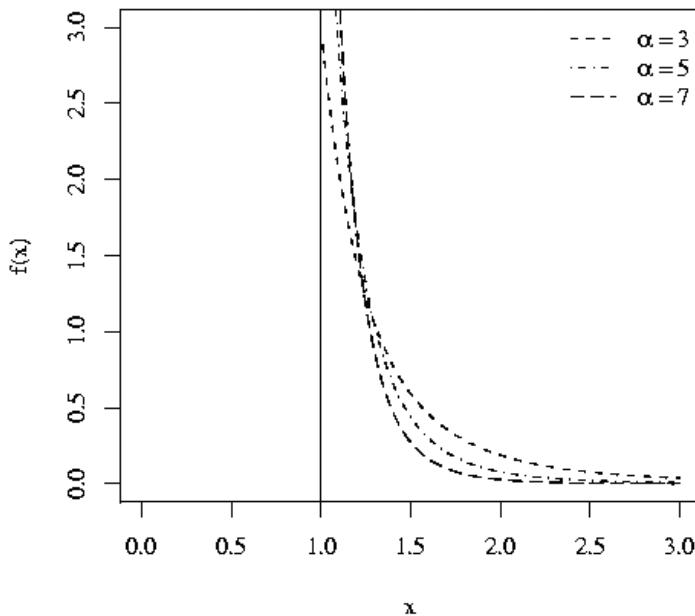


Figura 1. Densidade Pareto para diferentes valores de  $\alpha$  e  $K=1,0$ .

## Estimação dos parâmetros

A estimativa dos parâmetros do modelo é uma das principais etapas no ajuste de um modelo. Existem muitos procedimentos sugeridos na literatura com essa finalidade. O objetivo desta seção é apresentar alguns procedimentos que são utilizados na estimativa dos parâmetros da distribuição Pareto.

Considerando a distribuição Pareto como completamente especificada através do vetor de parâmetros  $\theta = (\alpha, k)^T$ , a seguir são apresentados os estimadores de máxima verossimilhança e do método dos momentos.

### Método de máxima verossimilhança

Seja  $x_1, \dots, x_n$  uma amostra aleatória de tamanho  $n$  da distribuição Pareto, ou seja, os  $x_j$ 's são variáveis aleatórias independentes com densidade

$$f(x_j; \alpha, k) = \frac{\alpha k^\alpha}{x_j^{\alpha+1}},$$

para  $j = 1, \dots, n$ . A função de verossimilhança é dada por

$$L(\alpha, k) = \prod_{j=1}^n \frac{\alpha k^\alpha}{x_j^{\alpha+1}}.$$

A função de log-verossimilhança é da forma

$$l(\alpha, k) = n \ln \alpha + n \alpha \ln k - (\alpha + 1) \sum_{j=1}^n \ln x_j.$$

O estimador de máxima verossimilhança de  $\theta = (\alpha, k)^T$  é o valor  $\hat{\theta} = (\hat{\alpha}, \hat{k})^T$  que maximiza a função de verossimilhança  $L(\alpha, k)$ . O valor de  $\theta$  que maximiza a função de verossimilhança  $L(\alpha, k)$ , também maximiza  $l(\alpha, k)$ . Assim, os estimadores de máxima verossimilhança de  $\alpha$  e  $k$  são obtidos através da maximização do logaritmo da função de verossimilhança.

Assim, diferenciando-se parcialmente  $l(\alpha, k)$  em relação à  $\alpha$  e em relação à  $k$  e igualando o resultado à zero, temos que

$$\hat{\alpha} = n \left[ \sum_{j=1}^n \log\left(\frac{x_j}{\hat{k}}\right) \right]^{-1}.$$

Por outro lado, pode-se mostrar que  $l(\alpha, k)$  é monótona crescente com relação a  $x$ , ou seja, quanto maior o valor de  $x$ , maior será o valor da função de log-verossimilhança. Logo, um estimador para o parâmetro  $k$  é dado por

$$\hat{k} = \min_j x_j.$$

Além disso, pode-se mostrar que

$$E(\hat{k}) = \frac{nk\alpha}{n\alpha - 1}, \quad n > \frac{1}{\alpha},$$

e

$$Var(\hat{k}) = \frac{nk\alpha^2}{(n\alpha - 1)^2(n\alpha - 2)}, \quad n > \frac{2}{\alpha},$$

e que

$$E(\hat{\alpha}) = \frac{n\alpha}{n-2}, \quad n > 2$$

e

$$Var(\hat{\alpha}) = \frac{n^2\alpha^2}{(n-2)^2(n-3)}, \quad n > 3.$$

## Método dos momentos

O método dos momentos é um dos métodos de estimação mais simples. Seja

$$m_r = \frac{1}{n} \sum_{j=1}^n x_j^r, \quad r \geq 1,$$

o  $r$ -ésimo momento amostral, e  $\mu_r = E(X^r)$ , o  $r$ -ésimo momento populacional.

O método dos momentos consiste na obtenção de estimadores pontuais a partir do procedimento de igualar os momentos teóricos pelos respectivos momentos amostrais, ou seja, resolvendo as equações  $m_r = \mu_r$ . Para  $\alpha > 1$ , a esperança da distribuição Pareto é dada por  $E(X) = \alpha k(\alpha - 1)^{-1}$ . Igualando a esperança com a média amostral, temos que

$$\tilde{\alpha} = \frac{\bar{x}}{\bar{x} - \tilde{k}}, \text{ em que } \bar{x} = \frac{1}{n} \sum_{j=1}^n x_j.$$

Um estimador para  $k$  pode ser encontrado utilizando o seguinte argumento encontrado em Quandt (1966): a probabilidade dos  $x_j$ 's serem maior do que um determinado valor  $x$  é

$$\left( \frac{k}{x} \right)^{\alpha n},$$

para todo  $n$ . Após algumas álgebras, um estimador para  $k$  é dado por

$$\tilde{k} = \frac{(\tilde{\alpha}n - 1)x_{(1)}}{\tilde{\alpha}n},$$

em que  $x_{(1)} = \min_j x_j$ , assim podemos escrever  $\tilde{\alpha}$  como

$$\tilde{\alpha} = \frac{n\tilde{x} - x_{(1)}}{n(\tilde{x} - x_{(1)})}.$$

## Correção de viés

Usualmente, o viés é ignorado na prática sob a justificativa de que é desprezível quando comparado com o erro padrão do estimador do parâmetro, dado que, em uma amostra aleatória de tamanho  $n$ , o viés do estimador de máxima verossimilhança, em geral, é de ordem  $O(n^{-1})$ , enquanto o seu respectivo erro padrão assintótico é de ordem  $O(n^{-1/2})$ . Entretanto,

<sup>1</sup> Se  $\{a_n\}_{n \geq 1}$  e  $\{b_n\}_{n \geq 1}$  são duas sequências de números reais, dizemos que  $a_n$  é de ordem menor que  $b_n$  e escrevemos  $a_n = o(b_n)$  se  $\lim_{n \rightarrow \infty} a_n/b_n = 0$ . Dizemos que  $a_n$  é de ordem no máximo igual a  $b_n$ , denotado por  $a_n^{\sim} = O(b_n)$ , se existe um número real  $M > 0$  tal que  $|a_n/b_n| \leq M$ , isto é, a razão  $|a_n/b_n|$  é limitada. Observamos que se  $a_n = o(b_n)$  e que quando  $b_n \rightarrow 0$  a ordem fornece noção da taxa de convergência de  $a_n$  para zero.

em pequenas amostras, o viés pode ser apreciável e pode ter magnitude comparável ao erro padrão da estimativa do parâmetro (LEMONTE, 2006).

Correções de viés têm sido bastante estudadas na literatura estatística. Um artigo precursor sobre correção de viés é de Bartlett (1953), o qual apresentou uma expressão simples para o viés de ordem  $O(n^{-1})$  do estimador de máxima verossimilhança no caso uniparamétrico. Cox & Snell (1968) apresentam uma expressão geral para o viés de ordem  $O(n^{-1})$  do estimador de máxima verossimilhança nos casos uniparamétrico e multiparamétrico.

Outra técnica que pode ser utilizada com o intuito de obter estimadores corrigidos é o método de reamostragem *bootstrap*, desenvolvido por Efron (1979). O *bootstrap* é um método computacional de inferência estatística que substitui cálculos analíticos por esforço computacional.

Neste estudo, vamos utilizar o método de *bootstrap* para corrigir o viés das estimativas dos parâmetros da distribuição Pareto pelos métodos de máxima verossimilhança e momentos. O método *bootstrap* e a correção de viés por *bootstrap* são apresentados nas próximas seções.

### Método *Bootstrap*

Os métodos estatísticos computacionais surgiram exatamente da necessidade de simular um experimento ou um fenômeno diversas vezes, com o objetivo de obter resultados que a teoria matemática não pode fornecer de maneira rápida e eficiente. O *bootstrap* é um método computacional de inferência estatística introduzido por Efron (1979) capaz de responder a questões reais sem a necessidade de exaustivos, complicados e muitas vezes inviáveis cálculos analíticos.

Considere uma amostra aleatória  $\mathbf{x} = (x_1, \dots, x_n)$  de uma determinada distribuição  $F$  desconhecida e seja  $\eta(\mathbf{x}, F)$  certa quantidade de interesse. Geralmente, estamos interessados em estimar alguma característica probabilística de  $\eta(\mathbf{x}, F)$ . *Bootstrap* é uma técnica apropriada para essa tarefa. O método pode ser descrito em três passos:

- Trocar  $F$  por alguma distribuição conhecida  $\hat{F}$  que, num certo sentido, está “próxima” de  $F$ ;
- Obter  $B$  amostras  $x^{*1}, \dots, x^{*B}$  de  $\hat{F}$  e calcular  $T^j = \eta(x^{*j}, \hat{F})$  para  $j = 1, \dots, B$ ;
- Calcular características de interesse, como média e variância, a partir da distribuição amostral dos valores  $T^1, \dots, T^B$ .

Existem duas variações de *bootstrap*: não-paramétrico e paramétrico. No *bootstrap* não-paramétrico, uma distribuição amostral é usada como  $\hat{F}$ .

No *bootstrap* paramétrico,  $F$  é assumido ser membro de alguma família paramétrica e  $\hat{F}$  é escolhido dessa família simplesmente trocando os parâmetros desconhecidos por suas respectivas estimativas. Neste trabalho, vamos utilizar o *bootstrap* não-paramétrico.

### Correção de viés por *Bootstrap*

O método *bootstrap* foi introduzido para obter a estimativa do erro padrão de um estimador arbitrário. A ideia inicial de Efron foi usar *bootstrap* para estimar o erro padrão de  $\hat{\theta}$ , sem se importar quão complicada é a forma funcional de  $\hat{\theta}$ . Entretanto, o método *bootstrap* pode ser adaptado para estimar o viés de um estimador e, portanto, pode ser aplicado como um procedimento para correção de viés.

Seja  $\theta = t(F)$  uma função de  $F$  denominada parâmetro e seja  $\hat{\theta} = s(\mathbf{x})$  um estimador de  $\theta$ .

Assim, denotamos o viés de  $\hat{\theta}$ ,  $B_F(\hat{\theta}, \theta) = E_F[s(\mathbf{x})] - t(F)$ , em que o subscrito  $F$  indica que a esperança matemática é calculada com base em  $F$ . Dessa forma, o estimador *bootstrap* de viés na versão não-paramétrica é dado por

$$B_{\hat{F}}(\hat{\theta}, \theta) = E_{\hat{F}}[s(\mathbf{x})] - t(\hat{F}).$$

Gerando  $B$  amostras *bootstrap* independentes,  $(\mathbf{x}^{*_1}, \dots, \mathbf{x}^{*B})$  de  $\mathbf{x}$ , calculando as respectivas réplicas *bootstrap*  $(\hat{\theta}^{*_1}, \dots, \hat{\theta}^{*B})$ , em que

$\hat{\theta}^{*i} = s(\mathbf{x}^{*i})$  para  $i = 1, B$ , podemos aproximar  $E_{\hat{F}}[s(\mathbf{x})]$  pela média

$$\hat{\theta}^{*(\cdot)} = \frac{1}{B} \sum_{i=1}^B \hat{\theta}^{*i}. \text{ Obtemos, assim, as estimativas } bootstrap \text{ de viés, dada por}$$

$$\hat{B}_{\hat{F}}(\hat{\theta}, \theta) = \hat{\theta}^{*(\cdot)} - s(\mathbf{x}).$$

Utilizando as estimativas *bootstrap* do viés, define-se o estimador corrigido até segunda ordem (MacKinnon & Smith (1998)) por *bootstrap* como

$$\hat{\theta}^* = 2s(\mathbf{x}) - \hat{\theta}^{*(\cdot)}. \quad (1)$$

## Metodologia

Através de simulações de Monte Carlo, avaliamos os desempenhos dos estimadores de máxima verossimilhança ( $\hat{\theta}$ ) e momentos ( $\tilde{\theta}$ ) da distribuição Pareto e suas versões corrigidas. Consideramos amostras de tamanho finito e de diferentes valores para o vetor paramétrico. Todo o processo de simulação de Monte Carlo foi realizado utilizando a linguagem de programação C, desenvolvida por Dennis Ritchie (1972).

Antes de gerar as observações da distribuição Pareto, utilizamos o gerador de números pseudo-aleatórios, que gera números uniformes entre zero e um. Esses números pseudo-aleatórios uniformes foram gerados a partir do gerador *multiply-with-carry*, com período de aproximadamente 260, proposto por George Marsaglia (1997).

A simulação de uma variável aleatória  $X$  com distribuição Pareto foi realizada através da transformação em que  $\xi \sim U(0,1)$

$$F(x) = 1 - \left(\frac{k}{x}\right)^\alpha = \xi \Rightarrow x = \frac{k}{(1-\xi)^{\frac{1}{\alpha}}},$$

Os tamanhos amostrais considerados foram  $n = 25, 50, 75$  e  $100$  e os valores considerados para o parâmetro  $\alpha$  sendo  $\alpha = 3,0, 5,0$  e  $7,0$ . Sem perda de generalidade, o parâmetro de escala foi fixado em  $1,0$ , isto é,  $k = 1,0$  em todos os experimentos de Monte Carlo. Consideramos  $R = 2000$  (número de réplicas de Monte Carlo) e  $B = 500$  (número de réplicas *bootstrap*).

Para cada réplica de Monte Carlo, ou seja, para cada estimativa de

$\hat{\theta} = (\hat{\alpha}, \hat{k})$  e  $\tilde{\theta} = (\tilde{\alpha}, \tilde{k})$ , geramos  $B$  réplicas *bootstrap* de forma não-paramétrica, isto é, geramos

$x^{*1}, K, x^{*B}$ , em que  $x^{*i} = (x_1^*, \dots, x_n^*)$ ,  $i = 1, \dots, B$ .

Com essas réplicas *bootstrap*, determinamos as estimativas *bootstrap* dos vieses de  $\hat{\theta} = (\hat{\alpha}, \hat{k})$  e  $\tilde{\theta} = (\tilde{\alpha}, \tilde{k})$  e calculamos, assim, as estimativas corrigidas por *bootstrap*  $\hat{\theta}^* = (\hat{\alpha}^*, \hat{k}^*)$  e  $\tilde{\theta}^* = (\tilde{\alpha}^*, \tilde{k}^*)$  de acordo com (1).

Para a análise dos resultados da estimação pontual, calculamos a média, o viés, a variância (Var) e o erro quadrático médio dos estimadores (EQM).

## Resultados e discussão

A avaliação dos desempenhos dos estimadores dos parâmetros da distribuição Pareto através de métodos empíricos é a motivação do estudo apresentado nesta seção.

Os resultados da simulação são expostos nas Tabelas 1, 2 e 3. Nessas, obtemos as estimativas de  $\alpha$  e  $k$ , considerando os tamanhos amostrais 25, 50, 75 e 100, além do viés, variância (Var) e erro quadrático médio (EQM) das estimativas. Vale salientar que  $\hat{\theta}(\bar{\theta})$  é o estimador de máxima verossimilhança (momentos) e  $\hat{\theta}^*(\tilde{\theta}^*)$  é o estimador de máxima verossimilhança (momentos) corrigido por *bootstrap*.

A Tabela 1 apresenta as estimativas para o caso em que  $\alpha = 3,0$  e  $k = 1,0$ . Nesse cenário, podemos observar que os estimadores  $\hat{\alpha}^*$  e  $\tilde{\alpha}^*$  apresentaram, em módulo, viés menor que o viés dos estimadores não corrigidos para todos os tamanhos amostrais. O estimador  $\tilde{\alpha}^*$  se destaca, apresentando menor viés. Por exemplo, para  $n = 25$ , o viés de  $\tilde{\alpha}^*$  é dado, em módulo, por 0,010, enquanto que o viés de  $\hat{\alpha}^*$  é 0,047.

Em termos de EQM, nota-se que o estimador que apresentou menor EQM para o parâmetro  $\alpha$  foi o estimador  $\tilde{\alpha}^*$ , entretanto,  $\hat{\alpha}^*$  apresentou menor EQM do que  $\hat{\alpha}$ . Nota-se também que, à medida que a amostra cresce, o EQM dos quatro estimadores diminuem.

Em relação ao parâmetro  $k$ , observa-se que o estimador  $\hat{k}$  apresentou menor viés do que  $\tilde{k}^*$ . Além disso, os resultados numéricos evidenciam que o estimador para o parâmetro  $k$  pelo método dos momentos não é viesado, uma vez que a estimativa do viés de  $\tilde{k}$  é sempre próxima de zero.

O estimador  $\tilde{k}^*$  apresentou menor EQM para todos os tamanhos amostrais. Ressalta-se que  $\tilde{k}^*$  apresentou maior EQM do que sua versão não corrigida.

Tabela 1. Estimativas dos parâmetros no caso de  $\alpha = 3,0$  e  $k = 1,0$ .

$n$	Estimador	Estimativas de $\alpha$				Estimativas de $K$			
		Média	Viés	Var	EQM	Média	Viés	Var	EQM
25	$\hat{\theta}$	3,277	0,277	0,644	0,721	1,014	0,014	0,006	0,006
	$\hat{\theta}^*$	3,047	0,047	0,445	0,447	1,005	0,006	0,007	0,007
	$\tilde{\theta}$	3,224	0,224	0,606	0,504	0,005	0,005	0,006	0,006
50	$\hat{\theta}$	2,989	-0,010	0,397	0,397	-0,008	0,006	0,006	0,006
	$\hat{\theta}^*$	3,127	0,127	0,546	0,562	1,006	0,006	0,006	0,006
	$\tilde{\theta}$	3,024	0,024	0,485	0,485	1,0003	0,003	0,007	0,007
75	$\hat{\theta}$	3,107	0,107	0,232	0,232	0,999	-0,001	0,006	0,006
	$\hat{\theta}^*$	2,992	-0,008	0,446	0,446	0,996	-0,004	0,006	0,006
	$\tilde{\theta}$	3,098	0,098	0,512	0,522	1,005	0,005	0,006	0,006
100	$\hat{\theta}$	3,030	0,030	0,508	0,509	1,002	0,002	0,007	0,007
	$\hat{\theta}^*$	3,089	0,089	0,507	0,191	1,000	0,000	0,006	0,006
	$\tilde{\theta}$	3,010	0,011	0,483	0,483	0,997	-0,003	0,007	0,007
100	$\hat{\theta}$	3,072	0,072	0,495	0,500	1,003	0,003	0,006	0,006
	$\hat{\theta}^*$	3,022	0,022	0,518	0,519	1,001	0,001	0,007	0,007
	$\tilde{\theta}$	3,067	0,067	0,492	0,143	1,000	0,000	0,006	0,006
100	$\hat{\theta}$	3,007	0,007	0,499	0,500	0,988	-0,002	0,006	0,006

As Tabelas 2 e 3 apresentam as estimativas para o caso em que  $\alpha = 5,0$  e  $k = 1,0$ , e  $\alpha = 7,0$  e  $k = 1,0$ , respectivamente. Entretanto, a interpretação do desempenho empírico dos estimadores é equivalente à da Tabela 1.

Analisando as três tabelas, percebe-se que, à medida que o parâmetro  $\alpha$  aumenta as estimativas do viés, variância e EQM também aumentam em relação às estimativas do parâmetro  $\alpha$ .

Portanto, de acordo com os resultados apresentados nas Tabelas 1, 2 e 3, dentre os estimadores analisados, os estimadores  $\hat{\alpha}^*$  e  $\tilde{\alpha}^*$ , para o parâmetro  $\alpha$ , foram os que apresentaram os melhores resultados, mostrando assim a eficácia das correções do viés por bootstrap. Para o parâmetro  $k$ , o estimador  $\tilde{k}$  apresentou melhor desempenho, pois os resultados empíricos evidenciam que o estimador do parâmetro  $k$  pelo método dos momentos é não viesado. Vale ressaltar que o estimador  $\hat{k}^*$  apresentou melhores resultados em termos gerais do que o estimador  $\hat{k}$ .

Tabela 2. Estimativas dos parâmetros no caso de  $\alpha = 5,0$  e  $k = 1,0$ .

<i>n</i>	Estimador	Estimativas de $\alpha$			Estimativas de $K$				
		Média	Viés	Var	EQM	Média	Viés	Var	EQM
25	$\hat{\theta}$	5,462	0,462	4,971	5,185	1,008	0,008	0,006	0,006
	$\hat{\theta}^*$	5,078	0,078	3,437	3,444	1,004	0,004	0,007	0,007
	$\tilde{\theta}$	5,303	0,303	4,429	1,310	1,000	0,000	0,006	0,006
	$\tilde{\theta}^*$	4,924	-0,075	2,979	2,985	0,995	-0,005	0,006	0,006
50	$\hat{\theta}$	5,212	0,212	4,210	4,255	1,004	0,004	0,006	0,006
	$\hat{\theta}^*$	5,039	0,039	3,741	3,742	1,002	0,002	0,007	0,007
	$\tilde{\theta}$	5,138	0,138	3,994	0,576	0,999	0,000	0,006	0,006
	$\tilde{\theta}^*$	4,958	-0,042	3,443	3,445	0,998	-0,002	0,006	0,006
75	$\hat{\theta}$	5,164	0,163	3,954	3,980	1,003	0,003	0,006	0,006
	$\hat{\theta}^*$	5,051	0,051	3,922	3,925	1,001	0,001	0,007	0,007
	$\tilde{\theta}$	5,117	0,117	3,814	0,488	1,000	0,000	0,006	0,006
	$\tilde{\theta}^*$	4,998	-0,002	3,723	3,723	0,998	-0,002	0,006	0,006
100	$\hat{\theta}$	5,119	0,119	3,819	3,833	1,002	0,002	0,006	0,006
	$\hat{\theta}^*$	5,036	0,037	4,000	4,002	1,000	0,000	0,006	0,007
	$\tilde{\theta}$	5,085	0,085	3,719	0,356	1,000	0,000	0,006	0,006
	$\tilde{\theta}^*$	4,997	-0,003	3,855	3,854	0,998	-0,001	0,006	0,006

Tabela 3. Estimativas dos parâmetros no caso de  $\alpha = 7,0$  e  $k = 1,0$ .

$n$	Estimador	Estimativas de $\alpha$				Estimativas de $K$			
		Média	Viés	Var	EQM	Média	Viés	Var	EQM
25	$\hat{\theta}$	7,647	0,647	19,099	19,518	1,005	0,006	0,006	0,006
	$\hat{\theta}^*$	7,109	0,109	13,207	13,218	1,005	0,003	0,007	0,007
	$\tilde{\theta}$	7,394	0,395	16,729	2,529	1,000	0,000	0,006	0,006
50	$\hat{\theta}$	6,870	0,129	11,380	11,397	0,996	-0,003	0,005	0,006
	$\hat{\theta}^*$	7,297	0,297	16,173	16,261	1,003	0,003	0,006	0,006
	$\tilde{\theta}$	7,055	0,055	14,371	14,374	1,001	0,001	0,007	0,007
75	$\hat{\theta}$	7,177	0,177	15,185	1,100	0,999	0,000	0,006	0,005
	$\hat{\theta}^*$	6,921	-0,069	13,249	13,254	0,998	-0,002	0,006	0,006
	$\tilde{\theta}$	7,229	0,229	15,189	15,242	1,002	0,002	0,006	0,005
100	$\hat{\theta}$	7,071	0,071	15,068	15,073	1,000	0,002	0,006	0,006
	$\hat{\theta}^*$	7,152	0,152	14,554	0,943	1,000	0,001	0,006	0,005
	$\tilde{\theta}$	6,990	-0,009	14,307	14,307	0,998	0,000	0,006	0,006
100	$\hat{\theta}$	7,167	0,168	14,672	14,700	1,001	-0,001	0,006	0,005
	$\hat{\theta}^*$	7,051	0,051	15,368	15,370	1,000	0,001	0,006	0,006
	$\tilde{\theta}$	7,111	0,111	14,214	0,684	1,000	0,001	0,006	0,005
100	$\hat{\theta}$	6,991	0,009	14,806	14,806	0,999	-0,001	0,006	0,006

## Conclusão

No desenvolvimento deste trabalho, foram apresentadas as características e propriedades da distribuição Pareto. É avaliado, via simulação de Monte Carlo, o comportamento das estimativas dessa distribuição utilizando o método de máxima verossimilhança, o método dos momentos e as suas respectivas versões corrigidas por *bootstrap* em diversos cenários.

Os resultados das simulações de Monte Carlo realizadas mostraram que as correções de viés por *bootstrap*, para os dois métodos de estimação, foram eficazes e apresentaram os melhores desempenhos em termos de viés e erro quadrático médio com relação ao parâmetro  $\alpha$ . Para o parâmetro  $k$ , o método dos momentos apresentou melhor desempenho do que as versões corrigidas por *bootstrap*, evidenciando que o estimador  $\tilde{k}$  não é viesado.

Dessa forma, para estimar os parâmetros  $\alpha$  e  $k$  da distribuição Pareto, recomendamos os estimadores corrigidos por *bootstrap*, visto que essas duas metodologias apresentam desempenhos favoráveis relativamente na estimação por máxima verossimilhança e pelo método dos momentos, quando o estimador é viesado.

## Agradecimento

À CAPES, Coordenação de Aperfeiçoamento de Pessoal de Nível Superior, pelo auxílio financeiro.

## Referências

- BARTLETT, M. S. Approximate confidence intervals II. *Biometrika*, n. 40, p.306–317, 1953.
- COX, D. R.; SNELL, E. A general definition of residuals. *Journal of the Royal Statistical Society B*, n. 30, p. 248–275, 1968.
- EFRON, B. Bootstrap methods: another look at the jackknife. *Annals of Statistics*, n. 7, p. 1–26, 1979.
- HAYAKAWA, M. The application of Pareto's law of income to Japanese data. *Econometrics*, n. 19, p. 174–183, 1951.
- JOHNSON, N. L.; KOTZ, S. N.; BALAKRISHNAN, N. **Continuous Univariate Distributions**. New York: Wiley, 1994. vol. 1, 2 ed.
- LEMONTE, A. J. **Inferência sobre os parâmetros da distribuição Birnbaum-Saunders Bi-paramétrica**. Dissertação de mestrado – Universidade Federal de Pernambuco, 2006. 191 p.
- MACKINNON, J. G. & SMITH, J. A. A. Approximate bias correction in econometrics. *Journal of Econometrics*, v. 85, p. 205-230, 1998.

PEREIRA, Z. L.; REQUEIJO, J. G. **Qualidade: Planejamento e Controle Estatístico de Processo.** Fundação Faculdade de Ciência e Tecnologia da Universidade Nova Lisboa. Lisboa, Portugal, 2008.

PIGOU, A. C. **The Economics of Welfare.** London: Macmillan.1932.

QUANDT, R. E. Old and new methods of estimation and the Pareto distribution. **Metrika**, n. 10, p. 55–82, 1966.

SHIRRAS, G. F. The Pareto law and the distribution of income. **Economic Journal**, n.45, p.663–681, 1935.

Submetido em: 09/09/2011  
Aceito em: 26/12/2011