

*Estimação de quantis elevados em estatística de extremos***Frederico Caeiro***CMA e DM, Universidade Nova de Lisboa***M. Ivette Gomes***DEIO e CEAUL, Universidade de Lisboa*

Resumo: Nas mais variadas áreas de aplicação, de entre as quais destacamos os *Seguros* e as *Finanças*, pretende-se frequentemente estimar um quantil elevado, ou equivalentemente, em linguagem de âmbito financeiro, o parâmetro VaR_p (de *Value at Risk*), em que p , pequeno, representa a probabilidade de excedência de VaR_p . A estimação semi-paramétrica deste parâmetro depende não só da estimação do índice de cauda γ , o principal parâmetro na área de *Estatística de Extremos*, mas também de um parâmetro de escala de primeira ordem, aqui denotado C . Recentemente, para além de classes interessantes de estimadores de γ , com viés reduzido, foram também estudadas novas classes do parâmetro de escala C , dependentes da estimação de parâmetros de segunda ordem. A utilização de novas classes de estimadores de γ e de C na estimação de VaR_p permite-nos introduzir novas classes de estimadores de quantis elevados. Essas novas classes são comparadas entre si e com algumas das classes já existentes na literatura, através da utilização de técnicas de Monte Carlo. Será ainda fornecida uma aplicação na área das finanças.

Palavras-chave: Estatística de Extremos, Caudas pesadas, Quantis elevados, Estimação semi-paramétrica.

Abstract: In the most diversified areas of application, like *Insurance* and *Finance*, we often need to estimate a high quantile, or equivalently, the *Value at Risk*, VaR_p , a value that is exceeded with a probability p , small. The semi-parametric estimation of this parameter depends not only on the estimation of the tail index γ , the primary parameter in *Statistics of Extremes*, but also of a first order scale parameter or functional, here denoted C . Recently, new interesting classes of reduced bias γ and C estimators have appeared in the literature, all dependent on the estimation of second order parameters. The use of one of those new classes of γ as well as of C estimators in the estimation of VaR_p enables us to introduce new classes of high quantile estimators. These new classes are compared among themselves and with previous ones, through the use of Monte Carlo simulation. An application in the area of finance is also provided.

Keywords: Statistics of Extremes, Heavy tails, High quantiles, Semi-parametric estimation.

1 Introdução e preliminares

Dizemos que um modelo F tem cauda direita pesada se a *função de cauda* $\bar{F} := 1 - F \in RV_{-1/\gamma}$, $\gamma > 0$, onde RV_α denota a classe das funções de variação regular no infinito com índice de variação regular α , i.e., funções mensuráveis e não-negativas $g(\cdot)$ tais que $g(tx)/g(t) \rightarrow x^\alpha$, quando $t \rightarrow \infty$, para qualquer $x > 0$ (Gnedenko, 1943). Denotemos $U(t) := F^{\leftarrow}(1 - 1/t) = \inf \{x : F(x) \geq 1 - 1/t\}$. Então, podemos equivalentemente dizer que um modelo F tem cauda direita pesada se e só se $U \in RV_\gamma$ (de Haan, 1970), i.e.

$$\lim_{t \rightarrow \infty} \frac{U(tx)}{U(t)} = x^\gamma, \quad \text{para qualquer } x > 0. \quad (1)$$

Para valores pequenos de p , queremos estimar um quantil elevado $\chi_p \equiv VaR_p$, um valor tal que $1 - F(VaR_p) = \mathbb{P}(X > VaR_p) = p$, um parâmetro típico nas áreas de seguros e finanças. De modo mais específico, pretendemos estimar o parâmetro,

$$VaR_p = U(1/p), \quad p = p_n \rightarrow 0, \quad n p_n \leq 1, \quad (2)$$

e admitiremos estar a trabalhar na classe de modelos de Hall (Hall and Welsh, 1985), em que existe $\gamma > 0$, $\rho < 0$, $C > 0$ e $\beta \neq 0$ tais que

$$U(t) = Ct^\gamma (1 + \gamma \beta t^\rho / \rho + o(t^\rho)), \quad \text{quando } t \rightarrow \infty. \quad (3)$$

A classe em (3) contém muitos dos modelos de caudas pesadas mais frequentemente usados em aplicações, tais como os modelos *Fréchet*, *Pareto Generalizado* e *t-de-Student*.

Vamos basear a inferência nas k estatísticas ordinais (e.o.) de topo, e admitiremos que k é uma sucessão *intermédia* de inteiros em $[1, n)$, i.e.,

$$k = k_n \rightarrow \infty, \quad k \in [1, n), \quad k = o(n), \quad \text{quando } n \rightarrow \infty. \quad (4)$$

Como a partir de (2) e de (3), $VaR_p = U(1/p) \sim C p^{-\gamma}$, quando $p \rightarrow 0$, um estimador óbvio de VaR_p é dado por $\hat{C} p^{-\hat{\gamma}}$, com \hat{C} e $\hat{\gamma}$ quaisquer estimadores consistentes de C e de γ , respectivamente. Dada uma amostra (X_1, X_2, \dots, X_n) , denotemos $(X_{i:n}, 1 \leq i \leq n)$, a amostra das e.o. ascendentes. Seja Y uma variável aleatória (v.a.) Pareto padrão, i.e, uma v.a. com função de distribuição (f.d.) $F_Y(y) = 1 - 1/y$, $y \geq 1$. A transformação uniformizante permite-nos então escrever $X_{n-k:n} \stackrel{d}{=} U(Y_{n-k:n})$, e como $Y_{n-k:n} \stackrel{p}{\sim} (n/k)$, para k intermédio, e admitimos estar a trabalhar com modelos em (3), podemos escrever $X_{n-k:n} \stackrel{p}{\sim} C Y_{n-k:n}^\gamma \stackrel{p}{\sim} C (n/k)^\gamma$, quando $n \rightarrow \infty$. Consequentemente, um possível estimador de C é $\hat{C}_{\hat{\gamma}}(k) = X_{n-k:n} (k/n)^{\hat{\gamma}}$, sendo

$$Q_{\hat{\gamma}}(k; p) := X_{n-k:n} \left(\frac{k}{np} \right)^{\hat{\gamma}} \quad (5)$$

o estimador óbvio de VaR_p (Weissman, 1978).

Para caudas pesadas, é usualmente o estimador clássico do índice de cauda que é inserido em (5). E esse estimador clássico é o estimador de Hill, $\hat{\gamma} = \hat{\gamma}(k) =: H(k)$ (Hill, 1975), com a expressão funcional,

$$H(k) := \frac{1}{k} \sum_{i=1}^k U_i, \tag{6}$$

onde

$$U_i := i (\ln X_{n-i+1:n} - \ln X_{n-i:n}), \quad 1 \leq i \leq k < n, \tag{7}$$

são os espaçamentos escalados das log-observações. Com base no estimador de Hill $H(k)$ em (6), obtemos então o chamado estimador clássico do parâmetro VaR_p , para o qual usamos a notação $Q_H(k; p)$.

Para derivar o comportamento assintótico não-degenerado de estimadores semi-paramétricos de parâmetros de acontecimentos extremos necessitamos de impor condições sobre a velocidade de convergência em (1). Admitiremos aqui que existe uma função $A(t)$ tal que

$$\lim_{t \rightarrow \infty} \frac{\ln U(tx) - \ln U(t) - \gamma \ln x}{A(t)} = \frac{x^\rho - 1}{\rho} \tag{8}$$

para qualquer $x > 0$. Então, a função A é uma função de sinal constante próximo do infinito (positivo ou negativo), e $\rho \leq 0$ é um parâmetro de “forma” de segunda ordem. A função limite em (8) tem de ser necessariamente desta forma, e $|A| \in RV_\rho$ (Geluk and de Haan, 1987). Admitiremos em todo o trabalho que $\rho < 0$, e que estamos a trabalhar com modelos em (3), para os quais temos a validade de (8), com

$$A(t) = \gamma \beta t^\rho, \quad \gamma > 0, \beta \neq 0, \rho < 0. \tag{9}$$

Sob a validade da condição de segunda ordem em (8), e para k intermédio, i.e., sempre que se verifica (4), conseguimos garantir a normalidade assintótica do estimador de Hill em (6), para valores adequados de k . Na realidade, podemos escrever (de Haan and Peng, 1998),

$$H(k) \stackrel{d}{=} \gamma + \frac{\gamma P_k}{\sqrt{k}} + \frac{A(n/k)}{1 - \rho} (1 + o_p(1)),$$

$P_k = \sqrt{k} \left(\sum_{i=1}^k E_i/k - 1 \right)$, com $\{E_i\}$ v.a.’s i.i.d. e exponenciais unitárias. Consequentemente, mesmo que escolhamos um nível k tal que $\sqrt{k} A(n/k) \rightarrow \lambda \neq 0$, finito, quando $n \rightarrow \infty$, $\sqrt{k} (H(k) - \gamma)$ é assintoticamente normal, com variância igual a γ^2 e um viés não nulo, dado por $\lambda/(1 - \rho)$. Muito frequentemente, este

tipo de estimativas exibem um viés elevado para valores de k moderados, e trajectórias amostrais com regiões de estabilidade diminutas, em torno do valor alvo γ . Este tipo de comportamento tem levado à construção de novos estimadores de viés reduzido, aqui genericamente denotados $\hat{\gamma}_R(k)$, amplamente discutidos em Peng (1998), Beirlant *et al.* (1999), Feuerverger and Hall (1999) e Gomes *et al.* (2000), entre outros. Então, para k intermédio, i.e., sempre que se verifica (4), e sob a validade da condição de segunda ordem em (8), podemos escrever, com P_k^R v.a. assintoticamente normal padrão,

$$\hat{\gamma}_R(k) \stackrel{d}{=} \gamma + \frac{\sigma_R P_k^R}{\sqrt{k}} + o_p(A(n/k)),$$

onde $\sigma_R > 0$, sendo $A(\cdot)$ a função em (8). Consequentemente, a sucessão de v.a.'s, $\sqrt{k}(\hat{\gamma}_R(k) - \gamma)$ é assintoticamente normal com variância σ_R^2 e um valor médio nulo, mesmo quando $\sqrt{k} A(n/k) \rightarrow \lambda \neq 0$, finito, quando $n \rightarrow \infty$.

Gomes and Figueiredo (2003) sugerem que se utilizem, em (5), estimadores do índice de cauda de viés reduzido, tais como os estudados em Gomes and Martins (2001, 2002) e Gomes *et al.* (2004), conseguindo assim também reduzir a componente dominante do viés assintótico do estimador clássico de quantis. Recentemente, Gomes *et al.* (2004) e Caeiro *et al.* (2005) consideram novas classes de estimadores do índice de cauda, dependentes de $(\hat{\beta}, \hat{\rho})$, um estimador consistente adequado de (β, ρ) , o vector dos parâmetros de segunda ordem em (9). A influência de estimadores deste tipo na estimação de quantis foi também estudada em Gomes and Pestana (2005).

Muito recentemente foram propostos, em Caeiro (2005), novos estimadores de C , em que, em vez de se usar unicamente $X_{n-k:n}$, se usa um espaçamento $X_{n-[\theta k]:n} - X_{n-k:n}$, $0 < \theta < 1$. De modo mais específico, podemos estimar C através de

$$\tilde{C}_{\hat{\gamma}}(k; \theta) := \frac{X_{n-[\theta k]:n} - X_{n-k:n}}{\theta^{-\hat{\gamma}} - 1} \left(\frac{k}{n} \right)^{\hat{\gamma}},$$

onde $\theta \in (0, 1)$ é um *parâmetro de controlo*, sendo $\hat{\gamma}$ um qualquer estimador do índice de cauda γ . Caeiro (2005) estuda o termo dominante do viés assintótico de $\tilde{C}_{\hat{\gamma}}(k; \theta)$ e, com a parametrização $A(t) = \gamma \beta t^\rho$, obtem explicitamente esse termo, o qual é dado por

$$B(\gamma, \beta, \rho, \theta) = \frac{(\theta^{-(\gamma+\rho)} - 1)}{\theta^{-\gamma} - 1} \frac{\gamma \beta (n/k)^\rho}{\rho}. \quad (10)$$

É pois sensato considerar o seguinte estimador semi-paramétrico de C :

$$\bar{C}_{\hat{\gamma}}(k; \theta) := \frac{X_{n-[\theta k]:n} - X_{n-k:n}}{\theta^{-\hat{\gamma}} - 1} \left(\frac{k}{n} \right)^{\hat{\gamma}} \times \left(1 - B(\hat{\gamma}, \hat{\beta}, \hat{\rho}, \theta) \right), \quad (11)$$

com $B(\gamma, \beta, \rho, \theta)$ dado em (10).

Iremos aqui considerar o estimador de viés reduzido introduzido em *Cairo et al.* (2005),

$$\overline{H}(k) \equiv \overline{H}_{\hat{\beta}, \hat{\rho}}(k) = H(k) \left(1 - \frac{\hat{\beta}}{1 - \hat{\rho}} \left(\frac{n}{k} \right)^{\hat{\rho}} \right), \quad (12)$$

com $(\hat{\beta}, \hat{\rho})$ estimadores adequados do parâmetro de segunda ordem (β, ρ) , baseados num número de estatísticas ordinais de topo k_1 superior ao número de e.o.'s de topo k usadas para a estimação do índice de cauda. Paralelamente, e por analogia com o que foi feito em *Gomes e Martins* (2002), iremos também considerar um outro possível estimador de γ , com uma expressão funcional semelhante à do estimador em (12), mas em que o estimador de β é calculado no mesmo nível k , que se usa para a estimação de γ . Mais especificamente, trabalharemos também com

$$\overline{\overline{H}}(k) \equiv \overline{\overline{H}}_{\hat{\rho}}(k) = H(k) \left(1 - \frac{\hat{\beta}_{\hat{\rho}}(k)}{1 - \hat{\rho}} \left(\frac{n}{k} \right)^{\hat{\rho}} \right), \quad (13)$$

com

$$\hat{\beta}_{\hat{\rho}}(k) := \left(\frac{k}{n} \right)^{\hat{\rho}} \frac{\left(\frac{1}{k} \sum_{i=1}^k \left(\frac{i}{k} \right)^{-\hat{\rho}} \right) N_n^{(1)}(k) - N_n^{(1-\hat{\rho})}(k)}{\left(\frac{1}{k} \sum_{i=1}^k \left(\frac{i}{k} \right)^{-\hat{\rho}} \right) N_n^{(1-\hat{\rho})}(k) - N_n^{(1-2\hat{\rho})}(k)}, \quad (14)$$

onde

$$N_n^{(\alpha)}(k) := \frac{1}{k} \sum_{i=1}^k \left(\frac{i}{k} \right)^{\alpha-1} U_i,$$

com U_i dado em (7).

Observação 1.1 *Note-se que $\overline{\overline{H}}(k)$ em (13) coincide com $\overline{H}_{\hat{\beta}_{\hat{\rho}}(k), \hat{\rho}}(k)$, sendo $\overline{H}_{\hat{\beta}, \hat{\rho}}(k)$ dado em (12).*

Observação 1.2 *Note-se ainda que, contrariamente ao que acontece na classe de funcionais de Drees (1998), em que a variância assintótica mínima de um estimador do índice de cauda de viés reduzido é dada por $(\gamma(1 - \rho)/\rho)^2$, tal como acontece com o estimador em (13), com o estimador em (12) consegue-se (Cairo et al., 2005) uma variância assintótica igual a γ^2 , a variância assintótica do estimador de Hill, que é por sua vez o estimador de máxima verosimilhança de γ para um modelo de Pareto standard.*

Neste trabalho, e em conjunto com os estimadores $Q_H \equiv Q_H(k; p)$, $Q_{\overline{H}} \equiv Q_{\overline{H}}(k; p)$ e $Q_{\overline{\overline{H}}} \equiv Q_{\overline{\overline{H}}}(k; p)$, com $Q_{\hat{\gamma}}(k; p)$ dado em (5), sendo H , \overline{H} e $\overline{\overline{H}}$, os estimadores do índice de cauda γ fornecidos em (6), (12) e (13), respectivamente, iremos também estudar, por simulação, as novas classe de estimadores

$$\overline{Q}_{\hat{\gamma}} \equiv \overline{Q}_{\hat{\gamma}}(k; p) := \overline{C}_{\hat{\gamma}}(k; 1/2) \times p^{-\hat{\gamma}}, \quad \text{com } \hat{\gamma} = H, \overline{H} \text{ e } \overline{\overline{H}}, \quad (15)$$

sendo $\overline{C}_{\hat{\gamma}}(k; \theta)$ dado em (11). Os estimadores $(\hat{\beta}, \hat{\rho})$ dos parâmetros de segunda ordem irão ser especificados na Secção 2 deste trabalho. Na Secção 3, comparamos os diferentes estimadores por simulação de Monte Carlo, para valores finitos de n . Finalmente, na Secção 4, ilustramos a “performance” destes estimadores através da análise dos log-retornos do Euro versus a Libra inglesa.

2 Estimação dos parâmetros de segunda ordem

Os estimadores de segunda ordem com viés reduzido, introduzidos em (12) e (13), necessitam da consideração prévia de estimadores do parâmetro de segunda ordem (β, ρ) em (9). Iremos aqui considerar mais uma vez os estimadores do parâmetro de segunda ordem ρ propostos em Fraga Alves *et al.* (2003). Essa classe de estimadores pode ser parametrizada num *parâmetro de controlo* real $\tau \in \mathbb{R}$, e depende das estatísticas

$$T_n^{(\tau)}(k) := \begin{cases} \frac{(M_n^{(1)}(k))^\tau - (M_n^{(2)}(k)/2)^{\tau/2}}{(M_n^{(2)}(k)/2)^{\tau/2} - (M_n^{(3)}(k)/6)^{\tau/3}} & \text{se } \tau \neq 0 \\ \frac{\ln(M_n^{(1)}(k)) - \frac{1}{2} \ln(M_n^{(2)}(k)/2)}{\frac{1}{2} \ln(M_n^{(2)}(k)/2) - \frac{1}{3} \ln(M_n^{(3)}(k)/6)} & \text{se } \tau = 0 \end{cases},$$

onde

$$M_n^{(\alpha)}(k) := \frac{1}{k} \sum_{i=1}^k (\ln X_{n-i+1:n} - \ln X_{n-k:n})^\alpha, \quad \alpha > 0.$$

As estatísticas $T_n^{(\tau)}(k)$ convergem para $3(1-\rho)/(3-\rho)$, independentemente de τ , sempre que for válida a condição de segunda ordem em (8), k verificar a condição (4) e $\sqrt{k} A(n/k) \rightarrow \infty$, quando $n \rightarrow \infty$. Os estimadores de ρ que aqui considerámos têm a expressão funcional,

$$\hat{\rho}_\tau(k) \equiv \hat{\rho}_n^{(\tau)}(k) := \min \left(0, \frac{3(T_n^{(\tau)}(k) - 1)}{T_n^{(\tau)}(k) - 3} \right). \quad (16)$$

Observação 2.1 Para um valor apropriado do parâmetro de controlo τ os estimadores de ρ em (16) exibem trajetórias amostrais estáveis para valores elevados de k , o número de e.o.'s de topo usadas. Esse valor de τ pode ser escolhido com base em qualquer critério de estabilidade, como os considerados em Gomes and Pestana (2004), mas a escolha $\tau = 0$ ou $\tau = 1$ conforme $|\rho| \leq 1$ or $|\rho| > 1$ tem sido uma das escolhas aconselhadas do ponto de vista prático. Neste trabalho iremos sempre utilizar o valor $\tau = 0$ em (16), uma vez que ele fornece genericamente bons resultados.

Observação 2.2 Os resultados teóricos e simulados obtidos em Fraga Alves et al. (2003), e o uso destes estimadores de ρ em diferentes estatísticas de viés reduzido, tem frequentemente levado ao aconselhamento da estimação de ρ no nível elevado $k_1 = \min(n - 1, [2n/\ln \ln n])$. Neste trabalho, de forma análoga ao feito em Gomes et al. (2004), e por razões que se prendem com as aí especificadas, iremos considerar a estimação de ρ em

$$k_1 := [n^{0.995}] + 1. \quad (17)$$

Denotaremos genericamente $\hat{\rho}$ qualquer dos estimadores em (16). Sempre que estivermos a usar $\tau = 0$ em (16), usaremos a notação $\hat{\rho}_0$. Para a estimação de β consideraremos o estimador estudado em Gomes and Martins (2002), e já referido em (14). Nas simulações, e ao usar o estimador do índice de cauda em (12), ou estimadores de quantis que lhe estejam associados, trabalharemos com $\hat{\beta}_0 = \hat{\beta}_{\hat{\rho}_0}(k_1)$, com k_1 dado em (17).

3 Comportamento simulado dos estimadores *Var*

Iremos aqui usar a notação $\hat{\rho}_0 = \hat{\rho}_0(k_1)$, $\hat{\beta}_0 = \hat{\beta}_{\hat{\rho}_0}(k_1)$, $\overline{H}_0(k) = \overline{H}_{\hat{\beta}_0, \hat{\rho}_0}(k)$ e $\overline{\overline{H}}_0(k) = \overline{\overline{H}}_{\hat{\rho}_0}(k)$, com $\overline{H}_{\hat{\rho}}(k)$, $\overline{\overline{H}}_{\hat{\beta}, \hat{\rho}}(k)$, $\hat{\beta}_{\hat{\rho}}(k)$, $\hat{\rho}_{\tau}(k)$ e k_1 dados em (12), (13), (14), (16) e (17), respectivamente.

Nas Figuras 1, 2 e 3 apresentamos, para $p = 1/n$ e com base em $N = 5000$ “runs”, os padrões simulados do valor médio, $E[\cdot]$, e da raiz quadrada do erro quadrático médio, $RMSSE[\cdot]$, de $Q_H \equiv Q_H(k; p)$, $Q_{\overline{H}_0} \equiv Q_{\overline{H}_0}(k; p)$, $Q_{\overline{\overline{H}}_0} \equiv Q_{\overline{\overline{H}}_0}(k; p)$, $\overline{Q}_{\overline{H}_0} \equiv \overline{Q}_{\overline{H}_0}(k; p)$ e $\overline{Q}_{\overline{\overline{H}}_0} \equiv \overline{Q}_{\overline{\overline{H}}_0}(k; p)$, com $Q_{\hat{\rho}}(k; p)$ e $\overline{Q}_{\hat{\rho}}(k; p)$ dados em (5) e (15), respectivamente. A Figura 1 está relacionada com o modelo Fréchet com $\gamma = 1$ ($\rho = -1$, $\beta = 0.5$). Os modelos simulados subjacentes às Figuras 2 e 3, são os modelos de Burr(γ, ρ), com ($\gamma = 1$, $\rho = -0.5$) e ($\gamma = 1$, $\rho = -2$), respectivamente. A f.d. Burr(γ, ρ) é dada por $F(x) = 1 - (1 + x^{-\rho/\gamma})^{1/\rho}$, $x > 0$, $\gamma > 0$, $\rho < 0$, sendo a f.d. Fréchet(γ) dada por $F(x) = \exp(-x^{-1/\gamma})$, $x \geq 0$, $\gamma > 0$.

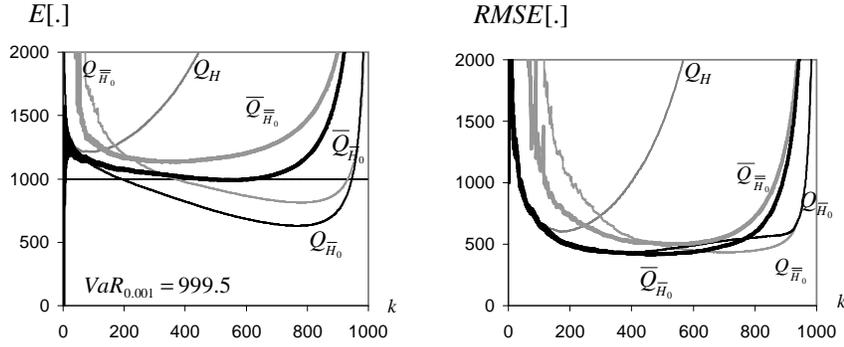


Figura 1: Modelo subjacente: *Fréchet* com $\gamma = 1$ ($\rho = -1$).

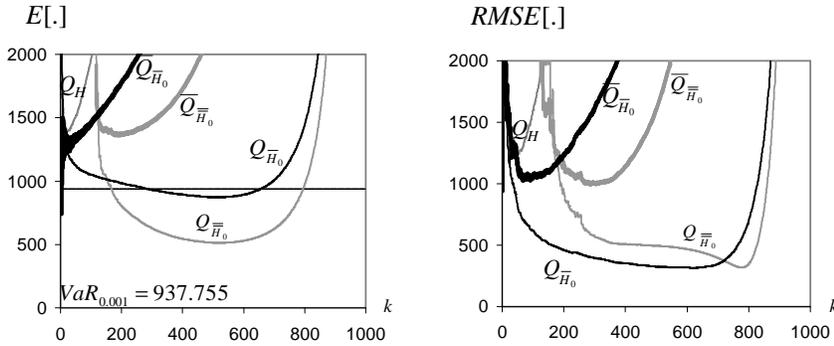


Figura 2: Modelo subjacente: *Burr* com $\gamma = 1$ e $\rho = -.5$.

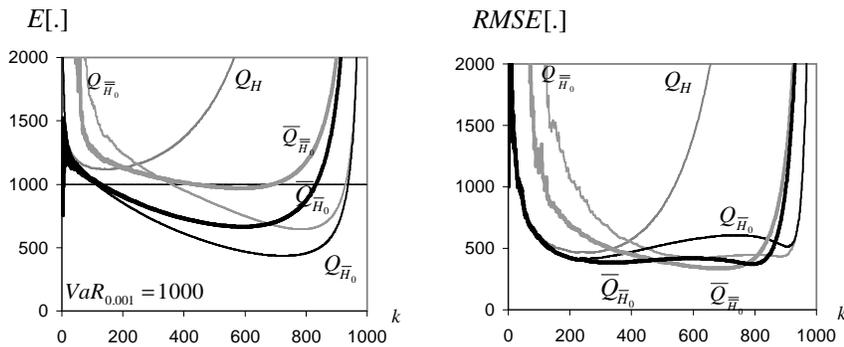


Figura 3: Modelo subjacente: *Burr* com $\gamma = 2$ e $\rho = -2$.

Os resultados obtidos levam-nos a aconselhar a utilização, para $|\rho| \geq 1$, do novo estimador de quantis em (15), com $\hat{\gamma}$ substituído por $\bar{H}_0 = \bar{H}_{\hat{\beta}_0, \hat{\rho}_0}$, o estimador em (12). Para ρ perto de zero, aqui ilustrado na Figura 2, com $\rho = -0.5$, é preferível usar o estimador “clássico” em (5), mas com $\hat{\gamma}$ substituído mais uma vez por \bar{H}_0 , estimador de viés reduzido.

Observação 3.1 *Note-se que, de modo semelhante ao que se passou anteriormente com a estimação do índice de cauda (Gomes et al., 2004; Caeiro et al., 2005), e para muitos dos modelos na classe de Hall em (3), a estimação dos dois parâmetros de segunda ordem β e ρ no nível k_1 em (17), permite-nos trabalhar com estimadores de quantis elevados com um erro quadrático médio inferior ao erro quadrático médio do estimador clássico Q_H , para quase todos os valores de k . Esses estimadores de quantis são fornecidos pelo uso em (5) e (15), do estimador de índice de cauda \bar{H} em (12). Note-se ainda que a utilização do estimador do índice de cauda em (13) nos estimadores de quantis elevados em (5) e (15) pode levar a a erros quadráticos médios mínimos mais pequenos para valores elevados de k , No entanto, estes estimadores são extraordinariamente voláteis para valores pequenos e moderados de k .*

Denotemos genericamente por $\tilde{Q}_\bullet(k; p)$ qualquer dos estimadores (ou v.a.’s) em estudo. Procedemos à simulação de propriedades de $\tilde{Q}_\bullet(\tilde{k}_0^*; p)$, com $\tilde{k}_0^* := \arg \min_k MSE [\tilde{Q}_\bullet(k; p)]$, usando o método multi-amostra com dimensão 5000×10 . Para detalhes sobre simulação multi-amostra, veja-se, por exemplo, Gomes and Oliveira (2001).

Tabela 1: Valores Médios / RMSE’s, escalados por VaR_p , dos estimadores indicados, calculados nos seus níveis óptimos, para $p = 1/n$ e populações subjacentes Fréchet com $\gamma = 1$.

n	Q_H	$Q_{\bar{H}_0}$	$Q_{\bar{H}_0}$	$\bar{Q}_{\bar{H}_0}$	$\bar{Q}_{\bar{H}_0}$
Fréchet ($\gamma = 1$)					
200	1.4015 / 1.0204	0.7402 / <u>0.6225</u>	0.7683 / 0.6628	<u>1.0016</u> / 0.6281	1.2667 / 0.8194
500	1.3156 / 0.7603	0.8016 / 0.5163	0.8058 / 0.5314	<u>0.9980</u> / <u>0.5075</u>	1.1985 / 0.6220
1000	1.2618 / 0.6111	0.8582 / 0.4319	0.8281 / 0.4402	<u>0.9985</u> / <u>0.4210</u>	1.1781 / 0.5056
2000	1.2162 / 0.4947	0.8899 / 0.3530	0.8419 / 0.3605	<u>0.9894</u> / <u>0.3366</u>	1.1435 / 0.4074
5000	1.1673 / 0.3790	0.9172 / 0.2655	0.8620 / 0.2670	<u>0.9798</u> / <u>0.2437</u>	1.1216 / 0.3083

Na Tabela 1 apresentamos uma ilustração parcial dos resultados obtidos, para um modelo subjacente Fréchet com $\gamma = 1$, para amostras de dimensão $n = 200, 500, 1000, 2000$ e 5000 , e $p = 1/n$. São aí apresentados os valores médios e

os *RMSE*'s dos estimadores Q_H , $Q_{\bar{H}_0}$, $Q_{\bar{H}_0}$, $\bar{Q}_{\bar{H}_0}$ e $\bar{Q}_{\bar{H}_0}$, divididos por Var_p , e calculados nos seus níveis óptimos. Por entre os estimadores e para cada valor de n , sublinhamos as entradas relacionadas com menor viés e menor erro médio quadrático.

4 Uma aplicação a dados reais

Iremos aqui ilustrar o comportamento dos estimadores em estudo através da análise das taxas de câmbio do Euro versus a Libra Inglesa desde Janeiro 4, 1999, até Dezembro 14, 2004. Estes dados foram coleccionados pelo Sistema Europeu de Bancos Centrais e foram obtidos em <http://www.bportugal.pt/rates/cambtx/>.

Na Figura 4, a análise dos $n_0 = 725$ log-retornos positivos, permitiu-nos obter as trajectórias amostrais seguintes, para os estimadores dos parâmetros de segunda ordem ρ e β em (16) e (14).

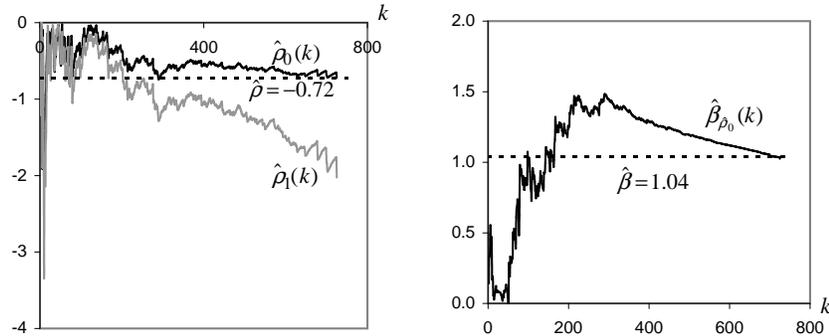


Figura 4: Estimativas do parâmetro de “forma” de segunda ordem ρ , através de $\hat{\rho}_\tau(k)$ em (16), $\tau = 0$ e 1 (*esquerda*) e do parâmetro de “escala” de segunda ordem β , através de $\hat{\beta}_{\hat{\rho}_0}(k)$ em (14) (*direita*), para os Log-retornos diários do Euro versus a Libra Inglesa.

As trajectórias amostrais dos estimadores de ρ associados a $\tau = 0$ e $\tau = 1$ em (16) levam-nos, com base em qualquer critério de estabilidade para valores elevados de k , à escolha da estimativa associada a $\tau = 0$. Obtivemos $\hat{\rho}_0 = \hat{\rho}_0(702) = -0.72$. A utilização de $\hat{\beta}_{\hat{\rho}_0}(k)$ em (14), calculado também no nível k_1 em (17), conduz-nos à estimativa $\hat{\beta}_0 = 1.04$.

As trajectórias amostrais associadas ao estimador clássico H em (6), e ao estimador de viés reduzido \bar{H} em (12), bem como dos estimadores Var associados, genericamente explicitados em (5) e em (15), e para $p = 0.001$, são apresentados na Figura 5. Para $p = 0.001$, qualquer critério de estabilidade

para valores moderados de k leva-nos à escolha do estimador $\overline{Q}_{\overline{H}}$ e à estimativa 2.87 for $VaR_{0.001}$, ilustrada também na Figura 5.

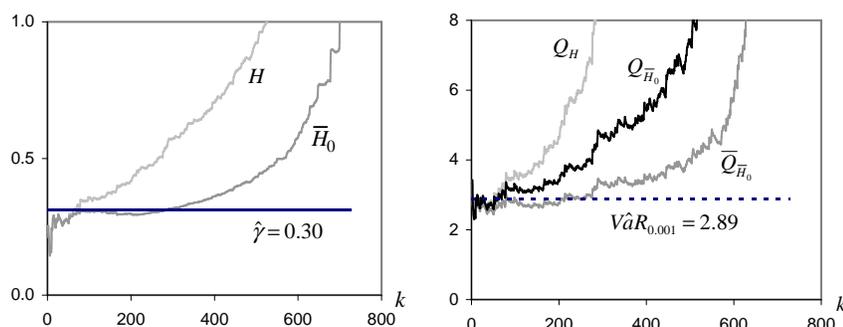


Figura 5: Estimativas fornecidas pelos estimadores de Hill, $H(k)$ e $\overline{H}(k)$ em (6) e (12), respectivamente (*esquerda*) e estimativas de $\ln-VaR_p$ fornecidas pelos estimadores em (5), com $\hat{\gamma}$ substituído por $H(k)$ e $\overline{H}(k)$ e pelo estimador de quantis em (15), com $\hat{\gamma}$ substituído por $\overline{H}(k)$, para os Log-retornos diários do Euro versus a Libra Inglesa, e para $p = 0.001$.

Agradecimentos

Investigação parcialmente financiada por FCT / POCTI e POCI / FEDER.

Referências

- [1] Beirlant, J., Dierckx, G., Goegebeur, Y. and Matthys, G. (1999). Tail index estimation and an exponential regression model. *Extremes* **2**, 177-200.
- [2] Caeiro, F. (2005). *Redução de viés na estimação semi-paramétrica de um parâmetro de escala*. Actas do XIII Congresso da SPE.
- [3] Caeiro, F., Gomes, M. I. and Pestana, D. (2005). Direct reduction of bias of the classical Hill estimator. *RevStat* **3**(2), 113-136.
- [4] Drees, H. (1998). A general class of estimators of the extreme value index. *J. Statist. Planning and Inference* **98**, 95-112.
- [5] Feuerverger, A. and P. Hall (1999). Estimating a tail exponent by modelling departure from a Pareto distribution. *Ann. Statist.* **27**, 760-781.
- [6] Fraga Alves, M. I., Gomes M. I. and de Haan, L. (2003). A new class of semi-parametric estimators of the second order parameter. *Portugaliae Mathematica* **60**:2, 194-213.
- [7] Geluk, J. and L. de Haan (1987). *Regular Variation, Extensions and Tauberian Theorems*. CWI Tract 40, Center for Mathematics and Computer Science, Amsterdam, Netherlands.

- [8] Gnedenko, B. V. (1943). Sur la distribution limite du terme maximum d'une série aléatoire. *Ann. Math.* **44**, 423-453.
- [9] Gomes, M. I., Caeiro, F. and Figueiredo, F. (2004). Bias reduction of a tail index estimator through an external estimation of the second order parameter. *Statistics* **38**(6), 497-510.
- [10] Gomes, M. I. and Figueiredo, F. (2003). Bias reduction in risk modelling: semi-parametric quantile estimation. To appear at *Test*.
- [11] Gomes, M. I. and M. J. Martins (2001). Alternatives to Hill's estimator — asymptotic versus finite sample behaviour. *J. Statist. Planning and Inference* **93**, 161-180.
- [12] Gomes, M. I. and M. J. Martins (2002). "Asymptotically unbiased" estimators of the tail index based on external estimation of the second order parameter. *Extremes* **5**:1, 5-31.
- [13] Gomes, M. I., Martins, M. J. and M. Neves (2000). Alternatives to a semi-parametric estimator of parameters of rare events — the Jackknife methodology. *Extremes* **3**:3, 207-229.
- [14] Gomes, M. I. and O. Oliveira (2001). The bootstrap methodology in Statistical Extremes — choice of the optimal sample fraction. *Extremes* **4**:4, 331-358.
- [15] Gomes, M. I. and Pestana (2004). A simple second order reduced bias' tail index estimator. To appear at *J. Statist. Comp. and Simul.*
- [16] Gomes, M. I. and Pestana (2005). *A sturdy second order reduced bias' Value at Risk estimator*. Notas e Comunicações 05/2005. Submitted.
- [17] Haan, L. de (1970). *On Regular Variation and its Application to the Weak Convergence of Sample Extremes*. Mathematical Centre Tract 32, Amsterdam.
- [18] Haan, L. de and Peng, L. (1998). Comparison of tail index estimators. *Statistica Neerlandica* **52**, 60-70.
- [19] Hall, P. and Welsh, A.H. (1985). Adaptive estimates of parameters of regular variation. *Ann. Statist.* **13**, 331-341.
- [20] Hill, B. M. (1975). A simple general approach to inference about the tail of a distribution. *Ann. Statist.* **3**, 1163-1174.
- [21] Peng, L.(1998). Asymptotically unbiased estimator for the extreme-value index. *Statistics and Probability Letters* **38**(2), 107-115.
- [22] Weissman, I. (1978). Estimation of parameters and large quantiles based on the k largest observations. *J. Amer. Statist. Assoc.* **73**, 812-815.