

Caudas pesadas: t de Student e variante assimétrica versus metodologia semi-paramétrica

Frederico Caeiro

DM e CMA, Universidade Nova de Lisboa - fac@fct.unl.pt

M. Ivette Gomes

DEIO e CEAUL, Universidade de Lisboa - ivette.gomes@fc.ul.pt

Resumo: Neste trabalho comparamos, por simulação, dois modelos usados para modelar o comportamento da cauda da distribuição dos log-retornos de dados financeiros. Numa perspectiva paramétrica utilizamos uma generalização assimétrica do modelo t de *Student* e fazemos a estimação por máxima verosimilhança. Num contexto semi-paramétrico, e de caudas pesadas, utilizamos alguns dos mais recentes estimadores de viés reduzido.

Palavras-chave: Distribuição t de *Student* assimétrica, Cauda pesada, Índice de valores extremos.

Abstract: In this work we compare, by simulation, two models used for the log returns distribution's tail. The first one is the a variant from the Student's t distribution which could be skewed and the inference will be done by maximum likelihood. The second models is a semi parametric method where we will use some of the most recent reduced bias estimators.

Keywords: Skew t distribution, Heavy tail, Extreme value index.

1 Introdução

Os log-retornos de muitos índices financeiros, como por exemplo taxas de câmbio, cotações de títulos ou taxas de juro, medidas em intervalos de tempo muito curto (diariamente ou semanalmente) não apresentam um comportamento normal nas caudas. A distribuição, F , subjacente aos log-retornos pode ser ligeiramente assimétrica, e as caudas são geralmente mais pesadas que as caudas da distribuição normal. Consequentemente, valores elevados ocorrem com maior probabilidade do que sob o modelo normal. Dizemos que F é um modelo com cauda (direita) pesada se e só se,

$$\lim_{t \rightarrow \infty} \frac{1 - F(tx)}{1 - F(t)} = x^\alpha, \quad \alpha > 0, \quad (1)$$

onde $\alpha > 0$ é o índice de cauda¹. A $\gamma = 1/\alpha$ dá-se o nome de índice de valores extremos. Propomos comparar duas metodologias distintas, usadas para estimar o índice de valores extremos da cauda direita. Por um lado uma metodologia semi-paramétrica que apenas pressupõe que o modelo tenha cauda pesada, por outro lado a metodologia paramétrica de estimação por máxima verosimilhança. Neste trabalho começamos por descrever as metodologias paramétrica e semi-paramétrica nas secções 2 e 3, respectivamente. Na secção 4 comparamos o comportamento das duas metodologias por simulação de Monte Carlo. Finalmente, na secção 5, apresentamos um aplicação a dados financeiros.

2 Estimação Paramétrica

Recentemente, Ferguson and Platen (2006) sugeriram que o clássico modelo t de *Student*, com aproximadamente 4 “graus de liberdade”, representa adequadamente o comportamento de muitos índices de retorno. Como existe a possibilidade dos log-retornos apresentarem uma ligeira assimetria, outros autores aconselham a utilização de outros modelos, como por exemplo, generalizações do modelo t que possuem um parâmetro adicional associado à sua assimetria. Podemos encontrar na literatura vários destes modelos, como por exemplo os modelos propostos por Fernandez and Steel (1998), Azzalini and Capitanio (2003) e Jones and Faddy (2003). Na metodologia paramétrica, vamos apenas considerar o modelo *Skew-t* introduzido em Jones and Faddy (2003), por ser o único que permite que os índices de valores extremos das duas caudas não sejam iguais, ou seja, vamos utilizar o modelo com função densidade:

$$f(x) = \frac{1}{2^{a+b-1} \mathcal{B}(a, b) \sqrt{a+b}} \left(1 + \frac{x}{\sqrt{a+b+x^2}}\right)^{a+1/2} \left(1 - \frac{x}{\sqrt{a+b+x^2}}\right)^{b+1/2},$$

onde $a > 0$ e $b > 0$ são parâmetros de forma, e \mathcal{B} é a função beta. Quando $a = b$, f é a densidade do modelo t de *Student* clássico com $2a$ graus de liberdade. A Figura 1 tem esta função densidade, representada para vários valores de (a, b) .

Observação: Trata-se de um modelo de cauda pesada com índice de valores extremos, da cauda direita, igual a $\gamma = \frac{1}{2b}$.

Vamos considerar o modelo *Skew-t* com parâmetros de localização e escala, isto é,

$$f(x; \mu, \sigma, a, b) = \frac{1}{\sigma} f\left(\frac{x-\mu}{\sigma}; a, b\right), \quad \mu \in \mathbb{R}, \quad \sigma > 0, \quad a > 0, \quad b > 0.$$

Como não existe solução explícita para os estimadores de máxima verosimilhança, as estimativas são obtidas através de métodos numéricos. Este método foi implementado no software R utilizando o comando `fitdistr` da biblioteca MASS. Devido à propriedade invariante, o estimador de máxima verosimilhança do índice de valores extremos é $\hat{\gamma}_{MLE} = \frac{1}{2b}$.

¹A definição de cauda pesada pode ser adaptada para a cauda esquerda

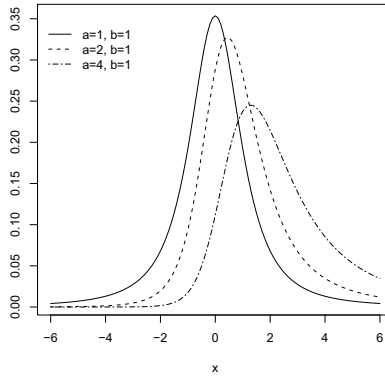


Figura 1: Função densidade do modelo *Skew-t*, para distintos valores de (a, b) .

3 Estimação Semi-paramétrica

Sejam X_1, X_2, \dots, X_n v.a.'s i.i.d. com função de distribuição F . Na abordagem semi-paramétrica, a inferência é feita a partir das $k + 1$ maiores observações, onde k representa uma sequência intermédia de valores inteiros em $[1, n]$, isto é,

$$k = k_n \rightarrow \infty, \quad k/n \rightarrow 0, \quad n \rightarrow \infty.$$

3.1 Estimadores do índice de valores extremos

Para caudas pesadas, o estimador clássico do índice de valores extremos é o estimador de Hill (1975),

$$H(k) = \hat{\gamma}_n^H(k) := \frac{1}{k} \sum_{i=1}^k \ln X_{n-i+1:n} - \ln X_{n-k:n},$$

onde $X_{1:n} \leq X_{2:n} \leq \dots \leq X_{n:n}$. Este estimador tem algumas propriedades, comuns a muitos outros estimadores clássicos do índice de valores extremos:

- variância elevada para pequenos valores de k ;
- viés elevado para pequenos valores de k ;
- pequena região de estabilidade da trajectória amostral de $H(k)$ vs k , junto ao verdadeiro valor do índice de valores extremos.

Devido a estas propriedades, nos últimos anos vários autores propuseram outros estimadores de viés reduzido, isto é, estimadores com componente dominante de viés assintótico nula. A trajectória das estimativas, fornecidas pelos novos estimadores (enquanto função de k), costuma ser mais estável, junto ao

valor do índice de valores extremos.

De modo a conseguir conhecer e remover o termo dominante de viés, é necessário considerar uma condição de segunda ordem relativa à velocidade de convergência de $\frac{1-F(tx)}{1-F(t)}$ para $x^{-1/\gamma}$, ou equivalentemente, relativa à velocidade de convergência de $U(tx)/U(t)$ para x^γ , onde $U(t) := F^\leftarrow(1-1/t) = \inf\{x : F(x) \geq 1-1/t\}$, $t > 1$ (de Haan, 1970). Assim, vamos admitir que

$$\lim_{t \rightarrow \infty} \frac{\ln U(tx) - \ln U(t) - \gamma \ln x}{A(t)} = \frac{x^\rho - 1}{\rho}, \quad \forall x > 0, \quad (2)$$

onde $A(t)$ é uma função tal que $|A(tx)/A(t)| \rightarrow x^\rho$, $\forall x > 0$ e $\rho \leq 0$ (Geluk and de Haan, 1987). Vamos ainda considerar que F pertence à classe de Hall (Hall 1982; Hall and Welsh, 1985) e assim, na equação (2), $A(t) := \gamma\beta t^\rho$ e $\rho < 0$.

Observação: A maioria dos modelos de cauda pesada, como por exemplo os modelos *Fréchet*, *Burr*, *GP*, t de *Student* e *Skew-t*, pertencem a esta classe.

Proposição 3.1. Usando a notação $A(t) = \gamma\beta t^\rho$, para k intermédio, e F na classe de Hall,

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

onde Z_k é uma v.a. assintoticamente normal padrão.

A componente dominante de viés do estimador de Hill é $A(n/k)/(1-\rho)$ e pode ser estimada e removida através de dos estimadores introduzidos em Caeiro *et al.* (2005) e em Gomes *et al.* (2007), respectivamente,

$$\overline{H}(k) := H(k) \left(1 - \frac{\hat{\beta}}{1-\hat{\rho}} \left(\frac{n}{k}\right)^{\hat{\rho}}\right), \quad (3)$$

e

$$\overline{M}(k) := H(k) - \hat{\beta} \left(\frac{n}{k}\right)^{\hat{\rho}} N_n^{(1-\hat{\rho})}(k) \quad (4)$$

onde $\hat{\rho}$ e $\hat{\beta}$ são estimadores consistentes dos parâmetros de segunda ordem de ρ e β , respectivamente,

$$N_n^{(\alpha)}(k) := \frac{1}{k} \sum_{i=1}^k \left(\frac{i}{k}\right)^{\alpha-1} U_i \quad (5)$$

são estimadores consistentes de $\frac{\gamma}{\alpha}$ e U_i são os espaçamentos escalados das log-observações definidos por $U_i := i(\ln X_{n-i+1:n} - \ln X_{n-i:n})$.

3.2 Escolha do nível k

Para modelos pertencentes à classe de Hall, o nível k óptimo que minimiza o erro quadrático médio assintótico, do estimador de Hill, é dado por:

$$k_0^H := \arg \min_k \left(\frac{\gamma^2}{k} + \frac{\gamma\beta(n/k)^\rho}{(1-\rho)^2} \right) = \left(\frac{(1-\rho)n^{-\rho}}{\beta\sqrt{-2\rho}} \right)^{2/(1-2\rho)} \quad (6)$$

Para os estimadores de viés reduzido, a escolha do nível k óptimo é um tema ainda em aberto. Propomos a utilização da heurística sugerida em Gomes and Pestana (2007), como escolha adaptativa. Esta escolha corresponde ao nível onde a trajectória do estimador de Hill sai da região de 95% de confiança de um dos estimadores de viés reduzido, que corresponde à solução da equação, $|\beta|(n/k)^\rho = 1.96/\sqrt{k}$, ou seja, ao nível

$$k_{01} = (1.96(1 - \rho)n^{-\rho}/|\beta|)^{2/(1-2\rho)}. \tag{7}$$

3.3 Estimação dos parâmetros de segunda ordem ρ e β

Para estimar ρ , vamos utilizar a classe de estimadores propostos em Fraga Alves *et al.* (2003),

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

onde $M_n^{(\alpha)}(k) := \frac{1}{k} \sum_{i=1}^k (\ln X_{n-i+1:n} - \ln X_{n-k:n})^\alpha$, $\alpha > 0$, e $\tau \in \mathbb{R}$ é um parâmetro de controlo. 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 (2), e $\sqrt{k} A(n/k) \rightarrow \infty$, quando $n \rightarrow \infty$. O estimador de ρ é,

$$\widehat{\rho}_\tau(k) \equiv \widehat{\rho}_n^{(\tau)}(k) := - \left| 3(T_n^{(\tau)}(k) - 1)/(T_n^{(\tau)}(k) - 3) \right|. \tag{8}$$

Observação 3.1. *O valor de τ pode ser escolhido com base em qualquer critério de estabilidade, 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.*

Observação 3.2. *Neste trabalho, iremos considerar a estimação de ρ em*

$$k_1 := [n^{0.995}]. \tag{9}$$

Para a estimação de β consideramos o estimador proposto em Gomes and Martins (2002),

$$\widehat{\beta}_{\widehat{\rho}}(k) := \left(\frac{k}{n}\right)^{\widehat{\rho}} \frac{\left(\frac{1}{k} \sum_{i=1}^k \left(\frac{i}{k}\right)^{-\widehat{\rho}}\right) N_n^{(1)}(k) - N_n^{(1-\widehat{\rho})}(k)}{\left(\frac{1}{k} \sum_{i=1}^k \left(\frac{i}{k}\right)^{-\widehat{\rho}}\right) N_n^{(1-\widehat{\rho})}(k) - N_n^{(1-2\widehat{\rho})}(k)}, \tag{10}$$

com $\widehat{\rho} = \widehat{\rho}_\tau(k)$ e $N_n^{(\alpha)}(k)$ definido em (5). Nas simulações, e ao usar os estimadores do índice de valores extremos (3) e (4), trabalhamos com $\widehat{\beta} = \widehat{\beta}_{\widehat{\rho}}(k_1)$, com k_1 dado em (9).

4 Estudo de simulação

Nesta secção comparamos, por simulações de Monte Carlo, o comportamento dos vários estimadores do índice de valores extremos (semi-paramétricos e o paramétrico de máxima verosimilhança para o modelo *Skew-t*). Os resultados foram obtidos com base em 1000 “runs” para amostras de dimensão $n = 500$, 1000 e 5000 do modelo *Skewt-t(a,b)* e do modelo *Skewt-t(a,b)* contaminado com 10% das observações proveniente de outro modelo com o mesmo índice de valores extremos. Escolhemos $a = 1$ e $b \in \{0.5, 1, 2\}$.

O índice de valores extremos foi estimado, no nível \hat{k}_0 , no estimador de Hill, e nos níveis \hat{k}_0 e $\hat{k}_{0,1}$, nos estimadores de viés reduzido \overline{H} e \overline{M} em (3) e (4), respectivamente. Consideramos os parâmetros de segunda ordem ρ e β desconhecidos e fizemos a sua estimação em k_1 , definido em (9).

Na Tabela 1 apresentamos os resultados do valor médio e raiz quadrada do erro quadrático médio (RMSE), para o modelo subjacente *Skew-t* e *Skew-t* com 10% de contaminação. Para facilitar a comparação, o valor médio associado à mais pequena RMSE está sublinhado. Quando o modelo subjacente é o *Skew-t*, o método paramétrico dá-nos quase sempre o mais pequeno valor de RMSE. Quando comparamos os resultados do modelo sem e com contaminação concluímos que a contaminação não afecta muito o desempenho dos estimadores semi-paramétricos. Para vários valores dos parâmetros do modelo com contaminação, os estimadores semi-paramétricos são os que têm melhor desempenho.

5 Aplicação a dados financeiros

Nesta secção analisamos os $n = 2049$ log-retornos da taxa de câmbio do Euro / Franco Suiço, entre 4 de Janeiro de 1999 e 29 de Dezembro de 2006. Na Figura 2, apresentamos os gráficos da taxa de câmbio diária, dos log-retornos diários e o gráfico de quantis normais dos log-retornos diários.

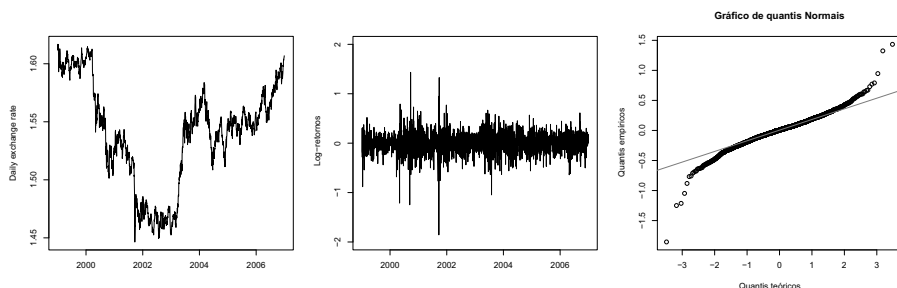


Figura 2: **Esquerda:** Taxa de câmbio diária do Euro/Franco Suiço; **Centro:** Log-retorno diário ($\times 100$) da taxa de câmbio Euro vs Franco Suiço; **Direita:** Gráfico de quantis normais dos log-retornos ($\times 100$) da taxa de câmbio Euro / Franco Suiço.

Tabela 1: Valor médio simulado (Raiz quadrada do erro quadrático médio simulado), dos estimadores indicados e populações subjacentes *Skew-t* e *Skew-t* contaminada.

$n =$	500		1000		5000	
Skew-t(1,0.5) ($\gamma = 1, \rho = -2$)						
<i>MLE</i>	<u>0.9964</u>	(0.0910)	<u>1.0006</u>	(0.0619)	<u>1.0020</u>	(0.0268)
$H(\hat{k}_0)$	1.0796	(0.1252)	1.0560	(0.0913)	1.0295	(0.0466)
$\overline{M}(\hat{k}_0)$	1.0307	(0.0991)	1.0205	(0.0737)	1.0118	(0.0378)
$\overline{M}(\hat{k}_{01})$	1.0839	(0.1178)	1.0626	(0.0882)	1.0330	(0.0445)
$\overline{H}(\hat{k}_0)$	1.0342	(0.1005)	1.0224	(0.0744)	1.0123	(0.0380)
$\overline{H}(\hat{k}_{01})$	1.1227	(0.1572)	1.0821	(0.1044)	1.0373	(0.0480)
Skew-t(1,0.5) com 10% de contaminação ($\gamma = 1$)						
<i>MLE</i>	<u>0.9537</u>	(0.0997)	<u>0.9583</u>	(0.0758)	0.9580	(0.0495)
$H(\hat{k}_0)$	1.0823	(0.1283)	1.0618	(0.0976)	1.0350	(0.0506)
$\overline{M}(\hat{k}_0)$	1.0334	(0.1018)	1.0260	(0.0785)	<u>1.0172</u>	(0.0402)
$\overline{M}(\hat{k}_{01})$	1.0876	(0.1224)	1.0676	(0.0932)	1.0393	(0.0493)
$\overline{H}(\hat{k}_0)$	1.0370	(0.1034)	1.0280	(0.0794)	1.0178	(0.0404)
$\overline{H}(\hat{k}_{01})$	1.1281	(0.1652)	1.0875	(0.1099)	1.0439	(0.0531)
Skew-t(1,1) ($\gamma = 0.5, \rho = -1$)						
<i>MLE</i>	0.4921	(0.0630)	<u>0.4970</u>	(0.0420)	<u>0.4995</u>	(0.0192)
$H(\hat{k}_0)$	0.5565	(0.1055)	0.5409	(0.0791)	0.5220	(0.0469)
$\overline{M}(\hat{k}_0)$	0.4762	(0.0817)	0.4788	(0.0648)	0.4857	(0.0415)
$\overline{M}(\hat{k}_{01})$	0.4800	(0.0596)	0.4800	(0.0491)	0.4821	(0.0332)
$\overline{H}(\hat{k}_0)$	0.4806	(0.0813)	0.4810	(0.0643)	0.4863	(0.0414)
$\overline{H}(\hat{k}_{01})$	<u>0.4921</u>	(0.0584)	0.4866	(0.0475)	0.4838	(0.0324)
Skew-t(1,1) com 10% de contaminação ($\gamma = 0.5$)						
<i>MLE</i>	0.4546	0.0748	0.4592	0.0571	0.4617	0.0426
$H(\hat{k}_0)$	0.5509	0.1017	0.5396	0.0800	0.5211	0.0480
$\overline{M}(\hat{k}_0)$	0.4726	0.0829	0.4783	0.0667	0.4853	0.0432
$\overline{M}(\hat{k}_{01})$	0.4695	0.0636	0.4736	0.0514	0.4799	0.0343
$\overline{H}(\hat{k}_0)$	0.4767	0.0820	0.4804	0.0663	0.4858	0.0431
$\overline{H}(\hat{k}_{01})$	<u>0.4802</u>	0.0607	<u>0.4795</u>	0.0492	<u>0.4815</u>	0.0334
Skew-t(1,2) ($\gamma = 0.25, \rho = -0.5$)						
<i>MLE</i>	<u>0.2437</u>	(0.0449)	<u>0.2474</u>	(0.0312)	<u>0.2494</u>	(0.0146)
$H(\hat{k}_0)$	0.4232	(0.1900)	0.3897	(0.1518)	0.3350	(0.0913)
$\overline{M}(\hat{k}_0)$	0.3389	(0.1087)	0.3276	(0.0923)	0.3025	(0.0605)
$\overline{M}(\hat{k}_{01})$	0.3754	(0.1359)	0.3596	(0.1171)	0.3237	(0.0772)
$\overline{H}(\hat{k}_0)$	0.3480	(0.1179)	0.3331	(0.0977)	0.3042	(0.0622)
$\overline{H}(\hat{k}_{01})$	0.4003	(0.1614)	0.3739	(0.1317)	0.3280	(0.0815)
Skew-t(1,2) com 10% de contaminação ($\gamma = 0.25$)						
<i>MLE</i>	<u>0.2218</u>	0.0536	<u>0.2250</u>	0.0415	<u>0.2277</u>	0.0268
$H(\hat{k}_0)$	0.3613	0.1311	0.3429	0.1074	0.3166	0.0738
$\overline{M}(\hat{k}_0)$	0.2922	0.0703	0.2900	0.0610	0.2861	0.0462
$\overline{M}(\hat{k}_{01})$	0.2889	0.0552	0.2897	0.0512	0.2900	0.0450
$\overline{H}(\hat{k}_0)$	0.2977	0.0750	0.2934	0.0638	0.2873	0.0473

O gráfico de quantis normais dos log-retornos diários, da Figura 2, mostra que a distribuição subjacente tem caudas mais pesadas que as da distribuição normal. Por isso, parece sensato considerar a hipótese da distribuição subjacente aos log-retornos ser a *Skew-t*.

5.1 Estimação paramétrica (Máxima Verosimilhança)

Vamos assim considerar que a distribuição *Skew-t* é adequada para modelar os log-retornos ($\times 100$). Utilizando o software estatístico R, para estimar por máxima verosimilhança os parâmetros, obtemos:

μ	σ	a	b
0.026551962	0.165135302	2.267396625	2.514174955
(0.012585340)	(0.004473282)	(0.241624579)	(0.280802537)

Aplicando a propriedade de invariância dos estimadores de máxima verosimilhança, obtemos para a estimativa do índice de valores extremos da cauda direita, $\hat{\gamma}_{MLE} = 0.1989$. Apresentamos, na Figura 3, a qualidade do ajustamento do modelo *Skew-t* estimado. Apesar deste modelo fornecer uma melhor aproximação do que o modelo normal, nas caudas continuam a existir algumas diferenças entre o modelo ajustado e as maiores/menores observações.

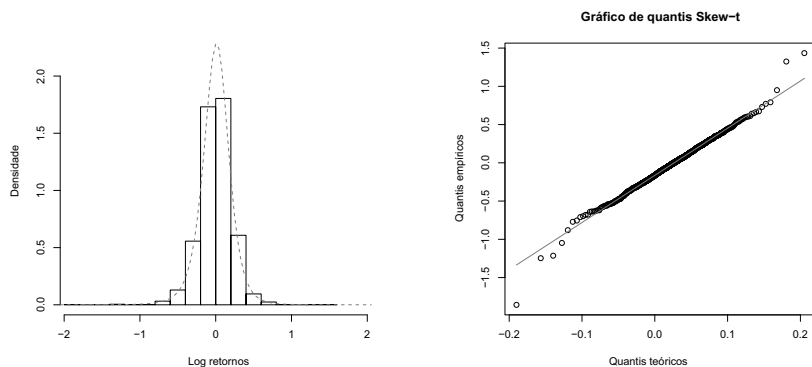


Figura 3: **Esquerda:** Histograma dos log-retornos da taxa de câmbio do Euro vs Franco Suiço, com a densidade *Skew-t* estimada; **Direita:** Gráfico de quantis do modelo *Skew-t* estimado e dos log-retornos da taxa de câmbio.

5.2 Estimação Semi-paramétrica

Como apenas pretendemos estimar o índice de valores extremos da cauda direita, vamos trabalhar com a sub-amostra dos log-retornos positivos. O número de log retornos positivos na amostra é $n^+ = 1040$. Apresentamos na Figura 4

as trajectórias amostrais das estimativas dos parâmetros de segunda ordem ρ e β . Escolhemos a trajectória amostral de $\hat{\rho}$ com $\tau = 0$ devido à sua maior estabilidade. Utilizando $\tau = 0$ e $k_1 = n^{0.995}$, obtemos as estimativas $(\hat{\rho}_0, \hat{\beta}_0) = (-0.650, 1.042)$.

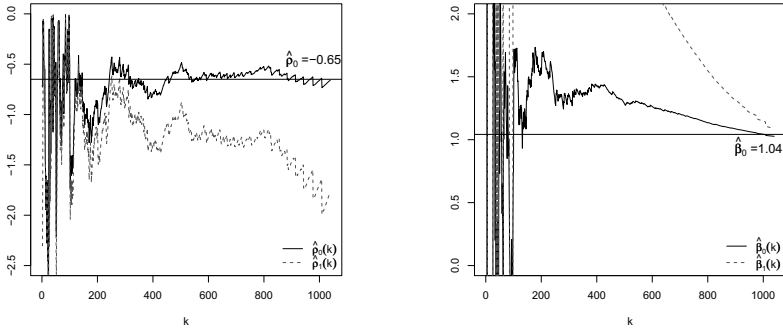
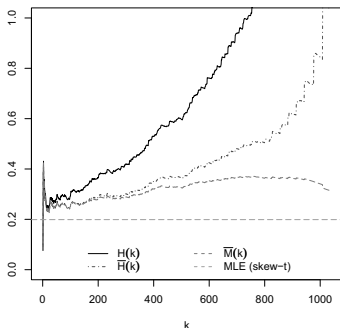


Figura 4: Trajectória amostral das estimativas dos parâmetros de segunda ordem ρ (esquerda) e β (direita) para os log-retornos positivos da taxa de câmbio do Euro vs Franco Suíço.

A Figura 5 contém as trajectórias amostrais dos estimadores semi-paramétricos do índice de valores extremos, para todos os possíveis valores de k . Incluímos também a estimativa obtida por máxima verosimilhança (que não depende do valor k), de modo a permitir a sua comparação. Junto ao gráfico da Figura 5, está a tabela com as estimativas fornecidas pelos vários estimadores semi-paramétricos. Podemos concluir que as estimativas fornecidas pelos estimadores semi-paramétricos não diferem muito e são um pouco superiores à estimativa obtida por máxima verosimilhança.



	H	\bar{H}	\bar{M}
$\hat{k}_0 = 67$	0.294	0.263	0.261
$\hat{k}_{01} = 136$		0.263	0.260

Figura 5: Trajectórias amostrais de H , \bar{H} e \bar{M} .

Referências

- [1] Azzalini, A. e Capitanio, A. (2003). Distributions generated by perturbation of symmetry with emphasis on a multivariate skew t-distribution. *Journal of Royal Statistical Society B*, 65(2), 367-389.
- [2] Caeiro, F., Gomes, M. I. e Pestana, D. (2005). Direct reduction of bias of the classical Hill estimator. *Revstat*, 3(2), 113-136.
- [3] Ferguson, K., e Platen, E. (2006), On the Distributional Characterization of daily Log-returns of a World Stock Index. *Applied Mathematical Finance*, 13(1), 19-38.
- [4] Fernandez, C. e Steel, M.F.J. (1998). On Bayesian modeling of fat tails and skewness. *Journal of the American Statistical Association*, 93, 359-371.
- [5] Fraga Alves, M.I., Gomes, M. I., e de Haan, L. (2003). A new class of semi-parametric estimators of the second order parameter. *Portugaliae Mathematica*, 60 (1), 193-213.
- [6] Geluk, J. e de Haan, L. (1987). *Regular Variation, Extensions and Tauberian Theorems*. Tech. Report CWI Tract 40, Centre for Mathematics and Computer Science, Amsterdam, Netherlands.
- [7] Gomes, M.I., e Martins, M.J. (2002). “Asymptotically Unbiased” Estimators of the Tail Index Based on External Estimation of the Second Order Parameter. *Extremes*, 5(1), 5-31.
- [8] Gomes, M.I., Martins, M.J., e Neves, M. (2007). Improving second order reduced-bias tail index estimation. *Revstat*, 5(2), 177-207.
- [9] Gomes, M.I. e Pestana, D. (2007). A Sturdy Reduced-Bias Extreme Quantile (VaR) Estimator. *Journal of the American Statistical Association*, 102(477), 280-292.
- [10] de Haan, L. (1970). *On Regular Variation and its Application to the Weak Convergence of Sample Extremes*. Mathematical Centre Tract 32, Amsterdam.
- [11] Hall, P. (1982). On some Simple Estimates of an Exponent of Regular Variation. *Journal of Royal Statistical Society B*, 44(1), 37-42.
- [12] Hall, P. e Welsh, A.H. (1985). Adaptive estimates of parameters of regular variation. *Annals of Statistics*, 13, 331-341.
- [13] Hill, B.M. (1975). A Simple General Approach to Inference About the Tail of a Distribution. *Annals of Statistics*, 3(5), 1163-1174.
- [14] Jones, M.C. e Faddy, M.J. (2003). A skew extension of the t distribution, with applications. *Journal of Royal Statistical Society B*, 65(2), 159-174.